

✓ Fake News Classifier Using LSTM

Dataset: <https://www.kaggle.com/c/fake-news/data#>

```
import pandas as pd
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
df=pd.read_csv('/content/drive/MyDrive/fake news detection.csv')
```

```
df.head()
```

	Unnamed: 0		title		text	label
0	8476		You Can Smell Hillary's Fear		Daniel Greenfield, a Shillman Journalism Fello...	FAKE
1	10294	Watch The Exact Moment Paul Ryan Committed Pol...			Google Pinterest Digg LinkedIn Reddit Stumbleu...	FAKE
2	3608	Kerry to go to Paris in gesture of sympathy			U.S. Secretary of State John F. Kerry said Mon...	REAL
3	10142	Bernie supporters on Twitter erupt in anger ag...			— Kaydee King (@KaydeeKing) November 9, 2016 T...	FAKE
4	875	The Battle of New York: Why This Primary Matters			It's primary day in New York and front-runners...	REAL

```
###Drop Nan Values
df=df.dropna()
```

```
## Get the Independent Features
```

```
X=df.drop('label',axis=1)
```

```
## Get the Dependent features
y=df['label']
```

```
X.shape
```

```
(6335, 3)
```

```
y.shape
```

```
(6335,)
```

```
import tensorflow as tf
```

```
tf.__version__
```

```
'2.15.0'
```

```
from tensorflow.keras.layers import Embedding
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.preprocessing.text import one_hot
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dense
```

```
### Vocabulary size
voc_size=5000
```

✓ Onehot Representation

```
messages=X.copy()

messages['title'][1]

'Watch The Exact Moment Paul Ryan Committed Political Suicide At A Trump Rally (VIDEO)'

messages.reset_index(inplace=True)

import nltk
import re
from nltk.corpus import stopwords

nltk.download('stopwords')

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
True

### Dataset Preprocessing
from nltk.stem.porter import PorterStemmer
ps = PorterStemmer()
corpus = []
for i in range(0, len(messages)):
    print(i)
    review = re.sub('[^a-zA-Z]', ' ', messages['title'][i])
    review = review.lower()
    review = review.split()

    review = [ps.stem(word) for word in review if not word in stopwords.words('english')]
    review = ' '.join(review)
    corpus.append(review)
```

Streaming output truncated to the last 5000 lines.

