Spoken Digit Recognition

In this notebook, You will do Spoken Digit Recognition.

Input - speech signal, output - digit number

It contains

1. Reading the dataset. and Preprocess the data set. Detailed instrctions are given below. You have to write the code in the same cell which contains the instrction.
2. Training the LSTM with RAW data
3. Converting to spectrogram and Training the LSTM network
4. Creating the augmented data and doing step 2 and 3 again.

instructions:

1. Don't change any Grader Functions. Don't manipulate any Grader functions. If you manipulate any, it will be considered as plagiarised.

2. Please read the instructions on the code cells and markdown cells. We will explain what to write.

3. please return outputs in the same format what we asked. Eg. Don't return List of we are asking for a numpy array.

4. Please read the external links that we are given so that you will learn the concept behind the code that you are writing.

5. We are giving instructions at each section if necessary, please follow them.

Every Grader function has to return True.

In [4]:
```
import numpy as np
from tensorflow.keras.layers import Input, LSTM, Dense
from sklearn.metrics import roc_auc_score
from sklearn.metrics import f1_score
from tensorflow.keras.models import Model
import tensorflow as tf
import pandas as pd
import librosa
import os
##if you need any imports you can do that here.
```

In [6]:
```
all_files = list()
data_files = os.listdir('recordings')
for i,sub_file in enumerate(data_files):
  if (sub_file.endswith("wav")):
    sub_file_path = 'recordings' + '/' + sub_file
    sub_file_path = str(sub_file_path)
    all_files.append(sub_file_path)
```

Grader function 1

In [8]:
```
def grader_files():
    temp = len(all_files)==2000
    temp1 = all([x[-3:]=="wav" for x in all_files])
    temp = temp and temp1
    return temp
grader_files()
```

Out[8]: True

Create a dataframe(name=df_audio) with two columns(path, label).

You can get the label from the first letter of name.

Eg: 0_jackson_0 --> 0

0_jackson_43 --> 0

In [9]:
```
#Create a dataframe(name=df_audio) with two columns(path, label).
#You can get the label from the first letter of name.
#Eg: 0_jackson_0 --> 0
#0_jackson_43 --> 0
label = [int(x[11]) for x in all_files]
label[0:5]

# intialise data of lists.
data = {'path':all_files,
        'label':label}
```

```
# Create DataFrame
df_audio = pd.DataFrame(data)
df_audio.head(5)
```

Out[9]:

	path	label
0	recordings/7_nicolas_39.wav	7
1	recordings/9_nicolas_32.wav	9
2	recordings/8_theo_9.wav	8
3	recordings/7_theo_20.wav	7
4	recordings/0_nicolas_8.wav	0

In [10]:
```
y = df_audio['label'].values
```

In [ ]:
```
#info
df_audio.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   path    2000 non-null   object
 1   label   2000 non-null   int64
dtypes: int64(1), object(1)
memory usage: 31.4+ KB
```

Grader function 2

In [11]:
```
def grader_df():
    #flag_shape = df_audio.shape==(2000,2)
    flag_columns = all(df_audio.columns==['path', 'label'])
    list_values = list(df_audio.label.value_counts())
    flag_label = len(list_values)==10
    flag_label2 = all([i==200 for i in list_values])
    final_flag = flag_columns and flag_columns and flag_label and flag_label2
    return final_flag
grader_df()
```

Out[11]: True

In [12]:
```
from sklearn.utils import shuffle
df_audio = shuffle(df_audio, random_state=33)#don't change the random state
```

Train and Validation split

In [14]:
```
#split the data into train and validation and save in X_train, X_test, y_train, y_test
#use stratify sampling
#use random state of 45
#use test size of 30%
# train test split
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(df_audio['path'],df_audio['label'] , test_size=0.3, stratify=df_audio['label'])
```

Grader function 3

In [17]:
```
def grader_split():
    flag_len = (len(X_train)==1400) and (len(X_test)==600) and (len(y_train)==1400) and (len(y_test)==600)
    values_ytrain = list(y_train.value_counts())
    flag_ytrain = (len(values_ytrain)==10) and (all([i==140 for i in values_ytrain]))
    values_ytest = list(y_test.value_counts())
    flag_ytest = (len(values_ytest)==10) and (all([i==60 for i in values_ytest]))
    final_flag = flag_len and flag_ytrain and flag_ytest
    return final_flag
grader_split()
```

Out[17]: True

Preprocessing

All files are in the "WAV" format. We will read those raw data files using the librosa

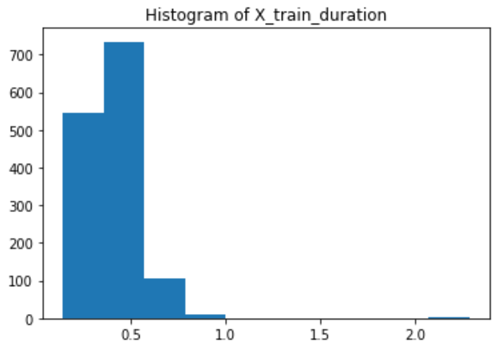
In [18]:
```
sample_rate = 22050
def load_wav(x, get_duration=True):
    '''This return the array values of audio with sampling rate of 22050 and Duration'''
    #loading the wav file with sampling rate of 22050
    samples, sample_rate = librosa.load(x, sr=22050)
    if get_duration:
        duration = librosa.get_duration(samples, sample_rate)
        return [samples, duration]
    else:
        return samples
```

In [19]:
```
X_train_process = [load_wav(row, get_duration=True) for row in X_train.tolist()]
X_test_process = [load_wav(row, get_duration=True) for row in X_test.tolist()]
```

In [ ]:
```
#use load_wav function that was written above to get every wave.
#save it in X_train_processed and X_test_processed
# X_train_processed/X_test_processed should be dataframes with two columns(raw_data, duration) with same index of X_train/y_train
```

In [36]:
```
from matplotlib import pyplot
#plot the histogram of the duration for trian

X_train_duration  = [i[1] for i in X_train_process]
# plot scores
pyplot.hist(X_train_duration)
pyplot.title("Histogram of X_train_duration")
pyplot.show()
```



In [ ]:
```
#print 0 to 100 percentile values with step size of 10 for train data duration.
```

