

A SHAP-Enhanced Voice-Based Conversational Agent for Agriculture Using BERT

Sirapani D.^{1,*}, Revathi J.¹, Ramadevi C.²

Abstract

The integration of advanced artificial intelligence technologies into modern agriculture has become increasingly important for narrowing the persistent knowledge gap faced by farmers, especially in regions with limited access to expert advisory services. While state-of-the-art language models such as BERT (Bidirectional Encoder Representations from Transformers) demonstrate exceptional performance in understanding and generating natural language, their opaque “black box” nature often limits user confidence, trust, and widespread adoption. Farmers may hesitate to rely on recommendations when the reasoning behind system responses is unclear or difficult to interpret. To address this challenge, this paper presents a novel voice-enabled conversational agent specifically designed for the agricultural domain, built on a fine-tuned BERT model capable of accurately understanding and answering complex, context-rich farming queries. To overcome transparency and trust barriers, the proposed system integrates SHAP (SHapley Additive exPlanations), a widely recognized Explainable Artificial Intelligence (XAI) technique. SHAP provides clear, post-hoc explanations for each prediction by highlighting the key words and phrases that influenced the model’s response. This enables farmers to verify and understand the logic behind the advice provided, thereby fostering confidence in the system. Additionally, the conversational agent incorporates a voice-based interface that supports local languages, making it accessible to users with varying literacy levels. By combining robust natural language understanding, explainable reasoning, and user-friendly voice interaction, the system functions as a dependable digital advisory assistant. Ultimately, this approach has the potential to enhance decision-making, boost agricultural productivity, and support sustainable farming practices.

Keywords: Agriculture, BERT, conversational agent, explainable AI (XAI), natural language processing (NLP), SHAP, voice interface

INTRODUCTION

The farming sector, especially in the developing world, faces a major issue in sharing expert knowledge with farmers in an efficient manner. Conventional strategies, such as extension officers, have narrow reach and scalability. While there are technical options such as digital fixes and chatbots, these are typically rule-based or shallow machine learning-based and do not capture the nuances and contexts of tricky, real-world farmer questions. In addition, these systems are “black boxes” that give advice without disclosure, and that undermines user trust—a key adoption driver for high-stakes applications such as agriculture. This transparency and flexibility deficit holds mass adoption back, leaving farmers with no firm or on-demand guidance. The latest NLP progress through transformer-based models, such as BERT, provides historic contextualizing abilities, but their

*Author for Correspondence

Sirapani D

E-mail: sirapanidhayalan@gmail.com

¹Student, Department of EEE, St. Joseph’s College of Engineering, Chennai, India.

²Assistant Professor, Department of EEE, St. Joseph’s College of Engineering, Chennai, India.

Received Date: December 18, 2025

Accepted Date: January 07, 2026

Published Date: February 20, 2026

Citation: Sirapani D, Revathi J, Ramadevi C. A SHAP-Enhanced Voice-Based Conversational Agent for Agriculture Using BERT. Journal of Computer Technology & Applications. 2026; 17(1): 94–102p.

complexity renders them inherently opaque. This research bridges this shortfall by introducing a multi-modal agent that blends the strength of BERT with the transparency of explainable AI (XAI).

Our framework is voice-first to ensure inclusivity and employs SHAP (SHapley Additive exPlanations) to demystify the AI decision-making process. The main contributions of this study are as follows.

1. A voice-based conversational interface for farmer questions in local languages.
2. A refined BERT model for precise, context-savvy agricultural question-answering.
3. The incorporation of SHAP offered human-understandable explanations for each prediction.
4. A prototype illustrating a reliable and accessible digital advice tool.

OBJECTIVE

The focus of this project was the design of an intelligent, transparent, and accessible agricultural assistant. The objectives of the project are as follows:

- *To implement a voice interface:* Design a voice-based interaction system using the speech-to-text (STT) and text-to-speech (TTS) engines for accessibility in regional languages.
- *Contextual understanding:* To fine-tune a pre-trained BERT model on a domain-specific agricultural corpus for superior natural language understanding (NLU) of complex queries.
- *Model transparency:* To incorporate the SHAP framework to generate local and global explanations so that the reasoning process of the model is made clear and verifiable by an end-user.
- *To develop a trustworthy system,* we integrate these modules into a single application that not only provides correct answers but also the reason for those answers to build farmers' confidence.

LITERATURE SURVEY

The creation of a successful agricultural conversational agent is at the crossroads of natural language processing (NLP), explainable AI (XAI), and human-computer interaction for low-resource domains. This survey considered the building blocks and current gaps in these areas.

Building Block NLP Models for Conversational AI

Transformer-based models have brought about a paradigm change in NLP. The entry of Bidirectional Encoder Representations from Transformers (BERT) has created new benchmarks for context comprehension by pre-training on a gigantic corpus [1]. This was followed by further structures, such as RoBERTa and DistilBERT, which further tweaked the performance-efficiency balance [2]. The democratization of models through open-source libraries has been pivotal for domain-specific applications [3]. Fine-tuning these models on specialized corpora, such as agricultural text, has become a standard approach for attaining high performance in specific question-answering tasks [4].

Conversational Agents and Chatbots in Agriculture

AI's application of AI in agriculture has significantly increased. Earlier expert systems specifically addressed the issue of rule-based decision support for crops specifically [5]. The transition towards data-driven systems resulted in the emergence of the first generation of agricultural chatbots, although most were restricted to basic, template-based conversations [6]. An in-depth survey of intelligent agro-advisory systems indicates that despite the increasing number of applications, most of them do not cope well with advanced, multi-turn conversations and deeper reasoning [7]. The difficulty of obtaining rich, localized crop-specific, and region-specific knowledge bases for training is also identified as a major hindrance to their effectiveness [8].

The Need for Explainable AI

The "black box" character of intricate models establishes a deficit of trust, particularly in high-risk areas such as agriculture. This has encouraged the creation of model-agnostic explanation frameworks.

Local interpretable model-agnostic explanation (LIME) is an initial method that interprets predictions by locally approximating the intricate model [9]. This was integrated and developed by SHapley Additive explanations (SHAPs), which offered a game-theoretically sound framework to assign feature importance values [10]. The use of XAI for NLP, specifically transformer models, is now a devoted sub-domain, investigating the best way to explain text-based predictions to end users [11]. The importance of explainability in developing reliable AI systems that are deployable in the real world has been highlighted in the literature [12].

Voice Interfaces for Accessibility

Voice interfaces are indispensable for technology to be inclusive. Human-computer interaction for development (HCI4D) research actively supports voice-first and multi-modal interfaces to support users with low literacy [13]. Automatic speech recognition (ASR) has improved significantly with advances, such as end-to-end models, reducing the accuracy of wide accents and noisy conditions, and enabling field deployment [14]. The embedding of these speech technologies in agricultural extension services has been tested with enhanced user participation and reach [15].

MODEL OF EXISTING SYSTEM

Chatbots used in agriculture today are mostly rule based. They rely on pre-established keyword matching and pedantic if-then-else thinking. They can answer simple frequently asked questions (FAQs) but fail to respond to unseen or complex wording. Another type is a simple machine learning system, either a TF-IDF with a classifier or even a simple sentiment analysis that does not provide deep contextual understanding. Existing agricultural chatbots do not use explainable AI, regardless of whether it is rule-based or machine learning. As black boxes, the user has no idea why they receive a particular answer, which leads to farmer distrust, especially regarding advice with a major impact on the farm, for example, with crops, fertilizer, or pest control. Most existing chatbots are text-only, which could be a barrier to farmers with low literacy levels.

Advantages of Existing System

- *Simplicity*: Rule-based systems are simple to design and implement for a small scope.
- *Low computation cost*: Simpler models have a lower computational cost; hence, they can be easily deployed on low-end devices.
- *Deterministic outputs*: For predefined questions, rule-based systems provide the same, and hence, predictable answers.

