Semi-Supervised Graph Learning Meets Dimensionality Reduction

Final Project for Unsupervised Learning in Fall 2021

Alex Morehead & Watchanan Chantapakul

November 2, 2021

1 Project Proposal

1.1 Abstract

Semi-supervised learning (SSL) has recently received increased attention by machine learning researchers [1]. By enabling effective propagation of known labels in graph-based deep learning (GDL) algorithms, SSL is poised to become an increasingly-used technique in GDL in the coming years. However, there are currently few explorations in the graph-based SSL literature on exploiting classical dimensionality reduction techniques for improved label propagation [2], [3]. In this work, we propose to investigate the use of dimensionality reduction techniques such as PCA, t-SNE, and UMAP to see their effect on the performance of graph neural networks (GNNs) designed for semi-supervised propagation of node labels. Our study will make use of benchmark semi-supervised GDL datasets such as the Cora and Citeseer datasets to allow meaningful comparisons of the representations learned by each algorithm when paired with a dimensionality reduction technique. Our node classification results will be reported in terms of cross entropy loss, precision, recall, and F1 score.

1.2 Responsibilities

1. Alex Morehead - Responsible for Graph Deep Learning Tasks
2. Watchanan Chantapakul - Responsible for Dimensionality Reduction Tasks

1.3 Timeline

- 11/11/2021: Have baseline graph deep learning environment for training models completed
- 11/16/2021: Have dimensionality reduction algorithms selected
- 11/18/2021: Have rough first versions of each dimensionality reduction algorithm implemented
- 11/23/2021: Have underway integration of dimensionality reduction algorithms into our deep learning environment
- 11/25/2021: Have rough first version of deep learning models training with *apriori* dimensionality reduction
- 11/30/2021: Have underway refinement of deep learning models and their results
- 12/02/2021: Have analysis of model results and presentation preparation underway
- 12/07/2021: Present study's results
- 12/09/2021: Have composition of final project report underway
- 12/14/2021: Have first few drafts of final project report completed
- 12/16/2021: Submit final project report

References

[1] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *CoRR*, abs/1609.02907, 2016.

[2] Guoxian Yu, Guoji Zhang, Carlotta Domeniconi, Zhiwen Yu, and Jane You. Semi-supervised classification based on random subspace dimensionality reduction. *Pattern Recognition*, 45(3):1119–1135, 2012.

[3] Danfeng Hong, Naoto Yokoya, Jocelyn Chanussot, Jian Xu, and Xiao Xiang Zhu. Learning to propagate labels on graphs: An iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction. *ISPRS journal of photogrammetry and remote sensing*, 158:35–49, 2019.