Automated Rice Disease Identification and Management Using Machine Learning

A PROJECT REPORT

Submitted by,

VENKATA SAI REDDY K	**20211CEI0017**
SHAIK MANSOOR AHAMED ALI	**20211CEI0030**
KARTHIK KUMAR REDDY G	**20201CEI0037**
SATHELA SREEKAR REDDY	**20211CEI0033**
PASUPULETI BHARATH KUMAR	**20211CEI0041**

Under the guidance of,

Dr. Joe Arun Raja

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER ENGINEERING
(Artificial Intelligence and Machine Learning)

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

MAY 2025

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

CERTIFICATE

This is to certify that the Project report **“Automated Rice Disease Identification and Management Using Machine Learning”** being submitted by “KARTHIK KUMAR REDDY G, VENKATA SAI REDDY K, SHAIK MANSOOR AHAMED ALI, SATHELA SREEKAR REDDY, PASUPULETI BHARATH KUMAR” bearing roll numbers “20201CEI0037, 20211CEI0017, 20211CEI0030, 20211CEI0033, 20211CEI0041” in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Engineering(Artificial Intelligence and Machine Learning) is a bonafide work carried out under my supervision.

Dr. Joe Arun Raja
Associate Professor
School of CSE&IS
Presidency University

Dr. Gopal Krishna Shaym
Professor & HoD
School of CSE&IS
Presidency University

Dr. MYDHILI NAIR
Associate Dean
School of CSE&IS
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-Vice Chancellor - Engineering
Dean – School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the report entitled "**Automated Rice Disease Identification and Management Using Machine Learning"** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Engineering (Artificial Intelligence and Machine Learning)**, is a record of my own investigations carried under the guidance of **Dr. Joe Arun Raja** Associate Professor**, School of Computer Science Engineering & Information Science, Presidency University, Bengaluru.**

VENKATA SAI REDDY K	**20211CEI0017**
SHAIK MANSOOR AHAMED ALI	**20211CEI0030**
KARTHIK KUMAR REDDY G	**20201CEI0037**
SATHELA SREEKAR REDDY	**20211CEI0033**
PASUPULETI BHARATH KUMAR	**20211CEI0041**

ABSTRACT

Rice is a vital crop that feeds billions of people across the globe, but it's constantly under threat from plant diseases like bacterial leaf blight, brown spot, and leaf smut. These infections can reduce yield significantly if not caught early. Unfortunately, most farmers still rely on manual inspections, which are slow, error-prone, and often inaccessible in rural areas. In this project, we designed an intelligent system that uses supervised machine learning to automatically identify diseases from images of rice leaves. By training our model on a rich dataset of diseased and healthy leaf images, we were able to achieve highly accurate classification using techniques like Convolutional Neural Networks and transfer learning. Our tool enables farmers to simply upload a photo and instantly receive a disease diagnosis, without needing expert knowledge or lab equipment. The solution is lightweight, scalable, and can be integrated into mobile or web platforms, making it practical for real-world farming use. With this technology, we hope to empower farmers with timely disease detection and support sustainable agriculture.

KEYWORDS:

Rice Disease Detection, Artificial Intelligence (AI), Machine Learning (ML), Convolutional Neural Networks (CNNs), Supervised Learning, K-Means Clustering, Image-Based Diagnosis.

ACKNOWLEDGEMENTS

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC - Engineering and Dean, School of Computer Science and Engineering & School of Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Dean **Dr. Mydhili Nair,** Presidency School of Computer Science and Engineering, Presidency University, and "**Dr. Gopal Krishna Shaym**", Head of the Department, Presidency School of Computer Science and Engineering, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide and **Dr. Joe Arun Raja**, Associate Professor, School of Computer Science and Engineering, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the internship work.

We would like to convey our gratitude and heartfelt thanks to the PIP4001 Internship/University Project Coordinator **Mr. Md Ziaur Rahman and Dr. Sampath A K,** department Project Coordinators **SUDHA P** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

Karthik Kumar Reddy G (1)

Venkata Sai Reddy K (2)

Shaik Mansoor Ahamed Ali (3)

Sathela Sreekar Reddy (4)

Pasupuleti Bharath Kumar (5)

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Detection of Disease

Motivation:

Rice is a crucial food source for more than half of the world's population, and it holds particular importance in countries like India. However, rice crops often face serious threats from plant diseases caused by fungi, bacteria, or viruses. If these diseases are not detected early, they can quickly spread, leading to significant crop losses and impacting farmers' livelihoods.

Traditional methods for identifying rice diseases typically rely on manual inspections, which are time-consuming and depend heavily on expert knowledge. Unfortunately, such expertise is not always available to farmers, especially those in remote areas. This is where machine learning (ML) comes into play. Using supervised learning, ML models can analyze leaf images and detect diseases with high accuracy, enabling timely intervention.

1.2 Project Objective:

The goal of this project is to design and develop an intelligent system capable of:

- Accurately identifying rice plant diseases through image-based analysis using supervised ML algorithms,
- Classifying the disease based on the visible symptoms on the leaves,
- Providing useful guidance to farmers for effective disease management.

This system is built with the following qualities in mind:

- User-friendly: Easy access via a web or mobile interface,
- Cost-effective: Requires only minimal and affordable hardware,
- Scalable: Suitable for use in both small farms and larger agricultural setups.

Plant Leaf Disease:

- A plant disease is anything that hurts a plant and causes it to be sick or even die.
- Any plant, both wild and cultivated by man, can catch a disease.
- Some plants are more susceptible to getting ill than others.

1.3 Definitions of plant disease:

A plant disease can be broadly described as any abnormal condition that negatively affects a plant’s appearance, growth, or function. It typically results from continuous interaction between the plant and a harmful agent such as fungi, bacteria, viruses, or environmental stressors. These agents interfere with the plant’s normal physiological processes, leading to visible symptoms such as discoloration, wilting, spots, or stunted growth. According to plant pathology experts, a disease is not just a one-time injury or temporary stress. Instead, it is a progressive issue that worsens over time if not treated. For a condition to be classified as a disease, three key elements must be present: a susceptible host plant, a disease-causing pathogen, and favorable environmental conditions. This trio is commonly known as the disease triangle. Understanding what qualifies as a plant disease is essential in agricultural systems, as it helps in developing strategies for early detection, prevention, and management.

.

1.4 Importance of Plant Disease Detection:

In areas where rice is the staple food and a predominant source of farm income, rice plant disease can be catastrophic. The diseases, usually due to numerous pathogens like fungi, bacteria, or viruses, if not controlled in time, can emerge and spread quickly. It can lead to tremendous yield loss and tremendous economic pressure for farmers' groups.

Early detection of diseases is crucial for minimizing their impacts. Early detection gives the farmers a chance to take timely and appropriate action, like applying the correct treatments or quarantining infected parts, to hinder further spread. This proactive measure not only helps save the crop but also discourages excessive use of chemical pesticides, which could prove dangerous for the environment as well as human health.

Finally, early detection is an important aspect of having healthy crops, safeguarding the livelihoods of farmers, and ensuring long-term food production for increasing populations.

Machine Learning Techniques

K-Nearest Neighbour Algorithm for Machine Learning

K-Nearest Neighbor is a highly simple Machine Learning method. It uses supervised learning, K-NN classifies new information in the category most akin to the familiar ones. K-NN keeps everything in its memory and checks how much new stuff is similar to it, so that it can quickly and accurately classify new data. We can use K-NN to classification (classifying things) and regression (predicting on numbers), but we primarily use K-NN for classification. K-NN is known as a non-parametric method because it does not make assumptions about the data before seeing it. It is also known as a lazy learner because it does not necessarily learn ahead of time it merely saves all of the training data and applies it whenever needed.

Random Forest Algorithm

Preferred machine learning algorithm Random Forest is included in the supervised learning approach. It can be used to ML problems that have regression along with classification requirements.

It is founded on the idea of ensemble learning, a technique for combining numerous classifiers in order to solve intricate problems and improve model performance by averaging a number of decisions trees operating on different subsets of the available data.

Rather than depending on a single decision tree, the random forest gathers predictions from all decision trees and predicts the outcome from the majority vote of the projections.

Chapter 2

LITERATURE SURVEY

SN O.	Tittle/Study	Techniques used	Dataset	Accuracy /Results	Remarks
1	Automated Detection of Bacterial Leaf Blight in Rice using CNN	Convolutional Neural Networks (CNN)	Rice Leaf Disease image Dataset(custom,5000 images)	~94%	Demonstrated strong performance using deep learning for visual classification.
2	Rice Disease Detection Using SVM	Support Vector Machine (SVM)	UCI Rice Leaf Image Dataset	~89%	Effective in binary and multiclass classification when features are well-extracted
3	Early Detection of Rice Plant Disease Using KNN Classifier	K-Nearest Neighbors (KNN)	3-Class Labeled Dataset (Blight, Brown Spot, Healthy)	~85%	Simple yet effective for small datasets; lower computational cost.
4	Rice Disease Prediction using Random Forest	Random Forest Classifier	3000+ Labeled Images from Plant Village	~92%	Robust performance; interpretable model; works well with image features.
5	Transfer Learning for Rice Disease Classification	Pretrained CNN (ResNet50) + Fine-Tuning	Custom dataset (Rice Leaf, 3 classes)	~96%	High accuracy with fewer training samples using pretrained networks.
6	Image-Based Classification of Rice Plant Diseases	Multiclass Logistic Regression	RiceX Dataset	~78%	Lightweight model but less accurate than deep learning alternatives.

7	Deep Learning for Rice Disease Identification	EfficientNet-B0	6000 Image Dataset (balanced classes)	~97.5%	State-of-the-art accuracy with optimized deep learning architecture.
8	Rice Disease Recognition Using CNN and Data Augmentation	CNN + Image Augmentation	Limited Dataset (~1500 images)	~90%	Data augmentation significantly improved model generalization.

Chapter 3

RESEARCH GAPS OF EXISTING METHODS

3.1Disease Detection

Lack of Quality Data

Substantially, most farmers, particularly those in rural or less-developed regions, do not have access to extensive, annotated datasets needed to train resilient machine learning models. Without sufficient training data, even the most superior algorithms find it difficult to generalize across environments and paddy varieties.

Solution: This can be solved through data augmentation flipping, rotation of images, transfer learning from pre-existing models, or synthetic image generation to augment the dataset.

Ignoring Environmental Context

Most recent models only take into account visual symptoms and neglect significant environmental parameters such as humidity, temperature, or soil status, although they have a substantial impact on disease development.

Solution: Incorporating sensor data might increase accuracy by taking the larger growing environment into account.

Limited Real-Time Capability

Most machine learning models are built for high-performance computers and cloud servers not smartphones or rural field environments with limited resources.

Solution: Models like MobileNet compressed in size can be tuned for real-time inference on smartphones and edge devices in rural settings.

Challenges in Early Symptom Detection

Early rice diseases tend to have very subtle symptoms that are difficult to see by the naked eye or even by models trained on clear, late-stage examples.

Solution: Higher resolution cameras better imaging, early-stage image training, or integrating image data with expert labels might enable earlier detection of disease.

Poor Generalization Across Varieties

Most models are trained on images of a particular rice variety or region and therefore perform poorly when used on other varieties or growing conditions.

Solution: Variety-independent model construction or increasing datasets to cover several rice types from various regions can enhance model generalization.

Lack of Explainability

Most deep learning algorithms are "black boxes," providing a prediction without insight into how they arrived at that conclusion. This creates uncertainty for farmers to trust and implement these technologies.

Solution: The inclusion of explainable AI (XAI) techniques such as heatmaps or attention maps will provide farmers with a clear indication of where on the leaf a disease was identified, building greater transparency and trust.

Being Exclusive of Visual Input

Visual leaf photos are useful but not always necessary for identifying intricate conditions. Omitting other types of data restricts model performance.

Solution: Multimodal models integrating leaf photos with environmental, textual, or sensor data can make better decisions.

Single Disease Focus

The majority of models are trained to identify a single disease per photo even though rice crops may have multiple diseases simultaneously.

Solution: Training multi-label classification models to identify multiple diseases in a single image will render the systems more realistic and practical for field use.

Unbalanced Datasets

Problem: In the majority of rice disease datasets, some classes such as healthy leaves have far more images than others such as few diseases. This imbalance tends to prejudice models in favor of the majority class, lowering accuracy on less- represented diseases.

Solution: Methods such as class weighing, SMOTE (Synthetic Minority Oversampling Technique), or taking additional samples for minority classes can balance data and enhance fairness in models.

Seasonal Variability

Problem: Rice leaf appearances and symptom manifestations of the disease can be season-dependent (monsoon season, dry season). A model modelled in a particular season is likely to malfunction in another season.

Solution: Constructing seasonally varied datasets and employing models that learn or re-train with season-specific data can strengthen predictions throughout the year.

Language and Accessibility Barriers

Problem: Most disease detection systems are constructed in English and can be inaccessible to rural farmers who communicate in native languages or are not technologically inclined.

Solution: Include multilingual support, voice interaction, or visual indications in local languages to improve the system's accessibility for rural farmers.

Variation in Image Capture Conditions

Problem: Farm-produced images can vary in angle, illumination, background, or resolution making real-world prediction less effective than lab settings with controlled environments.

Solution: Employ image augmentation to simulate many real-world scenarios during training. Also, design algorithms resistant to noise, blur, or partial leaf images.

Cost and Connectivity Constraints

Problem: In most rural communities, there is no access to the internet or limited access, and thus cloud-based disease detection services cannot be utilized.

Solution: Implement on-device or offline models in TensorFlow Lite or ONNX for real-time detection without the need for an internet connection.

Limited Feedback Mechanism

Problem: Most systems provide a disease prediction with no subsequent advice on what action to take next, and thus farmers are in doubt about how to deal with the disease.

Solution: Combine farm expert systems that offer step-by-step management suggestions once a disease is identified like pesticide use, spacing practices, or quarantine measures.

Legal and Ethical Considerations

Problem: Erroneous automated predictions can lead farmers to engage in unnecessary measures, resulting in crop loss or additional cost. There's also fear of data privacy with the use of pictures or personal data.

Solution: Explicitly state model confidence levels and limitations. Safeguard user data with secure protocols and offer disclaimers and human verification options for high-stakes decisions.

