BDL Lab 8 | Madhur Jindal | ME18B059

1. We start with installing the required dependencies and importing the required packages.
2. Moving on to importing the dataset, we used Keras.datasets for importing the CIFAR 10 dataset.
3. We then create a pandas dataframe with columns - content containing byte array form of the image array, label containing the label. This is then used to create a spark dataframe, using 25 rows of the dataset.
4. We then create a ImageNetDataset class inherited from the Dataset class which bring the data into usable form and performs the required transformations.
5. We then proceed to create a pandas_udf function wrapping a predict function which takes in input as the model instance and uses it to predict the label from input image.
6. The following models along with their accuracies (checked manually) were used:
    a. Mobile Net V2 – (12/25)
    b. Mobile Net V3 Large – (10/25)
    c. GoogLeNet – (8/25)
    d. Efficient Net – (13/25)
    e. VGG 16 – (13/25)
7. Thus we see that the Efficient net and VGG 16 perform the best with Googlenet performing the worst.