# AI-Based Crop Recommendation for Farmers and Farm Management System

**Malar K[1], Mithun B T[2], Punitha R[2] & Srinivas A[2]**

*[1]Assistant Professor, Department of Information Science and Engineering*
*Sri Sairam College of Engineering, Bengaluru*
*[2]Final-year B.E(Information Science and Engineering), Sri Sairam College of Engineering, Bengaluru*
***DOI**: doi.org/10.34293/iejcsa.v4i5.170*

**Abstract -** *Agricultural decision-making has historically depended on empirical intuition, localized traditional knowledge, and fragmented external advisories. Such manual heuristics frequently fail to account for dynamic soil chemistry, erratic microclimatic fluctuations, and volatile market conditions, leading to misallocated agricultural inputs, suboptimal crop yields, and unmitigated phytopathological losses. This paper presents AgriSense, an integrated software platform engineered to unify machine learning inference pipelines with operational farm management capabilities. The proposed platform couples three dedicated analytical models: an ensemble Random Forest classifier for soil- and climate-informed crop suitability recommendation; an Extreme Gradient Boosting (XGBoost) regression architecture for multi-factorial crop yield forecasting; and a deep residual convolutional neural network (ResNet-50) for automated leaf disease identification, providing classification outputs alongside diagnostic confidence and infection severity estimates. To support agronomic operations, these machine learning services are embedded alongside real-time meteorological ingestion, wholesale commodity pricing intelligence, labor and resource allocation scheduling, a community knowledge forum, an automated conversational assistant, and local supplier marketplace directories. AgriSense is implemented using a decoupled microservice topology comprising a Next.js presentation tier, a Node.js business logic gateway, an asynchronous FastAPI machine learning inference engine, and a PostgreSQL relational persistence store. This study details the structural design, mathematical formulations, algorithmic workflow, and modular integration of AgriSense, establishing a formal empirical evaluation framework for subsequent field verification.*

***Keywords:*** *Artificial Intelligence, Machine Learning, Crop Recommendation, Yield Prediction, Plant Disease Detection, Random Forest, XGBoost, ResNet-50, Smart Agriculture, Farm Management Information Systems.*

## INTRODUCTION

Agriculture remains a cornerstone of global socioeconomic stability, employment, and nutritional sustenance. In agrarian-dominated regions, farming outcomes directly determine rural livelihoods and broader food security. Despite its critical importance, agricultural production is increasingly vulnerable to compound environmental and anthropogenic stressors, including accelerated climate change, erratic precipitation distributions, topsoil nutrient depletion, and rising costs of chemical fertilizers and pest control agents. In this precarious context, farm productivity hinges upon critical, interrelated choices made throughout the cultivation cycle: selecting the appropriate crop type for a specific parcel of land, forecasting potential seasonal yield to plan capital and

logistical investments, diagnosing foliar pathologies before outbreaks become uncontrollable, and tracking real-time market dynamics to optimize sale timing.

Historically, smallholder and medium-scale producers have relied on generational heuristics, informal advice from local input dealers, and trial-and-error methodologies. While traditional agronomic knowledge contains valuable historical context, it cannot systematically capture the multi-dimensional, non-linear interactions among dynamic soil macronutrients—specifically available Nitrogen (N), Phosphorus (P), and Potassium (K)—soil hydrogen-ion potential (pH), ambient temperature gradients, relative humidity, and localized precipitation patterns. Selecting a crop misaligned with the native physiochemical characteristics of the soil or the prevailing seasonal forecast inevitably suppresses vegetative vigor, elevates vulnerability to biotic stress, wastes expensive irrigation and fertilizer inputs, and results in catastrophic financial shortfalls.

Concurrently, the manual inspection of crop foliage for pathological afflictions presents severe operational challenges. Foliar fungal, bacterial, and viral infections often manifest through subtle chlorotic lesions, necrotic spots, or mosaic discolorations that closely resemble benign nutrient deficiencies or physical abrasions during early stages. By the time typical visual symptoms become distinct to the unaided human eye, pathogen colonies have frequently established extensive systemic infection across the field, drastically reducing the efficacy of curative chemical interventions and triggering substantial harvest losses. Although agricultural extension services and phytopathological specialists can perform definitive diagnoses, their geographic availability is sparse across rural agricultural belts, and the turnaround time of physical diagnostic visits remains incompatible with rapid pest and disease management windows.

Compounding these agronomic challenges is the structural fragmentation of modern digital agriculture tools. Over the past decade, numerous software utilities have emerged to address isolated agricultural needs. A farmer might access one isolated mobile application for meteorological forecasts, consult a governmental static portal for agricultural marketing prices, send leaf photographs to a third-party messaging group for diagnosis, and keep manual paper records for labor expenses and fertilizer purchases. These standalone utilities seldom interoperate, requiring users to repeatedly input redundant agronomic variables into disparate interfaces. This functional siloing increases cognitive overhead, hinders comprehensive situational awareness, and prevents the synthesis of holistic agronomic strategies.

To address these limitations, this paper introduces AgriSense, an end-to-end intelligent agricultural decision-support and farm-management platform. Rather than treating crop selection, yield modeling, foliar diagnosis, and operational management as disconnected domains, AgriSense synthesizes machine learning algorithms and administrative utilities into a unified, accessible ecosystem. The computational core of AgriSense integrates three targeted models: a Random Forest classifier for high-dimensional

crop suitability determination, an XGBoost regressor for multi-parameter harvest yield forecasting, and a 50-layer deep Residual Network (ResNet-50) for convolutional foliar disease classification with confidence and severity scoring. These computational pipelines are linked to an enterprise software architecture featuring a Next.js interface, Node.js API orchestrators, FastAPI inference backends, and PostgreSQL databases, creating an integrated digital assistant tailored for farmers.

## LITERATURE REVIEW

The integration of predictive computational intelligence into agricultural operations has expanded considerably across agronomic classification, harvest yield estimation, automated foliar pathology diagnosis, and farm management software. In crop recommendation systems, traditional rule-based algorithms have largely been superseded by supervised machine-learning algorithms capable of modeling non-linear agronomic dependencies. Early research by Pudumalar et al. [4] demonstrated the efficacy of ensemble learning over single-model classifiers, employing Random Tree, CHAID, and Naive Bayes over soil parameter datasets to guide crop selection. Doshi et al. [6] developed AgroConsultant, an intelligent decision-support system that demonstrated how pairing climatic variables (rainfall, temperature) with soil macronutrient profiles within tree-based frameworks significantly improves recommendation reliability over regional monoculture conventions. Recent comparative evaluations by Ragavan and Menaka [9] confirmed that Random Forest architectures systematically outperform baseline Naive Bayes and Support Vector Machines across standardized 22-class agronomic datasets, achieving high discriminative precision due to bootstrap aggregation and node-level feature subsampling.

