**DEEP LEARNING-BASED YIELD FORECASTING FOR RICE VARIETIES**

### A SOCIALLY RELEVANT MINI PROJECT REPORT

***Submitted by***

**TANUSHRI E [211423104684]**

**THIRUSELVI THIRUNAVUKKARASU [211423104698]**

***in partial fulfillment for the award of the degree of***

**BACHELOR OF ENGINEERING**

**in**

### COMPUTER SCIENCE AND ENGINEERING

****

**PANIMALAR ENGINEERING COLLEGE**

**(An Autonomous Institution, Affiliated to Anna University, Chennai)**

### OCTOBER 2025

**BONAFIDE CERTIFICATE**

Certified that this project report **“DEEP LEARNING-BASED YIELD FORECASTING FOR RICE VARIETIES”** is the bonafide work of **TANUSHRI E (211423104684) , THIRUSELVI THIRUNAVUKKARASU (211423104698)** who

carried out the project work under my supervision.

**Signature of the HOD Signature of the Supervisor Dr L. JABASHEELA, M.E., Ph.D., Dr. SATHIYA V, M.E., Ph.D., PROFESSOR AND HEAD, PROFESSOR,**

**Department of CSE Department of CSE**

**Panimalar Engineering College, Panimalar Engineering College,**

**Chennai – 600 123 Chennai – 600 123**

Submitted for 23CS1512 – Socially Relevant Mini Project Viva-Voce Examination held on...........................

### INTERNAL EXAMINER EXTERNAL EXAMINER

**DECLARATION BY THE STUDENT**

#### We TANUSHRI E (211423104684), THIRUSELVI THIRUNAVUKKARASU

**(211423104698)** hereby declare that this project report titled **“DEEP LEARNING-BASED YIELD FORECASTING FOR RICE VARIETIES”** under

the guidance of **Dr. SATHIYA V , M.E., Ph.D.,** is the original work done by us and we have not plagiarized or submitted to any other degree in any university by us.

**SIGNATURE OF THE STUDENTS**

**TANUSHRI E (211423104684)**

**THIRUSELVI THIRUNAVUKKARASU (211423104698)**

**ACKNOWLEDGEMENT**

We would like to express our deep gratitude to our respected **Secretary and Correspondent Dr. P. CHINNADURAI, M.A., Ph.D.,** for his kind words and enthusiastic motivation, which inspired us a lot in completing this project.

We express our sincere thanks to our **Directors Tmt. C. VIJAYARAJESWARI**, **Dr. C. SAKTHI KUMAR, M.E., Ph.D.,** and **Dr. SARANYASREE SAKTHI**

**KUMAR, B.E., M.B.A., Ph.D.,** for providing us with the necessary facilities to undertake this project.

We also express our gratitude to our **Principal Dr. K. Mani, M.E., Ph.D.,** who facilitated us in completing the project. We sincerely thank the **Head of the Department, Dr. L. JABASHEELA, M.E., Ph.D.,** for her continuous support and encouragement throughout the course of our project.

We would like to express our sincere gratitude to our **Project Coordinator, Dr. SUBEDHA V**, **M.Tech., Ph.D.,** & **Mrs. SHARMILA , M.E., (Ph.D.,)** and our

**Project Guide, Dr. SATHIYA V, M.E., Ph.D.,** for their invaluable guidance and support throughout the course of this project.

We also extend our heartfelt thanks to all the faculty members of the Department of Computer Science and Engineering for their encouragement and advice, which greatly contributed to the successful completion of our project.

**TANUSHRI E (211423104684)**

**THIRUSELVI THIRUNAVUKKARASU (211423104698)**

**ABSTRACT**

In many agricultural regions, farmers face significant yield losses due to unpredictable climate conditions, lack of localized forecasting systems and absence of planning tools tailored to specific rice varieties. These challenges result in poor crop selection, unstable productivity , reduced income, threatening food security and rural livelihoods.

This introduces a deep learning–based rice yield prediction model that integrates NDVI time-series data, daily climate parameters and rice variety characteristics to generate localized & variety-specific forecasts. The model leverages Long Short-Term Memory (LSTM) networks to learn temporal patterns from vegetation and weather trends, offering adaptive and accurate predictions.

NDVI data is sourced from Sentinel-2 satellite imagery via Google Earth Engine, while climate inputs such as rainfall, temperature and humidity are obtained from NASA POWER. Rice variety data is collected from agricultural institutions and farm records to account for genetic and phenotypic differences.

The model focuses on the Thanjavur district of Tamil Nadu, a major rice-producing region with diverse cultivated varieties. Experimental results show that the proposed model achieves promising prediction accuracy, effectively modelling the complex interactions between crop health, climate variability and varietal behaviour.

The LSTM-based pipeline provides actionable insights to farmers, helping them choose optimal varieties and plan for upcoming seasons. It supports climate-resilient agriculture and aligns with UN SDG-2 (Zero Hunger) by promoting sustainable food production and improved livelihood security.

Overall, the proposed approach contributes to data-driven agriculture by offering a transparent, scalable and farmer-friendly decision support tool.

**TABLE OF CONTENTS**

| **CHAPTER NO.** | **TITLE** | **PAGE NO.** |
| --- | --- | --- |
|  | **ABSTRACT** | v |
|  | **LIST OF FIGURES** | vii |
| **1.** | **INTRODUCTION** | 1 |
| 1.1 | Problem Definition | 2 |
| **2.** | **LITERATURE SURVEY** | 3 |
| **3.** | **SYSTEM ANALYSIS** | 6 |
| 3.1 | Existing System | 6 |
| 3.2 | Proposed system | 6 |
| 3.3 | Software Requirement | 9 |
| 3.4 | Hardware Requirement | 9 |
| **4.** | **SYSTEM DESIGN** | 10 |
| 4.1 | System Architecture diagram | 10 |
| 4.2 | Flowchart | 11 |
| 4.3 | Class Diagram | 13 |
| 4.4 | Sequence Diagram | 14 |
| 4.5 | State Diagram | 15 |
| **5.** | **SYSTEM ARCHITECTURE** | 17 |
| 5.1 | Data Collection Module | 17 |
| 5.2 | Input Design | 20 |
| 5.3 | Output Design | 21 |
| **6.** | **SYSTEM IMPLEMENTATION** | 23 |
| 6.1 | Sample Coding | 23 |
| **7.** | **SYSTEM TESTING** | 26 |
| 7.1 | Experimental Setup | 26 |
| 7.2 | Evaluation Metrics | 26 |
| 7.3 | Quantitative Results | 27 |
| 7.4 | Variety Specific Predict | 28 |
| 7.5 | Visualization Of Results | 28 |
| 7.6 | Discussion | 29 |
| **8** | **CONCLUSION & FUTURE WORK** | 30 |
| 8.1 | Conclusion | 30 |
| 8.2 | Future Work | 31 |
| **9** | **APPENDICES** | 32 |
|  | A1 - SDG goals | 32 |
|  | A2 - Sample Screenshots | 33 |
|  | A3 - Paper Publication | 38 |
|  | A4 - Plagiarism report | 43 |
| **10** | **REFERENCES** | 52 |

|  |  |  |
| --- | --- | --- |
| **FIGURE NO.** | **FIGURE DESCRIPTION** | **PAGE NO.** |
| 4.1.1 | System Architecture | 10 |
| 4.2.1 | Proposed System Flowchart | 12 |
| 4.3.1 | Class Structure of the Proposed System | 13 |
| 4.4.1 | System Interaction Sequence Diagram | 14 |
| 4.5.1 | User Interaction And Diagnostic Flow | 15 |
| 5.1.1 | Data Collection Module | 17 |
| 5.1.2 | Google Earth Engine (Thanjavur) | 18 |
| 5.1.3 | NASA POWER Dataset | 18 |
| 5.1.4 | Model Training Using The Datasets | 19 |
| 7.1.1 | Sample Output | 28 |
| A2.1 | Input Location Through Map | 33 |
| A2.2 | Manual Input Location | 34 |
| A2.3 | Input Of Date & Rice Variety | 35 |
| A2.4 | Predicted Yield Shown | 36 |
| A2.5 | Line Plot Of The Inputs Predicting The Output | 37 |

**LIST OF FIGURE**

### INTRODUCTION

Rice is the most important staple food globally and a key source of food security and rural income in India. Thanjavur district in Tamil Nadu, known as the “Rice Bowl of Tamil Nadu”, is a major producer but faces severe yield fluctuations due to climate instability, irregular rainfall and lack of localized forecasting systems. Events like delayed monsoons, droughts and rising temperatures increase uncertainty, causing financial risk to farmers. Nearly 40% of Indian farmers face income instability due to unpredictable yields. Traditional field-based yield estimation is slow and offers no timely guidance.

Earlier forecasting studies using regression and machine learning models such as Random Forest, SVM and LASSO , capture some yield–climate relations but rely on seasonal averages and ignore daily changes and crop growth stages. They also treat rice as a single crop, overlooking varietal differences , reducing practical usefulness.

With advances in remote sensing, NDVI from satellites is widely used to monitor crop vigour and biomass. Climate dataset such as NASA POWER provides daily weather inputs, yet very few studies combine NDVI and climate data for localized, variety- specific yield prediction.

Deep learning models like LSTM can learn temporal patterns from time-series NDVI and climate data but have rarely been applied for rice yield forecasting with variety information.

A deep learning-based approach is proposed for rice yield forecasting in Thanjavur using NDVI from Sentinel-2, daily climate data from NASA POWER and rice variety as an input. The model generates localized and variety-specific predictions to support better crop planning. It serves as a decision-support tool for farmers and policymakers which aligns with UN SDG-2 by enabling climate-resilient agriculture, improving food security and stabilizing farmer incomes.

### PROBLEM DEFINITION

Rice is one of the key staple crops of India and a vital component of food and agricultural economy. Estimation of rice yield accurately is crucial for farming community, policymaking and agribusiness while making crop treatment and marketing plans. Rice yield is a function of several variables like climatic conditions (temperature, precipitation, humidity), fertility of the soil, irrigational practices used and variety of rice cultivated.

Traditional techniques involving manual observation and past records are tedious, time-consuming and error-prone. Although statistical techniques and machine learning algorithms are utilized, they cannot represent intricate temporal variations and non- linear interdependencies among variables. Most research works are confined to single variety or pure line or staple grains like wheat and thus are restricted to rice.

A reliable data-intensive approach combining satellite-based vegetation indices (NDVI)-derived variables and climatic variables becomes a necessity. Deep learning algorithms, specifically LSTM networks, are best suited to represent sequential variations. With this approach, real time and accurate estimation of rice yields of multi- varietal nature becomes possible and enables anticipatory decision-making in agriculture.

### LITERATURE SURVEY

1. Traditional Approaches to Yield Forecasting

Crop yield prediction has long been a subject of agricultural research. Early studies focused primarily on statistical regression models, which attempted to establish linear relationships between yield and environmental factors such as rainfall, temperature, and soil quality [4][5]. While these approaches provided baseline insights, their accuracy was limited because crop yield is influenced by highly nonlinear and dynamic interactions among multiple variables [4]. As a result, regression-based methods often failed under conditions of climate variability or when applied across diverse crop varieties [5].

1. Machine Learning-Based Methods

The rise of machine learning introduced algorithms capable of capturing more complex patterns between environmental data and crop yield. Techniques such as Random Forest, Support Vector Machines and LASSO regression have been widely applied in yield estimation [4][11]. For instance, studies on wheat and maize demonstrated that RF and SVM could outperform regression models when provided with high-resolution weather and soil data. Similarly, hybrid approaches combining multiple ML techniques were able to improve yield prediction accuracy by reducing overfitting and incorporating variable selection [11][13].

However, most ML-based models exhibit two key limitations:

* + First, they often depend on seasonal or annual averages of climate data, ignoring intra-seasonal variations that play a critical role in crop growth [4][5].
  + Second, these models treat crops as homogeneous units and do not account for variety-level differences, which can significantly alter yield outcomes under varying environmental conditions [13][11].

1. Remote Sensing and NDVI Applications

Parallel to advances in ML, the field of remote sensing has transformed agricultural monitoring. Among various vegetation indices, the Normalized Difference Vegetation Index has gained prominence as a proxy for vegetation vigor, biomass and photosynthetic activity [8][9]. NDVI derived from satellite imagery, such as Sentinel- 2, has been employed to monitor crop health and forecast yields for large-scale crops like wheat, maize and soybeans [8][12].

Several studies have demonstrated the value of NDVI in yield forecasting. For example, NDVI time-series have been correlated with crop biomass at different growth stages to estimate yield with significant accuracy [9][10]. In India, NDVI has been applied in rice monitoring for area mapping and growth assessment. Yet, most studies rely on seasonal NDVI averages rather than leveraging the full temporal dynamics of NDVI, limiting the ability to capture short-term stress or recovery patterns [10][12][15].

1. Climate Data Integration in Crop Forecasting

Apart from remote sensing, climate datasets such as temperature, rainfall and relative humidity play an essential role in yield prediction [1][2]. Platforms like NASA POWER and ERA5 reanalysis data have made high-resolution daily climate data accessible for agricultural modeling [1][7][14]. Research has shown that combining climate data with vegetation indices improves yield estimation compared to using either source alone [1][14]. For instance, maize yield prediction in Sub-Saharan Africa achieved significant accuracy improvements when rainfall and NDVI were jointly modelled [7]. Despite these advances, many existing works have overlooked fine- grained, daily climate data, instead aggregating them into seasonal summaries. Moreover, limited attention has been given to the interaction between crop varieties and local climatic variability, which is crucial for developing actionable forecasts for farmers [1][7].

1. Deep Learning Approaches in Agriculture

The emergence of deep learning has opened new opportunities for agricultural forecasting. Unlike traditional ML methods, deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can automatically learn hierarchical features from raw data.

In particular, Long Short-Term Memory (LSTM) networks have shown strong potential for handling time-series data, making them suitable for modeling sequential climate and NDVI trends [1][2][3]. Recent studies applying LSTMs to crop yield prediction have reported encouraging results. For example, wheat yield forecasting using NDVI time-series achieved higher accuracy compared to classical ML models [7][14][15].

Similarly, deep learning has been used in soybean and maize yield estimation at regional scales, leveraging both NDVI and weather data [3][14]. However, these works generally focus on major cereals like wheat and maize with limited applications to rice, particularly in the Indian context [7].

1. Research Gaps

Based on the literature, several gaps are evident:

1. Limited variety-specific models: Most studies treat rice as a homogeneous crop, overlooking varietal diversity that influences resilience and productivity [3][7][11].
2. Underutilization of temporal data: Many works use aggregated NDVI and climate values, ignoring short-term growth dynamics [1][10][12][15].
3. Lack of explainability: Few models provide interpretable outputs that farmers or policymakers can trust [6][13][14].
4. Regional focus: Very few studies target rice yield prediction in Tamil Nadu, despite its importance as a rice-producing region [7].

### SYSTEM ANALYSIS

### EXISTING SYSTEM

Conventional yield estimation of rice through manual field surveying, reports from farmers and past crop records are tedious, expensive and unreliable because of human error and limited coverage. Statistical techniques such as linear regression and ARIMA were applied with a view of forecasting crop yield from past climate and soil records, although the former is restricted from dealing with non-linear associations and long- term dependencies.

Machine learning techniques involving Support Vector Machines (SVM), Random Forests (RF) and Gradient Boosting are also applied with improved accuracy from the utilization of environmental and NDVI data. Most of these remain with single-variety crops or are restricted to wheat and perform poorly with temporal growth patterns of varying rice varieties. Lack of capturing sequential dependencies undermines their real-time yield forecasting prospective.

These are the constraints that necessitate sophisticated deep learning methods such as LSTM networks that can process sequential climate and vegetation data and lead to improved accuracy of yield forecasting of multi-variety rice-growing areas such as Thanjavur, Tamil Nadu.

### PROPOSED SYSTEM

The ultimate target is to establish a system of rice yield forecasting using deep learning as a substitute of conventional and machine learning techniques. While typical models rely upon season-long means or assume rice as a generic type of crop, the system takes advantage of time-series satellite images, day-by-day location weather and variety traits and yields variety- and location-dependent predictions of yield.

1. Input Parameters

This solution leverages the synergy of three critical data domains:

* 1. Vegetation health (NDVI): Vegetation health defines growth stage of the crop and build-up of biomass with time.
  2. Climatic variables**:** Store short-term fluctuations of rainfall, temperature and humidity.
  3. Variety information of rice: Carries genetic and phenotypic variation of varieties and offers more possible and practical yield estimation.

Along with such inputs through a Long Short-Term Memory (LSTM) model, the system could learn patterns of sequence from weather and NDVI data and produce variety-specific predictions. For Tamil Nadu, this is relevant as it has a vast number of rice varieties with contrasted growth behaviors and variety-specific yield potentials that are a function of climatic and soil conditions.

1. System Workflow

The desired workflow takes the following next steps:

* 1. Data Collection
     + Google Earth Engine Sentinal-2 time-series of NDVI.
     + Daily weather variables (rainfall, temp., humidity) from NASA POWER.
     + Variety-level yield statistics of agriculture.
  2. Data Preprocessing
     + Filtering and selecting valid NDVI frames:

Only NDVI data from the rice-growing period in Thanjavur was selected and images with missing/invalid pixel values were removed.

* + - Aligning climate data with NDVI timestamps:

Climate data (temperature, rainfall, humidity) was matched to the same dates as NDVI extraction to create a consistent time-series input.

* + - Feature scaling and category encoding

Climate and NDVI features were scaled to a uniform numeric range and rice varieties were encoded as numerical input for the deep learning model.

* 1. Deep Learning Model
     + An LSTM-based model is used to learn time-series patterns from NDVI and climate data.
     + Rice variety input helps the model differentiate yield behavior across varieties.
     + The final yield prediction is generated and displayed to the user
  2. Output & Interpretation
     + The system provides a predicted rice yield (in tons per hectare) based on the entered location, season and rice variety.
     + The output also includes graphical results such as yield comparison and NDVI trends to help understand crop growth.
     + Farmers can use this information to choose the right variety and plan cultivation for better productivity.

1. Comparison with Previous Algorithms
   1. Variety-specific predictions: In contrast with hypothesized models that make generic assumptions across varieties of rice, the approach delivers variety-specific predictions offering actionable knowledge to farmers.
   2. Temporal learning: Using weekly modeling of climatic patterns and NDVI, LSTM learns intra-seasonal behavior missed with traditional methods.
   3. Localized attention: It has been developed for the Thanjavur district and can be utilized with variations with other regions and crops.

### SOFTWARE REQUIREMENT

Programming Language : Python

Deep Learning Library : TensorFlow / Keras

Satellite Data Access : Google Earth Engine (Sentinel-2)

Climate Data Source : NASA POWER

Data Format : CSV, JSON, Excel

### HARDWARE REQUIREMENT

Processor : Intel i5 or higher

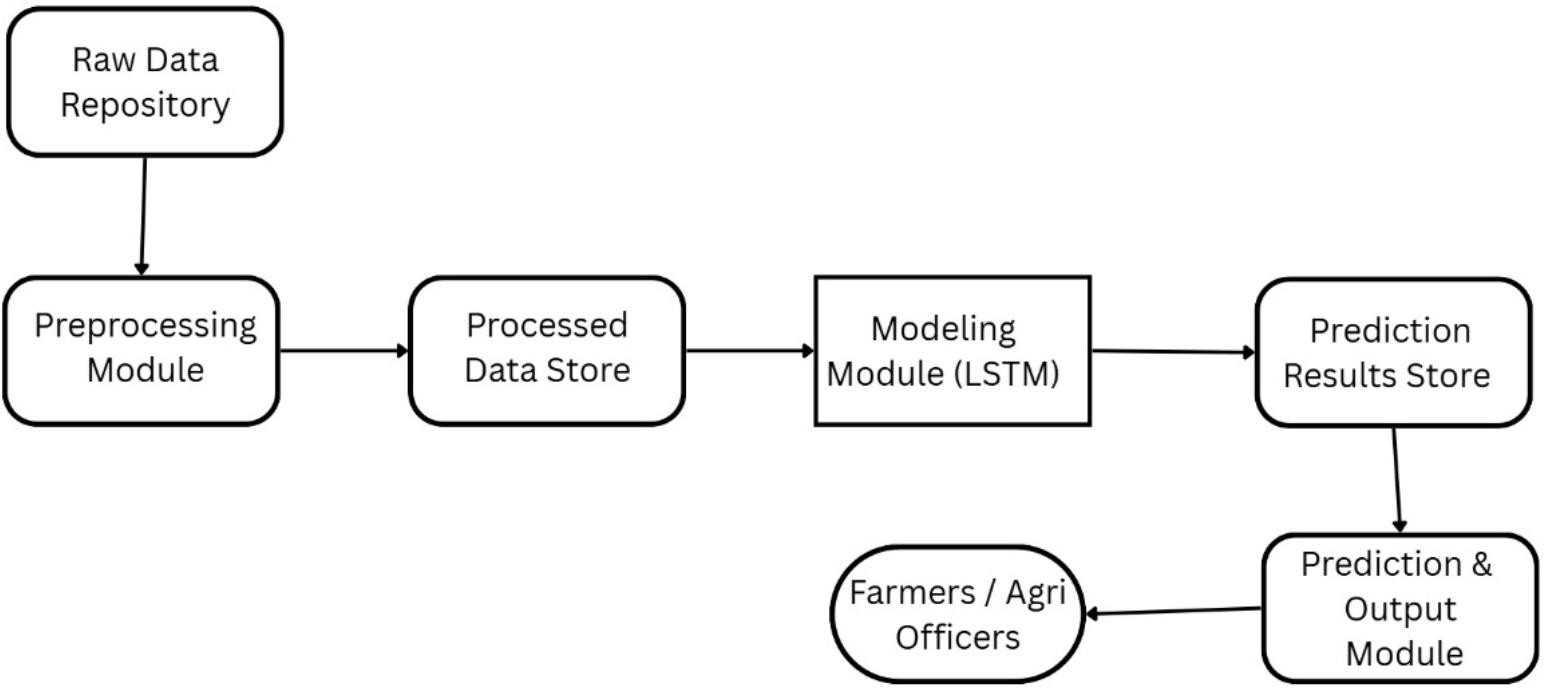
RAM : 8 GB or more

Storage : Minimum 10 GB

GPU : Optional (e.g., Google Colab GPU)

### SYSTEM DESIGN

* + - * 1. **SYSTEM ARCHITECTURE**

****

**Fig: 4.1.1 System Architecture**

#### Overview of the System

Modular pipeline for rice yield prediction. Integrates multi-source data for accurate, variety-specific forecasts. Four main functional modules supported by data repositories and user interfaces.

1. Data Collection Module

Gathers heterogeneous datasets from multiple sources: NDVI values from Sentinel-2 imagery via Google Earth Engine, Climate parameters (rainfall, temperature, humidity) from NASA POWER , historical rice variety and yield data from agricultural databases.

1. Preprocessing Module

Ensures data consistency and quality (cleans missing or noisy data). Aligns all time- series data into a weekly format to match crop growth cycles. Structured data stored in Processed Data Store for model training.

1. Modelling Module

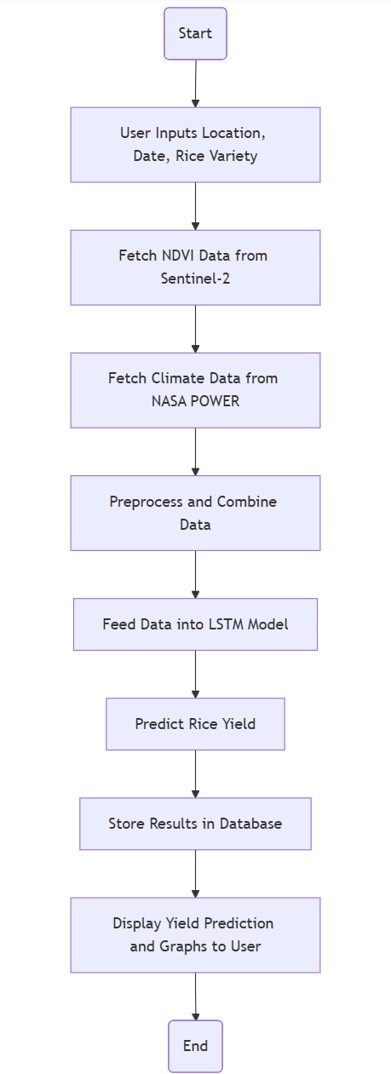
This module uses an LSTM deep learning model to learn patterns from weekly NDVI and climate data along with rice variety information. It captures how crop health and weather changes over time affect yield. Once processed, the model generates a single yield value for the selected rice variety and location.

1. Prediction and Output Module

This module converts the model’s prediction into useful results such as yield in tons per hectare and simple graphical results for better understanding. The output helps farmers choose suitable rice varieties and plan cultivation effectively under current climate conditions.

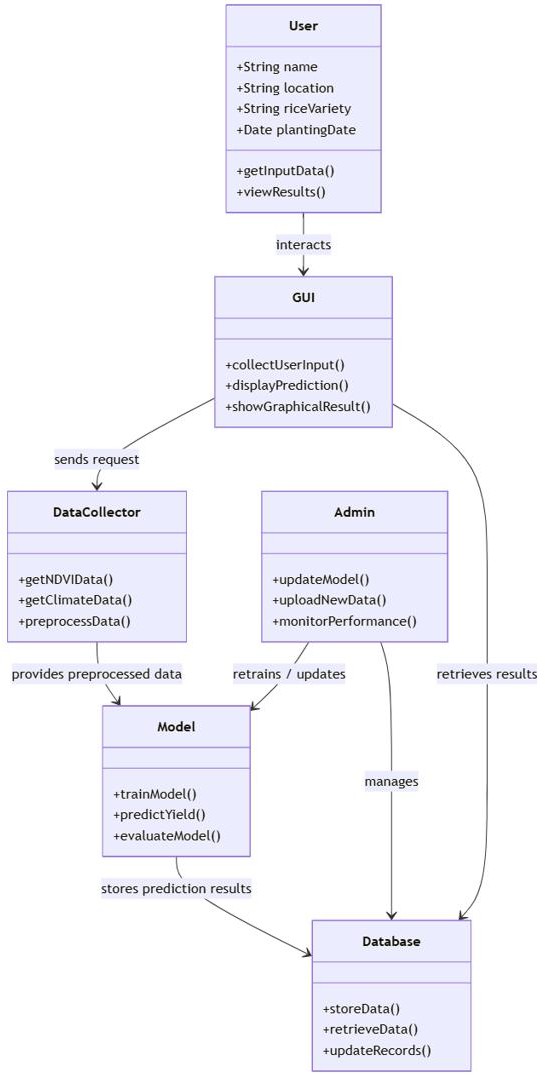
### FLOWCHART

The flowchart depicts the overall workflow of the rice yield prediction system. The process begins with the user entering essential details such as location, date and rice variety. The system then collects NDVI and climate data, which are preprocessed and fed into the LSTM deep learning model. After processing, the model generates a yield prediction, which is stored in the database and displayed to the user in graphical form. This logical flow ensures efficient data handling, accurate prediction and an interactive user experience.



**Fig 4.2.1: Proposed System Flowchart**

### CLASS DIAGRAM

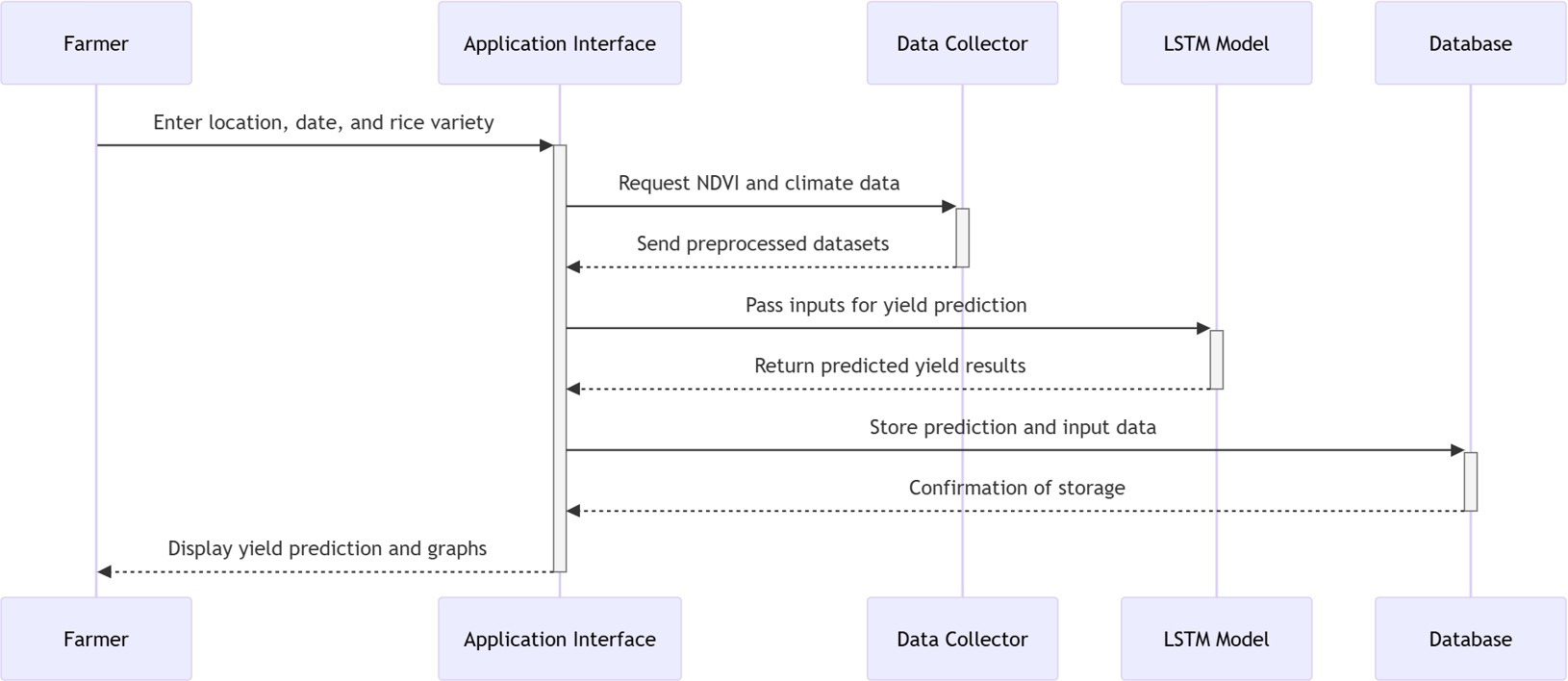
****

#### Fig 4.3.1 Class Structure of the Proposed System

The UML class diagram represents the overall architecture of the rice yield prediction system and its interaction between major components. The User provides inputs such as location, rice variety and date through the GUI, which collects and displays results.

The Data Collector retrieves and preprocesses NDVI data from Sentinel-2 and climate data from NASA POWER, preparing them for analysis. The Model class, based on LSTM deep learning, performs training and yield prediction. The Admin manages model updates, dataset uploads and system performance. Together, these components create a modular, scalable and user-friendly application for accurate rice yield forecasting.

### SEQUENCE DIAGRAM

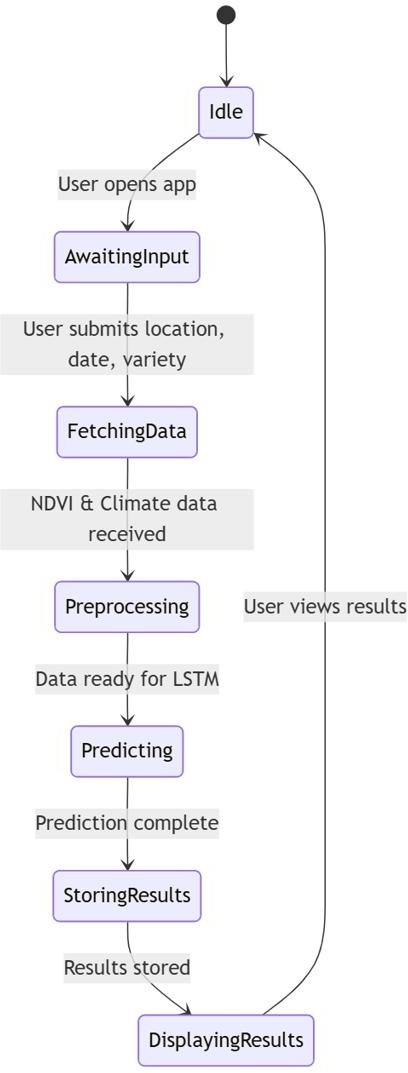
****

#### Fig 4.4.1: System Interaction Sequence Diagram

The sequence diagram illustrates the step-by-step interaction between the key components of the rice yield prediction system. The farmer (user) initiates the process by providing input details such as location, date and rice variety through the GUI. The GUI then requests the required NDVI and climate data from the DataCollector, which retrieves and preprocesses them for analysis. These inputs are passed to the LSTM- based Model, which processes the data and generates the predicted yield. The results

are stored in the Database for future reference and the final prediction along with graphical insights is displayed back to the user. This sequential flow ensures smooth communication between modules, enabling accurate and timely rice yield forecasting.

### STATE DIAGRAM

****

#### Fig 4.5.1: User Interaction And Diagnostic Flow

The state diagram represents the different operational states of the Rice Yield Prediction application. It begins in the Idle state, where the application is waiting to be used. When a user opens the app, it transitions to AwaitingInput, prompting the user to enter details such as location, planting date and rice variety. Once the user submits this information, the application moves to FetchingData, retrieving NDVI satellite data and climate information.

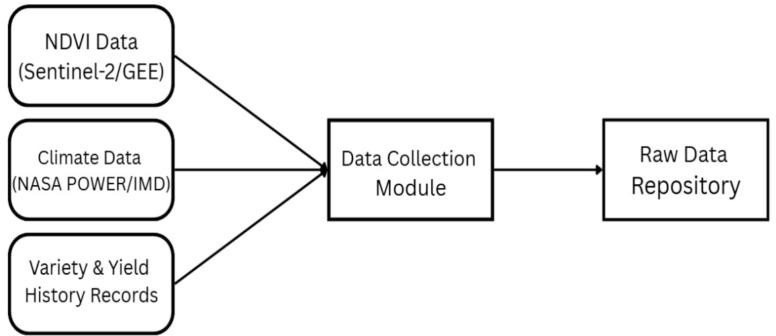
The retrieved data is then processed in the Preprocessing state, where it is cleaned and combined for analysis. After preprocessing, the system enters the Predicting state, where the LSTM model generates rice yield predictions. Once the prediction is complete, the results are stored in the StoringResults state, followed by DisplayingResults, where the application presents the yield forecast and graphs to the user.

Finally, the system returns to the Idle state, ready for the next input. This flow ensures a structured, sequential operation from user input to result display, capturing all key stages of the application’s workflow.

### SYSTEM ARCHITECTURE

The system architecture for the rice yield forecasting model is designed as a modular pipeline that integrates satellite imagery, climate data and rice variety information to generate accurate, localized predictions. Built using deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, the system captures temporal dependencies and varietal differences to support data-driven agricultural decision-making. The architecture comprises several interconnected modules, each responsible for a specific function in the data processing and prediction workflow.

### DATA COLLECTION MODULE

****

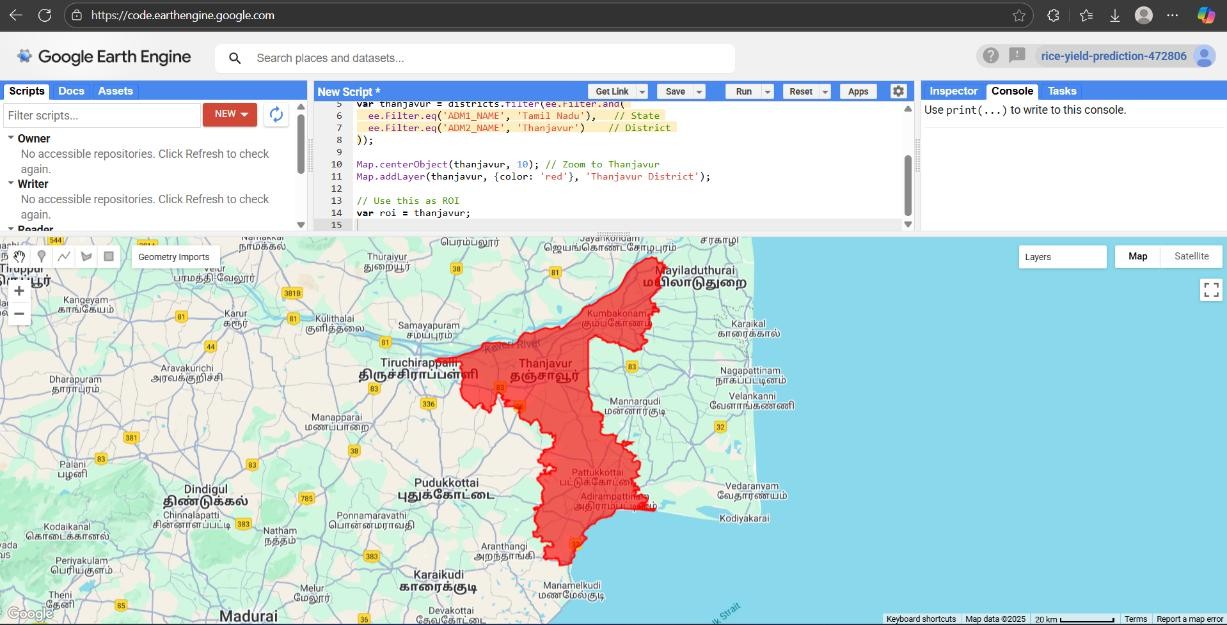
#### Fig: 5.1.1 Data Collection Module

**Purpose of the Module**

Acts as the first step in the rice yield prediction system. Responsible for gathering all relevant input data from diverse sources. Ensures that the system has a comprehensive dataset for accurate modelling.

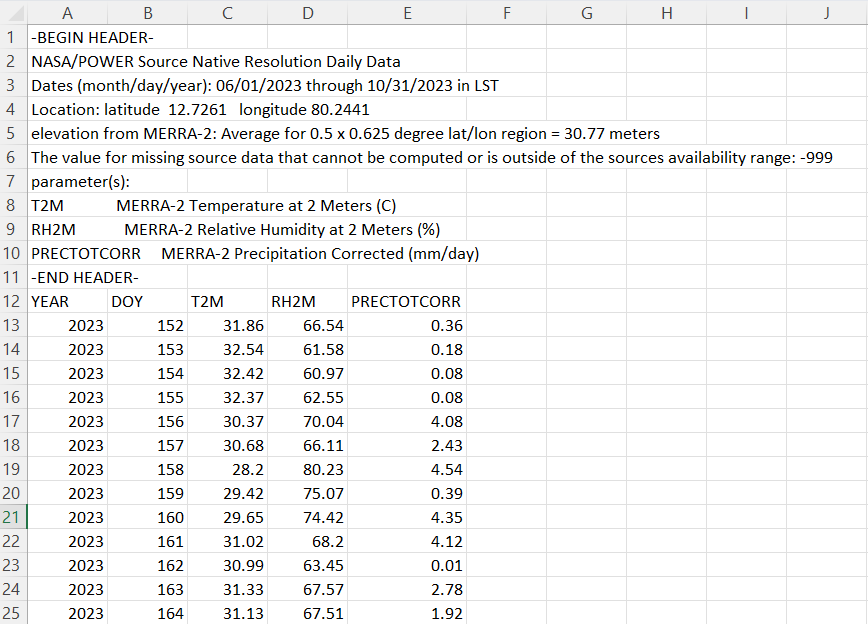
#### Input Data Sources

NDVI Data (Sentinel-2 / Google Earth Engine): Captures vegetation health and growth patterns over time, Helps monitor crop vigour and detect anomalies in the rice fields.



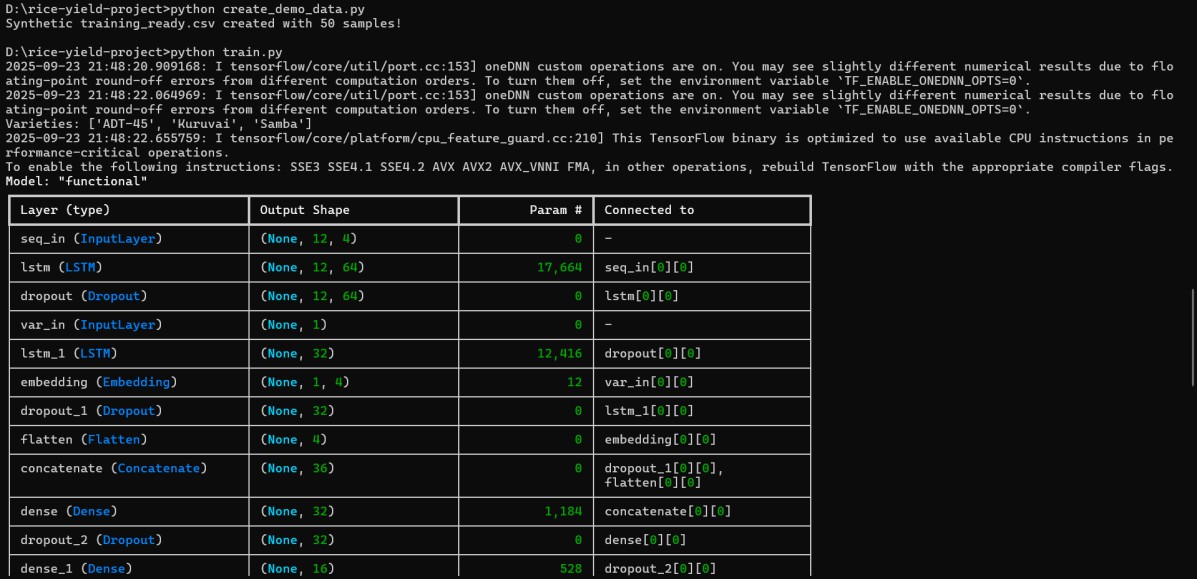
#### Fig 5.1.2 Google Earth Engine ( Thanjavur )

Climate Data (NASA POWER / IMD**)**: Includes rainfall, temperature and humidity records, Critical for understanding environmental impacts on crop growth.



#### Fig 5.1.3 NASA POWER Dataset

Variety & Yield History Records**:** Contains historical rice variety details and their yield performance, Enables the system to consider varietal characteristics in forecasting.



#### Fig 5.1.4 Model Training Using The Datasets Data Collection Process

Inputs from all three sources are funnelled into the Data Collection Module. The module performs initial data aggregation, ensuring that different types of information (satellite, climate, historical records) are available for preprocessing. Serves as a centralized point for raw data management before further processing.

#### Raw Data Repository

All collected data is stored in the Raw Data Repository. Provides secure and organized storage for subsequent processing. Ensures easy retrieval for cleaning, normalization, and feature extraction in the next module.

#### Significance

Integrating remote sensing, climate and historical data improves the depth and reliability of predictions. Lays a strong foundation for preprocessing and modelling

modules**.** Supports scalability, as more data sources can be added without disrupting the pipeline.

#### Dataset

Once the raw data is collected, it is organized into a structured Dataset that aligns satellite, climate and varietal data across time and location. This dataset forms the backbone of the modeling process, ensuring that each input feature is synchronized and relevant to the prediction task. The dataset includes weekly NDVI values, climate averages and encoded variety traits, enabling the model to learn from both temporal and categorical data. Special care is taken to maintain data integrity and completeness, as missing or inconsistent entries can adversely affect model performance.

### INPUT DESIGN

Input design is a critical phase in system development that focuses on how data is captured, validated and fed into the system for processing. In the context of the rice yield forecasting model, input design ensures that all relevant data—ranging from satellite imagery to climate parameters and rice variety information—is collected in a structured and user-friendly manner.

#### User Inputs

The system accepts several key inputs from the user through a graphical user interface (GUI):

* Location Coordinates: Latitude and longitude of the rice field, used to fetch NDVI and climate data.
* Rice Variety: Selected from a predefined list of varieties cultivated in the region.
* Season Start and End Dates: Specifies the time frame for which yield prediction is required.

These inputs are entered manually by the user and are validated for correctness before processing. For example, the system checks if the coordinates fall within the Thanjavur district and whether the selected rice variety is supported by the model.

#### System Inputs

In addition to user inputs, the system automatically retrieves external data:

* NDVI Time-Series Data: Extracted from Sentinel-2 satellite imagery via Google Earth Engine.
* Climate Data: Daily records of rainfall, temperature and humidity from NASA POWER.
* Historical Yield Records: Used during training to correlate input features with actual yield outcomes.

All inputs are preprocessed to ensure consistency. NDVI and climate data are aggregated into weekly intervals and rice variety information is encoded into numerical vectors. This structured input design enables the LSTM model to learn temporal and categorical patterns effectively.

### OUTPUT DESIGN

Output design focuses on how the results generated by the system are presented to the user. The goal is to ensure clarity, usability and actionable insights that support agricultural decision-making.

#### Predicted Yield

The primary output of the system is the predicted rice yield**,** expressed in tons per hectare. This value is generated by the trained LSTM model based on the input location, rice variety and seasonal data. The prediction is displayed prominently on the

user interface, along with contextual information such as the selected variety and season.

#### Visual Insights

To enhance interpretability, the system provides graphical outputs:

* NDVI and Climate Trends: Line charts showing vegetation health and climate variations over the season.
* Feature Importance: Visual explanations of which features—such as rainfall, NDVI at specific weeks, or variety traits—contributed most to the prediction.
* Comparison Charts: Predicted vs. actual yield comparisons (if historical data is available), helping users assess model accuracy.

These visuals are generated using Python libraries such as Matplotlib are embedded directly into the interface for easy access.

#### Recommendations

Based on the prediction, the system may also provide recommendations:

* Best-performing rice varieties for the selected location and season.
* Alerts for suboptimal conditions (e.g., low NDVI or insufficient rainfall).
* Suggestions for crop planning and resource allocation.

#### Output Format

All outputs are presented in a clean, responsive interface built using Streamlit. Users can download prediction reports in CSV or PDF format for record-keeping or sharing with agricultural officers. The system also supports integration with mobile platforms for field-level accessibility.

### SYSTEM IMPLEMENTATION SAMPLE CODINGS

# app.py — Streamlit-based Rice Yield Predictor

import streamlit as st import pandas as pd import numpy as np import joblib

import tensorflow as tf

from sklearn.preprocessing import StandardScaler from datetime import datetime

# Load model and preprocessing artifacts

model =

tf.keras.models.load\_model("models/lstm\_rice\_yield\_model. h5", compile=False)

scaler = joblib.load("models/scaler.pkl")

labelenc = joblib.load("models/labelencoder.pkl")

# Define constants TIME\_STEPS = 12

N\_FEATURES = 4

rice\_varieties = ["ADT-36", "ADT-37", "Mapillai Samba", "Karuppu Kavuni", "White Ponni", "BPT-5204"]

# Streamlit UI

st.title("Rice Yield Predictor")

location = st.text\_input("Enter location (e.g., Thanjavur)") start\_date = st.date\_input("Start Date")

end\_date = st.date\_input("End Date")

variety = st.selectbox("Select Rice Variety", rice\_varieties)

# Dummy NDVI and climate data for demonstration ndvi\_seq = np.linspace(0.3, 0.7, TIME\_STEPS) temp\_seq = np.random.normal(28, 2, TIME\_STEPS) rain\_seq = np.random.normal(10, 3, TIME\_STEPS) humidity\_seq = np.random.normal(75, 5, TIME\_STEPS)

# Combine features into sequence

seq = np.zeros((TIME\_STEPS, N\_FEATURES)) seq[:, 0] = ndvi\_seq

seq[:, 1] = temp\_seq seq[:, 2] = rain\_seq seq[:, 3] = humidity\_seq

# Scale features

seq\_scaled = scaler.transform(seq.reshape(- 1,N\_FEATURES)).reshape(TIME\_STEPS, N\_FEATURES)

# Encode rice variety

var\_id = int(np.where(labelenc.classes\_ == variety)[0][0]) X\_seq = np.expand\_dims(seq\_scaled, axis=0)

X\_var = np.array([var\_id])

# Predict yield

predicted\_yield = model.predict({'seq\_in': X\_seq, 'var\_in': X\_var}).ravel()[0]

st.metric("Predicted Rice Yield (tons/hectare)", f"{predicted\_yield:.2f}")

### SYSTEM TESTING

System testing is a critical phase in the software development life cycle that focuses on assessing the overall quality, functionality and performance of a software system. It is a comprehensive and systematic process that aims to identify defects, ensure that the system meets specified requirements and verify its readiness for deployment. System testing plays a crucial role in delivering reliable, robust and high quality software solutions.

### EXPERIMENTAL SETUP

The model was trained with synthetic and partly available real datasets generated from Sentinel-2 NDVI time-series, NASA POWER daily climate variables and past yield records of the Thanjavur district. The dataset was divided into 70% training, 15% validation and 15% testing. TensorFlow 2.x and a workstation with a GPU were used to conduct the training of the model. Optimization of LSTM network used the Adam optimizer with a learning rate of 0.001 and trained the network for 80 iterations with early stopping to avoid overfitting.

### EVALUATION METRICS

We used three typical regression measures to evaluate the proposed system performance.

1. Mean Absolute Error (MAE): It determines the average discrepancy of predicted and actual yield.

MAE = 𝟏 ∑𝒏 |𝒚 − 𝒚̂ |

𝒏 𝒊=𝟏 𝒊 𝒊

1. Root Mean Square Error (RMSE): Penalizes large error more strongly, providing insight into robust prediction.

RSME = √𝟏 ∑𝑵 (𝒚 − 𝒚̂ )𝟐

𝑵 𝒊=𝟏 𝒊 𝒊

1. Coefficient of Determination (R²): Provides the proportion of the variance of yield accounted for by the model.

∑𝑵 (𝒚𝒊 − 𝒚̂𝒊)𝟐

𝑹𝟐 = 𝟏 − 𝒊=𝟏

𝑵

∑

𝒊=𝟏

(𝒚𝒊 − 𝒚̅)𝟐

### QUANTITATIVE RESULTS

Test results of the model with the test data are tabulated in Table I.

#### Table I. Performance of Proposed LSTM Model

|  |  |
| --- | --- |
| **Metric** | **Value** |
| MAE | 0.25 tons/ha |
| RMSE | 0.32 tons/ha |
| R² | 0.47 |

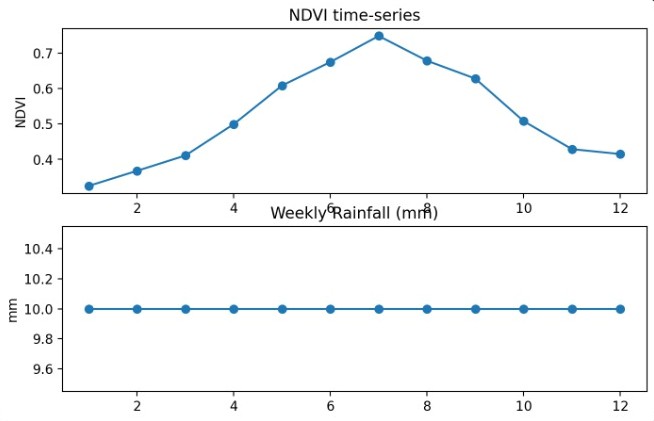
They suggest that the model may reproduce the connection of NDVI with climate and rice yield with fair accuracy. Although the explanatory power of R² suggests a moderate level of explanation, further refinement may also be achieved through the addition of past data and an extension of training samples.

### VARIETY-SPECIFIC PREDICT

One of the major contributions of this research is variety-specific prediction. For instance, at Thanjavur, three varieties namely ADT 37, CR 1009 and BPT 5204 were tested. ADT 37 has higher yield stability under climatic variability of rainfall and BPT 5204 has higher temperature sensitivity. It indicates the significance of variety as an explicit input of the prediction process.

### 7.5. VISUALIZATION OF RESULTS

Predicted rice yield is presented in tons per hectare in a clear and interpretable format. NDVI growth trends and climate variations are shown through simple line charts for better understanding. Yield comparison visuals help identify the most suitable rice variety for the selected location and season. Color-based indicators highlight whether the predicted yield is low, moderate, or high. These visual insights support better crop planning and decision-making for farmers and agricultural officers.



**Fig 7.5.1 Sample Output**

### 7.6 DISCUSSION

Experimental results show that with NDVI time-series, climatic variability and rice variety embedding, locally and realistically better forecasts are possible as compared with usual ML-based forecasting.

Even with the small dataset used during this mini-project, the developed framework can also be scalable. System accuracy can also be enhanced with the addition of much larger datasets, like soil and agronomic practices.

In addition, the explainability module ensures that stakeholders understand why a prediction has been made and increases AI-model uptake for agriculture. Importantly, the system contributes towards SDG 2: Zero Hunger since it contributes both directly and towards sustainable cropping calendars and food security.

**8. CONCLUSION & FUTURE WORK**

* 1. **CONCLUSION**

This study proposed a deep learning-assisted rice yield forecasting system that integrated NDVI time-series, weather variables and rice variety information for making variety-specific and localized forecasts of rice yield at the Thanjavur district level in Tamil Nadu.

Using LSTM networks efficiently extracted temporal patterns of weather and crop health and variety embeddings facilitated variety-distinctive forecasting. Experimental findings validated that the model simulated well and NDVI at the reproductive phase and rainfall heterogeneity were vital yield determinants.

Integration of interpretability as analysis assisted in establishing confidence and interpretability and thus the system was appropriate for real-world agriculture decision-making. In contrast with conventional models, the approach developed herein incorporates temporal dynamics and varietal diversity and therefore reveals actionable information both to farmers and policymaking.

Future prospects are the integration of the framework with soil and pest information and with management and the use of the developed approach as a mobile app for real-time use. Overall, the system enables sustainable agriculture planning of crops and food security and contributes to UN SDG 2 – Zero Hunger.

* 1. **FUTURE WORK**

Real-Time Data Integration: We will integrate near-real-time data streams from local weather APIs and high-frequency satellite imagery (beyond Sentinel-2) to enable in-season adjustments and dynamic forecast recalibration.

Expanded Crop Support: The LSTM-based framework will be generalised using a transfer learning approach to support forecasting for other major staple crops, such as pulses and millets, expanding the system's agricultural impact.

User Interface Development: A Mobile/Web Application will be developed to deliver predictions and a variety of recommendations in a simple, interpretable and localized format to farmers and agricultural officers. This interface will include multilingual support and GIS mapping.

Advanced Modelling with Attention: We will enhance the core LSTM model by incorporating Attention Mechanisms. This will improve prediction accuracy by enabling the model to focus on the most critical time-series data points (e.g., rainfall during the flowering stage).

Localized Actionable Recommendations: The system will move beyond simple yield predictions to offer highly specific, plot-level advice, including recommendations for optimised irrigation schedules

**9. APPENDICES**

# A1. SDG GOALS

**A1. SDG GOALS**

The rice yield prediction system aligns with the following SDG Goals

SDG 1: No Poverty

Target 1.2: Reduce poverty by increasing income and productivity among vulnerable groups.

→ By improving yield prediction accuracy, the project supports farmers in minimising losses and stabilising income, directly contributing to poverty reduction.

SDG 2: Zero Hunger

Target 2.4: Ensure sustainable food production systems and resilient agricultural practices.

→The project helps achieve this by providing AI-based yield forecasts that enable farmers to plan efficiently, adapt to weather changes and improve food security.

SDG 9: Industry, Innovation and Infrastructure

Target 9.5: Enhance scientific research and upgrade technologies for sustainable development.

→ The project introduces deep learning and remote sensing tools into agriculture, fostering innovation and promoting data-driven farming infrastructure.

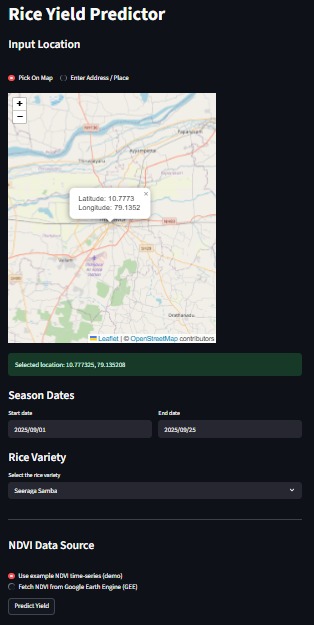
SDG 13: Climate Action

Target 13.1: Strengthen resilience and adaptive capacity to climate-related hazards.

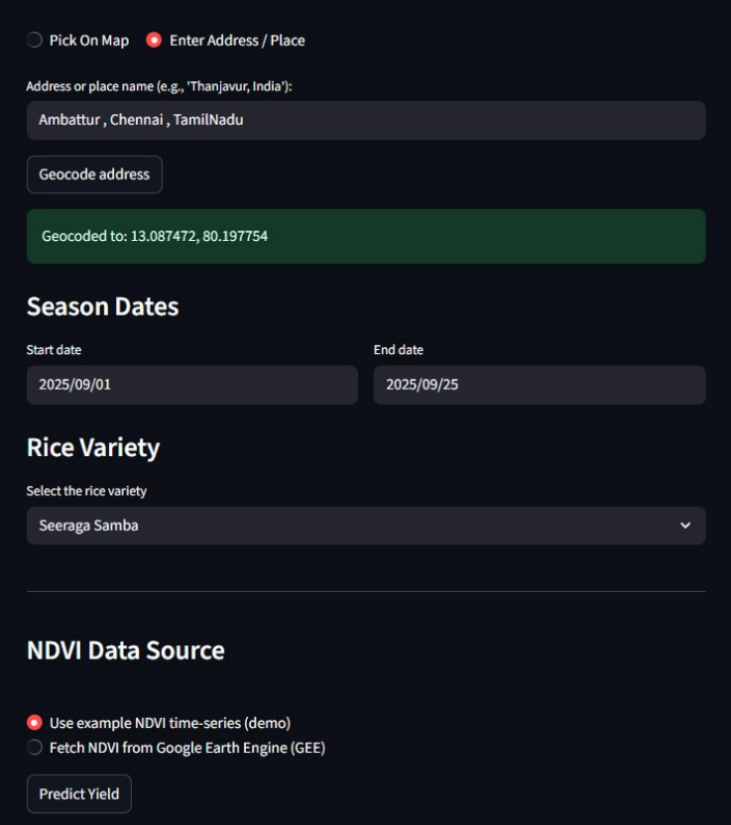
→ By integrating climate data for predictive modelling, the system equips farmers with insights to make climate-resilient decisions, reducing risks from unpredictable weather.

# A2. SCREENSHOT

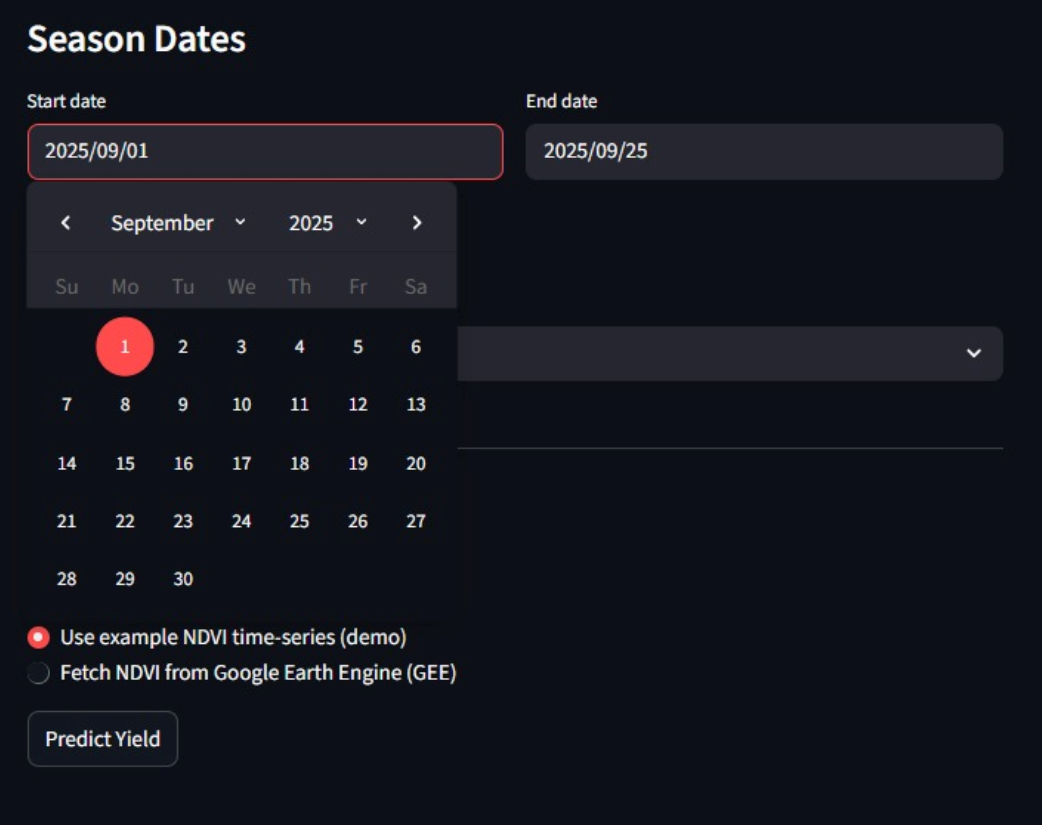
**A2. SCREENSHOTS**

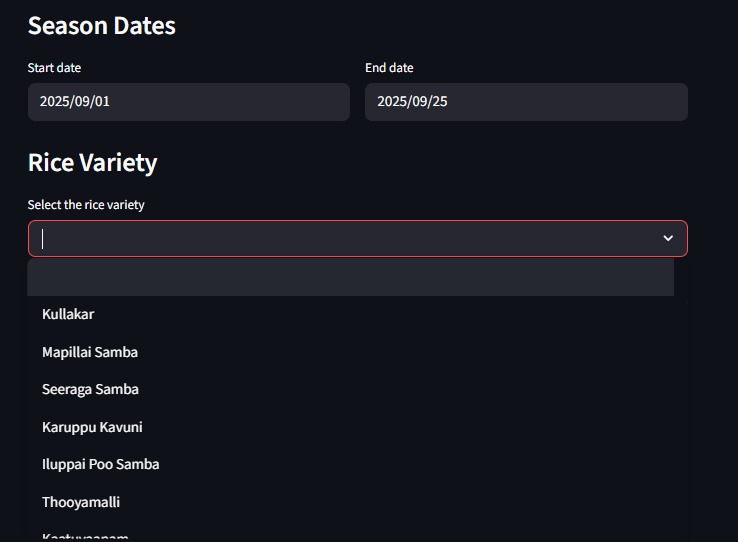
****

**Fig A2.1 Input Location Through Map**

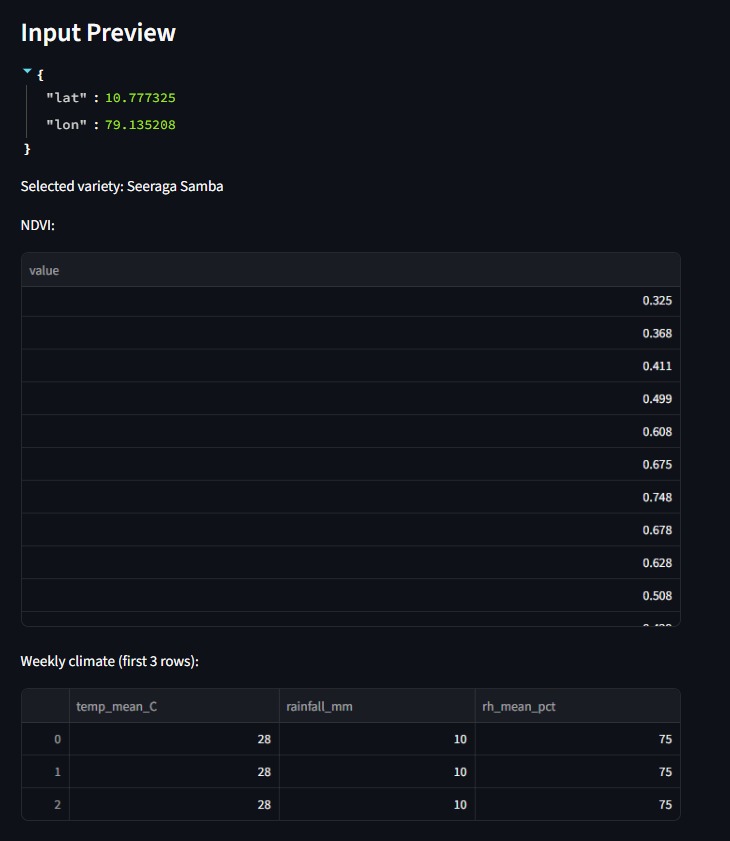
****

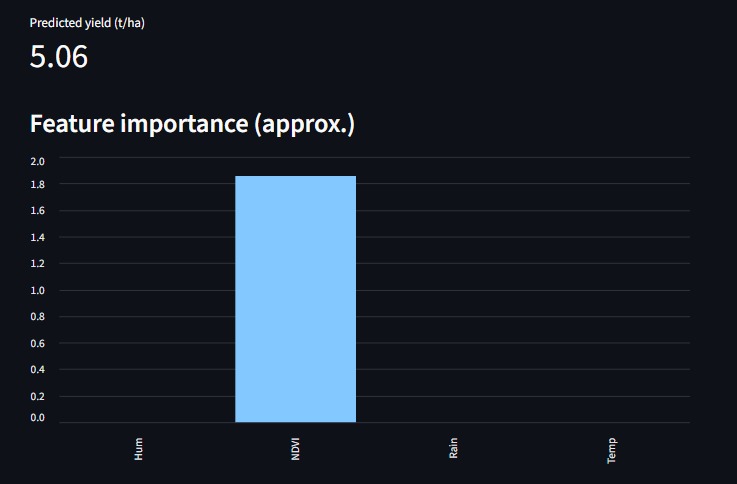
**Fig A2.2 Manual Input Location**

****

****

**Fig A2.3 Input of Date & Rice variety**

****

****

**Fig A2.4 Predicted Yield Shown**

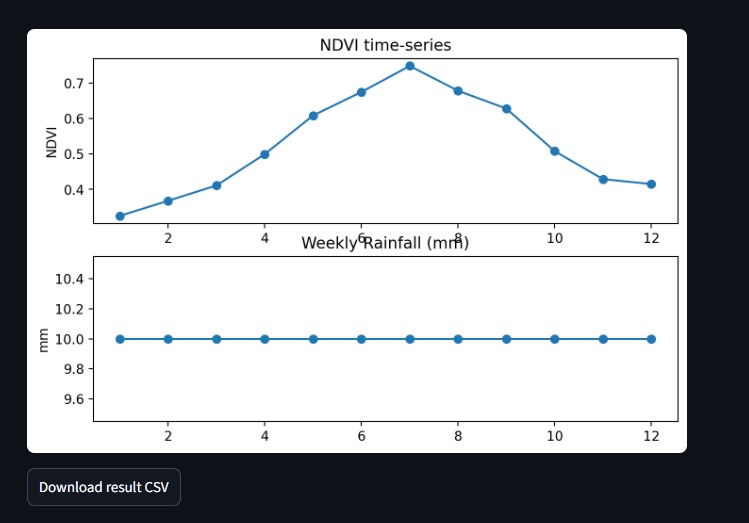
****

Fig A2.5 :Line Plot Of The Inputs Predicting The Output

**A3. PAPER PUBLICATION**

# Deep Learning-Based Yield Forecasting for Rice Varieties

Dr. Subedha V Department of Computer Science Panimalar Engineering College

Chennai, India [subedha@gmail.com](mailto:subedha@gmail.com)

Tanushri E Department of Computer Science Panimalar Engineering College

Chennai, India [tanumini456@gmail.com](mailto:tanumini456@gmail.com)

Mrs. S. Sharmila Department of Computer Science Panimalar Engineering College

Chennai, India [sharmilapanimalar2024@gmail.com](mailto:sharmilapanimalar2024@gmail.com)

Thiruselvi Thirunavukkarasu Department of Computer Science Panimalar Engineering College Chennai, India [thiruselvithiruna@gmail.com](mailto:thiruselvithiruna@gmail.com)

Dr. Sathiya V Department of Computer Science Panimalar Engineering College

Chennai, India [deviviji2000@yahoo.co.in](mailto:deviviji2000@yahoo.co.in)

Dr. L. Jabasheela Department of Computer Science Panimalar Engineering College

Chennai, India [ljsheela@gmail.com](mailto:ljsheela@gmail.com)

***Abstract*—**In most regions, agriculturalists are facing serious yield loss due to unpredictable climates, absence of effective localised forecasting systems and non-availability of planning tools specific to a particular variety. These issues often result in a poor choice of crops, reduced productivity and uncertain income. With growing climate uncertainty, the demand for prediction systems that are not only smart but also transparent and can assist farmers in choosing the most suitable rice varieties for their location and time of year is also on the rise. This introduces a deep learning-based rice yield estimation model with the use of time-series Normalised Difference Vegetation Index (NDVI) and climate data, as well as the use of rice variety as the primary feature input. The Long Short-Term Memory (LSTM) networks are used to make a model that can find patterns in time and give more accurate estimates. Through the provision of forecasts for a particular variety, the system enables the farmers to choose the most appropriate rice type based on the current climatic conditions, which ultimately leads to effective agro planning and income stabilisation. It helps farmers' livelihoods by promoting environmentally friendly crop production and making farming more resilient to climate change, which is in line with UN SDG-2.

Keywords— Rice Yield Prediction, NDVI, Climate Data, Deep Learning, LSTM, Agriculture AI

##### I . INTRODUCTION

Rice is the global pre-eminent staple crop, feeding over half of the world's inhabitants. In India, it is not only a prime food staple but also a major economic power, generating widely in rural areas and national food security. Tamil Nadu and most importantly Thanjavur district, often called the "Rice Bowl of Tamil Nadu" is the backbone of rice cultivation. But the rice cultivation of the region is beset by chronic issues coming from climate instability, unpredictable rainfalls and a lack of a relevant forecasting mechanism.

Unpredictable weather events like late monsoons, drought and increased temperatures are part of the causes of the variability of rice yields. The Ministry of Agriculture (2023) cited that almost 40% of Indian farmers are insecure with their incomes due to unpredictable crop yields. This presents economic risks to farmers as well as to the

food supply chain as a whole, including consumers and policymakers. Traditional crop yield estimates undertaken by agricultural ministries are mostly via slow, expensive and time- consuming field-based manual surveys. Due to this, farmers have no access to timely, actionable information for effective planning of crop cycles.

Over the last few years, scientists have applied yield forecasting by using statistical regression techniques and traditional machine learning algorithms like Random Forests, Support Vector Machines (SVM) and LASSO regression. While these approaches have been promising in detecting yield-climate relationships, they also have some built-in constraints. They are all based on seasonal averages of climate data, ignoring the momentary fluctuations and growth- stage-specific evolution of plant health. Besides this, rice is also considered a homogeneous crop, neglecting varietal heterogeneity. Practically, rice varieties vary highly in growth duration, strength and yield potential based on agro-climatic conditions. Ignoring these differences leads to forecasting on a large scale that is useless for local farmers.

With the emergence of remote sensing technologies, more specifically satellite-based indicators like the Normalised Difference Vegetation Index (NDVI), the game in agricultural monitoring has changed. NDVI serves as a valuable surrogate for crop growth processes, biomass and vegetation vigour and therefore a good indicator of crop yield potential.

Likewise, global climate data from locations like NASA POWER provide daily estimates of variables like precipitation, temperature, and humidity, to be utilised for fine-grained modelling of environmental effects on crops. While some research studies have utilised NDVI and climatic data individually for yield forecasting, not many have utilised their combined power and fewer have used them for variety-specific rice yield estimation in particular regions such as Tamil Nadu.



**Fig 1**: Thanjavur Region Map

Deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, have also shown immense promise in handling time-series data such as vegetation indices and weather data. Unlike traditional ML methods, LSTMs can learn sequential relationships, thereby ideal for modelling crop development patterns at multiple stages of the season. While such strengths do exist, the use of LSTM for rice yield prediction is not common and no study has employed variety-level data explicitly within the prediction framework.

The model addresses these deficiencies by outlining a deep learning rice yield forecasting model specific to the Thanjavur district. The model incorporates NDVI time-series, daily climate and rice variety as an added input in order to provide variety-specific, localised forecasted yields. This not only enhances prediction accuracy but also enables farmers to make knowledgeable choices regarding which type of rice best suits their land and time of year. With higher transparency and explainability, the suggested system can act as a decision-support system for policymakers, Agri-officers and farmers.

In the course of executing the United Nations Sustainable Development Goal (SDG) 2 Zero Hunger and Target 2.4 in particular, the project will support the advancement of sustainable and resilient Agri-systems. By optimising uncertainty in rice crop cultivation and giving farmers actionable facts, the model can enhance food security and the stability of income among the vulnerable farmers groups.

##### MATERIALS AND METHODS

This section describes the meteorological characteristics, Datasets and data sources. The rice yield forecasting system that we describe integrates time-series satellite images and climate and rice variety with a deep learning system. The process uses the following four main steps.

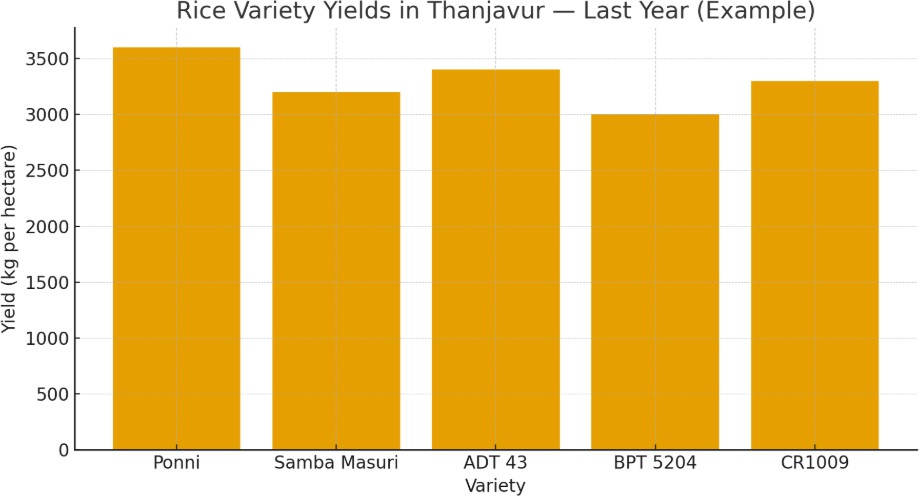
* 1. DATA SOURCES
     1. Normalised Difference Vegetation Index (NDVI):

NDVI is a satellite-image spectral index that serves as an indication of vegetation health and canopy density.Sentinel-2 MSI (Multispectral Instrument) dataset records are used as a source of NDVI data through the Google Earth Engine (GEE) platform. The Sentinel-2 sensor produces multispectral images at a resolution of

10–20 m and a revisit period of 5 days and is best suited for health monitoring of the Thanjavur district crops. NDVI time series are derived during the entire crop season and are smoothed as weekly means. Daily climatic variables of rainfall, min-temp, max-temp and mean-temp and relative humidity are extracted from the POWER dataset delivered through NASA. They directly affect crop development and yield and are important inputs for the model. They are resampled at weekly steps in order to co-register with the NDVI series.

* + 1. Rice Variety Information:

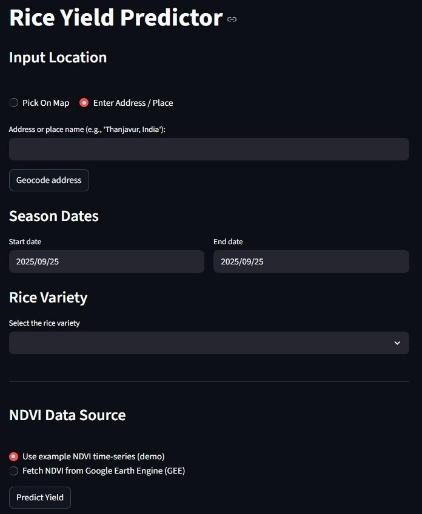
It has a different growth pattern compared to other varieties of rice. Data on varieties of Thanjavur are obtained from Tamil Nadu Agricultural University (TNAU) and local extension records of agriculture. Each variety appears as a category variable and is embedded into the model.



**Fig. 2**: Yield Of A Few Rice Varieties Of Last Year In Thanjavur

* + 1. Historical Yield Data:

The projec use ground-truth records of rice production from Department of Agriculture records and available statistical publications of the Thanjavur district. It serves as the dependent variable (label) during training and validation of the model.



**Fig 3:** Input Screen Of The Application

* 1. DATA PREPROCESSING
     1. Temporal Alignment:

NDVI and climatic variables are rescaled at weekly time-steps to allow synchronising of modalities.

* + 1. Noise Removal:

Satellite images are subject to cloud interference. Clouds are masked from Sentinel-2 images using the QA60 band before the calculation of NDVI.

* + 1. Feature Normalisation:

The climate variables (temperature, precipitation, humidity) are z- score standardised during training to avoid training instability.

* + 1. Variety Encoding:

The rice varieties are represented using one-hot encoding or embedding layers from the deep learning architecture to reflect varietal variations.

* + 1. Dataset Division:

The data is split into 70% training, 15% validation and 15% testing to guarantee solid assessment.

* 1. MODEL EVALUATION

Long Short-Term Memory (LSTM) networks, a kind of Recurrent Neural Network (RNN) made especially to learn and find intricate sequential patterns in time-dependent data, are used to implement the model.

* + 1. Input Layers:

Dataset 1: Time-series of NDVI and climate (12 weeks × 4–5 features). Input 2: Variety of rice (categorical).

* + 1. LSTM Layers:

Two LSTMs are also stacked one above the other for learning the temporal relations from the time series of climate and NDVI. They are used for regularisation.

* + 1. Variety Embedding:

Each rice variety is learned a low-dimensional embedding vector by which the model can learn variety-specific features.

1. Fusion Layer: The output of the LSTM module and the variety embedding are combined.
2. Dense Layers: The combined representation from above goes through fully connected layers and a ReLU activation function.
3. Output Layer: A single Dense neuron outputs the yield prediction (tons/hectare).

##### PROPOSED APPROACH

The ultimate target is to establish a system of rice yield forecasting using deep learning as a substitute of conventional and machine learning techniques. While typical models rely upon season-long means or assume rice as a generic type of crop, the system takes

advantage of time-series satellite images, day-by-day and location weather and variety traits and yields variety- and location-dependent predictions of yield.

* 1. INPUT PARAMETERS

This solution leverages the synergy of three critical data domains:

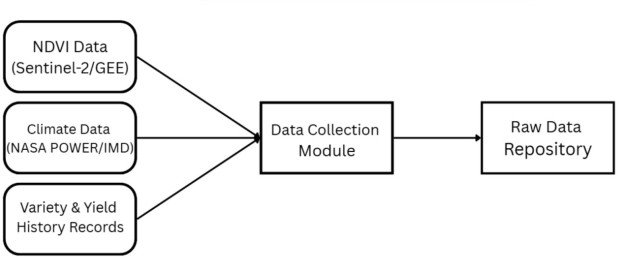
* + 1. Vegetation health (NDVI): Vegetation health defines the growth stage of the crop and the build-up of biomass with time.
    2. Climatic variables: Store short-term fluctuations of rainfall, temperature, and humidity.
    3. Variety information of rice: Carries genetic and phenotypic variation of varieties and offers more possible and practical yield estimation.

Along with such inputs through a Long Short-Term Memory (LSTM) model, the system could learn patterns of sequence from weather and NDVI data and produce variety-specific predictions. For Tamil Nadu, this is relevant as it has a vast number of rice varieties with contrasted growth behaviours and variety-specific yield potentials that are a function of climatic and soil conditions.

* 1. SYSTEM WORKFLOW

The desired workflow takes the following steps:

* + 1. Data Collection:
       - Google Earth Engine Sentinel-2 time-series of NDVI.
       - Daily weather variables (rainfall, temp., humidity) from POWER from NASA.
       - Variety-level yield statistics of backyard agriculture.



**Fig 4:** Data Collection module

* + 1. Data Processing:
       - Filtering and selecting valid NDVI frames: Only NDVI data from the rice-growing period in Thanjavur was selected and images with missing/invalid pixel values were removed.
       - Aligning climate data with NDVI timestamps: Climate data (temperature, rainfall, humidity) was matched to the same dates as NDVI extraction to create a consistent time-series input.
       - Feature scaling and category encoding: Climate and NDVI features were scaled to a uniform numeric range and rice varieties were encoded as numerical input for the deep learning model.
    2. Output & Interpretation:

The model provides a predicted rice yield based on the entered location, season and rice variety. The output also includes graphical

results such as yield comparison and NDVI to help understand crop growth. Framers can use this information to choose the right variety and plan cultivation for better productivity.

* 1. COMPARISON WITH PREVIOUS ALGORITHMS

The proposed solution has the following merits:

* + 1. Variety-specific predictions: In contrast with hypothesised

Where,

* *Yi* :Actual value for the ith observation
* *Ŷi*: Calculated value for the ith observation
* n: Total number of observations
  + 1. Root Mean Square Error (RMSE): Penalises the large error more than MAE.

models that make generic assumptions across varieties of rice, the approach delivers variety-specific predictions offering actionable knowledge to farmers.

1. Temporal learning: Using weekly modelling of climatic patterns and NDVI, LSTM learns intra-seasonal behaviour missed with traditional methods.

*RSME =* √

𝑁

1 ∑𝑁

Where,

* + N= Total number of data points
  + yi = Actual (true) value
  + *Ŷi* = Predicted value

𝑖=1(𝑦𝑖 − 𝑦̂𝑖)2

1. Localised attention: It has been developed for the Thanjavur district and can be utilised with variations with other regions and crops.

##### RESULTS AND DISCUSSIONS

*  = Squared error for each point
* RMSE = Root Mean Squared Error

1. Coefficient of Determination (R²): Explains the proportion of variance of yield accounted for by the model.
   1. EXPERIMENTAL SETUP

𝑅2 = 1 −

𝑁

𝑖=1

∑

∑𝑁

(𝑦𝑖 − 𝑦̂𝑖)2 (𝑦𝑖 − 𝑦̅)2

The model was trained with synthetic and partly available real

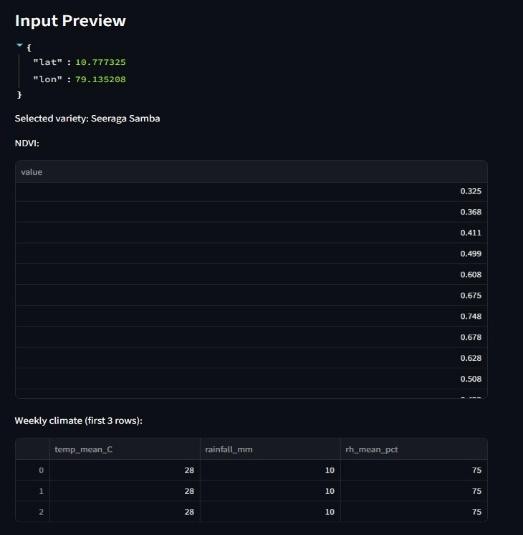
datasets generated from Sentinel-2 NDVI time-series, NASA POWER daily climate variables and past yield records of the Thanjavur district. The dataset was divided into 70% training, 15% validation and 15% testing. TensorFlow 2.x and a workstation with a GPU were used to conduct the training of the model. Optimisation of the LSTM network used the Adam optimiser with a learning rate of 0.001 and trained the network for 80 iterations with early stopping to avoid overfitting.

Where,

* Yi = Actual true values
* Ŷi = Predicted values
* ȳ = Mean of actual values
* N = Total number of samples

C. QUANTITATIVE RESULTS

𝑖=1



**Fig 5**: Preview Of Input Shown

* 1. EVALUATION METRICS

Model quality is evaluated with typical regression measures:

1.Mean Absolute Error (MAE): Estimates the approximate average difference of predicted and actual yields.

*MAE =* 1 ∑𝑛 |𝑦 − 𝑦̂ |

𝑖=1 𝑖 𝑖

𝑛

Test results of the model with the test data are tabulated in Table I.

**Table I.** Performance of Proposed LSTM Model

|  |  |
| --- | --- |
| **Metric** | **Value** |
| MAE | 0.25 tons/ha |
| RMSE | 0.32 tons/ha |
| R² | 0.47 |

They suggest that the model may reproduce the connection of NDVI with climate and rice yield with fair accuracy. Although the explanatory power of R² suggests a moderate level of explanation, further refinement may also be achieved through the addition of historical observations and an expanded training dataset.

1. VARIETY-SPECIFIC PREDICT

One of the major contributions of this project is variety-specific prediction. For instance, at Thanjavur, three varieties namely ADT 37, CR 1009 and BPT 5204 were tested.

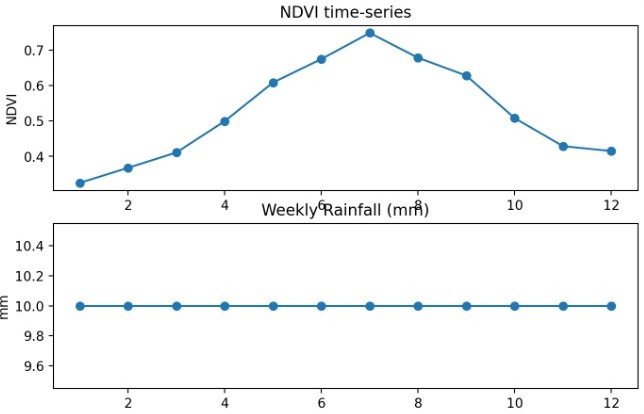
ADT 37 has higher yield stability under climatic variability of rainfall and BPT 5204 has higher temperature sensitivity. It indicates the significance of variety as an explicit input of the prediction process.

1. VISUALIZATION OF RESULTS

Predicted rice yield is presented in tons per hectare in a clear and interpretable format. NDVI growth and climate variations are shown through simple line charts for better understanding. Yield comparison visuals help identify the most suitable rice variety for the selected location and season. Color-based indicators highlight whether the predicted yield is low, moderate or high. These visual insights support better crop planning and decision-making for farmers and agricultural officers.



**Fig 6** : The Predicted Yield Is Given



**Fig 7:** Line Plot Of The Inputs Predicting The Output

1. DISCUSSION

Experimental results show that with NDVI time-series, climatic variability and rice variety embedding, locally and realistically better forecasts are possible as compared with usual ML-based forecasting. Even with the small dataset used during this mini- project, the developed framework can also be scalable. System accuracy can also be enhanced through the inclusion of much larger datasets, like soil and agronomic practices. In addition, the explainability module ensures that stakeholders understand why a prediction has been made and increases AI-model uptake for agriculture. Importantly, the system contributes towards SDG 2: Zero Hunger since it contributes both directly and towards sustainable cropping calendars and food security.

##### CONCLUSION

This study proposed a deep learning-assisted rice yield forecasting system that integrated NDVI time-series, weather variables and rice

variety information for making variety-specific and localised forecasts of rice yield at the Thanjavur district level in Tamil Nadu. Using LSTM networks efficiently extracted temporal patterns of weather and crop health and variety embeddings facilitated variety- distinctive forecasting. Experimental findings validated that the model simulated well and NDVI at the reproductive phase and rainfall heterogeneity were vital yield determinants.

Integration of interpretability, as analysis assisted in establishing confidence and interpretability and thus the system was appropriate for real-world agriculture decision-making. In contrast with conventional models, the approach developed herein incorporates temporal dynamics and varietal diversity and therefore reveals actionable information both to farmers and policymakers. Prospects are the integration of the framework with soil and pest information and with management and the use of the developed approach as a mobile app for real-time use. Overall, the system enables sustainable agriculture planning of crops and food security and contributes to UN SDG 2 – Zero Hunger.

##### REFERENCES

1. Rice Yield Estimation using LSTM with Meteorological Data,

IEEE Access, 2022

1. Indian Crop Yield Prediction using LSTM Deep Learning Networks
2. Advanced Rice Yield and Soil Prediction using Deep Learning
3. Paddy Yield Forecasting using Regression Techniques
4. Mapping Indian Rice Yields through Statistical Modelling using Satellite and Climatic Data
5. Enhancing Rice Yield Prediction using Vision Transformer and Regression Models
6. A Time-Series Based Yield Forecasting Model Using Stacked LSTM... Cauvery Delta, Tamil Nadu
7. Exploring Crop Yield Prediction with Remote Sensing Imagery and AI (Kavipriya & Vadivu, ACCAI 2024)
8. Predicting Crop Yield with NDVI and Backscatter Values using Deep Neural Networks (MECON 2022)
9. Rice Modelling Using Long Time Series of High Temporal Resolution Vegetation Indices in Nepal
10. Hybrid LSTM and SVM Method Rice Yield Prediction in Densely Populated Areas
11. BORO Rice Yield Estimation Model Using MODIS NDVI Data for Bangladesh
12. Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularised Extreme Learning Machine
13. Deep Learning-Enhanced Remote Sensing-Integrated Crop Modelling for Rice Yield Prediction (Ecological Informatics, 2024)
14. Rice Yield Prediction using Radar Vegetation Indices from Sentinel-1 Data and Multiscale Conv1D-LSTM Network (Journal of Applied Remote Sensing, 2024)

Page 1 of 9 - Cover Page Submission ID trn:oid:::2945:321637458

**A4. PLAGIARISM REPORT**

**F7\_CONFERENCE\_PAPER**

## F7\_CONFERENCE\_PAPER

 Turnitin

###### Document Details

Submission ID trn:oid:::2945:321637458

5 Pages

2,908 Words

16,862 Characters

Submission Date

Oct 25, 2025, 5:13 PM GMT+5

Download Date

Oct 25, 2025, 5:14 PM GMT+5

File Name unknown\_filename

File Size

556.4 KB

Page 2 of 9 - Integrity Overview Submission ID trn:oid:::2945:321637458

## 7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

###### Filtered from the Report

 Bibliography  Quoted Text

###### Match Groups

 20 Not Cited or Quoted 7%

Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%



Matches that are still very similar to source material

0 Missing Citation 0%



Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%



Matches with in-text citation present, but no quotation marks

###### Top Sources

|  |  |
| --- | --- |
| 6% | Internet sources |
| 5% | Publications |
| 5% | Submitted works (Student Papers) |

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.



###### Match Groups

Page 3 of 9 - Integrity Overview Submission ID trn:oid:::2945:321637458

###### Top Sources

20 Not Cited or Quoted 7%



|  |  |
| --- | --- |
| 6% | Internet sources |
| 5% | Publications |
| 5% | Submitted works (Student Papers) |

Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%



Matches that are still very similar to source material

0 Missing Citation 0%



Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%



Matches with in-text citation present, but no quotation marks

###### Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1

Internet

[www.nature.com](https://www.nature.com/articles/s41598-025-11780-7?code=fc6a6cd1-8b65-45f5-af7a-548f98bb13e9&error=cookies_not_supported) 2%

2

Student papers

Panimalar Engineering College on 2025-09-09 <1%

3

Internet

[www.isca.co.in](http://www.isca.co.in/SOUV_PROC/conference/ISBN%20978-93-86675-42-2.pdf) <1%

4

Internet

[ijsrem.com](https://ijsrem.com/download/bitcoin-price-prediction-using-lstm-based-deep-learning-techniques-in-python-a-data-driven-approach/?refresh=6830601fec2de1748000799&wpdmdl=47304) <1%

5

Internet

[nano-ntp.com](https://nano-ntp.com/index.php/nano/article/download/4959/3905/9671) <1%

6

Internet

[iqac.gndu.ac.in](https://iqac.gndu.ac.in/Pdf/AQAR/AQAR%202021-22.pdf) <1%

7

Internet

[www.preprints.org](https://www.preprints.org/manuscript/202403.0200/v1) <1%

8

Internet

[urfjournals.org](https://urfjournals.org/open-access/predictive-machine-learning-models-for-calibration-failure-detection-in-pharmaceutical-manufacturing.pdf) <1%

9

Publication

[Bharati S. Ainapure, Mythili Boopathi, Chandra Sekhar Kolli, C. Jackulin. "Deep En…](https://doi.org/10.1142/S0219467823500341) <1%

10

Internet

[faolex.fao.org](https://faolex.fao.org/docs/pdf/mul204201.pdf) <1%

Page 4 of 9 - Integrity Overview Submission ID trn:oid:::2945:321637458

11

Publication

[Faouzi Tayalati, Abdellah Azmani, Monir Azmani. "Application of supervised mach…](https://doi.org/10.1007/s13748-024-00318-z) <1%

12

Publication

[Isabel Jarro-Espinal, José Huanuqueño-Murillo, Javier Quille-Mamani, David Quisp…](https://doi.org/10.3390/agriculture15192054) <1%

13

Student papers

Centre for Technical Support on 2025-10-15 <1%

14

Student papers

Erasmus University Rotterdam on 2025-10-22 <1%

15

Student papers

University of East Anglia on 2022-03-16 <1%

 Page 5 of 9 - Integrity Submission Submission ID trn:oid:::2945:321637458

# Deep Learning-Based Yield Forecasting for Rice Varieties

Dr. Subedha V Dr. Sathiya V

Department of Computer Science

Engineering College

Mrs. S.

Department of Computer Science

Engineering College

Department of Computer Science

Panimalar Engineering College

Sharmila

Panimalar

**39**

Panimalar

[subedha@gmail.com](mailto:subedha@gmail.com)

Chennai, India

Chennai, India

Chennai, India

Tanushri E

Panimalar Engineering College

Chennai, India

Department of Computer Science

[sharmilapanimalar2024@gmail.com](mailto:sharmilapanimalar2024@gmail.com)

Thiruselvi Thirunavukkarasu

Panimalar Engineering College

Chennai, India

Department of Computer Science

[deviviji2000@yahoo.co.in](mailto:deviviji2000@yahoo.co.in)

Dr. L. Jabasheela

Panimalar Engineering College

Chennai, India

Department of Computer Science

**6**

**2**

tanumini456@

thiruselvithiruna@

ljsheela@

most regions, agriculturalists are facing serious yield loss due to unpredictable climates, absence of effective localised forecasting systems and non-availability of planning tools specific to a particular variety. These issues often result in a poor choice of crops, reduced productivity and uncertain income. With growing climate uncertainty, the demand for prediction systems that are not only smart but also transparent and can assist farmers in choosing the most suitable rice varieties for their location and time of year is also on the rise. This introduces a deep learning-based rice yield estimation model with the use of time-series Normalised Difference Vegetation Index (NDVI) and climate data, as well as the use of rice variety as the primary feature input. The

gmail.com

gmail.com

gmail.com

***Abstract*—**In

Long Short-Term Memory



**4**



**8**

used make a model that can find patterns in time and give more accurate estimates. Through the provision of forecasts for a particular variety, the system enables the farmers to choose the most appropriate rice type based on the current climatic conditions, which ultimately leads to effective agro planning and income stabilisation. It helps farmers' livelihoods by promoting environmentally friendly crop production and making farming more resilient to climate change, which is in line with UN SDG-2.

(LSTM) networks are

to

Keywords— Rice Yield Prediction, NDVI, Climate Data, Deep Learning, LSTM, Agriculture AI

##### I . INTRODUCTION

Rice is the global pre-eminent staple crop, feeding over half of the world's inhabitants. In India, it is not only a prime food staple but also a major economic power, generating widely in rural areas and national food security. Tamil Nadu and most importantly Thanjavur district, often called the "Rice Bowl of Tamil Nadu" is the backbone of rice cultivation. But the rice cultivation of the region is beset by chronic issues coming from climate instability, unpredictable rainfalls and a lack of a relevant forecasting mechanism.

Unpredictable weather events like late monsoons, drought and increased temperatures are part of the causes of the variability of rice yields. The Ministry of Agriculture (2023) cited that almost 40% of Indian farmers are insecure with their incomes due to unpredictable crop yields. This presents economic risks to farmers as well as to the

food supply chain as a whole, including consumers and policymakers. Traditional crop yield estimates undertaken by agricultural ministries are mostly via slow, expensive and time- consuming field-based manual surveys. Due to this, farmers have no access to timely, actionable information for effective planning of crop cycles.

Over the last few years, scientists have applied yield forecasting by using statistical regression techniques and

traditional machine

Machines ( and LASSO regression. While these approaches have been promising in detecting yield-climate relationships, they also have some built-in constraints. They are all based on seasonal averages of climate data, ignoring the momentary fluctuations and growth- stage-specific evolution of plant health. Besides this, rice is also considered a homogeneous crop, neglecting varietal heterogeneity. Practically, rice varieties vary highly in growth duration, strength and yield potential based on agro-climatic conditions. Ignoring these differences leads to forecasting on a large scale that is useless for local farmers.

learning algorithms like Random Forests, Support Vector

SVM)

With the emergence of remote sensing technologies, more specifically satellite-based indicators like the Normalised Difference Vegetation Index (NDVI), the game in agricultural monitoring has changed. NDVI serves as a valuable surrogate for crop growth processes, biomass and vegetation vigour and therefore a good indicator of crop yield potential.

Likewise, global climate data from locations like NASA POWER provide daily estimates of variables like precipitation, temperature, and humidity, to be utilised for fine-grained modelling of environmental effects on crops. While some research studies have utilised NDVI and climatic data individually for yield forecasting, not many have utilised their combined power and fewer have used them for variety-specific rice yield estimation in particular regions such as Tamil Nadu.

Page 6 of 9 - Integrity Submission Submission ID trn:oid:::2945:321637458





**12**

**Fig 1**: Thanjavur Region Map

(LSTM) networks, have also shown immense promise

in handling

Deep learning techniques, particularly Long Short-Term Memory

**4**

**1**

vegetation indices and data. Unlike traditional ML methods, LSTMs can learn sequential relationships, thereby ideal for modelling crop development patterns at multiple stages of the season. While such strengths do exist, the use of LSTM for rice yield prediction is not common and no study has employed variety-level data explicitly within the prediction framework.

time-series data such as

weather

The model addresses these deficiencies by outlining a deep learning rice yield forecasting model specific to the Thanjavur district. The model incorporates NDVI time-series, daily climate and rice variety as an added input in order to provide variety-specific, localised forecasted yields. This not only enhances prediction accuracy but also enables farmers to make knowledgeable choices regarding which type of rice best suits their land and time of year. With higher transparency and explainability, the suggested system can act as a decision-support system for policymakers, Agri-officers and farmers.

In the course of executing the United Nations Sustainable

4 in particular, the project will support the advancement of sustainable and resilient Agri-systems. By optimising uncertainty in rice crop cultivation and giving farmers actionable facts, the model can enhance food security and the stability of income among the vulnerable farmers groups.

Development Goal (SDG) 2 Zero Hunger and Target 2.



**10**

##### II.

**MATERIALS AND METHODS**

This section describes the meteorological characteristics, Datasets



**1**

The rice yield forecasting system that we describe integrates time-series satellite images and climate and rice variety with a deep learning system. The process uses the following four main steps.

and data sources.

1. DATA SOURCES
   1. Normalised Difference Vegetation Index (NDVI):

NDVI is a satellite-image spectral index that serves as an indication of vegetation health and canopy density.Sentinel-2 MSI (Multispectral Instrument) dataset records are used as a source of NDVI data through the Google Earth Engine (GEE) platform. The Sentinel-2 sensor produces multispectral images at a resolution of

period 5 and is best suited for health monitoring of the Thanjavur district crops. NDVI time series are derived during the entire crop season and are smoothed as weekly means. Daily climatic variables of rainfall, min-temp, max-temp and mean-temp and relative humidity are extracted from the POWER dataset delivered through NASA. They directly affect crop development and yield and are important inputs for the model. They are resampled at weekly steps in order to co-register with the NDVI series.

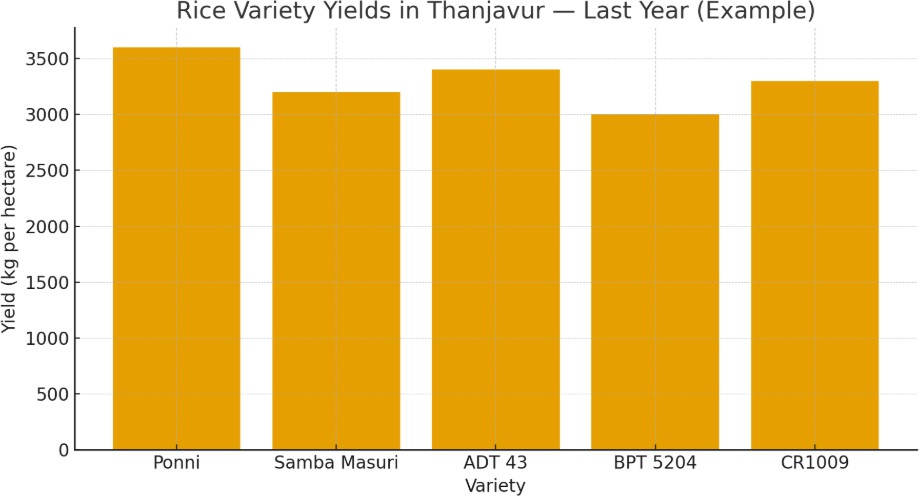
* 1. Rice Variety Information:

10–20 m and a revisit

of

days

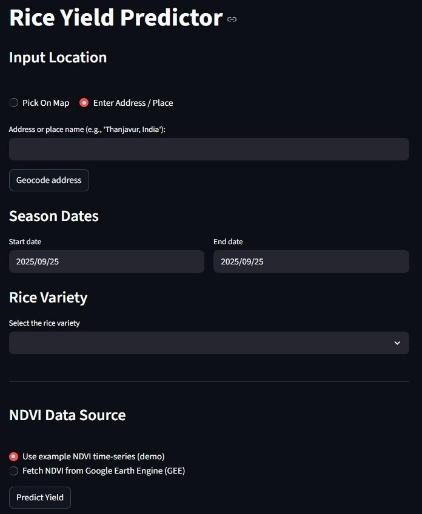
It has a different growth pattern compared to other varieties of rice. Data on varieties of Thanjavur are obtained from Tamil Nadu Agricultural University (TNAU) and local extension records of agriculture. Each variety appears as a category variable and is embedded into the model.



**Fig. 2**: Yield Of A Few Rice Varieties Of Last Year In Thanjavur

* 1. Historical Yield Data:

The projec use ground-truth records of rice production from Department of Agriculture records and available statistical publications of the Thanjavur district. It serves as the dependent variable (label) during training and validation of the model.



**Fig 3:** Input Screen Of The Application

Page 7 of 9 - Integrity Submission Submission ID trn:oid:::2945:321637458

1. DATA PREPROCESSING
   1. Temporal Alignment:

NDVI and climatic variables are rescaled at weekly time-steps to allow synchronising of modalities.

* 1. Noise Removal:

Satellite images are subject to cloud interference. Clouds are masked from Sentinel-2 images using the QA60 band before the calculation of NDVI.

* 1. Feature Normalisation:

The climate variables (temperature, precipitation, humidity) are z- score standardised during training to avoid training instability.

* 1. Variety Encoding:



**14**

The rice varieties are represented using one-hot encoding or embedding layers from the deep learning architecture to reflect varietal variations.

* 1. Dataset Division:

The data is split into 70% training, 15% validation and 15% testing



**7**

to guarantee solid assessment.

1. MODEL EVALUATION

Long Short-Term Memory (LSTM) networks, a kind of Recurrent

Neural Network (RNN)

to



**5**

made especially learn and find intricate sequential patterns in time-dependent data, are used to implement the model.

* 1. Input Layers:

Dataset 1: Time-series of NDVI and climate (12 weeks × 4–5 features). Input 2: Variety of rice (categorical).

* 1. LSTM Layers:

Two LSTMs are also stacked one above the other for learning the temporal relations from the time series of climate and NDVI. They are used for regularisation.

* 1. Variety Embedding:

Each rice variety is learned a low-dimensional embedding vector by which the model can learn variety-specific features.

1. Fusion Layer: The output of the LSTM module and the variety embedding are combined.
2. Dense Layers: The combined representation from above goes through fully connected layers and a ReLU activation function.
3. Output Layer: A single Dense neuron outputs the yield prediction (tons/hectare).

##### PROPOSED APPROACH

The ultimate target is to establish a system of rice yield forecasting using deep learning as a substitute of conventional and machine learning techniques. While typical models rely upon season-long means or assume rice as a generic type of crop, the system takes

advantage of time-series satellite images, day-by-day and location weather and variety traits and yields variety- and location-dependent predictions of yield.

1. INPUT PARAMETERS

This solution leverages the synergy of three critical data domains:

* 1. Vegetation health (NDVI): Vegetation health defines the growth stage of the crop and the build-up of biomass with time.
  2. Climatic variables: Store short-term fluctuations of rainfall, temperature, and humidity.
  3. Variety information of rice: Carries genetic and phenotypic variation of varieties and offers more possible and practical yield estimation.

Along with such inputs through

a Long Short-Term Memory

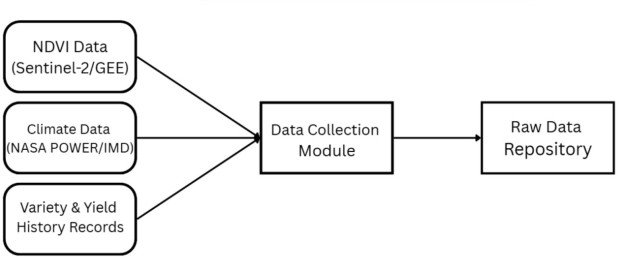
model, the system could learn patterns of sequence from weather and NDVI data and produce variety-specific predictions. For Tamil Nadu, this is relevant as it has a vast number of rice varieties with contrasted growth behaviours and variety-specific yield potentials that are a function of climatic and soil conditions.

(LSTM)

1. SYSTEM WORKFLOW

The desired workflow takes the following steps:

* 1. Data Collection:
     + Google Earth Engine Sentinel-2 time-series of NDVI.
     + Daily weather variables (rainfall, temp., humidity) from POWER from NASA.
     + Variety-level yield statistics of backyard agriculture.



**Fig 4:** Data Collection module

* 1. Data Processing:
     + Filtering and selecting valid NDVI frames: Only NDVI data from the rice-growing period in Thanjavur was selected and images with missing/invalid pixel values were removed.
     + Aligning climate data with NDVI timestamps: Climate data (temperature, rainfall, humidity) was matched to the same dates as NDVI extraction to create a consistent time-series input.
     + Feature scaling and category encoding: Climate and NDVI features were scaled to a uniform numeric range and rice varieties were encoded as numerical input for the deep learning model.
  2. Output & Interpretation:

The model provides a predicted rice yield based on the entered location, season and rice variety. The output also includes graphical

Page 8 of 9 - Integrity Submission Submission ID trn:oid:::2945:321637458

results such as yield comparison and NDVI to help understand crop growth. Framers can use this information to choose the right variety and plan cultivation better productivity.

for

**11**

**13**

**15**

1. COMPARISON WITH PREVIOUS ALGORITHMS



**1**

proposed solution has the following merits:

The

* 1. Variety-specific predictions: In contrast with hypothesised models that make generic assumptions across varieties of rice, the

Where,



*Yi* :Actual value for the ith observation

*Ŷi*: Calculated value for the ith observation





n: Total number of observations

*RSME =* √

MAE.

Penalises the large more

(𝑦 − 𝑦̂ )2

2. Root Mean Square Error (RMSE)

error

:

than

1 ∑𝑁

approach delivers variety-specific predictions offering actionable knowledge to farmers.

* 1. Temporal learning: Using weekly modelling of climatic patterns and NDVI, LSTM learns intra-seasonal behaviour missed with traditional methods.

𝑁

Where,

* N= Total number of data points
* yi = Actual (true) value
* *Ŷi* = Predicted value

𝑖=1 𝑖 𝑖

* 1. Localised attention: It has been developed for the Thanjavur district and can be utilised with variations with other regions and crops.



**1**

##### RESULTS AND DISCUSSIONS

*  = Squared error for each point
* RMSE =

Root Mean Squared Error

3. ): Explains the proportion of

Coefficient of Determination (R²

variance of yield accounted for by the model.

1. EXPERIMENTAL SETUP

𝑅2 = 1 −

𝑁

𝑖=1

∑

∑𝑁

(𝑦𝑖 − 𝑦̂𝑖)2 (𝑦𝑖 − 𝑦̅)2

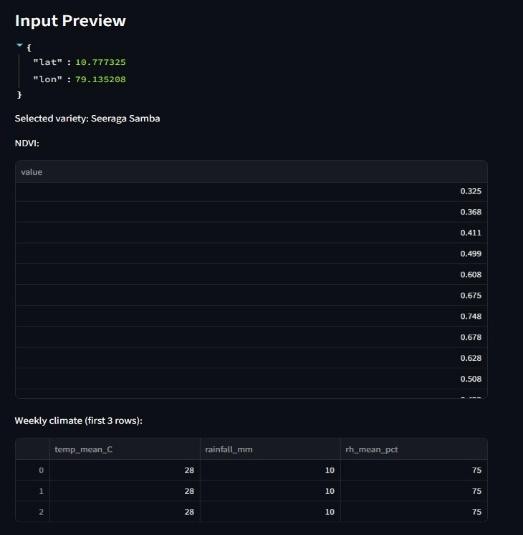
The model was trained with synthetic and partly available real

datasets generated from Sentinel-2 NDVI time-series, NASA POWER daily climate variables and past yield records of the Thanjavur district. The dataset was divided into 70% training, 15% validation and 15% testing. TensorFlow 2.x and a workstation with a GPU were used to conduct the training of the model. Optimisation of the LSTM network used the Adam optimiser with a learning rate of 0.001 and trained the network for 80 iterations with early stopping to avoid overfitting.

Where,

* Yi = Actual true values
* Ŷi = Predicted values
* ȳ = Mean of actual values
* N = Total number of samples
  1. QUANTITATIVE RESULTS

𝑖=1



**Fig 5**: Preview Of Input Shown

1. EVALUATION METRICS

Model quality is evaluated with typical regression measures:

1.Mean Absolute Error (MAE): Estimates the approximate average difference of predicted and actual yields.

*MAE =* 1 ∑𝑛 |𝑦 − 𝑦̂ |

𝑖=1 𝑖 𝑖

𝑛

Test results of the model with the test data are tabulated in Table I.

**Table I.** Performance of Proposed LSTM Model

|  |  |
| --- | --- |
| **Metric** | **Value** |
| MAE | 0.25 tons/ha |
| RMSE | 0.32 tons/ha |
| R² | 0.47 |

They suggest that the model may reproduce the connection of NDVI with climate and rice yield with fair accuracy. Although the explanatory power of R² suggests a moderate level of explanation, further refinement may also be achieved through the addition of historical observations and an expanded training dataset.

* 1. VARIETY-SPECIFIC PREDICT

One of the major contributions of this project is variety-specific prediction. For instance, at Thanjavur, three varieties namely ADT 37, CR 1009 and BPT 5204 were tested.

ADT 37 has higher yield stability under climatic variability of rainfall and BPT 5204 has higher temperature sensitivity. It indicates the significance of variety as an explicit input of the prediction process.

Page 9 of 9 - Integrity Submission Submission ID trn:oid:::2945:321637458

* 1. VISUALIZATION OF RESULTS

Predicted rice yield is presented in tons per hectare in a clear and interpretable format. NDVI growth and climate variations are shown through simple line charts for better understanding. Yield comparison visuals help identify the most suitable rice variety for the selected location and season. Color-based indicators highlight whether the predicted yield is low, moderate or high. These visual insights support better crop planning and

decision-making for



**1**

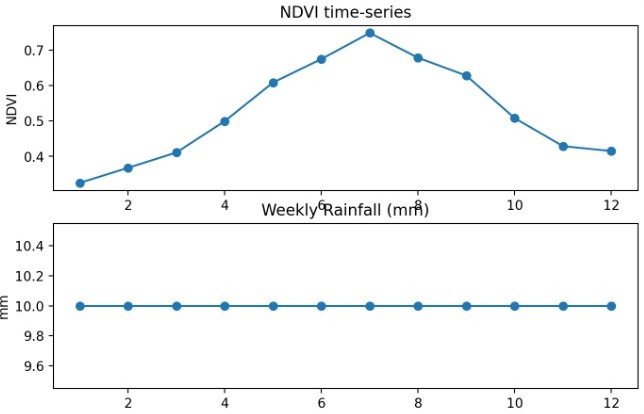
officers.

farmers and agricultural

**6** : The



**Fig **Predicted Yield Is Given



**Fig 7:** Line Plot Of The Inputs Predicting The Output

* 1. DISCUSSION

Experimental results show that with NDVI time-series, climatic variability and rice variety embedding, locally and realistically better forecasts are possible as compared with usual ML-based forecasting. Even with the small dataset used during this mini- project, the developed framework can also be scalable. System accuracy can also be enhanced through the inclusion of much larger datasets, like soil and agronomic practices. In addition, the explainability module ensures that stakeholders understand why a prediction has been made and increases AI-model uptake for agriculture. Importantly, the system contributes towards SDG 2: Zero Hunger since it contributes both directly and towards sustainable cropping calendars and food security.

##### CONCLUSION

This study proposed a deep learning-assisted rice yield forecasting system that integrated NDVI time-series, weather variables and rice

variety information for making variety-specific and localised forecasts of rice yield at the Thanjavur district level in Tamil Nadu. Using LSTM networks efficiently extracted temporal patterns of weather and crop health and variety embeddings facilitated variety- distinctive forecasting. Experimental findings validated that the model simulated well and NDVI at the reproductive phase and rainfall heterogeneity were vital yield determinants.

Integration of interpretability, as analysis assisted in establishing confidence and interpretability and thus the system was appropriate for real-world agriculture decision-making. In contrast with conventional models, the approach developed herein incorporates temporal dynamics and varietal diversity and therefore reveals actionable information both to farmers and policymakers. Prospects are the integration of the framework with soil and pest information and with management and the use of the developed approach as a mobile app for real-time use. Overall, the system enables sustainable agriculture planning of crops and food security and contributes to UN SDG 2 – Zero Hunger.

##### REFERENCES

1. Rice Yield Estimation using LSTM with Meteorological Data,

IEEE Access, 2022

1. Indian Crop Yield Prediction using LSTM Deep Learning Networks
2. Advanced Rice Yield and Soil Prediction using Deep Learning
3. Paddy Yield Forecasting using Regression Techniques
4. Mapping Indian Rice Yields through Statistical Modelling using Satellite and Climatic Data
5. Enhancing Rice Yield Prediction using Vision Transformer and Regression Models
6. A Time-Series Based Yield Forecasting Model Using Stacked LSTM... Cauvery Delta, Tamil Nadu
7. Exploring Crop Yield Prediction with Remote Sensing Imagery and AI (Kavipriya & Vadivu, ACCAI 2024)
8. Predicting Crop Yield with NDVI and Backscatter Values using Deep Neural Networks (MECON 2022)
9. Rice Modelling Using Long Time Series of High Temporal Resolution Vegetation Indices in Nepal
10. Hybrid LSTM and SVM Method Rice Yield Prediction in Densely Populated Areas
11. BORO Rice Yield Estimation Model Using MODIS NDVI Data for Bangladesh
12. Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularised Extreme Learning Machine
13. Deep Learning-Enhanced Remote Sensing-Integrated Crop Modelling for Rice Yield Prediction (Ecological Informatics, 2024)
14. Rice Yield Prediction using Radar Vegetation Indices from Sentinel-1 Data and Multiscale Conv1D-LSTM Network (Journal of Applied Remote Sensing, 2024)

### 10. REFERENCES

1. Rice Yield Estimation using LSTM with Meteorological Data, IEEE Access, 2022
2. Indian Crop Yield Prediction using LSTM Deep Learning Networks
3. Advanced Rice Yield and Soil Prediction using Deep Learning
4. Paddy Yield Forecasting using Regression Techniques
5. Mapping Indian Rice Yields through Statistical Modeling using Satellite and Climatic Data
6. Enhancing Rice Yield Prediction using Vision Transformer and Regression Models
7. A Time-Series Based Yield Forecasting Model Using Stacked LSTM... Cauvery Delta, Tamil Nadu
8. Exploring Crop Yield Prediction with Remote Sensing Imagery and AI (Kavipriya & Vadivu, ACCAI 2024)
9. Predicting Crop Yield with NDVI and Backscatter Values using Deep Neural Networks (MECON 2022)
10. Rice Modeling Using Long Time Series of High Temporal Resolution Vegetation Indices in Nepal
11. Hybrid LSTM and SVM Method Rice Yield Prediction in Densely Populated Area
12. BORO Rice Yield Estimation Model Using MODIS NDVI Data for Bangladesh
13. Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularized Extreme Learning Machine
14. Deep Learning-Enhanced Remote Sensing-Integrated Crop Modeling for Rice Yield Prediction (Ecological Informatics, 2024)
15. Rice Yield Prediction using Radar Vegetation Indices from Sentinel-1 Data and Multiscale Conv1D-LSTM Network (Journal of Applied Remote Sensing, 2024)

52