

Examen Final de Aprendizaje

(8 de enero de 2007)

Duración: 3 horas

Aprendizaje Clásico (5 puntos)

1. (2 puntos) Responde brevemente a **tres** de las siguientes cinco preguntas:

- El espacio de versiones es un algoritmo general de aprendizaje ¿por qué es inabordable si no imponemos restricciones al lenguaje de representación del concepto que queremos aprender? ¿qué perdemos al restringir el lenguaje de representación? ¿qué deberíamos hacer para que el coste del aprendizaje sea menor?
- ¿De qué manera afecta el ruido en el conjunto de ejemplos de entrenamiento a los árboles de decisión? ¿qué método se utiliza para reducir este problema?
- ¿Por qué es un problema que los árboles de decisión utilicen particiones paralelas a los ejes para representar las fronteras entre los conceptos que aprenden? ¿Por qué al utilizar los métodos de combinación de clasificadores puede mejorar el resultado de los árboles de decisión?
- ¿Por qué en el algoritmo de K vecinos cercanos aumentar demasiado el número de vecinos puede dar resultados peores? En el caso de tener sólo dos clases y el mismo número de ejemplos para cada clase en el conjunto de entrenamiento, ¿crees que pasaría lo mismo si utilizáramos todos los ejemplos para hacer la predicción y ponderáramos según la distancia a cada ejemplo?
- ¿Cómo crees que afecta el ruido en el conjunto de ejemplos de entrenamiento al algoritmo de K vecinos cercanos? ¿Crees que aumentando el número de vecinos que se utilizan para clasificar se reduciría el efecto del ruido?

2. (2 puntos) Responde brevemente a **tres** de las siguientes cinco preguntas:

- El algoritmo de naive bayes aproxima los conceptos mediante una distribución de probabilidad conjunta ¿qué supone este algoritmo sobre esta distribución de probabilidad? Las redes bayesianas también pueden usarse como clasificadores, ¿que diferencia existe entre la distribución de probabilidad que aprendería una red bayesiana y la que aprende el naive bayes?
- ¿Cómo crees que le afecta el ruido en los ejemplos de entrenamiento al algoritmo de naive bayes? ¿y los atributos no relevantes?
- ¿Cual es el objetivo de los métodos de selección de atributos? ¿que diferencias hay entre los métodos de selección de atributos basados en wrappers y los métodos de filtro?
- ¿A que algoritmo crees que le afectarán más los atributos irrelevantes a los árboles de decisión o a K vecinos cercanos? ¿por qué?
- ¿Cual es el objetivo del aprendizaje inductivo no supervisado? ¿Qué ventajas tienen los algoritmos no supervisados incrementales como el de vecinos cercanos respecto a los no incrementales como K-means? ¿y que inconvenientes?

3. (1 punto) Responde brevemente a **dos** de las siguientes cuatro preguntas:

- ¿Por qué se dice que el aprendizaje deductivo no es realmente aprendizaje en sentido estricto? Y si realmente no es estrictamente aprendizaje ¿cual es el beneficio que se obtiene utilizando el aprendizaje deductivo?
- ¿Por qué todos los métodos de EBL estan asociados a un método de resolución de problemas?
- ¿Qué se pretende obtener en el método de EBL que se utiliza en STRIPS al convertir todas las constantes de la primera columna de la tabla triangular en variables?
- ¿Es el aprendizaje por refuerzo aprendizaje inductivo? ¿Por qué? ¿Describe los elementos que forman parte de un problema de aprendizaje por refuerzo y el objetivo de aprendizaje?