Ethical Statement

Our submission does not involve any ethical issues, including but not limited to privacy, security, etc.

Appendix

A. Discussion and Comparison on SSL Tasks

We would like to highlight the contributions of the proposed two self-supervised tasks, namely *Label Distribution Preservation* and *Global-Path Prediction*, and how they differ from existing work from the following three aspects:

- *(1) Learning Objectives.* Most existing work on graph self-supervised learning has focused on applying pretext tasks to train a well-performing feature extractor, that is, to learn transferable feature knowledge from unlabeled data. However, the two self-supervised tasks proposed in this paper, e.g., *Label Distribution Preservation* and *Global-Path Prediction*, aim to improve graph augmentation, especially for heterophily graphs, by incorporating two important contextual information, e.g., *global position and label distribution*. Taking the task of Global-Path Prediction as an example, it enables graph augmentation to consider not only the similarity of node pairs but the long-range dependencies between nodes, which is tailored for structure learning, but simply applying it to feature extractor cannot maximize its potential.
- *(2) Method Design.* [1] proposes to take label distributions as supervision, but they simply take label distributions as pseudo-labels. These pseudo-labels are pre-calculated and kept fixed before training starts, making it only a means to expand the annotated data for learning an expressive feature extractor. In contrast, we treat label distributions as important contextual information and re-calculate the label distributions of the original and augmented graphs at each epoch. Finally, we take the consistency of their label distributions as a constraint to prevent overly drastic graph structure perturbations.
- *(3) Experimental Evaluation.* The results of the ablation study in Fig. 2 have demonstrated that both two self-supervised tasks help to improve the graph structure augmentation across various datasets and their performance gains for heterophily graphs are much larger than for homophily graphs.

References

1. Jin, W., Derr, T., Liu, H., Wang, Y., Wang, S., Liu, Z., Tang, J.: Self-supervised learning on graphs: Deep insights and new direction. arXiv preprint arXiv:2006.10141 (2020)