1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379

1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391

corpus

```
[ 'smell hillari fear',
  'watch exact moment paul ryan commit polit suicid trump ralli video',
  'kerri go pari gestur sympathi',
  'berni support twitter erupt anger dnc tri warn',
  'battl new york primari matter',
  'tehran usa',
  'girl horrifi watch boyfriend left facetim',
  'britain schindler die',
  'fact check trump clinton command chief forum',
  'iran reportedli make new push uranium concess nuclear talk',
  'three clinton iowa glimps fire elud hillari clinton campaign',
  'donald trump shockingli weak deleg game somehow got even wors',
  'strong solar storm tech risk today news oct video',
  'way america prepar world war',
  'trump take cruz lightli',
  'women lead differ',
  'shock michel obama hillari caught glamor date rape promot',
  'hillari clinton huge troubl america notic sick thing hidden pictur liberti writer news',
  'iran bill obama like',
  'chart explain everyth need know partisanship america',
  'slipperi slope trump propos ban muslim',
  'episod sunday wire hail deplor special guest randi j',
  'hillari clinton make bipartisan appeal staten island',
  'new senat major leader main goal gop scari',
  'inferno overpopul myth',
  'anti trump forc seek last ditch deleg revolt',
  'sander trounc clinton w va make differ',
  'donald trump chang campaign slogan prove racist',
  'pure chao donald trump campaign manag offer glimps govern style spoiler alert terribl',
  'syrian war report novemb syrian militari deploy advanc battl tank aleppo',
  'gop insid carli crush',
  'jeffrey sewel et al metabiolog face face artifici intellig video',
  'necessari relax stretch',
  'brexit encourag uk trade non eu state includ russia',
  'interview socrat anoth day life empir near endgam fool ever hope chang',
  'mike penc drape shawl immodest ladi justic statu onion america finest news sourc',
  'first ever hindu woman elect congress',
  'donald grope hillari trump weiner sext',
  'ex assist fbi director clinton crime famili',
  'hillari want aggress interventionist foreign polici',
  'parti want craft populist messag',
  'first take wall street bid goodbye june hike',
  'real disclosur secret alien base found moon tycho crater',
  'homeless woman protect trump walk fame star violent leftist',
  'tweet storm saturday clinton continu mock trump friday meltdown',
  'signific pipelin spill already year expos troubl safeti record',
  'obama make right call tough afghanistan',
  'senat race rank dem attack gop lay swing state groundwork',
  'know boy want rape court throw muslim migrant child sex charg german shock',
  'pieczenik rogu fbi agent wikileak spearhead movement stop clinton steal white hous truthfe',
  'american polit reach peak polar',
  'anti muhammad cartoon contest free speech deliber provoc video',
  'effect substanc abus individu famili commun',
  'tree shape vertic farm grow acr urban crop',
  'new comment featur ad',
  'world newspap react hebdo attack',
  'ying yang gold silver set',
  'close primari warp democraci',
```

```
onehot_repr=[one_hot(words,voc_size)for words in corpus]
onehot_repr
```

```
[[2931, 697, 4204],
 [4469, 2489, 1816, 4567, 2011, 1599, 2051, 3890, 4616, 2548, 3063],
 [2611, 3367, 2293, 857, 4225],
 [1883, 295, 2463, 84, 2692, 2768, 2003, 2016],
 [3617, 4113, 82, 3856, 3182],
 [3796, 126],
```

```
[46, 4916, 4469, 175, 1746, 824],
[3812, 1212, 2345],
[4061, 426, 4616, 4033, 816, 2922, 2255],
[3549, 995, 4318, 4113, 769, 1894, 943, 1637, 4676],
[4597, 4033, 3688, 2806, 2886, 1918, 697, 4033, 4858],
[2229, 4616, 1872, 2964, 1102, 1295, 4195, 2642, 536, 292],
[1666, 422, 742, 2834, 2785, 2737, 2855, 2520, 3063],
[4222, 3044, 2570, 2438, 262],
[4616, 4568, 646, 452],
[3369, 180, 3086],
[3272, 4658, 4807, 697, 2614, 3041, 223, 1248, 1251],
[697, 4033, 4822, 2603, 3044, 1390, 1831, 2171, 488, 4188, 2701, 3168, 2855],
[3549, 1903, 4807, 1553],
[2133, 4197, 4409, 258, 432, 3054, 3044],
[4141, 2870, 4616, 4954, 4850, 4033],
[4765, 1392, 1442, 4613, 1308, 1035, 3712, 3941, 482],
[697, 4033, 4318, 205, 1786, 297, 1777],
[4113, 2740, 1295, 4947, 2422, 62, 1124, 4222],
[3737, 3807, 1133],
[2087, 4616, 1052, 1296, 3757, 1950, 1102, 4609],
[3671, 2489, 4033, 4381, 2252, 4318, 3086],
[2229, 4616, 2728, 4858, 4807, 1637, 301],
[4687, 3291, 2229, 4616, 4858, 2748, 1935, 2806, 4033, 4214, 3897, 294, 355],
[3655, 262, 3361, 508, 3655, 4314, 4663, 2163, 3617, 4296, 860],
[1124, 3153, 942, 4307],
[1028, 3625, 3104, 4647, 1288, 3603, 3603, 702, 4407, 3063],
[3634, 3218, 477],
[3170, 3993, 1295, 1004, 169, 1917, 3907, 3845, 2094],
[2296, 2352, 1851, 3676, 3048, 328, 707, 533, 1464, 4909, 3016, 2728],
[404, 4724, 3567, 4848, 3151, 228, 1552, 3678, 570, 3044, 1371, 2855, 3255],
[1837, 4909, 2484, 683, 1762, 2800],
[2229, 4301, 697, 4616, 133, 1184],
[4685, 4391, 1272, 3590, 4033, 2776, 517],
[697, 3509, 2642, 2144, 3743, 4745],
[3104, 3509, 1288, 3524, 4571],
[1837, 4568, 472, 4292, 4282, 175, 2222, 4874],
[4865, 1285, 3517, 1675, 4983, 428, 559, 1572, 1655],
[2064, 683, 659, 4616, 4235, 4547, 1039, 263, 4361],
[2445, 742, 4601, 4033, 1724, 2620, 4616, 244, 2182],
[4598, 2581, 3158, 3361, 1982, 3270, 2603, 4299, 98],
[4807, 4318, 1024, 23, 2355, 91],
[2740, 3887, 3061, 3372, 974, 1124, 2285, 170, 3907, 2029],
[432, 3206, 3509, 1248, 3420, 3791, 4033, 2011, 4744, 4043, 2594, 2318, 3272],
[3958, 757, 1272, 3627, 388, 1697, 1088, 449, 4033, 1878, 2437, 178, 2406],
[696, 2051, 1747, 429, 2214],
[2087, 3975, 3525, 3943, 2637, 1722, 908, 2419, 3063],
[4340, 4868, 2822, 3297, 517, 3450],
[4262, 2617, 2333, 3748, 823, 450, 4226, 1832],
[4113, 276, 2355, 2134],
[2438, 3151, 2282, 4052, 974],
[3180, 4887, 4196, 328, 1977],
...]
```