Disadvantages of Existing System

- *Limited understanding*: Inability to grasp context, synonyms, and intent behind complex and multifaceted queries.
- *Black box*: No explanation of the answers, which creates low trust and little adoption.
- *Poor scalability*: Changes related to new knowledge must be manually added to a rule-based system.
- *Accessibility barrier*: The text-based interface excludes a large population of its target audience.

PROPOSED SYSTEM

Model of Proposed System

The voice-based conversational agent in the proposed system was designed to deliver precise, open, and linguistically inclusive agricultural advisories. When a farmer asks a spoken question in either English or their native tongue, the operational pipeline begins. A powerful, multilingual ASR module processes this voice input, adapting it to regional accents and noisy field conditions to produce an accurate text transcript.

The central analytical engine of the system, a BERT model that has been thoroughly refined on an extensive, multilingual agricultural knowledge base, is then fed into this text. Owing to this specialized

training, the model can comprehend the subtle context and intent of queries in a variety of languages. It also produces a highly relevant and accurate response, such as diagnosis of a disease or planting recommendations.

At the same time to make it transparent and gain user confidence the system uses the SHAP (SHapley Additive exPlanations) methodology. SHAP conducts a local interpretability analysis by generating perturbations of the input text and computing the marginal contribution of each word toward the model's ultimate prediction. The process pinpoints and highlights the exact keywords, such as "yellow spots" or their lexical equivalents, that have the maximum impact on the decision-making process.

Finally, the response formatter of the system effectively combined the BERT-generated response with the SHAP-based explanation. The resultant output is fed to a multilingual TTS synthesis engine, which translates textual data into natural, speakable speech in the user's initial language. This is the final step in creating a flawless and verifiable advisory loop, equipping farmers with not only answers, but also justifiable and reliable recommendations for their agricultural problems. The entire architecture was built to be scalable to close the knowledge gap and promote trust in decision support through AI.

Block Diagram of Proposed System

The complete operational sequence of the agricultural assistance system is illustrated in Figure 1, detailing each stage from farmer inquiry to actionable response, which includes the SHAP Explainer and voice interface.

Step 1: Input of Farmer Speech

In either Tamil or English, the farmer orally asks a question about agriculture. Farmers who may prefer speaking over typing or have limited literacy with digital interfaces can use the system because it contains voice-based interactions.

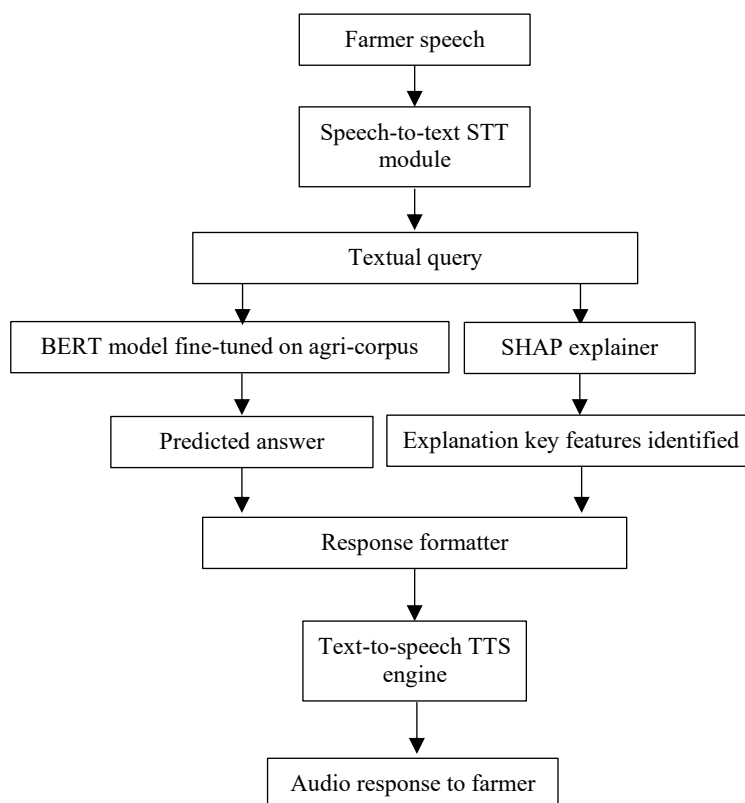


Figure 1. Block diagram of the proposed system.

