Sentiment_Analysis-binary-classification-BRNN-CuDNNGRU-Batchnormalization-AttentionLayer

January 24, 2020

1 Sentiment Analysis with an RNN

Run in Google Colab

View source on GitHub

http://www.polyvista.com/blog/wp-content/uploads/2015/06/sentiment-customer-exp-large.png

1.1 What is Sentiment Analysis?

Sentiment Analysis also know as opinion mining refers to the identification, extraction and study of sentiment states by using natural language processing, text analysis, computational linguistics and biometrics.

1.2 Sentiment Analysis with an Recurrent Neural Network

We will use a RNN for sentiment analysis because we care for the sequence in the data.

1.2.1 Imports

[1]:
```
import re
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt

from tensorflow.keras.models import Sequential, load_model
from tensorflow.compat.v1.keras.layers import CuDNNGRU, Embedding,␣
 ↪Dropout,Dense, Bidirectional, BatchNormalization
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.optimizers import RMSprop, Adamax , Adam

from attention.layers import AttentionLayer
```

```
# import keras
# from keras.models import Sequential, load_model
# from keras.layers import Dense, Embedding, Dropout
# from keras.preprocessing.text import Tokenizer
# from keras.preprocessing.sequence import pad_sequences
import tensorflow as tf
from tensorflow.python.client import device_lib
```

[2]:
```
from tensorflow.compat.v1 import ConfigProto
from tensorflow.compat.v1 import InteractiveSession

config = ConfigProto()
config.gpu_options.per_process_gpu_memory_fraction = 0.6
config.gpu_options.allow_growth = True
session = InteractiveSession(config=config)
```

[3]:
```
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all" #This is for multiple print␣
↪statements per cell
```

[4]:
```
value = tf.test.is_gpu_available(
    cuda_only=False,
    min_cuda_compute_capability=None
)
print ('***If TF can access GPU: ***\n\n',value) # MUST RETURN True IF IT CAN!!
```

```
WARNING:tensorflow:From <ipython-input-4-cb50da41978a>:3: is_gpu_available (from
tensorflow.python.framework.test_util) is deprecated and will be removed in a
future version.
Instructions for updating:
Use `tf.config.list_physical_devices('GPU')` instead.
***If TF can access GPU: ***

 True
```

[5]:
```
value = tf.config.list_physical_devices('GPU')
print(value)
```

```
[PhysicalDevice(name='/physical_device:GPU:0', device_type='GPU')]
```

[6]:
```
print(device_lib.list_local_devices())
```

```
[name: "/device:CPU:0"
device_type: "CPU"
memory_limit: 268435456
locality {
}
```

```
incarnation: 11803569409837407690
, name: "/device:XLA_CPU:0"
device_type: "XLA_CPU"
memory_limit: 17179869184
locality {
}
incarnation: 13130751114115519302
physical_device_desc: "device: XLA_CPU device"
, name: "/device:XLA_GPU:0"
device_type: "XLA_GPU"
memory_limit: 17179869184
locality {
}
incarnation: 14798385028866007146
physical_device_desc: "device: XLA_GPU device"
, name: "/device:GPU:0"
device_type: "GPU"
memory_limit: 1259942707
locality {
  bus_id: 1
  links {
  }
}
incarnation: 7457501116628703537
physical_device_desc: "device: 0, name: GeForce MX150, pci bus id: 0000:02:00.0,
compute capability: 6.1"
]
```

[7]:
```
tf.debugging.set_log_device_placement(True)
```

[8]:
```
 tf
print("Num GPUs Available: ", len(tf.config.experimental.
 ↪list_physical_devices('GPU')))
```

[8]:
```
<module 'tensorflow' from '/home/erolerten/anaconda3/envs/venv-
tensorflow/lib/python3.7/site-packages/tensorflow/__init__.py'>

Num GPUs Available:  1
```

2 Place tensors on the CPU

3 with tf.device(‘/GPU:0’):

a = tf.constant([[1.0, 2.0, 3.0], [4.0, 5.0, 6.0]]) b = tf.constant([[1.0, 2.0], [3.0, 4.0], [5.0, 6.0]])

c = tf.matmul(a, b) print(c)

