WEEK-1  
**Plant Disease Prediction Using Deep Learning**

**1. Problem Statement**

Agriculture forms the foundation of most economies, but crop yield and food production are severely affected by plant diseases. It is essential to have early and precise plant disease detection to aid farmers, reduce losses, and promote food security.  
Traditional diagnosis methods are labor-intensive, require a high level of technical knowledge, and are not affordable to everyone.  
Hence, there exists a necessity to design an automated system capable of reliably predicting plant diseases from images by utilizing Deep Learning methods.

## 2. About the Pipeline

The entire pipeline of plant disease prediction consists of the following steps:  
• **Dataset Collection:**  
Collect a vast dataset consisting of both diseased and healthy plant leaf images from publicly available sources, such as PlantVillage.

• **Data Preprocessing:**

* Resizing an image to a uniform size (such as 224x224 pixels).
* Normalization of pixel values.
* Data augmentation by rotation, flipping, and zooming to enhance dataset diversity.

• **Model Building and Selection:**  
Apply transfer learning using already-pre-trained networks such as ResNet50, VGG16, or MobileNet.  
Optimize the model using the plant disease dataset.  
Add more fully connected layers to perform classification.

• **Model Training:**

* Split the dataset into training, validation, and test sets.
* Utilize proper optimizers (e.g., Adam), loss functions (e.g., categorical cross-entropy), and evaluation metrics (e.g., accuracy, F1-score).
* Train the model using appropriate batch size and epochs to reach high accuracy.

• **Model Evaluation:**

* Test the model on unseen data to assess its performance.
* Utilize confusion matrix, precision, recall, and F1-score for thorough evaluation.

• **Deployment:**

* Deploy the model to a mobile or a web application to make it readily accessible.
* Make real-time predictions of diseases through upload or capture of leaf images.

• **Monitoring and Updation:**  
Regularly monitor the performance of the model in real-world settings.  
Periodically retrain using new data to maximize accuracy.