

Métodos Probabilísticos en Inteligencia Artificial

Septiembre 99. Examen Teórico. 5 puntos

1. A partir de la tabla de contingencia adjunta, estudiar la independencia entre las variables X e Y , con un error de tipo I, $\alpha = 0.05$.

	x_1	x_2	x_3
y_1	100	20	80
y_2	20	80	300

2. A partir de los 8 casos de la tabla adjunta, caracterizados en dos dimensiones, explica los algoritmos Forg y k -medias de McQueen de clasificación no supervisada.

<i>Caso</i>	X_1	X_2
1	1	1
2	2	4
3	3	2
4	3	5
5	4	4
6	4	7
7	6	4
8	6	6

3. Sean n items y una mochila capaz de soportar un peso máximo C . Denotamos por b_j el beneficio obtenido al introducir el item j en la mochila, mientras que w_j denota el peso asociado a dicho item j . Se trata de seleccionar un subconjunto de items que maximicen $Z = \sum_j b_j x_j$, sujeto a la restricción $\sum_j w_j x_j \leq C$, siendo

$$x_j = \begin{cases} 1 & \text{si el item } j \text{ seleccionado} \\ 0 & \text{si el item } j \text{ no seleccionado.} \end{cases}$$

4. Enunciar el algoritmo general de inducción de árboles de clasificación. Explicar los conceptos de entropía, ganancia en información, *prepruning* y *postpruning*.
5. La tabla adjunta contiene 6 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN. Los tres últimos casos -numerados del 7 al 9- constituyen el conjunto de testeo. Construye el clasificador k -NN a partir de la distancia euclídea y de dos valores distintos de k ($k = 1, 3$).

<i>Caso</i>	X	Y	<i>Clase</i>
1	1	1	0
2	2	4	0
3	3	2	0
4	3	5	0
5	4	4	0
6	4	7	1
7	2	7	?
8	3	4	?
9	4	6	?