3.0.1 Loading in Dataset

[9]:
```
data1 = pd.read_csv('Tweets.csv')
data2 = pd.read_csv('stanford-tweets.csv',sep=',')
# data1 = data1.sample(frac=1).reset_index(drop=True)
# data2 = data2.sample(frac=1).reset_index(drop=True)
print(data1.shape)
print(data2.shape)

data1.head()
data2.head()
```

```
(14640, 15)
(1600000, 2)
```

[9]:

	tweet_id	airline_sentiment	airline_sentiment_confidence	negativereason	negativereason_confidence	airline	airline_sentiment_gold	name	negativereason_gold	retweet_count	text	tweet_coord	tweet_created	tweet_location	user_timezone
0	570306133677760513	neutral	1.0000	NaN	NaN	Virgin America	NaN	cairdin	NaN	0	@VirginAmerica What @dhepburn said.	NaN	2015-02-24 11:35:52 -0800	NaN	Eastern Time (US & Canada)
1	570301130888122368	positive	0.3486	NaN	0.0000	Virgin America	NaN	jnardino	NaN	0	@VirginAmerica plus you've added commercials t…	NaN	2015-02-24 11:15:59 -0800	NaN	Pacific Time (US & Canada)
2	570301083672813571	neutral	0.6837	NaN	NaN	Virgin America	NaN	yvonnalynn	NaN	0	@VirginAmerica I didn't today… Must mean I n…	NaN	2015-02-24 11:15:48 -0800	Lets Play	Central Time (US & Canada)
3	570301031407624196	negative	1.0000	Bad Flight	0.7033	Virgin America	NaN	jnardino	NaN	0	@VirginAmerica it's really aggressive to blast…	NaN			
4	570300817074462722	negative	1.0000	Can't Tell	1.0000	Virgin America	NaN	jnardino	NaN	0	@VirginAmerica and it's a really big bad thing…	NaN			

```
3  2015-02-24 11:15:36 -0800            NaN  Pacific Time (US & Canada)
4  2015-02-24 11:14:45 -0800            NaN  Pacific Time (US & Canada)
```

[9]:

	sentiment	text
0	negative	@switchfoot http://twitpic.com/2y1zl - Awww, t…
1	negative	is upset that he can't update his Facebook by …
2	negative	@Kenichan I dived many times for the ball. Man…
3	negative	my whole body feels itchy and like its on fire
4	negative	@nationwideclass no, it's not behaving at all…

Removing all columns except the airline_sentiment and text column.

[10]:
```
data1 = data1[['airline_sentiment', 'text']]
new_columns = ['sentiment','text']
data1.columns = new_columns
data1.head()
```

[10]:

	sentiment	text
0	neutral	@VirginAmerica What @dhepburn said.
1	positive	@VirginAmerica plus you've added commercials t…
2	neutral	@VirginAmerica I didn't today… Must mean I n…
3	negative	@VirginAmerica it's really aggressive to blast…
4	negative	@VirginAmerica and it's a really big bad thing…

[11]:
```
df = data1.append(data2, ignore_index = True)
print(df.shape)
df
```

```
(1614640, 2)
```

[11]:

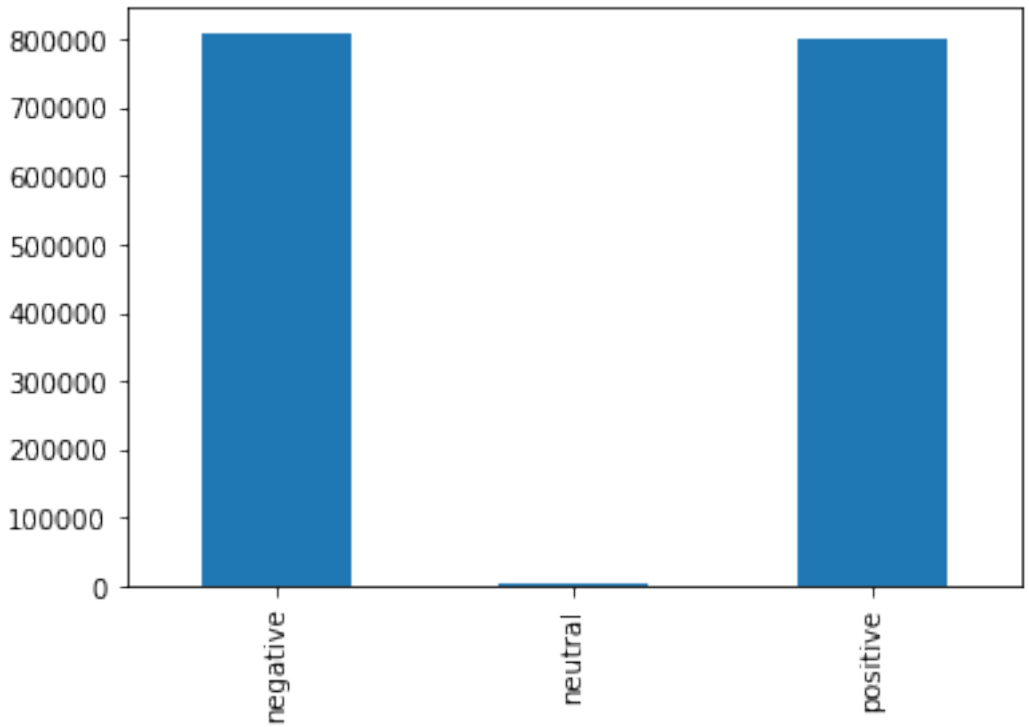
	sentiment	text
0	neutral	@VirginAmerica What @dhepburn said.
1	positive	@VirginAmerica plus you've added commercials t…
2	neutral	@VirginAmerica I didn't today… Must mean I n…
3	negative	@VirginAmerica it's really aggressive to blast…
4	negative	@VirginAmerica and it's a really big bad thing…
…	…	…
1614635	positive	Just woke up. Having no school is the best fee…
1614636	positive	TheWDB.com - Very cool to hear old Walt interv…
1614637	positive	Are you ready for your MoJo Makeover? Ask me f…
1614638	positive	Happy 38th Birthday to my boo of alll time!!! …
1614639	positive	happy #charitytuesday @theNSPCC @SparksCharity…

[1614640 rows x 2 columns]

3.0.2 Data exploration

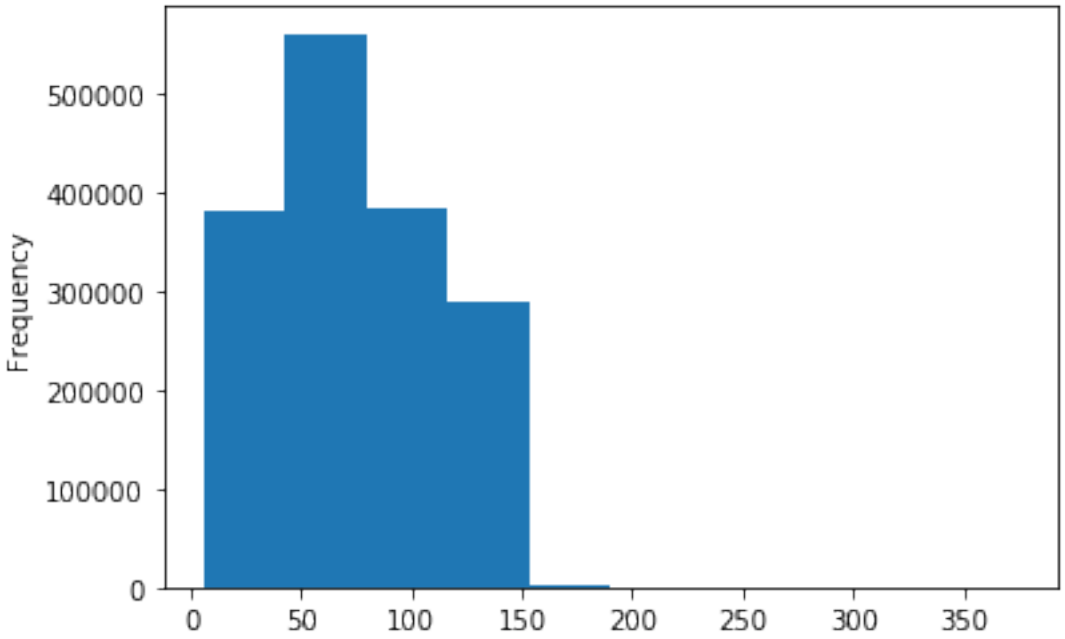
[12]:
```
df['sentiment'].value_counts().sort_index().plot.bar()
```

[12]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f20504d0710>
```



[13]:
```
df['text'].str.len().plot.hist()
```

[13]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2050510e10>
```