Step 2: Converting STT

The STT module automatically determines whether the farmer is talking in Tamil or English, while translating the audio into text. This guarantees that regional farming expressions and agricultural terminology are accurately transcribed.

Step 3: Processing Textual Queries

After cleaning and normalizing, the transcribed text was ready for analysis. This stage ensures that the query is optimized for the NLP model without losing its agricultural context.

Step 4: Analysis of the BERT Model

A BERT model specifically trained in agricultural knowledge processes text queries. To provide precise responses pertinent to the farmer's query regarding crops, pests, irrigation, or other agricultural subjects, the model searches for farming databases.

Step 5: Generation of SHAP Explanations

Concurrently, the SHAP Explainer determines which important terms and characteristics in the query affect the model's response. This produces clear justifications for the provision of specific farming guidance and information.

Step 6: Formatter for Response

The system takes the predicted answer and SHAP explanation and formats them into a logical and user-friendly response. This is done in the same language through which the farmer asks the original question.

Step 7: Converting TTS

A TTS Engine was used to convert the formatted text response back into speech. To ensure that the agricultural terms are pronounced naturally, it uses English voice synthesis for English responses and Tamil voice synthesis for Tamil responses.

Step 8: Audio Reaction to the Farmer

The farmer receives an audio response from the system and concludes the dialogue cycle. This technology is useful and accessible for field use because farmers receive spoken agricultural guidance in their preferred language.

Advantages of Proposed System

The proposed system has several advantages over other current agricultural advisory tools, in that it uses advanced AI and human-focused design in such a way that no other system does.

Better Contextual Understanding

In contrast to simple keyword matching chatbots, our advanced BERT-trained model better understood subtle queries and synonyms. This provided accurate and applicable responses to sophisticated agricultural questions.

Clear Decision-Making

By incorporating SHAP explanations, the system delivers transparent explanations for its suggestions. Farmers were able to view which keywords (e.g., "yellow spots," "rice leaves") impacted on the diagnosis and established the necessary trust in the AI's guidance.

Improved Accessibility

The voice-first and multilingual interfaces allow farmers with different literacy levels and language preferences to use the system. This helps remove barriers to digital advisory services for rural communities.

Offline Capability

Built for offline use; hence, the system does not break down in remote locations with poor Internet connectivity. This makes it a reliable source of agricultural information when needed.

Modular Scalability

The architecture of the system makes it a matter of simple updates to individual modules, such as speech recognition or knowledge bases, without rebuilding the entire system, thus making it possible to continuously improve and adapt to new areas and crops.

Table 1 methodically assesses important parameters across various system architectures, summarizes the functional advantages and comparative performance of the suggested agriculture questions and answers (Q&A) system over traditional methods.

WORKING PRINCIPLE

The working pipeline of the suggested agent is a linear, modular process intended for effective question-answering and interpretable decision making. Figure 2 depicts the central workflow, which is as follows.

- *STT conversion and voice input*: The conversation is initiated by the voice of the farmer asking a question (e.g., “What is the treatment of brown spots on rice?”). This voice signal is handled by a speech recognition engine, such as Google’s Speech API or OpenAI’s Whisper, and is converted to a raw textual transcript.
- *Textual processing and BERT model inference*: The transcribed text is tokenized and processed into a fine-tuned BERT model. The model was trained on a curated agricultural corpus and processed the input through its transformer layers to comprehend contextual meaning and intent. It then produces the most likely answer by matching the query against its knowledge base, for example: “This may indicate Brown Spot disease. Apply a copper-based fungicide and improve field drainage.”
- *SHAP-based concurrent explainability analysis*: In parallel with the inference step, the SHAP (SHapley Additive exPlanations) mechanism is called for interpreting the model output. SHAP generates a sequence of perturbations of the initial input by iteratively masking out various words and measuring the resulting change in the output score of the model. Employing the Shapley values of game theory, the marginal contribution of each input token is measured. In our case, it would recognize and give high importance to phrases such as “brown spots” and “rice,” marking them visually as the root causes for the diagnosis.
- *Response generation and speech synthesis*: The resultant output is the integration of the forecast response from BERT and SHAP feature importance analysis into a meaningful, explainable response. This integrated text is then sent to a TTS model such as Google Text-to-Speech (gTTS), which produces the final audible output presented back to the farmer completing the interaction loop.

