THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

MASTER'S THESIS PRESENTATION

SOWON JEONG
Department of Statistics
The University of Chicago

Understanding Graph Neural Network as a Dimension Reduction Technique

WEDNESDAY, February 15, 2023, at 11:00 AM
Zoom Meeting

ABSTRACT

Dimension reduction is an essential procedure dealing with high-dimensional data--it allows us to visualize the high-dimensional data in 2D or 3D space, reveal high-dimensional structures such as clusters, and select the critical features of ambient space. From traditional methods like Principal Component Analysis(PCA) to t-distributed Stochastic Neighbor Embedding(t-SNE), Uniform Manifold Approximation & Projection(UMAP), and dens-MAP, there are numerous attempts to better represent the given data in a lower-dimensional space. On top of these methods, we want to approach the Graph Neural Network(GNN) as a dimension reduction method and highlight the advantages of GNN over these methods. We also define the multiple criteria to compare the dimension reduction methods in diverse aspects: global geometry(structure), local geometry, density preservation, scalability with the sample size, scalability with the dimension, robustness to the initialization, and robustness to noise. In the end, we provide an extensive empirical evaluation on both synthetic and real datasets. This paper shows the strength and limitations of widely-used dimension reduction techniques and sheds light on what we can learn from GNN to improve the dimension reduction techniques.