{"cells":[{"cell_type":"markdown","metadata":{"_cell_guid":"79c7e3d0-c299-4dcb-8224-4455121ee9b0","_uuid":"d629ff2d2480ee46fbb7e2d37f6b5fab8052498a"},"source":["### Import the necessary libraries"]},{"cell_type":"code","execution_count":1,"metadata":{"_uuid":"d6fb32fd69316596e236eab5fb8cf77c848508c3","execution":{"iopub.execute_input":"2022-10-21T09:50:52.962404Z","iopub.status.busy":"2022-10-21T09:50:52.961986Z","iopub.status.idle":"2022-10-21T09:50:56.841064Z","shell.execute_reply":"2022-10-21T09:50:56.840150Z","shell.execute_reply.started":"2022-10-21T09:50:52.962263Z"},"trusted":true},"outputs":[{"name":"stderr","output_type":"stream","text":["Using TensorFlow backend.\n"]}],"source":["import pandas as pd\n","import numpy as np\n","import matplotlib.pyplot as plt\n","import seaborn as sns\n","from sklearn.model_selection import train_test_split\n","from sklearn.preprocessing import LabelEncoder\n","from keras.models import Model\n","from keras.layers import LSTM, Activation, Dense, Dropout, Input, Embedding\n","from keras.optimizers import RMSprop\n","from keras.preprocessing.text import Tokenizer\n","from keras.preprocessing import sequence\n","from keras.utils import to_categorical\n","from keras.callbacks import EarlyStopping\n","%matplotlib inline"]},{"cell_type":"markdown","metadata":{"_uuid":"f674695f1742479cefdeec0e81ab469f7b6ec90f"},"source":["### Download Dataset and Load into Dataframe"]},{"cell_type":"code","execution_count":2,"metadata":{"_uuid":"aca2f1d9da3f35d104763166fe4d25448410d8f2","execution":{"iopub.execute_input":"2022-10-21T09:51:16.165499Z","iopub.status.busy":"2022-10-21T09:51:16.165022Z","iopub.status.idle":"2022-10-21T09:51:16.221908Z","shell.execute_reply":"2022-10-21T09:51:16.220894Z","shell.execute_reply.started":"2022-10-21T09:51:16.165454Z"},"trusted":true},"outputs":[{"data":{"text/html":["<div>\n","<style>\n"," .dataframe thead tr:only-child th {\n"," text-align: right;\n"," }\n","\n"," .dataframe thead th {\n"," text-align: left;\n"," }\n","\n"," .dataframe tbody tr th {\n"," vertical-align: top;\n"," }\n","</style>\n","<table border=\"1\" class=\"dataframe\">\n"," <thead>\n"," <tr style=\"text-align: right;\">\n"," <th></th>\n"," <th>v1</th>\n"," <th>v2</th>\n"," <th>Unnamed: 2</th>\n"," <th>Unnamed: 3</th>\n"," <th>Unnamed: 4</th>\n"," </tr>\n"," </thead>\n"," <tbody>\n"," <tr>\n"," <th>0</th>\n"," <td>ham</td>\n"," <td>Go until jurong point, crazy.. Available only ...</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," </tr>\n"," <tr>\n"," <th>1</th>\n"," <td>ham</td>\n"," <td>Ok lar... Joking wif u oni...</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," </tr>\n"," <tr>\n"," <th>2</th>\n"," <td>spam</td>\n"," <td>Free entry in 2 a wkly comp to win FA Cup fina...</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," </tr>\n"," <tr>\n"," <th>3</th>\n"," <td>ham</td>\n"," <td>U dun say so early hor... U c already then say...</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," </tr>\n"," <tr>\n"," <th>4</th>\n"," <td>ham</td>\n"," <td>Nah I don't think he goes to usf, he lives aro...</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," <td>NaN</td>\n"," </tr>\n"," </tbody>\n","</table>\n","</div>"],"text/plain":[" v1 v2 Unnamed: 2 \\\n","0 ham Go until jurong point, crazy.. Available only ... NaN \n","1 ham Ok lar... Joking wif u oni... NaN \n","2 spam Free entry in 2 a wkly comp to win FA Cup fina... NaN \n","3 ham U dun say so early hor... U c already then say... NaN \n","4 ham Nah I don't think he goes to usf, he lives aro... NaN \n","\n"," Unnamed: 3 Unnamed: 4 \n","0 NaN NaN \n","1 NaN NaN \n","2 NaN NaN \n","3 NaN NaN \n","4 NaN NaN "]},"execution_count":2,"metadata":{},"output_type":"execute_result"}],"source":["df = pd.read_csv('../input/spam.csv',delimiter=',',encoding='latin-1')\n","df.head()"]},{"cell_type":"markdown","metadata":{},"source":["### Data Analysis"]},{"cell_type":"markdown","metadata":{"_uuid":"53083ccecf39523cff290495a6cc768061b

a9b46"},"source":["Drop the columns that are not required for the neural network."]},{"cell_type":"code","execution_count":3,"metadata":{"_uuid":"95a8b5d6f19cf42d4f55c6d
2842faf1d0d55c1d0","execution":{"iopub.execute_input":"2022-10-
21T09:51:26.379846Z","iopub.status.busy":"2022-10-
21T09:51:26.379480Z","iopub.status.idle":"2022-10-
21T09:51:26.396044Z","shell.execute_reply":"2022-10-
21T09:51:26.395172Z","shell.execute_reply.started":"2022-10-
21T09:51:26.379761Z"},"scrolled":true,"trusted":true},"outputs":[{"name":"stdout","output_type":"str
eam","text":["<class 'pandas.core.frame.DataFrame'>\n","RangeIndex: 5572 entries, 0 to
5571\n","Data columns (total 2 columns):\n","v1   5572 non-null object\n","v2   5572 non-null
object\n","dtypes: object(2)\n","memory usage: 87.1+ KB\n"]}],"source":["df.drop(['Unnamed: 2',
'Unnamed: 3', 'Unnamed:
4'],axis=1,inplace=True)\n","df.info()"]},{"cell_type":"code","execution_count":4,"metadata":{"_uuid":"
a12002f521dd8eaeb0f69a932cbf23815ffd09d7","execution":{"iopub.execute_input":"2022-10-
21T09:51:32.380579Z","iopub.status.busy":"2022-10-
21T09:51:32.380202Z","iopub.status.idle":"2022-10-
21T09:51:32.636159Z","shell.execute_reply":"2022-10-
21T09:51:32.635211Z","shell.execute_reply.started":"2022-10-
21T09:51:32.380508Z"},"trusted":true},"outputs":[{"data":{"text/plain":["Text(0.5,1,'Number of ham
and spam
messages')"]},"execution_count":4,"metadata":{},"output_type":"execute_result"},{"data":{"image/pn
g":"iVBORw0KGgoAAAANSUhEUgAAAY4AAAEWCAYAAABxMXBSAAAABHNCSVQICAgIfAhkiAAA
AAlwSFlzAAALEgAACxIB0t1+/AAAADl0RVh0U29mdHdhcmUAbWF0cGxvdGxpYiB2ZXJzaW9uIDI
uMS4wLCBodHRwOi8vbWF0cGxvdGxpYi5vcmcvpW3flQAAGctJREFUeJzt3Xm0HWWd7vHvQ8Kg
ggISERLa0IrdzlMEbO2WdgDEVli2KF6HqCi2V1v7Ltux75VBXWrrFWe7aUVAW5F2jCNGBYerAokTIi
q5gCQGIZgwOV3B3/2j3iNFPOfkFGSfIef7WWrrFWe7aUVAW5F2jCNGBYerAokTIi
BIkgYxOCRJgxgckqRBDA5J0iAGhyRpEINDU5LklCSvnaHXTpL3J9mU5Nxxpj8zyTdmom1bU5Kdk
qyb6XZIW2JwzFFJLk1yRZLb9cqek+TsGWjOPZWpKr2n+nGSPOdwTG3LQRePNONGCrJg
oGz3AW4tKp+NYr2SBrG4Jjb3gT8c5JdN5+QZGmSSrKwV3Z2kue04Wcm+T9JTkxydZKLk/xVK1+
b5Mokyzdb7B5JVia5LslXk9ylt+y/bNM2JvlJkif2pc408F7gI
UmuT3L8RBsjyZtbd9YlSR7TK9+jjZa+m+Tat/t2O+c9WJ7ksiRXJfmX3v
t2OG+c9WJ7ksiRXJfmX3vTbtG2zKcmPgAdP8jpp79WVSa5J8oMk9+5t33+b5L15W2vbtUlWJ/nr3
rTjkvxXkg+2ec9Pcvckr0hyZZK1SQ7tzX92kjckObn/V9x71tz0Sg3Dk3yfbhYtXyxLxyXpw+z/J4k/1bHOWzSBX6MyamG5bHxZYS/aUY8QHQh
vfnuvYePrWV362t0zVtG35kiut2mySntu17YXuP1/Wm753kY0k2tNd7UW/a/klWteVekeQtE22Dea
mqfMzBB3Ap8Cjg48BrW9lzgLPGnwncADwLWAC8FrgMeBewI3AwcB2wc6t/
Shv/mzb9bcA32rTbAWvbshYCDwSuAu7Vm/ca4KF0P1Z2Gmd9vgq8G9gJuD+wAXhkr63fmGRbPP
BP4PfDcti7PB9YDadMfC9wVCPBw4NfAA9u0g9p2eDWwfVvGBuBDwC7AvYDfAn8+wWsfBNynrd
d9gSuAIzZ7D/4DuA1wP+B3wD3a9DcAXwd2B/YBfgism+gEWbq
78xLgF2PvAXBcW79D2vTTgEuAf+ltj0u28Dn8NrAnsBi4EvgO8O8IDWlq8Ax7a6i4FfAoe17fXoNr6I7jN
0LfAXre5e3PT5+XBrz3bt8/GwKa7bG+g+V7sBS4AfjG3ftqzV7X3fAfhz4GLgkDb9W2/DOwIEz/
Tc/mx4z3gAft/CNuyk47k33pbyI4cFxUW/afVr9PXtlvwTu34ZPAU7vTdsZuJHuC+/JwNc3a9+/974
wTgFOm2Rd9mnL2qVX9nrglF5bx0hEbPgj4DbCgje/S5j2gV381LQym8L68F
Thxs/dgSW/6ucBRbfhi4NDetGOYODgeAfwUOBDYbrNpE743EyxrE3C/NnwcsLI37/NnwcsLI37XHA9eNsj10n+
Rw+tTf+MeA9vfF/BD7Zhl8OfGCz+c8EltMFx9XA3wO32azazOacBJ/e04yfbvr9sfg6CNP4ebguMA4L
LN5n0l8P42/DXgeGCPrfl3u6087Kqa46rqh8BngFfcjWs8nnm8LFrb697Lh8n5G9wt1qn7Ra9uv2d+DOAEkek+TbrdvparpfvHv05v1lVd3Yhn/T
nifbDn+U5IAkZ7XujmuAf9hs2TdrG93eztiy9ubm2+VnEVSU5KcvteIYneG5K8pH
XVXNPW/w6btXHzdb1qnO0x7vpPMP9E2+4uwJGbfU4eRrfn9Cu6HyD/AFye5LNJ/rLN9zK6vaxzk1
yQ5NljC9/Cum2+ffvDdwH23qwtr6Lbcw4Grg78OMk5yX+X94xOLYN
b/Ir8l9hkbSLIzXRfLero/yK9W1a69x85V9fzevJPdhnk9sHuSXXplfwb8/Fa2lyQ70v0KfjPd3tSuw
voi2hg8BK+h+3d8B+LcBy76c3jalW+cJVdXbq+pBdN1N1ndwe2ps87nvT+vxfDjwJ2K2t/zUD2rg1raX
b4+h/Tm5XVW8AqKozq+rRdHjAe9uxz22tG6X03VRjelv67V0XxWhqv0x8tpHlwJ2k3dEVTv07
VYe117yoqp4C3AI4I/DR9M5gnO8qqFe2ge6L92lS91WJKHJdkBeA1w
TlWtpdvjuXtuXuPpyfZvj0enOQeU2z/WuCbwOuT7JJTkvnS/+P7zVrYXuv7rHemOW9yQ7qD5hAd7b4Fd
6PaWfptkf+C/DZj3DOCVSXtLsoSuW2dcbXkjVcu8zyxpaNdL0SEQoboz1fifNLBVtPX9r6Z6LzYlUUn
SSebs8LTmOIMkXsP+xBFs4vDwHqF7NJ+X2tsm/T1ZM3VWDjeauMzN1rp4ecHnyLWNxDw2czlJJ/VQWhHRgyTMBrbAZ+gk8JjRNsNsNqhA==","text/plain":["<matplotlib.axes._subplots.AxesSubplot at 0x7f1ab1dbbcc0>"]},"metadata":{},"output_type":"display_data"}],"source":["sns.countplot(df.v1)\n","plt.xlabel('Label')\n","plt.title('Number of ham and spam messages')"]}

8Lgkh7TP5E7pTk5Y0g6oP759Of+OrrvsRoAkR7btA11XVLVpW1q3/vZdDLywN+1c4NokL28H0Rck
uXeSB7fXfFqSRVX1B7ouNLj59p7XDI5txwl0/cR9z6X7VfpLul+p37yVr/Ehur2bjcCD6LqjaF1MBwN
H0e09/ILuV9qOA5b9FLpjAuuBT9AdH1l5K9s71rYX0X2JbKL7Yl9xa5fb89+BE5JcR3eg9YwB8x5P
1z11CfBF4AOT1L093S/wTW2eX9LtRY0Z972hO4bwebrjIz+jC5zJug1HpgXZ4XRdQhtaO15K9z20
Hd3B7fV06/Bwum0L3dlm5yS5nu69e3FVXcKW1+0EYB3d9v0S8FG6UKJ1xT2O7kSMS+hO5ngvX
VcXwKHABe0130Z3XOq3W29rzG1jZ51ImqOSnEJ30Pd/znRbZrMkz6cLgIfPdFvmOvc4JG2TkuyV
5KFJtkvyF3R7NJ+Y6XZtCxZuuYokzUk70J0Wvi/dcYrT6a4V0q1kV5UkaRC7qiRJg4y0qyrJpXS3Qr
gRuKGqliXZne7U0aV0V50+qao2JQnd2QuH0V0k9cyq+k5bznJg7MDfa6vq1Mled4899qilS5du9fW
RpG3Z6tWrr6qqRVuqNx3HOP62qq7qjb8C+HJVvSHJK9r4y4HHAPu1xwHAe+iuRt6d7jTDZXTnb69
OsqKqNk30gkuXLmXVqlWjWRtJ2kYlmfDuBX0z0VV1ODC2x3AqcESv/LTqfBvYNcledDdfW1lVG1t
YrKQ7x1qSNANGHRwFfLHd7viYVrZnVV0O0J7v1MoXc/OLd9a1sonKbybJMe02yKs2bNiwlVdDkjR
m1F1VD62q9UnuBKxM8uNJ6o5375yapPzmBVUn0d1Bk2XLlnmqmCSNyEj3OKpqfXu+ku7Cm/3p
7uy5F3QX6NDdvx+6PYn+TciW0N1+YKJySdIMGFlwJLnd2N1O243LDqb7RzUr6O6/T3v+VBteATwj
nQOBa1pX1pnAwe1GZbu15Zw5qnZLkiY3yq6qPYFPdGfZshD4UFV9Icl5wBnp/iXoZcCRrf7n6E7FX
UN3Ou6zAKpqY5LXAOe1eidU1cYRtluSNIlt8srxZcuWlafjStIwSVZX1bIt1fPKcUnSIAaHJGkQ7447g
Qe99LSZboJmodVvesZMN0Gace5xSJIGMTgkSYMYHJKkQQwOSdIgBockaRCDQ5I0iMEhSRrE4J
AkDWJwSJIGMTgkSYMYHJKkQQwOSdIgBockaRCDQ5I0iMEhSRrE4JAkDWJwSJIGMTgkSYMYH
JKkQQwOSdIgBockaRCDQ5I0iMEhSRrE4JAkDWJwSJIGMTgkSYMYHJKkQQwOSdIgBockaRCDQ5
I0iMEhSRpk5MGRZEGS7yb5TBvfN8k5SS5K8pEkO7TyHdv4mjZ9aW8Zr2zlP0lyyKjbLEma2HTscb
wYuLA3/kbgxKraD9gEHN3KjwY2VdXdgBNbPZLcEzgKuBdwKPDuJAumod2SpHGMNDiSLAEeC7y
3jQd4BPDRVuVU4Ig2fHgbp01/ZKt/OHB6Vf2uqi4B1gD7j7LdkqSJjXqP463Ay4A/tPE7AldX1Q1tfB
2wuA0vBtYCtOnXtPp/LB9nnj9KckySVUlWbdiwYWuvhySpGVlwJPk74MqqWt0vHqdqbWHaZPPcV
FB1UlUtq6plixYtGtxeSdLULBzhsh8KPD7JYcBOwO3p9kB2TbKw7kVUsAda3+uuAfYB29Rsfw2qFR
srH9OeRJE2zke1xVNUrq2pJVS2lO7j9lap6KnAW8MRWbTnwqTa8Too3Tpn+lqqqVH9XOutoX2A84
d1TtliRNbpR7HBN5OXB6ktcC3wXe18rfB3wgyRq6PY2jAKrqgiRnAD8CbgBeUFU3Tn+zJUkwTcFR
VWcDZ7fhixnnrKiq+i1w5ATzvw543ehaKEmaKq8clyQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyRpEI
NDkjSIwSFJGsTgkCQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyRpEINDkjSIwSFJGsTgkCQNYnBIkgY
xOCRJgxgckqRBDA5J0iAGhyRpEINDkjSIwSFJGsTgkCQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyR
pEINDkjSIwSFJGsTgkCQNYnBIkgYZWXAk2SnJuUm+l+SCJMe38rsm+U6SNUk+nGSHVr5jG1/Tpi/t
LeuVrfwnSR4zqjZLkjZNjCGOL4jSxymMm2gZCVOIsCW8JZEw3T+Vd
xZtoGDiSWnxXIKNvcbH2xa3AxXjTm+eZbcQFBrRmA16Ihy9q3j/aTknTvIOu97rjoB+N/48l0UO
yd9Iswb1bwrpP4IKjOoqpLAGj0MQ+ZJcrKYxIHv+Pmd9ueUYMNhlwZ3VqspYLTNeq4V8cUD+/
fzHiHZ3BaPK58yv4P/y+XhM6gGE3GVB/J3b6wU+3H4vEQKnz2P8eTiB8klahJRLP84f1EhHdYkzz
UsMKrY1ZC5bfCqk4BpKG6bbJYvg8zyhoeBSXVp/vcX0wGyN5D8w9uZfrSxwk5KhN/7PGT17i7WZ3
Ttc3YqbIz+1m8fA9Hkqbg0kOWpj4A/fKjOrCkXGV3fBfEf3MT8eXjLjfUZWIPHN+2ecQSNvZ7PG
fWYWkpjHMCgSVWJYhSmtgLdU4yvhkrqTMfBd3WI3GKtTVqcXVyTPQp8VDlxCVJQjS0YTtZ4rWxsp
DRuJhlGJt7B5mS6F7S1pZcpMtJoFHdAV+Zr8Tjao7lGGeQ9xO8QNTKHZQbh+3p4ctCCMwf5lPC
Z+qnAiYN3cxbtSplKX3tWOeJFiUTvFDf3kA/ePHjbZeijBpXBQqSBqVWDf5e8JVHMBXVTVr7mnZ9
xhzs2GQ9D6H26CF+Sj1UQsXUT8oyYTAXl4OWdPGm6dDvN1hdfMbH4WPPBf+l5Fa6fApR8FS3
wC1Mg4+v2ePDFzRFdM4KiUwzSNgOrNqGkJbbkBKU1NlDpDa6Ldd+PT6AHTeVUgyRr+IJu8FMQ
HA0pK3BQVYj0ajYoMPCQ3ueg8lkgjo1I7aN9ezvtcgdaNnnuUhpRk5MaPiCNgAVj/4sxqOkNxKa
8Lgkh7TP5E7pTk5Y0g6oP759Of+OrrvsRoAkR7btA11XVLVpW1q3/vZdDLywN+1c4NokL28H0Rck

jSIwSFJGsTgkCQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyRpEINDkjSIwSFJGmRkwZHk5CRXJvlhr2
z3JCuTXNSed2vlSfL2JGuS/CDJA3vzLG/1L0qyfFTtlSRNzSj3OE4BDt2s7BXAl6tqP+DLbRzgMcB+
7XEM8B7oggY4FjgA2B84dixsJEkzY2TBUVVfAzZuVnw4cGobPhU4old+WnW+DeyaZC/gEGBlVW2
sqk3ASv40jCRJ02i6j3HsWVWXA7TnO7XyxcDaXr11rWyi8j+R5Jgkq5Ks2rBhw1ZvuCSpM1sOjmec
spqk/E8Lq06qqmVVtWzRokVbtXGSpJtMd3Bc0bqgaM9XtvJ1wD69ekuA9ZOUS5JmyHQHxwpg7
Myo5cCneuXPaGdXHQhc07qyzgQOTrJbOyh+cCuTJM2QhaNacJIPAwcBeyRZR3d21BuAM5IcDV
wGHNmqfw44DFgD/Bp4FkBVbUzyGuC8Vu+Eqtr8gLskaRqNLDiq6ikTTHrkOHULeMEEyzkZOHkrN
k2SdCvMloPjkqQ5wuCQJA1icEiSBjE4JEmDGBySpEFGdlaVpNG47IT7zHQTNAv92avPn7bXco9Dkj
SIwSFJGsTgkCQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyRpEINDkjSIwSFJGsTgkCQNYnBIkgYxOC
RJgxgckqRBDA5J0iAGhyRpEINDkjSIwSFJGsTgkCQNYnBIkgYxOCRJgxgckqRBDA5J0iAGhyRpEIN
DkjSIwSFJGsTgkCQNYnBIkgaZM8GR5NAkP0myJskrZro9kjRfzYngSLIAeBfwGOCewfwFOS3HNmWy
VJ89OcCA5gf2BNVV1cVf8POB04fIbbJEnz0sKZbsAUlQbW9sbXAQf0W9sbXAQf0KyQ5BjimjV6f5CfT1Lb5YA/
gqpluxGyQNy+f6Sbo5vxsjjk2W2Mpd5lKpbkSHONtkbrZSNVJwEnT05z5Jcmqqlo20+2QNudnc2b
Mla6qdcA+vfElwPoZaoskzWtzJTjOA/ZLsm+SHYCjgBUz3CZJmpfmRFdVVd2Q5IXAmcAC4OSqu
mCGmzWf2AWo2crP5gxIVW25liRJzVzpqpIkzRIGhyRpEINjHkuyNMkPZ7odkuYWg0OSNIjBoQVJ/i
PJBUm+mOQ2SZ6b5Lwk30/ysSS3BUhySpL3JDkrycVJHp7k5CQXJjllhtdDc1yS2yX5bPvc/TDJk5
NcmuSNSc5tj7u1uo9Lck6S7yb5UpI9W/lxSU5tn+VLkzwhyb/OT/JF5JsP7NruW0wOLQf8K6quh
dwNfD3wMer6sFVdT/gQuDoXv3dgEcA/wP4NHAicC/gPknuP60t17bmUGB9Vd2vqu4NfKGVX1tV
+wPvBN7ayr4BHFhVD6C7d93Lesu5K/BYuvzZfRA4q6ruA/ymletWjh0SVV9rw2vBpYC907y9STn
A0+lC4Yxn67uHO7zgSuq6vyq+gNwQZtXuqXOBx7V9jD+uqquae+Uf7j0/pA0vAc5sn9GXcvPP6Oer6
vdteQu4KYDOx8/oVmFw6He94RvpLgo9BXhh+5V2PLDTOPX/sNm8f2COXFCq2amqfgo8iO4L/vV
JXj02qV+tPb8DeGf7jD6Pc7ji9X8jG4lBofGswtweesPfupMN0bzQ5K9gV9X1QeBNwM
PbJOe3Hv+Vhu+A/DzNn7UJSfTpakzOKSt7x3AW6pqRZK/P0Mkx89As6UpMzikrW
+yO7fel+4OrQ8BXp1k7yQH092leH/g/sCDkvzNNLdZmjJvOSJtfUuAjyTZ9C9gBuKQ37VNV9RvgN0n
OoguLhwEHA99tdXamC5KvTV+TpakzOKSt7x3AW6pqRZKD6O6rNGbzK24LCPD6qvr36WqgtKU8VSVtf
VSVtfZPdufXwJDsluSNwEHAecCbw7CQ7AyRZnGSP6Wu2tGUcvS0vF+QKaI3w9oQ7YJ4M26J84UD8HsFj
VSVtfZPdufXwJDsluSNwEHAecCbw7CQ7AyRZnORO09VYaSj3OKRb57ZJ1vXG38Lkd249F/gs8Gf
Aa6pqPbA+yT2AbyUBuB54GnDl6JsvDee9qiRJg9hVJUkaxOCQJA1icEiSBjE4JEmDGBySpEEMDkn
SIAaHJGmQ/w8HGwM7lDAXhgAAAABJRU5ErkJggg==","text/plain":["<matplotlib.figure.Figure at
0x7f781c41f7f0>"]},"metadata":{},"output_type":"display_data"}],"source":["sns.countplot(df.v1)\n","
plt.xlabel('Label')\n","plt.title('Number of ham and spam
messages')"]},{"cell_type":"code","execution_count":5,"metadata":{"_uuid":"a1a345c1683e2fcc7173
ecae867a5da87f2dde24","execution":{"iopub.execute_input":"2022-10-
21T09:54:29.911092Z","iopub.status.busy":"2022-10-
21T09:54:29.910704Z","iopub.status.idle":"2022-10-
21T09:54:29.919823Z","shell.execute_reply":"2022-10-
21T09:54:29.918839Z","shell.execute_reply.started":"2022-10-
21T09:54:29.911029Z"},"trusted":true},"outputs":[],"source":["X = df.v2\n","Y = df.v1\n","le =
LabelEncoder()\n","Y = le.fit_transform(Y)\n","Y = Y.reshape(-
1,1)"]},{"cell_type":"code","execution_count":6,"metadata":{"_uuid":"aa3386af09469682c66cc53a18
30a4e42f0e70b6","execution":{"iopub.execute_input":"2022-10-
21T09:54:33.438115Z","iopub.status.busy":"2022-10-
21T09:54:33.437528Z","iopub.status.idle":"2022-10-
21T09:54:33.444835Z","shell.execute_reply":"2022-10-
21T09:54:33.443737Z","shell.execute_reply.started":"2022-10-
21T09:54:33.438044Z"},"scrolled":true,"trusted":true},"outputs":[],"source":["X_train,X_test,Y_train,Y
_test =
train_test_split(X,Y,test_size=0.15)"]},{"cell_type":"markdown","metadata":{"_uuid":"c5378d55c271
e01480c1ac07f94ff99a80f900d6"},"source":["### Pre-
Processing"]},{"cell_type":"code","execution_count":7,"metadata":{"_uuid":"bdca14f2b8cd7bd7cb5e
e66fd40ea522217c03c6","execution":{"iopub.execute_input":"2022-10-
21T09:54:38.389836Z","iopub.status.busy":"2022-10-
21T09:54:38.389540Z","iopub.status.idle":"2022-10-
21T09:54:38.626342Z","shell.execute_reply":"2022-10-

21T09:54:38.625512Z","shell.execute_reply.started":"2022-10-
21T09:54:38.389784Z"},"trusted":true},"outputs":[],"source":["max_words = 1000\n","max_len =
150\n","tok = Tokenizer(num_words=max_words)\n","tok.fit_on_texts(X_train)\n","sequences =
tok.texts_to_sequences(X_train)\n","sequences_matrix =
sequence.pad_sequences(sequences,maxlen=max_len)"]},{"cell_type":"markdown","metadata":{"_
uuid":"ad8706caa7a447fb49b44919fd109129e4082a93"},"source":["### Create Model and add
Layers"]},{"cell_type":"code","execution_count":9,"metadata":{"_uuid":"78fff25b8be1de575bff071a2
027f3dd2b11b911","execution":{"iopub.execute_input":"2022-10-
21T09:55:53.694120Z","iopub.status.busy":"2022-10-
21T09:55:53.693762Z","iopub.status.idle":"2022-10-
21T09:55:53.707856Z","shell.execute_reply":"2022-10-
21T09:55:53.706969Z","shell.execute_reply.started":"2022-10-
21T09:55:53.694044Z"},"trusted":true},"outputs":[],"source":["def RNN():\n"," inputs =
Input(name='inputs',shape=[max_len])\n"," layer =
Embedding(max_words,50,input_length=max_len)(inputs)\n"," layer = LSTM(64)(layer)\n","
layer = Dense(256,name='FC1')(layer)\n"," layer = Activation('relu')(layer)\n"," layer =
Dropout(0.5)(layer)\n"," layer = Dense(1,name='out_layer')(layer)\n"," layer =
Activation('sigmoid')(layer)\n"," model = Model(inputs=inputs,outputs=layer)\n"," return
model"]},{"cell_type":"markdown","metadata":{"_uuid":"9d7c489e32bff6d12b8c08c07a91e9ba5d30
2e0e"},"source":["### Compile the
Model"]},{"cell_type":"code","execution_count":10,"metadata":{"_uuid":"a0ede32d4127e8b4990fd74
fe97fadef9e565d17","execution":{"iopub.execute_input":"2022-10-
21T09:56:00.085329Z","iopub.status.busy":"2022-10-
21T09:56:00.084753Z","iopub.status.idle":"2022-10-
21T09:56:00.416307Z","shell.execute_reply":"2022-10-
21T09:56:00.415584Z","shell.execute_reply.started":"2022-10-
21T09:56:00.085265Z"},"trusted":true},"outputs":[{"name":"stdout","output_type":"stream","text":["__
____\n","Layer (type)
Output Shape Param #
\n","===\n","inputs
(InputLayer) (None, 150) 0
\n","____\n","embedding_1
(Embedding) (None, 150, 50) 50000
\n","____\n","lstm_1 (LSTM)
(None, 64) 29440
\n","____\n","FC1 (Dense)
(None, 256) 16640
\n","____\n","activation_1
(Activation) (None, 256) 0
\n","____\n","dropout_1 (Dropout)
(None, 256) 0
\n","____\n","out_layer (Dense)
(None, 1) 257
\n","____\n","activation_2
(Activation) (None, 1) 0
\n","===\n","Total
params: 96,337\n","Trainable params: 96,337\n","Non-trainable params:
0\n","____\n"]}],"source":["model =
RNN()\n","model.summary()\n","model.compile(loss='binary_crossentropy',optimizer=RMSprop(),
metrics=['accuracy'])"]},{"cell_type":"markdown","metadata":{"_uuid":"bc2e0a3ec50d14c790b82d66
f9255456ec6a69da"},"source":["### Fit the
Model"]},{"cell_type":"code","execution_count":11,"metadata":{"_uuid":"98f6d6318352420ea49c532
cda158f715f940f4b","execution":{"iopub.execute_input":"2022-10-
21T09:56:05.437010Z","iopub.status.busy":"2022-10-
21T09:56:05.436629Z","iopub.status.idle":"2022-10-

21T09:56:56.241469Z","shell.execute_reply":"2022-10-
21T09:56:56.240644Z","shell.execute_reply.started":"2022-10-
21T09:56:05.436940Z"},"trusted":true},"outputs":[{"name":"stdout","output_type":"stream","text":["Tr
ain on 3788 samples, validate on 948 samples\n","Epoch 1/10\n","3788/3788
[==============================] - 8s 2ms/step - loss: 0.3312 - acc: 0.8746 - val_loss:
0.1460 - val_acc: 0.9504\n","Epoch 2/10\n","3788/3788 [==============================] -
8s 2ms/step - loss: 0.0860 - acc: 0.9789 - val_loss: 0.0666 - val_acc: 0.9768\n","Epoch
3/10\n","3788/3788 [==============================] - 8s 2ms/step - loss: 0.0447 - acc:
0.9873 - val_loss: 0.0465 - val_acc: 0.9895\n","Epoch 4/10\n","3788/3788
[==============================] - 9s 2ms/step - loss: 0.0353 - acc: 0.9892 - val_loss:
0.0459 - val_acc: 0.9863\n","Epoch 5/10\n","3788/3788 [==============================] -
8s 2ms/step - loss: 0.0258 - acc: 0.9918 - val_loss: 0.0437 - val_acc: 0.9884\n","Epoch
6/10\n","3788/3788 [==============================] - 8s 2ms/step - loss: 0.0196 - acc:
0.9947 - val_loss: 0.0468 - val_acc: 0.9905\n"]},{"data":{"text/plain":["<keras.callbacks.History at
0x7f780f71ad68>"]},"execution_count":11,"metadata":{},"output_type":"execute_result"}],"source":["
model.fit(sequences_matrix,Y_train,batch_size=128,epochs=10,\n","
validation_split=0.2,callbacks=[EarlyStopping(monitor='val_loss',min_delta=0.0001)])"]},{"cell_type
":"markdown","metadata":{"_uuid":"ccca7839445a7d663ee7bc425a16e247df3e0e5b"},"source":["#
Test the
Model"]},{"cell_type":"code","execution_count":12,"metadata":{"_uuid":"80036135a11387d952becaf
2fecf653a65c02328","execution":{"iopub.execute_input":"2022-10-
21T09:57:38.153001Z","iopub.status.busy":"2022-10-
21T09:57:38.152577Z","iopub.status.idle":"2022-10-
21T09:57:38.186674Z","shell.execute_reply":"2022-10-
21T09:57:38.185725Z","shell.execute_reply.started":"2022-10-
21T09:57:38.152936Z"},"trusted":true},"outputs":[],"source":["test_sequences =
tok.texts_to_sequences(X_test)\n","test_sequences_matrix =
sequence.pad_sequences(test_sequences,maxlen=max_len)"]},{"cell_type":"code","execution_cou
nt":13,"metadata":{"_uuid":"0db183049b59d96388812a98efedfc865b7cc141","execution":{"iopub.
execute_input":"2022-10-21T09:57:43.980227Z","iopub.status.busy":"2022-10-
21T09:57:43.979749Z","iopub.status.idle":"2022-10-
21T09:57:44.671215Z","shell.execute_reply":"2022-10-
21T09:57:44.670527Z","shell.execute_reply.started":"2022-10-
21T09:57:43.980161Z"},"trusted":true},"outputs":[{"name":"stdout","output_type":"stream","text":["8
36/836 [==============================] - 1s 821us/step\n"]}],"source":["accr =
model.evaluate(test_sequences_matrix,Y_test)"]},{"cell_type":"code","execution_count":14,"metada
ta":{"_uuid":"3e121ab83f4a0b9f7376ab24aa25d67051171f89","execution":{"iopub.execute_input":"
2022-10-21T09:57:48.188706Z","iopub.status.busy":"2022-10-
21T09:57:48.188440Z","iopub.status.idle":"2022-10-
21T09:57:48.193811Z","shell.execute_reply":"2022-10-
21T09:57:48.193187Z","shell.execute_reply.started":"2022-10-
21T09:57:48.188672Z"},"scrolled":true,"trusted":true},"outputs":[{"name":"stdout","output_type":"str
eam","text":["Test set\n"," Loss: 0.057\n"," Accuracy: 0.986\n"]}],"source":["print('Test set\\n
Loss: {:0.3f}\\n Accuracy:
{:0.3f}'.format(accr[0],accr[1]))"]}],"metadata":{"kernelspec":{"display_name":"Python
3","language":"python","name":"python3"},"language_info":{"codemirror_mode":{"name":"ipython","v
ersion":3},"file_extension":".py","mimetype":"text/x-
python","name":"python","nbconvert_exporter":"python","pygments_lexer":"ipython3","version":"3.6.
4"}},"nbformat":4,"nbformat_minor":4}