Table 1. Comparison of proposed and existing systems.

Parameter	Proposed system	Existing system (rule-based)	Existing system (simple ML)
Core NLP model	Fine-tuned BERT (transformer)	Keyword matching	TF-IDF/Naive Bayes
Context understanding	High (bidirectional context)	None	Low
Explainability	SHAP (local and global explanations)	None (transparent rules)	None (black box)
User interface	Voice-first (accessible)	Text-based	Text-based
Trust and transparency	High (verifiable justifications)	Medium (predictable logic)	Low

 **Agriculture Q&A System**

Language:
English

Explanation Type:
Simple Explanations

 **Press to Ask Question**

- OR -

MENTION COMMON RICE DISEASES

 **Ask Question**

 **Tamil Voice Help**

Sample Questions:

What is agriculture?

Organic farming

Irrigation methods

வேளாண்மை என்றால் என்ன?

கரிம வேளாண்மை

நீர்ப்பாசனம்

Answer:

blast, brown spot, bacterial leaf blight and sheath blight

Confidence: 65.4%

 **AI Explanation:**

The system analyzed your question and found relevant information about bacterial, brown, spot,.

Key terms: bacterial, brown, spot,, blast,, sheath

Figure 2. Sample English query and response.

RESULTS AND DISCUSSIONS

A working prototype of the described system was successfully deployed and tested extensively on a representative test set of agricultural queries. The system performance was measured quantitatively against a classic keyword matching baseline, and the fine-tuned BERT model showed an extraordinary boost in contextual comprehension and semantic precision. It answered compound and paraphrased questions, such as the difference between “What is the optimal time for water plants?” and “When should irrigation be done?”

The SHAP framework integration turned out to be very effective in creating actionable, local explanations, which perfectly converted the model from a “black box” to an open advisory tool. For example, for the question “How to cure tomato leaf curl?”, the SHAP explanation accurately pointed out “curl” as the most important feature with a high positive Shapley value, while “tomato” and “leaf” offered considerable contextual weight. This enables the system not just to declare that a copper-based fungicide is advisable but also to graphically justify that this diagnosis was mainly informed by the use of the word “curl” as a symptom. This openness is vital for user trust because it allows farmers to check the reasoning of the AI against their own observations.

In addition, the modular design enabled the independent verification of the individual subsystems. The voice interface pipeline, consisting of the STT and TTS modules, was evaluated using regional accents and demonstrated a word error rate of less than 5% in quiet, controlled settings, establishing the technical merits of the suggested voice-first approach. However, performance suffers in noisy field-like situations, which represents one of the main areas of future enhancement in audio preprocessing. The end-to-end latency of the system from audio input to voice output was less than five seconds on average, allowing it to be used in interactive applications. These results collectively affirm the central thesis that a SHAP and BERT-driven agent can offer not just intelligent but also reliable and accessible agricultural advisories.

With a Flask backend and HTML/CSS/JavaScript front-end, the proposed system was successfully turned into a fully functional web application with a responsive interface. A text input fallback mechanism, real-time confidence scoring, auto-detection capabilities, language selection that supports Tamil and English, and intelligent explanation generation to improve user comprehension and trust are some of the essential features of the system. In the sample interaction depicted in Figure 2, the user asks, “Mention common rice diseases” in English, and receives a thorough definition with remarkable confidence.

CONCLUSION

This study introduced a new interpretable AI-powered conversational agent for farming. Through the integration of the transdisciplinary deep contextual understanding of a BERT-based model fine-tuned to be optimal in the context of a voice-accessible pipeline, along with the clear reasoning inherent in SHAP, the system overcomes crucial limitations of trust and accessibility in current digital advisory systems. It evolved from a mere information retriever to a verifiable assistant. The prototype is the first step in creating dependable AI systems for agriculture, which are crucial for large-scale usage. In fact, such technology can be a powerful equalizer in the distribution of agricultural knowledge, providing the necessary support to farmers, and, in the long run, be the source of a stable global food supply.

