

Literature Review of GAT for SAiDL Spring Assignment 2024

Shanay Mehta

March 2024

1 Main Body

Graph Attention Network(GAT) are graph neural network layers which use attention in a **computationally efficient** manner to address various shortcomings of its predecessors. GAT's are applicable to both transductive and inductive problems due to which they are effective in a variety of applications such as citation network analysis, PPI prediction, social network analysis. The key feature of GAT is its ability to assign distinct attention to neighbouring nodes in an implicit manner. This is a big improvement over GCN because this improves model capacity and also brings interpretability. GAT's achieve all this while maintaining a time-complexity similar to GCN. This is because the operation of the self-attentional layer can be parallelized across all edges, and the computation of output features can be parallelized across all nodes. Additionally, GATs do not depend on upfront access to the global graph structure because they allow every node to attend to every other node, dropping all structural information. Masked attention is used to input structural information in the model and using multi-head attention stabilises the learning process of self-attention. The authors evaluated GAT on 4 benchmark datasets(3 transductive and 1 inductive) where GAT models achieved comparable/ better performance against baselines such as GCN and GraphSAGE. The feature representations learned by the model show clear groupings in a 2D view, corresponding to the classes of the dataset, thus verifying the model's discriminative power. GAT has further been improved by using adaptive attention which can capture dynamic information during training leading to enhanced performance and robustness.