



Systeme de recommandation de films/séries

PROJET IA E14



Réalisé par:
Papa Talla DIOUM
Nadir KAREM
Sekouba TOURE
Yulin REN

Professeurs encadrents:
Nizar OUARTI
Maxime EGEA



Introduction

Dans cette ère de surabondance numérique, les utilisateurs sont souvent submergés par les choix disponibles, qu'il s'agisse de films, de séries, ou d'autres contenus. Comment simplifier leurs décisions tout en augmentant leur satisfaction ? L'intelligence artificielle ouvre une nouvelle voie. Aujourd'hui, nous vous présentons une solution innovante qui transforme la découverte de contenus grâce à des recommandations personnalisées et pertinentes.



Presentation du projet

Objectif ?

Créer un système de recommandation qui propose des films similaires à ceux choisis par les utilisateurs, en utilisant des techniques avancées d'intelligence artificielle.





Besoins répondu ?

Aider les utilisateurs à trouver
rapidement des films
correspondant à leurs goûts.

Proposer un système Transversale
à toutes les plateformes de
diffusion

Réduire le temps consacré à la
recherche et maximiser la satisfaction
des utilisateurs sur les plateformes.

Collecte et Traitement des Données

Les jeux de données constituent la base fondamentale des systèmes de recommandation de films. Ils sont comme les fondations d'une maison. Sans de bonnes données, le système de recommandation ne peut pas prédire avec précision les préférences des utilisateurs

Sources de Données

IMDB : Importation de bases de données donnant des méta données comme le titre, la durée, le résumé etc.

Processus de Traitement

- Préparation des embeddings textuels pour les descriptions de films.
- Normalisation des caractéristiques numériques.
- Fusion des caractéristiques et des embeddings dans un vecteur unique pour chaque film.

Comment traiter les jeux de données

Nettoyage des données

Avant de former le modèle, nous devons nettoyer et standardiser les données

Extraction de caractéristiques

Pour améliorer les capacités de recommandation du modèle, nous devons extraire des caractéristiques des données brutes

Standardisation des données

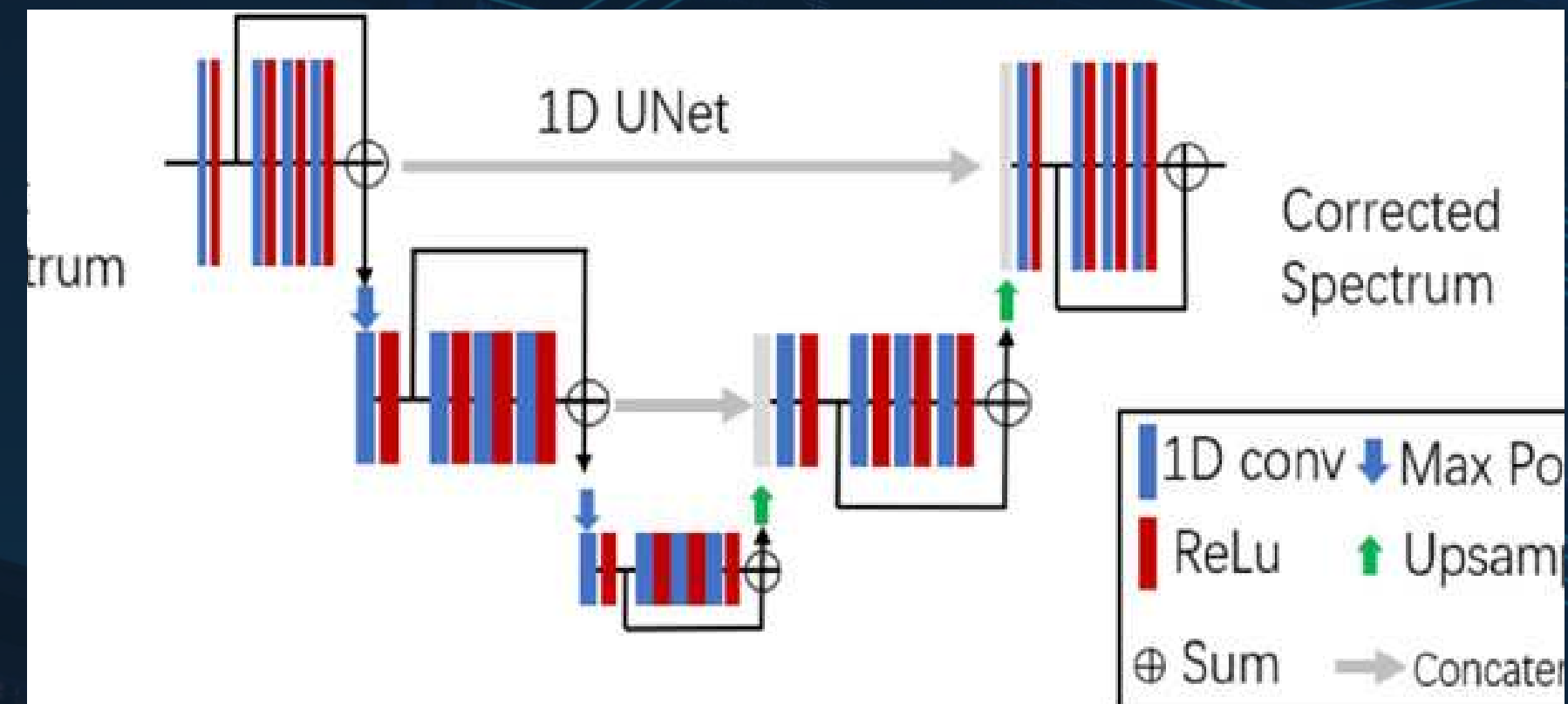
Pour éviter que les différences d'échelle entre les caractéristiques n'affectent le modèle

$$x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

Entrainement de l'IA avec UNET 1D

Fonctionnement du Modèle

1. Encodeur : Extraction des caractéristiques importantes à travers plusieurs couches convolutives.
2. Bottleneck : Réduction dimensionnelle pour capturer l'essence des données.
3. Decodeur : Reconstruction à partir des embeddings compressés.
4. Output Layer : Génération des vecteurs compressés utilisés pour les recommandations.



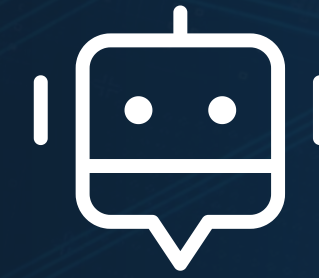


Systeme de Recommandation



Vue d'ensemble

- Objectif : Offrir des recommandations pertinentes.
- Approche :
 - Utilisation des embeddings textuels pour représenter les résumés de films.
 - Combinaison avec d'autres critères pour enrichir les résultats.



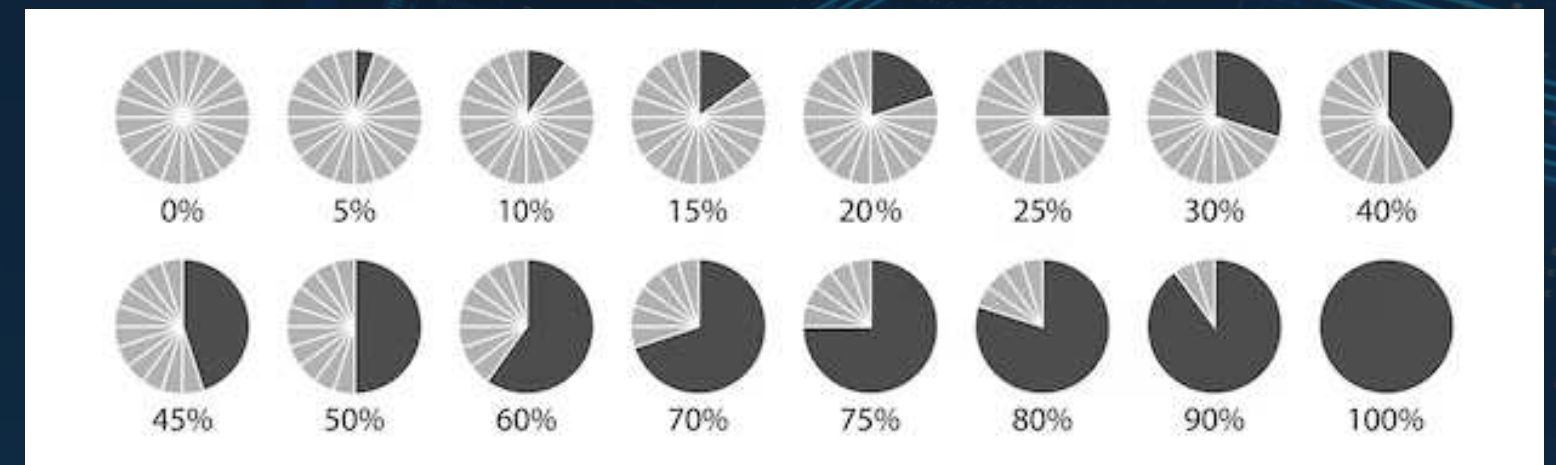


Embeddings Textuels

- Modèle utilisé : paraphrase-MiniLM-L6-v2 de Sentence Transformers.
- Caractéristiques des embeddings :
 - Vecteurs de 384 dimensions.
 - Capturent la sémantique des résumés.
 - Compact et performant pour des calculs rapides.
- Utilisation : Comparaison des films basée sur la similarité cosinus.

Critères de Recommandation

- Objectif : Pondération pour équilibrer l'impact de chaque critère.
- Critères principaux :
 - Collections (40%) : Films appartenant à la même franchise ou série.
 - Embeddings textuels (30%) : Mesure de la similarité des résumés.
 - Genres (15%) : Valorisation des genres communs.
 - Date de publication (10%) : Bonus pour les films proches en période de sortie.
 - Durée (5%) : Films de durée similaire.



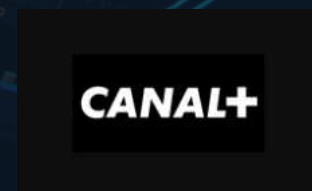
Calcul des Recommandations



- Étape 1 : Comparaison des embeddings textuels pour un score initial.
- Étape 2 : Évaluation des critères complémentaires (genres, collections, durée, etc.).
- Étape 3 : Application des pondérations pour un score global.
- Étape 4 : Tri des films selon le score et affichage des meilleures recommandations.

Perspectives d'Évolution

- Enrichissement de la base de données
- Tarification :
 - Abonnement gratuit avec publicité ou premium à 5 euros sans publicité, avec un mois d'essai gratuit.
 - Valorisation des données utilisateurs auprès de partenaires, dans le respect des réglementations.
- Partenariats stratégiques
- Diversification dans d'autres domaines: Musique, E-commerce, ...



Conclusion



Thank You

FOR YOUR ATTENTION

