

```
!pip install --upgrade tensorflow
```

```
Requirement already satisfied: tensorflow in  
/usr/local/lib/python3.10/dist-packages (2.13.0)
```

```
Collecting tensorflow
```

```
  Downloading tensorflow-2.14.0-cp310-cp310-  
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (489.8 MB)  
_____ 489.8/489.8 MB 2.8 MB/s eta
```

```
0:00:00
```

```
Requirement already satisfied: absl-py>=1.0.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.4.0)  
Requirement already satisfied: astunparse>=1.6.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.6.3)  
Requirement already satisfied: flatbuffers>=23.5.26 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (23.5.26)  
Requirement already satisfied: gast!=0.5.0,!0.5.1,!0.5.2,>=0.2.1  
in /usr/local/lib/python3.10/dist-packages (from tensorflow) (0.4.0)  
Requirement already satisfied: google-pasta>=0.1.1 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (0.2.0)  
Requirement already satisfied: h5py>=2.9.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.9.0)  
Requirement already satisfied: libclang>=13.0.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (16.0.6)  
Collecting ml-dtypes==0.2.0 (from tensorflow)
```

```
  Downloading ml_dtypes-0.2.0-cp310-cp310-  
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.0 MB)  
_____ 1.0/1.0 MB 62.3 MB/s eta
```

```
0:00:00
```

```
Requirement already satisfied: numpy>=1.23.5 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.23.5)  
Requirement already satisfied: opt-einsum>=2.3.2 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.3.0)  
Requirement already satisfied: packaging in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (23.2)  
Requirement already satisfied: protobuf!=4.21.0,!4.21.1,!4.21.2,!  
=4.21.3,!4.21.4,!4.21.5,<5.0.0dev,>=3.20.3 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.20.3)  
Requirement already satisfied: setuptools in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (67.7.2)  
Requirement already satisfied: six>=1.12.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.16.0)  
Requirement already satisfied: termcolor>=1.1.0 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (2.3.0)  
Requirement already satisfied: typing-extensions>=3.6.6 in  
/usr/local/lib/python3.10/dist-packages (from tensorflow) (4.5.0)  
Collecting wrapt<1.15,>=1.11.0 (from tensorflow)
```

```
  Downloading wrapt-1.14.1-cp310-cp310-  
manylinux_2_5_x86_64.manylinux1_x86_64.manylinux_2_17_x86_64.manylinux  
2014_x86_64.whl (77 kB)
```

```
_____ 77.9/77.9 kB 10.6 MB/s eta
```

```
0:00:00
ent already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (0.34.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.59.0)
Collecting tensorboard<2.15,>=2.14 (from tensorflow)
  Downloading tensorboard-2.14.1-py3-none-any.whl (5.5 MB)
  _____ 5.5/5.5 MB 126.3 MB/s eta
0:00:00
ator<2.15,>=2.14.0 (from tensorflow)
  Downloading tensorflow_estimator-2.14.0-py2.py3-none-any.whl (440
kB)
  _____ 440.7/440.7 kB 44.1 MB/s eta
0:00:00
tensorflow)
  Downloading keras-2.14.0-py3-none-any.whl (1.7 MB)
  _____ 1.7/1.7 MB 95.4 MB/s eta
0:00:00
ent already satisfied: wheel<1.0,>=0.23.0 in
/usr/local/lib/python3.10/dist-packages (from astunparse>=1.6.0-
>tensorflow) (0.41.2)
Requirement already satisfied: google-auth<3,>=1.6.3 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.15,>=2.14-
>tensorflow) (2.17.3)
Requirement already satisfied: google-auth-oauthlib<1.1,>=0.5 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.15,>=2.14-
>tensorflow) (1.0.0)
Requirement already satisfied: markdown>=2.6.8 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.15,>=2.14-
>tensorflow) (3.5)
Requirement already satisfied: requests<3,>=2.21.0 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.15,>=2.14-
>tensorflow) (2.31.0)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0
in /usr/local/lib/python3.10/dist-packages (from
tensorboard<2.15,>=2.14->tensorflow) (0.7.1)
Requirement already satisfied: werkzeug>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.15,>=2.14-
>tensorflow) (3.0.0)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.15,>=2.14->tensorflow) (5.3.1)
Requirement already satisfied: pyasn1-modules>=0.2.1 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.15,>=2.14->tensorflow) (0.3.0)
Requirement already satisfied: rsa<5,>=3.1.4 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.15,>=2.14->tensorflow) (4.9)
Requirement already satisfied: requests-oauthlib>=0.7.0 in
```

```

/usr/local/lib/python3.10/dist-packages (from google-auth-
oauthlib<1.1,>=0.5->tensorboard<2.15,>=2.14->tensorflow) (1.3.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.15,>=2.14->tensorflow) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.15,>=2.14->tensorflow) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.15,>=2.14->tensorflow) (2.0.6)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.15,>=2.14->tensorflow) (2023.7.22)
Requirement already satisfied: MarkupSafe>=2.1.1 in
/usr/local/lib/python3.10/dist-packages (from werkzeug>=1.0.1-
>tensorboard<2.15,>=2.14->tensorflow) (2.1.3)
Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in
/usr/local/lib/python3.10/dist-packages (from pyasn1-modules>=0.2.1-
>google-auth<3,>=1.6.3->tensorboard<2.15,>=2.14->tensorflow) (0.5.0)
Requirement already satisfied: oauthlib>=3.0.0 in
/usr/local/lib/python3.10/dist-packages (from requests-
oauthlib>=0.7.0->google-auth-oauthlib<1.1,>=0.5-
>tensorboard<2.15,>=2.14->tensorflow) (3.2.2)
Installing collected packages: wrapt, tensorflow-estimator, ml-dtypes,
keras, tensorboard, tensorflow
  Attempting uninstall: wrapt
    Found existing installation: wrapt 1.15.0
    Uninstalling wrapt-1.15.0:
      Successfully uninstalled wrapt-1.15.0
  Attempting uninstall: tensorflow-estimator
    Found existing installation: tensorflow-estimator 2.13.0
    Uninstalling tensorflow-estimator-2.13.0:
      Successfully uninstalled tensorflow-estimator-2.13.0
  Attempting uninstall: ml-dtypes
    Found existing installation: ml-dtypes 0.3.1
    Uninstalling ml-dtypes-0.3.1:
      Successfully uninstalled ml-dtypes-0.3.1
  Attempting uninstall: keras
    Found existing installation: keras 2.13.1
    Uninstalling keras-2.13.1:
      Successfully uninstalled keras-2.13.1
  Attempting uninstall: tensorboard
    Found existing installation: tensorboard 2.13.0
    Uninstalling tensorboard-2.13.0:
      Successfully uninstalled tensorboard-2.13.0
  Attempting uninstall: tensorflow
    Found existing installation: tensorflow 2.13.0
    Uninstalling tensorflow-2.13.0:

```

```
Successfully uninstalled tensorflow-2.13.0
Successfully installed keras-2.14.0 ml-dtypes-0.2.0 tensorboard-2.14.1
tensorflow-2.14.0 tensorflow-estimator-2.14.0 wrapt-1.14.1
```

```
!pip install plotly
!pip install --upgrade nbformat
!pip install nltk
!pip install spacy # spaCy is an open-source software library for
advanced natural language processing
!pip install WordCloud
!pip install gensim # Gensim is an open-source library for
unsupervised topic modeling and natural language processing
!pip install jupyterthemes
import nltk
nltk.download('punkt')

import tensorflow as tf
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud, STOPWORDS
import nltk
import re
from nltk.stem import PorterStemmer, WordNetLemmatizer
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize, sent_tokenize
import gensim
from gensim.utils import simple_preprocess
from gensim.parsing.preprocessing import STOPWORDS
# import keras
from tensorflow.keras.preprocessing.text import one_hot, Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten, Embedding, Input,
LSTM, Conv1D, MaxPool1D, Bidirectional
from tensorflow.keras.models import Model
```

```
Requirement already satisfied: plotly in
/usr/local/lib/python3.10/dist-packages (5.15.0)
Requirement already satisfied: tenacity>=6.2.0 in
/usr/local/lib/python3.10/dist-packages (from plotly) (8.2.3)
Requirement already satisfied: packaging in
/usr/local/lib/python3.10/dist-packages (from plotly) (23.2)
Requirement already satisfied: nbformat in
/usr/local/lib/python3.10/dist-packages (5.9.2)
Requirement already satisfied: fastjsonschema in
/usr/local/lib/python3.10/dist-packages (from nbformat) (2.18.1)
Requirement already satisfied: jsonschema>=2.6 in
/usr/local/lib/python3.10/dist-packages (from nbformat) (4.19.1)
```

Requirement already satisfied: jupyter-core in
/usr/local/lib/python3.10/dist-packages (from nbformat) (5.4.0)
Requirement already satisfied: traitlets>=5.1 in
/usr/local/lib/python3.10/dist-packages (from nbformat) (5.7.1)
Requirement already satisfied: attrs>=22.2.0 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat) (23.1.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat) (2023.7.1)
Requirement already satisfied: referencing>=0.28.4 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat) (0.30.2)
Requirement already satisfied: rpds-py>=0.7.1 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat) (0.10.6)
Requirement already satisfied: platformdirs>=2.5 in
/usr/local/lib/python3.10/dist-packages (from jupyter-core->nbformat)
(3.11.0)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-
packages (3.8.1)
Requirement already satisfied: click in
/usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in
/usr/local/lib/python3.10/dist-packages (from nltk) (1.3.2)
Requirement already satisfied: regex>=2021.8.3 in
/usr/local/lib/python3.10/dist-packages (from nltk) (2023.6.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-
packages (from nltk) (4.66.1)
Requirement already satisfied: spacy in
/usr/local/lib/python3.10/dist-packages (3.6.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in
/usr/local/lib/python3.10/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.10/dist-packages (from spacy) (2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.10/dist-packages (from spacy) (3.0.9)
Requirement already satisfied: thinc<8.2.0,>=8.1.8 in
/usr/local/lib/python3.10/dist-packages (from spacy) (8.1.12)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in
/usr/local/lib/python3.10/dist-packages (from spacy) (1.1.2)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in
/usr/local/lib/python3.10/dist-packages (from spacy) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in
/usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)

Requirement already satisfied: typer<0.10.0,>=0.3.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (0.9.0)
Requirement already satisfied: pathy>=0.10.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (0.10.2)
Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in
/usr/local/lib/python3.10/dist-packages (from spacy) (6.4.0)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (4.66.1)
Requirement already satisfied: numpy>=1.15.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (1.23.5)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (2.31.0)
Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in
/usr/local/lib/python3.10/dist-packages (from spacy) (1.10.13)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.10/dist-packages (from spacy) (3.1.2)
Requirement already satisfied: setuptools in
/usr/local/lib/python3.10/dist-packages (from spacy) (67.7.2)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (23.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in
/usr/local/lib/python3.10/dist-packages (from spacy) (3.3.0)
Requirement already satisfied: typing-extensions>=4.2.0 in
/usr/local/lib/python3.10/dist-packages (from pydantic!=1.8,!
=1.8.1,<3.0.0,>=1.7.4->spacy) (4.5.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0-
>spacy) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0-
>spacy) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0-
>spacy) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0-
>spacy) (2023.7.22)
Requirement already satisfied: blis<0.8.0,>=0.7.8 in
/usr/local/lib/python3.10/dist-packages (from thinc<8.2.0,>=8.1.8-
>spacy) (0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in
/usr/local/lib/python3.10/dist-packages (from thinc<8.2.0,>=8.1.8-
>spacy) (0.1.3)
Requirement already satisfied: click<9.0.0,>=7.1.1 in
/usr/local/lib/python3.10/dist-packages (from typer<0.10.0,>=0.3.0-
>spacy) (8.1.7)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.10/dist-packages (from jinja2->spacy) (2.1.3)
Requirement already satisfied: WordCloud in

```
/usr/local/lib/python3.10/dist-packages (1.9.2)
Requirement already satisfied: numpy>=1.6.1 in
/usr/local/lib/python3.10/dist-packages (from WordCloud) (1.23.5)
Requirement already satisfied: pillow in
/usr/local/lib/python3.10/dist-packages (from WordCloud) (9.4.0)
Requirement already satisfied: matplotlib in
/usr/local/lib/python3.10/dist-packages (from WordCloud) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(1.1.1)
Requirement already satisfied: cycler>=0.10 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(4.43.1)
Requirement already satisfied: kiwisolver>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(1.4.5)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(23.2)
Requirement already satisfied: pyparsing>=2.3.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.10/dist-packages (from matplotlib->WordCloud)
(2.8.2)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7-
>matplotlib->WordCloud) (1.16.0)
Requirement already satisfied: gensim in
/usr/local/lib/python3.10/dist-packages (4.3.2)
Requirement already satisfied: numpy>=1.18.5 in
/usr/local/lib/python3.10/dist-packages (from gensim) (1.23.5)
Requirement already satisfied: scipy>=1.7.0 in
/usr/local/lib/python3.10/dist-packages (from gensim) (1.11.3)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.10/dist-packages (from gensim) (6.4.0)
Collecting jupyterthemes
  Downloading jupyterthemes-0.20.0-py2.py3-none-any.whl (7.0 MB)
  7.0/7.0 MB 16.9 MB/s eta
0:00:00
Requirement already satisfied: jupyter-core in /usr/local/lib/python3.10/dist-
packages (from jupyterthemes) (5.4.0)
Requirement already satisfied: notebook>=5.6.0 in
/usr/local/lib/python3.10/dist-packages (from jupyterthemes) (6.5.5)
Requirement already satisfied: ipython>=5.4.1 in
/usr/local/lib/python3.10/dist-packages (from jupyterthemes) (7.34.0)
```

Requirement already satisfied: matplotlib>=1.4.3 in
/usr/local/lib/python3.10/dist-packages (from jupyterthemes) (3.7.1)
Collecting lesscpy>=0.11.2 (from jupyterthemes)

Downloading lesscpy-0.15.1-py2.py3-none-any.whl (46 kB)
46.7/46.7 kB 5.7 MB/s eta

0:00:00

Requirement already satisfied: setuptools>=18.5 in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (67.7.2)

Collecting jedi>=0.16 (from ipython>=5.4.1->jupyterthemes)

Downloading jedi-0.19.1-py2.py3-none-any.whl (1.6 MB)
1.6/1.6 MB 29.5 MB/s eta

0:00:00

Requirement already satisfied: decorator in /usr/local/lib/python3.10/dist-
packages (from ipython>=5.4.1->jupyterthemes) (4.4.2)

Requirement already satisfied: pickleshare in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (0.7.5)

Requirement already satisfied: traitlets>=4.2 in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (5.7.1)

Requirement already satisfied: prompt-toolkit!=3.0.0,!
=3.0.1,<3.1.0,>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from
ipython>=5.4.1->jupyterthemes) (3.0.39)

Requirement already satisfied: pygments in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (2.16.1)

Requirement already satisfied: backcall in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (0.2.0)

Requirement already satisfied: matplotlib-inline in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (0.1.6)

Requirement already satisfied: pexpect>4.3 in
/usr/local/lib/python3.10/dist-packages (from ipython>=5.4.1-
>jupyterthemes) (4.8.0)

Collecting ply (from lesscpy>=0.11.2->jupyterthemes)

Downloading ply-3.11-py2.py3-none-any.whl (49 kB)
49.6/49.6 kB 6.3 MB/s eta

0:00:00

Requirement already satisfied: contourpy>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (1.1.1)

Requirement already satisfied: cycler>=0.10 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (0.12.1)

Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (4.43.1)

Requirement already satisfied: kiwisolver>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (1.4.5)

Requirement already satisfied: numpy>=1.20 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (1.23.5)

Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (23.2)

Requirement already satisfied: pillow>=6.2.0 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (9.4.0)

Requirement already satisfied: pyparsing>=2.3.1 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (3.1.1)

Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.10/dist-packages (from matplotlib>=1.4.3-
>jupyterthemes) (2.8.2)

Requirement already satisfied: jinja2 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (3.1.2)

Requirement already satisfied: tornado>=6.1 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (6.3.2)

Requirement already satisfied: pyzmq<25,>=17 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (23.2.1)

Requirement already satisfied: argon2-cffi in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (23.1.0)

Requirement already satisfied: jupyter-client<8,>=5.3.4 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (6.1.12)

Requirement already satisfied: ipython-genutils in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (0.2.0)

Requirement already satisfied: nbformat in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (5.9.2)

Requirement already satisfied: nbconvert>=5 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (6.5.4)

Requirement already satisfied: nest-asyncio>=1.5 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (1.5.8)

Requirement already satisfied: ipykernel in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (5.5.6)

Requirement already satisfied: Send2Trash>=1.8.0 in

```
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (1.8.2)
Requirement already satisfied: terminado>=0.8.3 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (0.17.1)
Requirement already satisfied: prometheus-client in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (0.17.1)
Requirement already satisfied: nbclassic>=0.4.7 in
/usr/local/lib/python3.10/dist-packages (from notebook>=5.6.0-
>jupyterthemes) (1.0.0)
Requirement already satisfied: platformdirs>=2.5 in
/usr/local/lib/python3.10/dist-packages (from jupyter-core-
>jupyterthemes) (3.11.0)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in
/usr/local/lib/python3.10/dist-packages (from jedi>=0.16-
>ipython>=5.4.1->jupyterthemes) (0.8.3)
Requirement already satisfied: jupyter-server>=1.8 in
/usr/local/lib/python3.10/dist-packages (from nbclassic>=0.4.7-
>notebook>=5.6.0->jupyterthemes) (1.24.0)
Requirement already satisfied: notebook-shim>=0.2.3 in
/usr/local/lib/python3.10/dist-packages (from nbclassic>=0.4.7-
>notebook>=5.6.0->jupyterthemes) (0.2.3)
Requirement already satisfied: lxml in /usr/local/lib/python3.10/dist-
packages (from nbconvert>=5->notebook>=5.6.0->jupyterthemes) (4.9.3)
Requirement already satisfied: beautifulsoup4 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (4.11.2)
Requirement already satisfied: bleach in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (6.1.0)
Requirement already satisfied: defusedxml in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.7.1)
Requirement already satisfied: entrypoints>=0.2.2 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.4)
Requirement already satisfied: jupyterlab-pygments in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.2.2)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (2.1.3)
Requirement already satisfied: mistune<2,>=0.8.1 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.8.4)
Requirement already satisfied: nbclient>=0.5.0 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.8.0)
```

Requirement already satisfied: pandocfilters>=1.4.1 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (1.5.0)

Requirement already satisfied: tinycss2 in
/usr/local/lib/python3.10/dist-packages (from nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (1.2.1)

Requirement already satisfied: fastjsonschema in
/usr/local/lib/python3.10/dist-packages (from nbformat-
>notebook>=5.6.0->jupyterthemes) (2.18.1)

