

## Assignment -4

### SMS SPAM Classification

Assignment Date	15 November 2022
Student Name	Mr.J.Sanjay
Student Roll Number	E1194031
Maximum Marks	2 Marks

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeGnjxWU58LvgPP8mDkoG#scrollTo=t9-guQg5-4Lh

+ Code + Text
RAM
Disk

Import required library

import pandas as pd
import numpy as np
import csv
import tensorflow as tf
import matplotlib.pyplot as plt

[115] from google.colab import drive
drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

Read dataset and do pre-processing

[116] import chardet
import pandas as pd
with open('/content/drive/MyDrive/spam.csv', 'rb') as f:
    result=chardet.detect(f.read())
df=pd.read_csv('/content/drive/MyDrive/spam.csv',encoding=result['encoding'])
print(df)

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

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeGnjxWU58LvgPP8mDkoG#scrollTo=t9-guQg5-4Lh

+ Code + Text
RAM
Disk

Read dataset and do pre-processing

import chardet
import pandas as pd
with open('/content/drive/MyDrive/spam.csv', 'rb') as f:
    result=chardet.detect(f.read())
df=pd.read_csv('/content/drive/MyDrive/spam.csv',encoding=result['encoding'])
print(df)

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

Unnamed: 3 Unnamed: 4
0 NaN NaN
1 NaN NaN
2 NaN NaN
3 NaN NaN
4 NaN NaN
... ..
5567 NaN NaN
5568 NaN NaN
5569 NaN NaN
5570 NaN NaN
5571 NaN NaN
```



```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text
x_train
[616, 60, 0],
[617, 60, 0],
[618, 60, 0],
[619, 60, 0],
[620, 60, 0],
[621, 60, 0],
[622, 60, 0],
[623, 60, 0],
[624, 60, 0],
[625, 60, 0],
[626, 60, 0],
[627, 60, 0],
[628, 60, 0],
[629, 60, 0],
[630, 60, 0],
[631, 60, 0],
[632, 60, 0],
[633, 60, 0],
[634, 60, 0],
[635, 60, 0],
[636, 60, 0],
[637, 60, 0],
[638, 60, 0],
[639, 60, 0],
[640, 60, 0],
[641, 60, 0],
[642, 60, 0],
[643, 60, 0],
[644, 60, 0],
[645, 60, 0],
[646, 60, 0],
[647, 60, 0],
[648, 60, 0],
0s completed at 23:43
```

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text
[1026, 60, 0],
[1027, 60, 0],
[1028, 60, 0],
[1029, 60, 0],
[1030, 60, 0],
[1031, 60, 0],
[1032, 60, 0],
[1033, 60, 0],
[1034, 60, 0],
[1035, 60, 0],
[1036, 60, 0],
[1037, 60, 0],
[1038, 60, 0],
[1039, 60, 0],
[1040, 60, 0],
[1041, 60, 0],
[1042, 60, 0],
[1043, 60, 0],
[1044, 60, 0],
[1045, 60, 0],
[1046, 60, 0],
[1047, 60, 0],
[1048, 60, 0],
[1049, 60, 0],
[1050, 60, 0],
[1051, 60, 0],
[1052, 60, 0],
[1053, 60, 0],
[1054, 60, 0],
[1055, 60, 0],
[1056, 60, 0],
[1057, 60, 0],
[1058, 60, 0],
[1059, 60, 0],
...]
```

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text
[128] y_train
[[1003, 0],
 [1004, 0],
 [1005, 0],
 [1006, 0],
 [1007, 0],
 [1008, 0],
 [1009, 0],
 [1010, 0],
 [1011, 0],
 [1012, 0],
 [1013, 0],
 [1014, 0],
 [1015, 0],
 [1016, 0],
 [1017, 0],
 [1018, 0],
 [1019, 0],
 [1020, 0],
 [1021, 0],
 [1022, 0],
 [1023, 0],
 [1024, 0],
 [1025, 0],
 [1026, 0],
 [1027, 0],
 [1028, 0],
 [1029, 0],
 [1030, 0],
 [1031, 0],
 [1032, 0],
 [1033, 0],
 [1034, 0],
 [1035, 0]]
0s completed at 23:43
```

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text
[128]
[[1025, 0],
 [1026, 0],
 [1027, 0],
 [1028, 0],
 [1029, 0],
 [1030, 0],
 [1031, 0],
 [1032, 0],
 [1033, 0],
 [1034, 0],
 [1035, 0],
 [1036, 0],
 [1037, 0],
 [1038, 0],
 [1039, 0],
 [1040, 0],
 [1041, 0],
 [1042, 0],
 [1043, 0],
 [1044, 0],
 [1045, 0],
 [1046, 0],
 [1047, 0],
 [1048, 0],
 [1049, 0],
 [1050, 0],
 [1051, 0],
 [1052, 0],
 [1053, 0],
 [1054, 0],
 [1055, 0],
 [1056, 0],
 [1057, 0],
 [1058, 0],
 [1059, 0],
 ...]]
0s completed at 23:43
```

Untitled2.ipynb - Colaboratory | AI-Assignment-4 LSTM for Text | sms spam classification lstm - Se | +

https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text

✓ [128] [1058, 0],  
[1059, 0],  
...]

{x} [129] x\_train,y\_train=np.array(x\_train),np.array(y\_train)

[130] x\_train.shape  
(1190, 3)

[131] x\_train.ndim  
2

[132] x\_train.shape[0]  
1190

[133] x\_train.shape[1]  
3

[134] p=np.arange(9).reshape(3,3,1)

<> [135] p  
array([[0],  
[1],  
[2]])

0s completed at 23:43

Untitled2.ipynb - Colaboratory | AI-Assignment-4 LSTM for Text | sms spam classification lstm - Se | +

https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text

✓ [134] p=np.arange(9).reshape(3,3,1)

{x} [135] p  
array([[0],  
[1],  
[2]],  
  
[[3],  
[4],  
[5]],  
  
[[6],  
[7],  
[8]]])

[136] x\_train=np.reshape(x\_train,(1190,3,1))

[137] x\_train  
array([[[ 60],  
[ 60],  
[ 0]],  
  
[[ 61],  
[ 60],  
[ 0]],  
  
[[ 62],  
[ 60],  
[ 0]],  
  
...])

0s completed at 23:43

Colaboratory interface showing two code cells and their outputs.

**Code Cell 1:**

```
[136] x_train=np.reshape(x_train,(1190,3,1))
```

**Output:**

```
x_train
array([[[[ 60],
          [ 60],
          [ 0]],

        [[ 61],
          [ 60],
          [ 0]],

        [[ 62],
          [ 60],
          [ 0]],

        ...,

        [[1247],
          [ 60],
          [ 0]],

        [[1248],
          [ 60],
          [ 0]],

        [[1249],
          [ 60],
          [ 0]]]])
```

**Code Cell 2:**

```
[138] x_train.shape
```

**Output:**

```
(1190, 3, 1)
```

**Code Cell 3:**

```
[139] y_train
y_train.shape
```

**Output:**

```
(1190, 2)
```

**Add Layers (LSTM, Dense-(Hidden Layers), Output)**

```
from keras.models import Sequential
from keras.datasets import mnist
from keras.layers import Dense
from keras.layers import LSTM
```

**Code Cell 4:**

```
[141] model=Sequential()
```

**Code Cell 5:**

```
[142] model.add(LSTM(units=50,input_shape=(x_train.shape[1],1),return_sequences=True))
```

**Code Cell 6:**

```
[143] x_train.shape[1]
```

**Output:**

```
3
```

**Code Cell 7:**

```
[144] model.add(LSTM(units=50,return_sequences=True))
```

**Code Cell 8:**

```
[146] model.add(LSTM(units=50))
```

```
Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Se x +
https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjso

+ Code + Text
[144] model.add(LSTM(units=50,return_sequences=True))
[146] model.add(LSTM(units=50))
[147] model.add(Dense(units=1))

Compile the Model
[148] model.compile(optimizer='adam',loss='mse')

Fit the Model
model.fit(x_train,y_train,epochs=5,batch_size=32)
Epoch 1/5
38/38 [=====] - 5s 6ms/step - loss: 271382.8125
Epoch 2/5
38/38 [=====] - 0s 7ms/step - loss: 266213.5312
Epoch 3/5
38/38 [=====] - 0s 7ms/step - loss: 264322.8438
Epoch 4/5
38/38 [=====] - 0s 7ms/step - loss: 262842.5000
Epoch 5/5
38/38 [=====] - 0s 7ms/step - loss: 261484.1875
<keras.callbacks.History at 0x7faf5704f8d0>

Save The Model
[150] model.save('spam.h5')

[152] with open('/content/drive/MyDrive/spam.csv','rb') as f:
      result=chardet.detect(f.read())

dataset_test=pd.read_csv('/content/drive/MyDrive/spam.csv',encoding=result['encoding'])
print(dataset_test)

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

Unnamed: 3 Unnamed: 4
0 NaN NaN
1 NaN NaN
2 NaN NaN
3 NaN NaN
4 NaN NaN
... .. NaN
5567 NaN NaN
5568 NaN NaN
5569 NaN NaN
```





Untitled2.ipynb - Colaboratory

AI-Assignment-4 LSTM for Text

sms spam classification lstm - Se

+

https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjs0

+ Code + Text

RAM

Disk

[156] dataset\_test.shape

(5572, 5)

[157] dataset\_total=pd.concat((df["v1"],dataset\_test["v1"]))

[159] dataset\_total.shape

(11144,)

dataset\_total

0	ham
1	ham
2	spam
3	ham
4	ham
...	...
5567	spam
5568	ham
5569	ham
5570	ham
5571	ham

Name: v1, Length: 11144, dtype: object

[161] inputs=dataset\_total[len(dataset\_total)-len(dataset\_test)-60:].values

[166] inputs=inputs.reshape(1,-2)

[163] inputs

0s completed at 23:43

Untitled2.ipynb - Colaboratory

AI-Assignment-4 LSTM for Text

sms spam classification lstm - Se

+

https://colab.research.google.com/drive/1-21UthSsDtTeJGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjs0

+ Code + Text

RAM

Disk

[166] inputs=inputs.reshape(1,-2)

[167] inputs

array([[ 'ham', 'ham', 'ham', ..., 'ham', 'ham', 'ham']], dtype=object)

[168] inputs.shape

(1, 5632)

Test The Model

x\_test=[]

for i in range(60,80):

x\_test.append(inputs[i-60:i,0])

[170] x\_test=np.array(x\_test)

/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:1: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of objects with ragged shapes).

[171] x\_test.shape

(20,)

[172] x\_test=np.reshape(x\_test,(20,1))

0s completed at 23:43

Untitled2.ipynb - Colaboratory x AI-Assignment-4 LSTM for Text x sms spam classification lstm - Sc x +

https://colab.research.google.com/drive/1-21UthSsDtTeIGnjxWU58LvgPP8mDkoG#scrollTo=umrc4knyAjs0

+ Code + Text

array([[ 'ham', 'ham', 'ham', ..., 'ham', 'ham', 'ham']], dtype=object)

[168] inputs.shape

(1, 5632)

Test The Model

[169] x\_test=[]  
for i in range(60,80):  
x\_test.append(inputs[i-60:i,0])

x\_test=np.array(x\_test)

/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:1: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of objects that may be ragged) is deprecated. It should instead be created from an array of objects that has been pre-allocated to the correct shape.

[171] x\_test.shape

(20,)

[172] x\_test=np.reshape(x\_test,(20,1))

[173] x\_test.shape

(20, 1)

0s completed at 23:43