Graph Neural Networks

Graph Attention Networks

Dr. Alireza Aghamohammadi

Graph Attention Networks (GATs)

❖ The attention mechanism is a key component in transformer architectures and can be adapted for graph neural networks (GNNs).

❖ In Graph Attention Networks (GATs), an **attention mechanism** is used to learn the importance of each neighboring node's message. The incoming messages are weighted by **attention coefficients** A_{nm}, which determine the contribution of node m to node n's representation.

❖ The aggregated representation of node n is computed as:

$$\text{agg}[n] = \sum_{m \in \text{ne}[n]} A_{nm} h_m,$$

$$A_{nm} \geq 0, \quad \sum_{m \in \text{ne}[n]} A_{nm} = 1,$$

where h_m is the feature vector of node m and $\text{ne}[n]$ represents the set of neighboring nodes of n.

❖ The attention coefficients A_{nm} indicate the importance of the neighboring nodes' messages. Larger coefficients imply higher importance.

❖ One common approach to compute the attention coefficients is using a softmax function:

$$A_{nm} = \frac{\exp(h_n^T W h_m)}{\sum_{m' \in \text{ne}[n]} \exp(h_n^T W h_{m'})},$$

where $W \in \mathbb{R}^{D \times D}$ is a learnable weight matrix.