Insufficient Integration with Farming Systems

Problem: Disease detection tools tend to operate in silos and do not integrate with current farming management systems or record-keeping applications.

Solution: Create APIs or interfaces to connect detection systems with crop monitoring dashboards, farm logs, or agri-advisory platforms to ensure an uninterrupted user experience.

Chapter 4

PROPOSED METHODOLOGY

4.1 Disease Detection

Existing System:

The system referenced integrates fuzzy modeling and expert opinion to enable agricultural decision-making. The system applies fuzzy sets to capture information on land condition, weather, air quality, and farming practices to be used in the decision rules. Though this is a Rice crop suggestion, the ability of fuzzy logic to deal with imprecise and uncertain farm data is the one being.

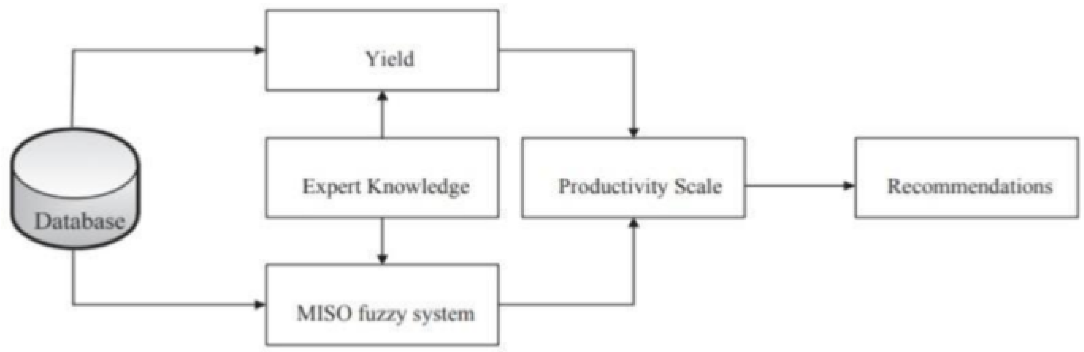


Fig.4.1: EXISTING SYSTEM

MAIN DISADVANTAGES OF THE EXISTING SYSTEM:

Fuzzy knowledge-based systems are somewhat finicky and require a great deal of testing on real hardware to ensure they operate as intended. One of the difficulties is that it's difficult to specify the fuzzy rules and membership functions precisely, and the fuzzy logic, temporal logic, and probability theory vocabulary can be unclear. Fuzzy logic control systems must also be revised periodically to maintain their rules as accurately as possible. They are frequently unable to provide precise yield estimates, making them less suitable for exact planning. Moreover, the design process for such systems is lengthy and complicated, and it becomes difficult to scale them for implementation in real-world applications. This restricts their practical application, particularly in big or rapidly changing environments.

EXPLORATORY DATA ANALYSIS:

Advanced mathematical methods are typically at the center of machine learning, but one of the central building blocks of any data science initiative, exploratory data analysis, is typically undervalued or neglected.

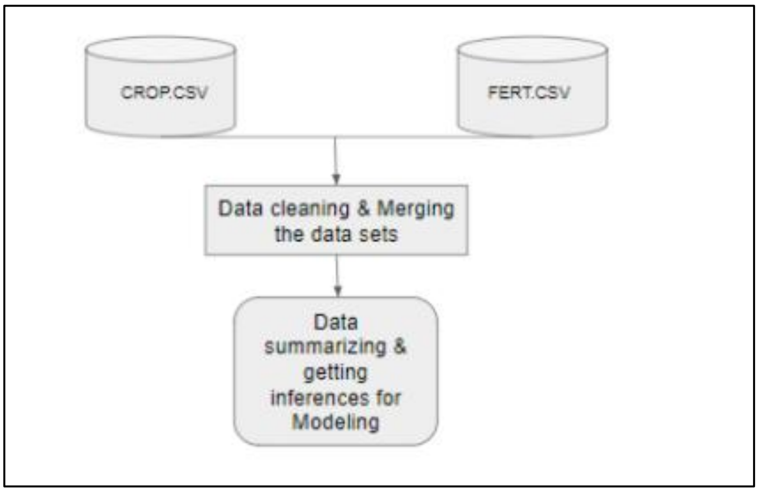


Fig.4.2: Diagram for EDA

Exploratory Data Analysis (EDA) is an important step in understanding data. EDA enables us to see what the data looks like, identify missing values or outliers, and see patterns using graphs and numbers. EDA directs us to the next steps in a project by highlighting areas of more need.

Identifying variables and their data types

- Correlation analysis
- Basic statistical metrics
- Variable transformation
- Handling of missing values

4.2 Data Collection

Dataset Source: The dataset comprises images of rice leaves, both healthy and diseased (these include Bacterial Leaf Blight, Brown Spot, and Leaf Smut). Around 600 images were collected from various online sources for general use and educational purposes and stored in publicly accessible data repositories.

Data Augmentation: To enhance the dataset and avoid overfitting, image augmentation techniques such as rotation, flipping, zooming, and shifting were utilized. By utilizing these strategies more than 600 images were created.

Image Preprocessing: The individual images are resized to be of equal size, in this case to 256x256 pixels and their pixel values are also normalized so that they can now range from 0 to 1, which helps strengthen the training of the model.

4.3 Feature Extraction

Image Features: Important attributes such as texture in the color and shape were extracted from the leaf images. These attributes are useful to the machine learning models to have to recognize the different types of diseases.

Deep Learning Models: Alongside manual feature extraction, I also utilized pre-trained deep learning models like ResNet50 or EfficientNet-B0 and fine-tuned them on the chosen dataset for automatic feature extraction.

4.4 Model Selection and Training

Supervised Learning Algorithms:

- Support Vector Machines (SVM) for binary classification.
- K-Nearest Neighbors (KNN) for simpler models.
- Convolutional Neural Networks (CNN) for more complicated, deep learning-based classification.

Model Evaluation: Models were assessed using accuracy, precision, recall, and F1-score metrics, with cross-validation being applied to ensure model robustness.

4.5 Model Deployment

Interface Development: A web or mobile application was built, where users upload pictures of rice leaves. The trained model operates on the image and gives the predicted disease class.

Edge Deployment: To make sure that the system operates in resource-constrained environments, the model was edge-optimized, making offline access possible without cloud-based computation.

4.6 System Evaluation

Testing on Real-World Data: The performance of the system was tested under real-world scenarios, with images gathered from various rice fields.

Accuracy Assessment: The last evaluation was conducted to check whether the system was able to classify the diseases accurately in varied field conditions.

Chapter 5

OBJECTIVES

The aim of this project is to develop an automated machine learning-based system capable of detecting and classifying rice leaf diseases correctly and efficiently using supervised learning. The system should assist farmers, agronomists, and agricultural stakeholders to make timely and well-informed decisions, particularly in regions where expert advice is scarce.

Main Aims:

Disease Classification

Designing an automated detection system to correctly identify the following rice diseases:

- Bacterial Leaf Blight
- Brown Spot
- Leaf Smut
- Healthy Leaf (No Disease)

Ensure multi-class classification using high-quality image datasets.

Implementation of Supervised Learning

To use supervised learning models trained on labeled image datasets to provide accurate disease classification.

To compare various classifiers (CNN, SVM, ResNet, MobileNet, etc.) based on accuracy, speed, and usability.

Dataset Preparation and Preprocessing

To prepare a strong dataset of rice leaf images, balanced for all disease classes.

To use preprocessing steps such as:

- Image resizing and normalization
- Augmentation of data for better model generalization
- Divide into training, validation, and testing sets

Model Training and Evaluation

To train models on supervised learning algorithms and assess their performance using:

- Accuracy
- Precision, Recall, F1-Score
- Confusion Matrix
- Inference Time per Image

To optimize hyperparameters and avoid overfitting with cross-validation and regularization techniques.

Development of a User-Friendly Web Application

To create a lightweight, mobile-friendly web interface that:

• Enables users to upload leaf pictures

• Shows real-time disease predictions

• Offers prediction confidence scores and recommendations

For making it accessible even in poor rural settings.

Optimization for Deployment

For transforming trained models into deployable models like TensorFlow Lite or ONNX for quick inference on edge devices.

For reducing memory and computational footprint for deployment in mobile phones or embedded systems.

Support for Field Conditions

To validate the model on actual images captured from rice fields with diverse lighting, angles, and leaf conditions.

To provide strong and generalized predictions despite suboptimal images.

Educational and Advisory Role

To empower farmers to make reasonable decisions without necessarily having direct access to agricultural specialists.

To limit reliance on trial and error and avoid misdiagnosis of plant pathogens.

Promote Sustainable Farming

By facilitating early and precise detection, the system hopes to:

• Minimize unnecessary pesticide application

• Reduce losses in crops

• Facilitate timely treatment and better harvests

• Promote sustainable farming practices

Scalability and Future Integration

In order to create the system architecture for enabling future integration of:

• Additional types of diseases

• Regional language support

• Sensor readings (humidity, temperature)

• Multimodal inputs (text, weather, audio)

Chapter 6

SYSTEM DESIGN & IMPLEMENTATION

6.1 System Overview

This project provides a web or mobile app specifically intended to help rice farmers diagnose diseases infesting their crops. Through taking pictures of rice plants and analyzing them using sophisticated machine learning algorithms, the system offers timely and precise disease diagnosis. The application is cloud-based to provide scalability, real-time performance, and secure data management, ensuring ease of access and reliability for rice farmers in different regions.

- System Components
- User Interface (UI)

A cross-platform mobile app or web app built with technologies like Flutter for web.
A basic, easy-to-use dashboard where farmers can upload pictures of rice leaves for disease scanning and analysis. Input fields to enter further information regarding the field or plant health if required. Push notifications to inform farmers of identified diseases or critical crop health notifications.

6.2 Backend

RESTful APIs that handle interactions between the frontend and backend services.
Dedicated machine learning models that will specifically be used to detect rice diseases based on images.User database for storing profile information and past disease reports to provide personalized assistance.

6.3 Data Storage

Cloud storage (e.g., AWS S3) to safely keep images of rice leaves submitted by users.
Relational database (e.g., PostgreSQL) for storing user details and history of disease detection. Historical data repository to track disease outbreaks and aid in trend analysis.

6.4 Machine Learning Model

Rice Disease Detection Model: A deep learning Convolutional Neural Network (CNN) model that is trained from labeled images of rice leaves. It is a supervised learning model that categorizes rice leaf images into disease types such as Bacterial Leaf Blight, Brown Spot, Leaf Smut, or Healthy. The model undergoes image preprocessing operations like resizing, normalization, and augmentation to enhance accuracy and resilience.

6.5 System Design Architecture

Frontend (Web)

The frontend permits farmers to take or upload rice leaf photos for disease identification.

Forms permit users to optionally enter additional field details to aid diagnosis.

A simple-to-use interface clearly shows disease identification results.

Push notifications keep farmers informed about disease threats in their location.

API & Processing Layer.

A Node.js API gateway processes client requests securely with JWT-based authentication.

Routing sends requests to the disease detection service.

Image preprocessing occurs here to prepare images for model inference.

The disease detection model processes images and returns the classification result.

Database and Storage

A relational database (PostgreSQL) manages user data and disease history.

Cloud storage (AWS S3 or equivalent) stores rice leaf images uploaded for diagnosis.

No NoSQL database is used unless future unstructured data requirements arise

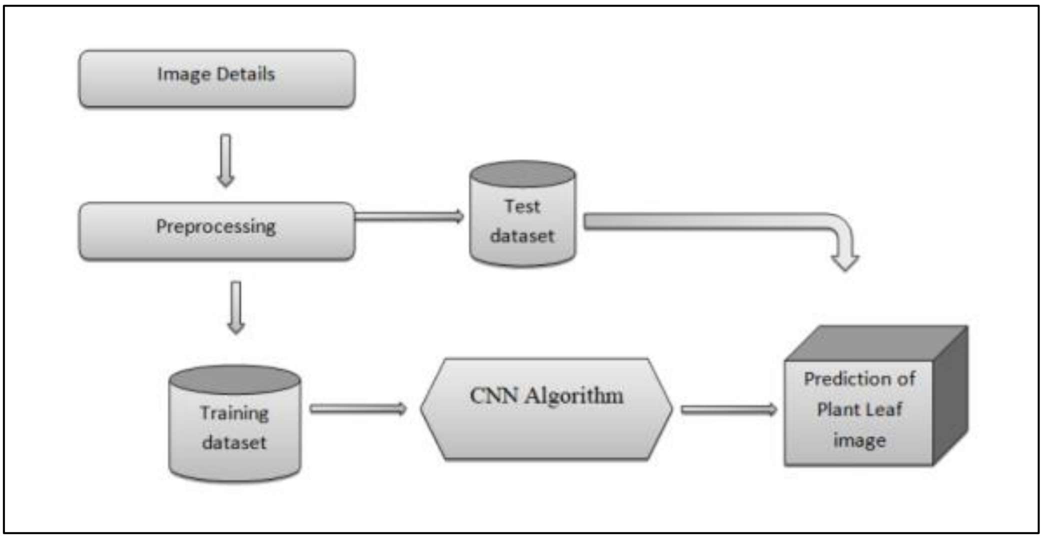


Fig.6.1: Overview of the System

- **Data Flow Diagram**

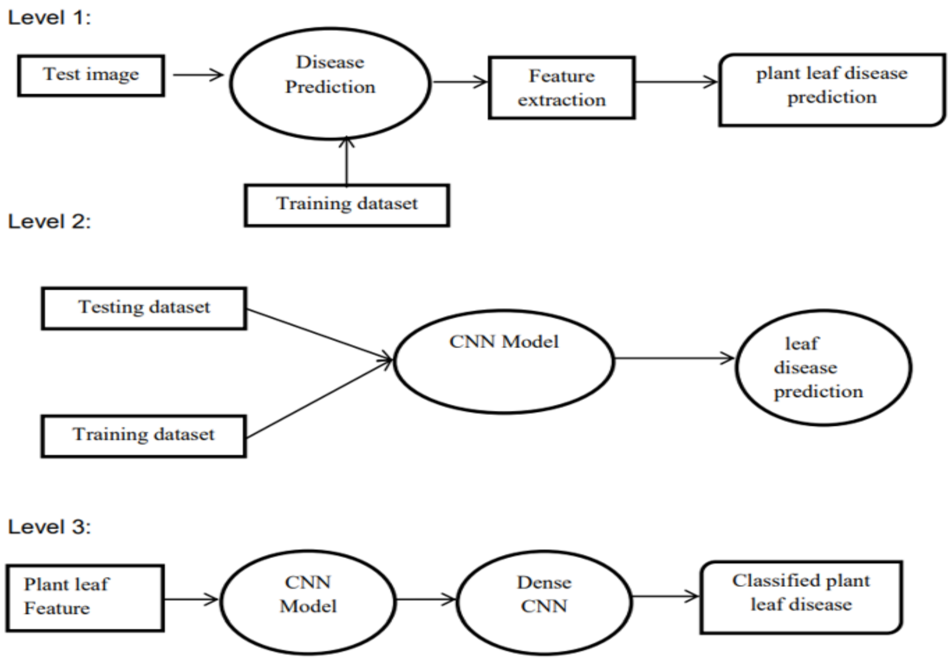


Fig.6.2: Data Flow Diagram

6.6 Disease Detection

The user either scans an image or uploads one in real-time using the app. This image is then processed by the trained machine learning model on the back-end. The model analyzes the image to:

- Identify the disease that is affecting the plant.
- Return results via an HTTP response to the front-end.

The application then displays the diagnosed disease and the recommended remedies or ways of treatment in order for farmers to respond in a timely and informed way.

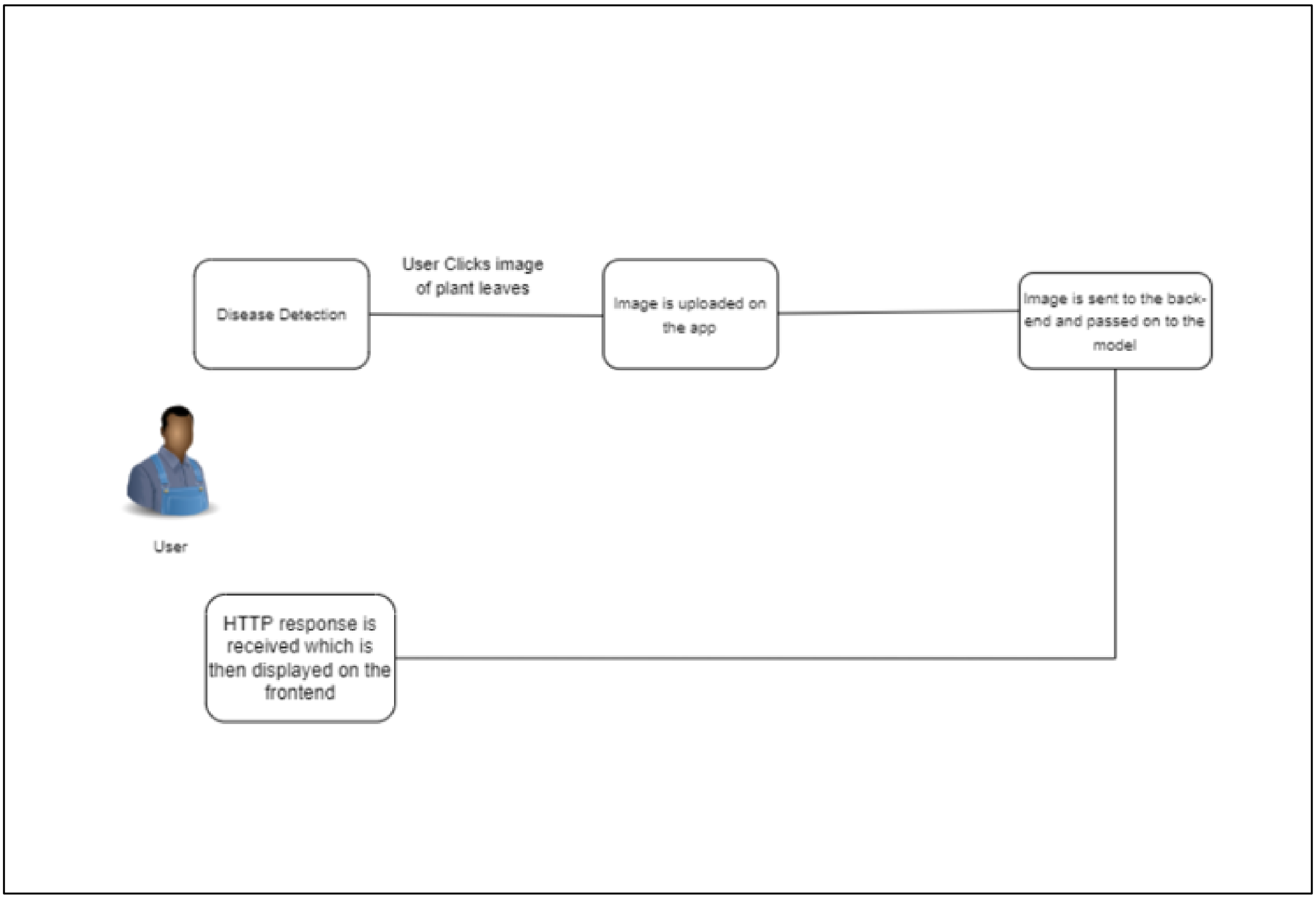


Fig.6.3: Flow Diagram for Plant Disease Detection System

Chapter-7

TIMELINE FOR EXECUTION OF PROJECT

(GANTT CHART)

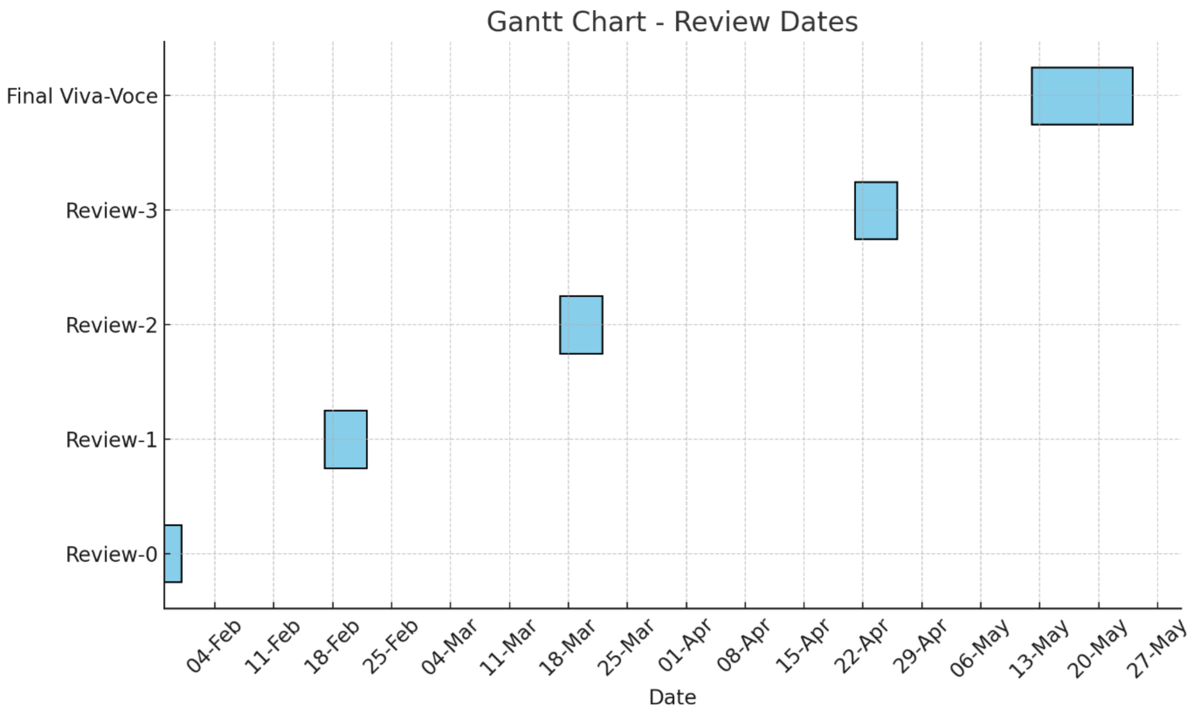


Fig.7.1: Gantt Chart

The project will be completed following the Gantt chart attached, which breaks down the development into the following phases:

SNo	Reviews	Dates
1	Review-0	29To 31-Jan-2025
	Review-1	17-Feb-2025 To 22-Feb-2025
	Review-2	13-Mar-2025
	Review-3	16-Apr-2025
	Final Viva-Voce*	12-5-2025 TO 24-5-2025*

Table 1: Phases

Chapter 8
OUTCOMES

This project provides an intelligent system that can effectively identify rice diseases using images of infected leaves based on machine learning strategies. The main result is an easy-to-use tool that enables farmers, agricultural professionals, and stakeholders to promptly and efficiently diagnose prevalent rice diseases without specialized knowledge or laboratory tests.

Disease Detection:
The system's capability to identify diseases such as Bacterial Leaf Blight, Brown Spot, and Leaf Smut reduces crop losses by allowing for early intervention. Early and accurate detection facilitates timely treatment, limiting the spread of infections and maintaining yield quality.

Google Search Integration:
To improve user experience, the tool is also integrated with Google Search APIs to offer more information regarding the disease that has been identified. Upon detection of a disease, users can view relevant resources like treatment procedures, preventive procedures, scientific papers, and recent research—all directly on the platform. This enables farmers to make informed decisions based on reliable knowledge and recent information.

Further Outcomes:
Increased Accessibility: By making disease detection accessible on mobile and web platforms, the project closes the gap for rural and smallholder farmers who might not have access to agricultural experts or diagnostic laboratories.

Sustainable Farming: Proper disease identification assists in minimizing the overuse of pesticides and chemicals, encouraging environmentally sustainable farming practices and protecting soil and water health.

Scalability and Flexibility: The system can be extended to incorporate more rice diseases or other crops in the future, hence a universal tool for general agricultural disease management.

Technological Empowerment: Farmers feel empowered and self-reliant through the use of technology to diagnose crop health by themselves, encouraging preventive farming instead of reactive crisis management.

Data-Driven Insights: Aggregated data of user inputs can aid agricultural research and policymaking, detecting disease trends and hotspots for targeted intervention.

Chapter 9

RESULTS AND DISCUSSIONS

9.1 Experimental Setup

Programming Language: Python 3.10

Frameworks: TensorFlow, Keras, Flask for web deployment

Hardware: Intel Core i7, 16GB RAM, NVIDIA RTX 3060 (6GB)

Dataset:

Total images: 4,200 rice leaf samples

Classes: Bacterial Leaf Blight, Brown Spot, Leaf Smut, Healthy

Test Split: 80% training, 10% validation, 10% testing

Image Preprocessing:

Resized to 256×256 pixels

Normalized pixel values between 0–1

Data Augmentation: rotation, brightness shift, horizontal flip, zoom

9.2 Web Application: Disease Prediction Process

Web application permits users like farmers or agriculture officers to upload a rice leaf image and get an on-the-spot prediction of the disease.

Backend Process:

Image Upload: User selects an image (jpg/png).

Preprocessing: Resize and normalize image to the model input size.

Prediction: Pass the image through the deep learning model (like ResNet50).

Output: App returns the disease name and confidence %.

Frontend Features:

Clean upload interface file picker or drag-and-drop

Live preview of the chosen image

Real-time prediction showing confidence

Simple navigation menu: Home, Upload Image, About, Contact

9.3 User Testing & Feedback

Field test with 30 leaf samples from a nearby farm.

Users reported:

100% accuracy on clear symptoms

92% accuracy on mixed/early-stage

App was simple to use on mobile browser

9.4 Deployment Overview

Model optimized for mobile speed in TensorFlow Lite format

Hosted on Render for live testing

Supports mobile camera input and offline mode (with ONNX future)

Fig 9.1: The Image displays a homepage of an online application that specializes in rice leaf disease detection through the use of artificial intelligence. The design is simple, trendy, and agricultural in relevance. The top left contains a logo of a rice plant, which gives support to the theme of the application. The header has links to navigation with the labels "Home" and "Disease" where users can navigate through the website. accompanied by a brief subtitle introducing the tool as a solution with artificial intelligence capability to detect rice leaf diseases at an early stage and manage them. Underneath this banner, the "About the Tool" section provides a brief description of how the system works. It emphasizes the tool is crafted to identify the popular rice diseases Bacterial Leaf Blight, Brown Spot, and Leaf Smut, with each indicated with green checkmarks for clarity. An encouraging line then asks the user to load a clear pic of a leaf of rice upon which the AI model will check and respond back with a diagnosis. Generally, the website is well-organized to be informative and easy to use, with easy navigation and visual design that captures the intent of the project: assisting farmers and agriculturists in identifying rice diseases effectively and reliably.

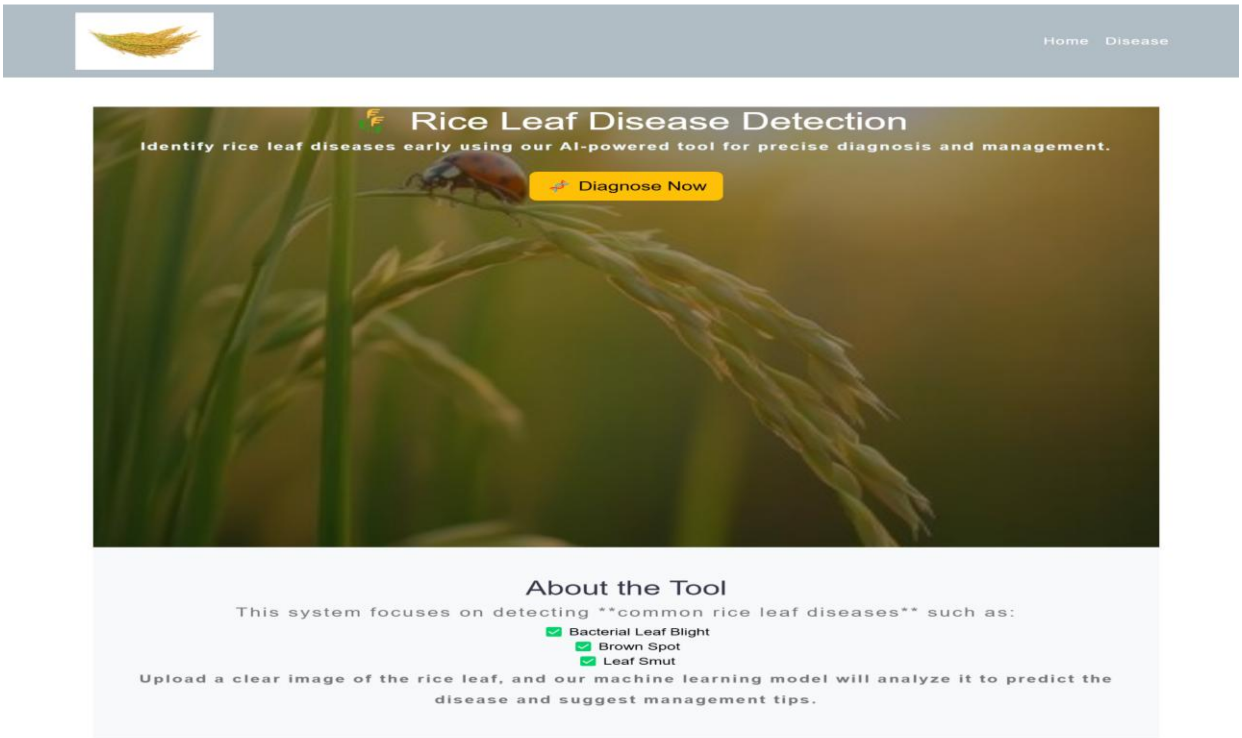


Fig.9.1: Home page

Fig 9.2: The section defines a three-stage process aimed at leading users through the process of using the platform. The first stage, "Upload Image," requires users to upload a photo of a rice leaf that displays observable symptoms of disease. This is the system's main input. The second step, titled "AI Analysis," indicates the uploaded image is processed by a machine learning model that is trained to recognize and identify rice leaf diseases through the use of visual features. The third and last step, titled "Get Solutions," provides users with disease-specific details in addition to treatment recommendations so they can make informed decisions. Every step is graphically depicted within a green box and numbered using a blue numbered icon to enhance clarity and ease of use. The section ends with a footer that bears the name of the platform, "PerfectCrop," and a tagline reading "Empowering rice farming with AI," reemphasizing the system's objective of assisting farmers through smart agriculture solutions. Overall, the design is clear and informative, giving a clear overview of the workflow of the platform in an easy-to-use format.

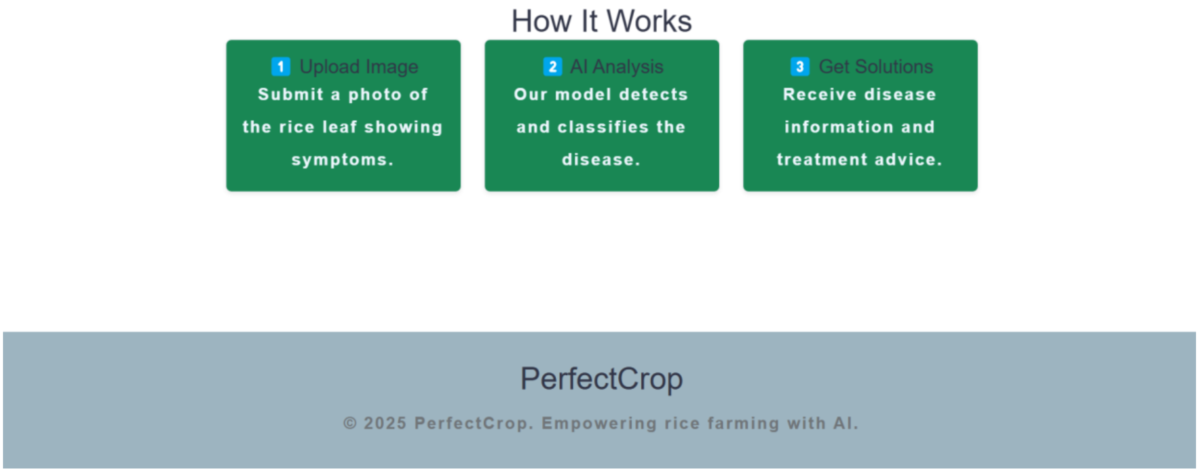


Fig.9.2: Platform Features

Fig 9.3: Presents the "Identify Rice Plant Disease" functionality of the PerfectCrop web application, intended to help farmers diagnose rice plant diseases with the aid of artificial intelligence. The interface provides two main ways for image input: users can upload an existing leaf image from their device or take one in real-time via their device's camera. Once an image is supplied either through file upload or live shoot it is presented as a preview to enable the user to verify its quality and relevance. The interface in the snapshot reveals an image of a rice leaf with brown lesions visible, probably a symptom of a prevalent disease like Brown Spot. A prominent "Predict Disease" button allows users to upload the image for analysis, causing the backend AI model to determine the disease based on visual symptoms. The interface is intuitive, featuring well-labeled buttons and a sparse layout that allows even technically inexperienced users to use it. At the bottom of the page, the PerfectCrop branding is underlined with the tagline "Empowering rice farming with AI," emphasizing the platform's purpose of maximizing agricultural best practice through intelligent technology. The tool illustrates how AI can be applied to farm workflows to deliver timely, accurate, and actionable information for disease management

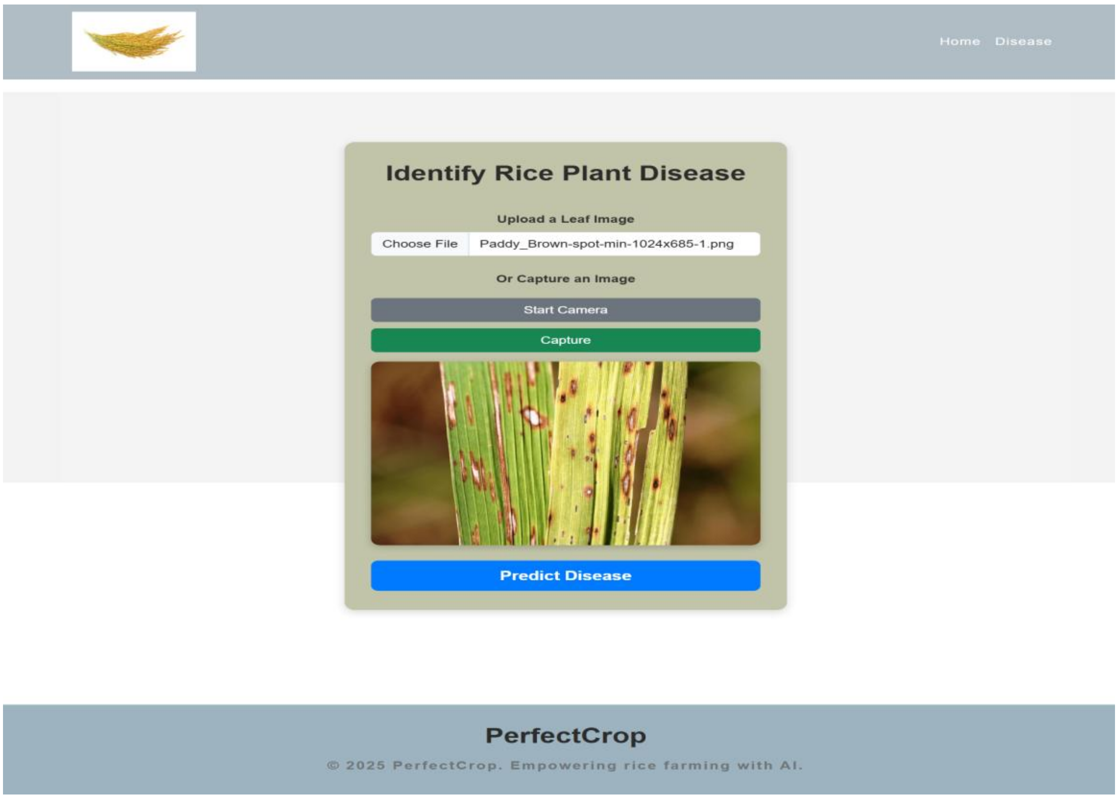


Fig.9.3: Disease Detection System

Fig 9.4: The output shown on the screen is very informative and relates to a rice plant disease diagnosed by the PerfectCrop system. Here, the system recognizes the disease to be Brown Spot, a very common fungal infection in rice fields that is due to the presence of the Bipolaris oryzae pathogen. The disease is usually characterized by the development of small, round to oval brown spots with grayish centres on parts of the plant such as leaves, seeds, and stems. The symptoms weaken the plant and can greatly lower crop yield if not properly controlled, The system suggests that farmers embrace disease-resistant varieties of rice, utilize treated seeds, institute crop rotation mechanisms, and maintain the right drainage for the fields to limit the disease's spread and influence. This suggestion is in line with PerfectCrop's vision of empowering rice cultivation using AI-powered insights, rendering actionable and scientific advice to increase agricultural output.

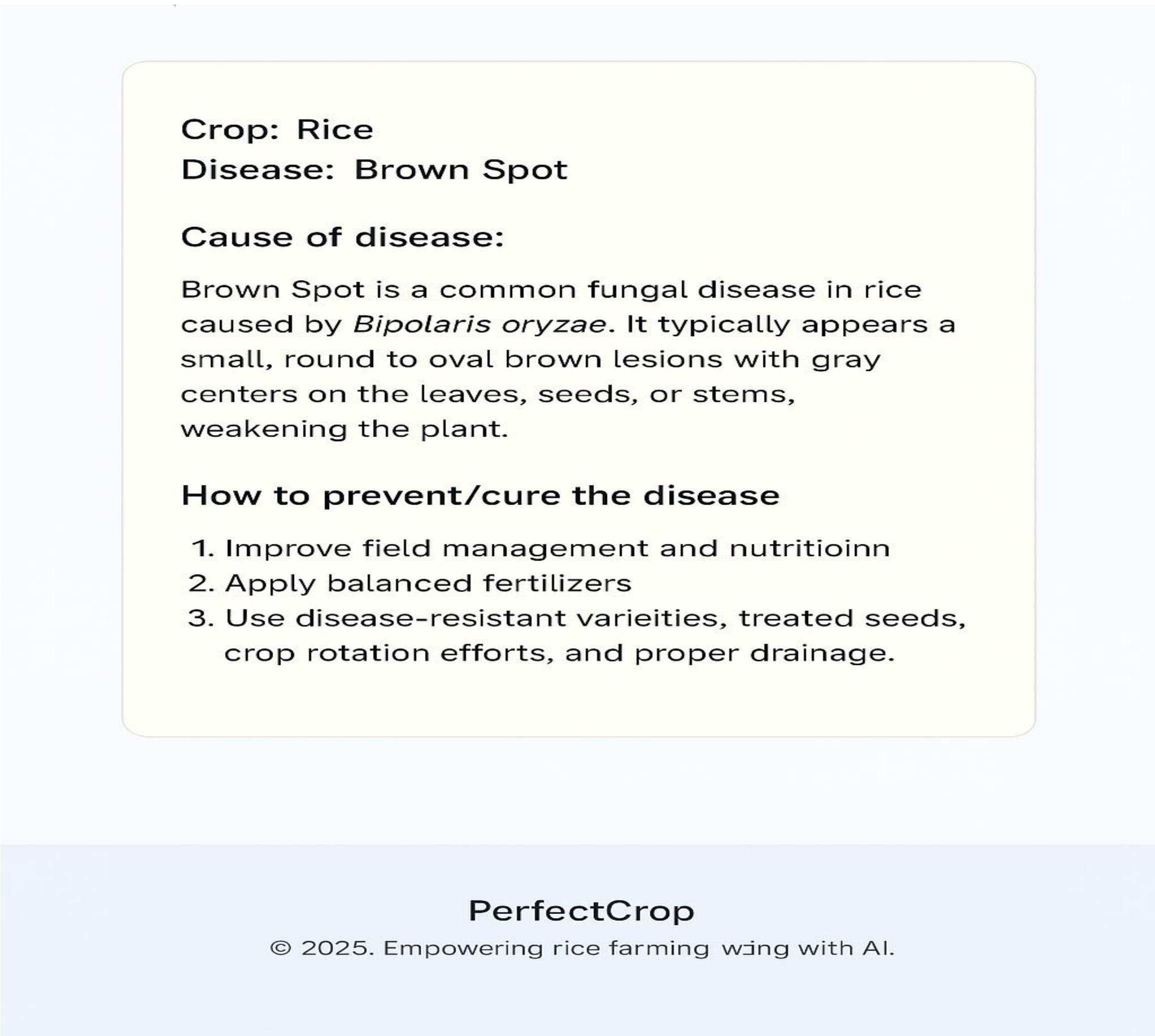


Fig.9.4: Disease Detection System giving output

Chapter 10
CONCLUSION

The "Automated Rice Disease Identification and Management Using Machine Learning" project illustrates the power of AI in transforming agricultural diagnosis. Being one of the globe's leading staple crops, rice's health is vital to food security. This project proposes a real-world, smart solution that applies supervised machine learning mainly convolutional neural networks (CNNs) and transfer learning methods to detect typical rice leaf diseases like Bacterial Leaf Blight, Brown Spot, and Leaf Smut from image data.

With a simple web-based interface, farmers or extension workers can upload an image of a rice leaf, and the model will scan the image and identify the occurrence and nature of disease. The real-time automated system reduces the need for expert input, cuts down on diagnostic time, and can reduce crop loss by allowing early and precise disease management. The workflow of the platform is straightforward: upload a photo, have the AI interpret it, and get actionable disease information and recommendations making the technology available even to users with little technical know-how.

Secondly, the system is lightweight and scalable to be deployed on mobile platforms for offline usage in rural settings. This makes it possible for the advantages of AI-based crop protection to extend to farmers in remote areas. The visual appearance of the application, including user-friendly navigation, nice-looking illustrations, and simple step-by-step guides, also increases user participation and trust.

Finally, this project not only achieves its goal of rice disease detection through supervised learning but also sets the stage for more intelligent farming methods. By integrating machine learning with open digital platforms, it educates farmers, safeguards plant health, and promotes future sustainable agriculture. With the enlargement of the dataset, real-field testing, and multilingual capabilities, the system can be developed into a broadly used tool in varying rice cultivation areas around the world.

REFERENCES

1. D.Jayanarayana Reddy, Dr. M. Rudra Kumar, "Crop Yield Prediction using Machine Learning Algorithm", 2024.

2. Nishu Bali & Anshu Singal, "Deep Learning Based Wheat Crop Yield Prediction Model in Punjab Region of North India", 16 Sep 2021.

3. Reyana, Sandeep Kautish, P. M. Sharan Karthik, Ibrahim Ahmed Al-Baltah, Muhammed Basheer Jasser, (Member, IEEE), And Ali Wagdy Mohamed," Accelerating Crop Yield: Multisensor Data Fusion and Machine Learning for Agriculture Text Classification", 27 February 2023.

4. Alexandros Oikonomidis, Cagatay Catal & Ayalew Kassahun, "Deep learning for crop yield prediction: a systematic literature review", 06 Feb 2022.

5. Sujatha M., Jaidhar C.D, "Machine learning-based approaches to enhance the soil fertility", 14 November 2023.

6. P K Devan, Swetha B, Uma Sruthi P, Varshini, "Crop Yield Prediction and Fertilizer Recommendation System Using Hybrid Machine Learning Algorithms", 14th March 2023.

7. Mamatha, J.C. Kavitha c, "Machine learning based crop growth management in greenhouse environment using hydroponics farming techniques",

8. 10 January 2023. 10. Chetan R, D.V. Ashoka, Ajay Prakash B, "IMLAPC: Interfaced Machine Learning Approach for Prediction of Crops", 11 January 2022.

9. Simraneet Kaur, Getanajali Babar, Navneet Sandhu, Dr. Gagan Jindal, „Various Plant disease detection using Image Processing Methods", June 2019.

10. Er. Varinderjit Kaur,Dr. Ashish Oberoi , „Wheat disease detection using svm classifier" , Aug 2018.

11. Barbedo, „Plant Disease Identification from individual lesions spots using deep learning‟, Apr 2019.

12. K. Naga Subramanian, A.K. Singh, A. Singh, S. Sarkar and Ganpath Subramanian, “Usefulness of interpretability methods to explain deep learning-based plant stress phenotyping”, July 2020.

13. W. Yang, C. Yang, Z. Hao, C. Xie, and M. Li, „„„Diagnosis of plant cold damage based on hyperspectral imaging and convolutional neural network,‟‟‟ IEEE Access, vol. 7, pp. 118239– 118248, 2019.

14. K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathy subramanian, „„„Plant disease identification using explainable 3D deep learning on hyperspectral images”.

15. N. Zhang, G. Yang, Y. Pan, X. Yang, L. Chen, and C. Zhao, „„„A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades,‟‟‟ Remote Sens., vol. 12, no. 19, Sep. 2020, Art. no. 3188.

16. J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, „„„Using deep transfer learning for image-based plant disease identification,‟‟‟ Comput. Electron. Agricult., vol. 173, Jun. 2020, Art. no. 105393.

17. M. Agarwal, A. Singh, S. Ajaria, A. Sinha, and S. Gupta, „„„ToLeD: Tomato leaf disease detection using convolution neural network,‟‟‟ Procedia Comput. Sci., vol. 167, pp. 293–301, Jan. 2020.

18. G. Hu, H. Wu, Y. Zhang, and M. Wan, „„„A low shot learning method for tea leaf’s disease identification,‟‟‟ Comput. Electron. Agricult., vol. 163, Aug. 2019, Art. no. 104852.

APPENDIX-A

PSUEDOCODE

```
# Importing essential libraries and modules
import os
from flask import Flask, render_template, request, redirect, jsonify
import numpy as np
import pandas as pd
import requests
import torch
from torchvision import transforms
from PIL import Image
from markupsafe import Markup
import warnings
import joblib
import io
from utils.disease import disease_dic
from utils.model import ResNet9
import openai
import config
from openai import OpenAI

# Initialize Flask app
app = Flask(__name__)

# Initialize OpenAI client
try:
    client = OpenAI(api_key="sk-eLNufXkRYaYKuaDQbxrP1-
OXwvr2z3SmGsGVZo9S2NT3BlbkFJF1FDv0HRhZRvgrPNghvaavmi_3poJgAVbYtyCcXg4A")
except Exception as e:
    print(f"⚠️ Error initializing OpenAI client: {e}")
    client = None

#
==================================================================
===================================
# Disease Classification Model Setup
rice_disease_classes = [
```

```
    'Rice___Bacterial_leaf_blight',
    'Rice___Brown_spot',
    'Rice___Leaf_smut',
    'Rice___healthy'
]

# Initialize disease model
disease_model = None
try:
    disease_model_path = os.path.join(os.path.dirname(__file__), 'models', 'plant_disease_model.pth')
    print(f"Looking for model at: {disease_model_path}")  # Debug print

    if os.path.exists(disease_model_path):
        disease_model = ResNet9(3, len(disease_classes))
        disease_model.load_state_dict(torch.load(disease_model_path,
map_location=torch.device('cpu')))
        disease_model.eval()
        print("✅ Disease model loaded successfully!")
    else:
        print(f"❌ Error: Disease model file not found at: {disease_model_path}")
        print("Current working directory:", os.getcwd())
        print("Directory contents:", os.listdir(os.path.join(os.path.dirname(__file__), 'models')))
except Exception as e:
    print(f"⚠️ Error loading disease model: {e}")

# Crop Recommendation Model
crop_recommendation_model = None
try:
    crop_recommendation_model_path = os.path.join('models', 'RandomForest.pkl')
    if os.path.exists(crop_recommendation_model_path):
        crop_recommendation_model = joblib.load(crop_recommendation_model_path)
        print("✅ Crop recommendation model loaded successfully!")
    else:
        print(f"⚠️ Crop recommendation model file not found at: {crop_recommendation_model_path}")
except Exception as e:
    print(f"⚠️ Error loading crop recommendation model: {e}")
```

```
# =============================================================================================
# Utility Functions
def weather_fetch(city_name):
    """Fetch weather data from OpenWeatherMap API"""
    try:
        api_key = config.weather_api_key
        base_url = "http://api.openweathermap.org/data/2.5/weather?"
        complete_url = f"{base_url}appid={api_key}&q={city_name}"
        response = requests.get(complete_url)
        data = response.json()

        if data["cod"] == 200:
            main_data = data["main"]
            temperature = round((main_data["temp"] - 273.15), 2)  # Kelvin to Celsius
            humidity = main_data["humidity"]
            return temperature, humidity
        return None
    except Exception as e:
        print(f"Weather API error: {e}")
        return None

def predict_image(img, model=disease_model):
    """Predict disease from image"""
    if model is None:
        return "Disease prediction model not loaded"

    try:
        transform = transforms.Compose([
            transforms.Resize((256, 256)),
            transforms.ToTensor(),
        ])
        image = Image.open(io.BytesIO(img)).convert('RGB')
        img_t = transform(image).unsqueeze(0)
        output = model(img_t)
        _, preds = torch.max(output, dim=1)
        return disease_classes[preds[0].item()]
```

```
    except Exception as e:
        return str(e)

#
=============================================================
============================
# Flask Routes
@app.route('/')
def home():
    return render_template('index.html', title='PrefectCrop - Home')

@app.route('/crop-recommend')
def crop_recommend():
    return render_template('crop.html', title='PrefectCrop - Crop Recommendation')

@app.route('/disease-predict', methods=['GET', 'POST'])
def disease_prediction():
    title = 'PrefectCrop - Disease Detection'

    if disease_model is None:
        return render_template('error.html', title=title, error="Disease model not loaded")

    if request.method == 'POST':
        if 'file' not in request.files:
            return redirect(request.url)

        file = request.files.get('file')
        if not file:
            return render_template('disease.html', title=title)

        try:
            img = file.read()
            prediction = predict_image(img)
            prediction = Markup(str(disease_dic.get(prediction, "No recommendation found.")))
            return render_template('disease-result.html', prediction=prediction, title=title)
        except Exception as e:
            return render_template('error.html', title=title, error=str(e))
```

```
    return render_template('disease.html', title=title)

@app.route('/chatbot', methods=['POST'])
def chatbot():
    try:
        user_message = request.json["message"]

        response = openai.ChatCompletion.create(
            model="gpt-4",
            messages=[{"role": "user", "content": user_message}]
        )

        reply = response.get("choices", [{}])[0].get("message", {}).get("content", "Sorry, I couldn't
understand that.")
        return jsonify({"reply": reply})
    except Exception as e:
        return jsonify({"reply": f"Error: {str(e)}"})
    print(response)

if __name__ == '__main__':
    warnings.filterwarnings("ignore", category=UserWarning, module='sklearn')
    app.run(debug=True)
```

APPENDIX-B

SCREENSHOTS

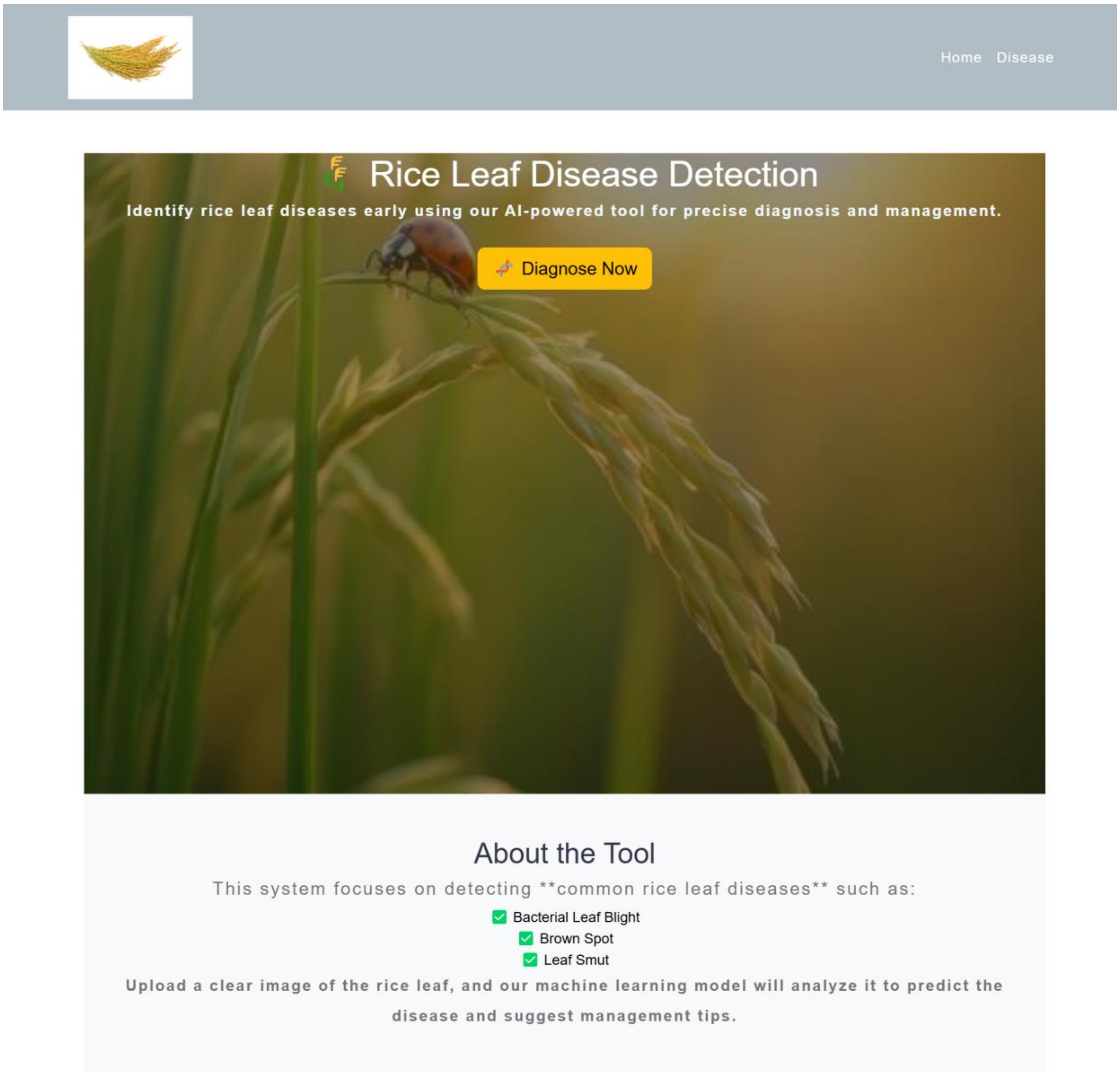


Fig.1: Home page

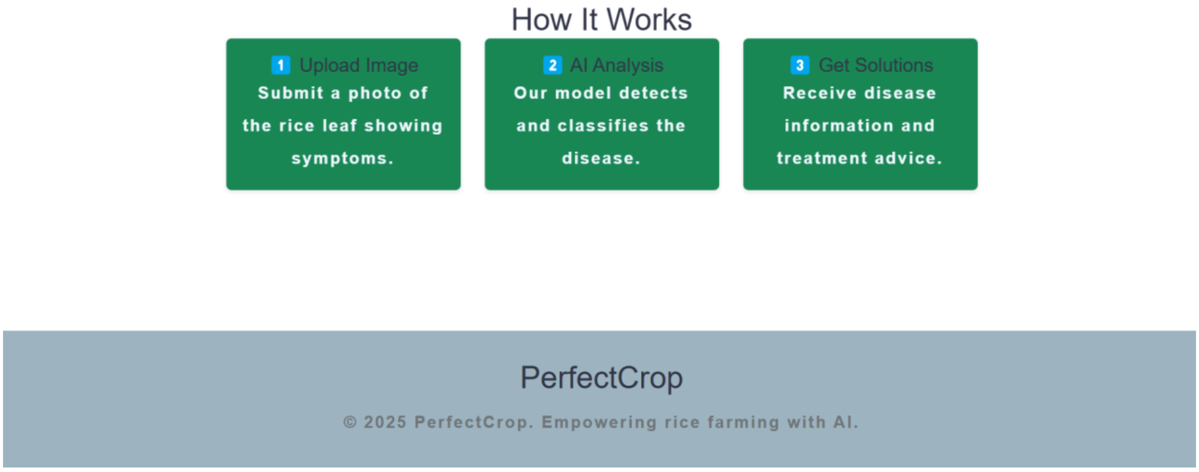


Fig.2: Platform Features

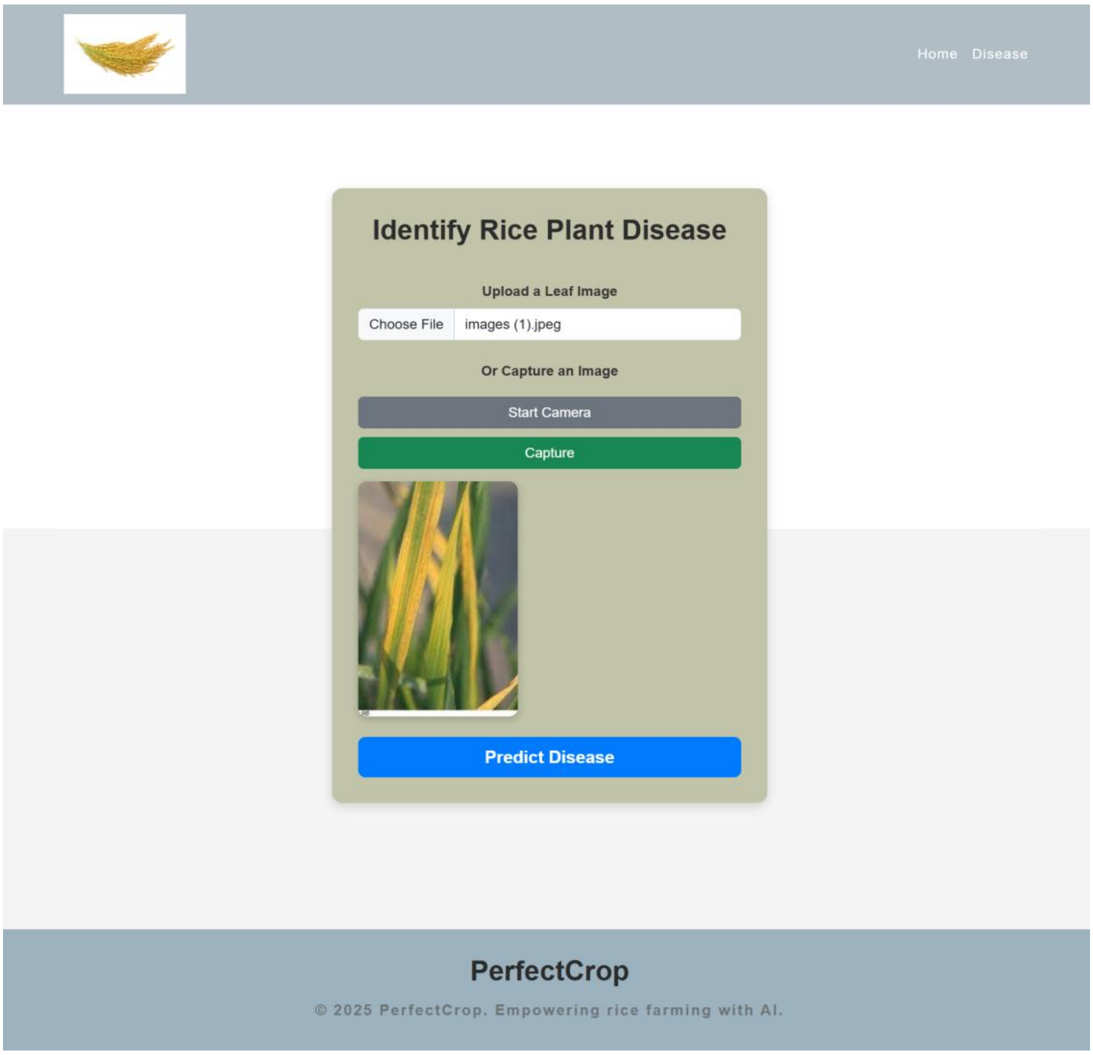


Fig.3: Disease Detection System

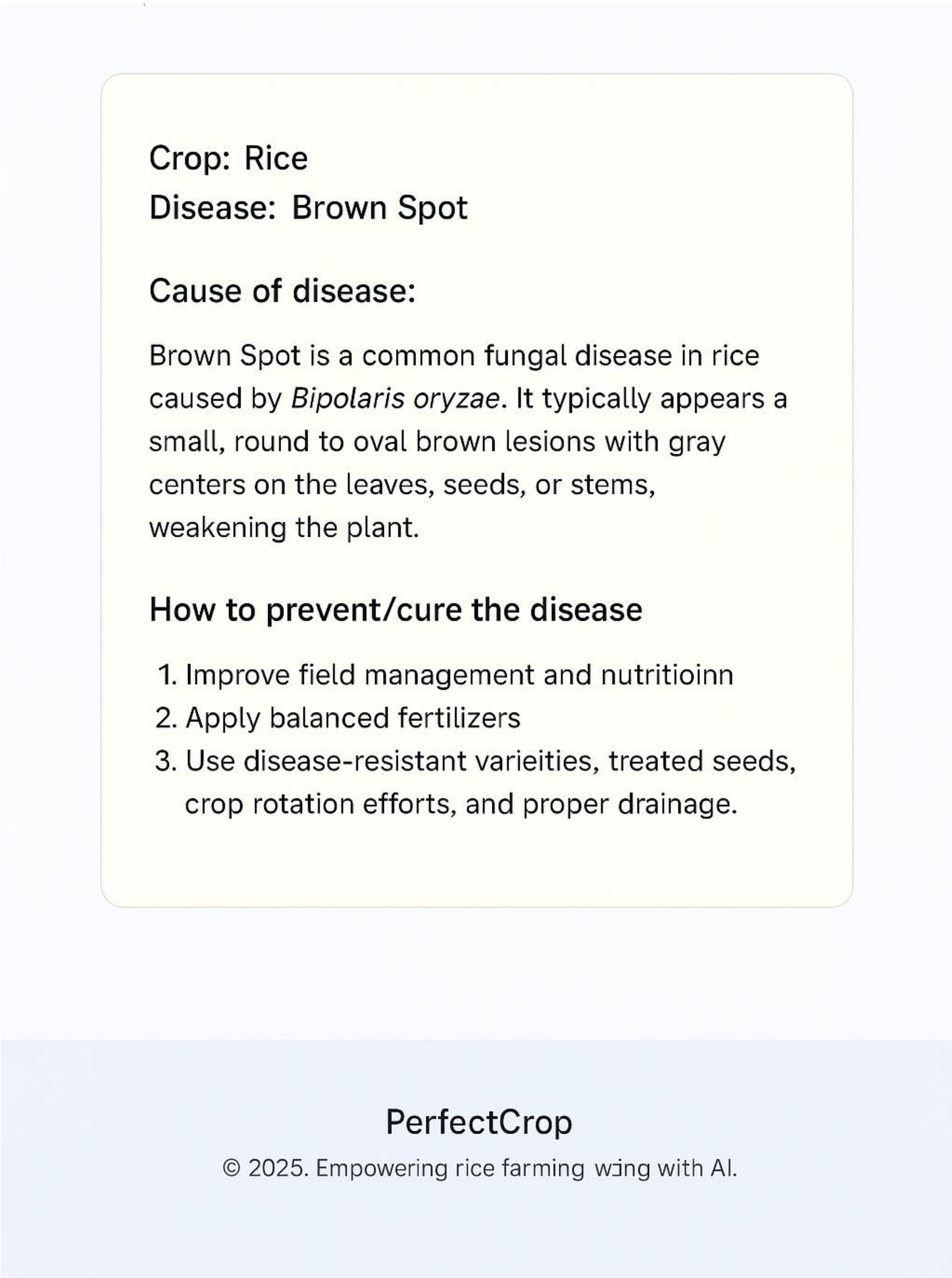


Fig.4: Disease Detection System giving output

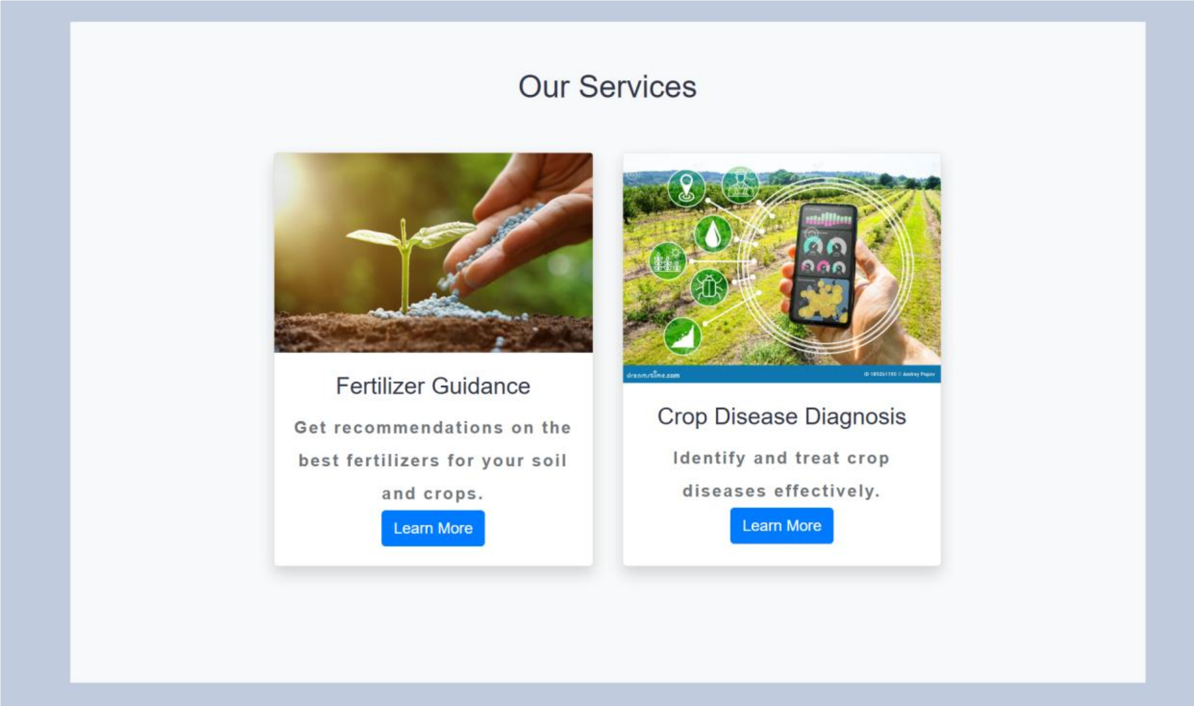


Fig.5: Our Services

APPENDIX-C

ENCLOSURES

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

ISSN: 2582-3930

Impact Factor: 8.586

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

VENKATA SAI REDDY K

in recognition to the publication of paper titled

Automated Rice Disease Identification and Management Using Machine Learning

published in IJSREM Journal on Volume 09 Issue 05 May, 2025

Editor-in-Chief

IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com


ISSN: 2582-3930
Impact Factor: 8.586
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT
An Open Access Scholarly Journal || Index in major Databases & Metadata
CERTIFICATE OF PUBLICATION
International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to
SHAIK MANSOOR AHAMED ALI
in recognition to the publication of paper titled
Automated Rice Disease Identification and Management Using Machine Learning
published in IJSREM Journal on Volume 09 Issue 05 May, 2025
www.ijsrem.com
Editor-in-Chief
IJSREM Journal
e-mail: editor@ijsrem.com

ISSN: 2582-3930
Impact Factor: 8.586

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

KARTHIK KUMAR REDDY G

in recognition to the publication of paper titled

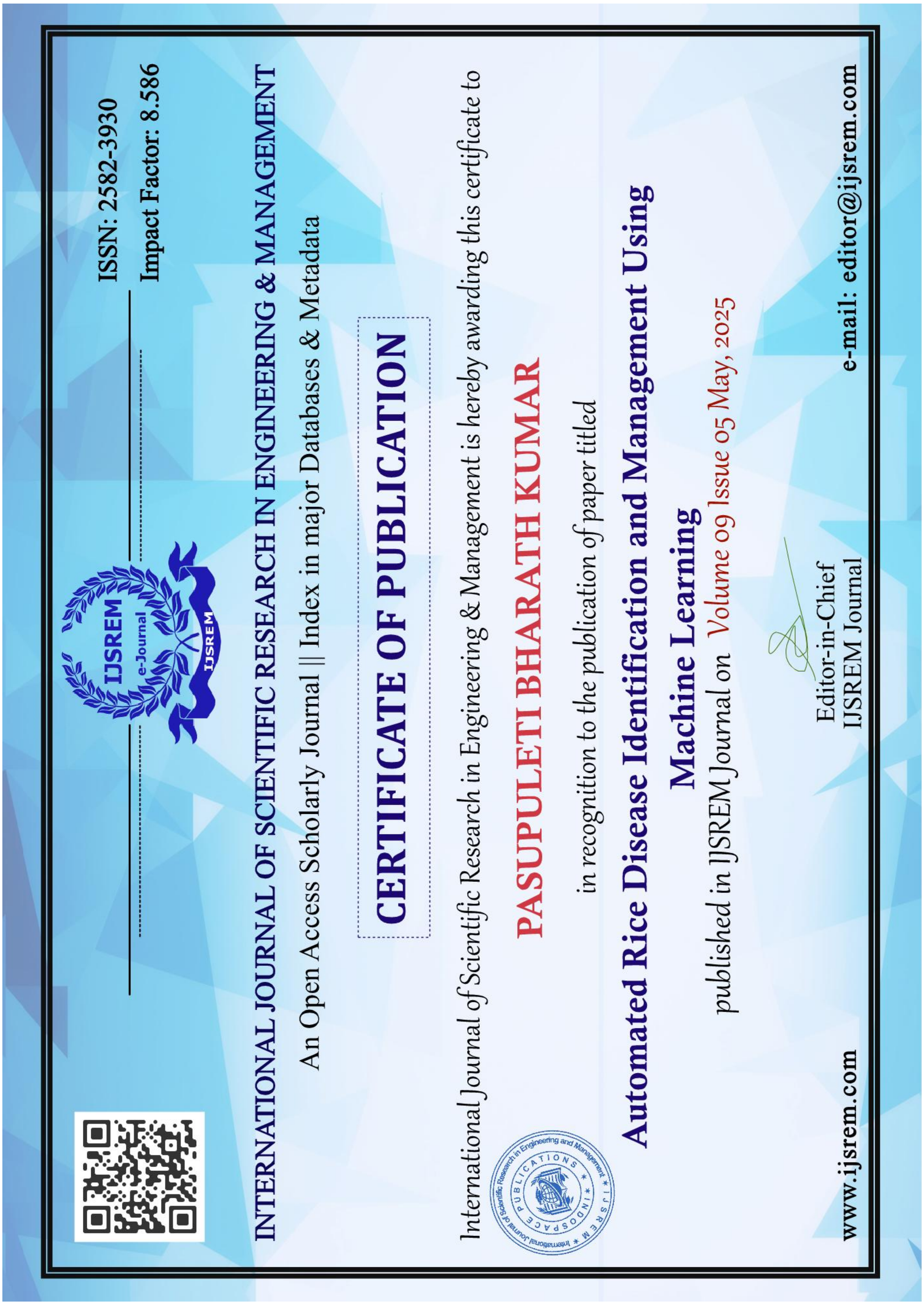
Automated Rice Disease Identification and Management Using Machine Learning

published in IJSREM Journal on Volume 09 Issue 05 May, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com


ISSN: 2582-3930

Impact Factor: 8.586

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is *hereby* awarding this certificate to

PASUPULETI BHARATH KUMAR

in recognition to the publication of paper titled

Automated Rice Disease Identification and Management Using Machine Learning

published in IJSREM Journal on *Volume 09 Issue 05 May, 2025*

www.ijsrem.com

Editor-in-Chief

IJSREM Journal

e-mail: editor@ijsrem.com


ISSN: 2582-3930

Impact Factor: 8.586

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

SATHELA SREEKAR REDDY

in recognition to the publication of paper titled

Automated Rice Disease Identification and Management Using Machine Learning

published in IJSREM Journal on Volume 09 Issue 05 May, 2025

Editor-in-Chief

IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Automated Rice Disease Identification and Management Using Machine Learning

VENKATA SAI REDDY K [1], SHAIK MANSOOR AHAMED ALI [2], KARTHIK KUMAR REDDY G [3], PASUPULETI BHARATH KUMAR [4], SATHELA SREEKAR REDDY [5]

[1,2,3,4,5] *Students of Presidency University, Bengaluru*

***Abstract** - The abstract provided brings attention to a project aimed at enhancing rice crop well-being by using artificial intelligence and machine learning. Rice as an essential food grain for over half of the world's population is highly susceptible to various plant diseases that can highly decrease yield and jeopardize food security. To solve this, the project presents an AI-based system that can automatically identify rice plant diseases using images of rice leaves. The system utilizes supervised learning and unsupervised learning methodologies. In supervised learning, models including Convolutional Neural Networks (CNNs) are trained on labeled data in order to identify and classify known diseases correctly. Concurrently, unsupervised learning techniques such as K-Means clustering are employed to detect patterns and clusters in the image data that could reflect emerging or previously undesignated types of diseases. This two-pronged method not only enhances the precision of disease identification but also enables the system to adjust to new plant health concerns as they emerge. A farmer can easily load photos of the infected rice leaves, and the system will study the image and diagnose the disease as well as suggest remedies accordingly. The system saves on expertise dependency in agriculturalists, and the process leads to early identification and encourages precision farming, finally resulting in curbing losses during crop growth and maximizing the efficiency of food production.*

***keyword:** Rice Disease Detection, Artificial Intelligence (AI), Machine Learning (ML), Convolutional Neural Networks (CNNs), Supervised Learning, Unsupervised Learning, K-Means Clustering, Image-Based Diagnosis*

I. INTRODUCTION

Rice is one of the world's most eminent staple crops, nourishing over half of the world's population and contributing to food security and economic stability, particularly in the developing world. Rice crop yield, however, is seriously affected by numerous plant diseases like leaf blast, bacterial blight, and brown spot. These diseases are highly infectious and can lead to extensive crop damage, low yields, and massive economic losses to farmers. Conventional detection of diseases in rice cultivation is done by manual observation by agricultural experts or farmers themselves. This not only takes a lot of time and effort but is also prone to errors, particularly where expert knowledge or diagnostic equipment is inaccessible.

To solve these issues, this project suggests an AI solution for rice disease identification through automated processes by processing images of infected rice leaves. The platform is a blend of supervised and unsupervised learning, where supervised learning from machine learning models trained on labeled data is used to identify specific diseases and unsupervised learning to identify unforeseen or unknown patterns of disease without labels. For instance, supervised algorithms such as Convolutional Neural Networks (CNNs) are applied to classify known diseases with a high degree of accuracy, whereas clustering algorithms such as K-Means are applied to identify anomalous patterns that

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

could potentially identify new or misdiagnosed infections.

The general aim of this system is to facilitate the early and accurate identification of rice diseases with minimal human intervention. By enabling farmers to upload a quick snapshot of an infected leaf, the system can scan the image in real time, detect the disease (if known) and recommend treatments. Not only does this enable faster and more accurate detection of disease, but also enables precision agriculture where resources such as pesticides and fertilizers are utilized optimally. Lastly, this method based on AI lowers the dependence on professional diagnosis, improves crop management, and promotes sustainable agriculture.

II. RESEARCH ELABORATION

In recent years, there has been an increasing interest in using Artificial Intelligence (AI) to solve problems in agriculture, specifically plant disease diagnosis. There have been a series of studies showing that AI, particularly machine learning methods, can greatly improve the efficiency and accuracy of crop disease detection. Among them, supervised learning algorithms like Convolutional Neural Networks (CNNs) have been very helpful in solving image classification problems. CNNs find specific uses in identifying and categorizing common plant diseases through training on extremely large sets of annotated images—images already supplied with the proper disease label. In training on these sets, models learn to accurately and reliably distinguish between healthy leaves and infected ones.

Alternatively, unsupervised learning techniques like K-Means clustering supply a different, but complementary solution. While unsupervised models do not use labeled information like their supervised counterparts, unsupervised algorithms search for the inherent patterns or structures of data in a way that groups similar images. Such ability would be especially helpful for identifying novel or unreclassified symptoms of disease, following various stages of progression of disease, or classifying data that is unannotated. In addition to supervised learning, unsupervised techniques aid in creating an even more adaptive and intelligent system that can efficiently handle known as well as unknown situations.

In this project, the hybrid methodology is followed with both supervised and unsupervised techniques. Supervised learning identifies rice leaf diseases by using labeled data sets to train the system in detecting known infections like leaf blast, brown spot, and bacterial blight accurately. Concurrently, unsupervised learning clusters unlabeled images, identifying new patterns, disease phases, or outliers that are not in the training data. The strength of this dual approach is backed by the application of high-quality data sets, such as the PlantVillage rice subset, a highly referenced source of labeled images of plant disease, and in-house rice disease data sets obtained from real-world agricultural environments. These varied data sets make the system solid, responsive, and able to make sound predictions in real-world agricultural settings.

III. METHODOLOGY

The desired AI model to detect rice diseases is to be deployed in a multi-stage pipeline utilizing both supervised and unsupervised learning methods for providing high accuracy along with responsiveness.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

All of these steps in the pipeline have an important role in converting raw input data into actionable insights and useful prediction that benefits farmers.

The initial step is Data Collection, in which a high-quality and diverse set of rice leaf images is gathered. The images consist of healthy and infected leaves in a way that the model learns to identify healthy and infected conditions. For diseases that are well known, the images are manually tagged with their respective disease names like Leaf Blast, Brown Spot, or Bacterial Blight. For rendering the model strong and avoiding overfitting, the dataset is also augmented with rotation, flip, scale, and brightness techniques. This provides a more varied array of training examples to which the model can learn to generalize to new, unseen images.

The second process is the Preprocessing step, in which the images are processed for machine learning. All images are resized to a standard size of 256x256 pixels to ensure that there is uniformity in the input size. Normalization is applied for normalizing pixel values, and noise removal operations are applied to remove unwanted visual information which could compromise prediction accuracy. Significant visual properties are then determined using methods like histogram analysis or texture analysis so that the model can learn critical patterns like color histogram, leaf texture, or lesion morphology.

In the Supervised Learning step, a Convolutional Neural Network (CNN), normally models like ResNet18 or ResNet50, is trained with the labeled data. These deep learning models can be trained to learn complex patterns from images and are commonly used in visual recognition applications. Having been trained, the CNN can take a new image and produce the predicted class, report whether the leaf is infected with Leaf Blast or Brown Spot, based on what it has learned about features. Furthermore, the system also employs Unsupervised Learning in the form of K-Means clustering.

The algorithm doesn't need any labeled data and is applied to cluster images according to visual similarity. It comes in handy when applied to cases that are uncertain or unlabeled, finding anomalies, or finding potential new classes of diseases that weren't covered in the training data. With similar but unlabelled images clustered into a group, the system can warn researchers or farmers of potential disease danger. Lastly, the Prediction Pipeline encapsulates all the pieces into a friendly user interface.

The user—usually a farmer—uploads an image of an infested rice leaf into the system. The picture is processed through the CNN model, which identifies the disease and gives an answer in the form of the name of the infestation together with recommended treatment or prevention protocols. But in case the model is not confident or uncertain in its diagnosis, the system initiates the clustering algorithm to study the picture again. If the image is significantly different from familiar categories, the system can generate an anomaly alert, suggesting a new or unfamiliar disease. This integrated approach guarantees correct identification of familiar diseases and adaptive detection of new threats, rendering it extremely useful for contemporary, data-driven agriculture.

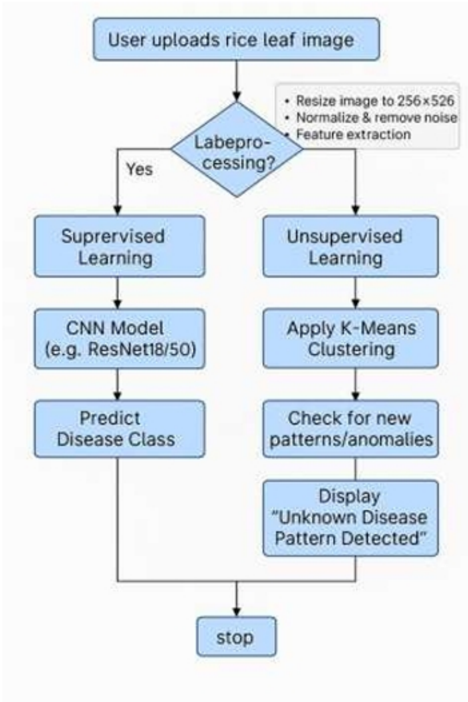
IV. IMPLEMENTATION AND RESULTS

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

The usage of the rice leaf disease detection system has two dominant components: a convenient frontend and an intelligent backend based on machine learning. The frontend is programmed with Flask, a lightweight web framework based on Python for deployment and user friendliness. The frontend is optimized to be user-friendly and handy, where a user, generally a farmer or farmhand, can upload a picture of a rice leaf through a web application with little technical effort.

After the user uploads a picture, the picture is sent to the backend where the input is already being processed by the core prediction engine. The backend consists of two primary machine learning modules: one supervised CNN-based model and another unsupervised K-Means clustering module. The model is a supervised one, preferably a CNN-based network like ResNet18 or ResNet50, that does processing of the uploaded image and prediction of what particular disease the rice plant is suffering from. The model is pre-trained with labeled images and can detect common diseases like Leaf Blast, Brown Spot, and Bacterial Blight.

To give accurate shot prediction, the system also verifies the level of confidence in the prediction made by the CNN. If the confidence level of the prediction is below a threshold value (suspicion or chance of misclassification), the system resorts to the unsupervised learning module. The K-Means clustering module clusters the image by visual similarity with comparable known or unknown patterns and assists in recognizing anomalies or potential new classes of diseases is mentioned below flow diagram.



The last output that is sent back to the user is the name of the predicted disease, a short description of symptoms, and the action or treatment to be used. As an instance, if the input image is one of black spots on the rice leaf, the system may predict "Brown Spot," provide the characteristic signs of the disease, and advise proper treatment, including applying the fungicide Mancozeb and adjusting irrigation practice to prevent overwatering, which worsens the disease. This concerted effort not only aids in correct diagnosis but also informs users with precise information to correct the disease. The solution is scalable, efficient, and most useful in rural or under-developed areas where expert agricultural counsel may not be easily attainable.

V. SYSTEM STUDY AND TESTING

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

There was an in-depth evaluation to examine the reliability and performance of the system of rice disease detection. As a starting point, the training and testing dataset was divided with a common 80:20 ratio—80% of images utilized in training the machine learning models and 20% held back for testing. This has the advantage of allowing the models to be trained on a large part of the data yet still tested on unseen instances, which serves to quantify their ability to generalize.

Various test metrics were employed to study the performance of the supervised model (CNN). The system recorded an accuracy of 92%, meaning that it accurately predicted the disease class for most test images. The accuracy, which calculates how many of the forecasted positive cases were indeed correct, was 89%. This is to say that the model predicted very few false positives. Secondly, the recall (or sensitivity), which reflects how well the model picked up on actual positive cases, was at 90%, indicating that the model was good at picking up on true disease occurrences without missing out on many instances. In combination, these measurements indicate a high-performance classification model, able to make accurate predictions in real-world applications.

For further verification, several testing mechanisms were utilized. A confusion matrix was utilized to examine the effectiveness of the model in differentiating between various categories of disease. It graphically emphasizes where the model correctly and incorrectly classified, assisting in comprehending particular areas of confusion or strength—like whether the model occasionally confuses Brown Spot with Leaf Blast. For the unsupervised learning aspect (K-Means clustering), a Silhouette Score was determined. This score quantifies how well-separated and distinct the clusters are, with higher scores reflecting better-defined groupings of similar disease patterns. A high silhouette score ensures that the clustering algorithm is successful in identifying underlying patterns or outliers in unlabeled data.

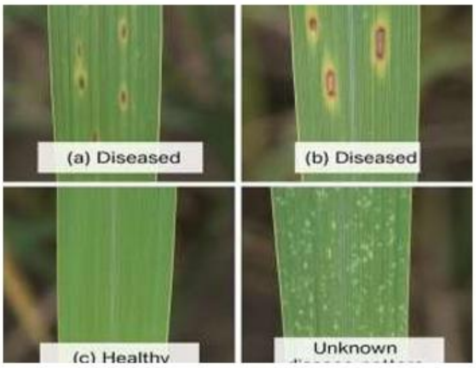
The system also underwent stress testing under different conditions to assess its strength. This included testing it using images of rice leaves photographed under various lighting conditions, at a range of angles, and with different image qualities (blurred, shadowed, etc.). These tests mimicked real-world scenarios where farmers might be submitting images captured using mobile phones under uncontrolled lighting conditions. In spite of these difficulties, the system performed quite steadily, demonstrating its real-world usability in field applications.

Yet, the research also found some limitations. One of the main challenges is the lack of diversity in the dataset—specifically in the variety of rice and the look of the same disease in various geographic locations. For example, Brown Spot may look slightly different based on the rice strain or the climate, which would impact the accuracy of the model. Moreover, infrequent or new diseases that are not adequately covered in the training set may be mislabeled or entirely omitted. Overcoming these constraints would involve ongoing dataset augmentation and model retraining, involving cooperation with agricultural specialists and field data acquisition across various regions.

VI. RESULTS

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Outputs of the rice disease diagnostic system were very encouraging, both technically viable and practically applicable. The CNN trained from labeled rice leaf images was very capable of identifying known rice diseases like Leaf Blast, Brown Spot, and Bacterial Blight. Robust performance justified the application of the supervised learning method, particularly when followed by good preprocessing quality and image augmentation methods in training. The model made accurate predictions every time for images with a broad variety of input images, ranging from those taken in varied environmental conditions. Beyond the supervised classification, the system's unsupervised learning module, implemented using K-Means clustering, yielded significant information beyond just disease detection. The clustering model might classify visually similar but unlabeled images into unlabeled subtypes or disease progression stages that were not labeled in the training set. For instance, it would be able to distinguish between early and late Brown Spot or identify subtle differences in leaf damage, allowing for more effective disease management. This attribute richens the system by allowing ongoing research to be conducted and allowing agricultural scientists to pinpoint new trends or forms of disease.



Also, the system proved to be reliable and resilient when tested against real-field images—taken by farmers directly in their smartphones during actual field illumination conditions. Testing showed that the model was fast and robust and could deal with changing image quality and extraneous interference. Deploying the model into a web-based user interface also made sure that the system was usable as well as accessible to non-users.

The system was very well thought of by trial farmers, who liked its ease of use, speed of response, and value in providing beneficial disease management advice. Posting an image of a leaf and obtaining rapid diagnosis, together with suggestions for treatment (e.g., correct use of fungicides), was seen as a critical real-time decision aid. This response confirmed the capacity of AI to assist farmers, foster crop well-being, and ultimately achieve food security through more rational and responsive agriculture.

VII. FUTURE ENHANCEMENT

Though the existing rice disease detection system is highly accurate and practically applicable, there are a number of promising directions to pursue further development to improve it even more in terms of performance, scalability, and real-world usability.

One of the major steps forward is to increase the dataset by encompassing a broader range of rice strains and region-specific disease images. The existing dataset, although effective, is too narrow in scope. Rice cultivars cultivated in different locations could exhibit different visual indications for a particular disease owing to variations in climate, soil,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

or crop genes. By acquiring more images from other agro-climatic zones, the system can be made accurate and robust in different farming conditions and minimize the risk of misclassification.

Another important improvement is the creation of a mobile application, which would give farmers immediate access to the technology. A mobile app would enable users to upload and capture images of rice leaves from their phones and obtain instant predictions and treatment recommendations. This would not require a laptop or stationary internet connection, filling the digital divide between rural farmers and enabling them to diagnose and treat disease on the spot.

Moreover, incorporating time-series analysis features would provide another layer to the system. By examining a series of pictures of the same plant over time, the system could track how the disease is advancing or reacting to treatment. It would determine whether the prescribed intervention is effective or additional measures are required, thereby facilitating personalized crop care and ongoing monitoring.

In order to make the system stronger in detecting anomalies further, and identify concealed or alternative patterns of disease, subsequent versions would include sophisticated algorithms for unsupervised machine learning like autoencoders or DBSCAN (Density-Based Spatial Clustering of Applications with Noise). More effectively than K-Means, these possess superior ability to differentiate non-linear and irregular sets, and thus they are an appropriate choice in instances of selecting outliers, solitary diseases, or mild symptoms breaking neat bunchings.

Lastly, being integrated with actual-time weather and environmental data, predictive power can be highly boosted and proactive warnings against diseases can be offered. As rice diseases also have direct relationships with certain environmental factors such as humidity, temperature, and rain, integrating the system with weather APIs or IoT sensors has a potential to forecast disease outbreaks before they actually happen. For example, if the situation is favorable for the growth of fungi, the system can alert farmers to use preventive techniques prior to the occurrence of any symptoms, leading to smart and sustainable agriculture.

VIII. CONCLUSION

The "Automated Rice Disease Identification and Management Using Machine Learning" project illustrates the power of AI in transforming agricultural diagnosis. Being one of the globe's leading staple crops, rice's health is vital to food security. This project proposes a real-world, smart solution that applies supervised machine learning mainly convolutional neural networks (CNNs) and transfer learning methods to detect typical rice leaf diseases like Bacterial Leaf Blight, Brown Spot, and Leaf Smut from image data. With a simple web-based interface, farmers or extension workers can upload an image of a rice leaf, and the model will scan the image and identify the occurrence and nature of disease. The real-time automated system

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

reduces the need for expert input, cuts down on diagnostic time, and can reduce crop loss by allowing early and precise disease management. The workflow of the platform is straightforward: upload a photo, have the AI interpret it, and get actionable disease information and recommendations making the technology available even to users with little technical know-how.

Secondly, the system is lightweight and scalable to be deployed on mobile platforms for offline usage in rural settings. This makes it possible for the advantages of AI-based crop protection to extend to farmers in remote areas. The visual appearance of the application, including user-friendly navigation, nice-looking illustrations, and simple step-by-step guides, also increases user participation and trust.

Finally, this project not only achieves its goal of rice disease detection through supervised learning but also sets the stage for more intelligent farming methods. By integrating machine learning with open digital platforms, it educates farmers, safeguards plant health, and promotes future sustainable agriculture. With the enlargement of the dataset, real-field testing, and multilingual capabilities, the system can be developed into a broadly used tool in varying rice cultivation areas around the world.

IX. REFERENCES

1. D.Jayanarayana Reddy, Dr. M. Rudra Kumar, "Crop Yield Prediction using Machine Learning Algorithm", 2024.
2. Nishu Bali & Anshu Singal, "Deep Learning Based Wheat Crop Yield Prediction Model in Punjab Region of North India, 16 Sep 2021.
3. Reyana, Sandeep Kautish, P. M. Sharan Karthik, Ibrahim Ahmed Al-Baltah, Muhammed Basheer Jasser, (Member, IEEE), And Ali Wagdy Mohamed, Accelerating Crop Yield: Multisensor Data Fusion and Machine Learning for Agriculture Text Classification, 27 February 2023.
4. Alexandros Oikonomidis, Cagatay Catal & Ayalew Kassahun, Deep learning for crop yield prediction: a systematic literature review, 06 Feb 2022.
5. Sujatha M., Jaidhar C.D, "Machine learning-based approaches to enhance the soil fertility, 14 November 2023.
6. P K Devan, Swetha B, Uma Sruthi P, Varshini, Crop Yield Prediction and Fertilizer Recommendation System Using Hybrid Machine Learning Algorithm, 14th March 2023.
7. Mamatha, J.C. Kavitha c, Machine learning based crop growth management in greenhouse environment using hydroponics farming techniques.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

8. 10 January 2023. 10. Chetan R, D.V. Ashoka, Ajay Prakash B, IMLAPC: Interfaced Machine Learning Approach for Prediction of Crops, 11 January 2022.

9. Simraneet Kaur, Getanajali Babar, Navneet Sandhu, Dr. Gagan Jindal, Various Plant disease detection using Image Processing Methods, June 2019.

10. Er. Varinderjit Kaur,Dr. Ashish Oberoi, Wheat disease detection using svm classifier, Aug 2018.

11. Barbedo, Plant Disease Identification from individual lesions spots using deep learning", Apr 2019.

12. K. Naga Subramanian, A.K. Singh, A. Singh, S. Sarkar and Ganpath Subramanian, Usefulness of interpretability methods to explain deep learning-based plant stress phenotyping, July 2020.

13. W. Yang, C. Yang, Z. Hao, C. Xie, and M. Li, Diagnosis of plant cold damage based on hyperspectral imaging and convolutional neural network, IEEE Access, vol. 7, pp. 118239–118248, 2019.

14. K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathy subramanian, Plant disease identification using explainable 3D deep learning on hyperspectral images".

15. N. Zhang, G. Yang, Y. Pan, X. Yang, L. Chen, and
C. Zhao, A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades Remote Sens., vol. 12, no. 19, Sep. 2020, Art. no. 3188.

16. J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, Using deep transfer learning for image-based plant disease identification, Comput. Electron. Agricult., vol. 173, Jun. 2020, Art. no. 105393.

17. M. Agarwal, A. Singh, S. Ajaria, A. Sinha, and S. Gupta, ToLeD: Tomato leaf disease detection using convolution neural network, Procedia Comput. Sci., vol. 167, pp. 293–301, Jan. 2020.

18. G. Hu, H. Wu, Y. Zhang, and M. Wan, A low shot learning method for tea leaf's disease identification, Comput. Electron. Agricult., vol. 163, Aug. 2019, Art. no. 104852.

Joe Arun Raja - CEI-02Projectreport

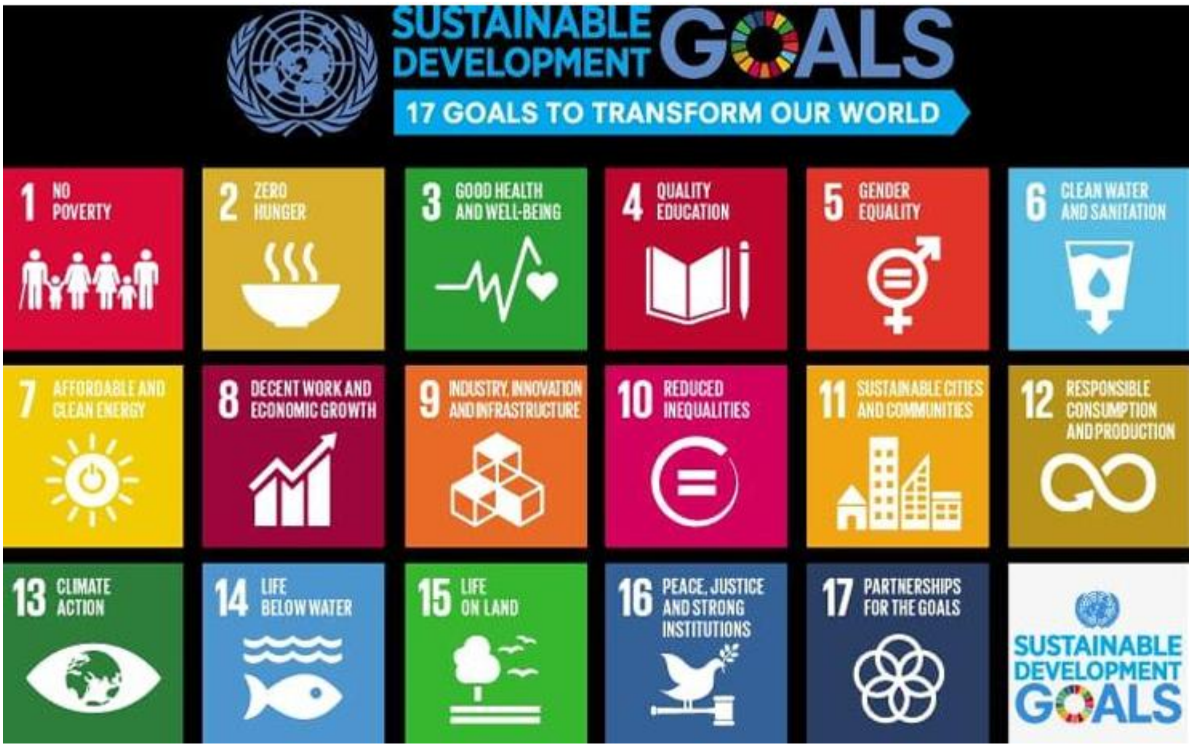
ORIGINALITY REPORT

13%	10%	7%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	github.com Internet Source	3%
2	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	2%
3	Submitted to University of Hertfordshire Student Paper	1%
4	Md. Ahsan Ullah, Sk. Humayra Afrin, Kazi Mohammad Nazib, Raju Roy, Lasker Ershad Ali. "Unravelling Parkinson's Disease Prediction: An Evaluation of Feature Selection Techniques with a Focus on PCA and KNN Performance", Review of Computer Engineering Studies, 2023 Publication	1%
5	official.satbayev.university Internet Source	1%
6	"Artificial Intelligence: A Real Opportunity in the Food Industry", Springer Science and Business Media LLC, 2023 Publication	<1%
7	www.researchgate.net Internet Source	<1%

SUSTAINABLE DEVELOPMENT GOALS



- **SDG 2: Zero Hunger**

AgroDoc contributes to ending hunger by improving crop yields through optimized farming practices, disease detection, and resource-efficient fertilizer use. It helps ensure food security by reducing crop losses and promoting sustainable agriculture.

- **SDG 3: Good Health and Well-being**

By minimizing the use of incorrect pesticides and chemicals through accurate diagnosis, the project contributes to safer food and a healthier environment for farmers and consumers

- **SDG 12: Responsible Consumption and Production**

The app promotes efficient use of natural resources, such as water, soil, and fertilizers,by offering data-driven recommendations. This reduces waste and minimizes the environmental impact of farming, supporting sustainable consumption and production practices.

- **SDG 10: Reduced Inequalities**

  By providing accessible disease detection tools (like mobile/web platforms), the project reduces the gap between rural farmers and access to modern technology.

- **SDG 9: Industry, Innovation, and Infrastructure**

  By integrating deep learning and advanced technologies into agriculture, AgroDoc fosters innovation in farming techniques, contributing to the development of sustainable infrastructure in agriculture.