ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

GNN Pretraining

By design Graph Neural Networks (GNNs) are well-suited for structured data including social networks, knowledge graphs and molecules. They are typically trained in a supervised manner, unfortunately for some domains like molecules task-specific labels are scarce as obtaining them is both time-consuming and expensive. Furthermore, GNNs in this domain also require good out-of-distribution generalization. Given these challenges and the success of pre-training in computer vision and natural language processing (NLP), there has been a lot of effort over the past few years to develop pre-training approaches for GNNs in the hope of similar results. Hu et al. [1] combines self-supervised node-level and supervised graph-level pre-training, which can be considered one of the earliest works to successfully transfer knowledge onto a downstream task. Many other approaches have been proposed since then, including graph-autoregressive-models like GPT-GNN [2], graph contrastive learning like GraphCL [3], L2P-GNN [4] which aims to learn to pre-train and thus can be considered meta-learning, as well as data-active pre-training [5] to select the most informative pre-training input out of a cross-domain dataset. Also well known concepts from NLP have been adapted to GNNs, MoleBert [6] adapts tokenization to molecular graphs to increase the vocabulary size for a harder and more informative pre-training task and GPPT [7] adapts prompting to graph data in order to narrow the training objective gap between pre-training task and downstream task.
Despite the variety of proposed approaches, we are still far away from a general purpose pre-training framework for GNNs that enables successful and meaningful knowledge transfer to downstream tasks independent of their domain and type, leading to a so-called foundation GNN model. In this thesis we therefore want to make steps in this direction.

Contact

- Béni Egressy: begressy@ethz.ch, ETZ G94

Literature

[1] Weihua Hu et al. *Strategies for Pre-training Graph Neural Networks*. 2020. arXiv: 1905.12265 [cs.LG].

[2] Ziniu Hu et al. *GPT-GNN: Generative Pre-Training of Graph Neural Networks*. 2020. arXiv: 2006.15437 [cs.LG].

[3] Yuning You et al. *Graph Contrastive Learning with Augmentations*. 2021. arXiv: 2010.13902 [cs.LG].

[4] Yuanfu Lu et al. "Learning to Pre-train Graph Neural Networks". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 35.5 (May 2021), pp. 4276–4284.

[5] Jiarong Xu et al. *Better with Less: Data-Active Pre-training of Graph Neural Networks*. 2023.

[6] Jun Xia et al. “Mole-BERT: Rethinking Pre-training Graph Neural Networks for Molecules”. In: *The Eleventh International Conference on Learning Representations*. 2023.

[7] Mingchen Sun et al. “GPPT: Graph Pre-Training and Prompt Tuning to Generalize Graph Neural Networks”. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. KDD ’22. Washington DC, USA: Association for Computing Machinery, 2022, pp. 1717–1727. ISBN: 9781450393850. DOI: 10.1145/3534678.3539249.