

# Métodos Matemáticos en Ciencias de la Computación

4 de Septiembre 2006. Parte Teórica. 5 puntos

1. no. Explica un operador de cruce y un operador de mutación válidos para generar nuevos individuos en el problema del viajante de comercio, cuando este se aborda desde un algoritmo genético con representación basada en la trayectoria.
2. La probabilidad a priori de recibir un e-mail calificado como *basura* es de 0.50. Vamos a considerar un modelo, muy simplificado, que sirva para clasificar un e-mail como *basura* o *no basura*. Para ello nos fijamos en dos variables relacionados con dicho e-mail. Una de dichas variables nos dirá si la dirección de correo electrónico del remitente pertenece o no a un conjunto de direcciones consideradas como *peligrosas*. La segunda variable se relaciona con la existencia en el e-mail enviado de ciertos *caracteres sospechosos*.

Si el e-mail es *basura*, la probabilidad de que se haya enviado desde una dirección *peligrosa* es de 0.80, mientras que si el e-mail es *no basura* dicha probabilidad baja a 0.10.

También sabemos que si el e-mail es *basura* la probabilidad de que contenga *caracteres sospechosos* es de 0.70. Además entre los e-mails que son *no basura* esta probabilidad baja a 0.40.

- Efectúa el diagnóstico para un e-mail enviado desde una dirección *peligrosa* y que *no* contenga *caracteres sospechosos*.
  - Explica las características del paradigma clasificatorio utilizado, analizando si falta alguna frase en el texto que haga que el problema se ajuste al paradigma
3. Tenemos dos urnas con diferente configuración. En la urna *A* hay 8 bolas blancas y 2 negras, mientras que en la urna *B* hay 4 bolas blancas y 6 negras. En función del resultado obtenido al lanzar un dado con caras numeradas del 1 al 6 seleccionamos la urna de la que se va a extraer la bola. Si el resultado obtenido con el dado es un uno o un dos, se selecciona la urna *A*. En caso contrario la urna seleccionada es la *B*.

Denotamos por  $X$  la variable que indica el valor obtenido al lanzar el dado, y por  $Y$  la variable relativa al color de la bola extraída.

Calcular:

- $H(X)$  y  $H(Y)$
  - $I(X, Y)$
4. Dados los 8 objetos de la siguiente tabla:

<i>Objeto</i>	$X_1$	$X_2$	$C$
$O_1$	2	0	1
$O_2$	4	2	1
$O_3$	1	1	1
$O_4$	2	4	2
$O_5$	1	2	1
$O_6$	2	3	2
$O_7$	3	4	2
$O_8$	3	6	2

Aplica los clasificadores:

- 1-NN básico y
- 3-NN con distancia media

para clasificar los casos:  $O_9 = (2, 1)$  y  $O_{10} = (4, 3)$

5. En relación con la evaluación de clasificadores explica los siguientes conceptos:

- Los métodos *K-fold cross-validation* y *bootstrap* para estimar la probabilidad de acierto de un clasificador
- La curva ROC