WORKFLOW FOR THE PROCESS:

1)Downloading the data files and then uploading them to google colab. Then cleaning of data and data visualization was carried out(2nd April-3rd April)

2)Deciding which kind of architecture to work with. (4th April)

3)Planning the architecture and deciding the input shape of the sound files. (5th April)

4)Writing the code for training the dataset and checking on validation data to determine the best possible arrangement of convolutional layers. (6th April-7th April)

5)Finally testing on the test dataset and making required submissions. (7th April)

6)I was planning to increase the efficiency by using resnet and tried to implement confusion matrices for better results. (7th April)

7)I was thinking about more fine-tuning the hyperparameters so that I could get better results. (7th April)

8)I was planning to use this model to construct sentences when a user is speaking, that is, I would give a user speech and my algorithm would recognize the speech. (8th April)

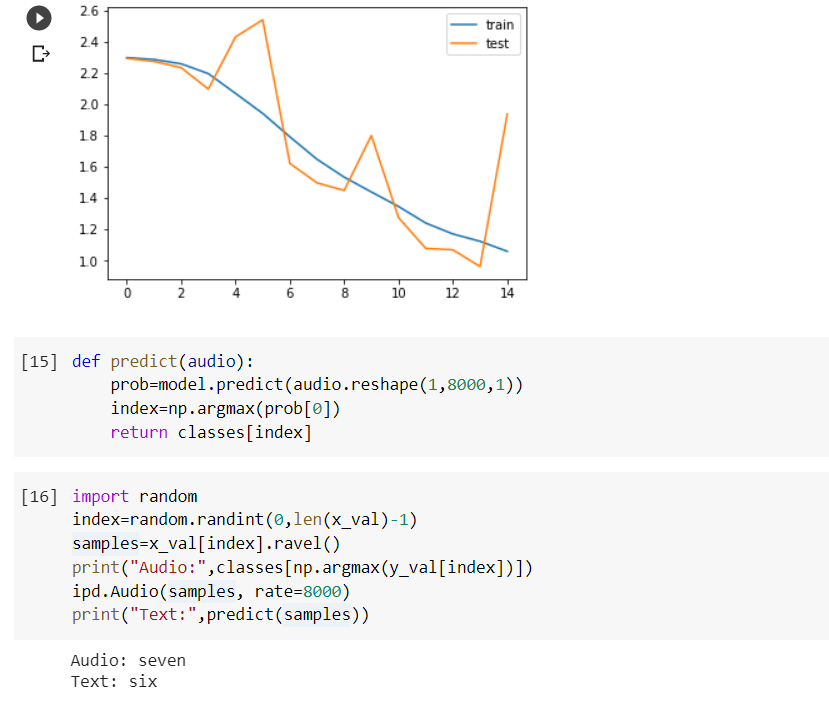
9)I built a basic layout and wrote some code but faced repeated errors, but given time I would work more on it and fix it. (8th April-9th April)

10)I have read several kaggle notebooks and liked the ideas used by many of them. I have tried to implement a few of them for my purpose.

11)I have read several medium articles on speech recognition which gave me insights on what king of CNN architecture might help.

12)Given, below is my work with various convolutional layer architecture. It layouts my motivation for selecting a particular architecture.

13) I tried several combinations and compared the efficiency across several models and finally decided to continue with one.



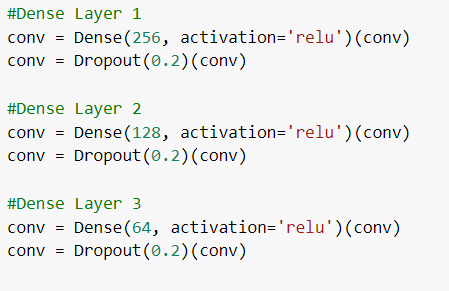
1)Here, I have used 4 convolutional layers and 2 dense layers. This was a purely experimental setup. I started with this and further made modifications in the layers to get various results.

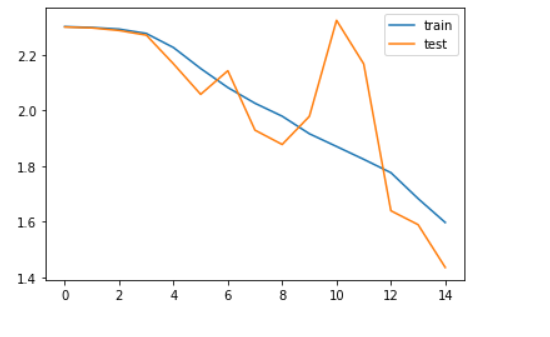
Based on the accuracy curves, I further made changes to get better efficiencies. I have shown the various graphs below:

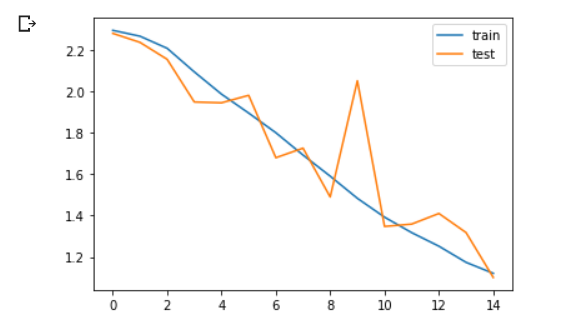
2)In the next figure, I have used six convolutional layers and 3 dense layers.

I have added the structure of convolutional and dense layers below for reference,





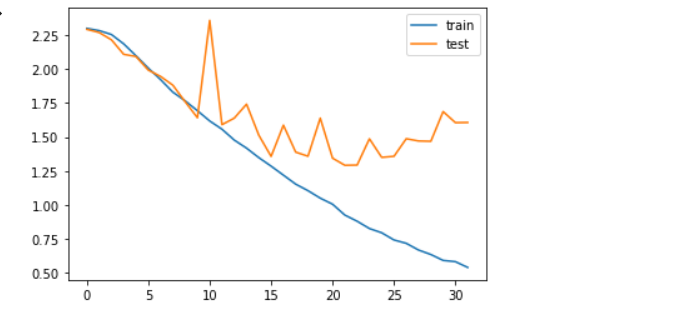




3)Here I used 2 convolutional layers with 3 dense layers.

4)So far I have been changing the number of convolutional and dense layers and had kept the number of iterations fixed to 15.

5)Now using the last combination I will run on 100 iterations, till I find the best model for my purpose.

 6)Iterations stopped at iteration number 32 giving the best accuracies. From the plot, it is visible that validation, as well as test error, has also reduced a lot.

7)Now I have saved the model and will use it to predict the test data.

8)Further, I have tested the given dataset on this model and achieved a score of 0.8019