1130532_Textsummarynlpproject_Final_Code

April 19, 2021

[ ]:
```
#step1 import all the required libraries
#install this version of transformers and pytorch
!pip install transformers==2.8.0
!pip install torch==1.4.0
from transformers import T5Tokenizer, T5ForConditionalGeneration
import tensorflow_datasets as tfds
import torch
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import tensorflow as tf
from tensorflow import keras
import nltk,spacy,re,string,random,time
import matplotlib.pyplot as plt
from gensim.parsing.preprocessing import STOPWORDS
from spacy.lang.en.stop_words import STOP_WORDS
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from sklearn.model_selection import train_test_split
from collections import Counter
from keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.layers import␣
↪Input,LSTM,Embedding,Dense,Concatenate,TimeDistributed,Bidirectional
from tensorflow.keras.models import Model
from tensorflow.keras.callbacks import EarlyStopping,ReduceLROnPlateau
from attension import AttentionLayer
from keras.initializers import Constant
from keras.optimizers import Adam
from keras import backend as K
from rouge import rouge_n,rouge
from bleau import compute_bleu
#ignore warnings
import warnings
warnings.filterwarnings("ignore")
#stopwords removal list
nltk.download('stopwords')
#punkt for tokenization
```

```
nltk.download('punkt')
#for tokenaizations
nltk.download('wordnet')
#combine all the stopwords and create one single list of stopwords
s1=stopwords.words('english')
s2=list(STOP_WORDS)
s3=list(STOPWORDS)
#final list of stopwords
stop_words = s1+s2+s3
#use cuda if available
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")

#step2
#contraction are used to replace words with their longer meaningfull counter␣
 ↪parts
contraction = {
"ain't": "am not / are not / is not / has not / have not",
"aren't": "are not / am not",
"can't": "cannot",
"can't've": "cannot have",
"'cause": "because",
"could've": "could have",
"couldn't": "could not",
"couldn't've": "could not have",
"didn't": "did not",
"doesn't": "does not",
"don't": "do not",
"hadn't": "had not",
"hadn't've": "had not have",
"hasn't": "has not",
"haven't": "have not",
"he'd": "he had / he would",
"he'd've": "he would have",
"he'll": "he shall / he will",
"he'll've": "he shall have / he will have",
"he's": "he has / he is",
"how'd": "how did",
"how'd'y": "how do you",
"how'll": "how will",
"how's": "how has / how is / how does",
"I'd": "I had / I would",
"I'd've": "I would have",
"I'll": "I shall / I will",
"I'll've": "I shall have / I will have",
"I'm": "I am",
"I've": "I have",
"isn't": "is not",
```

```
"it'd": "it had / it would",
"it'd've": "it would have",
"it'll": "it shall / it will",
"it'll've": "it shall have / it will have",
"it's": "it has / it is",
"let's": "let us",
"ma'am": "madam",
"mayn't": "may not",
"might've": "might have",
"mightn't": "might not",
"mightn't've": "might not have",
"must've": "must have",
"mustn't": "must not",
"mustn't've": "must not have",
"needn't": "need not",
"needn't've": "need not have",
"o'clock": "of the clock",
"oughtn't": "ought not",
"oughtn't've": "ought not have",
"shan't": "shall not",
"sha'n't": "shall not",
"shan't've": "shall not have",
"she'd": "she had / she would",
"she'd've": "she would have",
"she'll": "she shall / she will",
"she'll've": "she shall have / she will have",
"she's": "she has / she is",
"should've": "should have",
"shouldn't": "should not",
"shouldn't've": "should not have",
"so've": "so have",
"so's": "so as / so is",
"that'd": "that would / that had",
"that'd've": "that would have",
"that's": "that has / that is",
"there'd": "there had / there would",
"there'd've": "there would have",
"there's": "there has / there is",
"they'd": "they had / they would",
"they'd've": "they would have",
"they'll": "they shall / they will",
"they'll've": "they shall have / they will have",
"they're": "they are",
"they've": "they have",
"to've": "to have",
"wasn't": "was not",
"we'd": "we had / we would",
```

```
"we'd've": "we would have",
"we'll": "we will",
"we'll've": "we will have",
"we're": "we are",
"we've": "we have",
"weren't": "were not",
"what'll": "what shall / what will",
"what'll've": "what shall have / what will have",
"what're": "what are",
"what's": "what has / what is",
"what've": "what have",
"when's": "when has / when is",
"when've": "when have",
"where'd": "where did",
"where's": "where has / where is",
"where've": "where have",
"who'll": "who shall / who will",
"who'll've": "who shall have / who will have",
"who's": "who has / who is",
"who've": "who have",
"why's": "why has / why is",
"why've": "why have",
"will've": "will have",
"won't": "will not",
"won't've": "will not have",
"would've": "would have",
"wouldn't": "would not",
"wouldn't've": "would not have",
"y'all": "you all",
"y'all'd": "you all would",
"y'all'd've": "you all would have",
"y'all're": "you all are",
"y'all've": "you all have",
"you'd": "you had / you would",
"you'd've": "you would have",
"you'll": "you shall / you will",
"you'll've": "you shall have / you will have",
"you're": "you are",
"you've": "you have",
"rec'd": "received"
}
#rec'd this is my addition to the list of contractions

#step3
#process_text function is used to remove unwanted characters, stopwords, and␣
 ↪format the text to create fewer nulls word embeddings
def process_text(text,contractions,remove_stopwords = True):
```

```
  #convert words to lower case
  text = text.lower()

  #replace contractions with their longer forms
  if True:
    text = text.split()
    new_text = []
    for word in text:
      if word in contractions:
        new_text.append(contractions[word])
      else:
        new_text.append(word)
    text = " ".join(new_text)

  #format words and remove unwanted characters
  text = re.sub(r'https?:\/\/.*[\r\n]*', '', text, flags=re.MULTILINE) #remove␣
↪https string
  text = re.sub(r'\<a href', ' ', text) #remove hyperlink
  text = re.sub(r'&amp;', '', text) #remove & in text
  text = re.sub(r'[_"\-;%()|+&=*%.,!?:#$@\[\]/]', ' ', text) #remove unwanted␣
↪charecters like puntuation and others
  text = re.sub(r'<br />', ' ', text) #remove new line spaces
  text = re.sub(r'\'', ' ', text) #remove slashes
  text = " ".join(text.split()) #remove trailing spaces
  #string.printable returns all sets of punctuation, digits, ascii_letters and␣
↪whitespace.
  printable = set(string.printable)
  #filter to remove punctuations,digits, ascii_letters and whitespaces
  text = "".join(list(filter(lambda x: x in printable, text)))
  #remove stop words is true then remove stopwords also
  if remove_stopwords:
    text = text.split()
    text = [w for w in text if not w in stop_words]
    text = " ".join(text)

  return text

#step4
#get_data function gets the data from gz file into a dataframe and process the␣
↪columns
#stops are not removed for summary they are only removed from text this is done␣
↪to get more human like summaries
#after processing it returns a dataframe
def get_data(contractions):
  st=time.time()
  #load the data into a dataframe
```

```
  df = pd.read_json('/content/drive/MyDrive/
 ↪reviews_Clothing_Shoes_and_Jewelry_5.json.gz', lines=True,␣
 ↪compression='gzip')
  #drop unwanted columns
  df.drop(columns=['reviewerID', 'asin', 'reviewerName',␣
 ↪'helpful','overall','unixReviewTime', 'reviewTime'],inplace=True)
  print("length of the data",len(df))
  #apply preprocess function on the columns of the dataframe
  df['reviewText'] = df['reviewText'].apply(lambda x:␣
 ↪process_text(x,contractions,remove_stopwords = True))
  df['summary'] = df[ 'summary'].apply(lambda x:␣
 ↪process_text(x,contractions,remove_stopwords = False))
  #write preprocesssed data to csv file
  df.to_csv("/content/drive/MyDrive/product_reviews.csv",index=False)
  print("total time to generate data and write in csv file ",time.time()-st)


#step5
#get_embeddings function is used to gett te word embeddings
#i am using conceptual number batch word embeddings
def get_embeddings():
  #get word embeddings
  embeddings_index = {}
  with open('/content/drive/MyDrive/numberbatch-en-19.08.txt',␣
 ↪encoding='utf-8') as f:
    for line in f:
      values = line.split(' ')
      word = values[0]
      embedding = np.asarray(values[1:], dtype='float32')
      embeddings_index[word] = embedding

  print('Word embeddings:', len(embeddings_index))
  return embeddings_index

#step6
#this function is used to build vocabulary
def get_vocab(embeddings_index,word_counts,threshold):
  #get the number of missing words
  missing_words={k:v for k,v in word_counts.items() if v >= threshold if k not␣
 ↪in embeddings_index.keys()}
  missing_ratio = round(len(missing_words)/len(word_counts),4)*100
  print("Number of words missing from word_embeddings:", len(missing_words))
  print("Percent of words that are missing from our vocabulary: {}%".
 ↪format(missing_ratio))

  #mapping vocab to index
```

```
  lr=iter([item for item in range(0,len(word_counts))])
  vocab_to_int={k:next(lr) for k,v in word_counts.items() if v >= threshold or␣
 ↪k in embeddings_index.keys()}

  #mapping index to vocab
  lr=iter([item for item in range(0,len(word_counts))])
  int_to_vocab={next(lr):k for k,v in word_counts.items() if v >= threshold or␣
 ↪k in embeddings_index.keys()}

  # Special tokens that will be added to our vocab
  codes = ["<UNK>","<PAD>","<EOS>","<GO>"]

  # Add codes to vocab
  for code in codes:
      vocab_to_int[code] = len(vocab_to_int)
      int_to_vocab[len(int_to_vocab)] = code

  #print usage of words in our model and their percent
  usage_ratio = round(len(vocab_to_int) / len(word_counts),4)*100
  print("Total number of unique words:", len(word_counts))
  print("Number of words we will use:", len(vocab_to_int))
  print("Percent of words we will use: {}%".format(usage_ratio))
  print("length vocab_to_int",len(vocab_to_int))
  print("length int_to_vocab",len(int_to_vocab))

  return vocab_to_int,int_to_vocab

#step7
#function to map words with its word embeddings
#if embeddings not found for the word then map it with a random number in␣
 ↪range(-1.0,1.0)
def word_embedding_index(vocab_to_int,embeddings_index):
  #using 300 for embedding dimensions to match CN's vectors.
  embedding_dim = 300
  nb_words = len(vocab_to_int)

  # Create matrix with default values of zero
  word_embedding_matrix = np.zeros((nb_words, embedding_dim), dtype=np.float32)
  for word, i in vocab_to_int.items():
    if word in embeddings_index:
      word_embedding_matrix[i] = embeddings_index[word]
    else:
      # If word not in CN, create a random embedding for it
      new_embedding = np.array(np.random.uniform(-1.0, 1.0, embedding_dim))
      #embeddings_index[word] = new_embedding
      word_embedding_matrix[i] = new_embedding
```

```
  # Check if value matches len(vocab_to_int)
  print("length of word embedding matrix",len(word_embedding_matrix))
  return word_embedding_matrix

#step8
#append unk and eos tokens
#if eos is equal to true then append go and eos token at begining and end of␣
 ↪the summary
#add unknown token for word not found in vocabulary
def convert_to_ints(text,vocab_to_int,eos=False):
  ints = []
  for word in text.split():
    if word in vocab_to_int:
      ints.append(vocab_to_int[word])
    else:
      ints.append(vocab_to_int["<UNK>"])
  if eos:
    ints.insert(0,vocab_to_int["<GO>"])
    ints.insert(len(ints),vocab_to_int["<EOS>"])
  return ints

#step9
#count unknown tokens
def count_unk(text):
  unk=0
  eos=0
  #print(text)
  for value in text:
    if 41413 in value:
      unk+=1
  return unk

#step10
def counts(val):
  c=0
  for i in val:
    try:
      if i==41413:
        c+=1
    except:
      pass
  return c

#step11
#remove rows from data frame that dosent staisfy the condition this is done so␣
 ↪model is trained with proper data
#redundancey is less and input text is accurate
```

```
def get_refined_output(df,max_rl,max_sl):
  unk_rl=1 #unknown token review limit
  unk_sl=0 #unknown token summary limit
  min_rl=2 #minimum review length
  #get the total length of reviewText this is used for sorting
  df["total_length"]=df['reviewText'].apply(lambda x: len(x))
  #get reviewText whose length is greater then minimum review length
  df=df[df['reviewText'].apply(lambda x: len(x)>=min_rl)]
  #get reviewText whose length is less than maximum review length
  df=df[df['reviewText'].apply(lambda x: len(x)<=max_rl)]
  #filter out the unknwon tokens based on unknown token reviewText limit
  df=df[df['reviewText'].apply(lambda x: counts(x)<=unk_rl)]
  #get summary whose length is less than maximum summary length
  df=df[df['summary'].apply(lambda x: len(x)<=max_sl)]
  #filter out the unkown tokens based on unkown token summary limit
  df=df[df['summary'].apply(lambda x: counts(x)<=unk_sl)]
  #sort the values in ascending order
  df.sort_values(by=["total_length"],ascending=True,inplace=True)
  #drop unwanted columns
  df.drop(columns=["total_length","word"],inplace=True)
  #reset index
  df.reset_index(drop=True,inplace=True)
  return df

#step12
#function to plot the length of training, validation and testing
def plot_tr_tval_tt_len(xtr,xval,xtt):
  names = ['Training','Validation','Testing']
  values = [len(xtr),len(xval),len(xtt)]
  plt.figure(figsize=(10,5))
  plt.subplot(131)
  plt.
 ↪bar(names,values,color=['darkorange','coral','coral'],edgecolor='darkblue')
  plt.suptitle('Categorical Plotting')
  plt.show()

#step13
#function to plot loss and accuracy curves on training and validation set
def plotgraph(history):
  plt.figure(figsize=[8,6])
  plt.plot(history.history['loss'],'firebrick',linewidth=3.0)
  plt.plot(history.history['accuracy'],'turquoise',linewidth=3.0)
  plt.plot(history.history['val_loss'],'midnightblue',linewidth=3.0)
  plt.legend(['Training loss','Training Accuracy','Validation␣
 ↪loss'],fontsize=18)
  plt.xlabel('Epochs',fontsize=16)
  plt.ylabel('Loss and Accuracy',fontsize=16)
```

```
  plt.title('Loss Curves and Accuracy Curves for text␣
 ↪summarization',fontsize=16)

#step14
#this function is used to get the preprocessed csv file for our text summarizer
def Get_the_data():
  #lower the string in contractions and convert it into dict
  contractions = dict((k.lower(), v.lower()) for k, v in contraction.items())
  #till this step all data is processed and we get our csv file of cleaned texts
  get_data(contractions)

  #free memory
  del contractions

#step15 is used to call function Get_the_data which get the preprocessed data␣
 ↪and writes it into a csv file
#Get_the_data()

#step16
#this function combines all the above ouput generated by the above function in␣
 ↪a proper squence of steps
def combining_all_steps():

  st=time.time()
  #get the final cleaned data
  df=pd.read_csv('/content/drive/MyDrive/product_reviews.csv')[:180000]
  print("The length of dataset is ",len(df))
  #combine reviewText and summary so common vocabulary can be created by␣
 ↪finding frequent words
  df["word"]=df[['reviewText','summary']].apply(lambda x : '{} {}'.
 ↪format(x[0],x[1]), axis=1)
  #get frequency of words
  word_counts=pd.Series(np.concatenate([x.split() for x in df.word])).
 ↪value_counts()
  word_counts=word_counts.to_dict()
  #print(type(word_counts))
  print("vocab length",len(word_counts))
  #set the threshold
  threshold = 20
  max_rl=80 #maximum review length
  max_sl=10 #maximum summary length
  #get the embeddings matrix
  embeddings_index= get_embeddings()
  #get vocab to index and index to vocab mapping of words
  vocab_to_int,int_to_vocab=get_vocab(embeddings_index,word_counts,threshold)
  #get word embedding for the words in vocab
```

```
 word_embedding_matrix=word_embedding_index(vocab_to_int,embeddings_index)
 #convert words to integers based on their index positions
 df['reviewText'] = df['reviewText'].apply(lambda x:␣
↪convert_to_ints(str(x),vocab_to_int,eos=False))
 df['summary'] = df[ 'summary'].apply(lambda x:␣
↪convert_to_ints(str(x),vocab_to_int,eos=True))
 print("after word to index for reviewText",df["reviewText"][0])
 print("after word to index for summary",df["summary"][0])
 rvunk=count_unk(df["reviewText"])
 smunk=count_unk(df["summary"])
 print("total number of unk token are",rvunk+smunk)
 #apply the filters and get the final preprocessed data
 df=get_refined_output(df,max_rl,max_sl)
 print("length of dataset that will be used",len(df))
 #split data into 75% train, 15% validation and 15% test datasets
␣
↪x_tr,x_val,y_tr,y_val=train_test_split(df['reviewText'],df['summary'],test_size=0.
↪3,random_state=1,shuffle=True)
 x_tt,x_val,y_tt,y_val=train_test_split(x_val,y_val,test_size=0.
↪5,random_state=1,shuffle=True)
 print("length of split datasets train {}, test {} and validation {}".
↪format(len(x_tr),len(x_tt),len(x_val)))
 print("Vocabulary Size: {}".format(len(vocab_to_int)))
␣
↪print("voc_to_int_",vocab_to_int['<UNK>'],vocab_to_int['<PAD>'],vocab_to_int['<EOS>'])
 #reset index
 x_tr=x_tr.reset_index()
 y_tr=y_tr.reset_index()
 x_tt=x_tt.reset_index()
 y_tt=y_tt.reset_index()
 x_val=x_val.reset_index()
 y_val=y_val.reset_index()
 #find max lenght just to verfiy the output of get refined function
 #max([len(sentence) for sentence in y_tt["summary"]])
 #pad the reviewText and summary to the specified max length
 xtr=pad_sequences(x_tr["reviewText"], padding='post',maxlen=max_rl,␣
↪value=vocab_to_int["<PAD>"])
 ytr=pad_sequences(y_tr["summary"], padding='post',maxlen=max_sl,␣
↪value=vocab_to_int["<PAD>"])
 xtt=pad_sequences(x_tt["reviewText"], padding='post',maxlen=max_rl,␣
↪value=vocab_to_int["<PAD>"])
 ytt=pad_sequences(y_tt["summary"], padding='post',maxlen=max_sl,␣
↪value=vocab_to_int["<PAD>"])
 xval=pad_sequences(x_val["reviewText"], padding='post',maxlen=max_rl,␣
↪value=vocab_to_int["<PAD>"])
```

```
  yval=pad_sequences(y_val["summary"], padding='post',maxlen=max_sl,␣
 ↪value=vocab_to_int["<PAD>"])
  #find the number of unique tokens in the list
  #flat_list_rt = [item for sublist in df["reviewText"] for item in sublist]
  #flat_list_s = [item for sublist in df["summary"] for item in sublist]
  #rt=len(np.unique(flat_list_rt))
  #st=len(np.unique(flat_list_s))
  #print("number of unique tokens reviewText {} and summary {}".format(rt,st))
  #plot the length of training, validation and testing
  plot_tr_tval_tt_len(xtr,xval,xtt)
  print("total time to complete all the above steps and get final data ",time.
 ↪time()-st)
  #free memory delete values stored in variables which are not required further
  del df,word_counts,embeddings_index,x_tr,x_val,y_tr,y_val,x_tt,y_tt

  return␣
 ↪xtr,ytr,xtt,ytt,xval,yval,vocab_to_int,int_to_vocab,word_embedding_matrix,max_rl,max_sl

#step17
#function to get summary given a sequence
def seq_to_summary(seq,vocab_to_int,int_to_vocab):
  newstring=''
  for i in seq:
    if ((i!=0 and i!=vocab_to_int['<GO>']) and i!=vocab_to_int['<EOS>']):
      newstring=newstring+int_to_vocab[i]+' '
  return newstring

#step18
#function to get text given a sequence
def seq_to_text(seq,int_to_vocab):
  newstring=''
  for i in seq:
    if (i!=0):
      newstring=newstring+int_to_vocab[i]+' '
  return newstring

#step19
#this function get the data for the pretrained model t5small
def combining_all_steps_t5():
  #get the final cleaned data
  df=pd.read_csv('/content/drive/MyDrive/product_reviews.csv')[:147799]
  print("The length of dataset is ",len(df))

  #set the threshold
  threshold = 20
  max_rl=80 #maximum review length
  max_sl=10 #maximum summary length
```

```
 #get reviewText whose length is less than maximum review length
 df['reviewText']=df['reviewText'].str.slice(0,max_rl)

 #get summary whose length is less than maximum summary length
 df['summary']=df['summary'].str.slice(0,max_rl)

 #split data into 75% train, 15% validation and 15% test datasets
␣
↪x_tr,x_val,y_tr,y_val=train_test_split(df['reviewText'],df['summary'],test_size=0.
↪3,random_state=1,shuffle=True)
 x_tt,x_val,y_tt,y_val=train_test_split(x_val,y_val,test_size=0.
↪5,random_state=1,shuffle=True)

 #reset index
 x_tr=x_tr.reset_index()
 y_tr=y_tr.reset_index()
 x_tt=x_tt.reset_index()
 y_tt=y_tt.reset_index()
 x_val=x_val.reset_index()
 y_val=y_val.reset_index()
 print("train {}, val {}, test {}".format(len(x_tr),len(x_val),len(x_tt)))
 return x_tr,y_tr,x_tt,y_tt,x_val,y_val
```

```
Requirement already satisfied: transformers==2.8.0 in
/usr/local/lib/python3.7/dist-packages (2.8.0)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-
packages (from transformers==2.8.0) (2.23.0)
Requirement already satisfied: sentencepiece in /usr/local/lib/python3.7/dist-
packages (from transformers==2.8.0) (0.1.95)
Requirement already satisfied: regex!=2019.12.17 in
/usr/local/lib/python3.7/dist-packages (from transformers==2.8.0) (2019.12.20)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==2.8.0) (3.0.12)
Requirement already satisfied: sacremoses in /usr/local/lib/python3.7/dist-
packages (from transformers==2.8.0) (0.0.44)
Requirement already satisfied: boto3 in /usr/local/lib/python3.7/dist-packages
(from transformers==2.8.0) (1.17.53)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from transformers==2.8.0) (1.19.5)
Requirement already satisfied: tokenizers==0.5.2 in
/usr/local/lib/python3.7/dist-packages (from transformers==2.8.0) (0.5.2)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.7/dist-
packages (from transformers==2.8.0) (4.41.1)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests->transformers==2.8.0) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
```

```
/usr/local/lib/python3.7/dist-packages (from requests->transformers==2.8.0)
(1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from requests->transformers==2.8.0)
(2020.12.5)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from requests->transformers==2.8.0)
(3.0.4)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==2.8.0) (1.0.1)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==2.8.0) (7.1.2)
Requirement already satisfied: six in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==2.8.0) (1.15.0)
Requirement already satisfied: s3transfer<0.4.0,>=0.3.0 in
/usr/local/lib/python3.7/dist-packages (from boto3->transformers==2.8.0) (0.3.7)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in
/usr/local/lib/python3.7/dist-packages (from boto3->transformers==2.8.0)
(0.10.0)
Requirement already satisfied: botocore<1.21.0,>=1.20.53 in
/usr/local/lib/python3.7/dist-packages (from boto3->transformers==2.8.0)
(1.20.53)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in
/usr/local/lib/python3.7/dist-packages (from
botocore<1.21.0,>=1.20.53->boto3->transformers==2.8.0) (2.8.1)
Requirement already satisfied: torch==1.4.0 in /usr/local/lib/python3.7/dist-
packages (1.4.0)
[nltk_data] Downloading package stopwords to /root/nltk_data…
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data…
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data…
[nltk_data]   Package wordnet is already up-to-date!
```

