

Mini-Projet – Systèmes de recommandation sociale

1. Contexte et ressources du mini-projet

1.1 Objectifs

Les objectifs de ce mini-projet sont :

- **Comprendre**, analyser un modèle de recommandation **à partir d'un article scientifique** (modèle vu en cours)
- **Comprendre en vue de réutiliser un code de modèle** existant comme brique de base.
- **Proposer une extension** de modèle sur la base de : 1) votre compréhension du cours, 2) des notions de groupes développées en cours.
- **Expérimenter un modèle** de système de recommandation, interpréter et analyser les résultats.

1.2. Organisation et rendus

Le mini-projet peut être réalisé en groupes de 2-3 (4 au max) étudiants.

Un rapport est restitué sur Moodle, qui comprend :

- Le code commenté (Github, collab, comme pour les TP).
- Un document écrit qui restitue : 1) la description de l'extension proposée du modèle, faire le lien avec les parties décrites du modèle dans l'article scientifique ; 2) La description de la démarche expérimentale (objectifs et méthodologie), résultats, analyse et comparaison des résultats avec les résultats présentés dans l'article scientifique.
- **Bonus** : toutes les analyses pertinentes et justifiées (comme ferait un data scientist) permettant d'étayer vos écrits, sont les bienvenues.

Un dépôt Moodle est prévu pour le projet avec une **date limite : 30/01/2024**

1.3 Ressources

- **Article scientifique (disponible sur moodle) :**

[1] Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu, and Irwin King. 2011. **Recommender systems with social regularization**. In Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11). pp 287–296.

- **Collections de données à utiliser :**

Utiliser 1 parmi 2 collections utilisées dans l'article (pour comparaison) et 1 autre collection vous donnant des possibilités d'utiliser éventuellement d'autres caractéristiques des utilisateurs pour la recommandation sociale :

- Epinions: <http://www.trustlet.org/epinions.html>
- Delicious: <https://grouplens.org/datasets/hetrec-2011/>

- Code du modèle SocREc et des métriques

<https://github.com/hongleizhang/RSAgorithms>

2. Travail demandé

2.1. Reproduction du modèle de recommandation sociale SocRec

Ce projet s'appuie sur le modèle SocRec publié dans l'article [1], voir références ci-bas :

- Lire et comprendre le modèle de recommandation en faisant le lien avec l'article scientifique
- Faire une analyse descriptive des réseaux sociaux Epinions et Delicious : taille, densité, loi de distribution des données, degré de sparsité des matrices d'interactions etc. Résumer vos résultats dans des tableaux les statistiques calculées.
- Lancer les évaluations expérimentales du modèle SocRec avec le code fourni avec les 2 jeux de données Epinions et Delicious.
- Comparer les résultats que vous avez obtenus avec les résultats obtenus par les auteurs sur le dataset Epinions. Reproduisez-vous ces résultats ? Vos résultats sur le jeu de données Delicious, suivent-ils les mêmes tendances que ceux sur le jeu de données Epinions ? Expliquer les différences (s'il y en a) en vous appuyant sur l'analyse descriptive des 2 réseaux sociaux ci-avant, autant que possible.

2.2. Extension du modèle

Dans l'article SocRec [1], une mesure de similarité entre utilisateurs est proposée dans la Section 5.3, équations (13) et (14).

- Proposer 2 autres mesures de similarité entre utilisateurs, en remplacement de celles proposées en équations (13) et (14) :
 - 1 mesure sera basée sur les mesures de similarités de nœuds dans graphes vues dans le chapitre 1.

- 1 mesure sera basée sur les notions développées par les groupes d'étudiants (ex. *nodetovec*). Donner une formalisation.
- Coder les mesures proposées (au minima celle basées sur les notions de groupes d'étudiants) en les injectant dans le code fourni de *SocRec*, donnant lieu à votre version *SocRec_nomgroupe*.

Remarque : l'extension proposée peut couvrir à la fois la représentation des utilisateurs et leurs mesures de similarités comme dans le cas des GNN, knowledge graphs, etc.

2.3. Expérimentation et analyse des résultats

- Refaire les évaluations faites en 2.1 en utilisant les mesures de similarités utilisateurs que vous avez proposées (code *SocRec_nomgroupe*).
- Interpréter et comparer les résultats obtenus avec *Soc-Rec* et *SocRec_nomgroupe* dans les mêmes conditions d'expérimentation (même jeu de données Epinions, mêmes paramètres du modèle etc.).
- Interpréter et comparer les résultats obtenus de *SocRec* et *SocRec_nomgroupe* sur le jeu de données Delicious ?
- Faire un bilan des expérimentations : auto-évaluer l'extension que vous avez proposée en indiquant les apports, les limites, les possibilités d'amélioration.