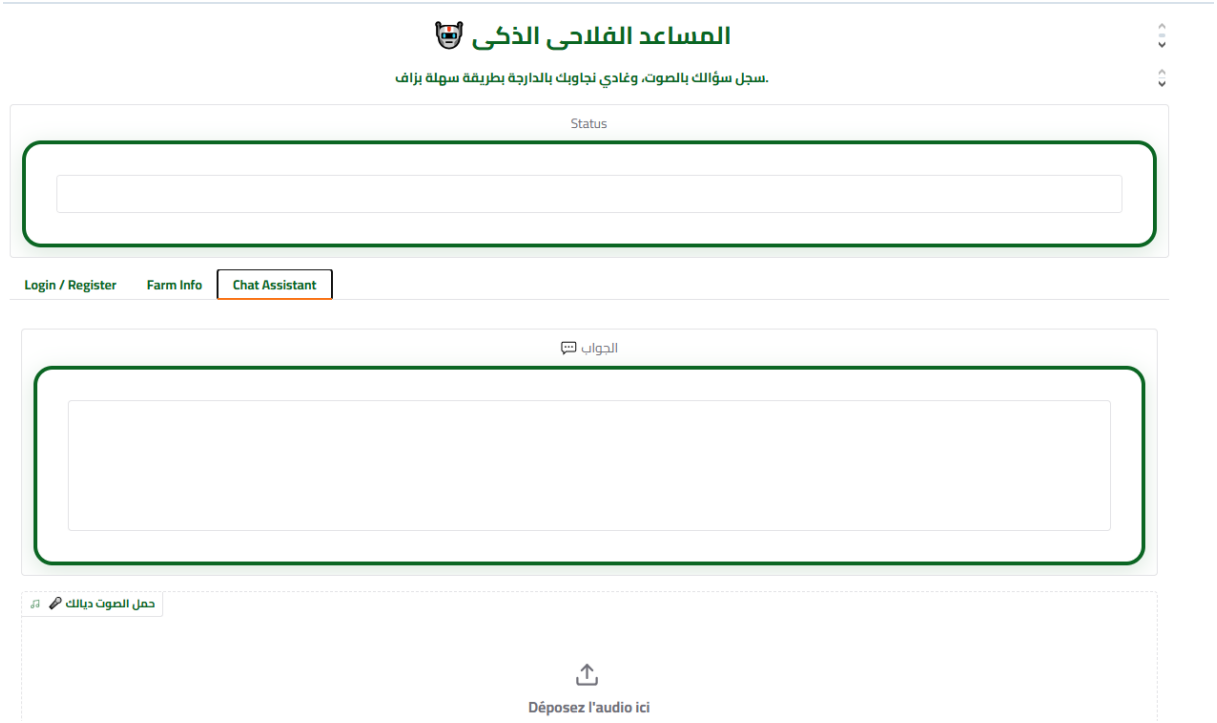
Integrated Project Report

ZiraaTech



Report by :
Anas Elkhabbaz
Hiba Benkaddour

Ilyas Boutayeb

Under the supervision of :
M. Hamza Gamouh

Academic Year 2024-2025

Contents

1. Project Overview

ZiraaTech is an innovative AI-powered agricultural assistant designed to empower Moroccan farmers. Given that agriculture accounts for **12% of Morocco's GDP** and employs **40% of the workforce**, ZiraaTech addresses key sector challenges: **low literacy rates** and **limited access to information**. The solution provides **personalized, voice-based agricultural advice** in simple Moroccan Darija, making it accessible to even the most illiterate farmers.

