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

```
dataset = pd.read_csv("/content/drive/MyDrive/CNN/spam.csv",encoding='ISO-8859-1')
```

```
dataset.head(200)
```

	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
...	...	...	...	...	...
195	ham	Gud mrng dear hav a nice day	NaN	NaN	NaN
196	ham	Did u got that persons story	NaN	NaN	NaN
197	ham	is your hamster dead? Hey so tmr i meet you at...	NaN	NaN	NaN
198	ham	Hi its Kate how is your evening? I hope i can	NaN	NaN	NaN

```
from sklearn.preprocessing import LabelEncoder
var = LabelEncoder()
dataset['v1']=var.fit_transform(dataset['v1'])
```

```
values_set = dataset.iloc[:,0:2]
values_set
```

	v1	v2
0	0	Go until jurong point, crazy.. Available only ...
1	0	Ok lar... Joking wif u oni...
2	1	Free entry in 2 a wkly comp to win FA Cup fina...
3	0	U dun say so early hor... U c already then say...
4	0	Nah I don't think he goes to usf, he lives aro...

```
nltk.download('stopwords')
```

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

```
from nltk.corpus import stopwords
from nltk.stem.porter import PorterStemmer
ps = PorterStemmer()
input = []
```

```
values_set.shape
```

```
(5572, 2)
```

```
from nltk.translate.ribes_score import word_rank_alignment
from numpy.lib.shape_base import split
for i in range(0,5572):
  review = values_set['v2'][i]

  # remove the punctuation
  review = re.sub('[^a-zA-Z]',' ',review)

  # lower case
  review = review.lower()

  #split the sentence

  review = review.split()

  #remove the stopwords and stemming
  review = [ps.stem(word) for word in review if not word in set(stopwords.words('english')

  review = ' '.join(review)

  input.append(review)
```

```
input
```

```
 'hi kate love see tonight ill phone tomorrow got sing guy gave card xxx',
 'happi new year dear brother realli miss got number decid send text wish happi
abiola'
```

```
 'mean get door',
 'opinion jada kusruthi lovabl silent spl charact matur stylish simpl pl repli',
 'hmmm thought said hour slave late punish',
 'beerag',
 'import custom servic announc premier call freephon',
 'dont think turn like randomlli within min open',
 'suppos make still town though',
 'time fix spell sometim get complet diff word go figur',
 'ever thought live good life perfect partner txt back name age join mobil commun
p 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',
```

```
from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer(max_features=2000)
```

```
 X = cv.fit_transform(input).toarray()
```

```
Y = values_set['v1'].values
```

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size = 0.2)
```

```
x_train = np.reshape(x_train,(4457,2000,1))
x_train.ndim
```

```
3
```

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM,Dense
```

```
model = Sequential()
model.add(LSTM(units=30,return_sequences= True,input_shape =(x_train.shape[1],1)))
```

```
model.add(LSTM(units = 30,return_sequences=True))
model.add(LSTM(units = 30,return_sequences=True))
model.add(LSTM(units=30))
model.add(Dense(units = 1))
```

```
model.compile(optimizer = "adam",loss = "mse",metrics = "mse")
```

```
model.fit(x_train,y_train,epochs=10,batch_size = 100)
```

```
Epoch 1/10
45/45 [==============================] - 195s 4s/step - loss: 0.1162 - mse: 0.1162
Epoch 2/10
45/45 [==============================] - 188s 4s/step - loss: 0.1131 - mse: 0.1131
Epoch 3/10
45/45 [==============================] - 188s 4s/step - loss: 0.1088 - mse: 0.1088
Epoch 4/10
45/45 [==============================] - 190s 4s/step - loss: 0.1035 - mse: 0.1035
Epoch 5/10
45/45 [==============================] - 187s 4s/step - loss: 0.1012 - mse: 0.1012
Epoch 6/10
45/45 [==============================] - 188s 4s/step - loss: 0.0991 - mse: 0.0991
Epoch 7/10
45/45 [==============================] - 188s 4s/step - loss: 0.0986 - mse: 0.0986
Epoch 8/10
45/45 [==============================] - 190s 4s/step - loss: 0.0998 - mse: 0.0998
Epoch 9/10
45/45 [==============================] - 186s 4s/step - loss: 0.0983 - mse: 0.0983
Epoch 10/10
45/45 [==============================] - 186s 4s/step - loss: 0.0963 - mse: 0.0963
<keras.callbacks.History at 0x7fbcfaad13d0>
```

```
model.fit(x_train,y_train,epochs=5,batch_size = 100)
```

```
Epoch 1/5
45/45 [==============================] - 189s 4s/step - loss: 0.0984 - mse: 0.0984
Epoch 2/5
45/45 [==============================] - 190s 4s/step - loss: 0.0978 - mse: 0.0978
Epoch 3/5
45/45 [==============================] - 197s 4s/step - loss: 0.0999 - mse: 0.0999
Epoch 4/5
45/45 [==============================] - 188s 4s/step - loss: 0.0978 - mse: 0.0978
Epoch 5/5
45/45 [==============================] - 189s 4s/step - loss: 0.0974 - mse: 0.0974
<keras.callbacks.History at 0x7fbcf633ebd0>
```

```
pred = model.predict(x_train) # predicted values
error = y_train - pred
SE = error*error
MSE = SE.mean()
MSE
```

```
140/140 [==============================] - 79s 555ms/step
0.12993789101326703
```

```
np.sqrt(MSE)
```

```
0.36046898758876195
```

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

TEST

```
text = "Hello, I am michkel. what abt you? I am fine here"
text = re.sub('[^a-zA-Z]',' ',text)
text = text.lower()
```

```
text = text.strip(' ')
text = text.split()
stopwords.words('english')
ref=[word for word in text if not word in stopwords.words('english')]
```

```
text = ' '.join(ref)
ans =['ham','spam']
```

```
a = cv.transform([text])
pred1= model.predict(a)
```

```
if pred1>0.5:
```

```
    if pred[0]>0.5:
      print(ans[1])
    else:
      print(ans[0])
```

```
ham
```

Colab paid products - Cancel contracts here

0s completed at 12:46 PM

Could not connect to the reCAPTCHA service. Please check your internet connection and reload to get a reCAPTCHA challenge.