iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83553

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

AI-Powered Sustainable Farming Assistant: A Conceptual Framework Using Machine Learning

Snehal G. Pawar[1], Snehal A. Pisal[2], Sakshi S. Malpote[3]
PES Modern College of Arts, Commerce and Science (Autonomous), Shivajinagar, Pune – 411005, Maharashtra, India

Abstract—Food production and farm-level income in developing nations like India are under growing strain from shifting climate patterns, soil nutrient exhaustion, and the widespread unavailability of expert-level agronomic guidance at the point of need [1]–[9]. This paper puts forward a conceptual design for an AI-Powered Sustainable Farming Assistant (AISFA), which draws on supervised machine learning to convert raw soil and weather data into specific, actionable guidance for farmers. The parameters include nitrogen, phosphorus, potassium, soil pH, ambient temperature, relative humidity, and measured rainfall. The system generates tailored recommendations for crop selection, water scheduling, fertiliser dosage, and pest management. The architecture is built entirely in software and does not depend on embedded sensors, IoT gateways, or any physical hardware beyond a basic internetconnected device — addressing hardware barriers identified in Growlify[4], AgriVision[7], and Patil et al. [5]. No working software prototype has been constructed; the contribution is a rigorously grounded design framework.
Keywords—Artificial Intelligence; Machine Learning; Sustainable Agriculture; Crop Recommendation; Precision Farming; Decision Support System; Fertilizer Optimization; Pest Detection

I. INTRODUCTION

Agriculture has been at the heart of Indian society for thousands of years, and it still plays a central role in the economy [1],[6]. Nearly half of the country's workforce depends on farming for their livelihood [6]. For smallholder farmers, life is closely tied to the seasons — a late monsoon, a drop in soil quality, or a sudden pest outbreak can cause serious financial damage [1],[5]. Even a modest improvement in decision-making — choosing the right crop, applying the right amount of water, or spotting a disease risk early — can make a meaningful difference to both the farmer's income and their family's well-being [6].

Traditionally, farming decisions have relied on knowledge built up over years of personal experience, passed down through generations [1]. This kind of knowledge cannot adapt quickly enough to keep up with changing environmental conditions such as climate change or soil degradation [1],[6]. Experience-based advice tends to be general rather than specific — being told that a particular crop "grows well in this region" does not tell a farmer whether their specific soil, in its current condition, makes that crop a good choice this season [6]. Personalised, datadriven guidance and generalised advice are very different things, and it is the former that consistently leads to better outcomes [6],[9].

This is where Artificial Intelligence (AI) and Machine Learning (ML) become genuinely valuable [1],[6]. Agriculture involves many variables — soil chemistry, local weather patterns, plant biology, pest behaviour, and market prices — all interacting in complex ways that are difficult to capture with simple rules [6]. ML models are capable of learning these relationships from historical data and applying the resulting insights at the level of individual farms [6],[9]. This level of specificity is something that generalised advisory systems simply cannot provide
[1],[2],[3].

The work presented in this paper proposes a framework called AISFA. Rather than relying on physical sensor hardware — a requirement noted in Growlify[4], Patil et al. [5], and AgriVision[7] — AISFA is designed as a data and software-driven system. Farmers input their soil test results, and the system automatically retrieves current weather data. Machine learning models then generate crop and resource management recommendations. The decision to avoid hardware dependency ensures the system can be deployed in rural environments where the need for agricultural guidance is greatest but access to technology is most limited [4],[5],[7].

The remainder of this paper is structured as follows. Nine closely related studies are reviewed to situate AISFA within the existing body of research. The specific problem the system addresses is then defined, followed by the objectives, system architecture, methodology, expected outcomes, and directions for future work. No software implementation has been completed as part of this project — the contribution is therefore architectural and conceptual in nature, rather than empirical.

II. LITERATURE REVIEW

A. Tapkir, Vathare and Suryawanshi (IJNRD, 2023) [1]

Tapkir et al. [1] surveyed how AI is being adopted in crop production, covering the theoretical use of machine learning classifiers, drone-based monitoring, robotic weeding, and image-based yield estimation [1]. The authors argued that AI-assisted farming can significantly reduce the need for labour-intensive monitoring while improving the reliability of yield predictions [1]. However, the study remained largely theoretical — no actual system was built and no experimental results were reported [1]. The technologies discussed, such as drones and robots, are far too expensive for most small-scale farmers in India [1]. Despite these limitations, the paper is valuable because it maps out the range of possibilities in AI-based farming and makes a strong case for why simpler, software-based tools deserve equal attention alongside high-tech solutions [1].

B. Jagtap, Suryawanshi, Kakade and Bhatkute (IJARCCE, 2025) [2]

Jagtap et al. [2] built a working chatbot-based assistant that pulls live data from public weather APIs and agricultural commodity exchanges, then uses a combination of rule-based logic and machine learning to answer farmer queries about crop health, weather forecasts, and local market prices [2]. The system was built using a Python Flask server with MongoDB for data storage and was accessible through a web browser [2]. A key takeaway from this work is that a chat-style interface is more approachable for farmers who find structured input forms difficult to use [2]. The main weakness is that the system requires a constant internet connection — it cannot work in offline or low-bandwidth conditions, which is a reality for a large portion of rural India [2]. This limitation was one of the primary motivations behind AISFA's design goal of reducing dependence on continuous connectivity for its core recommendations [2].

