

Examen Parcial

24 de Abril de 2007

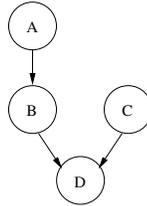
Aprendizaje de Maquina - 2007-I
Maestría en Ing. de Sistemas y Computación

1. Considere el siguiente problema de clasificación de dos clases, C_1 y C_2 . La probabilidad de que una observación $x \in [0, 1]$ pertenezca a cada una de las clases está dada por

$$P(x|C_1) = 2 - 2x$$
$$P(x|C_2) = \begin{cases} 8x - 4 & \text{si } x > 0,5 \\ 0 & \text{o.c.} \end{cases}$$

Además, las probabilidades *a priori* de las clases son: $P(C_1) = 4/5$ y $P(C_2) = 1/5$.

- a) ¿Qué valores de x deberían clasificarse en C_1 y qué valores en C_2 ?
- b) Suponga ahora que es posible rechazar una entrada, es decir no clasificarla en ninguna de las clases. Además, hay un costo asociado con clasificar mal un elemento dado por $\lambda_e = 2$ y un costo por rechazar dado por $\lambda_r = 1/2$. Calcule los intervalos en que la entrada es clasificada como C_1 o C_2 , o es rechazada.
2. Sean A , B , C y D variables binarias relacionadas de acuerdo con la siguiente red de creencia Bayesiana:



- a) Establezca las probabilidades condicionales y absolutas necesarias para especificar totalmente la red basándose en los siguientes ejemplos:

A	B	C	D
1	0	0	1
1	1	0	1
1	0	1	1
0	1	0	0
0	1	1	0
1	0	1	0
0	1	1	1
0	0	0	1
0	0	1	0

- b) Calcule la siguiente probabilidad condicional $P(A = 0|C = 1, D = 0)$

c) Dada la siguiente evidencia $A = 0$ y $C = 1$, ¿qué valor le daría a D ?

3. Dada la muestra $X_i = \{x_i^t, r_i^t\}$, definimos $g_i(x) = r_i^1$, es decir nuestro estimado para cualquier x es el valor r de la primera instancia en el conjunto de datos (desordenado) X_i . ¿Qué puede decir acerca de su sesgo y varianza, cuando se compara con $g_i(x) = 2$ y $g_i(x) = \sum_t r_i^t / N$? ¿Qué pasa si la muestra está ordenada, de manera que $g_i(x) = \min_t r_i^t$?
4. Considere 3 objetos x_1, x_2 y x_3 tal que $\langle x_1, x_1 \rangle = 1$, $\langle x_2, x_2 \rangle = 20$, $\langle x_3, x_3 \rangle = 13$, $\langle x_1, x_2 \rangle = 6$, $\langle x_1, x_3 \rangle = 5$ y $\langle x_2, x_3 \rangle = 14$.
 - a) ¿Cuál es el valor de $\|x_i - x_j\|^2$ para todos los valores $i \neq j$?
 - b) ¿Cuál es el valor de $\|x_i - x_j\|_\phi^2$ para todos los valores $i \neq j$? Donde ϕ es el espacio de características inducido por el kernel $k(x, y) = (\langle x, y \rangle + 1)^2$.