

Métodos Probabilísticos en Inteligencia Artificial

Junio 98. Examen Teórico. 5 puntos

- Diseñar un Algoritmo Genético (función de coste, representación de individuos, operadores de cruce y mutación) para el *problema de agrupamiento* siguiente:

Dado un conjunto finito de N números enteros positivos, se trata de agruparlos en k grupos disjuntos minimizando la suma de las diferencias entre los grupos.

- Explicar detalladamente los siguientes métodos de *clustering* particional: Forgy, k -medias de McQueen, explicando también las diferencias entre los mismos.
- Apoyándote en un ejemplo de tu invención explica el funcionamiento del algoritmo ID3 de inducción de árboles de clasificación. Los conceptos a explicar incluyen: entropía, ganancia en información, sobreajuste, algoritmo general de inducción de árboles de clasificación (T.D.I.D.T.).
- La tabla adjunta contiene 8 casos bidimensionales que constituyen el conjunto de entrenamiento para un clasificador k -NN. Los cuatro últimos casos -numerados del 9 al 12- constituyen el conjunto de testeo.

Caso	X	Y	Clase
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	6	4	1
8	6	6	1
9	2	7	?
10	3	4	?
11	4	6	?
12	5	5	?

Construye el clasificador k -NN a partir de la distancia euclídea y de dos valores distintos de k ($k = 1, 2$).

- Dado el grafo acíclico dirigido de la figura adjunta, testear -explicando los distintos pasos-, a partir del criterio de D -separación, las siguientes independencias condicionales:
 - $I(\{B, C\}, G \mid E)$
 - $I(\{A, C, D\}, G \mid \{E, F\})$.