j5hcllyrn

February 1, 2023

```
[3]: import pandas as pd
     import numpy as np
     import tensorflow as tf
     import string
     import re
```

```
[4]: data = pd.read_csv('US_Crime_Data.csv')
     data.sample(10)
```

```
[4]:             Date                                               Title  \
     1504    5/16/17 21:07                   Enough with the hate crime hoaxes
     3014     6/29/17 8:56  Blue Lives Matter Law Goes Into Effect In Kent…
     4475    8/31/17 13:37  Two high schoolers edited a photo to show blac…
     5777     11/2/17 8:52  Student Arrested For Purposely Infecting Roomm…
     5209    10/11/17 8:57  Sunderland man facing hate crime charges after…
     6699   12/13/17 12:37  Former KKK Priest Pays Restitution to Victims …
     4652    9/12/17 10:48  Man facing felonies after he spit, punched peo…
     3391     7/17/17 2:45  Trump and Modi's silence on hate crimes provid…
     1136     4/20/17 9:34  2 Teenage Girls Charged With Hate Crime for Li…
     1568    5/18/17 19:07  District Attorney Announces Conviction of Susp…

                       Organization         City State  \
     1504           The College Fix     Hillsdale    MI
     3014     LEX18 Lexington KY News     Lexington    KY
     4475           Tri-City Herald     Kennewick    WA
     5777             Opposing Views   Los Angeles    CA
     5209              The Recorder     Greenfield    MA
     6699        World Religion News  Santa Monica    CA
     4652        New York Daily News           NaN   NaN
     3391                 TRT World           NaN   NaN
     1136  Independent Journal Review           NaN   NaN
     1568             EastCountyToday           NaN   NaN

                                                    URL  \
     1504         https://www.thecollegefix.com/post/32630/
     3014  http://www.lex18.com/story/35779526/blue-lives…
     4475  http://www.tri-cityherald.com/news/nation-worl…
```

```
5777   http://www.opposingviews.com/i/society/crime/s…
5209   http://www.recorder.com/Race-driven-home-invas…
6699   http://www.worldreligionnews.com/religion-news…
4652   http://www.nydailynews.com/news/crime/chicago-…
3391   http://www.trtworld.com/opinion/trump-and-modi…
1136   http://ijr.com/2017/04/852326-2-teenage-girls-…
1568   http://eastcountytoday.net/district-attorney-a…
```

```
                                              Keyword  \
1504   american campus capital college crime hate hoa…
3014   bill blue classifying emts harsher hate kentuc…
4475   albuquerque bannerman highschool kkk kob kqre …
5777   backpack brochu channel conn facebook hartford…
5209   apartments assault chief cliffside deflumeri d…
6699   aitcheson butler catholic charlottesville cons…
4652   boucher chicago counts gerald man mitchell pun…
3391   beef cow criminals donald india islamic junaid…
1136   anne asia baltimore east fire joy maryland per…
1568   assailants attack attorney chase colton crime …
```

```
                                              Summary
1504   We are in the midst of a college hoax epidemic…
3014   KENTUCKY (LEX 18) - A new Kentucky law goes in…
4475   Two students, including one who played for the…
5777   A former student at Connecticut's University o…
5209   By ANDY CASTILLO\tSUNDERLAND - A local man who…
6699   By Armand Wagner, Bivels [CC BY 3.0], via Wiki…
4652   BY Brett Bodner\tWilliam Boucher is facing fel…
3391   Both US President Donald Trump and Indian Prim…
1136   By Parker Lee\tPresident Trump may be the new …
1568   On the night of September 25, 2016 Mr. Khalsa …
```

[3]: `data.isna().sum()`

```
[3]: Date              0
     Title             1
     Organization      0
     City           1167
     State          1245
     URL               0
     Keyword        1176
     Summary        2256
     dtype: int64
```

For this model we just need headlines

```
[2]: df = data[['Title']]
     df = df.dropna().reset_index(drop = True)
     df.head()
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In [2], line 1
----> 1 df = data[['Title']]
      2 df = df.dropna().reset_index(drop = True)
      3 df.head()

NameError: name 'data' is not defined
```

```
[5]: df['Title'][123]
```

```
[5]: 'St. Louis suburb victimized by cemetery vandalism mulling hate crimes registry'
```

```
[6]: df.shape
```

```
[6]: (6782, 1)
```

