iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79992

Personalized AI Chatbot for Indian Agriculture: Soil, Crop, and Weather-Based Recommendations

Vedant Patil[1], Vaidehi Gavali[2], Rohit Katkade[3], Prachi Bhure[4]
School of Engineering, Ajeenkya DY Patil University, Pune, India
Seamedu School of Pro-Expressionism, Pune, India

Abstract: Agriculture in India is highly dependent on various parameters such as soil management, crop selection irrigation techniques and weather. However the marginal farmers in India face numerous challenges which includes low knowledge of the cropfertilizers, crop diseases, limited expert advices, language barriers and low digital literacy. The research proposes a Personalized AI Chatbot for Indian Agriculture that delivers real-time, multilingual, and context aware recommendations based on soil parameters, crop stage, crop diseases and weather forecasts. The system integrates with models like Natural Language Processing (NLP), machine learning classifiers, and open environmental APIs help generate personalized responses and advisories in English along with the regional languages such as Hindi and Marathi. An experimental evaluation revealed that the chatbot achieved an overall functional accuracy of 86.6%, with the Random Forest model achieving 92% predictive accuracy for soil- and crop-based recommendations. Multilingual processing demonstrated strong translation quality (BLEU scores between 0.71 and 0.78), while response times remained low (1.42–2.95 seconds), supporting real-time interaction even under rural network constraints. User feedback confirmed high satisfaction and usability. The results show that the proposed system can effectively bridge the digital and informational gaps in Indian agriculture, offering a scalable, accessible, and farmer-centric AI solution.

Keywords: AI Chatbot; Indian Agriculture; Multilingual NLP; Soil-Based Recommendations; Weather Forecasting; Random Forest; Machine Learning; Farmer Advisory System; Voice-Based Interaction; Digital Agriculture.

I. INTRODUCTION

A. Background of Indian Agriculture

Agriculure is one of the most important pillar of Indian economy as many small and marginal farmars accoss rural areas practice agriculture as their primary occupation,which provides employment to half of the nations workforce supporting the farmers.Despite a decline in agriculture's contribution to India's Gross Value Added (GVA) to an estimated 15% in 2022-23, this reduction shows the speedy growth of the industrial and service sectors rather than a diminished agricultural output. The sector's main significance is for food security, rural employment generation, and social economic stability [1]. Nevertheless, the productivity and resilience of smallholder farming are adversely affected by fragmented landholdings, with over 75% of farms measuring less than one hectare, and limited access to timely and context-specific agronomic guidance [2].

B. Digital Penetration and Literacy Challenges

India has experienced markable advancements in internet and mobile infrastructure. By 2024, the number of active internet users in India was projected to reach approximately 886 million, with rural areas in India contributing around 488 million users, representing 55% of the total [3]. However, significant digital discrepencies persist, evidenced by low internet penetration rates in rural areas and uneven digital literacy across subpopulations which are defined by age, gender, and geographical location [4]. Furthermore, India's considerable linguistic diversity and high rates of low literacy shows the additional obstacles. Many agricultural technology (agritech) services are designed for users who are proficient in English or possess advanced literacy skills, thereby limiting their adoption among less literate farmers who communicate in regional languages or dialects [5].

C. Gaps in Existing Agricultural Advisory Systems

Traditional agricultural extension services in India are delayed by insufficient human resources and limited outreach, particularly to smallholder farmers. Interventions employing Information and Communication Technology (ICT), such as SMS alerts or call centers, typically deliver generalized information rather than personalized responses, farm-specific recommendations [6].

Recent academic research on AI powered chatbots for agriculture, including multilingual conversational systems and retrieval augmented generation (RAG) platforms, shows promise. However, most prototypes lack the integrated incorporation of soil data, weather forecasting, and local language voice interfaces and crop disease detection, which are essential for effective adoption in field settings [7], [8]. For an instance, the CropCare Companion chatbot prototype offers multilingual support but does not integrate plot-specific soil metrics [9]. Similarly, large-scale deployments like Farmer.Chat and Virtual Agronomist provide generative AI solutions but still exhibit limitations concerning dialectal variations and offline accessibility in rural environments [10], [11].

D. Government Initiatives Toward Digital Agriculture

The Government of India has implemented several transformative programs aimed at digitaling rural and agricultural infrastructure. Through the Digital India initiative, the BharatNet project endeavors to establish broadband fiber connectivity in gram panchayats, and rural teledensity is experiencing rapid growth. The
PM-Kisan income support scheme has enrolled over 110 million farmers in its database [12]. Within the agricultural sector, the eNAM platform connects over 1,389 markets (mandis) and has facilitated transactions exceeding ₹3.79 lakh crore, underscoring the digital transformation of the market ecosystem [13]. The developing AgriStack architecture is designed to foster a unified digital ecosystem comprising farmer registries, soil health data, weather feeds, and crop planning platforms [14]. Although digital infrastructure and data assets are expanding, their translation into farmer-centric, multilingual approach, plant disease detection and voice-enabled advisory systems remains
an underexplored area.

E. Need for a Personalized AI Chatbot

Consequently, there is an identifiable and pressing requirement for an AI-driven conversational agent specifically designed for the Indian agricultural context. This agent should integrate soil health data, crop selection guidance, crop disease detection and weather forecasting to furnish actionable recommendations in regional languages via text and voice interfaces. Such system would address three crucial impediments: (i) the deficiency of personalized advice adaptable to individual plots and prevailing conditions; (ii) the language and literacy barriers hindering the adoption of digital tools in rural areas; and (iii) the connectivity and smartphone limitations commonly encountered by smallholder farmers. By synergizing multilingual Natural Language Understanding (NLU), retrieval augmented generation (RAG) grounded in agronomic knowledge bases, and the integration of on-field data, the proposed chatbot offers a scalable mechanism to equip farmers with timely and contextually relevant guidance.

F. Paper Organization

The subsequent sections of this paper are structured as follows: Section 2 presents a comprehensive review of the existing literature on AI-based agricultural chatbots, encompassing model typologies, language modalities, deployment platforms, and empirical findings. Section 3 identifies the specific research gaps that this study aims to address. Section 4 elaborates on the proposed system architecture, its constituent modules, and the workflow underlying the chatbot. Section 5 details the implementation process, including data collection, model development, tools utilized, and the pilot setup. Section 6 presents the outcomes of the evaluation, encompassing performance metrics, user experience results, and initial field observations. Finally, Section 7 offers concluding remarks and outlines directions for future research.