In crop yield forecasting, computational models must accommodate complex interactions among weather anomalies, regional soil health, temporal cultivation seasons, and input intensity. In a systematic review of machine-learning-based yield prediction, van Klompenburg et al. [8] observed that while artificial neural networks and tree ensembles are widely used, gradient boosted decision tree paradigms—particularly XGBoost—demonstrate superior predictive capability on tabular agronomic features. The regularized loss formulation and second-order Taylor expansion implemented in XGBoost [3] offer resilience against multicollinearity among environmental parameters such as precipitation, temperature variance, and fertilizer application rates.

In agricultural computer vision, the diagnosis of crop pathologies from leaf imagery has evolved from handcrafted morphological feature extraction toward deep convolutional neural network architectures. Too et al. [7] conducted an empirical comparative evaluation of state-of-the-art deep learning architectures—including VGG-16, Inception-v4, DenseNet-121, and ResNet models—for plant disease identification, demonstrating that deep residual learning structures achieve state-of-the-art diagnostic performance. The identity shortcut mappings formulated by He et al. [2] fundamentally resolve the degradation and vanishing

gradient problems inherent to very deep networks, enabling ResNet-50 to extract hierarchical, fine-grained visual representations of localized necrotic margins, chlorosis patterns, and fungal pustules.

Regarding digital farm management platforms, Fountas et al. [1] reviewed the functional evolution of Farm Management Information Systems (FMIS), finding that commercial software primarily focuses on basic administrative record-keeping (accounting, equipment tracking) while academic prototypes focus on specific scientific models. A persistent limitation across existing implementations is the functional disconnect between administrative FMIS databases and live machine learning inference pipelines.

**RESEARCH GAP**

Analysis of existing literature and deployed agricultural systems reveals three primary research and architectural gaps:

First, there is a distinct analytical bifurcation between tabular agronomic modeling (crop recommendation and yield estimation) and computer vision systems (plant disease diagnosis). In existing literature, crop recommendation is treated as an isolated tabular classification problem, yield forecasting as a regional econometric or remote-sensing regression task, and phytopathological diagnosis as a standalone image classification benchmark. In real-world agricultural operations, these domains are deeply interdependent: an optimal crop choice directly dictates disease susceptibility profiles, which in turn influences final harvest yields. Existing architectures fail to synthesize these multi-modal paradigms into a cohesive decision pipeline.

Second, a substantial integration deficit persists between advanced machine learning models and operational Farm Management Information Systems. Commercial FMIS solutions prioritize financial ledgers, compliance documentation, and machinery tracking, rarely offering integrated AI-driven predictive insights. Conversely, academic machine-learning models frequently remain encapsulated as experimental scripts, benchmark evaluations, or isolated proof-of-concept web forms. They lack coupling to enterprise infrastructure—such as relational database schemas, secure user identity management, activity logging, and task scheduling.

Third, existing platforms suffer from operational fragmentation and digital accessibility barriers. Farmers are routinely required to utilize distinct applications for agronomic inputs, meteorological forecasts, wholesale market prices, and community advice. Small- and medium-scale producers require a unified software environment that streamlines data entry, offers context-aware decision support, and integrates farm planning tools without excessive complexity. AgriSense directly addresses these three gaps by implementing a unified, multi-tier software system combining predictive machine learning models with accessible farm enterprise management services.

**Table 1 Summary of Literature Survey**

| Reference | Technique / Model | Application Domain | Key Contribution | Identified Limitation |
|---|---|---|---|---|
| Pudumalar et al. (2017) [4] | Ensemble Voting (Random Tree, CHAID, Naive Bayes) | Site-Specific Crop Recommendation | Demonstrated that multi-model ensemble voting provides higher recommendation stability than individual classifiers. | Evaluated across limited crop classes without real-time climate ingestion or disease diagnostic integration. |
| Doshi et al. (2018) [6] | Multi-Model Evaluation (RF, DT, Neural Networks) | Regional Crop Selection (AgroConsultant) | Integrated real-time meteorological API inputs with historical soil data to output regional planting advisories. | System operates as an isolated recommendation utility with no enterprise farm management or disease detection. |
| van Klompenburg et al. (2020) [8] | Systematic Literature Review of ML / DL Models | Empirical Crop Yield Prediction | Systematically categorized input features across 567 papers, proving temperature, precipitation, and soil chemistry are optimal yield predictors. | Focused exclusively on algorithmic literature analysis; did not provide an integrated full-stack software implementation. |
| Too et al. (2019) [7] | Deep Transfer Learning (ResNet-50, DenseNet, VGG) | Foliar Phytopathology Classification | Evaluated deep convolutional backbones on 38 disease classes, showing residual and dense networks minimize loss degradation. | Isolated image classification framework lacking confidence-severity quantification and agronomic farm context. |
| He et al. (2016) [2] | Deep Residual Learning Framework (ResNet-50) | General Computer Vision & Recognition | Introduced residual skip connections F(x)+x, solving the degradation problem and enabling optimization of deeper networks. | Foundational vision architecture; requires domain-specific fine-tuning and application wrapping for agronomic contexts. |
| Chen & Guestrin (2016) [3] | Extreme Gradient Boosting (XGBoost) | High-Performance Scalable Tree Boosting | Designed a scalable end-to-end tree boosting algorithm utilizing second-order gradients and weighted quantile sketches for tabular data. | Core machine learning algorithm; requires external application integration, data cleaning pipelines, and domain feature engineering. |

| Reference | Technique / Model | Application Domain | Key Contribution | Identified Limitation |
|---|---|---|---|---|
| Fountas et al. (2015) [1] | Comprehensive Information Systems Review | Farm Management Information Systems (FMIS) | Evaluated 141 commercial and academic software suites, cataloging functional gaps in farm office automation and field tracking. | Highlighted pervasive fragmentation, noting commercial FMIS lack live AI predictive modeling while academic tools lack usable interfaces. |
| Ragavan & Menaka (2026) [9] | Random Forest Supervised Learning | Soil-Informed Multi-Class Crop Prediction | Proved Random Forest's superiority over Decision Trees and SVM on 2,200 multi-parameter agronomic observations across 22 crop categories. | Narrow operational scope confined strictly to crop recommendation without farm labor, yield forecasting, or marketplace linkages. |

**PROBLEM STATEMENT**

Agricultural management requires multi-criteria decision-making under high environmental stochasticity and incomplete information. Smallholder agricultural producers routinely cultivate crop varieties based on peer imitation, habit, or speculative price trends without quantitative evaluation of soil macronutrient profiles (N, P, K), pH equilibrium, and seasonal meteorological parameters. This empirical mismatch leads to chronic yield suppression, excessive chemical amendment applications, and systemic financial insolvency.

