

PRE-DRAFT: Do your graph's curves affect performance?: A unifying framework for quantifying the data propagation bottlenecks of graph representation learning methods - motivation

By: Joshua Victor Niemelä and Mustafa Hekmat Al-Abdelamir

Introduction

To alleviate various limitations of Graph Neural Networks for use in Graph Representation Learning, Topological Deep Learning is emerging as a promising field of machine learning. TDL has shown promising results in tackling over-smoothing and over-squashing, still, perhaps because of the young nature of the field, there is limited theoretical foundations for quantifying these possible advantages. This poses a problem in making quantitative comparisons between various architectures for GRL, in this paper we attempt to lay the foundations for various metrics to provide insights on over-squashing and over-smoothing and how they relate to the model's ability to learn. To verify and support our theoretical foundations we also run benchmarks against other approaches used for learning on graph data.

Data

We will be using the QM9, Benchmark dataset for graph classification and the Human Body segmentation datasets to benchmark our own implementations against other models cited in various papers.

Additionally, we will be constructing synthetic datasets to compare the expressivity between various models for the task of identifying various topological structures in graphs.

Methods

We will be using Pytorch Geometric to build our models and to replicate models cited in other papers as a comparison and empirical exploration of GRL.

Learning Objectives

1. Gain a solid understanding of the theoretical foundations of GNNs.
2. Investigate the oversquashing problem in GNNs and how it affects the ability for the model to learn long-range dependencies.[2]
3. Gain insights and understanding about TDL, and how topology is leveraged for learning and specifically how it relates to the aforementioned bottlenecks.[3]
4. Construct generalisable metrics to quantify various geometric and topological properties of GNNs and the datasets they are trained on.
5. Construct a model, for instance a transformer or CCNN[1], that can learn topological features in data and benchmark against non-topological approaches.

References

- [1]: [Topological Deep Learning](#)
- [2]: [On the Bottleneck of Graph Neural Networks and its Practical Implications](#)
- [3]: [Topological Graph Neural Networks](#)