**About Pipeline**

**Pipeline can be defined as the step-by-step procedure in which the model will work. So, according to our problem statement- Plant Disease prediction system for sustainable agriculture, we need to follow a pipeline which was discussed in the class itself. The pipeline is-**

* **Data Collection and Data Loading- In this step we are going to have a dataset for all the plants, which will be divided as Train data, Valid data and Test data. The train data is used to train the model, the valid data is the true values, and the test data is used to test the model.**
* **Uploading of dataset- The dataset is provided which is required to be uploaded in the Google Drive, so that while using Google Colab we can mount the dataset from drive and then can work with the dataset in the code.**
* **Image Processing and Image Augmentation- In this step the image which is uploaded is being verified and processed. The sizes of the images are adjusted and also different angle of the image is checked for the verification.**
* **Use of CNN Model- Then the CNN model is used for the prediction. At first, it checks whether there is a disease or not, if not then the prediction is given otherwise it classifies the disease.**
* **Testing- The last step for validating the predicted answer and the answer is display to the user.**