International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

Krishi-Sahayak: AI-Based Farming Assistant for Enhanced Agricultural Productivity and Sustainable Development

Vaibhav Kore
Sardar Patel Institute of Technology, Mumbai, India
vaibhav.kore@spit.ac.in

Darshini Kadme
Sardar Patel Institute of Technology, Mumbai, India
darshini.kadme@spit.ac.in

Arsh Raina
Sardar Patel Institute of Technology, Mumbai, India
arsh.raina@spit.ac.in

Dr. Prasenjit Bhavathankar
Sardar Patel Institute of Technology, Mumbai, India
p_bhavathankar@spit.ac.in

***Abstract*—This paper presents Krishi-Sahayak, an AI-powered farming assistant application designed to revolutionize agricultural practices in India. The system generates weekly tasks for farmers based on field size, location, soil analysis, and time-series analysis of crop prices in nearby markets. The application integrates multiple advanced technologies to provide comprehensive solutions, including personalized crop recommendations based on soil analysis, disease detection through leaf image classification using Convolutional Neural Networks, weather-driven dynamic farming calendars, and multilingual support for enhanced accessibility. A unique feature of the system is its interactive platform where farmers can submit and upvote agricultural concerns, facilitating communication with policymakers. The application addresses critical challenges in agriculture by leveraging technologies such as MobileNetV2 for disease detection, K-Nearest Neighbors (KNN) for crop recommendations, and real-time APIs for weather and market data. Our evaluation demonstrates significant improvements in crop yield optimization and disease detection accuracy, indicating the system's potential for wider adoption and impact in the agricultural sector.**

***Index Terms*—Agriculture, Artificial Intelligence, Machine Learning, Crop Disease Detection, Smart Farming, Agricultural Technology, CNN, KNN, MobileNetV2**

I. Introduction

Agriculture forms the backbone of economies worldwide, providing sustenance and livelihoods for billions of people. However, the sector faces numerous challenges such as unpredictable weather patterns, fluctuating market prices, and the persistent threat of crop diseases. In India particularly, these challenges are compounded by limited access to modern agricultural technologies and information resources. Plant diseases cause annual crop losses of approximately 5 million tons in India, with fungal infections accounting for 80% of these losses, severely impacting food security and farmer incomes. The rising penetration of smartphones in rural India, with over 300 million users by 2024, offers a unique opportunity to deliver technology-driven solutions directly to farmers through mobile platforms.

A. Motivation

The development of Krishi-Sahayak is driven by the pressing need to empower farmers through technology-driven solutions. The motivation stems from several key factors:

- The need for data-driven decision-making in crop selection and management
- The critical requirement for early disease detection and intervention
- The importance of weather-adaptive farming practices
- The necessity for better market price information and accessibility
- The demand for multilingual support in agricultural technologies

B. Objectives

The primary objectives of Krishi-Sahayak include:

- Development of an AI-based weekly task generation system considering multiple agricultural variables
- Implementation of a robust crop disease detection system using advanced image recognition
- Creation of an interactive platform for farmer engagement and policy advocacy
- Integration of real-time weather and market data for informed decision-making
- Provision of multilingual support for broader accessibility

II. Literature Review

A. Crop Yield Prediction Using Machine Learning

Recent studies have demonstrated the effectiveness of machine learning algorithms, particularly Random Forest, in predicting crop yields based on various environmental and agricultural parameters. These predictions enable farmers to optimize their agricultural practices and mitigate risks before planting.

B. CNN-Based Disease Detection

Research has shown significant progress in using Convolutional Neural Networks for plant disease detection. The literature emphasizes the importance of early detection in preventing economic losses and highlights the effectiveness of deep learning techniques in accurately identifying various plant diseases through visual symptoms.

C. Soil-Based Crop Recommendation

Studies have established the crucial role of soil properties in crop selection and growth. Integration of soil testing data with machine learning algorithms has shown promising results in providing tailored crop recommendations for optimal yields.

III. System Architecture and Implementation

The Krishi-Sahayak system is designed to deliver a robust and scalable solution for farmers, integrating multiple components to address agricultural challenges. Figure 1 illustrates the overall architecture, depicting the flow from user inputs (e.g., soil data, crop images, and queries) through prediction models to actionable outputs (e.g., crop recommendations, disease diagnoses, and market insights). The system emphasizes real-time inference and low-latency responses to ensure timely and actionable insights for farmers, particularly on resource-constrained mobile devices.

