

Asignatura: **CI5218 – Ingeniería del Conocimiento**  
Cuatrimestre: **2º** Examen: **Final** Convocatoria: **Ordinaria**  
Grupo: **5IT1** Curso: **2005/2006** Fecha: **6 del 6 del 06**

## PREGUNTAS

1. Considera el siguiente conjunto de entrenamiento con un atributo  $X$  real y una clase  $Y$  binaria. Se pretende usar el algoritmo K-NN con distancia euclídea sin usar ningún tipo de peso.

$X$	$Y$
-0,1	-
0,7	+
1,0	+
1,6	-
2,0	+
2,5	+
3,2	-
3,5	-
4,1	+
4,9	+

**Sin calcular las distancias para resolver los 2 primeros apartados** (en vez de eso representar los ejemplos etiquetados en una dimensión, es decir, en una recta) responder a las siguientes cuestiones:

- a) **(2 puntos)** Suponiendo que hacemos el test usando el mismo conjunto de entrenamiento, cuál sería el porcentaje de errores usando el algoritmo 1-NN en este conjunto de datos? ¿Y si usamos 3-NN? Explica el razonamiento que has usado para llegar a la solución.
- b) **(2 puntos)** Construir un perceptrón multicapa que tenga dos unidades de salida, usando el algoritmo de aprendizaje por retropropagación del error y la regla delta generalizada. Los demás parámetros son de libre elección. **NOTA:** Sólo entrenar con los dos primeros patrones.
- c) **(2 puntos)** Convertir el atributo  $X$  que tiene valores reales en discretos, dividiendo en 2 intervalos. **IMPORTANTE:** para hacer la división es necesario tener en cuenta el cambio de clase entre dos puntos contiguos.

d) **(2 puntos)** Suponiendo que no conocemos la clase para cada uno de los ejemplos, construir aplicando el algoritmo apropiado un sistema de clustering o agrupamiento jerárquico. Dibuja también el dendograma resultante.

2. **(2 puntos)** Sea la figura siguiente; en ella se representa un nodo de un árbol de decisión y 21 ejemplos de un conjunto de test. Además aparece la distribución de cómo han caído durante la fase de evaluación. Con todo ello decidir si se debe o no podar dicho nodo.