[ ]:
```
#step20
#function to design and evaluate the model
def␣
 ↪design_model_fit_eval(xtr,ytr,xval,yval,vocab_to_int,word_embedding_matrix,max_rl):
 ↪
  K.clear_session()
  latent_dim = 80
  embedding_dim=300

  # Encoder
  encoder_inputs = Input(shape=(max_rl,))

  #embedding layer
```

```
enc_emb =  Embedding(len(vocab_to_int),
                     embedding_dim,
                     embeddings_initializer=Constant(word_embedding_matrix),
                     trainable=False)(encoder_inputs)


#LSTM 1
encoder_lstm1 = LSTM(latent_dim,return_sequences=True,return_state=True)
encoder_output1, state_h1, state_c1 = encoder_lstm1(enc_emb)

#LSTM 2
encoder_lstm2 = LSTM(latent_dim,return_sequences=True,return_state=True)
encoder_output2, state_h2, state_c2 = encoder_lstm2(encoder_output1)

#LSTM 3
encoder_lstm3=LSTM(latent_dim, return_state=True, return_sequences=True)
encoder_outputs, state_h, state_c= encoder_lstm3(encoder_output2)

# Set up the decoder, using `encoder_states` as initial state.
decoder_inputs = Input(shape=(None,))

#embedding layer
dec_emb_layer = Embedding(len(vocab_to_int),
                          embedding_dim,
                          ␣
↪embeddings_initializer=Constant(word_embedding_matrix),
                          trainable=False)

#decoder
dec_emb = dec_emb_layer(decoder_inputs)

decoder_lstm = LSTM(latent_dim, return_sequences=True,␣
↪return_state=True,dropout=0.4,recurrent_dropout=0.2)
decoder_outputs,decoder_fwd_state, decoder_back_state =␣
↪decoder_lstm(dec_emb,initial_state=[state_h, state_c])

# Attention layer
attn_layer = AttentionLayer(name='attention_layer')
attn_out, attn_states = attn_layer([encoder_outputs, decoder_outputs])

# Concat attention input and decoder LSTM output
decoder_concat_input = Concatenate(axis=-1,␣
↪name='concat_layer')([decoder_outputs, attn_out])

#dense layer
decoder_dense =  TimeDistributed(Dense(len(vocab_to_int),␣
↪activation='softmax'))
```

```
  decoder_outputs = decoder_dense(decoder_concat_input)

  # Define the model
  model = Model([encoder_inputs, decoder_inputs], decoder_outputs)

  #print model summary
  model.summary()

  model.
↪compile(optimizer='rmsprop',loss='sparse_categorical_crossentropy',metrics=['accuracy'])
  #reduce_lr method is used to reduce the learning rate if the learning rate is␣
↪stagnant or if there are no major improvements in training
  reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2,
                                patience=5, min_lr=0.001)
  #early stopping condition
  es = EarlyStopping(monitor='val_loss', mode='min', verbose=1, patience=10)

  st=time.time()

  #fit te model
  history=model.fit([xtr,ytr[:,:-1]], ytr.reshape(ytr.shape[0],ytr.shape[1],␣
↪1)[:,1:] ,epochs=100,callbacks=[es],batch_size=512,␣
↪validation_data=([xval,yval[:,:-1]], yval.reshape(yval.shape[0],yval.
↪shape[1], 1)[:,1:]))

  #plot loss and accuracy curves
  plotgraph(history)
  print("total time required for training ",time.time()-st)
  return encoder_inputs,encoder_outputs, state_h,␣
↪state_c,decoder_inputs,decoder_lstm,attn_layer,decoder_dense,dec_emb_layer
```

[ ]:
```
#step21
#design of inference function
def design_inference(encoder_inputs,encoder_outputs, state_h,␣
 ↪state_c,decoder_inputs,decoder_lstm,attn_layer,decoder_dense,max_rl,dec_emb_layer):
 ↪
  #latent dimension
  latent_dim = 80

  #encode the input sequence to get the feature vector
  encoder_model = Model(inputs=encoder_inputs,outputs=[encoder_outputs,␣
 ↪state_h, state_c])

  #decoder setup
  #below tensors will hold the states of the previous time step
  decoder_state_input_h = Input(shape=(latent_dim,))
  decoder_state_input_c = Input(shape=(latent_dim,))
```

```
  decoder_hidden_state_input = Input(shape=(max_rl,latent_dim))

  #get the embeddings of the decoder sequence
  dec_emb2= dec_emb_layer(decoder_inputs)
  #to predict the next word in the sequence, set the initial states to the␣
↪states from the previous time step
  decoder_outputs2, state_h2, state_c2 = decoder_lstm(dec_emb2,␣
↪initial_state=[decoder_state_input_h, decoder_state_input_c])

  #attention inference
  attn_out_inf, attn_states_inf = attn_layer([decoder_hidden_state_input,␣
↪decoder_outputs2])
  decoder_inf_concat = Concatenate(axis=-1, name='concat')([decoder_outputs2,␣
↪attn_out_inf])

  #a dense softmax layer to generate prob dist. over the target vocabulary
  decoder_outputs2 = decoder_dense(decoder_inf_concat)

  #final decoder model
  decoder_model = Model([decoder_inputs] +␣
↪[decoder_hidden_state_input,decoder_state_input_h, decoder_state_input_c],
                        [decoder_outputs2] + [state_h2, state_c2])

  return encoder_model,decoder_model
```

[ ]:
```
#step22
#fucntion to get the decoded squence for the given review
def␣
↪decode_sequence(input_seq,encoder_model,decoder_model,vocab_to_int,int_to_vocab,max_sl):
↪
  # Encode the input as state vectors.
  e_out, e_h, e_c = encoder_model.predict(input_seq)

  # Generate empty target sequence of length 1.
  target_seq = np.zeros((1,1))

  # Populate the first word of target sequence with the start word.
  target_seq[0, 0] = vocab_to_int['<GO>']

  stop_condition = False
  decoded_sentence = ''
  while not stop_condition:
    output_tokens, h, c = decoder_model.predict([target_seq] + [e_out, e_h,␣
↪e_c])

    # Sample a token
```

```
    sampled_token_index = np.argmax(output_tokens[0, -1, :])
    sampled_token = int_to_vocab[sampled_token_index]

    if (sampled_token!="<EOS>"):
      decoded_sentence += ' '+sampled_token

      # Exit condition: either hit max length or find stop word.
      if (sampled_token == '<EOS>'  or len(decoded_sentence.split()) >=␣
↪(max_sl-1)):
        stop_condition = True

    # Update the target sequence (of length 1).
    target_seq = np.zeros((1,1))
    target_seq[0, 0] = sampled_token_index

    # Update internal states
    e_h, e_c = h, c

  return decoded_sentence
```

