Please read the entire contents of this file, for instructions on running the code smoothly.

1. Open Spyder .
2. With in Spyder, open the sample.py file in the same folder that this ReadMe file resides in.
3. Once the file has been loaded, please make sure you set the console working directory to the path this file is in as below. This way you don't have to give the entire path, just the relative path of the dataset file.
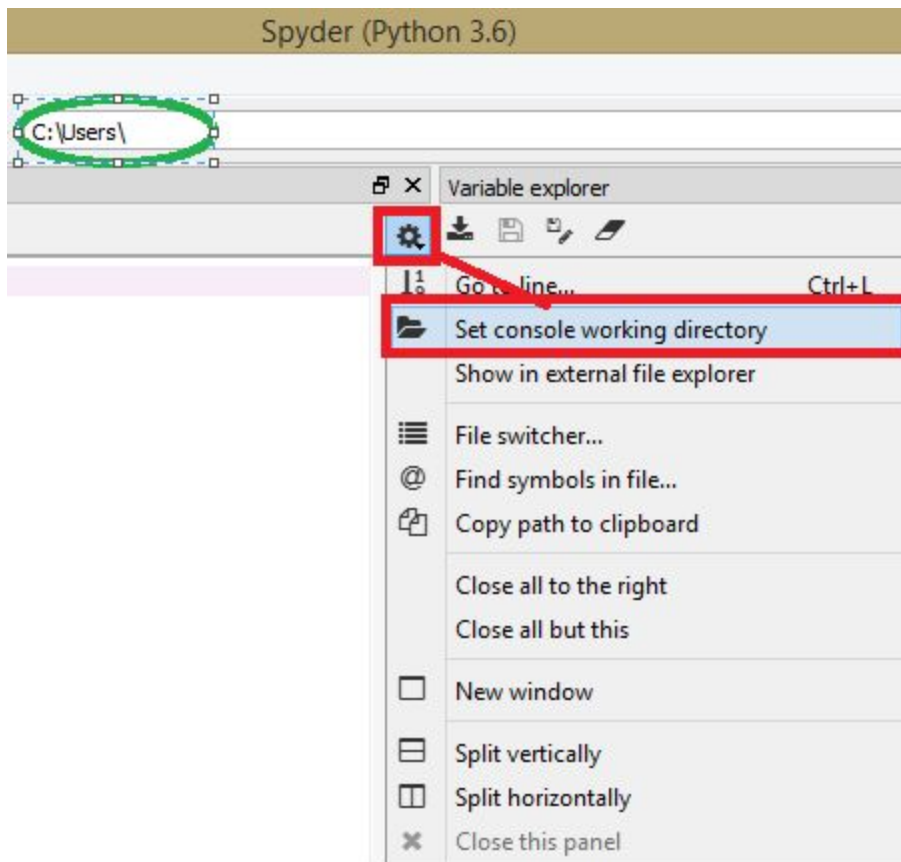
4. Now, run the sample code all at one or 1 part at a time.
5. In **Part 1**, dataset used was KDDTrain+.txt which is in the same folder that this ReadMe file resides in.
6. The dataset doesn't have a header, so header=None is passed to pandas in line 21.
7. Now, since the last 2 columns are label related columns, they have been excluded in the X variable.
8. The label_column should be made to hold the label column, which is the 2nd column from right (-2).
9. However, the label column has string values, I changed them to be binary (0 => Normal, 1 => Attack).
10. The code is set to run for only 10 epochs, and you would see the epoch results printing while the model is being trained as below. Since the model in the sample code is small, it will take less than 5 mins to train the model (**Part 2**). With expected out to be binary, I

used binary cross entropy as the loss function, and you can see through the eopch results that the loss decreased as the model gets trained over epochs.

```
Epoch 1/10
94479/94479 [==============================] - 17s 181us/step - loss: 0.0815 - acc: 0.9727
Epoch 2/10
94479/94479 [==============================] - 21s 219us/step - loss: 0.0402 - acc: 0.9805
Epoch 3/10
94479/94479 [==============================] - 26s 273us/step - loss: 0.0332 - acc: 0.9857
Epoch 4/10
94479/94479 [==============================] - 28s 292us/step - loss: 0.0312 - acc: 0.9868
Epoch 5/10
94479/94479 [==============================] - 36s 382us/step - loss: 0.0301 - acc: 0.9871
Epoch 6/10
94479/94479 [==============================] - 40s 421us/step - loss: 0.0296 - acc: 0.9874 - ETA: 48s - loss: 0.0308 - acc: 0.9871
Epoch 7/10
94479/94479 [==============================] - 41s 430us/step - loss: 0.0293 - acc: 0.9874
Epoch 8/10
94479/94479 [==============================] - 42s 441us/step - loss: 0.0290 - acc: 0.9877
Epoch 9/10
94479/94479 [==============================] - 40s 419us/step - loss: 0.0286 - acc: 0.9880
Epoch 10/10
94479/94479 [==============================] - 42s 440us/step - loss: 0.0273 - acc: 0.9893

In [28]:
```

11. Now, it is test the model, with the test subset you obtained in line 45. You can also use the KDDTest+.txt file (you need to do the same data pre-processing on this dataset file too - repeat Part 1). You will see the confusion matrix results printed as below. Change the threshold from 0.9 to 0.5 on **line 89** and you will see the confusion matrix results change too.

```
[16801    25]
[  630 14038]
```

Representing positives and negatives as below
[TN   FN]
[FP   TP]

12. Then you will see plots showing the accuracy and loss history as below upon running **Part 4** -