

Asignatura: **CI5128 – Ingeniería del conocimiento**

Cuatrimestre: **2º**

Examen: **Final**

Convocatoria: **Ordinaria**

Grupo: **5IT1**

Curso: **2006/2007**

Fecha: **28-5-2007**

ALUMNO: _____

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

Supongamos este conjunto de ejemplos etiquetados, donde la tarea a realizar consiste en clasificar si una película merece la pena o no ser vista.

	<i>¿MajorStudio?</i>	<i>Coste de producción</i>	<i>Género</i>	<i>¿Merece la pena?</i>
P1	Sí	146	Ciencia Ficción	sí
P2	No	28	Terror	sí
P3	No	52	Romántica	sí
P4	No	74	Ciencia Ficción	no
P5	Sí	22	Romántica	no

- a) **(2 puntos)** ¿Qué atributo ofrece mayor información respecto a la pregunta ¿Esta película merece la pena ser vista? Utiliza un método para discretizar los atributos que tenga en cuenta el cambio de clase.
- b) **(2 puntos)** Haciendo uso del algoritmo Naive Bayes, debes predecir si una película 6 cuyos valores para los atributos son MajorStudio=no, Coste de producción=100 y Género= Romántica merece la pena ser vista y con qué probabilidad.
- c) **(1.5 puntos)** ¿Cómo clasificaría mediante 1-NN a esta película 6?
- d) **(2. puntos)** Construye un perceptrón multicapa para este conjunto de datos y prepara el conjunto de datos de forma que sea aplicable un perceptrón multicapa. Para ello:
- Primero se deben preprocesar los datos de forma que los valores continuos estén normalizados.
 - Dibuja la estructura completa de la red. Es necesario utilizar dos unidades de entrada y

tres en la capa oculta.

- Indica cómo la regla delta generalizada cambiaría la red después de procesar el primer ejemplo **pero sin calcularlo**. Es decir, detalla con letras y **sin valores concretos** cómo cambiaría el peso de una sinapsis cualquiera entre la capa oculta y la de salida, otro cualquiera de la capa oculta y la de entrada y un umbral que afecte a una unidad de la capa oculta.

e) **(2.5 puntos)** Supón que tenemos un sitio web dedicado a buscar parejas a las personas que están suscritas (como en Match.com) Un criterio para buscar personas afines podría ser los gustos en cine, por ejemplo. Debes diseñar un sistema que permita hacer esto, y para ello se debe especificar:

- ¿Qué algoritmo de aprendizaje automático usarías y por qué?
- ¿Qué datos te harían falta? ¿Cómo los estructurarías en una base de datos? Haz un diseño a grosso modo de la base de datos y los campos que pedirías por ejemplo en un formulario web para construirla.
- ¿Cómo los deberías tener para aplicarle el algoritmo elegido? Especifica la tabla y dos posibles ejemplos.
- ¿Cómo lo integrarías en tu sitio web? Es decir, ¿qué deberías usar para llevarlo a cabo? Especifica tu elección en cuanto a tipos de servidores (por su función, no por el sistema operativo ni hardware), lenguaje de programación elegido para implementarlo, integración del motor de aprendizaje con el sitio web, etc...
- ¿Cómo harías para ofrecer funcionalidades extras como por ejemplo recomendar una película según tus gustos y los de tus afines? ¿Qué datos te harían falta? ¿Qué método de aprendizaje automático utilizarías? Razona la respuesta.