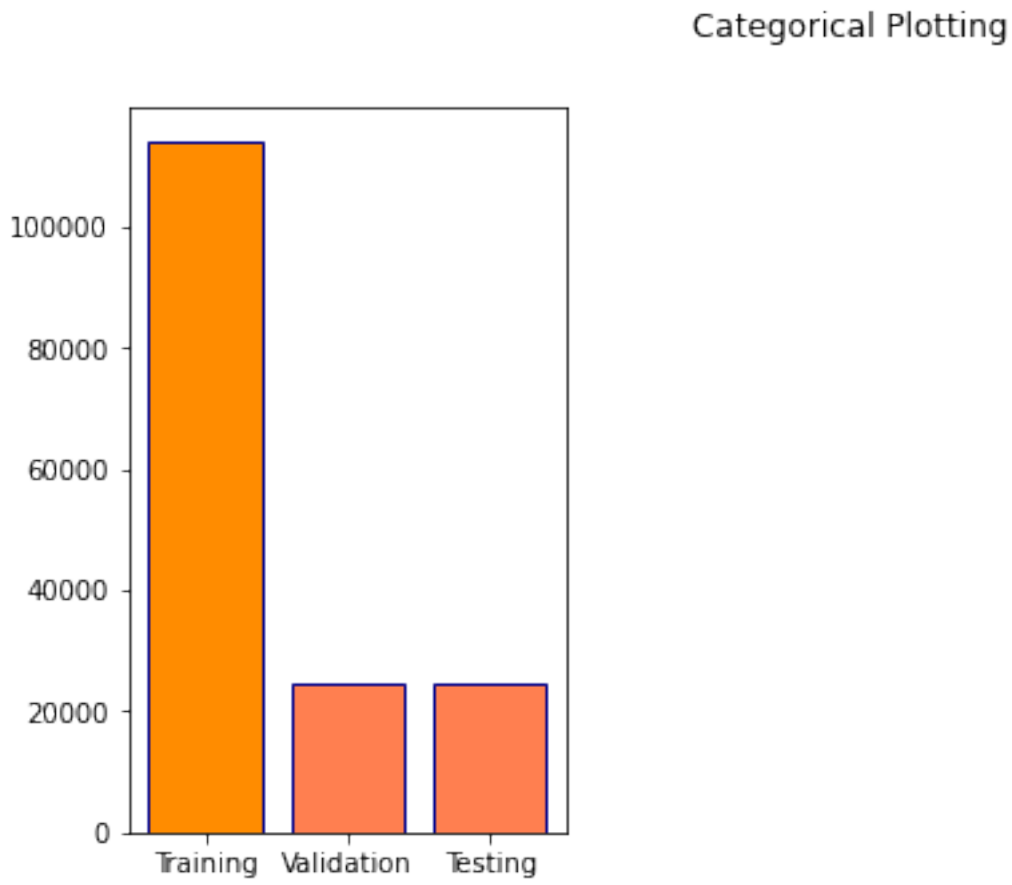
[ ]:
```
#step23
#this function is used to get the score for LSTM scratch model designed and␣
↪puts output in a txt file
def␣
↪test_scratch(xtt,ytt,int_to_vocab,vocab_to_int,encoder_model,decoder_model,max_sl,max_rl):
↪
  st=time.time()
  predictions = []
  real_og=[]
  pred_op=[]
  c=0
  b=50
  for i in range(0,len(xtt)):
    #review
    review=seq_to_text(xtt[i],int_to_vocab)
    review=review.replace("<PAD>",'')
    #original summary
    og_summary=seq_to_summary(ytt[i],vocab_to_int,int_to_vocab)
    og_summary=og_summary.replace("<PAD>",'')
    real_og.append(str(og_summary))
    #predicted summary
    predict_summary=decode_sequence(xtt[i].
↪reshape(1,max_rl),encoder_model,decoder_model,vocab_to_int,int_to_vocab,max_sl)
    predict_summary=predict_summary.replace("<PAD>",'')
    pred_op.append(str(predict_summary))
    #write to a text file name review_og_pred.txt
```

```
    predictions.append("review:"+review+"\t"+"orignal:
↪"+og_summary+"\t"+"predicted:"+predict_summary+"\n")
    #this part is used to print output if the size of c is greater than b
    #limited output is print as only 5000 lines can be printed in colab whole␣
↪output is written in a text file
    if c>b:
      print("Review: {}".format(review))
      print("Original Summary: {}".format(og_summary))
      print("Predicted Summary: {}".format(predict_summary))
      b+=b
    c+=1

  print("total time to complete {}".format(time.time()-st))
  file = open("/content/drive/MyDrive/LSTMscore.txt","w")
  file.writelines(predictions)
  file.close()

  bleau=compute_bleu(real_og,pred_op, max_order=4,smooth=False)
  bscore=nltk.translate.bleu_score.corpus_bleu(real_og,pred_op)
  rougen=rouge_n(pred_op, real_og, n=2)
  ro=rouge(pred_op, real_og)

  print("bleu, precisions, bp, ratio, translation_length,␣
↪reference_length",bleau)
  print("bleau score",bscore)
  print("rouge2",rougen)
  print("rouge",ro)
```

[ ]:
```
#step24
def lstmmodel():
  #this the model designed by me for text summarization
  st=time.time()
  #get the data
 ␣
↪xtr,ytr,xtt,ytt,xval,yval,vocab_to_int,int_to_vocab,word_embedding_matrix,max_rl,max_sl=com
  #call the model
  encoder_inputs,encoder_outputs, state_h,␣
↪state_c,decoder_inputs,decoder_lstm,attn_layer,decoder_dense,dec_emb_layer=design_model_fit
  #get the inference output
  encoder_model,decoder_model=design_inference(encoder_inputs,encoder_outputs,␣
↪state_h,␣
↪state_c,decoder_inputs,decoder_lstm,attn_layer,decoder_dense,max_rl,dec_emb_layer)
  #call test
 ␣
↪test_scratch(xtt,ytt,int_to_vocab,vocab_to_int,encoder_model,decoder_model,max_sl,max_rl)
  print("total time required for completing whole process ",time.time()-st)
```

```
[ ]: lstmmodel()
```

```
The length of dataset is  180000
vocab length 68861
Word embeddings: 516783
Number of words missing from word_embeddings: 728
Percent of words that are missing from our vocabulary: 1.06%
Total number of unique words: 68861
Number of words we will use: 37429
Percent of words we will use: 54.35%
length vocab_to_int 37429
length int_to_vocab 37429
length of word embedding matrix 37429
after word to index for reviewText [0, 3920, 0, 17, 12, 119, 278, 209, 79, 905,
3920, 1532]
after word to index for summary [37428, 0, 3920, 70, 1154, 565, 37427]
total number of unk token are 0
length of dataset that will be used 162996
length of split datasets train 114097, test 24449 and validation 24450
Vocabulary Size: 37429
voc_to_int_ 37425 37426 37427
```



```
total time to complete all the above steps and get final data  55.95032453536987
Model: "model"
__________________________________________________________________________________
__________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================
==================
input_1 (InputLayer)            [(None, 80)]         0
__________________________________________________________________________________
__________________
embedding (Embedding)           (None, 80, 300)      11228700    input_1[0][0]
__________________________________________________________________________________
__________________
lstm (LSTM)                     [(None, 80, 80), (No 121920      embedding[0][0]
__________________________________________________________________________________
__________________
input_2 (InputLayer)            [(None, None)]       0
__________________________________________________________________________________
__________________
lstm_1 (LSTM)                   [(None, 80, 80), (No 51520       lstm[0][0]
__________________________________________________________________________________
__________________
embedding_1 (Embedding)         (None, None, 300)    11228700    input_2[0][0]
__________________________________________________________________________________
__________________
lstm_2 (LSTM)                   [(None, 80, 80), (No 51520       lstm_1[0][0]
__________________________________________________________________________________
__________________
lstm_3 (LSTM)                   [(None, None, 80), ( 121920
embedding_1[0][0]
                                                                 lstm_2[0][1]
                                                                 lstm_2[0][2]
__________________________________________________________________________________
__________________
attention_layer (AttentionLayer ((None, None, 80), ( 12880       lstm_2[0][0]
                                                                 lstm_3[0][0]
__________________________________________________________________________________
__________________
concat_layer (Concatenate)      (None, None, 160)    0           lstm_3[0][0]
attention_layer[0][0]
__________________________________________________________________________________
__________________
time_distributed (TimeDistribut (None, None, 37429)  6026069
concat_layer[0][0]
==================================================================================
```

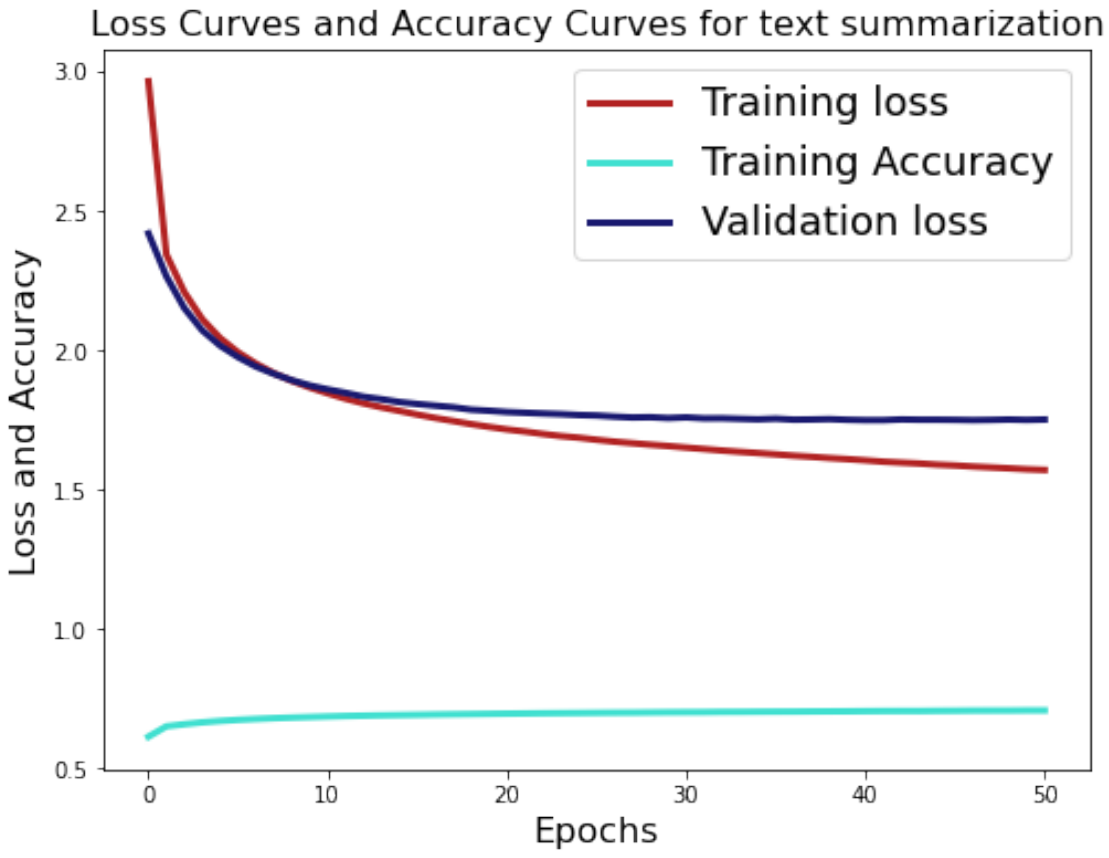
```
==================
Total params: 28,843,229
Trainable params: 6,385,829
Non-trainable params: 22,457,400
_________________________________________________________________________________
_________________
Epoch 1/100
223/223 [==============================] - 1248s 6s/step - loss: 4.1427 -
accuracy: 0.5584 - val_loss: 2.4180 - val_accuracy: 0.6474
Epoch 2/100
223/223 [==============================] - 1244s 6s/step - loss: 2.3772 -
accuracy: 0.6480 - val_loss: 2.2641 - val_accuracy: 0.6533
Epoch 3/100
223/223 [==============================] - 1243s 6s/step - loss: 2.2321 -
accuracy: 0.6559 - val_loss: 2.1500 - val_accuracy: 0.6627
Epoch 4/100
223/223 [==============================] - 1239s 6s/step - loss: 2.1236 -
accuracy: 0.6631 - val_loss: 2.0704 - val_accuracy: 0.6679
Epoch 5/100
223/223 [==============================] - 1235s 6s/step - loss: 2.0543 -
accuracy: 0.6677 - val_loss: 2.0154 - val_accuracy: 0.6720
Epoch 6/100
223/223 [==============================] - 1235s 6s/step - loss: 1.9877 -
accuracy: 0.6733 - val_loss: 1.9743 - val_accuracy: 0.6753
Epoch 7/100
223/223 [==============================] - 1245s 6s/step - loss: 1.9583 -
accuracy: 0.6751 - val_loss: 1.9404 - val_accuracy: 0.6789
Epoch 8/100
223/223 [==============================] - 1243s 6s/step - loss: 1.9236 -
accuracy: 0.6780 - val_loss: 1.9141 - val_accuracy: 0.6813
Epoch 9/100
223/223 [==============================] - 1258s 6s/step - loss: 1.8837 -
accuracy: 0.6812 - val_loss: 1.8914 - val_accuracy: 0.6834
Epoch 10/100
223/223 [==============================] - 1260s 6s/step - loss: 1.8665 -
accuracy: 0.6826 - val_loss: 1.8726 - val_accuracy: 0.6853
Epoch 11/100
223/223 [==============================] - 1254s 6s/step - loss: 1.8435 -
accuracy: 0.6851 - val_loss: 1.8585 - val_accuracy: 0.6866
Epoch 12/100
223/223 [==============================] - 1247s 6s/step - loss: 1.8204 -
accuracy: 0.6865 - val_loss: 1.8460 - val_accuracy: 0.6876
Epoch 13/100
223/223 [==============================] - 1253s 6s/step - loss: 1.8093 -
accuracy: 0.6872 - val_loss: 1.8321 - val_accuracy: 0.6892
Epoch 14/100
223/223 [==============================] - 1250s 6s/step - loss: 1.7929 -
accuracy: 0.6886 - val_loss: 1.8233 - val_accuracy: 0.6904
```

```
Epoch 15/100
223/223 [==============================] - 1249s 6s/step - loss: 1.7777 -
accuracy: 0.6905 - val_loss: 1.8132 - val_accuracy: 0.6914
Epoch 16/100
223/223 [==============================] - 1251s 6s/step - loss: 1.7641 -
accuracy: 0.6912 - val_loss: 1.8061 - val_accuracy: 0.6922
Epoch 17/100
223/223 [==============================] - 1246s 6s/step - loss: 1.7502 -
accuracy: 0.6922 - val_loss: 1.8002 - val_accuracy: 0.6931
Epoch 18/100
223/223 [==============================] - 1247s 6s/step - loss: 1.7366 -
accuracy: 0.6935 - val_loss: 1.7945 - val_accuracy: 0.6936
Epoch 19/100
223/223 [==============================] - 1251s 6s/step - loss: 1.7273 -
accuracy: 0.6936 - val_loss: 1.7857 - val_accuracy: 0.6947
Epoch 20/100
223/223 [==============================] - 1268s 6s/step - loss: 1.7194 -
accuracy: 0.6945 - val_loss: 1.7820 - val_accuracy: 0.6948
Epoch 21/100
223/223 [==============================] - 1260s 6s/step - loss: 1.7147 -
accuracy: 0.6950 - val_loss: 1.7773 - val_accuracy: 0.6956
Epoch 22/100
223/223 [==============================] - 1257s 6s/step - loss: 1.6994 -
accuracy: 0.6964 - val_loss: 1.7747 - val_accuracy: 0.6960
Epoch 23/100
223/223 [==============================] - 1259s 6s/step - loss: 1.6993 -
accuracy: 0.6956 - val_loss: 1.7721 - val_accuracy: 0.6960
Epoch 24/100
223/223 [==============================] - 1253s 6s/step - loss: 1.6862 -
accuracy: 0.6969 - val_loss: 1.7699 - val_accuracy: 0.6964
Epoch 25/100
223/223 [==============================] - 1253s 6s/step - loss: 1.6839 -
accuracy: 0.6964 - val_loss: 1.7665 - val_accuracy: 0.6968
Epoch 26/100
223/223 [==============================] - 1253s 6s/step - loss: 1.6738 -
accuracy: 0.6979 - val_loss: 1.7649 - val_accuracy: 0.6973
Epoch 27/100
223/223 [==============================] - 1253s 6s/step - loss: 1.6647 -
accuracy: 0.6987 - val_loss: 1.7608 - val_accuracy: 0.6972
Epoch 28/100
223/223 [==============================] - 1255s 6s/step - loss: 1.6560 -
accuracy: 0.6996 - val_loss: 1.7578 - val_accuracy: 0.6977
Epoch 29/100
223/223 [==============================] - 1259s 6s/step - loss: 1.6501 -
accuracy: 0.7000 - val_loss: 1.7587 - val_accuracy: 0.6980
Epoch 30/100
223/223 [==============================] - 1258s 6s/step - loss: 1.6475 -
accuracy: 0.7007 - val_loss: 1.7556 - val_accuracy: 0.6981
```

```
Epoch 31/100
223/223 [==============================] - 1257s 6s/step - loss: 1.6468 -
accuracy: 0.6999 - val_loss: 1.7581 - val_accuracy: 0.6977
Epoch 32/100
223/223 [==============================] - 1265s 6s/step - loss: 1.6340 -
accuracy: 0.7015 - val_loss: 1.7545 - val_accuracy: 0.6985
Epoch 33/100
223/223 [==============================] - 1261s 6s/step - loss: 1.6342 -
accuracy: 0.7012 - val_loss: 1.7548 - val_accuracy: 0.6984
Epoch 34/100
223/223 [==============================] - 1268s 6s/step - loss: 1.6290 -
accuracy: 0.7014 - val_loss: 1.7529 - val_accuracy: 0.6992
Epoch 35/100
223/223 [==============================] - 1268s 6s/step - loss: 1.6208 -
accuracy: 0.7028 - val_loss: 1.7511 - val_accuracy: 0.6993
Epoch 36/100
223/223 [==============================] - 1264s 6s/step - loss: 1.6151 -
accuracy: 0.7038 - val_loss: 1.7535 - val_accuracy: 0.6984
Epoch 37/100
223/223 [==============================] - 1259s 6s/step - loss: 1.6168 -
accuracy: 0.7028 - val_loss: 1.7499 - val_accuracy: 0.6994
Epoch 38/100
223/223 [==============================] - 1265s 6s/step - loss: 1.6097 -
accuracy: 0.7039 - val_loss: 1.7506 - val_accuracy: 0.6991
Epoch 39/100
223/223 [==============================] - 1275s 6s/step - loss: 1.6091 -
accuracy: 0.7029 - val_loss: 1.7518 - val_accuracy: 0.6990
Epoch 40/100
223/223 [==============================] - 1262s 6s/step - loss: 1.5974 -
accuracy: 0.7042 - val_loss: 1.7494 - val_accuracy: 0.6997
Epoch 41/100
223/223 [==============================] - 1272s 6s/step - loss: 1.5930 -
accuracy: 0.7050 - val_loss: 1.7481 - val_accuracy: 0.6996
Epoch 42/100
223/223 [==============================] - 1266s 6s/step - loss: 1.5941 -
accuracy: 0.7044 - val_loss: 1.7481 - val_accuracy: 0.6996
Epoch 43/100
223/223 [==============================] - 1272s 6s/step - loss: 1.5894 -
accuracy: 0.7047 - val_loss: 1.7507 - val_accuracy: 0.6996
Epoch 44/100
223/223 [==============================] - 1277s 6s/step - loss: 1.5861 -
accuracy: 0.7049 - val_loss: 1.7498 - val_accuracy: 0.7001
Epoch 45/100
223/223 [==============================] - 1272s 6s/step - loss: 1.5869 -
accuracy: 0.7049 - val_loss: 1.7497 - val_accuracy: 0.7003
Epoch 46/100
223/223 [==============================] - 1281s 6s/step - loss: 1.5802 -
accuracy: 0.7056 - val_loss: 1.7491 - val_accuracy: 0.7001
```

```
Epoch 47/100
223/223 [==============================] - 1270s 6s/step - loss: 1.5774 -
accuracy: 0.7062 - val_loss: 1.7482 - val_accuracy: 0.7000
Epoch 48/100
223/223 [==============================] - 1273s 6s/step - loss: 1.5757 -
accuracy: 0.7063 - val_loss: 1.7489 - val_accuracy: 0.6999
Epoch 49/100
223/223 [==============================] - 1275s 6s/step - loss: 1.5671 -
accuracy: 0.7072 - val_loss: 1.7504 - val_accuracy: 0.6997
Epoch 50/100
223/223 [==============================] - 1270s 6s/step - loss: 1.5673 -
accuracy: 0.7069 - val_loss: 1.7490 - val_accuracy: 0.7004
Epoch 51/100
223/223 [==============================] - 1268s 6s/step - loss: 1.5697 -
accuracy: 0.7066 - val_loss: 1.7504 - val_accuracy: 0.6997
Epoch 00051: early stopping
total time required for training  64167.45829248428
Review: compared hanes partner company champion hoodie exactly needed cool
winter spring fall nights fabric heavy cumbersome pulling head product
complaints value compared 34 branded 34 sweats usual service amazon
Original Summary: sweat price
Predicted Summary:  great quality
Review: briefs gift feel wear loves looks amazing complaints
Original Summary: full support in the briefest of briefs
Predicted Summary:  great
Review: took chance shoes match champagne colored dress perfect looking small
heel exactly looking quick delivery
Original Summary: wedding accessories
Predicted Summary:  love these shoes
Review: fit like years ago cheaper quality materials gravity extra weight
comfortable socks price
Original Summary: love them but
Predicted Summary:  good socks
Review: received compliments pair shoes run bit small mind love getting colors
Original Summary: very cute
Predicted Summary:  great shoes
Review: elegant perfect height beautiful black velvet love necklaces display
easy buy necklaces nice good price homework best priced places looked
Original Summary: elegant very nice way to display your necklaces
Predicted Summary:  beautiful
Review: styles choose happy got wife said look good block sunlight happy
purchase
Original Summary: cool sunglasses
Predicted Summary:  great
Review: dockers belt quality leather soft touch edging adds extra touch quality
attractiveness belt husband happy
Original Summary: top quality
Predicted Summary:  great belt
```

```
Review: boot cold weather sole little stiff need wear minute warm shoe strings
look bad tied tie tuck bow tongue shown size runs tad small maybe 1 4 size
fleece lining wear 8 5 ordered 9 perfect socks boot ready snow
Original Summary: boot for snow fun
Predicted Summary:  great boots
total time to complete 15829.7639939785
bleu, precisions, bp, ratio, translation_length, reference_length (0.0,
[0.2900551776136539, 0.0, 0.0, 0.0], 1.0, 19.09509591394331, 466856, 24449)
bleau score 0.7338717254431542
rouge2 (0.06599443828484215, 0.8312236286919831, 0.03436126421544687)
rouge {'rouge_1/f_score': 0.36915419958439477, 'rouge_1/r_score':
0.36583216309687744, 'rouge_1/p_score': 0.41662537566051006, 'rouge_2/f_score':
0.2881554948360078, 'rouge_2/r_score': 0.33457293997806903, 'rouge_2/p_score':
0.2751092653511975, 'rouge_l/f_score': 0.6180779054304887, 'rouge_l/r_score':
0.6641052933641588, 'rouge_l/p_score': 0.5875250521493721}
total time required for completing whole process  80087.67300844193
```