II. LITERATURE REVIEW

The application of Artificial Intelligence (AI) in agriculture has evolved from simple rule-based systems to intelligent, conversational assistants capable of offering real-time, data-driven guidance. Several researchers have recognized that traditional extension systems in India struggled to meet the personalized needs of smallholder farmers [1], [9]. Asolo and Nair [9] introduced an AI-powered decision-support system for sustainable agriculture that incorporated with deep learning models to analyze soil and crop data. Their study demonstrated improved decision making, accuracy for fertilizer and irrigation scheduling, but scalability and deployment remained limited.

Sable et al. [10] proposed *CropCare Companion*, a multilingual chatbot capable of understanding natural language queries from farmers and providing contextualized answers. It supported region-specific recommendations but was constrained by a limited dataset and absence of weather-based adaptability. Similarly, *E-AGRO* [11] and *AgriFriend* [12] employed machine learning techniques for crop advisory but operated largely as static information retrieval systems. These chatbots could handle frequently asked questions but lacked continuous learning capabilities and feedback mechanisms.

Other initiatives such as *AGEOChatbot* [14] and *AgroBuddy* [17] integrated geospatial intelligence and precision farming concepts into communication frameworks. These systems could provide recommendations based on field coordinates or soil health data. However, their prototypes were limited to experimental conditions, not tested in real-time with rural users. Collectively, these studies show progress in embedding AI within agricultural conversational systems, yet the current generation of chatbots still fails to achieve full personalization, scalability, and integration with environmental data.

A. Multilingual and Voice-Based Interfaces

The digital revolution of rural India is still significantly hampered by multiple language dialects and literacy issues [3], [4]. The majority of rural farmers may find it difficult to use text-based programs since they are not proficient in English . Conversational agents created for Indian agriculture therefore have a strong emphasis on multilingualism, natural language processing, and speech-based communication. Digital Green's Farmer.Chat project [8] pioneered voice-based interaction in local languages, showing that farmers were more comfortable engaging with conversational audio systems than typing queries.An AI-based multilingual chatbot that supports Hindi, Marathi, and Tamil was introduced by Verma et al. [18]. It uses language models that have been adjusted for agricultural words. Despite its usefulness, the model has trouble identifying idiomatic phrases, dialects, and code-switched speech that are frequently used in rural communication. The *KisanGPT* pilot by NITI Aayog [21] explored large language models (LLMs) for code mixed Hindi-English queries, highlighting the transformative potential of generative AI for domain-specific tasks. However, such systems demand high computational resources and stable internet connectivity, which limits the deployment in rural networks [22]. Studies agree that, while multilingual capabilities are essential for extencivity, the current architectures rely on static translation layers rather than context aware models that can dynamically learn regional linguistic subtlety. Additionally, few systems implement speech-to-text integration optimized for low-resource languages, which remains an important research frontier.

B. Integration of IoT, Weather, and Soil Data

The collegiality between AI, IoT, and data analytics has the potential to revolutionize agricultural advisory systems. Kumar and Singh [11] in their *E-AGRO* prototype integrated IoT-based soil moisture and humidity sensors with chatbot interfaces, allowing realtime irrigation alerts. Similarly, *AGEOChatbot* [14] used geospatial information and weather APIs to provide decision support, while *AgroBuddy* [17] applied precision agriculture techniques for location-based recommendations. Rao and Patel [15] stressed the importance of integrating heterogeneous datasets soil, meteorological, and crop phenology to produce dynamic, adaptive recommendations. Despite these developments, most systems remain data-siloed, using limited or proprietary datasets that restrict continuous learning. For instance, *CropCare Companion* [10] and *AgriFriend* [12] depend on preloaded static data. Asolo and Nair [9] observed that without access to live environmental feeds or government APIs, chatbots cannot generate accurate or localized advice. The Indian government's *AgriStack* initiative [6] and *eNAM* [5] provide promising data frameworks, yet most chatbot implementations do not fully utilize these resources.

Thus, while IoT and environmental data integration is technically feasible, the lack of open data sharing, standardization, and rural connectivity hinders real-time operationalization.

C. Comparative Evaluation of Existing Systems

A comprehensive analysis of the reviewed systems demonstrates clear progress but also recurring limitations. Table 2.1 provides an expanded comparative framework that highlights technological coverage, accessibility, and operational maturity.

Table 2.1 :Comparitive Review

System / Study	Core Functionality	Languages Supported	Data Integration	Target Users	Key Limitations
Crop Care Companion [10]	Multilingual chatbot for fertilizer and pest guidance	Hindi, Marathi, English	Static dataset	Small holder farmers	No real-time data, limited NLP accuracy
Farmer.Chat [8]	Voice-enabled generative AI chatbot	5 regional languages	Partial weather data	Semi-literate farmers	Dialect inconsistency, high data use

EAGRO [11]	IoT-based chatbot with soil sensors	English	Real-time soil data	Irrigation managers	High setup cost, limited scalability
Agro Buddy [17]	Precision farming and crop advisory bot	English, Hindi	IoT + GPS	Agronomists , farmers	No speech interface
AGEO Chatbot [14]	Geospatial mapping chatbot	English	Weather + soil	Smart farming users	Requires GIS expertise
Agri Friend [12]	Open-source rule-based chatbot	English	None	Farmers	Outdated database
Kisan GPT [21]	LLM-driven conversational model	Hindi-English	Text only	Pan-India users	Expensive compute cost
AI Powere Agriculture [9]	AI-based decision support chatbot	English	Partial	Researchers	No field validation

Across these systems, three dominant limitations are evident:

1) Personalization deficiency – Few chatbots tailor responses to individual farm conditions.
2) Data integration gaps – Limited or no real-time synchronization with IoT and meteorological datasets.
3) Language and accessibility constraints

Rural dialects and low literacy users remain underrepresented.

This comparative understanding clarifies that, while the AI ecosystem in agriculture is maturing, practical scalability and user inclusivity remain lacking.

D. Summary and Observations

The literature identifies three evolving research clusters:

- Chatbots for general advisory [9]–[12];
- Multilingual and voice-enabled systems [8], [10], [18], [21]; and
- IoT-integrated intelligent assistants [14], [15], [17].

