AI_Phase4-Copy1

October 20, 2023

1 Build an NLP model to differentiate real news from fake news

Phase 4: Development Part 2 . In this part you will continue building your project.

Continue building the fake news detection model by applying NLP techniques and training a classification model.

```
Text Preprocessing and Feature Extraction
```

```
Model training and evaluation
```

```
Dataset Link: https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset
```

So, Phase 4 continues the development that began in Phase 3, Part 1, where we further enhance the fake news detection model. This involves applying NLP techniques and training a classification model. We'ev already done applying NLP techniques in phase 3, so we continue in training our model. If you want to read through steps involved in applying NLP techniques jump to "Preprocessing the textual data Preprocessing the textual data or else,

Jump to continuation of phase 3 (Phase 4)

Model Training

Evaluation

1.1 Importing required libraries

```python
[3]: import warnings
     warnings.filterwarnings('ignore')
```

```python
[4]: import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     import seaborn as sns

     import nltk
     import re
     import string

     from sklearn.model_selection import train_test_split
     from sklearn.metrics import classification_report
```

```python
import keras
from keras.preprocessing import text,sequence
from keras.models import Sequential
from keras.layers import Dense,Embedding,LSTM,Dropout

import os
for dirname, _, filenames in os.walk('dataset/'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

2023-10-15 16:27:23.172675: I tensorflow/core/util/port.cc:111] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2023-10-15 16:27:23.380239: E tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered
2023-10-15 16:27:23.380269: E tensorflow/compiler/xla/stream_executor/cuda/cuda_fft.cc:609] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered
2023-10-15 16:27:23.380945: E tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered
2023-10-15 16:27:23.464522: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
To enable the following instructions: AVX2 AVX512F AVX512_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
2023-10-15 16:27:24.306371: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT

dataset/Fake.csv
dataset/archive(1).zip
dataset/True.csv

1.2 Loading Data

```python
[5]: real_data = pd.read_csv('dataset/True.csv')
     fake_data = pd.read_csv('dataset/Fake.csv')
```

```python
[6]: real_data.head()
```

```
[6]:                                      title  \
    0   As U.S. budget fight looms, Republicans flip t…
    1   U.S. military to accept transgender recruits o…
    2   Senior U.S. Republican senator: 'Let Mr. Muell…
    3   FBI Russia probe helped by Australian diplomat…
    4   Trump wants Postal Service to charge 'much mor…

                                           text        subject  \
    0   WASHINGTON (Reuters) - The head of a conservat…  politicsNews
    1   WASHINGTON (Reuters) - Transgender people will…  politicsNews
    2   WASHINGTON (Reuters) - The special counsel inv…  politicsNews
    3   WASHINGTON (Reuters) - Trump campaign adviser …  politicsNews
    4   SEATTLE/WASHINGTON (Reuters) - President Donal… politicsNews

                    date
    0   December 31, 2017
    1   December 29, 2017
    2   December 31, 2017
    3   December 30, 2017
    4   December 29, 2017
```

```
[7]: fake_data.head()
```

```
[7]:                                      title  \
    0   Donald Trump Sends Out Embarrassing New Year'…
    1   Drunk Bragging Trump Staffer Started Russian …
    2   Sheriff David Clarke Becomes An Internet Joke…
    3   Trump Is So Obsessed He Even Has Obama's Name…
    4   Pope Francis Just Called Out Donald Trump Dur…

                                           text subject  \
    0   Donald Trump just couldn t wish all Americans …    News
    1   House Intelligence Committee Chairman Devin Nu…    News
    2   On Friday, it was revealed that former Milwauk…    News
    3   On Christmas day, Donald Trump announced that …    News
    4   Pope Francis used his annual Christmas Day mes…    News

                    date
    0   December 31, 2017
    1   December 31, 2017
    2   December 30, 2017
    3   December 29, 2017
    4   December 25, 2017
```

```
[8]: #add column
    real_data['target'] = 1
    fake_data['target'] = 0
```

```
[9]: real_data.tail()
```

```
[9]:                                                        title  \
     21412  'Fully committed' NATO backs new U.S. approach…
     21413  LexisNexis withdrew two products from Chinese …
     21414  Minsk cultural hub becomes haven from authorities
     21415  Vatican upbeat on possibility of Pope Francis …
     21416  Indonesia to buy $1.14 billion worth of Russia…


