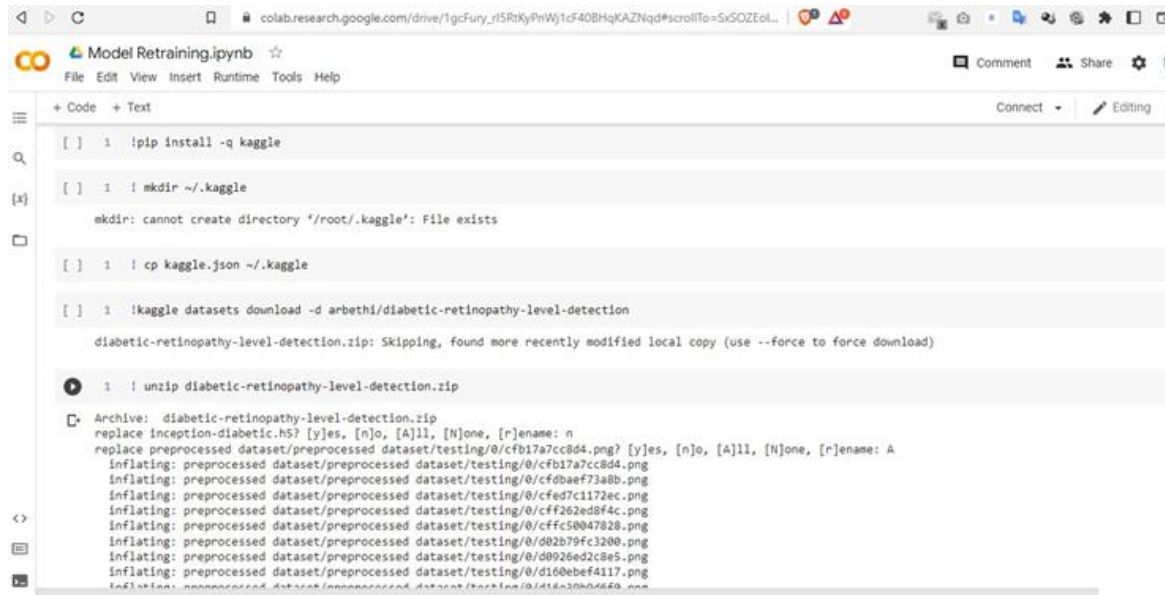


# Dataset Collection

## Dataset Download

The dataset is downloaded from kaggle link :

<https://www.kaggle.com/datasets/arbethi/diabetic-retinopathy-level-detection?select=preprocessed+dataset>



```
[ ] 1 !pip install -q kaggle

[ ] 1 !mkdir ~/.kaggle

mkdir: cannot create directory '/root/.kaggle': File exists

[ ] 1 !cp kaggle.json ~/.kaggle

[ ] 1 !kaggle datasets download -d arbethi/diabetic-retinopathy-level-detection

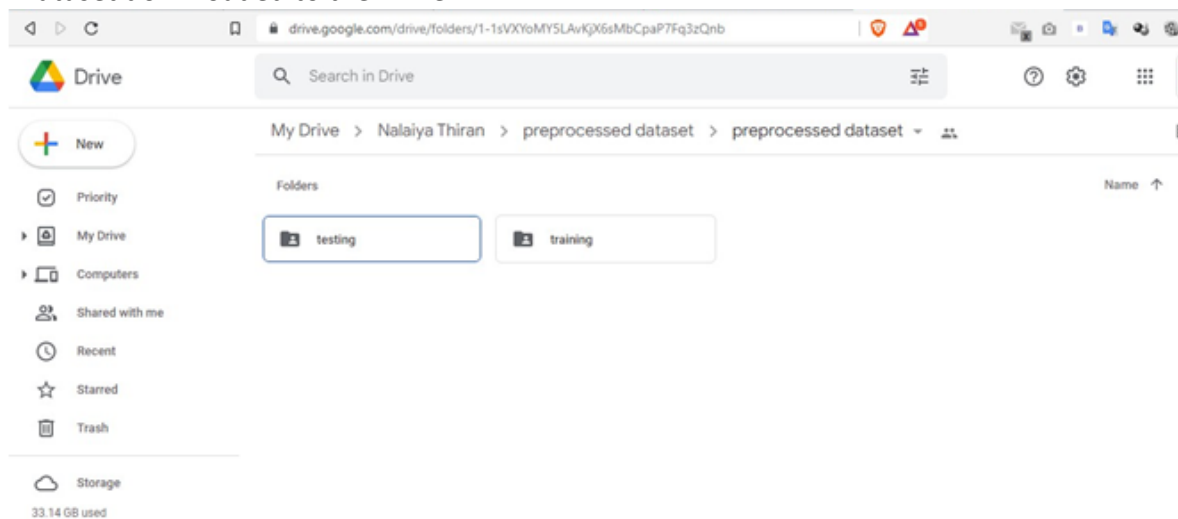
diabetic-retinopathy-level-detection.zip: Skipping, found more recently modified local copy (use --force to force download)

[ ] 1 !unzip diabetic-retinopathy-level-detection.zip

Archive: diabetic-retinopathy-level-detection.zip
  replace inception-diabetic.h5? [y]es, [n]o, [A]ll, [N]one, [r]ename: n
  replace preprocessed dataset/preprocessed dataset/testing/0/cfb17a7cc8d4.png? [y]es, [n]o, [A]ll, [N]one, [r]ename: A
  inflating: preprocessed dataset/preprocessed dataset/testing/0/cfb17a7cc8d4.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/cfb0baef73a8b.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/cfe7c1172ec.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/cff62ed8f4c.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/cffc50047828.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/d02b79fc3200.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/d0926ed2c8e5.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/d160ebef4117.png
  inflating: preprocessed dataset/preprocessed dataset/testing/0/d16a30b046d0.png
```

Kaggle package in Python is installed to collect the dataset from the source.

Dataset downloaded to the Drive



### Create Training and Testing path:

```
[ ] 1  img_size = [299,299]
    2  trainPath = "/preprocessed dataset/preprocessed dataset/training"
    3  testPath = "/preprocessed dataset/preprocessed dataset/testing"
```

To build a DL model we have to split training and testing data into two separate folders. But In the project dataset folder training and testing folders are presented. So, in this case we just have to assign a variable and pass the folder path to it.

Four different transfer learning models are used in our project and the best model (Xception) is selected.

The image input size of xception model is 299, 299.