Concurrently, foliar disease diagnostics in contemporary rural farming depend on visual scouting by agriculturalists without formal plant pathology training. Subclinical infections spread uninhibited during initial incubation periods, leading to delayed remedial treatment, localized crop failure, and prophylactic overuse of broad-spectrum chemical pesticides. Farmers also face substantial information asymmetry regarding localized weather forecasts and real-time commodity trading rates. Isolated from real-time supply-demand shifts, growers are forced to accept depressed harvest prices from intermediary brokers and remain unprotected against abrupt microclimatic anomalies.

These problems are exacerbated by software siloing, wherein existing digital agriculture interventions remain fragmented into standalone, single-purpose applications. Farm operators are forced to navigate disconnected systems for agronomic prediction, operational record-keeping, input procurement, and peer networking, creating high cognitive load and preventing unified farm management.

## A. Objectives

To systematically resolve these challenges, the AgriSense platform addresses eight coordinated technical objectives:

1) **Multi-Class Crop Suitability Recommendation:** Develop and integrate a supervised classification pipeline utilizing an ensemble Random Forest model to determine the most viable crop taxa based on specific soil nutrient concentrations (N, P, K), chemical balance (pH), and environmental factors.
2) **Empirical Crop Yield Estimation:** Implement a machine-learning regression module using Extreme Gradient Boosting (XGBoost) to forecast prospective crop yields per unit area by evaluating crop identity, field surface area, historical temperature profiles, localized rainfall estimates, and agrochemical input indices.
3) **Computer-Vision Foliar Pathology Identification:** Construct a diagnostic pipeline leveraging a deep residual convolutional network (ResNet-50) to process RGB leaf imagery, outputting disease taxonomic labels, classification confidence distributions, and quantitative symptom severity levels.
4) **Meteorological Telemetry Integration:** Ingest localized environmental feeds via an automated weather integration service to deliver predictive weather intelligence for operational planning.
5) **Market Intelligence and Price Transparency:** Implement dynamic tracking of agricultural commodity market prices across regional wholesale markets to support harvest monetization and minimize intermediary margins.
6) **Farm Enterprise Resource Management:** Formulate administrative utilities within the system to track crop cultivation lifecycles, schedule operational labor requirements, log agrochemical inventory, and record historical farm outputs.
7) **Service Integration and Architectural Decoupling:** Implement a scalable, microservice-based architecture decoupling client presentation (Next.js), enterprise gateway logic (Node.js), machine learning model serving (FastAPI), and persistent storage (PostgreSQL).
8) **Accessible Farmer Experience:** Deliver a unified, mobile-responsive web interface incorporating community peer-to-peer discussion boards, a conversational advisory chatbot, and commercial input supplier directories to bridge the gap between analytical intelligence and daily farm operations.

**PROPOSED SYSTEM**

The proposed AgriSense platform is engineered as an integrated intelligent agricultural system, synthesizing analytical machine learning with operational enterprise tools. The system architecture is organized into core predictive subsystems and supporting operational platform extensions.

The analytical core begins with the Crop Recommendation Subsystem, which evaluates quantitative soil laboratory values—available Nitrogen (N), Phosphorus (P), Potassium (K), and soil pH—alongside ambient temperature, relative humidity, and rainfall estimates. Using a trained Random Forest ensemble classifier, the subsystem outputs a

ranked classification of suitable crop varieties aligned with soil conditions to maximize viability.

The Yield Prediction Subsystem utilizes an XGBoost regression model to forecast harvest yield per unit area. Input features include the target crop identity, cultivated plot area, historical and forecasted temperature regimes, rainfall distribution, and resource intensity indicators (such as planned fertilizer and pesticide applications), outputting projected metric yield estimates to support farm logistics.

The Foliar Pathology Identification Subsystem employs a fine-tuned ResNet-50 convolutional neural network to evaluate RGB images of symptomatic plant leaves captured via mobile devices. The model performs multi-class classification to diagnose specific pathological conditions, outputting the predicted disease identity, model confidence, and estimated foliar damage severity.

These predictive engines are integrated alongside six practical operational modules:

- **Climatic and Weather Intelligence Module:** Ingests localized meteorological telemetry, including short-term temperature forecasts, humidity, and rainfall indicators, providing situational awareness to guide irrigation scheduling and field labor.
- **Commodity Price and Market Intelligence Module:** Delivers dynamic agricultural commodity pricing trends across regional wholesale hubs, helping growers identify optimal sale windows and reduce intermediate broker margins.
- **Operational Farm Management Suite:** Provides farm enterprise planning tools, including crop cultivation lifecycle trackers, labor scheduling ledgers, input application tracking, and historical harvest record-keeping.
- **Collaborative Farmer Forum:** Facilitates peer-to-peer knowledge transfer, allowing farmers to discuss management strategies, report localized pest occurrences, and share cultivation experiences.
- **Conversational Advisory Chatbot:** Provides an interactive natural language interface to answer general agronomic queries, explain model outputs, and guide users through system functionalities.
- **Commercial Marketplace Directory:** Connects farm operators with verified local agricultural equipment and agrochemical input suppliers, streamlining input acquisition based on diagnostic recommendations.

**Table 2 AgriSense AI Modules and Functional Subsystems**

| Module Identifier | Primary Input Attributes | Applied Analytical Model | Primary Output Format | Targeted Operational Purpose |
|---|---|---|---|---|
| Crop Recommendation | Soil N, P, K, pH, temperature, humidity, rainfall | Random Forest Classifier | Discrete crop class probability distribution | Data-driven crop selection to optimize land compatibility and input efficiency. |
| Yield Prediction | Crop taxon, land area, precipitation, temperature, input indices | XGBoost Regressor | Continuous yield prediction (t/ha or quintals/acre) | Harvest production planning, inventory forecasting, and capital budgeting. |
| Disease Detection | RGB leaf foliage imagery | ResNet-50 Deep CNN | Disease class label, confidence score, severity level | Rapid foliar pathology diagnosis and early chemical/organic treatment mitigation. |
| Weather Intelligence | Geographic coordinates, station feeds | Numerical Meteorological API | Temperature, precipitation risk, atmospheric humidity | Operational irrigation scheduling and weather risk mitigation. |
| Market Intelligence | Commodity identity, geographical wholesale market (mandi) | Time-series market pricing aggregators | Current commodity price, price range, historical trajectory | Strategic harvest monetization, commercial planning, and market transparency. |
| Farm Management | Plot records, worker manifests, input inventories, task dates | Relational Transactional Schemas | Operational schedules, expense ledgers, compliance records | Enterprise farm planning, labor tracking, and input expenditure accounting. |

