Machine Learning & Machine Learning (extended)

Practice Exercise Sheet – SVM Classification

Question 1: Given the equation of a straight line $y = -0.75x + 4$: (i) draw this line, (ii) represent it in the form $\boldsymbol{w}^T\boldsymbol{x} + b = 0$ by finding the values of $\boldsymbol{w}$ and b, (iii) draw parallel line $\boldsymbol{w}^T\boldsymbol{x} + b = 1$, (iv) draw parallel line $\boldsymbol{w}^T\boldsymbol{x} + b = -1$, (v) show the mathematical working and graphical visualization of a couple of example points for which $\boldsymbol{w}^T\boldsymbol{x} + b > 0$, (vi) show the mathematical working and graphical visualization of a couple of example points for which $\boldsymbol{w}^T\boldsymbol{x} + b < 0$.

Question 2: The following 2-dimensional data is to be used for training an SVM:

Class one: (1,1), (2,2), (2,0)

Class two: (0,0), (1,0), (0,1)

(i) Plot the training points and, by visual inspection, determine the position of the optimal margin decision boundary.

(ii) List the support vectors.

Question 3: Consider training data of 1-dimensional points from two classes:

Class 1: -5,5

Class 2: -2,1

(i) Are the two classes linearly separable?

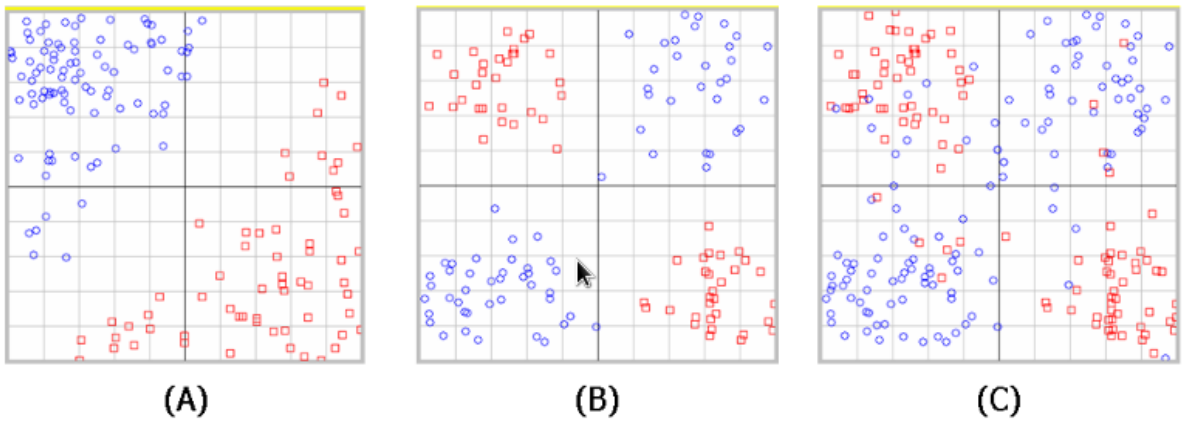
(ii) Consider the transformation φ: $R \rightarrow R^2$, $\varphi(x) = (x, x^2)$. Transform the data and plot these transformed points. Are these linearly separable?

(iii) Draw the optimal separating hyper-plane in the transformed space, and explain in one or two sentences how this linear boundary helps us to separate the original data points.

Question 4: What is the main idea behind a linear SVM? Illustrate your explanation by drawing a figure.

Question 5: How can a SVM classifier work for learning to classify data that is not linearly separable?

Question 6: Consider the three data sets illustrated below:



Each point has two numeric features (i.e. the x and y coordinates of the points). The circles and the squares represent two different classes.

(i) If you were to use SVM on these data sets, which data set will require a kernel to be used?

(ii) If you were to use SVM with Gaussian kernel on these data sets, how would you set the width parameter of the kernel?

Question 7: The XOR problem is to learn the function from the following input points to their class labels:

Class : (1,1), (-1,-1)

Class 2: (1,-1), (-1,1)

We know that support vector machine (SVM) with a kernel can solve this problem by mapping the points in a higher dimensional space. But higher dimensional spaces are difficult to visualise, and we would like to construct a support vector machine that classifies these points correctly in a 2-dimensional input space. Is this possible? If so, how? If not, explain why not.

Hint: Try to come up with a feature-transformation that stays in 2D (i.e. maps R^2 into R^2) and makes the classes linearly separable.