2306AML117 MAHESH BABU M ASS17

In [1]:
```python
#IMPORTING REQUIRED LIBRERIES
import pandas as pd#for data manipulation
import numpy as np#for data mathematical caliculation
import seaborn as sns# data visuaization
import matplotlib.pyplot as plt# data visualization
from sklearn.model_selection import train_test_split #train ,test of data
from sklearn.metrics import accuracy_score# compare the scores ,accuracy of model
from sklearn.metrics import classification_report# compare  the scores
import re        # read the  text  data ,and manipulation
import string    # text data
import warnings
warnings.filterwarnings("ignore")


from collections import Counter
import tensorflow as tf
from tensorflow import keras# for layers
from keras.layers import Dense
keras.layers.Dropout
#metrics
from sklearn.metrics import accuracy_score, f1_score
from sklearn.metrics import classification_report, confusion_matrix

import tensorflow as tf
from tensorflow.keras.layers import Embedding, Dense, GlobalAveragePooling1D # deep learning techniques
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
```

In [2]:
```python
#Read the data using pandas
datafake=pd.read_csv("fake.csv")
```

In [3]:
```python
datafake# uploded text data frame
```

Out[3]:

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017
...	...	...	...	...
23476	McPain: John McCain Furious That Iran Treated ...	21st Century Wire says As 21WIRE reported earl...	Middle-east	January 16, 2016
23477	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...	21st Century Wire says It s a familiar theme. ...	Middle-east	January 16, 2016
23478	Sunnistan: US and Allied 'Safe Zone' Plan to T...	Patrick Henningsen 21st Century WireRemember ...	Middle-east	January 15, 2016
23479	How to Blow $700 Million: Al Jazeera America F...	21st Century Wire says Al Jazeera America will...	Middle-east	January 14, 2016
23480	10 U.S. Navy Sailors Held by Iranian Military ...	21st Century Wire says As 21WIRE predicted in ...	Middle-east	January 12, 2016

23481 rows × 4 columns

In [4]:
```python
datafake.isnull().sum()  # to find out any null or missg vales
```

Out[4]:
```
title      0
text       0
subject    0
date       0
dtype: int64
```

In [5]:
```python
datafake.shape  # data measurements like rows ,columns
```

Out[5]:
```
(23481, 4)
```

data consisting of 23481 rows and 4 columns

In [6]:
```python
#insert a column "class"
datafake['class']=0
```

In [7]:
```python
datafake.shape# after adding a column fake value is zero  ,size of data
```

Out[7]:
```
(23481, 5)
```

In [8]:
```python
datafake.head() # data consist of fake news with value zero
```

Out[8]:

	title	text	subject	date	class
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017	0
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017	0
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017	0
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017	0
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017	0

In [9]:
```python
datatrue=pd.read_csv('True.csv')  #uploading data of true news
```

In [10]:
```python
datatrue['class']=1 #assigning the true news as one value
```

In [11]:
```python
datatrue.head()# data rame oftrue news with value one
```

Out[11]:

	title	text	subject	date	class
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017	1
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017	1
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017	1
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017	1
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017	1

In [12]:
```python
datatrue.shape # data  measurement
```

Out[12]:
```
(21417, 5)
```

true news consist of 21417 rows and 5 columns

In [13]:
```python
datatrue.shape,datafake.shape #  compare both true and fake news data shapes
```

Out[13]:
```
((21417, 5), (23481, 5))
```

In [14]:
```python
# merging the true ,fake files
fakenewsdetection=pd.concat([datatrue,datafake], ignore_index=True, sort=False)
df=fakenewsdetection
df.head()
```

Out[14]:

	title	text	subject	date	class
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017	1
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017	1
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017	1
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017	1
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017	1

In [15]:
```python
df.tail(10)#  fake news consisting the  tail column
```

Out[15]:

	title	text	subject	date	class
44888	Seven Iranians freed in the prisoner swap have...	21st Century Wire says This week, the historic...	Middle-east	January 20, 2016	0
44889	#Hashtag Hell & The Fake Left	By Dady Chery and Gilbert MercierAll writers ...	Middle-east	January 19, 2016	0
44890	Astroturfing: Journalist Reveals Brainwashing ...	Vic Bishop Waking TimesOur reality is carefull...	Middle-east	January 19, 2016	0
44891	The New American Century: An Era of Fraud	Paul Craig RobertsIn the last years of the 20t...	Middle-east	January 19, 2016	0
44892	Hillary Clinton: 'Israel First' (and no peace ...	Robert Fantina CounterpunchAlthough the United...	Middle-east	January 18, 2016	0
44893	McPain: John McCain Furious That Iran Treated ...	21st Century Wire says As 21WIRE reported earl...	Middle-east	January 16, 2016	0
44894	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...	21st Century Wire says It s a familiar theme. ...	Middle-east	January 16, 2016	0
44895	Sunnistan: US and Allied 'Safe Zone' Plan to T...	Patrick Henningsen 21st Century WireRemember ...	Middle-east	January 15, 2016	0
44896	How to Blow $700 Million: Al Jazeera America F...	21st Century Wire says Al Jazeera America will...	Middle-east	January 14, 2016	0
44897	10 U.S. Navy Sailors Held by Iranian Military ...	21st Century Wire says As 21WIRE predicted in ...	Middle-east	January 12, 2016	0

In [16]:
```python
#random mixing of two samples  in random
df = df.sample(frac = 1)
```

In [17]:
```python
df#  data frame of true and fake news
```

Out[17]:

	title	text	subject	date	class
19630	Senator urges U.S. airlines to cap fares for p...	WASHINGTON (Reuters) - U.S. Senator Bill Nelso...	worldnews	September 18, 2017	1
29717	'Responsible' Texas Father Shoots His 9-Year-...	Republicans talk a good game about their suppo...	News	February 6, 2016	0
669	Trump distances himself from remarks on Putin ...	HANOI (Reuters) - U.S. President Donald Trump ...	politicsNews	November 12, 2017	1
41060	WHY Ugly HATE And DIVISION In America Is Actua...	This is Obama s legacy of hate and division, a...	left-news	Nov 11, 2016	0
17638	Tillerson consulted Britain, China, France, Ru...	WASHINGTON (Reuters) - U.S. Secretary of State...	worldnews	October 12, 2017	1
...	...	...	...	...	...
803	U.S. commerce chief Ross says nothing improper...	LONDON (Reuters) - U.S. Commerce Secretary Wil...	politicsNews	November 6, 2017	1
1368	Weinstein on indefinite leave as company inves...	LOS ANGELES (Reuters) - Co-chairman Harvey Wei...	politicsNews	October 6, 2017	1
23709	'Go Back To Your Own Country': Right-Wing Big...	Donald Trump s presidency has inspired many bi...	News	March 4, 2017	0
776	Democrat Phil Murphy wins New Jersey governor'...	WASHINGTON (Reuters) - Democrat Phil Murphy wa...	politicsNews	November 8, 2017	1
10760	Obama digs into research on potential Supreme ...	WASHINGTON (Reuters) - President Barack Obama ...	politicsNews	February 19, 2016	1

44898 rows × 5 columns

In [18]:
```python
df.head() #data frame showing the total data of real ansfake news
```

Out[18]:

	title	text	subject	date	class
19630	Senator urges U.S. airlines to cap fares for p...	WASHINGTON (Reuters) - U.S. Senator Bill Nelso...	worldnews	September 18, 2017	1
29717	'Responsible' Texas Father Shoots His 9-Year-...	Republicans talk a good game about their suppo...	News	February 6, 2016	0
669	Trump distances himself from remarks on Putin ...	HANOI (Reuters) - U.S. President Donald Trump ...	politicsNews	November 12, 2017	1
41060	WHY Ugly HATE And DIVISION In America Is Actua...	This is Obama s legacy of hate and division, a...	left-news	Nov 11, 2016	0
17638	Tillerson consulted Britain, China, France, Ru...	WASHINGTON (Reuters) - U.S. Secretary of State...	worldnews	October 12, 2017	1

In [19]:
```python
df.duplicated().sum()
```

Out[19]:
```
209
```

In [20]:
```python
df=df.drop_duplicates() #remove duplicates
```

In [21]:
```python
df.duplicated().sum()
```

Out[21]:
```
0
```

In [22]:
```python
df.head()
```

Out[22]:

	title	text	subject	date	class
19630	Senator urges U.S. airlines to cap fares for p...	WASHINGTON (Reuters) - U.S. Senator Bill Nelso...	worldnews	September 18, 2017	1
29717	'Responsible' Texas Father Shoots His 9-Year-...	Republicans talk a good game about their suppo...	News	February 6, 2016	0
669	Trump distances himself from remarks on Putin ...	HANOI (Reuters) - U.S. President Donald Trump ...	politicsNews	November 12, 2017	1
41060	WHY Ugly HATE And DIVISION In America Is Actua...	This is Obama s legacy of hate and division, a...	left-news	Nov 11, 2016	0
17638	Tillerson consulted Britain, China, France, Ru...	WASHINGTON (Reuters) - U.S. Secretary of State...	worldnews	October 12, 2017	1

In [23]:
```python
df['class'].value_counts()
```

Out[23]:
```
0    23478
1    21211
Name: class, dtype: int64
```

In [24]:
```python
#drop the un necessary columns
df1=df.drop(['title','subject','date'],axis=1)
```

In [25]:
```python
df1.reset_index(inplace = True)# resetting the index after deleting the  columns title, subject, date
df1.drop(["index"], axis = 1, inplace = True)# droping the indexing column
```

In [26]:
```python
df1.head() #data of required data which have to train and test
```

Out[26]:

	text	class
0	WASHINGTON (Reuters) - U.S. Senator Bill Nelso...	1
1	Republicans talk a good game about their suppo...	0
2	HANOI (Reuters) - U.S. President Donald Trump ...	1
3	This is Obama s legacy of hate and division, a...	0
4	WASHINGTON (Reuters) - U.S. Secretary of State...	1

In [27]:
```python
df1.tail()
```

Out[27]:

	text	class
44684	LONDON (Reuters) - U.S. Commerce Secretary Wil...	1
44685	LOS ANGELES (Reuters) - Co-chairman Harvey Wei...	1
44686	Donald Trump s presidency has inspired many bi...	0
44687	WASHINGTON (Reuters) - Democrat Phil Murphy wa...	1
44688	WASHINGTON (Reuters) - President Barack Obama ...	1

In [28]:
```python
def strip_all_entities(text):
    # replacing values (\r \n etc.) with space
    text = text.replace('\r', '').replace('\n', ' ').replace('\n', ' ').lower()
    #replacing the links
    text = re.sub('http://\S+|https://\S+', "", text)
    return text

def text_cleaning(text):
    # removing the characters that are not belong to character inside brackets
    text = re.sub(r'[^A-Za-zÀ-ú ]+', '', text)

    # Convert to lower case
    text = text.lower()

    # remove scores
    text = text.translate(str.maketrans('', '', string.punctuation))

    # Remove extra whitespace
    text = re.sub(r'\s+', ' ', text).strip()
    return text


df1['text']=df1['text'].apply(strip_all_entities)
df1['text']=df1['text'].apply(text_cleaning)
```

In [29]:
```python
df1.head()
```

Out[29]:

