Cover sheet

***Faculty of computers and AI Helwan U***

Selected Topics from CS1:

**Team no : 14**

|  |  |
| --- | --- |
| **Amira wael** |  |
| **Thomas fathy** |  |
| **Joseph issa** |  |
| **John halim** |  |
| **Marina ezzat** |  |
| **Maer reda** |  |
| **Ramy Emad** |  |

# **Description** document:

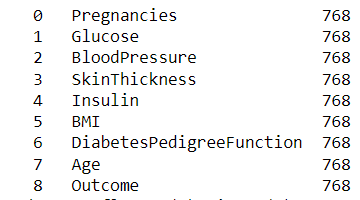
**1 : LR**

***A-General Information on dataset:***

## **-NAME :** Pima Indians Diabetes Database

https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

**-CLASSES** : 8 classes and the outcome column



**-the number of samples used in training, validation and testing.**

NO of training samples : 90 %

NO of testig samples : 10%

***B-Implementation details:***

* **Data Cleansing :**

1. **detecting Missing Values and replacing it with np.nan**
2. **- defining data thresholds and replacing the outliers with the min\max possible values**
3. **Replacing missing values with mean\mode**

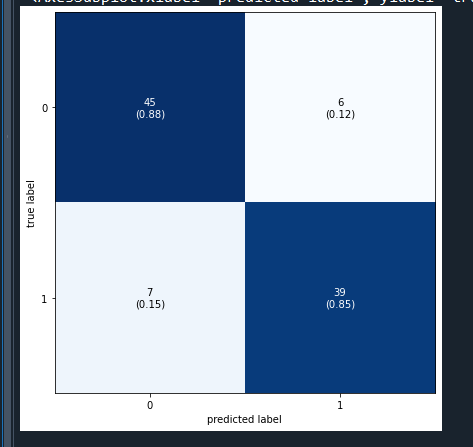
# **Data preprocessing :**

1. **extracting input and output features**
2. **Handling Class Imbalance Problem**

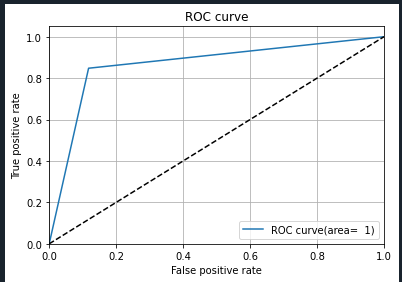
* **Splitting data**
* **Apply LR model**

****

**Confusion matrix :**

****

**ROC curve :**

****

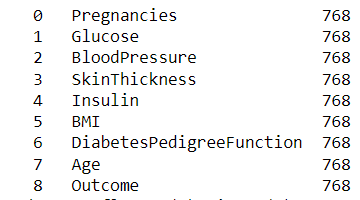
**2:SVM**

***A-General Information on dataset:***

## **-NAME :** Pima Indians Diabetes Database

https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

**-CLASSES** : 8 classes and the outcome column



**-the number of samples used in training, validation and testing.**

NO of training samples : 90 %

NO of testig samples : 10%

***B-Implementation details:***

* **Data Cleansing :**

1. **detecting Missing Values and replacing it with np.nan**
2. **- defining data thresholds and replacing the outliers with the min\max possible values**
3. **Replacing missing values with mean\mode**

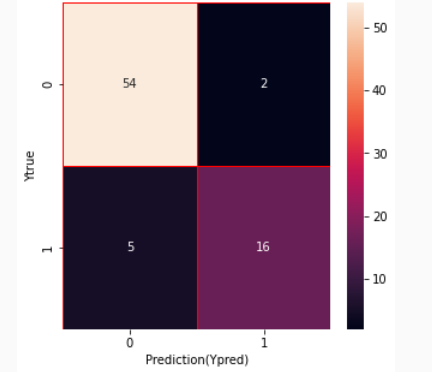
# **Data preprocessing :**

1. **extracting input and output features**
2. **applying normalization on data by using minmax scaling**

* **Splitting data**
* **Applying SVM model**

***C- Results details:***

****

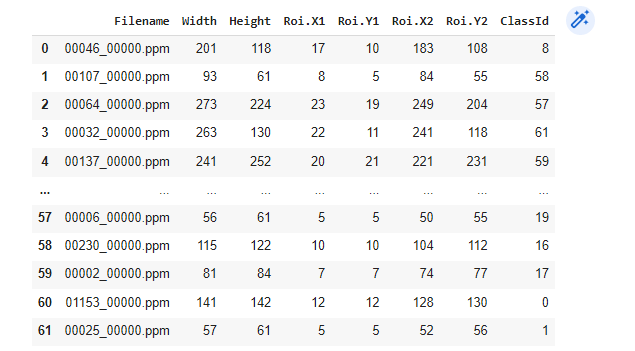
****

**3: SVM**

***A-General Information on dataset:***

## **-NAME : Traffic sign dataset**

**-CLASSES** : 62 classes and the outcome column



Total number of samples :2534

Their size :



**-the number of samples used in training, validation and testing.**

NO of training samples : 90 %

NO of testig samples : 10%

***B-Implementation details:***

1. **Calling the training datasets**
2. **- read the CSV of the training**
3. **read the CSV of the testing**

# **Data preprocessing :**

1. **Return it to hog**
2. **Make the feature extraction**

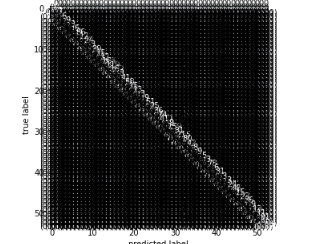
* **Apply SVC model**

***C- Results details:***

**Accuracy :**

**Screenshot 2022-12-19 234917**

**Confusion matrix :**

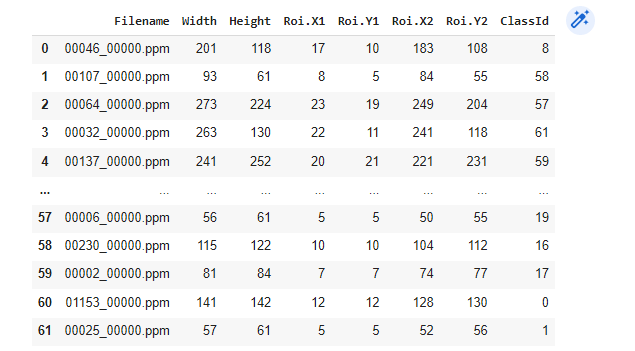
****

**4: ANN Model**

***A-General Information on dataset:***

## **-NAME : Traffic sign dataset**

**-CLASSES** : 62 classes and the outcome column



Total number of samples :2534

Their size :



**-the number of samples used in training, validation and testing.**

NO of training samples : 90 %

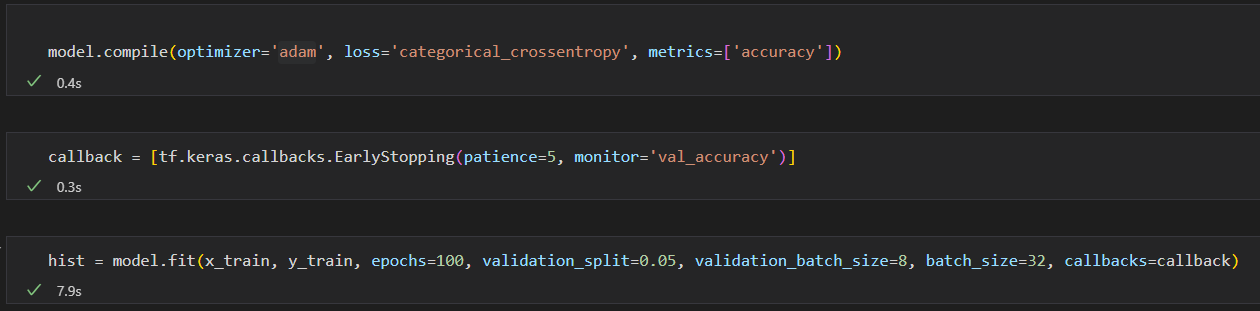
NO of testing samples : 10%

***B-Implementation details:***

* **Feature Extraction:**

**Text

Description automatically generated**

* **Hyperparameters:  
  **

***C- Results details:***

* **Accuracy** (87.54%)

Text

Description automatically generated

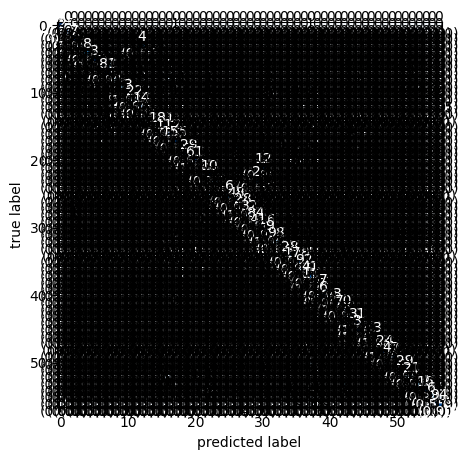
* **Accuracy Curve:**

**Chart, line chart

Description automatically generated**

* **Loss Curve:  
  Chart, line chart

  Description automatically generated**
* **Confusion Matrix:**

****