                                                         text     subject  \
     21412  BRUSSELS (Reuters) - NATO allies on Tuesday we…  worldnews
     21413  LONDON (Reuters) - LexisNexis, a provider of l…  worldnews
     21414  MINSK (Reuters) - In the shadow of disused Sov…  worldnews
     21415  MOSCOW (Reuters) - Vatican Secretary of State …  worldnews
     21416  JAKARTA (Reuters) - Indonesia will buy 11 Sukh…  worldnews


                       date  target
     21412  August 22, 2017       1
     21413  August 22, 2017       1
     21414  August 22, 2017       1
     21415  August 22, 2017       1
     21416  August 22, 2017       1
```

```
[10]: #Merging the 2 datasets
      data = pd.concat([real_data, fake_data], ignore_index=True, sort=False)
      data.head()
```

```
[10]:                                                 title  \
      0  As U.S. budget fight looms, Republicans flip t…
      1  U.S. military to accept transgender recruits o…
      2  Senior U.S. Republican senator: 'Let Mr. Muell…
      3  FBI Russia probe helped by Australian diplomat…
      4  Trump wants Postal Service to charge 'much mor…


                                                      text        subject  \
      0  WASHINGTON (Reuters) - The head of a conservat…  politicsNews
      1  WASHINGTON (Reuters) - Transgender people will…  politicsNews
      2  WASHINGTON (Reuters) - The special counsel inv…  politicsNews
      3  WASHINGTON (Reuters) - Trump campaign adviser …  politicsNews
      4  SEATTLE/WASHINGTON (Reuters) - President Donal…  politicsNews


                        date  target
      0  December 31, 2017       1
      1  December 29, 2017       1
      2  December 31, 2017       1
      3  December 30, 2017       1
      4  December 29, 2017       1
```

4

```
[11]:  data.isnull().sum()
```

```
[11]:  title      0
       text       0
       subject    0
       date       0
       target     0
       dtype: int64
```
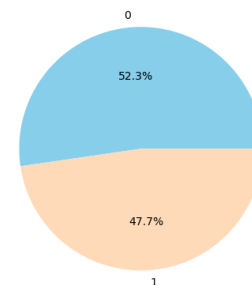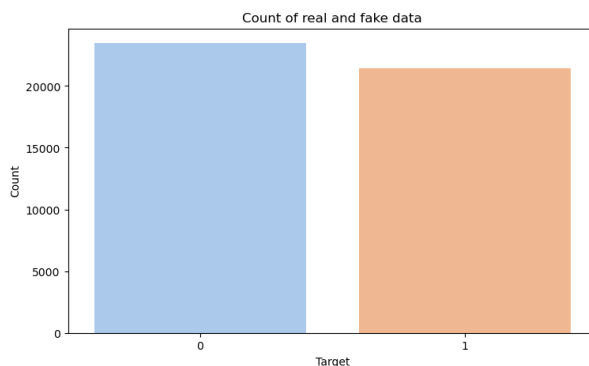
1.Count of Fake and Real Data

```
[12]:  print(data["target"].value_counts())
       fig, ax = plt.subplots(1,2, figsize=(19, 5))
       g1 = sns.countplot(data.target,ax=ax[0],palette="pastel");
       g1.set_title("Count of real and fake data")
       g1.set_ylabel("Count")
       g1.set_xlabel("Target")
       g2 = plt.pie(data["target"].value_counts().values,explode=[0,0],labels=data.
        ↪target.value_counts().index, autopct='%1.
        ↪1f%%',colors=['SkyBlue','PeachPuff'])
       fig.show()
```

```
0    23481
1    21417
Name: target, dtype: int64
```


2.Distribution of The Subject According to Real and Fake Data

```
[13]:  print(data.subject.value_counts())
       plt.figure(figsize=(10, 5))