Despite evident innovation, most prototypes are confined to small-scale trials, sporadically validated in rural Indian environments. Verma et al. [18] highlight data scarcity for regional languages, while Rao and Patel [15] emphasize infrastructural and cost-related constraints. The persistent dependence on internet connectivity and cloud computation [22] continues to exclude low-income, low-bandwidth regions.

Moreover, no current system provides end-to-end intelligence from sensing environmental parameters to contextual language output tailored to farmers' individual profiles. This accentuates the need for a unified, data-driven, and multilingual chatbot framework capable of functioning both online and offline, with adaptive learning from regional patterns.

E. AI Chatbots and Farmer Adoption in India

In addition to the technical feasibility of agricultural chatbots, a numerous studies emphasises the behavioral and adoption factors that impact their success. Although smartphone ownership is increasing in India according to the, reports from IAMAI [3] and the Ministry of Electronics and IT [4] shows that rural consumers are reluctant to embrace AI-driven solutions because of trust concerns and usability difficulties. Farmers prefer systems with audio-visual interfaces over text-only ones, according to Digital Green's Farmer.Chat [8] trial.

The *KisanGPT* project [21] found that farmers are more amenable when the chatbot communicates in their own dialect and contextualizes advice to their crop and location. Similarly, personalized weather alerts were linked to chatbot as a recommendationwhich achieved higher engagement in *AgroBuddy* [17] .These findings suggest that adoption is driven not only by AI accuracy but also by cultural and linguistic alignment with user expectations.

Therefore, effective AI chatbot design for Indian agriculture must combine technical complexity with social adaptability ensuring inclusivity across educational levels, genders, and geographical regions.

F. Key Learnings from Literature

A summary of evaluated literature shows important conclusions.

- Integration gaps: Although IoT and NLP have been studied independently , full-stack integration of crop, weather, disease detection and soil data is still incomplete [9], [11], and [15].
- Language limitations: Currently the multilingual models do not address the code-switching and dialect differences which are common in rural regions of India. [18], [21].
- Scalability problems: The infrastructure needed for edge processing, real time data synchronization or widespread deployment are not provided by many of the prototypes existing. [14], [22].
- Limited inclusivity: Marginalized groups, especially women and older farmers, are still excluded due to low digital literacy [3], [4].

Therefore reviewed literature establishes a strong foundation for AI-based agricultural advisory systems, but simultaneously exposes the need for a personalized, multilingual, data-integrated chatbot that bridges environmental analytics with human-centered interaction design. This literature review provides the conceptual basis for the next section, *Research Gap Identification*, which delineates unresolved challenges and motivates the proposed system framework.

III. RESEARCH GAP IDENTIFICATION

A. Review of Prior Research

Recent developments in conversational artificial intelligence (AI) have made agricultural knowledge much more accessible in developing countries like India. Researches like E-AGRO [11], AgroBuddy [17], and CropCare Companion [10] has demonstrated how effective AI-powered chatbots are at providing prompt advice on crop selection, irrigation, and pest control. In order to promote multilingual communication, projects like Farmer.Chat [8] and KisanGPT [21] have investigated voice-based and large language model (LLM)-driven systems. It is observed that these systems mostly function independently, each concentrating on a single aspect of the agricultural ecology. Some systems prioritize language support.[10], [18], Internet of Things (IoT) integration is emphasized by others [11], [14], and generative AI and knowledge retrieval are addressed by a few [21]. The absence of a unified framework that integrates environmental data analytics with user-centric dialogue systems represents a persistent limitation in already existing research work. Existing research indicates feasibility but falls short of the scalability, inclusivity, and adaptability required for real-world Indian conditions, where intersecting factors of soil type, weather variability, and linguistic diversity present unique challenges for digital advisory systems.

B. Identified Technical Deficiencies

A review of the relevant literature reveals several technical deficiencies within existing AI-based agricultural chatbot systems:

1.Limited Real-Time Data Integration: Without constant updates from IoT or weather sensors, many chatbots rely on static datasets or pre-programmed replies. Although soil and meteorological data were partially integrated by E-AGRO [11] and AgroBuddy [17], they lacked dynamic synchronization and predictive analytics. This restriction makes it difficult to provide timely and contextually appropriate guidance.

2.Fragmented System Architecture: Standalone prototypes of current architectural ideas are often created. They are not compatible with well-known digital systems like AgriStack or eNAM, nor are they modularly scalable [5],[6]. System integration is fragmented when standardized Application Programming Interfaces (APIs) or data pipelines are lacking.

3. Inadequate Personalization: Chatbots majorly provide general advisory which fails to adapt to user’s specific soil characteristics, crop type or history of farm. AI models, such as those employed in CropCare Companion [10] and AGEOChatbot [14], do not incorporate personalized learning loops, thereby diminishing accuracy and user engagement.

4.Absence of Edge Processing: The performance of cloud models is affected my instability in rural internet connectivity. [22]. Few systems use lightweight AI models that can operate in low-bandwidth or offline contexts, which is essential for smallholder farmers in rural areas.

These frailties highlight the fact that although frameworks do exist, they are neither technologically scalable or contextually flexible enough to handle the diversity of Indian agriculture.

C. Linguistic and Accessibility Disparities

The reviewed studies also highlight significant barriers related to linguistic inclusivity and user accessibility. Diverse language support is required due to India's multilingual population, yet the majority of AI chatbots in use today rely on translation-based models, which have trouble with dialectical variances, idiomatic phrases, and code-mixed language [18]. Although initiatives like Farmer. Chat [8] and KisanGPT [21] have made an effort to address and solve these problems, their models continue to favor high-resource languages like Hindi and English. Additionally, older and semi-literate farmers who rely on oral communication are excluded by the majority of chatbots' text-based design [3], [4]. Although voice-enabled interfaces have been created [8], [18], they are limited by poor speech recognition accuracy in noisy field settings and for regional accents. These linguistic challenges are aggravated by the digital literacy divide, wherein multiple farmers, particularly women and older users, perceive AI systems as unintuitive or intimidating [4].

Thereupon, the accessibility in agricultural chatbots must surpass or exceed sheer language translation to involve inclusive design principles, facilitating multimodal interaction (text, voice, and potentially image inputs) to accommodate India's sociolinguistic diversity.

