Graph Neural Networks

Graph Convolutional Networks

Dr. Alireza Aghamohammadi

Graph Convolutional Networks (GCNs)

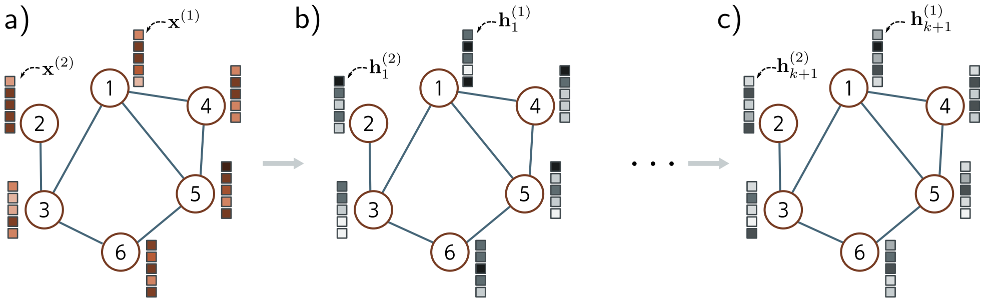
- GCNs are convolutional in nature: they update each node's embedding by aggregating information from neighboring nodes.
- This structure introduces a *relational inductive bias*, encouraging the model to prioritize information from a node's local neighborhood.
- Each layer in the GCN is a function that, given node embeddings and the adjacency matrix, outputs new node embeddings:

$$
\begin{aligned}
H_1 &= F(X, A, \phi_0) \\
H_2 &= F(H_1, A, \phi_1) \\
H_3 &= F(H_2, A, \phi_2) \\
&\vdots \\
H_K &= F(H_{K-1}, A, \phi_{K-1})
\end{aligned}
$$

where H_i represents the node embeddings at layer i, A is the adjacency matrix, and ϕ_i denotes parameters for mapping from layer i to $i+1$.

Message Passing in GCNs

- A basic GCN layer takes an input graph with adjacency matrix A and node embeddings. The adjacency matrix is hidden in the function aag$[n]$.
- Intuitively, each node aggregates messages from its neighbors to update its embedding.



$$\mathbf{h}_1^{(n)} = \mathbf{a}\left[\boldsymbol{\beta}_0 + \boldsymbol{\Omega}_0 \mathbf{x}_1^{(n)} + \boldsymbol{\Omega}_0 \mathbf{agg}[n]\right]$$

$$\mathbf{h}_{k+1}^{(n)} = \mathbf{a}\left[\boldsymbol{\beta}_k + \boldsymbol{\Omega}_k \mathbf{h}_k^{(n)} + \boldsymbol{\Omega}_k \mathbf{agg}[n]\right]$$

Image credit: Understanding Deep Learning

Aggregation and Update in GCNs

- There are various forms of the aggregation function, but it must be invariant to the ordering of inputs.
- For instance, at each node n in layer k, we can aggregate information by summing the embeddings of neighboring nodes:

$$\text{agg}[n,k] = \sum_{m \in \text{ne}[n]} h_k^{(m)}$$

  where $\text{ne}[n]$ denotes the set of neighboring nodes of node n.
- Using the adjacency matrix A, we can represent this aggregation more concisely:

$$\begin{aligned} H_{k+1} &= a\left[\beta_k \mathbf{1}^T + \Omega_k H_k + \Omega_k H_k A\right] \\ &= a\left[\beta_k \mathbf{1}^T + \Omega_k H_k (A + I)\right] \end{aligned}$$