Assignment: Plant Disease Detection System for Sustainable Agriculture (AI/ML)

# 1. Problem Statement

In agriculture, timely detection of plant diseases and efficient management of crop productivity are crucial for sustainable farming. However, traditional methods of disease identification often require manual inspection, which is time-consuming, labor-intensive, and prone to errors. Additionally, farmers often lack real-time data on soil quality and environmental conditions, which affects their ability to make informed decisions about crop selection and fertilizer use.

This project aims to build an **AI-powered plant disease detection system** that:

* Automatically detects diseases from **images of plant leaves** using advanced **machine learning and image processing** techniques.
* Considers **environmental factors** such as soil quality, temperature, and humidity to optimize crop recommendations.
* Provides **fertilizer suggestions** based on the disease status and environmental conditions to enhance crop productivity.

# 2. Project Pipeline

The overall project workflow follows a step-by-step process to achieve the final disease classification model. Below is an expanded explanation of each stage:

## Step 1: Data Collection & Data Loading

- Purpose: Gather a diverse and labeled dataset of leaf images for different crops and disease categories.  
- Structure: The dataset is typically divided into three parts:  
 - Training Set – used to train the model to learn patterns,  
 - Validation Set – used to fine-tune the model and avoid overfitting,  
 - Test Set – used to evaluate the model's final performance.  
Each image is associated with a category label such as “Grape\_\_\_Black\_rot” or “Corn\_\_\_healthy.”

## Step 2: ZIP & Mounting in Google Colab

- The dataset, usually large in size, is compressed (zipped) and stored in cloud storage platforms like Google Drive.  
- Using Google Colab, the drive is mounted and the dataset is extracted for processing.  
- This step is essential for efficient resource utilization, especially for students and researchers with limited computational resources.

## Step 3: Image Processing & Augmentation

- Image Preprocessing:  
 - Images are resized to a fixed dimension.  
 - Images are normalized to enhance learning efficiency.  
  
- Image Augmentation:  
 - Artificially increases the dataset size and diversity.  
 - Involves applying transformations like rotation, flipping, zooming, and shearing.  
 - Helps the model generalize better and perform well on unseen data.

## Step 4: CNN Model (Convolutional Neural Network)

- CNN is ideal for image classification as it extracts hierarchical features.  
- The architecture includes:  
 - Convolutional Layers for feature extraction,  
 - Pooling Layers for dimensionality reduction,  
 - Flatten and Dense Layers for output classification.  
- CNNs automatically learn features, eliminating the need for manual extraction.

## Step 5: Testing & Evaluation

- The model is evaluated using a separate test dataset.  
- Evaluation metrics:  
 - Accuracy: Overall correctness,  
 - Precision: Correctness of positive predictions,  
 - Recall: Ability to find all positive cases,  
 - F1-Score: Balance between precision and recall.

# 3. Flowchart of the Project Pipeline

