

## Tag2M- Task-Agnostic Knowledge Distillation Framework for Distilling Gnn- MLP

Graph Neural Networks (GNNs) have achieved remarkable success in various downstream tasks, such as node classification and link prediction. Yet, efficiently deploying GNNs remains a challenge due to their computational complexity and the overhead introduced by message passing. Graph knowledge distillation aims to address this by transferring task-specific structural knowledge from teacher GNNs to lightweight student GNNs or Multi-Layer Perceptrons (MLPs). Despite its promise, existing distillation approaches suffer from several limitations: (i) they require extensive task-specific supervision (ii) they must be retrained separately for each downstream task, and (iii) they often struggle in heterophilous settings. To overcome these challenges, we propose TAG2M, a Task-Agnostic Gnn-to-MLP distillation framework designed for efficient and accurate few-shot inference. TAG2M introduces several novel strategies, including a self-supervised contrastive loss that captures topological information solely from node attributes. Additionally, it leverages Lipschitz embeddings to encode positional information with provable distortion bounds, ensuring robust representation learning. To further enhance adaptability for few-shot inference, TAG2M incorporates a learnable prompt head, which facilitates rapid task adaptation even in label-scarce settings. Figure 1 illustrates the pipeline of Tag2M, which comprises of three key components: (1) task-agnostic training of a teacher GNN, (2) student MLP training to replicate the teacher’s topological knowledge using solely node attributes, and (3) an inference-time prompting layer that enhances the student with positional information for downstream tasks.

Unlike prior methods, TAG2M generalizes well across both homophilous and heterophilous datasets while delivering a significant computational advantage, achieving up to a 20X-200X speed-up. Extensive evaluations on 11 public datasets demonstrate its superior accuracy across diverse tasks, including node classification, link prediction, and node regression, outperforming state-of-the-art approaches.

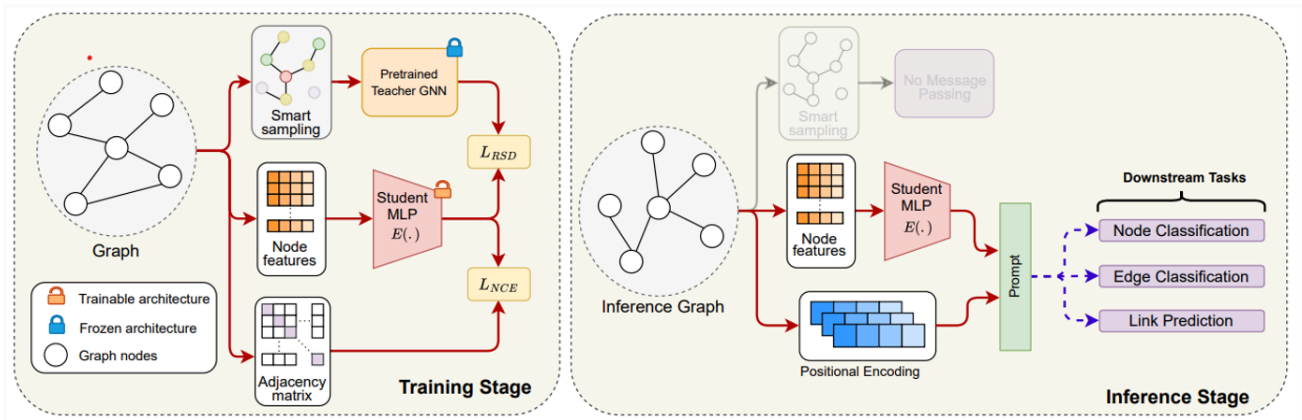


Figure 1: The overall framework of the proposed Tag2M. During inference, Tag2M does not use the test graph.