

Théorie de l'information

L'Information Mutuelle au service des Recommandations Personnalisées

Présenter par :



**Mouhamadou
GUEYE**



Aujourd'hui , dans un monde saturé d'informations, recommander le bon contenu à la bonne personne devient de plus en plus difficile . Les plateformes de Streaming, les sites de e-commerce et les médias sociaux exploitent des algorithmes sophistiqués pour améliorer la pertinence de leurs suggestions(Recommandations).

Parmi ces approches, **l'information mutuelle**, issue de la théorie de l'information, se distingue. Elle mesure **la relation entre les préférences des utilisateurs et les caractéristiques des objets recommandés**. Contrairement aux techniques classiques comme le filtrage collaboratif, elle capte des dépendances complexes et améliore la personnalisation, même avec des données bruitées.



Comment ?

Dans cette présentation, nous allons démontrer **comment utiliser l'information mutuelle pour recommander des films**, en simulant un système basé sur cette approche. À travers des graphiques interactifs, nous verrons comment les relations entre utilisateurs et contenus influencent les recommandations.





**Prêt à découvrir une méthode puissante et
intuitive pour révolutionner les
recommandations ? 🚀**

Formalisation mathématique

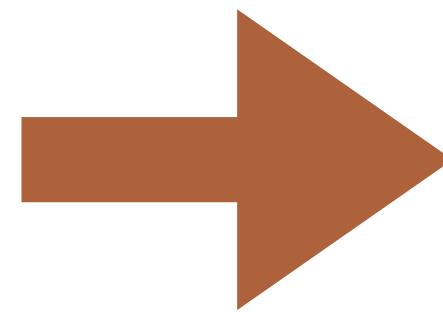
A- Information mutuelle entre utilisateur et objet

L'information mutuelle (**I**) mesure la dépendance entre deux variables aléatoires

Dans le contexte des systèmes de recommandation :

- **U** représente les **préférences** ou le profil de l'utilisateur,
- **O** représente les **caractéristiques** des objets (par ex., films, produits, articles).

L'information mutuelle entre
un utilisateur **u** et un objet **o**



$$I(U; O) = \sum_{u \in U} \sum_{o \in O} P(u, o) \log \frac{P(u, o)}{P(u)P(o)},$$

- **P (u, o)** est la probabilité jointe que l'utilisateur u interagisse avec l'objet o,
- **P (u)** et **P (o)** sont les probabilités marginales.

Une forte information mutuelle indique que l'utilisateur et l'objet partagent une relation significative.

Formalisation mathématique

B- Profil utilisateur et caractéristiques des objets

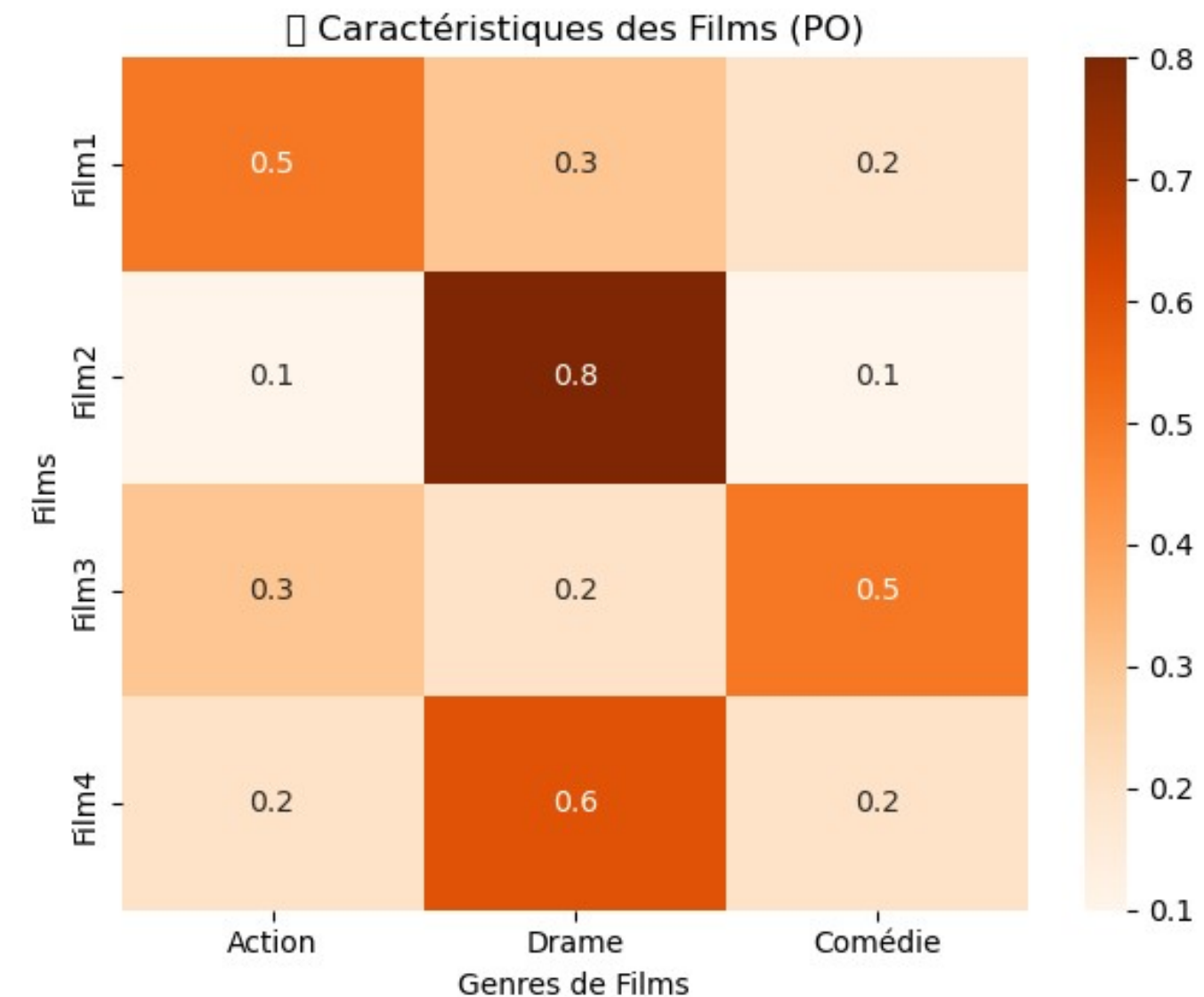
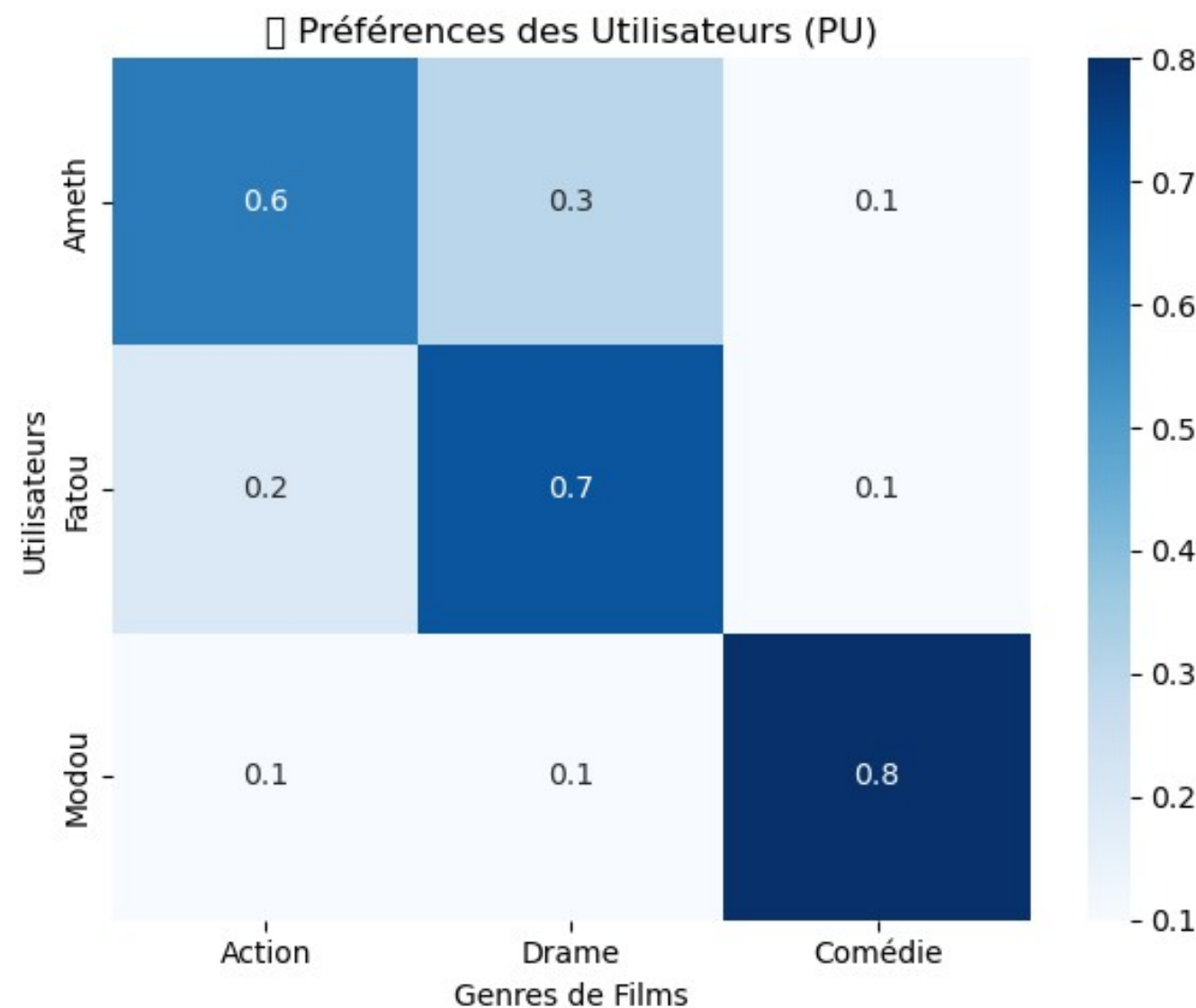
Le profil d'un utilisateur u peut être représenté comme une distribution de probabilité sur un ensemble de caractéristiques \mathbf{T} (par Ex: genres de films, catégories de produits). De même, chaque objet o peut être décrit par une distribution similaire.

$$P_U(t) = \frac{\text{fréquence de } t \text{ dans les interactions de } u}{\text{total des fréquences}},$$

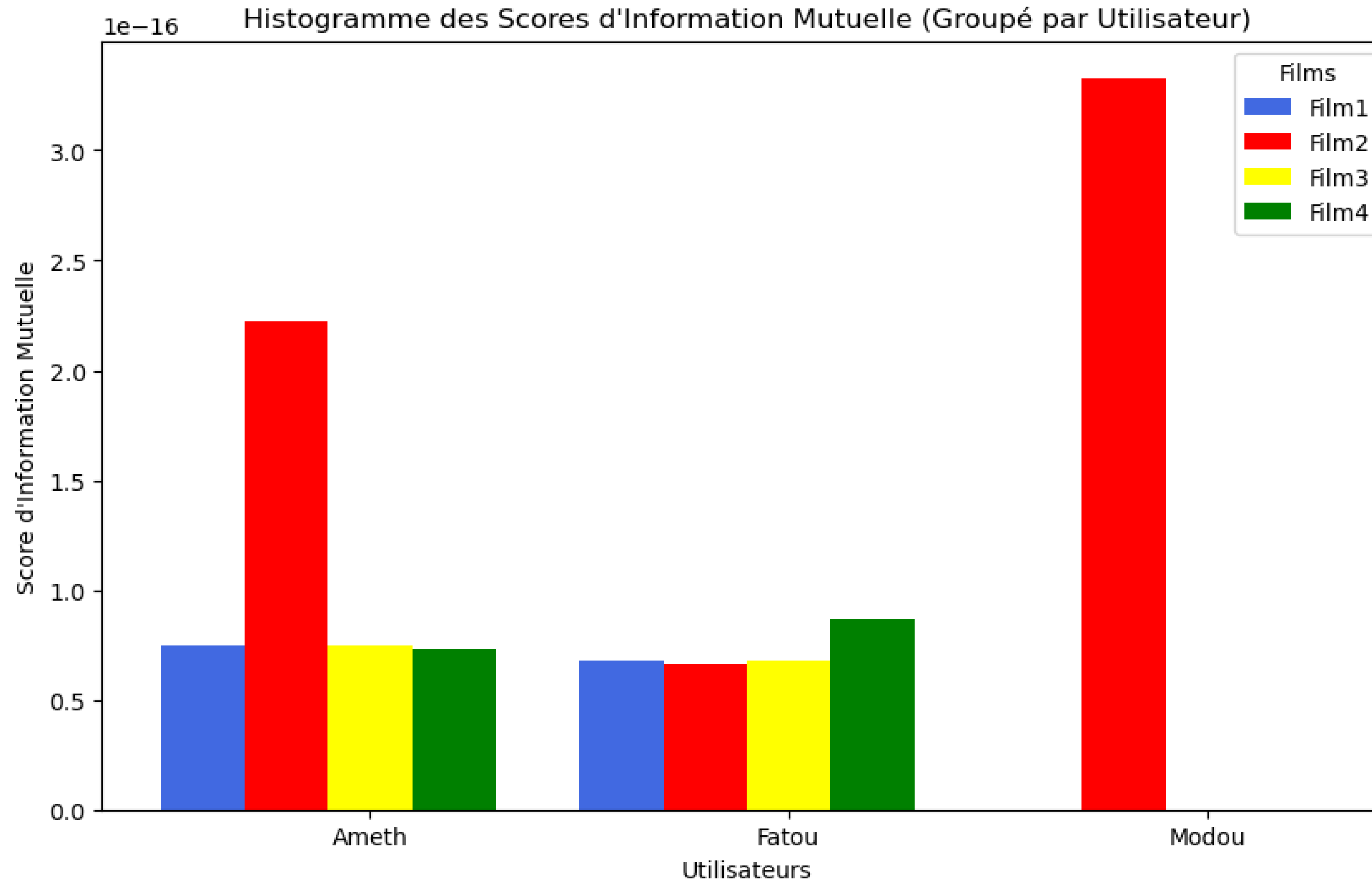
$$P_O(t) = \frac{\text{fréquence de } t \text{ dans les caractéristiques de } o}{\text{total des fréquences}}.$$

Simulation de recommandations basées sur l'Information Mutuelle

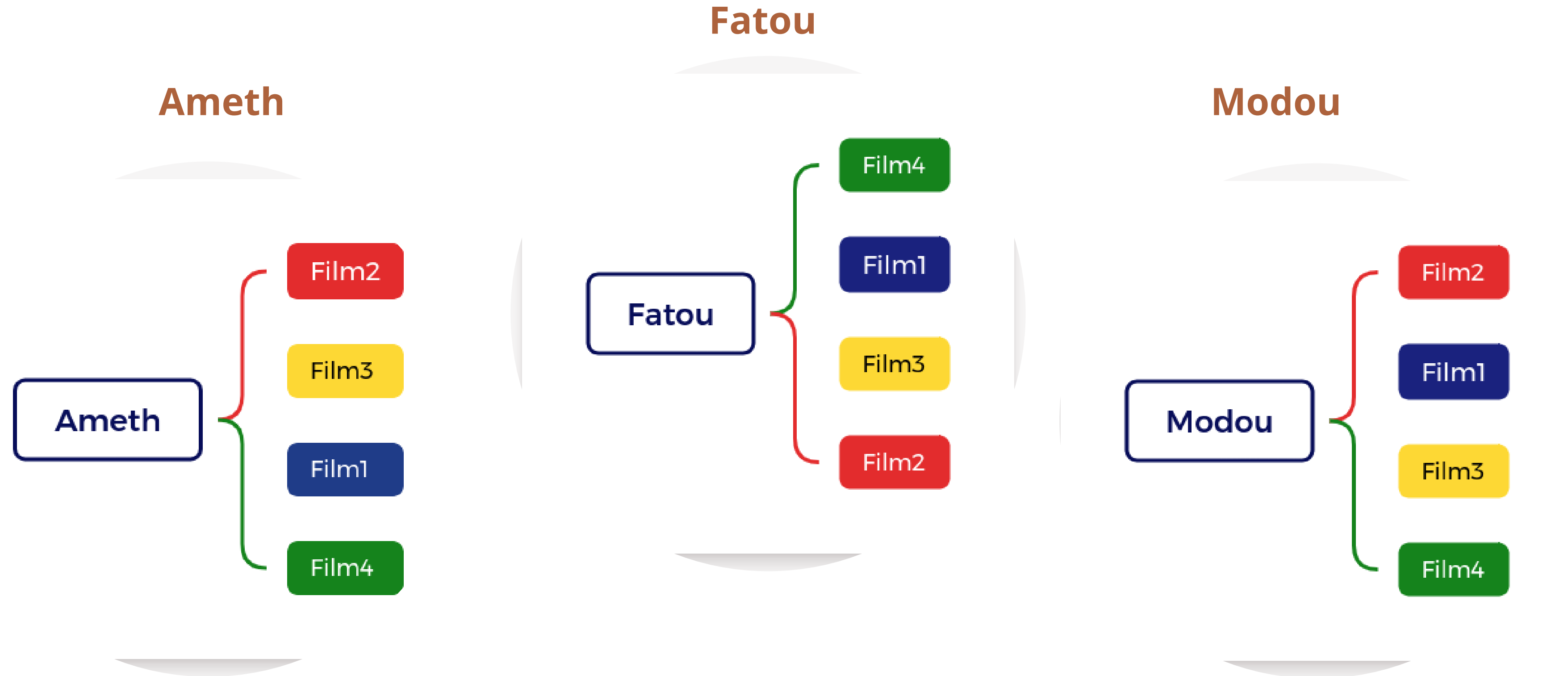
Considérons un système de recommandation de films avec **3 utilisateurs** (Ameth , Fatou , Modou) et **4 films**(film1,film2,film3,film4). Les distributions de probabilité des genres(Action,Drame,Comédie) pour chaque utilisateur et chaque film sont données par :



Visualisation en histogramme 🔥 pour montrer la force des relations



Résultats des recommandations



Code Python

```
# Données simulées : préférences utilisateur et caractéristiques des objets
users = ["Ameth", "Fatou", "Modou"]
items = ["Film1", "Film2", "Film3", "Film4"]

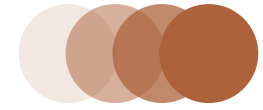
# Fonction pour calculer l'information mutuelle
def info_mutuelle(pu, po):
    pu = pu[:, np.newaxis] # Transformer en matrice colonne (2D)
    po = po[np.newaxis, :] # Transformer en matrice ligne (2D)

    matrice_pro = pu * po # Produit matriciel
    marginal_u = np.sum(matrice_pro, axis=1, keepdims=True)
    marginal_o = np.sum(matrice_pro, axis=0, keepdims=True)

    with np.errstate(divide='ignore', invalid='ignore'):
        mi_matrix = matrice_pro * np.log(matrice_pro / (marginal_u * marginal_o))
        mi_matrix = np.nan_to_num(mi_matrix) # Gérer les NaN

    return np.sum(mi_matrix)

# Calcul de l'information mutuelle avec valeur absolue
mi_scores = np.abs(np.array([[info_mutuelle(PU[u], PO[o])
                               for o in range(len(items))] for u in range(len(users))]))
```

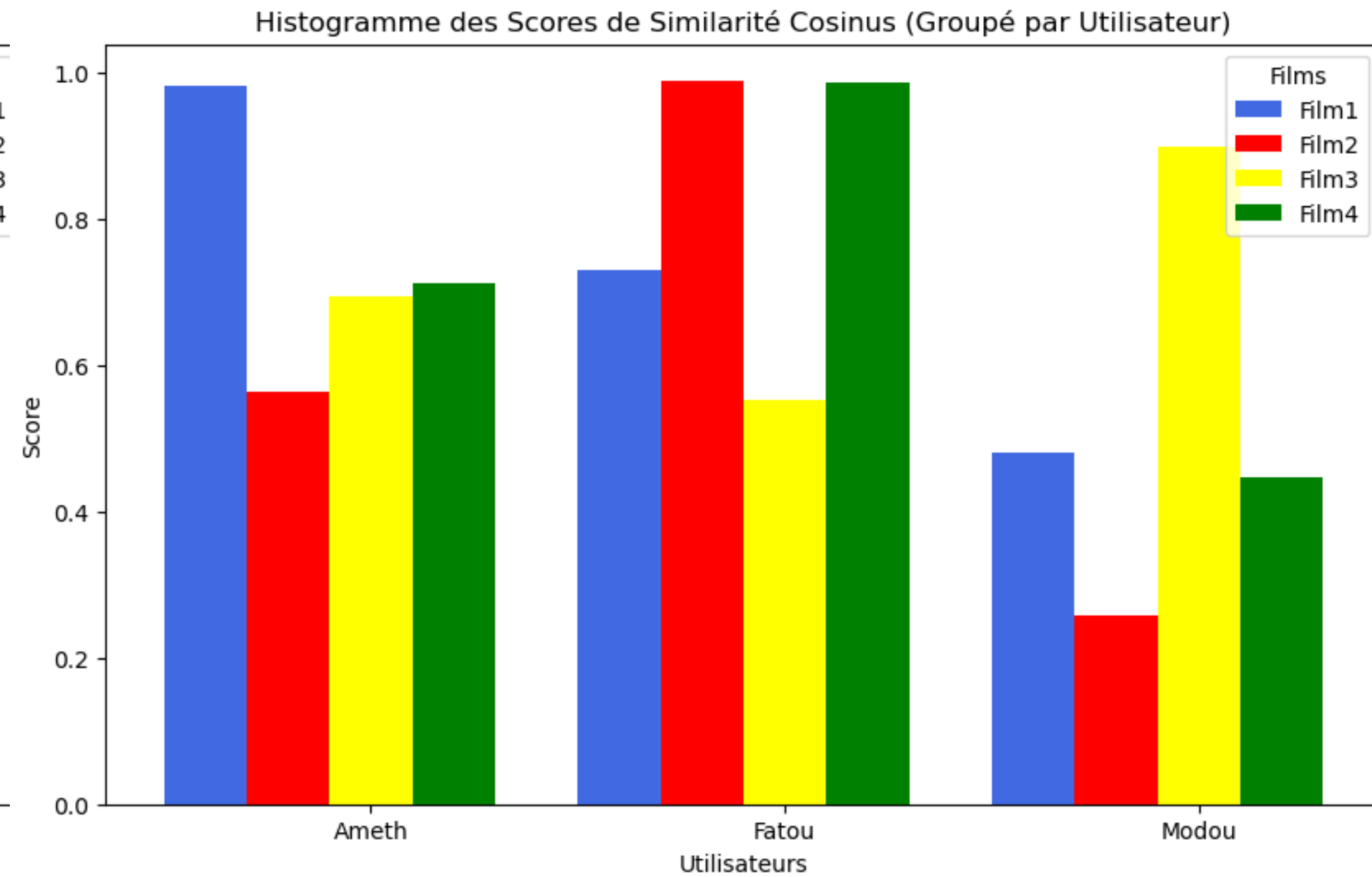
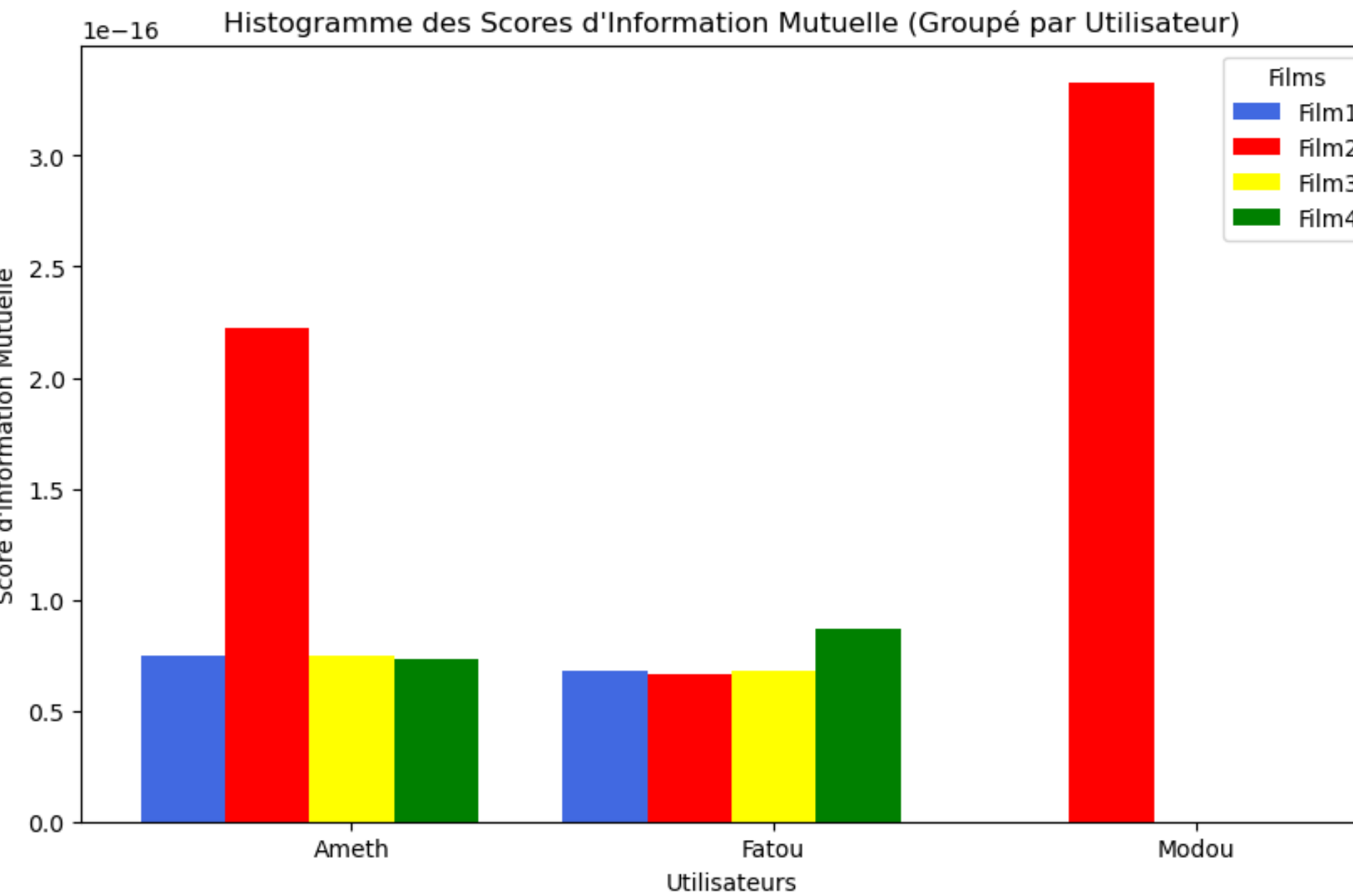


Comparaison avec la Similarité Cosinus

L'information mutuelle capte **les dépendances non linéaires**, mais comment se compare-t-elle à une méthode plus simple comme la **similarité cosinus** utilisée dans de nombreux systèmes de recommandation ?

Nous allons implémenter la similarité cosinus et comparer les recommandations générées par chaque méthode.

Information Mutuelle VS Similarité Cosinus

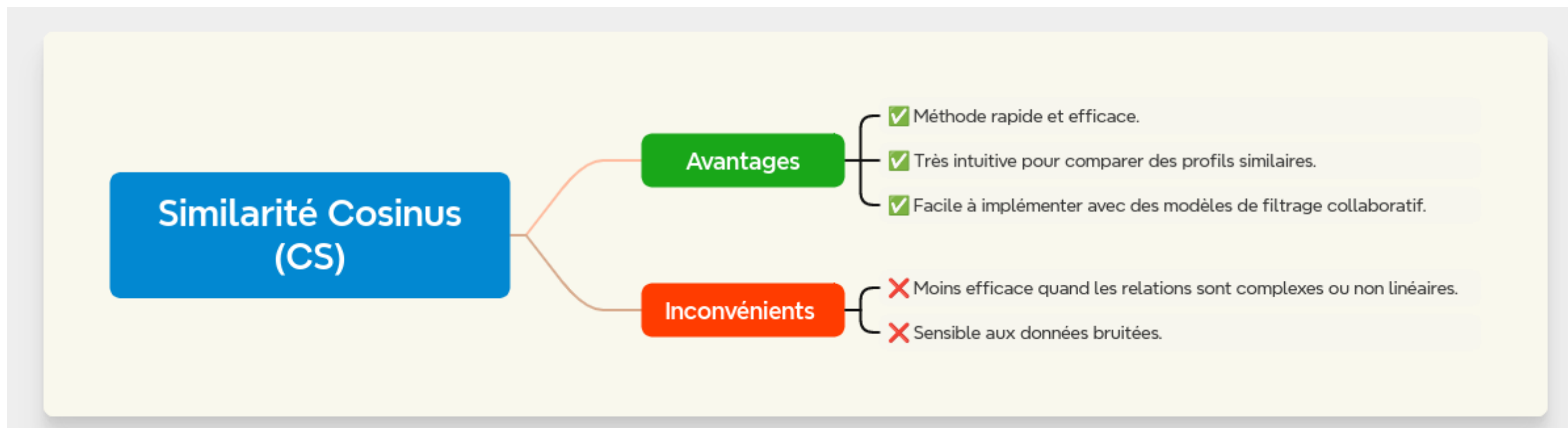
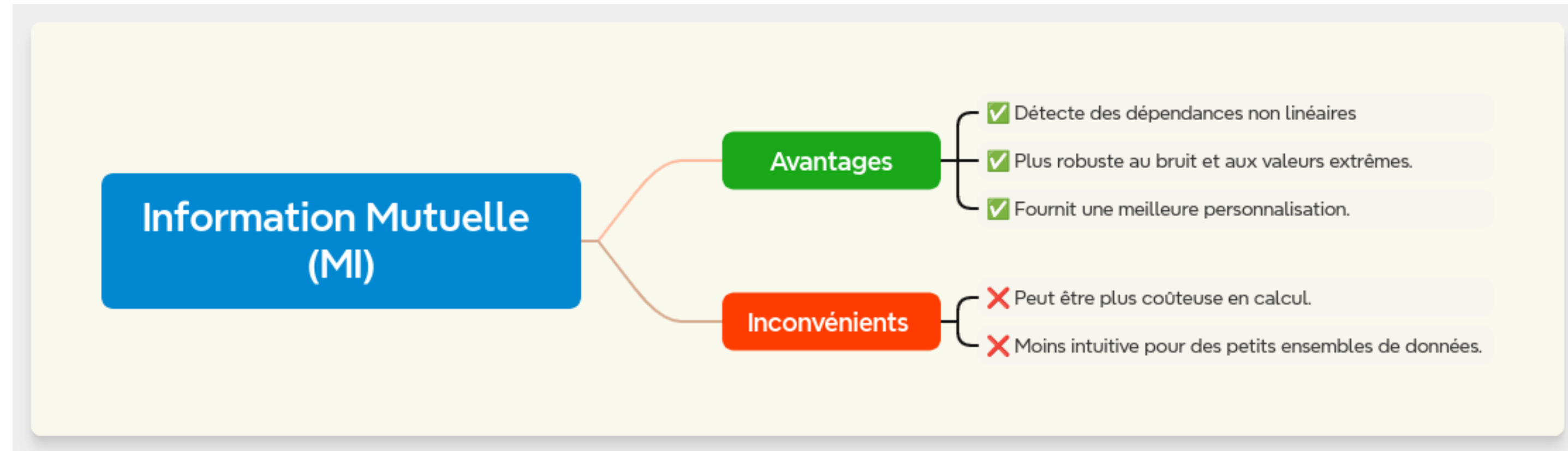


Analyse des Résultats

Ces graphiques permettent d'évaluer l'efficacité de chaque méthode dans l'association des utilisateurs aux films recommandés.

- L'**Information Mutuelle** identifie des relations **complexes et non linéaires** entre utilisateurs et films.
- La **Similarité Cosinus** est plus sensible aux valeurs absolues et fonctionne bien lorsque les profils sont directement proportionnels.

Interprétation des Résultats



Conclusion et Recommandation

L'utilisation de l'**information mutuelle** dans les **systèmes de recommandation** permet de capturer des *relations subtiles* entre les profils des utilisateurs et les caractéristiques des objets. Cette approche offre une **personnalisation avancée, robuste et adaptée** aux besoins individuels de chaque utilisateur, améliorant ainsi la **pertinence** et la **qualité** des recommandations.

- ✓ Si vous avez des relations **complexes** entre utilisateurs et films, privilégiez l'**Information Mutuelle**.
- ✓ Si vous voulez une approche **rapide et classique**, la **Similarité Cosinus** est un bon choix.
- ✓ Une approche **hybride** peut être envisagée, combinant les forces des deux méthodes.



Théorie de l'information

**Merci pour
votre attention**



Jërëjëf !