

## Importing necessary libraries

```
In [1]: import pandas as pd
import numpy as np
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
nltk.download('stopwords')
from nltk.stem import PorterStemmer, WordNetLemmatizer
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing import sequence
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, SimpleRNN, LSTM, Dropout, Embedding
import warnings
warnings.filterwarnings("ignore")
```

## Reading the dataset

```
In [2]: df=pd.read_csv('news.csv')
df
```

```
Out[2]:
```

|       |  | title   | text | label |
|-------|--|---|------|-------|
| 0     | Donald Trump Sends Out Embarrassing New Year'...   | Donald Trump just couldn't wish all Americans ... | fake |       |
| 1     | Drunk Bragging Trump Staffer Started Russian ...   | House Intelligence Committee Chairman Devin Nu... | fake |       |
| 2     | Sheriff David Clarke Becomes An Internet Joke...   | On Friday, it was revealed that former Milwauk... | fake |       |
| 3     | Trump Is So Obsessed He Even Has Obama's Name...   | On Christmas day, Donald Trump announced that ... | fake |       |
| 4     | Pope Francis Just Called Out Donald Trump Dur...   | Pope Francis used his annual Christmas Day mes... | fake |       |
| ...   | ...  | ...   | ...  |       |
| 44893 | 'Fully committed' NATO backs new U.S. approach...  | BRUSSELS (Reuters) - NATO allies on Tuesday we... | true |       |
| 44894 | LexisNexis withdrew two products from Chinese ...  | LONDON (Reuters) - LexisNexis, a provider of l... | true |       |
| 44895 | Minsk cultural hub becomes haven from authorities  | MINSK (Reuters) - In the shadow of disused Sov... | true |       |
| 44896 | Vatican upbeat on possibility of Pope Francis ...  | MOSCOW (Reuters) - Vatican Secretary of State ... | true |       |
| 44897 | Indonesia to buy \$1.14 billion worth of Russia... | JAKARTA (Reuters) - Indonesia will buy 11 Sukh... | true |       |

44898 rows × 3 columns

## Checking for null values

In [3]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 44898 entries, 0 to 44897
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype  
---  -
 0   title   44898 non-null   object 
 1   text    44898 non-null   object 
 2   label   44898 non-null   object 
dtypes: object(3)
memory usage: 1.0+ MB
```

## Checking if there are any empty spaces in string

In [4]:

```
a=[]
for indx,title,text,label in df.itertuples():
    if title.isspace():
        a.append(indx)
    if text.isspace():
        a.append(indx)
```

In [5]: a

Out[5]:

```
[10923,
11041,
11190,
11225,
11236,
11241,
11247,
11249,
11267,
11268,
11283,
11284,
11285,
11289,
11290,
11292,
11295,
11296,
11301,
11302]
```

In [6]: df['label'].value\_counts()

Out[6]:

```
fake      23481
true       21417
Name: label, dtype: int64
```

```
In [7]: df.drop(index=a,inplace=True) #removing those rows which have empty spaces in string
```

```
In [8]: df
```

```
Out[8]:
```

|       |  | title   | text | label |
|-------|--|---|------|-------|
| 0     | Donald Trump Sends Out Embarrassing New Year'...   | Donald Trump just couldn t wish all Americans ... |      | fake  |
| 1     | Drunk Bragging Trump Staffer Started Russian ...   | House Intelligence Committee Chairman Devin Nu... |      | fake  |
| 2     | Sheriff David Clarke Becomes An Internet Joke...   | On Friday, it was revealed that former Milwauk... |      | fake  |
| 3     | Trump Is So Obsessed He Even Has Obama's Name...   | On Christmas day, Donald Trump announced that ... |      | fake  |
| 4     | Pope Francis Just Called Out Donald Trump Dur...   | Pope Francis used his annual Christmas Day mes... |      | fake  |
| ...   | ...  | ...   | ...  | ...   |
| 44893 | 'Fully committed' NATO backs new U.S. approach...  | BRUSSELS (Reuters) - NATO allies on Tuesday we... |      | true  |
| 44894 | LexisNexis withdrew two products from Chinese ...  | LONDON (Reuters) - LexisNexis, a provider of l... |      | true  |
| 44895 | Minsk cultural hub becomes haven from authorities  | MINSK (Reuters) - In the shadow of disused Sov... |      | true  |
| 44896 | Vatican upbeat on possibility of Pope Francis ...  | MOSCOW (Reuters) - Vatican Secretary of State ... |      | true  |
| 44897 | Indonesia to buy \$1.14 billion worth of Russia... | JAKARTA (Reuters) - Indonesia will buy 11 Sukh... |      | true  |

44267 rows × 3 columns

## Creating a function to automatically preprocess the text according to requirements

```
In [13]: lem=WordNetLemmatizer()
```

```
In [14]: stp=stopwords.words('english')
```

```
In [15]: def clean(x):
          x=word_tokenize(x.lower())
          x=[i for i in x if i.isalpha()]
          x=[i for i in x if i not in stp]
          x=[lem.lemmatize(i,pos='v')for i in x]
          x=' '.join(x)
          return x
```

## Applying function for preprocessing title and text columns

```
In [16]: df['title']=df['title'].apply(clean)
```

```
In [17]: df
```

```
Out[17]:
```

|       |   | title   | text | label |
|-------|---|---|------|-------|
| 0     | donald trump send embarrass new year eve messa... | Donald Trump just couldn t wish all Americans ... |      | fake  |
| 1     | drink brag trump staffer start russian collusi... | House Intelligence Committee Chairman Devin Nu... |      | fake  |
| 2     | sheriff david clarke become internet joke thre... | On Friday, it was revealed that former Milwauk... |      | fake  |
| 3     | trump obsess even obama name cod website image    | On Christmas day, Donald Trump announced that ... |      | fake  |
| 4     | pope francis call donald trump christmas speech   | Pope Francis used his annual Christmas Day mes... |      | fake  |
| ...   | ...   | ...   | ...  | ...   |
| 44893 | commit nato back new approach afghanistan         | BRUSSELS (Reuters) - NATO allies on Tuesday we... |      | true  |
| 44894 | lexisnexis withdraw two products chinese market   | LONDON (Reuters) - LexisNexis, a provider of l... |      | true  |
| 44895 | minsk cultural hub become authorities             | MINSK (Reuters) - In the shadow of disused Sov... |      | true  |
| 44896 | vatican upbeat possibility pope francis visit ... | MOSCOW (Reuters) - Vatican Secretary of State ... |      | true  |
| 44897 | indonesia buy billion worth russian jet           | JAKARTA (Reuters) - Indonesia will buy 11 Sukh... |      | true  |

44267 rows × 3 columns

```
In [18]: df['text']=df['text'].apply(clean)
```

In [19]: df

Out[19]:

|       |   | title   | text | label |
|-------|---|---|------|-------|
| 0     | donald trump send embarrass new year eve messa... | donald trump wish americans happy new year lea... | fake |       |
| 1     | drink brag trump staffer start russian collusi... | house intelligence committee chairman devin nu... | fake |       |
| 2     | sheriff david clarke become internet joke thre... | friday reveal former milwaukee sheriff david c... | fake |       |
| 3     | trump obsess even obama name cod website image    | christmas day donald trump announce would back... | fake |       |
| 4     | pope francis call donald trump christmas speech   | pope francis use annual christmas day message ... | fake |       |
| ...   | ...   | ...   | ...  |       |
| 44893 | commit nato back new approach afghanistan         | brussels reuters nato ally tuesday welcome pre... | true |       |
| 44894 | lexisnexis withdraw two products chinese market   | london reuters lexisnexis provider legal regul... | true |       |
| 44895 | minsk cultural hub become authorities             | minsk reuters shadow disused factories minsk s... | true |       |
| 44896 | vatican upbeat possibility pope francis visit ... | moscow reuters vatican secretary state cardina... | true |       |
| 44897 | indonesia buy billion worth russian jet           | jakarta reuters indonesia buy sukhoi fighter j... | true |       |

44267 rows × 3 columns

## Splitting input and output columns and encoding the output column

In [20]: `x=df['text']`  
`y=df['label']`

In [21]: `from sklearn.preprocessing import LabelEncoder`

In [22]: `le=LabelEncoder()`

In [23]: `y=le.fit_transform(df['label'])`

In [24]: `y`

Out[24]: `array([0, 0, 0, ..., 1, 1, 1])`

## Splitting the dataset into training and testing data

```
In [25]: from sklearn.model_selection import train_test_split
xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.3,random_state=0)
```

## Finding out the length of tokens/words for each text and representing them into a column

```
In [26]: senlen=[]
for i in df['text']:
    senlen.append(len(word_tokenize(i)))
```

```
In [27]: df['sentlen']=senlen
```

```
In [28]: df
```

```
Out[28]:
```

|       | title   | text  | label | sentlen |
|-------|---|---|-------|---------|
| 0     | donald trump send embarrass new year eve messa... | donald trump wish americans happy new year lea... | fake  | 242     |
| 1     | drink brag trump staffer start russian collusi... | house intelligence committee chairman devin nu... | fake  | 166     |
| 2     | sheriff david clarke become internet joke thre... | friday reveal former milwaukee sheriff david c... | fake  | 288     |
| 3     | trump obsess even obama name cod website image    | christmas day donald trump announce would back... | fake  | 226     |
| 4     | pope francis call donald trump christmas speech   | pope francis use annual christmas day message ... | fake  | 209     |
| ...   | ...   | ...   | ...   | ...     |
| 44893 | commit nato back new approach afghanistan         | brussels reuters nato ally tuesday welcome pre... | true  | 248     |
| 44894 | lexisnexis withdraw two products chinese market   | london reuters lexisnexis provider legal regul... | true  | 75      |
| 44895 | minsk cultural hub become authorities             | minsk reuters shadow disused factories minsk s... | true  | 163     |
| 44896 | vatican upbeat possibility pope francis visit ... | moscow reuters vatican secretary state cardina... | true  | 117     |
| 44897 | indonesia buy billion worth russian jet           | jakarta reuters indonesia buy sukhoi fighter j... | true  | 133     |

44267 rows × 4 columns

## Taking into account 95 percentile value of length and defining that length as the maximum length of all texts

```
In [29]: max(senlen)
```

```
Out[29]: 4606
```

```
In [30]: np.quantile(senlen,0.95)
```

```
Out[30]: 503.0
```

```
In [31]: lenmax=500
```

## Using the tokenizer function to analyse the text and assign a no to each word based on their importance

```
In [32]: tok=Tokenizer()  
tok.fit_on_texts(xtrain)
```

```
In [33]: tok.index_word
```

```
Out[33]: {1: 'say',  
          2: 'trump',  
          3: 'state',  
          4: 'would',  
          5: 'president',  
          6: 'people',  
          7: 'make',  
          8: 'one',  
          9: 'go',  
          10: 'also',  
          11: 'take',  
          12: 'new',  
          13: 'tell',  
          14: 'house',  
          15: 'reuters',  
          16: 'clinton',  
          17: 'donald',  
          18: 'time',  
          19: 'get',  
          20: 'the',  
          21: 'and',  
          22: 'of',  
          23: 'a',  
          24: 'in',  
          25: 'on',  
          26: 'to',  
          27: 'for',  
          28: 'with',  
          29: 'at',  
          30: 'by',  
          31: 'from',  
          32: 'as',  
          33: 'but',  
          34: 'or',  
          35: 'so',  
          36: 'that',  
          37: 'which',  
          38: 'who',  
          39: 'what',  
          40: 'when',  
          41: 'where',  
          42: 'how',  
          43: 'why',  
          44: 'if',  
          45: 'then',  
          46: 'and',  
          47: 'or',  
          48: 'so',  
          49: 'that',  
          50: 'which',  
          51: 'who',  
          52: 'what',  
          53: 'when',  
          54: 'where',  
          55: 'how',  
          56: 'why',  
          57: 'if',  
          58: 'then',  
          59: 'and',  
          60: 'or',  
          61: 'so',  
          62: 'that',  
          63: 'which',  
          64: 'who',  
          65: 'what',  
          66: 'when',  
          67: 'where',  
          68: 'how',  
          69: 'why',  
          70: 'if',  
          71: 'then',  
          72: 'and',  
          73: 'or',  
          74: 'so',  
          75: 'that',  
          76: 'which',  
          77: 'who',  
          78: 'what',  
          79: 'when',  
          80: 'where',  
          81: 'how',  
          82: 'why',  
          83: 'if',  
          84: 'then',  
          85: 'and',  
          86: 'or',  
          87: 'so',  
          88: 'that',  
          89: 'which',  
          90: 'who',  
          91: 'what',  
          92: 'when',  
          93: 'where',  
          94: 'how',  
          95: 'why',  
          96: 'if',  
          97: 'then',  
          98: 'and',  
          99: 'or',  
          100: 'so',  
          101: 'that',  
          102: 'which',  
          103: 'who',  
          104: 'what',  
          105: 'when',  
          106: 'where',  
          107: 'how',  
          108: 'why',  
          109: 'if',  
          110: 'then',  
          111: 'and',  
          112: 'or',  
          113: 'so',  
          114: 'that',  
          115: 'which',  
          116: 'who',  
          117: 'what',  
          118: 'when',  
          119: 'where',  
          120: 'how',  
          121: 'why',  
          122: 'if',  
          123: 'then',  
          124: 'and',  
          125: 'or',  
          126: 'so',  
          127: 'that',  
          128: 'which',  
          129: 'who',  
          130: 'what',  
          131: 'when',  
          132: 'where',  
          133: 'how',  
          134: 'why',  
          135: 'if',  
          136: 'then',  
          137: 'and',  
          138: 'or',  
          139: 'so',  
          140: 'that',  
          141: 'which',  
          142: 'who',  
          143: 'what',  
          144: 'when',  
          145: 'where',  
          146: 'how',  
          147: 'why',  
          148: 'if',  
          149: 'then',  
          150: 'and',  
          151: 'or',  
          152: 'so',  
          153: 'that',  
          154: 'which',  
          155: 'who',  
          156: 'what',  
          157: 'when',  
          158: 'where',  
          159: 'how',  
          160: 'why',  
          161: 'if',  
          162: 'then',  
          163: 'and',  
          164: 'or',  
          165: 'so',  
          166: 'that',  
          167: 'which',  
          168: 'who',  
          169: 'what',  
          170: 'when',  
          171: 'where',  
          172: 'how',  
          173: 'why',  
          174: 'if',  
          175: 'then',  
          176: 'and',  
          177: 'or',  
          178: 'so',  
          179: 'that',  
          180: 'which',  
          181: 'who',  
          182: 'what',  
          183: 'when',  
          184: 'where',  
          185: 'how',  
          186: 'why',  
          187: 'if',  
          188: 'then',  
          189: 'and',  
          190: 'or',  
          191: 'so',  
          192: 'that',  
          193: 'which',  
          194: 'who',  
          195: 'what',  
          196: 'when',  
          197: 'where',  
          198: 'how',  
          199: 'why',  
          200: 'if',  
          201: 'then',  
          202: 'and',  
          203: 'or',  
          204: 'so',  
          205: 'that',  
          206: 'which',  
          207: 'who',  
          208: 'what',  
          209: 'when',  
          210: 'where',  
          211: 'how',  
          212: 'why',  
          213: 'if',  
          214: 'then',  
          215: 'and',  
          216: 'or',  
          217: 'so',  
          218: 'that',  
          219: 'which',  
          220: 'who',  
          221: 'what',  
          222: 'when',  
          223: 'where',  
          224: 'how',  
          225: 'why',  
          226: 'if',  
          227: 'then',  
          228: 'and',  
          229: 'or',  
          230: 'so',  
          231: 'that',  
          232: 'which',  
          233: 'who',  
          234: 'what',  
          235: 'when',  
          236: 'where',  
          237: 'how',  
          238: 'why',  
          239: 'if',  
          240: 'then',  
          241: 'and',  
          242: 'or',  
          243: 'so',  
          244: 'that',  
          245: 'which',  
          246: 'who',  
          247: 'what',  
          248: 'when',  
          249: 'where',  
          250: 'how',  
          251: 'why',  
          252: 'if',  
          253: 'then',  
          254: 'and',  
          255: 'or',  
          256: 'so',  
          257: 'that',  
          258: 'which',  
          259: 'who',  
          260: 'what',  
          261: 'when',  
          262: 'where',  
          263: 'how',  
          264: 'why',  
          265: 'if',  
          266: 'then',  
          267: 'and',  
          268: 'or',  
          269: 'so',  
          270: 'that',  
          271: 'which',  
          272: 'who',  
          273: 'what',  
          274: 'when',  
          275: 'where',  
          276: 'how',  
          277: 'why',  
          278: 'if',  
          279: 'then',  
          280: 'and',  
          281: 'or',  
          282: 'so',  
          283: 'that',  
          284: 'which',  
          285: 'who',  
          286: 'what',  
          287: 'when',  
          288: 'where',  
          289: 'how',  
          290: 'why',  
          291: 'if',  
          292: 'then',  
          293: 'and',  
          294: 'or',  
          295: 'so',  
          296: 'that',  
          297: 'which',  
          298: 'who',  
          299: 'what',  
          300: 'when',  
          301: 'where',  
          302: 'how',  
          303: 'why',  
          304: 'if',  
          305: 'then',  
          306: 'and',  
          307: 'or',  
          308: 'so',  
          309: 'that',  
          310: 'which',  
          311: 'who',  
          312: 'what',  
          313: 'when',  
          314: 'where',  
          315: 'how',  
          316: 'why',  
          317: 'if',  
          318: 'then',  
          319: 'and',  
          320: 'or',  
          321: 'so',  
          322: 'that',  
          323: 'which',  
          324: 'who',  
          325: 'what',  
          326: 'when',  
          327: 'where',  
          328: 'how',  
          329: 'why',  
          330: 'if',  
          331: 'then',  
          332: 'and',  
          333: 'or',  
          334: 'so',  
          335: 'that',  
          336: 'which',  
          337: 'who',  
          338: 'what',  
          339: 'when',  
          340: 'where',  
          341: 'how',  
          342: 'why',  
          343: 'if',  
          344: 'then',  
          345: 'and',  
          346: 'or',  
          347: 'so',  
          348: 'that',  
          349: 'which',  
          350: 'who',  
          351: 'what',  
          352: 'when',  
          353: 'where',  
          354: 'how',  
          355: 'why',  
          356: 'if',  
          357: 'then',  
          358: 'and',  
          359: 'or',  
          360: 'so',  
          361: 'that',  
          362: 'which',  
          363: 'who',  
          364: 'what',  
          365: 'when',  
          366: 'where',  
          367: 'how',  
          368: 'why',  
          369: 'if',  
          370: 'then',  
          371: 'and',  
          372: 'or',  
          373: 'so',  
          374: 'that',  
          375: 'which',  
          376: 'who',  
          377: 'what',  
          378: 'when',  
          379: 'where',  
          380: 'how',  
          381: 'why',  
          382: 'if',  
          383: 'then',  
          384: 'and',  
          385: 'or',  
          386: 'so',  
          387: 'that',  
          388: 'which',  
          389: 'who',  
          390: 'what',  
          391: 'when',  
          392: 'where',  
          393: 'how',  
          394: 'why',  
          395: 'if',  
          396: 'then',  
          397: 'and',  
          398: 'or',  
          399: 'so',  
          400: 'that',  
          401: 'which',  
          402: 'who',  
          403: 'what',  
          404: 'when',  
          405: 'where',  
          406: 'how',  
          407: 'why',  
          408: 'if',  
          409: 'then',  
          410: 'and',  
          411: 'or',  
          412: 'so',  
          413: 'that',  
          414: 'which',  
          415: 'who',  
          416: 'what',  
          417: 'when',  
          418: 'where',  
          419: 'how',  
          420: 'why',  
          421: 'if',  
          422: 'then',  
          423: 'and',  
          424: 'or',  
          425: 'so',  
          426: 'that',  
          427: 'which',  
          428: 'who',  
          429: 'what',  
          430: 'when',  
          431: 'where',  
          432: 'how',  
          433: 'why',  
          434: 'if',  
          435: 'then',  
          436: 'and',  
          437: 'or',  
          438: 'so',  
          439: 'that',  
          440: 'which',  
          441: 'who',  
          442: 'what',  
          443: 'when',  
          444: 'where',  
          445: 'how',  
          446: 'why',  
          447: 'if',  
          448: 'then',  
          449: 'and',  
          450: 'or',  
          451: 'so',  
          452: 'that',  
          453: 'which',  
          454: 'who',  
          455: 'what',  
          456: 'when',  
          457: 'where',  
          458: 'how',  
          459: 'why',  
          460: 'if',  
          461: 'then',  
          462: 'and',  
          463: 'or',  
          464: 'so',  
          465: 'that',  
          466: 'which',  
          467: 'who',  
          468: 'what',  
          469: 'when',  
          470: 'where',  
          471: 'how',  
          472: 'why',  
          473: 'if',  
          474: 'then',  
          475: 'and',  
          476: 'or',  
          477: 'so',  
          478: 'that',  
          479: 'which',  
          480: 'who',  
          481: 'what',  
          482: 'when',  
          483: 'where',  
          484: 'how',  
          485: 'why',  
          486: 'if',  
          487: 'then',  
          488: 'and',  
          489: 'or',  
          490: 'so',  
          491: 'that',  
          492: 'which',  
          493: 'who',  
          494: 'what',  
          495: 'when',  
          496: 'where',  
          497: 'how',  
          498: 'why',  
          499: 'if',  
          500: 'then',  
          501: 'and',  
          502: 'or',  
          503: 'so',  
          504: 'that',  
          505: 'which',  
          506: 'who',  
          507: 'what',  
          508: 'when',  
          509: 'where',  
          510: 'how',  
          511: 'why',  
          512: 'if',  
          513: 'then',  
          514: 'and',  
          515: 'or',  
          516: 'so',  
          517: 'that',  
          518: 'which',  
          519: 'who',  
          520: 'what',  
          521: 'when',  
          522: 'where',  
          523: 'how',  
          524: 'why',  
          525: 'if',  
          526: 'then',  
          527: 'and',  
          528: 'or',  
          529: 'so',  
          530: 'that',  
          531: 'which',  
          532: 'who',  
          533: 'what',  
          534: 'when',  
          535: 'where',  
          536: 'how',  
          537: 'why',  
          538: 'if',  
          539: 'then',  
          540: 'and',  
          541: 'or',  
          542: 'so',  
          543: 'that',  
          544: 'which',  
          545: 'who',  
          546: 'what',  
          547: 'when',  
          548: 'where',  
          549: 'how',  
          550: 'why',  
          551: 'if',  
          552: 'then',  
          553: 'and',  
          554: 'or',  
          555: 'so',  
          556: 'that',  
          557: 'which',  
          558: 'who',  
          559: 'what',  
          560: 'when',  
          561: 'where',  
          562: 'how',  
          563: 'why',  
          564: 'if',  
          565: 'then',  
          566: 'and',  
          567: 'or',  
          568: 'so',  
          569: 'that',  
          570: 'which',  
          571: 'who',  
          572: 'what',  
          573: 'when',  
          574: 'where',  
          575: 'how',  
          576: 'why',  
          577: 'if',  
          578: 'then',  
          579: 'and',  
          580: 'or',  
          581: 'so',  
          582: 'that',  
          583: 'which',  
          584: 'who',  
          585: 'what',  
          586: 'when',  
          587: 'where',  
          588: 'how',  
          589: 'why',  
          590: 'if',  
          591: 'then',  
          592: 'and',  
          593: 'or',  
          594: 'so',  
          595: 'that',  
          596: 'which',  
          597: 'who',  
          598: 'what',  
          599: 'when',  
          600: 'where',  
          601: 'how',  
          602: 'why',  
          603: 'if',  
          604: 'then',  
          605: 'and',  
          606: 'or',  
          607: 'so',  
          608: 'that',  
          609: 'which',  
          610: 'who',  
          611: 'what',  
          612: 'when',  
          613: 'where',  
          614: 'how',  
          615: 'why',  
          616: 'if',  
          617: 'then',  
          618: 'and',  
          619: 'or',  
          620: 'so',  
          621: 'that',  
          622: 'which',  
          623: 'who',  
          624: 'what',  
          625: 'when',  
          626: 'where',  
          627: 'how',  
          628: 'why',  
          629: 'if',  
          630: 'then',  
          631: 'and',  
          632: 'or',  
          633: 'so',  
          634: 'that',  
          635: 'which',  
          636: 'who',  
          637: 'what',  
          638: 'when',  
          639: 'where',  
          640: 'how',  
          641: 'why',  
          642: 'if',  
          643: 'then',  
          644: 'and',  
          645: 'or',  
          646: 'so',  
          647: 'that',  
          648: 'which',  
          649: 'who',  
          650: 'what',  
          651: 'when',  
          652: 'where',  
          653: 'how',  
          654: 'why',  
          655: 'if',  
          656: 'then',  
          657: 'and',  
          658: 'or',  
          659: 'so',  
          660: 'that',  
          661: 'which',  
          662: 'who',  
          663: 'what',  
          664: 'when',  
          665: 'where',  
          666: 'how',  
          667: 'why',  
          668: 'if',  
          669: 'then',  
          670: 'and',  
          671: 'or',  
          672: 'so',  
          673: 'that',  
          674: 'which',  
          675: 'who',  
          676: 'what',  
          677: 'when',  
          678: 'where',  
          679: 'how',  
          680: 'why',  
          681: 'if',  
          682: 'then',  
          683: 'and',  
          684: 'or',  
          685: 'so',  
          686: 'that',  
          687: 'which',  
          688: 'who',  
          689: 'what',  
          690: 'when',  
          691: 'where',  
          692: 'how',  
          693: 'why',  
          694: 'if',  
          695: 'then',  
          696: 'and',  
          697: 'or',  
          698: 'so',  
          699: 'that',  
          700: 'which',  
          701: 'who',  
          702: 'what',  
          703: 'when',  
          704: 'where',  
          705: 'how',  
          706: 'why',  
          707: 'if',  
          708: 'then',  
          709: 'and',  
          710: 'or',  
          711: 'so',  
          712: 'that',  
          713: 'which',  
          714: 'who',  
          715: 'what',  
          716: 'when',  
          717: 'where',  
          718: 'how',  
          719: 'why',  
          720: 'if',  
          721: 'then',  
          722: 'and',  
          723: 'or',  
          724: 'so',  
          725: 'that',  
          726: 'which',  
          727: 'who',  
          728: 'what',  
          729: 'when',  
          730: 'where',  
          731: 'how',  
          732: 'why',  
          733: 'if',  
          734: 'then',  
          735: 'and',  
          736: 'or',  
          737: 'so',  
          738: 'that',  
          739: 'which',  
          740: 'who',  
          741: 'what',  
          742: 'when',  
          743: 'where',  
          744: 'how',  
          745: 'why',  
          746: 'if',  
          747: 'then',  
          748: 'and',  
          749: 'or',  
          750: 'so',  
          751: 'that',  
          752: 'which',  
          753: 'who',  
          754: 'what',  
          755: 'when',  
          756: 'where',  
          757: 'how',  
          758: 'why',  
          759: 'if',  
          760: 'then',  
          761: 'and',  
          762: 'or',  
          763: 'so',  
          764: 'that',  
          765: 'which',  
          766: 'who',  
          767: 'what',  
          768: 'when',  
          769: 'where',  
          770: 'how',  
          771: 'why',  
          772: 'if',  
          773: 'then',  
          774: 'and',  
          775: 'or',  
          776: 'so',  
          777: 'that',  
          778: 'which',  
          779: 'who',  
          780: 'what',  
          781: 'when',  
          782: 'where',  
          783: 'how',  
          784: 'why',  
          785: 'if',  
          786: 'then',  
          787: 'and',  
          788: 'or',  
          789: 'so',  
          790: 'that',  
          791: 'which',  
          792: 'who',  
          793: 'what',  
          794: 'when',  
          795: 'where',  
          796: 'how',  
          797: 'why',  
          798: 'if',  
          799: 'then',  
          800: 'and',  
          801: 'or',  
          802: 'so',  
          803: 'that',  
          804: 'which',  
          805: 'who',  
          806: 'what',  
          807: 'when',  
          808: 'where',  
          809: 'how',  
          810: 'why',  
          811: 'if',  
          812: 'then',  
          813: 'and',  
          814: 'or',  
          815: 'so',  
          816: 'that',  
          817: 'which',  
          818: 'who',  
          819: 'what',  
          820: 'when',  
          821: 'where',  
          822: 'how',  
          823: 'why',  
          824: 'if',  
          825: 'then',  
          826: 'and',  
          827: 'or',  
          828: 'so',  
          829: 'that',  
          830: 'which',  
          831: 'who',  
          832: 'what',  
          833: 'when',  
          834: 'where',  
          835: 'how',  
          836: 'why',  
          837: 'if',  
          838: 'then',  
          839: 'and',  
          840: 'or',  
          841: 'so',  
          842: 'that',  
          843: 'which',  
          844: 'who',  
          845: 'what',  
          846: 'when',  
          847: 'where',  
          848: 'how',  
          849: 'why',  
          850: 'if',  
          851: 'then',  
          852: 'and',  
          853: 'or',  
          854: 'so',  
          855: 'that',  
          856: 'which',  
          857: 'who',  
          858: 'what',  
          859: 'when',  
          860: 'where',  
          861: 'how',  
          862: 'why',  
          863: 'if',  
          864: 'then',  
          865: 'and',  
          866: 'or',  
          867: 'so',  
          868: 'that',  
          869: 'which',  
          870: 'who',  
          871: 'what',  
          872: 'when',  
          873: 'where',  
          874: 'how',  
          875: 'why',  
          876: 'if',  
          877: 'then',  
          878: 'and',  
          879: 'or',  
          880: 'so',  
          881: 'that',  
          882: 'which',  
          883: 'who',  
          884: 'what',  
          885: 'when',  
          886: 'where',  
          887: 'how',  
          888: 'why',  
          889: 'if',  
          890: 'then',  
          891: 'and',  
          892: 'or',  
          893: 'so',  
          894: 'that',  
          895: 'which',  
          896: 'who',  
          897: 'what',  
          898: 'when',  
          899: 'where',  
          900: 'how',  
          901: 'why',  
          902: 'if',  
          903: 'then',  
          904: 'and',  
          905: 'or',  
          906: 'so',  
          907: 'that',  
          908: 'which',  
          909: 'who',  
          910: 'what',  
          911: 'when',  
          912: 'where',  
          913: 'how',  
          914: 'why',  
          915: 'if',  
          916: 'then',  
          917: 'and',  
          918: 'or',  
          919: 'so',  
          920: 'that',  
          921: 'which',  
          922: 'who',  
          923: 'what',  
          924: 'when',  
          925: 'where',  
          926: 'how',  
          927: 'why',  
          928: 'if',  
          929: 'then',  
          930: 'and',  
          931: 'or',  
          932: 'so',  
          933: 'that',  
          934: 'which',  
          935: 'who',  
          936: 'what',  
          937: 'when',  
          938: 'where',  
          939: 'how',  
          940: 'why',  
          941: 'if',  
          942: 'then',  
          943: 'and',  
          944: 'or',  
          945: 'so',  
          946: 'that',  
          947: 'which',  
          948: 'who',  
          949: 'what',  
          950: 'when',  
          951: 'where',  
          952: 'how',  
          953: 'why',  
          954: 'if',  
          955: 'then',  
          956: 'and',  
          957: 'or',  
          958: 'so',  
          959: 'that',  
          960: 'which',  
          961: 'who',  
          962: 'what',  
          963: 'when',  
          964: 'where',  
          965: 'how',  
          966: 'why',  
          967: 'if',  
          968: 'then',  
          969: 'and',  
          970: 'or',  
          971: 'so',  
          972: 'that',  
          973: 'which',  
          974: 'who',  
          975: 'what',  
          976: 'when',  
          977: 'where',  
          978: 'how',  
          979: 'why',  
          980: 'if',  
          981: 'then',  
          982: 'and',  
          983: 'or',  
          984: 'so',  
          985: 'that',  
          986: 'which',  
          987: 'who',  
          988: 'what',  
          989: 'when',  
          990: 'where',  
          991: 'how',  
          992: 'why',  
          993: 'if',  
          994: 'then',  
          995: 'and',  
          996: 'or',  
          997: 'so',  
          998: 'that',  
          999: 'which',  
          1000: 'who',  
          1001: 'what',
```

```
In [36]: trainseq=tok.texts_to_sequences(xtrain)
```

```
In [37]: trainseq
```

```
Out[37]: [[17,
           2,
           175,
           1,
           138,
           32,
           720,
           597,
           162,
           2386,
           2633,
           97,
           783,
           5790,
           276,
           1328,
           9148,
           852,
           1448,
           ^
```

## Inserting zeros to make all the text length same as per defined length

```
In [38]: trainpad=sequence.pad_sequences(trainseq,maxlen=lenmax)
```

```
In [39]: trainpad
```

```
Out[39]: array([[ 0,  0,  0, ..., 100, 877, 742],
                 [ 0,  0,  0, ..., 19, 135, 398],
                 [ 0,  0,  0, ..., 6, 30, 1087],
                 ...,
                 [ 0,  0,  0, ..., 1048, 233, 77],
                 [ 0,  0,  0, ..., 5387, 548, 1],
                 [ 0,  0,  0, ..., 155, 877, 742]])
```

## Performing the same preprocess options on testing data

```
In [40]: testseq=tok.texts_to_sequences(xtest)
         testpad=sequence.pad_sequences(testseq,maxlen=lenmax)
```

## Building the neural network to train the model



```
In [42]: rnn = Sequential()

rnn.add(Embedding(vocablen+1,1000, input_length=lenmax, mask_zero=True))
rnn.add(LSTM(units=32, activation="tanh"))
rnn.add(Dense(units=32, activation="relu"))
rnn.add(Dropout(0.2))

rnn.add(Dense(units=1, activation="sigmoid"))

rnn.compile(optimizer="adam", loss="binary_crossentropy")

rnn.fit(trainpad, ytrain, batch_size=512, epochs=5)

ypred = rnn.predict(testpad)

ypred = ypred>0.5
```

```
Epoch 1/5
61/61 [=====] - 201s 3s/step - loss: 0.2953
Epoch 2/5
61/61 [=====] - 206s 3s/step - loss: 0.0566
Epoch 3/5
61/61 [=====] - 209s 3s/step - loss: 0.0228
Epoch 4/5
61/61 [=====] - 208s 3s/step - loss: 0.0079
Epoch 5/5
61/61 [=====] - 212s 3s/step - loss: 0.0057
416/416 [=====] - 34s 79ms/step
```

## Checking the accuracy

```
In [43]: from sklearn.metrics import classification_report
```

```
In [44]: print(classification_report(ytest,ypred))
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.99      | 0.98   | 0.99     | 6818    |
| 1            | 0.98      | 0.99   | 0.99     | 6463    |
| accuracy     |           |        | 0.99     | 13281   |
| macro avg    | 0.99      | 0.99   | 0.99     | 13281   |
| weighted avg | 0.99      | 0.99   | 0.99     | 13281   |

```
In [ ]:
```