

# Lab 4

Santiago Bernal  
Rodrigo Arias

January 7, 2018

## Exercise 1

In the file `ex1.py` three generators are used to provide 3 different random datasets. The SVM algorithm is applied to the train set of each dataset, and is plotted in the figures 1, 2 and 3. The train data is shown with circle markers, while the test data uses crosses. The SVM decision regions are also filled with color, in order to visually see the behavior of the classifier.

We see that the datasets 1 and 3 are linearly separable, so a linear function should be used. The Gaussian kernel is used in the latter, but is behaving similarly to a linear kernel, so the classification is correct. The dataset 2, in contrast, is clearly not linearly separable, so a linear function should not be used. In the figure 4 a Gaussian kernel was used, to see the difference. We see a better classification as expected.

The figure 2 shows a correct classification of 100%, even if the classifier is not accurate. That can be explained because the test set provided with the generator, is only choosing elements that are placed in the outermost groups.

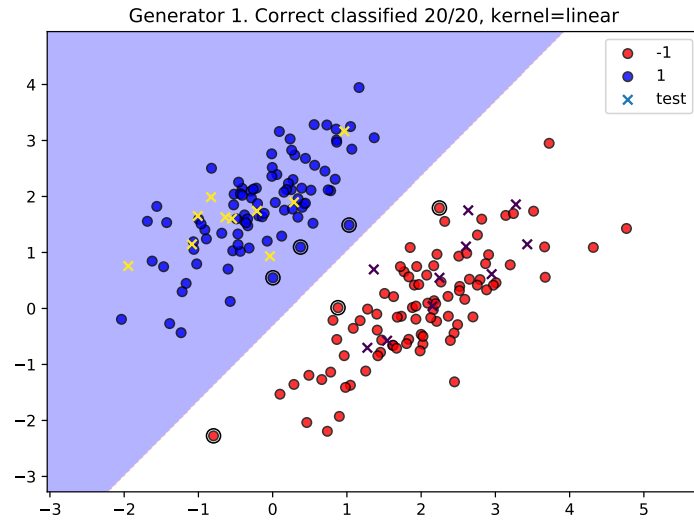


Figure 1: Dataset of the generator 1

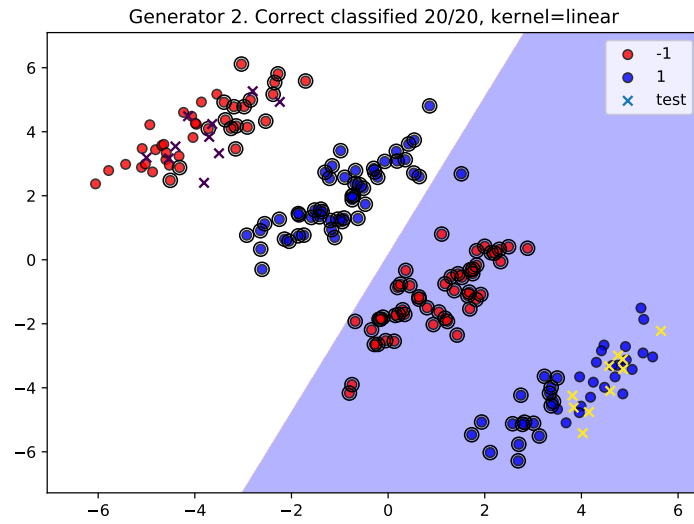


Figure 2: Dataset of the generator 2

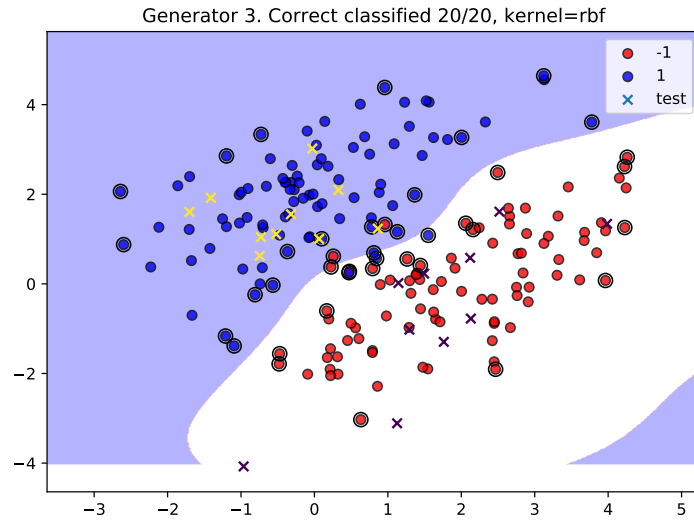


Figure 3: Dataset of the generator 3

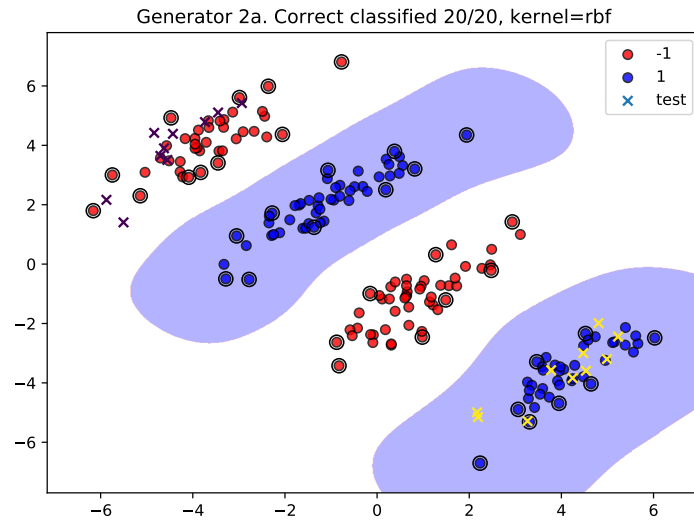


Figure 4: Dataset of the generator 2 using a Gaussian kernel