**TP – Analyse de règles d’association dans un magasin agroalimentaire**

**Contexte**

Un magasin agroalimentaire souhaite optimiser l’attribution de promotions sur ses produits alimentaires. L’objectif est de mieux comprendre les habitudes d’achat des clients, en identifiant les produits qui sont fréquemment achetés ensemble. Grâce à ces informations, le magasin pourra proposer des promotions ciblées et augmenter les ventes croisées (cross-selling).

**Objectifs pédagogiques**

* Découvrir et manipuler les techniques de fouille de règles d’association.
* Comprendre les notions de **support**, **confiance** et **lift**.
* Comparer deux algorithmes populaires : **Apriori** et **FP-Growth**.
* Interpréter les règles obtenues et formuler des recommandations marketing.

**Travail demandé**

**1. Formulation du problème**

* Représenter les tickets de caisse du magasin sous forme de transactions (chaque transaction = un panier d’achat).
* Exprimer le besoin sous forme d’un problème de fouille de règles d’association : *« Trouver des ensembles de produits fréquemment achetés ensemble, et les règles qui permettent de prédire l’achat d’un produit à partir de la présence d’autres produits. »*
* Exemple : {Pain, Fromage} → {Vin}.

**2. Application de l’algorithme Apriori**

* Définir les seuils de **support minimal** (ex. 20%) et de **confiance minimale** (ex. 60%).
* Exécuter l’algorithme Apriori pour générer :
  + Les **ensembles fréquents** de produits.
  + Les **règles d’association** correspondantes.
* Interpréter quelques règles obtenues.
  + Exemple : {Lait, Céréales} → {Jus} avec support = 25%, confiance = 70%.

**3. Comparaison avec l’algorithme FP-Growth**

* Exécuter FP-Growth avec les mêmes paramètres.
* Comparer les résultats avec ceux d’Apriori :
  + Nombre de règles obtenues.
  + Temps d’exécution.
  + Similarités et différences dans les ensembles fréquents générés.
* Discuter des avantages de FP-Growth (plus efficace, pas besoin de générer tous les candidats comme Apriori).

**4. Interprétation et recommandations**

* Identifier les règles les plus intéressantes à partir du **lift** (valeur > 1 = corrélation positive utile).
* Formuler des stratégies marketing :
  + Exemple : si {Pâtes, Sauce tomate} → {Fromage râpé}, alors proposer une promotion groupée « pack spaghetti ».
  + Si {Bière, Chips} apparaissent souvent ensemble, proposer une réduction combinée.
* Relever aussi les règles surprenantes ou inattendues, qui peuvent ouvrir de nouvelles opportunités.

**Attendus du TP**

* Un compte-rendu écrit avec :
  1. La formulation du problème en termes de règles d’association.
  2. Les ensembles fréquents et règles générés par Apriori et FP-Growth.
  3. Une comparaison critique des deux approches.
  4. Une interprétation des règles retenues avec propositions de promotions concrètes.