

### **Projet : Partie 3**

## **Extraction de motifs fréquents, règles d'associations et corrélations**

L'exploration des motifs fréquents conduit à la découverte d'associations et de corrélations entre les éléments d'un dataset transactionnel ou relationnel. Avec des quantités massives de données collectées et stockées en permanence, de nombreuses industries s'intéressent à l'extraction de ces modèles dans leurs bases de données. La découverte de relations de corrélation intéressantes entre d'énormes quantités d'enregistrements de transactions commerciales peut aider dans de nombreux processus décisionnels commerciaux tels que la conception de catalogues, le marketing croisé et l'analyse du comportement d'achat ou de consommation des clients.

Ainsi, dans cette troisième partie du projet, il vous est demandé de procéder à l'analyse et l'extraction des motifs fréquents, des règles d'associations et de corrélation du **dataset 2**. Cette étude a pour but d'analyser la consommation d'un ensemble d'individus en termes de contenu youtube. Les règles d'associations et de corrélation à extraire permettront d'établir un système de recommandation de vidéos youtube aux différents utilisateurs. Il vous est alors demandé de :

- A. Jeu de données :
  - a. Analyser le nouveau jeu de données: **dataset 2**.
  - b. Appliquer un prétraitement adapté aux données manipulées.
- B. Application de l'algorithme ***Apriori***:
  - a. Extraction de règles d'association (support, confiance).
  - b. Extraction de règles de corrélation (support, confiance, lift).
  - c. Expérimenter quelques valeurs de Min\_Supp et Min\_Conf.
- C. Options avancées de l'IHM:
  - a. Exécution de la méthode de DataMining à appliquer (Apriori).
  - b. Choix du Min\_Supp et Min\_Conf.
  - c. Insertion des données sur un utilisateur et en déduire une recommandation.

#### **Notes:**

- Le rapport doit être bien structuré et comporter **au moins**: une section analyse, une section pré-traitement, une section Extraction des motifs fréquents décrivant en détail le fonctionnement de l'algorithme Apriori, une section Extraction des règles d'associations et corrélations expliquant leur intérêt, equation, etc ainsi que les résultats de l'expérimentation des différentes valeurs de Min\_Supp et Min\_Conf.
- Chaque binôme est tenu d'envoyer la version électronique du rapport au plus tard: **Samedi 19 Novembre 2022 à 23h59**.
- Chaque binôme devra présenter son interface et code source le **Dimanche 20 Novembre 2022** durant la séance de TP.

Bon courage !