

## Exercice 1 : Les règles associatives avec apriori [5pts]

1. Donner le **principe** algorithmique de la méthode apriori. [1pt]

- **définir un seuil  $s$  de l'indice support**
- **déterminer les sous-ensembles les plus fréquents selon la valeur de  $s$**
- **construire l'ensemble des règles associatives correspondant aux ensembles conservés**
- **trier les règles par confiance ou par lift**

2. Rappeler le **type de données** fouillées pour l'algorithme apriori, justifier votre réponse. [0.5pt]

**Données transactionnelles transformées en données binaires car pour cette méthode on travaille sur des transactions : tickets de caisse, tableau de symptômes, etc.**

3. Donner les formules de chaque indice et expliquer son rôle dans le choix final des règles associatives :

Indice support : **probabilité absolue  $P(X \cup Y)$**  [0.5pt]

Indice confiance : **probabilité conditionnelle  $P(Y/X)$**  [0.5pt]

Indice Lift : **mesure d'intérêt confiance (règle) /  $p$  (résultat)** [0.5pt]

4. On considère le ticket de caisse suivant, et la règle associative :  **$B \rightarrow E$**

T26	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
T1245	<b>B</b>	<b>C</b>	<b>E</b>	<b>F</b>	
T156	<b>B</b>	<b>E</b>			
T2356	<b>A</b>	<b>B</b>	<b>D</b>		
T145	<b>C</b>	<b>D</b>			

Donner le calcul pour :  
(Formule complète et calcul)

Support :  **$3/5$**  [0.5pt]

Confiance :  **$(3/5)/(4/5) = 3/4$**  [0.5pt]

LIFT :  **$(3/4)/(3/5) = 5/4$**  [0.5pt]

5. Comment peut-on juger la **pertinence** de cette règle ? Expliquer votre décision. [0.5pt]

**C'est une règle pertinente car son lift est supérieur strictement à 1**