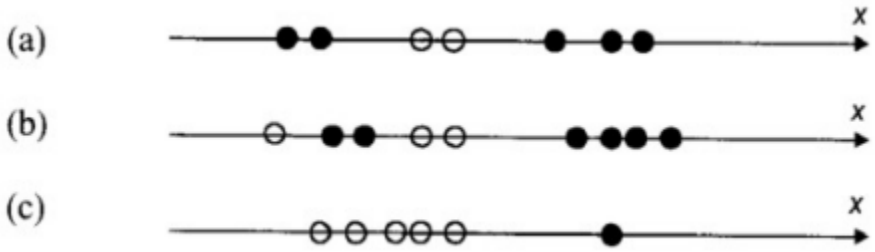
Exercise 7 (17.11.2014)

1) Read chapter 6 from Haykin's book (2nd edition). Summarize or sketch your insights in mind-map or an outline or a summary.

2) The graphs below represent three different one-dimensional classification (dichotomization) tasks (along a sketched x-axis, dash means "no data point").
What is the lowest-order polynomial decision function that can correctly classify the given data? Black dots denote class 1 with target function value y1 = +1 and white dots depict class 2 with targets y2 = -1. What are the decision boundaries?



If you wanted to classify the data sets (a), (b), (c) using SVM's with Gaussian basis functions, how many hidden layer neurons would you need for each problem?

3) Download LIVSVM from http://www.csie.ntu.edu.tw/~cjlin/libsvm/. Study how to use it with MATLAB or python. In this exercise you have to use this.
Figure 1.8 on next page shows a pair of "moons" facing each other in an asymmetrically arranged manner. The moon labeled "Region A" is positioned symmetrically with respect to the y-axis, whereas the moon labeled "Region B" is displaced to the right of the y-axis by an amount equal to the radius r and below the x-axis by the distance d. The two moons have identical parameters:

Radius of each moon, r = 10,
Width of each moon, w = 6.

The vertical distance d separating the two moons is adjustable; it is measured with respect to the x-axis, as indicated in the figure 1.8.

- Increasingly positive values of d signify increased separation between the two moons;
- Increasingly negative values of d signify the two moons' coming closer to each other.

The training sample consists of 1000 pairs of data points, with each pair consisting of one point picked from region A and another point picked from region B, both randomly. The test sample consists of 3,000 pairs of data points, again picked in a random manner.

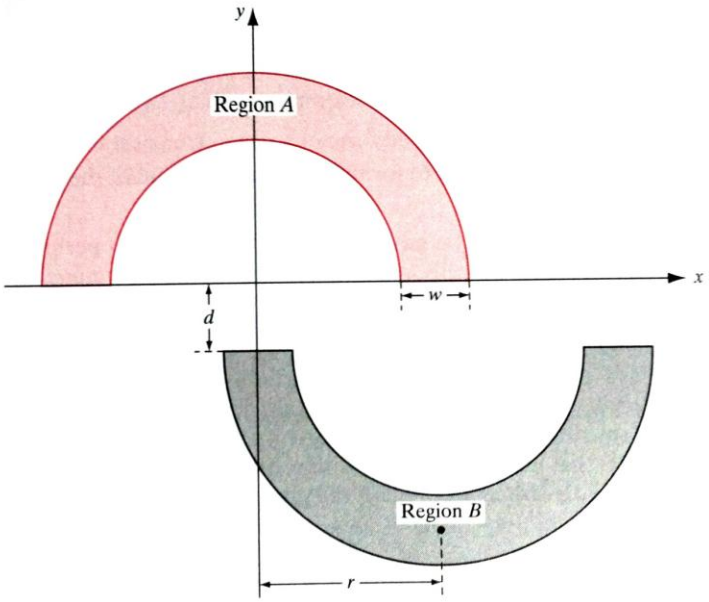


FIGURE 1.8 The double-moon classification problem.

Tasks:

Your task is to classify the dataset using SVM (Support Vector Machine) for some cases given below. Generate the dataset for each case and classify using different kernels (e.g. linear, polynomial, radial basis etc.) Show the decision boundary (Plotting the classified points using different color will be enough)

- Case 1: d = 0
- Case 2: |d| = 1/2 * (radius of moon's inner half-circle) and d is negative i.e. d is in the upper side of x-axis.
- Case 3: Increase d negatively such that both of the moons touch each other.
- Case 4: Both moons overlap each other
- Case 5: Add some noise in the training set

Try to experiment with different options in svmtrain (LIBSVM). Comment on your findings.

This experiment is taken from Haykin's book (3rd edition) which is introduced in the first chapter and continued throughout the book (chapter 2, 3, 4). For intuition, you can take a look in it.

4) Please upload 3 questions and their brief answers on the reading material in a separate txt file.