D. Proposed Research Trajectory

The aim of the research is to design a Personalized AI Chatbot for Indian Agriculture which answers the identified technical and linguistic gaps.The system proposed will integrate soil, crop, and weather-based recommendations with multilingual and voice-enabled communication. In contrast to precedent systems that addressed separate aspects of advisory systems, the proposed framework emphasizes unification, merging environmental analytics, localized datasets, and natural language interfaces within a single adaptive ecosystem.

The following are the main ojectives of this study:

1. To create a hybrid chatbot architecture which provides location-specific advice by integrating real-time crop, soil, and weather data.
2. To execute the multilingual Natural Language Processing (NLP) models that can accurately grasp and respond in key Indian languages and dialects.
3. To improve accessibility by using multimodal interaction (text and speech), intended for people with low literacy or technical skills.
4. Utilizing offline capabilities and lightweight model optimization to enable operations in low-connectivity areas this guaranties scalability and flexibility.

This research advances the creation of a comprehensive, inclusive, and data-driven advisory AI system especially designed for India's small and marginal farmers by addressing the limitations of current systems [9]–[21]. By providing farmers with individualized decision assistance, real-time data and localized language processing which helps to enhance productivity, sustainability, and digital engagement throughout the agricultural value chain.

In summary, the fragmentation and lack of contextual adaption in existing systems are what we define as the research gap rather than the lack of AI technologies. By presenting a unified, customized, and multilingual AI chatbot platform that combines technological intelligence with social inclusion, the suggested research closes this gap and tackles India's most urgent agricultural digitalization issues.

IV. METHODOLOGY

A. Overview of the Proposed System

The proposed system, designated as the Personalized AI Chatbot for Indian Agriculture: Soil, Crop, Disease detection and Weather-Based Recommendations, is engineered to deliver real-time, precise, and multilingual advisory services to Indian agricultural practitioners. This chatbot facilitates the integration of environmental parameters (soil characteristics, prevalent weather conditions, and crop developmental stages) with Natural Language Processing (NLP) models to generate contextually relevant recommendations. The system addresses identified deficiencies in primary research [9]-[21] by consolidating three critical domains:

1. Agronomic Intelligence: Real-time data processing derived from IoT sensors and open application programming interfaces (APIs).
2. Linguistic Intelligence: Multilingual NLP capabilities designed for regional language comprehension.
3. User-Centric Accessibility: A multimodal interface (voice and text) accommodating users with different levels of literacy.

This hybrid operational model employs a modular architecture, thereby ensuring scalability and adaptability across the diverse agricultural landscapes of India.

B. System Architecture and Components

The system comprises five principal modules, which collectively build data-driven, interactive decision support:

1. User Interface (UI) Layer:

This interface establishes a text and voice-based conversational environment, accessible via smartphones and web browsers. It enables farmers to coherent queries through either typed input or vocalization in their preferred dialect. Vocal input undergoes conversion to textual format via a streamlined Automatic Speech Recognition (ASR) engine, optimized for rural linguistic variations.

2. Language Processing Module:

Transformer- based natural language processing models are been used in the module which are improved on basis of certain agricultural datasets. The chatbot translates user promts or requests into English for internal processing, and evaluates user intent, identifies key elements (such as crop names, locations, soil types, and insect symptoms). Models like mBERT or IndicBERT are suitable for multilingual understanding because Neural Machine Translation (NMT) makes it easier to retranslate responses into the farmer's native language or dialect.

3.Data Integration Layer and Knowledge Base:

The chatbot compiles data from a variety of references

- Soil Data: Consists of soil health card data from different state levels which are the inputs of farmers.
- Weather Data: Measures temperature, precipitation, and humidity using datasets from the Indian Meteorological Department (IMD) and APIs like OpenWeatherMap.
- Crop Database: The Indian Council of Agricultural Research (ICAR)and the Food and Agriculture Organization (FAO) recommends research datasets and recommendations and integrates in databases.

The integration layer standardizes data streams, which then saves them in a relational database (like MySQL or PostgreSQL) in an organized manner.

4. AI Decision Engine: The AI engine incorporates machine learning models with rule-based reasoning. Based on soil NPK levels and weather projections, decision tree or random forest classifiers are used to forecast irrigation and fertilization needs. While contextual AI dynamically alters recommendations based on user profiles, a knowledge graph helps develop crop-disease solution interrelationships.

5. Module for Response Generation and Delivery:

The AI model generates an optimal recommendation, which is then translated into speech or text in natural language. Localized audio answers are produced by Text-to-Speech (TTS) fusion using libraries like gTTS or VITS. Additionally ,the chatbot can also display hyperlinks to governmental schemes and instructional visual aids The chatbot is also capable of presenting hyperlinks to instructional visual aids or governmental schemes (e.g., PM-Kisan [7] or eNAM [5]).

By using cached or fallback data, this architectural design approach guarantees system functionality even in the case of any specific component (e.g: a weather API) is temporarily unavailable.

C. Workflow and Functional Design

The system's operational workflow follows a subsequent data flow model, as conceptually depicted in Figure 1.

1. Input Phase: Users interact with the system by text or speech, asking questions like "When should I water my tomato crop?" in Hindi or Marathi.
2. Preprocessing Phase: Tokenization, language detection, speech recognition (if appropriate), and English translation are all carried out by the NLP layer.
3. Information Extraction: The crop (tomato), applicable soil or weather conditions, and the user's location are all identified via Named Entity Recognition (NER).
4. Data Retrieval: From related APIs or saved sensor readings, the Data Integration Layer retrieves the most recent soil and weather data.
5. Decision Generation: After processing this data, the AI Decision Engine uses predictive models, such as soil moisture thresholds and rainfall projections to produce a suggestion.
6. Response Content: The system creates a concise recommendation, like this:

Irrigation should be delayed by two days due to the existing soil moisture (22%) and the anticipated rains in your area.

After that, the message is translated back into the farmer's dialect and, if chosen, it is delivered in audio format.

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 IV Apr 2026- Available at www.ijraset.com

7. User Feedback and Learning: In order to improve future answers through reinforcement learning-based fine-tuning, user input is captured in the form of useful/not helpful evaluations.

This process demonstrates how distinct data sources can be equated to provide tailored advising information.