       ax = sns.countplot(x="subject",  hue='target', data=data, palette="pastel")
       plt.title("Distribution of The Subject According to Real and Fake Data")
```
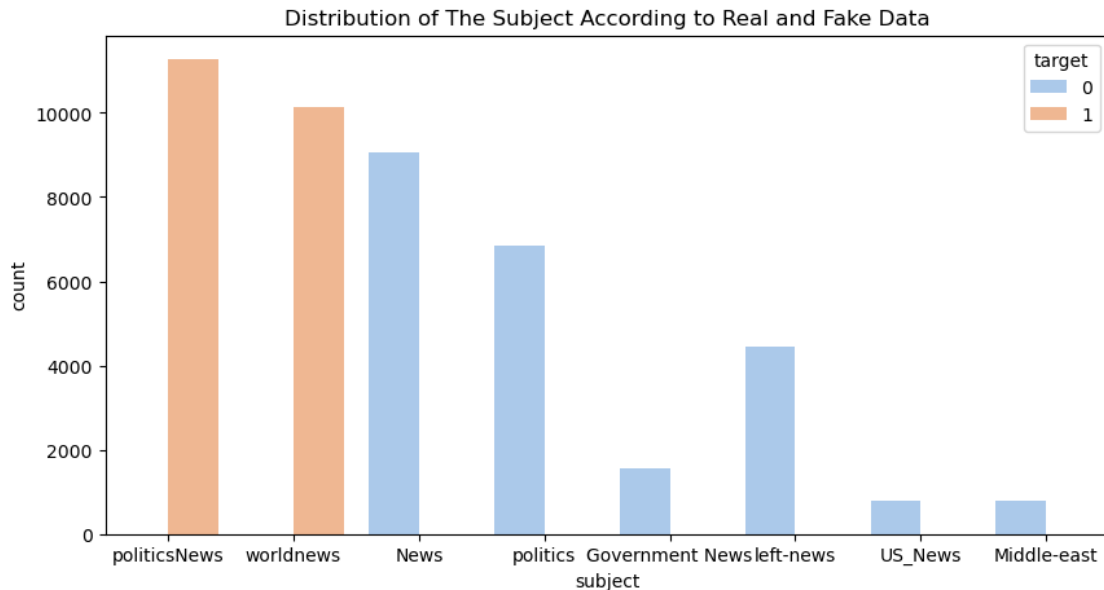
```
politicsNews      11272
worldnews         10145
```

```
News               9050
politics           6841
left-news          4459
Government News     1570
US_News             783
Middle-east         778
Name: subject, dtype: int64
```

[13]: Text(0.5, 1.0, 'Distribution of The Subject According to Real and Fake Data')

2 Preprocessing the textual data

```python
data['text']= data['subject'] + " " + data['title'] + " " + data['text']
del data['title']
del data['subject']
del data['date']
data.head()
```

[14]:
```
                                                text  target
0  politicsNews As U.S. budget fight looms, Repub…       1
1  politicsNews U.S. military to accept transgend…       1
2  politicsNews Senior U.S. Republican senator: '…       1
3  politicsNews FBI Russia probe helped by Austra…       1
4  politicsNews Trump wants Postal Service to cha…       1
```

[15]:
```python
first_text = data.text[10]
first_text
```

[15]: 'politicsNews Jones certified U.S. Senate winner despite Moore challenge
(Reuters) - Alabama officials on Thursday certified Democrat Doug Jones the
winner of the state's U.S. Senate race, after a state judge denied a challenge
by Republican Roy Moore, whose campaign was derailed by accusations of sexual
misconduct with teenage girls. Jones won the vacant seat by about 22,000 votes,
or 1.6 percentage points, election officials said. That made him the first
Democrat in a quarter of a century to win a Senate seat in Alabama. The seat
was previously held by Republican Jeff Sessions, who was tapped by U.S.
President Donald Trump as attorney general. A state canvassing board composed of
Alabama Secretary of State John Merrill, Governor Kay Ivey and Attorney General
Steve Marshall certified the election results. Seating Jones will narrow the
Republican majority in the Senate to 51 of 100 seats. In a statement, Jones
called his victory "a new chapter" and pledged to work with both parties. Moore
declined to concede defeat even after Trump urged him to do so. He stood by
claims of a fraudulent election in a statement released after the certification
and said he had no regrets, media outlets reported. An Alabama judge denied
Moore's request to block certification of the results of the Dec. 12 election in
a decision shortly before the canvassing board met. Moore's challenge alleged
there had been potential voter fraud that denied him a chance of victory. His
filing on Wednesday in the Montgomery Circuit Court sought to halt the meeting
scheduled to ratify Jones' win on Thursday. Moore could ask for a recount, in
addition to possible other court challenges, Merrill said in an interview with
Fox News Channel. He would have to complete paperwork "within a timed period"
and show he has the money for a challenge, Merrill said. "We've not been
notified yet of their intention to do that," Merrill said. Regarding the claim
of voter fraud, Merrill told CNN that more than 100 cases had been reported.
"We've adjudicated more than 60 of those. We will continue to do that," he said.
Republican lawmakers in Washington had distanced themselves from Moore and
called for him to drop out of the race after several women accused him of sexual
assault or misconduct dating back to when they were teenagers and he was in his
early 30s. Moore has denied wrongdoing and Reuters has not been able to
independently verify the allegations. '

First, let's remove HTML content.

[16]:
```
from bs4 import BeautifulSoup

