

TD n°10

Apprentissage Supervisé : introduction aux Arbres de Décision

Exercice 1 Construction d'un arbre de décision

Nous supposons qu'il n'y a que 2 classes + et - de sorte que l'entropie d'un ensemble puisse être donnée par un tableau à 2 entrées, X et Y, X étant le nombre d'éléments de la classe + et Y le nombre d'éléments de la classe -.

X	Y	1	2	3	4	5
5		3.90	6.04	7.63	8.92	10.
4		3.61	5.51	6.90	8.	
3		3.25	4.85	6.		
2		2.75	4			
1		2				

Nous considérons l'ensemble E d'exemples suivant :

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
A	a1	a1	a1	a2	a2	a2	a1	a2	a3	a3
B	b1	b2	b2	b1	b2	b2	b1	b1	b1	b2
C	c1	c2	c3	c1	c1	c1	c1	c2	c3	c2
D	d2	d2	d1	d1	d1	d2	d1	d2	d1	d2
Classe	+	+	-	-	-	+	+	-	+	+

1. Construisez l'arbre de décision en utilisant les attributs ABCD dans cet ordre.
2. Construisez l'arbre de décision en utilisant les attributs DCBA dans cet ordre.
3. Construisez l'arbre de décision qui discrimine tous les exemples de E pour les classes + et - en faisant appel à la méthode de construction employée pour l'algorithme ID3 (C4.5). Justifiez le choix des attributs aux différents noeuds de l'arbre.
4. Construisez l'arbre de décision qui discrimine tous les exemples de E pour les classes + et - en utilisant la mesure de Gini.

Remarque : La mesure de Gini telle qu'elle est utilisée dans l'algorithme CART cherche lors de la segmentation le regroupement binaire qui maximise le gain. Ce n'est pas le cas ici.

Exercice 2 Construction d'un arbre de décision (2)

Attribut	Exemple	E1	E2	E3	E4	E5
	F1	a	c	a	b	c
	F2	c	a	a	c	c
	F3	a	c	c	a	b
	Classe	+	+	-	-	-

- Construisez l'arbre de décision qui discrimine tous les exemples les classes + et - en faisant appel à la méthode de construction employée pour l'algorithme ID3 (C4.5).