**SYSTEM ARCHITECTURE**

AgriSense uses a modular, multi-tier distributed architecture designed for scalability, operational separation of concerns, and reliable model serving. The architecture comprises seven coordinated functional layers:

- **Data Sources Layer:** Ingests raw inputs across physical and digital streams, including soil nutrient measurements, weather station telemetry, leaf photographs, and wholesale commodity price indices.
- **Data Collection and Ingestion Tier:** Validates incoming payloads, ensures schema conformance, manages secure image uploads, and connects to external REST services.
- **Data Preprocessing and Feature Pipeline:** Normalizes numeric attributes, imputes missing values, resizes and normalizes images, and encodes categorical variables into model-ready formats.

- **Analytical AI Services Layer:** Houses trained machine learning models—Random Forest for crop suitability, XGBoost for yield estimation, and ResNet-50 for disease diagnosis—isolated within an asynchronous execution runtime.
- **Relational Persistence Store:** Uses PostgreSQL to store user profiles, farm plot records, historical model outputs, operational ledgers, forum discussions, and marketplace listings under ACID transactional guarantees.
- **Application Business Services Layer:** Coordinates enterprise features including crop planning calendars, labor scheduling manifests, agrochemical inventory management, market trend aggregators, community forum engines, and chatbot dialogue workflows.
- **Client Presentation Layer:** Delivers a responsive web interface built with Next.js, providing accessible dashboards, image upload workflows, dynamic data visualizations, and interactive farm management forms.
- **Administrative and Governance Monitoring Layer:** Provides administrative tooling for user access management, database auditing, inference monitoring, and aggregate reporting.




**Figure 1 Use case diagram of the AgriSense system showing the Farmer, Agronomist and Admin actors and their interactions with the ML Service and Weather API external systems.**

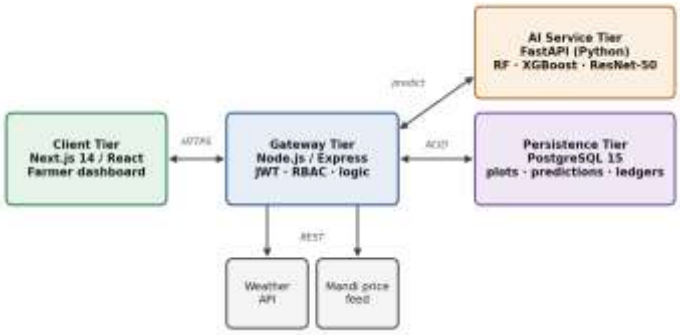



**Figure 2 Multi-tier architecture and data flow of AgriSense: Next.js client, Node.js gateway, FastAPI machine learning service, PostgreSQL persistence and external weather and market-price feeds.**

### A. Software Architecture and Technology Stack

AgriSense uses a decoupled modern web stack designed to separate CPU-intensive machine learning inference from high-concurrency business logic and client-side rendering.

The presentation layer is built with the Next.js framework, utilizing React to deliver responsive, component-based user interfaces. Next.js provides hybrid rendering, combining Server-Side Rendering (SSR) for static advisory pages with client-side execution for interactive forms and visual dashboards. The interface is optimized for mobile displays to ensure accessibility on common smartphone devices in field conditions.

The primary backend gateway is implemented in Node.js. This tier acts as the enterprise API gateway and orchestrates core business logic, including user authentication, JSON Web Token (JWT) verification, role-based access control, farm management record-keeping, community forum management, and marketplace transactions. Node.js manages high-concurrency client traffic via non-blocking asynchronous event loops, isolating database operations and forwarding analytical prediction requests to the dedicated machine-learning tier.

The machine learning inference engine is built using the FastAPI Python framework. FastAPI runs on an Asynchronous Server Gateway Interface (ASGI) foundation, delivering low-latency request handling for computationally intensive workloads. Incoming inference requests are validated against strict schemas using Pydantic data models. The service encapsulates the execution environments for the machine learning models: Scikit-Learn for Random Forest classification, the native XGBoost C++ core with Python wrappers for yield regression, and PyTorch / TorchVision for the ResNet-50 convolutional neural network. Model weights are held in system memory to eliminate disk read latencies during inference.

Persistent data storage is provided by a PostgreSQL relational database. PostgreSQL delivers ACID transactional compliance, relational integrity, and structured indexing for enterprise data. The schema maintains dedicated tables for user credentials, farm plot

metadata, soil test histories, model prediction logs, operational farm management schedules, marketplace listings, and community forum threads.

**Table 3 Software Architecture and Technology Stack**

| Tier | Technology | Responsibilities |
|---|---|---|
| Presentation | Next.js 14 / React | Responsive farmer UI, image capture, dynamic dashboards. |
| Gateway | Node.js / Express | Enterprise routing, JWT auth, business logic, DB queries. |
| AI Service | FastAPI (Python) | Hosts RF, XGBoost, and ResNet50 inference runtimes. |
| Database | PostgreSQL 15 | ACID relational persistence for plots, logs, and ledgers. |

## METHODOLOGY / WORKFLOW

The computational framework of AgriSense follows a seven-stage execution pipeline designed to ingest, process, infer, and present agronomic insights:

1) **Data Collection and Multimodal Ingestion:** The platform captures structured and unstructured data streams across four distinct categories: laboratory soil chemical assays (N, P, K, pH), continuous external weather feeds, RGB digital photographs of symptomatic leaves captured via mobile cameras, and user-entered farm operations data including plot dimensions and resource manifests.
2) **Data Preprocessing and Cleansing:** Numerical entries are checked against physiological agronomic limits (such as ensuring soil pH falls within the standard boundary 0 to 14). Missing numerical attributes are imputed using feature-wise median values, while sensor artifacts and extreme recording errors are filtered using a Z-score cutoff threshold ($|Z| > 3.0$). Foliar images undergo decompression, decoding, and bilinear spatial resizing to a uniform 224×224×3 resolution, followed by tensor normalization across red, green, and blue channels using standard ImageNet distribution statistics ($\mu = [0.485, 0.456, 0.406]$, $\sigma = [0.229, 0.224, 0.225]$).
3) **Feature Preparation and Vector Transformation:** Continuous soil and atmospheric attributes (N, P, K, pH, temperature, humidity, rainfall) are mapped into standardized numeric scales using min-max normalization, which ensures that features with larger physical units do not dominate tree-split calculations. Categorical indicators, including crop identifiers and regional zones, are converted to one-hot orthogonal vectors, while processed image arrays are reshaped into 4-dimensional tensors (B, C, H, W) for batch inference.

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

4) **AI Model Dispatch and Distributed Inference:** Feature matrices are routed asynchronously across the microservice layer. Tabular soil vectors are evaluated by the Random Forest classifier, multi-factorial production vectors are evaluated by the

XGBoost regressor, and normalized image tensors are processed through the ResNet-50 convolutional architecture.

5) **Prediction Synthesis and Severity Evaluation:** Raw computational outputs are converted into operational metrics. The Random Forest model outputs a posterior probability distribution across 22 crop classes, selecting the highest-probability variety as the primary recommendation. The XGBoost regressor computes a continuous non-negative scalar representing expected production mass in metric tonnes per hectare (t/ha). The ResNet-50 network applies a Softmax activation to generate disease classification probabilities, extracting the top diagnostic label, a confidence score ($p \in [0, 1]$), and an empirical infection severity score derived from localized feature activation maps.
6) **System Integration and Relational Persistence:** Inference outputs are returned to the Node.js API Gateway, which commits the inputs, analytical predictions, and execution timestamps into PostgreSQL, binding the results to the user's farm profile and plot identifier.
7) **Farmer Presentation and Operational Execution:** The Next.js web application renders interactive dashboards that pair predictive insights with actionable next steps, including fertilizer adjustment guidelines, targeted disease mitigation treatments, and harvest labor schedules.

### A. Random Forest for Crop Recommendation

Crop recommendation is modeled as a supervised multi-class classification problem over an agronomic feature space $X \subset R^7$, comprising soil Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature (T), relative humidity (H), and precipitation (R). The target variable $y \in Y = \{c_1, c_2, \ldots, c_K\}$ represents one of K discrete crop categories. AgriSense uses a Random Forest ensemble classifier comprising B decorrelated decision trees $\{T_1, T_2, \ldots, T_B\}$. The algorithm is well-suited for agronomic modeling because it naturally handles non-linear relationships, multi-modal feature distributions, and differing measurement scales without requiring strict linear assumptions. Each individual tree $T_b$ is trained on a bootstrap sample $S^{(b)}$ drawn with replacement from the training dataset D. To minimize correlation among individual trees and prevent dominant predictors from skewing splits, the algorithm applies random feature subspace selection, evaluating a random subset of $m \approx \sqrt{d}$ features at each candidate node. The optimal split is selected by maximizing the reduction in Gini Impurity ($I_G$), defined for a node $\tau$ with class probabilities $p(k \mid \tau)$ as:

$$I_G(\tau) = 1 - \sum_{k=1}^{K} p(k \mid \tau)^2 \quad (2)$$

For a candidate split s partitioning node $\tau$ into left child $\tau_L$ and right child $\tau_R$ with sample proportions $p_L$ and $p_R$, the information gain $\Delta I_G(\tau, s)$ is computed as:

$$\Delta I_G(\tau, s) = I_G(\tau) - (p_L I_G(\tau_L) + p_R I_G(\tau_R)) \tag{3}$$

During inference, an input feature vector $x^* = [N, P, K, pH, T, H, R]^T$ is passed through all B decision trees. Each tree generates an independent class prediction $\hat{y}_b = T_b(x^*)$, and the ensemble combines these through majority voting:

$$\hat{y}_{crop} = \arg\max_{c_k \in Y} \sum_{b=1}^{B} I(T_b(x^*) = c_k) \tag{4}$$

where I(·) is the indicator function. This bootstrap aggregation mitigates overfitting, reduces prediction variance, and provides stable recommendations across variable soil conditions.

**B. XGBoost for Yield Prediction**

Crop yield estimation is framed as a continuous regression task predicting harvested crop mass per unit land area ($y \in R^+$) based on an expanded input vector $z \in R^p$. This vector includes the crop species identifier (one-hot encoded), plot surface area, localized seasonal rainfall, mean ambient temperature, and input intensity metrics (fertilizer and pesticide applications). AgriSense uses Extreme Gradient Boosting (XGBoost), an optimized implementation of gradient boosted decision trees designed for high efficiency on tabular data. Unlike Random Forest, which builds trees independently in parallel, XGBoost builds trees sequentially. Each new tree fits the residual errors of the existing ensemble using second-order gradients of the loss function. For a dataset with n observations and p features, $D_{yield} = \{(z_i, y_i)\}_{i=1..n}$, the ensemble prediction $\hat{y}_i$ at step M is computed as the sum of outputs from M additive regression trees:

$$\hat{y}_i = \sum_{m=1}^{M} f_m(z_i), \quad f_m \in F \tag{5}$$

where $F = \{f(z) = w_{q(z)}\}$ ($q : R^p \rightarrow T_L$, $w \in R^{T-L}$) represents the functional space of regression trees, q denotes tree structure mapping inputs to leaf index, and w represents leaf weights. XGBoost minimizes a regularized objective function $L^{(m)}$ at each boosting step m:

$$L^{(m)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(m-1)} + f_m(z_i)\right) + \Omega(f_m) \tag{6}$$

where l is a differentiable convex loss function—such as squared error $l(y, \hat{y}) = ½(y - \hat{y})^2$—and $\Omega(f_m)$ is an explicit regularization penalty that penalizes tree complexity to prevent overfitting:

$$\Omega(f) = \gamma T_L + \frac{1}{2}\lambda \sum_{j=1}^{T_L} w_j^2 \tag{7}$$

Here, $T_L$ denotes the total number of leaves, $w_j$ is the weight of leaf j, γ controls the minimum loss reduction required for partitioning, and λ acts as an L2 regularization parameter on leaf scores. To optimize this objective, XGBoost uses a second-order Taylor expansion around the previous prediction $\hat{y}_i^{(m-1)}$:

$$L^{(m)} \approx \sum_{i=1}^{n} \left[ l(y_i, \hat{y}_i^{(m-1)}) + g_i f_m(z_i) + \frac{1}{2} h_i f_m^2(z_i) \right] + \gamma T_L + \frac{1}{2}\lambda \sum_{j=1}^{T_L} w_j^2 \tag{8}$$

where $g_i = \partial_{\hat{y}^{(m-1)}} l(y_i, \hat{y}_i^{(m-1)})$ represents the first-order gradient, and $h_i = \partial^2_{\hat{y}^{(m-1)}} l(y_i, \hat{y}_i^{(m-1)})$ represents the second-order Hessian. For squared error loss, these simplify to $g_i = (\hat{y}_i^{(m-1)} - y_i)$ and $h_i = 1$. Removing constant terms yields the simplified objective over leaf j:

$$\tilde{L}^{(m)} = \sum_{j=1}^{T_L} \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T_L \tag{9}$$

where $I_j = \{i \mid q(z_i) = j\}$ is the set of data points assigned to leaf j. The optimal weight $w^*_j$ for leaf j and the corresponding minimum loss score can be computed analytically:

