

Série 3 des Travaux pratiques en Fouille de données Règles d'association

Travail préparatoire

1. Créer un document Word et le sauvegarder sous le nom *TP3_nom1_nom2.doc(x)* où *nom1* et *nom2* désignent vos noms respectifs.
2. Les réponses aux questions doivent être rédigées dans le document Word et regroupées par exercice et numérotées conformément à la série.

Exercice 1 : Sous Tanagra, créer un nouveau diagramme *TP3_nom1_nom2.tdm* et l'associer au fichier *heart.txt* téléchargeable depuis la plateforme de elearning.

- Glisser le composant *view dataset* sur l'icône *dataset (heart.txt)* et visualiser le contenu du fichier.
- Définir un nouveau status *status 1* et sélectionner les attributs « sex », « cp », « fbs » et « resecg ».

Question 1 : Expliquer pourquoi ces attributs sont adéquats pour la recherche de règles d'association.

Question 2 : Quels sont les différents items ?

- Glisser le composant *Frequent Itemsets* à partir du groupe *Association*.
- Exécuter la recherche des itemsets fréquents pour un support minimal nul et les autres paramètres par défaut.

Question 3 : Quel est le nombre d'itemsets fréquents ?

Question 4 : Quel est le meilleur itemset fréquent de cardinalité maximale et le moins bon de cardinalité Minimale ? Justifier la réponse.

- Glisser à nouveau le composant *Frequent Itemsets* à partir du groupe *Association* sur le statut précédent.
- Exécuter la recherche des itemsets fréquents pour un support minimal de 60%

Question 5 : Quel est le nombre d'itemsets fréquents. Justifier la différence avec le nombre précédent

- Refaire le travail précédent avec un support minimal de 40%

Question 6 : Quel est le nombre d'itemsets fréquents ?

Question 7 : Comment peut-on utiliser le meilleur itemset fréquent cardinalité maximale de la question précédente dans un cas pratique ?

Question 8 : quelle est l'influence des paramètres *max_support*, *max_length*, *min_length* sur le nombre d'itemsets fréquents générés ? Justifier la réponse.

Exercice 2 :

- Définir un nouveau statut *status* et sélectionner les attributs «cp», «sex» et « chol ».

- Glisser le composant *Apriori* à partir du groupe *Association* sur *Define status*.
- Exécuter la recherche des règles d'associations avec les paramètres par défaut.

Question 9 : Quel est le problème rencontré ? Justifier la réponse.

Question 10 : Comment peut-on rechercher des règles d'association dans ce cas ? (donner la solution sans l'implémenter)

- Définir un nouveau statut *status* et sélectionner les attributs «slope », «thal» et «num».
- Glisser à nouveau le composant *Apriori* à partir du groupe *Association* sur le statut précédent et garder les paramètres par défaut.
- Exécuter la recherche des règles d'associations.

Question 11 : Quel est le nombre de règles obtenues ?

Question 12 : Quelle(s) est/sont la/les meilleure(s) règle(s) en justifiant la réponse? Comment peut-on les utiliser dans un cas pratique ?

- On veut obtenir le minimum de règles d'association en se basant sur le filtrage des règles par support minimal seulement.

Question 13 : Quelle est la valeur de ce support minimal arrondie à deux chiffres décimaux qui donne le minimum des règles? Justifier la réponse dans Tanagra.

Question 14 : Quel est le contenu de l'ensemble minimal des règles ?

Modalité d'envoi du travail : Le fichier Word et le fichier tanagra doivent être envoyés à l'adresse email suivante : doukifli.boukraa@univ-jijel.dz. La date limite d'envoi sera communiquée ultérieurement.