In [32]:
```
import numpy as np
p = [0,10,20,30,40,50,60,70,80,90,100]
range =  np.percentile(X_train_duration, p)

for i , j  in enumerate(range):
  print(f'{p[i]} th percentile is {range[i]}')
```

```
0 th percentile is 0.1435374149659864
10 th percentile is 0.25988208616780045
20 th percentile is 0.30080725623582766
30 th percentile is 0.33424489795918366
40 th percentile is 0.36007256235827667
50 th percentile is 0.3915873015873016
60 th percentile is 0.418639455782313
70 th percentile is 0.44988662131519275
80 th percentile is 0.48596825396825394
90 th percentile is 0.5549160997732426
100 th percentile is 2.282766439909297
```

In [ ]:
```
##print 90 to 100 percentile values with step size of 1.
```

In [33]:
```
import numpy as np
p = [90,91,92,93,94,95,96,97,98,99,100]
range =  np.percentile(X_train_duration, p)

for i , j  in enumerate(range):
  print(f'{p[i]} th percentile is {range[i]}')
```

```
90 th percentile is 0.5549160997732426
91 th percentile is 0.5659854875283448
92 th percentile is 0.5779083900226759
93 th percentile is 0.5933292517006803
94 th percentile is 0.609092970521542
95 th percentile is 0.6231496598639454
96 th percentile is 0.6420553287981859
97 th percentile is 0.6635741496598639
98 th percentile is 0.6956090702947844
99 th percentile is 0.7831392290249433
100 th percentile is 2.282766439909297
```

In [20]:
```
X_trian_raw_data = [i[0] for i in X_train_process]
X_train_duration = [i[1] for i in X_train_process]

dict = {'raw_data':X_trian_raw_data , 'duration':X_train_duration}

X_train_processed = pd.DataFrame(dict)
```

In [21]:
```
X_test_raw_data = [i[0] for i in X_test_process]
X_test_duration = [i[1] for i in X_test_process]

dict = {'raw_data':X_test_raw_data , 'duration':X_test_duration}

X_test_processed = pd.DataFrame(dict)
```

Grader function 4

In [22]:
```
def grader_processed():
    flag_columns = (all(X_train_processed.columns==['raw_data', 'duration'])) and (all(X_test_processed.columns==['raw_data', 'duration']))
    flag_shape = (X_train_processed.shape ==(1400, 2)) and (X_test_processed.shape==(600,2))
    return flag_columns and flag_shape
grader_processed()
```

Out[22]: True

```
Based on our analysis 99 percentile values are less than 0.8sec so we will limit maximum length of X_train_processed
and X_test_processed to 0.8 sec. It is similar to pad_sequence for a text dataset.

While loading the audio files, we are using sampling rate of 22050 so one sec will give array of length 22050. so, our
maximum length is 0.8*22050 = 17640

Pad with Zero if length of sequence is less than 17640 else Truncate the number.

Also create a masking vector for train and test.

masking vector value = 1 if it is real value, 0 if it is pad value. Masking vector data type must be bool.
```

In [23]:
```
X_train_sample = [i[0] for i in X_train_processed]
X_test_sample = [i[0] for i in X_test_processed]
```

In [24]:
```
max_length  = 17640
```

In [ ]:
```
## as discussed above, Pad with Zero if length of sequence is less than 17640 else Truncate the number.
## save in the X_train_pad_seq, X_test_pad_seq
## also Create masking vector X_train_mask, X_test_mask

## all the X_train_pad_seq, X_test_pad_seq, X_train_mask, X_test_mask will be numpy arrays mask vector dtype must be bool.
```

In [25]:
```
X_train_processed.head()
```

Out[25]:

	raw_data	duration
0	[-8.8885805e-05, 0.00012398604, 0.0003388623, ...	0.427029
1	[3.4486034e-06, -4.19734e-05, -8.592559e-05, -...	0.419410
2	[-0.00036025772, 0.0002468019, 0.0006729113, 0...	0.460136
3	[0.00057921023, 0.0002459513, -0.00026787468, ...	0.286259
4	[-0.00022238896, -0.0002804552, -0.0003052361,...	0.156417

In [26]:
```
X_train_sample = X_train_processed['raw_data'].values
X_test_sample = X_test_processed['raw_data'].values
```

In [27]:
```
X_train_sample.shape,X_test_sample.shape
```

Out[27]: ((1400,), (600,))

In [28]:
```
X_train_pad_seq = []
for li in X_train_sample:
  if len(li) < max_length:
    li = li.tolist()
    a = [0]*(max_length - len(li))
    li.extend(a)
    X_train_pad_seq.append(li)

  else:
    li = li.tolist()
    X_train_pad_seq.append(li[0:max_length])

X_train_pad_seq = np.array(X_train_pad_seq)


X_test_pad_seq = []
for li in X_test_sample:
  if len(li) < max_length:
    li = li.tolist()
    a = [0]*(max_length - len(li))
    li.extend(a)
    X_test_pad_seq.append(li)

  else:
    li = li.tolist()
    X_test_pad_seq.append(li[0:max_length])

X_test_pad_seq = np.array(X_test_pad_seq)
```

In [29]:
```
X_train_mask = np.array([(i > 0).tolist()for i in X_train_pad_seq])
X_test_mask = np.array([(i > 0).tolist()for i in X_test_pad_seq])
```

In [30]:
```
X_train_mask.dtype , X_test_mask.dtype
```

Out[30]: (dtype('bool'), dtype('bool'))

Grader function 5

In [33]:
```
def grader_padoutput():
    flag_padshape = (X_train_pad_seq.shape==(1400, 17640)) and (X_test_pad_seq.shape==(600, 17640)) and (y_train.shape==(1400,))
    flag_maskshape = (X_train_mask.shape==(1400, 17640)) and (X_test_mask.shape==(600, 17640)) and (y_test.shape==(600,))
    flag_dtype = (X_train_mask.dtype==bool) and (X_test_mask.dtype==bool)
    return flag_padshape and flag_maskshape and flag_dtype
    return flag_padshape
grader_padoutput()
```

Out[33]: True

In [34]:
```
X_train_mask = X_train_mask.astype('float')
X_test_mask = X_test_mask.astype('float')
```

1. Giving Raw data directly.

```
Now we have

Train data: X_train_pad_seq, X_train_mask and y_train
Test data: X_test_pad_seq, X_test_mask and y_test

We will create a LSTM model which takes this input.

Task:

1. Create an LSTM network which takes "X_train_pad_seq" as input, "X_train_mask" as mask input. You can use any number
of LSTM cells. Please read LSTM documentation(https://www.tensorflow.org/api_docs/python/tf/keras/layers/LSTM) in
tensorflow to know more about mask and also https://www.tensorflow.org/guide/keras/masking_and_padding
2. Get the final output of the LSTM and give it to Dense layer of any size and then give it to Dense layer of size
10(because we have 10 outputs) and then compile with the sparse categorical cross entropy( because we are not
converting it to one hot vectors).
3. Use tensorboard to plot the graphs of loss and metric(use micro F1 score as metric) and histograms of gradients.
4. make sure that it won't overfit.
5. You are free to include any regularization
```

In [36]:
```
X_train_pad_seq.shape[1]
```

Out[36]: 17640

In [37]:
```
Y_train = tf.keras.utils.to_categorical(y_train, 10)
Y_test = tf.keras.utils.to_categorical(y_test, 10)
```

MODEL :- 1

In [38]:
```
## as discussed above, please write the LSTM

lstm = LSTM(units = 25,activation="tanh",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))

input_layer = Input(shape=(X_train_pad_seq.shape[1],1),dtype=float )
input_mask = Input(shape=(X_train_mask.shape[1],1),dtype=bool)
LSTM_layer  = lstm(inputs=input_layer,mask=input_mask)
dense = Dense(50,activation="relu",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))(LSTM_layer)
output_1 = Dense(10,activation='softmax',kernel_initializer=tf.keras.initializers.GlorotUniform(seed=0))(dense)

model = Model(inputs = [input_layer,input_mask],outputs = output_1)
model.compile(optimizer='adam',loss = tf.keras.losses.sparse_categorical_crossentropy,metrics='accuracy')
model.summary()
```

```
Model: "model"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
input_1 (InputLayer)            [(None, 17640, 1)]   0
__________________________________________________________________________________________________
input_2 (InputLayer)            [(None, 17640, 1)]   0
__________________________________________________________________________________________________
lstm (LSTM)                     (None, 25)           2700        input_1[0][0]
                                                                 input_2[0][0]
__________________________________________________________________________________________________
dense (Dense)                   (None, 50)           1300        lstm[0][0]
__________________________________________________________________________________________________
dense_1 (Dense)                 (None, 10)           510         dense[0][0]
==================================================================================================
Total params: 4,510
Trainable params: 4,510
Non-trainable params: 0
__________________________________________________________________________________________________
```

In [39]:
```
class f1_score_and_auc_Callback(tf.keras.callbacks.Callback):
```

```
    def  on_train_begin(self,logs={}):
      self.f1_micro=[]
      self.auc_score=[]

    def on_epoch_end(self, epoch, logs=None):
      y_pred=self.model.predict([X_test_pad_seq,X_test_mask])
      y_pred = np.argmax(y_pred, axis = 1)

      y_true=y_test
      score=f1_score(y_true, y_pred, average='micro')

      self.f1_micro.append(score)
      print(" F1 micro :",score)

metrics=f1_score_and_auc_Callback()
```

In [40]:

```
model.fit(x=[X_train_pad_seq,X_train_mask],y=y_train,validation_data=([X_test_pad_seq,X_test_mask],y_test),
            epochs=5,batch_size=10,steps_per_epoch=len(X_train_mask)//10 , callbacks=metrics)
```

```
Epoch 1/5
140/140 [==============================] - 63s 227ms/step - loss: 2.3092 - accuracy: 0.1023 - val_loss: 2.3015 - val_accuracy: 0.0917
 F1 micro : 0.10499999999999998
Epoch 2/5
140/140 [==============================] - 30s 217ms/step - loss: 2.2991 - accuracy: 0.0888 - val_loss: 2.2991 - val_accuracy: 0.1400
 F1 micro : 0.115
Epoch 3/5
140/140 [==============================] - 31s 218ms/step - loss: 2.2981 - accuracy: 0.1119 - val_loss: 2.2984 - val_accuracy: 0.1100
 F1 micro : 0.095
Epoch 4/5
140/140 [==============================] - 30s 217ms/step - loss: 2.2939 - accuracy: 0.1097 - val_loss: 2.2948 - val_accuracy: 0.1317
 F1 micro : 0.10333333333333333
Epoch 5/5
140/140 [==============================] - 30s 218ms/step - loss: 2.2918 - accuracy: 0.1166 - val_loss: 2.2915 - val_accuracy: 0.1250
 F1 micro : 0.09333333333333334
```

Out[40]: `<tensorflow.python.keras.callbacks.History at 0x7fa4f57bf7d0>`

2. Converting into spectrogram and giving spectrogram data as input

```
We can use librosa to convert raw data into spectrogram. A spectrogram shows the features in a two-dimensional
representation with the
intensity of a frequency at a point in time i.e we are converting Time domain to frequency domain. you can read more
about this in https://pnsn.org/spectrograms/what-is-a-spectrogram
```

In [41]:

```
def convert_to_spectrogram(raw_data):
    '''converting to spectrogram'''
    spectrum = librosa.feature.melspectrogram(y=raw_data, sr=sample_rate, n_mels=64)
    logmel_spectrum = librosa.power_to_db(S=spectrum, ref=np.max)
    return logmel_spectrum
```

In [42]:

```
X_train_spectrogram = [convert_to_spectrogram(row) for row in X_train_pad_seq]
X_train_spectrogram = np.array(X_train_spectrogram)

X_test_spectrogram = [convert_to_spectrogram(row) for row in X_test_pad_seq]
X_test_spectrogram = np.array(X_test_spectrogram)
```

In [ ]:

```
##use convert_to_spectrogram and convert every raw sequence in X_train_pad_seq and X_test_pad-seq.
## save those all in the X_train_spectrogram and X_test_spectrogram ( These two arrays must be numpy arrays)
```

Grader function 6

In [44]:

```
def grader_spectrogram():
    flag_shape = (X_train_spectrogram.shape==(1400,64, 35)) and (X_test_spectrogram.shape == (600, 64, 35))
    return flag_shape
grader_spectrogram()
```

Out[44]: True

MODEL :- 2

In [45]:

```
## as discussed above, please write the LSTM

lstm = LSTM(units = 100,activation="tanh",kernel_initializer=tf.keras.initializers.he_uniform(seed=0),return_sequences=True)

input_layer = Input(shape=(X_test_spectrogram[0].shape),dtype=float)
LSTM_layer  = lstm(inputs=input_layer)
glo_avg = tf.keras.layers.GlobalAveragePooling1D()(LSTM_layer)
dense = Dense(50,activation="relu",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))(glo_avg)
output_1 = Dense(10,activation='softmax',kernel_initializer=tf.keras.initializers.GlorotUniform(seed=0))(dense)

model = Model(inputs = [input_layer],outputs = output_1)
model.compile(optimizer='adam',loss = tf.keras.losses.sparse_categorical_crossentropy,metrics='accuracy')
model.summary()
```

```
Model: "model_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
```

```
input_3 (InputLayer)         [(None, 64, 35)]          0
_________________________________________________________________
lstm_1 (LSTM)                (None, 64, 100)           54400
_________________________________________________________________
global_average_pooling1d (Gl (None, 100)               0
_________________________________________________________________
dense_2 (Dense)              (None, 50)                5050
_________________________________________________________________
dense_3 (Dense)              (None, 10)                510
=================================================================
Total params: 59,960
Trainable params: 59,960
Non-trainable params: 0
_________________________________________________________________
```

In [47]:

```
X_train_spectrogram.shape , X_test_spectrogram.shape ,y_train.shape,y_test.shape
```

Out[47]: ((1400, 64, 35), (600, 64, 35), (1400,), (600,))

In [50]:

```
model.fit(x=X_train_spectrogram,y=y_train,validation_data=(X_test_spectrogram,y_test),
            epochs=60,batch_size=10,steps_per_epoch=len(X_train_spectrogram)//10, callbacks=metrics)
```

```
Epoch 1/60
140/140 [==============================] - 2s 8ms/step - loss: 2.3180 - accuracy: 0.1404 - val_loss: 2.0487 - val_accuracy: 0.3483
 F1 micro : 0.34833333333333333
Epoch 2/60
140/140 [==============================] - 1s 5ms/step - loss: 2.0085 - accuracy: 0.3272 - val_loss: 1.8431 - val_accuracy: 0.3800
 F1 micro : 0.38
Epoch 3/60
140/140 [==============================] - 1s 5ms/step - loss: 1.8225 - accuracy: 0.4102 - val_loss: 1.7608 - val_accuracy: 0.4100
 F1 micro : 0.41
Epoch 4/60
140/140 [==============================] - 1s 5ms/step - loss: 1.6993 - accuracy: 0.4173 - val_loss: 1.5852 - val_accuracy: 0.4633
 F1 micro : 0.4633333333333333
Epoch 5/60
140/140 [==============================] - 1s 5ms/step - loss: 1.5721 - accuracy: 0.4665 - val_loss: 1.5581 - val_accuracy: 0.4433
 F1 micro : 0.44333333333333336
Epoch 6/60
140/140 [==============================] - 1s 5ms/step - loss: 1.5315 - accuracy: 0.4791 - val_loss: 1.4707 - val_accuracy: 0.4817
 F1 micro : 0.4816666666666667
Epoch 7/60
140/140 [==============================] - 1s 5ms/step - loss: 1.4253 - accuracy: 0.4983 - val_loss: 1.4195 - val_accuracy: 0.5250
 F1 micro : 0.525
Epoch 8/60
140/140 [==============================] - 1s 5ms/step - loss: 1.3464 - accuracy: 0.5494 - val_loss: 1.2990 - val_accuracy: 0.5583
 F1 micro : 0.5583333333333333
Epoch 9/60
140/140 [==============================] - 1s 6ms/step - loss: 1.3082 - accuracy: 0.5603 - val_loss: 1.2673 - val_accuracy: 0.5950
 F1 micro : 0.595
Epoch 10/60
140/140 [==============================] - 1s 5ms/step - loss: 1.2744 - accuracy: 0.5833 - val_loss: 1.2807 - val_accuracy: 0.5750
 F1 micro : 0.575
Epoch 11/60
140/140 [==============================] - 1s 5ms/step - loss: 1.1988 - accuracy: 0.6096 - val_loss: 1.2071 - val_accuracy: 0.6217
 F1 micro : 0.6216666666666667
Epoch 12/60
140/140 [==============================] - 1s 5ms/step - loss: 1.1114 - accuracy: 0.6225 - val_loss: 1.1720 - val_accuracy: 0.6100
 F1 micro : 0.61
Epoch 13/60
140/140 [==============================] - 1s 5ms/step - loss: 1.1402 - accuracy: 0.6309 - val_loss: 1.1122 - val_accuracy: 0.6450
 F1 micro : 0.645
Epoch 14/60
140/140 [==============================] - 1s 5ms/step - loss: 1.0350 - accuracy: 0.6601 - val_loss: 1.1288 - val_accuracy: 0.6333
 F1 micro : 0.6333333333333333
Epoch 15/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9975 - accuracy: 0.6840 - val_loss: 1.0192 - val_accuracy: 0.6650
 F1 micro : 0.665
Epoch 16/60
140/140 [==============================] - 1s 5ms/step - loss: 1.0196 - accuracy: 0.6736 - val_loss: 1.0432 - val_accuracy: 0.6883
 F1 micro : 0.6883333333333334
Epoch 17/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9571 - accuracy: 0.7004 - val_loss: 0.9979 - val_accuracy: 0.6517
 F1 micro : 0.6516666666666666
Epoch 18/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9421 - accuracy: 0.6942 - val_loss: 0.9802 - val_accuracy: 0.6700
 F1 micro : 0.67
Epoch 19/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9200 - accuracy: 0.6860 - val_loss: 0.9754 - val_accuracy: 0.6633
 F1 micro : 0.6633333333333333
Epoch 20/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9642 - accuracy: 0.6923 - val_loss: 0.9786 - val_accuracy: 0.6950
 F1 micro : 0.695
Epoch 21/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9103 - accuracy: 0.7146 - val_loss: 0.9158 - val_accuracy: 0.7133
 F1 micro : 0.7133333333333335
Epoch 22/60
140/140 [==============================] - 1s 5ms/step - loss: 0.8507 - accuracy: 0.7205 - val_loss: 0.9479 - val_accuracy: 0.6767
 F1 micro : 0.6766666666666666
Epoch 23/60
140/140 [==============================] - 1s 5ms/step - loss: 0.9133 - accuracy: 0.6884 - val_loss: 1.0044 - val_accuracy: 0.6467
 F1 micro : 0.6466666666666666
Epoch 24/60
140/140 [==============================] - 1s 5ms/step - loss: 0.8296 - accuracy: 0.7205 - val_loss: 1.0393 - val_accuracy: 0.6583
 F1 micro : 0.6583333333333333
Epoch 25/60
140/140 [==============================] - 1s 5ms/step - loss: 0.8318 - accuracy: 0.6998 - val_loss: 0.8756 - val_accuracy: 0.7217
 F1 micro : 0.7216666666666668
```

```
Epoch 26/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7740 - accuracy: 0.7452 - val_loss: 0.8733 - val_accuracy: 0.7217
 F1 micro : 0.7216666666666668
Epoch 27/60
140/140 [==============================] - 1s 5ms/step - loss: 0.8224 - accuracy: 0.7007 - val_loss: 0.8673 - val_accuracy: 0.7300
 F1 micro : 0.7299999999999999
Epoch 28/60
140/140 [==============================] - 1s 5ms/step - loss: 0.8227 - accuracy: 0.7343 - val_loss: 0.9158 - val_accuracy: 0.6950
 F1 micro : 0.695
Epoch 29/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7489 - accuracy: 0.7591 - val_loss: 0.8509 - val_accuracy: 0.7333
 F1 micro : 0.7333333333333333
Epoch 30/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7242 - accuracy: 0.7606 - val_loss: 0.8522 - val_accuracy: 0.7183
 F1 micro : 0.7183333333333334
Epoch 31/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7573 - accuracy: 0.7537 - val_loss: 0.9058 - val_accuracy: 0.6950
 F1 micro : 0.695
Epoch 32/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7342 - accuracy: 0.7519 - val_loss: 0.8327 - val_accuracy: 0.7400
 F1 micro : 0.74
Epoch 33/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6316 - accuracy: 0.8089 - val_loss: 0.8992 - val_accuracy: 0.7200
 F1 micro : 0.72
Epoch 34/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7447 - accuracy: 0.7324 - val_loss: 0.8749 - val_accuracy: 0.7167
 F1 micro : 0.7166666666666667
Epoch 35/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6652 - accuracy: 0.7891 - val_loss: 0.7466 - val_accuracy: 0.7533
 F1 micro : 0.7533333333333333
Epoch 36/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6539 - accuracy: 0.7832 - val_loss: 0.7692 - val_accuracy: 0.7567
 F1 micro : 0.7566666666666667
Epoch 37/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5808 - accuracy: 0.8124 - val_loss: 0.7494 - val_accuracy: 0.7783
 F1 micro : 0.7783333333333333
Epoch 38/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6691 - accuracy: 0.7614 - val_loss: 0.7841 - val_accuracy: 0.7450
 F1 micro : 0.745
Epoch 39/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6618 - accuracy: 0.7777 - val_loss: 0.8492 - val_accuracy: 0.7117
 F1 micro : 0.7116666666666667
Epoch 40/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6941 - accuracy: 0.7603 - val_loss: 0.8244 - val_accuracy: 0.7300
 F1 micro : 0.7299999999999999
Epoch 41/60
140/140 [==============================] - 1s 5ms/step - loss: 0.7014 - accuracy: 0.7590 - val_loss: 0.7705 - val_accuracy: 0.7550
 F1 micro : 0.755
Epoch 42/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6046 - accuracy: 0.7901 - val_loss: 0.7908 - val_accuracy: 0.7450
 F1 micro : 0.745
Epoch 43/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6440 - accuracy: 0.7858 - val_loss: 0.7527 - val_accuracy: 0.7583
 F1 micro : 0.7583333333333333
Epoch 44/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5902 - accuracy: 0.7969 - val_loss: 0.6963 - val_accuracy: 0.7767
 F1 micro : 0.7766666666666666
Epoch 45/60
140/140 [==============================] - 1s 5ms/step - loss: 0.6025 - accuracy: 0.7965 - val_loss: 0.7241 - val_accuracy: 0.7700
 F1 micro : 0.7699999999999999
Epoch 46/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5739 - accuracy: 0.8191 - val_loss: 0.7043 - val_accuracy: 0.7767
 F1 micro : 0.7766666666666666
Epoch 47/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5780 - accuracy: 0.8004 - val_loss: 0.6934 - val_accuracy: 0.7950
 F1 micro : 0.795
Epoch 48/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5650 - accuracy: 0.7988 - val_loss: 0.7111 - val_accuracy: 0.7567
 F1 micro : 0.7566666666666667
Epoch 49/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5743 - accuracy: 0.8041 - val_loss: 0.7490 - val_accuracy: 0.7800
 F1 micro : 0.78
Epoch 50/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5596 - accuracy: 0.8127 - val_loss: 0.7667 - val_accuracy: 0.7517
 F1 micro : 0.7516666666666667
Epoch 51/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5679 - accuracy: 0.8108 - val_loss: 0.6756 - val_accuracy: 0.8033
 F1 micro : 0.8033333333333333
Epoch 52/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5520 - accuracy: 0.8093 - val_loss: 0.7182 - val_accuracy: 0.7700
 F1 micro : 0.7699999999999999
Epoch 53/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5380 - accuracy: 0.8156 - val_loss: 0.7209 - val_accuracy: 0.7700
 F1 micro : 0.7699999999999999
Epoch 54/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5252 - accuracy: 0.8342 - val_loss: 0.6695 - val_accuracy: 0.7817
 F1 micro : 0.7816666666666666
Epoch 55/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5133 - accuracy: 0.8361 - val_loss: 0.6721 - val_accuracy: 0.7850
 F1 micro : 0.785
Epoch 56/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5073 - accuracy: 0.8239 - val_loss: 0.7259 - val_accuracy: 0.7667
 F1 micro : 0.7666666666666667
Epoch 57/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5239 - accuracy: 0.8192 - val_loss: 0.7132 - val_accuracy: 0.7567
 F1 micro : 0.7566666666666667
Epoch 58/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5107 - accuracy: 0.8183 - val_loss: 0.7335 - val_accuracy: 0.7583
```

```
 F1 micro : 0.7583333333333333
Epoch 59/60
140/140 [==============================] - 1s 5ms/step - loss: 0.5017 - accuracy: 0.8294 - val_loss: 0.6552 - val_accuracy: 0.7900
 F1 micro : 0.79
Epoch 60/60
140/140 [==============================] - 1s 5ms/step - loss: 0.4873 - accuracy: 0.8293 - val_loss: 0.6485 - val_accuracy: 0.7983
 F1 micro : 0.7983333333333333
```

Out[50]:
```
<tensorflow.python.keras.callbacks.History at 0x7fa4f23f5250>
```

```
Now we have

Train data: X_train_spectrogram and y_train
Test data: X_test_spectrogram and y_test

We will create a LSTM model which takes this input.

Task:

1. Create an LSTM network which takes "X_train_spectrogram" as input and has to return output at every time step.
2. Average the output of every time step and give this to the Dense layer of any size.
(ex: Output from LSTM will be  (#., time_steps, features) average the output of every time step i.e, you should get
(#.,time_steps)
and then pass to dense layer )
3. give the above output to Dense layer of size 10( output layer) and train the network with sparse categorical cross
entropy.
4. Use tensorboard to plot the graphs of loss and metric(use micro F1 score as metric) and histograms of gradients.
5. make sure that it won't overfit.
6. You are free to include any regularization
```

3. data augmentation

```
Till now we have done with 2000 samples only. It is very less data. We are giving the process of generating augmented
data below.

There are two types of augmentation:
1. time stretching - Time stretching either increases or decreases the length of the file. For time stretching we move
the file 30% faster or slower
2. pitch shifting - pitch shifting moves the frequencies higher or lower. For pitch shifting we shift up or down one
half-step.
```

In [51]:
```
## generating augmented data.
def generate_augmented_data(file_path):
    augmented_data = []
    samples = load_wav(file_path,get_duration=False)
    for time_value in [0.7, 1, 1.3]:
        for pitch_value in [-1, 0, 1]:
            time_stretch_data = librosa.effects.time_stretch(samples, rate=time_value)
            final_data = librosa.effects.pitch_shift(time_stretch_data, sr=sample_rate, n_steps=pitch_value)
            augmented_data.append(final_data)
    return augmented_data
```

In [54]:
```
x_train_data_aug = []

for path in X_train:
  aug_temp = generate_augmented_data(path)
  x_train_data_aug.extend(aug_temp)

x_train_data_aug = np.array(x_train_data_aug)

y_train_data_aug = []

for i in y_train:
  c = [i]*9
  y_train_data_aug.extend(c)

y_train_data_aug = np.array(y_train_data_aug)
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:7: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequenc
es (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you
must specify 'dtype=object' when creating the ndarray
  import sys
```

In [55]:
```
X_train_pad_seq1 = []
for li in x_train_data_aug:
  if len(li) < max_length:
    li = li.tolist()
    a = [0]*(max_length - len(li))
    li.extend(a)
    X_train_pad_seq1.append(li)

  else:
    li = li.tolist()
    X_train_pad_seq1.append(li[0:max_length])

X_train_pad_seq1 = np.array(X_train_pad_seq1)
```

In [56]:
```
X_train_mask1 = np.array([(i > 0).tolist()for i in X_train_pad_seq1])
```

MODEL :- 3

In [57]:
```
## as discussed above, please write the LSTM

lstm = LSTM(units = 100,activation="tanh",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))

input_layer = Input(shape=(X_train_pad_seq.shape[1],1),dtype=float)
input_mask = Input(shape=(X_train_mask.shape[1],1),dtype=bool)
LSTM_layer  = lstm(inputs=input_layer,mask=input_mask)
dense = Dense(50,activation="relu",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))(LSTM_layer)
output_1 = Dense(10,activation='softmax',kernel_initializer=tf.keras.initializers.GlorotUniform(seed=0))(dense)

model = Model(inputs = [input_layer,input_mask],outputs = output_1)
model.compile(optimizer='adam',loss = tf.keras.losses.sparse_categorical_crossentropy,metrics='accuracy')
model.summary()
```

```
Model: "model_2"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
input_4 (InputLayer)            [(None, 17640, 1)]   0
__________________________________________________________________________________________________
input_5 (InputLayer)            [(None, 17640, 1)]   0
__________________________________________________________________________________________________
lstm_2 (LSTM)                   (None, 100)          40800       input_4[0][0]
                                                                 input_5[0][0]
__________________________________________________________________________________________________
dense_4 (Dense)                 (None, 50)           5050        lstm_2[0][0]
__________________________________________________________________________________________________
dense_5 (Dense)                 (None, 10)           510         dense_4[0][0]
==================================================================================================
Total params: 46,360
Trainable params: 46,360
Non-trainable params: 0
__________________________________________________________________________________________________
```

In [58]:
```
X_train_pad_seq1.shape, X_train_mask1.shape, y_train_data_aug.shape
```

Out[58]:
```
((12600, 17640), (12600, 17640), (12600,))
```

In [59]:
```
X_test_pad_seq.shape,X_test_mask.shape,y_test.shape
```

Out[59]:
```
((600, 17640), (600, 17640), (600,))
```

In [ ]:
```
model.fit(x=[X_train_pad_seq1,X_train_mask1],y=y_train_data_aug,validation_data=([X_test_pad_seq,X_test_mask],y_test),
            epochs=2,batch_size=10,steps_per_epoch=len(X_train_mask)//10 , callbacks=metrics)
```

```
Epoch 1/2
140/140 [==============================] - 37s 249ms/step - loss: 2.3044 - accuracy: 0.1021 - val_loss: 2.3211 - val_accuracy: 0.0933
 F1 micro : 0.10000000000000002
Epoch 2/2
140/140 [==============================] - 34s 246ms/step - loss: 2.3003 - accuracy: 0.1071 - val_loss: 2.3110 - val_accuracy: 0.1167
 F1 micro : 0.10000000000000002
```

Out[ ]:
```
<tensorflow.python.keras.callbacks.History at 0x7fd90c6d1c50>
```

As discussed above, for one data point, we will get 9 augmented data points.

Split data into train and test (80-20 split)

We have 2000 data points(1600 train points, 400 test points)

Do augmentation only on train data, after augmentation we will get 14400 train points.

do the above steps i.e training with raw data and spectrogram data with augmentation.