[ ]:
```
#summary using T5small pretrained model
```

[ ]:
```
#step26
#function is used to return the loss
def step(inputs_ids, attention_mask, y, pad_token_id, model):
  y_ids = y[:, :-1].contiguous()
  lm_labels = y[:, 1:].clone()
  lm_labels[y[:, 1:] == pad_token_id] = -100
  output = model(inputs_ids, attention_mask=attention_mask,␣
 ↪decoder_input_ids=y_ids, lm_labels=lm_labels)
  # loss
  return output[0]
```

[ ]:
```
#step25
#this function is used to train the pretrained t5small model
def t5train(train_loader,val_loader,pad_token_id,model,EPOCHS,log_interval):
  #initialize empty list for train_loss and val_loss
  train_loss = []
  val_loss = []
  #optimizer
  optimizer = torch.optim.Adam(model.parameters(),lr=1e-4, weight_decay=1e-4/25)
  #iterate for number of epochs
  for epoch in range(EPOCHS):
    model.train()
    #start time
    start_time = time.time()
    #for data in train_loader train the model
    for i, (inputs_ids, attention_mask, y) in enumerate(train_loader):
      inputs_ids = inputs_ids.to(device)
      attention_mask = attention_mask.to(device)
      y = y.to(device)

      optimizer.zero_grad()
      loss = step(inputs_ids, attention_mask, y, pad_token_id, model)
      train_loss.append(loss.item())
      loss.backward()
      torch.nn.utils.clip_grad_norm_(model.parameters(), 0.5)
      optimizer.step()

      if (i + 1) % log_interval == 0:
        with torch.no_grad():
          x, x_mask, y = next(iter(val_loader))
          x = x.to(device)
          x_mask = x_mask.to(device)
          y = y.to(device)

          v_loss = step(x, x_mask, y, pad_token_id, model)
          v_loss = v_loss.item()
```

```
          elapsed = time.time() - start_time
          print('| epoch {:3d} | [{:5d}/{:5d}] | '
                'ms/batch {:5.2f} | '
                'loss {:5.2f} | val loss {:5.2f}'.format(
                  epoch, i, len(train_loader),
                  elapsed * 1000 / log_interval,
                  loss.item(), v_loss))
          start_time = time.time()
          val_loss.append(v_loss)

  return model
```

[ ]:
```
#step26
#function to test the model it writes original and predicted summary in txt file
def testT5(model,tokenizer,test_loader):
  #intialize the empty lists
  predictions = []
  real_og=[]
  pred_op=[]
  c=0
  b=1000
  #for data in test loader
  for i, (input_ids, attention_mask, y) in enumerate(test_loader):
    input_ids = input_ids.to(device)
    attention_mask = attention_mask.to(device)
    y = y.to(device)
    #generate summaries
    #store real and predicted summary in a list and write in txt file
    summaries = model.generate(input_ids=input_ids,␣
↪attention_mask=attention_mask,max_length=10)
    pred = [tokenizer.decode(g, skip_special_tokens=True,␣
↪clean_up_tokenization_spaces=False) for g in summaries]
    real = [tokenizer.decode(g, skip_special_tokens=True,␣
↪clean_up_tokenization_spaces=False) for g in y]
    #this part is used to print output if the size of c is greater than b
    #limited output is print as only 5000 lines can be printed in colab whole␣
↪output is written in a text file
    for pred_sent, real_sent in zip(pred, real):
      if c>b:
        print("Original: {}".format(real_sent))
        print("Predicted: {}".format(pred_sent))
        print("\n")
        b+=b
      real_og.append(real_sent)
      pred_op.append(pred_sent)
      predictions.append(str("pred sentence: " + pred_sent + "\t\t real␣
↪sentence: " + real_sent+"\n"))
```

```
    c+=1
  file1 = open("/content/drive/MyDrive/TFIVE.txt","w")
  file1.writelines(predictions)
  file1.close()
  #calculate scores
  bleau=compute_bleu(real_og,pred_op, max_order=4,smooth=False)
  bscore=nltk.translate.bleu_score.corpus_bleu(real_og,pred_op)
  rougen=rouge_n(pred_op, real_og, n=2)
  ro=rouge(pred_op, real_og)

  print("bleu, precisions, bp, ratio, translation_length, reference_length",bleau)
  print("bleau score",bscore)
  print("rouge2",rougen)
  print("rouge",ro)
```

[ ]:
```
#step27
#fucntion to get the data and call all the functions in a squence
def tf5token():
  class MyDataset(torch.utils.data.Dataset):
    def __init__(self, articles, highlights):
      self.x = articles
      self.y = highlights

    def __getitem__(self,index):
      x = tokenizer.encode_plus(model.config.prefix + str(self.x[index]), max_length=80, return_tensors="pt", pad_to_max_length=True)
      y = tokenizer.encode(str(self.y[index]), max_length=10, return_tensors="pt", pad_to_max_length=True)
      return x['input_ids'].view(-1), x['attention_mask'].view(-1), y.view(-1)

    def __len__(self):
      return len(self.x)

  #get the data
  x_tr,y_tr,x_tt,y_tt,x_val,y_val=combining_all_steps_t5()
  BATCH_SIZE = 128
  SHUFFEL_SIZE = 1024
  EPOCHS = 150
  log_interval = 200
  #get the pretrained model t5-small
  tokenizer = T5Tokenizer.from_pretrained('t5-small')
  model = T5ForConditionalGeneration.from_pretrained('t5-small').to(device)

  task_specific_params = model.config.task_specific_params
  if task_specific_params is not None:
    model.config.update(task_specific_params.get("summarization", {}))
```

```
#create train,test and validation datasets
train_ds = MyDataset(x_tr["reviewText"],y_tr["summary"])
val_ds = MyDataset(x_val["reviewText"],y_val["summary"])
test_ds = MyDataset(x_tt["reviewText"],y_tt["summary"])

train_loader = torch.utils.data.DataLoader(train_ds, batch_size=BATCH_SIZE)
val_loader = torch.utils.data.DataLoader(val_ds, batch_size=BATCH_SIZE)
test_loader = torch.utils.data.DataLoader(test_ds, batch_size=BATCH_SIZE)

x, x_mask, y = next(iter(val_loader))
print(x.shape, x_mask.shape, y.shape)
pad_token_id = tokenizer.pad_token_id

#call the train function
model=t5train(train_loader,val_loader,pad_token_id,model,EPOCHS,log_interval)
#call the test function
testT5(model,tokenizer,test_loader)
```

