**Plant Disease Detection System for Sustainable Agriculture**

**Problem Statement:**

The detection of plant diseases in agriculture is often delayed and inaccurate due to reliance on traditional methods, leading to significant crop losses, inefficient pesticide use, and environmental harm. There is a need for an efficient, automated system that can accurately identify plant diseases at an early stage, enabling timely intervention and promoting sustainable farming practices.

**Pipeline:**

1. Data Collection and Data Loading
2. Train, Valid, Test of a Model
3. Format of CNN (DATA SET)
4. Zip folder
5. Image Processing and Image Augmentation
6. Build CNN model
7. Test / Evaluate

**1.Data Collection and Data Loading:**

**Data Collection** refers to the process of gathering raw data from various sources for the purpose of analysis. This data set should contain the information of the plants to predict the disease of the plants. We can download the dataset from the Kaggle which contains the dataset of different problem statements. After collecting the data set, we need to load that data into the drive which we need to perform the action.

**Data Loading** refers to the process of transferring or importing the data into any storage device so it can be analyzed or processed. It is used to clean or manipulate the raw data.

Here in our project, we are using the data set and unzip it into the drive and performing the further steps in the google colab.

**2. Train, Valid, Test of a Model:**

Train

**Model**

Valid

Test

The model is trained from the data we are giving, while training the model we validate the data of the model. The model converts the data cleaning, modification of data takes place in validation process. After validation when the model completes learning then the model is tested.

**3. Format of CNN (DATA SET):**

The dataset should include the train, test and valid folders for training the data testing and validation of the model.

1. Train
2. Test
3. Valid

**4.Zip folder:**

1. Upload the zip folder into the google drive.

2. Mount drive on google colab.

3. Python code🡪unzip🡪dataset.

**5. Image Processing and Image Augmentation:**

**Image Processing** refers to the techniques and methods used to manipulate and analyse digital images in order to extract useful information or enhance the image for specific applications. It involves applying algorithms and operations to images to improve quality, enhance features, or prepare images for further analysis or machine learning tasks.

**Image Augmentation** refers to the process of artificially increasing the size of an image dataset by applying various transformations to the original images. This helps to improve the performance and generalization of machine learning models, particularly in tasks like image classification and object detection, by providing more diverse training examples.

**6. Build CNN Model:**

After completing the above process, we need to build the CNN model for the data.

**7.Test / Evaluate Model:**

Testing refers to the phase where the trained model is applied to unseen data (test data) in order to determine how well it performs in real-world scenarios. The main goal is to understand how well the model generalizes to new, unseen data, not just how well it performed on the training set

Evaluation is a broader term that refers to the entire process of assessing how well a model or system performs according to defined objectives. It not only includes testing but also involves analysing the results, interpreting the model's effectiveness, and identifying areas for improvement.