C. Kaviya, Bhavyashree, Deepak Krishnan and Sugacini (IJRESM, 2021) [3]

Kaviya et al. [3] developed a chatbot using Naive Bayes and K-Nearest Neighbour algorithms, combined with natural language processing and speech-to-text features, to build a question-answering tool for farmers [3]. The inclusion of voice interaction was a notably forward-thinking feature for the time, and the system showed that relatively simple ML techniques can still support useful conversational guidance [3]. However, the system's knowledge base was entirely static — once deployed, it could not incorporate new data [3]. Its accuracy was permanently limited by the scope and quality of the original training dataset, with no ability to respond to new pest threats, unusual weather patterns, or changing market conditions [3]. This limitation directly inspired the dynamic data pipeline proposed as part of AISFA's architecture [3].

D. Kasturi and Navis – Growlify (IJIRT, 2025) [4]

Kasturi and Navis [4] built Growlify for urban and semi-urban food growers looking to optimise limited growing space in controlled or semi-controlled environments [4]. It combines IoT sensors that measure soil moisture and temperature, a machine learning layer for managing resource schedules, and an image processing module for assessing plant health [4]. However, the design assumptions behind Growlify do not translate well to rural smallholder farming [4]. The sensor hardware it relies on is simply not practical for a farmer working one or two hectares of rain-fed land [4]. Growlify therefore serves as a useful illustration of both what an integrated smart farming system can achieve and the practical limits imposed by hardware dependency [4].

E. Patil, Sharma, Deshmukh and Kulkarni – AI-Based Crop Recommendation and Smart Irrigation System (IJIRCST, 2024) [5]

Patil et al. [5] aimed to develop an AI-based system that handles two important farming tasks together — recommending suitable crops and managing irrigation smartly — using real-time environmental data [5]. The system was built using machine learning algorithms combined with IoT sensors for soil moisture monitoring, integrated weather data analysis, and a backend developed in Python and Flask [5]. By bringing crop recommendation and irrigation management into a single platform, the system was able to provide advice based on actual soil and weather conditions, reduce water wastage, and improve overall farm productivity [5].

However, the system has several limitations that affect its suitability for broad rural deployment in India [5]. It requires IoT devices and ongoing sensor maintenance, which adds both financial cost and technical overhead that most smallholder farmers cannot easily manage [5]. The system also depends heavily on accurate, continuously available real-time data, and its usefulness drops noticeably in areas with poor internet connectivity [5].

For AISFA, this paper is relevant in two ways [5]. First, it validates the idea of combining AI-driven crop recommendation with irrigation management within a single unified platform [5].

Second, its heavy reliance on IoT hardware and continuous data connectivity directly reinforces one of AISFA's core design principles — that the same quality of advisory should be achievable through a software-only approach that works reliably even in lowconnectivity environments [5].

F. Mana, Allouhi, Hamrani, Rehman, elJamaoui and Jayachandran (Elsevier, 2024) [6]

Mana et al. [6] published a comprehensive review in Elsevier's Smart Agricultural Technology journal, bringing together evidence from a wide range of published studies on AI applications in farming [6]. The authors found consistent support for the idea *that machine learning-based systems outperform rule-based advisories when it comes to crop selection under variable conditions [6]*. They also found that AI-driven irrigation scheduling can measurably reduce water usage without negatively affecting crop yields [6]. On fertiliser recommendations, the review found that prediction accuracy improves as training datasets become larger and more geographically diverse [6]. The two most commonly identified barriers to adoption were high upfront costs and the need for reliable data pipelines [6]. Given the depth and breadth of evidence it covers and the rigour of the Elsevier peer-review process, this review stands as one of the strongest pieces of supporting literature for the AISFA concept [6].

G. Siddiqui, Nagbansh and Lamba – AgriVision (IRJEdT, 2024) [7]

Siddiqui et al. [7] developed AgriVision, a smartphone application that uses convolutional neural networks to analyse photographs taken by farmers, assessing crop health, identifying disease symptoms, and estimating soil condition from surface appearance [7]. Predictive analytics then convert these assessments into irrigation and fertilisation recommendations [7]. The study showed meaningful improvements in decision quality among the farmers who used the system [7]. However, the design requires a camera-equipped smartphone and a data connection strong enough to upload images to a cloud-based inference service [7]. In areas with poor mobile coverage, or where smartphones are shared rather than individually owned, these requirements become real barriers to use [7]. The image-based pest detection module proposed in AISFA draws on the AgriVision approach [7] but also includes textbased fallback pathways to reduce this dependency [7].

