

# Projet : CinéGraph - Système Avancé de Recommandation de Films

## Contexte

Dans l'industrie du streaming, la personnalisation des recommandations est cruciale. CinéGraph vise à créer un système de recommandation de films sophistiqué, utilisant le filtrage collaboratif et l'analyse de graphes pour offrir des suggestions précises et variées aux utilisateurs.

## Objectifs

1. Développer un système de recommandation de films basé sur le filtrage collaboratif ([What is collaborative filtering? | IBM](#)).
2. Utiliser des algorithmes de graphes pour améliorer la qualité des recommandations.
3. Implémenter des visualisations interactives pour analyser les relations entre films et utilisateurs.
4. Optimiser les performances du système pour gérer de grands ensembles de données.

## Outils et Bibliothèques

- **NumPy** : Pour les calculs numériques efficaces
- **pandas** : Pour la manipulation et l'analyse des données
- **NetworkX** : Pour la modélisation et l'analyse des graphes
- **Matplotlib/Seaborn** : Pour la visualisation des résultats

## Fonctionnalités Requises

1. Gestion et Préparation des Données :
  - Utiliser pandas pour importer et nettoyer un large dataset de films et d'évaluations (ex: MovieLens : [MovieLens | GroupLens](#))
  - Créer des DataFrames efficaces pour stocker les informations sur les films, utilisateurs et évaluations
2. Implémentation du Filtrage Collaboratif :
  - Utiliser NumPy pour calculer efficacement les similarités entre utilisateurs et entre films
  - Implémenter les méthodes de filtrage collaboratif "user-based" et "item-based"
3. Analyse de Graphes :
  - Utiliser NetworkX pour créer un graphe biparti utilisateurs-films
  - Implémenter des algorithmes de recommandation basés sur la structure du graphe (ex: PageRank personnalisé)
4. Visualisation et Analyse :

- Créer des visualisations interactives avec Matplotlib pour :
    - Afficher les relations entre films (graphe de similarité)
    - Visualiser les clusters d'utilisateurs
  - Utiliser Seaborn pour générer des heatmaps de corrélation entre genres de films
5. Optimisation des Performances :
- Utiliser les capacités de calcul vectoriel de NumPy pour accélérer les calculs de similarité
  - Implémenter des techniques de réduction de dimensionnalité (ex: SVD avec NumPy) pour gérer de grands ensembles de données
6. Interface Utilisateur :
- Développer une interface en ligne de commande pour interagir avec le système
  - Permettre aux utilisateurs de recevoir des recommandations et de visualiser les analyses

## Étapes du Projet

1. Préparation et Exploration des Données :
  - Charger le dataset avec pandas
  - Effectuer une analyse exploratoire des données (EDA) avec pandas et Seaborn
2. Implémentation du Filtrage Collaboratif :
  - Développer les fonctions de calcul de similarité avec NumPy
  - Implémenter les algorithmes de recommandation basés sur les utilisateurs et les items
3. Modélisation par Graphes :
  - Construire le graphe utilisateurs-films avec NetworkX
  - Implémenter des algorithmes de recommandation basés sur la structure du graphe
4. Visualisation et Analyse :
  - Créer des visualisations de graphes avec NetworkX et Matplotlib
  - Générer des graphiques d'analyse avec Seaborn (distributions, corrélations)
5. Optimisation et Évaluation :
  - Optimiser les calculs avec NumPy
  - Évaluer les performances des différentes méthodes de recommandation
6. Développement de l'Interface :
  - Créer une interface CLI pour interagir avec le système

- Intégrer les visualisations dans l'interface utilisateur

### **Livrables Attendus**

1. Code source Python bien structuré et commenté
2. Notebook Jupyter détaillant l'analyse exploratoire des données et les résultats
3. Rapport technique incluant :
  - Description des algorithmes implémentés
  - Analyse des performances des différentes approches
  - Interprétation des visualisations et des résultats d'analyse de graphes
4. Présentation du projet avec démonstration des visualisations

### **Défis Techniques**

- Gestion efficace de grandes matrices de données avec NumPy
- Optimisation des algorithmes de graphe pour passer à l'échelle
- Création de visualisations informatives et interactives