Data-Efficient Graph Learning

Kaize Ding

Northwestern University
kaize.ding@northwestern.edu

My research strives to develop fundamental graph-centric learning algorithms to reduce the need for human supervision in low-resource scenarios. The focus is on achieving effective and reliable data-efficient learning on graphs, which can be summarized into three facets:

- **Graph Weakly-supervised Learning:** Existing Graph ML algorithms are mainly developed for the supervised or semi-supervised setting. However, weakly-supervised Graph ML, especially when dealing with *incomplete, inaccurate, and indirect* supervision signals, remains understudied. To address this, I proposed a series of works (Ding, Li, and Liu 2019; Ding et al. 2022, 2024, 2023a) that focus on learning with the aforementioned three types of weak supervision. For example, I developed a novel graph self-training framework – Meta Propagation Networks (Ding et al. 2022), which is able to adjust its label propagation strategy and leverage large receptive fields for inferring accurate pseudo labels on unlabeled nodes. By augmenting the scarce training data with these pseudo labels, one can learn a label-efficient GNN with only a few labels per class.
- **Graph Few-shot Learning:** Despite that humans are capable of learning new tasks rapidly by utilizing what they learned in the past, current AI algorithms cannot rapidly generalize from a few examples. To bridge this gap between Graph ML models and humans, I have developed a series of graph few-shot learning algorithms (Ding et al. 2020; Wang et al. 2022; Xu et al. 2022). For example, to handle the *never-before-seen* node classes, I proposed Graph Prototypical Networks (GPN) (Ding et al. 2020) to extrapolate the meta-knowledge from many-shot seen node classes to few-shot unseen node classes using graph meta-learning. By leveraging the graph property information, GPN supports estimating the informativeness of each labeled node and derives highly robust and representative class prototypes for classifying those *never-before-seen* node classes with even a few labeled nodes.
- **Graph Self-supervised Learning:** My research also investigates graph self-supervised learning to alleviate the demand for massive labeled data and provide sufficient supervision, which can be categorized into two lines: (1) *graph contrastive learning*, including (Ding et al. 2023b, 2024; Liu* et al. 2023). For example, I designed a simple yet effective graph contrastive learning model S^3-CL (Ding et al. 2023b) to elicit global structural and semantic knowledge from the input graph, which can address the "global unawareness" of existing graph contrastive learning methods; and (2) *graph generative modeling*, such as (Ding et al. 2019). Utilizing input graph reconstruction as a self-supervision pretext, I developed the first graph generative model Dominant (Ding et al. 2019) for graph anomaly detection. It can conduct both structure and feature reconstruction to capture the patterns of the majority nodes in a self-supervised fashion.

References

Ding, K.; Li, J.; Bhanushali, R.; and Liu, H. 2019. Deep Anomaly Detection on Attributed Networks. In *SDM*.

Ding, K.; Li, J.; and Liu, H. 2019. Interactive anomaly detection on attributed networks. In *WSDM*.

Ding, K.; Nouri, E.; Zheng, G.; Liu, H.; and White, R. 2024. Toward Robust Graph Semi-Supervised Learning against Extreme Data Scarcity. *TNNLS*.

Ding, K.; Wang, J.; Caverlee, J.; and Liu, H. 2022. Meta Propagation Networks for Few-shot Semi-supervised Learning on Graphs. In *AAAI*.

Ding, K.; Wang, J.; Li, J.; Caverlee, J.; and Liu, H. 2023a. Robust Graph Meta-learning for Weakly-supervised Few-shot Node Classification. *TKDD*.

Ding, K.; Wang, J.; Li, J.; Shu, K.; Liu, C.; and Liu, H. 2020. Graph prototypical networks for few-shot learning on attributed networks. In *CIKM*.

Ding, K.; Wang, Y.; Yang, Y.; and Liu, H. 2023b. Eliciting Structural and Semantic Global Knowledge in Unsupervised Graph Contrastive Learning. In *AAAI*.

Liu*, Y.; Ding*, K.; Liu, H.; and Pan, S. 2023. GOOD-D: On Unsupervised Graph Out-Of-Distribution Detection. In *WSDM*.

Wang, S.; Ding, K.; Zhang, C.; Chen, C.; and Li, J. 2022. Task-adaptive few-shot node classification. In *KDD*.

Xu, Z.; Ding, K.; Wang, Y.-X.; Liu, H.; and Tong, H. 2022. Generalized Few-Shot Node Classification. In *ICDM*.