Cairo University
Faculty of Computers and Artificial Intelligence

Learning project

Menna-Allah Hamdy	20190558

Course Name**:** *Learning from data*

Under supervision of

Dr. Elshimaa Elgendi

Spring 2021-2022

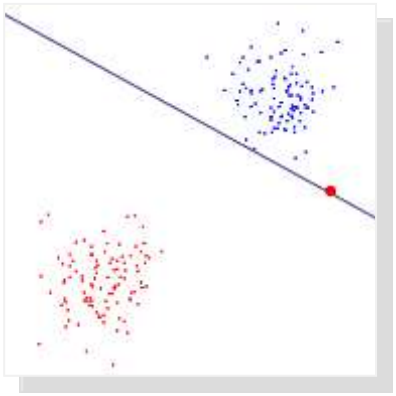
Developing computer programs coded by python that detects the genuine banknote using:

- SVM with hard margins coded by Scikit
- SVM with soft margins coded by Scikit

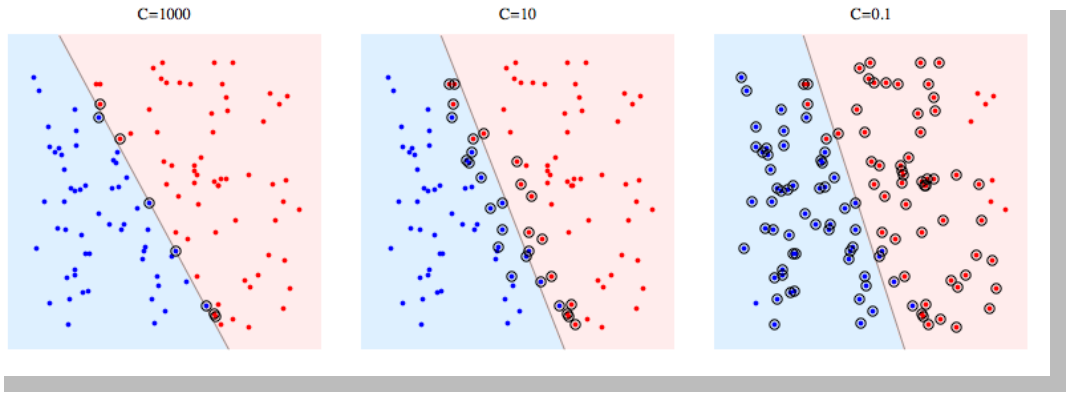
Support Vector Machine -SVM-

- For two-group classification problems, a support vector machine (SVM) which is a supervised machine learning model that employs classification or regression problems. SVM models can categorize fresh text after being given sets of labelled training data for each category.
  - They have two key benefits over newer algorithms like neural networks: greater speed and greater performance with a limited number of samples (in the thousands). This makes the algorithm particularly well suited to text classification problems, where it's usual to only have access to a few thousand labelled samples in dataset.
  - There're two types of SVM (Soft Margin SVM & Hard Margin SVM) there are differences between them, like that:
    - Hard Margin SVM is quite sensitive to outliers.
    - Soft Margin SVM avoids iterating over outliers.

    In this diagram you can notice overfitting of hard margin SVM.



  - Soft-margin SVM can choose a decision boundary that has non-zero training error even if the dataset is linearly separable, and is less likely to overfit. You can notice that decreasing C value causes the classifier to leave linear separability in order to gain stability.

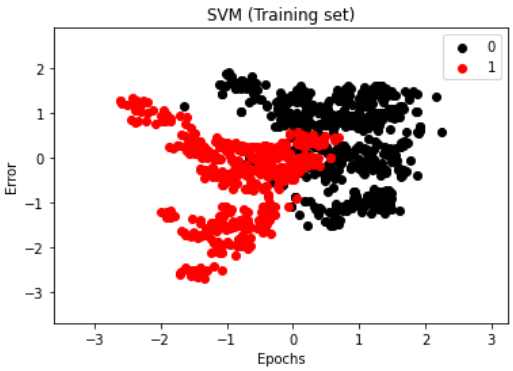


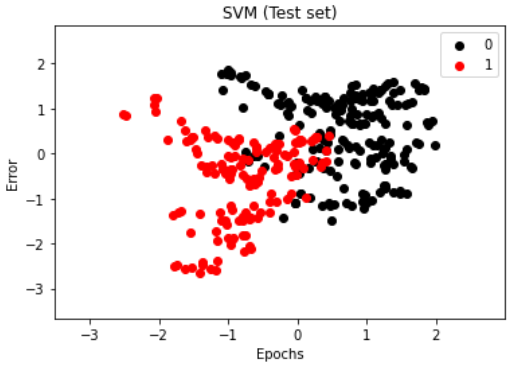
  - You need to set the best value of C in SVM. The more optimal value of C will give you more accuracy.

SVM Hard Margin

- Kernal type : Linear
- C = 1

```
[[189   6]
 [  0 148]]
```





SVM Soft Margin

- Kernal type : Linear
- C = 1

```
[[587  19]
 [  2 490]]
              precision    recall  f1-score   support

           0       1.00      0.97      0.98       606
           1       0.96      1.00      0.98       492

    accuracy                           0.98      1098
   macro avg       0.98      0.98      0.98      1098
weighted avg       0.98      0.98      0.98      1098
```