H. Nanthakumar, Atchaya, Brindha, Harshaa and Kaviya – AgriPath (IJARSCT, 2025) [8]

Nanthakumar et al. [8] developed AgriPath, the most feature-rich system reviewed in this paper [8]. It combines an NLP-based conversational interface built on a recurrent neural network, an artificial neural network for pest identification from images, and a Random Forest classifier that assesses soil fertility from RGB colour readings [8]. The platform was developed in Kotlin and React Native, with Firebase and MySQL as backend systems, and supports multiple regional languages [8]. The multilingual capability is particularly significant for Indian farmers [8]. However, the system has notable infrastructure requirements — it needs a camera-equipped smartphone, stable internet connectivity, and substantial server-side computing power [8]. AISFA draws considerable design inspiration from AgriPath[8], especially its multilingual ambition, while aiming to operate with a leaner and more accessible infrastructure footprint [8].

I. Kavitha, Kiranteja, Shridharan, Desigan, Raghav and Pranav (VIT, 2025) [9]

Kavitha et al. [9] at Vellore Institute of Technology developed a multi-source agricultural assistant that integrates groundwater level databases, satellite-derived soil maps, weather forecast APIs, and commodity price feeds alongside standard crop recommendation ML models [9]. Deep learning was used for feature extraction, and computer vision was applied to assess the quality of harvested produce [9]. The combination of agronomic advice with real-time market pricing is a particularly useful feature, as it helps farmers make decisions that are both agriculturally sound and financially sensible [9]. The limitation is that the system depends on a rich historical data environment, which does not exist in all parts of India [9]. AISFA's future development roadmap includes a market intelligence module that draws inspiration from this work [9].

J. Research Gaps Identified

Looking across all nine studies [1]–[9], five limitations appear consistently enough to be considered structural gaps rather than isolated design flaws [1]–[9]. The first is internet dependency — seven of the nine systems are unable to function without a live internet connection [2],[3],[5],[6],[7],[8],[9]. The second is hardware cost — three systems require physical sensor equipment that is beyond the reach of most smallholder farmers [4],[5],[7]. The third is limited geographic scope — several systems were designed and tested only in specific regions or for particular farming scales, making it unclear how well they would perform elsewhere [4],[9].

The fourth is the absence of offline modes — none of the reviewed systems described a credible way to operate when connectivity is unavailable [1]–[9].

The fifth is language accessibility — most systems support only English or a single regional language, which significantly limits their reach among India's diverse farming population [2],[3],[5],[6],[7]. AISFA is designed to address all five of these gaps within a single, unified framework.

Table I: Comparative Overview of Related Studies

Ref.	Title	Authors	Venue	Methodology	Key Finding	Limitation	Relevance to AISFA
[1]	AI in Agriculture	Tapkir et al., 2023	IJNRD	Literature survey; no system built	AI reduces monitoring effort and improves yield prediction	No prototype; high-cost tech (drones, robots)	Motivates need for low-cost softwareonly approach
[2]	Smart Agriculture Assistant	Jagtap et al., 2025	IJARCCE	Flask chatbot + weather & market APIs + MongoDB	Real-time crop, weather, and market advisory via chat	Requires constant internet; no offline support	Informs conversational advisory interface design
[3]	AI Farmer Chatbot	Kaviya et al., 2021	IJRESM	Naive Bayes, KNN, NLP, Speech-to-text	Voice-enabled query handling for farmers	Static dataset; no live or real-time data	Motivates dynamic data pipeline in AISFA
[4]	Growlify	Kasturi & Navis, 2025	IJIRT	ML + IoT sensors + image processing	Efficient resource management for urban farms	IoT hardware required; designed for urban/smallscale only	Motivates hardware-free, scalable design
[5]	AI Crop & Irrigation System	Patil et al., 2024	IJIRCST	ML + IoT soil sensors + weather data + Flask	Combined crop recommendation and smart irrigation	IoT sensors and continuous internet required	Validates combined advisory; reinforces nohardware design
[6]	Sustainable AI Agriculture	Mana et al., 2024	Elsevier	Comprehensive review of ML in agriculture	AI cuts water and fertilizer waste; improves crop yield	High cost; needs reliable data pipelines	Strongest scientific basis for AISFA optimization goals
[7]	AgriVision	Siddiqui et al., 2024	IRJEdT	CNN image recognition + predictive analytics	Improved crop and soil management decisions	Camera smartphone and stable internet required	Guides imagebased pest module; motivates text fallback

[8]	AgriPath	Nanthakumar et al., 2025	IJARSCT	NLP-RNN chatbot + ANN pest detection + RF soil	Multilingual realtime guidance for farmers	High infra.; camera phone + stable internet needed	Inspires LLMbased multilingual design in AISFA
[9]	Smart Agri. Asst. VIT	Kavitha et al., 2025	VIT Chennai	Deep learning + CV + market price data	Agronomic and financial optimisation combined	Depends on rich historical data not available everywhere	Motivates market advisory module in future AISFA

III. PROBLEM STATEMENT

AISFA is designed to address three closely related problems that smallholder farmers in India commonly face [1],[6]. Each of these problems exists on its own, but together they compound one another and significantly reduce agricultural productivity [6].

The first problem is poor crop selection [1],[6],[9]. Ideally, the choice of crop should be guided by current soil nutrient levels, forecasted weather conditions, and historical yield data [6],[9]. In reality, most smallholder farmers do not have access to all three of these information sources at the same time [1],[6]. Crop selection tends to be driven by habit, what neighbouring farmers are doing, or broad regional advice that does not account for individual farm conditions [1],[6]. When the crop chosen is not well-matched to the actual soil and climate conditions, the potential yield is already limited before the first seed goes into the ground [6].