Removing Punctuations

```
[7]: def clean_text(df):
         df['Title'] = df['Title'].apply(lambda x : x.lower())
         tokens = df['Title'].str.replace('[{}]'.format(string.punctuation), '')
         return tokens
```

```
[8]: tokens = clean_text(df)
```

```
C:\Users\Nuwanga Wijamuni\AppData\Local\Temp\ipykernel_20704\363633391.py:3:
FutureWarning: The default value of regex will change from True to False in a
future version.
  tokens = df['Title'].str.replace('[{}]'.format(string.punctuation), '')
```

```
[9]: tokens
```

```
[9]: 0       pizza hut driver who killed coworker with shot…
     1       residents of nj township receiving kkk promoti…
     2       house oks bill to expand kentuckys hate crimes…
     3       amid protests blue lives matter bill passes ke…
     4       lafourche inmates charged with hate crimes in …
                                   …
     6777    police rash of vandalism on margate menorahs i…
     6778    exstudent accused of smearing used tampon on r…
     6779    queensbury woman charged with hate crime again…
     6780    hate crime hartford student accused of smearin…
```

```
6781    police make arrest following alleged hate crim…
Name: Title, Length: 6782, dtype: object
```

[10]: `len(set(tokens))`

[10]: 6563

Tokenizing

[11]:
```python
tokenizer = tf.keras.preprocessing.text.Tokenizer()
tokenizer.fit_on_texts(tokens)
seq = tokenizer.texts_to_sequences(tokens)
```

[12]: `seq[:10]`

[12]:
```
[[3973, 3974, 450, 47, 88, 3975, 12, 2211, 103, 1370, 2803],
 [830, 4, 423, 1558, 2212, 470, 3976, 672],
 [257, 1846, 75, 5, 593, 3977, 1, 6, 36],
 [721, 898, 480, 274, 301, 75, 481, 634, 257],
 [3978, 2804, 16, 12, 1, 6, 3, 424, 19],
 [13, 722, 1, 2, 48, 594, 3979, 5, 401],
 [3980, 480, 274, 301, 75, 21, 1371, 2805, 5, 214, 3981, 173],
 [194, 320, 18, 402, 2806, 239, 15, 98, 38],
 [194,
  320,
  18,
  402,
  2806,
  239,
  15,
  98,
  38,
  635,
  174,
  556,
  557,
  14,
  498,
  558,
  275,
  42],
 [33, 174, 673, 302, 2, 1372, 2807, 3, 2213]]
```

[13]:
```python
x = []
y = []
total_words_drop = 0
for i in seq:
```

```
        if len(i) > 1:
            for j in range(1, len(i)):
                x.append(i[:j])
                y.append(i[j])

        else :
            total_words_drop +=1
print('Total Words Dropped : {}'.format(total_words_drop))
```

Total Words Dropped : 12

[14]:
```
y[: 10]
```

[14]: [3974, 450, 47, 88, 3975, 12, 2211, 103, 1370, 2803]

Padding sequences

[15]:
```
x = tf.keras.preprocessing.sequence.pad_sequences(x)
```

[16]:
```
x.shape
```

[16]: (64701, 49)

Shaping y same as x

[17]:
```
y = tf.keras.utils.to_categorical(y)
```

[18]:
```
y.shape
```

[18]: (64701, 7569)

Vocab Size : total no. of unique words

[19]:
```
vocab_size = len(tokenizer.word_index) + 1
vocab_size
```

[19]: 7569

LSTM Model

[20]:
```
model = tf.keras.Sequential([tf.keras.layers.Embedding(vocab_size,49 ),
                             tf.keras.layers.LSTM(100, return_sequences = True),
                             tf.keras.layers.LSTM(100),
                             tf.keras.layers.Dense(100, activation = 'relu'),
                             tf.keras.layers.Dense(vocab_size, activation =
 'softmax')])
```