soup = BeautifulSoup(first_text, "html.parser")
first_text = soup.get_text()
first_text
```

[16]: 'politicsNews Jones certified U.S. Senate winner despite Moore challenge
(Reuters) - Alabama officials on Thursday certified Democrat Doug Jones the
winner of the state's U.S. Senate race, after a state judge denied a challenge
by Republican Roy Moore, whose campaign was derailed by accusations of sexual
misconduct with teenage girls. Jones won the vacant seat by about 22,000 votes,
or 1.6 percentage points, election officials said. That made him the first
Democrat in a quarter of a century to win a Senate seat in Alabama. The seat

was previously held by Republican Jeff Sessions, who was tapped by U.S. President Donald Trump as attorney general. A state canvassing board composed of Alabama Secretary of State John Merrill, Governor Kay Ivey and Attorney General Steve Marshall certified the election results. Seating Jones will narrow the Republican majority in the Senate to 51 of 100 seats. In a statement, Jones called his victory "a new chapter" and pledged to work with both parties. Moore declined to concede defeat even after Trump urged him to do so. He stood by claims of a fraudulent election in a statement released after the certification and said he had no regrets, media outlets reported. An Alabama judge denied Moore's request to block certification of the results of the Dec. 12 election in a decision shortly before the canvassing board met. Moore's challenge alleged there had been potential voter fraud that denied him a chance of victory. His filing on Wednesday in the Montgomery Circuit Court sought to halt the meeting scheduled to ratify Jones' win on Thursday. Moore could ask for a recount, in addition to possible other court challenges, Merrill said in an interview with Fox News Channel. He would have to complete paperwork "within a timed period" and show he has the money for a challenge, Merrill said. "We've not been notified yet of their intention to do that," Merrill said. Regarding the claim of voter fraud, Merrill told CNN that more than 100 cases had been reported. "We've adjudicated more than 60 of those. We will continue to do that," he said. Republican lawmakers in Washington had distanced themselves from Moore and called for him to drop out of the race after several women accused him of sexual assault or misconduct dating back to when they were teenagers and he was in his early 30s. Moore has denied wrongdoing and Reuters has not been able to independently verify the allegations. '

Let's now remove everything except uppercase / lowercase letters using Regular Expressions.

```
[17]: first_text = re.sub('\[[^]]*\]', ' ', first_text)
      first_text = re.sub('[^a-zA-Z]',' ',first_text)  # replaces non-alphabets with␣
       ↪spaces
      first_text = first_text.lower() # Converting from uppercase to lowercase
      first_text
```

```
[17]: 'politicsnews jones certified u s  senate winner despite moore challenge
      reuters    alabama officials on thursday certified democrat doug jones the
      winner of the state s u s  senate race  after a state judge denied a challenge
      by republican roy moore  whose campaign was derailed by accusations of sexual
      misconduct with teenage girls  jones won the vacant seat by about        votes
      or      percentage points  election officials said  that made him the first
      democrat in a quarter of a century to win a senate seat in alabama    the seat
      was previously held by republican jeff sessions  who was tapped by u s
      president donald trump as attorney general  a state canvassing board composed of
      alabama secretary of state john merrill  governor kay ivey and attorney general
      steve marshall certified the election results  seating jones will narrow the
      republican majority in the senate to    of     seats  in a statement  jones
      called his victory  a new chapter  and pledged to work with both parties  moore
```

declined to concede defeat even after trump urged him to do so he stood by
claims of a fraudulent election in a statement released after the certification
and said he had no regrets media outlets reported an alabama judge denied
moore s request to block certification of the results of the dec   election in
a decision shortly before the canvassing board met moore s challenge alleged
there had been potential voter fraud that denied him a chance of victory his
filing on wednesday in the montgomery circuit court sought to halt the meeting
scheduled to ratify jones win on thursday moore could ask for a recount in
addition to possible other court challenges merrill said in an interview with
fox news channel he would have to complete paperwork within a timed period
and show he has the money for a challenge merrill said   we ve not been
notified yet of their intention to do that   merrill said regarding the claim
of voter fraud merrill told cnn that more than   cases had been reported   we
ve adjudicated more than   of those we will continue to do that   he said
republican lawmakers in washington had distanced themselves from moore and
called for him to drop out of the race after several women accused him of sexual
assault or misconduct dating back to when they were teenagers and he was in his
early   s moore has denied wrongdoing and reuters has not been able to
independently verify the allegations '

Let's remove stopwords like is,a,the... Which do not offer much insight.

