****

Faculty of Computers and Artificial Intelligence

Computer Science Department

2021/2022

**CS 395 Selected Topics in CS-1**

**Research Project**

Report Submitted for Fulfillment of the Requirements and ILO’s for Selected Topics in CS-1 course for Fall 2021

Team No. 24

|  |  |  |  |
| --- | --- | --- | --- |
|  | ID | Name | Grade |
|  | 201900439 | عبدالرحمن محمود محمود عبدالباقي |  |
|  | 201900419 | عبدالرحمن سيد فتحي |  |
|  | 201900384 | ضحى علي مصطفي عبدالرحمن |  |
|  | 201900600 | ماريان نادر موريس جرجس |  |
|  | 201900149 | اسماء محمد عابدين توفيق |  |
|  | 201900176 | امنيه محمد احمد محمد جلال |  |
|  | 201900298 | رضوان أحمد محمد عاطف |  |

Delivered to:

**Eng. Islam Gamal**

I. NUMERICAL DATASET

1. Project Introduction

* 1. **Dataset Name**

(What is the dataset used?)

Diabetes dataset.

* 1. **Number of classes and their labels**

(Specify number of classes and their labels.)

The number of classes is 7 classes,

Their labels are [Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, Age, Outcome].

* 1. **Dataset Samples Numbers**

(The total number of samples in dataset)

The total number of samples is 768 samples.

* 1. **Training, Validation and Testing**

(The number of samples used in training, validation and testing.)

The number of samples used in testing is 20%

x-test (144,7) , y-test (144,)

The number of samples used in training is 80%

x-train (575,7) , y-train (575,)

The number of samples used in validation is:

x-test (144,7) , y-test (144,)

1. Implementation Details
   * 1. **Extracted Features**

(How many features were extracted, their names, the dimension of resulted features)

The number of features is 6 features,

Their names are [Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, Age],

And the dimension of resulted features is (768,1)

* + 1. **Cross-validation**

(Is cross-validation is used in any of implemented models? If yes, specify the number of fold and ratio of training/validation)

None

* + 1. **Artificial Neural Network (ANN)**
* **Hyper-parameters**

(Specify all the hyper-parameters (initial learning rate, optimizer, regularization, batch size, no. of epochs…) with their specified value in implementation)

Optimizer is ‘Adam’

Learning rate is default

Regularization is default

Batch size is 64

No.epochs is 1000

* + 1. **Support Vector Machine** **(SVM)**
* **Hyper-parameters**

(Specify all the hyper-parameters (optimizer, regularization, …) with their specified value in implementation)

regularization is default

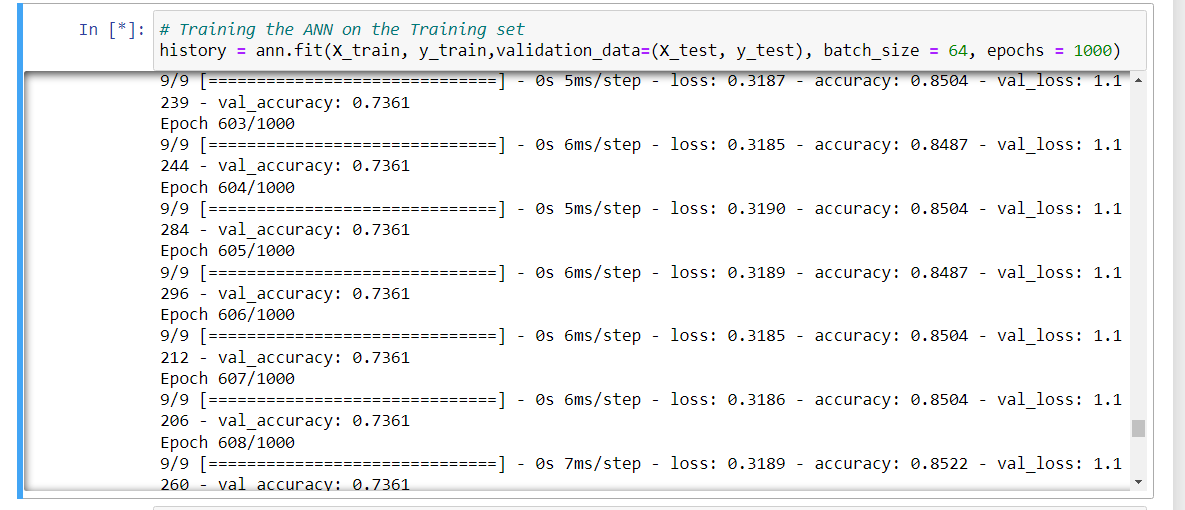
Kernel :linear

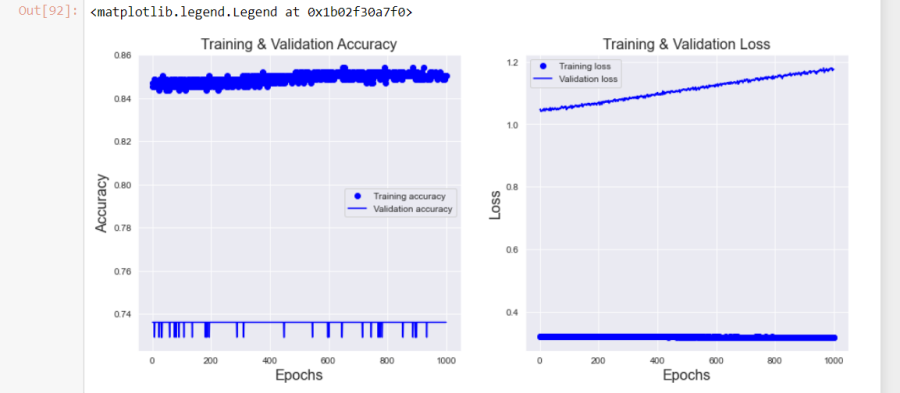
random\_state =0

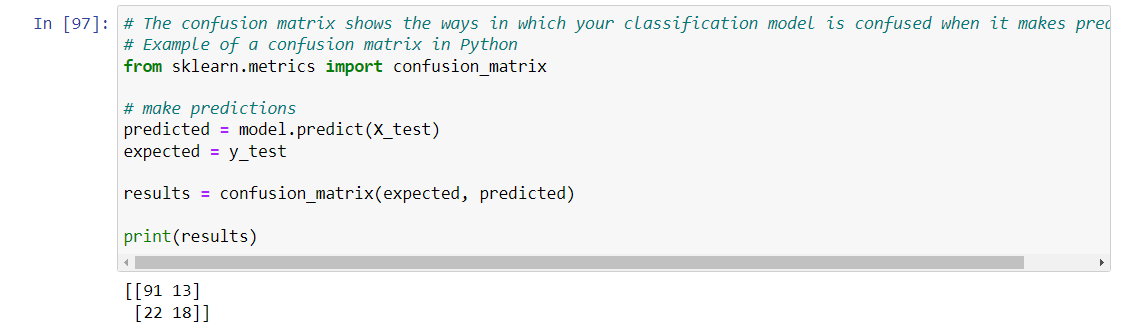
1. Models Results

**For each model you should show all these results for your model on testing data** (loss curve, accuracy, confusion matrix, ROC curve)

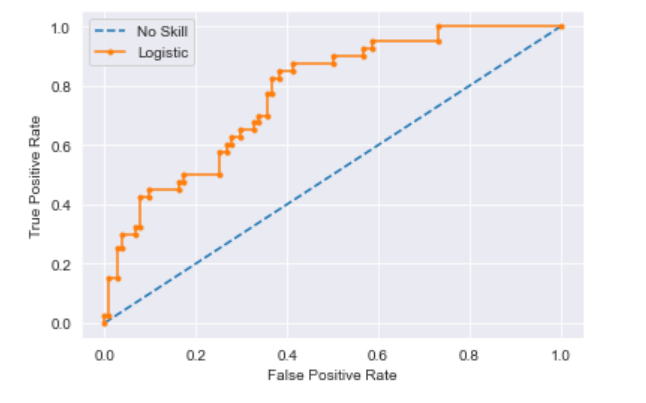
* 1. **ANN Results**
     + Accuracy is 85%

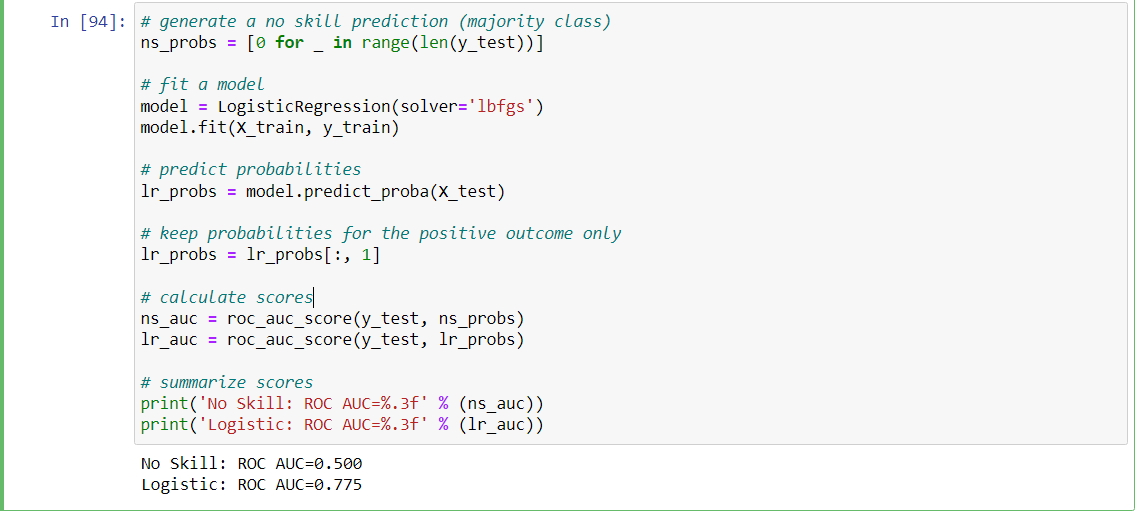


* + - Loss curve:
    - Confusion matrix:



* + - ROC curve:





* 1. **SVM Results**

**-** Accuracy is 77,08333333333334%

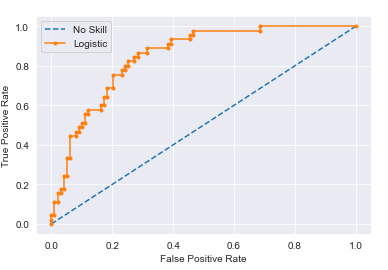
****

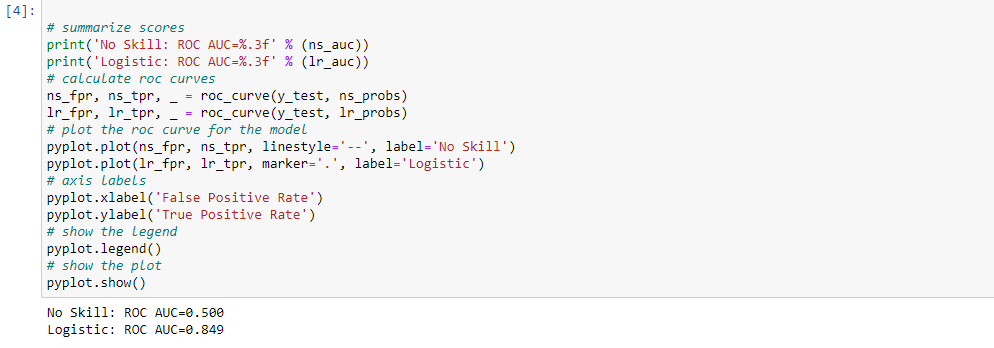
**-** Confusion matrix:

[[85 14]

[19 26]]

* ROC curve:





II. IMAGE DATASET

1. Project Introduction

* 1. **Dataset Name**

(What is the dataset used?)

Dogs vs Cats dataset.

* 1. **Number of classes and their labels**

(Specify number of classes and their labels.)

**//**

* 1. **Dataset Images Numbers and size**

(The total number of images in dataset and the size of each.)

1000 images.

Around 20 KB.

* 1. **Training, Validation and Testing**

(The number of images used in training, validation and testing.)

1000 images for training.

200 images for validation.

100 images for testing

2. Implementation Details

* + 1. **Extracted Features**

(How many features were extracted, their names, the dimension of resulted features)

//

* + 1. **Cross-validation**

(Is cross-validation is used in any of implemented models? If yes, specify the number of fold and ratio of training/validation)

No, not used.

* + 1. **Artificial Neural Network (ANN)**
* **Hyper-parameters**

(Specify all the hyper-parameters (initial learning rate, optimizer, regularization, batch size, no. of epochs…) with their specified value in implementation)

Initial learning rate: 0.0001

Optimizer: Adam algorithm

Regularization: default

Batch size :100

No. Of epochs: 10

* + 1. **Support Vector Machine** **(SVM)**
* **Hyper-parameters**

(Specify all the hyper-parameters (optimizer, regularization, …) with their specified value in implementation)

optimizer: Adam

regularization: default

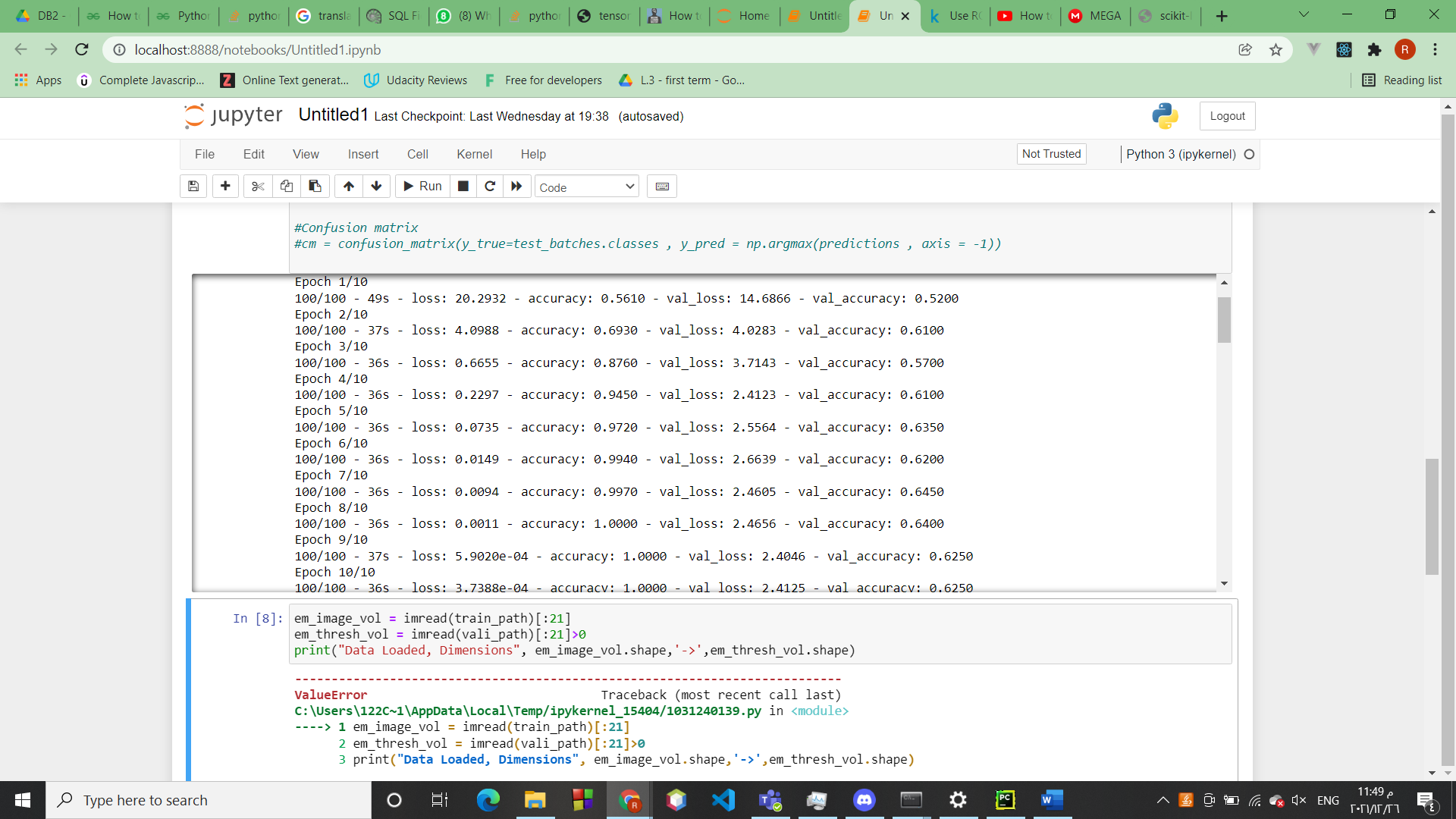
3. Models Results

**For each model you should show all these results for your model on testing data** (loss curve, accuracy, confusion matrix, ROC curve)

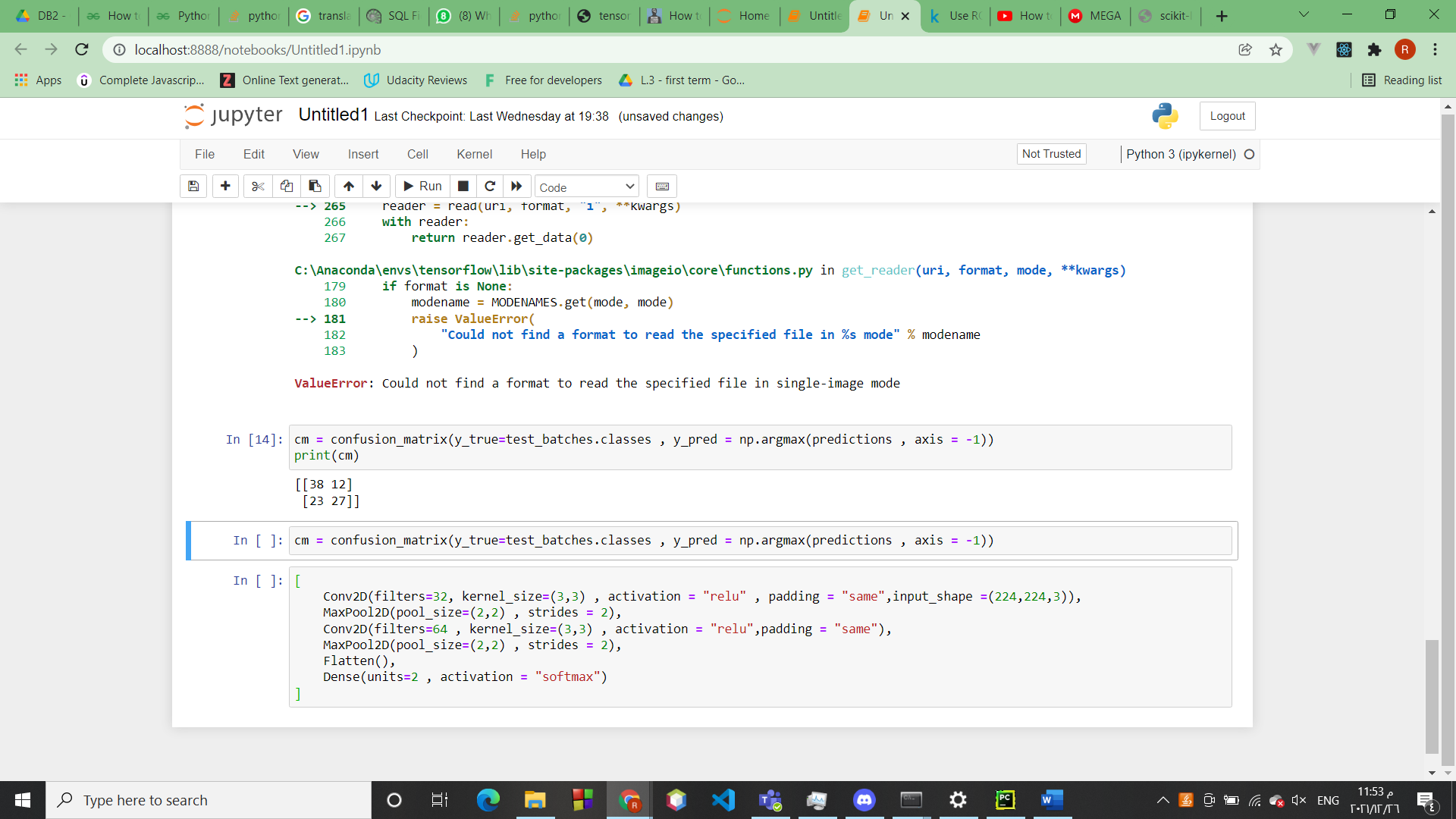
* 1. **ANN Results**

Loss curve: //

Accuracy: 100%



Confusion matrix:



ROC Curve:

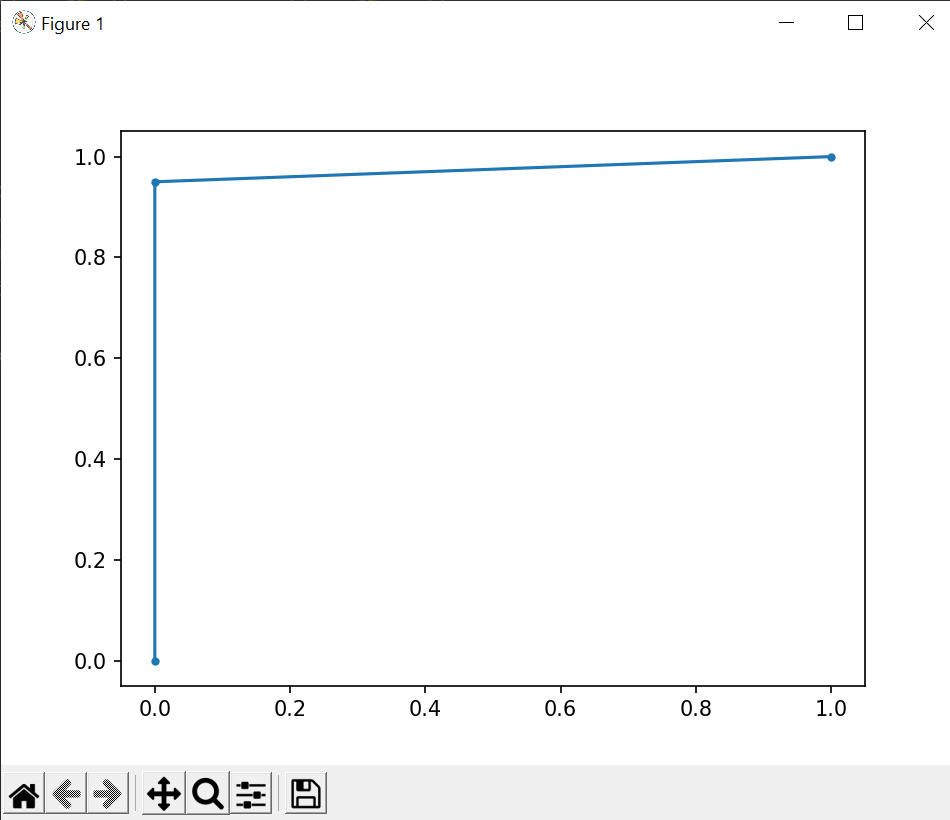
* 1. **SVM Results**

Accuracy:

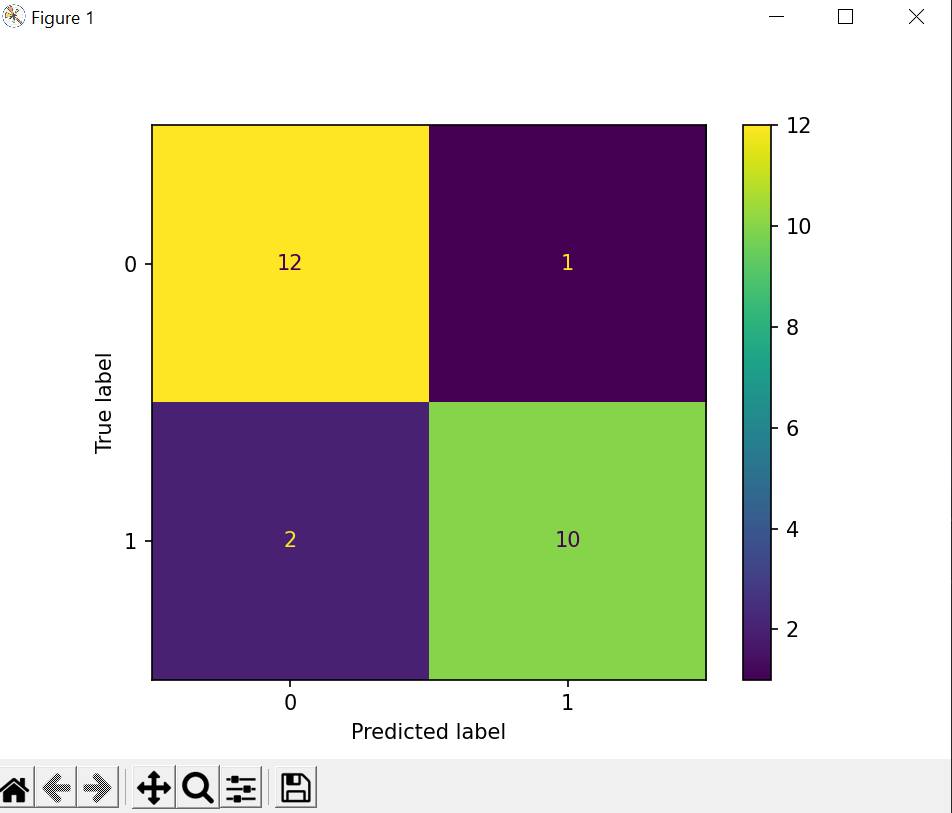
Text

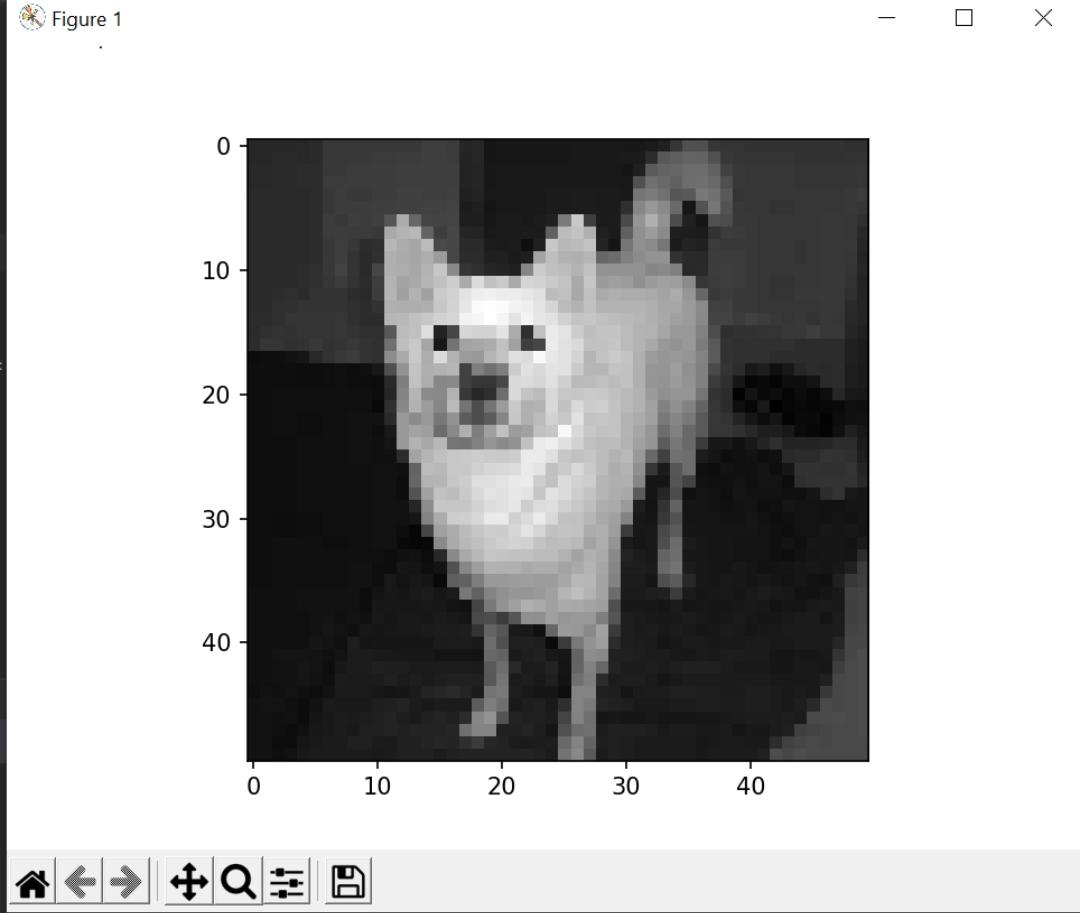
Description automatically generated

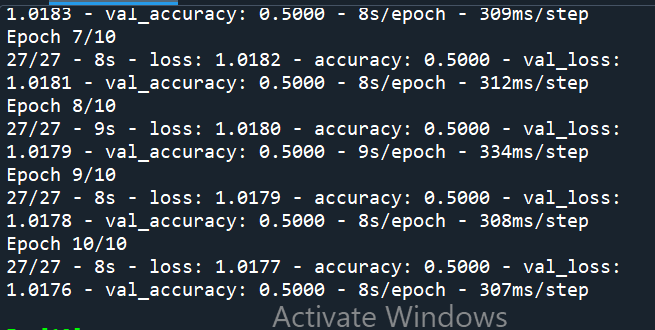
ROC Curve:

****

Matrix:





Another svm model:

Thank You