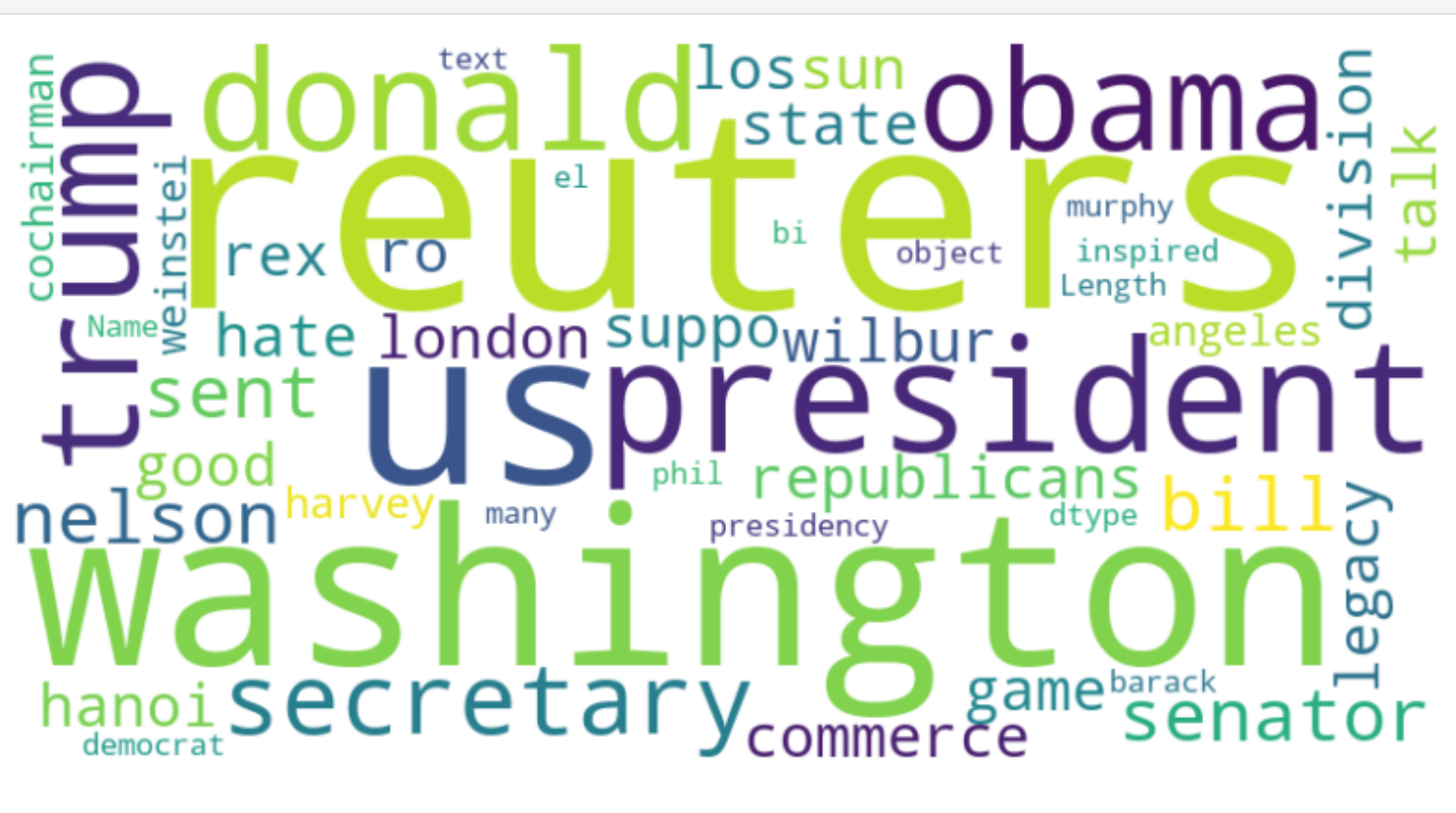
	text	class
0	washington reuters us senator bill nelson sent...	1
1	republicans talk a good game about their suppo...	0
2	hanoi reuters us president donald trump on sun...	1
3	this is obama s legacy of hate and division an...	0
4	washington reuters us secretary of state rex t...	1

In [30]:
```
!pip install wordcloud matplotlib nltk
```

```
Requirement already satisfied: wordcloud in c:\anaconda\lib\site-packages (1.9.2)
Requirement already satisfied: matplotlib in c:\anaconda\lib\site-packages (3.7.0)
Requirement already satisfied: nltk in c:\anaconda\lib\site-packages (3.7)
Requirement already satisfied: pillow in c:\anaconda\lib\site-packages (from wordcloud) (9.4.0)
Requirement already satisfied: numpy>=1.6.1 in c:\anaconda\lib\site-packages (from wordcloud) (1.23.5)
Requirement already satisfied: cycler>=0.10 in c:\anaconda\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: packaging>=20.0 in c:\anaconda\lib\site-packages (from matplotlib) (22.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\anaconda\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: contourpy>=1.0.1 in c:\anaconda\lib\site-packages (from matplotlib) (1.0.5)
Requirement already satisfied: fonttools>=4.22.0 in c:\anaconda\lib\site-packages (from matplotlib) (4.25.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\anaconda\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: pyparsing>=2.3.1 in c:\anaconda\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: click in c:\anaconda\lib\site-packages (from nltk) (8.0.4)
Requirement already satisfied: regex>=2021.8.3 in c:\anaconda\lib\site-packages (from nltk) (2022.7.9)
Requirement already satisfied: joblib in c:\anaconda\lib\site-packages (from nltk) (1.1.1)
Requirement already satisfied: tqdm in c:\anaconda\lib\site-packages (from nltk) (4.64.1)
Requirement already satisfied: six>=1.5 in c:\anaconda\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Requirement already satisfied: colorama in c:\anaconda\lib\site-packages (from click->nltk) (0.4.6)
```

In [31]:
```python
import nltk
import matplotlib.pyplot as plt
import seaborn as sn
from wordcloud import WordCloud
from collections import Counter
from keras.layers import LSTM

from nltk.corpus import stopwords #removing stepwords
from nltk.tokenize import word_tokenize # tokenize the sentence
from nltk.stem import WordNetLemmatizer, SnowballStemmer
```

In [32]:
```python
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import nltk
from nltk.corpus import stopwords

# Sample text for the word cloud
text = df1['text']

# Make sure the text is a string
if not isinstance(text, str):
    text = str(text)

# Tokenize the text using nltk
tokens = nltk.word_tokenize(text)

# Remove common stopwords
stop_words = set(stopwords.words('english'))
filtered_tokens = [word for word in tokens if word.lower() not in stop_words]

# Combine the tokens into a single string
text = ' '.join(filtered_tokens)

# Generate the word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)

# Display the word cloud using Matplotlib
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
```



training the model

In [33]:
```python
 y= df1["class"].values
#Converting X to format acceptable by gensim, removing annd punctuation stopwords in the process
X = []
stop_words = set(nltk.corpus.stopwords.words("english"))
tokenizer = nltk.tokenize.RegexpTokenizer(r'\w+')
for par in df1["text"].values:
    tmp = []
    sentences = nltk.sent_tokenize(par)
    for sent in sentences:
        sent = sent.lower()
        tokens = tokenizer.tokenize(sent)
        filtered_words = [w.strip() for w in tokens if w not in stop_words and len(w) > 1]
        tmp.extend(filtered_words)
    X.append(tmp)
```

In [34]:
```python
import gensim# gensim for topic modeling, word embeddings, and document similarity analysis.
```

In [35]:
```python
EMBEDDING_DIM = 100
w2v_model = gensim.models.Word2Vec(sentences=X, vector_size=EMBEDDING_DIM, window=5, min_count=1)
```

In [36]:
```python
len(w2v_model.wv.index_to_key)
```

Out[36]:
```
209869
```

In [37]:
```python
w2v_model.wv["match"]
```

Out[37]:
```
array([-0.85682017,  0.08574206, -0.12064003,  0.79403996, -0.35490164,
       -1.4354153 ,  0.79993236, -0.05141838,  0.65020084, -0.09870437,
       -0.9601121 , -0.0191504 , -0.07034373,  0.7142804 , -0.46105844,
        0.05468215,  0.01809822, -0.43793368, -0.5138375 ,  0.10336709,
        0.6779992 , -0.15689029,  0.36783624, -0.10619244, -0.20108637,
       -0.7068752 , -0.44216537, -0.13267277,  0.18250272, -1.9088554 ,
       -0.5462787 ,  0.4008284 , -0.9909133 , -0.489065  , -0.19645575,
       -0.48185226,  0.17678791, -0.03895253,  1.4100319 , -0.60995805,
       -0.24704537, -0.9614032 , -0.3919433 , -1.2689296 ,  0.7722162 ,
       -0.13928236, -0.54405457,  0.22575094,  0.27479848, -0.3438843 ,
        1.4655007 ,  0.7969257 , -0.7386859 ,  0.58176583,  0.31538686,
        0.11029206,  1.0605781 ,  0.00499073,  0.01754118, -0.256999  ,
        0.61209774,  0.01209338, -0.35631394,  0.09695489, -0.44617215,
        0.7039269 , -0.8467834 ,  0.5423378 , -0.62872237,  0.14470063,
       -1.540308  ,  0.15004262, -1.2728742 , -0.7642892 , -1.0633048 ,
       -0.66027313,  0.300896  , -0.31916347,  0.32803786, -0.04779035,
        0.05893159,  0.14918162,  0.03928469,  1.3631997 , -0.1656198 ,
       -0.53774506, -0.18537995, -0.0549104 , -0.23642851,  0.16396119,
       -0.09828877, -0.99348104, -0.25665456, -0.889005  ,  1.2295768 ,
        0.15580578, -1.1009908 , -0.40602934, -0.07766868,  0.7717875 ],
      dtype=float32)
```

In [38]:
```python
tokenizer = Tokenizer()#Let's convert reviews above to numeric by tokenizing
tokenizer.fit_on_texts(X)

X = tokenizer.texts_to_sequences(X)
```

In [39]:
```python
word_index = tokenizer.word_index
for word, num in word_index.items():
    print(f"{word} -> {num}")
    if num == 10:
        break
```

```
said -> 1
trump -> 2
us -> 3
would -> 4
president -> 5
people -> 6
one -> 7
state -> 8
also -> 9
new -> 10
```

In [40]:
```python
from tensorflow.keras.preprocessing.sequence import pad_sequences
```

In [41]:
```python
maxlen = 50
X = pad_sequences(X, maxlen=maxlen)#pad the tokenized reviews
```

In [42]:
```python
vocab_size = len(tokenizer.word_index) + 1#  vocabulary size defined
```

In [43]:
```python
def get_weight_matrix(model, vocab):
    vocab_size = len(vocab) + 1
    weight_matrix = np.zeros((vocab_size, EMBEDDING_DIM))
    for word, i in vocab.items():
        weight_matrix[i] = model.wv[word]
    return weight_matrix
```

In [113…
```python
embedding_vectors = get_weight_matrix(w2v_model, word_index)# words in a continuous vector space. They capture semantic relationships between words,
```

RNN with TENSOR FLOW

building the model

In [45]:
```python
from keras import Sequential
from tensorflow.keras.layers import Embedding
```

In [46]:
```python
model = Sequential()

model.add(Embedding(vocab_size, output_dim=EMBEDDING_DIM, weights=[embedding_vectors], input_length=maxlen, trainable=False))

model.add(LSTM(units=128))

model.add(Dense(16, activation='relu'))

model.add(Dense(1, activation='sigmoid'))
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['acc'])

#del embedding_vectors
```

In [47]:
```python
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 50, 100)           20987000

 lstm (LSTM)                 (None, 128)               117248

 dense (Dense)               (None, 16)                2064

 dense_1 (Dense)             (None, 1)                 17

=================================================================
Total params: 21106329 (80.51 MB)
Trainable params: 119329 (466.13 KB)
Non-trainable params: 20987000 (80.06 MB)
_________________________________________________________________
```

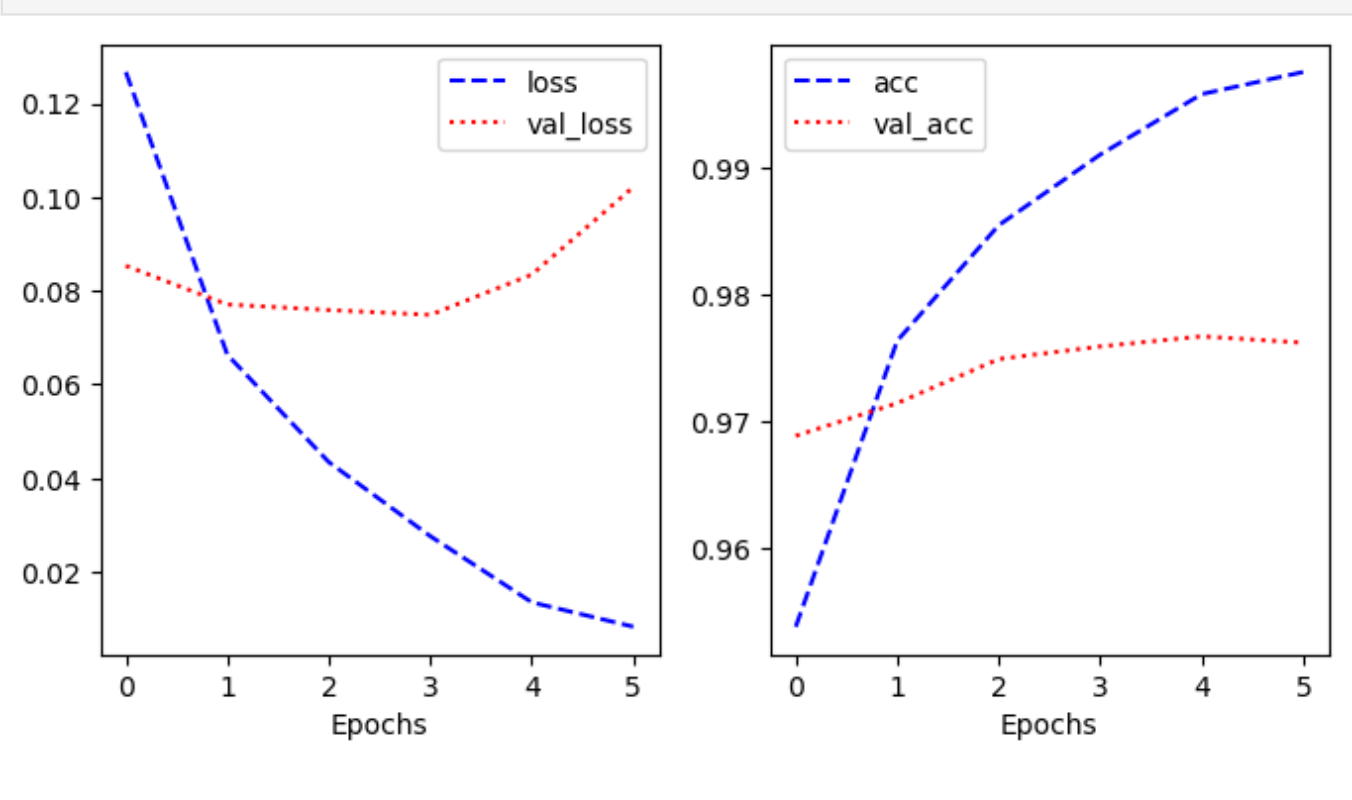
In [48]:
```python
X_train, X_test, y_train, y_test = train_test_split(X, y) # splitting the data
```

In [49]:
```python
history = model.fit(X_train, y_train, validation_split=0.3, epochs=6) # fitting the model
```

```
Epoch 1/6
734/734 [==============================] - 117s 154ms/step - loss: 0.1267 - acc: 0.9538 - val_loss: 0.0853 - val_acc: 0.9689
Epoch 2/6
734/734 [==============================] - 118s 161ms/step - loss: 0.0664 - acc: 0.9764 - val_loss: 0.0771 - val_acc: 0.9715
Epoch 3/6
734/734 [==============================] - 143s 194ms/step - loss: 0.0434 - acc: 0.9855 - val_loss: 0.0759 - val_acc: 0.9749
Epoch 4/6
734/734 [==============================] - 139s 189ms/step - loss: 0.0276 - acc: 0.9911 - val_loss: 0.0749 - val_acc: 0.9759
Epoch 5/6
734/734 [==============================] - 133s 181ms/step - loss: 0.0135 - acc: 0.9959 - val_loss: 0.0835 - val_acc: 0.9767
Epoch 6/6
734/734 [==============================] - 135s 183ms/step - loss: 0.0083 - acc: 0.9976 - val_loss: 0.1021 - val_acc: 0.9762
```

values are increasing loss functin as well as value losss accuracy is increasing while the epochs are increased hence the model treated as best model

In [50]:
```python
plt.figure(figsize=(8, 4))# plotig the graph of model pdeformance

plt.subplot(1, 2, 1)
plt.plot(history.history['loss'], 'b--', label = 'loss')
plt.plot(history.history['val_loss'], 'r:', label = 'val_loss')
plt.xlabel('Epochs')
plt.legend()

plt.subplot(1, 2, 2)
plt.plot(history.history['acc'], 'b--', label = 'acc')
plt.plot(history.history['val_acc'], 'r:', label = 'val_acc')
plt.xlabel('Epochs')
plt.legend()

plt.show()
```



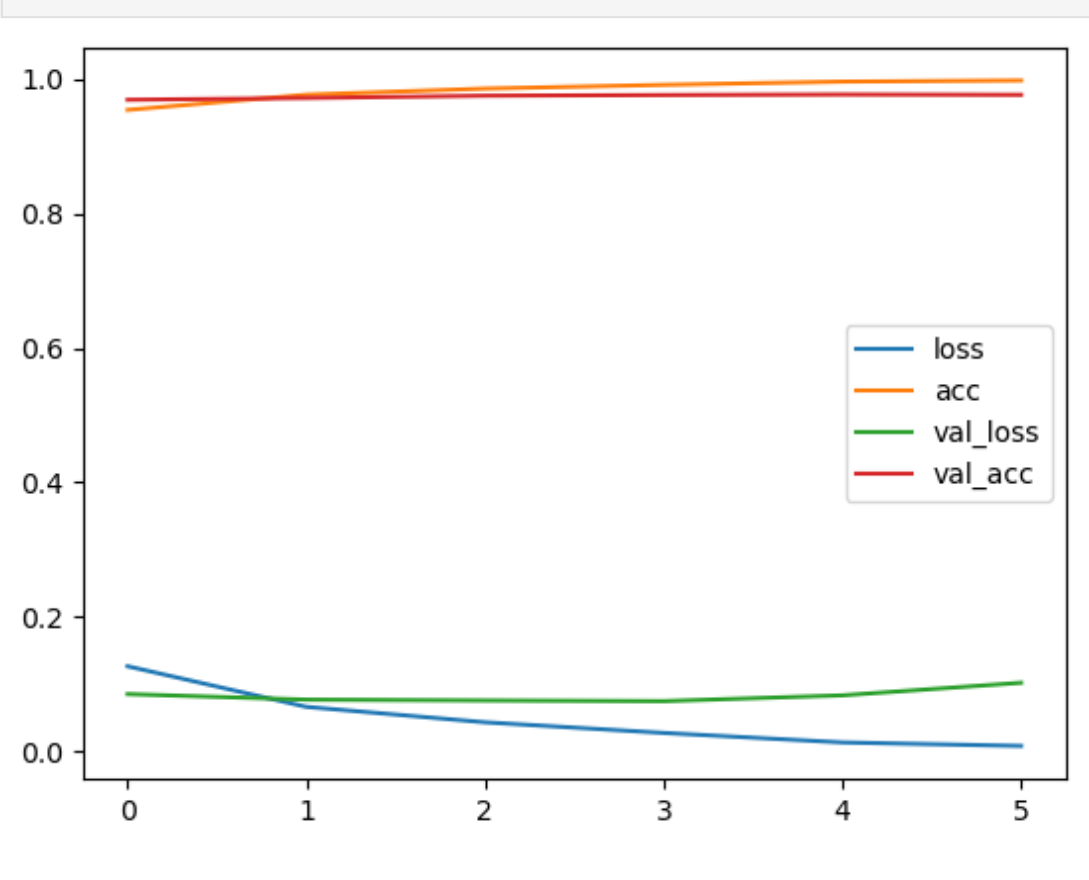
model evaluation

In [51]:
```python
model_loss = pd.DataFrame(model.history.history) # model evaluation of each statement
model_loss.head()
```

Out[51]:

	loss	acc	val_loss	val_acc
0	0.126724	0.953796	0.085342	0.968871
1	0.066356	0.976386	0.077080	0.971457
2	0.043415	0.985508	0.075897	0.974938
3	0.027603	0.991092	0.074867	0.975932
4	0.013476	0.995865	0.083478	0.976728

In [52]:
```python
model_loss.plot();
```



relatio between the values is near

Evaluating the model

In [53]:
```python
model.evaluate(X_train, y_train)
```

```
1048/1048 [==============================] - 52s 50ms/step - loss: 0.0356 - acc: 0.9912
```

Out[53]:
```
[0.03556814044713974, 0.9911982417106628]
```

In [54]:
```python
model.evaluate(X_test, y_test)
```

```
350/350 [==============================] - 16s 45ms/step - loss: 0.0978 - acc: 0.9774
```

Out[54]:
```
[0.09775781631469727, 0.9773561358451843]
```

In [55]:
```python
from sklearn.metrics import confusion_matrix, classification_report, accuracy_score, f1_score, roc_auc_score

y_train_pred = (model.predict(X_train) >= 0.5).astype("int32")

print(confusion_matrix(y_train, y_train_pred))
print("-------------------------------------------------------")
print(classification_report(y_train, y_train_pred))
```

```
1048/1048 [==============================] - 56s 53ms/step
[[17446   104]
 [  191 15775]]
-------------------------------------------------------
              precision    recall  f1-score   support

           0       0.99      0.99      0.99     17550
           1       0.99      0.99      0.99     15966

    accuracy                           0.99     33516
   macro avg       0.99      0.99      0.99     33516
weighted avg       0.99      0.99      0.99     33516
```

scores are real and fake news are near to compare good results

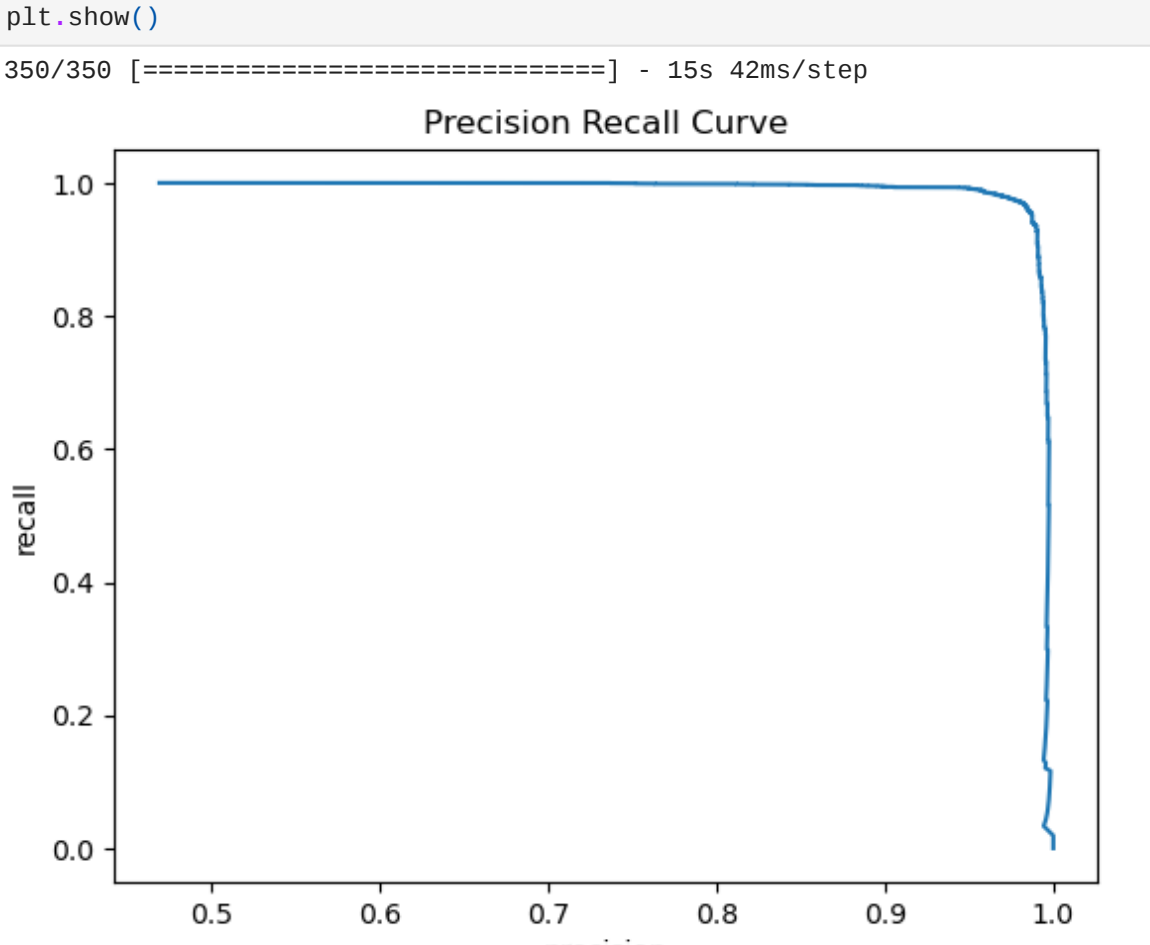
In [56]:
```python
y_pred = (model.predict(X_test) >= 0.5).astype("int32")

print(confusion_matrix(y_test, y_pred))
print("-------------------------------------------------------")
print(classification_report(y_test, y_pred))
```

```
350/350 [==============================] - 15s 44ms/step
[[5836   92]
 [ 161 5084]]
-------------------------------------------------------
              precision    recall  f1-score   support

           0       0.97      0.98      0.98      5928
           1       0.98      0.97      0.98      5245

    accuracy                           0.98     11173
   macro avg       0.98      0.98      0.98     11173
weighted avg       0.98      0.98      0.98     11173
```

confusion matrix also giving the good scores of so given model is good to predict

In [57]:
```python
from sklearn.metrics import precision_recall_curve, average_precision_score

y_pred_proba = model.predict(X_test)
precision, recall, thresholds = precision_recall_curve(y_test, y_pred_proba)

# plt.plot([1, 0], [0, 1],'k--')
plt.plot(precision, recall)
plt.xlabel('precision')
plt.ylabel('recall')
plt.title('Precision Recall Curve')
plt.show()
```

```
350/350 [==============================] - 15s 42ms/step
```



In [58]:
```python
from sklearn.metrics import precision_recall_curve, average_precision_score, recall_score

DL_AP = average_precision_score(y_test, y_pred_proba)
DL_f1 = f1_score(y_test, y_pred)
DL_rec = recall_score(y_test, y_pred)
```

prediction

In [95]:
```python
review1 = "A courts committee says it lacks authority to modify a broadcasting ban in time for the historic criminal cases against the former president."


review2 = "I am not just running against Crooked Hillary Clinton, I am running against the very dishonest and totally biased media - but I will win!"
review3 = "Heading to New Hampshire - will be talking about Hillary saying her brain SHORT CIRCUITED, and other things!"
review4 = "Trump is right about violent crime: Itâ€™s on the rise in major cities"
review5 ="Hillary Clinton has bad judgment and is unfit to serve as President"
review6 = "There is great unity in my campaign, perhaps greater than ever before. I want to thank everyone for your tremendous support. Beat Crooked H!"
review7 = "donald trump is indian"
review8 = "donalld trumpspeaks telugu"
review9="donald trump  takes milk "
review10="donald trump visited mahabub nagar "
review11= "donald trump was president "
review12="donald trump was former president "
review13="Vast numbers of manufacturing jobs in Pennsylvania have moved to Mexico and other countries. That will end when I win!"
review14="The following statementsÂ were posted to the verified Twitter accounts of U.S. President Donald Trump, @realDonaldTrump and @POTUS.  The opinions exp

review15="(Reuters) - The U.S. State Department has told refugee agencies it will sharply pare back the number of offices across the country authorized to rese
review16="Many people have raised the alarm regarding the fact that Donald Trump is dangerously close to becoming an autocrat. The thing is, democracies become

reviews = [review1, review2, review3, review4, review5, review6, review7, review8, review9, review10,review11,review12,review13,review14,review15,review16]
```

the prediction purpose data collected from data world Tweets from @realdonaldtrump. This includes the ID, the text of the tweet, and the source application. Inspired by the blog post http://varianceexplained.org/r/trump-tweets/

abvoe review 1 to 6 are taken from trumptweet source and 7 to12are not related to trumptweet source 14 to 16 from source

Let's convert reviews above to numeric by tokenizing

In [105…
```python
tokens = tokenizer.texts_to_sequences(reviews)
```

convert reviews above to numeric by tokenizing.

In [106…
```python
tokens = tokenizer.texts_to_sequences(reviews)
```

pad the tokenized reviews

In [107…
```python
tokens_pad = pad_sequences(tokens, maxlen=1000)
tokens_pad.shape
```

Out[107]:
```
(16, 1000)
```

*predict the sentiment of our reviews.**

In [108…
```python
mod_pred = model.predict(tokens_pad)
```

```
1/1 [==============================] - 0s 423ms/step
```

In [109…
```python
mod_pred
```

Out[109]:
```
array([[3.0681636e-02],
       [8.7762557e-05],
       [9.8575678e-05],
       [4.2360663e-04],
       [8.4935006e-04],
       [1.0132841e-05],
       [2.6730809e-03],
       [5.3098826e-03],
       [6.4844341e-04],
       [2.0647321e-03],
       [5.4644267e-03],
       [3.9108299e-02],
       [1.4126438e-02],
       [2.8334305e-02],
       [9.7664535e-01],
       [5.6837486e-09]], dtype=float32)
```

reate DataFrame for visually a better understanding.

In [110…
```python
df_pred = pd.DataFrame(mod_pred, index=reviews)
df_pred.rename(columns={0: 'Pred_Proba'}, inplace=True)
```

In [111…
```python
df_pred["Predicted_Feedbaack"] = df_pred["Pred_Proba"].apply(lambda x: "Recommended" if x>=0.5 else "Not Recommended")
```

In [112…
```python
df_pred
```

Out[112]:

	Pred_Proba	Predicted_Feedbaack
A courts committee says it lacks authority to modify a broadcasting ban in time for the historic criminal cases against the former president.	3.068164e-02	Not Recommended
I am not just running against Crooked Hillary Clinton, I am running against the very dishonest and totally biased media - but I will win!	8.776256e-05	Not Recommended
Heading to New Hampshire - will be talking about Hillary saying her brain SHORT CIRCUITED, and other things!	9.857568e-05	Not Recommended
Trump is right about violent crime: Itâ€™s on the rise in major cities	4.236066e-04	Not Recommended
Hillary Clinton has bad judgment and is unfit to serve as President	8.493501e-04	Not Recommended
There is great unity in my campaign, perhaps greater than ever before. I want to thank everyone for your tremendous support. Beat Crooked H!	1.013284e-05	Not Recommended
donald trump is indian	2.673081e-03	Not Recommended
donalld trumpspeaks telugu	5.309883e-03	Not Recommended
donald trump takes milk	6.484434e-04	Not Recommended
donald trump visited mahabub nagar	2.064732e-03	Not Recommended
donald trump was president	5.464427e-03	Not Recommended
donald trump was former president	3.910830e-02	Not Recommended
Vast numbers of manufacturing jobs in Pennsylvania have moved to Mexico and other countries. That will end when I win!	1.412644e-02	Not Recommended
The following statementsÂ were posted to the verified Twitter accounts of U.S. President Donald Trump, @realDonaldTrump and @POTUS. The opinions expressed are his own.Â Reuters has not edited the statements or confirmed their accuracy. @realDonaldTrump : - Our big and very popular Tax Cut and Reform Bill has taken on an unexpected new source of â€œloveâ€ - that is big companies and corporations showering their workers with bonuses. This is a phenomenon that nobody even thought of, and now it is the rage. Merry Christmas! [0747 EST] - At some point, and for the good of the country, I predict we will start working with the Democrats in a Bipartisan fashion. Infrastructure would be a perfect place to start. After having foolishly spent $7 trillion in the Middle East, it is time to start rebuilding our country! [0805 EST] - â€œThe President has accomplished some absolutely historic things during this past year.â€ Thank you Charlie Kirk of Turning Points USA. Sadly, the Fake Mainstream Media will NEVER talk about our accomplishments in their end of year reviews. We are compiling a long & beautiful list. [0917 EST] - With all my Administration has done on Legislative Approvals (broke Harry Trumanâ€™s Record), Regulation Cutting, Judicial Appointments, Building Military, VA, TAX CUTS & REFORM, Record Economy/Stock Market and so much more, I am sure great credit will be given by mainstream news? [1004 EST] - Will be signing the biggest ever Tax Cut and Reform Bill in 30 minutes in Oval Office. Will also be signing a much needed 4 billion dollar missile defense bill. [1007 EST] -- Source link: (bit.ly/2jBh4LU) (bit.ly/2jpEXYR)	2.833430e-02	Not Recommended
(Reuters) - The U.S. State Department has told refugee agencies it will sharply pare back the number of offices across the country authorized to resettle people in 2018 as President Donald Trump cuts the number of refugees allowed into the United States. The announcement was made at a Dec. 1 meeting in Washington with State Department officials and representatives from nine major refugee agencies, several executives of the agencies said. Advocates said the decision is likely to lead to the closure of dozens of resettlement offices around the country, potentially leaving some refugees without access to services that help them integrate into American life. Several state refugee coordinators said they had also been made aware of the closures. Â Â Â Refugee resettlement in the United States is handled by nine non-profit agencies that receive funding from the federal government for some of their refugee work. They partner with, or oversee, hundreds of local offices in nearly every state that help new arrivals with basic tasks like enrolling children in school, arranging doctorsâ€™ visits and applying for Social Security cards and other documents. Though the agencies are independent, they must get government approval for where they will resettle new refugees. Aid workers and state officials involved in refugee resettlement said the agencies were informed by the State Department in the Dec. 1 meeting that offices expected to handle fewer than 100 refugees in fiscal year 2018 will no longer be authorized to resettle new arrivals, which means many of them will have to close. There are about 300 resettlement offices spread across 49 states, and advocates estimate several dozen are at risk, though shuttering plans will not be finalized until next year. The Trump administration has said it wants refugees to assimilate quickly, both to promote national security and so that they can become self-sufficient. Refugee advocates say the closure of local offices will undermine that goal. They say the offices play a crucial role in helping newcomers traumatized from having fled conflict or persecution. Even if no new refugees are resettled by the offices they still have an obligation to help those already here, they say. If refugees lose access to â€œservices to help them navigate the processes of registering for school, and English classes and finding a job, that will mean that it will take longer for them to navigate life in the United States and contribute to our economy,â€ said Robert Carey, who directed the Office of Refugee Resettlement under former President Barack Obama. A State Department official confirmed the Dec. 1 meeting and said the agency is looking to â€œreduce costs and simplify management structures to help the U.S. Refugee Admissions Program run in a way that is fiscally responsible and sustainable.â€ Some conservative groups that favor lower immigration said they would welcome curbs on the agenciesâ€™ activities. â€œThese organizations have to adapt when their services are no longer needed as much,â€ said Jessica Vaughan, director of policy studies at the Center for Immigration Studies. â€œThere is no reason to keep funneling money to them.â€ Joshua Meservey, a senior policy analyst at the conservative Heritage Foundation who formerly worked in refugee resettlement, said that costs need to be balanced against benefits. â€œIt is unclear to me if the assimilation gains are great enough to justify the extra expenseâ€ of funding the smaller agencies, he said. The nine agencies are now trying to coordinate closures so that they can maintain at least one resettlement agency in as many states as possible, several agency executives said. â€œWeâ€™re hoping that they (the State Department) only close sites where there is possible duplication,â€ said Mark Hetfield, president of HIAS, one of the nine agencies. â€œThis is going to have to be a negotiation and a process.â€ Since taking office in January, Trump has moved to sharply reduce refugee admissions to the United States, because of national security concerns and a belief that money could be better spent resettling people closer to their original homes. Soon after taking office, he slashed the 2017 U.S. refugee cap to 50,000 from the 110,000 ceiling set by Obama. In September, he announced a cap of 45,000 for 2018, the lowest number since the modern U.S. refugee program was established in 1980. The resettlement office in Chattanooga, Tennessee is at risk of shutting down, because it is only projected to receive about 85 refugees, said Holly Johnson, the stateâ€™s refugee coordinator. â€œSmall doesnâ€™t necessarily mean weak or subpar,â€ Johnson said. â€œThey spend more time with folks, they have really well-established connections to the community, so people feel welcomed, which really helps.â€ Until this year, Idaho had four resettlement offices - three in Boise and one in Twin Falls, said Jan Reeves, director of the Idaho Office for Refugees, a non-profit which administers resettlement in the state. Earlier this year one of the sites in Boise shut down, he said. â€œIt was disruptive, and weâ€™ve lost a really valuable partner and weâ€™ve lost some capacity to do the job,â€ he said.	9.766454e-01	Recommended
Many people have raised the alarm regarding the fact that Donald Trump is dangerously close to becoming an autocrat. The thing is, democracies become autocracies right under the people s noses, because they can often look like democracies in the beginning phases. This was explained by Republican David Frum just a couple of months into Donald Trump s presidency, in a piece in The Atlantic called How to Build an Autocracy. In fact, if you really look critically at what is happening right now the systematic discrediting of vital institutions such as the free press and the Federal Bureau of Investigation as well the direct weaponization of the Department of Justice in order to go after Trump s former political opponent, 2016 Democratic nominee Hillary Clinton, and you have the makings of an autocracy. We are more than well on our way. Further, one chamber of Congress, the House of Representatives, already has a rogue band of Republicans who are running a parallel investigation to the official Russian collusion investigation, with the explicit intent of undermining and discrediting the idea that Trump could have possibly done anything wrong with the Russians in order to swing the 2016 election in his favor.All of that is just for starters, too. Now, we have Trump making United Nations Ambassador Nikki Haley bully and threaten other countries in the United Nations who voted against Trump s decision to change U.S. policy when it comes to recognition of Jerusalem as the capital of the Jewish State. Well, one expert, who is usually quite measured, has had enough of Trump s autocratic antics: Former CIA Director John O. Brennan. The seasoned spy took to Trump s favorite platform, Twitter, and blasted the decision:Trump Admin threat to retaliate against nations that exercise sovereign right in UN to oppose US position on Jerusalem is beyond outrageous. Shows @realDonaldTrump expects blind loyalty and subservience from everyone qualities usually found in narcissistic, vengeful autocrats. John O. Brennan (@JohnBrennan) December 21, 2017Director Brennan is correct, of course. Trump is behaving just like an autocrat, and so many people in the nation are asleep when it comes to this dangerous age, in which the greatest threat to democracy and the very fabric of the republic itself is the American president. Fellow Americans, we know the GOP-led Congress will not be the check on Trump that they are supposed to be. It s time to get out and flip the House and possibly the Senate in 2018, and resist in the meantime, if we want to save our country from devolving into something that looks more like Russia or North Korea than the America we have always know. We re already well on our way.Featured image via BRENDAN SMIALOWSKI/AFP/Getty Images	5.683749e-09	Not Recommended

all are fakenews 15 th one is taken from true source

In [ ]:

In [ ]: