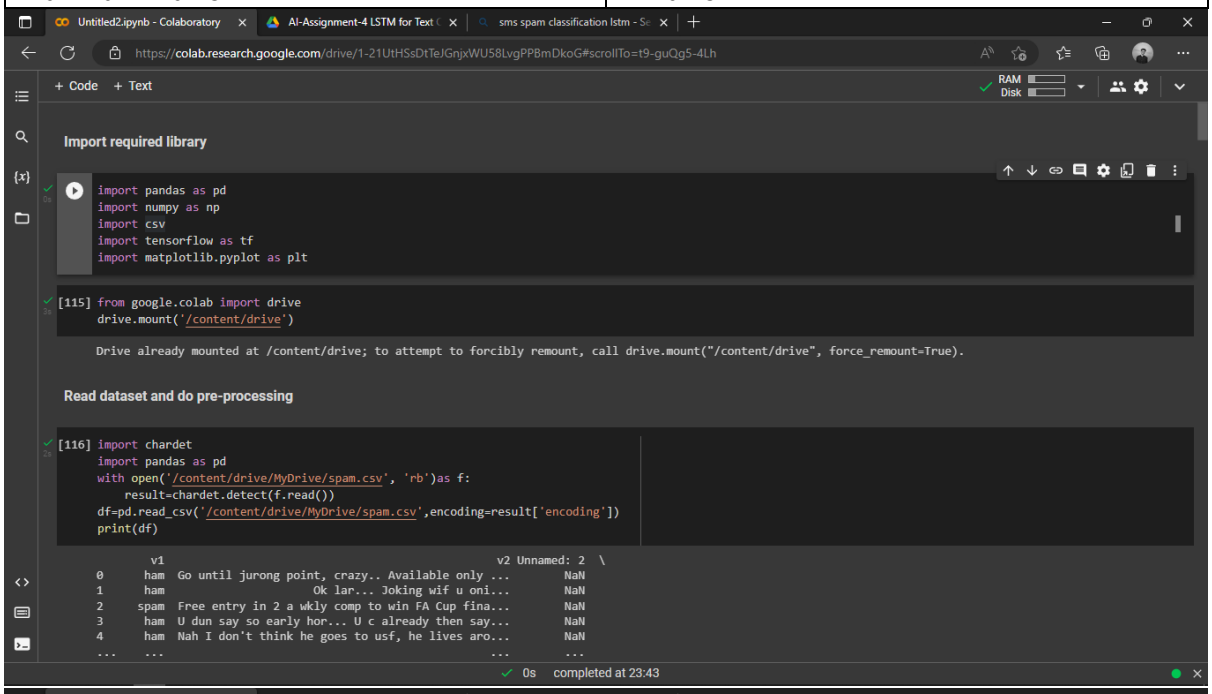
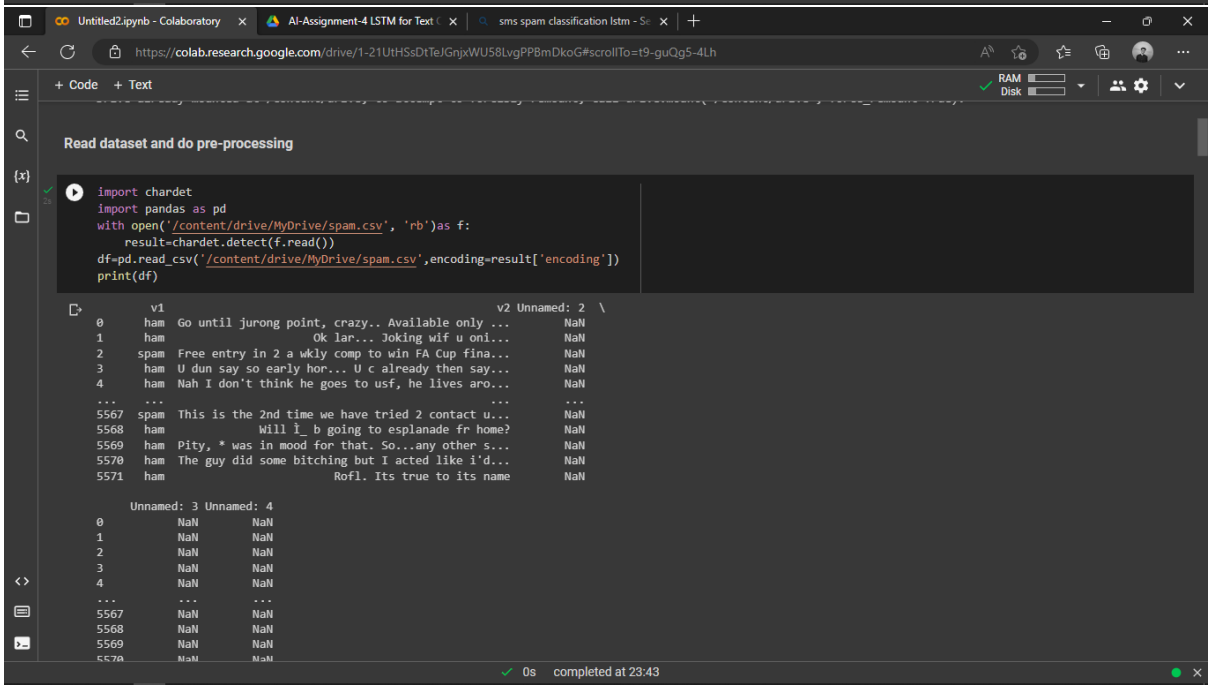
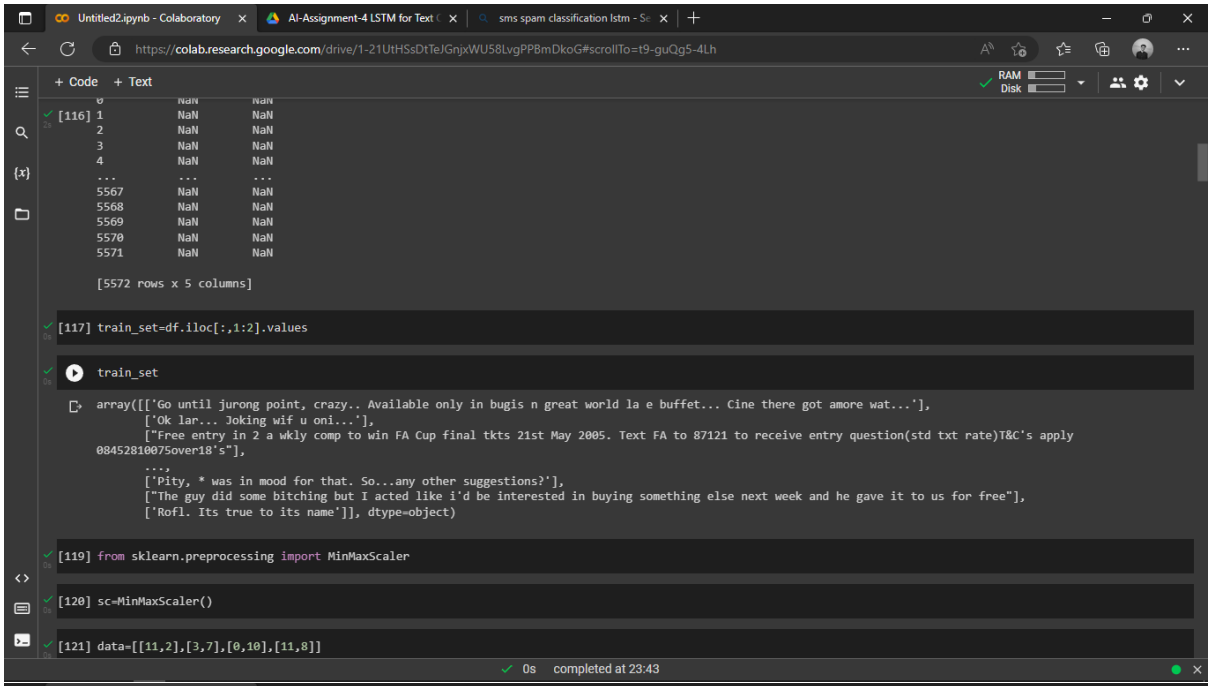
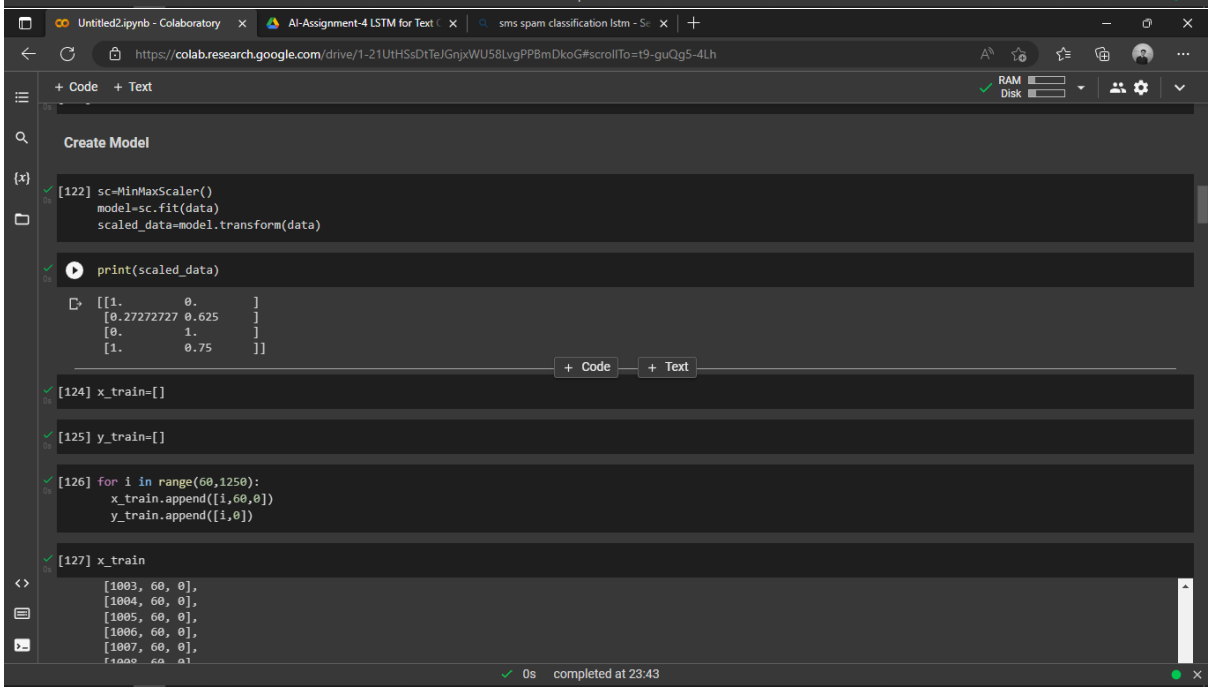
Assignment -4
SMS SPAM Classification

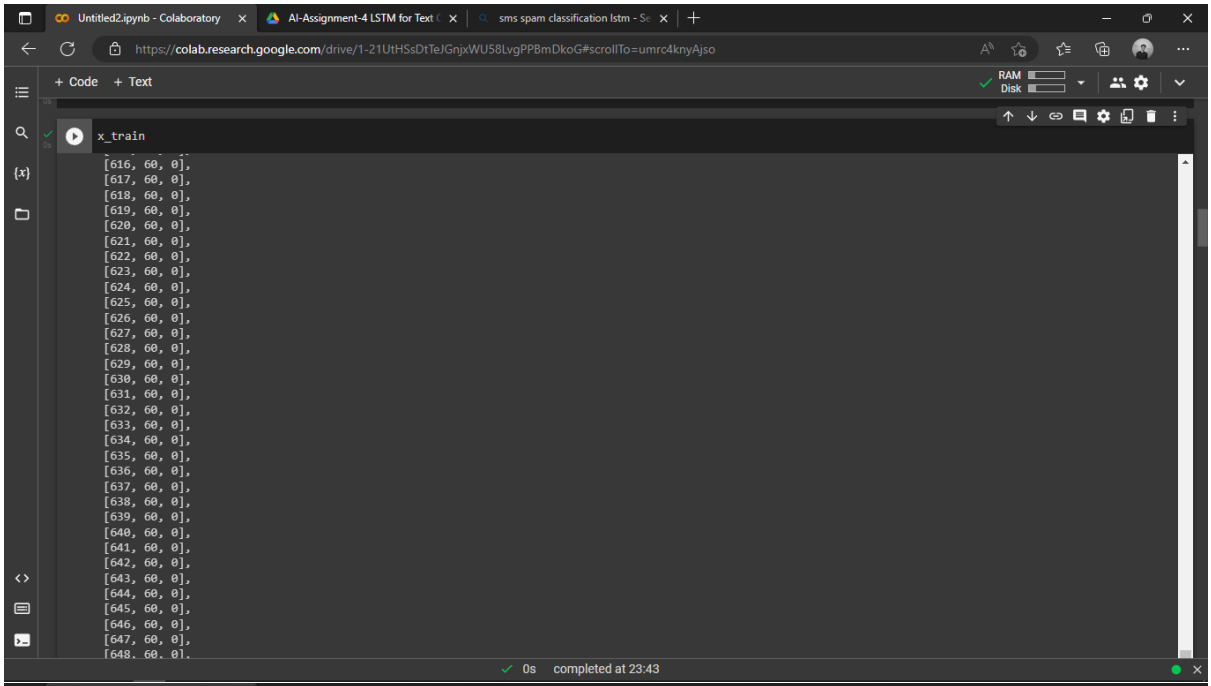
Assignment Date	13 November 2022
Student Name	Mr.C.Senthil Murugan
Student Roll Number	E1194033
Maximum Marks	2 Marks

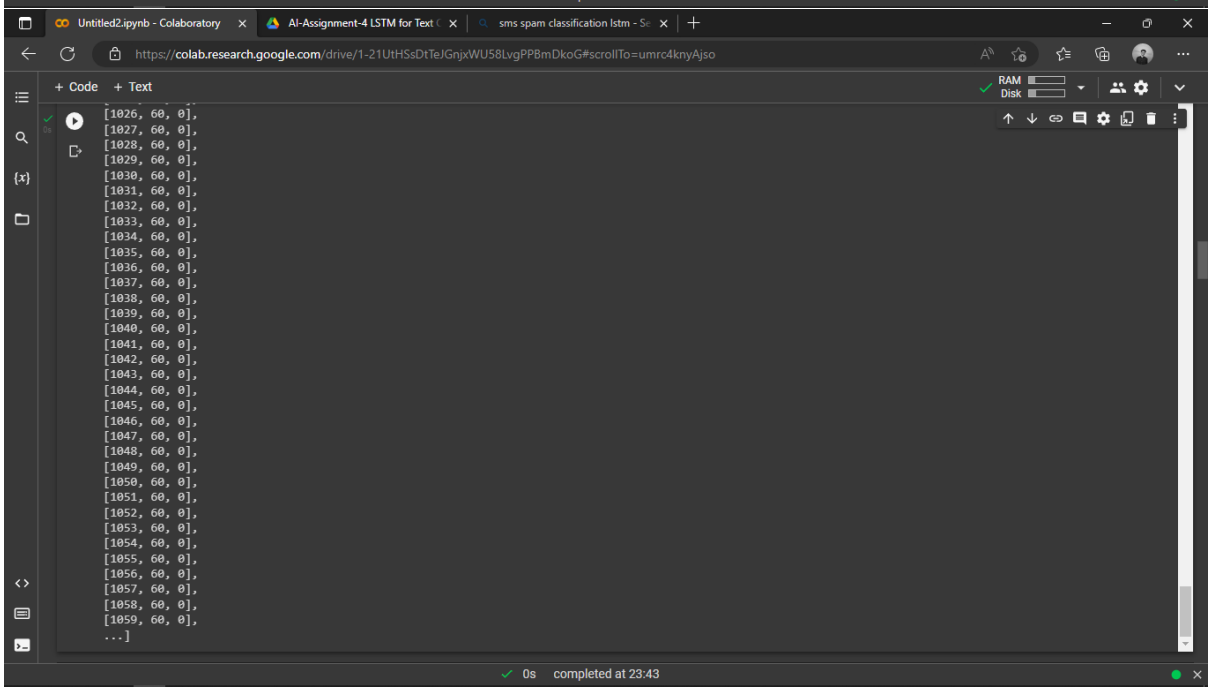


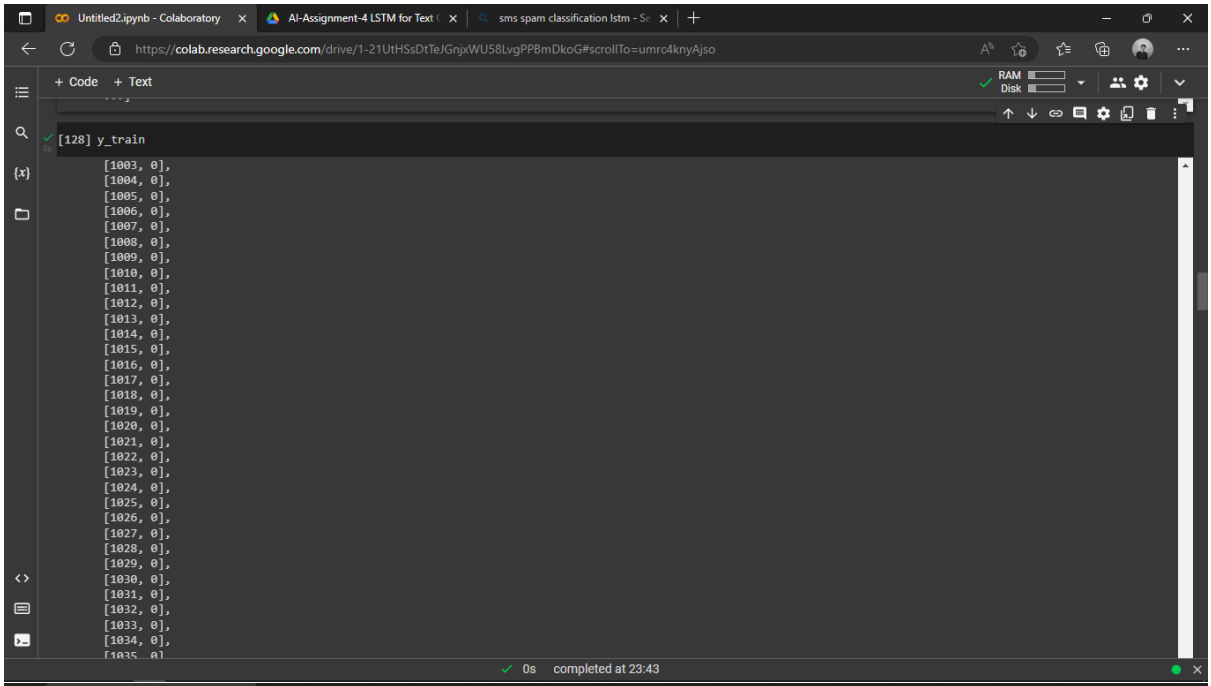


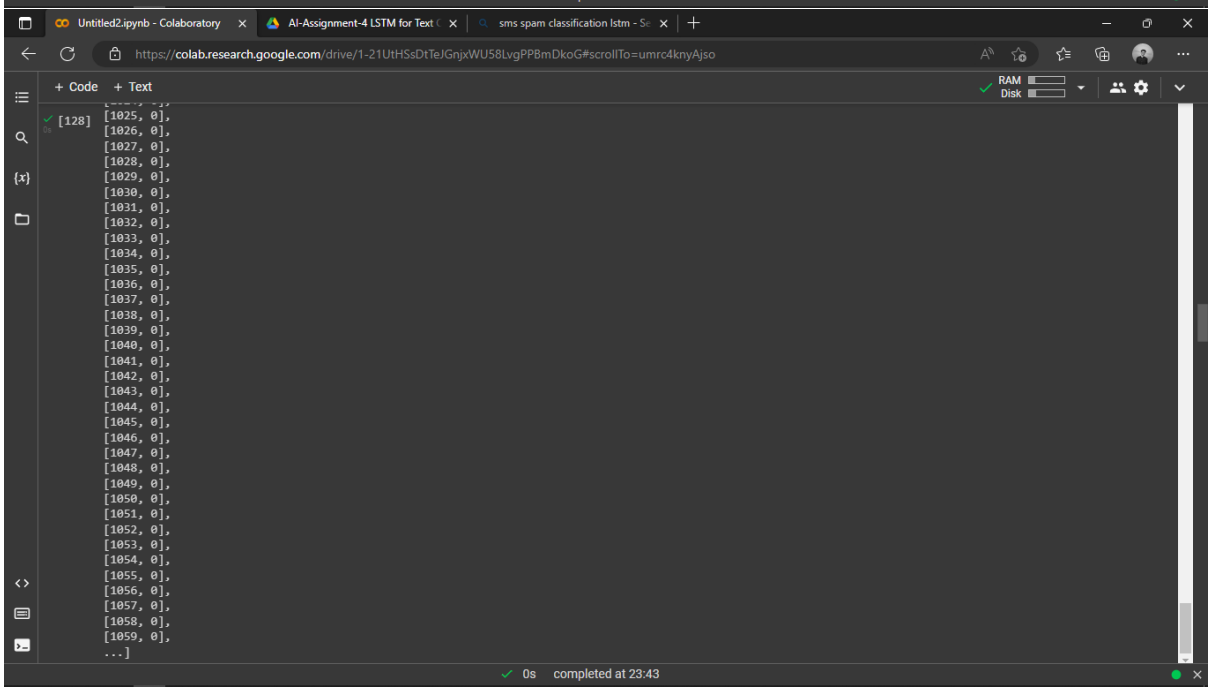

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=t9-guQg5-4Lh
+ Code + Text
RAM
Disk
[116] 1 NaN NaN
2 NaN NaN
3 NaN NaN
4 NaN NaN
... ...
5567 NaN NaN
5568 NaN NaN
5569 NaN NaN
5570 NaN NaN
5571 NaN NaN
[5572 rows x 5 columns]
[117] train_set=df.iloc[:,1:2].values
train_set
array([['Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...'],
['Ok lar... Joking wif u oni...'],
["Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)T&C's apply
08452810075over18's"],
...,
['Pity, * was in mood for that. So...any other suggestions?'],
["The guy did some bitching but I acted like i'd be interested in buying something else next week and he gave it to us for free"],
['Rofl. Its true to its name']], dtype=object)
[119] from sklearn.preprocessing import MinMaxScaler
[120] sc=MinMaxScaler()
[121] data=[[11,2],[3,7],[0,10],[11,8]]
0s completed at 23:43