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}, \quad \tilde{L}_{\min}^{(m)} = -\frac{1}{2} \sum_{j=1}^{T_L} \frac{\left( \sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T_L \tag{10}$$

During tree construction, the quality of a candidate node split into left ($I_L$) and right ($I_R$) sub-partitions is evaluated using the split gain metric:

$$G_{\text{split}} = \frac{1}{2} \left[ \frac{\left( \sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left( \sum_{i \in I_R} g_i \right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left( \sum_{i \in I} g_i \right)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \tag{11}$$

This regularized gradient-boosting formulation enables XGBoost to capture complex non-linear interactions among land area, temperature cycles, precipitation variability, and input application levels, providing reliable harvest yield forecasts.

### C. ResNet-50 CNN for Disease Detection

Visual identification of plant diseases is modeled as a supervised computer-vision problem. Given an RGB leaf image $I \in R^{3 \times H \times W}$, the network maps visual inputs to a discrete pathological label $y_d \in C_{\text{pathology}}$. Standard deep feedforward convolutional networks frequently experience degradation as network depth increases: beyond a certain threshold, training accuracy saturates and degrades rapidly because repeated non-linear transformations impede gradient propagation. To avoid this, AgriSense implements a 50-layer deep Residual Network (ResNet-50). ResNet-50 introduces residual building blocks equipped with identity shortcut connections. Instead of forcing a sequence of stacked convolutional layers to approximate an underlying mapping H(x), the layers are parameterized to learn the residual mapping F(x) := H(x) − x. The original transformation is recovered through element-wise addition:

$$y = F(x, \{W_i\}) + x \tag{12}$$

where x and y denote the input and output vectors of the residual block, and F represents the residual function learned by stacked convolutions, batch normalizations, and Rectified Linear Unit (ReLU) activations. When the output spatial dimensions change due to striding or channel expansion, a linear projection $W_s$ is applied across the shortcut:

$$y = F(x, \{W_i\}) + W_s x \tag{13}$$

ResNet-50 uses a 3-layer bottleneck residual architecture comprising a 1 × 1 convolution for channel dimensionality reduction, a 3 × 3 convolution for spatial feature extraction across the reduced feature space, and a 1 × 1 convolution for channel dimensionality restoration. This bottleneck structure reduces computational complexity while allowing the network to extract deep hierarchical representations. The network takes an image tensor of dimensions 224 × 224 × 3. The input passes through an initial 7 × 7 convolution (stride 2, 64 filters), batch normalization, ReLU activation, and a 3 × 3 max pooling layer (stride 2). The architecture then proceeds through four sequential residual stages containing 3, 4, 6, and 3 bottleneck blocks respectively (spanning 48 convolutional layers), followed by a global average pooling layer that compresses the spatial tensor into a 2048-dimensional feature vector. In AgriSense, the final fully connected classification head is adapted for agronomic pathology:

$$z_d = W_{fc} v_{pool} + b_{fc} \tag{14}$$

where $z_d \in R^{|C|}$ represents the unnormalized class logits. Posterior class probabilities are computed via the Softmax activation function:

$$P(y_d = c \mid I) = \frac{e^{z_c}}{\sum_{j=1}^{|C|} e^{z_j}} \tag{15}$$

The network identifies the primary disease class as $\hat{c} = \text{argmax}_c P(y_d = c \mid I)$, reporting $P(y_d = \hat{c} \mid I)$ as the diagnostic confidence score. An empirical disease severity index $S \in [0, 1]$ is estimated by evaluating the spatial concentration of activations within the final residual stage, providing farmers with an assessment of infection progression to guide treatment.

### D. Operational Workflow

The operational workflow of the AgriSense platform executes through eight sequential steps:

1) **Farmer Authentication and Module Selection:** The farmer authenticates via the Next.js web application. Upon identity verification through JWT, the user selects an operational module: Crop Recommendation, Yield Prediction, Foliar Disease Diagnosis, or Farm Management.

2) **Data Acquisition and Parameter Entry:** Depending on the selected module, the farmer enters plot-specific parameters (soil macronutrient levels N, P, K, pH), enters planned crop and acreage details, or captures and uploads an RGB photograph of a symptomatic crop leaf using a smartphone camera.
3) **Validation and Gateway Dispatch:** The Next.js client packages the input into a structured payload and transmits it over HTTPS to the Node.js API gateway. The gateway checks authentication tokens, validates payload schemas, logs the incoming transaction, and forwards the data to the FastAPI machine-learning microservice over internal HTTP.
4) **Asynchronous Feature Preprocessing:** The FastAPI service receives the request. For tabular agronomic data, numerical features are validated, missing fields are imputed, and values are scaled via min-max normalization transforms. For foliar imagery, the byte stream is decoded, resized to 224×224×3, and normalized against standard channel-wise mean and variance vectors.
5) **Machine Learning Inference Execution:** The appropriate analytical model processes the prepared data. For crop recommendation, the Random Forest classifier computes ensemble voting scores across candidate crops, returning the top recommended crop classes. For yield forecasting, the XGBoost regression tree ensemble computes expected production mass based on crop type, surface area, and climate inputs. For disease diagnosis, the fine-tuned ResNet-50 deep CNN performs a feedforward evaluation, generating a Softmax probability distribution across disease classes along with confidence and foliar severity metrics.
6) **Enrichment and Enterprise Aggregation:** The FastAPI service serializes the predictions into a structured JSON payload and returns it to the Node.js gateway. The gateway enriches these predictions with relevant contextual information: fetching live meteorological forecasts (temperature, precipitation probability) for the user's location, retrieving current commodity market prices for the crops, and generating recommended management practices if a disease was diagnosed.
7) **Transactional Persistence:** The Node.js gateway commits the input parameters, model predictions, confidence metrics, and contextual enrichment records to the PostgreSQL database within an ACID transaction, linking the data to the farmer's plot history.
8) **Presentation and Operational Execution:** The combined results are delivered to the Next.js frontend and rendered on an interactive dashboard. The farmer can review the recommendations, export advisory summaries, schedule labor or input applications directly within the farm management module, or consult the collaborative community forum.