[21]:
```
model.summary()
```

```
Model: "sequential"

_____
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, None, 49)          370881

 lstm (LSTM)                 (None, None, 100)         60000

 lstm_1 (LSTM)               (None, 100)               80400

 dense (Dense)               (None, 100)               10100

 dense_1 (Dense)             (None, 7569)              764469


=================================================================
Total params: 1,285,850
Trainable params: 1,285,850
Non-trainable params: 0

_____
```

```python
[22]: model.compile(loss  = 'categorical_crossentropy',
                optimizer = 'adam',
                metrics = ['accuracy'],
                )
```

```python
[24]: history = model.fit(x,y,
                    epochs = 100,
                    batch_size = 256,
                    callbacks = tf.keras.callbacks.EarlyStopping(monitor =␣
      ↪'val_loss',
                                                          patience = 5,
                                                              ␣
      ↪restore_best_weights = True))
```

```
Epoch 1/100
253/253 [==============================] - ETA: 0s - loss: 6.2247 - accuracy:
0.1225WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 72s 286ms/step - loss: 6.2247 -
accuracy: 0.1225
Epoch 2/100
253/253 [==============================] - ETA: 0s - loss: 6.0780 - accuracy:
0.1297WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 78s 310ms/step - loss: 6.0780 -
accuracy: 0.1297
Epoch 3/100
253/253 [==============================] - ETA: 0s - loss: 5.9329 - accuracy:
```

```
0.1324WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 73s 288ms/step - loss: 5.9329 -
accuracy: 0.1324
Epoch 4/100
253/253 [==============================] - ETA: 0s - loss: 5.7830 - accuracy:
0.1412WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 278ms/step - loss: 5.7830 -
accuracy: 0.1412
Epoch 5/100
253/253 [==============================] - ETA: 0s - loss: 5.6355 - accuracy:
0.1506WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 153s 607ms/step - loss: 5.6355 -
accuracy: 0.1506
Epoch 6/100
253/253 [==============================] - ETA: 0s - loss: 5.4947 - accuracy:
0.1592WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 120s 473ms/step - loss: 5.4947 -
accuracy: 0.1592
Epoch 7/100
253/253 [==============================] - ETA: 0s - loss: 5.3637 - accuracy:
0.1679WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 67s 263ms/step - loss: 5.3637 -
accuracy: 0.1679
Epoch 8/100
253/253 [==============================] - ETA: 0s - loss: 5.2324 - accuracy:
0.1773WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 272ms/step - loss: 5.2324 -
accuracy: 0.1773
Epoch 9/100
253/253 [==============================] - ETA: 0s - loss: 5.0800 - accuracy:
0.1873WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 68s 269ms/step - loss: 5.0800 -
accuracy: 0.1873
Epoch 10/100
253/253 [==============================] - ETA: 0s - loss: 4.9379 - accuracy:
0.1962WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 276ms/step - loss: 4.9379 -
accuracy: 0.1962
Epoch 11/100
253/253 [==============================] - ETA: 0s - loss: 4.8041 - accuracy:
```

```
0.2065WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 272ms/step - loss: 4.8041 -
accuracy: 0.2065
Epoch 12/100
253/253 [==============================] - ETA: 0s - loss: 4.6843 - accuracy:
0.2159WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 272ms/step - loss: 4.6843 -
accuracy: 0.2159
Epoch 13/100
253/253 [==============================] - ETA: 0s - loss: 4.5663 - accuracy:
0.2244 WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 11957s 47s/step - loss: 4.5663 -
accuracy: 0.2244
Epoch 14/100
253/253 [==============================] - ETA: 0s - loss: 4.4572 - accuracy:
0.2313WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 66s 260ms/step - loss: 4.4572 -
accuracy: 0.2313
Epoch 15/100
253/253 [==============================] - ETA: 0s - loss: 4.3533 - accuracy:
0.2400WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 276ms/step - loss: 4.3533 -
accuracy: 0.2400
Epoch 16/100
253/253 [==============================] - ETA: 0s - loss: 4.2541 - accuracy:
0.2480WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 65s 258ms/step - loss: 4.2541 -
accuracy: 0.2480
Epoch 17/100
253/253 [==============================] - ETA: 0s - loss: 4.1571 - accuracy:
0.2549WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 64s 251ms/step - loss: 4.1571 -
accuracy: 0.2549
Epoch 18/100
253/253 [==============================] - ETA: 0s - loss: 4.0653 - accuracy:
0.2632WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 66s 259ms/step - loss: 4.0653 -
accuracy: 0.2632
Epoch 19/100
253/253 [==============================] - ETA: 0s - loss: 3.9739 - accuracy:
```

```
0.2704WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 271ms/step - loss: 3.9739 -
accuracy: 0.2704
Epoch 20/100
253/253 [==============================] - ETA: 0s - loss: 3.8844 - accuracy:
0.2791WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 65s 256ms/step - loss: 3.8844 -
accuracy: 0.2791
Epoch 21/100
253/253 [==============================] - ETA: 0s - loss: 3.7985 - accuracy:
0.2873WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 64s 253ms/step - loss: 3.7985 -
accuracy: 0.2873
Epoch 22/100
253/253 [==============================] - ETA: 0s - loss: 3.7132 - accuracy:
0.2952WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 65s 257ms/step - loss: 3.7132 -
accuracy: 0.2952
Epoch 23/100
253/253 [==============================] - ETA: 0s - loss: 3.6349 - accuracy:
0.3049WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 64s 254ms/step - loss: 3.6349 -
accuracy: 0.3049
Epoch 24/100
253/253 [==============================] - ETA: 0s - loss: 3.5583 - accuracy:
0.3147WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 67s 267ms/step - loss: 3.5583 -
accuracy: 0.3147
Epoch 25/100
253/253 [==============================] - ETA: 0s - loss: 3.4829 - accuracy:
0.3256WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 73s 287ms/step - loss: 3.4829 -
accuracy: 0.3256
Epoch 26/100
253/253 [==============================] - ETA: 0s - loss: 3.4123 - accuracy:
0.3345WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 75s 296ms/step - loss: 3.4123 -
accuracy: 0.3345
Epoch 27/100
253/253 [==============================] - ETA: 0s - loss: 3.3399 - accuracy:
```

