**Shell-Edunet Skills4Future Internship**

**Project title: Plant Disease Detection System for Sustainable Agriculture**

**1. Data Collection and Loading**

The first step involves collecting the dataset which contains images of plants categorized by disease types or health status.

* **Dataset Structure:**
  + train/
    - category1/
    - category2/
  + test/
    - category1/
    - category2/
  + valid/
    - cat1/
    - cat2/

These directories contain the images used for training, testing, and validation respectively.

**2. Preprocessing and Directory Setup**

* **Compression & Storage:**
  + The dataset is zipped for efficient upload.
  + It is uploaded to **Google Drive**.
* **Colab Integration:**
  + Google Drive is mounted on **Google Colab** using Python code.
  + The dataset is **unzipped** inside Colab for further usage.

**3. Image Processing & Augmentation**

* **Preprocessing:**
  + All input images are resized to the same dimension (e.g., 128x128 pixels) for consistency.
* **Augmentation Techniques:**
  + Rotation, flipping, zooming, brightness adjustment etc., may be applied to increase the size and diversity of the training data.

**4. Building the CNN Model**

* A **Convolutional Neural Network (CNN)** is built for the classification of plant diseases.
* This model processes the images and learns patterns corresponding to healthy or diseased plants.

**5. Training the Model**

* The CNN model is trained using the train dataset.
* The valid dataset is used to monitor and fine-tune the training process.

**6. Testing and Evaluation**

* The model is evaluated using the test dataset.
* Performance metrics like accuracy, precision, recall, and F1 score may be used.

**7. Model Deployment and Operation**

* Once trained and tested, the model is ready for deployment.
* Input plant images (e.g., size 100x100 or 400x400) are first preprocessed to match the input dimension of the CNN (128x128).
* The model predicts the disease or health status.
* The model is integrated into a **Python-based interface** for real-time disease detection operations.