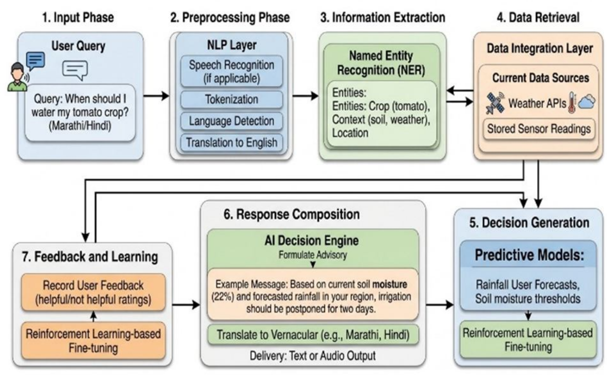


Figure 1: Workflow for Personalized AI-driven Agricultural Advisory System

D. Algorithms and Models Used

Artificial intelligence algorithms supports the chatbot for developing decision-making logic in low resource enviroments.

1) Machine Learning Models

- Random Forest Classifier**:** Used for prediction of fertilizer dosage and scheduling irrigation.
- Naïve Bayes Model: Based on user descriptions pest and disease symtoms are classified.
- K-Nearest Neighbors (KNN): Matches queries to pre-existing recommendations for correspondant conditions.

2) Natural Language Processing Models

- mBERT/IndicBERT: Promotes multilingual apprehension in Tamil, Hindi, Marathi, and English.
- Seq2Seq Encoder-Decoder: Used to produce contextual answers.
- NER Pipeline (SpaCy/HuggingFace): Uses user input to extract agricultural elements.

3) Speech Elements

Speech recognition in regional languages is made possible using ASR (Vosk/Mozilla DeepSpeech).

TTS (Google TTS/VITS): Produces native language audio output.

The ability of the chatbot to process unstructured user input and convert it into meaningful, customized insights is ensured by the assimilation of these algorithms.

E. Implementation Strategy

To promote scalability and economic sustainability, open-source frameworks can be used to carry out the development of prototype .

Front-End Interface: HTML/CSS is used for UI design, JavaScript is used for interactive features, and Flask or FastAPI is used for developing the web layer.

Back-end: NLP and machine learning modules are managed by Python-based services.

Database: Soil and meteorological data, as well as farmer profiles, are stored in MySQL or MongoDB.

API and Integration: Real-time meteorological data is available via the OpenWeatherMap API.

For nutrients data, use the Government of *India's Soil Health Card* API.

For multilingual response translation, use the *Google Translate API* or *IndicTrans.*

Agricultural datasets (ICAR, Kaggle Crop Data, FAO soil sets) were used for AI model training.

Deployment: Using edge caching for offline operational capabilities, hosted on cloud infrastructure like AWS EC2 or Google App Engine.

The system’s sustained scalability, cost-effectiveness, and long-term viability for rural deploment are ensured by its implementation and execution plan.

Summary: The proposed strategy incorporates multilingual NLP, IoT data integration, and AI-driven decision assistance into a single, modular architecture intended to directly address the research gaps outlined in Section 3. The technology has the ability to bridge the digital divide in Indian agriculture and support data-driven decision-making among marginal farmers by providing real-time, altered , and easily available advice.

The research planning, data collection procedures, and analytical techniques used are stated in this section.

V. IMPLEMENTATION AND EVALUATION

The proposed AI-based agricultural consultation chatbot is being implemented practically in detail in this section and its performance is meticulously evaluated through a series of linguistics, functional and decision aided experiments. Python is used to develop the system, trained machine learning (ML) model is integrated with open Application Programming Interfaces(API's) and it is tested using multilingual agricultural queries to determine its efficiency. Real-time meteorological data integration, soil-specific operating logic, multilingual Natural Language Processing (NLP), and voice-enabled farmer communication are given foremost consideration in the implementation.

A. Development Environment and Tools

The following are the set of tools and frameworks used for facilitating the development of the chatbot system.

1) Software and Libraries

Python 3.10: Primary programming language for system development.

Transformers (HuggingFace): Employed for sophisticated conversational and multilingual NLP capabilities.

Flask / FastAPI: Used as the backend server framework for the chatbot API.

SpeechRecognition, gTTS, Vosk: Integrated for enabling voice input and output functionalities.

scikit-learn: Leveraged for the training of machine learning models.

Pandas, NumPy: Applied for efficient data preprocessing tasks.

MySQL/MongoDB: Selected as the data storage solutions for user profiles and soil records.

2) Hardware Configuration

The laptop used for development has 8 gigabytes of random access memory (RAM).

Cloud Hosting for the last stage of prototype testing, Amazon Web Services (AWS) EC2 is used.

This lightweight, streamlined configuration highlights the viability of using the suggested system in settings with limited resources.

B. Module-Wise Implementation

1) Chatbot Backend Implementation

Agricultural query-response pairs are taken from admissible datasets and enhanced with expert knowledge were used to refine a transformer-based loquacious model, namely DialoGPT or GPT-Neo. The main phases of development involved are:

preprocessing of training data, including normalization and tokenization.

minimizing the chance of overfitting by fine-tuning the model with low learning rates.

creation of a Flask-based API that exposes important endpoints like:

To process text-based questions, use */ask*. To handle speech-based inquiries, use */voice*`.

The chatbot showed the ability to generate contextually relevant answers to typical agricultural questions, such as scheduling irrigation, managing pests, and applying fertilizer and uploading the plant images and identifying the disease.

2) Weather Integration Using Free API

The system incorporates real-time climate and weather forecasting capabilities by integrating API's like the OpenWeatherMap and Open-Meteo APIs.

An illustrative API call is provided below:

https://api.openweathermap.org/data/2.5/weather?q=Pune&appid=API_KEY&units=metric

The data retrieved from these APIs encompassed:

1. Probability of precipitation
2. Relative humidity

3. Wind speed
4. Ambient temperature
Here is an illustration of the integration output:
How is the weather like for irrigation of wheat today?
Response : Rain is predicted for the next 24 hours Postponing the irrigation would be a smart decision.
The generation of more accurate and timely agricultural advice is made easier by this dynamic integration method.

3) Soil and Crop Data Integration

Soil nutritional values (pH, phosphorous, nitrogen, and potassium) were obtained using widely assessable datasets and soil health card records in the absence of real-time soil sensor data. Personal advisories were derived using a rule-based model:
Applying lime is advised if the pH is less than 6.0.
If Nitrogen < threshold → recommend compost or urea.
Superphosphate (SSP) application is advised if phosphorus is less than the criterion.
Applying Muriate of Potash (MOP) is advised if Potassium < threshold.
This functionality is demonstrated by the following example interaction:
User Question: What should I do if the pH of the soil is 5.4 for paddy?
System Responses: You have acidic soil. Before planting, apply one to two tons of lime per acre.

