Đại học Quốc gia Thành phố Hồ Chí Minh
Trường Đại học Công nghệ Thông tin (UIT)

BÁO CÁO LAB

Môn học: Phương pháp học máy trong an toàn thông tin

LAB 3 : Advanced Malware Detection

Nhóm: 05

THÔNG TIN CHUNG:

Lớp: NT522.N21.ATCL

STT	Họ và tên	MSSV	Phân công
1	Đỗ Quang Thắng	20521893	1,3,5,7
2	Nguyễn Đoàn Thiên Cung	20521146	2,4,6

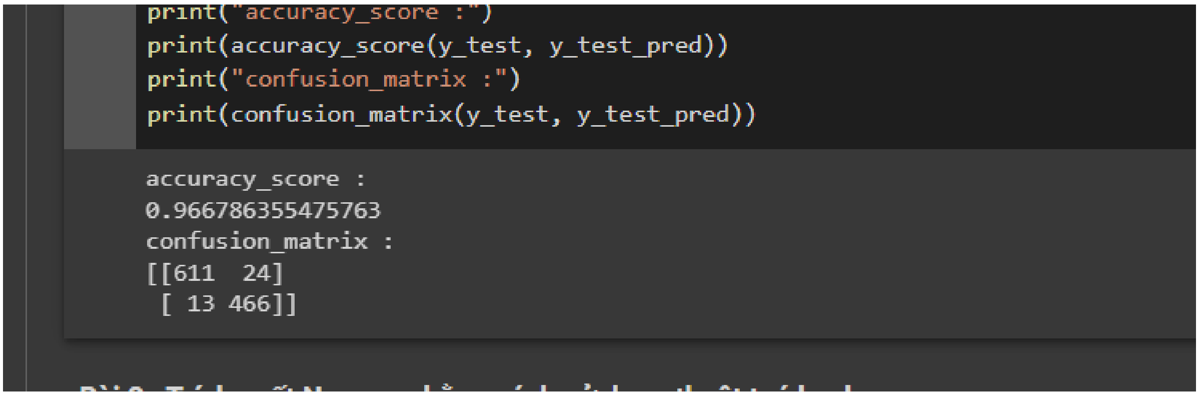
1. NỘI DUNG THỰC HIỆN:

BÁO CÁO CHI TIẾT

Bài 1,3,5,7:
https://colab.research.google.com/drive/1F5qypKbiEr4cehSiTp-w9dzIWBmHlx8M?usp=sharing

Bài 2,4,6 :
https://drive.google.com/drive/folders/1QFDKTIcZEGCxJfpHo8MYoUgJsCOnzlVi?fbclid=IwAR2m9eHeLpjlS7BaoQu5zMNIHOjh-xStTp5kX826-mi5KSrEuDFGXZpurEM

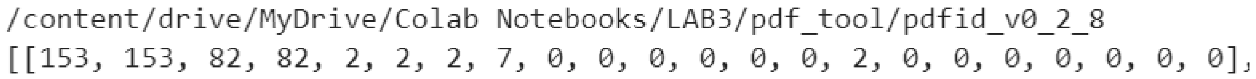
Câu 1 : Cho biết kết quả accuracy và confusion matrix



```
print("accuracy_score :")
print(accuracy_score(y_test, y_test_pred))
print("confusion_matrix :")
print(confusion_matrix(y_test, y_test_pred))

accuracy_score :
0.966786355475763
confusion_matrix :
[[611  24]
 [ 13 466]]
```

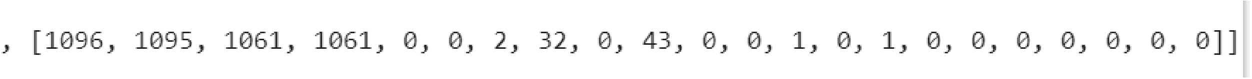
Câu 2: Cho biết kết quả vector X

File thứ nhất:



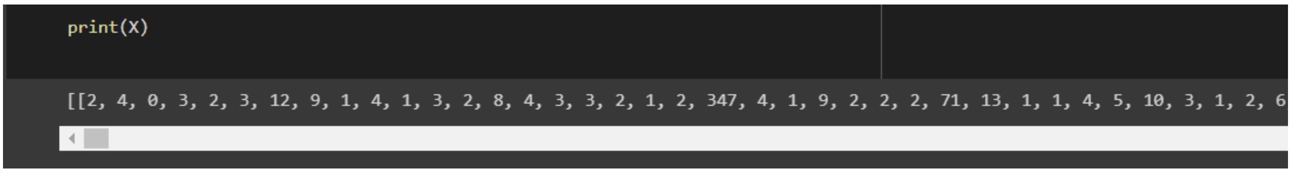
```
/content/drive/MyDrive/Colab Notebooks/LAB3/pdf_tool/pdfid_v0_2_8
[[153, 153, 82, 82, 2, 2, 2, 7, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0],
```

File thứ hai:



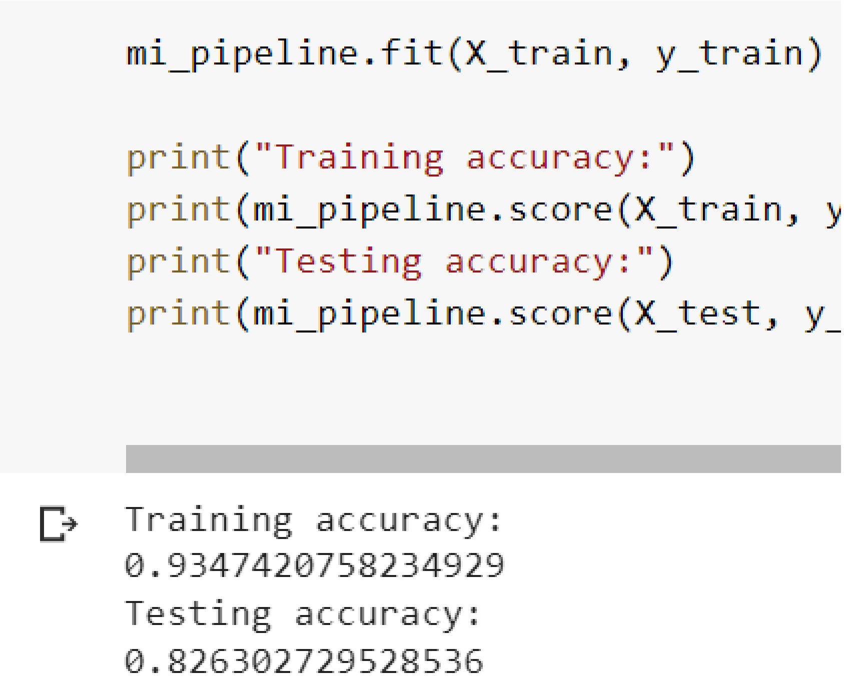
```
, [1096, 1095, 1061, 1061, 0, 0, 2, 32, 0, 43, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0]]
```

Câu 3 : Cho biết kết quả vector X



```
print(X)

[[2, 4, 0, 3, 2, 3, 12, 9, 1, 4, 1, 3, 2, 8, 4, 3, 3, 2, 1, 2, 347, 4, 1, 9, 2, 2, 2, 71, 13, 1, 1, 4, 5, 10, 3, 1, 2, 6
```

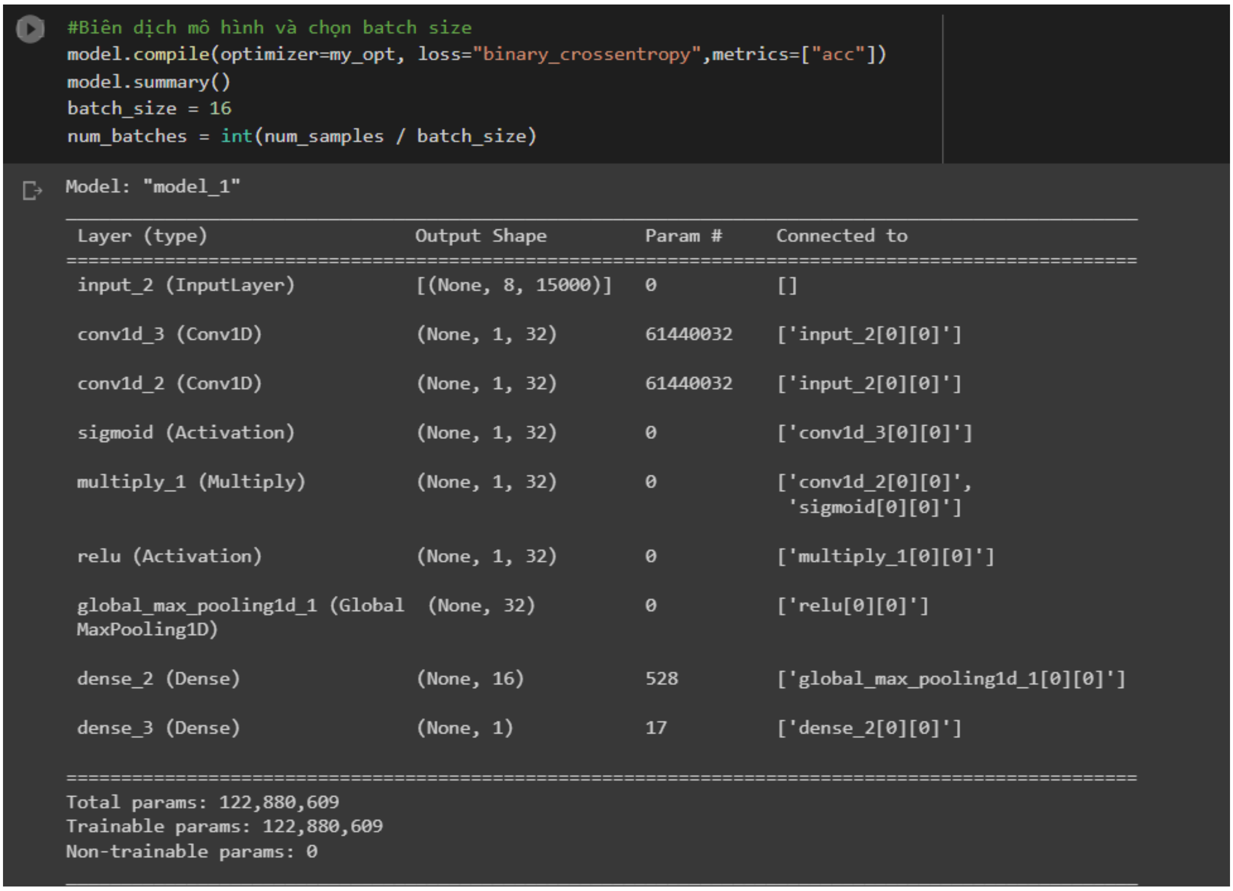
Câu 4: Cho biết kết quả đánh giá.



```
mi_pipeline.fit(X_train, y_train)

print("Training accuracy:")
print(mi_pipeline.score(X_train, y
print("Testing accuracy:")
print(mi_pipeline.score(X_test, y_
```

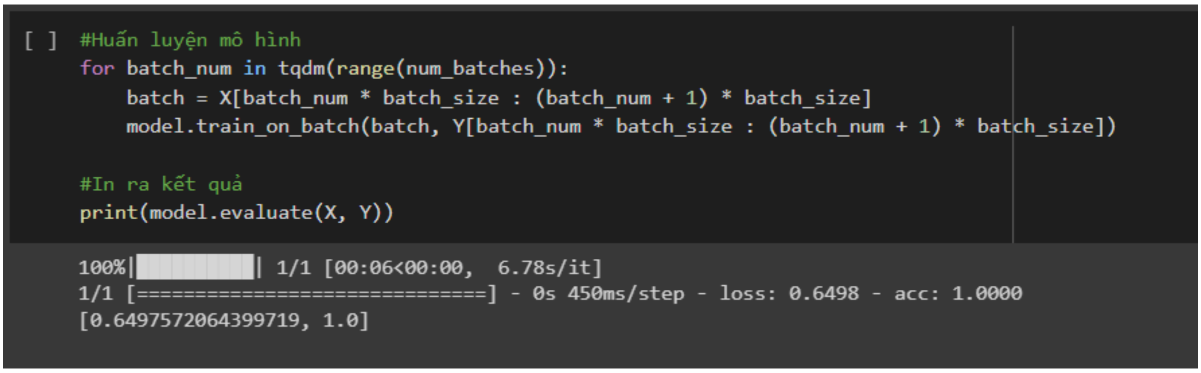
```
Training accuracy:
0.9347420758234929
Testing accuracy:
0.826302729528536
```

Câu 5 : Cho biết kết quả đánh giá mô hình qua tập test.



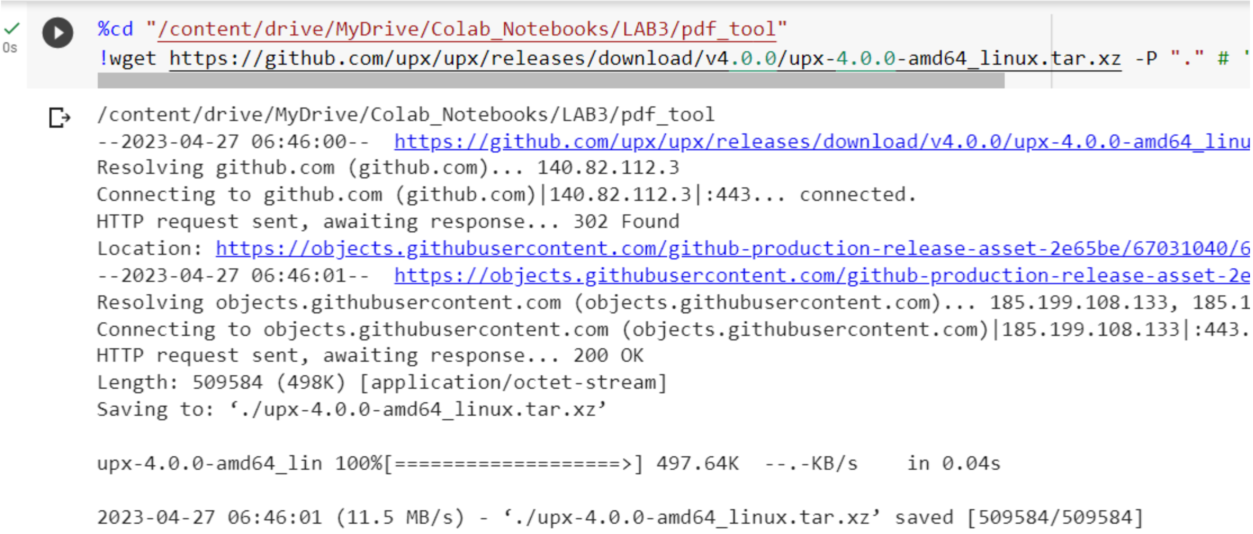
```
#Biên dịch mô hình và chọn batch size
model.compile(optimizer=my_opt, loss="binary_crossentropy",metrics=["acc"])
model.summary()
batch_size = 16
num_batches = int(num_samples / batch_size)
```

```
Model: "model_1"
__________________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_2 (InputLayer)           [(None, 8, 15000)]   0           []

 conv1d_3 (Conv1D)              (None, 1, 32)        61440032    ['input_2[0][0]']

 conv1d_2 (Conv1D)              (None, 1, 32)        61440032    ['input_2[0][0]']

 sigmoid (Activation)           (None, 1, 32)        0           ['conv1d_3[0][0]']

 multiply_1 (Multiply)          (None, 1, 32)        0           ['conv1d_2[0][0]',
                                                                  'sigmoid[0][0]']

 relu (Activation)              (None, 1, 32)        0           ['multiply_1[0][0]']

 global_max_pooling1d_1 (Global  (None, 32)          0           ['relu[0][0]']
 MaxPooling1D)

 dense_2 (Dense)                (None, 16)           528         ['global_max_pooling1d_1[0][0]']

 dense_3 (Dense)                (None, 1)            17          ['dense_2[0][0]']

==================================================================================================
Total params: 122,880,609
Trainable params: 122,880,609
Non-trainable params: 0
__________________________________________________________________________________________________
```

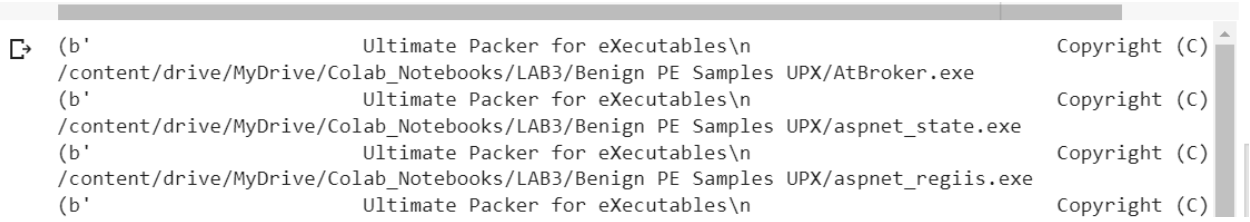


```
[ ] #Huấn luyện mô hình
    for batch_num in tqdm(range(num_batches)):
        batch = X[batch_num * batch_size : (batch_num + 1) * batch_size]
        model.train_on_batch(batch, Y[batch_num * batch_size : (batch_num + 1) * batch_size])

    #In ra kết quả
    print(model.evaluate(X, Y))

    100%|██████████| 1/1 [00:06<00:00,  6.78s/it]
    1/1 [==============================] - 0s 450ms/step - loss: 0.6498 - acc: 1.0000
    [0.6497572064399719, 1.0]
```

Câu 6: Cài đặt UPX từ https://github.com/1.upx/upx/releases, và tiến hành đóng gói các tập tin pe tại Benign PE Samples UPX

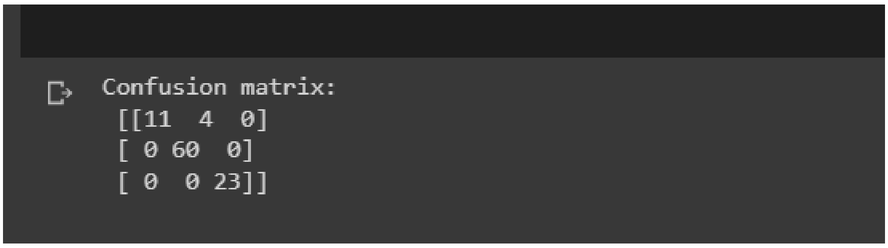


```
%cd "/content/drive/MyDrive/Colab_Notebooks/LAB3/pdf_tool"
!wget https://github.com/upx/upx/releases/download/v4.0.0/upx-4.0.0-amd64_linux.tar.xz -P "." # '

/content/drive/MyDrive/Colab_Notebooks/LAB3/pdf_tool
--2023-04-27 06:46:00--  https://github.com/upx/upx/releases/download/v4.0.0/upx-4.0.0-amd64_linu
Resolving github.com (github.com)... 140.82.112.3
Connecting to github.com (github.com)|140.82.112.3|:443... connected.
HTTP request sent, awaiting response... 302 Found
Location: https://objects.githubusercontent.com/github-production-release-asset-2e65be/67031040/6
--2023-04-27 06:46:01--  https://objects.githubusercontent.com/github-production-release-asset-2e
Resolving objects.githubusercontent.com (objects.githubusercontent.com)... 185.199.108.133, 185.1
Connecting to objects.githubusercontent.com (objects.githubusercontent.com)|185.199.108.133|:443.
HTTP request sent, awaiting response... 200 OK
Length: 509584 (498K) [application/octet-stream]
Saving to: ‘./upx-4.0.0-amd64_linux.tar.xz’

upx-4.0.0-amd64_lin 100%[===================>] 497.64K  --.-KB/s    in 0.04s

2023-04-27 06:46:01 (11.5 MB/s) - ‘./upx-4.0.0-amd64_linux.tar.xz’ saved [509584/509584]
```



```
(b'                       Ultimate Packer for eXecutables\n                         Copyright (C)
/content/drive/MyDrive/Colab_Notebooks/LAB3/Benign PE Samples UPX/AtBroker.exe
(b'                       Ultimate Packer for eXecutables\n                         Copyright (C)
/content/drive/MyDrive/Colab_Notebooks/LAB3/Benign PE Samples UPX/aspnet_state.exe
(b'                       Ultimate Packer for eXecutables\n                         Copyright (C)
/content/drive/MyDrive/Colab_Notebooks/LAB3/Benign PE Samples UPX/aspnet_regiis.exe
(b'                       Ultimate Packer for eXecutables\n                         Copyright (C)
```

Câu 7 : Cho biết kết quả đánh giá.



```
Confusion matrix:
 [[11  4  0]
 [ 0 60  0]
 [ 0  0 23]]
```