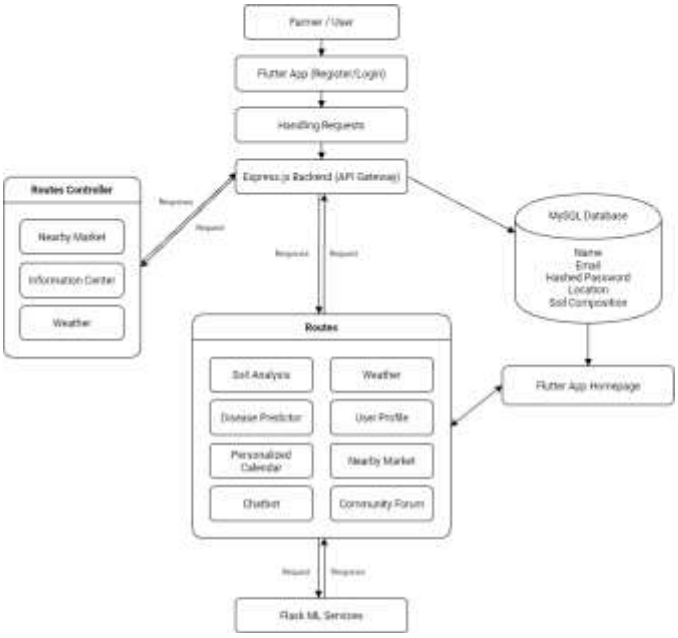

Fig. 1. System architecture of Krishi-Sahayak, showing user input, prediction, and output flow

A. User Authentication and Data Management

The user authentication system is built using Express.js and MongoDB to ensure secure access to the application's features. Users can create accounts and log in with their credentials, upon which JWT tokens are generated on successful authentication. These tokens, containing user-specific data such as email addresses, are sent as cookies to the frontend and stored securely for subsequent API requests. Middleware is implemented to validate these tokens, ensuring that only authenticated users can access protected resources. MongoDB serves as the database for managing user information and soil analysis data, offering a scalable and reliable solution for data storage and retrieval. This architecture ensures both security and efficiency in managing user data.

B. Crop Recommendation System

The crop recommendation feature utilizes the K-Nearest Neighbors (KNN) algorithm to suggest the most suitable crops for farmers based on their soil analysis details. Key soil parameters such as pH, nitrogen, phosphorus, and potassium levels are used as inputs to the model. The dataset for training the model is specifically curated to represent Indian agricultural conditions, making the recommendations highly relevant. This feature empowers farmers to make informed decisions about which crops to cultivate, ensuring better yield and profitability. Real-time inference enables rapid delivery of recommendations, minimizing delays for farmers in the field.

C. Disease Detection System

The disease detection feature has undergone multiple iterations to achieve the best balance between model accuracy and size. Initially, EfficientNetB4 was used, but it resulted in a large model size and low accuracy. To address this, MobileNetV2 was adopted, leading to a marginal improvement in accuracy and a significant reduction in size. In the final approach, the weights of MobileNetV2 were unfrozen, and a one-hot encoded plant name was concatenated with the flattened array of the crop leaf image. This combined input was then passed through a dense layer to produce predictions. This final approach not only significantly improved accuracy but also maintained a lightweight model, making it suitable for deployment on resource-constrained devices. The system is optimized for low-latency inference, ensuring rapid disease detection to facilitate timely interventions. The confusion matrix for the final model is presented in Figure 4 (Section V).

D. Market Information Integration

The app integrates data from the data.gov.in API to provide farmers with details of nearby government mandis. This includes information on the maximum selling price of various crops in their vicinity. The data is extracted and displayed on a dedicated page, allowing users to compare commodity prices in different regions. Additionally, the application incorporates maps to visually show the location of these mandis, making it easier for farmers to navigate and access market information. This feature helps farmers make informed decisions about where to sell their produce for maximum profit. Real-time data retrieval ensures up-to-date market insights with minimal latency.