0.3446WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 76s 300ms/step - loss: 3.3399 - accuracy: 0.3446
Epoch 28/100
253/253 [==============================] - ETA: 0s - loss: 3.2772 - accuracy: 0.3515WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 76s 302ms/step - loss: 3.2772 - accuracy: 0.3515
Epoch 29/100
253/253 [==============================] - ETA: 0s - loss: 3.2163 - accuracy: 0.3613WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 77s 303ms/step - loss: 3.2163 - accuracy: 0.3613
Epoch 30/100
253/253 [==============================] - ETA: 0s - loss: 3.1539 - accuracy: 0.3691WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 79s 313ms/step - loss: 3.1539 - accuracy: 0.3691
Epoch 31/100
253/253 [==============================] - ETA: 0s - loss: 3.0964 - accuracy: 0.3767WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 77s 306ms/step - loss: 3.0964 - accuracy: 0.3767
Epoch 32/100
253/253 [==============================] - ETA: 0s - loss: 3.0426 - accuracy: 0.3838WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 76s 300ms/step - loss: 3.0426 - accuracy: 0.3838
Epoch 33/100
253/253 [==============================] - ETA: 0s - loss: 2.9874 - accuracy: 0.3929WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 276ms/step - loss: 2.9874 - accuracy: 0.3929
Epoch 34/100
253/253 [==============================] - ETA: 0s - loss: 2.9395 - accuracy: 0.3998WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 66s 261ms/step - loss: 2.9395 - accuracy: 0.3998
Epoch 35/100
253/253 [==============================] - ETA: 0s - loss: 2.8896 - accuracy:

