

OPEN ACCESS

Volume: 13

Special Issue: 2

Month: January

Year: 2026

E-ISSN: 2582-0397

P-ISSN: 2321-788X

Citation:

A, Shajil Kumar P., et al. "An Integrated AI Based Smart Crop Advisory System for Holistic Agricultural Decision Making in Maharashtra." *Shanlax International Journal of Arts, Science and Humanities*, vol. 13, no. 2, 2026, pp. 105–19.

DOI:

<https://doi.org/10.34293/sijash.v13iS2-i1-Jan.11184>

An Integrated AI Based Smart Crop Advisory System for Holistic Agricultural Decision Making in Maharashtra

Shajil Kumar P A

Assistant Professor

Vidyalankar School of Information Technology, Mumbai, Maharashtra, India

Sumit Shinde

Student

Vidyalankar School of Information Technology, Mumbai, Maharashtra, India

Ayush Mithbavkar

Student

Vidyalankar School of Information Technology, Mumbai, Maharashtra, India

Abstract

Agricultural productivity in Maharashtra is constrained by fragmented advisory services, limited access to real time information, and the absence of integrated decision support systems for smallholder farmers. This paper presents an AI based Smart Crop Advisory System that implements a unified and multi modal decision support framework by integrating machine learning, deep learning, and rule-based reasoning. The system comprises seven interconnected components, namely crop recommendation, yield prediction, fertilizer advisory, pest risk estimation, market driven profit prediction, image-based crop disease detection for cotton, maize, and tomato, and treatment recommendation. Real time weather and agricultural market data are incorporated through external application programming interfaces (APIs) to enable dynamic and context aware predictions. The system employs a random forest classifier for crop recommendation, regression-based models for yield prediction, and EfficientNet based deep learning architectures for disease classification. Rule based inference mechanisms and structured expert knowledge are utilized to generate fertilizer and treatment advisories. The framework is implemented as a lightweight web-based application, ensuring accessibility and usability for farmers and users with minimal technical expertise. Experimental observations indicate that the integrated advisory system effectively supports data driven agricultural decision making by combining predictive analytics, real time environmental inputs, and expert driven recommendations. The system is scalable, regionally adaptable, and demonstrates strong potential for enhancing sustainable agricultural practices and improving productivity in diverse agro climatic conditions.

Keywords: Smart Crop Advisory System, Machine Learning, Deep Learning, Decision Support Systems, Precision Agriculture

Introduction

Agriculture remains a primary source of livelihood and food security for a significant portion of the population in Maharashtra. Despite its importance, farmers across the state continue to face persistent structural challenges that adversely affect sustainable agricultural production and

economic viability. Variability in monsoon onset and intensity, increasing frequency of droughts, declining soil fertility, and rising occurrences of extreme weather events have substantially increased uncertainty in crop planning and management. These climatic instabilities, coupled with temperature fluctuations, have made traditional agricultural decision making increasingly unreliable. In addition to climatic constraints, farmers encounter several agronomic and economic challenges, including nutrient imbalanced soils, delayed or inaccurate disease identification, escalating pest infestations, and highly volatile market prices. Limited access to timely, reliable, and localized advisory services further aggravates these issues. Although several digital agricultural tools are available, they are often fragmented, task specific, and inadequately validated for field level decision support. As a result, many farmers continue to depend largely on indigenous knowledge and experience rather than data driven insights, widening the knowledge and productivity gap.

Existing digital agriculture solutions typically focus on isolated functionalities such as crop recommendation, disease detection, or weather forecasting. The lack of integrated analysis across multiple stages of the crop lifecycle restricts their effectiveness and limits adoption among farmers who require cohesive, end-to-end decision support. The absence of a unified platform capable of synthesizing agronomic, environmental, and economic factors remains a significant barrier to the large scale adoption of advanced agricultural technologies. To address these limitations, this paper presents an Integrated AI Based Smart Crop Advisory System, designed as a unified, multi modal decision support platform. The system integrates machine learning, deep learning, rule based reasoning, and real time open data sources to enable comprehensive agricultural advisory services. By combining analytical insights across multiple layers of the crop production cycle, the system aims to support informed decision making and reduce production risk. The developed model consists of seven interconnected components: (i) a crop recommendation model that evaluates suitable crops based on soil nutrients and prevailing weather conditions; (ii) a yield prediction model that estimates potential crop yield using state level agricultural statistics, soil characteristics, weather data, and geographic information; (iii) a fertilizer advisory model that applies rule based expert knowledge to generate nutrient specific fertilizer recommendations; (iv) a pest risk assessment model that estimates pest susceptibility using environmental parameters; (v) a profit estimation model that predicts economic returns based on expected yield and real time or historical market prices; (vi) a disease detection model that employs an EfficientNet based deep learning architecture for leaf image based disease identification; and (vii) a treatment advisor model that provides chemical, biological, and cultural management recommendations tailored to the identified disease and crop type. By integrating these components into a single platform and interfacing with real-time weather and market application programming interfaces (APIs), the model delivers timely, contextual, and actionable recommendations to farmers. This integrated approach is intended to enhance agricultural decision making, mitigate production risks, and promote sustainable farming practices across the diverse agro climatic regions of Maharashtra.

Literature Review

A systematic literature review was conducted on agricultural decision support systems, machine learning approaches for crop recommendation, yield forecasting methodologies, deep learning techniques for plant disease detection, and rule based agricultural advisory frameworks. Existing studies demonstrate substantial progress across individual components of agricultural decision making, including soil and nutrient based crop selection, climate driven yield prediction, image based disease diagnosis, fertilizer advisory mechanisms, and weather dependent pest risk assessment. However, the review reveals that most prior research efforts have primarily focused on isolated or task specific models, addressing individual stages of the agricultural lifecycle independently. The absence of integrated, multi modal advisory frameworks limits the practical applicability of these systems in real world farming environments. Consequently, there exists a clear research gap for a unified decision support system that combines predictive analytics, expert knowledge, and real time environmental data to provide holistic and continuous agricultural guidance across all stages of crop planning, cultivation, and management.

Recent studies have extensively investigated machine learning approaches for crop recommendation using soil nutrient and climatic attributes. Research reported in [1] and [16] establishes that ensemble learning models, particularly random forest and support vector machine classifiers, consistently achieve high predictive accuracy by capturing complex nonlinear interactions among nitrogen, phosphorus, potassium, soil pH, and meteorological variables. These findings indicate the suitability of ensemble classifiers for region specific crop advisory applications. Further investigations in [2] reinforce these observations by demonstrating that decision tree and random forest models offer a favorable balance between prediction performance and interpretability, which is critical for adoption in farmer facing decision support systems.

Crop yield prediction has been addressed using a variety of statistical, regression, and ensemble learning techniques. Studies presented in [3] and [17] show that ensemble regression models outperform conventional linear regression methods in terms of accuracy and robustness under heterogeneous climatic conditions. A systematic review in [4] consolidates these findings by identifying random forest, gradient boosting, convolutional neural networks, and long short term memory models as the most effective yield prediction approaches, with rainfall, temperature, and soil characteristics emerging as dominant explanatory variables. Deep learning based disease detection has become a prominent research area due to advancements in image recognition. The work in [5] demonstrates that convolutional neural networks significantly outperform traditional machine learning techniques in plant disease classification tasks. Further improvements are reported in [6] and [18], where EfficientNet and transfer learning based architectures achieve high classification accuracy while reducing model complexity and training requirements, thereby enhancing feasibility for real time agricultural applications. Hybrid deep learning frameworks have been explored to capture both spatial and temporal characteristics of agricultural data. Research in [7] illustrates that combining convolutional neural networks with long short term memory networks enables effective modeling of weather sequences and crop growth dynamics, resulting in improved predictive performance across multiple agricultural forecasting tasks.

The integration of Internet of Things infrastructure with machine learning models has been shown to enhance real time agricultural decision making. The study in [8] highlights that sensor driven data streams significantly improve irrigation management, fertilizer application strategies, and crop monitoring efficiency. Research in [9] emphasizes the role of explainable artificial intelligence techniques, such as feature importance analysis and SHAP based methods, in improving transparency and trust in machine learning driven agricultural advisory systems. Remote sensing and satellite imagery have further expanded the scope of agricultural analytics. Studies in [10] and [11] demonstrate that machine learning and deep learning models trained on multispectral imagery and vegetation indices enable accurate large scale yield prediction and early detection of crop stress across diverse agro climatic regions. Weather driven forecasting models have also received considerable attention in agricultural research. Findings reported in [12] confirm that incorporating real time temperature, rainfall, and humidity data substantially enhances prediction accuracy and strengthens early warning mechanisms for agricultural risk management. Rule based and knowledge driven decision support systems continue to play an essential role in agricultural advisory services. Research presented in [13] and [19] shows that expert systems provide transparent and reliable fertilizer and disease treatment recommendations based on soil nutrient analysis, thereby improving interpretability and farmer confidence. Data driven fertilizer recommendation systems discussed in [14] further contribute to sustainable agriculture by optimizing nutrient usage and reducing long term soil degradation. Recent efforts have focused on developing integrated smart crop advisory platforms that combine multiple analytical components. The studies in [15] and [20] demonstrate that unified systems integrating crop recommendation, yield prediction, disease detection, and real time weather and market data significantly improve decision accuracy and system usability. These findings collectively indicate that integrated multi modal frameworks are more effective than isolated analytical models in supporting comprehensive agricultural decision making.

Existing literature demonstrates significant progress in the development of individual components of agricultural decision support systems. However, the findings collectively indicate a persistent research gap in the design and implementation of scalable, integrated, and end-to-end advisory platforms that can effectively support farmers across all stages of the crop lifecycle. Despite the extensive body of literature on individual agricultural prediction and advisory models, empirical evaluations and real-world case studies addressing several critical aspects remain limited. Existing studies largely lack unified and multimodal advisory frameworks, comprehensive integration of machine learning, deep learning, and rule-based reasoning, and systematic incorporation of real-time application programming interfaces. Furthermore, there is a notable absence of datasets and experimental evaluations tailored to the diverse agro-climatic regions of Maharashtra, as well as end-to-end decision support systems that span the complete crop lifecycle from crop selection to disease diagnosis and treatment. This research addresses these gaps by proposing an integrated Smart Crop Advisory System that combines seven predictive and rule-based modules within a scalable and cohesive framework. The system is specifically designed to operate under the agro-climatic conditions of Maharashtra and provides continuous decision support across all major stages of agricultural planning, cultivation, and crop health management.

Research Objectives

The primary objective of this research is to design and implement an integrated, data-driven Smart Crop Advisory System that supports informed agricultural decision-making across multiple stages of the crop lifecycle. The study aims to combine predictive modeling, rule-based reasoning, and real-time data integration within a unified framework tailored to the agro-climatic conditions of Maharashtra. This research initiative advances the stated objectives as follows:

- To develop a machine learning-based model that recommends suitable crops based on soil nutrient composition and prevailing weather conditions.
- To design a regression-based yield prediction model that estimates expected crop yield using district-specific agro-climatic parameters.
- To develop a rule-based fertilizer advisory system that generates nutrient-specific recommendations to address soil nutrient imbalances.
- To design a rule-based pest risk assessment model that evaluates pest susceptibility using crop growth stage, infestation percentage, crop variety, fertilizer application, pesticide usage history, alerts from neighboring fields, prevailing weather conditions, and trap count observations.
- To develop an image-based crop disease detection model using an EfficientNet deep learning architecture for identifying leaf diseases in cotton, maize, and tomato crops.
- To design a rule-based treatment advisory module that provides chemical, biological, and cultural control recommendations for diseases detected in crops.
- To integrate application programming interfaces for real-time weather alerts and agricultural market information to dynamically generate advisory recommendations and profit estimations.
- To deploy and integrate all seven modules within a unified Flask-based web application that functions as a single decision support platform for farmers.
- To evaluate and compare the performance of each model using appropriate accuracy metrics, numerical error measures, and classification evaluation criteria.
- To assess the overall system workflow with respect to scalability, usability, and practical relevance for farmers across the agro-climatic regions of Maharashtra.

Research Methodology

The Smart Crop Advisory System was developed following a systematic research methodology that integrates theoretical analysis, experimental modeling, and rule-based system design. The development

process comprised multiple stages, including dataset exploration, model formulation, model training, rule construction, application programming interface development, and software deployment. In addition, several established digital agriculture and machine learning research repositories were examined to identify effective modeling strategies and to guide the overall system architecture. The following methodological steps were adopted:

- Investigation of key factors influencing crop suitability, including soil nutrient composition, weather parameters, and regional agro climatic conditions.
- Review of regression-based crop yield prediction techniques and identification of relevant features aligned with district specific cropping patterns in Maharashtra.
- Examination and selection of deep learning architectures, with particular emphasis on EfficientNet models, for image-based crop disease recognition.
- Design and implementation of rule-based systems for fertilizer advisory, pest risk assessment, and treatment recommendation using agricultural extension literature, standard agronomic guidelines, and expert defined rules.
- Definition of the logical workflow and decision pipeline to enable seamless integration of the seven analytical modules into a unified and coherent advisory system.

Data Preparation and Validation

Dataset preparation was performed as a structured methodological step to ensure data quality, consistency, and suitability for machine learning and deep learning-based analysis. Tabular datasets related to soil properties, crop characteristics, and agro climatic conditions were collected from multiple agricultural sources and consolidated into standardized formats. Data cleaning procedures were applied to remove inconsistencies and missing values, followed by normalization and scaling of numerical attributes to support stable model training. Categorical features were standardized and encoded to maintain uniform representation across regions and crops, ensuring compatibility with the crop recommendation and yield prediction models.

For disease detection, crop leaf image datasets corresponding to cotton, maize, and tomato were organized by crop and disease class and processed through a dedicated preprocessing pipeline. Images were resized to a uniform resolution and augmented using geometric and photometric transformations to address class imbalance and improve model generalization under varying field conditions. Additionally, external weather and market price application programming interfaces were evaluated for data reliability, temporal consistency, and integration feasibility. These validated data sources were incorporated into the system to support real time advisory generation. Collectively, these dataset preparation and validation procedures established a reliable data foundation for the integrated Smart Crop Advisory System.

API Integration

To support real time advisory services and enable dynamic data driven decision making, multiple external application programming interfaces were integrated into the model workflows of the proposed system. These APIs provide access to essential external data sources, including weather forecast information obtained through the OpenWeatherMap API and agricultural commodity price data at the mandi level retrieved from the official Data.gov.in API. For weather data acquisition, users select a district through a district to city mapping layer that links administrative regions with valid geographic coordinates supported by the OpenWeatherMap service. Upon selection, the system retrieves a five-day weather forecast at three-hour intervals and aggregates the data to derive mean temperature, mean relative humidity, and cumulative rainfall values. These standardized environmental variables are subsequently utilized by the Crop Recommendation Model, Fertilizer Advisory Model, and Yield Prediction Model. In addition, the mapping layer supports approximate string matching of district names, thereby validating user inputs and reducing errors caused by spelling variations or invalid geographic entries.

Agricultural commodity price integration follows a hierarchical query mechanism like standard mandi price systems. The Data.gov.in API enables filtered retrieval of price records across states, districts, markets, and commodities, providing minimum, maximum, and modal price values for each crop. The modal price is used as the representative market price for profit estimation. All API requests are internally validated and standardized to ensure consistent response formats and to prevent failures arising from incomplete or invalid input parameters. In scenarios where real time API access is unavailable or no matching records are returned; the system automatically switches to a fallback mechanism that utilizes locally stored historical price datasets or user specified price inputs. This layered integration strategy ensures uninterrupted system operation under conditions of network instability and variable data availability. By combining standardized API protocols with trained predictive models, rule based advisory components, and web-based interfaces, the system delivers timely, reliable, and location specific decision support to farmers across diverse agricultural regions.

System Implementation Framework

The Smart Crop Advisory System was developed as a modular web-based application using the Flask framework to ensure scalability, maintainability, and seamless integration of multiple analytical components. Each functional component, including crop recommendation, yield prediction, fertilizer advisory, pest risk assessment, disease detection, treatment advisory, and profit estimation, was implemented as an independent backend module. This modular design enables isolated development, testing, and enhancement of individual models while maintaining interoperability within the unified system architecture. Flask route handlers were designed to manage user input processing, invoke the relevant predictive or rule-based modules, and aggregate multi model outputs for presentation to the user. The front-end interface was developed using HyperText Markup Language, Cascading Style Sheets, and Flask template rendering to provide an intuitive and interactive user experience. The interface supports structured data entry through text fields and dropdown menus, image uploads for disease detection, and real time display of advisory outputs. In addition, auxiliary Python utility scripts were implemented to support image preprocessing, application programming interface communication, rule engine execution, and file handling operations required across the system.

The overall methodology employed in this research encompasses theoretical model design, experimental machine learning and deep learning model training, rule-based system formulation, real time application programming interface integration, and Flask based software development. This multistage methodological approach ensures that the Smart Crop Advisory System is both scientifically rigorous and practically deployable, while remaining computationally efficient and contextually relevant for agricultural decision making in the agro climatic regions of Maharashtra.

AI Based Smart Crop Advisory System Architecture

The system level architecture of the proposed AI Based Smart Crop Advisory System illustrates an integrated decision support framework that consolidates multiple data sources, analytical models, and advisory outputs within a unified platform, as shown in Fig. 1. The system ingests heterogeneous inputs including soil nutrient data, weather information, farm management practices, market price data obtained through external APIs, and crop leaf images provided by users. These inputs are pre-processed and routed to the appropriate analytical modules to ensure seamless interaction between machine learning models, deep learning models, and rule-based components.

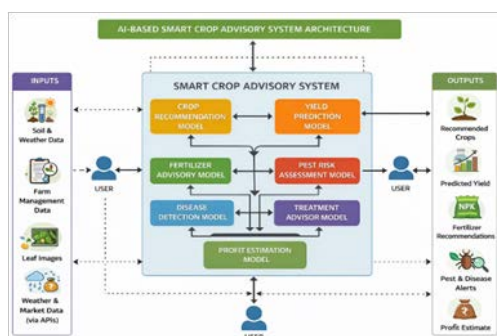


Figure 1 System Level Architecture of the AI Based Smart Crop Advisory System

The core of the architecture comprises seven interconnected modules, namely the Crop Recommendation Model, Yield Prediction Model, Fertilizer Advisory Model, Pest Risk Assessment Model, Disease Detection Model, Treatment Advisor Model, and Profit Estimation Model. Each module performs a specialized analytical function while sharing contextual information with other components to support multi-stage agricultural decision making. The outputs generated by these modules include recommended crops, predicted yields, fertilizer guidance, pest and disease alerts, treatment recommendations, and profit estimates, which are collectively presented to the farmer through a unified user interface. This integrated architecture enables end-to-end advisory support, enhances decision accuracy, and ensures that recommendations are timely, context-aware, and relevant to diverse agro-climatic conditions.

Crop Recommendation Model

A Random Forest classification model was developed to recommend suitable crop types based on soil nutrient composition and prevailing climatic conditions. Multiple datasets obtained from the agricultural sector of Maharashtra were integrated to construct a consolidated crop recommendation dataset. Crop nomenclature was standardized through collaborative agricultural data collection processes and missing relative humidity values were imputed using crop specific mappings derived from a prior crop research study. The final dataset comprised seven continuous predictor variables representing key soil and environmental attributes. The Random Forest algorithm was selected due to its ability to effectively model complex nonlinear interactions among multiple agricultural parameters without requiring feature scaling. The dataset was partitioned using an 80:20 train test split for model development and evaluation. The final model consisted of 100 decision trees and achieved an overall classification accuracy of 99.3%, with all crop classes exhibiting consistently high precision and statistically comparable performance. The crop recommendation process begins with the collection and pre-processing of soil nutrients and weather data, which are analysed using a Random Forest classifier to determine the suitability of different crops, as illustrated in Fig. 2. The model produces ranked crop recommendations with associated probability scores that support informed crop selection decisions.

The analysis indicated that soil moisture and soil nutrient parameters were the most influential factors in crop type selection. The trained model generates class probability scores for all potential crop categories, enabling the system to provide ranked crop recommendations based on prevailing environmental conditions. The model is deployed as an active component of the Smart Crop Advisory System and is continuously updated. It is accessible to farmers through a REST based application programming interface or via manual input interfaces, allowing real time crop recommendation based on soil and weather conditions.

Yield Prediction Model

The yield prediction module estimates crop productivity in terms of tons per hectare prior to harvest by incorporating both environmental and management related factors influencing crop production. To develop

the prediction model, a structured dataset titled `crop_yield.csv` was constructed, containing attributes including region, soil type, crop type, temperature, days to harvest, prevailing weather conditions, and management indicators representing fertilizer usage and irrigation practices. Categorical variables, namely region, soil type, crop type, and weather condition, were encoded using label encoding, while binary management indicators were converted to integer representations. The target variable was defined as yield expressed in tons per hectare, with all remaining attributes treated as predictor variables. Numerical features including rainfall, temperature, and days to harvest were standardized using a `StandardScaler` to address differences in numeric scales. The dataset was subsequently partitioned into training and testing subsets using an 80:20 split. A Random Forest regression model was selected due to its ability to capture nonlinear relationships between environmental variables, crop management practices, and yield outcomes, as well as its robustness to outliers and heterogeneous data types. The model was configured with 100 decision trees and a maximum tree depth of 10, using a fixed random state for reproducibility. Model performance was evaluated using mean absolute error, mean squared error, root mean squared error, and the coefficient of determination. Experimental results indicate that the Random Forest regressor achieved low error values and a high coefficient of determination, demonstrating strong agreement between predicted and observed yield values. Feature importance analysis revealed that rainfall, fertilizer usage, and irrigation practices were the most influential factors affecting yield, underscoring the critical role of water availability and input management, while crop type, soil characteristics, and weather conditions provided additional contextual refinement to the yield estimates. The yield prediction process integrates the selected crop with regional, climatic, and farm management parameters to estimate expected productivity using a Random Forest regression model, as depicted in Fig. 2. The resulting yield values expressed in tons per hectare assist farmers in effective planning and resource management.

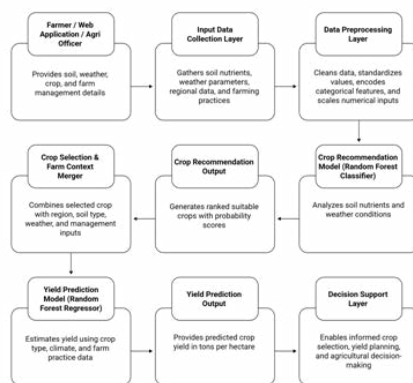


Figure 2 Architecture of the Crop Recommendation and Yield Prediction Model in the Smart Crop Advisory System

For deployment, the trained regression model, fitted scaler, label encoders, and ordered feature mappings were serialized into a unified package (`crop_yield_predictor.pkl`). During inference, farmer provided inputs are converted into a structured data format, with categorical attributes transformed using the stored label encoders, boolean inputs cast to integer values, and numerical features scaled using the pretrained scaler prior to prediction. This deployment strategy enables farmers to obtain real time yield estimates for specific crop, region, and management practice combinations, thereby supporting informed decision making related to production planning, resource allocation, and market strategy.

Fertilizer Advisory Model

The Fertilizer Advisory Model is implemented as a rule-based system rather than a learned predictive model, with the objective of providing transparent and context specific fertilizer recommendations. The rule base is defined in a structured JavaScript Object Notation format and consists of conditional if then statements that map soil test results, including nitrogen, phosphorus, potassium, and soil pH, to crop specific fertilizer advisories. In addition, recipe-oriented weather heuristics are incorporated to enhance decision reliability. For the supported crops, namely rice, wheat, cotton, maize, soybean, tomato, and selected horticultural crops, the rule base specifies recommended fertilizer sources and application strategies corresponding to three nutrient availability levels, defined as low, moderate, and high, and three soil pH categories, defined as acidic, neutral, and alkaline. The Fertilizer Advisory Model follows a structured rule-based workflow in which crop specific soil nutrient data and weather information are processed through validation, nutrient band classification, and conditional inference mechanisms to generate fertilizer recommendations, as illustrated in Fig. 3. The integration of weather safeguard rules and decision-making support logic ensures that the advisory output provides nutrient wise guidance along with precautionary warnings to support safe and effective fertilizer application.

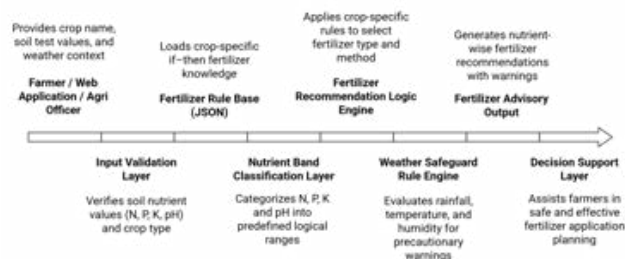


Figure 3 Architecture of the Rule-Based Fertilizer Advisory Model

During execution, the model accepts crop type and quantitative soil parameters as inputs and assigns each nutrient value to predefined threshold-based bands. Based on these classifications, the system retrieves the corresponding textual fertilizer recommendations from the rule base, such as split applications of nitrogen fertilizers, basal phosphorus application, potassium supplementation at critical growth stages, or indications that additional nutrient application is not required. In addition to soil driven rules, the model incorporates meteorological safeguard meta rules that generate advisory warnings under adverse weather conditions, including heavy rainfall, extreme temperatures, and high humidity. These safeguards advise farmers to postpone fertilizer applications to minimize nutrient loss and reduce the risk of unfavorable application conditions. The rule-based design ensures deterministic and interpretable outputs and enables straightforward updates to the advisory logic by modifying the underlying JSON file, thereby supporting real time deployment and adaptability to localized agronomic knowledge within the web-based application framework.

Pest Risk Assessment Model

The Pest Risk Assessment Model is implemented as a transparent, rule-based scoring framework rather than a learned predictive model, with the objective of estimating pest susceptibility using agronomic and field level inputs provided by the user. The model computes a composite risk index derived from multiple contributing factors, including crop growth stage with respect to susceptibility, observed infestation percentage, varietal resistance classification, fertilizer usage practices, recent pesticide application history, reports of pest incidence from neighboring fields, prevailing weather conditions favorable for pest proliferation, and trap count observations. Each contributing factor is mapped to a predefined integer score ranging from one to

three, where higher values represent increased biological risk, while protective factors such as recent pesticide application contribute a negative score to the overall risk index. The Pest Risk Assessment Model operates through a sequential workflow in which agronomic field inputs and environmental data are transformed into quantitative risk scores and aggregated into a composite pest risk index, as illustrated in Fig. 4. The resulting risk classification and decision support outputs assist farmers in identifying appropriate intervention strategies based on the assessed level of pest risk.

The composite pest risk score is calculated by aggregating the individual factor scores and is subsequently classified into three qualitative risk categories, namely low risk for scores less than or equal to four, medium risk for scores between five and eight, and high risk for scores greater than or equal to nine.



Figure 4 Architecture of the Rule-Based Pest Risk Assessment Model

In addition to providing an intuitive categorical output, the model retains the quantitative risk score to facilitate future refinement and sensitivity analysis. The rule-based formulation ensures interpretability and enables straightforward customization through adjustment of scoring weights and threshold values. Furthermore, the modular design of the scoring framework allows for progressive enhancement as additional agronomic data, pest specific information, or region-specific crop calendars become available.

Profit Estimation Model

The Profit Estimation Model integrates predicted crop yield information with market price data to enable farmers to estimate potential revenue prior to harvest. The model supports both real time price retrieval through external application programming interfaces and a fallback mechanism based on user supplied or locally stored price datasets, thereby ensuring robustness under varying connectivity conditions. In automatic mode, the model interfaces with the official agricultural market price API provided by api.data.gov.in to retrieve commodity specific price records. Helper functions are employed to query state, district, market, and commodity attributes to identify relevant price entries. The retrieved records typically include minimum, maximum, and modal prices corresponding to the selected market context, with the modal price used as the representative market value. Through the predictable profit endpoint implemented in the Flask framework, users provide expected yield values expressed in quintals along with location and commodity details. When valid pricing records are available, the system computes the estimated profit using the relation $\text{Profit} = \text{Yield (quintals)} \times \text{Modal Price (₹ per quintal)}$. The Profit Estimation Model follows a structured workflow in which predicted crop yield and location specific inputs are processed through market price retrieval, validation, and fallback mechanisms to compute expected revenue, as illustrated in Fig. 5. The integration of real time market APIs with manual price overrides ensures robust profit estimation and supports informed crop planning and marketing decisions.

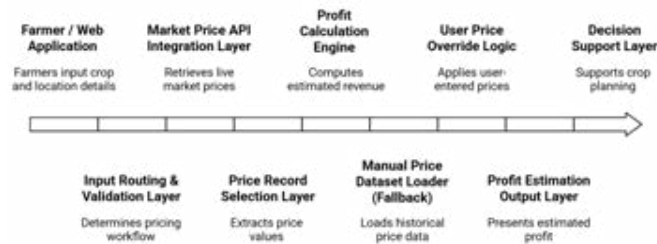


Figure 5 Architecture of the Market Driven Profit Estimation Model

To ensure operational continuity, the model also provides a manual pricing mode that is activated when real time API data are unavailable or when no matching records are returned. In this mode, a pre-processed historical price dataset specific to Maharashtra is loaded, and a separate endpoint enables cascading selections of district, market, commodity, and grade. The filtered grade level record supplies minimum, maximum, and modal price values derived from historical data. If pricing information is still unavailable, the system allows direct user specified price input or defaults to stored modal price values, while maintaining the same profit computation formula. This dual mode pricing strategy enhances transparency, minimizes computational overhead, and ensures reliable profit estimation even in scenarios with limited or intermittent network connectivity.

Disease Detection Model

The Disease Detection Model of the Smart Crop Advisory System employs three separate Convolutional Neural Network architectures developed using EfficientNet based transfer learning, with individual models trained for cotton, maize, and tomato crops. For each crop, leaf images were organized into class specific directories representing healthy leaves and the most prevalent diseases affecting the crop. The cotton dataset included powdery mildew, target spot, and bacterial blight; the maize dataset included common rust, grey leaf spot, and blight; and the tomato dataset included multiple foliar and viral diseases. Class balancing strategies were applied to improve model generalization, including geo and photometric data augmentation for underrepresented maize and tomato classes, and class weighting for the cotton dataset to address imbalance between healthy and diseased samples. All images were resized to 224×224 pixels and pre-processed according to the input requirements of the selected EfficientNet architectures. EfficientNetB3 was selected as the backbone network for cotton disease classification, incorporating a global average pooling layer, dropout regularization, and a softmax based classification head. Initially, the pre-trained base layers were frozen and only the classification head was trained, followed by gradual unfreezing and fine tuning of deeper layers. For maize and tomato disease detection, EfficientNetB0 was employed using TensorFlow image datasets loaded from directory with on-the-fly augmentation, including random flipping, rotation, zooming, contrast variation, and dropout regularization. A two-stage transfer learning strategy was similarly adopted, where the classification layers were trained first, followed by progressive fine tuning of the pre-trained convolutional layers. The Disease Detection Model follows a deep learning-based workflow in which crop leaf images are collected, pre-processed, and analysed using EfficientNet based transfer learning to extract discriminative visual features, as illustrated in Fig. 6. The classification output provides disease labels along with confidence scores, enabling timely and reliable disease identification to support informed crop health management decisions.

Model performance was evaluated using accuracy metrics, confusion matrices, and class wise precision, recall, and F1 scores computed on held out validation or test datasets. The maize model achieved the highest overall accuracy of approximately 94%, while the cotton and tomato models achieved accuracies of approximately 82% and 83%, respectively. Performance degradation was primarily observed in visually

similar disease classes, particularly among tomato blight categories, indicating the need for additional high quality training samples. Upon completion of training, each model was saved in Hierarchical Data Format and integrated into the advisory system through a prediction interface that allows users to upload leaf images.

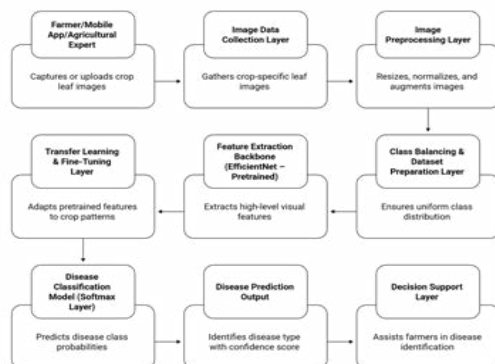


Figure 6 Architecture of the Disease Detection Model using Deep Learning

The system processes the input image and returns the predicted disease class along with an associated confidence score, enabling near real time disease identification within the Smart Crop Advisory System.

Treatment Advisor Model

The Treatment Advisory Model is implemented using a structured rule-based knowledge repository stored in JavaScript Object Notation format to enable efficient indexing and retrieval of crop and disease specific treatment information. Each crop is represented by a dedicated record containing treatment details for all supported crop disease combinations, allowing targeted queries using crop type and disease category as search parameters. For each crop disease combination, the advisory framework provides three complementary treatment strategies, namely chemical control, organic or biological control, and recommended cultural practices. The chemical control strategy specifies appropriate fungicides or insecticides along with recommended application rates, while the organic or biological strategy includes approved bio-based products such as neem oil, *Trichoderma* species, and *Pseudomonas fluorescens*. Cultural practice recommendations include the use of resistant varieties, field sanitation, crop rotation, optimized plant spacing, and vector management to reduce disease incidence. The Treatment Advisory Model operates through a rule-based workflow in which crop and disease information is validated and matched against a structured knowledge base to retrieve appropriate treatment strategies, as illustrated in Fig. 7.

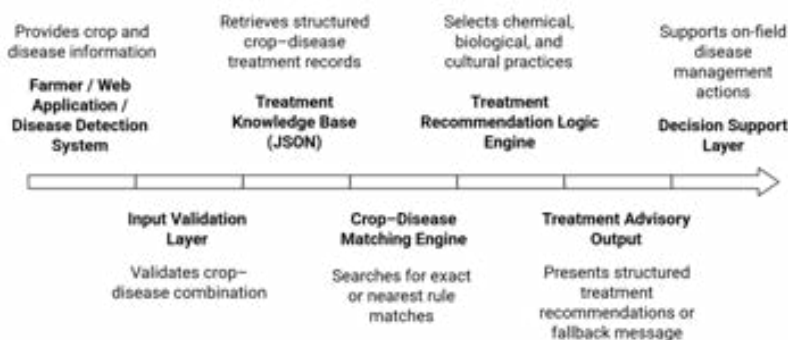


Figure 7 Architecture of the Rule-Based Treatment Advisory Model

The system generates structured recommendations encompassing chemical, biological, and cultural practices, thereby enabling timely and actionable disease management decisions at the field level. During execution, the model accepts the crop type and predicted disease label generated by the Disease Detection Model or selected by the user and retrieves the corresponding treatment recommendations in a structured and interpretable format. In cases where a matching crop disease entry is unavailable, the system returns a standardized no data available message. Each crop record also includes a dedicated healthy category to indicate the absence of disease and to provide preventive and integrated pest management guidance. The use of a standalone JSON based knowledge base ensures transparency, ease of maintenance, and rapid updates without the need for model retraining, thereby effectively bridging the gap between image-based disease diagnosis and actionable field level intervention.

Model Assessment

The comparative performance of the system developed with machine learning and deep learning models is summarized in Table 1, highlighting the algorithms, task types, input features, and evaluation metrics used across different modules of the Smart Crop Advisory System. The crop recommendation model, implemented using a Random Forest classifier, achieves high accuracy when analyzing soil and weather attributes, while the yield prediction model demonstrates strong regression performance using agro climatic and farm management features.

Table 1 Comparative Performance Summary of the Model

Characteristic	Algorithm	Task Type	Input	Performance
Crop Recommendation	Random Forest Classifier	Multi class Classification	Soil and weather attributes	Accuracy: 99.33%
Yield Prediction	Random Forest Regressor	Regression	Agro climatic and management features	MAE: 0.4002, RMSE: 0.5016, R ² : 0.9172
Cotton Disease Detection	EfficientNetB3 CNN	Image Classification	Cotton disease leaf images	Validation Accuracy: ~82%
Maize Disease Detection	EfficientNetB0 CNN	Image Classification	Maize leaf disease images	Validation Accuracy: ~94%
Tomato Disease Detection	EfficientNetB0 CNN	Image Classification	Tomato leaf disease dataset	Test Accuracy: ~83%

The disease detection models employ EfficientNet based convolutional neural networks and exhibit reliable image classification accuracy for cotton, maize, and tomato leaf disease datasets. Collectively, the results reported in Table 1 indicate that the selected models are well suited for their respective tasks and contribute effectively to the overall decision support capability of the system.

Suggestion and Recommendations

Drawing upon the results and insights obtained from all seven models of the Smart Crop Advisory System, several key observations and recommendations are identified to further enhance agricultural decision making and system effectiveness, as listed below:

- The integrated advisory framework demonstrates that multi modal recommendations achieved by combining predictive models, rule based reasoning, and real time data sources provide more robust and reliable guidance than isolated tools.
- The accuracy of crop recommendation outcomes improves significantly when updated soil nutrient and weather information is available. It is recommended that farmers periodically update soil nutrient data, preferably at least once per cropping season, to ensure higher precision in advisory outputs.

- Yield prediction performance benefits from the availability of consistent and complete environmental datasets. Expanding region specific datasets to cover additional districts within Maharashtra is expected to further enhance prediction accuracy and model stability.
- Disease detection performance can be improved through the inclusion of a larger number of locally relevant leaf images captured under diverse lighting and field conditions. Increasing the size and diversity of image datasets for cotton, maize, and tomato crops is likely to strengthen model robustness and generalization capability.
- Rule based fertilizer advisory systems should be updated on a regular basis to reflect the latest agricultural extension guidelines, soil health studies, and fertilizer usage norms to maintain recommendation relevance and agronomic validity.
- The integration of additional field reported trap count data and pest incidence alerts from neighboring farms is expected to improve the accuracy of pest risk assessment and provide a more comprehensive understanding of real time outbreak dynamics.
- Incorporating locally available and cost effective organic and biological pest management alternatives within treatment recommendations is essential for promoting sustainable and environmentally responsible farming practices.
- The overall effectiveness of the advisory system is closely linked to the availability of real time weather and market data services. The use of multiple backup application programming interfaces is recommended to minimize service disruptions and ensure continuous delivery of advisory outputs.
- Enhancing the web based user interface to support multilingual access, including regional languages such as Marathi and Hindi, is expected to improve system usability and adoption among farmers across Maharashtra.
- Conducting periodic field validation studies is essential to assess the real world impact of advisory recommendations on farmer decision making, crop yield outcomes, and efficiency of pest management expenditures.

These recommendations collectively aim to improve the accuracy, scalability, and long term sustainability of the Smart Crop Advisory System, while also supporting future research efforts toward the development of comprehensive and integrated digital agriculture platforms.

Conclusion

The objective of this research is to develop a unified Smart Crop Advisory System that supports data driven agricultural decision making for farmers. The study began with an examination of the key factors influencing crop selection, yield estimation, fertilizer requirements, pest risk, disease identification, and treatment planning. Based on these considerations, individual models were designed and implemented, including a machine learning based crop recommendation model, a regression based yield prediction model, a rule based fertilizer advisory module, a multi parameter pest risk assessment framework, an EfficientNet driven disease detection model, and profit and treatment estimation modules. The system integrates machine learning, deep learning, rule based reasoning, and real time weather and market data through application programming interfaces within a single decision support platform. All modules were developed using the Python programming language and deployed using the Flask web framework, with each component validated using appropriate classification, regression, or rule based evaluation methodologies. Experimental results indicate that the combined use of predictive analytics, expert knowledge, and real time data sources produces more accurate, timely, and context aware recommendations compared to conventional standalone advisory tools.

The findings further demonstrate that multi modal advisory systems significantly enhance the quality and responsiveness of agricultural decision making, particularly for farmers operating in the diverse agro climatic regions of Maharashtra. The developed framework is scalable and flexible, allowing for the incorporation of additional datasets, regional adaptations, and enhanced predictive models in future extensions. Overall, this

research contributes a practical and scientifically grounded approach to digital agriculture by demonstrating the effectiveness of integrating machine learning, deep learning, expert rules, and real time data sources in the development of farmer centric advisory systems that promote sustainable and efficient agricultural practices.

References

1. Shastri, R., Mehta, P., Kulkarni, S., & Deshpande, A. (2025). Machine learning based crop recommendation using soil nutrients and climatic parameters. *Scientific Reports*, 15(1), 10421–10435.
2. Patel, J., Kumar, V., & Singh, P. (2021). Crop recommendation systems using machine learning classifiers. *International Journal of Advanced Computer Applications*, 12(6), 45–52.
3. Nikhil, R., Verma, A., & Sinha, P. (2024). Crop yield prediction using regression and ensemble learning techniques. *Journal of Agriculture and Food Research*, 15, 100721.
4. Shawon, A. R., Hasan, M. M., & Hossain, M. A. (2024). Crop yield prediction using machine learning and deep learning: A systematic literature review. *Smart Agricultural Technology*, 6, 100328.
5. Ferentinos, K. P. (2020). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 170, 105251.
6. Atila, Ü. (2021). Plant leaf disease classification using EfficientNet deep learning models. *Ecological Informatics*, 61, 101182.
7. Li, Y., Zhang, H., Shen, D., & Wang, Z. (2022). Hybrid CNN–LSTM deep learning models for agricultural prediction tasks. *Computers and Electronics in Agriculture*, 198, 107047.
8. Akhter, R., Sofi, S. A., & Bashir, I. (2022). Internet of Things and machine learning based smart agriculture: A systematic review. *Computers and Electronics in Agriculture*, 197, 106902.
9. Bhadra, S., Ghosh, A., & Banerjee, S. (2023). Explainable artificial intelligence in smart agriculture: Applications and challenges. *Artificial Intelligence in Agriculture*, 7, 100112.
10. Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2020). Development of a satellite based crop yield prediction model using vegetation indices. *Remote Sensing of Environment*, 247, 111935.
11. Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2020). A review on the practice of deep learning in agriculture. *Computers and Electronics in Agriculture*, 147, 70–90.
12. Kim, D., Park, S., & Lee, J. (2021). Real-time weather integrated advisory systems for precision agriculture. *Sensors*, 21(5), 1803.
13. Orozco, R., Fernandez, D., & Lopez, J. (2020). Knowledge driven rule based decision support systems for sustainable agriculture. *Expert Systems with Applications*, 159, 113584.
14. Singh, R., Sharma, A., & Kaur, P. (2022). Machine learning based fertilizer recommendation systems for sustainable agriculture. *Journal of Cleaner Production*, 345, 131092.
15. Yadav, L., Singh, A., & Bansal, R. (2023). Integrated smart crop advisory platforms combining machine learning, IoT and real-time APIs. *Journal of Intelligent Systems*, 32(1), 412–428.
16. Verma, S., Mishra, A., & Tripathi, R. (2020). Machine learning based crop recommendation using soil and weather parameters. *International Journal of Agricultural Engineering*, 13(2), 145–152.
17. Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. (2021). Ensemble learning methods for crop yield prediction using climatic data. *Computers and Electronics in Agriculture*, 185, 106176.
18. Selvaraj, M. G., Vergara, A., Ruiz, H., Safari, N., Elayabalan, S., Ocimati, W., & Blomme, G. (2021). AI-powered plant disease detection using transfer learning. *Plant Methods*, 17(1), 1–15.
19. Patil, S. R., Deshmukh, R. B., & Jadhav, P. D. (2022). Rule based fertilizer recommendation system using soil nutrient data. *Journal of Agricultural Informatics*, 13(3), 23–34.
20. Sharma, R., Gupta, A., & Meena, S. (2023). Smart crop advisory system integrating machine learning models and real time weather APIs. *Smart Agricultural Technology*, 4, 100198.