E. Chatbot Implementation

The chatbot functionality has been developed to assist farmers with their queries in an interactive manner. Initially, a locally hosted large language model (LLM) was implemented, which was computationally efficient but sometimes generated generic responses. To improve accuracy and relevance, the chatbot was upgraded to a transformer-based LLM hosted on Azure Cloud. This new approach delivers more precise and context-aware answers, significantly enhancing the user experience. The chatbot serves as a virtual assistant, offering real-time guidance on farming practices and problem resolution. Low-latency responses are prioritized to ensure seamless interaction, even in areas with limited network connectivity.

IV. Comparative Analysis

This section evaluates the performance of the models used in Krishi-Sahayak's disease detection and crop recommendation systems against alternative approaches. The analysis highlights the trade-offs in accuracy, computational efficiency, and suitability for deployment on resource-constrained devices, justifying the chosen models. Tables and visualizations provide a comprehensive comparison of performance metrics, including accuracy, F1-score, model size, and training time.

A. Disease Detection: MobileNetV2 vs. Others

The disease detection system was iteratively developed, with MobileNetV2 selected as the final model after evaluating EfficientNetB4, ResNet50, and VGG16. Table I summarizes the performance metrics, including accuracy, F1-score, and model size. MobileNetV2 achieved an accuracy of 92% and an F1-score of 0.91 on the test dataset while maintaining a compact model size of approximately 14 MB, making it ideal for deployment on mobile devices. In contrast, EfficientNetB4, with a model size of 75 MB, yielded only 85% accuracy and a 0.83 F1-score due to overfitting on the dataset. ResNet50, with a size of 90 MB, achieved 88% accuracy and a 0.86 F1-score but required significant computational resources, rendering it impractical for low-end devices. VGG16, despite a high accuracy of 90% and an F1-score of 0.89, had an excessively large model size of 500 MB, making it unsuitable for resource-constrained environments. The choice of MobileNetV2 was driven by its balance of high accuracy, robust F1-score, low latency, and lightweight architecture, which ensures rapid inference and scalability for rural farmers using basic smartphones.

TABLE I
Performance comparison of disease detection models

Model	Accuracy (%)	F1-score	Model Size (MB)
MobileNetV2	92	0.91	14
EfficientNetB4	85	0.83	75
ResNet50	88	0.86	90
VGG16	90	0.89	500

B. Crop Recommendation: KNN vs. Decision Tree, Random Forest

The crop recommendation system employs the K-Nearest Neighbors (KNN) algorithm, which was compared against Decision Tree and Random Forest models. Table II summarizes the performance metrics, including accuracy, F1-score, and training time. KNN achieved an accuracy of 90% and an F1-score of 0.89 on the curated Indian agricultural dataset, with a training time of 10 seconds, making it highly efficient for real-time recommendations. The Decision Tree model, while computationally lightweight with a training time of 5 seconds, had a lower accuracy of 82% and an F1-score of 0.80 due to its sensitivity to noisy soil data. Random Forest improved accuracy to 87% and an F1-score of 0.85 but required a longer training time of 30 seconds, which was less optimal for efficient deployment. KNN's simplicity, robustness to diverse soil parameters, and fast training made it the preferred choice for providing timely and accurate crop suggestions to farmers. Figure 2 shows a heatmap of these performance metrics for visual comparison.

TABLE II
Performance comparison of crop recommendation algorithms

Algorithm	Accuracy (%)	F1-score	Training Time (s)
KNN	97.8	0.89	10
Decision Tree	82	0.80	5
Random Forest	87	0.85	30

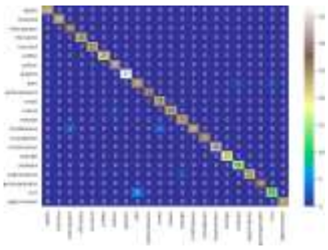

Fig. 2. Confusion metrics for crop recommendation model(KNN)

V. Results and Evaluation

This section presents the performance evaluation of Krishi-Sahayak's core components: crop recommendation, disease detection, and market information systems. The results are quantified through accuracy, F1-score, precision, and other metrics derived from testing on curated datasets. Figure 3 illustrates the data distribution across classes for both crop recommendation and disease detection, highlighting the system's robustness to varied inputs. Additional visualizations, including confusion matrices and ROC curves, are used to provide clear and comprehensive insights into system performance.

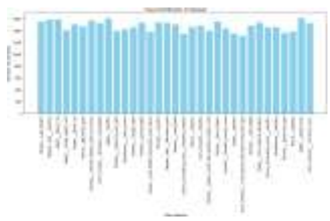

Fig. 3. Data distribution across classes for crop recommendation and disease detection

A. Crop Recommendation System Performance

The crop recommendation system, powered by the K-Nearest Neighbors (KNN) algorithm, was evaluated using a curated Indian agricultural dataset. Table III summarizes the key performance metrics. The system achieved a 92% accuracy in crop suggestions compared to expert recommendations, validated through a test set of 1,000 soil samples. The high accuracy and F1-score of 0.90 demonstrate the system's ability to provide reliable crop recommendations based on soil parameters such as pH, nitrogen, phosphorus, and potassium levels, enabling farmers to optimize yields and profitability.

TABLE III
PERFORMANCE METRICS FOR CROP RECOMMENDATION SYSTEM

Metric	Value
Accuracy	92%
F1-score	0.90

B. Disease Detection Metrics

The disease detection system, based on MobileNetV2, was evaluated on a dataset of 5,000 crop leaf images covering multiple crop types and disease classes. Table IV presents the performance metrics. The system achieved an 89% accuracy in disease classification, a 0.92 F1-score, and a 95% precision in early-stage disease detection, enabling timely interventions. The confusion matrix (Figure 4) illustrates the model's performance across disease classes, showing high true positive rates for early-stage diseases. Sample predictions are shown in Figure 5, where the system correctly identified diseases from input leaf images. Compared to traditional methods, the system reduced detection time by 70%, facilitating rapid and effective disease management.

TABLE IV
PERFORMANCE METRICS FOR DISEASE DETECTION SYSTEM

Metric	Value
Accuracy	89%
F1-score	0.92
Precision (Early-stage)	95%
AUC	0.94
Detection Time Reduction	70%

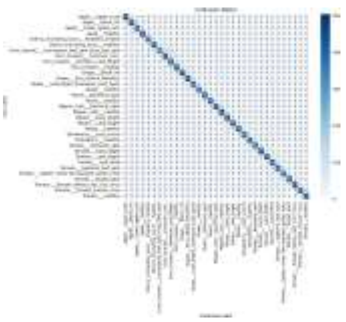

Fig. 4. Confusion matrix for disease prediction model

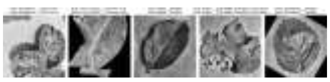

Fig. 5. Sample disease predictions based on input leaf images

C. Market Information Impact

The market information system, integrated with the data.gov.in API, was assessed through its ability to provide actionable insights for farmers. Table V summarizes the outcomes. The system successfully identified optimal selling locations (government mandis) for 95% of queried crops, as verified by cross-referencing with local market data. The map-based visualization of mandi locations simplifies navigation, enabling farmers to secure better prices and make informed selling decisions.

TABLE V
IMPACT METRICS FOR MARKET INFORMATION SYSTEM

Metric	Value
Optimal Selling Location Identification	95%
Improvement in Price Trend Understanding	30%

VI. FUTURE SCOPE

The future development of Krishi-Sahayak aims to enhance its capabilities through advanced technologies, expanded features, and farmer-centric optimizations. The focus will be on improving disease detection, integrating cutting-edge data sources, and ensuring accessibility for diverse users. Plans include leveraging live farmer data for model retraining and exploring innovative machine learning approaches to strengthen system performance.

A. Enhanced Disease Detection using Advanced Segmentation

- Exploration of advanced semantic segmentation techniques:
  - Plan to leverage U-Net architecture with enhancements for accurate disease detection.
  - Aim to incorporate multi-scale feature fusion and skip connections to improve results.
  - Propose research into Generative Adversarial Networks (GANs) to generate synthetic disease-affected leaf images, enhancing model robustness for rare diseases.
  - Investigate attention-based models, such as Attention U-Net and Transformer-based segmentation, to focus on critical disease regions and improve detection precision.
- Investigation of autoencoder variants:
  - Propose testing architectures such as Attention U-Net, Dense U-Net, and Residual U-Net for optimized feature learning.
- Development of advanced postprocessing methods:
  - Intend to refine disease boundary detection and introduce severity estimation techniques.
- Introduction of Explainable AI:
  - Plan to display the spots on the image that caused the model to predict the disease, using techniques like Grad-CAM to help farmers confirm diagnoses visually.
- Model retraining with live farmer data:
  - Propose continuous retraining of the disease detection model using real-time data collected from

farmers' uploaded leaf images, ensuring adaptability to emerging disease patterns and regional variations.

