

Examen « apprentissage automatique »  
1er juin 2012 - Durée 2h  
Documents autorisés

**Exercice 1 (arbres de décision)**

Un objet possède 3 attributs:  $a_0$ ,  $a_1$  et  $a_2$ . Les attributs prennent leur valeur dans  $\{0, 1, 2\}$ . Un objet appartient à une classe  $c$  valant + ou -.

Soient l'ensemble d'apprentissage et l'ensemble de test suivants:

learn	$a_0$	$a_1$	$a_2$	$c$
l1	1	1	0	+
l2	2	1	1	-
l3	0	1	2	+
l4	2	1	2	-
l5	0	0	1	-
l6	2	2	0	+
l7	2	2	2	+
l8	1	1	2	-
l9	0	0	2	+

test	$a_0$	$a_1$	$a_2$	$c$
t1	1	1	1	-
t2	0	0	0	+
t3	2	2	1	+
t4	0	2	2	-
t5	0	2	0	-
t6	2	0	2	+
t7	2	0	0	+
t8	2	1	0	-
t9	2	0	1	-

En utilisant l'ensemble d'apprentissage,

1° Construire l'arbre de décision  $T_1$  en prenant les attributs dans l'ordre:  $a_0$ , puis  $a_1$ , puis  $a_2$ . Les noeuds terminaux de l'arbre sont libellés +, - ou ? selon les cas.

2° Construire le premier niveau de l'arbre de décision  $T_2$  en prenant la minimisation de l'entropie pour critère de choix d'attribut. On pourra faire les approximations suivantes:  $\log(2) = 0.7$ ,  $\log(3) = 1.1$ ,  $\log(5) = 1.6$  et  $\log(7) = 2$ .

3° Terminer la construction de  $T_2$  avec ce critère.

En utilisant l'ensemble de test,

4° Donner le nombre d'erreurs de  $T_1$  et  $T_2$ . Une erreur + pour - ou - pour + sera comptée 1 et un ? sera compté 0.5.

5° Conclure.

