CSCE 52803: Assignment 4

Due 11:59 pm Saturday, October 24, 2026

1 Neural Networks and Backpropagation

Consider a neural network with one m-dimensional hidden layer:

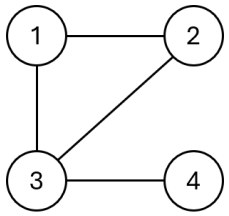
$$h = \mathrm{ReLU}(Wx + b), \qquad \hat{y} = \sigma(v^\top h + c),$$

where $x \in \mathbb{R}^d$ is an input column vector, $W \in \mathbb{R}^{m \times d}$ is the parameter matrix, $b, h, v \in \mathbb{R}^m$, $y \in \{0,1\}$ is the true label, $\hat{y} \in (0,1)$ is the predicted probability, and σ is the sigmoid function and the loss is binary cross-entropy:

$$L = -y\log\hat{y} - (1-y)\log(1-\hat{y}).$$

1. Explain why backpropagation is an application of the chain rule.
2. Derive $\partial L / \partial v$ and $\partial L / \partial W$.

2 Message Passing in Graph Neural Networks



Consider an undirected graph $G = (V, E)$ where $V = \{1,2,3,4\}$ and $E = \{\{1,2\},\{1,3\},\{2,3\},\{3,4\}\}$. A Message-Passing Neural Network (MPNN) updates each node embedding as

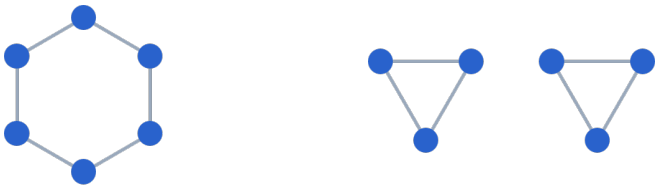
$$h_v^{(k+1)} = \mathrm{ReLU}\left(W_{\mathrm{self}} h_v^{(k)} + W_{\mathrm{nbr}} \sum_{u \in N(v)} h_u^{(k)} \right).$$

The initial node embeddings are

$$h_v^{(0)} = [1,2,3]^\top \text{ for every } v \qquad W_{\mathrm{self}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad W_{\mathrm{nbr}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

1. Compute the node embeddings after one message-passing layer.
2. Compute the embeddings after two layers.
3. Explain why a k-layer GNN with local message passing uses information from at most k hops away. Assume the initial features contain only local node information.

3 Expressive Limitations of GNNs



Graph A Graph B

Consider two undirected graphs: Graph A is the cycle C_6, and Graph B consists of two disjoint triangles. All six nodes in each graph have the same initial node embedding, which is also identical across the two graphs. Consider

$$h_v^{(k+1)} = \phi\left(h_v^{(k)}, \sum_{u \in N(v)} \psi\left(h_u^{(k)}\right)\right),$$

where the functions are shared across nodes and the same model parameters are used for both graphs.

1. Show that all nodes in both graphs have identical embeddings after any number of layers.