



## **Cahier des charges**

**Option : Master 2 Technologie de l'hypermédia**

**Cours : Théorie et conception des hypermédias**

### **Thème**

---

**Conception et réalisation d'un système de recommandation à  
filtrage collaboratif de contenu cinématographique**

---

***Réaliser par :***

AREZKI Mohamed Nadji

## Table des matières

<b>1. Introduction</b>	<b>3</b>
a. Définition du projet :	3
b. Public ciblé :	3
c. Acteur du projet :	3
d. Elément clés du projet :	3
<b>2. Etat de l'art :</b>	<b>4</b>
a. Introduction :	4
Les systèmes de recommandation sont devenus un domaine de recherche indépendant de la recherche d'information au milieu des années 1990. Ces dernières années, avec le développement du Web et plus particulièrement des plateformes e-commerce, l'intérêt pour les systèmes de recommandation a considérablement augmenté. A l'origine, les systèmes de recommandation sont apparus pour essayer de résoudre les problèmes liés à la surcharge informationnelle (surcharge cognitive) (Soualah-Alila F., 2015).	4
b. Définitions des systèmes de recommandation :	4
c. Classification des systèmes de recommandation :	4
i. Recommandation basée sur le filtrage collaboratif :	5
ii. Avantages et inconvénients du filtrage collaboratif	6

## 1. Introduction

Dans le cadre de la validation du cours Théories et conception des hypermédias, on nous a demandé de réaliser un cahier des charges pour un projet que nous avons proposé.

Dans ce projet nous avons proposé de concevoir et de développer un système de recommandation à filtrage collaboratif pour le contenu cinématographique.

### a. Définition du projet :

Le projet consiste en le développement d'une plateforme web interactive où les utilisateurs peuvent accéder à une liste non-exhaustive de contenu cinématographique.

L'accès est effectué via authentification de l'utilisateur par le biais de son email et son mot de passe.

L'utilisateur a le choix entre deux catégories majeures Films/Série, après avoir choisie l'une d'elle, un panel de produit cinématographique qui lui est recommandé en tenant compte des paramètres de choix fait précédemment ainsi que des préférences de profils utilisateur qui sont similaires à ces films regarder précédemment.

L'utilisateur a la possibilité d'émettre un avis sur chaque film disponible sur la plateforme, il peut aussi visionner la bande-annonce sur le site via une API reliée à un site tiers (Youtube).

### b. Public ciblé :

Le projet cible tous les amateurs de contenu cinématographique, disponible sur plusieurs supports numériques (Ordinateur, téléphone et tablette).

### c. Acteur du projet :

Ce projet se déroule sous format individuel, et sera réalisé par moi AREZKI Mohamed Nadji, le travail sera jugé et noté par le professeur SALEH Imad.

### d. Éléments clés du projet :

- Durée du projet : 8 semaines.
- Coût du projet : 0 euro, projet pédagogique.
- Qualité : Approche Agile et suivie continue de la réalisation.

## 2. Etat de l'art :

### a. Introduction :

Les systèmes de recommandation sont devenus un domaine de recherche indépendant de la recherche d'information au milieu des années 1990. Ces dernières années, avec le développement du Web et plus particulièrement des plateformes e-commerce, l'intérêt pour les systèmes de recommandation a considérablement augmenté. A l'origine, les systèmes de recommandation sont apparus pour essayer de résoudre les problèmes liés à la surcharge informationnelle (surcharge cognitive) (Soualah-Alila F., 2015).

Les systèmes de recommandation sont utilisés aujourd'hui dans divers domaines et les items recommandés varient selon les domaines. Un item peut donc être un livre, une vidéo, un article ou encore un morceau de musique (Lémdani, 2016). Les différents acteurs du Web qui développent ces outils de recommandations sont donc de plus en plus nombreux. Le site web de vente en ligne Amazon a en effet accru significativement le nombre de ses visites et de ses ventes grâce à cet outil. Certains utilisateurs visitent ce site principalement dans le but de bénéficier de ces recommandations, même s'ils n'ont pas l'intention d'y acheter de produits. Et il en va de même pour des sites comme Allociné pour les cinéphiles, Deezer pour la musique, ou Goodreads pour la recommandation de livre (Laurent Candillier, 2012).

### b. Définitions des systèmes de recommandation :

Il existe plusieurs définitions des systèmes de recommandation dans la littérature, l'une d'elles explique que les systèmes de recommandation sont des outils qui assistent les utilisateurs avec les problèmes de surcharge croissante d'informations et améliorent la gestion de la relation client en fournissant aux utilisateurs des recommandations personnalisées de produits ou de services, en d'autre terme les systèmes de recommandation tentent de prédire les évaluations pour les produits inconnus, pour chaque utilisateur, souvent en utilisant les évaluations des autres utilisateurs et recommandent les N meilleurs items ayant la plus haute valeur d'évaluation prédite (CAMELIA DADOUCHI, 2017).

### c. Classification des systèmes de recommandation :

Il est possible de classer les systèmes de recommandation de différentes manières. La classification la plus connue est une classification selon trois approches : la recommandation basée sur le contenu, la recommandation par filtrage collaboratif et les systèmes hybrides.

- Recommandation basée sur le contenu.
- Recommandation basée sur le filtrage démographique.
- Recommandation basée sur le filtrage hybride.
- **Recommandation basée sur le filtrage collaboratif.**

Dans notre projet nous allons utiliser le système de **recommandation basée sur le filtrage collaboratif**.

#### i. Recommandation basée sur le filtrage collaboratif :

Le filtrage collaboratif, exploite les appréciations des utilisateurs sur les ressources. La représentation des appréciations se fait en règle générale par des notes. À la différence de l'approche basée sur le contenu, qui utilise les ressources précédemment notés par un seul utilisateur  $u$ , l'approche de filtrage collaboratif s'appuie sur les notes de tous les utilisateurs du système.

Ces notes sont soit attribuées de façon explicite par les utilisateurs, soit de façon implicite déduite à partir des actions réalisées par cet utilisateur. Le filtrage collaboratif est une des techniques les plus explorées du domaine de la recommandation (Das *et al.*, 2007).

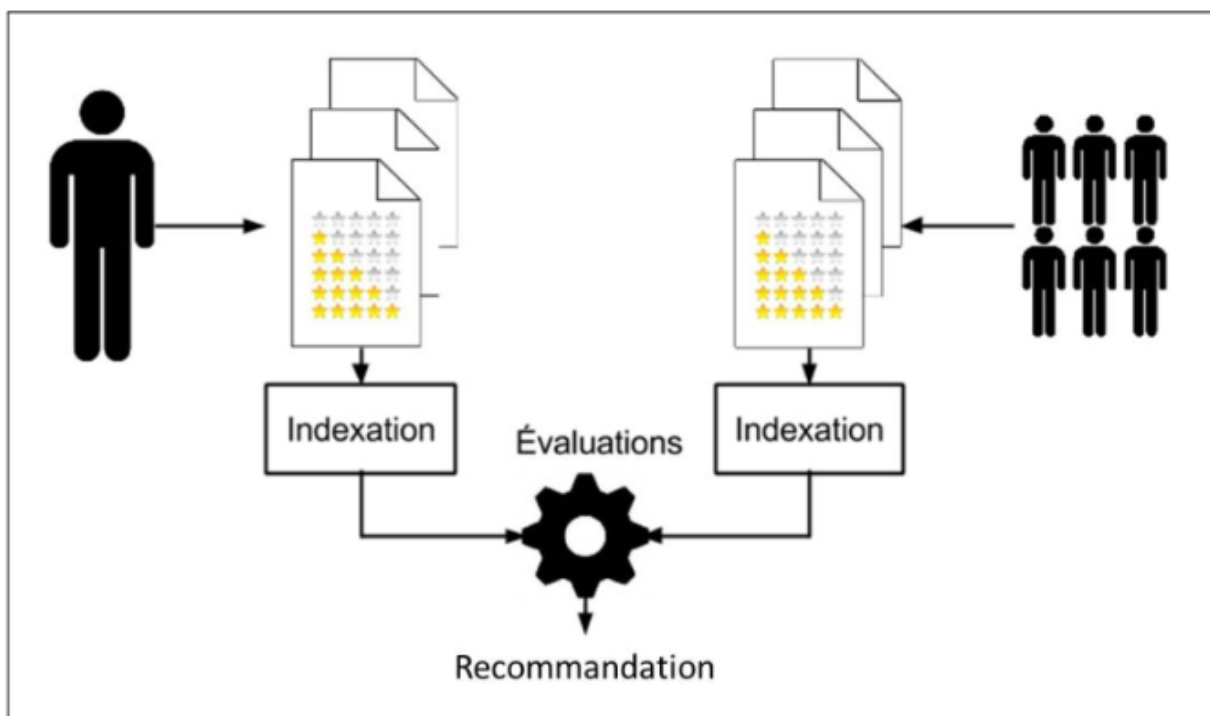


Figure 1 Filtrage collaboratif

Il existe plusieurs approches pour le filtrage collaboratif :

- **Approche mémoire (basée sur l'utilisateur).**
- Approche mémoire (basée sur les items).
- Approche modèle.

## ii. Avantages et inconvénients du filtrage collaboratif

- Les algorithmes de filtrage collaboratif sont particulièrement applicables et utiles dans des domaines où l'analyse de contenu est très coûteuse ou difficile, comme la musique et la suggestion de film (Amir Hossein Nabizadeh Rafsanjani, 2013).
- Le filtrage collaboratif se distingue par sa capacité à recommander à un utilisateur ce qui est hors du familier, un utilisateur peut se voir recommander des items de genres différents (Yammine, 2005).
- Le problème du « système débutant » survient lorsque la matrice d'usage est vide. Les méthodes de filtrage collaboratif ne peuvent fonctionner que s'il existe des informations dans cette matrice d'usage.
- Problème de démarrage à froid, cas du nouvel utilisateur : Un problème commun au filtrage basé sur le contenu et au filtrage collaboratif, est qu'un nouvel utilisateur qui n'a pas encore accumulé suffisamment d'évaluations ne peut pas avoir de recommandations pertinentes (Yammine, 2005).
- Problème de démarrage à froid cas du nouvel item, il doit avoir suffisamment d'évaluations pour que celui-ci soit pris en considération dans le processus de recommandation.
- Problème de démarrage à froid cas système débutant, Le cas du système débutant provient lors du lancement d'un nouveau service de recommandation. Le système ne possède alors aucune information sur les utilisateurs et sur les items.

### d. Exemples de systèmes de recommandation :

#### i. Systèmes de recommandation d'Amazon :

Le système de recommandation d'amazon.com utilise une méthode collaborative Item-item qui vise à suggérer à un utilisateur les items les plus adaptés à ses préférences. Ces suggestions dépendent du comportement de l'utilisateur. Les techniques de filtrage collaboratives traditionnelles cherchent les utilisateurs les plus similaires à un utilisateur donné et recommande à cet utilisateur les items préférés et les mieux évalués de ces utilisateurs similaires. Par contre le système d'Amazon calcule les similarités entre les items aux lieux des utilisateurs. Cette méthode consiste à prendre tous les items qu'un utilisateur a évalués et calcule leurs similarités avec d'autres items en produisant une matrice item-item pour calculer la similarité entre chaque pair d'items, les suggestions données à l'utilisateur seront basées sur cette matrice calculée.

Le moteur de recommandation d'Amazon utilise aussi le machine Learning, depuis plus de 20 ans, Amazon investit massivement dans le domaine de l'intelligence artificielle.

#### e. Système de recommandations de Netflix :

Netflix est une entreprise américaine proposant des films et séries télévisés en flux continu sur internet, implantée à travers le monde. Une des grandes préoccupations de cette entreprise était de capturer l'attention de l'utilisateur en améliorant leurs systèmes de recommandation, car les chercheurs de Netflix ont constaté à travers leur recherche sur les consommateurs que la plupart des utilisateurs perdent leur intérêt dans le site dans les premières 60 à 90 secondes de recherche s'ils n'arrivent pas à trouver ce qui les intéresse (Stenovec, 2016). Pour résoudre ce problème, ils ont développé et utilisé plusieurs algorithmes de recommandations. et voici quelques exemples :

##### **Personalized Video Ranker PVR :**

*Cet algorithme vise à organiser les items recommandés dans différentes rangées dans la page de recommandation, chaque rangée représente un genre de films ou une catégorie spéciale et personnalisée comme "Tendances" ou "continue à regarder" pour les séries ou films qu'on a déjà commencé à regarder (HUNT & NEIL, 2015).*

##### **Top-N Video Ranker :**

*C'est un algorithme qui utilise et combine plusieurs d'autres algorithmes et métriques de recommandation afin de produire une liste des N meilleures recommandations possibles qui sont les mieux adaptés et personnalisés pour les préférences de l'utilisateur (HUNT & NEIL, 2015).*

#### f. Systèmes de recommandation sur les réseaux sociaux

Les systèmes de recommandation sont très importants pour améliorer l'expérience des utilisateurs des réseaux sociaux, exemple : la file d'actualité de Facebook qui organise et trie les actualités et publications les plus intéressantes pour l'utilisateur de haut vers le bas et pour faire ça Facebook utilise une approche collaborative, donc une publication qui est suivie par plusieurs amis d'un utilisateur aura plus de chance d'apparaître en haut de sa file d'actualité, mais l'algorithme prend en considération d'autres facteurs comme la nouveauté d'une publication. Cette même approche est utilisée dans les réseaux sociaux pour la recommandation des amis, des groupes, des pages et même des événements (Mmanish Argwal).

### 3. Missions et découpage des taches :

La méthode de gestion de projet que nous allons appliquer pour le processus de développement sera en méthode classique en cycle V, qui va nous permettre de valider chaque étape de développement et du faite que je suis seul sur le projet.

#### a. Découpage du projet :

Le projet est découpé en 4 Catégories :

- Documentations.
- Données et traitement des données.
- Conception et modélisation.
- Développement Back-end et Front-end.

#### b. Taches de chaque catégorie :

##### Documentations :

Taches	Description
Cahier des charges	Rédaction des cahiers des charges
Documentation du code	Rédaction de la documentation du code base
Documentation de la Platform	Rédaction de la documentation pour l'utilisation de la Platform
Rapport final	Rédaction du document du rapport final noté

##### Données et traitements des données :

Taches	Description
Recherche du dataset	Choix du jeu de données le plus optimales
Nettoyage du dataset	Suppression des doublons, compléter le dataset
Normalisation du dataset	Standardisé les valeurs du afin de gagner du temps de traitement et d'avoir des valeurs significatif
Mise en place de la matrice de similarité	
Liste de scores miliaires	Calcule des scores de similarité entre film
Liste de score similaire pour de multiples contenus	Calcule des scores de similarité entre différente catégories de films
Calcule des avis utilisateurs	Calculer du score de l'avis utilisateur
Calcule des moments les plus visionné	
Test des résultats	



## Conception et modélisation :

Taches	Description
Diagrammes de cas d'utilisation	Décrire un cas d'utilisation et un scénario utilisateur
Diagramme de classe	Définir et découper les éléments du projet en objets
Entité relation	Définir les entités et leur relation afin de construire la base de données
Diagrammes d'activité	Décrire le déroulement de certaine fonctionnalité

## Développement Back-end et Front-end :

Taches	Description	Catégories
Mise en place des classes en django	<b>implémentation des modèles ORM django</b> <ul style="list-style-type: none"><li>• Utilisateur</li><li>• Contenu<ul style="list-style-type: none"><li>○ Film</li><li>○ Série</li></ul></li><li>• Profile</li><li>• List de favoris</li><li>• Commentaires</li></ul>	Back
Mise en place des View	<b>Implémentation des traitements de chaque vue utilisateur</b> <ul style="list-style-type: none"><li>• Authentification</li><li>• List du contenu</li><li>• Profile</li><li>• API Bande-Annonce</li><li>• Systèmes d'avis</li><li>• List de favoris</li><li>• Commentaires</li></ul>	Back
Mise en place des templates	<b>Implémentation des templates utilisateurs</b> <ul style="list-style-type: none"><li>• Authentification</li><li>• List du contenu</li><li>• Profile</li><li>• API bande-Annonce</li><li>• Systèmes d'avis</li><li>• List de favoris</li><li>• commentaires</li></ul>	Front
Mise en place des filtres utilisateur	<b>Implémentation des filtres de recherches</b> <ul style="list-style-type: none"><li>• Recherche par genre</li><li>• Recherche par catégorie</li></ul>	

	<ul style="list-style-type: none"> <li>• Recherche par similarité de profile</li> <li>• Recherche par nombre d'avis</li> <li>• Recherche par avis positive</li> </ul>	
--	---	--

### c. Technologie utilisé :

Les langages utilisés dans ce projet seront comme suit :

#### Traitement des données :

la partie des traitement de données elle sera faite en python en utilisant la librairie SKLearn.



#### Back-End :

Django pour la partie back-end WEB et SQLite pour la base de données afin de simplifier les requêtes et le déploiement du projet.



#### Front-end :

Une première partie sera faite en BOOTSTRAP afin d'avoir un rendu le plus rapidement possible, ensuite si j'ai le temps je passerais sur du ReactNative.



## Bibliographie

- Abhinandan Das, M. D. (2007, Mai). *Google News Personalization: Scalable Online Collaborative Filtering*. Google Inc.
- Adomavicius, & al. (2005).
- Amir Hossein Nabizadeh Rafsanjani, N. S. (2013). Recommendation Systems: a review. *International Journal of Computational Engineering Research*, 47-52.
- Arnautu, O. R. (2012, Juillet). *Mures : Un système de recommandation de musique*. Université de Montréal.
- Baeza-Yates, R.-N. (1999). *Modern Information Retrieval*. Addison-Wesley.
- Benoît CHARROUX, A. O.-M. (2009). *UML2 : pratiquer la modélisation*. Pearson Education France.
- Benouaret, I. (2017). *Un système de recommandation contextuel et composite pour la visite personnalisée de sites culturels*. Université de Technologie de Compiègne.
- Bonnin, G. (2010). *Vers des systèmes de recommandation robustes pour la navigation Web : inspiration de la modélisation statistique du langage*. Université Nancy II.
- Bouchindhomme, & Rochlitz. (1992).
- Burke, R. (2002). *Hybrid Recommender Systems: Survey and Experiments*. California State University, Fullerton Department of Information Systems and Decision Sciences.
- CAMELIA DADOUCHI, B. A. (2017). *État de l'art sur les systèmes de recommandation*. Département de mathématiques et génie industriel, École Polytechnique de Montréal.
- Chickering D.M., H. D. (1997). *A Bayesian approach to learning Bayesian networks with local structure*. Thirteenth Conference on Uncertainty in Artificial Intelligence.
- Damien Poirier, F. F. (2010). *De la Classification d'Opinion à la Recommandation : l'Apport des Textes Communautaires*. Récupéré sur ATALA: <http://www.atala.org/IMG/pdf/1-Poirier-TAL51-3.pdf>
- HUNT, C. A.-U., & NEIL. (2015). The Netflix Recommender System : Algorithms, Business Value, and innovation. Netflix, Inc.