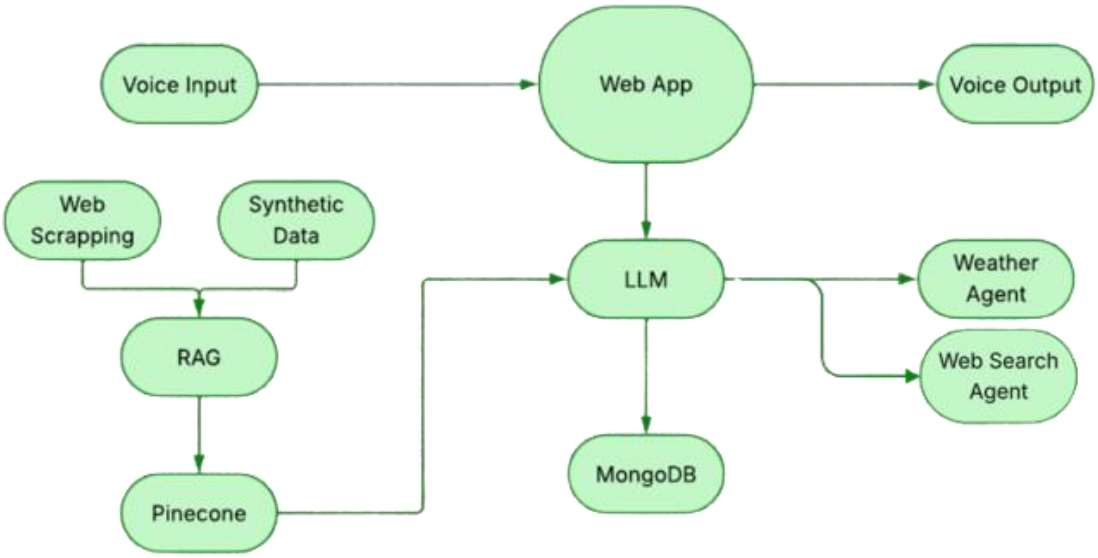
2. Objectives

- **Empower Farmers**: Deliver accessible, reliable, and personalized advice in Moroccan Darija.
- **Bridge Literacy Gaps**: Enable voice-based interactions rather than relying on text input.
- **Leverage AI**: Integrate advanced AI technologies for relevant and up-to-date recommendations.
- **Affordable Model**: Offer a **30 MAD/month** subscription, with the first month free, ensuring sustainability and accessibility.
- **Support SDGs**: Contribute to **SDG 1 (No Poverty)**, **SDG 2 (Zero Hunger)**, **SDG 4 (Quality Education)**, and **SDG 8 (Decent Work and Economic Growth)**.

3. System Architecture Overview

The system architecture is designed with modular components that interact seamlessly to deliver accurate and context-aware responses. The main components are:

1. Voice Input Module
2. Web Application Frontend
3. Large Language Model (LLM)
4. Data Sources (Synthetic Data & Web Scraping)
5. Retrieval-Augmented Generation (RAG) & Pinecone
6. Weather Agent
7. Web Search Agent
8. MongoDB (Database)
9. Voice Output Module



3. Module Descriptions

3.1 Voice Input Module

***Function**: Captures user queries in audio format using the Gradio interface.*

***Technology**: Gradio's Audio input component.*

***Format**: WAV file uploaded to server.*

3.2 Web Application

***Technology**: Gradio Blocks with custom CSS for Arabic support.*

***Features**:*

User Authentication (username/password)

Farm Data Management

Audio-based query submission

3.3 Large Language Model (LLM)

***Function**: Processes natural language queries and generates advice.*

***Technology**: Google Gemini API (gemini-2.0-flash).*

***Input**:*

Voice transcription (via Gemini STT)

Context block (retrieved via Pinecone + Web Scraping)

***Output**:*

Moroccan Darija advice text, structured for clarity and simplicity.

3.4 Data Sources

***Synthetic Data**: Curated dataset of agricultural knowledge.*

***Web Scraping**:*

Technology: SerpAPI for real-time search.

Purpose: To enrich context with up-to-date online information.

3.5 RAG (Retrieval-Augmented Generation)

***Technology**:*

SentenceTransformers (all-MiniLM-L6-v2) for embedding user queries.

Pinecone for fast vector similarity search.

***Logic**:*

Queries encoded and matched against dataset embeddings in Pinecone.

Top-k results are filtered by relevance threshold and keyword overlap.

Augmented with web search snippets if needed.

3.6 Pinecone

***Function**: Fast vector database for semantic search.*

***Integration**: Stores embeddings of agricultural dataset snippets.*

***API Key**: Configured securely in environment variables.*

3.7 Weather Agent

***Function**: Fetches real-time weather data.*

***Technology**: OpenWeatherMap API.*

***Integration**: Injects weather context into the prompt.*

3.8 MongoDB Database

***Purpose**: Stores user credentials and farm data.*

***Technology**: MongoDB Atlas cloud instance.*

***Security**: Passwords stored as SHA-256 hashes.*

3.9 Web Search Agent

***Function**: Enriches the assistant's response with the latest web data.*

***Technology**: SerpAPI.*

***Integration**: Used in the RAG pipeline to enhance answer relevance.*

3.10 Voice Output Module

***Function**: Converts AI-generated text answers into speech.*

***Technology**: Gradio TTS integration with a Moroccan Darija voice.*

***Process**: Uses Gradio's hosted TTS API with customized voice tone.*

4. Technology Used

4.1 Programming Language

Python

- *Used for backend development, including data processing, LLM integration, and API management.*
- *Chosen for its rich ecosystem and support for AI and web frameworks.*

4.2 Frontend Framework

Gradio

- *Provides an interactive web interface for farmers to upload voice queries and receive responses.*
- *Supports audio input/output, user authentication, and farm data forms, with custom CSS for Arabic localization.*

4.3 Large Language Model

Google Gemini API

- *Handles transcription (speech-to-text) and generates text-based agricultural advice in Moroccan Darija.*
- *Offers high-quality, context-aware responses using prompt engineering.*

4.4 Semantic Search

SentenceTransformers

- *Used to embed user queries into vector space for semantic similarity search.*
- *Facilitates context retrieval from the internal agricultural dataset.*

Pinecone

- *A vector database that stores embeddings of agricultural knowledge snippets.*
- *Allows fast, scalable similarity search to find relevant information for each query.*

4.5 Web Search

SerpAPI

- *Provides real-time web search results to enrich answers with up-to-date agricultural information.*
- *Adds dynamic external knowledge to complement internal datasets.*

4.6 Data Storage

MongoDB Atlas

- *Cloud-hosted NoSQL database that stores user credentials, farm data, and optionally query logs.*

- *Chosen for its scalability, flexibility, and easy integration with Python.*

4.7 Weather Data

OpenWeatherMap API

- *Retrieves current weather information for the farmer's region.*
- *Adds context to recommendations for planting, irrigation, etc.*

4.8 Text-to-Speech (TTS)

Gradio TTS (via Hugging Face)

- *Converts AI-generated text responses into spoken Moroccan Darija.*
- *Ensures accessibility for illiterate farmers by delivering advice in audio format.*

4.9 Additional Libraries

- ***hashlib**: For secure password hashing.*
- ***shutil, os**: For file handling and temporary storage of audio files.*
- ***requests**: For API calls to external services (weather, search).*
- ***datetime**: For timestamping logs and usage metrics.*

5. Interfaces

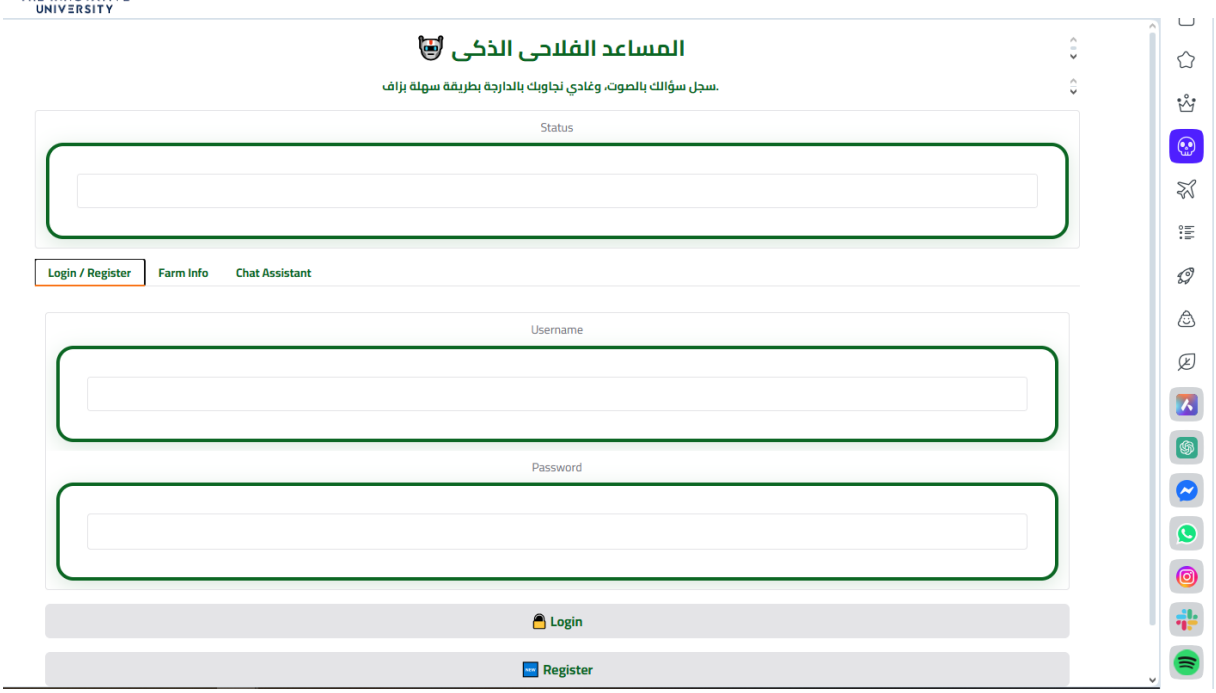
This section outlines the user interfaces of the ZiraaTech application.

5.1 Login/Register Interface

Purpose:
Enables users to authenticate or create an account securely before accessing the system's features.

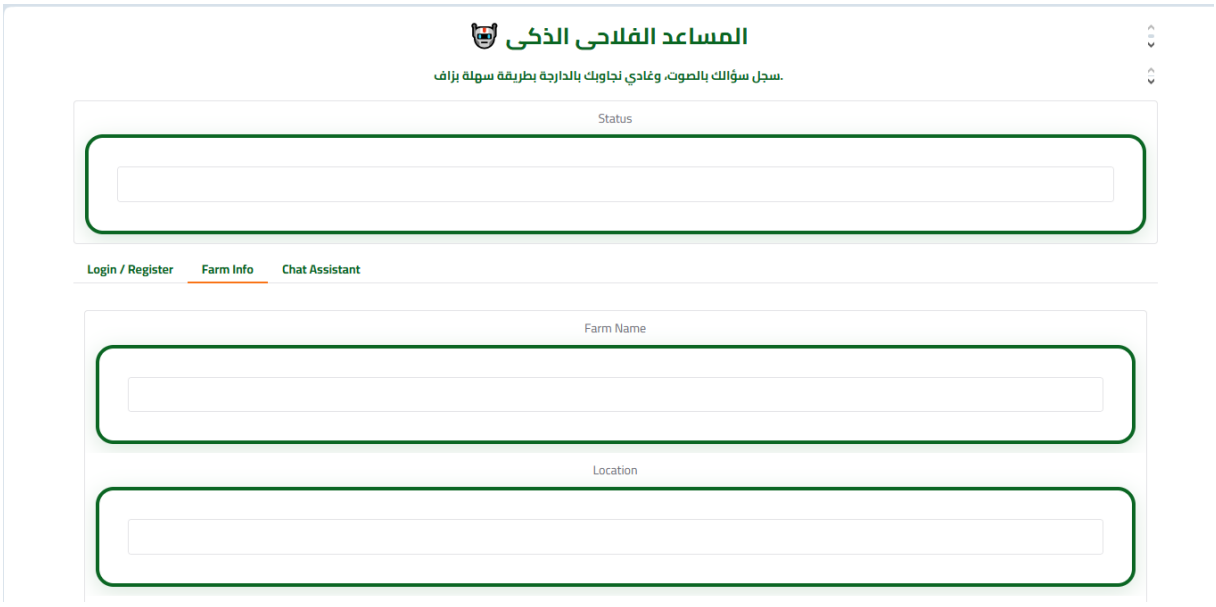
Features:

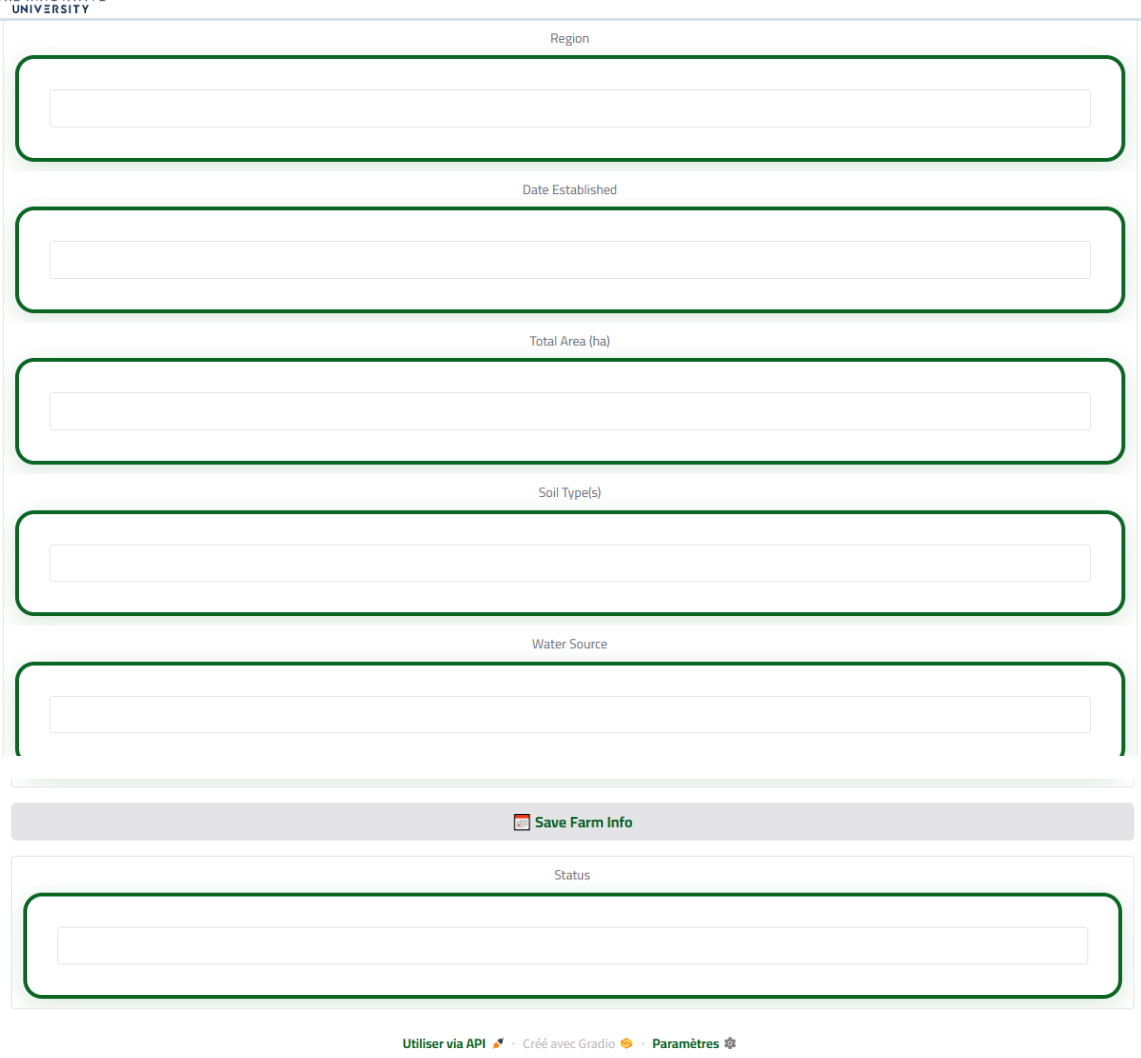
- ***Username & Password Fields**: Collect user credentials.*
- ***Status Box**: Displays feedback on login or registration attempts (e.g. success or error).*



5.2 Farm Info Interface

Purpose:
Collects essential farm-related data from the user to personalize AI responses.



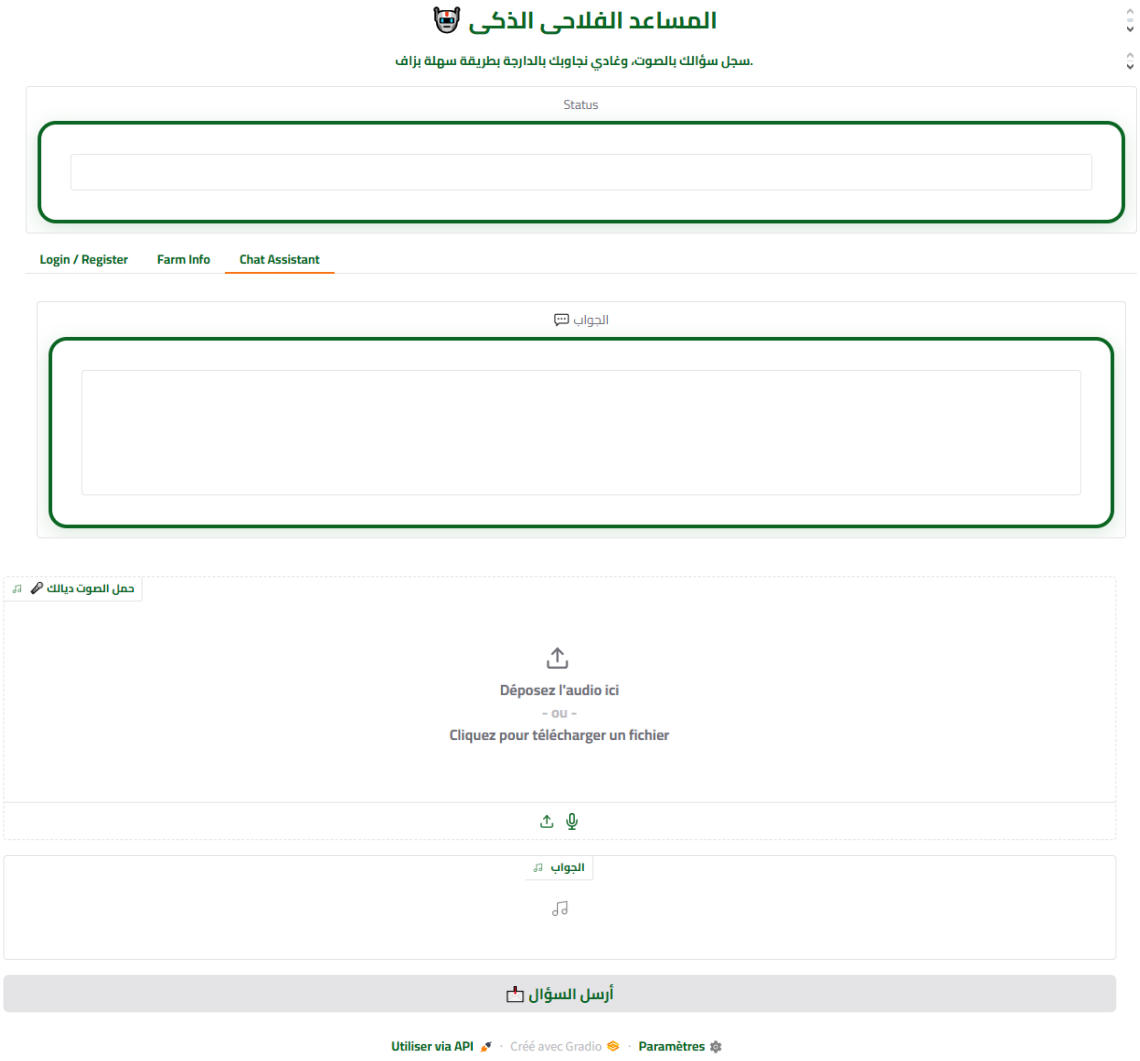


5.3 Chatbot Interface

Purpose:
Facilitates interaction between farmers and the AI assistant, allowing voice queries and receiving answers in Moroccan Darija.

Features:

- ***Audio Upload****: Farmers can record their question and upload it.*
- ***Answer Display****: Text response appears in a large, green-bordered box for readability.*
- ***Voice Playback****: AI answer is also played back via a TTS component, ensuring accessibility for illiterate users.*
- ***Submit Button****: sends the question to the backend for processing.*



6.Performance:

- *Pinecone query latency: <500 ms on average.*
- *Gemini response latency: varies, typically <3 seconds.*

7. Future Improvements

***IoT Integration**: Real-time sensor data (soil, water, temperature).*

***Computer Vision**: Disease and pest detection via images.*

***Expanded Dataset**: More region-specific vocabularies and farming practices.*

***Dynamic Prompting**: Adaptive prompts for seasonal crops and regional variations.*

***Edge Deployment**: For low-connectivity areas using local processing.*

8. Conclusion

ZiraaTech demonstrates a robust, modular AI architecture tailored to the needs of Moroccan farmers. By leveraging advanced AI models (LLMs, RAG, vector search) and ensuring cultural and linguistic adaptation, it empowers farmers with accessible and actionable agricultural advice.

9. Acknowledgments

*We would like to extend our heartfelt thanks to **Hamza Gamouh** and **Hakim Hafidi** for their invaluable support and guidance throughout the development of ZiraaTech. Their expertise, encouragement, and commitment have been instrumental in shaping the project's technical design, implementation, and overall success. We are truly grateful for their generous contributions and unwavering support, which helped us turn our vision into reality.*