MODEL :- 4

In [61]:
```
X_train_spectrogram1 = [convert_to_spectrogram(row) for row in X_train_pad_seq1]
X_train_spectrogram1 = np.array(X_train_spectrogram1)

X_test_spectrogram = [convert_to_spectrogram(row) for row in X_test_pad_seq]
X_test_spectrogram = np.array(X_test_spectrogram)
```

In [62]:
```
## as discussed above, please write the LSTM

lstm = LSTM(units = 100,activation="tanh",kernel_initializer=tf.keras.initializers.he_uniform(seed=0),return_sequences=True)

input_layer = Input(shape=(X_test_spectrogram[0].shape),dtype=float)
LSTM_layer  = lstm(inputs=input_layer)
glo_avg = tf.keras.layers.GlobalAveragePooling1D()(LSTM_layer)
dense = Dense(50,activation="relu",kernel_initializer=tf.keras.initializers.he_uniform(seed=0))(glo_avg)
output_1 = Dense(10,activation='softmax',kernel_initializer=tf.keras.initializers.GlorotUniform(seed=0))(dense)

model = Model(inputs = [input_layer],outputs = output_1)
```

```
model.compile(optimizer='adam',loss = tf.keras.losses.sparse_categorical_crossentropy,metrics='accuracy')
model.summary()
```

```
Model: "model_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
input_6 (InputLayer)         [(None, 64, 35)]          0         
_________________________________________________________________
lstm_3 (LSTM)                (None, 64, 100)           54400     
_________________________________________________________________
global_average_pooling1d_1 ( (None, 100)               0         
_________________________________________________________________
dense_6 (Dense)              (None, 50)                5050      
_________________________________________________________________
dense_7 (Dense)              (None, 10)                510       
=================================================================
Total params: 59,960
Trainable params: 59,960
Non-trainable params: 0
_________________________________________________________________
```

In [63]:
```
X_train_spectrogram1.shape , X_test_spectrogram.shape ,y_train_data_aug.shape,y_test.shape
```

Out[63]:
```
((12600, 64, 35), (600, 64, 35), (12600,), (600,))
```

In [65]:
```
model.fit(x=X_train_spectrogram1,y=y_train_data_aug,validation_data=(X_test_spectrogram,y_test),
            epochs=60,batch_size=10,steps_per_epoch=len(X_train_spectrogram1)//10, callbacks=metrics)
```

```
Epoch 1/60
1260/1260 [==============================] - 7s 5ms/step - loss: 2.0855 - accuracy: 0.2457 - val_loss: 1.5155 - val_accuracy: 0.4583
 F1 micro : 0.4583333333333333
Epoch 2/60
1260/1260 [==============================] - 5s 4ms/step - loss: 1.5582 - accuracy: 0.4557 - val_loss: 1.2428 - val_accuracy: 0.5833
 F1 micro : 0.5833333333333334
Epoch 3/60
1260/1260 [==============================] - 5s 4ms/step - loss: 1.2903 - accuracy: 0.5446 - val_loss: 1.0119 - val_accuracy: 0.6450
 F1 micro : 0.645
Epoch 4/60
1260/1260 [==============================] - 5s 4ms/step - loss: 1.1230 - accuracy: 0.6023 - val_loss: 0.9224 - val_accuracy: 0.6883
 F1 micro : 0.6883333333333334
Epoch 5/60
1260/1260 [==============================] - 5s 4ms/step - loss: 1.0422 - accuracy: 0.6267 - val_loss: 0.9214 - val_accuracy: 0.6533
 F1 micro : 0.6533333333333333
Epoch 6/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.9522 - accuracy: 0.6666 - val_loss: 0.7895 - val_accuracy: 0.7150
 F1 micro : 0.715
Epoch 7/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.9025 - accuracy: 0.6783 - val_loss: 0.7251 - val_accuracy: 0.7550
 F1 micro : 0.755
Epoch 8/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.8666 - accuracy: 0.6966 - val_loss: 0.8019 - val_accuracy: 0.7250
 F1 micro : 0.7250000000000001
Epoch 9/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.8373 - accuracy: 0.7045 - val_loss: 0.8586 - val_accuracy: 0.6917
 F1 micro : 0.6916666666666667
Epoch 10/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.8224 - accuracy: 0.7088 - val_loss: 0.8406 - val_accuracy: 0.7000
 F1 micro : 0.7
Epoch 11/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.7888 - accuracy: 0.7188 - val_loss: 0.6461 - val_accuracy: 0.7600
 F1 micro : 0.76
Epoch 12/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.7530 - accuracy: 0.7380 - val_loss: 0.7026 - val_accuracy: 0.7300
 F1 micro : 0.7299999999999999
Epoch 13/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.7452 - accuracy: 0.7375 - val_loss: 0.6318 - val_accuracy: 0.7783
 F1 micro : 0.7783333333333333
Epoch 14/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.7103 - accuracy: 0.7460 - val_loss: 0.6584 - val_accuracy: 0.7700
 F1 micro : 0.7699999999999999
Epoch 15/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6912 - accuracy: 0.7605 - val_loss: 0.6429 - val_accuracy: 0.7733
 F1 micro : 0.7733333333333333
Epoch 16/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6919 - accuracy: 0.7479 - val_loss: 0.6890 - val_accuracy: 0.7600
 F1 micro : 0.76
Epoch 17/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6716 - accuracy: 0.7659 - val_loss: 0.6571 - val_accuracy: 0.7517
 F1 micro : 0.7516666666666667
Epoch 18/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6319 - accuracy: 0.7726 - val_loss: 0.5319 - val_accuracy: 0.8317
 F1 micro : 0.8316666666666667
Epoch 19/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6001 - accuracy: 0.7868 - val_loss: 0.5885 - val_accuracy: 0.7867
 F1 micro : 0.7866666666666666
Epoch 20/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.6089 - accuracy: 0.7851 - val_loss: 0.5743 - val_accuracy: 0.7833
 F1 micro : 0.7833333333333333
Epoch 21/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5938 - accuracy: 0.7907 - val_loss: 0.6104 - val_accuracy: 0.7800
 F1 micro : 0.78
Epoch 22/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5948 - accuracy: 0.7931 - val_loss: 0.6007 - val_accuracy: 0.7900
 F1 micro : 0.79
Epoch 23/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5758 - accuracy: 0.7937 - val_loss: 0.5304 - val_accuracy: 0.8217
```

```
 F1 micro : 0.8216666666666665
Epoch 24/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5898 - accuracy: 0.7930 - val_loss: 0.5519 - val_accuracy: 0.7983
 F1 micro : 0.7983333333333333
Epoch 25/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5295 - accuracy: 0.8160 - val_loss: 0.5145 - val_accuracy: 0.8167
 F1 micro : 0.8166666666666667
Epoch 26/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5089 - accuracy: 0.8204 - val_loss: 0.5805 - val_accuracy: 0.7933
 F1 micro : 0.7933333333333333
Epoch 27/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5366 - accuracy: 0.8127 - val_loss: 0.5055 - val_accuracy: 0.8250
 F1 micro : 0.825
Epoch 28/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5215 - accuracy: 0.8168 - val_loss: 0.5169 - val_accuracy: 0.8217
 F1 micro : 0.8216666666666665
Epoch 29/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5099 - accuracy: 0.8148 - val_loss: 0.5228 - val_accuracy: 0.8100
 F1 micro : 0.81
Epoch 30/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4917 - accuracy: 0.8275 - val_loss: 0.5667 - val_accuracy: 0.8000
 F1 micro : 0.8000000000000002
Epoch 31/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.5039 - accuracy: 0.8186 - val_loss: 0.5312 - val_accuracy: 0.8083
 F1 micro : 0.8083333333333333
Epoch 32/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4862 - accuracy: 0.8289 - val_loss: 0.4811 - val_accuracy: 0.8400
 F1 micro : 0.8399999999999999
Epoch 33/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4817 - accuracy: 0.8289 - val_loss: 0.4957 - val_accuracy: 0.8167
 F1 micro : 0.8166666666666667
Epoch 34/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4720 - accuracy: 0.8272 - val_loss: 0.6818 - val_accuracy: 0.7817
 F1 micro : 0.7816666666666666
Epoch 35/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4751 - accuracy: 0.8297 - val_loss: 0.4849 - val_accuracy: 0.8250
 F1 micro : 0.825
Epoch 36/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4349 - accuracy: 0.8449 - val_loss: 0.4520 - val_accuracy: 0.8333
 F1 micro : 0.8333333333333334
Epoch 37/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4352 - accuracy: 0.8425 - val_loss: 0.4677 - val_accuracy: 0.8483
 F1 micro : 0.8483333333333335
Epoch 38/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4615 - accuracy: 0.8428 - val_loss: 0.4991 - val_accuracy: 0.8217
 F1 micro : 0.8216666666666665
Epoch 39/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4285 - accuracy: 0.8511 - val_loss: 0.5710 - val_accuracy: 0.8100
 F1 micro : 0.81
Epoch 40/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4580 - accuracy: 0.8349 - val_loss: 0.4461 - val_accuracy: 0.8383
 F1 micro : 0.8383333333333334
Epoch 41/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4423 - accuracy: 0.8452 - val_loss: 0.5938 - val_accuracy: 0.7767
 F1 micro : 0.7766666666666666
Epoch 42/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4223 - accuracy: 0.8497 - val_loss: 0.5558 - val_accuracy: 0.8067
 F1 micro : 0.8066666666666665
Epoch 43/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4126 - accuracy: 0.8539 - val_loss: 0.4821 - val_accuracy: 0.8333
 F1 micro : 0.8333333333333334
Epoch 44/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4294 - accuracy: 0.8488 - val_loss: 0.4448 - val_accuracy: 0.8383
 F1 micro : 0.8383333333333334
Epoch 45/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4031 - accuracy: 0.8549 - val_loss: 0.6177 - val_accuracy: 0.7967
 F1 micro : 0.7966666666666665
Epoch 46/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4316 - accuracy: 0.8463 - val_loss: 0.4597 - val_accuracy: 0.8417
 F1 micro : 0.8416666666666667
Epoch 47/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3952 - accuracy: 0.8611 - val_loss: 0.5491 - val_accuracy: 0.8117
 F1 micro : 0.8116666666666666
Epoch 48/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4199 - accuracy: 0.8492 - val_loss: 0.4631 - val_accuracy: 0.8450
 F1 micro : 0.845
Epoch 49/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3927 - accuracy: 0.8628 - val_loss: 0.4732 - val_accuracy: 0.8500
 F1 micro : 0.85
Epoch 50/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3937 - accuracy: 0.8601 - val_loss: 0.4906 - val_accuracy: 0.8283
 F1 micro : 0.8283333333333334
Epoch 51/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4084 - accuracy: 0.8517 - val_loss: 0.5264 - val_accuracy: 0.8233
 F1 micro : 0.8233333333333334
Epoch 52/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.4010 - accuracy: 0.8553 - val_loss: 0.4466 - val_accuracy: 0.8500
 F1 micro : 0.85
Epoch 53/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3757 - accuracy: 0.8658 - val_loss: 0.4432 - val_accuracy: 0.8300
 F1 micro : 0.83
Epoch 54/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3870 - accuracy: 0.8642 - val_loss: 0.4712 - val_accuracy: 0.8350
 F1 micro : 0.835
Epoch 55/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3714 - accuracy: 0.8685 - val_loss: 0.5282 - val_accuracy: 0.8200
 F1 micro : 0.82
Epoch 56/60
```

```
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3839 - accuracy: 0.8665 - val_loss: 0.5950 - val_accuracy: 0.8067
 F1 micro : 0.8066666666666665
Epoch 57/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3690 - accuracy: 0.8726 - val_loss: 0.4216 - val_accuracy: 0.8567
 F1 micro : 0.8566666666666667
Epoch 58/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3383 - accuracy: 0.8821 - val_loss: 0.4311 - val_accuracy: 0.8517
 F1 micro : 0.8516666666666667
Epoch 59/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3480 - accuracy: 0.8788 - val_loss: 0.4634 - val_accuracy: 0.8633
 F1 micro : 0.8633333333333333
Epoch 60/60
1260/1260 [==============================] - 5s 4ms/step - loss: 0.3351 - accuracy: 0.8768 - val_loss: 0.5140 - val_accuracy: 0.8217
 F1 micro : 0.8216666666666665
```

Out[65]:
```
<tensorflow.python.keras.callbacks.History at 0x7fa38ce99750>
```