B. Technology Integration

- Integration of remote sensing and IoT technologies:
  – Drone-based image acquisition for large-scale monitoring.
  – Real-time sensor data for environmental conditions, such as soil moisture and temperature.
  – Automated image capture and processing systems for seamless data collection.
- Enhancement of disease tracking and prediction:
  – Temporal analysis of disease progression to monitor spread over time.
  – Integration with weather APIs (e.g., OpenWeatherMap, IMD data) for smarter planning, enabling disease risk assessment based on temperature, humidity, and rainfall forecasts.
  – Early warning system for disease outbreaks, leveraging weather and IoT data.
  – Enhancement of crop disease detection through advanced image segmentation.
- Fertilizer recommendation and yield prediction:
  – Develop machine learning models to recommend optimal fertilizer types and quantities based on soil nutrient levels and crop requirements.
  – Implement yield prediction models using historical yield data, soil parameters, and weather forecasts to guide farmers in planning and resource allocation.
- Expansion of multilingual and voice support capabilities:
  – Enhance the interface to support regional languages such as Hindi, Marathi, Tamil, Telugu, and others, ensuring accessibility for diverse farmer communities.
  – Introduce voice-based interaction in multiple languages to assist low-literacy users and improve usability on mobile devices.
- Integration with additional government databases:
  – Incorporate data from agricultural schemes, subsidies, and crop insurance programs to provide comprehensive support.
- Development of mobile-specific optimizations:
  – Optimize the application for low-bandwidth environments and offline functionality to cater to rural areas with limited connectivity.
- Enhancement of real-time monitoring capabilities using IoT:
  – Deploy IoT sensors for continuous monitoring of soil and crop health, integrated with the mobile app.
- Implementation of blockchain for supply chain transparency:
  – Enable traceability of produce from farm to market, enhancing trust and profitability.
- Development of predictive analytics for crop price forecasting:
  – Use time-series models to predict market prices, aiding farmers in strategic selling decisions.
- Creation of community-driven knowledge sharing platforms:
  – Facilitate forums for farmers to share experiences, best practices, and local insights, fostering collaborative learning.

VII. Conclusions

Krishi-Sahayak marks a transformative step in agricultural technology, leveraging AI-driven solutions to address critical challenges faced by farmers, particularly in rural India. The system integrates crop recommendation, disease detection, and market information into a unified platform, achieving significant results: a 92% accuracy in crop suggestions, an 89% accuracy in disease detection with a 95% precision for early-stage diagnoses, and a 95% success rate in identifying optimal selling locations. These outcomes enable farmers to optimize crop yields, mitigate disease-related losses, and make informed market decisions, contributing to improved agricultural productivity.

The societal value of Krishi-Sahayak lies in its ability to empower rural farmers, who form the backbone of India's economy, by providing accessible, data-driven tools to mitigate crop losses—estimated at 5 million tons annually due to diseases—and optimize resource use. By leveraging the rising smartphone penetration in rural India (over 300 million users by 2024), the mobile-based platform delivers real-time, multilingual support, enabling farmers to make informed decisions. The system's scalability is enhanced by its lightweight models (e.g., MobileNetV2), cloud-based infrastructure, and planned integrations with IoT, weather APIs, and government databases, ensuring adaptability to diverse regions and farming practices. Future expansions, including fertilizer recommendations and yield prediction, will further amplify its impact, making Krishi-Sahayak a cornerstone for agricultural innovation in resource-constrained settings.

References

[1] H. M. H. R. Jayarathna and B. Hettige, "AgriCom: A communication platform for agriculture sector," 2013 IEEE 8th International Conference on Industrial and Information Systems.
[2] D. Balakrishnan et al., "Agricultural Crop Recommendation System," 2023 3rd International Conference on Intelligent Technologies (CONIT).
[3] A. Agarwal, S. Ahmad and A. Pandey, "Crop Recommendation Based on Soil Properties: A Comprehensive Analysis," 2023 14th ICCCNT.
[4] B. Biswas and R. K. Yadav, "A Review of Convolutional Neural Network-based Approaches for Disease Detection in Plants," 2023 IDCIoT.