Graph Neural Networks for Recommendation: Reproducibility, Graph Topology, and Node Representation

Daniele Malitesta
Polytechnic University of Bari, Italy
daniele.malitesta@poliba.it

Claudio Pomo
Polytechnic University of Bari, Italy
claudio.pomo@poliba.it

Tommaso Di Noia
Polytechnic University of Bari, Italy
tommaso.dinoia@poliba.it

Abstract

Graph neural networks (GNNs) have gained prominence in recommendation systems in recent years. By representing the user-item matrix as a bipartite and undirected graph, GNNs have demonstrated their potential to capture short- and long-distance user-item interactions, thereby learning more accurate preference patterns than traditional recommendation approaches. In contrast to previous tutorials on the same topic, this tutorial aims to present and examine three key aspects that characterize GNNs for recommendation: (i) the reproducibility of state-of-the-art approaches, (ii) the potential impact of graph topological characteristics on the performance of these models, and (iii) strategies for learning node representations when training features from scratch or utilizing pre-trained embeddings as additional item information (e.g., multimodal features). The goal is to provide three novel theoretical and practical perspectives on the field, currently subject to debate in graph learning but long been overlooked in the context of recommendation systems.

1 Learning objectives

With the current tutorial, we plan to cover both theoretical and practical aspects in GNNs-based recommendation. First, we investigate the current challenges in experimentally-reproducing the results of state-of-the-art approaches, and compare their performance to other shallower models for recommendation (with unexpected outcomes) [1, 2]. For the experimental study, we adopt Elliot [3], our framework for the rigorous reproducibility and evaluation of recommender systems. Indeed, such an investigation paves the way to understanding whether, how, and why topological graph properties (conceptually related to node degree) may influence the performance of GNNs-based recommendation systems [4]. Finally, the tutorial delves into the representation strategies for node embeddings [5, 6]; while most of the techniques learn node representation from scratch, our focus goes to the adoption of pre-trained item's side information (such as multimodal ones [7]) by analysing the main experimental implications of such a design choice [8].

2 Tutorial schedule

Total tutorial duration → *180 minutes*

Introduction and background (Tommaso Di Noia) → *20 minutes*

- Introduction and motivations of the tutorial → *5 minutes*
- Basics concepts of recommender systems & GNNs-based recommendation → *15 minutes*

Reproducibility → *60 minutes*

- **[Hands-on]** Implementation of GNNs-based recsys in Elliot with PyG and reproducibility issues (Daniele Malitesta) → *25 minutes*
- **[Hands-on]** Reproducing the results of state-of-the-art GNNs-based recommender systems (Claudio Pomo) → *25 minutes*
- Performance comparison of GNNs-based approaches to traditional recommendation systems (Claudio Pomo) → *10 minutes*

Break and Q&A → *15 minutes*

Graph topology → *30 minutes*

- Concepts and formulations of graph topological properties of the user-item graph (Tommaso Di Noia) → *15 minutes*
- Impact of topological graph properties on the performance of GNNs-based recommender systems (Daniele Malitesta) → *15 minutes*

Node representation → *45 minutes*

- Design choices to train node embeddings from scratch (Claudio Pomo) → *20 minutes*
- **[Hands-on]** Leveraging item's side-information (e.g., multimodal features) to represent node embeddings (Daniele Malitesta) → *25 minutes*

Closing remarks and Q&A → *10 minutes*

3 Relevance to LoG

The tutorial covers highly-related topics to the ones of the LoG conference and community, namely, the application of graph neural networks (GNNs) to the task of personalized recommendation. In particular, and differently from previous similar tutorials presented at other venues (see later), the main focus of this tutorial is on the reproducibility of state-of-the-art recommendation systems leveraging GNNs, the possible impact of graph topological characteristics on the performance of such approaches, and the modeling of node features (trained from scratch or freezed to use additional side information). Since the outlined three aspects have been widely debated in the graph representation learning literature over the last few years, we believe the LoG conference could benefit from an analysis of these aspects also in the field of recommender systems.

4 Previous related tutorials

Tutorials on graph-based recommendation. Previous tutorials on graph-based recommendation [9–12] are reported in Table 1, along with the reference, website, slides, and video recording.

Differences with previous tutorials. The majority of previous tutorials on graph-based recommendation address the topic of GNNs in recommendation from a general perspective. Conversely, our tutorial intends to approach the same topic from three main research aspects which are highly popular in the graph learning field but have not been previously investigated in graph-based recommendation, namely: reproducibility issues, the influence of graph topology on model's performance, and the modeling of node features. Indeed, the tutorial by Purificato et al. [12] is the closest to ours in the intention of providing a more specific analysis of GNNs-based recommendation, by considering user modeling and beyond-accuracy evaluation; however, as already highlighted, our investigation and contributions are different.

Other related tutorials. Scalable Graph Neural Networks with Deep Graph Library (KDD 2020); Deep Graph Learning: Foundations, Advances and Applications (KDD 2020); Learning Graph Neural Networks with Deep Graph Library (The Web Conf 2020); Graph Representation Learning: Foundations, Methods, Applications and Systems (KDD 2021); Advanced Deep Graph Learning: Deeper, Faster, Robuster, Unsupervised (The Web Conf 2021); Learning from Graphs: From Mathematical Principles to Practical Tools (The Web Conf 2021); Scalable Graph Neural Networks with

Table 1: Tutorials about graph-based recommendation from 2020-2023.

Title	Venue	Website	Slides	Video
Learning and Reasoning on Graph for Recommendation [9]	WSDM 2020	link	link	✗
Graph-based Representation Learning for Web-scale Recommender Systems [10]	KDD 2022	link	link	✗
Graph Neural Networks for Recommender System [11]	WSDM 2022	link	link	link
Tutorial on User Profiling with Graph Neural Networks and Related Beyond-Accuracy Perspectives [12]	UMAP 2023	link	link	✗

Deep Graph Library (WSDM 2021); Frontiers of Graph Neural Networks with DIG (KDD 2022); Graph Neural Networks: Foundation, Frontiers and Applications (KDD 2022); Graph Neural Networks: Foundation, Frontiers and Applications (The Web Conf 2023); Self-supervised Learning and Pre-training on Graphs (The Web Conf 2023); When Sparse Meets Dense: Learning Advanced Graph Neural Networks with DGL-Sparse Package (The Web Conf 2023).

5 Tutorial speakers

Daniele Malitesta is a PhD candidate at the Polytechnic University of Bari (Italy). During his research career so far, he has been studying and developing recommendation algorithms leveraging side information, with a specific focus on graph- and multimodal-based recommender systems. He has published at top-tier conferences, such as SIGIR, ECIR, RecSys, and MM, and has served as a reviewer at SIGIR 2023, RecSys 2023 (outstanding reviewer), NeurIPS 2023, LoG 2022-2023, ICLR 2024, and ECIR 2024. He is one of the organizers of the First International Workshop on Graph-Based Approaches in Information Retrieval (IRonGraphs), co-located with ECIR 2024. He is among the developers of Elliot, a framework for the rigorous evaluation and reproducibility of recommender systems, where he contributed with the implementation of more than 15 algorithms from the state-of-the-art. Recently, he has visited Dr. Pasquale Minervini at the University of Edinburgh as part of the internship period of his PhD.

Claudio Pomo is a research fellow at the Polytechnic University of Bari (Italy). He graduated with a doctorate in computer engineering. His research concerns the issues of responsible AI for personalization, with a particular interest in results' reproducibility and multi-objective performance evaluation. Contributions on these topics have been accepted in influential area conferences (SIGIR, RecSys, ECIR, UMAP) and journals (Information Science, IPM). Claudio is among the authors and contributors of Elliot.

Tommaso Di Noia is Professor of Computer Science at the Polytechnic University of Bari (Italy). His research activities focus on AI and Data Management. They were initially devoted to knowledge representation and automated reasoning. Then, he studied how to apply knowledge representation techniques to automated negotiations. Following these ideas, he has devoted his interest to applying knowledge graphs and Linked Open Data to RSs with papers published in international journals, conferences, and book chapters. During the last years, he moved his research into the Trustworthy AI topic with a particular interest in adversarial ML, explainability, fairness and privacy protection of RSs. He is serving as program co-chair at RecSys 2023 and general co-chair at RecSys 2024.

6 Intended audience and level

The proposed tutorial deals with both intermediate and advanced theoretical/practical topics spanning recommendation, graph representation learning, reproducibility in machine learning, graph topology, and multimodal learning. For these reasons, the tutorial might ideally be of interest to a wide range of researchers and practitioners working on (even a subset of) such aspects. Programming knowledge of Python and PyTorch (with a specific focus on PyTorch Geometric) would be a good-to-have skill,

even though the tutorial will guide the attendees step-by-step, especially during the hands-on sessions. Given the virtual nature of the tutorial, we do not set a strong requirement for the maximum number of participants (ideally up to 200).

References

[1] Vito Walter Anelli, Daniele Malitesta, Claudio Pomo, Alejandro Bellogín, Eugenio Di Sciascio, and Tommaso Di Noia. Challenging the myth of graph collaborative filtering: a reasoned and reproducibility-driven analysis. In *RecSys*, pages 350–361. ACM, 2023. 1

[2] Daniele Malitesta, Claudio Pomo, Vito Walter Anelli, Tommaso Di Noia, and Antonio Ferrara. An out-of-the-box application for reproducible graph collaborative filtering extending the elliot framework. In *UMAP (Adjunct Publication)*, pages 12–15. ACM, 2023. 1

[3] Vito Walter Anelli, Alejandro Bellogín, Antonio Ferrara, Daniele Malitesta, Felice Antonio Merra, Claudio Pomo, Francesco Maria Donini, and Tommaso Di Noia. Elliot: A comprehensive and rigorous framework for reproducible recommender systems evaluation. In *SIGIR*, pages 2405–2414. ACM, 2021. 1

[4] Daniele Malitesta, Claudio Pomo, Vito Walter Anelli, Alberto Carlo Maria Mancino, Eugenio Di Sciascio, and Tommaso Di Noia. A topology-aware analysis of graph collaborative filtering. *CoRR*, abs/2308.10778, 2023. 1

[5] Vito Walter Anelli, Yashar Deldjoo, Tommaso Di Noia, Daniele Malitesta, Vincenzo Paparella, and Claudio Pomo. Auditing consumer- and producer-fairness in graph collaborative filtering. In *ECIR (1)*, volume 13980 of *Lecture Notes in Computer Science*, pages 33–48. Springer, 2023. 1

[6] Vito Walter Anelli, Yashar Deldjoo, Tommaso Di Noia, Eugenio Di Sciascio, Antonio Ferrara, Daniele Malitesta, and Claudio Pomo. How neighborhood exploration influences novelty and diversity in graph collaborative filtering. In *MORS@RecSys*, volume 3268 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2022. 1

[7] Daniele Malitesta, Giandomenico Cornacchia, Claudio Pomo, Felice Antonio Merra, Tommaso Di Noia, and Eugenio Di Sciascio. Formalizing multimedia recommendation through multimodal deep learning. *CoRR*, abs/2309.05273, 2023. 1

[8] Daniele Malitesta, Giandomenico Cornacchia, Claudio Pomo, and Tommaso Di Noia. On popularity bias of multimodal-aware recommender systems: a modalities-driven analysis. *CoRR*, abs/2308.12911, 2023. 1

[9] Xiang Wang, Xiangnan He, and Tat-Seng Chua. Learning and reasoning on graph for recommendation. In *WSDM*, pages 890–893. ACM, 2020. 2, 3

[10] Ahmed El-Kishky, Michael M. Bronstein, Ying Xiao, and Aria Haghighi. Graph-based representation learning for web-scale recommender systems. In *KDD*, pages 4784–4785. ACM, 2022. 3

[11] Chen Gao, Xiang Wang, Xiangnan He, and Yong Li. Graph neural networks for recommender system. In *WSDM*, pages 1623–1625. ACM, 2022. 3

[12] Erasmo Purificato, Ludovico Boratto, and Ernesto William De Luca. Tutorial on user profiling with graph neural networks and related beyond-accuracy perspectives. In *UMAP*, pages 309–312. ACM, 2023. 2, 3