

# **SRI SIDDHARTHA ACADEMY OF HIGHER EDUCATION**

(DEEMED TO BE UNIVERSITY)

**Accredited A+ Grade by NAAC**

## **SRI SIDDHARTHA INSTITUTE OF TECHNOLOGY**

(A Constituent College of SSAHE)

KUNIGAL ROAD, MARALURU, TUMAKURU-572105



**Dept. Skill Lab-III (Machine Learning Laboratory) – 22IS507**

### **Mini Project Report**

**on**

### **“GREEN INTELLIGENCE”**

Submitted in partial fulfillment of the requirement for the completion of V semester of

### **BACHELOR OF ENGINEERING**

**Submitted**

**by**

ADITHYA S BHAT

22IS002

BHARGAV M

22IS015

**Under the Guidance**

**of**

**NANDINI G R** GUIDE QUALIFICATION

Guide designation, Dept. of ISE



**DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING**

**“Accredited by NBA, NEW DELHI”**

**2022-25**

# SRI SIDDHARTHA ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

Accredited A<sup>+</sup> Grade by NAAC

## SRI SIDDHARTHA INSTITUTE OF TECHNOLOGY

(A Constituent College of SSAHE)

KUNIGAL ROAD, MARALURU, TUMAKURU-572105

### DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING



## CERTIFICATE

Certified that mini project entitled “**Green Intelligence**”, is a bonafide work carried out by Mr./Ms. **ADITHYA S BHAT (22IS002)** and Mr./Ms **BHARGAV M (22IS015)** in partial fulfillment of the requirement for the completion of V semester B.E. in **Information Science and Engineering** of Sri Siddhartha Institute of Technology-Tumakuru during the year 2024-25. It is certified that all corrections/suggestions indicated have been incorporated in the report. The report has been approved as it satisfies the academic requirements with respect to Mini Project work.

### Signature of the Guide

Nandini G R **GUIDE**

**QUALIFICATION**

designation,  
Dept. of Information Science & Engg.

### Signature of the HoD

Dr. Annapurna H S **Ph.D**

Professor and Head,  
Dept. of Information Science & Engg.

### Name of the Examiners

1. \_\_\_\_\_

2. \_\_\_\_\_

### Signature

\_\_\_\_\_

\_\_\_\_\_

## DECLARATION

We, **ADITHYA S BHAT (22IS002)** and **BHARGAV M (22IS015)** of V semester, Department of Information Science and Engineering of Sri Siddhartha Institute of Technology, Tumakuru, hereby declare that this Mini project titled, “**Green Intelligence**”, has been carried out by us under the supervision of Nandini G R, designation, Department of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru in partial fulfillment of the requirement for the completion of V semester in Information Science and Engineering.

Date: 26/12 /2024

Place: Tumakuru

**ADITHYA S BHAT (21IS002)**

**BHARGAV M (21IS015)**

# TABLE OF CONTENTS

## ABSTRACT

i

## LIST OF FIGURES

ii

Sl. No.	Description	Page No.
1.	Introduction	
	1.1 Background	1
	1.2 Problem Statement	2
	1.3 Aim and Objective	3
2.	Literature Survey	4
3.	System Requirements	5
4.	System Design	
	4.1 System Architecture	6-7
	4.2 Flow Diagram	8
5.	Implementation	9-10
6.	Results	11
7.	Conclusion	12
8.	Future Scope	13
	References	14

## LIST OF FIGURES

Fig.No.	Figure Name	Page No.
1	Flow diagram	8
2	Results	11

## ABSTRACT

"Green Intelligence" is an innovative machine learning project focused on enhancing agricultural productivity through the prediction of crop yields. By harnessing the power of data analytics and machine learning algorithms, the project aims to transform the way farmers approach crop management. The core objective is to develop a robust prediction model that utilizes diverse data sources, including historical climate conditions, soil health metrics, and past crop performance, to forecast future crop yields. Through techniques such as regression analysis, decision trees, and deep learning, the project builds predictive models that can adapt to a variety of crops and environmental conditions. The primary goal is to assist farmers in making informed decisions on resource allocation, irrigation, and cultivation practices, thereby optimizing overall crop production and improving food security.

The project also focuses on sustainability, utilizing technology to promote environmentally friendly farming practices. By improving the accuracy of yield predictions, "Green Intelligence" helps minimize the overuse of resources such as water and fertilizers, promoting sustainable farming. The machine learning models are designed to continuously improve as new data is collected, providing real-time insights that farmers can act upon. Ultimately, "Green Intelligence" aims to bridge the gap between technology and traditional farming methods, enabling farmers to use data-driven insights to adapt to climate change, reduce environmental impact, and increase agricultural productivity, thereby contributing to long-term global food security.

## CHAPTER 1

# INTRODUCTION

### 1.1 Background:

Agriculture plays a crucial role in ensuring food security for the growing global population. However, the sector faces numerous challenges, including climate change, resource constraints, and the need to produce more food with fewer environmental impacts. Traditional farming practices, while effective, often lack the precision needed to address these challenges in an efficient and sustainable manner. With increasing environmental unpredictability and the demand for higher crop yields, there is a critical need for innovation in agricultural practices that can enhance decision-making and resource management.

Machine learning technologies offer promising solutions to these challenges by enabling data-driven insights and predictive analytics. By analyzing large datasets encompassing climate patterns, soil health, crop varieties, and historical yields, machine learning can uncover complex relationships and trends that human observation may overlook. The ability to forecast crop yields with high accuracy allows farmers to optimize their practices, reduce waste, and minimize environmental impact. The "Green Intelligence" project is born from this need to integrate modern machine learning techniques into agriculture, aiming to improve crop yield predictions and drive more sustainable farming practices. Through this initiative, we seek to empower farmers with tools that promote efficiency, reduce resource overuse, and ultimately contribute to sustainable food production.

## **1.2 Problem Statement:**

The agricultural sector is grappling with the increasing challenges of climate change, resource scarcity, and the pressure to meet the demands of a rapidly growing global population. One of the primary issues faced by farmers is the inability to accurately predict crop yields, which leads to suboptimal resource allocation, inefficient farming practices, and increased risks of crop failure. Traditional yield prediction methods are often based on historical data, but they fail to account for the complex, dynamic nature of environmental factors, such as changing weather patterns, soil conditions, and pest infestations. This uncertainty makes it difficult for farmers to make informed decisions about crop management, irrigation, and the use of fertilizers and pesticides.

Despite advances in technology, there remains a significant gap in the use of predictive analytics within the agricultural sector. Existing models often lack the precision needed to handle the vast array of variables influencing crop growth, resulting in inaccurate yield predictions. The problem is further exacerbated by the increasing unpredictability of weather and environmental conditions. Therefore, there is a pressing need for an intelligent, data-driven system that can provide more accurate and real-time yield predictions, enabling farmers to make better-informed decisions, optimize resource use, and adapt to changing environmental factors. "Green Intelligence" aims to address this problem by leveraging machine learning technologies to improve crop yield forecasting, ultimately contributing to sustainable farming practices and enhanced food security.



### 1.3 Aim and Objective:

#### Aim

The aim of the "Green Intelligence" project is to develop a machine learning-based system that accurately predicts agricultural crop yields by analyzing historical, climatic, and soil data. This system seeks to empower farmers with data-driven insights to optimize crop management, improve resource allocation, and enhance sustainability, ultimately contributing to food security and efficient farming practices.

#### Objectives

- **Data Collection and Integration:** To gather and integrate diverse datasets, including historical climate data, soil conditions, crop performance, and geographical information, for comprehensive analysis.
- **Model Development:** To develop and train machine learning models, such as regression analysis, decision trees, and neural networks, to predict crop yields under various environmental and agricultural conditions.
- **Model Evaluation and Optimization:** To evaluate the performance of the predictive models and refine them for accuracy and efficiency, ensuring they can handle the complexities of diverse agricultural environments.
- **Real-Time Insights:** To design the system to provide real-time predictions, enabling farmers to make timely and informed decisions about irrigation, fertilizer usage, and crop management.
- **Sustainability Promotion:** To ensure the system contributes to sustainable farming practices by promoting the efficient use of resources, minimizing waste, and reducing the environmental impact of farming activities.
- **User-Friendly Interface:** To develop an intuitive interface that allows farmers to easily interact with the system and apply the predictions to their farming practices without requiring advanced technical knowledge.

## CHAPTER 2

### LITERATURE SURVEY

The integration of machine learning in agriculture has gained significant attention in recent years, with various studies exploring its potential to improve farming efficiency, sustainability, and productivity. Machine learning techniques, such as supervised learning, decision trees, and deep learning, have been widely applied to predict crop yields, monitor plant health, and optimize resource management.

Several studies have demonstrated the effectiveness of machine learning models for yield prediction. For instance, a study by Chlingaryan et al. (2018) reviewed various machine learning applications in agriculture, highlighting the success of regression-based models for predicting crop yields based on environmental variables. These models have proven particularly useful in understanding the complex relationships between soil properties, weather patterns, and crop productivity. Similarly, Liu et al. (2020) explored the use of neural networks to predict crop yield in response to changing climate conditions, showcasing promising results in terms of both accuracy and adaptability to diverse geographic locations.

Other research, such as Kamilaris and Prenafeta-Boldú (2018), emphasized the use of big data and remote sensing in precision agriculture, where machine learning models analyze data from satellites, drones, and IoT sensors to enhance yield predictions. These approaches offer the advantage of real-time data collection, providing more timely and accurate insights for farmers. Additionally, Pérez et al. (2021) demonstrated the potential of deep learning techniques, particularly convolutional neural networks (CNNs), in analyzing complex agricultural datasets, further pushing the boundaries of prediction accuracy.

While these studies show significant promise, challenges remain, particularly in the scalability and robustness of models across different farming environments. Many existing systems rely on centralized data, which can limit their applicability in regions with limited access to modern technology. Moreover, most models fail to account for the dynamic nature of farming, where real-time decisions are critical, and external factors such as pests and sudden weather changes play a crucial role.

"Green Intelligence" builds upon this body of work by integrating diverse datasets—including climate, soil health, and crop-specific data—into a unified system that can provide actionable insights in real time. This approach aims to overcome the limitations of existing models by delivering more accurate, dynamic, and scalable solutions for crop yield prediction.

## CHAPTER 3

# SYSTEM REQUIREMENTS

- **Hardware Requirements:**

- 1) Processor: Intel Core i7 or AMD Ryzen 5 (or higher) for efficient computation
- 2) RAM: Minimum 16 GB (recommended 32 GB for handling large datasets)
- 3) Storage: SSD with at least 512 GB of storage (1 TB recommended for datasets and model files)
- 4) Graphics Card: NVIDIA GPU with CUDA support (e.g., NVIDIA GTX 1660 or higher) for accelerating model training
- 5) Network: Stable internet connection for accessing cloud resources and datasets
- 6) Additional Hardware (Optional): IoT devices or sensors for real-time data collection in agricultural environments

- **Software Requirements:**

- 1) Operating System: Windows 10/11, macOS, or Linux (Ubuntu 20.04 or later)
- 2) Programming Language: Python 3.8 or higher
- 3) Libraries and Frameworks:
  - 4) TensorFlow or PyTorch (for machine learning model development)
  - 5) Scikit-learn (for regression and decision tree models)
  - 6) Pandas and NumPy (for data processing and analysis)
  - 7) Matplotlib and Seaborn (for data visualization)
- 8) Database: MySQL or PostgreSQL for data storage
- 9) Development Tools:
  - 10) Jupyter Notebook or VS Code (for coding and experimentation)
  - 11) Git (for version control)
- 12) Cloud Platforms (Optional): AWS or Google Cloud for scalability and remote computation
- 13) Visualization Tools: Tableau or Power BI for advanced data visualization (optional)

## CHAPTER 4

# SYSTEM DESIGN

### 4.1 System Architecture:

The system architecture of "Green Intelligence" is designed to integrate multiple components seamlessly, ensuring efficient data processing, model training, and user interaction. The architecture is divided into the following layers:

#### 1. Data Collection Layer

- Input Sources:
- Historical climate data (temperature, rainfall, humidity) from meteorological databases.
- Soil health data (pH, moisture, nutrient levels) from IoT sensors or agricultural databases.
- Crop yield data (historical production records).
- Tools: APIs, IoT devices, and CSV/Excel data files.
- Functionality: Aggregates raw data from various sources for preprocessing.

#### 2. Data Preprocessing Layer

- Data cleaning (handling missing values and outliers).
- Feature engineering (normalization, one-hot encoding for categorical data).
- Data splitting (training, validation, and testing sets).
- Tools: Python libraries (Pandas, NumPy, Scikit-learn).
- Functionality: Ensures data consistency and prepares it for modeling.

#### 3. Machine Learning Layer

- Model selection (e.g., Decision Trees, Regression Models, Neural Networks).
- Training and validation of models on preprocessed data.
- Hyperparameter tuning for optimization.
- Tools: TensorFlow, PyTorch, Scikit-learn.
- Functionality: Develops a predictive model capable of forecasting crop yields with high accuracy.

#### 4. Data Storage Layer

- Centralized relational database for storing processed data and model outputs.
- Data backup and retrieval systems.
- Tools: MySQL, PostgreSQL, or cloud-based storage solutions (AWS S3).
- Functionality: Ensures secure and efficient storage and retrieval of data.

#### 5. User Interaction Layer

- Web-based dashboard or mobile app for presenting insights.
- Visualization tools for displaying yield predictions, resource suggestions, and environmental analytics.
- Tools: Flask/Django for backend, React.js/Flutter for front-end, Tableau/Plotly for visualizations.
- Functionality: Provides an intuitive interface for farmers and stakeholders to access predictions and insights.

#### 6. Real-Time Processing Layer (Optional)

- Integration of IoT devices for real-time data updates (e.g., soil sensors, weather stations).
- Continuous monitoring and prediction updates based on live data streams.
- Tools: MQTT protocol, IoT platforms like Arduino or Raspberry Pi.
- Functionality: Supports adaptive decision-making based on live agricultural conditions.

4.2 Flow Diagram:

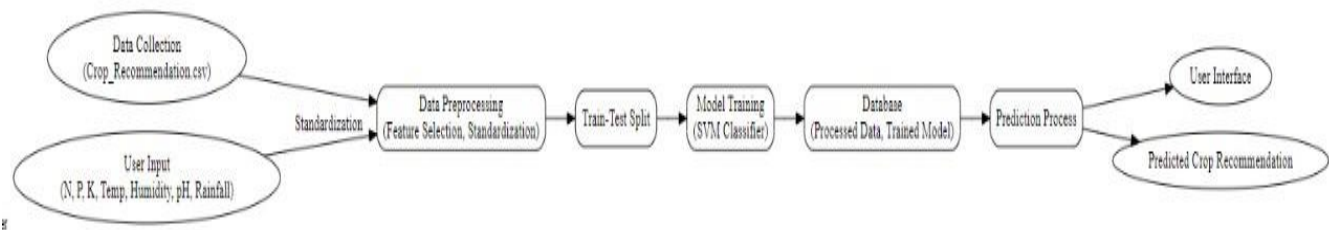


Fig1: Flow Diagram

## CHAPTER 5

# IMPLEMENTATION

```
pip install gradio

import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, classification_report

import gradio as gr

file_path = '/mnt/data/Crop_Recommendation.csv'
data = pd.read_csv("Crop_Recommendation.csv")

X = data.iloc[:, :-1]
y = data.iloc[:, -1]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Standardize the features
scaler = StandardScaler()

X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

svm_classifier = SVC(kernel='rbf', random_state=42)

svm_classifier.fit(X_train, y_train)

def recommend_crop(nitrogen, phosphorus, potassium, temperature, humidity, ph, rainfall):
    try:
        # Combine inputs into a list
        user_input = [[nitrogen, phosphorus, potassium, temperature, humidity, ph, rainfall]]
```

```
# Standardize the input

user_input_scaled = scaler.transform(user_input)


# Predict the crop

predicted_crop = svm_classifier.predict(user_input_scaled)

return f"Recommended Crop: {predicted_crop[0]}"

except Exception as e:

    return f"Error: {e}"

# Create Gradio Interface

inputs = [

    gr.Number(label="Nitrogen (N)"),

    gr.Number(label="Phosphorus (P)"),

    gr.Number(label="Potassium (K)"),

    gr.Number(label="Temperature (°C)"),

    gr.Number(label="Humidity (%)"),

    gr.Number(label="pH Level"),

    gr.Number(label="Rainfall (mm)"]

outputs = gr.Textbox()

# Launch Gradio interface

interface = gr.Interface(

    fn=recommend_crop,

    inputs=inputs,

    outputs=outputs,

    title="Crop Recommendation System",

    description="Enter soil contents to get a recommended crop.

interface.launch(share=True)
```



CHAPTER 6

RESULTS

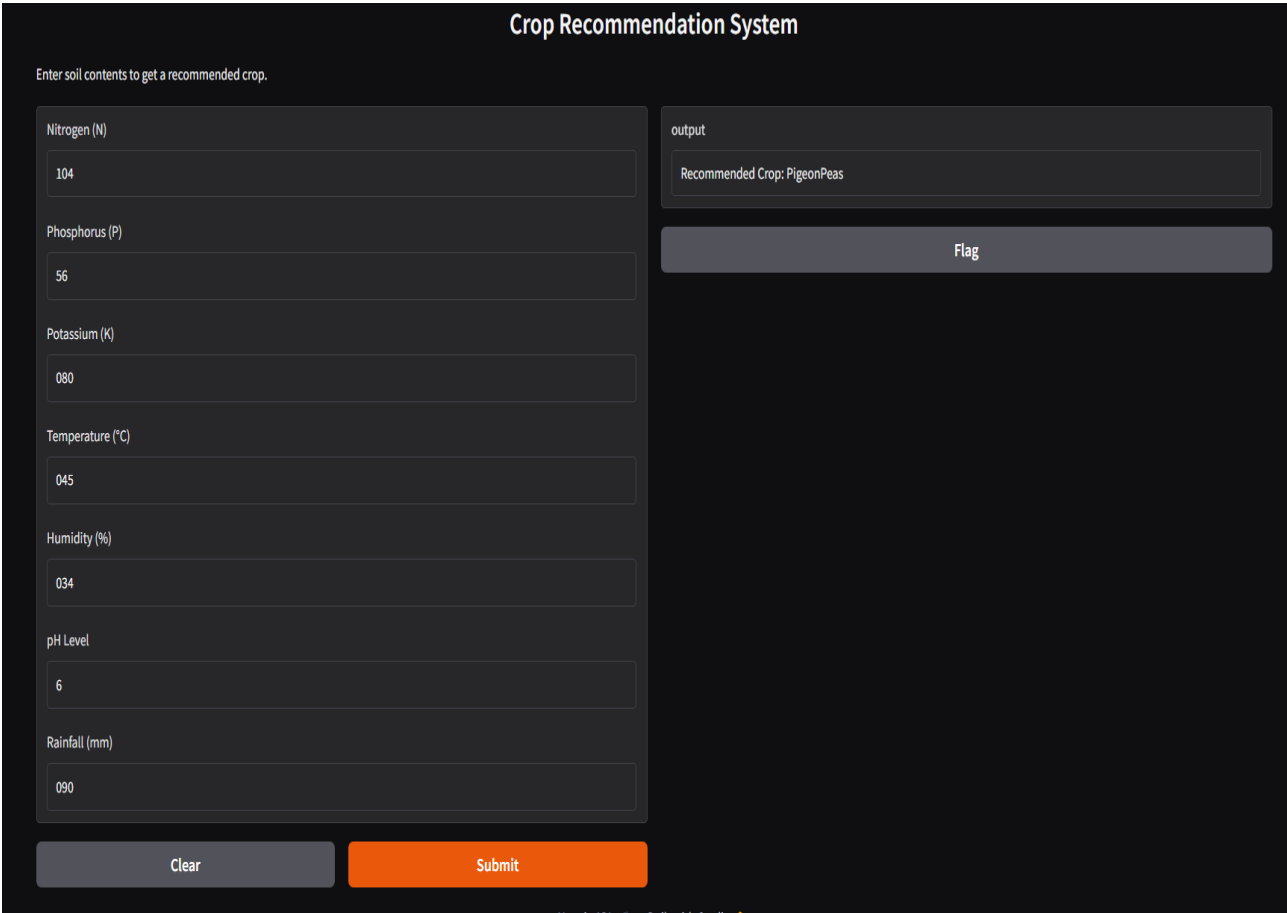


fig2: Results

## CHAPTER 7

### CONCLUSION

This project uses a **Support Vector Machine (SVM)** to predict suitable crops based on soil and environmental parameters like nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall. The data is preprocessed using **StandardScaler**, and the model is trained on a split dataset (80% training, 20% testing). A user-friendly interface is created using **Gradio**, allowing users to input values and receive crop recommendations interactively. The system ensures scalability and ease of use, aiding agricultural decision-making efficiently.

## CHAPTER 8

### FUTURE SCOPE

#### Future Scope of the Crop Recommendation System

**1. Integration with Real-Time Data:**

- Incorporate real-time weather and soil data from IoT sensors or APIs to improve prediction accuracy and usability.

**2. Incorporation of Advanced Models:**

- Experiment with advanced machine learning models like Random Forest, XGBoost, or deep learning for enhanced performance.

**3. Mobile Application Development:**

- Develop a mobile app for wider accessibility, especially for farmers in remote areas.

**4. Sustainability Recommendations:**

- Suggest sustainable farming practices and companion crops to improve soil health and yield.

**5. Integration with Agricultural Platforms:**

- Link the system with platforms that provide fertilizer recommendations, pest control tips, or irrigation schedules.

**6. Predictive Analysis:**

- Incorporate predictive analysis to foresee future farming trends based on climate change and soil degradation patterns.

## REFERENCES

- **Gradio Documentation:** <https://gradio.app/>
- **Kaggle Datasets:** <https://www.kaggle.com/>
- **ChatGPT**