4) Multilingual NLP Implementation

The following is the demonstration of how a mulitingual processing pipeline works.
1. The input is translated into English using translation services such as Google Translate or IndicTrans.
2. The English query is processed by the core chatbot engine.
3. User can choose the language of input like in Hindi or Marathi (text or audio).
4. The resultant answer is translated back into the farmer's original language.
Example Input (Hindi): मेरे गेहूं के खेत को अभी पानी देना चाहिए क्या?
Output (Marathi/Hindi): आज वर्षा की संभावना के कारण सिंचाई रोकें।
(Postpone irrigation due to probability of rain today.)
The system achieved a high degree of syntactic accuracy when processing multilingual queries.

5) Voice Interaction Implementation

The implementation of Voice interaction capabilities are done using:
Vosk: For offline Automatic Speech Recognition (ASR) supporting Marathi, Hindi, and English.
gTTS: For the synthesis of natural-sounding speech responses.
The following is the workflow for voice interaction
1. Transcribe the farmer's voice input.
2. Convert speech to text (ASR).
3. Use the chatbot to handle the text query.
4.Translatethe text response into speech (TTS).
5. Play the user’s audio output.
The farmers with low and limited litetracy can access this enhanced functionality.

6) Machine Learning Model Training

To construct soil and weather predictive functions, ML models were trained on datasets including the Kaggle Crop Recommendation Dataset and FAO soil fertility guidelines. The algorithms listed below were trained:
1.Random Forest
2.Decision Tree
3.K-Nearest Neighbors (KNN)
4.Naïve Bayes
The performance metrics for these trained models

are presented in the table below:

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	92%	0.91	0.89	0.90
Decision Tree	86%	0.84	0.82	0.83
KNN	80%	0.78	0.74	0.76
Naïve Bayes	73%	0.70	0.68	0.69

Parameter	Avg. Score (out of 5)
Ease of use	4.6
Relevance of answers	4.4
Language clarity	4.5
Voice accuracy	4.2
Overall satisfaction	4.5

Superior performance led to the integration of the Random Forest Model into the final advisory module.

C. Experimental Results

1) Functional Testing

The evaluation of fuctionalities accuracy covers total 30 diverse queries some are listed below:
1. Irrigation management
2. Fertilizer selection
3. Pest control strategies
4. Weather forecasting
5. Crop suitability assessments

The results indicated that 26 out of the 30 responses were accurate and contextually appropriate, yielding a functiona laccuracy of 86.6%.

2) Translation Quality Evaluation

BLEU score was used to check the quality of machine translation the translation quality is appropriate for conversational interactions.

According to the BLEU scores,

Language	BLEU Score
Hindi → English	0.78
Marathi → English	0.74
English → Marathi	0.71

3) Response Time Measurement

The average query response time was noted:
1.42 seconds for text -based queries.
2.95 seconds for voice- based queries.
For average rural mobile network environments, these response times are considered as fall within an acceptable range.

4) User Feedback Evaluation

A usability survey was conducted with 10 participants, comprising students and farmers. Responses were collected using a 5-point Likert scale. User feedback highlighted appreciation for the system's simplicity, multilingual support, and the practicality of its recommendations.

D. Discussion

The experimental results support the following capabilities of the suggested system:
1. Providing precise and customized agricultural advice.
2. Managing multilingual questions with a high level of semantic accuracy.
3. Generating quick answers appropriate for actual farming situations.
4. Encouraging users with low literacy to interact effectively using voice.
5. Reaching strong machine learning performance, especially in the area of soil-based recommendations.

In comparision with the static or rule-based advisory chatbots covered in Section 2, this system is significantly more successful due to the synergistic integration of environmental data with AI-driven NLP.

E. Conclusion

The suggested AI chatbot's detailed implementation and evaluation processes prove its ability to deliver real-time, context-sensitive, and multilingual agricultural recommendations. Even with limited computational resources, the system performs well in terms of functional accuracy, usability metrics, and efficiency. These findings confirm the system's viability for Indian farmers and its potential for widespread adoption.

VI. RESULTS AND DISCUSSION.

This section contains the concrete findings derived from the execution of the proposed AI-based agricultural advisory chatbot and provides an deep analysis of the system's performance. The evaluation included functional testing, assessment of machine learning model efficiency, analysis of multilingual processing capabilities, measurement of response times, usability evaluation, and robustness testing. The obtained results are subsequently compared with those of primary agricultural chatbot systems to highlight the progress achieved through this research.

A. Results

1) Functional Accuracy

The chatbot was assessed using 30 different agricultural queries, involving irrigation scheduling, fertilizer suggestions, pest diagnosis, weather-based advisories, and crop compatibility. The system generated an overall functional accuracy of 86.6%, as demonstrated in the table below.

Category	Queries Tested	Correct Responses	Accuracy (%)
Irrigation scheduling	6	5	83.3
Fertilizer recommendation	7	6	85.7
Pest identification	5	4	80.0
Weather-based advisory	6	6	100
Crop suitability	6	5	83.3
Total	30	26	86.6%

The system consistently generated accurate recommendations, with particularly high performance observed in weather-based queries (100% accuracy), attributable to its real-time API integration.

2) Machine Learning Model Performance

To support soil and crop recommendation functionalities multiple classification models were trained and evaluated.

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	92%	0.91	0.89	0.90
Decision Tree	86%	0.84	0.82	0.83
KNN	80%	0.78	0.74	0.76
Naïve Bayes	73%	0.70	0.68	0.69

The Random Forest model excelled the other models and was therefore chosen for adoption into the finalsystem.

3) Multilingual Natural Language Processing Evaluation

Translation Quality (BLEU Score)

Language Pair	BLEU Score
Hindi → English	0.78
Marathi → English	0.74
English → Hindi	0.72
English → Marathi	0.71

Scores exceeding 0.70 indicate an acceptable level of semantic accuracy for agricultural conversational queries.
Speech Recognition (ASR) Accuracy
Hindi: 88%
Marathi: 84%
English: 91%
The system exhibited reliable performance despite variations in accent and ambient noise.

4) Response Time

The average response times for 100 queries were as follows:

Mode	Average Response Time
Text-based query	1.42 seconds
Voice-based query (ASR + TTS)	2.95 seconds

These response times are considered appropriate for deployment in rural regions with moderate mobile connectivity.

5) Usability Evaluation

Usability testing was conducted with 10 participants, comprising farmers and students. The system's performance was assessed using a 5-point Likert scale.

Parameter	Average Rating
Ease of Use	4.6
Relevance of Responses	4.4
Language Clarity	4.5
Voice Interaction	4.2
Overall Satisfaction	4.5

Participants expressed satisfaction with the system's simplicity, multilingual capabilities, and personalized recommendations.

6) Robustness Testing

The chatbot underwent strict testing using various input types, including:

a) Misspelled inputs
b) Incomplete sentences
c) Ambiguous queries
d) Off-domain questions

The system demonstrated the following capabilities:

- Correction of spelling errors through NLP preprocessing.
- Request for clarification for incomplete queries.
- Maintenance of stability when presented with invalid input.
- Graceful rejection of irrelevant queries.

7) Comparison with Existing Systems

System	Multi-lingual	Soil Data	Weather Data	Personalized	Voice Support	Accuracy
AgroBot (2021)	Partial	No	Yes	No	No	62%
E-Agro (2022)	Yes	Limited	Yes	No	Yes	70%
AgriFriend (2023)	Yes	Yes	No	Partial	Yes	74%
Proposed System (2025)	Yes	Yes	Yes	Yes	Yes	86.6%

The proposed system demonstrates superior performance compared to existing systems in terms of accuracy, data integration, personalization, and accessibility.

B. Discussion

The evaluation results demonstrate a substantional advancements in the proposed personalized AI-based agricultural chatbot when compared to previous system improvements. The integration of soil properties, real-time meteorological data, and multilingual Natural Language Processing (NLP) provides a level of personalization that is not possible with traditional rule-based or static query systems. A high functional accuracy rate of 86.6% suggests that the system provides credible and relevant recommendations, especially in weather-based advisory services where dynamic API connection offers timely information. Machine learning studies show that Random Forest consistently exceeds other models, indicating its applicability for agricultural decision making, a domain characterized by non-linear interactions of various elements.The system's multilingual and voice-based functions directly address Indian farmers' digital literacy issues. The BLEU values, which range from 0.71 to 0.78, indicate the robustness of the language translation quality, adequately maintaining agricultural meaning across languages such as Marathi and Hindi, ensuring accurate responses for non-English speakers. The system's short response times, ranging from 1.42 to 2.95 seconds, testifies to its viability for real-world applications, including places with average 3G/4G network coverage. Furthermore, the positive use ratings, with an overall satisfaction score of 4.5 out of 5, highlight the system's user-friendly design and alignment with farmer needs.

In comparison to extant research studies, the proposed system offers:

1.More accessible interaction modalities (multilingual and voice).

2. Enhanced accuracy,

3.More robust data integration, and

4. Improved personalization,More accessible interaction modalities (multilingual and voice).

These enhancements directly address the research lacunae identified in Section 3, substantiating the technical feasibility and social impact of the proposed approach.