▼ Embedding Representation

```
sent_length=20
embedded_docs=pad_sequences(onehot_repr,padding='pre',maxlen=sent_length)
print(embedded_docs)

[[ 0  0  0 ... 2931  697 4204]
 [ 0  0  0 ... 4616 2548 3063]
 [ 0  0  0 ... 2293  857 4225]
 ...
 [ 0  0  0 ... 3648 4577 1283]
 [ 0  0  0 ... 4837 1328 3168]
 [ 0  0  0 ...  974 4616 3182]]

embedded_docs[0]

array([ 0,  0,  0,  0,  0,  0,  0,  0,  0,  0,
        0,  0,  0,  0,  0,  0, 2931,  697, 4204], dtype=int32)
```

```
## Creating model
embedding_vector_features=40
model=Sequential()
model.add(Embedding(voc_size,embedding_vector_features,input_length=sent_length))
model.add(LSTM(100))
model.add(Dense(1,activation='sigmoid'))
model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
print(model.summary())
```

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 20, 40)	200000
lstm (LSTM)	(None, 100)	56400
dense (Dense)	(None, 1)	101

=====
Total params: 256501 (1001.96 KB)
Trainable params: 256501 (1001.96 KB)
Non-trainable params: 0 (0.00 Byte)
=====

None

```
len(embedded_docs),y.shape

(6335, (6335,))
```

```
import numpy as np
X_final=np.array(embedded_docs)
y_final=np.array(y)
```

```
X_final.shape,y_final.shape

((6335, 20), (6335,))
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X_final, y_final, test_size=0.33, random_state=42)
```

Model Training

```
### Finally Training
model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=15, batch_size=64)
```

Epoch 1/15

```
-----
UnimplementedError                                Traceback (most recent call last)
```

```
<ipython-input-30-fb2f3329c019> in <cell line: 2>()
```

```
1 ### Finally Training
```

```
----> 2 model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=15, batch_size=64)
```

```
1 frames
```

```
/usr/local/lib/python3.10/dist-packages/tensorflow/python/eager/execute.py in quick_execute(op_name, num_outputs, inputs, attrs,
ctx, name)
```

```
51     try:
52         ctx.ensure_initialized()
--> 53         tensors = pywrap_tfe.TFE_Py_Execute(ctx._handle, device_name, op_name,
54                                             inputs, attrs, num_outputs)
55     except core._NotOkStatusException as e:
```

```
UnimplementedError: Graph execution error:
```

```
Detected at node binary_crossentropy/Cast defined at (most recent call last):
```

```
File "/usr/lib/python3.10/runpy.py", line 196, in _run_module_as_main
```

```
File "/usr/lib/python3.10/runpy.py", line 86, in _run_code
```

```
File "/usr/local/lib/python3.10/dist-packages/colab_kernel_launcher.py", line 37, in <module>
```

```
File "/usr/local/lib/python3.10/dist-packages/traitlets/config/application.py", line 992, in launch_instance
```

```
File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelapp.py", line 619, in start
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/platform/asyncio.py", line 195, in start
```

```
File "/usr/lib/python3.10/asyncio/base_events.py", line 603, in run_forever
```

```
File "/usr/lib/python3.10/asyncio/base_events.py", line 1909, in _run_once
```

```
File "/usr/lib/python3.10/asyncio/events.py", line 80, in _run
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/ioloop.py", line 685, in <lambda>
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/ioloop.py", line 738, in _run_callback
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/gen.py", line 825, in inner
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/gen.py", line 786, in run
```

```
File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelbase.py", line 361, in process_one
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/gen.py", line 234, in wrapper
```

```
File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelbase.py", line 261, in dispatch_shell
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/gen.py", line 234, in wrapper
```

```
File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelbase.py", line 539, in execute_request
```

```
File "/usr/local/lib/python3.10/dist-packages/tornado/gen.py", line 234, in wrapper
```

✓ Adding Dropout

```
from tensorflow.keras.layers import Dropout
## Creating model
embedding_vector_features=40
model=Sequential()
model.add(Embedding(voc_size,embedding_vector_features,input_length=sent_length))
model.add(Dropout(0.3))
model.add(LSTM(100))
model.add(Dropout(0.3))
model.add(Dense(1,activation='sigmoid'))
model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
```

```
File "/usr/lib/python3.10/runpy.py", line 196, in _run_module_as_main
```

✓ Performance Metrics And Accuracy

```
File "/usr/local/lib/python3.10/dist-packages/keras/engine/training.py", line 1807, in fit
y_pred=model.predict_classes(X_test)
```

```
-----  
AttributeError                                Traceback (most recent call last)  
<ipython-input-32-66f7fe571d01> in <cell line: 1>()  
----> 1 y_pred=model.predict_classes(X_test)  
  
AttributeError: 'Sequential' object has no attribute 'predict_classes'
```

```
from sklearn.metrics import confusion_matrix
```

```
File "/usr/local/lib/python3.10/dist-packages/keras/src/losses.py", line 2521, in binary_crossentropy  
confusion_matrix(y_test,y_pred)
```

```
-----  
NameError                                Traceback (most recent call last)  
<ipython-input-41-e5b86ff07486> in <cell line: 1>()
```