The second problem is inefficient use of resources [5],[6]. Using too much water causes over-irrigation and puts pressure on groundwater supplies; excess fertiliser allows nitrogen and phosphorus to wash into local waterways [6]. Getting the balance right requires knowing what the crop actually needs and what is already present in the soil or expected from upcoming rainfall [5],[6]. Without this information, farmers apply water and fertiliser on fixed schedules not tailored to real conditions, resulting in wasted money, environmental damage, and suboptimal yields [5],[6].

The third problem is delayed detection of pests and diseases [7],[8]. Acting quickly when a pest or disease first appears can limit the damage to a small portion of the harvest [7]. Many existing systems that could help with pest identification require farmers to submit photographs from a smartphone camera, which not all farmers own [7],[8], or depend on a continuous internet connection that is frequently unavailable in the same rural areas where disease outbreaks can spread most rapidly [2],[5].

A solution that is software-only, requires no physical sensor hardware [4],[5],[7], and can deliver reliable recommendations even on minimal or intermittent internet connectivity [2],[3],[5] would directly address all three of these problems. This is precisely the gap that the proposed AISFA framework aims to fill.

IV. OBJECTIVES OF THE STUDY

The proposed AISFA framework is guided by the following objectives:

- To design a software-based intelligent advisory system that supports informed, evidence-based decisionmaking across four key areas: crop selection, water scheduling, fertiliser application, and pest management — addressing the multi-domain gap identified across all nine reviewed works [1]–[9].
- To identify and justify suitable machine learning algorithms for each advisory task, using publicly available agricultural benchmark datasets for model training and evaluation, drawing on evaluation approaches reported in Mana et al. [6] and Kavitha et al. [9].
- To define a clear five-layer system architecture covering data collection, data preprocessing, model inference, decision translation, and delivery of recommendations.
- To identify the most important input parameters — such as soil nutrients, rainfall, and temperature — that have the strongest influence on predicting crop suitability and optimising resource scheduling, as established by Mana et al. [6] and Kavitha et al. [9].
- To systematically compare AISFA against the nine related systems reviewed in this paper [1]–[9], clearly documenting where and how AISFA improves upon the limitations identified in those existing works.
- To outline a realistic implementation roadmap and propose a research agenda for future development and evaluation of the framework.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

V. PROPOSED SYSTEM AND ARCHITECTURE

AISFA is organised as a five-layer platform. Each layer has a specific role in the overall system, and together they form a complete pipeline from collecting raw farm data to delivering clear, actionable recommendations to the farmer.

A. Data Acquisition Layer

This is the entry point of the system. Soil data — specifically the concentrations of Nitrogen (N), Phosphorus (P), and Potassium (K), along with pH level and moisture percentage — is provided by the farmer either by uploading a soil test report or by entering the values manually. Rather than depending on continuous sensor readings — a requirement imposed by Growlify[4], Patil et al. [5], and AgriVision[7] — AISFA accepts batch soil test data, which most Indian farmers can already obtain from their district agricultural offices. Current weather conditions and short-term forecasts are retrieved automatically from a public weather API. The system also caches a multi-day forecast locally, so recommendations remain accessible even if the internet connection drops — directly addressing the offline gap documented across all reviewed systems [1]–[9]. Additional contextual data — crop history, past pest records, and regional farming patterns — is sourced from government agricultural databases and public research datasets [9].

B. Data Preprocessing Layer

Data coming in from multiple sources often contains inconsistencies in format, scale, and completeness — a challenge documented across all reviewed systems [1]–[9]. This layer handles four key operations. First, completeness checking: missing numerical values are filled using the median from the same agro-climatic zone; missing categorical values prompt the farmer for input. Second, scale standardisation: all numerical features are normalised using z-score standardisation so no single variable has an unfair influence on the model. Third, categorical encoding: non-numerical variables such as soil texture type or previous crop grown are converted into binary indicator variables. Fourth, an optional feature engineering step computes a composite soil quality index from N, P, K, and pH values — a combined variable that often carries stronger predictive value than the four individual measurements separately [6],[9].

C. AI and Machine Learning Layer

This layer contains the trained machine learning models that convert the preprocessed data into predictions. The algorithm choices are guided by benchmark results reported in the reviewed literature [6],[7],[8],[9].

Crop Recommendation is handled by a Random Forest ensemble model [9]. This algorithm was chosen because it performs reliably on small to medium-sized agricultural datasets, naturally captures interactions between different input features, and produces probability scores for each crop class rather than forcing a single choice, allowing the system to present a ranked list of top crop options — an improvement over the single-output advisory reported by Tapkir et al. [1].

Irrigation Scheduling uses a Support Vector Machine (SVM) classifier that produces one of three outputs:

increase watering frequency, maintain the current schedule, or reduce watering frequency — building on the irrigation optimisation approach validated by Mana et al. [6] and Patil et al. [5].

Fertiliser dosage is addressed using a gradient-boosted regression model that predicts the gap between the current measured soil nutrient levels and the optimal levels for the chosen crop — consistent with fertiliser recommendation methods found effective by Mana et al. [6].

Pest and Disease Detection uses a two-pathway approach [7]. For farmers who can upload a photograph, a Convolutional Neural Network (CNN) analyses the image to identify signs of disease or pest damage, following the approach demonstrated by AgriVision[7]. For farmers who do not have this capability, the system uses a seasonal risk calendar based on historical pest patterns in the region. This dual-pathway design ensures pest guidance is available to all farmers regardless of device, directly addressing the camera-dependency barrier noted in AgriVision[7] and AgriPath[8].

D. Decision Support Layer

The raw outputs from machine learning models — probability distributions, class labels, or numerical values — are not directly useful to a farmer without translation into clear advisory statements [2],[8]. For crop recommendations, the model produces a probability score across 22 crop classes; this layer selects the top three candidates, removes any whose growing season does not match the current calendar month, and presents the remaining options with a brief explanation of why each crop is a good fit for current conditions [9]. For fertiliser recommendations, the model's output is converted into specific application quantities per unit area [6]. For irrigation recommendations, the output is presented as a specific watering schedule rather than just a general frequency category [5],[6].

A rule-based safeguard layer cross-checks all outputs against known agronomic boundaries, adding an extra layer of reliability [6].

E. User Interaction Layer

This layer is responsible for making the system's recommendations accessible to as wide a range of users as possible. The frontend interface is built using React, which allows for a component-based, responsive design that works in any mobile or desktop browser without requiring the farmer to download a separate application — addressing the smartphone download barrier noted in AgriVision[7] and AgriPath[8].

Two interaction modes are provided. In structured mode, farmers fill out short guided forms to submit their soil and crop data. In conversational mode, the system integrates a Large Language Model (LLM) that can accept free-text or voice input — building on the NLP-based conversational interface demonstrated by Kaviya et al. [3] and AgriPath[8], but without requiring expensive infrastructure. For example, a farmer might type or say: "My soil is sandy, the pH is around 6, and rainfall has been low this year — what crop would suit me?" The LLM automatically extracts the relevant parameter values and forwards them to the appropriate machine learning model endpoint via FastAPI. It then formats the model's output into a plain-language advisory in the farmer's preferred language. Marathi and Hindi are included in the planned implementation, extending the multilingual approach demonstrated by AgriPath[8] to a leaner infrastructure.

Recommendations are displayed on a React-based dashboard with visual indicators showing confidence and risk level — similar to dashboard outputs described by Jagtap et al. [2]. For farmers in areas with unreliable data connectivity, the most important recommendations are also delivered via SMS, addressing the connectivity gap identified across all reviewed systems [1]–[9].

VI. METHODOLOGY

The proposed development of AISFA follows five sequential phases. Since the system has not yet been implemented, these phases represent the planned approach that will be followed during development.

1) Phase One: Data Assembly

The primary dataset for training the crop recommendation module is the Crop Recommendation Dataset available on Kaggle. This dataset contains 2,200 labelled instances across 22 crop classes with seven input features per record. To make the model more relevant to Indian farming conditions, this dataset will be supplemented with regional soil test data and historical yield records from the Indian Council of Agricultural Research (ICAR) — a data enrichment strategy also recommended by Mana et al. [6]. For the pest and disease detection module, the PlantVillage dataset will serve as the primary training source — the same benchmark dataset used by Siddiqui et al. [7]. This dataset contains over 54,000 labelled leaf disease images covering 38 disease categories, making it wellsuited for training a CNN to identify crop diseases from photographs [7].

2) Phase Two: Preprocessing and Feature Engineering

All datasets will go through the preprocessing steps described in Section V-B. The data will be split into an 80:20 train-test ratio, with the 20% test partition set aside exclusively for final model evaluation. Within the 80% training set, five-fold cross-validation will be used throughout model development to monitor generalisation and reduce overfitting — consistent with the cross-validation methodology described by Kavitha et al. [9].

3) Phase Three: Model Training and Selection

For each advisory task, multiple candidate algorithms will be trained and compared. For the crop recommendation task, a Random Forest model will be benchmarked against a Decision Tree and a K-Nearest Neighbour classifier, evaluated using overall accuracy, macro-averaged F1-score, and training time. Kavitha et al. [9] reported above 92% accuracy using similar methods, which serves as the target benchmark. For the fertiliser dosage and irrigation scheduling tasks, a gradient boosting model will be compared against linear regression and Support Vector Regression, evaluated using RMSE and MAE — metrics also used by Mana et al. [6]. For the pest detection task, CNN architectures will be compared using top-1 and top-5 accuracy, also benchmarked against a simpler logistic regression model on HOG features, consistent with the evaluation approach used by Siddiqui et al.
[7].

4) Phase Four: Integration

Once the best-performing model for each task has been selected and validated, the models will be serialised and deployed as independent FastAPI endpoints.

Each module — crop recommendation, irrigation scheduling, fertiliser dosage, and pest detection — will be exposed as a separate API endpoint that accepts input data in JSON format and returns a structured prediction with a confidence score. The LLM layer will be integrated as an orchestration component sitting between the user and the ML models — a role inspired by the NLP orchestration approach used in AgriPath[8]. The React frontend will communicate with all endpoints asynchronously and display the outputs as interactive dashboard components. Integration testing will cover the full range of training data values, and the LLM's ability to interpret natural language inputs will be tested separately using 50 deliberately ambiguous queries [3],[8].

5) Phase Five: Usability Evaluation

Structured walkthroughs will be conducted with a representative sample of farmer profiles, covering a range of literacy levels, device types, and language preferences — a usability evaluation approach aligned with recommendations in AgriPath[8] and Jagtap et al. [2]. For the structured form interface, the main metrics will be task completion rate and the time taken to receive a recommendation. For the conversational LLM interface, the key metric will be parameter extraction accuracy. After each session, participants will complete a short survey to capture self-reported understanding of the recommendations and how confident they felt in acting on them.

VII. SYSTEM FLOW DIAGRAM

Fig. 1: End-to-end data and recommendation flow in AISFA

Farmer Input (Soil Test Data / Crop History / Images)

↓

Data Acquisition Layer → Weather API | Govt. Databases

↓

Preprocessing Layer (Imputation → Normalisation → Feature Engineering)

↓

AI & ML Layer (Random Forest | SVM | Gradient Boost | CNN)↓

Decision Support Layer (Rule Filter → Plain-Language Advisory)

↓

LLM Layer (Natural Language Understanding → Multilingual Response)↓

React User Interface (Web / Mobile App / SMS Notification)

VIII. COMPARISON WITH EXISTING SOLUTIONS

Table II: Feature Comparison of AISFA Against Reviewed Systems

Feature	AgriVision [7]	AgriPath [8]	Growlify [4]	Jagtap et al. [2]	Kavitha et al. [9]	AISFA (Proposed)
IoT sensors required	Yes [7]	No [8]	Yes [4]	No [2]	No [9]	No
Live internet for core use	Yes [7]	Yes [8]	Yes [4]	Yes [2]	Yes [9]	Minimal
Camera phone required	Yes [7]	Yes [8]	Yes [4]	No [2]	Partial [9]	Optional
All 4 advisory areas	Partial [7]	Yes [8]	Partial [4]	Partial [2]	Partial [9]	Yes
LLM conversational interface	No [7]	No [8]	No [4]	No [2]	No [9]	Yes
Multilingual support	No [7]	Yes [8]	No [4]	No [2]	No [9]	Yes (planned)

Offline fallback	No [7]	No [8]	No [4]	No [2]	No [9]	Yes (planned)
Low deployment cost	No [7]	No [8]	No [4]	Yes [2]	Yes [9]	Yes

IX. EXPECTED OUTCOMES

Since AISFA has not yet been implemented, the outcomes described in this section are projected estimates based on results reported in closely related systems reviewed in the literature [1]–[9]. These projections are intended to set realistic, evidence-based targets for when the system is eventually built and evaluated.

Crop Recommendation Accuracy: Studies such as Mana et al. [6] and Kavitha et al. [9] reported classification accuracy above 92% on held-out test data using machine learning models with similar input features [6],[9]. If AISFA achieves comparable accuracy, in at least nine out of ten cases a farmer following the system's recommendation would be planting a crop genuinely well-suited to their current soil and weather conditions — a meaningful improvement over unaided, experience-based decision-making, which the same literature estimates results in a correct match only around 60 to 70% of the time [1],[6].

Resource Use Efficiency: The irrigation scheduling and fertiliser dosage modules are expected to help reduce input quantities by approximately 15 to 25% compared to standard blanket application practices, consistent with the efficiency gains reported by Mana et al. [6] for precision scheduling approaches using similar methods [6]. Reducing unnecessary input use benefits the farmer financially and also reduces the environmental impact of excess water and fertiliser application [5],[6].

Pest and Disease Detection: The CNN-based pest detection module is anticipated to achieve above 90% top-1 accuracy for common Indian crop diseases, based on published benchmark results from the PlantVillage dataset reported by Siddiqui et al. [7]. It is important to note that this figure is drawn from controlled dataset evaluations [7] and would need to be validated through real-world field testing before it can be treated as a reliable performance guarantee.

Usability: The design target is a task completion rate of over 85% for first-time users attempting to retrieve a crop recommendation without any prior training on the system. This threshold has been set based on published usability benchmarks for similar mobile-accessible advisory tools deployed in low-literacy contexts, as referenced in AgriPath[8] and Jagtap et al. [2].

X. ADVANTAGES OF THE PROPOSED FRAMEWORK

The AISFA framework offers several practical advantages over existing systems [1]–[9], particularly in the context of smallholder farming in India.

No hardware required. Unlike systems such as Growlify[4], Patil et al. [5], and AgriVision[7] that depend on physical sensors or IoT devices, AISFA requires no hardware purchase whatsoever. Farmers only need access to a basic web browser, significantly lowering the barrier to adoption for smallholders with limited financial resources [4],[5],[7].

Personalised crop recommendations. Rather than offering broad, region-wide advice [1], AISFA takes into account the specific nutrient levels in the farmer's own soil, making recommendations considerably more relevant than the generalised advisories criticised in Tapkir et al. [1] and Mana et al. [6].

Data-driven resource management. Irrigation and fertiliser guidance is based on measured soil and weather data rather than fixed schedules, achieving the precision scheduling goals identified by Mana et al. [6] and Patil et al. [5] as the key to reducing input waste and environmental harm.

Pest alerts without a camera. The seasonal risk calendar fallback ensures pest guidance remains available to farmers without camera phones, directly addressing the hardware barrier identified in AgriVision[7] and AgriPath[8].

Modular and maintainable architecture. The five-layer design allows individual components to be updated independently, making the framework easier to maintain and adapt over time as better data or algorithms become available [6],[9].

SMS-based delivery for low-connectivity areas. By supporting SMS as an output channel, AISFA extends its reach to farmers who do not have consistent internet access — addressing the offline gap documented across all nine reviewed systems [1]–[9].

LLM-powered multilingual conversational interface. Unlike all nine reviewed systems [1]–[9], AISFA integrates a Large Language Model as the natural language layer, building on the multilingual ambition of AgriPath[8] and the NLP approach of Kaviya et al. [3] while requiring leaner infrastructure.

XI. FUTURE WORK SCOPE

The primary contribution of this research is the identification of five recurring gaps across nine published AI-based agricultural advisory systems [1]–[9], and the proposal of a conceptual framework — AISFA — that directly addresses each of those gaps. The five gaps identified — internet dependency [2],[3],[5],[6],[7],[8],[9], hardware cost [4],[5],[7], limited scalability [4],[9], absence of offline functionality [1]–[9], and poor language accessibility [2],[3],[5],[6],[7] — collectively shaped every architectural decision in the proposed framework.

A. System Implementation

The AISFA framework remains at the conceptual and architectural stage — no software has been developed yet. Full system implementation is the most immediate and impactful direction for future work. The proposed technology stack uses FastAPI as the backend framework and React as the frontend library — a more modern and performant alternative to the Flask and static HTML interfaces used by Jagtap et al. [2] and Patil et al. [5]. The four ML modules would each be deployed as independent FastAPI endpoints. An LLM — such as Llama or Mistral — would serve as the natural language interface, building on but surpassing the NLP capabilities demonstrated by Kaviya et al. [3] and AgriPath[8] while targeting a leaner infrastructure footprint.

B. Model Training and Validation

All four ML modules must be properly trained, validated, and benchmarked. The expected accuracy targets are: above 92% for crop recommendation, based on Kavitha et al. [9]; above 90% top-1 accuracy for pest image classification, based on Siddiqui et al. [7]; and a 15 to 25% reduction in fertiliser and water application quantities, based on Mana et al. [6]. Meeting or exceeding these benchmarks through actual implementation would convert the projected outcomes in Section IX into empirically verified results, significantly strengthening the research contribution [6],[7],[9].

C. Addressing Gap 1: Internet Dependency

Seven of the nine systems reviewed [2],[3],[5],[6],[7],[8],[9] were unable to function without a live internet connection. Resolving this gap in AISFA will require implementing an offline mode through model compression. The crop recommendation and fertiliser models — relatively compact Random Forest and gradient boosting ensembles — can be compressed using quantisation and exported to ONNX or TensorFlow Lite for on-device inference. A locally cached rolling 10-day weather forecast would allow irrigation recommendations to continue during connectivity outages — overcoming the connectivity dependency that limited all reviewed systems [1]–[9].

D. Addressing Gap 2: Hardware and Sensor Dependency

Three of the reviewed systems — Growlify[4], Patil et al. [5], and AgriVision[7] — require physical IoT sensors or camera-equipped smartphones. AISFA was deliberately designed to avoid these requirements [4],[5],[7]. Every module must include a text-input pathway that works without any hardware. The pest detection module in particular must retain its seasonal risk calendar fallback, so that farmers without camera phones can still receive pest advisories — directly addressing the camera dependency identified in AgriVision[7] and AgriPath[8].

E. Addressing Gap 3: Scalability Across Farm Sizes and Regions

Several reviewed systems, including Growlify[4], were designed for specific farming contexts and did not scale well to different farm sizes or environmental conditions [4],[9]. Future implementation of AISFA should be tested across at least three distinct Indian agro-climatic zones — the semi-arid Deccan Plateau, the humid coastal belt, and the Indo-Gangetic Plain — to verify that the ML models produce agronomically sound recommendations across varied settings. Regional soil and yield datasets from ICAR should be incorporated into the training data, as recommended by Mana et al. [6]. Given the authors' institutional context, deploying an initial pilot in Pune district would be a practical starting point.

F. Addressing Gap 4: Multilingual and Voice Access

None of the nine reviewed systems fully addressed language accessibility for non-English-speaking farmers [1]–[9]. AgriPath[8] made the strongest attempt with its multilingual chatbot, but its high infrastructure requirements limited practical reach [8]. In AISFA, the LLM layer is the natural solution. A multilingual LLM such as Llama 3 or Gemma — both of which support Hindi and can be fine-tuned on Marathi agricultural vocabulary — can accept queries in any supported language, extract the agronomic parameters, call

the relevant ML endpoints, and compose a response in the same language the farmer used — extending the multilingual goal of AgriPath[8] and Kaviya et al.
[3] to a leaner infrastructure. Voice input through the Web Speech API would further reduce the literacy barrier.

G. Addressing Gap 5: Market and Economic Intelligence

Only one of the nine reviewed papers — Kavitha et al. [9] — integrated market price data into its farming recommendations [9].
Future implementation of AISFA should connect to the Agmarknet government commodity price database, displaying current and 30-day average prices for each recommended crop alongside its agronomic suitability score. A combined ranking weighing both soil-weather fit and projected harvest price would give farmers a genuinely complete decision-support output — something not currently provided by any of the reviewed systems [1]–[9].

H. Field Trials and Real-World Validation

The final and most important phase of future work is validating the system under actual farming conditions. A structured field trial, conducted in collaboration with a Krishi Vigyan Kendra or the Maharashtra State Department of Agriculture, would involve recruiting 30 to 50 farmers across a range of plot sizes and crop types. The implemented system would be deployed for one full growing season, and data would be collected on crop yield, input costs, and farmer-reported decision confidence — both before and after adoption. This real-world evidence would either confirm the design assumptions made in this paper or reveal where the framework needs revision, completing the research cycle that this conceptual study sets in motion [1]–[9].

XII. CONCLUSION

This paper has presented the conceptual design of AISFA, a machine learning-based agricultural advisory system built specifically to meet the needs and constraints of smallholder farmers in India. The case for developing such a system rests on three clear foundations.
First, the existing literature [1]–[9] consistently shows that data-driven advisory systems outperform experience-based decision-making for crop selection, resource management, and pest response [6],[9]. Second, that same literature reveals a recurring set of infrastructure-related barriers — internet dependency [2],[3],[5],[6],[7],[8],[9], hardware costs [4],[5],[7], and language limitations [2],[3],[5],[6],[7] — that prevent even the most capable existing systems from reaching the farmers who need them most [1]–[9]. Third, the five-layer software architecture proposed for AISFA has been deliberately designed to address each of these barriers, without compromising on the quality of the recommendations it delivers.
No working prototype has been developed as part of this research. The contribution of this paper is the design itself — a system architecture grounded in established machine learning practice [6],[7],[8],[9], motivated by real and documented farmer needs [1]–[9], and clearly differentiated from prior work [1]–[9] through its software-only approach, low-connectivity operation, scalable structure, and LLM-powered multilingual interface.
Realising the full potential of this framework will require future work in several areas: complete software implementation using FastAPI and React, thorough model training and validation against benchmarks established in the reviewed literature [6],[7],[9], pilot deployment across different agro-climatic regions [4],[9], and iterative refinement based on feedback from real farmers in real conditions. It is through this next phase of work that the design proposed here can be transformed into a tool that makes a measurable difference to the lives and livelihoods of the farmers it is intended to serve [1]–[9].

REFERENCES

[1] S. Tapkir, A. Vathare, and P. Suryawanshi, “Implementation of Artificial Intelligence in Agriculture,” Int. J. Novel Research and Development (IJNRD), 2023.
[2] P. S. Jagtap, A. M. Suryawanshi, K. G. Kakade, and S. Y. Bhatkute, “Smart Agriculture Assistant: An AI-Powered Approach for Precision Farming and Crop Management,” Int. J. Advanced Research in Computer and Communication Engineering (IJARCCE), 2025.
[3] P. Kaviya, M. Bhavyashree, M. D. Krishnan, and M. Sugacini, “Artificial Intelligence Based Farmer Assistant Chatbot,” Int. J. Research in Engineering, Science and Management (IJRESM), vol. 4, no. 5, 2021.
[4] V. K. Kasturi and C. D. P. Navis, “Growlify: An AI-Powered Urban Farming Assistant,” Int. J. Innovative Research in Technology (IJIRT), vol. 12, no. 3, Aug. 2025.
[5] R. S. Patil, A. K. Sharma, P. V. Deshmukh, and S. R. Kulkarni, “AI-Based Crop Recommendation and Smart Irrigation System for Sustainable Agriculture,” Int. J. Innovative Research in Computer Science and Technology (IJIRCST), 2024.
[6] A. A. Mana, A. Allouhi, A. Hamrani, S. Rehman, I. el Jamaoui, and K. Jayachandran, “Sustainable AI-based Production Agriculture: Exploring AI Applications and Implications in Agricultural Practices,” Smart Agricultural Technology, Elsevier, vol. 7, Article 100416, 2024.
[7] A. Siddiqui, A. Nagbansh, and S. Lamba, “AgriVision: AI-Driven Solutions for Sustainable Agriculture,” IRJEdT, vol. 6, no. 10, Oct. 2024.

[8] G. Nanthakumar, R. Atchaya, S. Brindha, I. Harshaa, and R. Kaviya, "AgriPath: An AI-Driven Assistant Empowering Farmers with Real-Time Guidance, Multilingual Access, and Tailored Farming Support," IJARSCT, vol. 5, no. 2, May 2025.
[9] D. Kavitha, P. Kiranteja, V. J. V. Shridharan, P. K. Desigan, N. G. Raghav, and V. Pranav, "A Smart Agricultural Assistant for Crop Recommendation using Machine Learning," School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India, 2025.