▾ 1. Import required library

```
import pandas as pd
import numpy as np
import nltk
import re

nltk.download('stopwords')
from nltk.corpus import stopwords
from nltk.stem.porter import PorterStemmer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.preprocessing import LabelEncoder
```

```
    [nltk_data] Downloading package stopwords to /root/nltk_data...
    [nltk_data]   Package stopwords is already up-to-date!
```

▾ 2. Read dataset and do pre-processing

─── + Code ─── + Text ───

```
df=pd.read_csv('spam.csv',encoding="ISO-8859-1")
df
```

	v1	v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
0	ham	Go until jurong point, crazy.. Available only ...	NaN	NaN	NaN
1	ham	Ok lar... Joking wif u oni...	NaN	NaN	NaN
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...	NaN	NaN	NaN
3	ham	U dun say so early hor... U c already then say...	NaN	NaN	NaN
4	ham	Nah I don't think he goes to usf, he lives aro...	NaN	NaN	NaN
...	...	...	...	...	...
5567	spam	This is the 2nd time we have tried 2 contact u...	NaN	NaN	NaN
5568	ham	Will Ì_ b going to esplanade fr home?	NaN	NaN	NaN
5569	ham	Pity, * was in mood for that. So...any other s...	NaN	NaN	NaN
5570	ham	The guy did some bitching but I acted like i'd...	NaN	NaN	NaN
5571	ham	Rofl. Its true to its name	NaN	NaN	NaN

```python
port=PorterStemmer()
cv = CountVectorizer()
dataset=[]

for i in range(0,len(df)):
  review = df['v2'][i] # Reading data
  review = re.sub('[^a-zA-Z]',' ',review) # Removing special character
  review = review.lower() # Convert capital letters into small letters
  review = review.split() # Split the input
  review = [port.stem(w) for w in review if w not in set(stopwords.words('english'))] # Stemm
  review = ' '.join(review) # join words
  dataset.append(review)

dataset
```

```
     sm',
      'free top polyphon tone call nation rate get toppoli tune sent everi week text
     subpoli per pole unsub',
      'gud mrng dear hav nice day',
      'hope enjoy game yesterday sorri touch pl know fondli bein thot great week abiola',
      'e best ur drive tmr',
      'u dogbreath sound like jan c al',
      'omg want scream weigh lost weight woohoo',
      'gener one uncount noun u dictionari piec research',
      'realli get hang around',
      'orang custom may claim free camera phone upgrad loyalti call offer end thmarch c
     appli opt availa',
      'petey boy wherear friendsar thekingshead come canlov nic',
      'ok msg u b leav hous',
      'gimm lt gt minut ago',
      'last chanc claim ur worth discount voucher today text shop savamob offer mobil cs
     savamob pobox uz sub',
      'appt lt time gt fault u listen told u twice',
      'free st week nokia tone ur mobil everi week txt nokia get txting tell ur mate www
     getz co uk pobox w wq norm p tone',
      'guarante award even cashto claim ur award call free stop getstop php rg jx',
      'k',
      'dled imp',
      'sure make sure know smokin yet',
      'boooo alway work quit',
      'take half day leav bec well',
      'ugh wanna get bed warm',
      'nervou lt gt',
      'ring come guy costum gift futur yowif hint hint',
      'congratul ur award either cd gift voucher free entri weekli draw txt music tnc www
     ldew com win ppmx age',
      'borrow ur bag ok',
      'u outbid simonwatson shinco dvd plyr bid visit sm ac smsreward end bid notif repli
     end',
      'boytoy miss happen',
      'lot use one babe model help youi bring match',
      'also bring galileo dobbi',
      'respond',
      'boo babe u enjoyin yourjob u seem b gettin well hunni hope ure ok take care
     llspeak u soonlot lovem xxxx',
```

```
 'good afternoon starshin boytoy crave yet ach fuck sip cappuccino miss babe teas
kiss',
 'road cant txt',
 'smsservic yourinclus text credit pl goto www comuk net login qxj unsubscrib stop
extra charg help comuk cm ae',
 'p alfi moon children need song ur mob tell ur txt tone chariti nokia poli chariti
poli zed profit chariti',
 'good even ttyl',
 'hmm bit piec lol sigh',
 'hahaha use brain dear',
 'hey got mail',
 'sorri light turn green meant anoth friend want lt gt worth may around',
 'thank yesterday sir wonder hope enjoy burial mojibiola',
 'u secret admir reveal think u r special call opt repli reveal stop per msg recd
cust care',
 'hi mate rv u hav nice hol messag say hello coz sent u age start drive stay road
rvx',
 'dear voucher holder claim week offer pc pleas go http www e tlp co uk expressoff
```

```
x = cv.fit_transform(dataset).toarray()
x
```

```
array([[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, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0]])
```

```
enc=LabelEncoder()
y= enc.fit_transform(df['v1'])
y
```

```
array([0, 0, 1, ..., 0, 0, 0])
```

3. Create Model

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
model = Sequential()
```

4. Add Layers (LSTM, Dense-(Hidden Layers), Output)

```
model.add(Dense(1500, activation='relu'))
model.add(Dense(3000,activation='relu'))
```

```
model.add(Dense(1, activation='sigmoid'))
```

▾ 5. Compile the Model

```
model.compile(optimizer="adam",loss='binary_crossentropy',metrics=['accuracy'])
```

▾ 6. Fit the Model

```
model.fit(x,y,epochs=10)
```

```
Epoch 1/10
175/175 [==============================] - 25s 135ms/step - loss: 0.1110 - accuracy: 0.9
Epoch 2/10
175/175 [==============================] - 22s 126ms/step - loss: 0.0074 - accuracy: 0.9
Epoch 3/10
175/175 [==============================] - 24s 135ms/step - loss: 0.0010 - accuracy: 0.9
Epoch 4/10
175/175 [==============================] - 22s 127ms/step - loss: 1.8880e-04 - accuracy
Epoch 5/10
175/175 [==============================] - 22s 127ms/step - loss: 6.7313e-05 - accuracy
Epoch 6/10
175/175 [==============================] - 22s 128ms/step - loss: 3.7799e-05 - accuracy
Epoch 7/10
175/175 [==============================] - 22s 127ms/step - loss: 2.4380e-05 - accuracy
Epoch 8/10
175/175 [==============================] - 22s 127ms/step - loss: 1.7118e-05 - accuracy
Epoch 9/10
175/175 [==============================] - 22s 127ms/step - loss: 1.2519e-05 - accuracy
Epoch 10/10
175/175 [==============================] - 22s 127ms/step - loss: 9.5173e-06 - accuracy
<keras.callbacks.History at 0x7fc247d05a50>
```

▾ 7. Save The Model

```
model.save('nlp.h5')
```

▾ 8. Test the Model

```
# Test 1
```

```python
text = 'Hi how are you'
print(text +"\n")
text = re.sub('[^a-zA-Z]',' ',text)
text = text.lower()
text = text.split()
text = [port.stem(w) for w in text if w not in set(stopwords.words('english'))]
text = ' '.join(text)
text = cv.transform([text]).toarray()
pred = model.predict(text)
print("\nResult : ",end="")
if pred>0.5:
  print('Spam')
else: print('Ham')
```

```
    Hi how are you

    1/1 [==============================] - 0s 105ms/step

    Result : Ham
```

```python
# Test 2
text='Free entry in 2 a wkly comp to win FA Cup fina'
print(text + "\n")
text = re.sub('[^a-zA-Z]',' ',text)
text = text.lower()
text = text.split()
text = [port.stem(w) for w in text if w not in set(stopwords.words('english'))]
text = ' '.join(text)
text = cv.transform([text]).toarray()
pred = model.predict(text)
print("\nResult : ",end="")
if pred>0.5:
  print('Spam')
else: print('Ham')
```

```
    Free entry in 2 a wkly comp to win FA Cup fina

    1/1 [==============================] - 0s 29ms/step

    Result : Spam
```

Colab paid products - Cancel contracts here

✓ 0s   completed at 12:19 PM   ● ✕