[ ]:
```
tf5token()
```

```
The length of dataset is  147799
train 103459, val 22170, test 22170
torch.Size([128, 80]) torch.Size([128, 80]) torch.Size([128, 10])
| epoch   0 | [  199/  809] | ms/batch 247.76 | loss  4.32 | val loss  4.37
| epoch   0 | [  399/  809] | ms/batch 248.13 | loss  3.74 | val loss  4.19
| epoch   0 | [  599/  809] | ms/batch 248.59 | loss  4.19 | val loss  4.01
| epoch   0 | [  799/  809] | ms/batch 247.97 | loss  3.83 | val loss  3.94
| epoch   1 | [  199/  809] | ms/batch 248.90 | loss  3.82 | val loss  3.87
| epoch   1 | [  399/  809] | ms/batch 249.66 | loss  3.50 | val loss  3.90
| epoch   1 | [  599/  809] | ms/batch 249.99 | loss  3.95 | val loss  3.82
| epoch   1 | [  799/  809] | ms/batch 250.72 | loss  3.66 | val loss  3.79
| epoch   2 | [  199/  809] | ms/batch 251.27 | loss  3.67 | val loss  3.76
| epoch   2 | [  399/  809] | ms/batch 250.84 | loss  3.42 | val loss  3.70
| epoch   2 | [  599/  809] | ms/batch 251.13 | loss  3.73 | val loss  3.77
| epoch   2 | [  799/  809] | ms/batch 250.01 | loss  3.54 | val loss  3.71
| epoch   3 | [  199/  809] | ms/batch 250.65 | loss  3.59 | val loss  3.73
| epoch   3 | [  399/  809] | ms/batch 250.62 | loss  3.37 | val loss  3.65
| epoch   3 | [  599/  809] | ms/batch 250.53 | loss  3.67 | val loss  3.66
| epoch   3 | [  799/  809] | ms/batch 250.37 | loss  3.46 | val loss  3.70
| epoch   4 | [  199/  809] | ms/batch 250.74 | loss  3.48 | val loss  3.58
| epoch   4 | [  399/  809] | ms/batch 250.33 | loss  3.30 | val loss  3.59
| epoch   4 | [  599/  809] | ms/batch 250.53 | loss  3.60 | val loss  3.61
| epoch   4 | [  799/  809] | ms/batch 249.99 | loss  3.38 | val loss  3.62
| epoch   5 | [  199/  809] | ms/batch 250.36 | loss  3.38 | val loss  3.62
| epoch   5 | [  399/  809] | ms/batch 249.20 | loss  3.19 | val loss  3.63
| epoch   5 | [  599/  809] | ms/batch 249.51 | loss  3.53 | val loss  3.58
```

```
| epoch   5 | [  799/  809] | ms/batch 249.32 | loss  3.39 | val loss  3.54
| epoch   6 | [  199/  809] | ms/batch 250.17 | loss  3.32 | val loss  3.61
| epoch   6 | [  399/  809] | ms/batch 249.12 | loss  3.11 | val loss  3.55
| epoch   6 | [  599/  809] | ms/batch 251.31 | loss  3.45 | val loss  3.52
| epoch   6 | [  799/  809] | ms/batch 250.64 | loss  3.29 | val loss  3.47
| epoch   7 | [  199/  809] | ms/batch 249.94 | loss  3.32 | val loss  3.47
| epoch   7 | [  399/  809] | ms/batch 249.94 | loss  3.12 | val loss  3.44
| epoch   7 | [  599/  809] | ms/batch 249.65 | loss  3.42 | val loss  3.52
| epoch   7 | [  799/  809] | ms/batch 250.05 | loss  3.28 | val loss  3.41
| epoch   8 | [  199/  809] | ms/batch 250.14 | loss  3.18 | val loss  3.44
| epoch   8 | [  399/  809] | ms/batch 250.49 | loss  3.06 | val loss  3.49
| epoch   8 | [  599/  809] | ms/batch 250.25 | loss  3.34 | val loss  3.41
| epoch   8 | [  799/  809] | ms/batch 249.93 | loss  3.24 | val loss  3.46
| epoch   9 | [  199/  809] | ms/batch 249.50 | loss  3.20 | val loss  3.37
| epoch   9 | [  399/  809] | ms/batch 249.73 | loss  3.02 | val loss  3.32
| epoch   9 | [  599/  809] | ms/batch 249.70 | loss  3.26 | val loss  3.46
| epoch   9 | [  799/  809] | ms/batch 250.00 | loss  3.25 | val loss  3.48
| epoch  10 | [  199/  809] | ms/batch 250.51 | loss  3.14 | val loss  3.45
| epoch  10 | [  399/  809] | ms/batch 249.99 | loss  2.96 | val loss  3.42
| epoch  10 | [  599/  809] | ms/batch 249.54 | loss  3.27 | val loss  3.37
| epoch  10 | [  799/  809] | ms/batch 250.41 | loss  3.13 | val loss  3.39
| epoch  11 | [  199/  809] | ms/batch 250.59 | loss  3.10 | val loss  3.38
| epoch  11 | [  399/  809] | ms/batch 249.91 | loss  2.84 | val loss  3.35
| epoch  11 | [  599/  809] | ms/batch 250.52 | loss  3.22 | val loss  3.41
| epoch  11 | [  799/  809] | ms/batch 250.03 | loss  3.21 | val loss  3.36
| epoch  12 | [  199/  809] | ms/batch 249.50 | loss  3.11 | val loss  3.32
| epoch  12 | [  399/  809] | ms/batch 250.16 | loss  2.88 | val loss  3.34
| epoch  12 | [  599/  809] | ms/batch 250.44 | loss  3.21 | val loss  3.27
| epoch  12 | [  799/  809] | ms/batch 249.76 | loss  3.09 | val loss  3.39
| epoch  13 | [  199/  809] | ms/batch 249.93 | loss  3.02 | val loss  3.35
| epoch  13 | [  399/  809] | ms/batch 248.74 | loss  2.84 | val loss  3.38
| epoch  13 | [  599/  809] | ms/batch 249.41 | loss  3.18 | val loss  3.40
| epoch  13 | [  799/  809] | ms/batch 249.34 | loss  3.11 | val loss  3.37
| epoch  14 | [  199/  809] | ms/batch 249.34 | loss  3.00 | val loss  3.39
| epoch  14 | [  399/  809] | ms/batch 249.09 | loss  2.74 | val loss  3.41
| epoch  14 | [  599/  809] | ms/batch 249.66 | loss  3.20 | val loss  3.38
| epoch  14 | [  799/  809] | ms/batch 251.43 | loss  3.08 | val loss  3.38
| epoch  15 | [  199/  809] | ms/batch 251.38 | loss  2.97 | val loss  3.37
| epoch  15 | [  399/  809] | ms/batch 250.15 | loss  2.76 | val loss  3.39
| epoch  15 | [  599/  809] | ms/batch 249.78 | loss  3.09 | val loss  3.37
| epoch  15 | [  799/  809] | ms/batch 249.48 | loss  3.04 | val loss  3.35
| epoch  16 | [  199/  809] | ms/batch 250.30 | loss  2.92 | val loss  3.42
| epoch  16 | [  399/  809] | ms/batch 249.65 | loss  2.69 | val loss  3.35
| epoch  16 | [  599/  809] | ms/batch 249.92 | loss  3.04 | val loss  3.39
| epoch  16 | [  799/  809] | ms/batch 250.85 | loss  3.06 | val loss  3.35
| epoch  17 | [  199/  809] | ms/batch 250.13 | loss  2.89 | val loss  3.42
| epoch  17 | [  399/  809] | ms/batch 250.30 | loss  2.70 | val loss  3.37
| epoch  17 | [  599/  809] | ms/batch 249.76 | loss  3.03 | val loss  3.32
```

```
| epoch  17 | [  799/  809] | ms/batch 250.98 | loss  2.97 | val loss  3.39
| epoch  18 | [  199/  809] | ms/batch 249.68 | loss  2.90 | val loss  3.32
| epoch  18 | [  399/  809] | ms/batch 247.22 | loss  2.67 | val loss  3.48
| epoch  18 | [  599/  809] | ms/batch 249.25 | loss  3.01 | val loss  3.36
| epoch  18 | [  799/  809] | ms/batch 249.10 | loss  3.01 | val loss  3.46
| epoch  19 | [  199/  809] | ms/batch 247.67 | loss  2.93 | val loss  3.30
| epoch  19 | [  399/  809] | ms/batch 249.34 | loss  2.70 | val loss  3.39
| epoch  19 | [  599/  809] | ms/batch 248.97 | loss  3.03 | val loss  3.32
| epoch  19 | [  799/  809] | ms/batch 249.18 | loss  2.95 | val loss  3.30
| epoch  20 | [  199/  809] | ms/batch 249.53 | loss  2.79 | val loss  3.38
| epoch  20 | [  399/  809] | ms/batch 249.33 | loss  2.58 | val loss  3.38
| epoch  20 | [  599/  809] | ms/batch 249.66 | loss  2.92 | val loss  3.32
| epoch  20 | [  799/  809] | ms/batch 249.57 | loss  2.96 | val loss  3.40
| epoch  21 | [  199/  809] | ms/batch 248.95 | loss  2.81 | val loss  3.43
| epoch  21 | [  399/  809] | ms/batch 248.85 | loss  2.64 | val loss  3.47
| epoch  21 | [  599/  809] | ms/batch 248.73 | loss  2.94 | val loss  3.38
| epoch  21 | [  799/  809] | ms/batch 247.84 | loss  2.90 | val loss  3.31
| epoch  22 | [  199/  809] | ms/batch 248.49 | loss  2.81 | val loss  3.38
| epoch  22 | [  399/  809] | ms/batch 248.16 | loss  2.50 | val loss  3.36
| epoch  22 | [  599/  809] | ms/batch 247.28 | loss  2.90 | val loss  3.43
| epoch  22 | [  799/  809] | ms/batch 247.45 | loss  2.92 | val loss  3.45
| epoch  23 | [  199/  809] | ms/batch 247.13 | loss  2.77 | val loss  3.43
| epoch  23 | [  399/  809] | ms/batch 247.20 | loss  2.53 | val loss  3.37
| epoch  23 | [  599/  809] | ms/batch 247.71 | loss  2.94 | val loss  3.37
| epoch  23 | [  799/  809] | ms/batch 250.64 | loss  2.87 | val loss  3.44
| epoch  24 | [  199/  809] | ms/batch 248.72 | loss  2.72 | val loss  3.48
| epoch  24 | [  399/  809] | ms/batch 248.19 | loss  2.52 | val loss  3.40
| epoch  24 | [  599/  809] | ms/batch 250.84 | loss  2.94 | val loss  3.37
| epoch  24 | [  799/  809] | ms/batch 251.39 | loss  2.80 | val loss  3.38
| epoch  25 | [  199/  809] | ms/batch 250.62 | loss  2.71 | val loss  3.43
| epoch  25 | [  399/  809] | ms/batch 250.08 | loss  2.42 | val loss  3.45
| epoch  25 | [  599/  809] | ms/batch 249.37 | loss  2.79 | val loss  3.34
| epoch  25 | [  799/  809] | ms/batch 250.09 | loss  2.87 | val loss  3.39
| epoch  26 | [  199/  809] | ms/batch 249.45 | loss  2.61 | val loss  3.38
| epoch  26 | [  399/  809] | ms/batch 248.94 | loss  2.43 | val loss  3.34
| epoch  26 | [  599/  809] | ms/batch 248.98 | loss  2.80 | val loss  3.41
| epoch  26 | [  799/  809] | ms/batch 248.05 | loss  2.80 | val loss  3.40
| epoch  27 | [  199/  809] | ms/batch 249.69 | loss  2.65 | val loss  3.41
| epoch  27 | [  399/  809] | ms/batch 249.15 | loss  2.43 | val loss  3.46
| epoch  27 | [  599/  809] | ms/batch 247.87 | loss  2.73 | val loss  3.39
| epoch  27 | [  799/  809] | ms/batch 247.51 | loss  2.80 | val loss  3.47
| epoch  28 | [  199/  809] | ms/batch 250.18 | loss  2.56 | val loss  3.35
| epoch  28 | [  399/  809] | ms/batch 251.69 | loss  2.43 | val loss  3.40
| epoch  28 | [  599/  809] | ms/batch 249.91 | loss  2.69 | val loss  3.45
| epoch  28 | [  799/  809] | ms/batch 250.29 | loss  2.73 | val loss  3.55
| epoch  29 | [  199/  809] | ms/batch 249.72 | loss  2.54 | val loss  3.44
| epoch  29 | [  399/  809] | ms/batch 249.77 | loss  2.37 | val loss  3.41
| epoch  29 | [  599/  809] | ms/batch 248.94 | loss  2.72 | val loss  3.42
```

```
| epoch  29 | [  799/  809] | ms/batch 248.99 | loss  2.74 | val loss  3.38
| epoch  30 | [  199/  809] | ms/batch 248.36 | loss  2.63 | val loss  3.38
| epoch  30 | [  399/  809] | ms/batch 248.94 | loss  2.34 | val loss  3.43
| epoch  30 | [  599/  809] | ms/batch 248.14 | loss  2.75 | val loss  3.56
| epoch  30 | [  799/  809] | ms/batch 248.70 | loss  2.70 | val loss  3.53
| epoch  31 | [  199/  809] | ms/batch 248.19 | loss  2.62 | val loss  3.38
| epoch  31 | [  399/  809] | ms/batch 249.17 | loss  2.31 | val loss  3.52
| epoch  31 | [  599/  809] | ms/batch 248.82 | loss  2.70 | val loss  3.43
| epoch  31 | [  799/  809] | ms/batch 249.40 | loss  2.70 | val loss  3.42
| epoch  32 | [  199/  809] | ms/batch 249.37 | loss  2.48 | val loss  3.47
| epoch  32 | [  399/  809] | ms/batch 249.09 | loss  2.24 | val loss  3.49
| epoch  32 | [  599/  809] | ms/batch 251.04 | loss  2.67 | val loss  3.47
| epoch  32 | [  799/  809] | ms/batch 250.67 | loss  2.73 | val loss  3.51
| epoch  33 | [  199/  809] | ms/batch 249.22 | loss  2.55 | val loss  3.40
| epoch  33 | [  399/  809] | ms/batch 249.69 | loss  2.22 | val loss  3.54
| epoch  33 | [  599/  809] | ms/batch 249.25 | loss  2.65 | val loss  3.26
| epoch  33 | [  799/  809] | ms/batch 248.94 | loss  2.71 | val loss  3.36
| epoch  34 | [  199/  809] | ms/batch 250.17 | loss  2.43 | val loss  3.39
| epoch  34 | [  399/  809] | ms/batch 250.01 | loss  2.17 | val loss  3.50
| epoch  34 | [  599/  809] | ms/batch 249.54 | loss  2.70 | val loss  3.38
| epoch  34 | [  799/  809] | ms/batch 249.44 | loss  2.61 | val loss  3.47
| epoch  35 | [  199/  809] | ms/batch 249.45 | loss  2.46 | val loss  3.50
| epoch  35 | [  399/  809] | ms/batch 249.14 | loss  2.27 | val loss  3.47
| epoch  35 | [  599/  809] | ms/batch 249.26 | loss  2.65 | val loss  3.43
| epoch  35 | [  799/  809] | ms/batch 249.05 | loss  2.62 | val loss  3.44
| epoch  36 | [  199/  809] | ms/batch 248.89 | loss  2.44 | val loss  3.45
| epoch  36 | [  399/  809] | ms/batch 250.10 | loss  2.28 | val loss  3.52
| epoch  36 | [  599/  809] | ms/batch 249.18 | loss  2.61 | val loss  3.42
| epoch  36 | [  799/  809] | ms/batch 250.16 | loss  2.60 | val loss  3.52
| epoch  37 | [  199/  809] | ms/batch 250.17 | loss  2.44 | val loss  3.40
| epoch  37 | [  399/  809] | ms/batch 251.38 | loss  2.14 | val loss  3.43
| epoch  37 | [  599/  809] | ms/batch 251.87 | loss  2.53 | val loss  3.58
| epoch  37 | [  799/  809] | ms/batch 251.01 | loss  2.62 | val loss  3.53
| epoch  38 | [  199/  809] | ms/batch 251.74 | loss  2.40 | val loss  3.47
| epoch  38 | [  399/  809] | ms/batch 251.53 | loss  2.14 | val loss  3.46
| epoch  38 | [  599/  809] | ms/batch 249.59 | loss  2.63 | val loss  3.48
| epoch  38 | [  799/  809] | ms/batch 250.49 | loss  2.52 | val loss  3.58
| epoch  39 | [  199/  809] | ms/batch 249.21 | loss  2.47 | val loss  3.52
| epoch  39 | [  399/  809] | ms/batch 249.33 | loss  2.01 | val loss  3.44
| epoch  39 | [  599/  809] | ms/batch 249.59 | loss  2.52 | val loss  3.48
| epoch  39 | [  799/  809] | ms/batch 249.37 | loss  2.53 | val loss  3.41
| epoch  40 | [  199/  809] | ms/batch 248.87 | loss  2.44 | val loss  3.47
| epoch  40 | [  399/  809] | ms/batch 248.40 | loss  2.08 | val loss  3.55
| epoch  40 | [  599/  809] | ms/batch 249.23 | loss  2.55 | val loss  3.43
| epoch  40 | [  799/  809] | ms/batch 249.45 | loss  2.51 | val loss  3.59
| epoch  41 | [  199/  809] | ms/batch 249.50 | loss  2.36 | val loss  3.50
| epoch  41 | [  399/  809] | ms/batch 249.26 | loss  2.11 | val loss  3.50
| epoch  41 | [  599/  809] | ms/batch 248.74 | loss  2.48 | val loss  3.45
```

```
| epoch  41 | [  799/  809] | ms/batch 247.65 | loss  2.50 | val loss  3.46
| epoch  42 | [  199/  809] | ms/batch 249.37 | loss  2.31 | val loss  3.52
| epoch  42 | [  399/  809] | ms/batch 249.26 | loss  2.03 | val loss  3.57
| epoch  42 | [  599/  809] | ms/batch 248.05 | loss  2.50 | val loss  3.59
| epoch  42 | [  799/  809] | ms/batch 247.18 | loss  2.43 | val loss  3.45
| epoch  43 | [  199/  809] | ms/batch 248.37 | loss  2.31 | val loss  3.52
| epoch  43 | [  399/  809] | ms/batch 248.68 | loss  1.92 | val loss  3.50
| epoch  43 | [  599/  809] | ms/batch 249.33 | loss  2.47 | val loss  3.57
| epoch  43 | [  799/  809] | ms/batch 249.54 | loss  2.44 | val loss  3.57
| epoch  44 | [  199/  809] | ms/batch 248.11 | loss  2.27 | val loss  3.55
| epoch  44 | [  399/  809] | ms/batch 251.18 | loss  2.01 | val loss  3.56
| epoch  44 | [  599/  809] | ms/batch 249.54 | loss  2.42 | val loss  3.55
| epoch  44 | [  799/  809] | ms/batch 249.55 | loss  2.45 | val loss  3.57
| epoch  45 | [  199/  809] | ms/batch 250.87 | loss  2.28 | val loss  3.65
| epoch  45 | [  399/  809] | ms/batch 248.98 | loss  2.01 | val loss  3.68
| epoch  45 | [  599/  809] | ms/batch 249.72 | loss  2.33 | val loss  3.51
| epoch  45 | [  799/  809] | ms/batch 249.42 | loss  2.36 | val loss  3.43
| epoch  46 | [  199/  809] | ms/batch 248.63 | loss  2.31 | val loss  3.50
| epoch  46 | [  399/  809] | ms/batch 248.87 | loss  1.94 | val loss  3.47
| epoch  46 | [  599/  809] | ms/batch 248.78 | loss  2.33 | val loss  3.52
| epoch  46 | [  799/  809] | ms/batch 248.61 | loss  2.37 | val loss  3.49
| epoch  47 | [  199/  809] | ms/batch 248.30 | loss  2.22 | val loss  3.51
| epoch  47 | [  399/  809] | ms/batch 250.30 | loss  1.94 | val loss  3.51
| epoch  47 | [  599/  809] | ms/batch 250.47 | loss  2.43 | val loss  3.63
| epoch  47 | [  799/  809] | ms/batch 249.65 | loss  2.40 | val loss  3.64
| epoch  48 | [  199/  809] | ms/batch 248.59 | loss  2.20 | val loss  3.44
| epoch  48 | [  399/  809] | ms/batch 248.36 | loss  1.96 | val loss  3.65
| epoch  48 | [  599/  809] | ms/batch 250.04 | loss  2.41 | val loss  3.64
| epoch  48 | [  799/  809] | ms/batch 251.69 | loss  2.35 | val loss  3.49
| epoch  49 | [  199/  809] | ms/batch 251.22 | loss  2.20 | val loss  3.67
| epoch  49 | [  399/  809] | ms/batch 250.87 | loss  1.97 | val loss  3.49
| epoch  49 | [  599/  809] | ms/batch 251.03 | loss  2.32 | val loss  3.68
| epoch  49 | [  799/  809] | ms/batch 251.03 | loss  2.31 | val loss  3.62
| epoch  50 | [  199/  809] | ms/batch 250.60 | loss  2.14 | val loss  3.64
| epoch  50 | [  399/  809] | ms/batch 250.77 | loss  1.88 | val loss  3.74
| epoch  50 | [  599/  809] | ms/batch 251.67 | loss  2.30 | val loss  3.64
| epoch  50 | [  799/  809] | ms/batch 250.49 | loss  2.30 | val loss  3.61
| epoch  51 | [  199/  809] | ms/batch 251.49 | loss  2.16 | val loss  3.60
| epoch  51 | [  399/  809] | ms/batch 251.12 | loss  1.83 | val loss  3.63
| epoch  51 | [  599/  809] | ms/batch 250.99 | loss  2.34 | val loss  3.64
| epoch  51 | [  799/  809] | ms/batch 250.68 | loss  2.30 | val loss  3.59
| epoch  52 | [  199/  809] | ms/batch 250.85 | loss  2.12 | val loss  3.64
| epoch  52 | [  399/  809] | ms/batch 250.44 | loss  1.83 | val loss  3.63
| epoch  52 | [  599/  809] | ms/batch 251.36 | loss  2.22 | val loss  3.63
| epoch  52 | [  799/  809] | ms/batch 250.84 | loss  2.29 | val loss  3.72
| epoch  53 | [  199/  809] | ms/batch 250.80 | loss  2.12 | val loss  3.56
| epoch  53 | [  399/  809] | ms/batch 250.53 | loss  1.76 | val loss  3.63
| epoch  53 | [  599/  809] | ms/batch 250.97 | loss  2.22 | val loss  3.60
```

```
| epoch  53 | [  799/  809] | ms/batch 250.99 | loss  2.26 | val loss  3.64
| epoch  54 | [  199/  809] | ms/batch 251.39 | loss  2.14 | val loss  3.59
| epoch  54 | [  399/  809] | ms/batch 250.82 | loss  1.79 | val loss  3.63
| epoch  54 | [  599/  809] | ms/batch 251.09 | loss  2.26 | val loss  3.73
| epoch  54 | [  799/  809] | ms/batch 251.81 | loss  2.28 | val loss  3.59
| epoch  55 | [  199/  809] | ms/batch 250.94 | loss  2.05 | val loss  3.66
| epoch  55 | [  399/  809] | ms/batch 251.10 | loss  1.81 | val loss  3.70
| epoch  55 | [  599/  809] | ms/batch 251.43 | loss  2.22 | val loss  3.65
| epoch  55 | [  799/  809] | ms/batch 251.10 | loss  2.30 | val loss  3.74
| epoch  56 | [  199/  809] | ms/batch 251.16 | loss  2.04 | val loss  3.66
| epoch  56 | [  399/  809] | ms/batch 250.41 | loss  1.72 | val loss  3.63
| epoch  56 | [  599/  809] | ms/batch 250.69 | loss  2.18 | val loss  3.68
| epoch  56 | [  799/  809] | ms/batch 251.34 | loss  2.20 | val loss  3.68
| epoch  57 | [  199/  809] | ms/batch 251.36 | loss  2.07 | val loss  3.63
| epoch  57 | [  399/  809] | ms/batch 251.73 | loss  1.72 | val loss  3.71
| epoch  57 | [  599/  809] | ms/batch 251.72 | loss  2.15 | val loss  3.61
| epoch  57 | [  799/  809] | ms/batch 251.35 | loss  2.23 | val loss  3.70
| epoch  58 | [  199/  809] | ms/batch 250.90 | loss  1.92 | val loss  3.64
| epoch  58 | [  399/  809] | ms/batch 250.57 | loss  1.81 | val loss  3.64
| epoch  58 | [  599/  809] | ms/batch 251.80 | loss  2.18 | val loss  3.74
| epoch  58 | [  799/  809] | ms/batch 251.72 | loss  2.14 | val loss  3.75
| epoch  59 | [  199/  809] | ms/batch 252.41 | loss  2.01 | val loss  3.69
| epoch  59 | [  399/  809] | ms/batch 253.04 | loss  1.74 | val loss  3.63
| epoch  59 | [  599/  809] | ms/batch 252.82 | loss  2.10 | val loss  3.70
| epoch  59 | [  799/  809] | ms/batch 252.12 | loss  2.18 | val loss  3.75
| epoch  60 | [  199/  809] | ms/batch 252.27 | loss  2.00 | val loss  3.75
| epoch  60 | [  399/  809] | ms/batch 250.79 | loss  1.76 | val loss  3.70
| epoch  60 | [  599/  809] | ms/batch 250.54 | loss  2.04 | val loss  3.87
| epoch  60 | [  799/  809] | ms/batch 250.77 | loss  2.19 | val loss  3.80
| epoch  61 | [  199/  809] | ms/batch 250.78 | loss  1.95 | val loss  3.68
| epoch  61 | [  399/  809] | ms/batch 251.30 | loss  1.71 | val loss  3.76
| epoch  61 | [  599/  809] | ms/batch 251.03 | loss  2.20 | val loss  3.85
| epoch  61 | [  799/  809] | ms/batch 250.74 | loss  2.14 | val loss  3.56
| epoch  62 | [  199/  809] | ms/batch 250.92 | loss  1.95 | val loss  3.63
| epoch  62 | [  399/  809] | ms/batch 250.84 | loss  1.63 | val loss  3.71
| epoch  62 | [  599/  809] | ms/batch 250.54 | loss  2.08 | val loss  3.75
| epoch  62 | [  799/  809] | ms/batch 250.86 | loss  2.10 | val loss  3.59
| epoch  63 | [  199/  809] | ms/batch 251.04 | loss  2.00 | val loss  3.67
| epoch  63 | [  399/  809] | ms/batch 250.26 | loss  1.67 | val loss  3.78
| epoch  63 | [  599/  809] | ms/batch 251.14 | loss  2.03 | val loss  3.83
| epoch  63 | [  799/  809] | ms/batch 251.67 | loss  2.05 | val loss  3.77
| epoch  64 | [  199/  809] | ms/batch 250.88 | loss  1.86 | val loss  3.73
| epoch  64 | [  399/  809] | ms/batch 250.49 | loss  1.66 | val loss  3.64
| epoch  64 | [  599/  809] | ms/batch 250.24 | loss  2.07 | val loss  3.79
| epoch  64 | [  799/  809] | ms/batch 250.18 | loss  1.99 | val loss  3.81
| epoch  65 | [  199/  809] | ms/batch 251.22 | loss  1.98 | val loss  3.76
| epoch  65 | [  399/  809] | ms/batch 250.88 | loss  1.58 | val loss  3.85
| epoch  65 | [  599/  809] | ms/batch 250.88 | loss  2.08 | val loss  3.87
```

```
| epoch  65 | [  799/  809] | ms/batch 252.70 | loss  2.05 | val loss  3.69
| epoch  66 | [  199/  809] | ms/batch 250.68 | loss  1.86 | val loss  3.96
| epoch  66 | [  399/  809] | ms/batch 249.50 | loss  1.59 | val loss  3.81
| epoch  66 | [  599/  809] | ms/batch 250.48 | loss  1.99 | val loss  3.86
| epoch  66 | [  799/  809] | ms/batch 249.53 | loss  2.02 | val loss  3.80
| epoch  67 | [  199/  809] | ms/batch 249.97 | loss  1.84 | val loss  3.68
| epoch  67 | [  399/  809] | ms/batch 250.17 | loss  1.62 | val loss  3.68
| epoch  67 | [  599/  809] | ms/batch 250.41 | loss  1.95 | val loss  3.90
| epoch  67 | [  799/  809] | ms/batch 249.55 | loss  1.95 | val loss  3.93
| epoch  68 | [  199/  809] | ms/batch 249.51 | loss  1.83 | val loss  3.73
| epoch  68 | [  399/  809] | ms/batch 250.31 | loss  1.59 | val loss  3.84
| epoch  68 | [  599/  809] | ms/batch 249.36 | loss  1.97 | val loss  3.92
| epoch  68 | [  799/  809] | ms/batch 249.80 | loss  1.92 | val loss  3.84
| epoch  69 | [  199/  809] | ms/batch 249.87 | loss  1.78 | val loss  3.74
| epoch  69 | [  399/  809] | ms/batch 249.25 | loss  1.47 | val loss  3.85
| epoch  69 | [  599/  809] | ms/batch 249.53 | loss  1.94 | val loss  3.93
| epoch  69 | [  799/  809] | ms/batch 250.05 | loss  1.98 | val loss  3.91
| epoch  70 | [  199/  809] | ms/batch 249.67 | loss  1.72 | val loss  3.83
| epoch  70 | [  399/  809] | ms/batch 249.57 | loss  1.59 | val loss  3.83
| epoch  70 | [  599/  809] | ms/batch 250.20 | loss  1.92 | val loss  3.95
| epoch  70 | [  799/  809] | ms/batch 249.19 | loss  1.94 | val loss  3.88
| epoch  71 | [  199/  809] | ms/batch 249.53 | loss  1.75 | val loss  3.84
| epoch  71 | [  399/  809] | ms/batch 249.53 | loss  1.52 | val loss  3.86
| epoch  71 | [  599/  809] | ms/batch 249.23 | loss  1.99 | val loss  3.97
| epoch  71 | [  799/  809] | ms/batch 249.74 | loss  1.91 | val loss  3.99
| epoch  72 | [  199/  809] | ms/batch 249.62 | loss  1.71 | val loss  3.96
| epoch  72 | [  399/  809] | ms/batch 250.31 | loss  1.49 | val loss  4.03
| epoch  72 | [  599/  809] | ms/batch 250.12 | loss  1.88 | val loss  3.86
| epoch  72 | [  799/  809] | ms/batch 251.47 | loss  1.90 | val loss  3.94
| epoch  73 | [  199/  809] | ms/batch 251.42 | loss  1.75 | val loss  3.96
| epoch  73 | [  399/  809] | ms/batch 251.33 | loss  1.46 | val loss  3.97
| epoch  73 | [  599/  809] | ms/batch 250.63 | loss  1.87 | val loss  3.93
| epoch  73 | [  799/  809] | ms/batch 250.89 | loss  1.95 | val loss  4.00
| epoch  74 | [  199/  809] | ms/batch 250.82 | loss  1.68 | val loss  3.89
| epoch  74 | [  399/  809] | ms/batch 251.37 | loss  1.54 | val loss  3.97
| epoch  74 | [  599/  809] | ms/batch 251.59 | loss  1.83 | val loss  3.92
| epoch  74 | [  799/  809] | ms/batch 251.05 | loss  1.82 | val loss  3.94
| epoch  75 | [  199/  809] | ms/batch 251.75 | loss  1.65 | val loss  3.95
| epoch  75 | [  399/  809] | ms/batch 251.72 | loss  1.46 | val loss  3.87
| epoch  75 | [  599/  809] | ms/batch 251.00 | loss  1.77 | val loss  3.99
| epoch  75 | [  799/  809] | ms/batch 250.83 | loss  1.78 | val loss  4.07
| epoch  76 | [  199/  809] | ms/batch 250.92 | loss  1.79 | val loss  3.85
| epoch  76 | [  399/  809] | ms/batch 251.62 | loss  1.53 | val loss  3.98
| epoch  76 | [  599/  809] | ms/batch 251.40 | loss  1.73 | val loss  3.94
| epoch  76 | [  799/  809] | ms/batch 251.58 | loss  1.76 | val loss  3.80
| epoch  77 | [  199/  809] | ms/batch 252.01 | loss  1.64 | val loss  3.93
| epoch  77 | [  399/  809] | ms/batch 252.35 | loss  1.42 | val loss  3.89
| epoch  77 | [  599/  809] | ms/batch 250.78 | loss  1.71 | val loss  4.13
```

```
| epoch  77 | [  799/  809] | ms/batch 250.63 | loss  1.81 | val loss  4.06
| epoch  78 | [  199/  809] | ms/batch 250.29 | loss  1.63 | val loss  3.93
| epoch  78 | [  399/  809] | ms/batch 249.55 | loss  1.44 | val loss  4.15
| epoch  78 | [  599/  809] | ms/batch 250.80 | loss  1.75 | val loss  4.01
| epoch  78 | [  799/  809] | ms/batch 251.55 | loss  1.77 | val loss  4.00
| epoch  79 | [  199/  809] | ms/batch 252.15 | loss  1.54 | val loss  4.00
| epoch  79 | [  399/  809] | ms/batch 253.33 | loss  1.32 | val loss  3.87
| epoch  79 | [  599/  809] | ms/batch 253.50 | loss  1.66 | val loss  3.78
| epoch  79 | [  799/  809] | ms/batch 252.37 | loss  1.74 | val loss  3.97
| epoch  80 | [  199/  809] | ms/batch 251.57 | loss  1.71 | val loss  3.96
| epoch  80 | [  399/  809] | ms/batch 250.72 | loss  1.30 | val loss  3.86
| epoch  80 | [  599/  809] | ms/batch 251.33 | loss  1.71 | val loss  4.08
| epoch  80 | [  799/  809] | ms/batch 252.12 | loss  1.78 | val loss  4.02
| epoch  81 | [  199/  809] | ms/batch 251.33 | loss  1.61 | val loss  4.03
| epoch  81 | [  399/  809] | ms/batch 250.66 | loss  1.37 | val loss  4.14
| epoch  81 | [  599/  809] | ms/batch 251.60 | loss  1.64 | val loss  3.93
| epoch  81 | [  799/  809] | ms/batch 251.08 | loss  1.73 | val loss  4.02
| epoch  82 | [  199/  809] | ms/batch 251.74 | loss  1.54 | val loss  4.02
| epoch  82 | [  399/  809] | ms/batch 250.84 | loss  1.30 | val loss  4.11
| epoch  82 | [  599/  809] | ms/batch 251.03 | loss  1.68 | val loss  4.00
| epoch  82 | [  799/  809] | ms/batch 251.12 | loss  1.73 | val loss  3.99
| epoch  83 | [  199/  809] | ms/batch 250.99 | loss  1.51 | val loss  4.09
| epoch  83 | [  399/  809] | ms/batch 251.23 | loss  1.36 | val loss  3.99
| epoch  83 | [  599/  809] | ms/batch 251.55 | loss  1.67 | val loss  4.11
| epoch  83 | [  799/  809] | ms/batch 251.49 | loss  1.74 | val loss  4.15
| epoch  84 | [  199/  809] | ms/batch 251.66 | loss  1.46 | val loss  4.18
| epoch  84 | [  399/  809] | ms/batch 251.31 | loss  1.27 | val loss  4.08
| epoch  84 | [  599/  809] | ms/batch 252.29 | loss  1.60 | val loss  4.13
| epoch  84 | [  799/  809] | ms/batch 251.58 | loss  1.61 | val loss  4.06
| epoch  85 | [  199/  809] | ms/batch 251.81 | loss  1.47 | val loss  4.08
| epoch  85 | [  399/  809] | ms/batch 251.35 | loss  1.30 | val loss  4.19
| epoch  85 | [  599/  809] | ms/batch 251.59 | loss  1.53 | val loss  4.10
| epoch  85 | [  799/  809] | ms/batch 251.28 | loss  1.67 | val loss  4.20
| epoch  86 | [  199/  809] | ms/batch 252.41 | loss  1.44 | val loss  4.23
| epoch  86 | [  399/  809] | ms/batch 251.66 | loss  1.33 | val loss  4.16
| epoch  86 | [  599/  809] | ms/batch 252.09 | loss  1.59 | val loss  4.19
| epoch  86 | [  799/  809] | ms/batch 251.43 | loss  1.65 | val loss  4.20
| epoch  87 | [  199/  809] | ms/batch 251.34 | loss  1.44 | val loss  4.07
| epoch  87 | [  399/  809] | ms/batch 251.70 | loss  1.16 | val loss  4.12
| epoch  87 | [  599/  809] | ms/batch 251.70 | loss  1.58 | val loss  4.15
| epoch  87 | [  799/  809] | ms/batch 251.56 | loss  1.71 | val loss  4.08
| epoch  88 | [  199/  809] | ms/batch 252.05 | loss  1.47 | val loss  4.17
| epoch  88 | [  399/  809] | ms/batch 251.66 | loss  1.26 | val loss  4.11
| epoch  88 | [  599/  809] | ms/batch 251.84 | loss  1.44 | val loss  4.19
| epoch  88 | [  799/  809] | ms/batch 251.90 | loss  1.63 | val loss  4.17
| epoch  89 | [  199/  809] | ms/batch 251.65 | loss  1.44 | val loss  4.13
| epoch  89 | [  399/  809] | ms/batch 251.53 | loss  1.25 | val loss  4.21
| epoch  89 | [  599/  809] | ms/batch 251.30 | loss  1.47 | val loss  4.16
```

```
| epoch  89 | [  799/  809] | ms/batch 250.90 | loss  1.59 | val loss  4.13
| epoch  90 | [  199/  809] | ms/batch 250.36 | loss  1.37 | val loss  4.07
| epoch  90 | [  399/  809] | ms/batch 251.18 | loss  1.21 | val loss  4.10
| epoch  90 | [  599/  809] | ms/batch 251.68 | loss  1.55 | val loss  4.24
| epoch  90 | [  799/  809] | ms/batch 251.56 | loss  1.64 | val loss  4.24
| epoch  91 | [  199/  809] | ms/batch 251.86 | loss  1.38 | val loss  4.28
| epoch  91 | [  399/  809] | ms/batch 251.44 | loss  1.21 | val loss  4.12
| epoch  91 | [  599/  809] | ms/batch 251.26 | loss  1.57 | val loss  4.21
| epoch  91 | [  799/  809] | ms/batch 250.69 | loss  1.54 | val loss  4.38
| epoch  92 | [  199/  809] | ms/batch 251.13 | loss  1.40 | val loss  4.18
| epoch  92 | [  399/  809] | ms/batch 250.88 | loss  1.18 | val loss  4.17
| epoch  92 | [  599/  809] | ms/batch 251.72 | loss  1.53 | val loss  4.30
| epoch  92 | [  799/  809] | ms/batch 251.44 | loss  1.53 | val loss  4.33
| epoch  93 | [  199/  809] | ms/batch 251.50 | loss  1.39 | val loss  4.31
| epoch  93 | [  399/  809] | ms/batch 251.31 | loss  1.21 | val loss  4.39
| epoch  93 | [  599/  809] | ms/batch 250.86 | loss  1.45 | val loss  4.34
| epoch  93 | [  799/  809] | ms/batch 253.53 | loss  1.53 | val loss  4.32
| epoch  94 | [  199/  809] | ms/batch 250.98 | loss  1.38 | val loss  4.22
| epoch  94 | [  399/  809] | ms/batch 250.44 | loss  1.16 | val loss  4.20
| epoch  94 | [  599/  809] | ms/batch 250.51 | loss  1.44 | val loss  4.27
| epoch  94 | [  799/  809] | ms/batch 251.86 | loss  1.47 | val loss  4.25
| epoch  95 | [  199/  809] | ms/batch 251.20 | loss  1.29 | val loss  4.18
| epoch  95 | [  399/  809] | ms/batch 249.93 | loss  1.13 | val loss  4.40
| epoch  95 | [  599/  809] | ms/batch 249.58 | loss  1.49 | val loss  4.38
| epoch  95 | [  799/  809] | ms/batch 249.34 | loss  1.46 | val loss  4.24
| epoch  96 | [  199/  809] | ms/batch 249.30 | loss  1.39 | val loss  4.16
| epoch  96 | [  399/  809] | ms/batch 249.10 | loss  1.11 | val loss  4.34
| epoch  96 | [  599/  809] | ms/batch 249.12 | loss  1.47 | val loss  4.36
| epoch  96 | [  799/  809] | ms/batch 248.94 | loss  1.49 | val loss  4.39
| epoch  97 | [  199/  809] | ms/batch 249.67 | loss  1.29 | val loss  4.42
| epoch  97 | [  399/  809] | ms/batch 249.54 | loss  1.21 | val loss  4.38
| epoch  97 | [  599/  809] | ms/batch 248.91 | loss  1.41 | val loss  4.36
| epoch  97 | [  799/  809] | ms/batch 248.61 | loss  1.52 | val loss  4.44
| epoch  98 | [  199/  809] | ms/batch 248.58 | loss  1.32 | val loss  4.33
| epoch  98 | [  399/  809] | ms/batch 248.54 | loss  1.09 | val loss  4.44
| epoch  98 | [  599/  809] | ms/batch 249.36 | loss  1.40 | val loss  4.41
| epoch  98 | [  799/  809] | ms/batch 248.79 | loss  1.46 | val loss  4.13
| epoch  99 | [  199/  809] | ms/batch 248.92 | loss  1.25 | val loss  4.37
| epoch  99 | [  399/  809] | ms/batch 249.64 | loss  1.20 | val loss  4.45
| epoch  99 | [  599/  809] | ms/batch 249.12 | loss  1.49 | val loss  4.39
| epoch  99 | [  799/  809] | ms/batch 250.05 | loss  1.36 | val loss  4.29
| epoch 100 | [  199/  809] | ms/batch 253.89 | loss  1.35 | val loss  4.44
| epoch 100 | [  399/  809] | ms/batch 251.82 | loss  1.06 | val loss  4.51
| epoch 100 | [  599/  809] | ms/batch 251.18 | loss  1.40 | val loss  4.42
| epoch 100 | [  799/  809] | ms/batch 251.25 | loss  1.47 | val loss  4.50
| epoch 101 | [  199/  809] | ms/batch 251.93 | loss  1.16 | val loss  4.47
| epoch 101 | [  399/  809] | ms/batch 252.32 | loss  1.09 | val loss  4.49
| epoch 101 | [  599/  809] | ms/batch 250.29 | loss  1.30 | val loss  4.41
```

```
| epoch 101 | [  799/  809] | ms/batch 249.74 | loss  1.50 | val loss  4.46
| epoch 102 | [  199/  809] | ms/batch 250.10 | loss  1.20 | val loss  4.51
| epoch 102 | [  399/  809] | ms/batch 249.39 | loss  1.00 | val loss  4.44
| epoch 102 | [  599/  809] | ms/batch 249.42 | loss  1.32 | val loss  4.73
| epoch 102 | [  799/  809] | ms/batch 249.96 | loss  1.46 | val loss  4.39
| epoch 103 | [  199/  809] | ms/batch 250.28 | loss  1.27 | val loss  4.47
| epoch 103 | [  399/  809] | ms/batch 253.46 | loss  1.03 | val loss  4.40
| epoch 103 | [  599/  809] | ms/batch 252.96 | loss  1.31 | val loss  4.57
| epoch 103 | [  799/  809] | ms/batch 251.83 | loss  1.30 | val loss  4.51
| epoch 104 | [  199/  809] | ms/batch 252.81 | loss  1.29 | val loss  4.79
| epoch 104 | [  399/  809] | ms/batch 253.54 | loss  1.06 | val loss  4.48
| epoch 104 | [  599/  809] | ms/batch 251.45 | loss  1.23 | val loss  4.69
| epoch 104 | [  799/  809] | ms/batch 251.20 | loss  1.33 | val loss  4.48
| epoch 105 | [  199/  809] | ms/batch 251.92 | loss  1.20 | val loss  4.54
| epoch 105 | [  399/  809] | ms/batch 253.12 | loss  1.03 | val loss  4.44
| epoch 105 | [  599/  809] | ms/batch 253.23 | loss  1.36 | val loss  4.66
| epoch 105 | [  799/  809] | ms/batch 253.36 | loss  1.42 | val loss  4.62
| epoch 106 | [  199/  809] | ms/batch 254.20 | loss  1.21 | val loss  4.52
| epoch 106 | [  399/  809] | ms/batch 253.41 | loss  1.00 | val loss  4.69
| epoch 106 | [  599/  809] | ms/batch 253.41 | loss  1.40 | val loss  4.55
| epoch 106 | [  799/  809] | ms/batch 253.21 | loss  1.30 | val loss  4.61
| epoch 107 | [  199/  809] | ms/batch 253.14 | loss  1.14 | val loss  4.53
| epoch 107 | [  399/  809] | ms/batch 252.68 | loss  0.94 | val loss  4.47
| epoch 107 | [  599/  809] | ms/batch 253.15 | loss  1.22 | val loss  4.62
| epoch 107 | [  799/  809] | ms/batch 253.20 | loss  1.38 | val loss  4.60
| epoch 108 | [  199/  809] | ms/batch 253.96 | loss  1.20 | val loss  4.61
| epoch 108 | [  399/  809] | ms/batch 254.74 | loss  0.96 | val loss  4.60
| epoch 108 | [  599/  809] | ms/batch 254.39 | loss  1.31 | val loss  4.74
| epoch 108 | [  799/  809] | ms/batch 254.70 | loss  1.32 | val loss  4.68
| epoch 109 | [  199/  809] | ms/batch 255.47 | loss  1.06 | val loss  4.77
| epoch 109 | [  399/  809] | ms/batch 255.64 | loss  0.98 | val loss  4.74
| epoch 109 | [  599/  809] | ms/batch 254.73 | loss  1.20 | val loss  4.82
| epoch 109 | [  799/  809] | ms/batch 253.91 | loss  1.26 | val loss  4.69
| epoch 110 | [  199/  809] | ms/batch 254.37 | loss  1.11 | val loss  4.59
| epoch 110 | [  399/  809] | ms/batch 254.40 | loss  0.93 | val loss  4.56
| epoch 110 | [  599/  809] | ms/batch 254.44 | loss  1.21 | val loss  4.52
| epoch 110 | [  799/  809] | ms/batch 254.45 | loss  1.25 | val loss  4.69
| epoch 111 | [  199/  809] | ms/batch 254.01 | loss  1.05 | val loss  4.81
| epoch 111 | [  399/  809] | ms/batch 254.18 | loss  0.94 | val loss  4.66
| epoch 111 | [  599/  809] | ms/batch 254.13 | loss  1.19 | val loss  4.62
| epoch 111 | [  799/  809] | ms/batch 254.31 | loss  1.25 | val loss  4.50
| epoch 112 | [  199/  809] | ms/batch 253.69 | loss  1.17 | val loss  4.57
| epoch 112 | [  399/  809] | ms/batch 255.82 | loss  0.85 | val loss  4.64
| epoch 112 | [  599/  809] | ms/batch 255.28 | loss  1.19 | val loss  4.75
| epoch 112 | [  799/  809] | ms/batch 254.37 | loss  1.16 | val loss  4.50
| epoch 113 | [  199/  809] | ms/batch 254.04 | loss  0.96 | val loss  4.69
| epoch 113 | [  399/  809] | ms/batch 253.47 | loss  0.89 | val loss  4.81
| epoch 113 | [  599/  809] | ms/batch 253.47 | loss  1.20 | val loss  4.57
```

```
| epoch 113 | [  799/  809] | ms/batch 253.57 | loss  1.13 | val loss  4.79
| epoch 114 | [  199/  809] | ms/batch 254.16 | loss  1.12 | val loss  4.66
| epoch 114 | [  399/  809] | ms/batch 254.66 | loss  0.98 | val loss  4.60
| epoch 114 | [  599/  809] | ms/batch 254.98 | loss  1.11 | val loss  4.73
| epoch 114 | [  799/  809] | ms/batch 254.46 | loss  1.23 | val loss  4.77
| epoch 115 | [  199/  809] | ms/batch 254.21 | loss  1.03 | val loss  4.71
| epoch 115 | [  399/  809] | ms/batch 254.50 | loss  0.94 | val loss  4.83
| epoch 115 | [  599/  809] | ms/batch 254.78 | loss  1.09 | val loss  4.62
| epoch 115 | [  799/  809] | ms/batch 254.55 | loss  1.29 | val loss  4.99
| epoch 116 | [  199/  809] | ms/batch 253.96 | loss  1.09 | val loss  4.66
| epoch 116 | [  399/  809] | ms/batch 254.55 | loss  0.90 | val loss  4.94
| epoch 116 | [  599/  809] | ms/batch 254.53 | loss  1.13 | val loss  4.82
| epoch 116 | [  799/  809] | ms/batch 255.17 | loss  1.15 | val loss  4.72
| epoch 117 | [  199/  809] | ms/batch 255.06 | loss  1.02 | val loss  4.99
| epoch 117 | [  399/  809] | ms/batch 254.73 | loss  0.81 | val loss  4.90
| epoch 117 | [  599/  809] | ms/batch 254.55 | loss  1.12 | val loss  4.61
| epoch 117 | [  799/  809] | ms/batch 254.72 | loss  1.24 | val loss  4.71
| epoch 118 | [  199/  809] | ms/batch 254.26 | loss  1.12 | val loss  4.96
| epoch 118 | [  399/  809] | ms/batch 254.32 | loss  0.79 | val loss  4.72
| epoch 118 | [  599/  809] | ms/batch 254.68 | loss  1.11 | val loss  4.90
| epoch 118 | [  799/  809] | ms/batch 253.77 | loss  1.15 | val loss  4.93
| epoch 119 | [  199/  809] | ms/batch 252.35 | loss  0.97 | val loss  4.96
| epoch 119 | [  399/  809] | ms/batch 251.46 | loss  0.84 | val loss  5.00
| epoch 119 | [  599/  809] | ms/batch 252.37 | loss  1.05 | val loss  4.91
| epoch 119 | [  799/  809] | ms/batch 252.33 | loss  1.23 | val loss  4.86
| epoch 120 | [  199/  809] | ms/batch 251.89 | loss  1.11 | val loss  4.85
| epoch 120 | [  399/  809] | ms/batch 251.59 | loss  0.83 | val loss  4.93
| epoch 120 | [  599/  809] | ms/batch 251.96 | loss  1.01 | val loss  5.01
| epoch 120 | [  799/  809] | ms/batch 251.72 | loss  1.14 | val loss  4.90
| epoch 121 | [  199/  809] | ms/batch 251.70 | loss  0.96 | val loss  4.91
| epoch 121 | [  399/  809] | ms/batch 253.58 | loss  0.77 | val loss  4.83
| epoch 121 | [  599/  809] | ms/batch 252.04 | loss  1.02 | val loss  4.95
| epoch 121 | [  799/  809] | ms/batch 251.62 | loss  1.15 | val loss  4.83
| epoch 122 | [  199/  809] | ms/batch 251.83 | loss  1.03 | val loss  4.79
| epoch 122 | [  399/  809] | ms/batch 251.84 | loss  0.79 | val loss  5.05
| epoch 122 | [  599/  809] | ms/batch 251.83 | loss  1.08 | val loss  4.97
| epoch 122 | [  799/  809] | ms/batch 251.10 | loss  1.02 | val loss  4.81
| epoch 123 | [  199/  809] | ms/batch 253.25 | loss  0.97 | val loss  4.90
| epoch 123 | [  399/  809] | ms/batch 253.74 | loss  0.78 | val loss  5.06
| epoch 123 | [  599/  809] | ms/batch 253.01 | loss  0.94 | val loss  4.85
| epoch 123 | [  799/  809] | ms/batch 251.54 | loss  1.05 | val loss  4.96
| epoch 124 | [  199/  809] | ms/batch 251.79 | loss  0.89 | val loss  4.90
| epoch 124 | [  399/  809] | ms/batch 251.67 | loss  0.83 | val loss  5.06
| epoch 124 | [  599/  809] | ms/batch 252.75 | loss  0.99 | val loss  5.10
| epoch 124 | [  799/  809] | ms/batch 252.40 | loss  1.11 | val loss  4.86
| epoch 125 | [  199/  809] | ms/batch 253.63 | loss  0.94 | val loss  4.95
| epoch 125 | [  399/  809] | ms/batch 252.45 | loss  0.76 | val loss  4.88
| epoch 125 | [  599/  809] | ms/batch 252.12 | loss  1.03 | val loss  4.84
```

```
| epoch 125 | [  799/  809] | ms/batch 253.25 | loss  1.00 | val loss  4.86
| epoch 126 | [  199/  809] | ms/batch 252.14 | loss  0.96 | val loss  4.96
| epoch 126 | [  399/  809] | ms/batch 251.44 | loss  0.78 | val loss  4.98
| epoch 126 | [  599/  809] | ms/batch 251.48 | loss  1.05 | val loss  4.68
| epoch 126 | [  799/  809] | ms/batch 251.87 | loss  1.09 | val loss  4.63
| epoch 127 | [  199/  809] | ms/batch 251.56 | loss  0.99 | val loss  5.05
| epoch 127 | [  399/  809] | ms/batch 250.74 | loss  0.77 | val loss  5.12
| epoch 127 | [  599/  809] | ms/batch 248.91 | loss  0.97 | val loss  4.99
| epoch 127 | [  799/  809] | ms/batch 249.92 | loss  0.99 | val loss  4.98
| epoch 128 | [  199/  809] | ms/batch 248.81 | loss  0.92 | val loss  4.84
| epoch 128 | [  399/  809] | ms/batch 251.36 | loss  0.71 | val loss  4.80
| epoch 128 | [  599/  809] | ms/batch 249.71 | loss  0.94 | val loss  5.08
| epoch 128 | [  799/  809] | ms/batch 250.72 | loss  1.07 | val loss  5.26
| epoch 129 | [  199/  809] | ms/batch 251.70 | loss  0.85 | val loss  4.82
| epoch 129 | [  399/  809] | ms/batch 249.34 | loss  0.84 | val loss  5.04
| epoch 129 | [  599/  809] | ms/batch 248.55 | loss  0.95 | val loss  5.12
| epoch 129 | [  799/  809] | ms/batch 249.89 | loss  1.04 | val loss  4.94
| epoch 130 | [  199/  809] | ms/batch 250.70 | loss  0.83 | val loss  5.03
| epoch 130 | [  399/  809] | ms/batch 250.10 | loss  0.75 | val loss  5.04
| epoch 130 | [  599/  809] | ms/batch 249.77 | loss  0.99 | val loss  5.31
| epoch 130 | [  799/  809] | ms/batch 251.69 | loss  0.99 | val loss  5.26
| epoch 131 | [  199/  809] | ms/batch 250.53 | loss  0.80 | val loss  5.08
| epoch 131 | [  399/  809] | ms/batch 251.14 | loss  0.70 | val loss  5.27
| epoch 131 | [  599/  809] | ms/batch 250.77 | loss  0.88 | val loss  4.89
| epoch 131 | [  799/  809] | ms/batch 249.82 | loss  0.92 | val loss  5.07
| epoch 132 | [  199/  809] | ms/batch 249.98 | loss  0.88 | val loss  4.98
| epoch 132 | [  399/  809] | ms/batch 250.55 | loss  0.62 | val loss  5.25
| epoch 132 | [  599/  809] | ms/batch 249.97 | loss  0.92 | val loss  5.06
| epoch 132 | [  799/  809] | ms/batch 250.20 | loss  0.92 | val loss  4.97
| epoch 133 | [  199/  809] | ms/batch 250.59 | loss  0.74 | val loss  5.01
| epoch 133 | [  399/  809] | ms/batch 249.84 | loss  0.66 | val loss  5.21
| epoch 133 | [  599/  809] | ms/batch 249.92 | loss  0.94 | val loss  5.16
| epoch 133 | [  799/  809] | ms/batch 250.18 | loss  0.94 | val loss  5.20
| epoch 134 | [  199/  809] | ms/batch 249.95 | loss  0.76 | val loss  5.05
| epoch 134 | [  399/  809] | ms/batch 249.90 | loss  0.75 | val loss  5.31
| epoch 134 | [  599/  809] | ms/batch 250.97 | loss  0.95 | val loss  5.29
| epoch 134 | [  799/  809] | ms/batch 250.33 | loss  1.01 | val loss  4.89
| epoch 135 | [  199/  809] | ms/batch 251.11 | loss  1.02 | val loss  5.27
| epoch 135 | [  399/  809] | ms/batch 250.68 | loss  0.66 | val loss  5.20
| epoch 135 | [  599/  809] | ms/batch 249.81 | loss  0.99 | val loss  5.40
| epoch 135 | [  799/  809] | ms/batch 250.57 | loss  0.89 | val loss  5.33
| epoch 136 | [  199/  809] | ms/batch 249.57 | loss  0.85 | val loss  5.39
| epoch 136 | [  399/  809] | ms/batch 250.12 | loss  0.64 | val loss  5.24
| epoch 136 | [  599/  809] | ms/batch 249.52 | loss  0.91 | val loss  5.50
| epoch 136 | [  799/  809] | ms/batch 250.59 | loss  0.88 | val loss  5.27
| epoch 137 | [  199/  809] | ms/batch 249.46 | loss  0.85 | val loss  5.08
| epoch 137 | [  399/  809] | ms/batch 250.02 | loss  0.73 | val loss  5.32
| epoch 137 | [  599/  809] | ms/batch 250.31 | loss  0.94 | val loss  5.19
```

```
| epoch 137 | [  799/  809] | ms/batch 249.81 | loss  0.97 | val loss  5.35
| epoch 138 | [  199/  809] | ms/batch 250.03 | loss  0.87 | val loss  5.16
| epoch 138 | [  399/  809] | ms/batch 250.17 | loss  0.68 | val loss  5.22
| epoch 138 | [  599/  809] | ms/batch 249.58 | loss  0.80 | val loss  5.23
| epoch 138 | [  799/  809] | ms/batch 250.02 | loss  0.82 | val loss  5.32
| epoch 139 | [  199/  809] | ms/batch 251.16 | loss  0.80 | val loss  5.35
| epoch 139 | [  399/  809] | ms/batch 249.76 | loss  0.60 | val loss  5.01
| epoch 139 | [  599/  809] | ms/batch 249.84 | loss  0.76 | val loss  5.34
| epoch 139 | [  799/  809] | ms/batch 250.25 | loss  0.88 | val loss  5.22
| epoch 140 | [  199/  809] | ms/batch 252.25 | loss  0.71 | val loss  5.32
| epoch 140 | [  399/  809] | ms/batch 251.51 | loss  0.56 | val loss  5.26
| epoch 140 | [  599/  809] | ms/batch 250.28 | loss  0.82 | val loss  5.58
| epoch 140 | [  799/  809] | ms/batch 249.29 | loss  0.93 | val loss  5.09
| epoch 141 | [  199/  809] | ms/batch 247.35 | loss  0.83 | val loss  5.43
| epoch 141 | [  399/  809] | ms/batch 247.52 | loss  0.68 | val loss  5.41
| epoch 141 | [  599/  809] | ms/batch 248.28 | loss  0.74 | val loss  5.46
| epoch 141 | [  799/  809] | ms/batch 248.08 | loss  0.88 | val loss  5.18
| epoch 142 | [  199/  809] | ms/batch 249.82 | loss  0.69 | val loss  5.22
| epoch 142 | [  399/  809] | ms/batch 250.52 | loss  0.67 | val loss  5.24
| epoch 142 | [  599/  809] | ms/batch 249.80 | loss  0.69 | val loss  5.21
| epoch 142 | [  799/  809] | ms/batch 252.28 | loss  0.77 | val loss  5.42
| epoch 143 | [  199/  809] | ms/batch 252.01 | loss  0.71 | val loss  5.50
| epoch 143 | [  399/  809] | ms/batch 249.89 | loss  0.65 | val loss  5.38
| epoch 143 | [  599/  809] | ms/batch 251.49 | loss  0.78 | val loss  5.43
| epoch 143 | [  799/  809] | ms/batch 252.72 | loss  0.86 | val loss  5.32
| epoch 144 | [  199/  809] | ms/batch 252.56 | loss  0.77 | val loss  5.26
| epoch 144 | [  399/  809] | ms/batch 252.21 | loss  0.61 | val loss  5.54
| epoch 144 | [  599/  809] | ms/batch 252.36 | loss  0.80 | val loss  5.62
| epoch 144 | [  799/  809] | ms/batch 251.27 | loss  0.80 | val loss  5.32
| epoch 145 | [  199/  809] | ms/batch 250.99 | loss  0.76 | val loss  5.29
| epoch 145 | [  399/  809] | ms/batch 250.10 | loss  0.63 | val loss  5.46
| epoch 145 | [  599/  809] | ms/batch 250.40 | loss  0.77 | val loss  5.45
| epoch 145 | [  799/  809] | ms/batch 248.92 | loss  0.98 | val loss  5.22
| epoch 146 | [  199/  809] | ms/batch 249.29 | loss  0.73 | val loss  5.31
| epoch 146 | [  399/  809] | ms/batch 249.52 | loss  0.57 | val loss  5.49
| epoch 146 | [  599/  809] | ms/batch 248.87 | loss  0.76 | val loss  5.29
| epoch 146 | [  799/  809] | ms/batch 249.39 | loss  0.78 | val loss  5.40
| epoch 147 | [  199/  809] | ms/batch 249.74 | loss  0.71 | val loss  5.46
| epoch 147 | [  399/  809] | ms/batch 249.19 | loss  0.62 | val loss  5.51
| epoch 147 | [  599/  809] | ms/batch 249.66 | loss  0.84 | val loss  5.53
| epoch 147 | [  799/  809] | ms/batch 248.74 | loss  0.77 | val loss  5.41
| epoch 148 | [  199/  809] | ms/batch 251.35 | loss  0.73 | val loss  5.56
| epoch 148 | [  399/  809] | ms/batch 249.05 | loss  0.66 | val loss  5.45
| epoch 148 | [  599/  809] | ms/batch 249.06 | loss  0.69 | val loss  5.33
| epoch 148 | [  799/  809] | ms/batch 248.31 | loss  0.87 | val loss  5.32
| epoch 149 | [  199/  809] | ms/batch 249.31 | loss  0.71 | val loss  5.50
| epoch 149 | [  399/  809] | ms/batch 248.14 | loss  0.51 | val loss  5.57
| epoch 149 | [  599/  809] | ms/batch 250.74 | loss  0.73 | val loss  5.53
```

```
| epoch 149 | [  799/  809] | ms/batch 248.44 | loss  0.84 | val loss  5.47
Original: snug fit feels great
Predicted: will be great for spring rides or


Original: cute
Predicted: tate hand earrings exactly what i


Original: the product description is inaccurate
Predicted: senzor senzor senzor vi viper ok but


Original: glasses
Predicted: linguri linguri linguris are great


Original: true to size
Predicted: always a great shoe fit is


bleu, precisions, bp, ratio, translation_length, reference_length (0.0,
[0.22144055961974557, 0.0, 0.0, 0.0], 1.0, 35.20658547586829, 780530, 22170)
bleau score 0.6859844838250574
rouge2 (0.16848225200975267, 0.15806677562456223, 0.18036717690370083)
rouge {'rouge_1/f_score': 0.04740721897690091, 'rouge_1/r_score':
0.06681084477092596, 'rouge_1/p_score': 0.04179674019660488, 'rouge_2/f_score':
0.005587580152682946, 'rouge_2/r_score': 0.009073214912186496,
'rouge_2/p_score': 0.004661490216294005, 'rouge_l/f_score': 0.036532753369628,
'rouge_l/r_score': 0.06430128659492665, 'rouge_l/p_score': 0.03582168095023305}
```

[ ]:
```python
#function to view the output scores of T5 model this can be done for manual␣
 ↪checking after txt file is generated
def view_t5_op():
  #get the final cleaned data
  df=pd.read_csv('/content/drive/MyDrive/product_reviews.csv')[:147799]
  print("The length of dataset is ",len(df))

  #set the threshold
  threshold = 20
  max_rl=80 #maximum review length
  max_sl=10 #maximum summary length

  #get reviewText whose length is less than maximum review length
  df['reviewText']=df['reviewText'].str.slice(0,max_rl)

  #get summary whose length is less than maximum summary length
```

```
  df['summary']=df['summary'].str.slice(0,max_rl)

  f = open("/content/drive/MyDrive/TFIVE.txt", "r")
  text=f.readlines()
  text=pd.DataFrame(text,columns=["value"])
  text=text["value"].str.split("\t",expand=True)
  text.columns=["predicted","value","original"]
  text.drop(columns=["value"],inplace=True)
  text["predicted"]=text["predicted"].str.split(":").str[1]
  text["original"]=text["original"].str.split(":").str[1]
  text["original"]=text["original"].replace('\n','', regex=True)
  f.close()

  bleau=compute_bleu(text["original"],text["predicted"],␣
↪max_order=4,smooth=False)
  bscore=nltk.translate.bleu_score.
↪corpus_bleu(text["original"],text["predicted"])
  rougen=rouge_n(text["predicted"], text["original"], n=2)
  ro=rouge(text["predicted"],text["original"])

  print("bleu, precisions, bp, ratio, translation_length,␣
↪reference_length",bleau)
  print("bleau score",bscore)
  print("rouge2",rougen)
  print("rouge",ro)
  return df,text
```