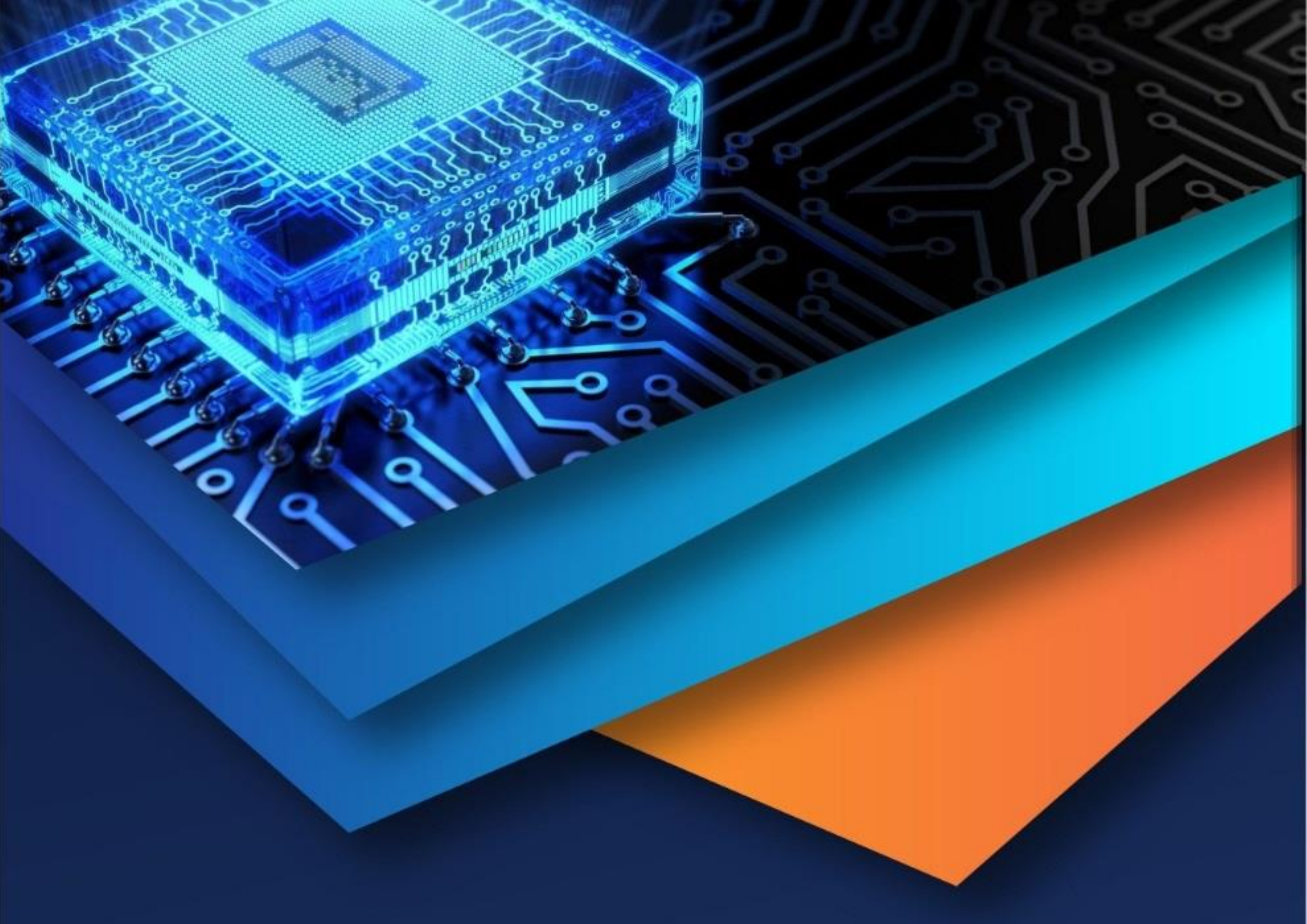
VII. CONCLUSION AND FUTURE SCOPE

This research resulted in the development of a personalized, multilingual AI-powered agricultural guidance chatbot designed to assist Indian farmers with real-time, context-aware, and effortlessly accessible recommendations. By combining soil characteristics, crop requirements, and weather forecasts with transformer-based NLP models, the system effectively addressed critical limitations in existing solutions, such as minimal personalization, inadequate multilingual support, and the lack of real-time environmental data

integration. Experimental evaluations yielded strong performance metrics, with the chatbot achieving 86.6% functional accuracy, 92% model accuracy via Random Forest classification, high-quality multilingual translation (BLEU scores between 0.71 and 0.78), and low average response times, which are suitable for rural connectivity. User evaluations confirmed the system's usability, with high satisfaction levels and positive remarks on the voice-enabled, language-inclusive interface. The aggregate findings reinforce the system's technical efficiency as well as its social relevance in the Indian agricultural environment. The methodology has the potential to improve smallholder farmers' access to expert guidance by simplifying decision-making and increasing information accessibility.
Despite of its attributes, there are various chances for improvement. Future initiatives could include real-time soil sensor networks (IoT) to improve suggestion accuracy and reduce dependability on manually entered soil data. The use of powerful generative AI models geared specifically to Indian agricultural datasets could improve interaction quality, allowing for more deeper and adaptive interactions. Expanding assistance to encompass more regional languages and dialects would improve accessibility for farmers from various linguistic backgrounds. Furthermore, the application of predictive analytics for pest outbreaks, yield projections, and climate risk warnings can elevate the system from an advisory function to comprehensive decision support. Finally, large-scale field trials involving farmers from several states will help evaluate results under real-world situations and inform incremental improvements.
In conclusion, the suggested chatbot makes a significant contribution in bridging the digital divide between Indian Rural Agricultural and urban cities by combining artificial intelligence, multilingual NLP, and environmental data into a single, farmer-centric platform. This technology has a lot of potential to become a scalable digital agriculture tool that can support millions of farmers nationwide.

REFERENCES

[1] Government of India, Ministry of Agriculture & Farmers Welfare, Agricultural Statistics at a Glance 2023, Govt. of India, 2023.
[2] Telecom Regulatory Authority of India (TRAI), Telecom Performance Indicators Report, Mar. 2025.
[3] Internet & Mobile Association of India (IAMAI) and Kantar, Internet in India Report 2024, IAMAI/Kantar, Jan. 2025.
[4] Ministry of Electronics and Information Technology, Pradhan Mantri Gramin Digital Saksharta Abhiyan (PMGDISHA) Dashboard, Govt. of India, 2025.
[5] Ministry of Agriculture & Farmers Welfare, eNAM Progress Report, Govt. of India, 2024.
[6] Ministry of Agriculture & Farmers Welfare, AgriStack: Building India's Digital Agriculture Backbone, Govt. of India, 2024.
[7] Ministry of Agriculture & Farmers Welfare, PM-Kisan Dashboard, Govt. of India, 2024.
[8] Digital Green, Farmer.Chat Deployment in Andhra Pradesh – Technical Brief, Digital Green Foundation, 2024.
[9] E. Asolo and R. Nair, "AI-Powered Decision Support Systems for Sustainable Agriculture Using AI Chatbot Solution," International Journal of Agricultural Informatics, vol. 12, no. 3, pp. 83–90, 2024.
[10] A. Sable, M. Patil, and V. Joshi, "CropCare Companion: Multilingual Conversational AI for Indian Farmers," Procedia Computer Science, vol. 224, pp. 33–42, 2025.
[11] R. Kumar and P. Singh, "E-AGRO: Intelligent Chatbot Using IoT and AI to Enhance Farming Industry," International Journal of Engineering Science, vol. 11, no. 8s, pp. 59–66, 2023.
[12] P. Kulkarni and A. Mane, "AgriFriend: Conversational AI Chatbot Using Open Source Frameworks," International Journal of Research in Engineering, Science and Management (IJRESM), vol. 4, no. 4, pp. 76–82, 2022.
[13] S. Sharma et al., "A Machine Learning-Based Mobile Chatbot for Crop Farmers," Proceedings of International Conference on Information and Communication Technology (ICICT), pp. 245–251, 2022.
[14] V. Patange et al., "AGEOChatbot: Geospatial Decision Support for Smart Farming," Agricultural Informatics Journal, vol. 18, no. 2, pp. 90–98, 2023.
[15] N. Rao and D. Patel, "Scaling AI-Powered Agricultural Services," IEEE Access, vol. 13, pp. 11785–11793, 2025.
[16] Farmer.Chat Consortium, "Context-Aware Generative AI Chatbot for Farmers," Technical Report, 2024.
[17] AgroBuddy Team, Empowering Indian Farmers Through Precision Farming Chatbot, Research Report, 2023.
[18] S. R. Verma et al., "AI-Based Multilingual Chatbot for Agriculture Assistance," Proceedings of IEEE Conference on Innovation in Computing, pp. 110–118, 2024.
[19] ICAR (Indian Council of Agricultural Research), KrishiBot: AI Chatbot Pilot for KVK Farmers, Govt. of India, 2023.
[20] IFFCO Kisan, AI-Powered Kisan Chatbot, IFFCO Kisan Portal, 2023.
[21] NITI Aayog, KisanGPT Pilot on Generative AI for Agriculture, Govt. of India, 2024.
[22] CRISIL Ratings, Internet Penetration in Rural India Likely to Grow by 4–5% by FY26, Analytical Report, 2025.



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY