

Examen

Questions de cours (3 pts)

1. Expliquer le principe de la méthode Bootstrap.
2. Donner la différence et la relation entre un motif fréquent, un motif fréquent fermé et un motif fréquent maximal.
3. Expliquer comment peut-on obtenir les règles de décision après la construction d'un arbre de décision.

Exercice 1 Motifs fréquents (8 pts : 4 + 2 + 2)

Soit les transactions suivantes :

TID	Items
T1	Pain, Lait, Beurre, Couches, Fromage
T2	Pain, Beurre, Eau, Confiture, Fromage
T3	Fromage, Couches, Pain, Beurre, Confiture
T4	Beurre, Lait, Jus
T5	Couches, Fromage, Jus, Eau

1. Avec un support minimum de 60%, trouver l'ensemble des motifs fréquents.
2. Trouver les motifs fréquents fermés et les motifs fréquents maximaux.
3. Avec un minimum de confiance de 70%, trouver les règles d'association de la forme $A \Rightarrow B, C$.

Exercice 2 Classification (9 pts : 4 + 2 + 3)

Nous considérerons l'ensemble E d'exemples suivant ayant les attributs A,B,C et D :

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
A	a_1	a_1	a_1	a_2	a_2	a_2	a_1	a_2	a_3	a_3
B	b_1	b_2	b_2	b_1	b_2	b_2	b_1	b_1	b_1	b_2
C	c_1	c_2	c_3	c_1	c_1	c_1	c_1	c_2	c_3	c_2
D	d_2	d_2	d_1	d_1	d_1	d_2	d_1	d_2	d_1	d_2
Classe	+	+	-	-	-	+	+	-	+	+

1. Construire l'arbre de décision correspondant à l'ensemble E en utilisant l'algorithme ID3.

2. Donner la précision de l'arbre obtenu sur la table d'entraînement. Calculer la moyenne harmonique sur la même table.
3. Trouver la classe de l'exemple ayant les attributs (a_2, b_1, c_3, d_1) en utilisant la classification bayésienne naïve.

NB : On donne le tableau suivant représentant les valeurs de la fonctions

$$H(x, y) = -\frac{x}{x+y} \log_2\left(\frac{x}{x+y}\right) - \frac{y}{x+y} \log_2\left(\frac{y}{x+y}\right)$$

$x \backslash y$	1	2	3	4	5	6
6	0,592	0,811	0,918	0,971	0,994	1
5	0,650	0,863	0,954	0,991	1	
4	0,722	0,918	0,985	1		
3	0,811	0,971	1			
2	0,918	1				
1	1					

★★★ Bonne chance ★★★

Dr A.Djeffal

Corrigé type

Questions de cours (3 pts)

- Voir le cours

Motifs fréquents (8 pts : 4 + 2 + 2)

- Avec un support minimum de 60%, trouver l'ensemble des motifs fréquents.

- Taille 1 :

$$F_1 = \{\text{Beurre, Couches, Fromage, Pain}\}$$

$$\text{Support} = \{4/5, 3/5, 4/5, 3/5\}$$

(1pt)

- Taille 2 :

$$F_2 = \{\text{Beurre Fromage, Beurre Pain, Couches Fromage, Fromage Pain}\}$$

$$\text{Support} = \{3/5, 3/5, 3/5, 3/5\}$$

(1pt)

- Taille 3 :

$$F_3 = \{\text{Beurre Fromage Pain}\}$$

$$\text{Support} = \{3/5\}$$

(1pt)

$$\text{Ensemble des motifs fréquents} = \{F_1 \cup F_2 \cup F_3\}$$

(1pt)

- Trouver les motifs fréquents fermés et les motifs fréquents maximaux.

- Motifs fréquents fermés : {Beurre, Fromage, Couches Fromage, Beurre Fromage Pain} (1pt)

- Motifs fréquents maximaux : {Couches Fromage, Beurre Fromage Pain} (1pt)

- Avec un minimum de confiance de 70%, trouver les règles d'association de la forme $A \Rightarrow B, C$.

- Pain \Rightarrow Beurre Fromage (100 %)

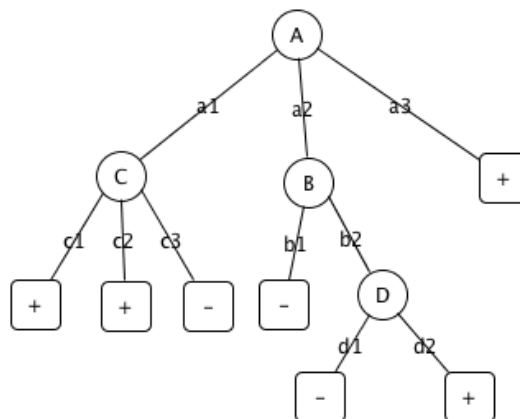
- Beurre \Rightarrow Pain Fromage (75 %)

- Fromage \Rightarrow Pain Beurre (75 %)

(2pts)

Classification (9 pts : 4 + 2 + 3)

- Construire l'arbre de décision correspondant à l'ensemble E en utilisant l'algorithme ID3. (4pts)



- La précision de l'arbre obtenu sur la table d'entraînement = 100 % (1pt)

Moyenne harmonique

$$sv = \frac{6}{6} = 1$$

$$sp = \frac{4}{4} = 2$$

$$\text{Moyenne harmonique} = \frac{2 \times 1 \times 1}{1+1} = 1 \quad (1\text{pt})$$

- (c) Trouver la classe de l'exemple ayant les attributs (a_2, b_1, c_3, d_1) en utilisant la classification bayésienne naïve.

$$P(+)=\frac{6}{10} \times \frac{1}{6} \times \frac{3}{6} \times \frac{1}{6} \times \frac{2}{6} = 0.0027 \quad (1\text{pt})$$

$$P(-)=\frac{4}{10} \times \frac{3}{4} \times \frac{2}{4} \times \frac{1}{4} \times \frac{3}{4} = 0.02 \quad (1\text{pt})$$

Donc l'exemple appartient à la classe négative. (1pt)