Towards Unsupervised Deep Graph Structure Learning

Zepeng Zhang

Existing supervised graph structure learning (GSL) methods have the following issues: (1) The reliance on label information limits the application of supervised GSL on more general cases where annotation is unavailable; (2) In the semi-supervised node classification task, the connections among the supervised nodes and their neighbors would receive more guidance than other nodes; (3) Existing methods only consider specific downstream tasks, which means the learned structure contains more task-specific information rather than general knowledge.

This paper [1] proposed a novel StrUcture Bootstrapping contrastive LearnIng fraMEwork (SUBLIME) for unsupervised graph structure learning with the aid of self-supervised contrastive learning. Specially, an "anchor graph" is generated from the original data and a contrastive loss is used to maximize the agreement between anchor graph and the learned graph. The overall pipeline of SUBLIME is depicted below. SUBLIME consists of two components: the graph structure learning module that models and regularizes the learning graph topology and the structure bootstrapping contrastive learning module that provides a self-optimized supervision signal for GSL.

1. In the graph structure learning module, a sketched adjacency matrix is first parameterized by a graph learner (e.g., full-parameterization, attentive network, MLP, and GNN), and then refined by a post-processor to be a sparse, non-negative, symmetric and normalized adjacency matrix.

2. In the structure bootstrapping contrastive learning module, an anchor graph is first constructed (identity matrix or original adjacency matrix if it is available). The feature masking and edge dropping schemes are used for data augmentation. Then with a GNN as the encoder and a MLP as the projector, the contrastive loss is computed for the resulting node representations.

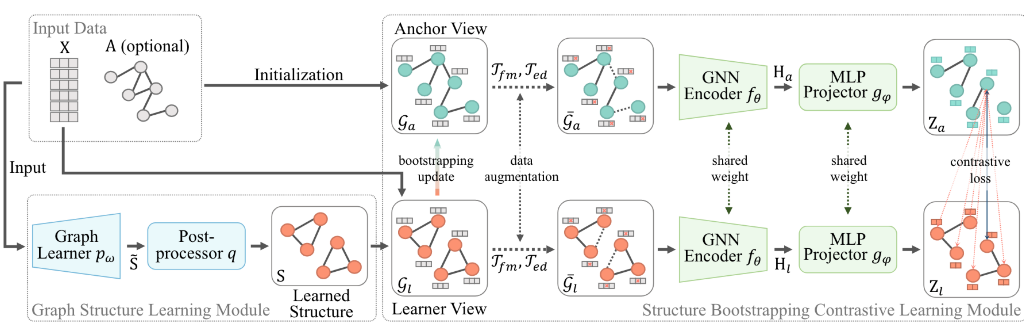

Figure 1: The overall pipeline of SUBLIME

However, the fixed anchor graph may lead to several issues: (1) Noises in the original graph will be inherited by the learned graph; (2) Once the learned graph captures the information of the anchor graph, it will be hard for the model to gain effective supervision in the following training steps; (3) The learned structure may overfit the anchor structure. To mitigate the above issues, a structure bootstrapping mechanism is used. The core idea is to update the anchor structure $\mathbf{A}$ with a slow-moving augmentation of the learned structure $\mathbf{S}$ every c iterations: $\mathbf{A} \leftarrow \tau\mathbf{A} + (1-\tau)\,\mathbf{S}$.

References

[1] Yixin Liu, Yu Zheng, Daokun Zhang, Hongxu Chen, Hao Peng, and Shirui Pan. Towards unsupervised deep graph structure learning. *arXiv preprint arXiv:2201.06367*, 2022. (document)