



Asignatura: **CI5128 – Ingeniería del conocimiento**

Cuatrimestre: **2º**

Examen: **Final**

Convocatoria: **Extraordinaria**

Grupo: **SIT1**

Curso: **2010/2011**

Fecha: **11-7-2010**

Alumno: \_\_\_\_\_

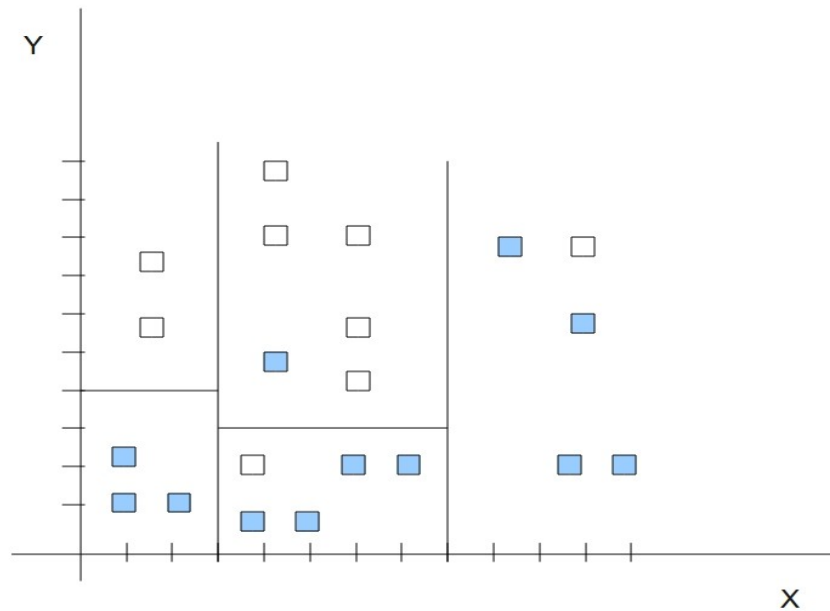
**NOTA: Se puede usar todo tipo de libros y apuntes de manera personal e intransferible. La duración del examen es de 2 horas y media.**

1. Una empresa vende D discos diferentes, L libros diferentes y V películas diferentes a sus clientes por Internet. Cada cliente, en cada venta, puede comprar varias cosas, pero, como máximo, una de cada tipo (disco, libro y/o película). Cada producto tiene un precio almacenado (entre 0 y M, siendo números enteros), y una categoría que lo describe. En el caso de los discos las categorías son ambient, hip-hop, metal, rock, pop o clásico. En el caso de los libros pueden ser de ficción o no, y, dentro de los de ficción, pueden ser romántico, policiaco, o histórico. En el caso de las películas, pueden ser acción, comedia o drama. Se pretende recomendar a cada cliente, en cada venta, un artículo para la próxima vez que se conecte. Se pide:
  - a) **(2 puntos)** Seleccionar una técnica de aprendizaje automático para resolver la tarea y diseñar cómo serían los atributos de entrada y salidas de un sistema que utilizara esa técnica.
  - b) **(1 punto)** Da la representación explícita de un patrón de entrada
  - c) Dada la técnica elegida en el apartado a y el ejemplo representado en el apartado b, se pide:
    - i) **(1 punto)** Realiza el preprocesado que sea necesario
    - ii) **(1,5 puntos)** Procesa una iteración del algoritmo elegido, desde que se presenta el patrón de entrada hasta que se obtiene la salida.

2. (1.5 puntos) Una región de decisión es una zona delimitada por un modelo y que clasificará a los ejemplos futuros con la clase mayoritaria de los ejemplos situados en esa región. En la figura se puede ver un gráfico con las 5 regiones que ha generado un árbol de decisión a partir del algoritmo C4.5. Se pide:

a) Dibuja el árbol que ha generado estas regiones de decisión, indicando el rango de valores que has considerado para los atributos X e Y ya discretizados.

b) ¿Cuál sería el porcentaje de errores que ha cometido si lo testeamos con el mismo conjunto de entrenamiento? ¿Cuál sería su matriz de confusión?



3. (3 puntos) Resuelve mediante el perceptrón multicapa de la figura un problema de clasificación, del que sabemos que para el patrón  $I_1=1$ , la salida debe ser (0,1,1). Elige para ello todos los parámetros de la red que creas necesarios, añadiéndolos al dibujo junto con los nombres dados a cada una de las unidades, y utiliza la regla delta generalizada para la actualización de los pesos sinápticos.