3.0.3 Preprocessing

[14]:
```
# How much of Dataset to be used
frac = 0.1
```

[15]:
```
# data['text'] = data['text'].str.replace('@VirginAmerica', '')
# data.head()
df = df.sample(frac=frac).reset_index(drop=True)
df
```

[15]:

	sentiment	text
0	positive	wait for some surprise
1	positive	@sydneyfamous Don't forget to tell your friend…
2	positive	watching videos
3	positive	Well . . . . . I'm finally sleepy so take …
4	positive	hmmm, should i stickam tonight?
…	…	…
161459	positive	@fishk8 Hahaha that's not necessarily true. P…
161460	positive	@beelarge ohhhhh um I hope mine comes soooooooo…
161461	negative	Great school tomorrow & cheer after
161462	positive	@meaningoftruth @chrys73 ahhahah u 2 are exagg…
161463	negative	@MaggieHizakata if i still had mine you could …

[161464 rows x 2 columns]

[16]:
```
df['text'].apply(lambda x: x.lower()) #transform text to lowercase
df['text'] = df['text'].apply(lambda x: re.sub('[^a-zA-z0-9\s]', '', x))
df['text'].head()
```

[16]:
```
0                                  wait for some surprise
1         @sydneyfamous don't forget to tell your friend…
2                                        watching videos
3         well . .  . . .  i'm finally sleepy   so take …
4                          hmmm, should i stickam tonight?
                                …
161459    @fishk8 hahaha that's not necessarily true.  p…
161460    @beelarge ohhhhh um i hope mine comes sooooooo…
161461             great school tomorrow &amp; cheer after
161462    @meaningoftruth @chrys73 ahhahah u 2 are exagg…
161463    @maggiehizakata if i still had mine you could …
Name: text, Length: 161464, dtype: object
```

[16]:
```
0                             wait for some surprise
1    sydneyfamous Dont forget to tell your friends …
2                                  watching videos
3    Well        Im finally sleepy   so take care a…
4                       hmmm should i stickam tonight
Name: text, dtype: object
```

[17]:
```
df['sentiment']
```

[17]:
```
0         positive
1         positive
2         positive
3         positive
4         positive
            …
161459    positive
161460    positive
161461    negative
161462    positive
161463    negative
Name: sentiment, Length: 161464, dtype: object
```

[18]:
```
df = df[df['sentiment'] != 'neutral']
```

[19]:
```
df
```

[19]:
```
       sentiment                                                text
0       positive                            wait for some surprise
1       positive  sydneyfamous Dont forget to tell your friends …
2       positive                                  watching videos
```

```
3       positive  Well        Im finally sleepy   so take care a…
4       positive                      hmmm should i stickam tonight
…          …                                                    …
161459  positive  fishk8 Hahaha thats not necessarily true  Plur…
161460  positive  beelarge ohhhhh um I hope mine comes soooooooo…
161461  negative            Great school tomorrow amp cheer after
161462  positive  meaningoftruth chrys73 ahhahah u 2 are exagger…
161463  negative  MaggieHizakata if i still had mine you could t…

[161113 rows x 2 columns]
```

[20]:
```
vocabulary_size = 12000
```

[21]:
```
tokenizer = Tokenizer(num_words=vocabulary_size, split=" ")
tokenizer.fit_on_texts(df['text'].values)

X = tokenizer.texts_to_sequences(df['text'].values)
X = pad_sequences(X) # padding our text vector so they all have the same length
X[:5]
```

[21]:
```
array([[   0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,  140,   10,   64, 1506],
       [   0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,   39,  748,    2,  255,   41,  198,   61,    3, 1697,
          16,   88,   62,   31,   11,   10,  364],
       [   0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,  135,  974],
       [   0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,   66,   13,  224,  764,   16,  159,  516,
          71,  163,  419,  271,  953,   82,   57],
       [   0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
           0,    0, 1068,  139,    1, 4822,  123]], dtype=int32)
```

3.0.4 Creating model

[22]:
```
model = Sequential()
model.add(Embedding(vocabulary_size, 256, input_length=X.shape[1]))
model.add(Dropout(0.3))
model.add(Bidirectional(CuDNNGRU(256, return_sequences=True)))
model.add(Dropout(0.3))
model.add(Bidirectional(CuDNNGRU(256, return_sequences=True)))
model.add(AttentionLayer(name='attention'))
model.add(BatchNormalization())
model.add(Dense(2, activation='sigmoid'))
```

```
Executing op RandomUniform in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op Sub in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op Mul in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op Add in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op VarIsInitializedOp in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op LogicalNot in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op Assert in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op AssignVariableOp in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op RandomUniform in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op Sub in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Mul in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Add in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarIsInitializedOp in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op LogicalNot in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Assert in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op AssignVariableOp in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op RandomStandardNormal in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op Qr in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op DiagPart in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Sign in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Transpose in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Reshape in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Fill in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
```

```
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
```

