

Sujet : Revisiting Graph Homophily Measures

Semaine II

Mehdi Naima

Description générale du problème

L'homophilie est une propriété des graphes qui décrit la tendance des arêtes à relier des nœuds similaires. Il existe plusieurs mesures utilisées pour évaluer l'homophilie, mais toutes présentent certains inconvénients : en particulier, elles ne peuvent pas être utilisées de manière fiable pour comparer des ensembles de données dont le nombre de classes et l'équilibre entre les tailles de classes varient. Pour le démontrer, des travaux antérieurs sur l'homophilie des graphes ont suggéré plusieurs propriétés souhaitables pour une bonne mesure de l'homophilie, tout en notant qu'aucune mesure existante de l'homophilie ne possède toutes ces propriétés.

Un article récent tente de remédier à ce problème. Nous allons implémenter les mesures présentées dedans et reproduire une partie des résultats.

Description détaillée

- Implémenter les mesures d'homophilie présentées à la Section 2. de [1].
- Tester les résultats d'incohérences de la Section 3 sur deux nouveaux datasets et discuter les résultats obtenus à la lumière de ceux de l'article.
- Implémenter la mesure d'homophilie non-biaisées présentée dans l'article.
- Tester les différentes mesures sur quelques nouveaux datasets. Les résultats sont-ils similaires à ceux du Tableau 5 ?

Évaluation du projet

Attendus de base.

- Une implémentation complète et un rapport détaillant votre travail.
- Le code peut être rédigé dans n'importe quel langage, mais un `makefile` (si c'est un langage compilé) doit permettre la compilation sur les machines de la PPTI.
- Les commandes lançant les différentes simulations doivent être groupées dans un fichier bash. Ainsi l'exécution du fichier bash doit pouvoir lancer toutes les simulations.
- Une documentation doit être fournie expliquant les options des différentes commandes.

Un projet ne respectant pas ces attendus de base n'aura pas la moyenne.

Attendus avancés. Discuter de quelques mesures d'homophilie dans des hypergraphes et voir si les propriétés 1 à 6 sont respectées.

Références

- [1] Mikhail Mironov and Liudmila Prokhorenkova. Revisiting graph homophily measures. *arXiv preprint 2412.09663*, 2024.