0.4078WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 65s 257ms/step - loss: 2.8896 -
accuracy: 0.4078
Epoch 36/100
253/253 [==============================] - ETA: 0s - loss: 2.8463 - accuracy:
0.4133WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 65s 257ms/step - loss: 2.8463 -
accuracy: 0.4133
Epoch 37/100
253/253 [==============================] - ETA: 0s - loss: 2.8011 - accuracy:
0.4208WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 64s 255ms/step - loss: 2.8011 -
accuracy: 0.4208
Epoch 38/100
253/253 [==============================] - ETA: 0s - loss: 2.7542 - accuracy:
0.4275WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 64s 254ms/step - loss: 2.7542 -
accuracy: 0.4275
Epoch 39/100
253/253 [==============================] - ETA: 0s - loss: 2.7111 - accuracy:
0.4346WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 272ms/step - loss: 2.7111 -
accuracy: 0.4346
Epoch 40/100
253/253 [==============================] - ETA: 0s - loss: 2.6699 - accuracy:
0.4400WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 78s 309ms/step - loss: 2.6699 -
accuracy: 0.4400
Epoch 41/100
253/253 [==============================] - ETA: 0s - loss: 2.6341 - accuracy:
0.4449WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 78s 308ms/step - loss: 2.6341 -
accuracy: 0.4449
Epoch 42/100
253/253 [==============================] - ETA: 0s - loss: 2.5951 - accuracy:
0.4520WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 82s 323ms/step - loss: 2.5951 -
accuracy: 0.4520
Epoch 43/100
253/253 [==============================] - ETA: 0s - loss: 2.5589 - accuracy:

```
0.4578WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 76s 299ms/step - loss: 2.5589 -
accuracy: 0.4578
Epoch 44/100
253/253 [==============================] - ETA: 0s - loss: 2.5258 - accuracy:
0.4619WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 75s 295ms/step - loss: 2.5258 -
accuracy: 0.4619
Epoch 45/100
253/253 [==============================] - ETA: 0s - loss: 2.4908 - accuracy:
0.4670WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 79s 311ms/step - loss: 2.4908 -
accuracy: 0.4670
Epoch 46/100
253/253 [==============================] - ETA: 0s - loss: 2.4595 - accuracy:
0.4729 WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 9284s 37s/step - loss: 2.4595 -
accuracy: 0.4729
Epoch 47/100
253/253 [==============================] - ETA: 0s - loss: 2.4275 - accuracy:
0.4785WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 275ms/step - loss: 2.4275 -
accuracy: 0.4785
Epoch 48/100
253/253 [==============================] - ETA: 0s - loss: 2.3939 - accuracy:
0.4844WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 108s 428ms/step - loss: 2.3939 -
accuracy: 0.4844
Epoch 49/100
253/253 [==============================] - ETA: 0s - loss: 2.3627 - accuracy:
0.4896WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 114s 451ms/step - loss: 2.3627 -
accuracy: 0.4896
Epoch 50/100
253/253 [==============================] - ETA: 0s - loss: 2.3341 - accuracy:
0.4944WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 91s 360ms/step - loss: 2.3341 -
accuracy: 0.4944
Epoch 51/100
253/253 [==============================] - ETA: 0s - loss: 2.3063 - accuracy:
```

```
0.4978WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 114s 453ms/step - loss: 2.3063 -
accuracy: 0.4978
Epoch 52/100
253/253 [==============================] - ETA: 0s - loss: 2.2809 - accuracy:
0.5031WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 122s 481ms/step - loss: 2.2809 -
accuracy: 0.5031
Epoch 53/100
253/253 [==============================] - ETA: 0s - loss: 2.2552 - accuracy:
0.5061WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 123s 488ms/step - loss: 2.2552 -
accuracy: 0.5061
Epoch 54/100
253/253 [==============================] - ETA: 0s - loss: 2.2287 - accuracy:
0.5097WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 123s 485ms/step - loss: 2.2287 -
accuracy: 0.5097
Epoch 55/100
253/253 [==============================] - ETA: 0s - loss: 2.1989 - accuracy:
0.5156WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 79s 313ms/step - loss: 2.1989 -
accuracy: 0.5156
Epoch 56/100
253/253 [==============================] - ETA: 0s - loss: 2.1742 - accuracy:
0.5204WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 107s 423ms/step - loss: 2.1742 -
accuracy: 0.5204
Epoch 57/100
253/253 [==============================] - ETA: 0s - loss: 2.1479 - accuracy:
0.5244WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 139s 551ms/step - loss: 2.1479 -
accuracy: 0.5244
Epoch 58/100
253/253 [==============================] - ETA: 0s - loss: 2.1267 - accuracy:
0.5295WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 67s 264ms/step - loss: 2.1267 -
accuracy: 0.5295
Epoch 59/100
253/253 [==============================] - ETA: 0s - loss: 2.1048 - accuracy:
```

```
0.5343WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 101s 399ms/step - loss: 2.1048 -
accuracy: 0.5343
Epoch 60/100
253/253 [==============================] - ETA: 0s - loss: 2.0810 - accuracy:
0.5362WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 140s 553ms/step - loss: 2.0810 -
accuracy: 0.5362
Epoch 61/100
253/253 [==============================] - ETA: 0s - loss: 2.0609 - accuracy:
0.5409WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 94s 371ms/step - loss: 2.0609 -
accuracy: 0.5409
Epoch 62/100
253/253 [==============================] - ETA: 0s - loss: 2.0352 - accuracy:
0.5449WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 105s 413ms/step - loss: 2.0352 -
accuracy: 0.5449
Epoch 63/100
253/253 [==============================] - ETA: 0s - loss: 2.0159 - accuracy:
0.5486WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 141s 558ms/step - loss: 2.0159 -
accuracy: 0.5486
Epoch 64/100
253/253 [==============================] - ETA: 0s - loss: 1.9910 - accuracy:
0.5531WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 119s 471ms/step - loss: 1.9910 -
accuracy: 0.5531
Epoch 65/100
253/253 [==============================] - ETA: 0s - loss: 1.9724 - accuracy:
0.5578WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 70s 275ms/step - loss: 1.9724 -
accuracy: 0.5578
Epoch 66/100
253/253 [==============================] - ETA: 0s - loss: 1.9506 - accuracy:
0.5612WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 69s 273ms/step - loss: 1.9506 -
accuracy: 0.5612
Epoch 67/100
253/253 [==============================] - ETA: 0s - loss: 1.9346 - accuracy:
```

```
0.5626WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 81s 319ms/step - loss: 1.9346 -
accuracy: 0.5626
Epoch 68/100
253/253 [==============================] - ETA: 0s - loss: 1.9135 - accuracy:
0.5650WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 143s 566ms/step - loss: 1.9135 -
accuracy: 0.5650
Epoch 69/100
253/253 [==============================] - ETA: 0s - loss: 1.8954 - accuracy:
0.5697WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 134s 527ms/step - loss: 1.8954 -
accuracy: 0.5697
Epoch 70/100
253/253 [==============================] - ETA: 0s - loss: 1.8719 - accuracy:
0.5759WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 142s 562ms/step - loss: 1.8719 -
accuracy: 0.5759
Epoch 71/100
253/253 [==============================] - ETA: 0s - loss: 1.8550 - accuracy:
0.5784WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 136s 539ms/step - loss: 1.8550 -
accuracy: 0.5784
Epoch 72/100
253/253 [==============================] - ETA: 0s - loss: 1.8358 - accuracy:
0.5824WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 143s 564ms/step - loss: 1.8358 -
accuracy: 0.5824
Epoch 73/100
253/253 [==============================] - ETA: 0s - loss: 1.8210 - accuracy:
0.5856WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 90s 355ms/step - loss: 1.8210 -
accuracy: 0.5856
Epoch 74/100
253/253 [==============================] - ETA: 0s - loss: 1.8004 - accuracy:
0.5893WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 74s 292ms/step - loss: 1.8004 -
accuracy: 0.5893
Epoch 75/100
253/253 [==============================] - ETA: 0s - loss: 1.7875 - accuracy:
```

```
0.5913WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 76s 299ms/step - loss: 1.7875 -
accuracy: 0.5913
Epoch 76/100
253/253 [==============================] - ETA: 0s - loss: 1.7684 - accuracy:
0.5942WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 112s 443ms/step - loss: 1.7684 -
accuracy: 0.5942
Epoch 77/100
253/253 [==============================] - ETA: 0s - loss: 1.7478 - accuracy:
0.5997WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 88s 349ms/step - loss: 1.7478 -
accuracy: 0.5997
Epoch 78/100
253/253 [==============================] - ETA: 0s - loss: 1.7334 - accuracy:
0.6023WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 79s 310ms/step - loss: 1.7334 -
accuracy: 0.6023
Epoch 79/100
253/253 [==============================] - ETA: 0s - loss: 1.7166 - accuracy:
0.6056WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 74s 291ms/step - loss: 1.7166 -
accuracy: 0.6056
Epoch 80/100
253/253 [==============================] - ETA: 0s - loss: 1.7038 - accuracy:
0.6081WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 100s 396ms/step - loss: 1.7038 -
accuracy: 0.6081
Epoch 81/100
253/253 [==============================] - ETA: 0s - loss: 1.6881 - accuracy:
0.6094WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 145s 573ms/step - loss: 1.6881 -
accuracy: 0.6094
Epoch 82/100
253/253 [==============================] - ETA: 0s - loss: 1.6736 - accuracy:
0.6154WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 129s 511ms/step - loss: 1.6736 -
accuracy: 0.6154
Epoch 83/100
253/253 [==============================] - ETA: 0s - loss: 1.6566 - accuracy:
```

0.6161WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 92s 365ms/step - loss: 1.6566 -
accuracy: 0.6161
Epoch 84/100
253/253 [==============================] - ETA: 0s - loss: 1.6410 - accuracy:
0.6194WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 125s 495ms/step - loss: 1.6410 -
accuracy: 0.6194
Epoch 85/100
253/253 [==============================] - ETA: 0s - loss: 1.6271 - accuracy:
0.6215WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 128s 504ms/step - loss: 1.6271 -
accuracy: 0.6215
Epoch 86/100
253/253 [==============================] - ETA: 0s - loss: 1.6131 - accuracy:
0.6240WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 74s 294ms/step - loss: 1.6131 -
accuracy: 0.6240
Epoch 87/100
253/253 [==============================] - ETA: 0s - loss: 1.5969 - accuracy:
0.6293WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 75s 296ms/step - loss: 1.5969 -
accuracy: 0.6293
Epoch 88/100
253/253 [==============================] - ETA: 0s - loss: 1.5824 - accuracy:
0.6322WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 75s 298ms/step - loss: 1.5824 -
accuracy: 0.6322
Epoch 89/100
253/253 [==============================] - ETA: 0s - loss: 1.5653 - accuracy:
0.6355WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 153s 605ms/step - loss: 1.5653 -
accuracy: 0.6355
Epoch 90/100
253/253 [==============================] - ETA: 0s - loss: 1.5534 - accuracy:
0.6375WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 146s 575ms/step - loss: 1.5534 -
accuracy: 0.6375
Epoch 91/100
253/253 [==============================] - ETA: 0s - loss: 1.5425 - accuracy:

0.6383WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 72s 285ms/step - loss: 1.5425 -
accuracy: 0.6383
Epoch 92/100
253/253 [==============================] - ETA: 0s - loss: 1.5285 - accuracy:
0.6426WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 72s 284ms/step - loss: 1.5285 -
accuracy: 0.6426
Epoch 93/100
253/253 [==============================] - ETA: 0s - loss: 1.5151 - accuracy:
0.6449WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 72s 283ms/step - loss: 1.5151 -
accuracy: 0.6449
Epoch 94/100
253/253 [==============================] - ETA: 0s - loss: 1.5008 - accuracy:
0.6475WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 71s 280ms/step - loss: 1.5008 -
accuracy: 0.6475
Epoch 95/100
253/253 [==============================] - ETA: 0s - loss: 1.4875 - accuracy:
0.6510 WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 4439s 18s/step - loss: 1.4875 -
accuracy: 0.6510
Epoch 96/100
253/253 [==============================] - ETA: 0s - loss: 1.4761 - accuracy:
0.6533WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 60s 239ms/step - loss: 1.4761 -
accuracy: 0.6533
Epoch 97/100
253/253 [==============================] - ETA: 0s - loss: 1.4616 - accuracy:
0.6545WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 61s 243ms/step - loss: 1.4616 -
accuracy: 0.6545
Epoch 98/100
253/253 [==============================] - ETA: 0s - loss: 1.4548 - accuracy:
0.6567WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 62s 244ms/step - loss: 1.4548 -
accuracy: 0.6567
Epoch 99/100
253/253 [==============================] - ETA: 0s - loss: 1.4346 - accuracy:

```
0.6632WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 80s 318ms/step - loss: 1.4346 -
accuracy: 0.6632
Epoch 100/100
253/253 [==============================] - ETA: 0s - loss: 1.4294 - accuracy:
0.6622WARNING:tensorflow:Early stopping conditioned on metric `val_loss` which
is not available. Available metrics are: loss,accuracy
253/253 [==============================] - 85s 337ms/step - loss: 1.4294 -
accuracy: 0.6622
```

[25]:
```python
model.save('model.h5')
```

[26]:
```python
vocab_array = np.array(list(tokenizer.word_index.keys()))
vocab_array
```

[26]:
```
array(['hate', 'crime', 'in', …, 'walks', 'lgbti', 'retires'],
      dtype='<U23')
```

Final Function for Predictions

[27]:
```python
def make_predictions(text, n_words):
    for i in range(n_words):
        text_tokenize = tokenizer.texts_to_sequences([text])
        text_padding  = tf.keras.preprocessing.sequence.
 ↪pad_sequences(text_tokenize, maxlen = 49)
        prediction = np.squeeze(np.argmax(model.predict(text_padding), axis =␣
 ↪-1))
        prediction = str(vocab_array[prediction - 1])
        text += " " + prediction
    return text
```

Testing Model

[28]:
```python
make_predictions('california',5)
```

```
1/1 [==============================] - 2s 2s/step
1/1 [==============================] - 0s 29ms/step
1/1 [==============================] - 0s 22ms/step
1/1 [==============================] - 0s 26ms/step
1/1 [==============================] - 0s 24ms/step
```

[28]:
```
'california hate crimes rose in 2016'
```

[29]:
```python
make_predictions('new york',8)
```

```
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 19ms/step
1/1 [==============================] - 0s 27ms/step
```

```
1/1 [==============================] - 0s 23ms/step
1/1 [==============================] - 0s 22ms/step
1/1 [==============================] - 0s 19ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 20ms/step
```

[29]: 'new york man charged with hate crime for allegedly yanking'

[30]: `make_predictions('highway',8)`

```
1/1 [==============================] - 0s 23ms/step
1/1 [==============================] - 0s 30ms/step
1/1 [==============================] - 0s 18ms/step
1/1 [==============================] - 0s 27ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 22ms/step
```

[30]: 'highway annual torres charged with racial slur allegations cut'

[33]: `make_predictions('home',8)`

```
1/1 [==============================] - 0s 29ms/step
1/1 [==============================] - 0s 29ms/step
1/1 [==============================] - 0s 32ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 27ms/step
1/1 [==============================] - 0s 25ms/step
1/1 [==============================] - 0s 24ms/step
1/1 [==============================] - 0s 20ms/step
```

[33]: 'home charged with hate crime in attack on elderly'

[36]: `make_predictions('pizza',12)`

```
1/1 [==============================] - 0s 35ms/step
1/1 [==============================] - 0s 28ms/step
1/1 [==============================] - 0s 24ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 27ms/step
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 25ms/step
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 20ms/step
1/1 [==============================] - 0s 24ms/step
1/1 [==============================] - 0s 22ms/step
1/1 [==============================] - 0s 23ms/step
```

[36]: 'pizza hut driver who killed coworker with shotgun found mentally competent to rant'

[37]: `make_predictions('allegations',9)`

```
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 24ms/step
1/1 [==============================] - 0s 22ms/step
1/1 [==============================] - 0s 24ms/step
1/1 [==============================] - 0s 28ms/step
1/1 [==============================] - 0s 27ms/step
1/1 [==============================] - 0s 21ms/step
1/1 [==============================] - 0s 30ms/step
1/1 [==============================] - 0s 26ms/step
```

[37]: 'allegations hate crimes are the start rate law but true'