



QCM Règles d'Association

Majeure Science de Données - UP2 : Apprentissage Automatique
Mines Saint-Étienne – le 11 décembre 2018

Nom et prénom :

.....

Parfois les questions peuvent comporter plusieurs bonnes réponses. Vous devez remplir la case de la bonne réponse avec un stylo de couleur foncée comme ceci : ■.

Question 1 Notons F l'ensemble d'itemsets fréquents, MF l'ensemble d'itemsets fréquents maximaux et FF l'ensemble d'itemsets fréquents et fermés. Il est évident que $MF \subset F$ et $FF \subset F$. Entre FF et MF il y a une relation d'inclusion :

- ☐ $MF \subset FF$
- ☐ $FF = MF$
- ☐ $FF \subset MF$
- ☐ il n'y a pas d'inclusion entre ces deux ensembles

Question 2 Que signifie que la règle d'association $\emptyset \rightarrow i$, avec i item et \emptyset ensemble vide, a été générée par l'algorithme Apriori pour un support min_sup et une confiance min_conf :

- ☐ i est un item dont le support dépasse min_conf
- ☐ le support de i est de 1
- ☐ l'algorithme Apriori peut se tromper
- ☐ le cas ne peut pas arriver

Question 3 Pendant l'exécution de l'algorithme Apriori pour les $k - 1$ premiers étapes aucun élagage n'a été nécessaire (ce qui se traduit par $L_j = C_j$, $1 \leq j \leq k - 1$). Le nombre de candidats C_k de l'étape k est :

- ☐ on ne peut pas le déterminer
- ☐ dépend de k et de nombre total d'items
- ☐ aucune réponse n'est correcte
- ☐ 2^k
- ☐ dépend de k et du nombre de transactions

Question 4 Le calcul des itemsets fréquents et des règles d'association est réalisé pour:

- ☐ faire directement des prédictions
- ☐ comprendre les données contenues dans les transactions
- ☐ en utilisant certaines règles d'association, il serait possible de faire de la prédiction
- ☐ Aucune de ces réponses n'est correcte.



Question 5 L'ensemble des items \mathcal{I} d'un ensemble de transactions est partagé en \mathcal{C} (cible) et \mathcal{D} . On note $|\mathcal{I}| = N$ et $|\mathcal{C}| = K$, évidemment $K < N$ et $|\mathcal{D}| = N - K$. Le nombre de règles d'association ayant la forme $L \rightarrow R$ avec $L \subset \mathcal{D}$ et $R \subset \mathcal{C}$, $L, R \neq \emptyset$ est :

- ☐ aucune de ces valeurs, car il dépend aussi du nombre de transactions
- ☐ $K(N - K)$
- ☐ $3^N - 2^N - 2^K + 2$
- ☐ $(2^{N-K} - 1) \times (2^K - 1)$