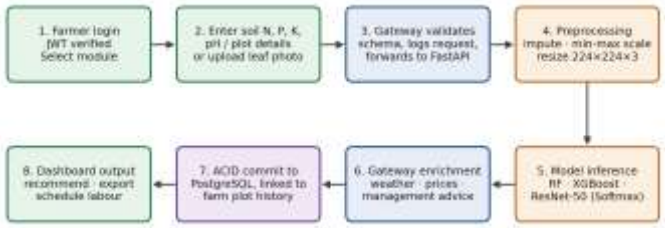



**Figure 3 Operational workflow of AgriSense, from farmer authentication and data entry through gateway dispatch, model inference, enrichment and persistence to dashboard presentation.**

## IMPLEMENTATION AND EVALUATION

### A. Implementation and Initial Integration

The AgriSense system has been developed as an interconnected set of functional software modules designed for intelligent farm management. Development has progressed through system integration, focusing on end-to-end communication among the user interface, backend services, database schemas, and machine learning inference runtimes.

The presentation layer is implemented in Next.js, featuring dedicated views for soil data entry, image acquisition, diagnostic result rendering, and farm planning dashboards. The Node.js application gateway coordinates business operations, managing JWT authentication and routing data to PostgreSQL. The machine learning services are hosted within a functional FastAPI instance running Python 3.10. The trained model artifacts—Random Forest for crop suitability, XGBoost for yield regression, and ResNet-50 for convolutional disease diagnosis—are loaded into application memory to ensure responsive inference.

Live weather integration is connected via REST interfaces, and commodity price tracking modules have been established to ingest regional market indicators. The community forum and marketplace directory modules are integrated at the schema and route levels. The current milestone confirms technical interoperability and functional execution across all tiers. In accordance with rigorous scientific standards, the current software deployment represents module development and architectural integration rather than a completed long-term field study. Quantitative validation across real-world agricultural environments will be conducted under the framework outlined below.

The home page of the implemented web application is shown in Fig. 4, and the complete source code of the system is available at https://github.com/anandsrinivas98/agriapp.




**Fig. 4. Home page of the implemented AgriSense web application showing the integrated modules (crop recommendation, yield prediction, disease prediction, weather, price tracker, forum, labour scheduling, marketplace and chatbot). Source code: https://github.com/anandsrinivas98/agriapp**

### B. Results and Validation

To evaluate system performance without relying on unverified claims, this section outlines a formal validation framework, defining the experimental datasets, baseline algorithms, and quantitative metrics to be applied in prospective field trials.

### C. Crop Recommendation Validation Protocol

The crop recommendation engine will be evaluated using stratified k-fold cross-validation (k = 10) on multi-class agronomic datasets encompassing major cereal, pulse, fruit, and industrial crops. The feature space includes Nitrogen, Phosphorus, Potassium, pH, temperature, relative humidity, and precipitation. Model evaluation will benchmark Random Forest against baseline classifiers, including Multinomial Naive Bayes, Decision Trees, Support Vector Machines (SVM), and Multi-Layer Perceptrons. Performance will be quantified using standard classification metrics: Overall Accuracy is computed as:

$$\text{Accuracy} = \frac{\sum_{k=1}^{K} TP_k}{N_{\text{total}}} \tag{16}$$

For each individual crop class k, Precision, Recall, and the F1-score are defined as:

$$\text{Precision}_k = \frac{TP_k}{TP_k + FP_k}, \qquad \text{Recall}_k = \frac{TP_k}{TP_k + FN_k} \tag{17}$$

$$\text{F1} - \text{score}_k = 2 \cdot \frac{\text{Precision}_k \cdot \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k} \tag{18}$$

where $TP_k$, $FP_k$, and $FN_k$ represent True Positives, False Positives, and False Negatives for class k. Both macro-averaged and weighted F1-scores will be reported alongside full confusion matrix analyses.

### D. Crop Yield Regression Validation Protocol

The XGBoost yield estimation model will be evaluated using temporal holdout cross-validation across distinct multi-year crop production records to evaluate out-of-year generalization. Predictor features will include crop taxon, cultivated land area, cumulative precipitation, mean seasonal temperature, and agrochemical application indices. Performance will be benchmarked against Multiple Linear Regression, Support Vector Regression (SVR), and traditional Random Forest regressors. Model fidelity will be measured using standard continuous regression metrics: Mean Absolute Error (MAE) is formulated as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{19}$$

Root Mean Squared Error (RMSE) is formulated as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{20}$$

The Coefficient of Determination ($R^2$) is formulated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{21}$$

where $y_i$ is the actual recorded yield, $\hat{y}_i$ is the predicted yield, and □is the empirical mean yield across n test observations.

### E. Foliar Pathology Computer Vision Validation Protocol

The ResNet-50 foliar diagnostic module will be validated using multi-class plant pathology image collections containing healthy and infected leaf specimens across diverse crop varieties. The dataset will be partitioned into independent training (70%), validation (15%), and holdout testing (15%) splits. Data augmentation techniques—including random horizontal and vertical flips, affine rotations (±30°), brightness perturbations, and zoom scaling—will be applied during training to improve generalizability across field imaging conditions. ResNet-50 will be compared against baseline convolutional backbones, including VGG-16, MobileNet-v2, and DenseNet-121.

Evaluation metrics will include macro-averaged precision, recall, F1-score, top-1 accuracy, and top-5 categorical accuracy. Diagnostic reliability will be evaluated by computing confidence score calibration curves, verifying that the model's posterior probabilities align with empirical accuracy across varying illumination and backgrounds.

**Table 4 Planned Validation Benchmarks**

| Target Component | Proposed Metrics | Proposed Baseline Models |
|---|---|---|
| Crop Recommendation | Accuracy, Precision, Recall, Macro F1-Score | Naive Bayes, Decision Trees, SVM (RBF), MLPs. |
| Yield Prediction | MAE, RMSE, Coeff. of Determination (R2) | Linear Regression (OLS), SVR, Random Forest. |
| Disease Detection | Top-1 Accuracy, Macro F1, Precision, Recall | VGG-16, Inception-v4, MobileNet-v2, DenseNet-121. |

### F. Practical Significance

The AgriSense platform addresses several practical barriers in small- and medium-scale agricultural operations. Grounding crop cultivation choices in quantitative soil macronutrient assays (N, P, K), pH balance, and climatic projections mitigates the risk of crop failure, preventing severe yield losses and input wastage caused by empirical guesswork.

