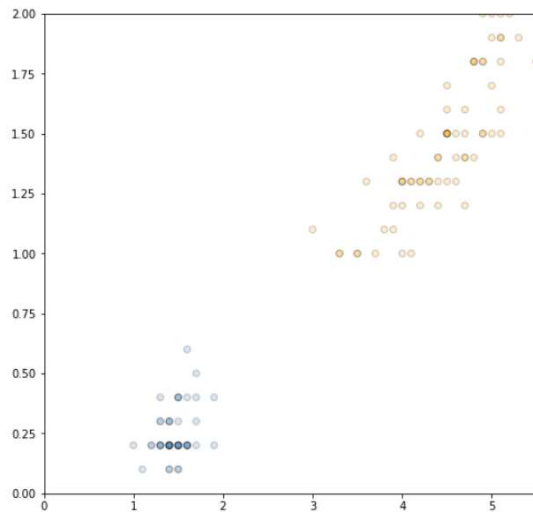


# Report for Assignment 2

-Debanshu Biswas (M20MA053)

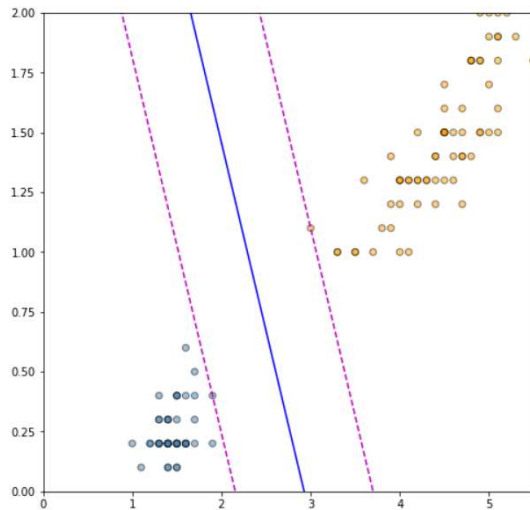
## Hard SVM using cvxopt

We used the iris dataset from sklearn.datasets. There we used only two linearly separable attributes “petal length” and “petal width. And our target column name is “target”. Then we plotted the points with axis petal length and petal width.



Then we used CVXOPT to compute  $w$  and  $b$ . And we got  $w = [1.29411773 \ 0.82352938]$  and  $b = -3.7882354695378924$ .

Below is the diagram of the points with optimal solution, margin and support vectors.



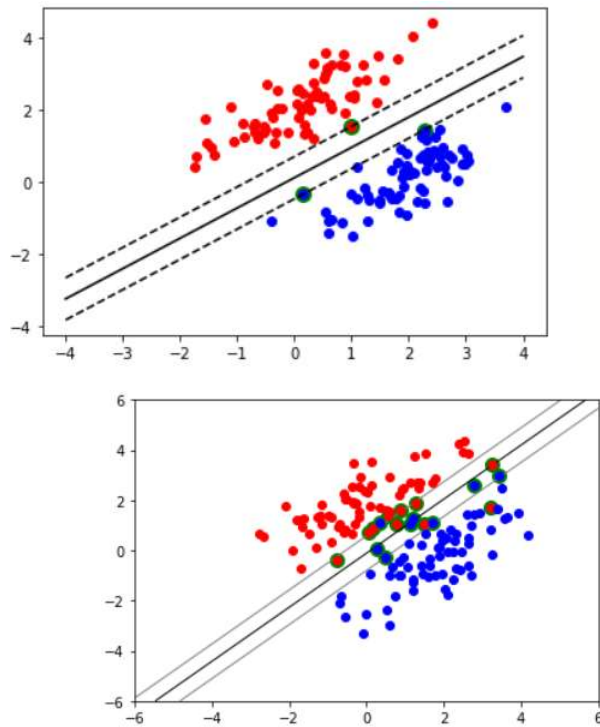
**Soft SVM using cvxopt**

We first defined svm class and defined three functions in t fit, project and predict.

Then we generated linearly separable data , linearly non separable data and linearly separable overlapping data with 100 entries each. The results are below.

| Type of Split | No of Support vectors for X_train, y_train |                             |                                     |
|---------------|--|-----------------------------|-------------------------------------|
|               | Linearly separable data                    | Linearly non separable data | Linearly separable overlapping data |
| 70:30         | 3 of 140                                   | 9 of 140                    | 18 of 140                           |
| 80:20         | 3 of 160                                   | 10 of 160                   | 14 of 160                           |
| 90:10         | 3 of 180                                   | 10 of 180                   | 9 of 180                            |

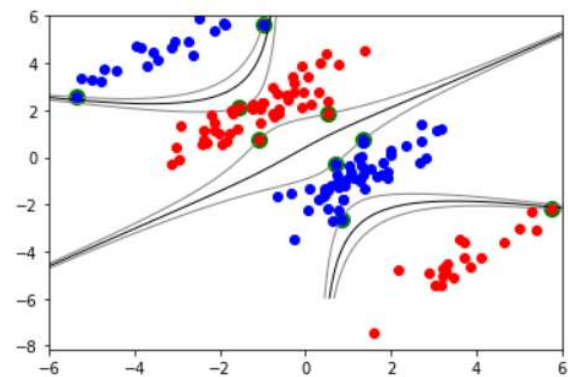
| Type of Split | Accuracy                |                             |                                     |
|---------------|-------------------------|-----------------------------|-------------------------------------|
|               | Linearly separable data | Linearly non separable data | Linearly separable overlapping data |
| 70:30         | 1                       | 0.95                        | 0.967                               |
| 80:20         | 1                       | 1                           | 1                                   |
| 90:10         | 1                       | 1                           | 1                                   |



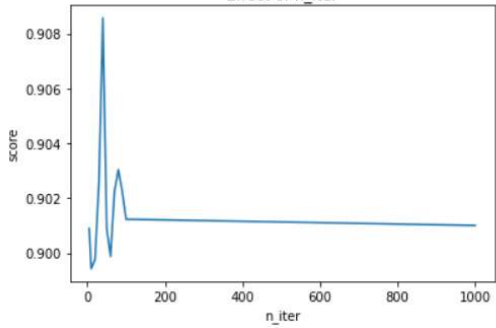
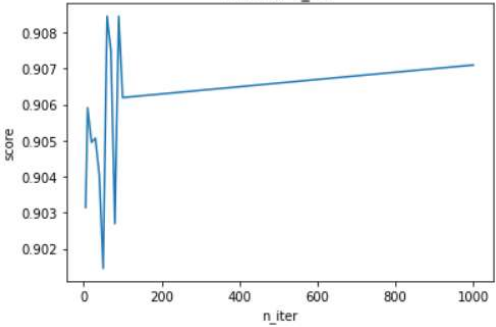
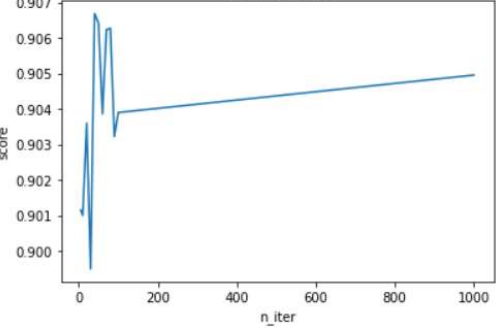
**Soft SVM using SGD**

We used one dataset “dataset.csv”. Then we used train test split to split the dataset in train and test sets. Then we used SGDClassifier from sklearn.

Then we are plotting score vs n\_iter for different splitting,



| Type of Split | Accuracy           |
|---------------|--------------------|
| 70:30         | 0.9003273507167852 |

|       |   |
|-------|---|
|       | <p>Effect of n_iter</p>  <p>score</p> <p>n_iter</p>                            |
| 80:20 | <p>0.9083417993001467</p> <p>Effect of n_iter</p>  <p>score</p> <p>n_iter</p>  |
| 90:10 | <p>0.9031869661737593</p> <p>Effect of n_iter</p>  <p>score</p> <p>n_iter</p> |