Graph Attention Networks

NUS CS6208 Paper Review

Runze Cai (A0212755H)

Introduction

- Previous spectral graph learning, e.g., ConvNets and MoNet, have limitations in handling different graph structures. While Spatial learning could overcome these limitations.
- The GAT, a spatial learning approach, proposes a self-attention mechanism that better learns the weights of each node's neighbor and generates node representations for graphs.
- GAT can be applied to many applications, such as node classification.

Method

- GAT leverages a self-attention mechanism that enables each node to learn from the most relevant neighbors with learned weights.
- GAT uses multi-heads to represent different types of node relationships, improving model complexity and efficiency through parallel computation.

Evaluation & Results

- The GAT model was evaluated on four classification benchmarks, including Cora, Citeseer, and Pubmed for transductive learning, and the PPI dataset for inductive learning.
- Results showed that GAT outperformed all other methods, including MoNet on the Cora and Citeseer datasets, and achieved comparable accuracy to GCN-64 on the Pubmed dataset.
- For inductive training, GAT improved by 20.5% compared to the SOTA GraphSAGE.

Conclusion & Improvement

- The GAT model's attention mechanism enables it to operate on groups of neighbors in graphs using spatial learning, making it flexible to different graph structures and outperforming other techniques on node classification benchmarks.
- There is still room for improvement, such as balancing the number of layers, hidden dimensions, and multi-head numbers.
- Considering global information, e.g., transformers, or global representations using master nodes [1].
- Incorporating edge features with a self-attention mechanism could also be helpful.

P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, *et al.*, "Relational inductive biases, deep learning, and graph networks," *arXiv preprint arXiv:1806.01261*, 2018.