

**Asignatura: CI5128 – Ingeniería del conocimiento**Cuatrimestre: **2º**Examen: **Final**Convocatoria: **Extraordinaria**Grupo: **5IT1**Curso: **2006/2007**Fecha: **9-7-2007****ALUMNO:** \_\_\_\_\_

**La duración del examen es de 2 horas. Se pueden utilizar todo tipo de apuntes y libros de forma personal e intransferible.**

1. Después de muchas discusiones y de experimentar distintos procedimientos de evaluación, los profesores de Ingeniería del Conocimiento han llegado a la conclusión de suprimir los exámenes finales. En su lugar, se usará el algoritmo ID3 para decidir si un alumno aprueba la asignatura. Para ello, considerarán los siguientes factores: participación en clase, trabajos realizados en casa, nota media de su expediente hasta el momento y si juega bien o no al fútbol. Los ejemplos que usan para construir el árbol de decisión se muestran en la siguiente tabla:

<i>Ejemplo</i>	<i>Participa</i>	<i>Trabajos</i>	<i>Nota</i>	<i>Fútbol</i>	<i>Aprueba</i>
1	si	si	baja	experto	no
2	no	si	alta	novato	si
3	no	no	baja	novato	no
4	si	no	baja	experto	no
5	si	si	alta	experto	si
6	si	no	baja	novato	no
7	no	no	alta	novato	no
8	no	si	alta	novato	si
9	si	no	baja	experto	no
10	si	si	baja	novato	no
11	no	no	alta	experto	no
12	si	si	baja	novato	no

<i>Ejemplo</i>	<i>Participa</i>	<i>Trabajos</i>	<i>Nota</i>	<i>Fútbol</i>	<i>Aprueba</i>
13	si	no	alta	experto	si
14	si	si	alta	novato	si
15	si	si	alta	experto	si

- a) **(2 puntos)** Construye el árbol de decisión basándose en el algoritmo ID3 y sin realizar cálculos de entropía. Justifícalo explicando el árbol resultante.
- b) **(2 puntos)** Aplica el algoritmo de eliminación de candidatos a los tres primeros ejemplos y al 5º ejemplo, considerando el primero y el tercero como ejemplos negativos (si no se hace así la pregunta se puntuará con un cero) ¿Se ha aprendido algún concepto? Si es así, ¿cuál?
- c) **(1 punto)** Codifica los ejemplos de forma que puedan ser usados para entrenar un perceptrón multicapa. El único requisito es que alguien puede ser experto y novato a la vez, mientras que los demás atributos son excluyentes. Detalla la descripción de todos los atributos que formarían el conjunto de entrenamiento.
- d) **(1 puntos)** Explica con esta tabla cómo sería un proceso de entrenamiento en el que se usase para testear la técnica de 5-fold cross validation. Para ello, explica primero el proceso y aplícalo a los ejemplos de esta tabla.
2. **(2 puntos)** Sea el siguiente conjunto de ejemplos. Se necesita agrupar los ejemplos en 2 clusters; di qué algoritmo utilizarías y aplícalo a este conjunto de datos. Se recomienda representar gráficamente los patrones y hacer uso de la distancia Manhattan.

$d1=(1,1)$

$d2=(2,2)$

$d3=(3,0)$

$d4=(7,3)$

$d5=(8,4)$

Una vez contruidos los clusters, ¿En qué grupo se incluiría el ejemplo  $d6=(5,2)$ ? ¿Cómo lo harías para saberlo (analíticamente, no de forma gráfica o visual)?

3. **(2 puntos)** Resuelve el ejercicio 2 mediante una red neuronal. Para ello:

- Elije el modelo de red neuronal que creas conveniente
- Elije el valor de todos los parámetros (incluidos los umbrales) que creas necesarios
- Entrena sólo con los dos primeros patrones.
- Sin hacer cálculos, cómo clasificaría al tercer ejemplo. Explícalo claramente.

Razona qué ha pasado en las sucesivas iteraciones y explica el resultado final obtenido.