Voice Enabled Assistance For Rural India

Under Gudiance of Dr. Manpreet Kaur

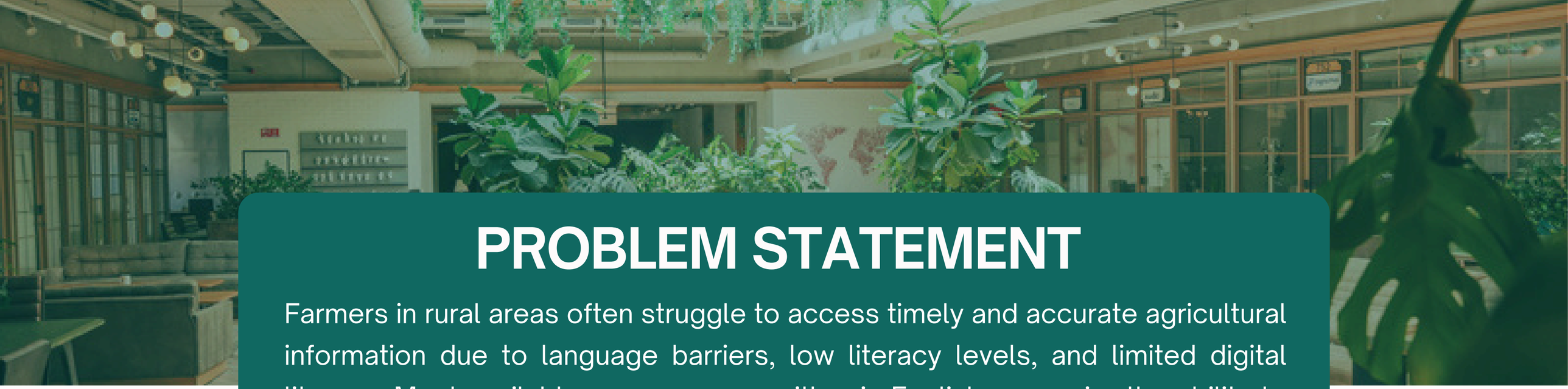
TEAM DETAILS

Dogga Pavan Sekhar - 2K22CSUN01167

Ketha Sathwik Reddy - 2K22CSUN01173

Kesanakurthi Naga Siddhartha - 2K22CSUN01172

Table of Contents



PROBLEM STATEMENT

Farmers in rural areas often struggle to access timely and accurate agricultural information due to language barriers, low literacy levels, and limited digital literacy. Most available resources are either in English or require the ability to read and understand complex agricultural terminology. Additionally, conventional FAQ systems lack personalization and fail to provide simplified explanations. There is a need for a solution that can understand a farmer's natural language query, find relevant information, simplify it, and present it in their native language preferably in spoken form. This project addresses these challenges by building an intelligent system that combines query understanding, semantic similarity, generative summarization, multilingual translation, and text-to-speech to make agricultural knowledge more accessible, understandable, and farmer-friendly.

CHALLENGES

Millions of rural individuals face significant barriers in accessing critical information about government schemes, healthcare services, and agricultural practices due to :

Low literacy levels

Many individuals cannot read or write, limiting access to digital or printed information

Limited internet connectivity

Many rural areas suffer from poor or no internet access, restricting their ability to search for information.

Language barriers

Most online resources are available in English or Hindi, making them inaccessible for native speakers of regional languages and dialects.

Complex government processes

People often struggle to navigate bureaucratic procedures, leading to missed benefits from welfare programs.

REAL CASE SCENARIO

A farmer in a rural village faces a pest infestation on his crops but struggles to find reliable solutions due to illiteracy and lack of internet access. With no nearbyagricultural experts, he relies on word-of-mouth advice, which may be inaccurate. The proposed voice-enabled NLP system allows him to ask questions in his local dialect and receive spoken, easy-to-understand responses with verified pest control methods. This ensures quick action, reduces crop loss, and improves productivity, even in offline conditions.

CHALLENGES

Cannot find reliable information in his native language.

Relies on local word-of-mouth advice, whichmay be inaccurate.

Government helplines are difficult to navigate due to language barriers.

SOLUTIONS

The farmer asks hisquestion via voice input in his native language.

The system recognizes the speech, processes the query, and retrieves the best answer from trusted sources.

The response is translated and spoken back in the farmer’s language for easy understanding.

LITERATURE SURVEY

https://docs.google.com/spreadsheets/d/1Ui8xLol27GnYdJsX8C7YAebHT-AilJmOtrYJ8QYivDE/edit?usp=sharing

NOVELITY OF APPROACH

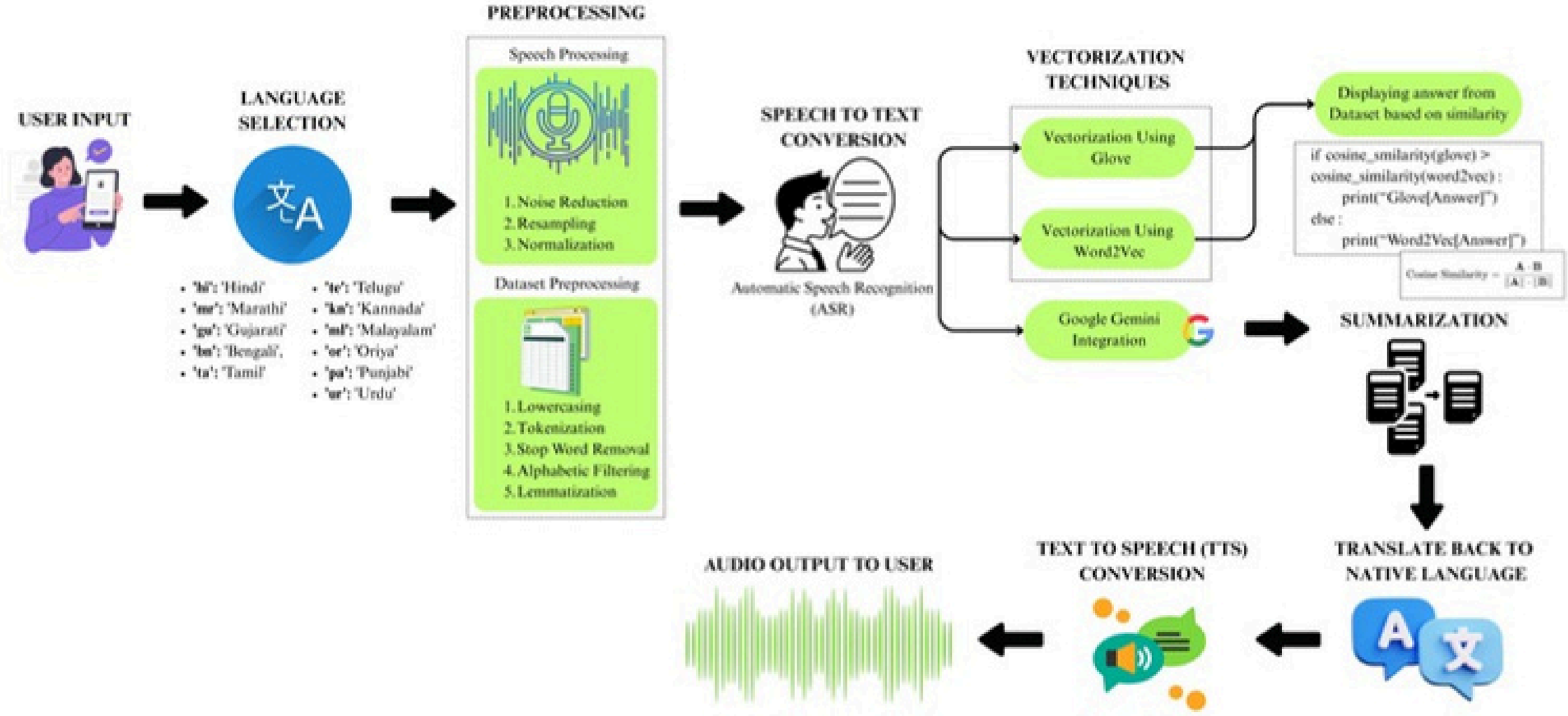
MULTILINGUAL VOICE ASSISTANCE

OFFLINE FUNCTIONALITY

CONTEXT-AWARE AI

ADAPTATION TO RURAL SPEECH

PROPOSED METHODOLOGY



DATASET PREPARATION

We curated a custom dataset by manually collecting frequently asked questions from official government websites. Each entry includes a category, topic, question, and answer, focused on areas like government schemes, healthcare, and agriculture. The questions were simplified and answers made concise to suit voice-based interactions. This domain-specific dataset significantly enhanced the assistant's accuracy and intent recognition, making it more effective and accessible for rural users.

Category	Topic	Question	Answer
Government Scheme	PM Kisan Scheme	What is Pradhan Mantri Kisan Samman Nidhi?	Pradhan Mantri Kisan Samman Nidhi (PM Kisan) is a new Central Sector Scheme to provide income support to all landholding farmers' families in the country to supplement their financial needs for procuring various inputs related to agriculture and allied activities as well as domestic needs. Under the PM Kisan Scheme, the entire financial liability towards transfer of benefit to targeted beneficiaries will be borne by Government of India.
Government Scheme	PM Kisan Scheme	Whether the benefits of the PM Kisan Scheme are admissible to only Small & Marginal Farmers' (SMF) families?	No. In the beginning when the PM Kisan Scheme was launched on 24th February, 2019, its benefits were admissible only to Small & Marginal Farmers' (SMF) families, with combined landholding upto 2 hectare. The PM Kisan Scheme was later on revised w.e.f. 1.6.2019 and extended to all farmer families irrespective of the size of their landholdings.
Government Scheme	PM Kisan Scheme	When was the PM Kisan Scheme launched?	The PM Kisan Scheme was launched by the Hon'ble Prime Minister on 24th February, 2019.
Government Scheme	PM Kisan Scheme	From which date the PM Kisan Scheme has come into effect?	The PM Kisan Scheme takes effect from 01.12.2018.
Government Scheme	PM Kisan Scheme	Who are eligible to get benefits under the PM Kisan Scheme?	All landholding farmers' families, which have cultivable landholding in their names are eligible to get benefit under the PM Kisan Scheme.
Government Scheme	PM Kisan Scheme	Who are not eligible to get benefits under the PM Kisan Scheme?	of Rs.6000/- per annum per family payable in three equal installments of Rs.2000/- each, every four months. The following categories are not eligible to get the benefits of the PM Kisan Scheme: (a) All Institutional Land holders; and (b) Farmer families in which one or more of its members belong to the following categories - i. Former and present holders of constitutional posts ii. Former and present Ministers/ State Ministers and former/present Members of Lok sabha/ Rajya sabha/ state Legislative Assemblies/ State Legislative councils, former and present Mayors of Municipal corporations, former and present Chairpersons of District Panchayats iii. All serving or retired officers and employees of
Government Scheme	PM Kisan Scheme	How many times the benefit will be given in a year under PM Kisan Scheme?	Under the PM Kisan Scheme, all landholding farmers' families shall be provided the financial benefit of Rs.6000/- per annum per family payable in three equal installments of Rs.2000/- each, every four months.
Government Scheme	PM Kisan Scheme	Government / PSU / Autonomous Organization, etc., who holds cultivable land in his/her own name in	Departments and their field units, Central or State PSEs and Attached offices / Autonomous Institutions under Government as well as regular employees of the Local Bodies are not eligible to get
Government Scheme	PM Kisan Scheme	Will any individual or farmer family owning more than 2 hectare of cultivable land get any benefit under the PM Kisan Scheme?	Yes. The ambit of the PM Kisan Scheme has been extended to cover all farmer families, irrespective of the size of their land holdings.

PREPROCESSING TECHNIQUES

PREPROCESSING FOR INPUT SPEECH

To ensure clean and standardized input for our voice-enabled assistant, we apply the following preprocessing steps to the raw audio

Noise Reduction

We use noisereduce to eliminate background noise, improving speech clarity.

Resampling

The audio is resampled to a standard sampling rate of 16 kHz, ensuring consistency across all inputs.

Normalization

The amplitude of the audio signal is normalized to scale between -1 and 1, enhancing model stability and performance.

PREPROCESSING FOR DATASET

To prepare the questions for effective vectorization and similarity matching, we applied the following NLP techniques:

Lowercasing

Converted all text to lowercase for consistency.

Stop Word Removal

Removed common words (e.g., “the”, “is”, “in”) that do not add meaning.

Alphabet Filtering

Kept only alphabetic tokens, removing numbers and symbols.

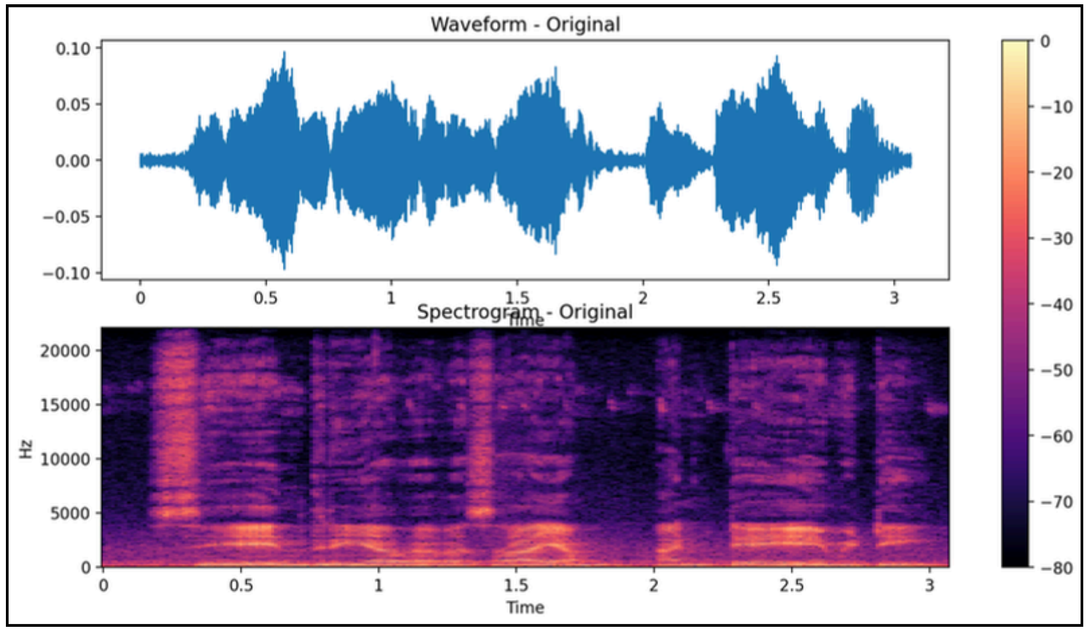
Tokenization

Split text into individual words (tokens) using NLTK.

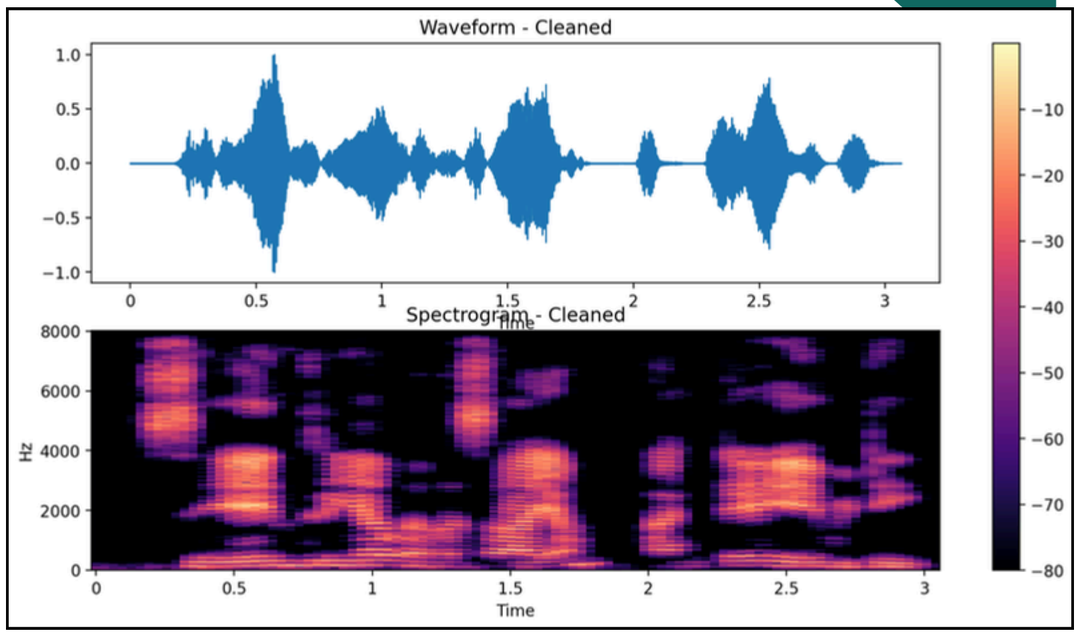
Lemmatization

Reduced words to their root form (e.g., “running” → “run”) for standardization.

OUTPUTS OF PREPROCESSING



Before Preprocessing



After Preprocessing

VECTORIZATION TECHNIQUES

To improve response accuracy, we implemented two word embedding techniques: Word2Vec and GloVe.

Word2Vec

Word2Vec learns word associations from large datasets using neural networks. It captures contextual meaning by considering surrounding words, making it effective for semantic similarity in conversations.

GloVe

GloVe (Global Vectors for Word Representation) uses word co-occurrence statistics from a global corpus. It excels at capturing overall semantic relationships but lacks deep contextual understanding.

We compute vectors using both methods and combine their cosine similarities to find the most relevant FAQ answer. This hybrid approach leverages the strengths of both techniques — Word2Vec's contextual awareness and GloVe's global understanding — ensuring accurate and meaningful responses.

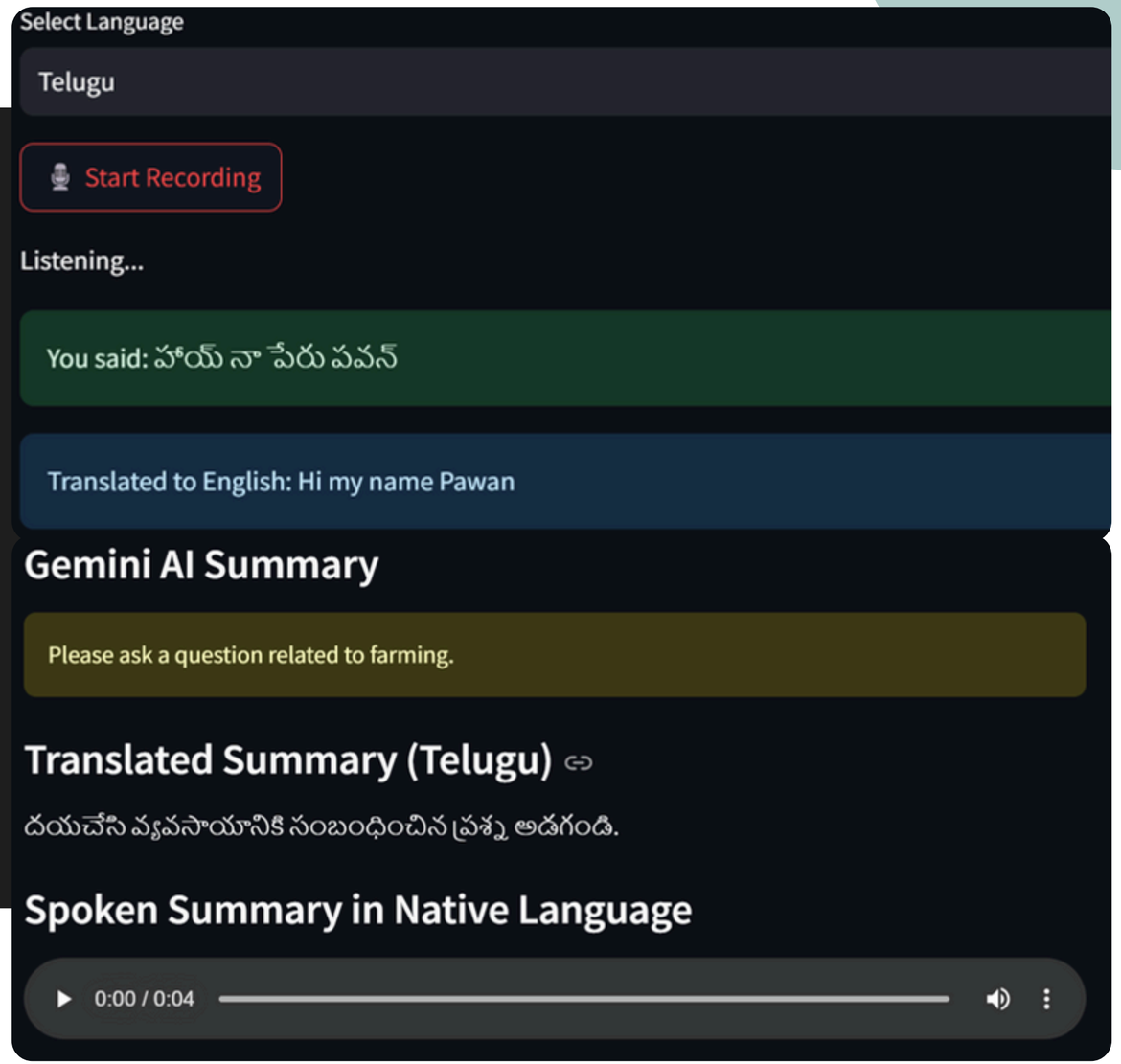
RESULTS AND DISCUSSION

```
--- Word2Vec Match ---
Matched Question: What are the major schemes and programs implemented by the Ministry of Agriculture and Farmers' Welfare?
Answer: The major schemes and programs implemented by the Ministry of Agriculture and Farmers' Welfare include: (i) Pradhan Mantr
Similarity Score: 0.5988507

--- GloVe Match ---
Matched Question: What are the major schemes and programs implemented by the Ministry of Agriculture and Farmers' Welfare?
Answer: The major schemes and programs implemented by the Ministry of Agriculture and Farmers' Welfare include: (i) Pradhan Mantr
Similarity Score: 0.7628881

✅ Final Answer (Based on GloVe):
The major schemes and programs implemented by the Ministry of Agriculture and Farmers' Welfare include: (i) Pradhan Mantri Kisan
```

```
✅ Final FAQ Answer (similarity ≥ 0.5):
The major crops grown in India include rice, wheat, pulses, oilseeds, sugarcane, cotton, jute, tea, coffee, and spices. India is also one of the largest producers of fi

**Translated Response in Native Language:** भारत में उगाई जाने वाली प्रमुख फसलों में चावल, गेहूं, दाल, तिलहन, गन्ने, कपास, जूट, चाय, कॉफी और मसाले शामिल हैं।भारत भी दुनिया के फलों और सबि

📅 Gemini API Response:
Rice needs nitrogen, phosphorus, and potassium - think NPK.  Urea is a good source of nitrogen.  Phosphorus can come from rock phosphate or DAP. Muriate of potash (MOP)


**Translated Response in Native Language:** బియ్యం నత్రజని, భాస్వరం మరియు పొటాషియం అవసరం - NPK గురించి ఆలోచించండి.యూరియా నత్రజని యొక్క మంచి మూలం.భాస్వరం ర
```

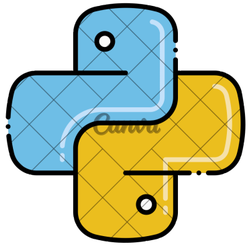
TEST CASES INCLUDED

```python
def is_farming_related(text):
    # Define keywords common in farming-related queries
    keywords = [
    # Core farming & agriculture
    'crop', 'soil', 'pesticide', 'fertilizer', 'weather', 'rain', 'market', 'wheat',
    'rice', 'harvest', 'plant', 'farm', 'irrigation', 'insect', 'yield', 'disease',
    'sowing', 'ploughing', 'organic', 'livestock', 'milk', 'tractor', 'seed',
    'paddy', 'maize', 'barley', 'agriculture', 'agronomy', 'drought', 'manure',
    'horticulture', 'poultry', 'farming', 'crop rotation', 'drip irrigation',
    'sprayer', 'greenhouse', 'weather forecast', 'pest', 'farm income', 'commodity prices',
    'barn', 'fodder', 'mulching', 'vermicompost', 'polyhouse', 'thresher',
    'soil health', 'crop insurance', 'farm loan', 'farm subsidy', 'climate change',
    'germination', 'nursery', 'land preparation', 'intercropping', 'weeding',
    'biopesticide', 'biofertilizer', 'water logging', 'crop disease', 'yield prediction',
    'precision farming', 'kharif', 'rabi', 'zayed', 'tillage', 'sprinkler irrigation',
    'irrigation canal', 'farmer', 'veterinary', 'agri input', 'market rate',

    # Government schemes and support
    'pm kisan', 'pm-kisan', 'pm fasal bima yojana', 'pradhan mantri fasal bima yojana',
    'soil health card', 'rashtriya krishi vikas yojana', 'e-nam', 'kisan credit card',
    'kcc', 'pmksy', 'pradhan mantri krishi sinchayee yojana', 'national food security mission',
    'paramparagat krishi vikas yojana', 'msp', 'minimum support price', 'agricultural subsidy',
    'dbt agriculture', 'farmers welfare', 'crop loan waiver', 'subsidy scheme', 'farmer pension scheme',
    'gramin bhandaran yojana', 'kisan samman nidhi', 'agrimarket', 'kisan call center',
    'agmarknet', 'rural employment', 'nrega', 'mnrega', 'rural development'
    'agriculture technology', 'agri-tech', 'digital agriculture', 'smart farming',
    'agriculture innovation', 'agriculture research', 'agriculture extension', 'agriculture education',]

    # Tokenize and lowercase user query
    tokens = [word.lower() for word in word_tokenize(text)]

    # Return True if any keyword is found
    return any(keyword in tokens for keyword in keywords)
```



TECHNOLOGIES :

Streamlit



Python

Google Sheets

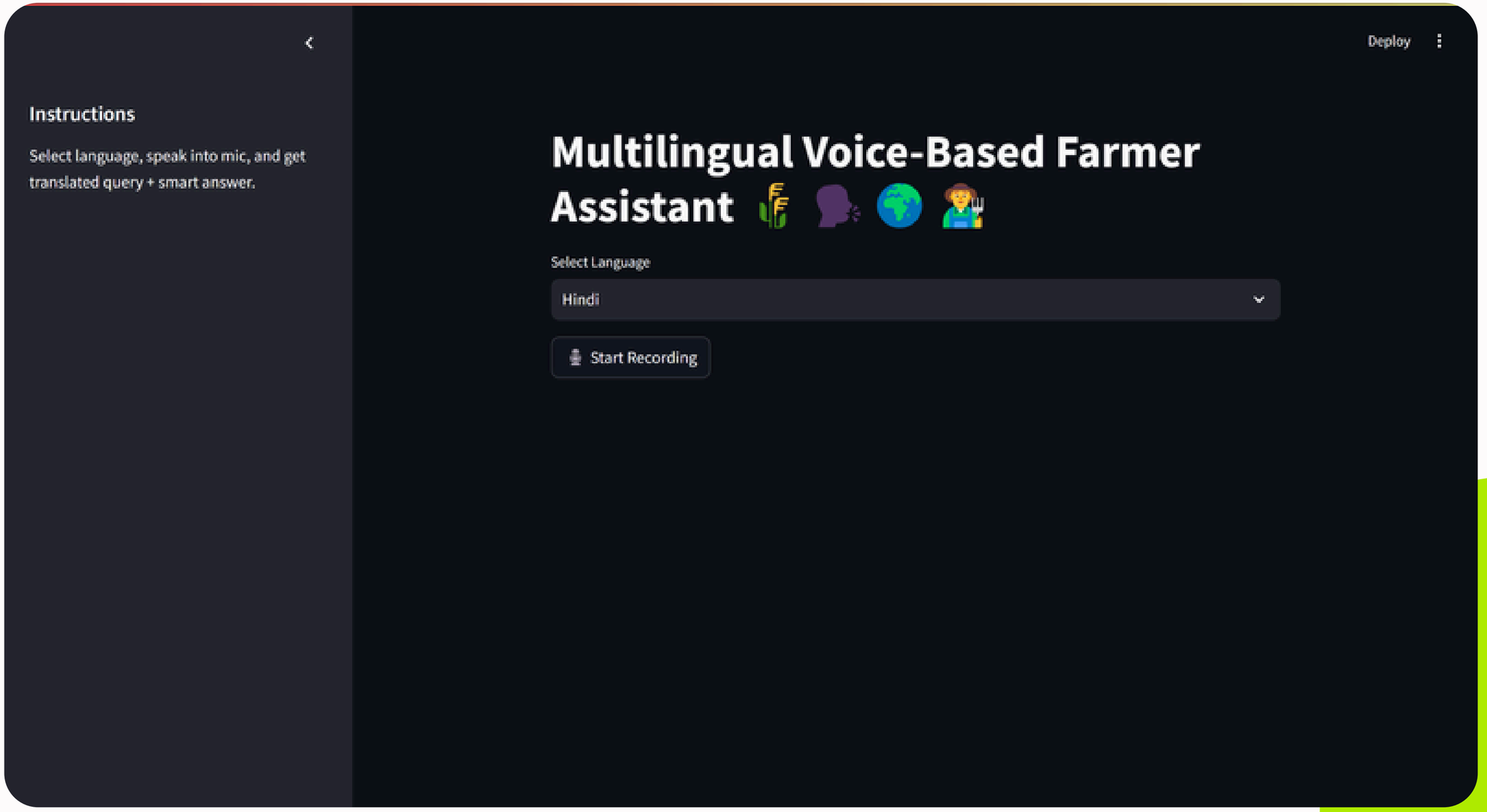
Natural Language Toolkit (NLTK)

Gemini

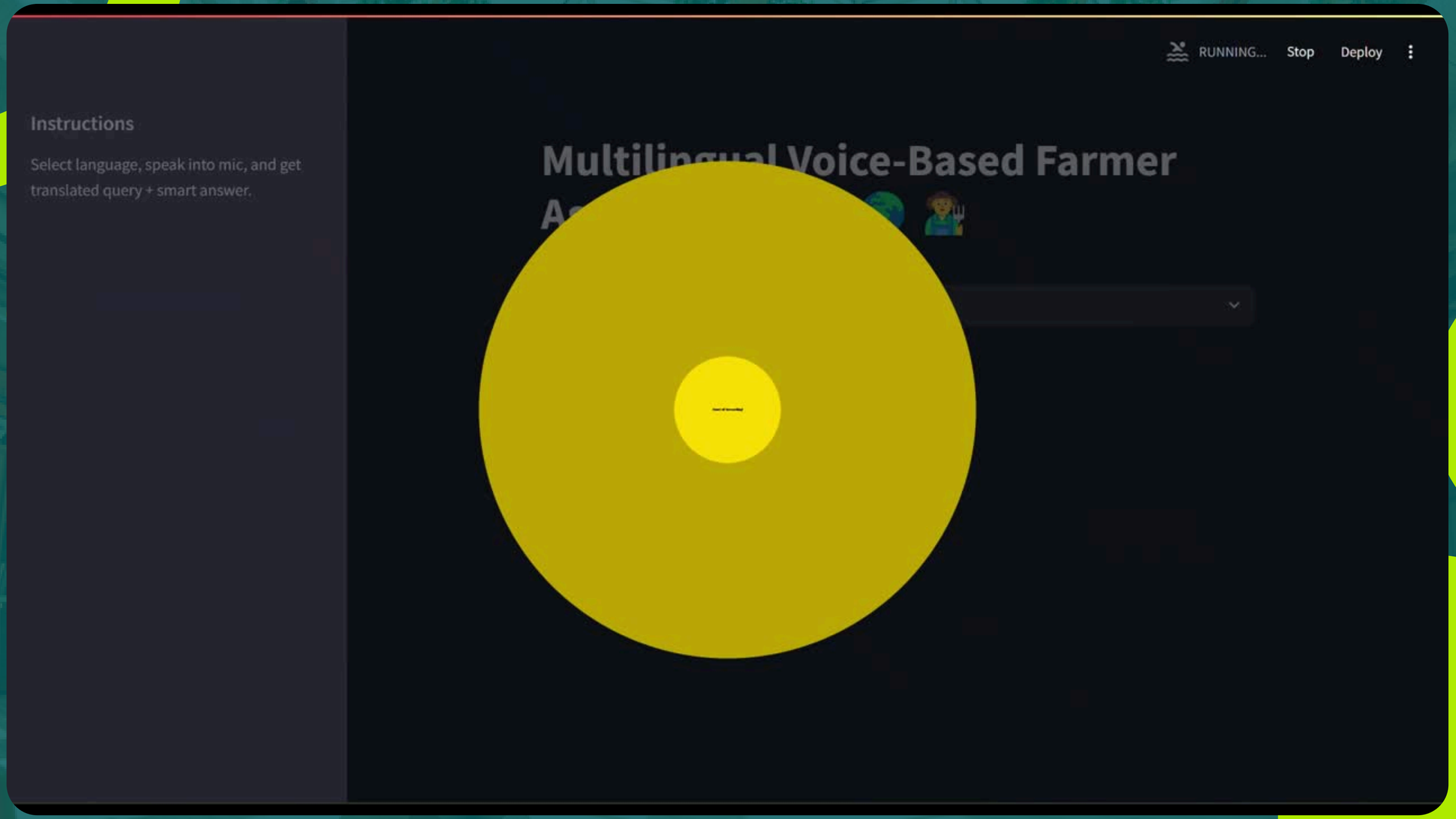
FUNCTIONALITIES

1. Voice input in Indian languages.
2. Translation to and from English.
3. Query cleaning and embedding (Word2Vec + GloVe).
4. Semantic matching with a local FAQ database.
5. AI-generated answers for unseen farming queries.
6. Text-to-speech feedback in the user’s language.
7. Audio and visual feedback of spoken query and cleaned version.

USER INTERFACE



DEMO VIDEO


RUNNING...
Stop
Deploy
Instructions
Select language, speak into mic, and get translated query + smart answer.
Multilingual Voice-Based Farmer

CONCLUSION

This project successfully bridges the communication gap between rural users and agricultural knowledge resources by combining natural language processing, semantic similarity techniques, and generative AI. By identifying farming-related queries, retrieving relevant FAQ responses, generating simple AI-driven summaries, and translating them into the user's native language with speech support, the system enhances accessibility and understanding for non-technical farmers. The integration of Word2Vec, GloVe, Gemini, and TTS technologies demonstrates the potential of multilingual, voice-enabled AI tools to empower rural communities and improve agricultural outcomes.

THANK YOU