Assignment - 4

Python Programming

Assignment Date	18-10-2022
Student Name	SURESH CHINMAYE
Student Roll Number	1 I 1519104155
Maximum Marks	2 Marks

EuestiOn 1:

Import the necessary libraries

Solution:

```
inprt pandas as pd
inport   nuapy  as  rip
import matp1ot1ib.pyplot as pit
1mport  seaborn  as  sns
from sklearn.mode1_se1ect1on 1mpor£ tra1n_test_sp11t
fcom  sk1earn .p r e¿ar oc es »1ng  1mpoyt   Labe1Encoder
from  keras.models  1aport  t1Dde1
from  kera s . 1ayer s  1mporf  L5Tfd,  Act1x*aQ1on,  Den se,  Dropout,  Inp1a6,  Embedd1ng
      ke ras.opt 1n1zers  ' nport RFtSprop
fFDm  ke ras.preproces sing. text  imporl  Token1»er
f.I-om ke ras.preproCRs sing  1mpot•t  sequence
f.rom ke ras . ut1ls  1morf  pad_sequences
fPDm  keras. ut11s  1mport  to_categor1ca 1
from  ke ras. caI1backs  iapoi%  Early5topp1ng
```

Question 2:

Download the Dataset

Solution:

Dataset Downloaded and uploaded to dnve https://www.kaggle.com/code/kredy10/simple-lstm-for-text-clas.xification/data

Question 3:

Read dataset and do pre-processing

Solution:

Read dataset

```
df = pd.read_csv('›'content/'drive, yDri'›e/spam.csv',delimiter=',',en coding='latin-1'j
df . head( )
```

	v1	v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
0	ham	Go until yurong 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 ie ‹vin FA Cup fina...	NaN	NaN	NaN
3	ham	U dun say so early hor... U c already then say..,	NaN	NaN	NaN
4	ham	Nah1 don't think he gses to usf, he lives aro...	NaN	NaN	NaN

Pre-processing the Dataset

```
df. drop ( ['Unnamed: 2”, 'unnamed: 3 ', unnamed: 4 '] , ax1s =1, inp1ace=True)
df. 1nfo ( )
```

```
Range1ndex: 5572 entri es, 6 to 5571
Data columns (total 2 columns ) :
 #    Column hon-Null Count    Dtype

 e    •i        s572 non-null    object
 1    v2        5572 non-null    object
dtypes: object(2)
memory usage: 87.2+ KB
```

```
X    df. v2
Y    df . v1
1e • Labe1Encoder ( )
v • 1c. f1k_transfom(Y J
Y    ¥.reshape{-1,1)
```

```
X_traln,x_test,'r_traln, Y_test • tra1n_tcst_cptGt 'X, Y, test_s1ze•0.15 j
```

```
eex_1en - lsa
tok • 7oken1zer (nuo_nords •eex_words ’›
tok.fit_on_texts(X_train)
sequences • tok . texts_to_scquenCes ( x_traLn )
sequences_eatrI x • ped_sequencec ( sequences ,aox Len• esx_l en)
```

```
Inputs • InpuE {nane- " i nput s ' , shape• [ cax_1cn ] j
layer • Enbcdd1ng(mgx_words,50,1nput_1en@h•wax_1en)(InputsI
1syer • LSTH{6A) ( 1ayen )
layer • Ocnse (2S6, naee• FCI ) ( 1ayer)
1eyer - Act Cvet ton( *rr 1u' ) I 1eyer)
layer - Dropout (o. s \ (1eyer I
1ayer - Dence ¿1, nane• "sut_1ayer’ ) (1ayer j
layer • Ac t I va t1on ( sig•o l d ) { layer )
atode1 - fJode I [ Input s•Input s , output s • 1ayer)
```

wode1: *——•1 1*

Layer (type)	&tpux Shape	Psra•s
1nput s (InputLayec)	[(None, 15e)]	0
eebedd1n$ 1 (E iedd1ng)	(lone, 150, 59)	50000
1sM l (LSW!)	(None, 64)	29440
FCf (Oense)	(Mone, 2M)	1664B
activation_2 (Activation)	(None, 256)	0
drspout_I {Dcopou6)	(None, 256)	0
out_lmyer {Oense)	(None, 1)	2S7
act1vet1on_3 { Act tvst on)	(None, 1)	0

Tgte1 96,337
Trs Lnab1e pcrmis : 96, 337

In [28]:

```
model.fit(sequences_matrix,Y_train,batch_size=128,epochs=10,
          validation_split=0.2)
```

Epoch 1/10
30/30 [=] - lbs 26A/st/ s£0p \ 055 - 6.3182 6CC uP6Cy ‹ ?. 0788 - v6T_ Tc•ss : O. 15?t - 'v0 J_6£ 6uP dC y : 9. 9715 36 '
38 [- - 7s 247•s/ step - Nos s‹ 6. B805 - eccur•c y › O. 97bd - e1_Loss : 8.6742 - vaS_ec curecy: 8. 9778
30/30 [=) - 7 237•s/•i*$ loss: 0.$wO& - •‹<••x<y. $,9O$a *l_lo*x. e.08?g - •*i *cc«r <y. o.e$Oi
E poc h J/ 1$t
30/30 [=] - 7s 245ms/step - loss: 0.0272 - accuracy: 0.9929 - val_loss: 0.0806 - val_accuracy: 0.9778
3e ' 3g (• • • • • •] • 7s 242ss.• step • I ass : B. az 28 - oc < ura-c y: O.e'9 47- '.s I Los i : G.T8Ze • v•I p•c cu r•c y * e. est
E poc h 6 'IN
36a 3-6 [• •] 7s >^^=•z step Nos s. 6. B178 •c cur•cy 6i.995'S vB} loss : o.e*gs - v4} accurpc y - e. e ss
3Gy SO (--] • 7t 241•s/ atep • loss i 6. B158 - •c cura‹y : 6. 995g • ma L_Lost : 0 .0g69 • ve1_ac curac y : 6. 99Gg
E poC h d.•' } 0'
30/30 [==] - 7s 241ms/step - loss: 0.0162 - accuracy: 0.9958 - val_loss: 0.0901 - val_accuracy: 0.9768
36.' 3e (••••••••• ! - 7s *A6a*/ step loss ‹ 8. 059 - secured y : 0. M68 •-s1_1oas . e. us4 - vaJ_eccur ac y : 6. 9769
§poCh 167 16
30/30 [] - 7s 247ms/step - loss: 0.0355 - accuracy: 0.9905 - val_loss: 0.1264 - val_accuracy: 0.9726

Savs Th« Uodef

«›ode1. save ('sus c1assi l se<.h5')

6hs$«ocescIng ths Teat Dataaat

test_sequrnces • tok . trxt s_to_sequences| x_test ›
test_seqvences_astr1x pad_ scquences { t e st_xcquencrs ,max ten =a ax ten)

Testing ths Modal

accr t>>de1. evat uatr i test_sequence s_aatrkx , Y_Eest)

27a' 27 [• -• •• - - •-- - - - --- - • - • • - - •••---•] - 1 s 2Ws 's l e p - l as s: 6. 66B6 - ac c u•ac y : e. 982 l

pctn¥ • Te s t s et ' n Los-t [: 6. 37)*.n Accu acy : : e. *l ' . fo••ost l acc r [e] , ac cr | y] T)

Test set
cor›: e.ess
dcc u^ acy : 0. 9B2