[ ]:
```
df,text=view_t5_op()
```

```
The length of dataset is  147799
bleu, precisions, bp, ratio, translation_length, reference_length (0.0,
[0.2185841237246996, 0.0, 0.0, 0.0], 1.0, 36.200090211998194, 802556, 22170)
bleau score 0.6837615158224546
rouge2 (0.22406421765542467, 0.19950063644374816, 0.2555259445054084)
rouge {'rouge_1/f_score': 0.23601519307003388, 'rouge_1/r_score':
0.313584763479892, 'rouge_1/p_score': 0.20924413443019668, 'rouge_2/f_score':
0.0052723712383178, 'rouge_2/r_score': 0.0072777097607950115, 'rouge_2/p_score':
0.00455244825339548, 'rouge_l/f_score': 0.1833967729912532, 'rouge_l/r_score':
0.30963190608319163, 'rouge_l/p_score': 0.1774901912351168}
```

[ ]:
```
#function to view LSTM output scores of LSTM model this can be done for manual␣
 ↪checking after txt file is generated
def view_lstm():
  f = open("/content/drive/MyDrive/LSTMscore.txt", "r")
  text=f.readlines()
  text=pd.DataFrame(text,columns=["value"])
  text=text["value"].str.split("\t",expand=True)
```

```
  text.columns=["value","original","predicted"]
  text["original"]=text["original"].str.split(":").str[1]
  text["predicted"]=text["predicted"].str.split(":").str[1]
  text["predicted"]=text["predicted"].replace('\n','', regex=True)
  f.close()
  bleau=compute_bleu(text["original"],text["predicted"],␣
↪max_order=4,smooth=False)
  bscore=nltk.translate.bleu_score.
↪corpus_bleu(text["original"],text["predicted"])
  rougen=rouge_n(text["predicted"], text["original"], n=2)
  ro=rouge(text["predicted"],text["original"])

  print("bleu, precisions, bp, ratio, translation_length,␣
↪reference_length",bleau)
  print("bleau score",bscore)
  print("rouge2",rougen)
  print("rouge",ro)
  return text
```

[ ]:
```
text=view_lstm()
```

```
bleu, precisions, bp, ratio, translation_length, reference_length (0.0,
[0.2900551776136539, 0.0, 0.0, 0.0], 1.0, 19.09509591394331, 466856, 24449)
bleau score 0.7338717254431542
rouge2 (0.06599443828484215, 0.8312236286919831, 0.03436126421544687)
rouge {'rouge_1/f_score': 0.36915419958439477, 'rouge_1/r_score':
0.36583216309687744, 'rouge_1/p_score': 0.41662537566051006, 'rouge_2/f_score':
0.2881554948360078, 'rouge_2/r_score': 0.33457293997806903, 'rouge_2/p_score':
0.2751092653511975, 'rouge_l/f_score': 0.6180779054304887, 'rouge_l/r_score':
0.6641052933641588, 'rouge_l/p_score': 0.5875250521493721}
```

[ ]: