Overview of Dataset

We have loaded data set into google drive and used google colab. The dataset contains a total of 12,446 CT kidney images, stored in JPEG format. Each image has a unique identifier and is associated with a label indicating the class of kidney abnormality present in the image. We visualized the dataset and found that the dataset has distribution of images across the four classes, with 5,077 normal images, 3,709 cyst images, 2,283 tumor images, and 1,377 stone images. The data is imbalanced and contains more images of normal kidney. The output can be [0,1,2,3] which signifies [0-Normal ,1-Cyst ,2-Tumor ,3-Stone].

The 'target' column ranges from 0 to 3, with a mean value of 1.18 and a standard deviation of 1.05. The majority of the values fall within the range of 0 to 2, with the median value being 1.0.

The dataset doesn’t contain any null or duplicate values. As a preprocessing step we planned to do normalization and data argumentation to increase the quality of images.

We want to do feature extraction to extract some images from the data. Although, there is data loss, train the model with feature extracted dataset and without feature extraction dataset, want to compare them to evaluate their performance.

Models we want to work with SVM, Random Forest, KNN, Logistic regression, Multi -layerperceptron and CNN.Evaluating the models by using the performance metrics Accuracy and find the model which gives the best accuracy.

We extracted images to the train, test, validation folders so that some images are left to test and validate after training the model is done from the input data set.

When observed about the shape of the images found that they are of different shapes. These differences in image size can potentially impact the performance of a algorithm so did image normalization by using preprocess\_input() function which ensures that the input data is normalized and centered around zero, which can improve the accuracy and stability of the model during training.We also performed data argumentation by considering horizontal\_flip=True.