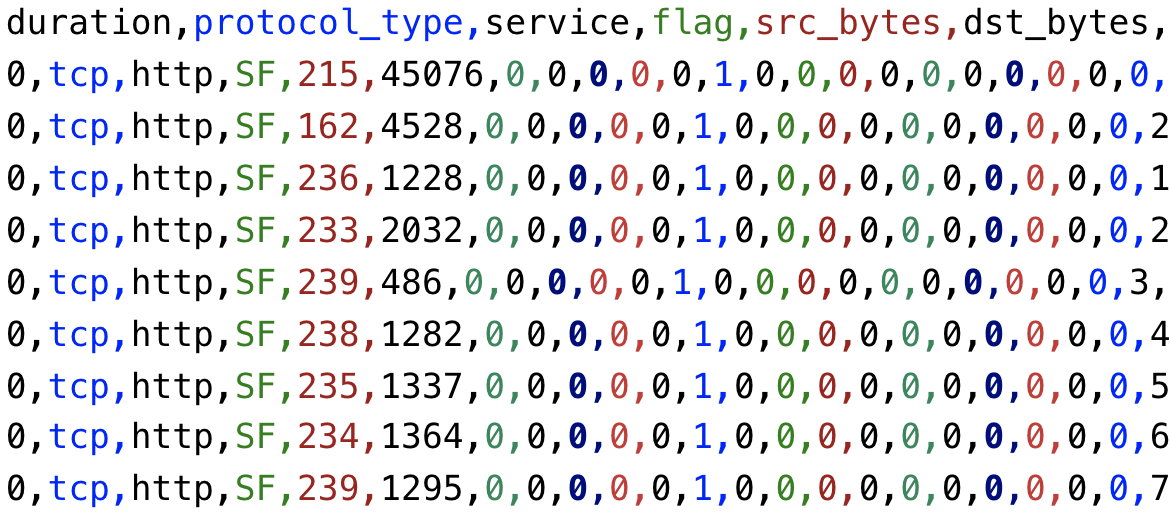
Machine learning cybersecurity **intrusion detection**

# LAB 1: Writing a classifier for kdd99 DATASET

**Lab Description:** This lab is to implement a binary classifier to distinguish normal connections from attacks. Since the dataset is a large dataset, you may need to employ GPU to run the code.



**Lab Environment:**

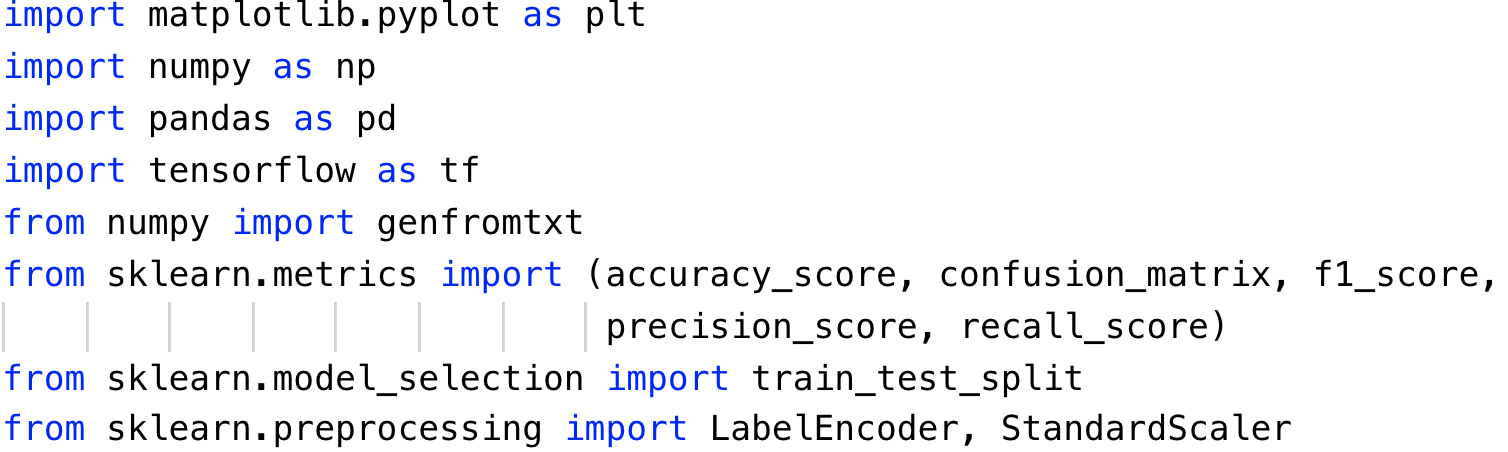
* The students should have access to a machine with Linux system
* The environment for python is required as well as some packages such as numpy, tensorflow and sklearn.
* GPU box

**Lab Files that are Needed:**

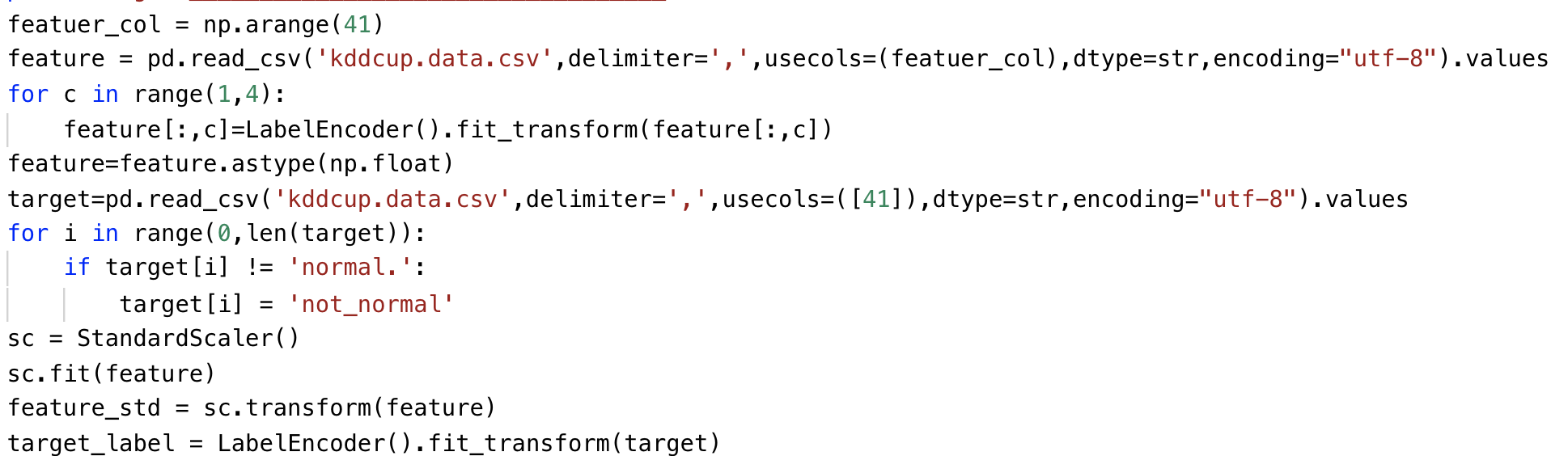
* For this lab you will need only one file (kdd.data.csv) for both WEKA and python script.
* The last column is the class value, others are the features.

### **Lab exercise 1**

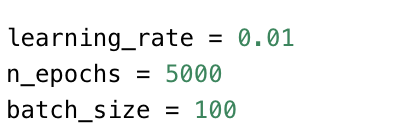
* In this exercise, you will implement an artificial neural network classifier based on Tensorflow
* Import the required libraries



* Repeat the same steps to preprocess the data as Exercise 2. Read the data, standard scale the feature and encode the labels.

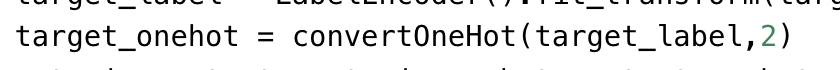


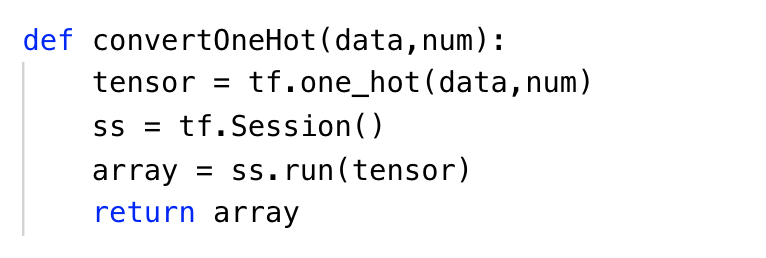
* Define the learning rate, number of epochs and batch size for artificial neural network



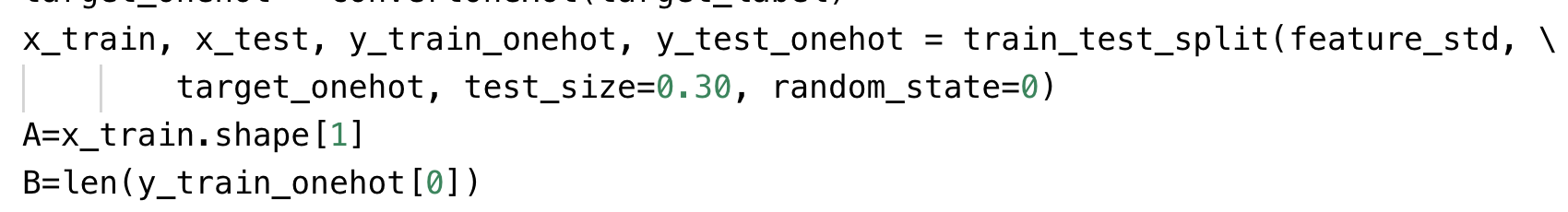
* An extra step in preprocess is to perform the one-hot encoding for the labels.

Since the type of result of tf.one\_hot is a tensor object, it has to be converted to a numpy array.

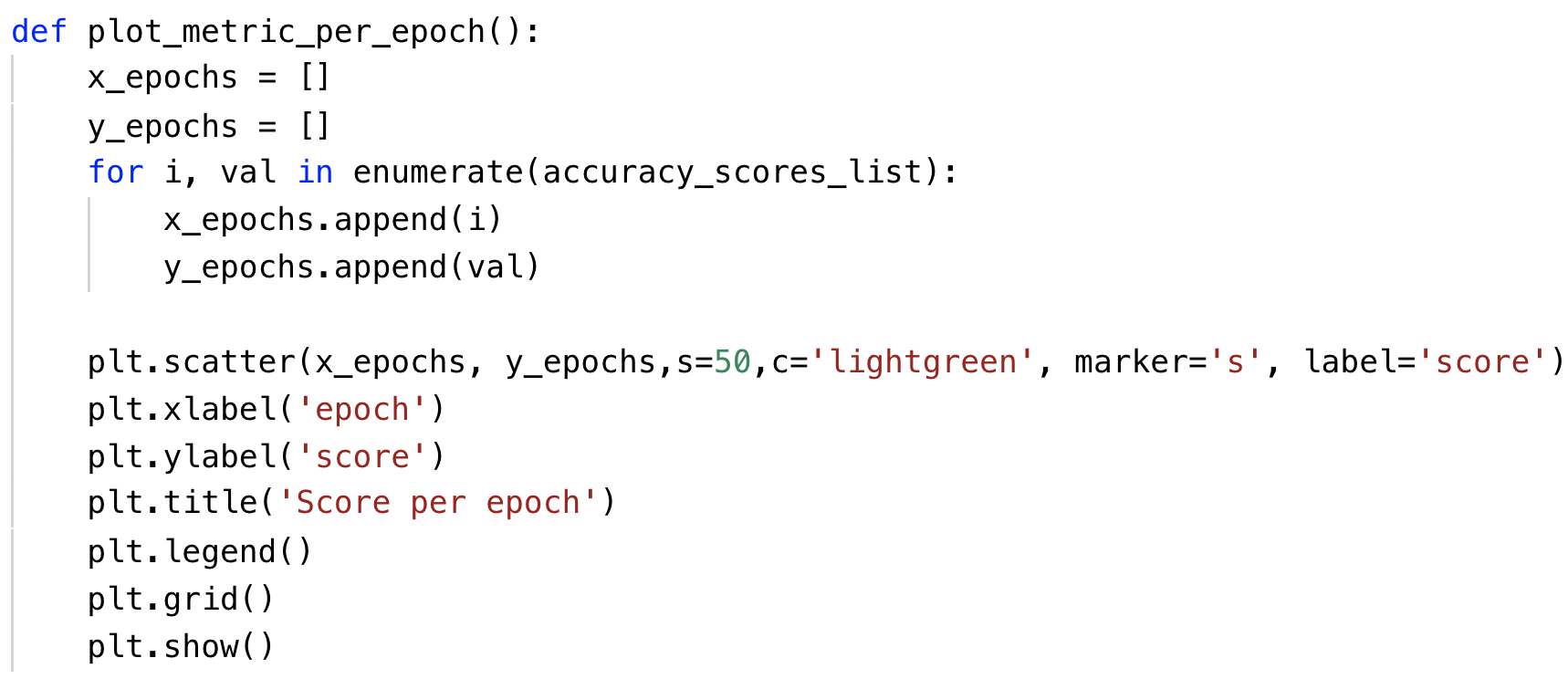




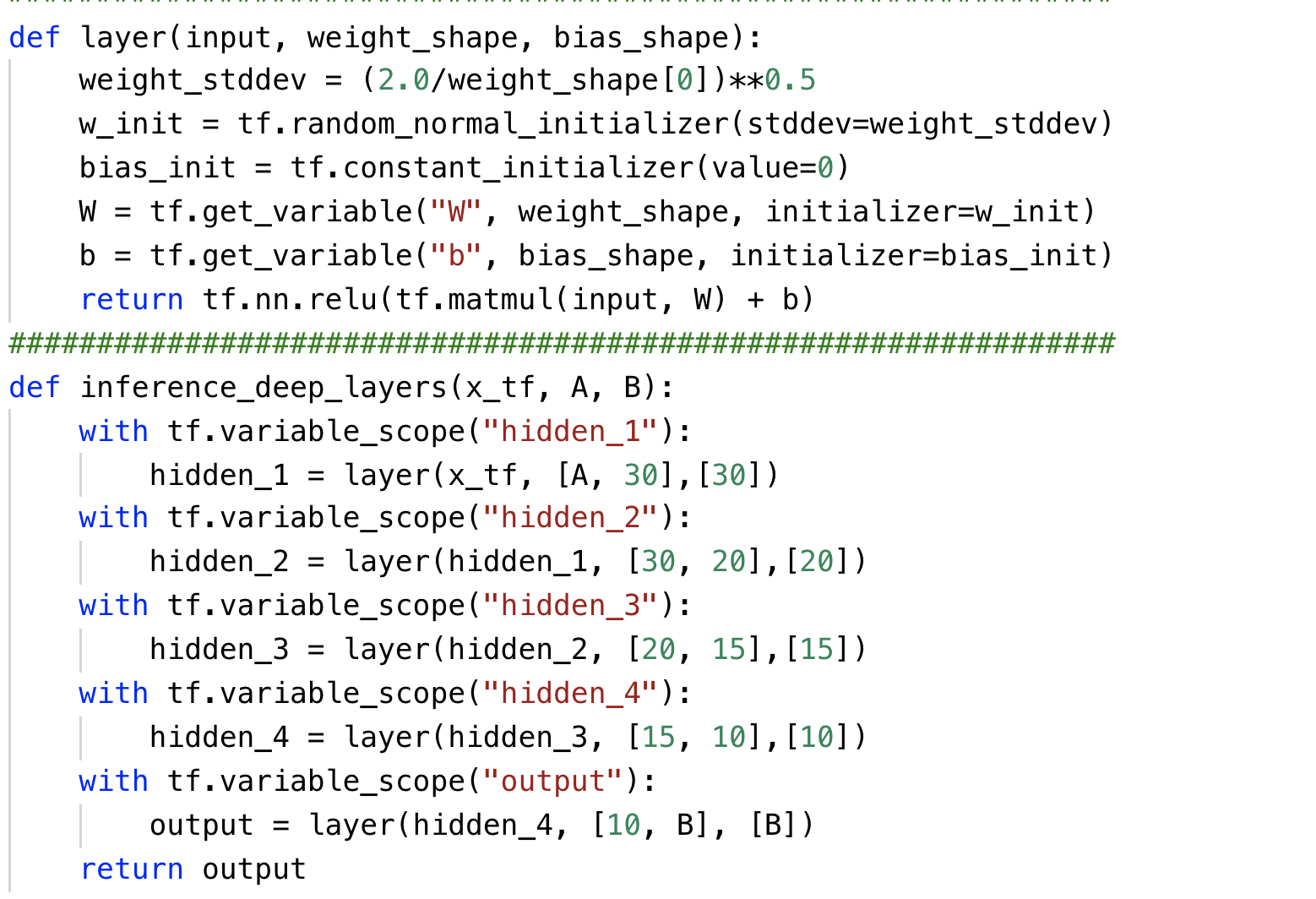
* Split the dataset after preprocessing and define the parameters to store the shape of placeholder.

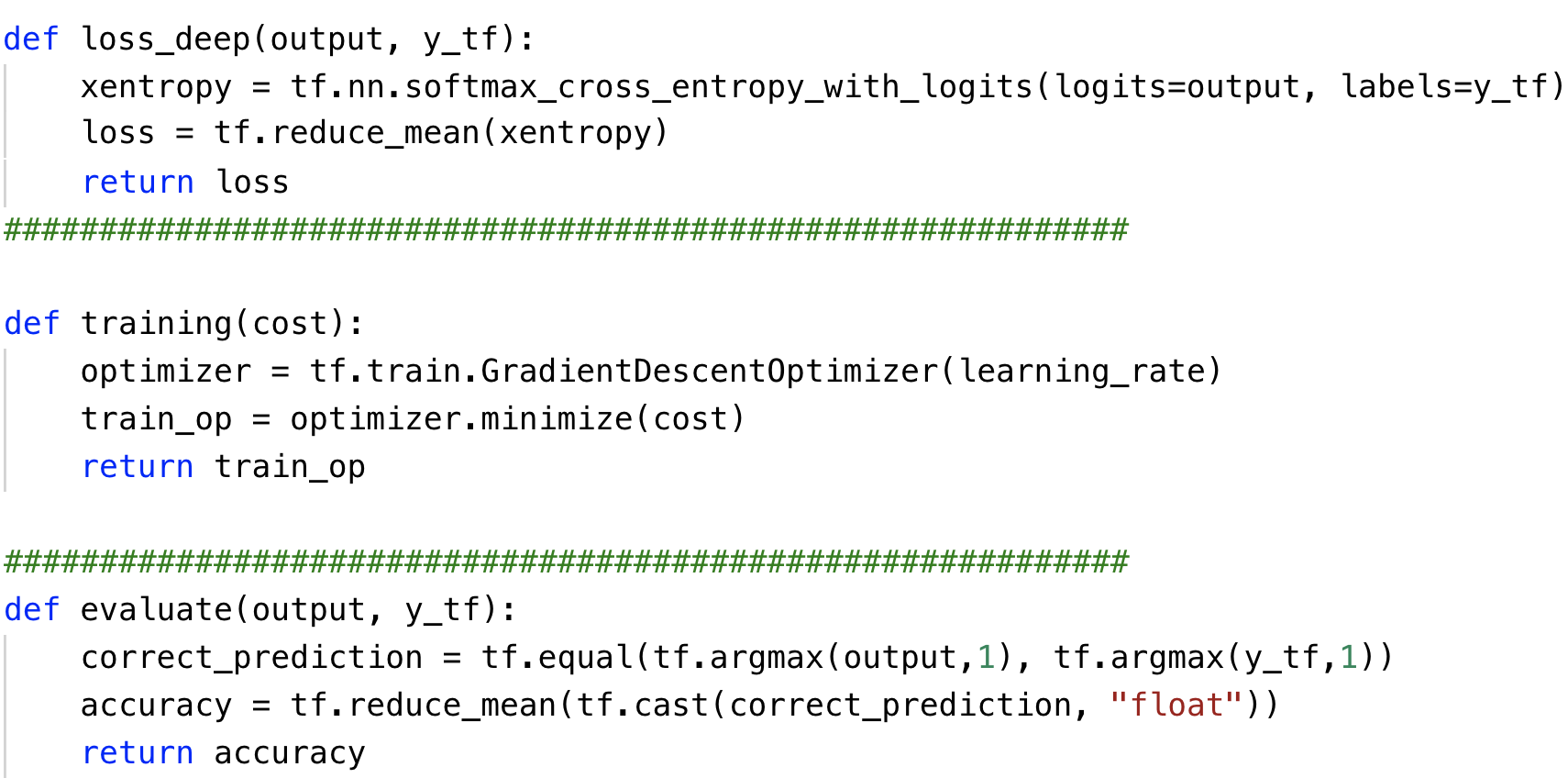


* Define the function to draw the plot of performance

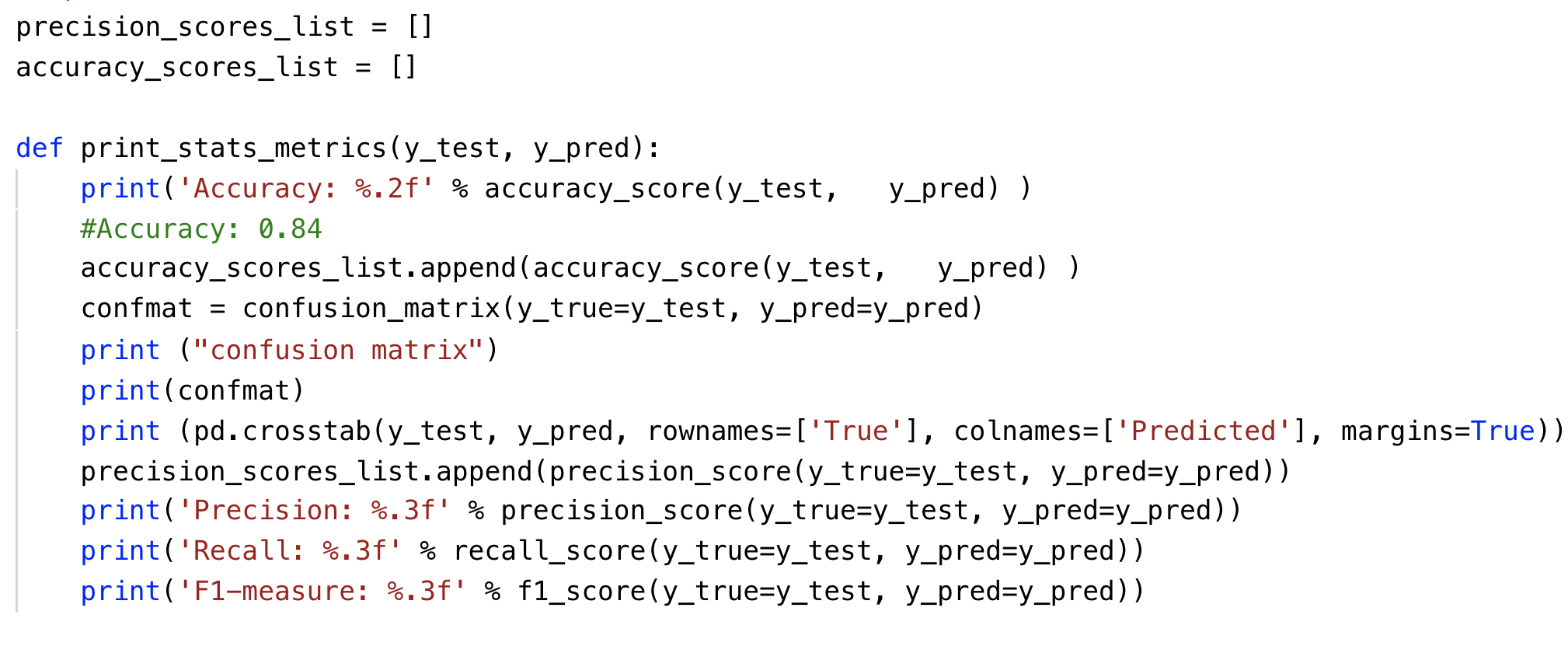


* Define your own architecture of neural network

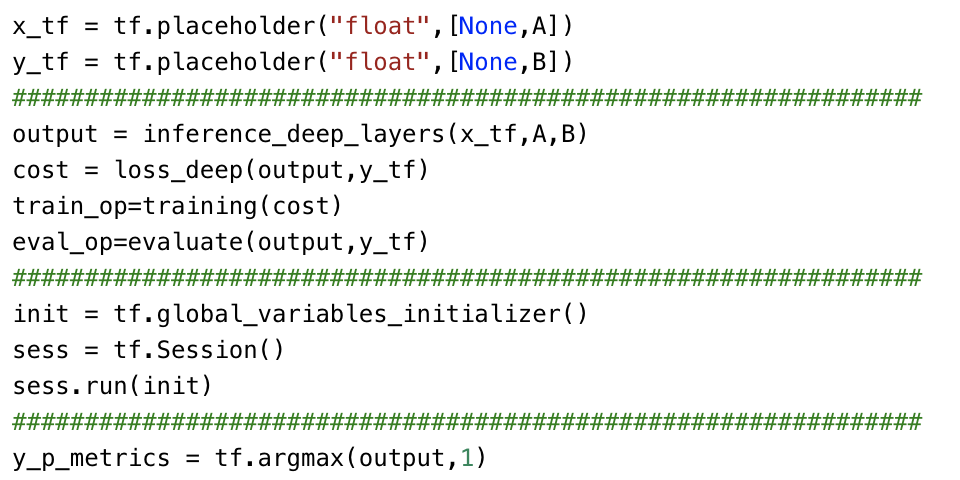




* Please print the statistics metrics such as accuracy, recall, precision and f1 score.



* Initialize the variables and placeholders. Then perform the training and testing on subset of kdd dataset.



## What to Submit

You should submit a lab report file which includes:

* + The steps for how you preprocessed data
  + The necessary code snippet of your classifier and architecture.
  + The screenshot of the results
  + You can name your report "Lab\_kdd\_yourname.doc".