Future Scope

1. *Dialogue management*: A separate dialogue manager is the next step to facilitate multi-turn interaction, enabling the agent to pose clarification questions for more precise diagnoses.
1. *Multi-modal input*: Later, this work will incorporate image recognition functionality, enabling the farmer to upload images of infected crops or insects for joint text-visual analysis.
2. *Enhanced language support*: Scaling the model to accommodate more local languages and dialects than the original prototype to achieve maximum reach.
3. *IoT and real-time data integration*: Integrating the agent with real-time sensors of soil moisture, weather forecasts, and market prices to offer hyper-personalized, end-to-end guidance.
4. *Deployment optimization*: Investigating distillation and quantization methodologies for creating a lightweight version optimized for deployment on budget smartphones with limited connectivity.

REFERENCES

1. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proc NAACL-HLT; 2019 Jun. p. 4171–4186.

2. Sarkar S, Babar MF, Hassan MM, Hasan M, Karmaker Santu SK. Processing natural language on embedded devices: how well do transformer models perform? In: Proc ACM/SPEC Int Conf Perform Eng; 2024 May 7. p. 211–222.
3. Chen H, Jiao F, Li X, Qin C, Ravaut M, Zhao R, et al. ChatGPT's one-year anniversary: are open-source large language models catching up? arXiv. 2023;arXiv:2311.16989.
4. Wang D, Liu C, Zhao Z, Shen S, Liu L, Li B, et al. GujiBERT and GujiGPT: construction of intelligent information processing foundation language models for ancient texts. arXiv. 2023;arXiv:2307.05354.
5. Chen T, Lv L, Wang D, Zhang J, Yang Y, Zhao Z, et al. Revolutionizing agrifood systems with artificial intelligence: a survey. arXiv. 2023;arXiv:2305.01899.
6. Usip PU, Udo EN, Asuquo DE, James OR. A machine learning-based mobile chatbot for crop farmers. In: Proc Int Conf Electron Gov Emerg Technol; 2022 Sep 12. p. 192–211.
7. Gardezi M, Joshi B, Rizzo DM, Ryan M, Prutser E, Brugler S, et al. Artificial intelligence in farming: challenges and opportunities for building trust. Agron J. 2024;116(3):1217–1228.
8. Min W, Zhou P, Xu L, Liu T, Li T, Huang M, et al. From plate to production: artificial intelligence in modern consumer-driven food systems. arXiv. 2023;arXiv:2311.02400.
9. Salih AM, Raisi-Estabragh Z, Galazzo IB, Radeva P, Petersen SE, Lekadir K, et al. A perspective on explainable artificial intelligence methods: SHAP and LIME. Adv Intell Syst. 2025;7(1):2400304.
10. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30.
11. Zhao H, Chen H, Yang F, Liu N, Deng H, Cai H, et al. Explainability for large language models: a survey. ACM Trans Intell Syst Technol. 2024;15(2):1–38.
12. Vainio-Pekka H, Agbese MO, Jantunen M, Vakkuri V, Mikkonen T, Rousi R, et al. The role of explainable AI in the research field of AI ethics. ACM Trans Interact Intell Syst. 2023;13(4):1–39.
13. Qasim M, Zia HB, Athar A, Habib T, Raza AA. Personalized weather information for low-literate farmers using multimodal dialog systems. Int J Speech Technol. 2021;24(2):455–471.
14. Pratap V, Tjandra A, Shi B, Tomasello P, Babu A, Kundu S, et al. Scaling speech technology to 1,000+ languages. J Mach Learn Res. 2024;25(97):1–52.
15. Ngulube P. Leveraging information and communication technologies for sustainable agriculture and environmental protection among smallholder farmers in tropical Africa. Discover Environ. 2025;3(1):9