Training Graph Neural Networks via Self-Supervised Learning: Experiments and Analysis

Ching-Ru Ho

Data Science Master Degree Program, National Taiwan University

Abstract

Supervised learning is a popular way for model training. However, its success relies a large number of labeled data. Recent advances in self-supervised learning has provided researchers another way to train models on data in which only few observations are required to be labeled. Self-supervised learning is efficient since it can carry out model training without a large amount of preprocessed data. State-of-the-art self-supervised models can achieve, even exceed, performance of the supervised ones.

While most of the research on self-supervised learning has been conducted in the fields of computer vision and natural language processing, self-supervised learning on graph data is still in its infancy. In this thesis, we explored self-supervised learning for graph neural networks. We conducted experiments by training graph neural network models on four molecular and bioinformatics datasets with different experimental setting. We also provided possible explanations for the experiment results. We found that models with deeper encoder structure can get better results. However, increasing the hidden dimension when model are trained on small or medium size datasets can only achieve little improvements. On the other hand, different data augmentation methods and different types of models can yield different results in molecular and bioinformatics datasets.

Keywords: self-supervised learning, graph neural network, SSL encoder