****

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. 38

|  |  |  |  |
| --- | --- | --- | --- |
|  | ID | Name | Grade |
|  | 201900547 | غيداء الطاهر احمد محمد | 3rd Grade |
|  | 201900534 | عمر نبيل صابر ندا | 3rd Grade |
|  | 201900573 | كاترين حبيب جورج ناداب | 3rd Grade |
|  | 201900554 | فاروق ابراهيم فاروق عبداللطيف | 3rd Grade |
|  | 201900568 | فريدة أحمد توكو سعيد | 3rd Grade |
|  | 201900688 | محمد شريف فتحي محمد سالم | 3rd Grade |
|  | 201900386 | طارق ايمن الدسوقي طلبة | 3rd Grade |

Delivered to:

**Dr. Wessam El-Behaidy**

**Eng. Islam Gamal**

**Eng. Muhammed Kamal**

I. NUMERICAL DATASET

1. Project Introduction

* 1. **Dataset Name**

(What is the dataset used?)

**Pima Indians Diabetes**

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

(Specify number of classes and their labels.)

**Classes = 1 , it’s labels (0 and 1)**

**1 = True (Diabetic)**

**0= False (not diabetic(**

* 1. **Dataset Samples Numbers**

(The total number of samples in dataset)

**786**

**Number of rows in the dataset 768**

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

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

**in training = 692**

**in testing = 76**

**validation = 0.39**

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

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

**Features = 8**

**Features’names :**

**Pregnancies**

**Glucose**

**BloodPressure**

**SkinThickness**

**Insulin**

**BMI**

**DiabetesPedigreeFunction**

**Age**

**dimension of resulted : one dimension**

* + 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)

**Yes , we did .**

**Number of fold = 10**

**We have a ratio of 35/65**

* + 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 =** 0.01

**optimizer =** opt from SGD

**regularization =** 0.01

**batch size =** 10

**epoch =** 100

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

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

**Regularization** = 1.0

**kernel**= linear

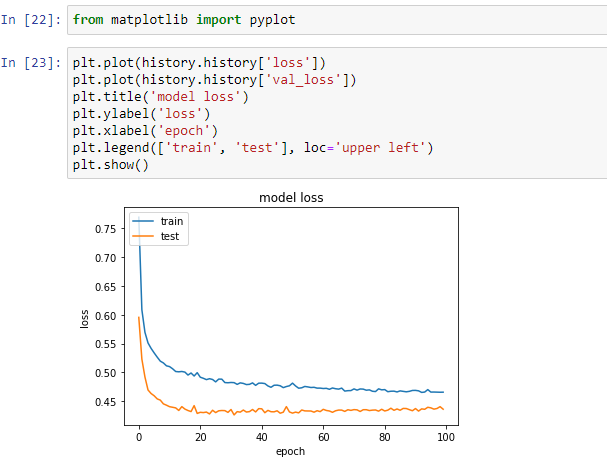
**degree**= 3

1. Models Results

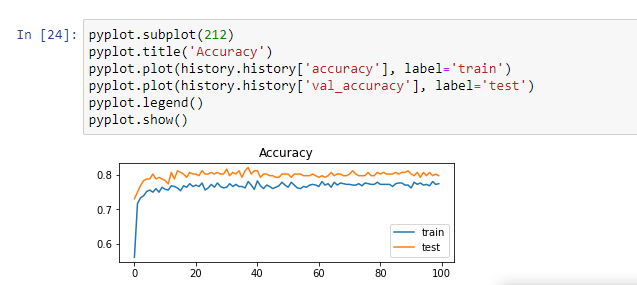
**For each model you should show all these results for your model on testing data** (loss curve, accuracy, confusion matrix, ROC curve)

* 1. **ANN Results**

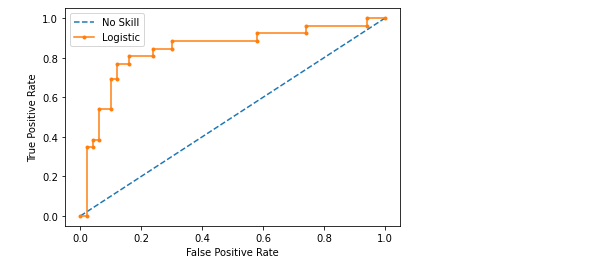
loss curve :

****

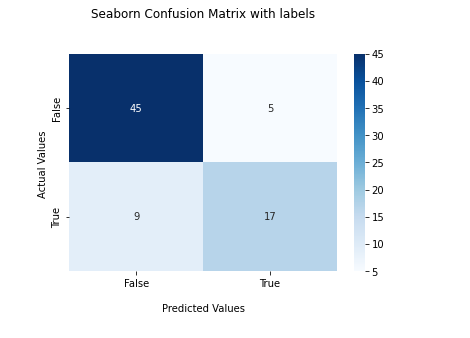
Accuracy :

****

ROC curve :

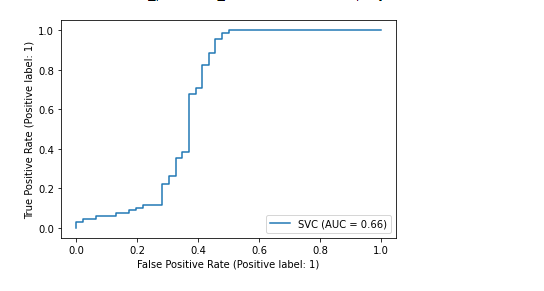
****

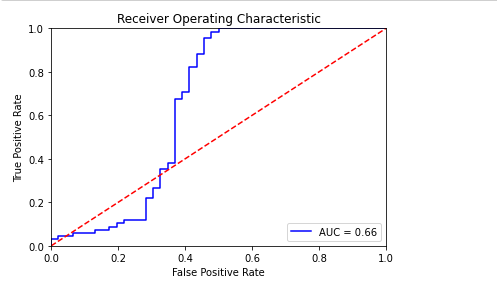
confusion matrix :

****

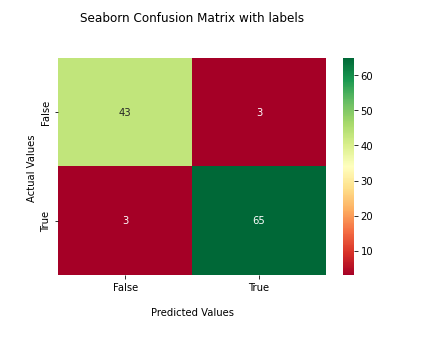
* 1. **SVM Results**

ROC & Auc curve :





confusion matrix :



II. IMAGE DATASET

1. Project Introduction

* 1. **Dataset Name**

(What is the dataset used?)

# Alien vs. Predator images

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

(Specify number of classes and their labels.)

**Classes :** 2

**Labels :** Alien , Predator

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

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

894 – (50 \* 50)

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

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

694

200

2. Implementation Details

* + 1. **Extracted Features**

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

694

(50 \* 50 \* 3)

* + 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)

* + 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:** 0.001

**Optimizer:** adam

**batch size:** 32

**no. of epochs:**15

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

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

Regularization:

train\_x, test\_x = train\_x / 50.0, test\_x / 50.0

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**

Graphical user interface

Description automatically generated

* 1. **SVM Results**

confusion matrix :

Graphical user interface

Description automatically generated with medium confidence

ROC curve:

Graphical user interface, application

Description automatically generated

Learning rat :

Graphical user interface, text, application

Description automatically generated