

DÉVELOPPEMENT ET ÉVALUATION DE SYSTÈMES DE RECOMMANDATION DE MUSIQUES

BOUCHERY Loïc, LE CONTEL Loïc, RIHET Nathan

Université Laval

ABSTRACT

Nous abordons le développement de systèmes de recommandation pour les données sparses, en particulier les ensembles de données de musique. Nous appliquons une approche basée sur les autoencodeurs superficiels, en démontrant une performance améliorée par rapport aux méthodes de filtrage collaboratif état-de-l'art, y compris les modèles non linéaires et profonds.

Index Terms— Système de recommandation, SVD, ALS, EASE, MultVAE, CDAE, NDCG

1. INTRODUCTION

Dans ce travail, nous introduisons une méthode simple mais efficace pour recommander des items musicaux en utilisant des autoencodeurs superficiels, qui se distinguent par leur capacité à gérer des données sparses et à améliorer l'exactitude des recommandations.

2. ÉTAT DE L'ART

Cette section offre un aperçu des avancées récentes dans le domaine des systèmes de recommandation de musique, en mettant l'accent sur les techniques d'autoencodeurs, de filtrage collaboratif, et d'autres méthodes pertinentes.

2.1. Autoencodeurs dans la Recommandation Musicale

Les autoencodeurs, en particulier les modèles variantes (VAE) et les autoencodeurs débruitants (DAE), ont gagné en popularité pour leur capacité à traiter des données sparses et à modéliser des préférences utilisateur complexes. Liang et al. (2018) ont introduit un modèle d'autoencodeur variationnel pour le filtrage collaboratif (MultVAE), montrant des améliorations significatives par rapport aux méthodes traditionnelles. De même, Sedhain et al. (2015) ont exploré l'utilisation des DAE dans les systèmes de recommandation, démontrant leur efficacité face aux données bruitées et incomplètes.

2.2. Filtrage Collaboratif et Techniques Avancées

Le filtrage collaboratif reste une technique dominante dans les systèmes de recommandation. Hu et al. (2008) ont proposé une approche basée sur les moindres carrés alternés (ALS) pour les données implicites, une méthode aujourd'hui largement utilisée dans les recommandations musicales. Par ailleurs, les travaux de Koren et al. (2009) sur la décomposition en valeurs singulières (SVD) et son application dans le système de recommandation de Netflix ont jeté les bases des méthodes de factorisation matricielle modernes.

2.3. Traitement des Données Sparses

Les ensembles de données musicales sont souvent caractérisés par leur caractère sparse. Rendle et al. (2019) ont introduit une méthode de factorisation tensorielle pour les données sparses, offrant un cadre plus flexible et efficace pour les recommandations musicales. De même, l'approche EASE (Embarrassingly Shallow Autoencoders) proposée par Steck (2019) a montré des performances prometteuses pour les données sparses, grâce à sa structure simple mais puissante.

2.4. Évolutions et Perspectives Futures

L'intégration de modèles basés sur le deep learning et les réseaux de neurones a marqué une évolution significative dans les systèmes de recommandation. Oord et al. (2013) ont exploré l'utilisation de réseaux de neurones convolutifs pour la recommandation de musique, ouvrant la voie à des méthodes plus complexes et personnalisées. Les recherches actuelles se concentrent sur l'amélioration de la personnalisation et de l'interprétabilité des recommandations, ainsi que sur l'intégration de modèles hybrides combinant plusieurs techniques pour optimiser la précision et la pertinence des recommandations.

3. MÉTHODOLOGIE

Après avoir introduit notre approche, nous détaillons maintenant la méthodologie adoptée pour mener à bien notre étude. Notre étude s'appuie sur le Million Song Dataset pour développer et évaluer des systèmes de recommandation musicaux. Nous avons adopté une approche en plusieurs étapes,

comprenant la construction d'une matrice utilisateur-item et l'application de diverses techniques de recommandation, telles que SVD, ALS et EASE. Chaque modèle a été formé et évalué en fonction de sa précision dans la recommandation de musique.

3.1. Prétraitement des Données

Nous avons commencé par normaliser la durée des chansons et leur indice de popularité ('hotness') pour garantir une comparaison juste entre les morceaux. Les valeurs manquantes ont été traitées avec soin pour maintenir l'intégrité de l'ensemble de données.

3.2. Sélection des Données Utilisateur Réelles

Contrairement à une approche basée sur des données synthétiques, nous avons choisi d'utiliser des données utilisateur réelles issues d'un dataset complémentaire. Cette démarche a permis d'incorporer des comportements d'écoute authentiques, essentiels pour la validité de notre étude.

Pour ce faire, nous avons d'abord identifié les utilisateurs ayant écouté les chansons de notre sous-ensemble de base, composé de 10 000 morceaux. Ensuite, nous avons filtré ces utilisateurs pour ne conserver que ceux ayant écouté au moins cinq chansons, optimisant ainsi le volume des données tout en ciblant des profils d'utilisateurs pertinents.

3.3. Transformation des Données d'Interactions

Les données d'interaction utilisateur-chanson ont été transformées en une matrice utilisateur-item. Cette matrice essentielle pour nos modèles a été élaborée en associant à chaque utilisateur des ensembles de chansons, reflétant ainsi les tendances d'écoute observées.

3.4. Construction de la Matrice de Cooccurrence des Items

Nous avons ensuite élaboré une matrice de cooccurrence des items à partir de la matrice utilisateur-item. Cette étape cruciale a permis de saisir les relations et les motifs d'écoute communs entre les chansons, offrant une perspective riche sur les préférences des utilisateurs.

3.5. Résolution de la Régression Ridge pour EASE

Pour le modèle EASE, une régression ridge a été appliquée. Cette méthode a permis de trouver une solution optimale qui favorise la parcimonie dans les poids attribués aux items, tout en garantissant d'excellentes performances de recommandation.

3.6. Génération des Recommandations

Enfin, nous avons exploité les matrices de poids dérivées des modèles pour générer des recommandations personnalisées. Nous avons pris soin d'exclure les chansons déjà écoutées par les utilisateurs, afin de simuler un scénario de recommandation réaliste et pratique.

Ayant établi notre cadre méthodologique, nous passons à la présentation détaillée des modèles de recommandation utilisés.

4. DÉFINITIONS DES MODÈLES

Cette section fournit un aperçu des modèles de recommandation utilisés dans notre étude, en expliquant leur fonctionnement et en présentant les formules mathématiques clés.

4.1. Autoencodeurs Superficiels pour Données Sparses (EASE)

EASE a été sélectionné pour sa simplicité et son efficacité à traiter des ensembles de données sparses, permettant une interprétabilité accrue. EASE représente une simplification des autoencodeurs classiques, axée sur une structure linéaire sans couches cachées. Ce modèle est particulièrement efficace pour traiter les données sparses, telles que nos données musicales, en mettant l'accent sur la découverte des liens entre les items plutôt que sur la reconstruction de l'entrée. La formule d'EASE est :

$$B = (X^T X + \lambda I)^{-1} X^T X$$

où X est la matrice utilisateur-item, λ le paramètre de régularisation, et I la matrice identité.

4.2. Singular Value Decomposition (SVD)

La SVD a été choisie pour sa capacité à décomposer les grandes matrices et à révéler des relations latentes, essentielle dans notre contexte de données sparses. La SVD est une méthode de factorisation matricielle qui nous a permis de décomposer la matrice utilisateur-item en ses composantes latentes, révélant ainsi des structures cachées et facilitant la prédiction des préférences utilisateur. La SVD est décrite par :

$$M = U \Sigma V^T$$

où M est la matrice utilisateur-item, U et V représentent les matrices orthogonales des utilisateurs et des items, et Σ est la matrice des valeurs singulières.

4.3. Alternating Least Squares (ALS)

L'ALS a été utilisé pour sa pertinence dans le traitement des données implicites, typiques des systèmes de recommandation musicale. L'ALS est une technique de factorisation pour

les données implicites, utilisée pour optimiser alternativement les matrices des facteurs latents des utilisateurs et des items. Cette approche est particulièrement adaptée pour les ensembles de données de grande taille. Les équations d'ALS sont :

$$U = (VV^T + \lambda I)^{-1}VM^T$$

$$V = (UU^T + \lambda I)^{-1}UM$$

où U et V représentent respectivement les facteurs latents des utilisateurs et des items.

4.4. Variational Autoencoders for Collaborative Filtering (MultVAE)

MultVAE a été préféré pour sa capacité à modéliser des distributions complexes de préférences utilisateur, une caractéristique cruciale pour la personnalisation des recommandations. MultVAE utilise un modèle probabiliste pour capturer la distribution des préférences utilisateur. En incorporant un mécanisme de génération, il offre des recommandations variées et personnalisées. Les formules clés de MultVAE sont :

$$q_\phi(z|u) = \mathcal{N}(z|\mu_\phi(u), \Sigma_\phi(u))$$

$$p_\theta(u|z) = \prod_i p_\theta(u_i|z)$$

où q_ϕ est l'encodeur, p_θ le décodeur, z les facteurs latents, et u les interactions utilisateur.

4.5. Collaborative Denoising Auto-Encoders (CDAE)

Le CDAE a été adopté pour sa robustesse face aux données bruitées, renforçant ainsi la fiabilité des recommandations. Le CDAE apprend des représentations résistantes au bruit des préférences utilisateur. Il intègre un mécanisme de denoising pour améliorer la généralisation. Le fonctionnement du CDAE est exprimé par :

$$y = f(Wx + W_u u + b)$$

$$r = g(Vy + c)$$

où x est l'entrée bruitée, u l'identifiant de l'utilisateur, W , W_u , V les matrices de poids, et f , g les fonctions d'activation.

Après avoir défini les modèles, nous expliquons maintenant comment ces derniers ont été entraînés pour garantir des recommandations précises.

5. ENTRAÎNEMENT DES MODÈLES

L'entraînement efficace de nos modèles est crucial pour garantir la précision et la pertinence des recommandations musicales proposées.

5.1. Entraînement d'EASE

Pour EASE, nous avons privilégié une fonction de perte orientée vers la maximisation des similitudes entre les items, tout en intégrant un paramètre de régularisation pour maintenir la 'sparsité'.

5.2. Entraînement de SVD

L'entraînement de la SVD correspond au calcul de la décomposition en valeurs singulières de notre large matrice utilisateur-item, avec un nombre arbitraire de valeurs singulières à retenir.

5.3. Entraînement d'ALS

Pour ALS, le défi majeur résidait dans l'optimisation alternée des facteurs des utilisateurs et des items, nécessitant une attention particulière à la convergence et à l'évitement des minima locaux.

5.4. Entraînement de MultVAE

Pour MultVAE, l'accent a été mis sur la maximisation de la log-vraisemblance des interactions utilisateur-item, ce qui requiert une compréhension détaillée des préférences et comportements des utilisateurs. La courbe de perte de l'entraînement de MultVAE illustre une diminution rapide de l'erreur, indiquant que le modèle a rapidement appris les structures latentes pertinentes des données. Ce type de convergence est typique des modèles qui parviennent à capturer efficacement les relations complexes dans les données. La figure ci-dessous illustre l'évolution de la fonction de perte pendant l'entraînement de MultVAE, montrant un apprentissage efficace dès les premières itérations.

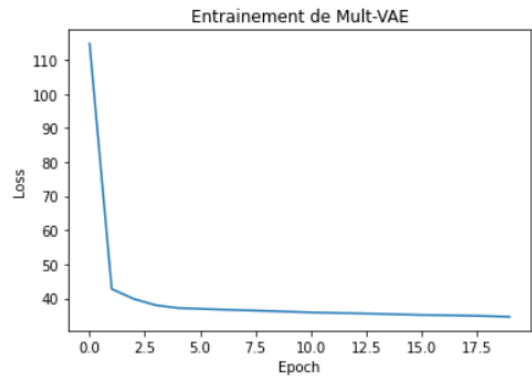


Fig. 1. Courbe de perte lors de l'entraînement de MultVAE. La décroissance rapide indique un modèle apprenant efficacement les caractéristiques pertinentes des données.

5.5. Entraînement de CDAE

Pour CDAE, son processus d'entraînement a inclus un mécanisme de denoising qui a pour objectif d'améliorer la généralisation du modèle. La courbe de perte montre une tendance similaire avec une diminution significative de l'erreur dès les premières époques, suggérant que l'introduction de bruit a permis au modèle de se focaliser sur les caractéristiques les plus saillantes des données, améliorant ainsi sa capacité à prédire les préférences des utilisateurs. La figure suivante dépeint la courbe de perte pour le CDAE, révélant une capacité de généralisation améliorée grâce au denoising intégré au processus d'entraînement.

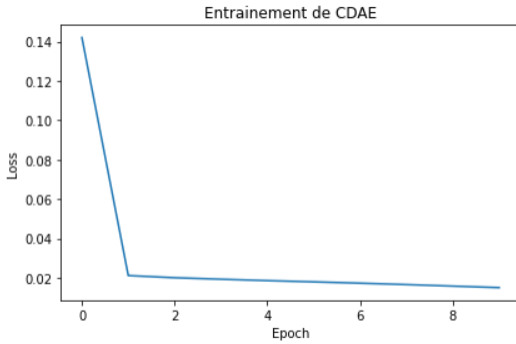


Fig. 2. Courbe de perte pour le CDAE montrant une forte réduction de l'erreur, ce qui témoigne d'un apprentissage robuste malgré le bruit introduit.

6. ÉVALUATION DES MODÈLES

Notre évaluation s'est concentrée sur la mesure de la pertinence des recommandations musicales à l'aide du score NDCG. Cette métrique nous a permis d'évaluer comment chaque modèle classait les chansons en fonction de leur pertinence pour les utilisateurs.

6.1. Évaluation de SVD, ALS et EASE

L'évaluation de SVD, ALS et EASE a révélé des nuances intéressantes dans leur capacité à traiter différentes dimensions des données musicales. Le score NDCG a mis en évidence la supériorité de SVD en termes de précision, tandis qu'ALS et EASE ont montré des performances compétitives malgré des approches distinctes en termes de complexité algorithmique et de traitement des données sparses.

6.2. Évaluation de MultVAE et CDAE

Les scores NDCG pour MultVAE et CDAE ont mis en exergue leur efficacité à modéliser des distributions complexes de préférences d'écoute. Leur capacité à intégrer des nuances

subtiles dans les recommandations musicales a été particulièrement remarquable, reflétant le potentiel de ces modèles dans l'environnement riche et varié de la musique.

L'entraînement des modèles étant achevé, nous évaluons à présent leur performance à travers le score NDCG.

7. RÉSULTATS

Les résultats obtenus soulignent la robustesse des différentes approches de recommandation musicale. Bien que la SVD ait démontré une supériorité en termes de précision, les modèles basés sur les autoencodeurs, tels que MultVAE et CDAE, ont excellé dans la personnalisation des recommandations.

Table 1. Scores NDCG pour les modèles de recommandation

Méthode	NDCG Moyen (en %)
SVD	96.308
ALS	95.363
EASE	95.669
MultVAE	95.535
CDAE	95.5819

8. CONCLUSION

Notre étude a révélé une tendance intéressante parmi les modèles de recommandation de musique : bien que variés dans leurs approches, ils produisent des résultats étonnamment similaires en termes de scores NDCG. Cette uniformité soulève des questions sur l'efficacité de ces modèles à refléter pleinement la diversité et la richesse des préférences musicales des utilisateurs. Elle suggère que nos modèles pourraient être optimisés spécifiquement pour ce dataset, ou que les caractéristiques intrinsèques des données musicales favorisent une certaine uniformité dans les performances.

Cette convergence des scores NDCG nous pousse à remettre en question non seulement la diversité des recommandations générées, mais aussi la capacité de la métrique NDCG à saisir l'intégralité de la qualité des systèmes de recommandation dans un contexte musical. Elle indique la nécessité d'une évaluation plus diversifiée, employant une gamme de métriques pour appréhender pleinement les performances de ces modèles. Cela ouvre la voie à de futures explorations des métriques de recommandation, pour une compréhension plus nuancée des performances des modèles dans le domaine complexe de la musique.

En conclusion, notre étude offre une perspective précieuse sur les performances de divers systèmes de recommandation musicaux. Toutefois, la limitation majeure réside dans le choix du dataset et la dépendance à la métrique NDCG. Des recherches futures pourraient bénéficier de l'exploration de

datasets plus complets et diversifiés, avec des données utilisateur plus riches. Ceci permettrait l'application d'autres métriques et une compréhension plus approfondie des nuances des recommandations musicales, contribuant ainsi à une évaluation plus complète et détaillée des systèmes de recommandation dans le domaine de la musique.

9. RÉFÉRENCES

1. Sumanth, R. (Date non spécifiée). "Understanding NDCG as a Metric for Your Recommendation System." Medium. Disponible sur : <https://medium.com/@readsumant/understanding-ndcg-as-a-metric-for-your-recomendation-system-5cd012fb3397>
2. Liang, D., Krishnan, R. G., Hoffman, M. D., & Jbara, T. (2019). "Variational Autoencoders for Collaborative Filtering." arXiv preprint arXiv:1802.05814. Disponible sur : <https://arxiv.org/pdf/1802.05814.pdf>
3. "The Million Song Dataset." Disponible sur : <http://millionsongdataset.com/pages/getting-dataset/>
4. The Echo Nest. "Taste Profile Subset." Million Song Dataset. Disponible sur : <http://millionsongdataset.com/tasteprofile/#getting>
5. Steck, H. (2019). "Embarrassingly Shallow Autoencoders for Sparse Data." arXiv preprint arXiv:1905.03375. Disponible sur : <https://arxiv.org/pdf/1905.03375.pdf>