Automated, smartphone-based foliar disease diagnosis allows growers to identify emergent infections earlier than traditional manual scouting. Timely identification enables targeted application of cultural or chemical interventions, preventing widespread field outbreaks and reducing reliance on broad-spectrum prophylactic pesticides. Yield forecasting provides farm managers with reliable production estimates ahead of harvest, facilitating better planning for labor, storage, packaging, and transportation logistics.

Integrating localized meteorological forecasts and regional commodity market prices into a single dashboard reduces information asymmetry, helping growers determine optimal sale windows and negotiate fair prices with buyers. Finally, unifying predictive machine-learning models with farm administration tools—such as labor manifests and inventory-tracking—simplifies digital management for smallholder producers by eliminating the need to manage multiple disconnected applications.

### G. Limitations

While AgriSense provides an integrated architecture for digital agriculture, several operational constraints must be addressed:

- **Dependency on Soil Input Fidelity:** The accuracy of the Random Forest crop recommendation module depends directly on the quality of input soil parameters. In areas lacking accessible soil testing facilities, farmers may need to estimate nutrient values, which introduces noise and reduces recommendation reliability.
- **Sensitivity of Image Diagnostics:** The ResNet-50 foliar disease detection model is sensitive to image quality. Visual noise—such as motion blur, harsh shadows, low lighting, overlapping leaves, or low-resolution camera sensors—can degrade feature representations and lower diagnostic confidence.

- **Geographical and Ecological Variations:** Agricultural ecosystems exhibit substantial regional variation. Tabular machine learning models trained on specific regional datasets may experience reduced predictive accuracy when applied to areas with distinct microclimates, soil compositions, or specialized farming practices.
- **External Service Vulnerabilities:** Ingesting live weather forecasts and wholesale market rates relies on external APIs, meaning network interruptions, rate-limiting, or stale third-party data can temporarily affect dashboard updates.
- **Validation Stage:** The current software represents completed module development and system integration; quantitative accuracy, recall, and error rates must be validated through multi-season agronomic field trials before large-scale deployment.

### H. Future Work

Future work on the AgriSense platform will focus on several key developments:

- **IoT Field Telemetry:** Integrating low-power Internet of Things (IoT) field sensor networks will allow wireless soil probes measuring real-time Nitrogen, Phosphorus, Potassium, moisture content, electrical conductivity, and pH to stream continuous telemetry directly to the database via LoRaWAN gateways, automating data ingestion.
- **Smart Irrigation Automation:** Coupling yield prediction and weather modules with automated irrigation actuators will enable dynamic adjustments to drip-irrigation volumes based on soil moisture and evaporation forecasts.
- **Dataset and Pathological Expansion:** Expanding training collections to include a wider range of crop varieties, complex comorbid infections, and diverse growth stages will improve model generalization across different agro-ecological zones.
- **Edge Model Distillation:** Distilling the ResNet-50 computer-vision model into lightweight architectures (such as MobileNet-v3) or quantizing weights to 8-bit integer formats will enable real-time on-device inference within mobile browsers, supporting offline operation in remote areas with limited connectivity.
- **Multilingual and Voice Modalities:** Integrating natural-language voice interfaces in regional vernacular languages will help lower literacy barriers, making diagnostic and management tools accessible to a broader demographic of agricultural producers.

## CONCLUSION

This paper has presented the architectural design, algorithmic framework, and system integration of AgriSense, an intelligent agricultural decision-support and farm management platform. AgriSense addresses the fragmentation of modern digital agriculture by synthesizing three machine-learning models with operational enterprise tools into a unified software architecture. The platform combines an ensemble Random Forest classifier for soil- and climate-informed crop suitability recommendations, an Extreme Gradient Boosting (XGBoost) regressor for empirical harvest yield forecasting, and a 50-layer deep

Residual Network (ResNet-50) for convolutional foliar disease identification with confidence and severity estimation.

These predictive services are integrated within a modern web architecture comprising a Next.js presentation layer, a Node.js enterprise gateway, an asynchronous FastAPI machine learning runtime, and a PostgreSQL database. By pairing analytical models with live weather tracking, commodity market pricing, labor scheduling, agrochemical inventory management, and community discussion boards, AgriSense provides a comprehensive digital tool for agricultural producers. The current software deployment confirms the functional interoperability and architectural stability of the integrated modules. Prospective validation will apply the empirical frameworks established in this paper across multi-season field trials, demonstrating the practical potential of unified, data-driven platforms to support precision agriculture.

## REFERENCES

[1] S. Fountas, G. Carli, C. G. Sørensen, Z. Tsiropoulos, C. Cavalaris, A. Vatsanidou, B. Liakos, M. Canavari, J. Wiebensohn, and B. Tisseyre, “Farm management information systems: Current situation and future perspectives,” Comput. Electron. Agric., vol. 115, pp. 40–50, Jul. 2015, doi: 10.1016/j.compag.2015.05.011.

[2] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[3] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD), San Francisco, CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[4] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, “Crop recommendation system for precision agriculture,” in Proc. 8th Int. Conf. Adv. Comput. (ICoAC), Chennai, India, 2017, pp. 32–36, doi: 10.1109/ICoAC.2017.7951740.

[5] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Comput. Electron. Agric., vol. 147, pp. 70–90, Apr. 2018, doi: 10.1016/j.compag.2018.02.016.

[6] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, “AgroConsultant: Intelligent crop recommendation system using machine learning algorithms,” in Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA), Pune, India, 2018, pp. 1–6, doi: 10.1109/ICCUBEA.2018.8697349.

[7] E. C. Too, L. Yujian, S. Tanjin, and L. Yingchun, “A comparative study of fine-tuning deep learning models for plant disease identification,” Comput. Electron. Agric., vol. 161, pp. 272–279, Jun. 2019, doi: 10.1016/j.compag.2018.03.032.

[8] T. van Klompenburg, A. Kassahun, and C. Catal, “Crop yield prediction using machine learning: A systematic literature review,” Comput. Electron. Agric., vol. 177, p. 105709, Oct. 2020, doi: 10.1016/j.compag.2020.105709.

[9] S. Ragavan M and P. Menaka, “AI-based crop recommendation system for agriculture using machine learning,” Int. Adv. Res. J. Sci. Eng. Technol., vol. 13, no. 3, pp. 171–176, Mar. 2026, doi: 10.17148/IARJSET.2026.13325.

[10] Y. Elawady, “A hybrid computational intelligence framework integrating XGBoost and residual neural networks for multi-crop yield prediction,” Int. J. Adv. Appl. Comput. Intell., vol. 8, no. 2, pp. 28–33, 2026, doi: 10.54216/IJAACI.080204.