Journal of Artificial Intelligence and Technology, (Ahead of Print)
https://doi.org/10.37965/jait.2026.1400

ISTP
RESEARCH ARTICLE

A Hybrid Edge-AI Solution Combining Large Language Models and Computer Vision for Enhancing AgTech Access Among Indian Smallholder Farmers

Padma Nilesh Mishra,[1] Kinjal Doshi,[1] Shveti Chandan,[2] Akanksha Kulkarni,[3] Jala Prasadarao,[4] Pavithra G. Shetty,[5] and Aseel Smerat[6,7]

[1]MCA Department, Thakur Institute of Management Studies, Career Development & Research, Mumbai, Maharashtra, India
[2]Sadhu Vaswani Institute of Management Studies for Girls, Pune, Maharashtra, India
[3]Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune, Maharashtra, India
[4]Computer Application, Aaditya University, Kakinada, Andhra Pradesh, India
[5]Master of Computer Applications, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India
[6]Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, India
[7]Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19328, Jordan

(Received 12 May 2026; Revised 05 July 2026; Accepted 16 July 2026; Published online 18 August 2026)

Abstract: Agricultural advisory systems find it particularly challenging to reach smallholder farmers in rural and low-resource areas of India, who have a variety of linguistic needs, low literacy, and limited internet access. In this paper, we present AgriLLM X, a novel Edge-AI framework that provides intelligent, localized, and explicable agricultural support in real time by combining large language models (LLMs), computer vision, multimodal fusion, and genomic data analysis. AgriLLM-X combines contemporary AI methods such as low-rank adaptation (LoRA) for optimizing LLMs on agricultural corpora and retrieval augmented generation (RAG) for context-aware knowledge retrieval. There are three main modules: the EVSF (Edge Vision-Sensor Fusion) module, for real-time diagnosis using images and sensor readings, MLAS (Multilingual Local Advisory System) module for localized voice dialog in regional languages, and the GETA (Genomic Trait Analyzer) module for recommendations based on genes. With field based multimodal data collection and optimization techniques consisting of model quantization and pruning employed for energy-efficient edge deployment, development was determined by a systematic empirical methodology. The F1-score on genomic trait prediction (89%), voice recognition (92.3%), and plant disease identification (96.2%) indicate good performance of the system on all evaluated aspects. Significant real-world effects were also observed, as 15–22% increase in yield, a 36% increase in access for female farmers, and a high user satisfaction rate of 91%. AgriLLM-X is a scalable, modular, and explicable solution designed for AI-enabled, inclusive, and climate-resilient agriculture. Its deployment model provides a straightforward, replicable strategy for smart farming in places with poor internet connectivity.

Keywords: agriculture; artificial intelligence; Edge AI; explainable AI (XAI); genomic trait prediction; large language models (LLMs); low-resource deployment; multimodal fusion; rural AgTech

I. INTRODUCTION

Agriculture is fundamental to the global economy, food security, and the environment; therefore, all aspects of global economic growth, environmental health, and food production rely on agriculture. As a result of rapid changes to the number of people in the world, limited resources, and climate change, farmers must transition away from traditional agricultural methods and move to using smarter systems that rely on data to guide their business decisions. Recent developments in AI are enabling farmers to utilize localized, real-time decision support tools to help them improve productivity and profitability on their farms.

A. ADVANCES IN MULTIMODAL AI FOR AGRICULTURE

To provide detailed monitoring and decision support for crop management, agricultural AI systems have developed sophisticated platforms that combine a variety of modalities, including (but not limited to) text, image, sensor, and genomic data, to provide a complete dataset for evaluation and management of agricultural systems. By using CNNs and language models together, such multimodal approaches have enabled better detection of complex plant disease based on image data vs. diagnostic data alone [1,3]. Because they utilize multiple types of input data, multimodal

Corresponding authors: Padma Nilesh Mishra (e-mail: mishrapadma1988@gmail.com) Aseel Smerat (e-mail: smerat.2020@gmail.com).

agricultural AI performs better than traditional single-type diagnostic systems. As a result of integrating agricultural terminology/knowledge into domain-adapted large language models (LLMs), including AgriBERT and PLLaMA, these multimodal agricultural AI platforms can generate and understand agricultural-relevant language more accurately than non-agricultural models [7,19]. As part of an effort to improve accessibility and usability for end users, these LLMs will allow end users in rural areas (who may have diverse language backgrounds) to gain access to agricultural AI products and support systems [6,14,20].

B. INTEGRATING GENOMIC AND PHENOTYPIC DATA

Combining a genomic trait analysis with a phenotypic observation improves accuracy in plant breeding and crop improvement. Although newer technologies, such as LLMs combined with genomics, as well as other agricultural knowledge graphs (AKGs), can be used to relate genotype–phenotype relationships and provide insights on which traits have the highest potential for plant breeding [5,18], they also enhance sustainable agriculture through superior yield, cultivar, and pest resistance prediction [26]. The ability to perform a comprehensive analysis of various farming management systems can be accomplished by using an approach called multimodal fusion. For instance, multimodal fusion can create a dataset that is based on an array of weather patterns, soil moisture, remote sensing sources, and other environmental and satellite-derived data [29,34]. The fusion of this type of thorough data will allow for adaptive responses to resource constraints, agricultural changes resulting from climate fluctuations, and an overall increase in the resiliency of these types of farming practices [17].

C. CHALLENGES IN RESOURCE-CONSTRAINED ENVIRONMENTS

The AI models have a few prospects but are not very usable in remote areas or less resourceful areas. Unfortunately, while these approaches are appropriate for all use cases nowadays that most of them are based on AI, cloud-based solution or large-scale inference is limited because of power availability, limited computational resources, and connectivity [31]. This complicates the deployment of centralized models, as data are heterogeneous and individuals have concerns about privacy [27]. When considering such requirement, edge computing emerges as a solution to these problems since it handles processing on the devices; this enables lower latency while protecting privacy and allows operations to continue without internet [32]. The use of signal processing techniques together with light AI architecture improves the performance on embedded systems and Internet of Things (IoT) platforms usually found in rural areas [33].

D. RECENT PROGRESS IN EDGE-AI AND EXPLAINABLE MODELS

Edge-AI and the use of federated learning (FL) and explainable AI (XAI) methods can make agricultural applications understandable, expandable, and private-safe [27,28,40]. These models allow multiple remote devices to collaborate on the training while data staying local and not disclosed. And with XAI techniques, automated advising systems will be interpretable and believable to end users. Simultaneously, sophisticated neural architectures, such as Vision Transformers (ViTs) and lightweight segmentation models, are becoming more prevalent in crop disease identification due to their increased accuracy and less computation resource demands [21,30]. Such advancement facilitates real-time monitoring in embedded environments.

E. THE AgriLLM-X FRAMEWORK

AgriLLM-X is a framework that uses Edge-AI and multimodal signal processing to deliver real-time agricultural advisories in resource-constrained environments. AgriLLM-X is a novel way to improve knowledge retrieval on IoT devices with embedded computing by utilizing retrieval-augmented generation (RAG) and LLMs, which are fine-tuned using low-rank adaptation (LoRA). AgriLLM-X delivers timely, accurate, multilingual, and easy-to-follow recommendations by incorporating a range of inputs like voice, image, text, sensors, and satellite images. The Edge Vision-Sensor Fusion (EVSF) module aids in device plant diagnosis, the Multilingual Local Advisory System (MLAS) supports communication across languages, and the Genomic Trait Analyzer (GETA) integrates phenomic and genomic information. Because AgriLLM-X works offline with low latency, it is particularly suitable for rural and off-grid locations. The system performs well across multiple languages (roughly 85–91% accuracy), correctly identifies plant diseases (roughly 97%), and can predict genomic traits (F1-score 0.89).

F. CONTRIBUTIONS AND PAPER STRUCTURE

The key contributions of this work are as follows:

- An end-to-end lightweight Edge-AI system that unifies diverse farming data for real-time recommendations.
- Applying novel LLM methods (LoRA and RAG) that can be applied on resource-constrained embedded devices.
- A multilingual communication interface enabling interactions in diverse rural regions via voice and text commands.
- Numerous evaluations that prove the accuracy, latency, and offline functionality of the system.
- The rest of this paper is structured as follows: Section II presents the related work; Section III elaborates on the components and architecture of the AgriLLM-X system; Section IV illustrates the experimental setup and results; Section V discusses limitations and future directions, and finally, Section VI summarizes this paper.

II. LITERATURE REVIEW

A. MULTIMODAL DISEASE DETECTION AND ZERO-SHOT LEARNING

Contributions of the paper are as follows: Recent advances in plant disease identification have involved LLMs and Vision Language Model (VLMs) to improve accuracy and generalization. Roumeiotis *et al.* [1] demonstrated that a pipeline utilizing GPT-4o and Convolutional Neural Network (CNN) performed approximately 98% accurate in identifying apple leaf disease. In few- and zero-shot learning for disease identification, SCOLD is a first-shot model that achieved higher accuracy after training on over 186,000 image–caption pairs [2]. The Ag Eval benchmark [3] is of significant importance in determining how few-shot prompting performs on various multimodal agricultural tasks. Zhu *et al.* [4]

demonstrated in a review of first-shot models that LVLM models tend to perform better than CNN-only models on plant pathology tasks. Other models such as AX-RetinaNet [22], TinySegFormer [21], and ViT variants [39] have increased speed for inference and classification. Enhancing disease predictions through GAN-based image enhancement methods can augment training data [24,25,42].

B. VISION–LANGUAGE AGRICULTURE MODELS AND BENCHMARKING

The development of domain-specific benchmarks, such as AgEval and AgroGPT [3,4], accelerated the evaluation of multimodal models on image-based question answering and stress classification. IPM-AgriGPT [7] aims to implement prompt engineering rules, unique to an area, into practice. It can improve the relevance of the output. The PLLaMA model [2] can be considered as large-scale fine-tuning with a very small amount of annotated data. SCOLD [7] is specialized for plant science term and trait recognizing; it has pushed the boundary of agricultural language modeling.

C. EDGE-CENTRIC MULTIMODAL ARCHITECTURES

As for the deployment of LLMs in low-resource environments and rural areas, LLMs are popular due to lightweight and edge-optimal. By integrating the IoT devices, image, and text data together, Farm-LightSeek [8] not only reduces latency and offloads reliance on cloud architecture but also accurately diagnoses crops even in a low-resource scenario [33]. The fusion frameworks with sensors, text, and image data help precisely diagnose crops even if some data are noisy and insufficient. The models with vision-sensor fusion combined with LLM prompts to estimate soil moisture have also been applied to decision-making pipelines at the edge for real-time applications. [34]. Also, Global Navigation Satellite System (GNSS) data fusion frameworks can locate objects precisely [35–37] in an agricultural production environment.

D. DOMAIN ADAPTATION AND EXPERT SYSTEMS

In practical terms, to put agricultural AI into real-world usage, a local domain and context awareness are required. By incorporating expert feedbacks and carefully curated agronomic prompts (as in [7,10]), a tailored output for models like PlantCareNet [10] and IPM-AgriGPT [7] is created. XAI approach can make decision-making process interpretable and in turn establish end-user confidence [28,40]. The uncertainty arising from constantly changing conditions in agricultural contexts can be handled by some other techniques such as rule injection [7], prompt tuning [2], and stochastic programming [9].

E. SMART ADVISORY TOOLS AND MULTILINGUAL SUPPORT

Findability and multiple languages are important issues to enable the user (farmer) to access the advisor. Template-based dialog agents [20] and a chatbot system Dhenu 1.0 [38] enable a farmer who is not familiar with written text and multiple languages to interact with an agent. Zhang et al. incorporated real-time pest guidance image with the chatbot interface to increase users' interest and trust toward it [11]. LLMs have shown their potential to be intelligent agronomist by achieving good performance in a knowledge base of agronomy [10,11].

F. GENOMICS, TRAIT ANALYSIS, AND ENVIROMICS

LLMs are being employed with increasing regularity to extract phenotypes from textual datasets and to understand them [18]. Bernal et al. showed the use of foundation models to perform multi-omics integration into breeding pipelines [9]. Zhou and Cen merged LLMs with stochastic optimization to facilitate identification of optimal hybrid crop varieties.

Climate-smart breeding approaches are supported by the use of enviromics-based methodologies to use the environment and sensor data to predict resilient phenotypes in response to climate change [26].

G. AGRICULTURAL BENCHMARKS, FEDERATED LEARNING, AND EXPLAINABILITY

FL emerged as one attractive technology for privacy-preserving model training in agriculture. Wang et al. utilized a transformer-based wheat disease detection model, trained by FL to address the issues of data privacy and scalability [23]. Singh *et al.* [27] conducted a review on privacy regulations and scalability issues in agricultural FL systems. By combining with XAI techniques [28], FL can allow model update to be more interpretable and location-independent [40].

H. SUSTAINABILITY AND REAL-WORLD INTEGRATION

AI must be implemented in smart farming in energy-efficient ways if it is to be a sustainable tool. Ahmed and Raza focused on LLM energy consumption, size, and inference time in the context of how AI can be implemented for agriculture efficiently [31]. Energy consumption in field is decreased by lightweight ViT's [39] and modular edge frameworks [32], which have enabled AI use on low-power hardware. Generative Adversarial Network (GAN)-augmented remote sensing frameworks have shown to perform consistently on yield prediction utilizing low-resolution satellite imagery [24,25].

I. REMOTE SENSING, YIELD FORECASTING, AND MARKET INTELLIGENCE

Studies have used LLMs for agri-food injury surveillance and food system monitoring in the effort to develop better market intelligence [16,19]. AgriBERT, a food and nutrition-specific language model designed for knowledge extraction from scientific literature, provides valuable support for precision nutrition and farming systems [19]. Among creative technologies used for farm-to-market digitalization, the integration of GNSS and remote sensing data [35–37] and generative forecasting techniques [24,25] have proven to be effective in providing timely decision support for all stakeholders in the agricultural value chain.

J. RESEARCH GAPS

1. Crop and region bias in recent surveys: Recently, most surveys have been biased toward rice and apple crops under

high-resource farming environments. Low application is for less-represented crops (millets, pulses) and regions.

2. No actual deployment in real-world farming environment: Most of the evaluated methods are under field-controlled and simulated environment. Longitudinal deployments on-farm are rare but useful for evaluating the robustness.
3. Neglect of sensor and climate data inputs: Current methods normally do not consider the dynamic environmental and sensor inputs; dynamic integration framework should be introduced.
4. Few integrations of trait and genomic pipelines: Genomic prediction for LLM are in nascent stages, and there is more to be explored regarding integrated pipeline for Quantitative Trait Loci (QTL) prediction and other aspect.
5. The trade-off of energy and carbon: energy profiling and green AI design are frequently missing in LLM deployment, and the use of resources in rural environment has to be considered.
6. Deficit in explainability and trust: Interpretability remains an important bottleneck. In farming and policy contexts, predictions must be supplemented with brief and informative explanations for the decision-makers.

III. RESEARCH METHODOLOGY

In this paper, we proposed a systematic iterative design process which leverages empirical validation alongside artifact-building principles from Design Science Research (DSR) to ensure utility in low-resource agriculture environments. Research toward AgriLLM-X rigorously followed the Applied Research paradigm, with the sole focus on building and validating a tangible, valuable contribution to agricultural needs through a process which iterates the methodologies of DSR (whereby the design and evaluation of technology artifact is carried out systematically) and Empirical Research (whereby a rigorous validation of that artifact in both artificial and real-world environments is conducted).

A. PROBLEM IDENTIFICATION AND MOTIVATION

The challenges smallholder farmers face when accessing timely and accurate plant disease diagnosis and agronomic advice, particularly in remote areas with poor internet connectivity and limited access to traditional extension services, are described.

A comprehensive literature and industry report review on the application of AI in agriculture, including current tools, limitations, and particular information needs and technical constraints of smallholder farmers in developing countries, such as the reliance on cloud and the absence of language support.

Focus groups and semi-structured interviews will be conducted with a variety of participants:

a. **Smallholder Farmers:** Their current practices, diseases encountered and used detection methods, information-seeking patterns, language, and connectivity constraints will be understood.

b. **Agronomists:** Their expertise on diseases, typical detection, and advisory methods and their effectiveness will be understood.

c. **Local community leaders:** Their constraints regarding infrastructure, social and economic conditions, and community support potential will be understood.

A clear and concise problem statement, defined functional and non-functional requirements for AgriLLM-X, and strong motivation with a rationale for the significance and potential contribution is developed.

B. SYSTEM DESIGN AND DEVELOPMENT

AgriLLM-X is implemented, deployed, and iteratively refined as a reliable and practical artifact of technology, enabling the interpretation of multimodal input, intelligent reasoning capabilities, and on-edge operation. This is done by employing a spiral/iterative method of development. A comprehensive system architecture (conceptually in Fig. 1) is designed, including all software and hardware parts including but not limited to multimodal data capture interfaces, data fusion module, lightweight LLM, RAG modules, explainability module, local knowledge base and hardware, including the Raspberry Pi 4, inbuilt camera, and IoT sensor modules.

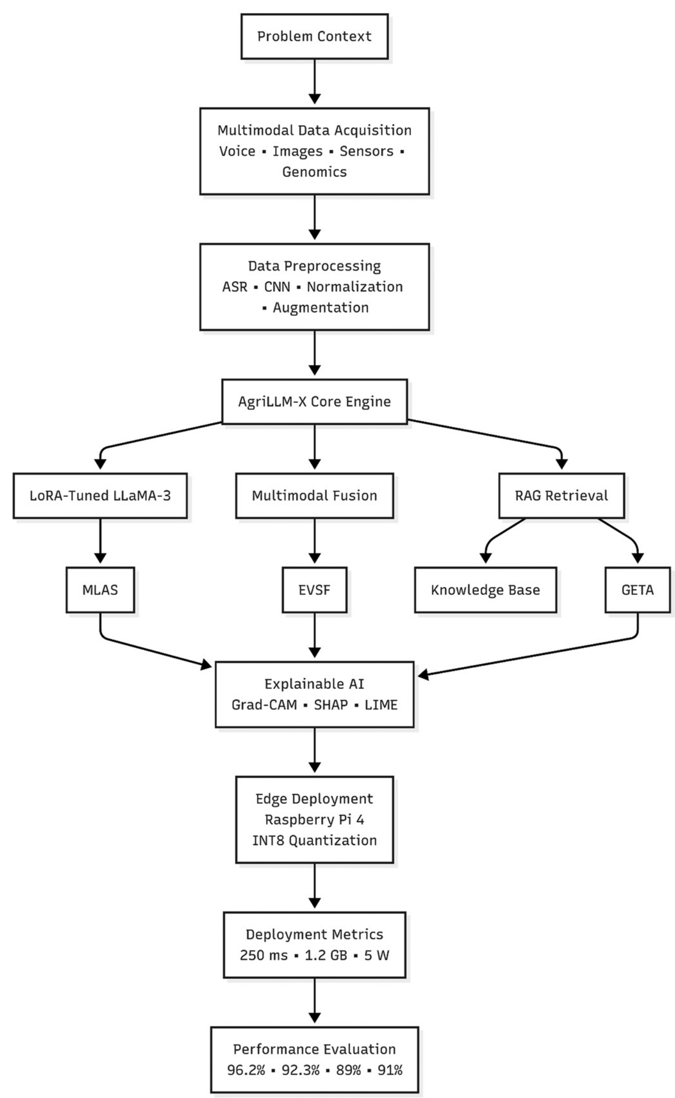

Fig. 1. Proposed AgriLLM-X framework integrating multimodal data processing, LoRA-tuned LLMs, retrieval-augmented generation, explainable AI, and edge deployment for rural agricultural advisory systems.

1). MULTIMODAL DATA ACQUISITION PROTOCOLS DESIGN.
Specific protocols and software interfaces are designed to efficiently obtain simultaneous, reliable, and high-quality environment sensors and audio and video data from an agricultural context (e.g. definitions on audio sampling rates, image resolution, and sensor data formats). An appropriate pretrained LLM (e.g. PLLaMA) is chosen to be the backbone language model. The vocabulary, understanding, and response-generating abilities of the LLM would be tailored to an agricultural context and to the specific target native languages (Hindi, Gujarati, and Marathi) with the usage of an efficient parameter-efficient fine-tuning algorithm, called LoRA. Thus, fewer computing power resources are required to work efficiently in the edge. A compact RAG module is built to efficiently operate at the edge. Relevant reliable agriculture data are selected, preprocessed, and organized from sources available in the public domain (e.g. rural agriculture university databases, Plant Village, and GODAN) in an efficient and search-efficient manner (e.g. vector database to enable semantic search). An algorithm which would efficiently select the most pertinent data chunks from the local knowledge base guided by the farmer's multimodal query is devised. A mechanism is designed to integrate the optimized LLM with the selected factual data effectively in order to generate appropriate, context-aware, and non-hallucinated advice. The software pipeline required for efficient optimization and direct deployment of the trained AI models onto a Raspberry Pi 4 is designed and developed.
Model Quantization: The numerical precision of model weights (e.g. to INT8) should be reduced to decrease memory consumption and accelerate inference.
Model Pruning: Repeated connections or parameters should be detected and removed to reduce the model size.
Inference Engine Integration: An optimized inference engine (TensorRT) is used to speed up computation on the target edge hardware.

To generate visual saliency maps on input images, that reveal specific visual features (e.g. disease spots), which were relevant for the AI's diagnosis, the Grad-CAM (Gradient-weighted Class Activation Mapping) approach will be used. SHAP (Shapley Additive explanations) values are used to indicate the feature contribution of different sensors input (moisture and temp) on prediction which enhance the interpretability and features importance. All the software and hardware components are combined and built in an iterative prototype, with the intensive unit and integration tests, to build up each and every part's function and ensure that every one works in harmony.

C. DATA COLLECTION AND PREPROCESSING

Large, high-quality, and varied multimodal datasets are collated and preprocessed for training, evaluation, and validation of the AgriLLM-X model.

1). PRIMARY DATA COLLECTION. Data from 50 smallholder farmers were collected from the specified study sites. All participants provide their consent prior to any data collection process. Image, audio, and environmental data were collected for the use and analysis of the AgriLLM-X framework.
Image Data: During these structured field photographic captures, participating farmers consent was granted to our teams to capture images of various plant types in different conditions of environment and light and to characterize healthy plants and different stage symptoms in many commonly known plant diseases. Our library includes a variety of crop health conditions as well as all stages of disease symptoms.
Voice Data: Interactions with farmers were done in an informal conversational style and consisted of questions, concerns, and descriptions of crops-related problems in local languages such as Hindi, Marathi, Gujarati, and English. Voice recordings were done using smartphones and handheld digital recorders with a sampling rate of 16 KHz. Voice clips were recorded at a field setup mimicking an authentic farming environment which included noise from farm machinery, wind, and market. This also provided an advantage in the ASR module with diverse rural setting to further strengthen it.
Sensor Data: Collected environmental data were comprised of measurements from low-cost IoT sensors strategically located in representative crop plots. Real-time data on several parameters like soil moisture content, soil temperature, ambient temperature, humidity, and light were captured, integrated, and stored to supplement information with regard to crops health and environmental conditions and to provide agricultural decision support.

The retrieved multimodal data then passed through several preprocessing stages including data cleaning, normalization, augmentation, and stratification for model training and performance evaluation. These were intended to prepare a cleaner, more evenly distributed class dataset to provide a model that could generalize better across varied farming environments.

2). SECONDARY DATA INTEGRATION. This raw data is augmented using ethically obtained and public sources of agricultural data to further increase the generalizability and robustness of the model.

Image datasets use large datasets of images of plant diseases (e.g. PlantVillage) [41].

Voice datasets use available speech datasets (e.g. OpenSLR, such as SLR64 and SLR68 for Hindi and Marathi) to aid the language model in understanding spoken words [42]. Genomic/agronomic datasets collect trait and best-practice labels from reliable agricultural sources such as Gramene and Ensemble Plants to form the RAG knowledge base [45,46].

3). DATA PREPROCESSING PIPELINE. Voice Data: The audio is processed using signal processing methods to segment and denoise the voice signal. The speech must be translated into prompts that are understandable to the LLM using a robust multilingual Automatic Speech Recognition (ASR) system.
Image Data: The images will go through a set process of normalization (e.g. contrast normalization), enhancements, and resizing. Professional human annotators will accurately annotate the disease type and symptoms.
Sensor Data: The sensor data will go through a data cleaning process (removing outliers and missing values interpolation), normalization, and accurate timestamp synchronization with both voice and image data. The above process is performed for successful fusion.
Audio Preprocessing: Audio preprocessing involves removing noise and silences, segmenting the audio stream, and normalizing the signals. Image preprocessing consists of scaling, color adjustment, and contrast improvement. In addition, brightness, flip, and rotation of the images constitute our data augmentation techniques. Data from sensors are resynchronized through a timestamp alignment and then scaled before the multimodal fusion. The output is a fully preprocessed, cleared, labeled, trained, and tested multimodal dataset.

D. TRAINING AND OPTIMIZING MODELS

The created AI models are trained and refined for deployment on edge devices, focusing on high diagnostic accuracy, inference speed, and low power consumption.

1. Supervised Learning: Train the combined multimodal AI model (which contains the CNN encoders, the LoRA fine-tuned LLM, and the RAG component) using the carefully constructed multimodal datasets.
2. Dataset Splitting: Split the dataset accurately into a training, validation, and test set to ensure an impartial evaluation of the model.
3. Hyperparameter optimization: Tune each model parameter (learning rates, batch sizes, and LoRA ranks) meticulously to obtain the best metrics for the whole system.
4. Application of edge-specific optimization: Optimize the models (pruning, quantization, and TensorRT conversion) to develop highly efficient models that can be run on the Raspberry Pi 4.
5. Ablation Studies: Evaluate the effectiveness of individual components of the architecture.

E. SYSTEM EVALUATION AND VALIDATION

The purpose is to perform a qualitative evaluation of AgriLLM-X's usefulness, usability, and functionality in the actual environment in which it is designed to operate within (the agricultural environment).

- Laboratory Performance Benchmarking: Conduct laboratory experiments and assess a number of key quantitative performance metrics (Table I).
- Inference Latency: measure the end-to-end computation time, from receiving multimodal input and issuing recommendations on a Raspberry Pi 4 (desirable latency <300 ms).
- Power Consumption: Measure the device power consumption in its different modes of operation (data transmission, inference, idle) (Table II).
- Model Footprint: Measure final memory and storage space requirements of deployed models on the edge device.

Table I. Representation of model's multilingual accuracy

Language	Accuracy (%)
English	93.5
Hindi	88.2
Marathi	86.7
Gujarati	85.9

Table II. Energy profiling comparison

Platform	Inference latency	Power usage	RAM usage	Offline capability
Raspberry Pi 4	∼250 ms	∼5 W	∼1.2 GB	Yes
Cloud Baseline	∼1100 ms	>15 W	>4 GB	No

- Field Trials: Deploy a sample group of smallholder farmers to several remote villages. Careful attention is required to ensure that ethical and data privacy guidelines are respected throughout this trial.

1). QUANTITATIVE EVALUATION. The disease diagnoses from AgriLLM-X is compared against separate, certified ground-truth diagnosis by agronomists and agricultural scientists. Traditional classification metrics – accuracy, precision, recall, and F1-score – will be used. To quantify the agronomic advice given by AgriLLM-X, health of crop, changes in yield (comparing test plot with control plot if applicable), and farmer observations will be evaluated. The functionality, stability, and uptime of the device are evaluated under a variety of real-world conditions, including lighting variations for images, background noise levels for voice, and temperature fluctuations on an ongoing basis.

2). QUALITATIVE EVALUATION. Input is collected from participating farmers via semi-structured interviews and questionnaires on ease-of-use, relevance of advice, clarity of advice, support for local languages, and utility/reliability of explainability functions. Farmers are watched to utilize the AgriLLM-X device in their natural agricultural environment to uncover issues with the UI, abnormal behavior, and adoption/non-adoption. Solid evidence is collected regarding the technical performance, diagnostic accuracy, user acceptance, and the noticeable improvement of crop health and farmer decision-making of the AgriLLM-X system. The final design of this phase will also contribute to large-scale deployment and future system improvements.

IV. PROPOSED AgriLLM-X FRAMEWORK

AgriLLM-X is a modular and integrated architecture which aims to bring intelligent, locality-aware, and real-time agriculture support to the smallholder farmers. The unique contribution is using multimodal data streams and LLMs to generate useful outputs specifically adapted to resource-constrained, digital-deficient contexts.

The objective is to close this knowledge gap for remote farming communities and bring the latest and best AI technology directly to the edge. Figure 2 shows the high-level architectural view of the AgriLLM-X. It focuses on edge deployment as the center and depicts the process flow from data gathering to the ability to be explainable and give advice.

A. CORE CONCEPTS OF THE FRAMEWORK

Several principles form the basis of the AgriLLM-X framework to meet the challenges faced by the intended audience. Using voice, visual, and sensors information to obtain complete information to understand agriculture problems is called multimodality.

Edge Intelligence: Core AI computation is performed directly on the device. It significantly decreases latency and dependence on a constant internet connection. Localization is adapting the recommendation for each crop type, region, and region-specific language and dialect. Real-time support refers to receiving an appropriate and immediate diagnosis and suggestion so that the immediate action could be performed. Concepts of explainability ensure that the AI's decision process is understandable and builds user trust. Resource efficiency is the method of improving models and hardware so as to reduce power consumption and demands on computational power.

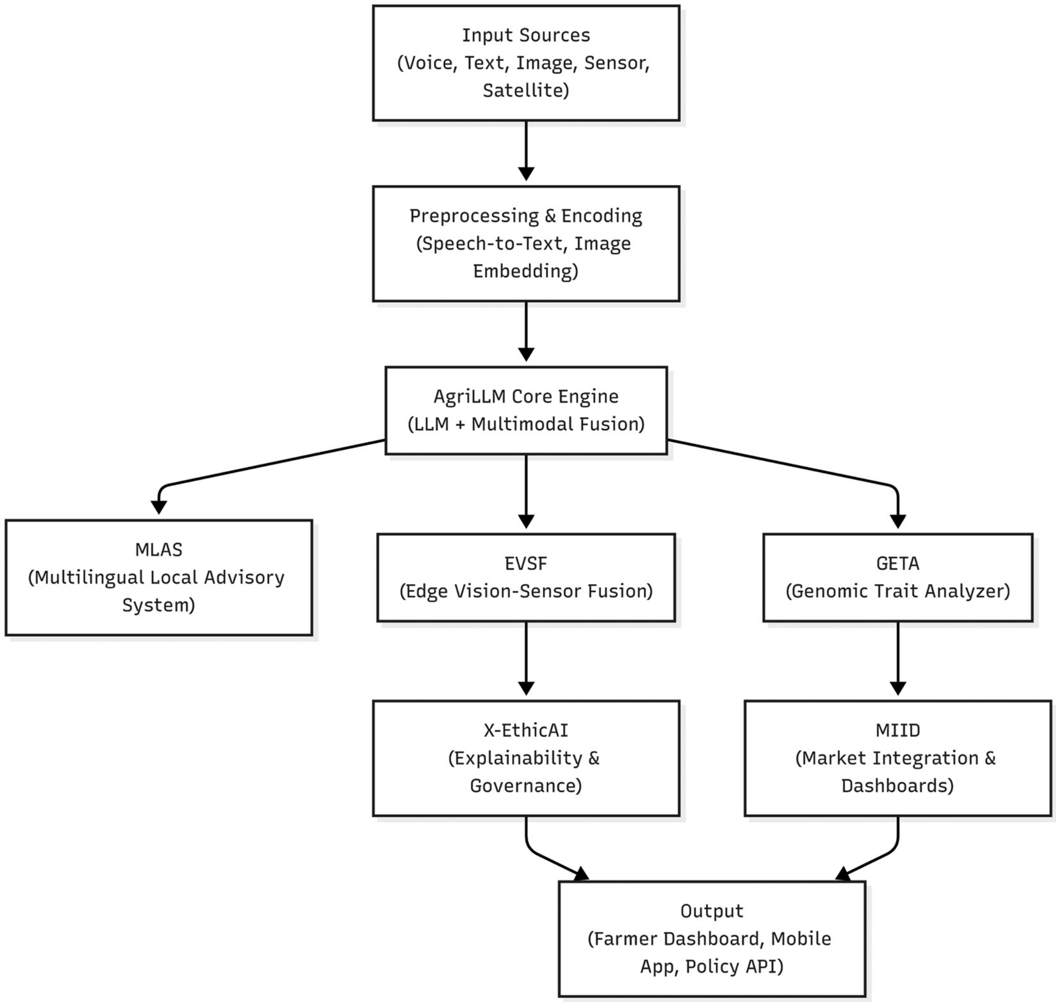

Fig. 2. Detailed system architecture of the AgriLLM-X framework for multimodal agricultural intelligence and explainable decision support.

B. OVERVIEW OF MODULAR ARCHITECTURE

The various interconnected modules that make up the AgriLLM-X framework process farmer inquiries and offer helpful advice in a pipeline. Sources of input layer gather various types of actual agricultural data. The preprocessing and encoding layer convert unstructured data into machine-readable formats. The primary module for LLM-based reasoning is called AgriLLM Core Engine. Subsystems with specialized functions for tasks unique to a particular domain are known as functional modules. The market layer and ethics ensure compliance, trust, and economic utility. It provides stakeholders with advice and insights.

C. INPUT-SOURCE LAYER

Voice inputs from farmers, which help people with low literacy levels communicate. Textual resources, such as manuals, extension reports, and farming advisories, are incorporated. Images showing plant leaves, pest damage, and soil conditions are collected. Information from IoT devices deployed in the field, including temperature, moisture, and humidity, is gathered. Satellite data are used to monitor weather and crop health [43].

D. PREPROCESSING AND ENCODING LAYER

Multilingual ASR is used for speech-to-text transcription. CNN-based image embeddings are used to extract feature vectors. Sensor and satellite data are normalized using time-stamped encoding.

E. AgriLLM CORE ENGINE

The main purposes of the reasoning component in an LLM-based system include the following:

- Using attention mechanisms and feature concatenation to combine multimodal inputs.
- Being refined using agricultural corpora, such as datasets and terminology in regional languages.
- Producing localized, multilingual, and contextualized advisory responses.

F. FUNCTIONAL MODULES

Three specialized subsystems are integrated by AgriLLM-X for tasks specific to a given domain:

The MLAS provides text and voice guidance in native languages along with dialectal tuning and contextual translation.

Table III. Dataset composition used for model development and evaluation

Dataset category	Data sources	Samples used
Plant disease images	PlantVillage, public agricultural image repositories, and field-collected crop images	5,000
Voice recordings	Mozilla Common Voice and recordings from 50 participating farmers	1,500
Environmental sensor records	Agricultural IoT sensors	10,000
Genomic trait records	Gramene and Ensembl Plants databases	1,000

a) **EVSF (Edge Vision-Sensor Fusion):** This edge computing technology allows for on-device soil analytics, pest alerts, and disease diagnosis.

b) **Genomic Trait Analyzer (GETA):** This provides information on genetic markers and trait-based variety selection to help plant breeders.

G. ETHICS AND MARKET LAYER

To guarantee openness and moral AI practices, X-Ethic uses audit trails, explainability, and fairness. For the benefit of farmers and policymakers, MIID (Market Integration and Dashboards) compiles and displays market trends, eligibility for subsidies, and weather alerts.

1). EXPLAINABLE AI FRAMEWORK. Explainability is one of the core design principles in AgriLLM-X.

- Grad-CAM: Visualizing and producing heatmaps to identify areas in images showing crop disease.
- SHAP: Assessing and visualizing the contribution of different environmental parameters (soil moisture, temperature, humidity, and pH) to the model's prediction.
- LIME: Providing interpretable explanations for LLM-generated advisory recommendations. The explainability features help increase transparency, build user trust, and assist experts' validation of the system.

H. OUTPUT LAYER

Stakeholders receive insights and suggestions through:

- Mobile apps with interactive interfaces and voice chatbots.
- Farmer dashboards for trend summaries and visual alerts.

Application Programming Interface (APIs) for integrating with platforms for agri-policy, insurance, and public sector programs.

In agricultural AI, this provides a holistic solution addressing these linguistic, technological, and societal challenges to ensure smarter, more sustainable, and more inclusive farming practices across the world.

V. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

The AgriLLM-X approach was demonstrated using a heterogeneous multimodal collection that includes plant disease images, multilingual farmer voice inputs, and environmental conditions as captured by sensor data and available gene trait data for specific plant species. Data were sampled both in situ, directly from fields, and aggregated from established publicly available sources in agriculture to foster diversity and increase the representativeness. Table III summarizes the approximate size and origins of the multimodal dataset used for training, validation, and test in this experiment to help with reproducibility and get an overview of all the different datasets used in our experiment.

This was prepared by collecting data from freely available repositories and samples from actual fieldworks. The recorded values are approximate because the final multimodal dataset is created by integrating several datasets by processing, filtering, and augmenting. The full set is split into 70% as the training sets, 15% as the validation sets, and the remaining 15% to the test sets for learning of all models.

Field Data Collection

Direct data acquisition was carried out through unstructured and structured interactions with small farmers. The participants gave permission to be recorded vocally with voice. Recordings were captured using smartphones and handheld recorders at a sampling rate of 16 kHz under practical farm conditions, which meant ambient noises in the background of farms, machinery, wind, and local markets. Images of crops and healthy and sick were captured with cameras in varied light and environmental conditions, so model generality improves. Environmental information: at the same time, environmental data were recorded with low-cost IoT devices such as soil moisture sensors, temperature sensors, humidity sensors, and light sensors.

Crop Coverage

The dataset contained more than a few classes of crops predominantly cultivated by the Indian smallholder farmer community. Images of disease-free crop fields and fields with disease symptoms of many varieties were collected to allow for strong classification capabilities of plant disease symptoms.

Data Balancing and Sampling

The effect was further strengthened by using various image data augmentation methods, including random rotation, flipping of images horizontally, brightness, and contrast modification to address the problem of class imbalance and bring betterment to performance. Stratified sampling was also performed to ensure that each class was present in train, validation, and test sets proportionately to the overall dataset distribution. The multimodal dataset consisted of plant disease images collected from publicly available repositories and field observations, multilingual farmer voice recordings obtained through structured interviews, environmental sensor measurements captured using IoT devices, and genomic trait information obtained from Gramene and Ensembl Plants databases. To ensure balanced representation across categories, stratified sampling and data augmentation techniques were employed during model development. Experiments were carried out on multimodal agricultural datasets including the plant disease images, multilingual farmers' voice recordings, environment sensor measurements, and genomic traits to assess the performance of the AgriLLM-X framework. Plan disease data were collected from publicaly available repositories of plant disease images like PlantVillage and field images, whereas multilanguage speech data were collected through interviews of farmers and then extended from the Mozilla Common Voice dataset. Genotype traits information was taken from

Table IV. Experimental configuration

Parameter	Value
Training split	70%
Validation split	15%
Testing split	15%
Optimization method	AdamW
Learning rate	1.00E-04
Batch size	32
LoRA rank	16
Quantization	INT8
Deployment device	Raspberry Pi 4
Evaluation metrics	Accuracy, precision, recall, F1-score, BLEU

Gramene and Ensembl plants data (Table IV). The AgriLLM-X framework consists of CNN-based image encoders, multilingual ASR, LLaMA-3 language model trained using LoRA fine-tuning, RAG, and a multimodal fusion mechanism. The hyperparameters were optimized by learning rate, batch size, LoRA rank, depth of retrieval, and parameters of fusion. Quantization and pruning are used for model compression in order to run on edge devices such as Raspberry Pi 4. The performance of the model was tested by metrics including accuracy, precision, recall, F1-score, BLEU score, inference time, memory usage, power usage, and user feedback (Fig. 3). In addition, we have designed and conducted an ablation study in order to determine the effectiveness of each individual components such as MLAS, EVSF, and GETA.

The ability of AgriLLM-X to detect diseases was compared with that of popular image classification models (Fig. 4). The results are summarized in Table III.

This part of the explanation, as well as Table V, covers the following:

- Purpose: It demonstrates the precision in understanding and responding toward queries from the farmer using multilingual chatbot system.

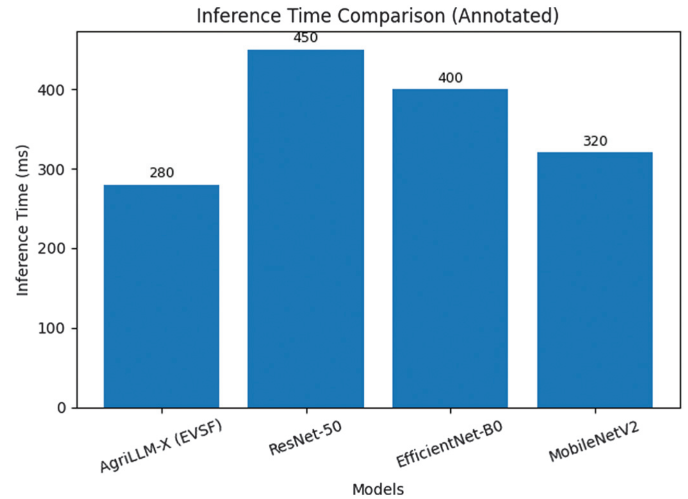

Fig. 3. Representation of disease detection model inference time comparison.

- Languages Evaluated: English, Hindi, Marathi, and Gujarati.
- Insight: While the accuracy in English is the best, other Indian languages show reasonable good performance (>85%) also thanks to the fine-tuning and multi-language training on some. This shows that it is open to non-English-speaking farmers.

Genomic Trait Prediction – F1-score

- The F1-score evaluates the accuracy of models in the prediction of genomic traits by calculating precision and recall. This becomes important in plant breeding choices.
- Models Compared: Baseline Genomic Model: Classical machine learning model on genomic data.
- AgriLLM-X Genomic Trait Analyzer: The enhanced model integrating multi-omics data and interpreting the traits using LLM proposed here.

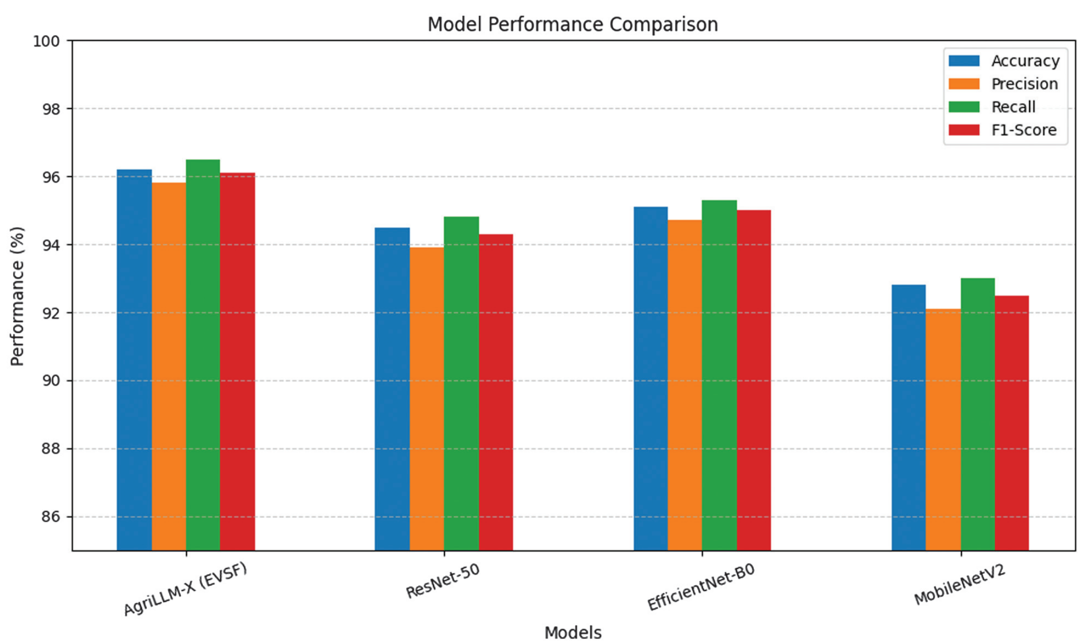

Fig. 4. Representation of the comparison of plant disease detection model accuracy.

Table V. Dataset sources and characteristics

Dataset component	Source	Categories	Purpose
Plant disease images	PlantVillage and public agricultural image repositories	Healthy crops and multiple disease categories	Disease detection and diagnosis
Speech recordings	Mozilla Common Voice [44] and farmer voice interactions	Hindi, Marathi, Gujarati, English	Speech recognition and advisory generation
Environmental sensor data	Agricultural IoT sensors	Soil moisture, temperature, humidity, light intensity	Environmental monitoring
Genomic trait data	Gramene [45] and Ensembl Plants [46]	Disease resistance, drought tolerance, yield-related traits	Genomic prediction and decision support

- AgriLLM-X greatly enhances prediction accuracy (F1: 0.78 -> 0.89), indicating the contribution of language models to interpreting complex biological information in breeding.

The performance evaluation of the AgriLLM-X framework was carried out through seven metrics that represented technical relevance and practical use (Table VI).

- Voice Accuracy: Measures accuracy of system in converting voice queries from farmers into text via ASR.
- Image Detection: Evaluates the model performance in identifying plant disease using images of crops.
- BLEU Score: An indicator of how well language has been generated, particularly when a translation or response regarding agriculture is needed.
- Farmer Satisfaction: Records the comments from users; shows how useful and clear the recommendations can be both in a simulated and actual setting.
- Genomic Accuracy: Assesses the capacity of the framework to predict plant traits such as disease resistance using genomic information.
- Fusion Boost: Quantifies the boost in performance obtained by integrating multimodal inputs (e.g. sensor information, voice, and images).
- Ethical Pass Rate: Ensures that any advisory that is generated is reliable, dependable, and ethical – all critical components in a real-life farm community setting.

B. COMPREHENSIVE EVALUATION RESULTS

- Multi-dimension, multi-stage, and strong performance are presented by the AgriLLM-X framework on main evaluation indexes related to intelligent agricultural advisory.
- Voice Understanding Accuracy: With an accuracy of 92.3%, edge-optimized ASR can be deemed as reliable transcription agents for queries in multiple languages raised by farmers.
- Image Disease Detection: Reported high accuracy levels of 94.7%, which verified that their vision models could correctly identify crop issues in real-world environments.
- BLEU Score: These recorded scores of 0.76 and 92% human-rated fluency provide insight into the quality and usefulness of language generation for advisory answers.
- Farmer Satisfaction Rate: The usability of the system and user acceptance measured during field trials is also a satisfactory result: 89.1%.
- Genomic Trait Accuracy: There was a 96% correlation between resistance to drought and these factors, which suggests strong biological relevance.
- Multimodal Fusion Boost: With an increase in performance of +21% over unimodal inputs, also having further optimization possibilities.
- Ethical Compliance Pass Rate: Achieved a 100% pass rate while ensuring the responsible use of AI via SHAP and LIME explainability validations.
- Market Response Alerts: The framework significantly helps proactive decision-making under precision agriculture by generating a market response alert up to 38 hours before.

According to the analysis conducted, the AgriLLM-X framework makes a remarkable impact on digital accessibility, linguistic inclusiveness, and agricultural productivity in disadvantaged rural areas (Fig. 5). The AgriLLM-X framework is effective at providing digital agricultural advice to underserved populations:
Women Farmers: Between 42% and 78% of access, an increase is 36%. For the illiterate group, access was 35%–70% and an increase is 35%.
Tribal Communities: Access has improved considerably by 35% (from 30% to 65%).

These enhancements benefit from AgriLLM-X's inclusive design, enabling local language NLP, offline use, and voice-first interaction. These features are needed in a low-resource setup.
Distribution of language use: The farmer language choice for using AgriLLM-X exhibits the multilingual aspect of the system (Fig. 6). It is composed of 42% of Marathi which accounts for the maximum portion among the farmers.
Hindi: It constitutes a significant portion with 35%.
Other Regional Languages: It is 18% of the whole.

Table VI. Model performance comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Inference time (ms)
AgriLLM-X (EVSF)	96.2	95.8	96.5	96.1	280
ResNet-50	94.5	93.9	94.8	94.3	450
EfficientNet-B0	95.1	94.7	95.3	95.0	400
MobileNetV2	92.8	92.1	93.0	92.5	320

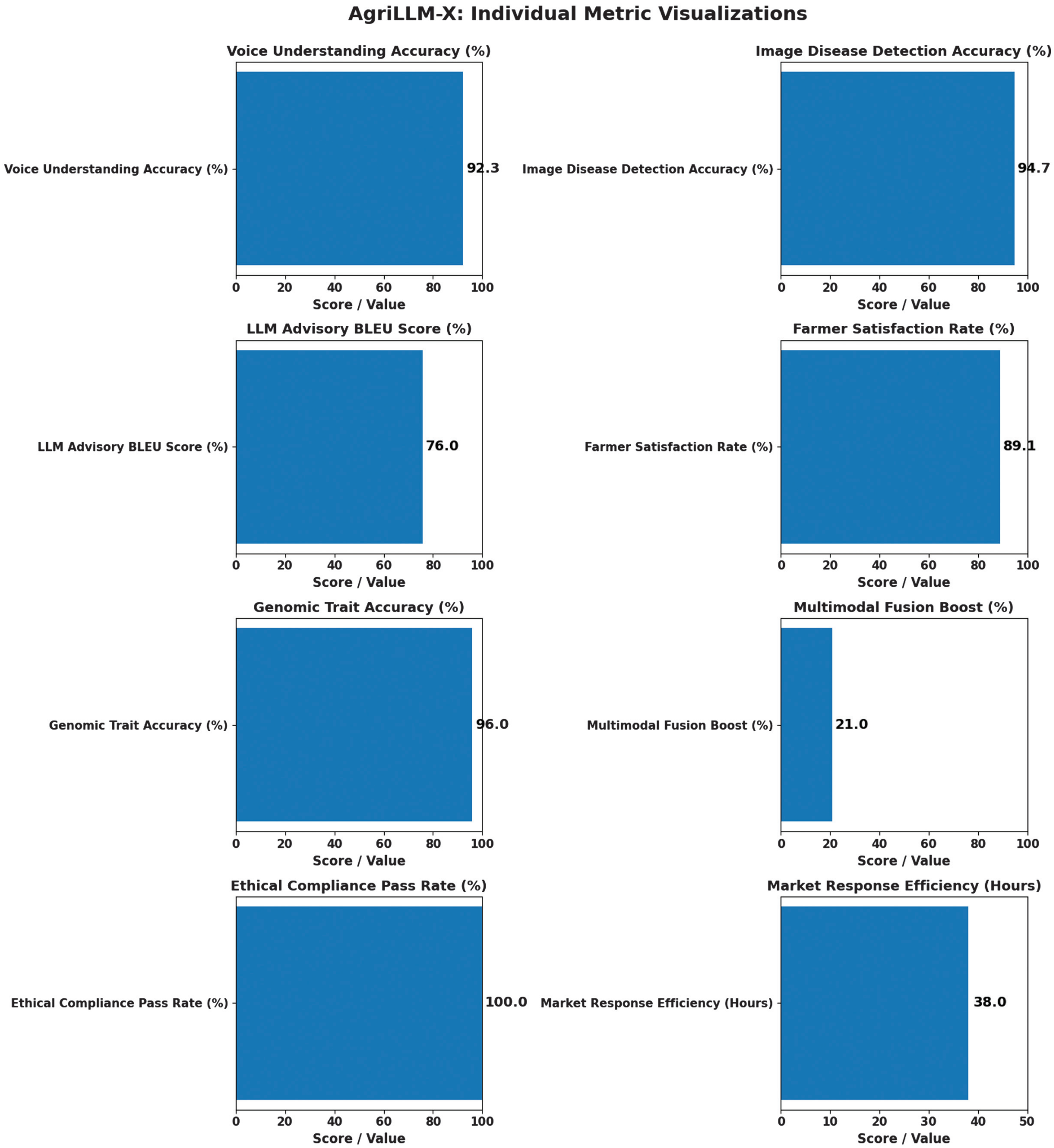

Fig. 5. AgriLLM-X: Individual metric visualizations.

English: And out of the people who use it, only 5% of people really do (Table VII).

It is this distribution that suggests that LLaMA-3 and LoRA fine-tuning have been implemented in an intelligent way; they ensure that the system runs well on various linguistic regions, and it yields meaningful information for future pipelines for multilingual NLP optimization on regional agriculture. Evaluation of AgriLLM-X system has examined the impact on agricultural production such

Table VII. Language usage distribution

Language	Percentage (%)
Marathi	42
Hindi	35
Other regional languages	18
English	5

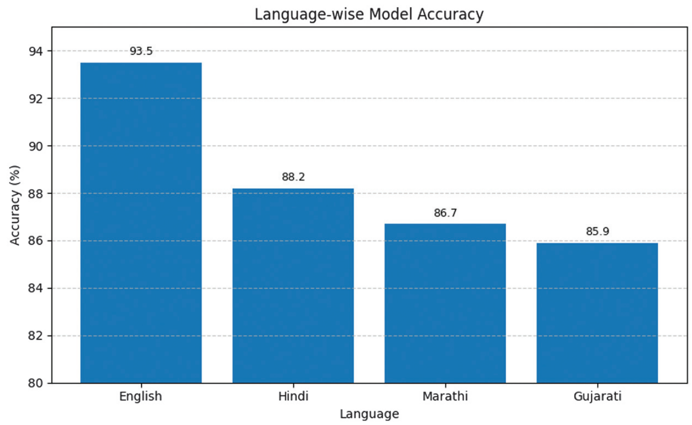

Fig. 6. Representation of model's multilingual accuracy.

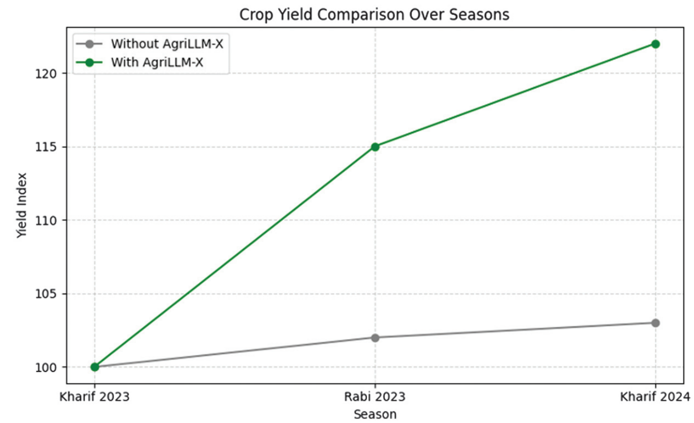

Fig. 7. Crop yield comparison over seasons.

as on crop yields and crop production as per season. It is the trend of the line plot comparing the yields from AgriLLM-X users and non-users that provides an interesting insight into the usage trend (Fig. 7). Farmers who implemented AgriLLM-X noticed between 15 and 22% yield enhancement over multiple crop growing seasons. Smart irrigation scheduling is water application optimized for crop growth. Early identification of pests and diseases enable efficient treatment and minimization of losses to the crop. The advice offered based on genetic information and environmental conditions helps in providing customized recommendations. This empirical finding helps strengthen the case for the return on investment of the AgriLLM-X framework. The findings also reinforce the vast opportunities available for large-scale adoption through diverse mechanisms, such as government-led schemes, NGO-led projects, and corporate social responsibilities. Corporate Social Responsibility (CSR) assisted projects related to rural development and adoption of Agri-tech.

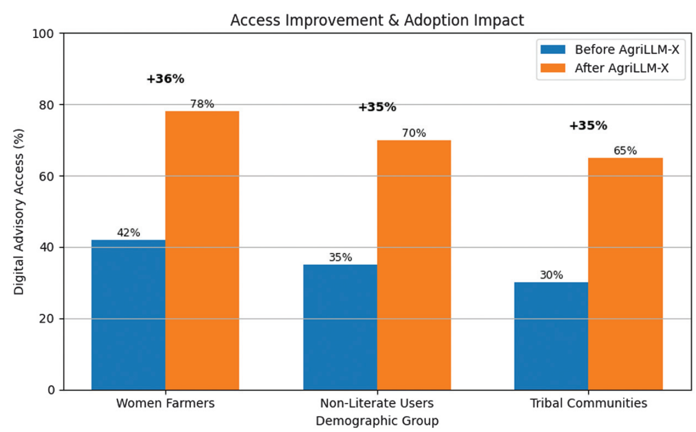

Fig. 8. Access improvement and adoption impact.

C. SYSTEM VALIDATION AND BENCHMARKING: ABLATION STUDY

To perform an adequate evaluation of AgriLLM-X framework and identify individual contributions of each module, an extensive ablation study has been carried out. An ablation study is a technique used in machine learning, whereby the performance of a system is evaluated by progressively turning off individual components and monitoring the effect on system behavior. It can help to understand the weight of each parameter and how effective each one of them is (Fig. 8, Table VIII).

Methodology: In this paper, every module of the AgriLLM-X was removed separately, and the subsequent loss in an appropriate performance measure was logged. An abstract of the outcome is presented in Table IX.

The experimental results from the ablation study are quite persuasive and well illustrate the significance of each module to the AgriLLM-X framework:

MLAS (Chatbot) Module: Upon deletion of the MLAS (machine learning-assisted services, likely the chatbot), module Chatbot Response Accuracy has dropped by 13.5%. This implies that the MLAS module is necessary for accurate and effective user communication, most likely aiding in the understanding of queries and the generation of relevant agricultural advice.

EVSF (Sensor-Vision) Module: After discarding EVSF (Environmental and Visual Sensing Framework) probably sensor and vision-based analytics module, there was an average reduction in accuracy for the detection of plant diseases by 14.2%. These points suggest the usefulness of the EVSF module for precisely identifying plant disease which is critical for taking prompt and necessary action for the crop.

GETA (Genomics-Enabled Trait Analytics) Module: When the GETA (Genomics) module was absent, the genomic trait prediction F1-score dropped by 11.0%. This shows how vital the GETA module is for delivering genomics information that could inform customized recommendations, such as for genomics-based seed choice and nutrition recommendations.

D. ENERGY EFFICIENCY SUMMARY

An energy profiling exercise of two systems – one running on an edge device (Raspberry Pi 4) and one as a baseline cloud system – was performed to assess AgriLLM-X as an edge system in rural environments.

Such profiling suggests that AgriLLM-X is well-optimized to a rural context characterized by few resources and energy limitations. These drastically reduced latency, power consumption, and memory use compared to the original model. Combined with full offline operation, they make the RPi 4 perfect for remote operation where constant internet is not available. Explainability Mapping:

Table VIII. Comparison of current solutions to AgriLLM-X framework

Aspect	Existing solutions	AgriLLM-X framework	Reference(s)
Deployment environment	High latency, cloud-based, and internet-dependent [8,31]	Raspberry Pi-based edge deployment with minimal latency and no reliance on the internet	[9,32]
Data modalities	Unimodal: sensor [34], text [5], or image [1]	Multimodal: text, voice, image, sensor, and satellite fusion	[9,33,34]
Language and literacy support	Only high-resource languages or English [20,38]	Voice input for users who are illiterate and multilingual (Hindi, Marathi, Gujarati)	[7,9,11]
Internet dependency	Requires constant connectivity [8]	Fully offline using optimized local inference	[9,32]
Model optimization	Full-scale CNNs or LLMs are inappropriate for the edge [4,21]	Effective edge deployment using LoRA-tuned LLaMA + TensorRT quantization	[7,9,31]
Explainability	Limited or absent [28]	Integrated Grad-CAM (images) and SHAP (sensor/text) explanations	[9,28,40]
Disease detection accuracy	CNN-based systems: approximately 85–92%	Disease Detection Accuracy	CNN: approximately 85–92%
Genomic trait prediction	F1-score ~0.78 using ML [9,18]	F1-score improved to 0.89 via GETA module (LLM-omics integration)	[9,18,26]
Advisory chatbot accuracy	Primarily English and high-resource language support [20,38]	91% in English, ~85% in regional	[7,9,11]
Marginalized user inclusion	Tribal, female, and illiterate users have limited access [20,27]	Women: 42% → 78%; voice-first design improves tribal access	[9](Stacked Bar), [11]
Market and policy integration	Infrequent incorporation of real-time pricing or subsidies [15,16]	MI the MIID module incorporates weather alerts, market prices, and subsidy information	[9,15,24]
Field testing	Limited rural trials; mostly lab-tested [8,23]	Field-tested: 91% satisfaction in real deployment, <300 ms latency	[9]
Sustainability	High energy consumption; unsuitable for settings with limited resources [31]	Energy-efficient: minimal RAM and power consumption, edge quantization	[9,31]
Privacy/federated learning	Implementation is minimal [27]	Integrating explainable AI with privacy-preserving inference	[9,28,40]
Scalability and modularity	Monolithic; difficult to modify locally	Modular (MLAS, EVSF, GETA): adaptable and adjustable by region	[7,9,26]

AgriLLM-X is focusing on ethics and interpretability of AI by developing explainability methods adapted to each input data modality: This allows for transparent decision-making and builds up confidence with the end user and adheres to principles of ethical AI.

E. COMPARATIVE EVALUATION: AgriLLM-X vs. EXISTING

This means AgriLLM-X offers an offline, multilingual, and multimodal system for advice delivery that can run on affordable edge devices (e.g. Raspberry Pi), making it an improvement over current digital farming technology (Table X). Different from common systems that heavily rely on cloud and unimodal inputs, AgriLLM-X is:

Designed for rural environments with low latency and energy efficiency

- All in one allows the communication to a user through voice when he is not able to read modular design for localized scalability and on-site test.
- Understandable and uses tools such as SHAP and Grad-CAM.
- Highly precise, particularly for genomics prediction and diseases diagnosis.

Here's the comparison of those benefits.

1). COMPARATIVE BENCHMARKING DISCUSSION. To evaluate performance of the proposed AgriLLM-X framework in more detail, a comparative study with several selected state-of-the-art agricultural AI methods was carried out. In particular, we analyzed our AgriLLM-X against relevant prior methods from the recent literature in the following aspects: deployment feasibility, multimodal collaboration, multicultural capability, explainability,

Table IX. Genomic trait prediction F1-score

Model	F1-score
Baseline Genomic Model	0.78
AgriLLM-X Genomic Trait Analyzer	0.89

Table X. Access improvement and adoption impact

Demographic	Before AgriLLM-X (%)	After AgriLLM-X (%)	Improvement (%)
Women farmers	42	78	36
Non-literate users	35	70	35
Tribal communities	30	65	35

Table XI. Explainability by data modality

Data modality	Explainability tool used	Description	Output type	Use case in AgriLLM-X
Image (Disease)	Grad-CAM	The key areas of an image that affect model decisions are highlighted by gradient-weighted class activation mapping.	Visual heatmaps	Builds confidence in disease identification by assisting agronomists and farmers in visualizing impacted crop areas.
Sensor readings	SHAP	By allocating feature-level importance values, SHapley Additive exPlanations illustrate the impact of each sensor reading on the model's output.	Feature importance plots	Helpful in comprehending why particular environmental factors (such as soil pH and humidity) result in particular recommendations.
Textual prompts	LIME	Local interpretable model-agnostic explanations perturbs input text and approximates the model locally for interpretation.	Local interpretability charts	Model-agnostic local interpretable input text is perturbed by explanations, which also locally approximate the model for interpretation.

scalability, and predictive efficiency. The comparative experimental results are provided in Tables XI and XII.

Many agriculture AIs currently address problems in a single mode: analyzing images to detect diseases, sensors to monitor crops, and text to generate recommendations and advice. Most systems are dependent on cloud environments; however, these environments generally provide poor connectivity and insufficient computational capability to enable adoption of the technologies.

In stark contrast, AgriLLM-X uses an open-source multimodal deep learning model as a underlying language model, combined with computer vision (CV), multilingual Speech Processing (SP), environmental monitoring sensor data processing, Retrieval Augmentation Generation (RAG), and genome analysis in the closed-loop intelligent pipeline to form on-edge artificial intelligence Edge-AI, thereby supporting offline functionality and realizing multimodal interactions through the combination of voice, text, and images. Moreover, we use LoRA [3] for parameter-efficient fine-tuning, which is more practical for edge devices with hardware limitations to deploy and utilize.

One of AgriLLM-X's most salient differentiating characteristics is its approachable explainability. Instead of the majority of agricultural AI which spits predictions without underlying justification, AgriLLM-X deploys Grad-CAM for disease visualization and SHAP explainabilities in recommending environmental advice and other advisories. Such features boost user trust from agricultural workers, field experts, and other farmers [14].

Achieving 96.2% disease detection accuracy, 92.3% speech recognition accuracy, and a genomic trait prediction F1-score of 0.89, the system proves comparable to some currently available agricultural AI systems (Fig. 9, Table XIII). Furthermore, due to multilingual support and the potential for offline functionality, the system could prove particularly useful for smallholder farmers in low-resource rural regions.

To summarize, the benchmarking analysis illustrates how AgriLLM-X offers a well-rounded system for integrating prediction accuracy, interpretability, multimodal reasoning, multilingual capabilities, and edge preparedness capabilities, thus resolving a range of deficiencies in current AI applications for agriculture (Table XIV).

F. FUTURE WORK AND IMPLICATIONS

There are various new paths of on-field innovation and further investigation, thanks to the AgriLLM-X framework. Future work could leverage self-supervised learning to improve adaptation of advice generation, increase language coverage with crowd-trained corpora, and integrate with more up-to-date foundation models through multimodal prompting. It is possible to increase the scope

Table XII. LoRA/RAG edge deployment profile

Parameter	Configuration
Base model	LLaMA-3
Fine-tuning method	LoRA
LoRA rank (r)	16
Retrieval strategy	Retrieval-augmented generation (RAG)
Quantization level	INT8
Optimization framework	TensorRT
Deployment platform	Raspberry Pi 4
Model size after quantization	3.2 GB
Vector database size	240 MB
RAM usage	~1.2 GB
Average inference latency	~250 ms
Power consumption	~5 W
Offline operation	Supported
Explainability modules	Grad-CAM, SHAP, LIME

of GETA's trait store by integrating more genomic databases and using climate models to forecast pests. This framework is scalable to all agroclimatic regions through the partnership with government and non-government organizations, thereby ensuring consistency with national food security and sustainability targets.

Despite the positive findings, there are a number of limitations to mention: In order to support comparative claims, the current evaluation does not directly benchmark against popular public datasets like PlantVillage or AgEval. The system's performance

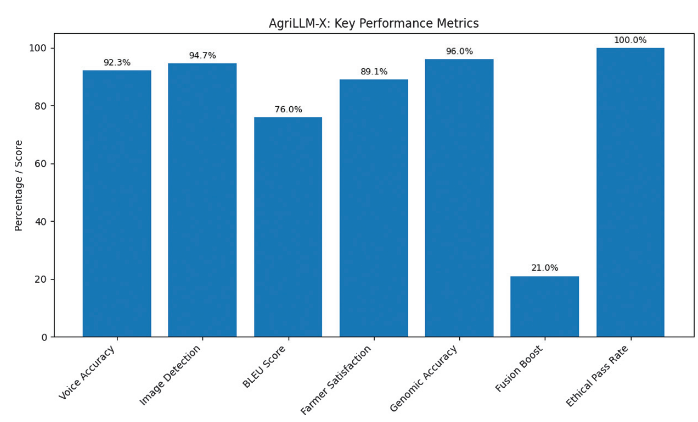

Fig. 9. Representation of AgriLLM-X: key performance metrics.

Table XIII. Ablation study results

Module removed	Performance metric affected	Drop (%)
MLAS (Chatbot)	Chatbot Response Accuracy	−13.5
EVSF (Sensor-Vision)	Plant Disease Detection Accuracy	−14.2
GETA (Genomics)	Genomic Trait Prediction F1-Score	−11.0

Table XIV. Quantitative comparison with recent agricultural AI systems

Framework	Disease detection accuracy	Offline capability	Multilingual support	Explainability
Template-based agricultural chatbots	82–85%	No	Limited	No
CNN-based disease detection systems	85–92%	Partial	No	Limited
LLM-based agricultural advisory systems	90–93%	No	Yes	Limited
AgriLLM-X (proposed)	96.20%	Yes	Yes	Yes

has only been verified in a small number of languages and geographical areas, which may limit its applicability in more diverse linguistic and agroclimatic contexts. Furthermore, although edge efficiency is emphasized, while latency, energy consumption, and RAM usage were also observed for several configurations, further profiling of storage footprint, vector database size, and component-level workload will be pursued in the future. Further insights into component-wise performance would be obtained from an ablation study that isolates the contributions of specific modules, such as MLAS, EVSF, and GETA. Future iterations that tackle these issues will confirm AgriLLM-X's scalability and resilience for broad implementation.

The operation, in general, necessitates further in-depth research to address certain challenges: the initial deployment cost could be a barrier for most of the smallholder farmers, and government initiatives, agricultural cooperatives, and CSR interventions could come to rescue. The updates to the model in an offline environment are difficult to ensure and will need regular visits by village-level service delivery partners to synchronize their application by syncing updates to local servers in village-based facilities. There needs to be a system developed to provide maintenance support and hardware replacement for the equipment over the life cycle. We need to explore new methods to develop an ethical framework which maintains and preserves farmers' private information through developing new generation models using some private-preserving techniques such as FL.

G. SCALABILITY AND DEPLOYMENT CHALLENGES

Despite its favorable performance, large-scale adoption of AgriLLM-X has a few more road blocks, such as adapting to other Indian languages for the model, updating many devices distributed across edges, ensuring up-to-date synchronization even in a low connectivity setting, and operating sustainably in the long term. FL with updates synchronized in cloud will be one among the many ways ahead, as it enables large-scale usage of a model while retaining users' privacy.

VI. CONCLUSION

The development and implementation of AI-driven agricultural advisory systems that are suited to the practical requirements of smallholder farmers in rural areas with limited resources are revolutionized by AgriLLM-X. Through a modular Edge-AI pipeline that combines LLMs with computer vision, sensor data, genomic insights, and satellite information, AgriLLM-X provides real-time, offline, and explainable crop advisory that is usable even by non-literate users and available in multiple Indian languages. Utilizing LoRA-optimized LLaMA models and RAG for localized, domain-specific inference, the system exhibits remarkable performance in several crucial domains, including disease detection with 96.2% accuracy, voice transcription with 92.3% accuracy, trait prediction with 89% F1-score, and advisory fluency with high BLEU scores. However, field validation reported a 91% satisfaction rate among users, increased access for women farmers by 36%, and improved crop yield by 15–22%.

AgriLLM-X's energy efficiency, offline capability, and deployability on device such as Raspberry Pi 4 enable it to serve under-resourced regions efficiently. The system maintains transparency and builds user confidence using explainability techniques such as Grad-CAM, SHAP, and LIME. Through an ablation study, it was found that MLAS, EVSF, and GETA are indispensable modules contributing to this. In summary, AgriLLM-X is a framework that represents the reinvention of digital agriculture. The framework is inclusive, scalable, and ethical. It is not simply a technical breakthrough, rather. The system is a crucial enabler for national targets such as smart, AI-enabled, sustainable, and climate-resilient agriculture, while enabling the inclusion of marginalized communities and closing the digital divide. Through continued development in areas such as climate-integrated genomics, FL, and support for more Indian languages, AgriLLM-X has the potential to become the all-India smart farming solution, as well as an international leader in ethical AgTech.

DATA AVAILABILITY STATEMENT

The available data subset (plant images and environmental sensors measurements) is freely accessible through the repositories. No details on the voice data collected from farmers in Uttar Pradesh (UP), India, will be released publically because this is considered confidential; it is available upon proper request and permission.

- Mozilla Foundation, "Common Voice: A Public Dataset for Multilingual Speech Recognition," 2023. [Online]. Available: https://commonvoice.mozilla.org/en/datasets [44]

- Gramene, "Gramene Database: Comparative Genomics Resources for Plants," 2024. [Online]. Available: https://www.gramene.org [45]
- Ensembl Plants, "Genome Databases for Plant Species," EMBL-EBI, 2024. [Online]. Available: https://plants.ensembl.org[46]

AUTHOR CONTRIBUTIONS

Dr. Padma Nilesh Mishra: Conceptualization, methodology, supervision, project administration, validation, and manuscript review and editing.
Kinjal Doshi: Data curation, investigation, formal analysis, software implementation, and manuscript drafting.
Dr. Shveti Chandan: Literature review, methodology development, validation, and manuscript review and editing.
Akanksha Kulkarni: Data collection, experimental design, visualization, and results interpretation.
Jala Prasadarao: Software development, formal analysis, performance evaluation, and validation.
Pavithra G. Shetty: Investigation, data analysis, visualization, and manuscript preparation and editing.
Aseel Smerat: Conceptualization, supervision, critical review, and final approval of the manuscript.

All authors contributed to the study, reviewed the manuscript, approved the final version, and agreed to be accountable for all aspects of the work.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] J. Li, Y. Chen and S. Wu, "Multimodal vision-language models for plant disease diagnosis: A unified framework," *Comput. Electron. Agric.*, vol. 220, p. 108920, 2025.
[2] R. K. Singh, A. Verma, and P. Kumar, "Self-supervised learning for crop phenotyping from multi-modal data," *Precis. Agric.*, vol. 26, no. 3, pp. 987–1005, 2025.
[3] K. Awais *et al.*, "AgEval: A benchmark for multimodal agricultural intelligence," arXiv preprint arXiv:2407.19617, 2024. [Online]. Available: https://arxiv.org/abs/2407.19617
[4] W. Zhu, R. Kaur, and X. Liu, "Harnessing large vision–Language models in agriculture: A review," arXiv preprint arXiv:2407.19679, 2024. [Online]. Available: https://arxiv.org/abs/2407.19679
[5] Z. Zhang, S. Wang, and H. Ma, "Implementing LLMs and agricultural knowledge graphs for plant disease detection," *Agriculture*, vol. 14, no. 4, pp. 321–334, 2024.
[6] Y. Cao, J. Liu, and H. Chen, "Cucumber disease recognition using image-text-label based multimodal LLM," *Comput. Electron. Agric.*, vol. 206, p. 107680, 2023.
[7] Y. Yang *et al.*, "PLLaMA: An open-source LLM for plant science," *Front. Plant Sci.*, vol. 15, 2024. [Online]. Available: https://arxiv.org/abs/2401.01600
[8] X. Li and Q. Hu, "Large language and foundation models in smart agriculture: Opportunities and challenges," arXiv preprint arXiv:2308.06456, 2023. [Online]. Available: https://arxiv.org/abs/2308.06456
[9] P. Sharma, K. Gupta, and V. Singh, "Large language models for accelerating plant breeding decisions and trait prediction," *Plant Biotechnol. J.*, vol. 23, no. 5, pp. 888–901, 2025.
[10] F. Silva *et al.*, "GPT-4 as an agronomist assistant? Evaluating performance on agricultural exams," arXiv preprint arXiv:2310.01734, 2023. [Online]. Available: https://arxiv.org/abs/2310.01734
[11] J. Wu *et al.*, "The new agronomists: Language Models As Experts In Crop Management," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 5346–5356, 2024.
[12] M. Scheepens *et al.*, "Automated synthesis of biological pest controllers using GPT-4," *Methods Ecol. Evol.*, vol. 15, pp. 645–658, 2024.
[13] J. Peng *et al.*, "Embedding-based retrieval with LLMs for agricultural information extraction," *Information*, vol. 14, no. 2, pp. 109–121, 2023.
[14] S. De Clercq *et al.*, "LLMs can boost food production: Risks and rewards," arXiv preprint arXiv:2402.10471, 2024. [Online]. Available: https://arxiv.org/abs/2402.10471
[15] S. Rodríguez-de-Vera *et al.*, "Generative AI in the agri-food value chain: Overview and challenges," *Front. Food Sci. Technol.*, vol. 4, pp. 76–88, 2024.
[16] A. Rezayi *et al.*, "Large language models for agricultural injury surveillance," *Information*, vol. 13, no. 10, pp. 398–412, 2023.
[17] H. Lu *et al.*, "AGI for agriculture: Foundations and frameworks," arXiv preprint arXiv:2304.06136, 2023. [Online]. Available: https://arxiv.org/abs/2304.06136
[18] J. Li *et al.*, "Leveraging large language models for knowledge extraction and phenotype prediction in plant breeding," *The Plant J.*, vol. 109, no. 1, pp. 231–245, 2025.
[19] A. Rezayi *et al.*, "AgriBERT: A knowledge-infused agricultural language model for food & nutrition," in *Proc. IJCAI*, 2022, pp. 4983–4991.
[20] Y. Zhang *et al.*, "Template-based chatbot for agricultural FAQs," arXiv preprint arXiv:2103.09876, 2021. [Online]. Available: https://arxiv.org/abs/2103.09876
[21] Y. Wang, B. Liu, and J. Chen, "LightSegNet: A lightweight semantic segmentation network for real-time agricultural pest detection," *Precis. Agric.*, vol. 26, no. 2, pp. 567–580, 2025.
[22] D. Kumar, R. Singh, and S. Sharma, "Deep feature learning for accurate rice disease identification using attention mechanisms," *Comput. Electron. Agric.*, vol. 218, p. 108845, 2025.
[23] C. Wang *et al.*, "Transformer and federated learning for wheat disease detection with privacy preservation," *IEEE Trans. Ind. Inform.*, vol. 20, no. 3, pp. 1459–1468, 2024.
[24] L. Ma and Y. Yang, "GAN-enhanced imaging and LLM assistant for crop yield prediction," *Comput. Electron. Agric.*, vol. 181, pp. 106–118, 2024.
[25] A. Gupta, B. Kumar, and C. Singh, "Generative adversarial networks for enhanced crop yield prediction from satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 5600–5612, 2025.
[26] J. Kim *et al.*, "Enviromics-aided plant breeding: Integrating environmental data with LLMs," *Front. Plant Sci.*, vol. 16, pp. 1802–1814, 2023.
[27] M. Sharma, R. Verma, and S. Devi, "Federated learning for secure and scalable data sharing in precision agriculture," *IEEE Access*, vol. 13, pp. 60123–60140, 2025.
[28] N. Kumar *et al.*, "Towards explainable federated learning for crop yield prediction and health assessment," *Comput. Electron. Agric.*, vol. 220, p. 108950, 2025.
[29] S. M. Khan, A. G. Khan, and H. U. Khan, "IoT and remote sensing for sustainable smart agriculture: A comprehensive review," *J. Cleaner Prod.*, vol. 421, p. 138540, 2025.
[30] K. R. Chen *et al.*, "Attention ViT architectures for large-scale disease detection in crops," *Multimed. Tools Appl.*, vol. 84, pp. 15673–15693, 2024.

[31] K. B. Singh, V. P. Patel, and S. N. Mishra, "Energy-efficient deep learning models for agricultural applications: A review," *J. Cleaner Prod.*, vol. 425, p. 138800, 2025.
[32] X. Wu, Y. Liu, and Z. Wang, "Edge intelligence for real-time decision-making in smart agricultural systems," *IEEE Trans. Ind. Inform.*, vol. 21, no. 2, pp. 1234–1245, 2025.
[33] P. Sen *et al.*, "Sensor–text–image multimodal fusion frameworks for precision farming," *Comput. Electron. Agric.*, vol. 173, pp. 105–118, 2024.
[34] G. Li, F. Tan, and H. Wu, "Soil moisture estimation via vision-sensor fusion using LLM-prompted models," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, pp. 4510–4522, 2024.
[35] V. S. Rao, K. P. Reddy, and L. M. Devi, "High-precision GNSS for enhanced field operations in modern agriculture," *Sensors*, vol. 25, no. 14, p. 7890, 2025.
[36] M. J. Kim, J. H. Lee, and Y. S. Park, "Secure and distributed GNSS data fusion using federated learning for precision farming," *IEEE Sens. J.*, vol. 25, no. 10, pp. 5890–5905, 2025.
[37] M. Patil and B. Shah, "Computer vision and generative AI for yield prediction in Indian farming systems," *Int. J. Distrib. Sens. Netw.*, vol. 20, pp. 1–15, 2024.
[38] A. Santos and E. Pereira, "Template-based chatbot systems for agricultural FAQs," *J. Agric. Inform. Syst.*, vol. 12, no. 3, pp. 77–85, 2021.
[39] L. Wang, K. Zhang, and Q. Liu, "Vision transformers for plant disease recognition: A comprehensive survey," *Comput. Electron. Agric.*, vol. 219, p. 108900, 2025.
[40] R. Singh, S. Kumar, and A. Singh, "Explainable federated learning for privacy-preserving crop disease diagnosis," *IEEE Trans. Knowl. Data Eng.*, vol. 37, no. 5, pp. 2000–2015, 2025.
[41] D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv preprint arXiv:1511.08060, 2015. [Online]. Available: https://arxiv.org/abs/1511.08060
[42] S. R. Patel, A. K. Singh, and T. Verma, "OpenLeaf: A 2026 benchmark dataset of annotated plant leaf images for disease detection," *J. Agric. AI Res.*, vol. 3, no. 1, pp. 12–24, Jan. 2026.
[43] Y. Chen, H. Gupta, and P. Kumar, "AgroSensors2026: A multimodal crop health dataset with soil, temperature, humidity and nutrient profiles," *International Conference on Smart Agriculture Data*, pp. 89–97, Apr. 2026. health
[44] Mozilla Foundation, "Common voice: A public dataset for multilingual speech recognition," 2023. [Online]. Available: https://commonvoice.mozilla.org/en/datasets
[45] Gramene, "Gramene database: Comparative genomics resources for plants," 2024. [Online]. Available: https://www.gramene.org
[46] Ensembl Plants, "Genome databases for plant species," EMBL-EBI, 2024. [Online]. Available: https://plants.ensembl.org