[23]:
```
model.compile(loss='binary_crossentropy', optimizer='adam',␣
 ↪metrics=['accuracy'])
model.summary()
```

```
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, 40, 256)           3072000
_________________________________________________________________
dropout (Dropout)            (None, 40, 256)           0
_________________________________________________________________
bidirectional (Bidirectional (None, 40, 512)           789504
_________________________________________________________________
dropout_1 (Dropout)          (None, 40, 512)           0
_________________________________________________________________
bidirectional_1 (Bidirection (None, 40, 512)           1182720
_________________________________________________________________
attention (AttentionLayer)   (None, 512)               263168
_________________________________________________________________
batch_normalization (BatchNo (None, 512)               2048
_________________________________________________________________
dense (Dense)                (None, 2)                 1026
=================================================================
Total params: 5,310,466
Trainable params: 5,309,442
Non-trainable params: 1,024
_________________________________________________________________
```

[24]:
```
y = pd.get_dummies(df['sentiment']).values
[print(df['sentiment'][i], y[i]) for i in range(0,5)]
```

```
positive [0 1]
positive [0 1]
positive [0 1]
positive [0 1]
positive [0 1]
```

[24]:
```
[None, None, None, None, None]
```

[25]:
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=0)
```

3.0.5 Training model

[29]:
```
batch_size = 32
epochs = 6
import time
from datetime import datetime
datetime = str(datetime.now())
csv_logger = tf.keras.callbacks.CSVLogger('training'+datetime+'.log')
start = time.time()
history = model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size,␣
 ↪verbose=2, callbacks=[csv_logger])
end = time.time()
elapsed = end - start
print(elapsed/60," minutes")
```

```
Executing op RangeDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op RepeatDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op MapDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op PrefetchDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op FlatMapDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op TensorDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op RepeatDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op ZipDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op ParallelMapDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op DatasetCardinality in device
/job:localhost/replica:0/task:0/device:CPU:0
Train on 128890 samples
Epoch 1/6
Executing op ModelDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op AnonymousIteratorV2 in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op MakeIterator in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op AssignVariableOp in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
```

```
Executing op LogicalNot in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op Assert in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op VarHandleOp in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op __inference_distributed_function_4267 in device
/job:localhost/replica:0/task:0/device:GPU:0
128890/128890 - 285s - loss: 0.4843 - accuracy: 0.7674
Epoch 2/6
128890/128890 - 269s - loss: 0.4159 - accuracy: 0.8105
Epoch 3/6
128890/128890 - 268s - loss: 0.3730 - accuracy: 0.8348
Epoch 4/6
```

```
128890/128890 - 269s - loss: 0.3290 - accuracy: 0.8580
Epoch 5/6
128890/128890 - 269s - loss: 0.2839 - accuracy: 0.8818
Epoch 6/6
128890/128890 - 268s - loss: 0.2446 - accuracy: 0.9003
Executing op DeleteIterator in device
/job:localhost/replica:0/task:0/device:CPU:0
27.110332651933035  minutes
```

3.0.6 Plotting Training History

[30]:
```
# print(history)
```

[31]:
```
import matplotlib.pyplot as plt

# Plot training & validation accuracy values
plt.plot(history.history['accuracy'])
# plt.plot(history.history['val_accuracy'])
plt.title('Model accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()