Requirement already satisfied: jsonschema>=2.6 in
/usr/local/lib/python3.10/dist-packages (from nbformat-
>notebook>=5.6.0->jupyterthemes) (4.19.1)

Requirement already satisfied: ptyprocess>=0.5 in
/usr/local/lib/python3.10/dist-packages (from pexpect>4.3-
>ipython>=5.4.1->jupyterthemes) (0.7.0)

Requirement already satisfied: wcwidth in
/usr/local/lib/python3.10/dist-packages (from prompt-toolkit!=3.0.0,!
=3.0.1,<3.1.0,>=2.0.0->ipython>=5.4.1->jupyterthemes) (0.2.8)

Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7-
>matplotlib>=1.4.3->jupyterthemes) (1.16.0)

Requirement already satisfied: argon2-cffi-bindings in
/usr/local/lib/python3.10/dist-packages (from argon2-cffi-
>notebook>=5.6.0->jupyterthemes) (21.2.0)

Requirement already satisfied: attrs>=22.2.0 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat->notebook>=5.6.0->jupyterthemes) (23.1.0)

Requirement already satisfied: jsonschema-specifications>=2023.03.6 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat->notebook>=5.6.0->jupyterthemes) (2023.7.1)

Requirement already satisfied: referencing>=0.28.4 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat->notebook>=5.6.0->jupyterthemes) (0.30.2)

Requirement already satisfied: rpds-py>=0.7.1 in
/usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6-
>nbformat->notebook>=5.6.0->jupyterthemes) (0.10.6)

Requirement already satisfied: anyio<4,>=3.1.0 in
/usr/local/lib/python3.10/dist-packages (from jupyter-server>=1.8-
>nbclassic>=0.4.7->notebook>=5.6.0->jupyterthemes) (3.7.1)

Requirement already satisfied: websocket-client in
/usr/local/lib/python3.10/dist-packages (from jupyter-server>=1.8-
>nbclassic>=0.4.7->notebook>=5.6.0->jupyterthemes) (1.6.4)

Requirement already satisfied: cffi>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from argon2-cffi-bindings-
>argon2-cffi->notebook>=5.6.0->jupyterthemes) (1.16.0)

Requirement already satisfied: soupsieve>1.2 in
/usr/local/lib/python3.10/dist-packages (from beautifulsoup4-
>nbconvert>=5->notebook>=5.6.0->jupyterthemes) (2.5)

Requirement already satisfied: webencodings in

```

/usr/local/lib/python3.10/dist-packages (from bleach->nbconvert>=5-
>notebook>=5.6.0->jupyterthemes) (0.5.1)
Requirement already satisfied: idna>=2.8 in
/usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0-
>jupyter-server>=1.8->nbclassic>=0.4.7->notebook>=5.6.0-
>jupyterthemes) (3.4)
Requirement already satisfied: sniffio>=1.1 in
/usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0-
>jupyter-server>=1.8->nbclassic>=0.4.7->notebook>=5.6.0-
>jupyterthemes) (1.3.0)
Requirement already satisfied: exceptiongroup in
/usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0-
>jupyter-server>=1.8->nbclassic>=0.4.7->notebook>=5.6.0-
>jupyterthemes) (1.1.3)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.10/dist-packages (from cffi>=1.0.1->argon2-
cffi-bindings->argon2-cffi->notebook>=5.6.0->jupyterthemes) (2.21)
Installing collected packages: ply, lesscpy, jedi, jupyterthemes
Successfully installed jedi-0.19.1 jupyterthemes-0.20.0 lesscpy-0.15.1
ply-3.11

```

```

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.

```

```

df_fake = pd.read_csv('/usr/Fake.csv')
df_true = pd.read_csv('/usr/True.csv')
df_true

```

		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...		
...			
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
\			
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
...
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
0      December 31, 2017
1      December 29, 2017
2      December 31, 2017
3      December 30, 2017
4      December 29, 2017
...
21412    August 22, 2017
21413    August 22, 2017
21414    August 22, 2017
21415    August 22, 2017
21416    August 22, 2017

```

```
[21417 rows x 4 columns]
```

```
df_true.isnull().sum()
```

```

title      0
text       0
subject    0
date       0
dtype: int64

```

```
df_fake.isnull().sum()
```

```

title      0
text       0
subject    0
date       0
dtype: int64

```

```
df_true.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21417 entries, 0 to 21416
Data columns (total 4 columns):

```

#	Column	Non-Null Count	Dtype
0	title	21417 non-null	object
1	text	21417 non-null	object
2	subject	21417 non-null	object
3	date	21417 non-null	object

dtypes: object(4)

memory usage: 669.4+ KB

df_fake.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 23481 entries, 0 to 23480

Data columns (total 4 columns):

#	Column	Non-Null Count	Dtype
0	title	23481 non-null	object
1	text	23481 non-null	object
2	subject	23481 non-null	object
3	date	23481 non-null	object

dtypes: object(4)

memory usage: 733.9+ KB

#add a target class column to indicate a news is true or fake

df_true['isfake'] = 1

df_true.head()

df_fake['isfake'] = 0

df_fake.head()

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

```
#add real and fake news
```

```
df = pd.concat([df_true, df_fake]).reset_index(drop = True)
df
```

	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...
...	...
44893	McPain: John McCain Furious That Iran Treated ...
44894	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...
44895	Sunnistan: US and Allied 'Safe Zone' Plan to T...
44896	How to Blow \$700 Million: Al Jazeera America F...
44897	10 U.S. Navy Sailors Held by Iranian Military ...

	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
...
44893	21st Century Wire says As 21WIRE reported earl...	Middle-east
44894	21st Century Wire says It s a familiar theme. ...	Middle-east
44895	Patrick Henningsen 21st Century WireRemember ...	Middle-east
44896	21st Century Wire says Al Jazeera America will...	Middle-east
44897	21st Century Wire says As 21WIRE predicted in ...	Middle-east

	date	isfake
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
...
44893	January 16, 2016	0
44894	January 16, 2016	0

```
44895    January 15, 2016    0
44896    January 14, 2016    0
44897    January 12, 2016    0
```

```
[44898 rows x 5 columns]
```

```
# drop unnecessary Date column # RUN ONLY ONCE
df.drop(columns = ['date'], inplace = True)
# combine title and text together
df['original'] = df['title'] + ' ' + df['text']
df.head()
```

```

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

```

                                     original
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...
```

```
df['original'][0]
```

```
{"type": "string"}
```

```
#DATA CLEANING
```

```
# download stopwords
```

```
nlTK.download("stopwords")
```

```
[nlTK_data] Downloading package stopwords to /root/nltk_data...
```

```
[nlTK_data] Package stopwords is already up-to-date!
```

```
True
```



```
# Obtain additional stopwords from nltk
from nltk.corpus import stopwords
stop_words = stopwords.words('english')
stop_words.extend(['from', 'subject', 're', 'edu', 'use']) # add
additional stop words to remove
```

```
stop_words
```

```
['i',
 'me',
 'my',
 'myself',
 'we',
 'our',
 'ours',
 'ourselves',
 'you',
 "you're",
 "you've",
 "you'll",
 "you'd",
 'your',
 'yours',
 'yourself',
 'yourselves',
 'he',
 'him',
 'his',
 'himself',
 'she',
 "she's",
 'her',
 'hers',
 'herself',
 'it',
 "it's",
 'its',
 'itself',
 'they',
 'them',
 'their',
 'theirs',
 'themselves',
 'what',
 'which',
 'who',
 'whom',
 'this',
 'that',
 "that'll",
```

'these',
'those',
'am',
'is',
'are',
'was',
'were',
'be',
'been',
'being',
'have',
'has',
'had',
'having',
'do',
'does',
'did',
'doing',
'a',
'an',
'the',
'and',
'but',
'if',
'or',
'because',
'as',
'until',
'while',
'of',
'at',
'by',
'for',
'with',
'about',
'against',
'between',
'into',
'through',
'during',
'before',
'after',
'above',
'below',
'to',
'from',
'up',
'down',
'in',

'out',
'on',
'off',
'over',
'under',
'again',
'further',
'then',
'once',
'here',
'there',
'when',
'where',
'why',
'how',
'all',
'any',
'both',
'each',
'few',
'more',
'most',
'other',
'some',
'such',
'no',
'nor',
'not',
'only',
'own',
'same',
'so',
'than',
'too',
'very',
's',
't',
'can',
'will',
'just',
'don',
"don't",
'should',
"should've",
'now',
'd',
'll',
'm',
'o',

```
're',  
've',  
'y',  
'ain',  
'aren',  
"aren't",  
'couldn',  
"couldn't",  
'didn',  
"didn't",  
'doesn',  
"doesn't",  
'hadn',  
"hadn't",  
'hasn',  
"hasn't",  
'haven',  
"haven't",  
'isn',  
"isn't",  
'ma',  
'mightn',  
"mightn't",  
'mustn',  
"mustn't",  
'needn',  
"needn't",  
'shan',  
"shan't",  
'shouldn',  
"shouldn't",  
'wasn',  
"wasn't",  
'weren',  
"weren't",  
'won',  
"won't",  
'wouldn',  
"wouldn't",  
'from',  
'subject',  
're',  
'edu',  
'use']
```

Remove stopwords and remove words with 2 or less characters using gensim

```
def preprocess(text):  
    result = []  
    for token in gensim.utils.simple_preprocess(text):
```

```

        if token not in gensim.parsing.preprocessing.STOPWORDS and
len(token) > 3 and token not in stop_words:
            result.append(token)

    return result

# Apply the function to the dataframe
df['clean'] = df['original'].apply(preprocess)

# Show original news
df['original'][0]

{"type": "string"}

# Show cleaned up news after removing stopwords
print(df['clean'][0])

['budget', 'fight', 'looms', 'republicans', 'flip', 'fiscal',
'script', 'washington', 'reuters', 'head', 'conservative',
'republican', 'faction', 'congress', 'voted', 'month', 'huge',
'expansion', 'national', 'debt', 'cuts', 'called', 'fiscal',
'conservative', 'sunday', 'urged', 'budget', 'restraint', 'keeping',
'sharp', 'pivot', 'republicans', 'representative', 'mark', 'meadows',
'speaking', 'face', 'nation', 'drew', 'hard', 'line', 'federal',
'spending', 'lawmakers', 'bracing', 'battle', 'january', 'return',
'holidays', 'wednesday', 'lawmakers', 'begin', 'trying', 'pass',
'federal', 'budget', 'fight', 'likely', 'linked', 'issues',
'immigration', 'policy', 'november', 'congressional', 'election',
'campaigns', 'approach', 'republicans', 'seek', 'control', 'congress',
'president', 'donald', 'trump', 'republicans', 'want', 'budget',
'increase', 'military', 'spending', 'democrats', 'want',
'proportional', 'increases', 'defense', 'discretionary', 'spending',
'programs', 'support', 'education', 'scientific', 'research',
'infrastructure', 'public', 'health', 'environmental', 'protection',
'trump', 'administration', 'willing', 'going', 'increase', 'defense',
'discretionary', 'spending', 'percent', 'meadows', 'chairman',
'small', 'influential', 'house', 'freedom', 'caucus', 'said',
'program', 'democrats', 'saying', 'need', 'government', 'raise',
'percent', 'fiscal', 'conservative', 'rationale', 'eventually',
'people', 'money', 'said', 'meadows', 'republicans', 'voted', 'late',
'december', 'party', 'debt', 'financed', 'overhaul', 'expected',
'balloon', 'federal', 'budget', 'deficit', 'trillion', 'years',
'trillion', 'national', 'debt', 'interesting', 'hear', 'mark', 'talk',
'fiscal', 'responsibility', 'democratic', 'representative', 'joseph',
'crowley', 'said', 'crowley', 'said', 'republican', 'require',
'united', 'states', 'borrow', 'trillion', 'paid', 'future',
'generations', 'finance', 'cuts', 'corporations', 'rich', 'fiscally',
'responsible', 'bills', 'seen', 'passed', 'history', 'house',
'representatives', 'think', 'going', 'paying', 'years', 'come',
'crowley', 'said', 'republicans', 'insist', 'package', 'biggest',

```

```
'overhaul', 'years', 'boost', 'economy', 'growth', 'house', 'speaker',  
'paul', 'ryan', 'supported', 'recently', 'went', 'meadows', 'making',  
'clear', 'radio', 'interview', 'welfare', 'entitlement', 'reform',  
'party', 'calls', 'republican', 'priority', 'republican', 'parlance',  
'entitlement', 'programs', 'mean', 'food', 'stamps', 'housing',  
'assistance', 'medicare', 'medicaid', 'health', 'insurance',  
'elderly', 'poor', 'disabled', 'programs', 'created', 'washington',  
'assist', 'needy', 'democrats', 'seized', 'ryan', 'early', 'december',  
'remarks', 'saying', 'showed', 'republicans', 'overhaul', 'seeking',  
'spending', 'cuts', 'social', 'programs', 'goals', 'house',  
'republicans', 'seat', 'senate', 'votes', 'democrats', 'needed',  
'approve', 'budget', 'prevent', 'government', 'shutdown', 'democrats',  
'leverage', 'senate', 'republicans', 'narrowly', 'control', 'defend',  
'discretionary', 'defense', 'programs', 'social', 'spending',  
'tackling', 'issue', 'dreamers', 'people', 'brought', 'illegally',  
'country', 'children', 'trump', 'september', 'march', 'expiration',  
'date', 'deferred', 'action', 'childhood', 'arrivals', 'daca',  
'program', 'protects', 'young', 'immigrants', 'deportation',  
'provides', 'work', 'permits', 'president', 'said', 'recent',  
'twitter', 'messages', 'wants', 'funding', 'proposed', 'mexican',  
'border', 'wall', 'immigration', 'changes', 'exchange', 'agreeing',  
'help', 'dreamers', 'representative', 'debbie', 'dingell', 'told',  
'favor', 'linking', 'issue', 'policy', 'objectives', 'wall',  
'funding', 'need', 'daca', 'clean', 'said', 'wednesday', 'trump',  
'aides', 'meet', 'congressional', 'leaders', 'discuss', 'issues',  
'followed', 'weekend', 'strategy', 'sessions', 'trump', 'republican',  
'leaders', 'white', 'house', 'said', 'trump', 'scheduled', 'meet',  
'sunday', 'florida', 'republican', 'governor', 'rick', 'scott',  
'wants', 'emergency', 'house', 'passed', 'billion', 'package',  
'hurricanes', 'florida', 'texas', 'puerto', 'rico', 'wildfires',  
'california', 'package', 'exceeded', 'billion', 'requested', 'trump',  
'administration', 'senate', 'voted']
```

```
# Obtain the total words present in the dataset
```

```
list_of_words = []  
for i in df.clean:  
    for j in i:  
        list_of_words.append(j)
```

```
list_of_words
```

```
['budget',  
'fight',  
'looms',  
'republicans',  
'flip',  
'fiscal',  
'script',  
'washington',  
'reuters',
```

'head',
'conservative',
'republican',
'faction',
'congress',
'voted',
'month',
'huge',
'expansion',
'national',
'debt',
'cuts',
'called',
'fiscal',
'conservative',
'sunday',
'urged',
'budget',
'restraint',
'keeping',
'sharp',
'pivot',
'republicans',
'representative',
'mark',
'meadows',
'speaking',
'face',
'nation',
'drew',
'hard',
'line',
'federal',
'spending',
'lawmakers',
'bracing',
'battle',
'january',
'return',
'holidays',
'wednesday',
'lawmakers',
'begin',
'trying',
'pass',
'federal',
'budget',
'fight',
'likely',

'linked',
'issues',
'immigration',
'policy',
'november',
'congressional',
'election',
'campaigns',
'approach',
'republicans',
'seek',
'control',
'congress',
'president',
'donald',
'trump',
'republicans',
'want',
'budget',
'increase',
'military',
'spending',
'democrats',
'want',
'proportional',
'increases',
'defense',
'discretionary',
'spending',
'programs',
'support',
'education',
'scientific',
'research',
'infrastructure',
'public',
'health',
'environmental',
'protection',
'trump',
'administration',
'willing',
'going',
'increase',
'defense',
'discretionary',
'spending',
'percent',
'meadows',

'chairman',
'small',
'influential',
'house',
'freedom',
'caucus',
'said',
'program',
'democrats',
'saying',
'need',
'government',
'raise',
'percent',
'fiscal',
'conservative',
'rationale',
'eventually',
'people',
'money',
'said',
'meadows',
'republicans',
'voted',
'late',
'december',
'party',
'debt',
'financed',
'overhaul',
'expected',
'balloon',
'federal',
'budget',
'deficit',
'trillion',
'years',
'trillion',
'national',
'debt',
'interesting',
'hear',
'mark',
'talk',
'fiscal',
'responsibility',
'democratic',
'representative',
'joseph',

'crowley',
'said',
'crowley',
'said',
'republican',
'require',
'united',
'states',
'borrow',
'trillion',
'paid',
'future',
'generations',
'finance',
'cuts',
'corporations',
'rich',
'fiscally',
'responsible',
'bills',
'seen',
'passed',
'history',
'house',
'representatives',
'think',
'going',
'paying',
'years',
'come',
'crowley',
'said',
'republicans',
'insist',
'package',
'biggest',
'overhaul',
'years',
'boost',
'economy',
'growth',
'house',
'speaker',
'paul',
'ryan',
'supported',
'recently',
'went',
'meadows',

'making',
'clear',
'radio',
'interview',
'welfare',
'entitlement',
'reform',
'party',
'calls',
'republican',
'priority',
'republican',
'parlance',
'entitlement',
'programs',
'mean',
'food',
'stamps',
'housing',
'assistance',
'medicare',
'medicaid',
'health',
'insurance',
'elderly',
'poor',
'disabled',
'programs',
'created',
'washington',
'assist',
'needy',
'democrats',
'seized',
'ryan',
'early',
'december',
'remarks',
'saying',
'showed',
'republicans',
'overhaul',
'seeking',
'spending',
'cuts',
'social',
'programs',
'goals',
'house',

'republicans',
'seat',
'senate',
'votes',
'democrats',
'needed',
'approve',
'budget',
'prevent',
'government',
'shutdown',
'democrats',
'leverage',
'senate',
'republicans',
'narrowly',
'control',
'defend',
'discretionary',
'defense',
'programs',
'social',
'spending',
'tackling',
'issue',
'dreamers',
'people',
'brought',
'illegally',
'country',
'children',
'trump',
'september',
'march',
'expiration',
'date',
'deferred',
'action',
'childhood',
'arrivals',
'daca',
'program',
'protects',
'young',
'immigrants',
'deportation',
'provides',
'work',
'permits',

'president',
'said',
'recent',
'twitter',
'messages',
'wants',
'funding',
'proposed',
'mexican',
'border',
'wall',
'immigration',
'changes',
'exchange',
'agreeing',
'help',
'dreamers',
'representative',
'debbie',
'dingell',
'told',
'favor',
'linking',
'issue',
'policy',
'objectives',
'wall',
'funding',
'need',
'daca',
'clean',
'said',
'wednesday',
'trump',
'aides',
'meet',
'congressional',
'leaders',
'discuss',
'issues',
'followed',
'weekend',
'strategy',
'sessions',
'trump',
'republican',
'leaders',
'white',
'house',

'said',
'trump',
'scheduled',
'meet',
'sunday',
'florida',
'republican',
'governor',
'rick',
'scott',
'wants',
'emergency',
'house',
'passed',
'billion',
'package',
'hurricanes',
'florida',
'texas',
'puerto',
'rico',
'wildfires',
'california',
'package',
'exceeded',
'billion',
'requested',
'trump',
'administration',
'senate',
'voted',
'military',
'accept',
'transgender',
'recruits',
'monday',
'pentagon',
'washington',
'reuters',
'transgender',
'people',
'allowed',
'time',
'enlist',
'military',
'starting',
'monday',
'ordered',
'federal',

'courts',
'pentagon',
'said',
'friday',
'president',
'donald',
'trump',
'administration',
'decided',
'appeal',
'rulings',
'blocked',
'transgender',
'federal',
'appeals',
'courts',
'washington',
'virginia',
'week',
'rejected',
'administration',
'request',
'hold',
'orders',
'lower',
'court',
'judges',
'requiring',
'military',
'begin',
'accepting',
'transgender',
'recruits',
'justice',
'department',
'official',
'said',
'administration',
'challenge',
'rulings',
'department',
'defense',
'announced',
'releasing',
'independent',
'study',
'issues',
'coming',
'weeks',

'litigate',
'interim',
'appeal',
'occurs',
'administration',
'decided',
'wait',
'study',
'continue',
'defend',
'president',
'lawful',
'authority',
'district',
'court',
'meantime',
'official',
'said',
'speaking',
'condition',
'anonymity',
'september',
'pentagon',
'said',
'created',
'panel',
'senior',
'officials',
'study',
'implement',
'directive',
'trump',
'prohibit',
'transgender',
'individuals',
'serving',
'defense',
'department',
'submit',
'plan',
'trump',
'lawyers',
'representing',
'currently',
'serving',
'transgender',
'service',
'members',
'aspiring',
'recruits',

'said',
'expected',
'administration',
'appeal',
'rulings',
'conservative',
'majority',
'supreme',
'court',
'hoping',
'happen',
'pentagon',
'spokeswoman',
'heather',
'babb',
'said',
'statement',
'mandated',
'court',
'order',
'department',
'defense',
'prepared',
'begin',
'accessing',
'transgender',
'applicants',
'military',
'service',
'applicants',
'meet',
'accession',
'standards',
'jennifer',
'levi',
'lawyer',
'lesbian',
'transgender',
'advocacy',
'group',
'glad',
'called',
'decision',
'appeal',
'great',
'news',
'hoping',
'means',
'government',

'come',
'justify',
'good',
'military',
'country',
'levi',
'said',
'glad',
'american',
'civil',
'liberties',
'union',
'represent',
'plaintiffs',
'lawsuits',
'filed',
'administration',
'appealed',
'hard',
'line',
'conservative',
'supporters',
'trump',
'announced',
'july',
'prohibit',
'transgender',
'people',
'serving',
'military',
'reversing',
'democratic',
'president',
'barack',
'obama',
'policy',
'accepting',
'trump',
'said',
'twitter',
'time',
'military',
'burdened',
'tremendous',
'medical',
'costs',
'disruption',
'transgender',
'military',

'entail',
'federal',
'judges',
'baltimore',
'washington',
'seattle',
'riverside',
'california',
'issued',
'rulings',
'blocking',
'trump',
'legal',
'challenges',
'republican',
'president',
'policy',
'proceed',
'judges',
'said',
'likely',
'violate',
'right',
'constitution',
'equal',
'protection',
'pentagon',
'issued',
'guidelines',
'recruitment',
'personnel',
'order',
'enlist',
'transgender',
'applicants',
'memo',
'outlined',
'medical',
'requirements',
'specified',
'applicants',
'identified',
'undergarments',
'wear',
'trump',
'administration',
'previously',
'said',
'legal',

'papers',
'armed',
'forces',
'prepared',
'train',
'thousands',
'personnel',
'medical',
'standards',
'needed',
'process',
'transgender',
'applicants',
'accept',
'individuals',
'medically',
'service',
'obama',
'administration',
'deadline',
'july',
'begin',
'accepting',
'transgender',
'recruits',
'trump',
'defense',
'secretary',
'james',
'mattis',
'postponed',
'date',
'president',
'indefinitely',
'trump',
'taken',
'steps',
'aimed',
'rolling',
'transgender',
'rights',
'october',
'administration',
'said',
'federal',
'banning',
'gender',
'based',
'workplace',

'discrimination',
'protect',
'transgender',
'employees',
'reversing',
'obama',
'position',
'february',
'trump',
'rescinded',
'guidance',
'issued',
'obama',
'administration',
'saying',
'public',
'schools',
'allow',
'transgender',
'students',
'restroom',
'corresponds',
'gender',
'identity',
'senior',
'republican',
'senator',
'mueller',
'washington',
'reuters',
'special',
'counsel',
'investigation',
'links',
'russia',
'president',
'trump',
'election',
'campaign',
'continue',
'interference',
'despite',
'calls',
'trump',
'administration',
'allies',
'republican',
'lawmakers',
'shut',

'prominent',
'republican',
'senator',
'said',
'sunday',
'lindsey',
'graham',
'serves',
'senate',
'armed',
'forces',
'judiciary',
'committees',
'said',
'department',
'justice',
'special',
'counsel',
'robert',
'mueller',
'needs',
'carry',
'russia',
'investigation',
'political',
'interference',
'investigation',
'forward',
'investigation',
'conducted',
'political',
'influence',
'graham',
'said',
'face',
'nation',
'news',
'program',
'need',
'mueller',
'think',
'right',
'right',
'time',
'question',
'russia',
'interfered',
'election',
'trump',

'campaign',
'links',
'ordinated',
'effort',
'loomed',
'white',
'house',
'trump',
'took',
'office',
'january',
'shows',
'sign',
'receding',
'trump',
'prepares',
'second',
'year',
'power',
'despite',
'intensified',
'rhetoric',
'trump',
'allies',
'recent',
'weeks',
'accusing',
'mueller',
'team',
'bias',
'republican',
'president',
'trump',
'undercut',
'supporters',
'interview',
'week',
'york',
'times',
'said',
'expected',
'mueller',
'going',
'fair',
'russia',
'role',
'election',
'question',
'possible',

'links',
'trump',
'campaign',
'focus',
'multiple',
'inquiries',
'washington',
'committees',
'senate',
'house',
'representatives',
'investigating',
'mueller',
'team',
'took',
'earlier',
'probe',
'launched',
'federal',
'bureau',
'investigation',
'members',
'trump',
'campaign',
'administration',
'convicted',
'indicted',
'investigation',
'trump',
'allies',
'deny',
'collusion',
'russia',
'campaign',
'kremlin',
'denied',
'meddling',
'election',
'graham',
'said',
'wants',
'examination',
'dossier',
'links',
'trump',
'russia',
'compiled',
'british',
'christopher',

'steele',
'prompted',
'trump',
'allies',
'republicans',
'question',
'mueller',
'inquiry',
'saturday',
'york',
'times',
'reported',
'dossier',
'triggered',
'early',
'probe',
'trump',
'campaign',
'foreign',
'policy',
'adviser',
'george',
'papadopoulos',
'australian',
'diplomat',
'russia',
'damaging',
'information',
'trump',
'rival',
'hillary',
'clinton',
'want',
'somebody',
'look',
'department',
'justice',
'dossier',
'bothers',
'greatly',
'want',
'somebody',
'look',
'graham',
'said',
'said',
'russia',
'investigation',
'continue',

'matter',
'fact',
'hurt',
'ignored',
'said',
'russia',
'probe',
'helped',
'australian',
'diplomat',
'washington',
'reuters',
'trump',
'campaign',
'adviser',
'george',
'papadopoulos',
'told',
'australian',
'diplomat',
'russia',
'political',
'dirt',
'democratic',
'presidential',
'candidate',
'hillary',
'clinton',
'york',
'times',
'reported',
'saturday',
'conversation',
'papadopoulos',
'diplomat',
'alexander',
'downer',
'london',
'driving',
'factor',
'decision',
'open',
'counter',
'intelligence',
'investigation',
'moscow',
'contacts',
'trump',
'campaign',

```

'times',
'reported',
'months',
'meeting',
'australian',
'officials',
'passed',
'information',
'came',
'papadopoulos',
...]

len(list_of_words)

9276947

# Obtain the total number of unique words (using set())
total_words = len(list(set(list_of_words)))
total_words

108704

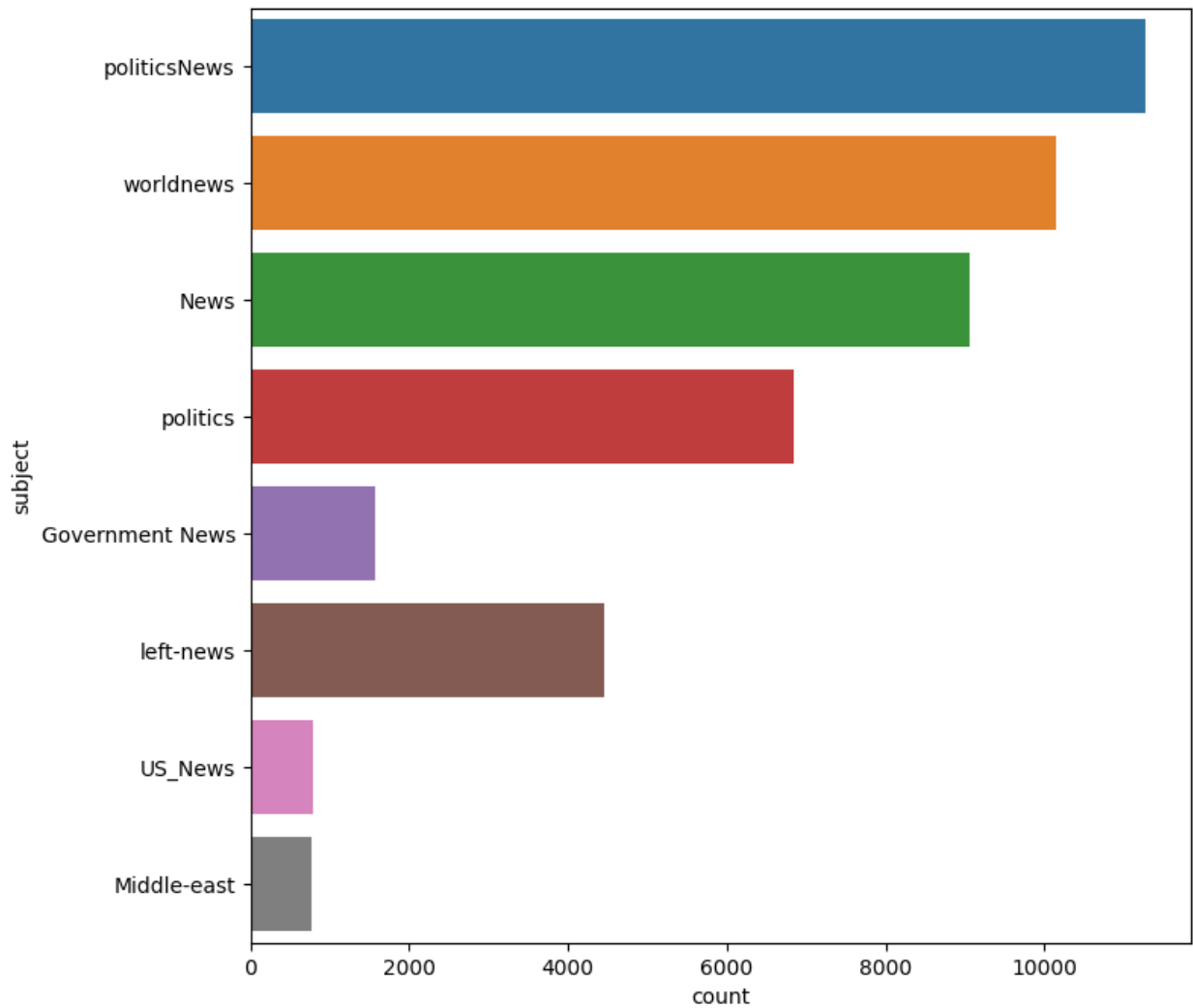
df['clean_joined'] = df['clean'].apply(lambda x: " ".join(x))
df['clean_joined'][0]

{"type": "string"}

plt.figure(figsize = (8, 8))
sns.countplot(y = "subject", data = df)

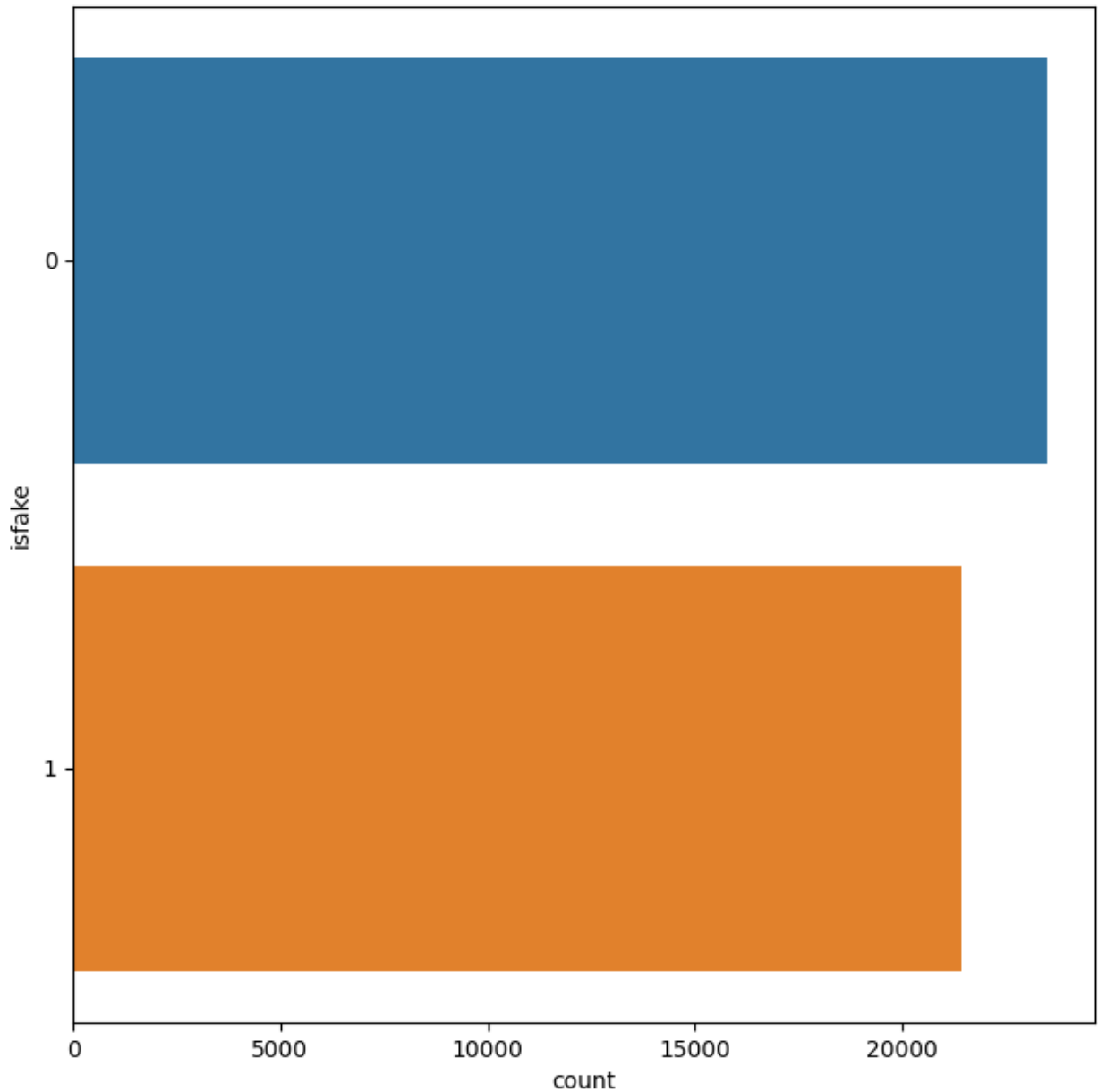
<Axes: xlabel='count', ylabel='subject'>

```



```
# plot the number of samples in 'isfake'
plt.figure(figsize = (8, 8))
sns.countplot(y = "isfake", data = df)

<Axes: xlabel='count', ylabel='isfake'>
```



```
# plot the word cloud for text that is Real
plt.figure(figsize = (20,20))
wc = WordCloud(max_words = 2000 , width = 1600 , height = 800 ,
stopwords = stop_words).generate(" ".join(df[df.isfake ==
1].clean_joined))
plt.imshow(wc, interpolation = 'bilinear')

<matplotlib.image.AxesImage at 0x790db2a81540>
```



```
# return all the tokens (each and every single word) as an array of tokens
```

```
nltk.word_tokenize(df['clean_joined'][1])
```

```
['military',  
'accept',  
'transgender',  
'recruits',  
'monday',  
'pentagon',  
'washington',  
'reuters',  
'transgender',  
'people',  
'allowed',  
'time',  
'enlist',  
'military',  
'starting',  
'monday',  
'ordered',  
'federal',  
'courts',  
'pentagon',  
'said',  
'friday',  
'president',  
'donald',  
'trump',  
'administration',  
'decided',  
'appeal',  
'rulings',  
'blocked',  
'transgender',  
'federal',  
'appeals',  
'courts',  
'washington',  
'virginia',  
'week',  
'rejected',  
'administration',  
'request',  
'hold',  
'orders',  
'lower',  
'court',  
'judges',  
'requiring',
```

'military',
'begin',
'accepting',
'transgender',
'recruits',
'justice',
'department',
'official',
'said',
'administration',
'challenge',
'rulings',
'department',
'defense',
'announced',
'releasing',
'independent',
'study',
'issues',
'coming',
'weeks',
'litigate',
'interim',
'appeal',
'occurs',
'administration',
'decided',
'wait',
'study',
'continue',
'defend',
'president',
'lawful',
'authority',
'district',
'court',
'meantime',
'official',
'said',
'speaking',
'condition',
'anonymity',
'september',
'pentagon',
'said',
'created',
'panel',
'senior',
'officials',

'study',
'implement',
'directive',
'trump',
'prohibit',
'transgender',
'individuals',
'serving',
'defense',
'department',
'submit',
'plan',
'trump',
'lawyers',
'representing',
'currently',
'serving',
'transgender',
'service',
'members',
'aspiring',
'recruits',
'said',
'expected',
'administration',
'appeal',
'rulings',
'conservative',
'majority',
'supreme',
'court',
'hoping',
'happen',
'pentagon',
'spokeswoman',
'heather',
'babb',
'said',
'statement',
'mandated',
'court',
'order',
'department',
'defense',
'prepared',
'begin',
'accessing',
'transgender',
'applicants',

'military',
'service',
'applicants',
'meet',
'accession',
'standards',
'jennifer',
'levi',
'lawyer',
'lesbian',
'transgender',
'advocacy',
'group',
'glad',
'called',
'decision',
'appeal',
'great',
'news',
'hoping',
'means',
'government',
'come',
'justify',
'good',
'military',
'country',
'levi',
'said',
'glad',
'american',
'civil',
'liberties',
'union',
'represent',
'plaintiffs',
'lawsuits',
'filed',
'administration',
'appealed',
'hard',
'line',
'conservative',
'supporters',
'trump',
'announced',
'july',
'prohibit',
'transgender',

'people',
'serving',
'military',
'reversing',
'democratic',
'president',
'barack',
'obama',
'policy',
'accepting',
'trump',
'said',
'twitter',
'time',
'military',
'burdened',
'tremendous',
'medical',
'costs',
'disruption',
'transgender',
'military',
'entail',
'federal',
'judges',
'baltimore',
'washington',
'seattle',
'riverside',
'california',
'issued',
'rulings',
'blocking',
'trump',
'legal',
'challenges',
'republican',
'president',
'policy',
'proceed',
'judges',
'said',
'likely',
'violate',
'right',
'constitution',
'equal',
'protection',
'pentagon',

'issued',
'guidelines',
'recruitment',
'personnel',
'order',
'enlist',
'transgender',
'applicants',
'memo',
'outlined',
'medical',
'requirements',
'specified',
'applicants',
'identified',
'undergarments',
'wear',
'trump',
'administration',
'previously',
'said',
'legal',
'papers',
'armed',
'forces',
'prepared',
'train',
'thousands',
'personnel',
'medical',
'standards',
'needed',
'process',
'transgender',
'applicants',
'accept',
'individuals',
'medically',
'service',
'obama',
'administration',
'deadline',
'july',
'begin',
'accepting',
'transgender',
'recruits',
'trump',
'defense',

```
'secretary',  
'james',  
'mattis',  
'postponed',  
'date',  
'president',  
'indefinitely',  
'trump',  
'taken',  
'steps',  
'aimed',  
'rolling',  
'transgender',  
'rights',  
'october',  
'administration',  
'said',  
'federal',  
'banning',  
'gender',  
'based',  
'workplace',  
'discrimination',  
'protect',  
'transgender',  
'employees',  
'reversing',  
'obama',  
'position',  
'february',  
'trump',  
'rescinded',  
'guidance',  
'issued',  
'obama',  
'administration',  
'saying',  
'public',  
'schools',  
'allow',  
'transgender',  
'students',  
'restroom',  
'corresponds',  
'gender',  
'identity']
```

```
# split data into test and train  
from sklearn.model_selection import train_test_split
```

```
x_train, x_test, y_train, y_test = train_test_split(df.clean_joined,
df.isfake, test_size = 0.2)
```

```
from nltk import word_tokenize
```

```
# Create a tokenizer to tokenize the words and create sequences of  
tokenized words
```

```
tokenizer = Tokenizer(num_words = total_words)
```

```
tokenizer.fit_on_texts(x_train)
```

```
train_sequences = tokenizer.texts_to_sequences(x_train)
```

```
test_sequences = tokenizer.texts_to_sequences(x_test)
```

```
len(train_sequences)
```

```
35918
```

```
len(test_sequences)
```

```
8980
```

```
print("The encoding for document\n",df.clean_joined[0],"\n is :  
",train_sequences[0])
```

The encoding for document

budget fight looms republicans flip fiscal script washington reuters
head conservative republican faction congress voted month huge
expansion national debt cuts called fiscal conservative sunday urged
budget restraint keeping sharp pivot republicans representative mark
meadows speaking face nation drew hard line federal spending lawmakers
bracing battle january return holidays wednesday lawmakers begin
trying pass federal budget fight likely linked issues immigration
policy november congressional election campaigns approach republicans
seek control congress president donald trump republicans want budget
increase military spending democrats want proportional increases
defense discretionary spending programs support education scientific
research infrastructure public health environmental protection trump
administration willing going increase defense discretionary spending
percent meadows chairman small influential house freedom caucus said
program democrats saying need government raise percent fiscal
conservative rationale eventually people money said meadows
republicans voted late december party debt financed overhaul expected
balloon federal budget deficit trillion years trillion national debt
interesting hear mark talk fiscal responsibility democratic
representative joseph crowley said crowley said republican require
united states borrow trillion paid future generations finance cuts
corporations rich fiscally responsible bills seen passed history house
representatives think going paying years come crowley said republicans
insist package biggest overhaul years boost economy growth house
speaker paul ryan supported recently went meadows making clear radio
interview welfare entitlement reform party calls republican priority
republican parlance entitlement programs mean food stamps housing
assistance medicare medicaid health insurance elderly poor disabled
programs created washington assist needy democrats seized ryan early

december remarks saying showed republicans overhaul seeking spending cuts social programs goals house republicans seat senate votes democrats needed approve budget prevent government shutdown democrats leverage senate republicans narrowly control defend discretionary defense programs social spending tackling issue dreamers people brought illegally country children trump september march expiration date deferred action childhood arrivals daca program protects young immigrants deportation provides work permits president said recent twitter messages wants funding proposed mexican border wall immigration changes exchange agreeing help dreamers representative debbie dingell told favor linking issue policy objectives wall funding need daca clean said wednesday trump aides meet congressional leaders discuss issues followed weekend strategy sessions trump republican leaders white house said trump scheduled meet sunday florida republican governor rick scott wants emergency house passed billion package hurricanes florida texas puerto rico wildfires california package exceeded billion requested trump administration senate voted is : [6, 16216, 706, 123, 887, 486, 358, 421, 213, 5089, 215, 5482, 3657, 388, 779, 56, 583, 388, 2270, 51, 3, 188, 23, 3344, 1162, 487, 11, 12434, 1941, 369, 486, 71, 44, 20, 1053, 118, 59, 17, 21, 2111, 44, 20, 3505, 217, 6, 41, 95, 2, 1053, 2357, 65, 1094, 31, 127, 65, 1123, 123, 479, 3669, 2616, 302, 4481, 250, 388, 59, 824, 145, 123, 59, 17, 22, 706, 61, 123, 1053, 782, 79, 627, 2435, 57, 5279, 30554, 3067, 123, 44, 20, 5862, 388, 66, 11, 3229, 2039, 1275, 731, 9690, 217, 123, 95, 2, 59, 322, 37, 1053, 771, 9194, 20, 1312, 50, 23, 1178, 4482, 1653, 6278, 24, 7, 19, 29, 59, 1464, 37, 1053, 3960, 4281, 44, 20, 1387, 3298, 382, 550, 1727, 5417, 170, 30, 1178, 1371, 1676, 7393, 71, 738, 1343, 5499, 516, 14309, 17, 4837, 487, 250, 70, 1589, 643, 4792, 844, 550, 118, 114, 5198, 899, 891, 1450, 13528, 2, 14309, 83, 5089, 215, 5482, 59, 70, 123, 732, 388, 3658, 358, 779]

Add padding can either be maxlen = 4406 or smaller number maxlen = 40 seems to work well based on results

Make sure all different samples have the same length (filled 0 s for missing)

```
padded_train = pad_sequences(train_sequences,maxlen = 40, padding = 'post', truncating = 'post')
```

```
padded_test = pad_sequences(test_sequences,maxlen = 40, truncating = 'post')
```

```
for i,doc in enumerate(padded_train[:2]):
```

```
    print("The padded encoding for document",i+1," is : ",doc)
```

```
The padded encoding for document 1 is : [ 6 16216 706 123
887 486 358 421 213 5089 215 5482
3657 388 779 56 583 388 2270 51 3 188 23
3344
1162 487 11 12434 1941 369 486 71 44 20 1053
118
```

```
59 17 21 2111]
```

```
The padded encoding for document 2 is : [ 27 1134 525 1975 5943
```

6	1312	64	85	6	41	1312	64	475								
	41	1116	486	29	345	883	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

```
pip install tensorflow
```

```
Requirement already satisfied: tensorflow in
/usr/local/lib/python3.10/dist-packages (2.13.0)
Requirement already satisfied: absl-py>=1.0.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.4.0)
Requirement already satisfied: astunparse>=1.6.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.6.3)
Requirement already satisfied: flatbuffers>=23.1.21 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (23.5.26)
Requirement already satisfied: gast<=0.4.0,>=0.2.1 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (0.4.0)
Requirement already satisfied: google-pasta>=0.1.1 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (0.2.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.59.0)
Requirement already satisfied: h5py>=2.9.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.9.0)
Requirement already satisfied: keras<2.14,>=2.13.1 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (2.13.1)
Requirement already satisfied: libclang>=13.0.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (16.0.6)
Requirement already satisfied: numpy<=1.24.3,>=1.22 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.23.5)
Requirement already satisfied: opt-einsum>=2.3.2 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.3.0)
Requirement already satisfied: packaging in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (23.2)
Requirement already satisfied: protobuf!=4.21.0,!
=4.21.1,!
=4.21.2,!
=4.21.3,!
=4.21.4,!
=4.21.5,<5.0.0dev,>=3.20.3 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (3.20.3)
Requirement already satisfied: setuptools in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (67.7.2)
Requirement already satisfied: six>=1.12.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.16.0)
Requirement already satisfied: tensorboard<2.14,>=2.13 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (2.13.0)
Requirement already satisfied: tensorflow-estimator<2.14,>=2.13.0
in /usr/local/lib/python3.10/dist-packages (from tensorflow) (2.13.0)
Requirement already satisfied: termcolor>=1.1.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (2.3.0)
Requirement already satisfied: typing-extensions<4.6.0,>=3.6.6 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (4.5.0)
Requirement already satisfied: wrapt>=1.11.0 in
/usr/local/lib/python3.10/dist-packages (from tensorflow) (1.15.0)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in
```



```
/usr/local/lib/python3.10/dist-packages (from tensorflow) (0.34.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in
/usr/local/lib/python3.10/dist-packages (from astunparse>=1.6.0-
>tensorflow) (0.41.2)
Requirement already satisfied: google-auth<3,>=1.6.3 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.14,>=2.13-
>tensorflow) (2.17.3)
Requirement already satisfied: google-auth-oauthlib<1.1,>=0.5 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.14,>=2.13-
>tensorflow) (1.0.0)
Requirement already satisfied: markdown>=2.6.8 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.14,>=2.13-
>tensorflow) (3.5)
Requirement already satisfied: requests<3,>=2.21.0 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.14,>=2.13-
>tensorflow) (2.31.0)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0
in /usr/local/lib/python3.10/dist-packages (from
tensorboard<2.14,>=2.13->tensorflow) (0.7.1)
Requirement already satisfied: werkzeug>=1.0.1 in
/usr/local/lib/python3.10/dist-packages (from tensorboard<2.14,>=2.13-
>tensorflow) (3.0.0)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.14,>=2.13->tensorflow) (5.3.1)
Requirement already satisfied: pyasn1-modules>=0.2.1 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.14,>=2.13->tensorflow) (0.3.0)
Requirement already satisfied: rsa<5,>=3.1.4 in
/usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3-
>tensorboard<2.14,>=2.13->tensorflow) (4.9)
Requirement already satisfied: requests-oauthlib>=0.7.0 in
/usr/local/lib/python3.10/dist-packages (from google-auth-
oauthlib<1.1,>=0.5->tensorboard<2.14,>=2.13->tensorflow) (1.3.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.14,>=2.13->tensorflow) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.14,>=2.13->tensorflow) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.14,>=2.13->tensorflow) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0-
>tensorboard<2.14,>=2.13->tensorflow) (2023.7.22)
Requirement already satisfied: MarkupSafe>=2.1.1 in
/usr/local/lib/python3.10/dist-packages (from werkzeug>=1.0.1-
>tensorboard<2.14,>=2.13->tensorflow) (2.1.3)
```

Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in
/usr/local/lib/python3.10/dist-packages (from pyasn1-modules>=0.2.1-
>google-auth<3,>=1.6.3->tensorboard<2.14,>=2.13->tensorflow) (0.5.0)
Requirement already satisfied: oauthlib>=3.0.0 in
/usr/local/lib/python3.10/dist-packages (from requests-
oauthlib>=0.7.0->google-auth-oauthlib<1.1,>=0.5-
>tensorboard<2.14,>=2.13->tensorflow) (3.2.2)

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten, Embedding, Input,
LSTM, Conv1D, MaxPool1D, Bidirectional
#building and training the model
#Picking up the sequential model
model = Sequential()
#embedding the layer
model.add(Embedding(total_words,output_dim=128))
#Using Bi-directional Rnn and LSTM
#128 no of neurons
model.add(Bidirectional(LSTM(128)))
#Dense layers
model.add(Dense(128,activation="relu"))
model.add(Dense(1,activation='sigmoid'))
model.compile(optimizer='adam',loss='binary_crossentropy',metrics=['ac
c'])
model.summary()
```

Model: "sequential_3"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, None, 128)	13914112
bidirectional (Bidirectional)	(None, 256)	263168
dense (Dense)	(None, 128)	32896
dense_1 (Dense)	(None, 1)	129
Total params: 14210305 (54.21 MB)		
Trainable params: 14210305 (54.21 MB)		
Non-trainable params: 0 (0.00 Byte)		

total_words

108704

```

y_train = np.asarray(y_train)
model.fit(padded_train, y_train,
batch_size=64,validation_split=0.1,epochs=4)

Epoch 1/4
506/506 [=====] - 246s 487ms/step - loss:
0.0094 - acc: 0.9978 - val_loss: 0.0044 - val_acc: 0.9997
Epoch 2/4
506/506 [=====] - 243s 481ms/step - loss:
6.5794e-04 - acc: 0.9999 - val_loss: 2.1264e-04 - val_acc: 1.0000
Epoch 3/4
506/506 [=====] - 258s 510ms/step - loss:
0.0025 - acc: 0.9994 - val_loss: 0.0010 - val_acc: 0.9997
Epoch 4/4
506/506 [=====] - 247s 488ms/step - loss:
4.3921e-04 - acc: 0.9999 - val_loss: 0.0029 - val_acc: 0.9992

<keras.src.callbacks.History at 0x790d822d47c0>

pred = model.predict(padded_test)
# if predicted value is greater than 0.5 it is real or else it is fake
prediction = []
for i in range(len(pred)):
    if pred[i].item()>0.5:
        prediction.append(1)
    else:
        prediction.append(0)

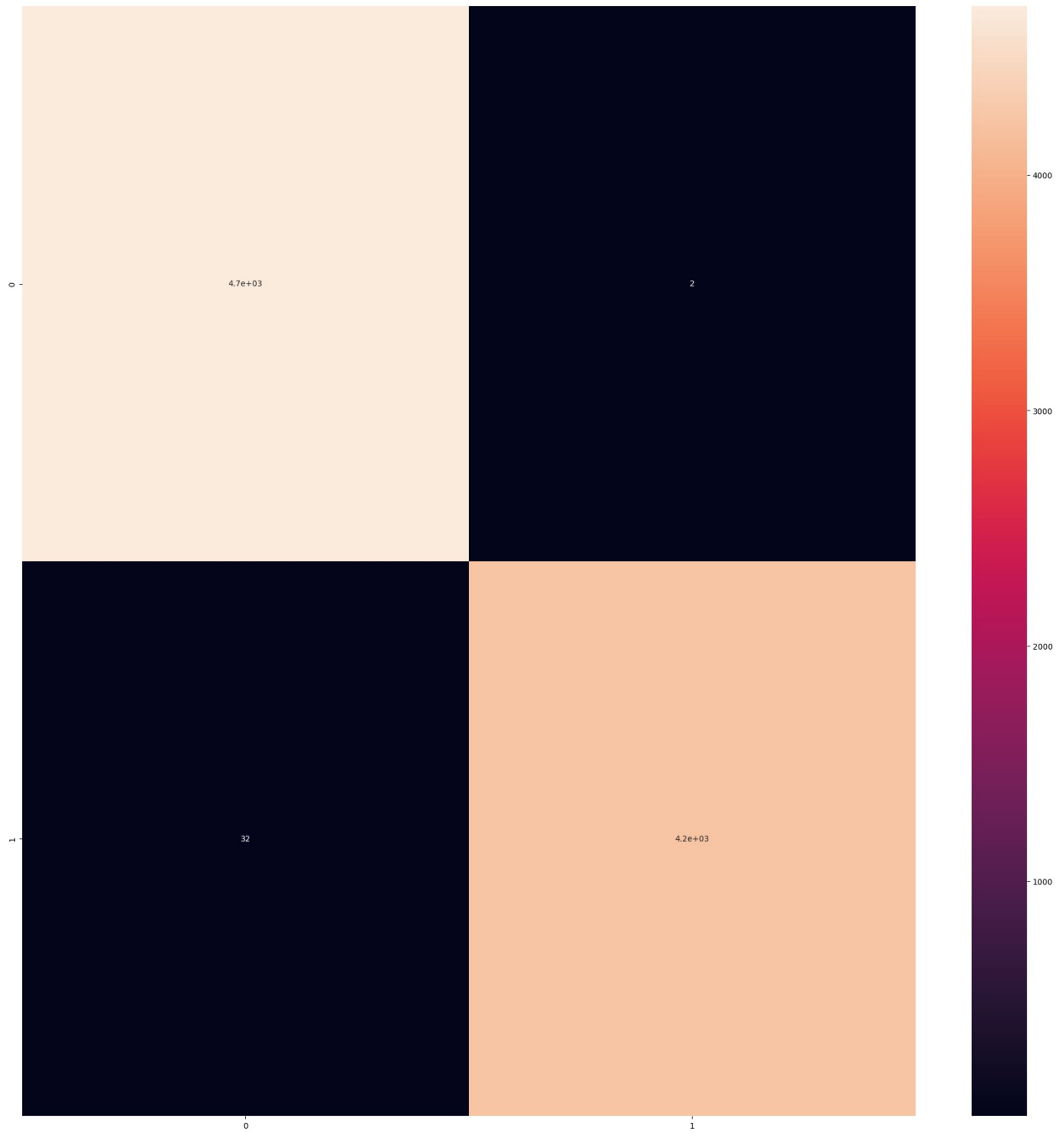
281/281 [=====] - 13s 45ms/step

from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix
accuracy = accuracy_score(list(y_test),prediction)
print('Model Accuracy: ',accuracy)
cm = confusion_matrix(list(y_test),prediction)
plt.figure(figsize=(25,25))
sns.heatmap(cm,annot=True)

Model Accuracy: 0.9962138084632517

<Axes: >

```



```
#category dictionary  
category = {0:"Fake News", 1 : "True News"}
```