



Diversification & Recommandation

Webinaire Culture-Inria 29 novembre 2023

Présentation du pass Culture

Une nouvelle politique culturelle

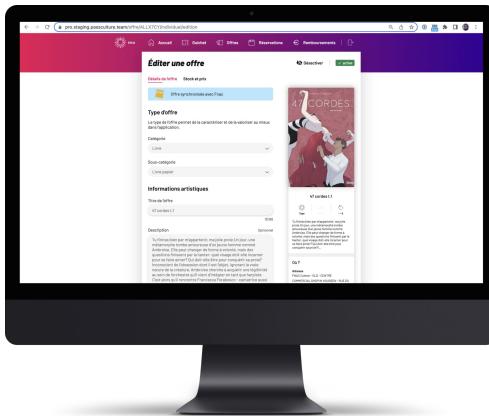


Une mission

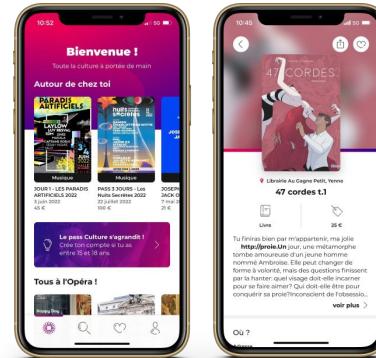
**intensifier et diversifier les pratiques
culturelles des jeunes.**

Quelle est la mesure de performance du pass Culture ?

L'évolution de la diversification des pratiques culturelles



Création d'un catalogue le plus diversifié possible



Mesure d'un score de diversification

Une appli au service des jeunes

Le pass Culture est une application gratuite et géolocalisée, destinée à favoriser l'accès à la culture aux jeunes de 15 à 18 ans

