T-61.231 Principles of Pattern Recognition

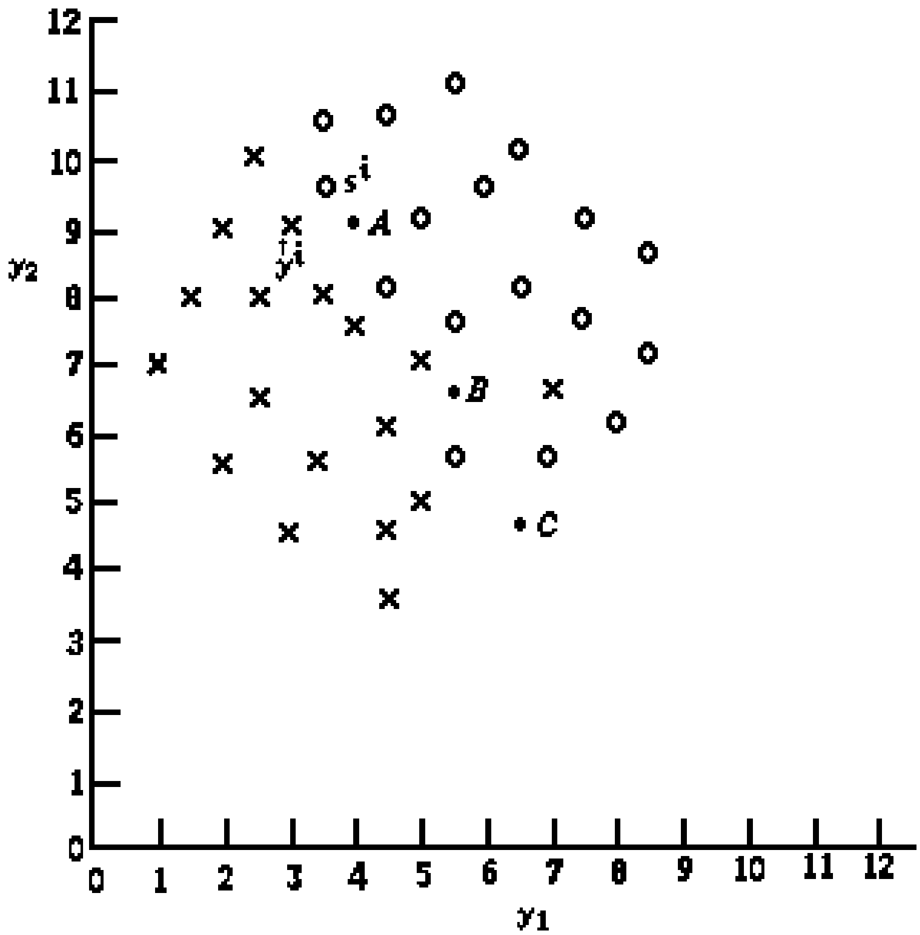
Exercise 3: 15.10.2001

1. Formulate a one-dimensional Parzen estimate for density function p_y from given samples

$$y^{(i)} : 2.5, 2.8, 3.4, 4.2, 4.5, 4.7, 5.2, 5.6, 7.5$$

   Use a rectangular window function.

2. Classify points A, B and C in figure using the 5-nearest neighbour classification rule and an Euclidean distance function.



3. Applying the nearest neighbour classification rule for two classes (figure), decision regions and boundary are received for these classes. If $\mathbf{y}^{(i)}$ are sample vectors from class 1 and $\mathbf{s}^{(i)}$ sample vectors from class 2, then according to the definition any point $\hat{\mathbf{y}}$ lying in the boundary satisfies the equation

$$\min_i d(\hat{\mathbf{y}}, \mathbf{y}^{(i)}) = \min_i d(\hat{\mathbf{y}}, \mathbf{s}^{(i)})$$

Assume distance function to be Euclidean.

a) What is the equation for a decision boundary in a small region near the boundary if constant vector $\mathbf{y}^{(i)}$ is nearest to the boundary from class 1 and $\mathbf{s}^{(j)}$ is constant vector nearest to the boundary from class 2?

b) Draw a perfect decision boundary for classes in figure.

c) Find those sample vectors, whose extraction doesn't change the decision boundary.

4. An attempt to estimate the error rate of the classifier is often made with the help of a test set (error percentage is calculated from the test vectors). Sometimes the test set is the same set with which the classifier was constructed. Show that this is not a good approach especially when the classifier is a 1-NN classifier.

5. Let us assume that two sets of points H_1 and H_2 in R^d are linearly separable. That means there is a vector $\underline{\omega}' \in R^d$ and a scalar $\omega_0 \in R^d$ so that

$$\underline{\omega}'^T \underline{x} - \omega_0 \begin{cases} > 0 & \forall \underline{x} \in H_1 \\ < 0 & \forall \underline{x} \in H_2 \end{cases}$$

Let us augment the vectors in the following way:

$$\underline{\hat{x}} = \begin{pmatrix} \underline{x} \\ -1 \end{pmatrix}, \qquad \underline{\omega} = \begin{pmatrix} \underline{\omega}' \\ \omega_0 \end{pmatrix}.$$

And let us finally change the sign of vectors $\underline{\hat{x}}$ of the set H_2.

a) Show that $\underline{\omega}^T \underline{\hat{x}} > 0 \quad \forall \underline{\hat{x}}$.

b) Show the situation graphically, when $d = 1$ and $H_1 = \{-3, -2, -1\}, H_2 = \{5, 6, 7\}$.

6. In the Perceptron Learning Rule, the updating $\underline{\omega}^{(n+1)} = \underline{\omega}^{(n)} + \alpha_n \underline{\hat{x}}_i$ is done, when $\underline{\hat{x}}_i$ has been classified incorrectly. That is $e_i = \underline{\omega}^{(n)T} \underline{\hat{x}}_i < 0$. How should α_n be chosen so that $\underline{\hat{x}}_i$ would surely be classified correctly after updating?

7. Let H_1 and H_2 be linearly separable, and the samples $x_1 = -2 \in H_1$, $x_2 = -1 \in H_2$ and $x_3 = 2 \in H_2$ are available for learning. Leth the initial weight vector in the equation of question 6 $\underline{w}^0 = [1, 1]^T$ and the parameter $\alpha = 0.5$. Teach the percepton using these samples augmented as in question 5.