Instituto Superior Técnico

Machine Learning (Aprendizagem Automática)

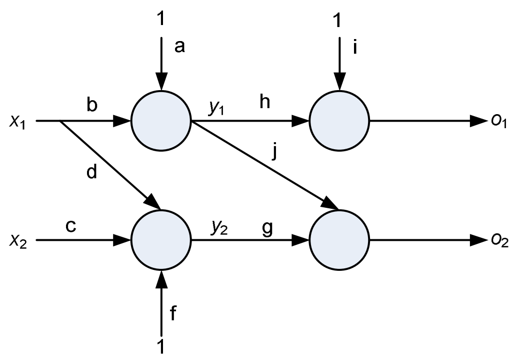
Exam of 3/2/2011. Duration: 3 hours

Notes:

- *Present all responses in a clear, ordered and detailed way, with a brief justification of each step.*
- *Present all calculations.*
- *Keep at least three significant digits in all calculations.*

Problem 1

Consider a multilayer perceptron with the structure indicated in the figure. In this figure, roman letters indicate weights and italic letters indicate network variables.



All weights have an initial value of 0.5. The output units are linear and the units in the first layer have, as activation function, the hyperbolic tangent.

The training set is:

x_1	x_2	d_1	d_2
–1	1	1	1
1	–1	2	–2

a) **(2 points)** Draw the backpropagation network that corresponds to the network shown in the figure. Indicate, by means of appropriate symbols, the gains of all branches. Also indicate, by means of appropriate symbols, all the variables (node values) that you will need to compute to perform the calculation requested in item b) below.

b) **(3 points)** Find, through the backpropagation method in **batch mode**, the value of weight *b* after one iteration with no momentum term and with a step size coefficient $\eta = 0.2$. The patterns are presented in the order given in the table, and the cost function is the total squared error.

Problem 2

In this problem, use (x_1, x_2) to denote coordinates in the input space and $(\tilde{x}_1, \tilde{x}_2)$ to denote coordinates in the feature space. You may use the same figure to represent the answers to several of the items below, as long as you clearly identify them. As in non-graphical responses, when you make graphical representations you should give a brief justification of what you do.

Consider a nonlinear, binary SVM classifier with output $y \in \{-1,1\}$, which uses the nonlinear mapping $\phi(x) = (x_1, x_1 x_2)$. The vector and the bias that define the classification boundary in the feature space are $\tilde{w} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{4} \end{bmatrix}$ and $b = 1$, respectively.

a) **(1 point)** Find the decision rule of the classifier, expressed in the input space.

b) **(1.5 points)** Find the equations of the margins of the classifier in the feature space and represent them graphically.

If you did not solve this item, use, in items c) and d) below, the following equations for the margins:

$$\tilde{x}_1 + \tilde{x}_2 = 0 \quad \text{and} \quad \tilde{x}_1 + \tilde{x}_2 = 4.$$

c) **(1.5 points)** Find the equations of the margins of the classifier in the input space and represent them graphically.

d) **(1 point)** Choose two vectors that could be support vectors of some training set, one for each class, and represent them graphically in both spaces (input and feature spaces).

Problem 3

In this problem you can use a graphical representation to perform distance comparisons without computing the actual distances, as long as the result of the comparison is clear from the graphical representation.

Consider the following training set: $\{(-3,6), (-1,0), (1,0), (3,6)\}$.

a) **(1.5 points)** Perform one iteration of the k-means algorithm with two centers, using as initial positions of the centers (–1,1) and (1,1).

   If you didn't solve this item, use the initial position, instead of the result of this item, in b) to d) below.

b) **(1 point)** Check whether the result that you obtained corresponds to a minimum (not necessarily a global minimum!) of the cost function of the algorithm.

c) **(1.5 points)** Compute the value of the cost function for the result that you obtained in a) above.

d) **(1 point)** Find a position of the centers that corresponds to a minimum of the cost function that is lower than the value of the cost function in the position found in a) above. Note that you don't have to *compute* that minimum. You just have to indicate what is the position of the centers, to show that it is a minimum and to show that the cost function is lower than in the position obtained in a).

Problem 4

Consider the training set $\boldsymbol{X} = \{(-5,1), (-3,-3), (3,5), (5,1)\}$.

a) **(1.5 points)** Find the first principal direction corresponding to this set. Indicate the direction by means of a vector.

   If you didn't solve this item, use the direction of the vector $(2,1)$ *as principal direction, in the following items.*

b) **(1 point)** Find the first principal component of the vector $(5,1)$.

c) **(1.5 points)** Find the reconstruction of the vector $(5,1)$ using only the first principal component.

d) **(1 point)** Assume that a new training set $\boldsymbol{Y}$ is formed through the equation $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x}$ are the vectors of $\boldsymbol{X}$ above, and

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

   Find the covariance matrix of the vectors of $\boldsymbol{Y}$, basing yourself on the computations done above, without actually computing the vectors of $\boldsymbol{Y}$.