# Plot training & validation loss values
plt.plot(history.history['loss'])
# plt.plot(history.history['val_loss'])
plt.title('Model loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()
```

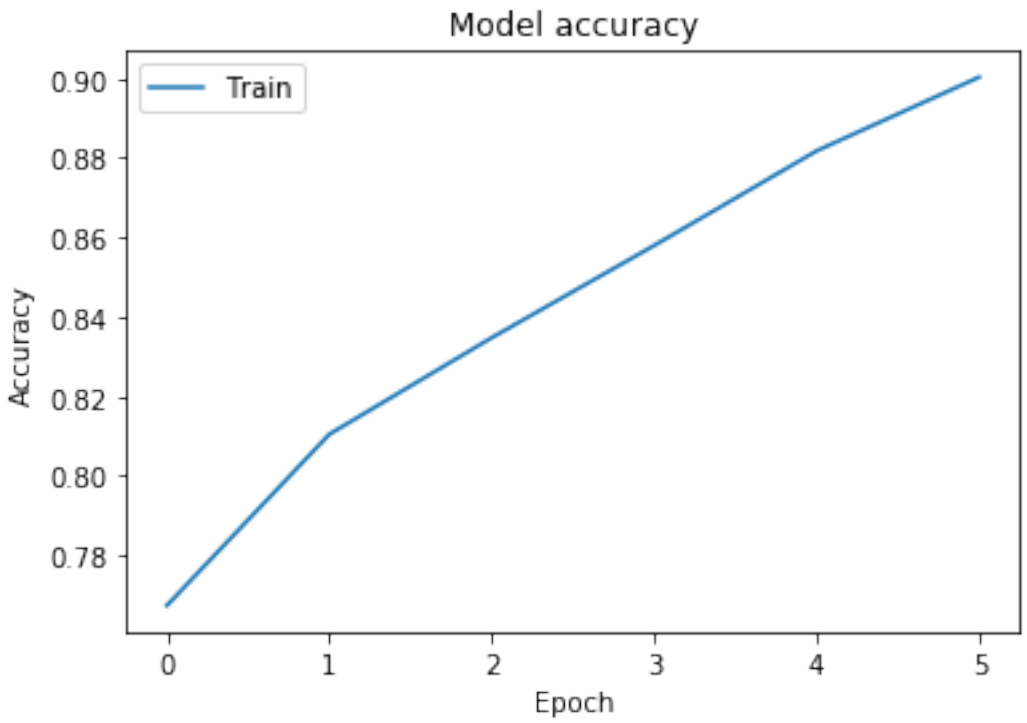
[31]:
```
[<matplotlib.lines.Line2D at 0x7f1f52950950>]
```

[31]:
```
Text(0.5, 1.0, 'Model accuracy')
```

[31]:
```
Text(0, 0.5, 'Accuracy')
```

[31]:
```
Text(0.5, 0, 'Epoch')
```

[31]:
```
<matplotlib.legend.Legend at 0x7f1f52950dd0>
```



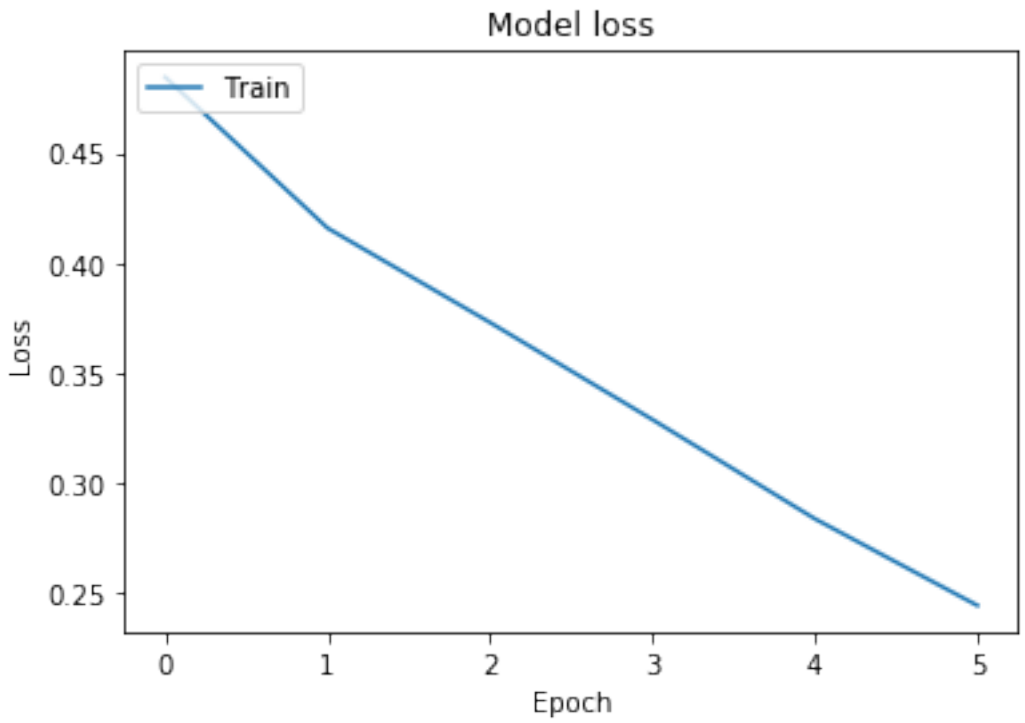
[31]: [<matplotlib.lines.Line2D at 0x7f1f482ab090>]

[31]: Text(0.5, 1.0, 'Model loss')

[31]: Text(0, 0.5, 'Loss')

[31]: Text(0.5, 0, 'Epoch')

[31]: <matplotlib.legend.Legend at 0x7f1f528e5090>



3.0.7 Testing model

[32]:
```
predictions = model.predict(X_test)