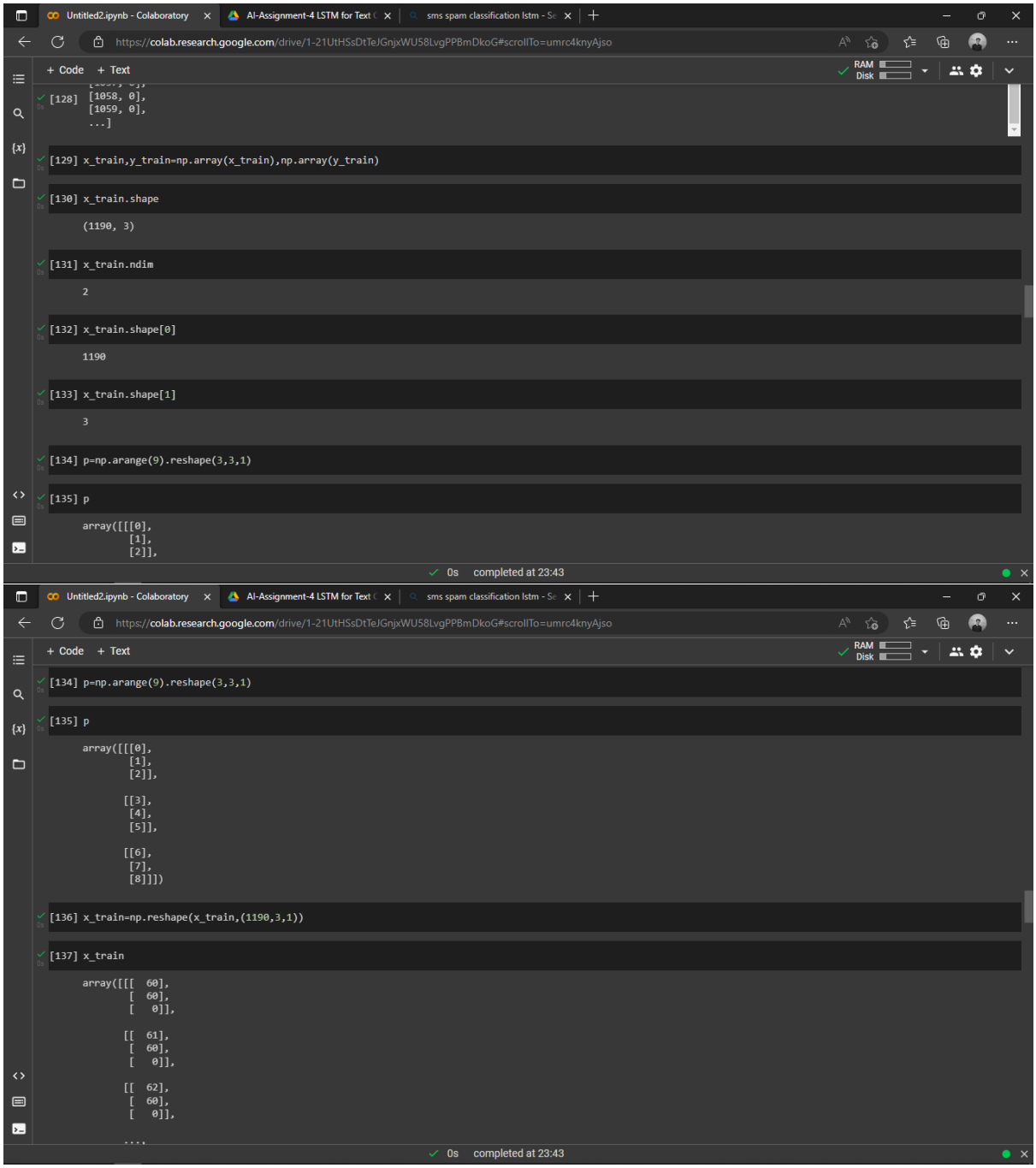

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=t9-guQg5-4Lh
+ Code + Text
RAM
Disk
Create Model
[122] sc=MinMaxScaler()
model=sc.fit(data)
scaled_data=model.transform(data)
print(scaled_data)
[[1. 0.]
[0.27272727 0.625]
[0. 1.]
[1. 0.75]]
+ Code + Text
[124] x_train=[]
[125] y_train=[]
[126] for i in range(60,1250):
x_train.append([i,60,0])
y_train.append([i,0])
[127] x_train
[1003, 60, 0],
[1004, 60, 0],
[1005, 60, 0],
[1006, 60, 0],
[1007, 60, 0],
0s completed at 23:43

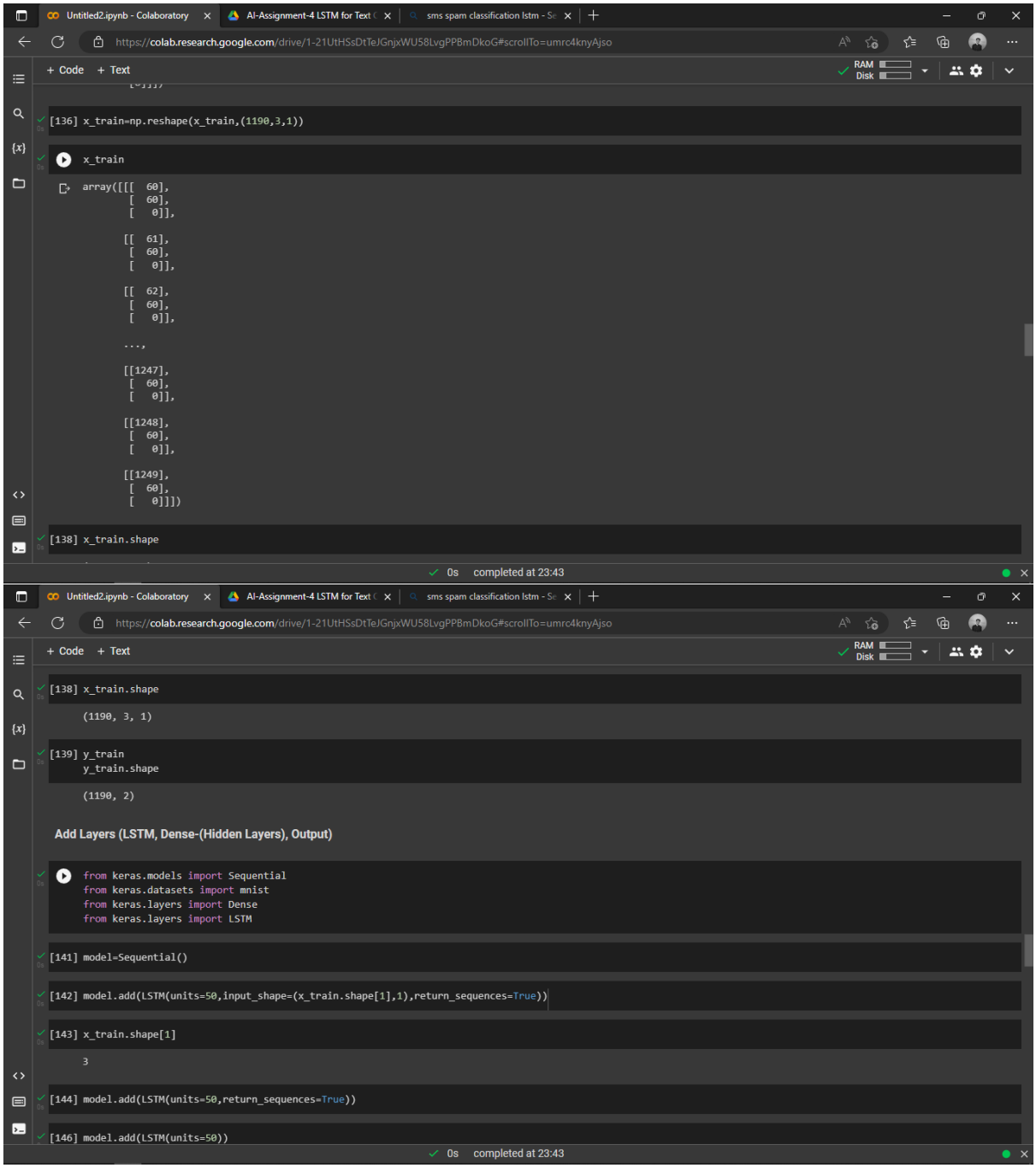

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
x_train
[616, 60, 0],
[617, 60, 0],
[618, 60, 0],
[619, 60, 0],
[620, 60, 0],
[621, 60, 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],
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[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],
...]
0s completed at 23:43

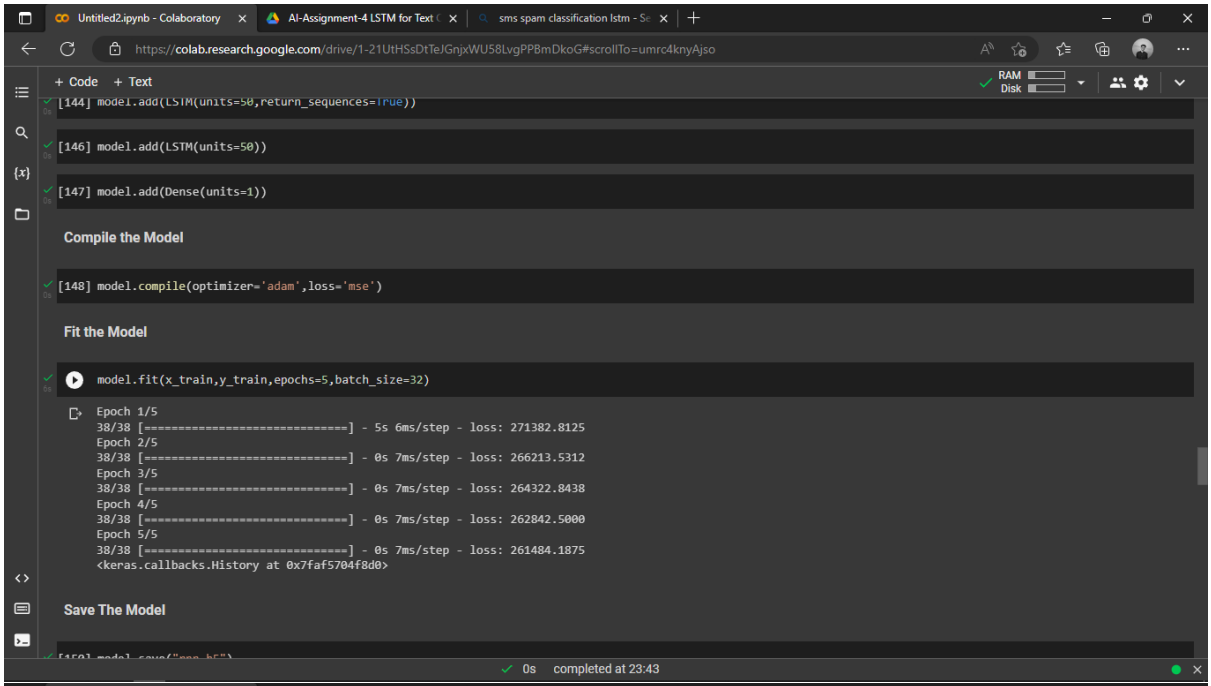

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[128] y_train
[1003, 0],
[1004, 0],
[1005, 0],
[1006, 0],
[1007, 0],
[1008, 0],
[1009, 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],
0s completed at 23:43

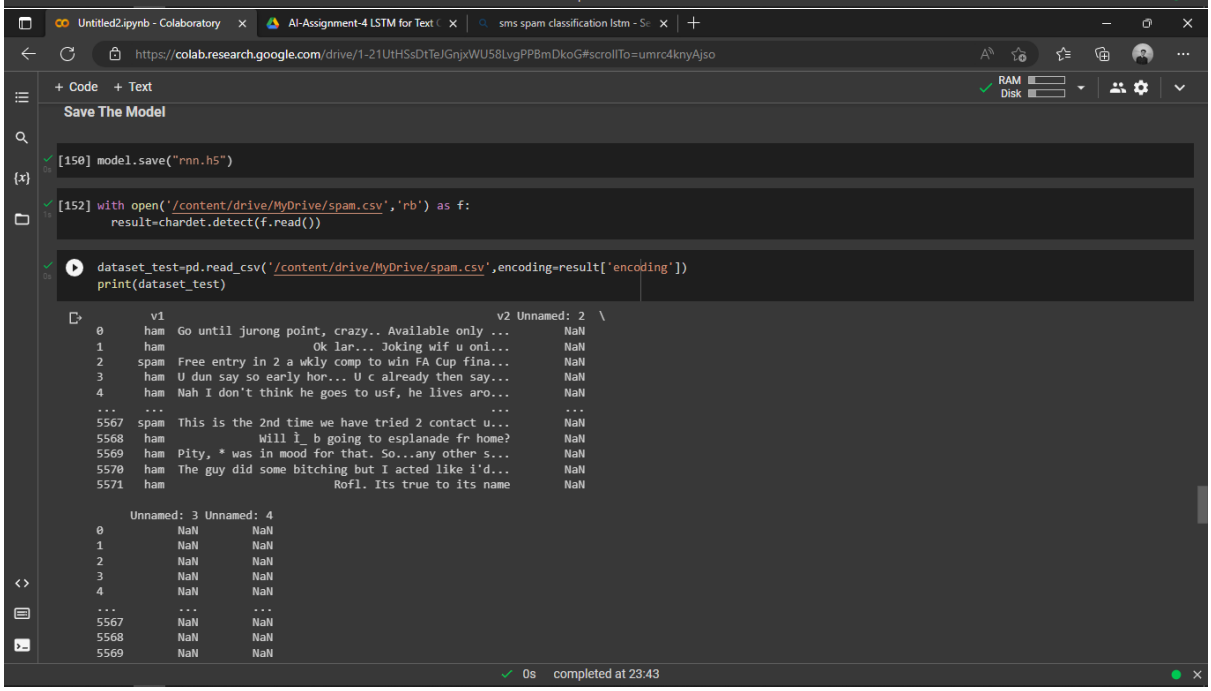

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[128]
[1025, 0],
[1026, 0],
[1027, 0],
[1028, 0],
[1029, 0],
[1030, 0],
[1031, 0],
[1032, 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

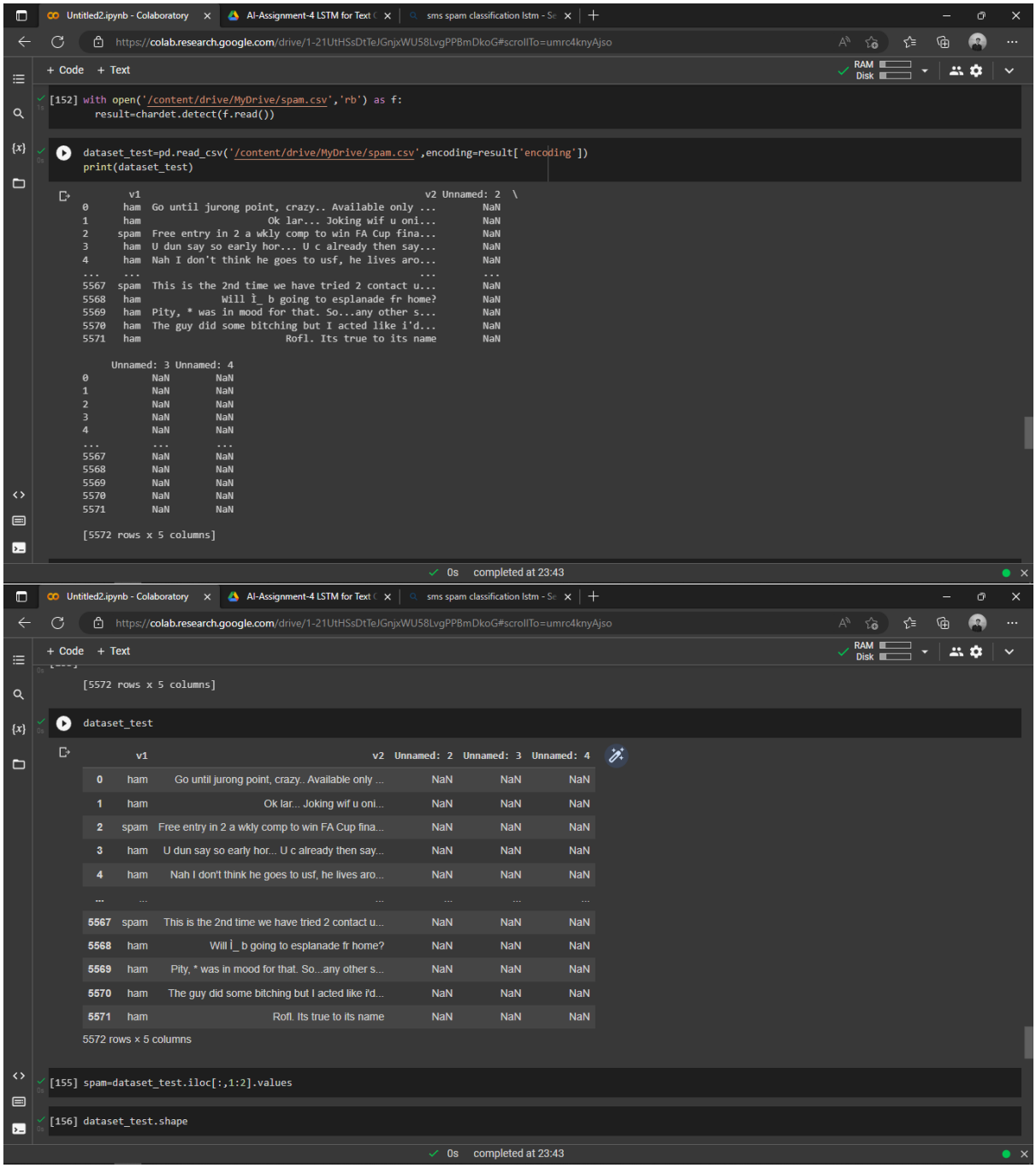

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[128] [1058, 0],
[1059, 0],
...]
[129] x_train,y_train=np.array(x_train),np.array(y_train)
[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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[134] p=np.arange(9).reshape(3,3,1)
[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


Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[136] x_train=np.reshape(x_train,(1190,3,1))
x_train
array([[[60],
[60],
[0]],
[[61],
[60],
[0]],
[[62],
[60],
[0]],
...,
[[1247],
[60],
[0]],
[[1248],
[60],
[0]],
[[1249],
[60],
[0]]])
[138] x_train.shape
0s
completed at 23:43

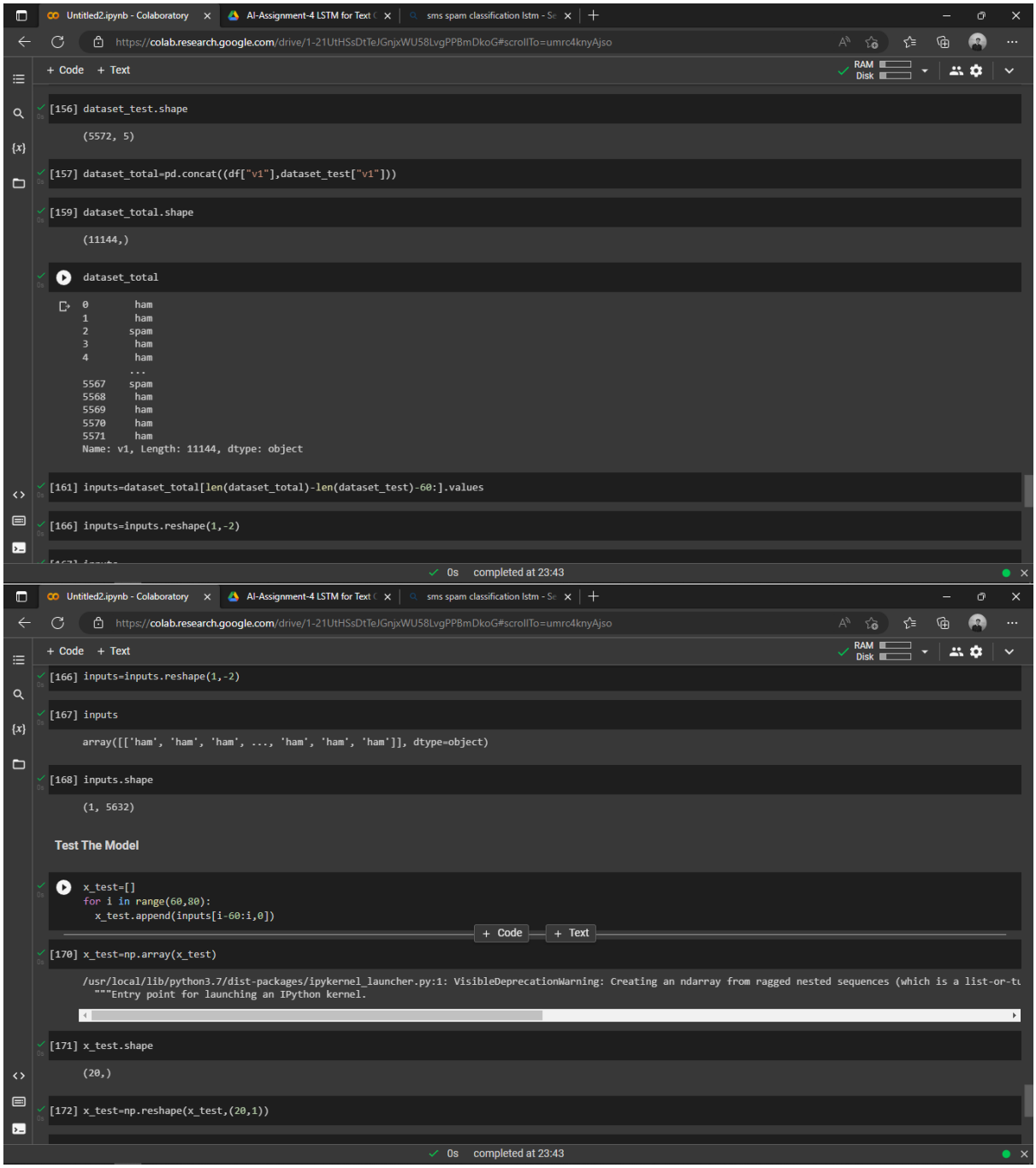
Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[138] x_train.shape
(1190, 3, 1)
[139] y_train
y_train.shape
(1190, 2)
Add Layers (LSTM, Dense-(Hidden Layers), Output)
from keras.models import Sequential
from keras.datasets import mnist
from keras.layers import Dense
from keras.layers import LSTM
[141] model=Sequential()
[142] model.add(LSTM(units=50,input_shape=(x_train.shape[1],1),return_sequences=True))
[143] x_train.shape[1]
3
[144] model.add(LSTM(units=50,return_sequences=True))
[146] model.add(LSTM(units=50))
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[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
0s completed at 23:43

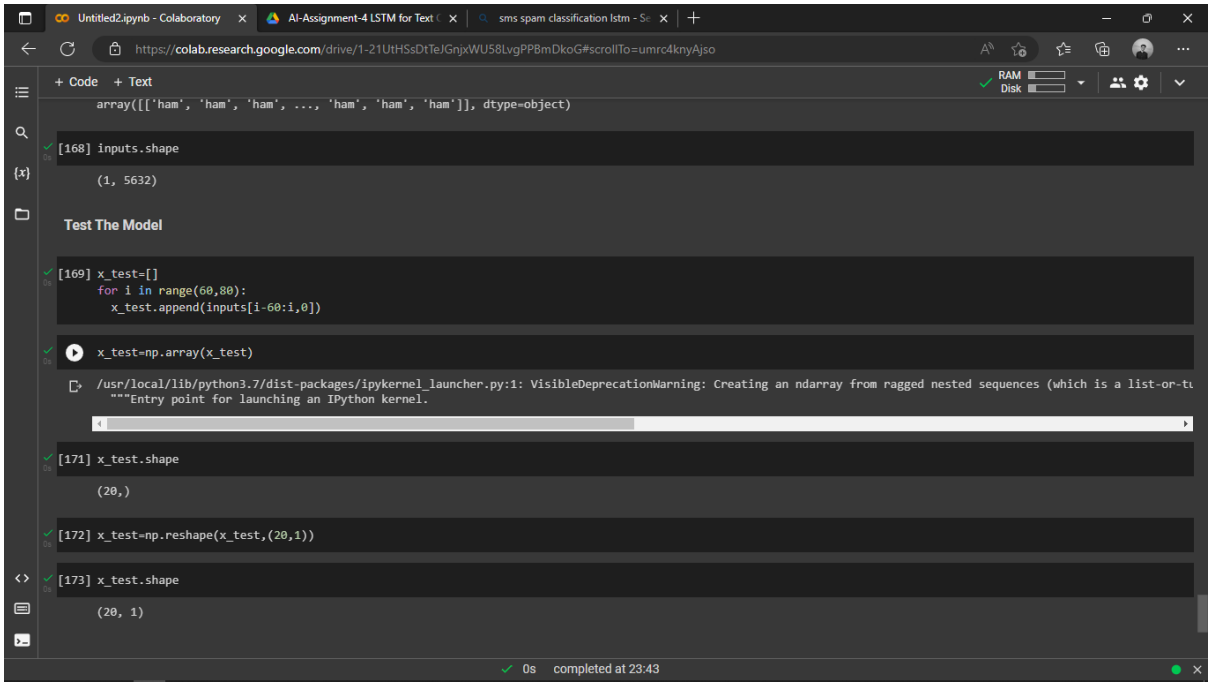

Untitled2.ipynb - Colaboratory
AI-Assignment-4 LSTM for Text
sms spam classification lstm - Se
https://colab.research.google.com/drive/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
Save The Model
[150] model.save("rnn.h5")
[152] with open('/content/drive/MyDrive/spam.csv','rb') 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
... ...
5567 spam This is the 2nd time we have tried 2 contact u... NaN
5568 ham Will Ì_ 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
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[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
... ...
5567 spam This is the 2nd time we have tried 2 contact u... NaN
5568 ham Will Ì_ 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
[5572 rows x 5 columns]
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
[5572 rows x 5 columns]
dataset_test
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
5572 rows × 5 columns
[155] spam=dataset_test.iloc[:,1:2].values
[156] dataset_test.shape
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ 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)
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ 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])
+ Code + Text
[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-tu
"""Entry point for launching an IPython kernel.
[171] x_test.shape
(20,)
[172] x_test=np.reshape(x_test,(20,1))
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/1-21UtHSsDtTeJGnjxWU58LvgPPBmDkoG#scrollTo=umrc4knyAjso
+ Code
+ Text
RAM
Disk
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-tu
"""Entry point for launching an IPython kernel.
[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