15 ans

16 ans

17 ans

18 ans

Part individuelle

20 €

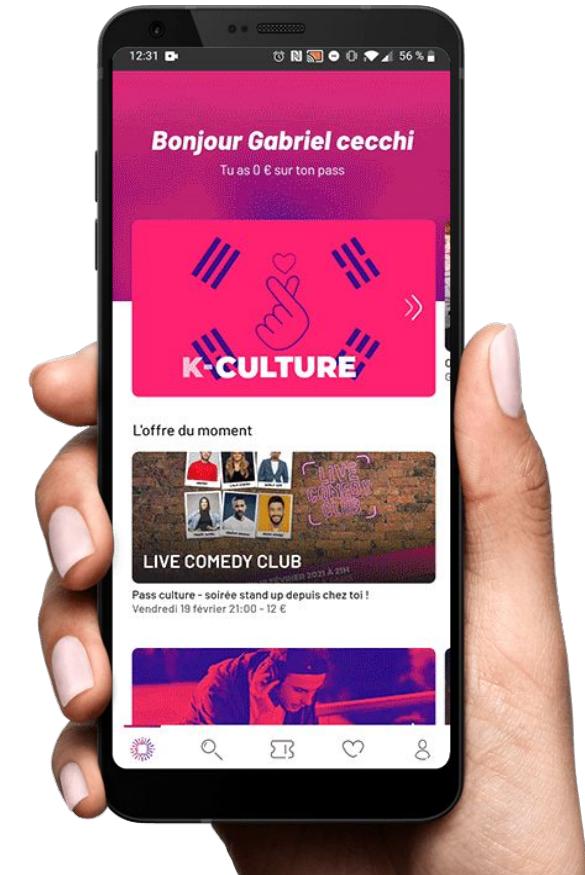
30 €

30 €

300 €

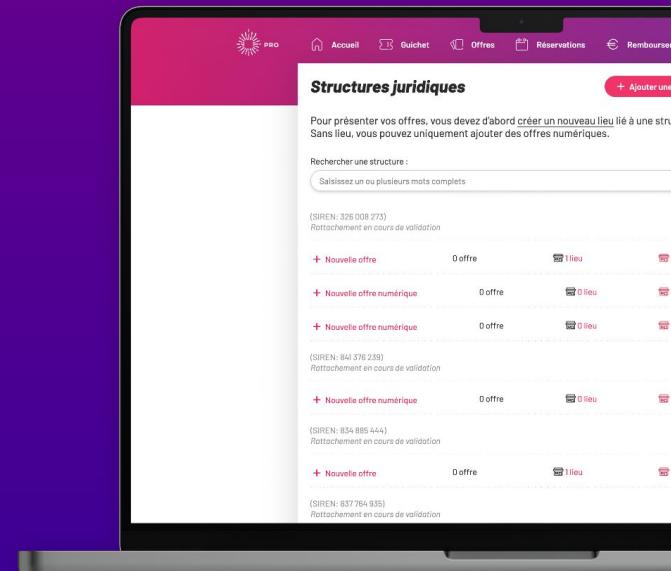
Part collective

120€



... au service des acteurs culturels et des collectivités

- Une **interface pour les acteurs culturels** qui y référencent leurs propositions.
- Un outil simple pour **accéder aux jeunes du territoire, communiquer directement auprès d'eux pour promouvoir leurs offres localement**.
- Un outil pour **connaître les pratiques et usages culturels de ces jeunes**, qui peut aider les acteurs culturels et les collectivités dans la construction de leurs offres auprès de ce public.
- Un outil qui peut ainsi **favoriser et renforcer l'accès des jeunes aux lieux et propositions culturels de leur territoire**.



La richesse de l'offre culturelle sur le pass



Places et abonnements

de spectacle vivant, de cinéma, de concert, de médiathèque, de festivals...



Biens culturels

livres, BD, DVD, CD, disques vinyles, instruments de musique...



Services numériques

ebooks, abonnements à la presse en ligne, jeux vidéo musique en ligne et SVOD*



Rencontres

avec les artistes, conférences, dédicaces, répétitions de spectacles, découverte des métiers...



Visites

de musées, de lieux historiques, de centres d'art...



Cours et ateliers

de danse, de théâtre, de musique, de chant, de dessin...



Matériel beaux arts

calligraphie, dessin, peinture, street art...

Un dispositif plébiscité par les jeunes et les acteurs culturels



2,6M

de jeunes inscrits actifs à ce jour



79%

des jeunes de 18 ans en France qui bénéficient du pass



31 000

partenaires culturels



20M

de réservations validées



Dans le top

des applications Apple Store et Android.

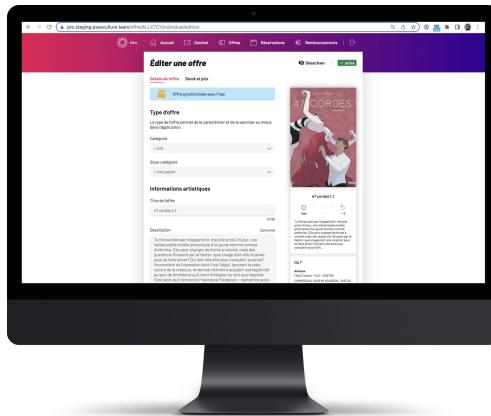


80M

d' offres créées

Quelle est la mesure de performance du pass Culture ?

L'évolution de la diversification des pratiques culturelles



Création d'un catalogue le plus diversifié possible



Mesure d'un score de diversification



Mesure de la diversification



Le score de diversification

La diversification

- Réserver une **offre qui possède des caractéristiques différentes** de celles réservées précédemment
- L'utilisateur a **effectué une découverte culturelle** dans la mesure où il a **initié une pratique** que nous n'avions pas observée jusqu'alors.

Comment quantifier la diversification ?

- Le jeune commence avec un score de 1
- À chaque nouvelle réservation, **un nombre de points est attribué à l'utilisateur** selon des critères liés à la nature de l'offre
- Plus le **score de diversification est élevé, plus la diversification est importante**



Le score de diversification

Comment quantifier la diversification ?

Critères globaux

- **Catégorie** : +1 point pour chaque nouvelle catégorie réservée par le jeune
- **Sous-catégorie** : +1 point pour chaque nouvelle sous-catégorie réservée par le jeune
- **Lieu** : +1 point pour chaque nouvelle librairie, salle de concert, salle de spectacle ou magasin de CDs visités, +2 point pour chaque nouveau lieu autre
- **Format** (*physique, numérique, évènement*) : +1 point pour chaque nouveau format de réservation

Critères particuliers (extra catégorie)

- **Macro rayon** (uniquement pour les livres) : + 1 point pour chaque nouveau livre acheté dans un rayon différent
- **Genre musical** (pour les concerts et les CDs / Vinyles) : +1 point pour chaque nouveau genre musical exploré
- **Type de spectacle** : +1 point pour chaque nouveau type de spectacle exploré (humour, cirque, théâtre...)

Exemple

Comment mesurer la diversification ?

1 pt



+2 pt



+5 pt



+0 pt



Catégorie : Livre
Sous-catégorie : Livre papier
Genre : Manga
Lieu : Fnac des Halles
Type : Offre physique

Livre
Livre papier
Littérature française +1
La Malle aux histoires +1
Offre physique

Concert +1
Spectacle représentation +1
Rap / hip hop +1
Zénith Paris +1
Événement +1

Livre
Livre papier
Manga
Fnac des Halles
Offre physique

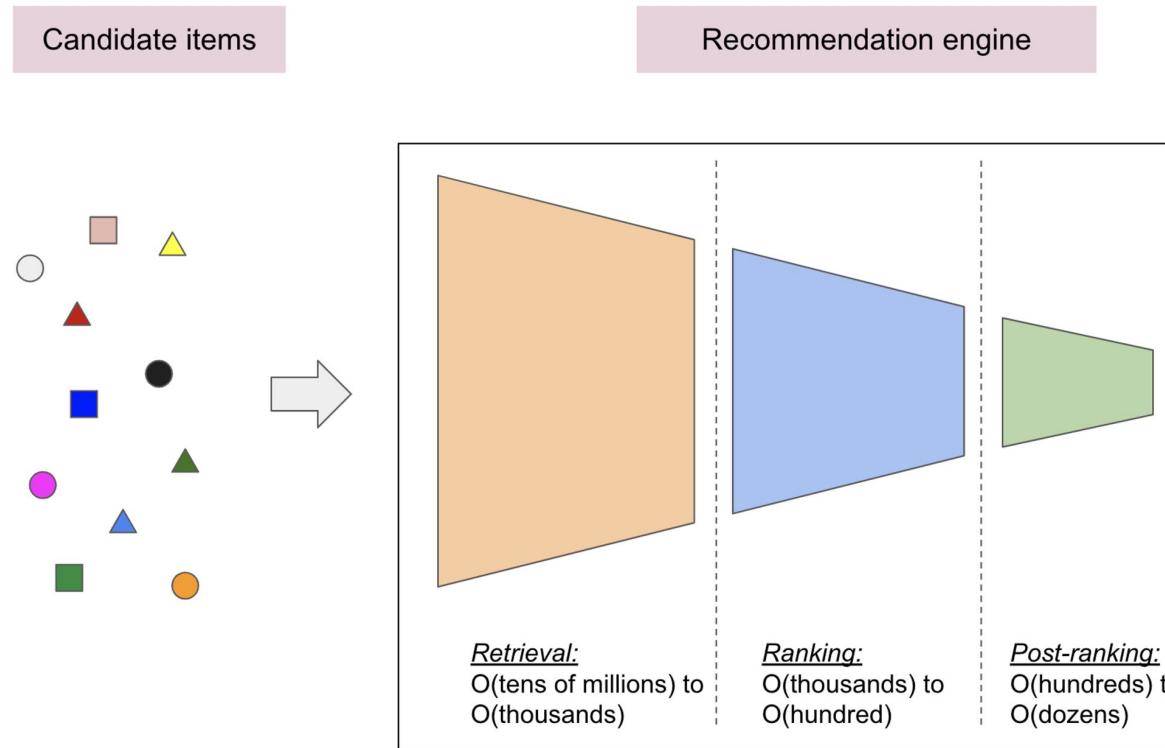
Après ces 4 réservations, le jeune obtient 8 points



Algorithme de recommandation

Recommandations personnalisées

Comment sélectionner les éléments les plus pertinents de notre catalogue pour chacun des utilisateurs



Retrieval :

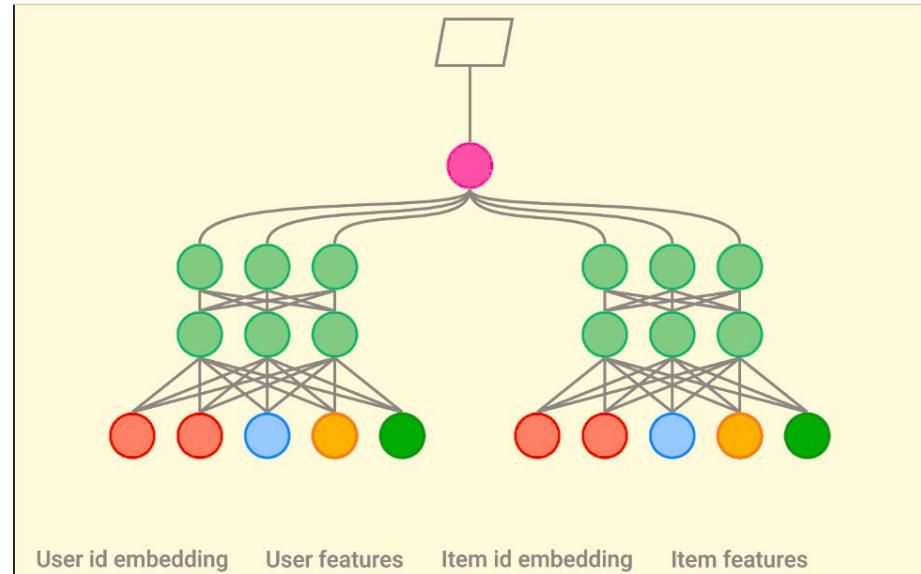
Trouver les X items les plus pertinents pour un utilisateur parmi tous les éléments du catalogue

Two-Tower Model

Modélisation d'un espace latent item et utilisateur qui représente la similarité/affinité entre les deux populations



TensorFlow



User et Item Tower

Exemple de caractéristiques en entrée

User features

- consultations d'offres
- réservations d'offres
- mise en favoris
- crédit restant
- ...

Item features

- catégories et sous-catégories (LIVRE, CINEMA, ...)
- **contenu sémantique de l'offre** (embedding)
- artiste, rayon (autres metadata)
- **image** (embedding)

Preprocessing

Certaines features sont transformées en amont via des modèles optimisés pour traduire les caractéristiques données dans un espace mathématique.

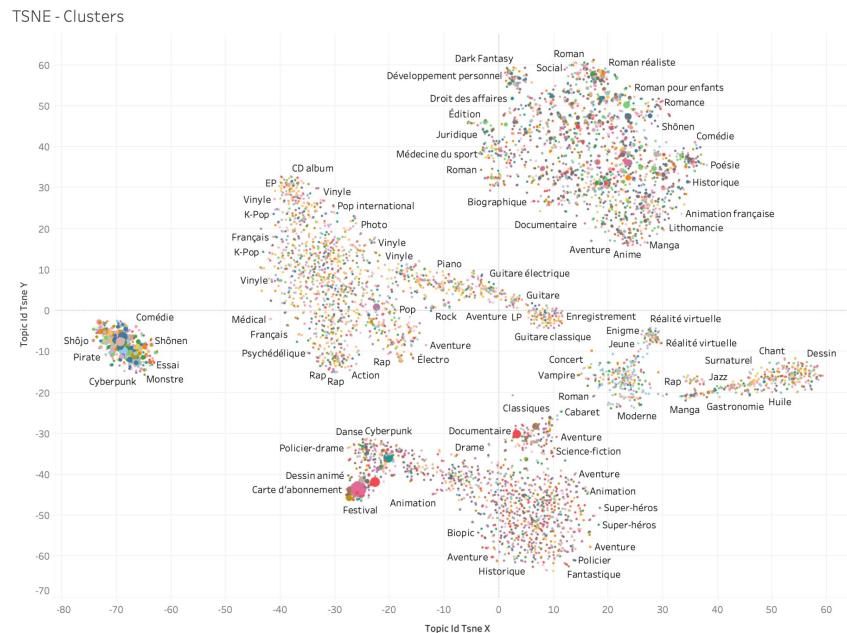


HUGGING FACE

Entraînement

- Entraînement du modèle basé sur les interactions passées (clics)
 - La métrique de validation ('top-k') est quant à elle calculée à partir des réservations
 - Output : Extraction des embeddings item × user

chaque item et utilisateur est maintenant représenté par un vecteur (64 dimensions)



TSNE sur des clusters sémantiques

Inférence

À partir des représentations mathématiques des utilisateurs et des items dans un espace commun il est possible de calculer la ‘proximité’ (score) entre une paire utilisateur/item (dot, L2).

Appliquer ces calculs à tous les éléments du catalogue à chaque recommandation est beaucoup trop coûteux pour être obtenu en direct

Pour accélérer ces calculs il est possible d'indexer l'espace en cases et ainsi ne plus avoir à calculer les scores pour la totalité du catalogue mais seulement ceux des cases adjacentes (ANN)



LanceDB

Applications

Mise en contexte

Offres similaires

Recherche d'items dans cet espace vectoriel

Algorithmes d'approximation de type plus proches voisins (ANN)

Recommandation

Produit matriciel entre matrice d'item × user, top items

Les items qui ont le plus de chance d'être réservés par l'utilisateur

Création de différents types des **playlists thématiques personnalisées**
les plus similaires, les plus proches de l'utilisateur, qui ont lieu bientôt, etc.

Ranking

Permet de réordonner les offres du retrieval en fonction du contexte de l'application (distance de l'utilisateur aux offres, crédit disponible...).

Elle permet par ailleurs d'inclure d'autres typologies d'offres qui peuvent être adaptées au contexte (nouveautés, tops).

L'objectif de cette phase est de passer de milliers d'offres à une centaine d'offres (150).

LightGBM

Régression linéaire basée sur les features de contexte pour prédire la cible suivante :

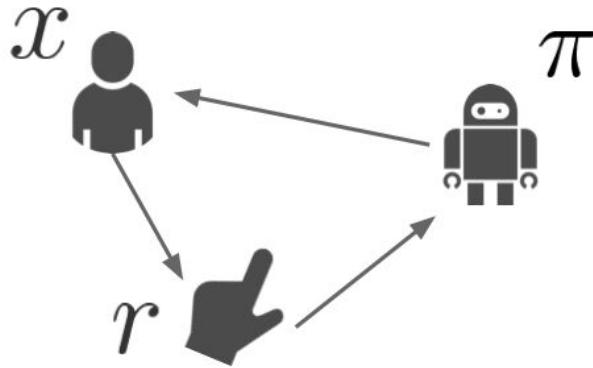
$$p(C = 1) + p(B = 1) \times (1 + \Delta_{div})$$

Avec C le clic, B le booking et Δ l'incrément moyen de diversification de l'offre

Travaux en cours avec Inria pour raffiner cette notion (modèles de bandit)

Contextual Bandit

Deploy policy online



Get the log data

$$\{(x_i, a_i, r_i)\}_{i=1}^n$$

$$\sim \prod_{i=1}^n p(x_i) \pi_e(a_i | x_i) p(r_i | x_i, a_i)$$

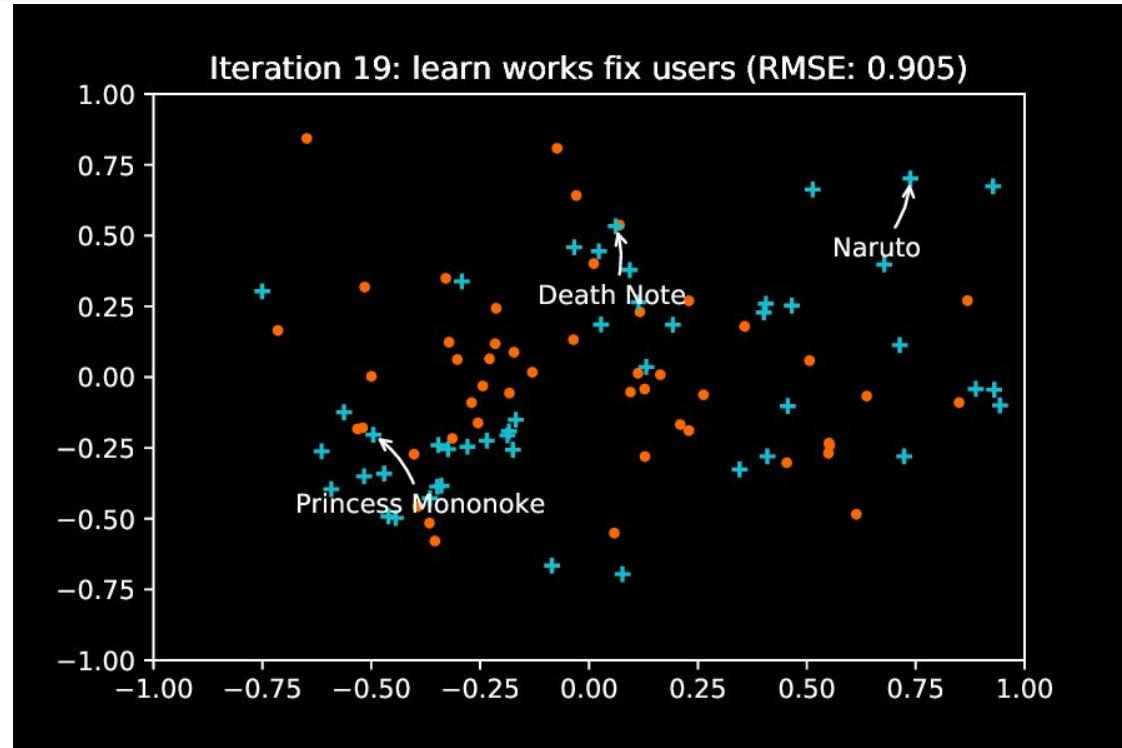
Estimate value as average reward

$$\hat{V}_{\text{A/B}}(\pi_e; \mathcal{D}_e) := \frac{1}{n} \sum_{i=1}^n r_i$$

Parcours dans l'espace de culture

Quelle métrique de diversité ?

Quels items permettent de rentrer facilement dans une nouvelle catégorie ?



Fin 🙌