```
[18]: nltk.download("stopwords")
      from nltk.corpus import stopwords

      # we can use tokenizer instead of split
      first_text = nltk.word_tokenize(first_text)
```

```
[nltk_data] Downloading package stopwords to /home/djoe/nltk_data…
[nltk_data]   Package stopwords is already up-to-date!
```

```
[19]: first_text = [ word for word in first_text if not word in set(stopwords.
      ↪words("english"))]
```

Lemmatization to bring back multiple forms of same word to their common root like
'coming', 'comes' into 'come'.

```
[20]: lemma = nltk.WordNetLemmatizer()
      first_text = [ lemma.lemmatize(word) for word in first_text]

      first_text = " ".join(first_text)
      first_text
```

```
[20]: 'politicsnews jones certified u senate winner despite moore challenge reuters
      alabama official thursday certified democrat doug jones winner state u senate
      race state judge denied challenge republican roy moore whose campaign derailed
      accusation sexual misconduct teenage girl jones vacant seat vote percentage
      point election official said made first democrat quarter century win senate seat
      alabama seat previously held republican jeff session tapped u president donald
```

trump attorney general state canvassing board composed alabama secretary state john merrill governor kay ivey attorney general steve marshall certified election result seating jones narrow republican majority senate seat statement jones called victory new chapter pledged work party moore declined concede defeat even trump urged stood claim fraudulent election statement released certification said regret medium outlet reported alabama judge denied moore request block certification result dec election decision shortly canvassing board met moore challenge alleged potential voter fraud denied chance victory filing wednesday montgomery circuit court sought halt meeting scheduled ratify jones win thursday moore could ask recount addition possible court challenge merrill said interview fox news channel would complete paperwork within timed period show money challenge merrill said notified yet intention merrill said regarding claim voter fraud merrill told cnn case reported adjudicated continue said republican lawmaker washington distanced moore called drop race several woman accused sexual assault misconduct dating back teenager early moore denied wrongdoing reuters able independently verify allegation'

Performing it for all the examples in the data.

```
[21]:  #Removal of HTML Contents
       def remove_html(text):
           soup = BeautifulSoup(text, "html.parser")
           return soup.get_text()

       #Removal of Punctuation Marks
       def remove_punctuations(text):
           return re.sub('\[[^]]*\]', '', text)

       # Removal of Special Characters
       def remove_characters(text):
           return re.sub("[^a-zA-Z]"," ",text)

       #Removal of stopwords
       def remove_stopwords_and_lemmatization(text):
           final_text = []
           text = text.lower()
           text = nltk.word_tokenize(text)

           for word in text:
               if word not in set(stopwords.words('english')):
                   lemma = nltk.WordNetLemmatizer()
                   word = lemma.lemmatize(word)
                   final_text.append(word)
           return " ".join(final_text)

       #Total function
       def cleaning(text):
```

```
    text = remove_html(text)
    text = remove_punctuations(text)
    text = remove_characters(text)
    text = remove_stopwords_and_lemmatization(text)
    return text

#Apply function on text column
data['text']=data['text'].apply(cleaning)
```

[22]: `data.head()`

[22]:
```
                                               text   target
0   politicsnews u budget fight loom republican fl…      1
1   politicsnews u military accept transgender rec…      1
2   politicsnews senior u republican senator let m…      1
3   politicsnews fbi russia probe helped australia…      1
4   politicsnews trump want postal service charge …      1
```

2.1 Train Test Split

[23]:
```
X_train, X_test, y_train, y_test = train_test_split(data['text'],␣
 ↪data['target'], random_state=0)
```

2.2 Tokenizing

[24]:
```
max_features = 10000
maxlen = 300
```

[25]:
```
tokenizer = text.Tokenizer(num_words=max_features)
tokenizer.fit_on_texts(X_train)
tokenized_train = tokenizer.texts_to_sequences(X_train)
X_train = sequence.pad_sequences(tokenized_train, maxlen=maxlen)
```

[26]:
```
tokenized_test = tokenizer.texts_to_sequences(X_test)
X_test = sequence.pad_sequences(tokenized_test, maxlen=maxlen)
```

[27]: `X_train.size, X_test.size`

[27]: `(10101900, 3367500)`

[28]: `X_train[3363]`

[28]:
```
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,    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,     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,     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,     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,     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,   133,   192,     1,
     612,  1520,  1575,  1477,    17,  4359,   873,  3351,     1,    17,  1520,
    1575,  1477,  2439,   343,   583,  1205,   182,  3466,   207,    15,  1355,
     274,  2766,  2626,  2369,  2866,   149,     3,    99,  1515,  2977,   932,
       2,   689,  1754,   153,  4468,   529,  2701,  1477,     2,   207,   308,
    1845,   321,   666,    17,  6085,  1477,    32,    11,  9895,  5786,   873,
       1,    95,     2,  1675,  3271,  1037,  3466,   211,   432,   503,    45,
       7,  3805,   240,  2436,   493,   868,   240,  5870,     2,   147,  2766,
       2,   859,   619], dtype=int32)
```

2.3 Training Model

```python
[29]: # Define key parameters for a machine learning model
      batch_size = 256 # Number of data samples processed in each step
      epochs = 10 # Number of complete passes through the dataset
      embed_size = 100 # Dimensionality of word embeddings
```

```python
[30]: # Create a sequential neural network model
      model = Sequential()

      # Add a non-trainable embedding layer
      model.add(Embedding(max_features, output_dim=embed_size, input_length=maxlen,
       ↪trainable=False))

      # Add two LSTM layers with specific configurations
      model.add(LSTM(units=128, return_sequences=True, recurrent_dropout=0.25,
       ↪dropout=0.25))
      model.add(LSTM(units=64, recurrent_dropout=0.1, dropout=0.1))

      # Add two fully connected (Dense) layers with ReLU and sigmoid activations
      model.add(Dense(units=32, activation='relu'))
      model.add(Dense(1, activation='sigmoid'))
```

```python
# Compile the model with specific settings
model.compile(optimizer=keras.optimizers.Adam(lr=0.01),
    loss='binary_crossentropy', metrics=['accuracy'])
```

```
2023-10-15 16:43:16.811302: I
tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:168] retrieving
CUDA diagnostic information for host: djoe
2023-10-15 16:43:16.811325: I
tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:175] hostname:
djoe
2023-10-15 16:43:16.811421: I
tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:199] libcuda
reported version is: NOT_FOUND: was unable to find libcuda.so DSO loaded into
this program
2023-10-15 16:43:16.811449: I
tensorflow/compiler/xla/stream_executor/cuda/cuda_diagnostics.cc:203] kernel
reported version is: 525.105.17
WARNING:absl:`lr` is deprecated in Keras optimizer, please use `learning_rate`
or use the legacy optimizer, e.g.,tf.keras.optimizers.legacy.Adam.
```

[31]: `model.summary()`

```
Model: "sequential"

_____
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 300, 100)          1000000

 lstm (LSTM)                 (None, 300, 128)          117248

 lstm_1 (LSTM)               (None, 64)                49408

 dense (Dense)               (None, 32)                2080

 dense_1 (Dense)             (None, 1)                 33

=================================================================
Total params: 1168769 (4.46 MB)
Trainable params: 168769 (659.25 KB)
Non-trainable params: 1000000 (3.81 MB)
_____
```

[32]: `history = model.fit(X_train, y_train, validation_split=0.3, epochs=10,`
    `batch_size=batch_size, shuffle=True, verbose = 1)`

```
Epoch 1/10
93/93 [==============================] - 91s 935ms/step - loss: 0.5028 -
accuracy: 0.7535 - val_loss: 0.4817 - val_accuracy: 0.8149
Epoch 2/10
```

```
93/93 [==============================] - 86s 929ms/step - loss: 0.6644 -
accuracy: 0.6189 - val_loss: 0.5567 - val_accuracy: 0.6679
Epoch 3/10
93/93 [==============================] - 87s 932ms/step - loss: 0.4758 -
accuracy: 0.7617 - val_loss: 1.5183 - val_accuracy: 0.6048
Epoch 4/10
93/93 [==============================] - 87s 937ms/step - loss: 0.5428 -
accuracy: 0.7181 - val_loss: 0.6267 - val_accuracy: 0.7598
Epoch 5/10
93/93 [==============================] - 80s 865ms/step - loss: 0.6954 -
accuracy: 0.5232 - val_loss: 0.6920 - val_accuracy: 0.5202
Epoch 6/10
93/93 [==============================] - 81s 871ms/step - loss: 0.6916 -
accuracy: 0.5247 - val_loss: 0.6916 - val_accuracy: 0.5202
Epoch 7/10
93/93 [==============================] - 80s 860ms/step - loss: 0.6907 -
accuracy: 0.5223 - val_loss: 0.6935 - val_accuracy: 0.4798
Epoch 8/10
93/93 [==============================] - 83s 898ms/step - loss: 0.6910 -
accuracy: 0.5245 - val_loss: 0.6910 - val_accuracy: 0.5202
Epoch 9/10
93/93 [==============================] - 87s 939ms/step - loss: 0.6769 -
accuracy: 0.5618 - val_loss: 0.7113 - val_accuracy: 0.5202
Epoch 10/10
93/93 [==============================] - 87s 941ms/step - loss: 0.5692 -
accuracy: 0.6793 - val_loss: 0.4762 - val_accuracy: 0.7544
```

2.4 Evaluation

```python
[33]: # Print the accuracy of the model on the training data
      print("Accuracy of the model on Training Data is - ", model.evaluate(X_train,␣
       ↪y_train)[1] * 100, "%")


      # Print the accuracy of the model on the testing data
      print("Accuracy of the model on Testing Data is - ", model.evaluate(X_test,␣
       ↪y_test)[1] * 100, "%")
```

```
1053/1053 [==============================] - 48s 46ms/step - loss: 0.4732 -
accuracy: 0.7583
Accuracy of the model on Training Data is -  75.82632899284363 %
351/351 [==============================] - 16s 45ms/step - loss: 0.4738 -
accuracy: 0.7601
Accuracy of the model on Testing Data is -  76.0089099407196 %
```

```python
[34]: # Create a figure for the plot
      plt.figure()


      # Plot the training accuracy over epochs and label it as "Train"
```

```python
plt.plot(history.history["accuracy"], label="Train")

# Plot the testing accuracy over epochs and label it as "Test"
plt.plot(history.history["val_accuracy"], label="Test")

# Set the title and labels for the plot
plt.title("Accuracy")
plt.ylabel("Accuracy")
plt.xlabel("Epochs")

# Display a legend to distinguish the "Train" and "Test" lines
plt.legend()

# Show the plot
plt.show()
```
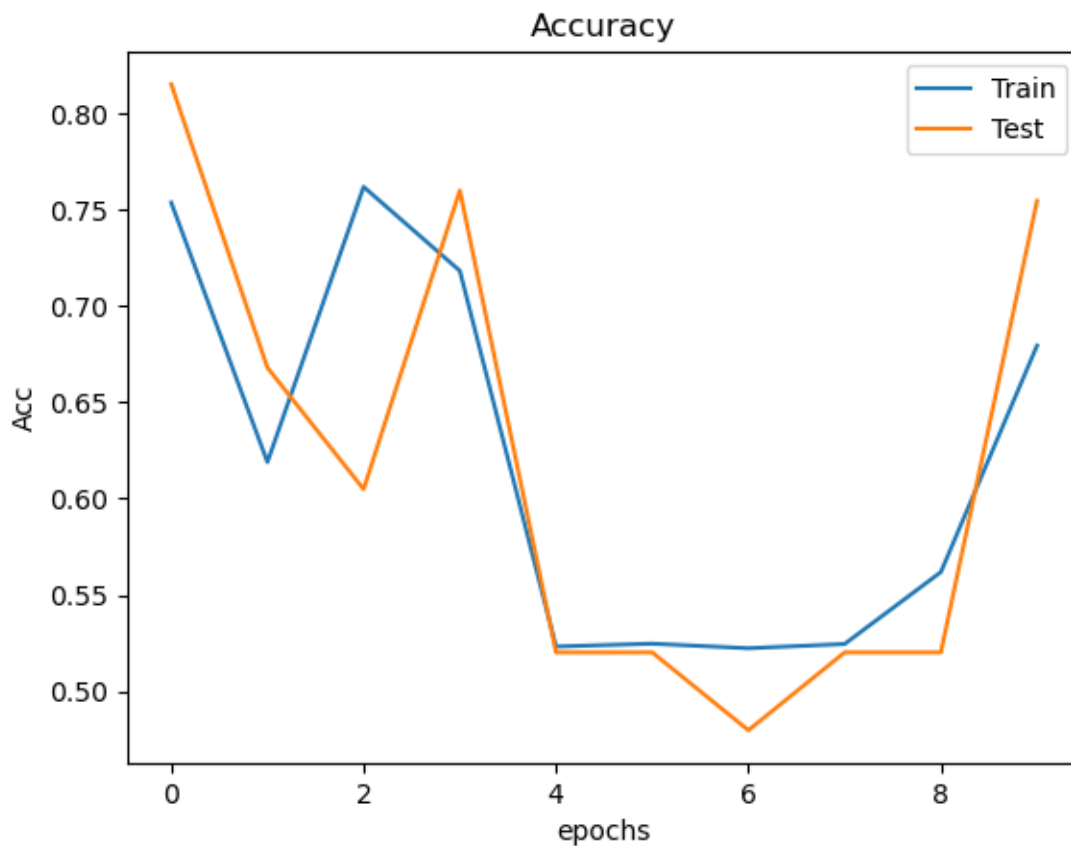


```python
[35]:  # Create a figure for the plot
       plt.figure()

       # Plot the training loss over epochs and label it as "Train"
```
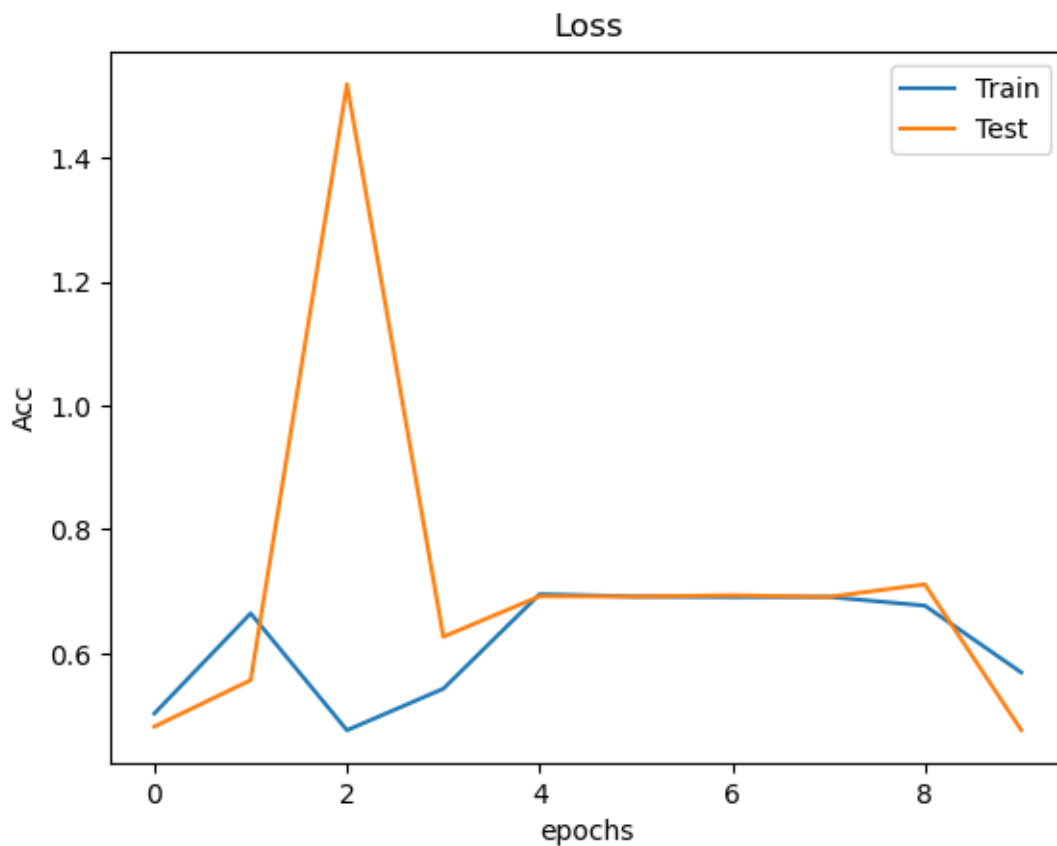
```python
plt.plot(history.history["loss"], label="Train")

# Plot the testing loss over epochs and label it as "Test"
plt.plot(history.history["val_loss"], label="Test")

# Set the title and labels for the plot
plt.title("Loss")
plt.ylabel("Loss")
plt.xlabel("Epochs")

# Display a legend to distinguish the "Train" and "Test" lines
plt.legend()

# Show the plot
plt.show()
```



```python
[50]: model.predict(X_test)
```

351/351 [==============================] - 14s 40ms/step

```
[50]: array([[0.40189645],
             [0.00880686],
             [0.00962298],
             ...,
             [0.01128194],
             [0.72117466],
             [0.7350749 ]], dtype=float32)
```

[ ]: