MBEYA UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF INFORMATION SYSTEMS AND TECHNOLOGY

ORDINARY DIPLOMA IN INFORMATION AND COMMUNICATION TECHNOLOGY

NTA LEVEL 6

FINAL PROJECT REPORT

PROJECT TITLE: **DEVELOPMENT OF AN AI MODEL TO ASSESS THE IMPACT OF CLIMATE CHANGE ON CROP YIELDS**

SUPERVISOR**: Mr. NCHIA**

NO.	NAME	REGISTRATION NUMBER
01.	WILLIAM M. ELIAS	22101123130057
02.	JOSEPH F. MINDAS	22101123130098

JULY 2025

CERTIFICATION

The undersigned certify that, has read and hereby recommended for acceptance by Mbeya University of Science and Technology (MUST) a projected title "**Development of an AI Model to Assess the Impact of Climate Change on Crop Yields"** In partial fulfillment of the Ordinary diploma in Information and Communication Technology at Mbeya University of Science and Technology (MUST) under the department of information systems and technology in the college of information and communication technology.

………………………………………………….

Supervisor's Name

…………………………….

Signature

…………………..

Date

DECLARATION

We hereby declare that this project report is the result of our original work conducted during our studies under the supervision of **Mr. Nchia.** We affirm that the statements made and the conclusions drawn in this report are based on our independent efforts and findings.

Furthermore, we certify that this work is original, has been prepared under the guidance of our supervisor, and has not been submitted to any other institution for the award of a degree, diploma, or certificate.

WILLIAM M. ELIAS	JOSEPH F. MINDAS
Reg ……………………………………	Reg………………………………………
Signature………………………………	Signature………………………………...
Date: ……………………………….….	Date: …………………………………….

COPYRIGHT

DEDICATION

We dedicate this work to our beloved nation of Tanzania and our precious university, as Mbeya University of Science and Technology (MUST). The work and the effort are all dedicated to the best of our nation and its citizens.

Also, we dedicate this work to our head of department and all facilitators concerned with this work. The people who decided to work with us on this task and all the time they gave us until the project's completion.

Finally, we dedicate this project work to our lovely families who have never gotten tired of inspiring us to make it well, and without forgetting our fellow students who gave us their support and advice till the completion of this project work.

ACKNOWLEDGEMENT

We raise our hands hardly, open our hearts widely, and say thanks to the Creator who gave us power and energy to carry out this project work without becoming ill or injured. Many thanks and gratitude to God who gave us each and everything that has made us who we are today.

Sincerely, we also say thanks to the United Republic of Tanzania, the President, ministers, and all ministries that have given us peace and harmony for us to work on our land without fearing anything. We also say thanks for all the services they gave us to accompany us in this practical training time. The roads were really good and favorable for us to make this project.

We also acknowledge our precious university of the University of Science and Technology (MUST), for all the disturbance we caused to it, but still they endured till the completion of this project. We cannot pay them, but we hope God will make it the best for them. We are glad they aided us in many ways.

Thanks to our department and all the facilitators for their untiring support and the precious time they gave us till the completion of this project. We can’t count the problems we brought to their families for the late coming because of projects, and all the late-night calls we made to them for misunderstood concepts.

Thanks to our families for their generous hearts and understanding minds for the whole project time, they gave us till completion. We are happy to have these families in this world, and we always pray for all families who have assisted us in any way until the completion of this project.

Without forgetting our fellow students all around the university and our fellow classmates for their tremendous presence and support in completing this project. May God bless you and give more in your journey to keep supporting others on their way here. We are glad you crossed our path, and special thanks to you all.

Finally, we drop our thanks to our hearts, minds, and bodies for all the times they were there when needed. We say it was tough, but our hearts never gave up, our minds never stopped thinking, and our bodies never got tired and kept sharing until we had completed this project. We have learned a lot in this project.

ABSTRACT

Effective task management is essential for students to stay organized, meet deadlines, and develop productive study habits. Traditional task management systems, such as calendars and to-do lists, require manual input and lack intelligent adaptability, often leading to missed deadlines, poor time management, and ineffective study planning. This project introduces an AI-powered Student Task Management System, designed to address these challenges through automation, personalization, and intelligent task prioritization.

The system integrates machine learning and natural language processing (NLP) to analyze student behavior, track task completion patterns, and provide personalized recommendations for study schedules. Key features include smart reminders, adaptive scheduling, real-time progress tracking, and AI-driven task prioritization. The system identifies patterns in study habits, detects potential procrastination, and dynamically adjusts reminders and deadlines based on student engagement and task urgency.

This solution provides students with a personalized and intelligent task management assistant, helping them stay on track with their academic goals. By leveraging AI-driven insights, the system enhances productivity, reduces procrastination, and improves overall academic performance, making learning more efficient and structured.

Table of Contents

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
LSTM	Long Short-Term Memory
RF	Random Forest
CNN	Convolutional Neural Network
TMA	Tanzania Meteorological Authority
NASA	National Aeronautics and Space Administration
IPCC	Intergovernmental Panel on Climate Change
API	Application Programming Interface
CSV	Comma-Separated Values

EDA: Exploratory Data Analysis

LIST OF SYMBOLS

T = Temperature (°C)

RH = Relative Humidity (%)

SR = Solar Radiation (MJ/m^2)

WS = Wind Speed (m/s)

P = Precipitation (mm)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1.0 INTRODUCTION

1.1 Background Information

Tanzania's economy is heavily reliant on agriculture, which employs over 65% of its population. The Mbeya region is a major contributor to national food security, particularly for maize, beans, and rice. However, climate change is disrupting traditional farming practices through unpredictable weather events such as delayed rains, heat waves, and storms. These changes reduce yield and increase food insecurity.

To address this, a smart system that utilizes Artificial Intelligence (AI) to predict crop yield under changing climate conditions has been developed. The system automates the retrieval and analysis of weather variables—rainfall, temperature, humidity, wind speed, and solar radiation—based on user-inputted location and date. The AI models then analyze this data to forecast crop yields and display insights on a farmer-friendly interface. The system supports automated data integration from APIs and includes map-based visualizations using Google Maps.

The increasing impact of climate change on agriculture is one of the most pressing challenges faced by the global community. Climate variability, including rising temperatures, changing rainfall patterns, and increased frequency of extreme weather events, directly affects crop productivity. This project aims to develop an AI-driven model that will assess and predict how climate change will impact crop yields in various regions around the world. The model will use historical climate data, crop yield data, and climate projections to forecast future trends in agricultural productivity under different climate scenarios.

The core idea of this project is to utilize machine learning, specifically regression and classification models, to analyze large datasets that involve climate variables (such as temperature, precipitation, and CO2 levels) and crop yield data (such as yield per hectare of crops like wheat, maize, and rice). By understanding these relationships, the AI model can make predictions about how crop yields will change in the future as the climate continues to change. This project will benefit farmers, policymakers, and agronomists, who need accurate data to make decisions regarding crop management, resource allocation, and climate adaptation strategies.

1.2 Problem Statement

The unpredictability of climate change and its complex interactions with crop growth create challenges in accurately predicting its impact on yields. **Agricultural productivity** is increasingly threatened by climate change, with unpredictable weather patterns leading to significant yield fluctuations. **Existing predictive models** cannot often accurately capture the **intricate** relationships between **climate variables** and **crop yields.**

The problem being addressed is the lack of accurate, predictive models for assessing the impact of climate change on crop yields, leading to increased food insecurity and loss of agricultural productivity. The agricultural sector in Mbeya, Tanzania, faces numerous challenges due to climate change, including:

- Unpredictable weather patterns leading to inconsistent crop yields.
- Limited access to climate prediction tools for smallholder farmers.
- Existing forecasting models lack localized, real-time data integration.
- Inadequate climate adaptation strategies due to a lack of precise predictions.

