Project Proposal (Research): Leveraging Global Information in Self-attention Graph Convolutional Networks

Chen Wang (wc497) Chengyuan Deng (cd751)

Graph Convolutional Networks (GCN) [1, 2] have been widely applied to multiple tasks such as graph classification, node label prediction and graph clustering. Thanks to its power in inferring and manipulating relational information, it has been regarded as one of the most promising techniques in pursuing Artificial General Intelligence [3]. Like other types of neural networks, it is a common practice to utilize attention or self-attention mechanism in GCNs to improve performance. Existing self-attention mechanisms in GCNs usually consider the feature information between neighboring vertices, and assign connection weights to each vertex accordingly [4, 5]. This type of technique can augment the information from neighboring nodes with similar features. Nevertheless, it only considers *local geometry* as the edge connections of the graph, and exclude possible scenarios when a vertex can have strong correlation/influence with another *without edge connection*. To date, as we can see from a comprehensive survey [6], there has now been any publication considering such issues. Hence, in this project, we intend to design a novel GCN to take global information into self-attention mechanism.

The motivation comes from the self-attention generative adversarial networks (SAGAN) [7], which proposed a global self-attention mechanism into convolutional GANs. The self-attention layer takes the image features x of previous layer as input, the attention is calculated by the following equation:

$$\beta_{j,i} = \mathbf{softmax}(s_{i,j}), \text{where } s_{i,j} = (W_{1,i}X_i)^{\mathrm{T}}(W_{2,j}X_j) \tag{1}$$

The output of the self-attention layer is:

$$o_j = W(\sum_{i=1}^{N} \beta_{j,i} W_h X_i) \tag{2}$$

Following this motivation, we propose Global Self-attention Graph Convolution Networks (GSA-GCN), which leverage self-attention mechanism into graph networks instead of vanilla convolutional networks.This work is supposed to be a generalization since convolutional networks could be regarded as a fully-connected graph networks within a convolution kernel. We intend to conduct experiments to prove that long-range, global-level dependencies could be learned in graph-structured data.

To evaluate our work, we can test the performance of the proposed self-attention GCN on different tasks involving graph data. For instance, we can test node classification tasks on citation data such as Cora, Citeseer and Pubmed. For another example, we can test unsupervised or semi-supervised node clustering on datasets like BlogCatalog. If GCNs with the proposed global self-attention mechanism shows consistent better performance than GCNs without it, we can therefore conclude the effectiveness of the mechanism we proposed.

References

[1] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2008.

[2] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[3] Peter W Battaglia, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

[4] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

[5] John Boaz Lee, Ryan Rossi, and Xiangnan Kong. Graph classification using structural attention. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1666–1674. ACM, 2018.

[6] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S Yu. A comprehensive survey on graph neural networks. *arXiv preprint arXiv:1901.00596*, 2019.

[7] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. *arXiv preprint arXiv:1805.08318*, 2018.