[print(df['text'][i], predictions[i], y_test[i]) for i in range(0, 5)]
```

```
Executing op RangeDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op RepeatDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op MapDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op PrefetchDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op FlatMapDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op TensorDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op RepeatDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op ZipDataset in device /job:localhost/replica:0/task:0/device:CPU:0
Executing op ParallelMapDataset in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op ModelDataset in device /job:localhost/replica:0/task:0/device:CPU:0
```

```
Executing op AnonymousIteratorV2 in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op __inference_distributed_function_77359 in device
/job:localhost/replica:0/task:0/device:GPU:0
wait for some surprise  [0.00311694 0.9967097 ] [0 1]
sydneyfamous Dont forget to tell your friends about the contest so they can get
in for free  [0.05295864 0.9462807 ] [0 1]
watching videos  [0.97296077 0.02676613] [1 0]
Well        Im finally sleepy   so take care amp please stay N touch Miss U
[0.98890734 0.01075463] [1 0]
hmmm should i stickam tonight  [0.98953295 0.01027828] [1 0]
```

[32]:
```
[None, None, None, None, None]
```

[33]:
```
accurate_prediction_count, inaccurate_prediction_count = 0, 0
for i, prediction in enumerate(predictions):
    if np.argmax(prediction)==np.argmax(y_test[i]):
        accurate_prediction_count += 1
    else:
        inaccurate_prediction_count += 1

total_predictions = accurate_prediction_count + inaccurate_prediction_count
print('Number of predictions: ', total_predictions)
print('Number of accurate predictions: ', accurate_prediction_count)
print('Number of false predictions: ', inaccurate_prediction_count)
print('Accuracy: ', accurate_prediction_count/total_predictions)
```

```
Number of predictions:  32223
Number of accurate predictions:  25128
Number of false predictions:  7095
Accuracy:  0.7798156596220092
```

[34]:
```
model.
 ↪save('Sentiment_Analysis-binary-classification-BRNN-CuDNNGRU-Batchnormalization-AttentionLa
 ↪h5')
```

```
Executing op ReadVariableOp in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op Identity in device /job:localhost/replica:0/task:0/device:GPU:0
Executing op ReadVariableOp in device
/job:localhost/replica:0/task:0/device:CPU:0
Executing op ReadVariableOp in device
/job:localhost/replica:0/task:0/device:GPU:0
Executing op Identity in device /job:localhost/replica:0/task:0/device:GPU:0
```

[ ]:
```
# pos_count, neu_count, neg_count = 0, 0, 0
# real_pos, real_neu, real_neg = 0, 0, 0
# for i, prediction in enumerate(predictions):
```

```
#     if np.argmax(prediction)==2:
#         pos_count += 1
#     elif np.argmax(prediction)==1:
#         neu_count += 1
#     else:
#         neg_count += 1

#     if np.argmax(y_test[i])==2:
#         real_pos += 1
#     elif np.argmax(y_test[i])==1:
#         real_neu += 1
#     else:
#         real_neg +=1

# print('Positive predictions:', pos_count)
# print('Neutral predictions:', neu_count)
# print('Negative predictions:', neg_count)
# print('Real positive:', real_pos)
# print('Real neutral:', real_neu)
# print('Real negative:', real_neg)
```

3.1 Improvements we could implement

Weight classes (because data is skew)

Train more epochs

Use bigger network

Try other word number

3.2 Resources

Recurrent Neural Networks Explained (my own post and video)

Sentiment Analysis (Wikipedia)

What is the best way to do sentiment analysis with Python? (Quora)

How to Do Sentiment Analysis (Siraj Raval)