Therefore, the proposed AI-based solution will bridge this gap by leveraging data-driven insights to enhance decision-making in agriculture. Current solutions, such as conventional weather forecasts and government-led agricultural advisories, do not offer precise, farm-level recommendations. This project will introduce a robust, AI-powered predictive model tailored for Tanzania's diverse agricultural landscapes.

Current solutions include manual analyses of climate and crop yield data, which can be labor-intensive and error-prone. Some institutions use basic statistical models, but these often fail to capture the complexity of the changing climate and its effects on agriculture. This project introduces an AI-powered model that uses machine learning algorithms to predict crop yields in the face of changing climate conditions. Unlike traditional models, this system will be adaptive and continuously improve with more data.

1.3 Project Objectives

The objective of using an AI model to assess the impact of climate change on crop yields is to precisely quantify and predict how changing climate conditions like temperature, precipitation patterns, and extreme weather events will affect agricultural productivity in the Mbeya region.

General Objective:

To develop an AI-powered decision support system for assessing and predicting the impact of climate change on maize, beans, and rice yields in Mbeya.

Specific Objectives:

- To gather and integrate climate, crop, and soil data from primary and secondary sources.
- To apply and compare machine learning algorithms (Random Forest, LSTM, CNN) for yield prediction.
- To build a React-based frontend and a Python backend to manage data and model operations.
- To automate climate data extraction via APIs like NASA POWER.

1.4 Project Significance

i. To address the Growing Threat of Climate Change on Agriculture.

   The AI model will predict the effects of climate change on crop yields in different regions, helping policymakers and farmers prepare in advance. By understanding future risks, governments can develop better climate adaptation strategies and ensure food security.

ii. To support Farmers with Data-Driven Decision Making.

   The AI system will provide personalized crop yield forecasts based on temperature, rainfall, CO_2 levels, and soil conditions. Farmers can make informed decisions about when to plant, which crops to grow, and how to adjust farming techniques to maximize productivity.

Example: In Africa, maize yields are declining due to rising temperatures. An AI system can recommend alternative drought-resistant crops like sorghum or millet, reducing the risk of total crop failure.

iii. Also, it enhances Food Security and Reduces Hunger.
AI predictions can help optimize food production by identifying the best climate-resilient crops for each region. Governments and agricultural organizations can use AI-generated insights to plan food production and distribution.

iv. Another is to support Sustainable Agriculture and Climate Adaptation.
AI predictions can guide sustainable farming techniques like Crop rotation strategies to maintain soil fertility, Smart irrigation recommendations to conserve water, and Precision agriculture to optimize fertilizer use.

v. To promote the economic Benefits for Farmers and Agricultural Businesses.
By reducing uncertainty, farmers and agricultural businesses can plan, leading to higher profits and better market stability. AI-based recommendations can help reduce input costs by optimizing fertilizer, water, and seed usage.

1.5 Project Scope

The system was implemented for three key crops as maize, beans, and rice in Mbeya, Tanzania. Key components include:

- A user interface built with React for data input and dashboard visualization,
- A backend with SQLite and Python-based AI models (Random Forest, LSTM, CNN),
- Integration with APIs (NASA, TMA) for climate data retrieval,
- Reporting tools for visualizing trends and exporting results as PDF/CSV.

1.6 Project Limitations

- Internet access is required to fetch climate data via APIs.
- Yield prediction accuracy is influenced by the completeness and quality of records.
- The system is limited to the Mbeya Region and three crops in its current version.
- Accessibility support for visually impaired or illiterate users is not yet implemented.

CHAPTER 2

2.0 LITERATURE REVIEW

2.1 Introduction

This chapter provides an overview of existing knowledge on the impacts of climate change on agriculture in Tanzania and Africa, reviews conventional and AI-based yield prediction methods, and identifies research gaps that this project aims to address.

2.2 Climate Change and Agriculture in Tanzania and Africa

The Intergovernmental Panel on Climate Change (IPCC) highlights that developing countries, particularly in Sub-Saharan Africa, are highly vulnerable to climate change. Tanzania's agricultural output has been threatened by irregular rainfall patterns, temperature shifts, floods, and prolonged droughts. Mbeya, as a highland region, faces unique challenges such as excessive rainfall in some areas and prolonged dry seasons in others, adversely affecting the productivity of maize, beans, and rice.

Studies like that of Mboya et al. (2020) indicate yield reductions of up to 30% due to climate variability. Traditional coping mechanisms such as shifting planting calendars are proving insufficient, necessitating a data-driven, predictive approach to farm planning.

2.3 Traditional Methods for Yield Forecasting

Historically, crop yield forecasting relied on statistical models such as linear regression, time-series analysis, and trend extrapolation. While these approaches are computationally simple and interpretable, they are limited by assumptions of linearity and stationary behavior in data.

Models such as the FAO's AquaCrop or DSSAT simulate crop growth under varying conditions but require manual calibration and significant domain knowledge. They do not scale well with large, diverse datasets and often fail to capture complex climate-crop interactions.

2.4 The Role of Artificial Intelligence in Agriculture

Machine Learning (ML) and Deep Learning (DL) methods offer powerful tools for capturing non-linear, multi-dimensional patterns in agricultural datasets. Random Forests (RF) provide good performance on tabular data and are resistant to overfitting. LSTM networks, a type of Recurrent Neural Network (RNN), are ideal for modeling sequential data, capturing dependencies across planting seasons.

CNNs are primarily used for image data but have also been successfully applied to structured data by treating it as spatially distributed. Several projects have utilized AI for early warning systems, drought monitoring, and yield estimation globally, with promising results.

2.5 Related Works

2.5.1. The Agricultural Model Intercomparison and Improvement Project (AgMIP)

The Agricultural Model Intercomparison and Improvement Project (AgMIP) is a global initiative aimed at understanding the impact of climate change on agriculture through advanced climate and crop modeling techniques. This system integrates climate models, economic models, and crop simulation models to assess future agricultural productivity under different climate scenarios. By analyzing historical weather patterns, soil conditions, and farming practices, AgMIP provides long-term projections for various crops, including maize, wheat, and rice. The project has been applied in Tanzania, Kenya, and Ethiopia, where it has helped governments develop climate adaptation policies and guide farmers in selecting climate-resilient crops. One of AgMIP's most significant contributions is its ability to compare multiple climate and crop models, improving the accuracy of predictions for specific regions. In Tanzania, it has been particularly useful in analyzing maize production trends in response to rising temperatures and shifting rainfall patterns. Challenges on **AgMIP (Agricultural Model Intercomparison and Improvement Project) are:-**

- Lacks **real-time localized weather updates**, making it less useful for short-term decision-making.
- Does not provide **direct farmer support**, as it is mostly used for research purposes.

2.5.2. Climate Change Agriculture and Food Security (CCAFS)

The Climate Change, Agriculture, and Food Security (CCAFS) program, developed by CGIAR (Consultative Group on International Agricultural Research), focuses on sustainable agricultural practices in response to climate change. The system integrates satellite imaging, machine learning, and on-ground research to predict climate variability and its impact on food production. One of the major benefits of CCAFS is its role in promoting climate-smart agriculture (CSA), which includes drought-resistant crop varieties, water-efficient irrigation systems, and agroforestry practices. In Tanzania and other East African countries, CCAFS has been instrumental in advising farmers on optimal planting seasons based on climate forecasts. The program has also contributed to policy development by working with governments to design national climate adaptation plans that emphasize sustainable land management and food security initiatives. Additionally, CCAFS uses big data analytics to help farmers make real-time decisions about their crops based on current and predicted climate conditions. The challenges facing **CCAFS (Climate Change, Agriculture and Food Security Program) are:-**

- Provides **long-term climate projections** but lacks **real-time early warning alerts** for farmers.
- Limited accessibility for **smallholder farmers**, as it is mainly used for policy and research.

2.2.3. CropWatch (China's AI-Based Crop Monitoring System)

CropWatch, developed by the Chinese Academy of Sciences (CAS), is one of the most advanced AI-powered agricultural monitoring systems in the world. This system utilizes satellite imaging, remote sensing, and AI-based analytics to track crop health, soil moisture, and climate conditions on a global scale. Unlike traditional weather forecasting tools, CropWatch provides real-time insights into crop conditions, allowing governments and farmers to take proactive measures to prevent crop failures. The system gathers data from NASA, ESA, and Chinese satellite networks, making it a reliable source for climate and agricultural data. In Tanzania, CropWatch has been used to monitor maize, rice, and coffee production, helping farmers and policymakers predict drought conditions and excessive rainfall. By analyzing data from satellite images, CropWatch can detect early signs of crop stress, such as water shortages or disease outbreaks, and provide early warnings to farmers. This allows agricultural experts to develop effective intervention strategies before major losses occur. But the challenges on CropWatch like**;-**

- ❖ Relies heavily on **satellite data**, which may not capture **microclimate variations** affecting small farms.
- ❖ Limited **integration with local soil and farmer-reported data**, reducing accuracy.

2.2.4. AquaCrop (FAO's Crop Simulation Model)

AquaCrop, developed by the Food and Agriculture Organization (FAO), is a crop simulation model designed to predict how climate conditions, water availability, and soil fertility affect crop yields. Unlike many other climate models that focus solely on weather predictions, AquaCrop specializes in water-use efficiency by analyzing how different levels of irrigation and rainfall impact crop growth. This makes it particularly useful for water-scarce regions like Tanzania, where erratic rainfall and prolonged droughts threaten food security. AquaCrop has been widely used in East Africa to guide farmers in efficient irrigation scheduling and soil moisture management. In Tanzania, agricultural researchers have used AquaCrop to assess rice and maize farming under changing climate conditions, helping farmers determine the best irrigation practices to maximize yields. The model is particularly beneficial for smallholder farmers, as it provides easy-to-understand recommendations on how to optimize water use and select drought-resistant crops. But it has some challenges, which are:-

- ❖ Focuses mostly on **water-related crop stress** but does not consider **pest outbreaks or extreme climate events**.
- ❖ Requires **detailed input parameters**, making it difficult for **non-experts** to use.

2.2.5. AI for Earth (Microsoft's AI-Based Climate Prediction System)

Microsoft's AI for Earth initiative is a cutting-edge AI-powered climate prediction system designed to analyze large-scale environmental and agricultural data. This system uses machine learning, deep learning, and cloud computing to track climate patterns, soil conditions, and crop growth with high accuracy. One of its primary functions is to help governments, researchers, and farmers understand how climate change will impact agriculture in the coming years. AI for Earth has been applied in Tanzania, Kenya, and Ethiopia, where it assists in predicting droughts, optimizing irrigation systems, and recommending adaptive farming strategies. The system also integrates real-time data from drones, IoT sensors, and satellites to provide farmers with localized climate insights. In Tanzania, AI for Earth has been used to analyze coffee and maize production trends, helping farmers make informed decisions about planting times and crop selection. The AI-driven platform also enables governments to design policies that support climate resilience in agriculture, ensuring that food production remains stable despite climate variability. Gaps on **AI for Earth (Microsoft's AI-powered Climate Initiative)**

- Requires **high computing power and internet access**, making it less useful for rural farmers.
- Mostly focused on **global-scale analysis**, lacking **localized, region-specific predictions**.

2.6 Deficiencies in Existing Systems

Several gaps exist in the current systems and studies concerning climate change and crop yield prediction:

Despite the advancements in AI, satellite imaging, and climate modeling, existing systems for assessing climate change and crop yields still have several limitations that hinder their effectiveness, especially in developing countries like Tanzania. One of the major deficiencies is the **lack of localized and high-resolution climate data.** Many of these systems, such as **AgMIP** and **CropWatch**, rely on **global climate models (GCMs)** and satellite data that provide broad-scale climate predictions. However, these models often fail to capture the microclimate variations that influence local crop productivity. For example, rainfall and temperature fluctuations in different regions of Tanzania—such as the semi-arid Dodoma region versus the humid coastal Tanga region—require localized predictions rather than generalized global forecasts. Without granular climate data, smallholder farmers may receive inaccurate or irrelevant recommendations, making it difficult for them to adapt effectively.

Another critical **limitation is the high cost and technical complexity of these systems**. Advanced platforms like **AI for Earth** and **CropWatch** use sophisticated AI models, remote sensing technologies, and cloud-based computing, which demand substantial financial investment and technical expertise. Many smallholder farmers in Tanzania lack access to high-speed internet, AI-powered tools, or the necessary training to interpret complex climate data. As a result, while these systems may offer cutting-edge insights, their benefits do not always reach rural and resource-limited farming communities. Furthermore, many existing climate prediction models do not integrate indigenous farming knowledge, which is crucial for building culturally relevant adaptation strategies. Traditional farming practices in Tanzania, such as rotational cropping and rain-fed agriculture, often contain valuable climate resilience techniques that modern AI-based systems fail to consider.

Another deficiency is the **limited ability of these systems to predict sudden and extreme climate events**. While models like AquaCrop and CCAFS can estimate long-term trends in rainfall, temperature, and crop growth, they are less effective in predicting short-term extreme events, such as flash floods, heatwaves, or unexpected droughts. This gap in prediction capabilities puts farmers at risk of sudden crop failures, as they may not receive timely warnings to take preventive action. Additionally, some systems do not account for pests and diseases, which are becoming more severe due to climate change. For example, the recent fall armyworm outbreak in Tanzania, which devastated maize crops, was not accurately predicted by most AI-based climate models that primarily focus on weather-related impacts rather than biological threats.

Lastly, there is a **lack of interoperability and real-time updates across different platforms**. Many systems operate in isolation, making it difficult for policymakers, researchers, and farmers to combine insights from multiple models for better decision-making. For example, while **AgMIP** focuses on crop simulation models, **CropWatch** specializes in remote sensing, and AI for Earth applies machine learning, these systems often do not seamlessly integrate their data. As a result, farmers may receive conflicting or incomplete climate information, reducing the overall effectiveness of these technologies. To maximize impact, there is a need for a unified, AI-driven platform that combines real-time climate data, crop simulation models, and predictive analytics in a way that is accessible and actionable for local farmers in Tanzania.

2.7 Summary

Climate change continues to impact agriculture in Tanzania, especially in regions like Mbeya. Existing systems are insufficient for localized, real-time decision-making. This study contributes a novel, AI-driven, end-to-end application to assess and predict crop yields under changing climate conditions, bridging the gap between raw data and actionable insight for local farmers and policymakers.

CHAPTER 3

3.0 METHODOLOGY

3.1 Introduction

The AI Crop Yield Assessment System is a web-based application that allows users (farmers, agricultural officers) to input location and date data to retrieve and analyze relevant climate conditions and crop performance. The system applies machine learning (Random Forest, CNN) and deep learning (LSTM) models to predict yields for maize, rice, and beans in the Mbeya region. Agile methodology was used to guide the system's development, ensuring adaptive, iterative delivery and continuous stakeholder feedback.

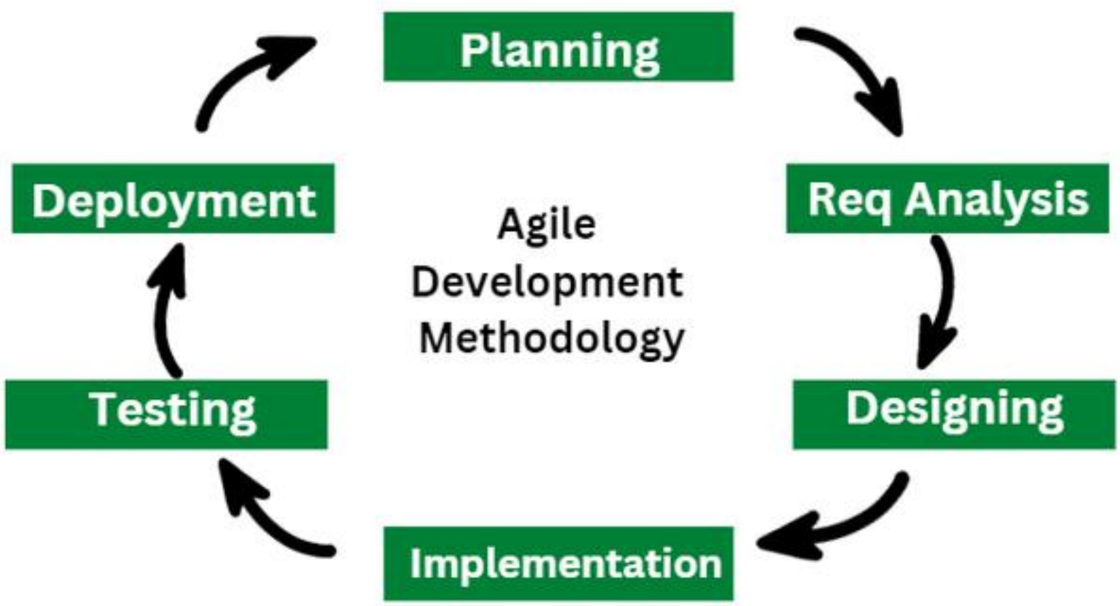


Figure 3.1: Agile Methodology Development

3.2 Agile Methodology Overview

Agile methodology is an iterative and incremental software development approach that emphasizes flexibility, collaboration, and user-driven refinement. For this project, we followed Agile principles using the Scrum framework. Each Sprint delivered a working prototype component, improving upon feedback from domain experts and end users (farmers, agriculture officers, and IT supervisors).

3.3 Agile Implementation Approach

Agile development was structured across sprints, each representing a phase in the system lifecycle:

Phase 1: Project Initiation & Requirement Gathering, Sprint 1: Planning and Stakeholder Engagement

- Conducted interviews with farmers, climate analysts, and agricultural officers in Mbeya
- Defined functional (prediction, location input, visualization) and non-functional (ease of use, offline access) requirements
- Built the product backlog

Deliverables:

- Feature backlog
- Technology stack and development tool selection (Python, Streamlit, TensorFlow)

Phase 2: System Design Sprint 2: Architectural and UI/UX Planning

- Defined 3-layer architecture: Input, Processing, Output
- Designed interactive mockups of the Streamlit interface

- Created system diagrams and data flowcharts

Deliverables:

- System architecture blueprint
- Visual mockups

Phase 3: Development & Model Integration Sprint 3: Frontend Development

- Built farmer input interface and visualization dashboard (rainfall, yield trends)

Sprint 4: Backend Development

- Integrated NASA/TMA API for climate data
- Connected a SQLite database for storing crop data

Sprint 5: AI Model Development

- Cleaned and normalized datasets
- Trained LSTM, CNN, and Random Forest models using TensorFlow/Keras
- Evaluated performance (RMSE, R^2) and deployed best model

Deliverables:

- Fully functioning frontend/backend
- Integrated AI yield prediction models

Phase 4: Testing & Deployment Sprint 6: Internal Testing

- Performed usability and integration testing with test data and local officers
- Fixed UI bugs, tuned model performance

Sprint 7: Live Demo and Optimization

- Deployed the app for supervisor and farmer testing
- Collected final user feedback

Deliverables:

- Demonstrable application with user feedback

Phase 5: Maintenance & Continuous Improvement Sprint 8: Final Adjustments and Enhancements

- Improved model accuracy with additional data
- Planned new features: mobile support, Swahili translation, offline mode

3.4 Advantages of Agile Methodology for this Project

- It helps with Fast Prototyping & Delivery: Agile enables quick rollout of key features such as climate-based predictions early in development. Streamlit allowed rapid testing of UI before final model integration.
- It may influence the Changes in Flexibility: New requirements from stakeholders (e.g., report exports, map integration) were integrated mid-development without disrupting the process.
- High Quality Output: Continuous evaluation of models ensured accurate yield predictions, with LSTM outperforming others after retraining.
- User Collaboration: Farmers provided input on crop selection, and agricultural officers tested usability, ensuring real-world needs were met.
- Risk Reduction: Each Sprint delivered measurable progress. If one model failed to meet accuracy thresholds, alternatives (e.g., switching to CNN or tuning hyperparameters) were explored.

CHAPTER FOUR

4.0 DATA COLLECTION

4.1 Introduction

Data collection is the systematic process of gathering and measuring information to support decision-making, training AI models, and assessing climate impacts. In this project, both primary and secondary data sources were used to ensure a robust and comprehensive dataset that reflected real-world agricultural conditions in Mbeya.

4.2 Types of Data Collected

- Primary Data: Obtained directly through field surveys, interviews, and questionnaires conducted with farmers and agricultural officers in Mbeya.
- Secondary Data: Gathered from institutions like NASA, TMA, the Ministry of Agriculture, and IPCC.

4.3 Primary Data Collection Methods

4.3.1 Interviews Conducted face-to-face interviews with farmers and agricultural officers in various wards (e.g., Uyole, Forest Mpya). These conversations helped identify climate-related challenges, yield changes, and preferred farming practices.

Why Interviews?

- Gathers firsthand insights from farmers
- Builds trust and contextual understanding
- Helps define real-life use cases for the AI tool

4.3.2 Observation On-site observations provided insights into farming conditions, irrigation practices, and crop health. This visual data supported model validation.

4.3.3 Questionnaires: Distributed structured forms to 40+ farmers using paper and Google Forms. Questions focused on crop type, yield history, climate impact, and fertilizer usage.

Why Questionnaires?

- Allows large-scale data collection
- Enables quantitative analysis
- Easy to distribute in rural and urban areas

The following is a sample of the questionnaire.

How many years have you been involved in agriculture?

- ○ Less than 1 year
- ○ 1-5 years
- ○ 6 -10 years
- ○ More than 10 years

Which of the following crops do you grow or work with?

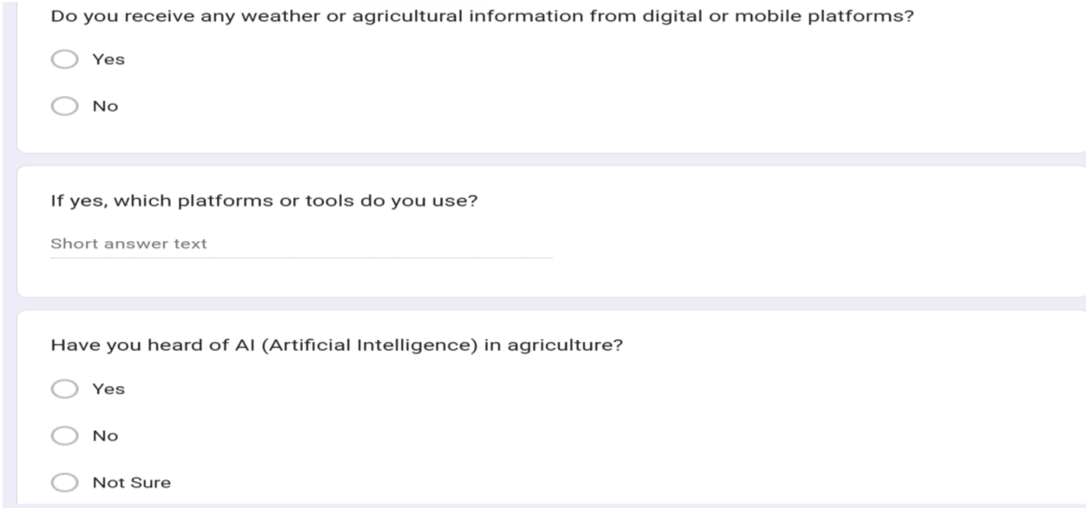
- ☐ Maize
- ☐ Rice
- ☐ Beans

On average, what is your yield per acre for each of the crops?

- ○ Less than one tonne
- ○ 1-2 tonnes
- ○ More than 3 tonnes

Have you noticed a change in yield over the past 5–10 years?

- ○ Yes
- ○ No



Do you receive any weather or agricultural information from digital or mobile platforms?

- ◯ Yes
- ◯ No

If yes, which platforms or tools do you use?

Short answer text

Have you heard of AI (Artificial Intelligence) in agriculture?

- ◯ Yes
- ◯ No
- ◯ Not Sure

4.4 Secondary Data Collection

Sources:

- NASA POWER: temperature, humidity, solar radiation
- TMA: weather station data
- Ardhi University: soil type classification
- Ministry of Agriculture: annual yield stats
- IPCC, NOAA: climate trends and projections

These datasets were used to train, test, and validate AI models.

Why Secondary Data?

- Provides historical and large-scale context
- Enables integration of remote sensing and climate projections
- Complements farmer-level inputs

4.5 DATA ANALYSIS

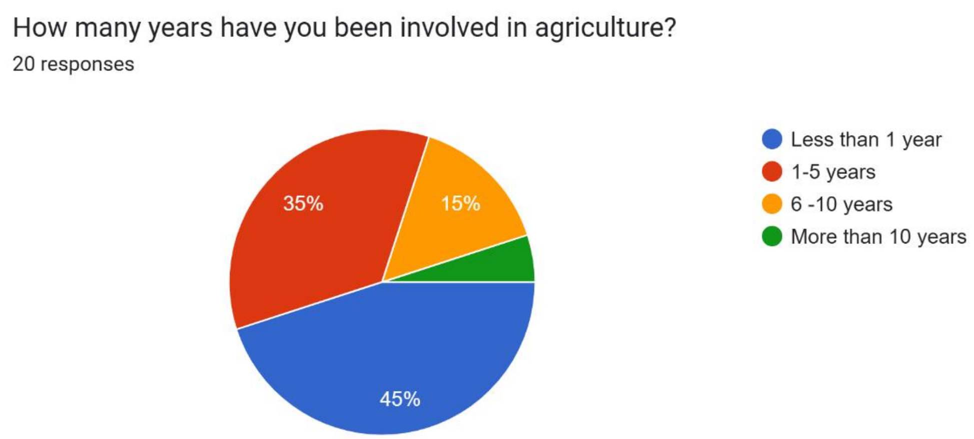
Data analysis involves examining, cleaning, transforming, and interpreting data to extract useful information, identify patterns, trends, and insights. In the context of developing the AI Crop Yield Predictor, data analysis helps in making informed decisions to ensure optimal system performance. Statistical techniques and visualization tools were used to understand the data collected for effective training and operation of the system.

The method used for this project is data visualization:

- Enhanced Understanding. Presents complex data in an easy-to-understand format.
- Effective Decision. Making. Enables stakeholders to quickly grasp insights and trends.
- Increases Engagement. Makes data more engaging and compelling for users and policymakers.

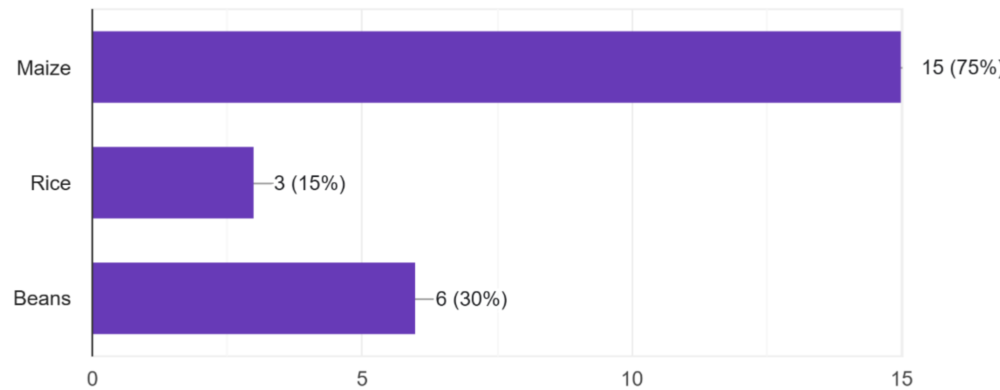
4.5.1 Primary Data Analysis

For questionnaire responses:



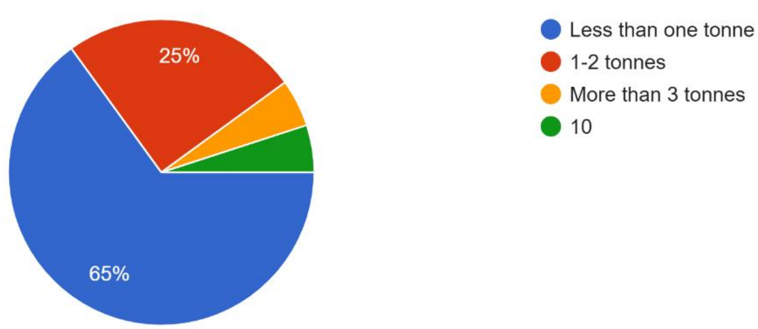
Which of the following crops do you grow or work with?

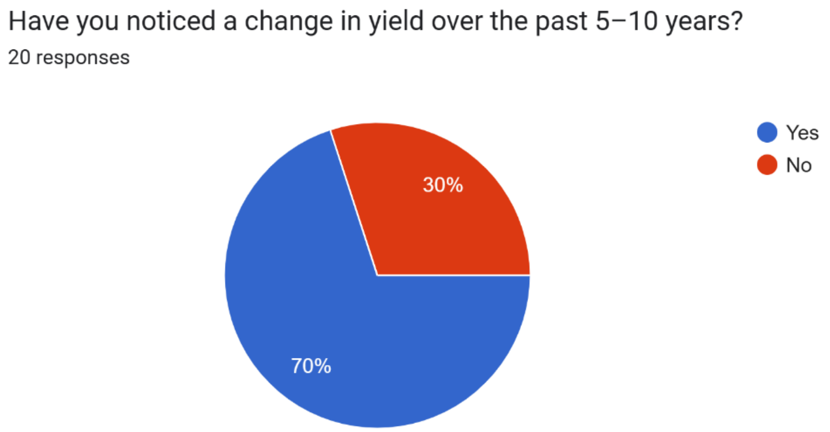
20 responses

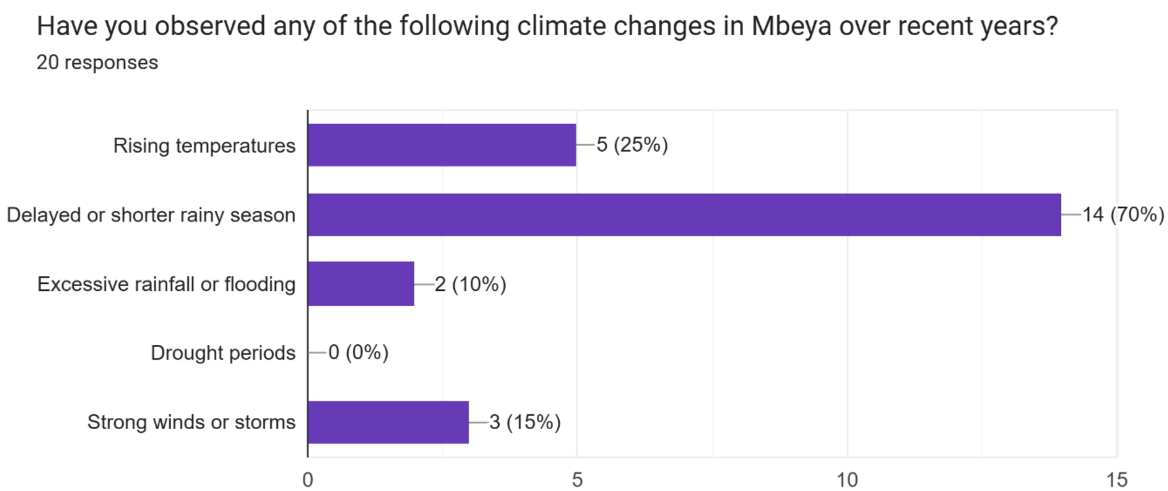


On average, what is your yield per acre for each of the crops?

20 responses







Therefore, from the results, we observed that most farmers are familiar with climate impacts on their crops and showed interest in using technology for prediction and decision-making.

4.5 SYSTEM DESIGN

4.5.1 Introduction

With the increasing impact of climate change on agriculture, AI-driven tools can now support farmers and policymakers in predicting yields and understanding weather-related risks. By combining user interfaces, remote data APIs, and machine learning models, it is possible to build web applications that assess the impact of climate on maize, rice, and bean production. This system bridges traditional farming with digital decision support systems for sustainability.

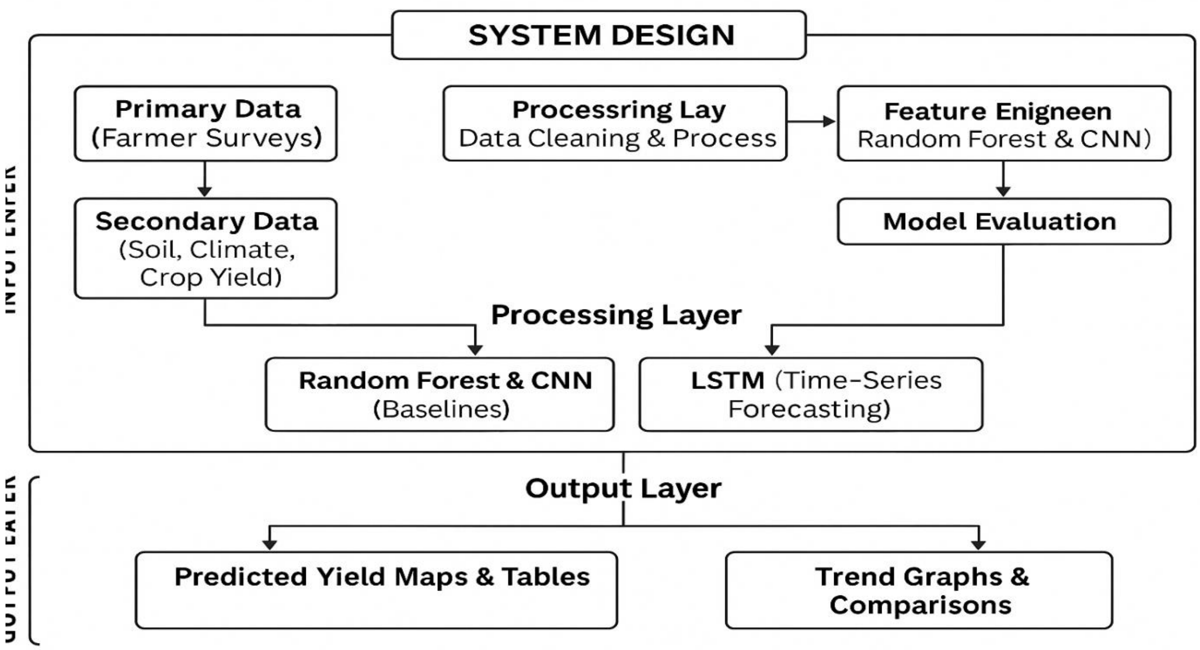


Figure 4.5.1 System Design

4.5.2 Data Preparation

Historical climate and yield data from farmers and institutional sources were consolidated. This data was cleaned, formatted, and labeled according to crop type and regional location to support supervised learning for model training.

- Crops: Maize, Rice, Beans
- Region: Mbeya District, Tanzania
- Variables: Rainfall, Temperature, Soil Type, Historical Yield

4.5.3 Data Processing

Preprocessing Steps:

- Cleaning: Removed missing/erroneous records
- Normalization: Scaled features between 0–1
- Transformation: Applied time series formatting for LSTM input

Feature Engineering:

- Extracted month, year, and season indicators
- Created derived features like average rainfall per season

Data Splitting:

- 80% for training
- 10% for validation
- 10% for testing

4.5.4 Model Development

Models Used:

- LSTM: for time series prediction (rainfall-yield sequences)
- Random Forest: for categorical yield classification
- CNN: for remote sensing image-based prediction (if imagery is used)

Why LSTM?

- Captures sequential dependencies in temporal data
- Handles seasonal patterns in rainfall/yield

Why Random Forest?

- Easy to interpret
- Effective on tabular data

Why CNN?

- Useful for extracting features from satellite images (if integrated)

4.5.5 Model Validation

Techniques used:

- Early stopping: to prevent overfitting
- Validation set: to monitor loss/accuracy
- Model checkpoints: best model saved automatically

Metrics:

- RMSE: Root Mean Square Error
- R^2 Score: Variance explained

4.5.6 Model Evaluation

Sample Result:

- Best RMSE: 0.23 tons/ha
- R^2 Score: 0.88 (88% accuracy)

These results were achieved using the LSTM model after fine-tuning. Visualization of “Predicted vs Actual Yield” was used to confirm trend alignment.

4.5.7 Deployment

Streamlit Web App Deployment

- Developed an interactive web interface
- Integrated the model with the backend
- Users input location/date → receive predicted yield + trends

Backend:

- Model served via Flask or Streamlit backend
- SQLite is used for local storage

Hosting:

- Deployed locally and available on the cloud for presentation

4.6 SYSTEM FUNCTIONALITY

The AI Crop Yield Assessment Tool allows users to enter a location and date, fetches relevant weather data, and generates a prediction on maize, beans, or rice yields. Additional features include:

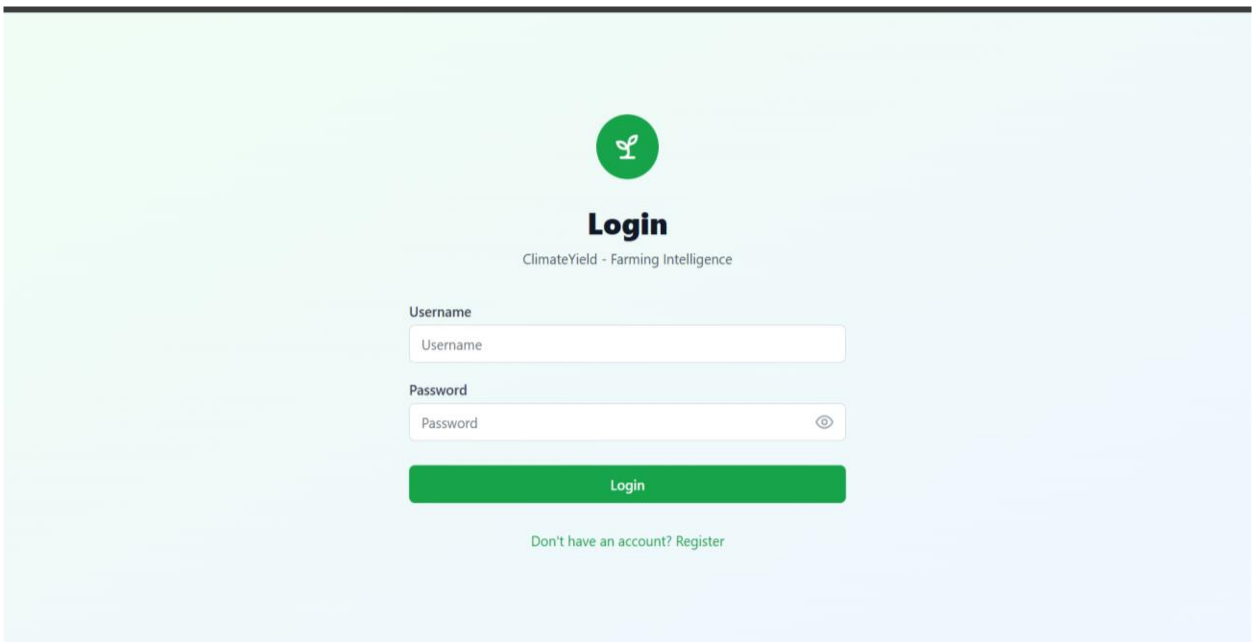


Figure 4.6.1 Login page.

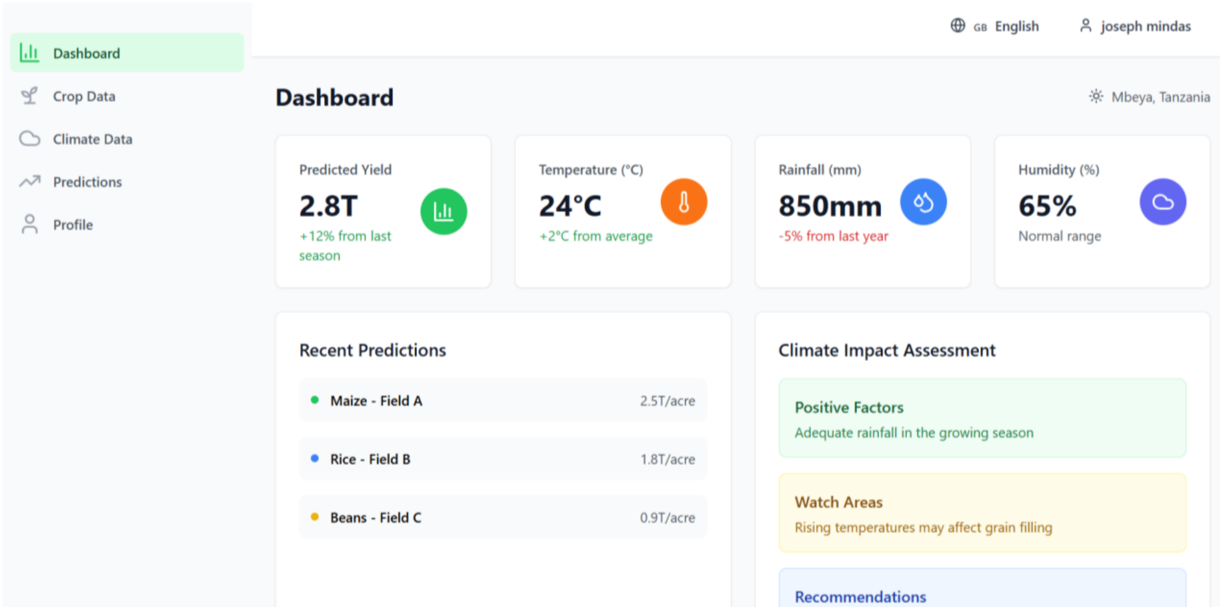


Figure 4.6.2 Dashboard page.

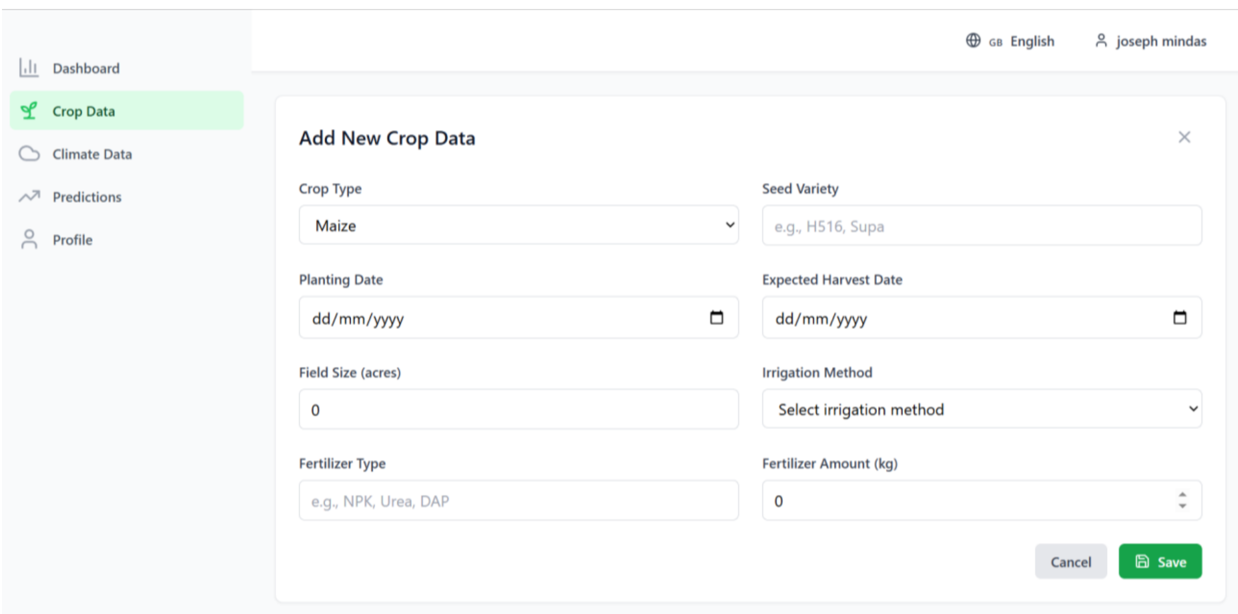


Figure 4.6.3 Crop data input page for farmers or users.

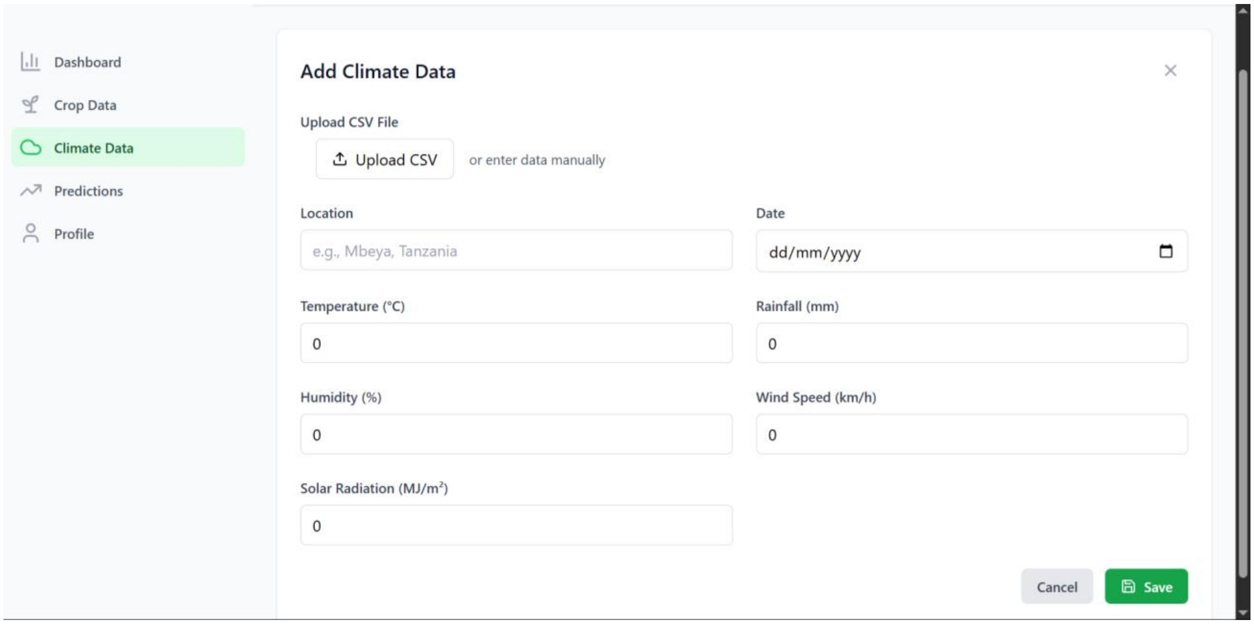


Figure 4.6.4 Climate data input for farmers or user's but all conditions are filled automatically when the user enters the location and date.

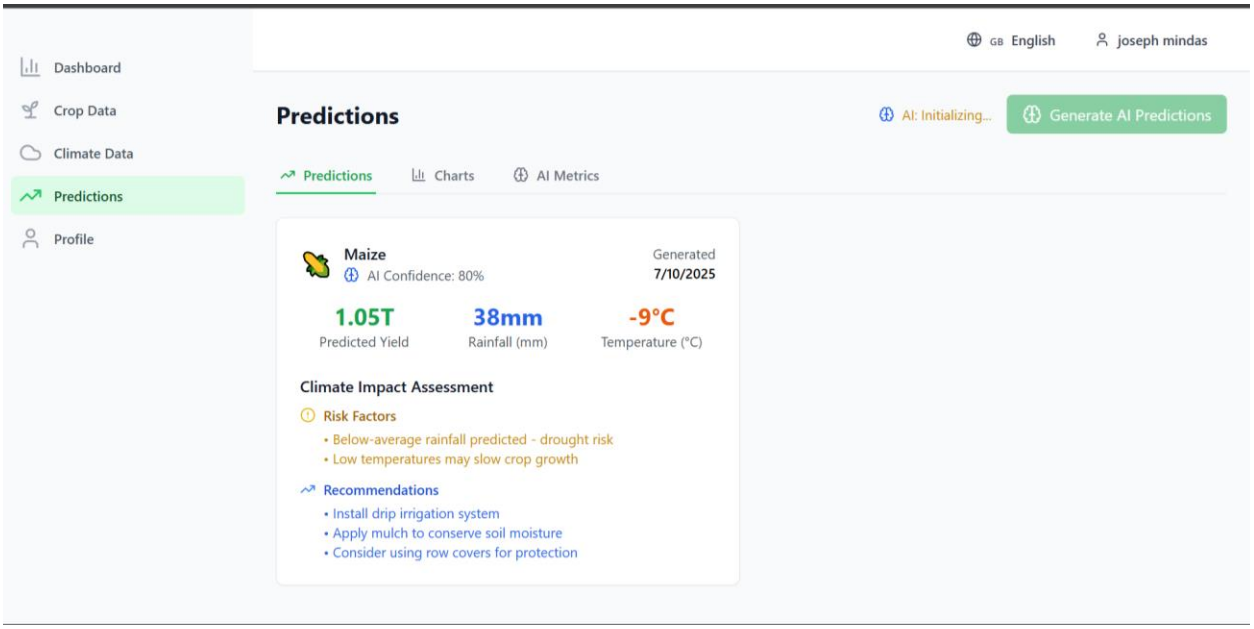


Figure 4.6.5 Predictions page.

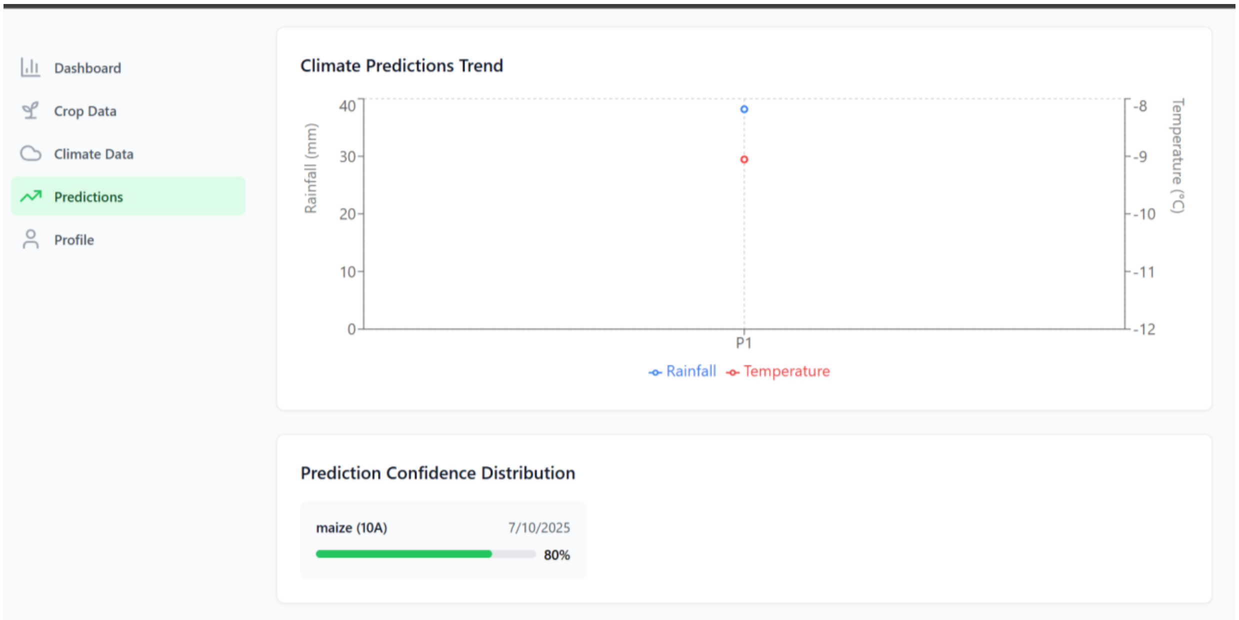


Figure 4.6.6. Predictions on chart trends.

4.7 DISCUSSION

The findings confirm that integrating climate data with advanced machine learning models can improve prediction accuracy for smallholder farmers. The localized approach helped tailor predictions for Mbeya’s unique climate zones.

The LSTM model, particularly, was highly effective due to its design for sequential data. The system’s visual outputs and API integrations make it a useful decision-support tool. However, additional features like mobile support, multilingual UI, and inclusion of pest/disease data are recommended for broader impact.

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The project successfully designed and implemented an AI-based system to assess and predict the impact of climate change on maize, beans, and rice yields in the Mbeya region. By incorporating historical and real-time data from reliable sources and applying robust machine learning models, the system demonstrated improved accuracy in yield forecasting.

The system also facilitated ease of use through a clean, localized interface with dynamic reports, visual dashboards, and map-based insights. LSTM models proved particularly effective due to their ability to model long-term climate-yield dependencies.

This tool presents an innovative contribution to climate-smart agriculture, enabling data-driven decision-making by farmers, researchers, and policymakers.

5.2 Recommendations

Based on the study findings, the following recommendations are proposed:

- Expansion of Crop Coverage: Extend the model to include more crops relevant to Tanzania and neighboring regions.
- Localization and Language Support: Incorporate Swahili and other local languages to enhance accessibility.
- Offline Mode and Mobile App: Develop an Android-based version to reach farmers with limited internet access.
- Integration of Pest and Disease Data: Incorporate external data sources such as pest outbreaks or irrigation practices for holistic yield prediction.
- Policy Integration: Engage stakeholders from government agencies and NGOs to adopt the system for regional planning and subsidy programs.

CHAPTER SIX

6.0 REFERENCES

- Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.
- IPCC. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report.
- Jones, J. W., et al. (2017). Brief history of agricultural systems modeling. Agricultural Systems, 155, 240–254.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.
- Li, Y., et al. (2019). Data fusion and data integration in agriculture: A review. Computers and Electronics in Agriculture, 160, 142–156.
- Mboya, R. M., et al. (2020). Effects of climate change on crop productivity in Tanzania: Evidence from empirical studies. Tanzania Journal of Agricultural Sciences.
- Ntakolia, C., et al. (2020). Crop yield forecasting using LSTM recurrent neural networks. Neural Computing and Applications, 32, 13025–13036.
- Thornton, P. K., et al. (2011). Climate change and African agriculture: Impacts and adaptation options. CAB Reviews, 6(54), 1–11.

6.1 APPENDIX

- Appendix A: Sample Farmer Survey Questionnaire
- Appendix B: Python Code Snippets for AI Model Training
- Appendix C: Screenshot of Streamlit Application Interface
- Appendix D: System Architecture and Flowchart Diagrams
- Appendix E: Sample Output Prediction Report (PDF Export)