**Support Vector Machines**

*Exercise Session 1 - Classification,*

*report by,*

*Marimuthu Ananthavelu (r0652832), KU Leuven.*

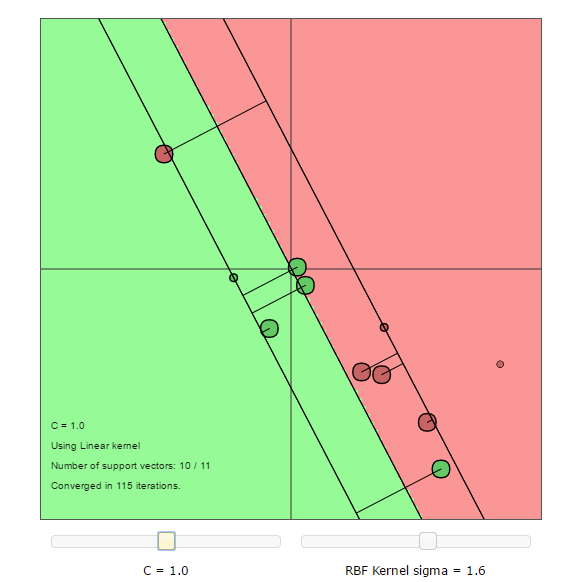
* 1. Two Gaussians

Geometric construction of a classifier line looks complicated given the data points are spread across each other. The classifying line should ideally be a decision boundary between the two classes namely ‘versicolor’ and ‘virginica. In our case, the decision boundary cannot be a straight line considering the nature of the data points which all are pertaining to two different classes. When the decision boundary separates these 2 classes with clearly indicating all the data points pertaining to ‘versicolor’ on one side and all the data points pertaining to‘virginica’ on the other side, then we can consider that the classifier is optimal/valid.

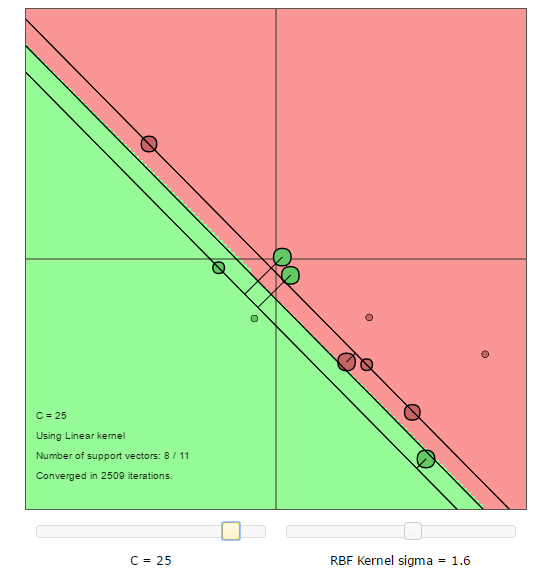
* 1. The Support Vector Machines
     1. When adding the new data points to the vector space the following happens;
* Each additional data point gets added to the dataset
* Decision boundary adjusts itself with respect to the modified dataset and accordingly the number of iterations gets modified with respect the inclusion of data point and re-optimization.
* When the new data point is added to the extreme left or right sides of the decision boundary, the existing decision boundary changes its position drastically (making a larger difference in the positional points in the vector space) considering the new data point.
  + 1. When adding an outlier, the following happens;
* With the same given parameters of ‘C’ and ‘kernel sigma’, the decision boundary changes with respect to the new data point and the new support vector is getting added to the side opposite of the class in which the outlier is present.
  + 1. C is the parameter for the soft margin cost function, which controls the influence of each individual support vector.

Changing the C value to different values the following are the observations;

* The increase in C parameter, makes the classifier to be more accurate as possible by penalising the misclassification heavily.
* The decrease in C parameter, makes the classifier to be leaving some of the data points under the wrong class which makes the classifier less robust.
* For example, with C parameter of ‘1’, the classifier is as below;



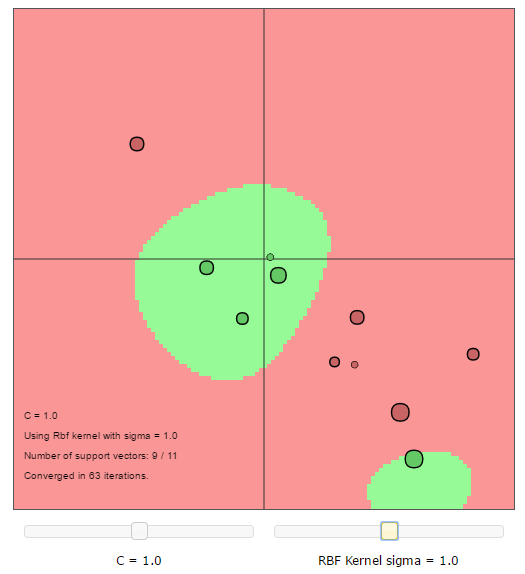
With C parameter of ‘25’, the classifier is as below;



* In Fig-1, with C=1, the classifier is not robust and classifies one or some of the data points on the wrong side of the decision boundary.
* In Fig-2, with C=25, the classifier attempts to be more robust and creates a very closely possible decision boundary especially giving importance to the data points which all are outliers.

***Using RBF kernel:***

* + 1. Switching over to RBF kernel gives the decision boundary which is not a straight line but closed over the different cluster(s) of data points. This clearly indicates that the RBF can be used for non-linearly separable data points. The classifier space looks as below;



Sigma is a parameter of the Gaussian kernel and defines the steepness of the rise around the landmark.

* 1. Using LS-SVMlab
  2. Homework Problems
     1. The Ripley Data-set
     2. Breast Cancer Dataset
     3. Diabetes Database