<u>Mid-term Exam (Graph Neural Networks –Fall 2024)</u>

Full Name:
Student ID:

1. (10pt) Given a graph with an adjacency matrix A and initial node feature matrix $H^{(0)}$ as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \qquad H^{(0)} = \begin{bmatrix} -2 & 1 \\ -3 & -3 \\ -1 & 4 \\ 2 & 2 \\ 4 & 0 \end{bmatrix}$$

Assume that the hidden layer of an GCN model of all nodes at layer (k) can be calculated as:

$$H^{(k)} = \sigma\left(A \cdot H^{(k-1)}\right),$$

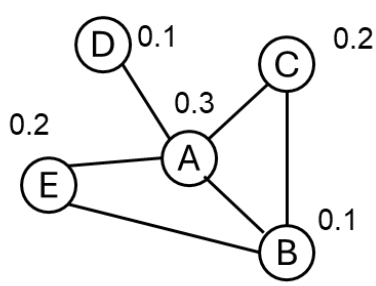
where $H^{(k)}$ denotes the output at layer k, σ is a ReLU function $\text{ReLU}(x) = \max(0, x)$. Calculate the output of the GCN model at layer $k = 1$.

<u>**Ans:**</u>

$$H^1 = \sigma(AH^0) = \text{ReLU}\left(\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} -2 & 1 \\ -3 & -3 \\ -1 & 4 \\ 2 & 2 \\ 4 & 0 \end{bmatrix}\right) = \text{ReLU}\left(\begin{bmatrix} 2 & 2 \\ 4 & 0 \\ -1 & -1 \\ -1 & 7 \\ -4 & 4 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 2 & 2 \\ 4 & 0 \\ 0 & 0 \\ 0 & 7 \\ 0 & 4 \end{bmatrix}$$

2. (10pt) Consider an undirected graph G of five nodes A, B, C, D, and E given in the following figure. Each node has initial features that are the numbers standing next to it (i.e., the initial feature of node 'A' is $h_A^{(0)} = 0.3$). According to GraphSAGE model with an AGGREGATE is a MEAN function, the feature of a node i at layer k can be updated as:



$$h_{N(i)}^{(k)} = \text{AGGREGATE}\left(\{h_u^{(k-1)}, \forall u \in N(i)\}\right)$$
$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} || h_{N(i)}^{(k)}\right)$$

where $||$ is a concatenation, $\text{ReLU}(x) = \max(0, x)$, *N(i)* is the neighbour nodes of node *i*.

a) Calculate the feature of each node at $k = 1$.
b) Calculate a graph-level embedding h_G by using a 'Mean' global pooling when $k = 1$

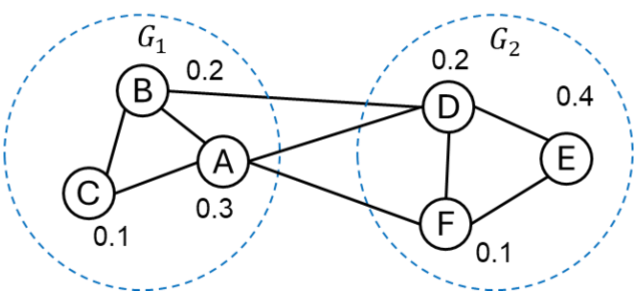
Ans:

$$h_{N(i)}^{(k)} = \text{AGGREGATE}\left(\{h_u^{(k-1)}, \forall u \in N(i)\}\right)$$
$$h_{N(A)}^{(1)} = MEAN(h_B, h_C, h_D, h_E) = MEAN(0.1, 0.2, 0.1, 0.2) = 0.15$$
$$h_{N(B)}^{(1)} = MEAN(h_A, h_C, h_E) = MEAN(0.3, 0.2, 0.2) \approx 0.23$$
$$h_{N(C)}^{(1)} = MEAN(h_A, h_B) = MEAN(0.3, 0.1) = 0.2$$
$$h_{N(D)}^{(1)} = MEAN(h_A) = MEAN(0.3) = 0.3$$
$$h_{N(E)}^{(1)} = MEAN(h_A, h_B) = MEAN(0.3, 0.1) = 0.2$$

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} || h_{N(i)}^{(k)}\right)$$
$$h_A^{(1)} = \text{ReLU}\left(h_A^{(0)} || h_{N(A)}^{(0)}\right) = Max(0, [0.3, 0.15]) = [0.3, 0.15]$$
$$h_B^{(1)} = \text{ReLU}\left(h_B^{(0)} || h_{N(B)}^{(0)}\right) = Max(0, [0.1, 0.23]) = [0.1, 0.23]$$
$$h_C^{(1)} = \text{ReLU}\left(h_C^{(0)} || h_{N(C)}^{(0)}\right) = Max(0, [0.2, 0.2]) = [0.2, 0.2]$$
$$h_D^{(1)} = \text{ReLU}\left(h_D^{(0)} || h_{N(D)}^{(0)}\right) = Max(0, [0.1, 0.3]) = [0.1, 0.3]$$
$$h_E^{(1)} = \text{ReLU}\left(h_E^{(0)} || h_{N(E)}^{(0)}\right) = Max(0, [0.2, 0.2]) = [0.2, 0.2]$$

$$h_G = MEAN\left(h_A^{(1)}, h_B^{(1)}, h_C^{(1)}, h_D^{(1)}, h_E^{(1)}\right) = [0.18, 0.216]$$

3. (10pt) Consider an undirected graph G of six nodes A, B, C, D, E and F given in the following figure. The graph G contains two cluster G_1 and G_2. Each node has initial features that are the numbers standing next to it. According to ClusterGCN model, the feature of a node i at layer k can be updated as:



$$h_{N(i)}^{(k)} = \text{MEAN}\left(\{h_u^{(k-1)}, \forall u \in N(i), G_u = G_i\}\right)$$
$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} || h_{N(i)}^{(k)}\right)$$

where || is a concatenation.

Calculate the output representations of all nodes at layer $k = 1$.

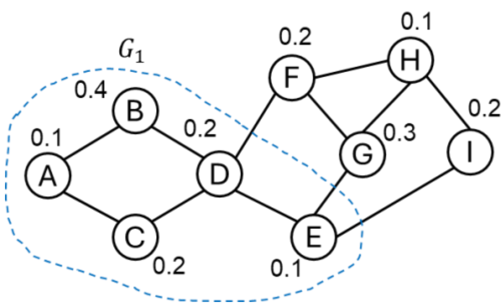
Ans:

$$h_{N(i)}^{(k)} = \text{MEAN}\left(\{h_u^{(k-1)}, \forall u \in N(i)\}, G_u = G_i\right)$$
$$h_{N(A)}^{(1)} = MEAN(h_B, h_C) = MEAN(0.2, 0.1) = 0.15$$
$$h_{N(B)}^{(1)} = MEAN(h_A, h_C) = MEAN(0.3, 0.1) = 0.2$$

$$h_{N(C)}^{(1)} = MEAN(h_A, h_B) = MEAN(0.3,0.2) = 0.25$$
$$h_{N(D)}^{(1)} = MEAN(h_E, h_F) = MEAN(0.4,0.1) = 0.25$$
$$h_{N(E)}^{(1)} = MEAN(h_D, h_F) = MEAN(0.2,0.1) = 0.15$$
$$h_{N(F)}^{(1)} = MEAN(h_D, h_E) = MEAN(0.4,0.2) = 0.3$$

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)}||h_{N(i)}^{(k)}\right)$$
$$h_A^{(1)} = \text{ReLU}\left(h_A^{(0)}||h_{N(A)}^{(0)}\right) = Max(0, [0.3,0.15]) = [0.3,0.15]$$
$$h_B^{(1)} = \text{ReLU}\left(h_B^{(0)}||h_{N(B)}^{(0)}\right) = Max(0, [0.2,0.2]) = [0.2,0.2]$$
$$h_C^{(1)} = \text{ReLU}\left(h_C^{(0)}||h_{N(C)}^{(0)}\right) = Max(0, [0.1,0.25]) = [0.1,0.25]$$
$$h_D^{(1)} = \text{ReLU}\left(h_D^{(0)}||h_{N(D)}^{(0)}\right) = Max(0, [0.2,0.25]) = [0.2,0.25]$$
$$h_E^{(1)} = \text{ReLU}\left(h_E^{(0)}||h_{N(E)}^{(0)}\right) = Max(0, [0.4,0.15]) = [0.4,0.15]$$
$$h_F^{(1)} = \text{ReLU}\left(h_F^{(0)}||h_{N(F)}^{(0)}\right) = Max(0, [0.1,0.25]) = [0.1,0.3]$$

4. (10pt) Consider an undirected graph G of nine nodes A, B, C, D, E, F, G, H and I given in the following figure. The graph G has subgraph sampling G_1. Each node has initial features that are the numbers standing next to it. According to GraphSAINT model, the feature of a node i at layer k can be updated as:



$$h_{N(i)}^{(k)} = \text{MEAN}\left(\{h_u^{(k-1)}, \forall u \in N(i), G_u = G_i\}\right)$$
$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)}||h_{N(i)}^{(k)}\right)$$

where || is a concatenation.

Calculate the output representations of node A, B, C, D, E at layer $k = 1$.

<u>**Ans:**</u>

$$h_{N(i)}^{(k)} = \text{MEAN}\left(\{h_u^{(k-1)}, \forall u \in N(i)\}, G_u = G_i\right)$$
$$h_{N(A)}^{(1)} = MEAN(h_B, h_C) = MEAN(0.4, 0.2) = 0.3$$
$$h_{N(B)}^{(1)} = MEAN(h_A, h_D) = MEAN(0.1,0.2) = 0.15$$
$$h_{N(C)}^{(1)} = MEAN(h_A, h_D) = MEAN(0.1,0.2) = 0.15$$
$$h_{N(D)}^{(1)} = MEAN(h_B, h_C, h_E) = MEAN(0.4,0.2, 0.1) = 0.23$$
$$h_{N(E)}^{(1)} = MEAN(h_D) = MEAN(0.2) = 0.2$$

$$h_i^{(k)} = \text{ReLU}\left(h_i^{(k-1)} || h_{N(i)}^{(k)}\right)$$
$$h_A^{(1)} = \text{ReLU}\left(h_A^{(0)} || h_{N(A)}^{(0)}\right) = Max(0, [0.1, 0.3]) = [0.1, 0.3]$$
$$h_B^{(1)} = \text{ReLU}\left(h_B^{(0)} || h_{N(B)}^{(0)}\right) = Max(0, [0.15, 0.4]) = [0.15, 0.4]$$
$$h_C^{(1)} = \text{ReLU}\left(h_C^{(0)} || h_{N(C)}^{(0)}\right) = Max(0, [0.15, 0.2]) = [0.15, 0.2]$$
$$h_D^{(1)} = \text{ReLU}\left(h_D^{(0)} || h_{N(D)}^{(0)}\right) = Max(0, [0.23, 0.2]) = [0.23, 0.2]$$
$$h_E^{(1)} = \text{ReLU}\left(h_E^{(0)} || h_{N(E)}^{(0)}\right) = Max(0, [0.2, 0.1]) = [0.2, 0.1]$$

5. (10pt) (JKNET) Given a graph with an adjacency matrix A and initial node feature matrix $H^{(0)}$ as follows:

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \qquad H^{(0)} = \begin{bmatrix} -2 & 0 \\ 2 & -2 \\ 2 & 4 \\ -2 & 2 \\ 4 & 3 \end{bmatrix}$$

Assume that the output of an JK network model of all nodes at layer (k) can be calculated as:

$$H^{(k)} = \max\left(\sigma\left(\tilde{A} \cdot H^{(0)}\right), \sigma\left(\tilde{A} \cdot H^{(1)}\right), \dots, \sigma\left(\tilde{A} \cdot H^{(k-1)}\right)\right)$$

where $H^{(k)}$ denotes the output at layer k, $\tilde{A}$ is the normalized matrix ($\tilde{A} = D^{-1}A$), σ is a ReLU function $\text{ReLU}(x) = \max(0, x)$.

a) Calculate $\tilde{A}$.
b) Calculate the output representations at layer $k = 2$.

Ans:

$$D = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

$$D^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

$$\tilde{A} = D^{-1}A = \begin{bmatrix} \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

$$H^{(k)} = \max\left(\sigma\left(\tilde{A}\ \cdot H^{(0)}\right), \sigma\left(\tilde{A}\ \cdot H^{(1)}\right), \dots, \sigma\left(\tilde{A}\ \cdot H^{(k-1)}\right)\right)$$

- Layer 1:

$$H^1 = \max\left(\sigma\left(\tilde{A}\ \cdot H^{(0)}\right)\right) = \mathrm{ReLU}\left(\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 2 & -2 \\ 2 & 4 \\ -2 & 2 \\ 4 & 3 \end{bmatrix}\right)$$

$$= \mathrm{ReLU}\left(\begin{bmatrix} \frac{-2}{3} & 0 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & \frac{7}{4} \\ 1 & \frac{5}{2} \\ 0 & 3 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & \frac{7}{4} \\ 1 & \frac{5}{2} \\ 0 & 3 \end{bmatrix}$$

- Layer 2:

$$H^2 = \max\left(\sigma\left(\tilde{A}\cdot H^{(0)}\right), \sigma\left(\tilde{A}\cdot H^{(1)}\right)\right)$$

$$= \max\left(\begin{bmatrix} 0 & 0 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & \frac{7}{4} \\ 1 & \frac{5}{2} \\ 0 & 3 \end{bmatrix}, ReLU\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & \frac{7}{4} \\ 1 & \frac{5}{2} \\ 0 & 3 \end{bmatrix}\right)$$

$$= \max\left(\begin{bmatrix} 0 & 0 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & \frac{7}{4} \\ 1 & \frac{5}{2} \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} \frac{2}{3} & \frac{5}{3} \\ \frac{1}{2} & \frac{11}{4} \\ \frac{7}{8} & \frac{39}{16} \\ \frac{1}{2} & \frac{11}{4} \\ \frac{5}{4} & \frac{17}{8} \end{bmatrix}\right) = \begin{bmatrix} \frac{2}{3} & \frac{5}{3} \\ 1 & \frac{11}{4} \\ \frac{3}{2} & \frac{39}{16} \\ 1 & \frac{11}{4} \\ \frac{5}{4} & 3 \end{bmatrix}$$

6. (15pt) Given a graph with an adjacency matrix A and initial node feature matrix $H^{(0)}$ as follows:

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix} \qquad H^{(0)} = \begin{bmatrix} 3 \\ -2 \\ 2 \\ 4 \\ 3 \end{bmatrix}$$

Assume that the output of an GCNII model of all nodes at layer (k) can be calculated as:

$$H^{(k)} = \sigma\left[\left((1-\beta)I_n\right)\cdot\left((1-\alpha)\tilde{A}\cdot H^{(k-1)} + \alpha H^{(0)}\right)\right]$$

where $H^{(k)}$ denotes the output at layer $k, \tilde{A}$ is the normalized matrix ($\tilde{A} = D^{-1}A$), I_n is the identity matrix, $\alpha = \beta = 0.5$, σ is a ReLU function $\text{ReLU}(x) = \max(0, x)$.

c) Calculate $\tilde{A}$.

d) Calculate the output representations at layer $k = 1$.

Ans:

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

$$D^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

$$\tilde{A} = D^{-1}A = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}$$

- $H^{(k)} = \sigma\left[\left((1-\beta)I_n\right)\left((1-\alpha)\tilde{A}\cdot H^{(k-1)} + \alpha H^{(0)}\right)\right]$

$$\tilde{A}\cdot H^{(k-1)} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ 2 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

$$I_n = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\left[\left((1-\beta)I_n\right)\left((1-\alpha)\tilde{A}\cdot H^{(k-1)} + \alpha H^{(0)}\right)\right]$$

$$= \left(0.5\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\right) \cdot \left(0.5\begin{bmatrix} 0 \\ 2 \\ 3 \\ 2 \\ 1 \end{bmatrix} + 0.5\begin{bmatrix} 3 \\ -2 \\ 2 \\ 4 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} \frac{3}{4} \\ 0 \\ \frac{5}{4} \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

$$H^{(k)} = \sigma\left(\begin{bmatrix} \frac{3}{4} \\ 0 \\ \frac{5}{4} \\ \frac{3}{2} \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \frac{3}{4} \\ 0 \\ \frac{5}{4} \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

7. (10pt) Given a graph with an adjacency matrix A and initial node feature matrix $H^{(0)}$ as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad H^{(0)} = \begin{bmatrix} 0 & -2 \\ -1 & 3 \\ 4 & -2 \\ 0 & -5 \end{bmatrix}$$

Assume that the hidden layer of an DeepGCNs model of all nodes at layer (k) can be calculated as:

$$H^{(k)} = \sigma\left(A \cdot H^{(k-1)}\right) + H^{(k-1)},$$

where $H^{(k)}$ denotes the output at layer k, σ is a ReLU function $\text{ReLU}(x) = \max(0, x)$. Calculate the output of the GCN model at layer $k = 2$.

Ans:

$$H^{(1)} = \sigma\left(AH^{(0)}\right) + H^{(0)} = \text{ReLU}\left(\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & -2 \\ -1 & 3 \\ 4 & -2 \\ 0 & -5 \end{bmatrix}\right) + \begin{bmatrix} 0 & -2 \\ -1 & 3 \\ 4 & -2 \\ 0 & -5 \end{bmatrix}$$

$$= \max\left(\begin{bmatrix} -1 & -2 \\ -1 & 3 \\ -1 & -4 \\ 3 & -4 \end{bmatrix}\right) + \begin{bmatrix} 0 & -2 \\ -1 & 3 \\ 4 & -2 \\ 0 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \\ 0 & 0 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ -1 & 3 \\ 4 & -2 \\ 0 & -5 \end{bmatrix}$$

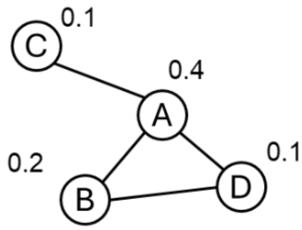
$$= \begin{bmatrix} 0 & -2 \\ -1 & 6 \\ 4 & -2 \\ 3 & -5 \end{bmatrix}$$

$$H^{(2)} = \sigma\left(AH^{(1)}\right) + H^{(1)} = \text{ReLU}\left(\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & -2 \\ -1 & 6 \\ 4 & -2 \\ 3 & -5 \end{bmatrix}\right) + \begin{bmatrix} 0 & -2 \\ -1 & 6 \\ 4 & -2 \\ 3 & -5 \end{bmatrix}$$

$$= \max\left(\begin{bmatrix} 2 & 1 \\ -1 & 6 \\ 2 & -1 \\ 6 & -1 \end{bmatrix}\right) + \begin{bmatrix} 0 & -2 \\ -1 & 6 \\ 4 & -2 \\ 3 & -5 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 6 \\ 2 & 0 \\ 6 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ -1 & 6 \\ 4 & -2 \\ 3 & -5 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 12 \\ 6 & -2 \\ 9 & -5 \end{bmatrix}$$

8. (15pt) Consider an undirected graph G of four nodes A, B, C, and D given in the following figure. Each node has initial features that are the numbers standing next to it (i.e., the initial

feature of node 'A' is $h_A^{(0)} = 0.4$). According to GAT model, the weight matrix W is randomly initialized as [0.5]. The feature of node 'i' at layer (k) can be updated as:



$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im} W h_m\right)$$

Where $\alpha_{im} = \frac{e_{im}}{\sum_{k \in N(i)} e_{ik}}$ and

$$e_{im} = \sigma\big(\mathrm{MEAN}(Wh_i, Wh_m)\big)$$

σ is a ReLU function $\mathrm{ReLU}(x) = \max(0, x)$.

a) Calculate the attention coefficients e_{AB}, e_{AC}, and e_{AD}

b) Calculate the feature of node 'A' at $k = 1$.

Ans:

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im} W h_m\right)$$

where: $\alpha_{im} = \frac{e_{im}}{\sum_{k \in N(i)} e_{ik}}$, and $e_{im} = \sigma\big(\mathrm{MEAN}(Wh_i, Wh_m)\big)$

$$e_{AC} = \sigma\big(\mathrm{MEAN}(Wh_A, Wh_C)\big) = \sigma\big(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.1)\big)$$
$$= \sigma\big(\mathrm{MEAN}(0.2, 0.05)\big) = 0.125$$
$$e_{AB} = \sigma\big(\mathrm{MEAN}(Wh_A, Wh_B)\big) = \sigma\big(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.2)\big)$$
$$= \sigma\big(\mathrm{MEAN}(0.2, 0.1)\big) = 0.15$$
$$e_{AD} = \sigma\big(\mathrm{MEAN}(Wh_A, Wh_D)\big) = \sigma\big(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.1)\big)$$
$$= \sigma\big(\mathrm{MEAN}(0.2, 0.05)\big) = 0.125$$

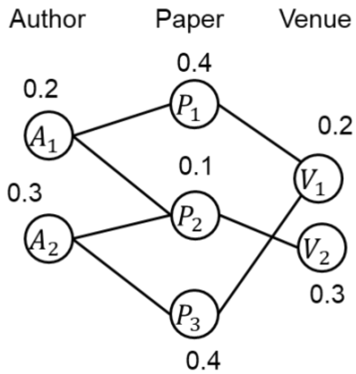
– –

$$\alpha_{AB} = \frac{e_{AB}}{e_{AB} + e_{AC} + e_{AD}} = \frac{0.15}{0.125 + 0.15 + 0.125} = 0.375$$
$$\alpha_{AC} = \alpha_{AD} = 0.3125$$

– –

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im} W h_m\right) = \sigma(\alpha_{AB} W h_B + \alpha_{AC} W h_C + \alpha_{AD} W h_D)$$
$$= \sigma(0.375 * 0.5 * 0.2 + 0.3125 * 0.5 * 0.1 + 0.3125 * 0.5 * 0.1) = 0.06875$$

9. (10pt) Consider a heterogeneous graph given in the following figure. There are three types of nodes in the academic network: Author (A), Paper (P), and Venue (V). Each node has initial features that are the numbers standing next to it (i.e., the initial feature of node 'A_1' is $h_{A_1}^{(0)} = 0.2$). According to HAN model, the weight matrix W is randomly initialized as [0.5]. The feature of node 'i' at layer (k) can be updated as:



$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im}^{\Phi} W h_m\right)$$

Where $\alpha_{im}^{\Phi} = \frac{e_{im}^{\Phi}}{\sum_{k \in N^{\Phi}(i)} e_{ik}^{\Phi}}$ and

$$e_{im}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_i^{\Phi}, W h_m^{\Phi}\right)\right)$$

σ is a ReLU function $\mathrm{ReLU}(x) = \max(0, x)$.

a) List all the meta-path PAP and PVP. Calculate the attention coefficients $\alpha_{P_1P_2}^{\Phi}, \alpha_{P_1P_3}^{\Phi}, and\ \alpha_{P_2P_3}^{\Phi}$.

b) Calculate the feature of node 'P_1' at $k = 1$.

Ans:

List meta-path:

- PAP: $P_1A_1P_2, P_2A_2P_3$.
- PVP: $P_1V_1P_3$.

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha_{im}^{\Phi} W h_m\right)$$

Where $\alpha_{im}^{\Phi} = \frac{e_{im}^{\Phi}}{\sum_{k \in N^{\Phi}(i)} e_{ik}^{\Phi}}$ and $e_{im}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_i^{\Phi}, W h_m^{\Phi}\right)\right)$

PAP:

$$e_{P_1P_2}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_{P_1}, W h_{P_2}\right)\right) = \sigma\left(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.1)\right)$$
$$= \sigma\left(\mathrm{MEAN}(0.2, 0.05)\right) = 0.125$$
$$e_{P_2P_3}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_{P_2}, W h_{P_3}\right)\right) = \sigma\left(\mathrm{MEAN}(0.5 * 0.1, 0.5 * 0.4)\right)$$
$$= \sigma\left(\mathrm{MEAN}(0.05, 0.2)\right) = 0.125$$

$$e_{P_1P_3}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_{P_1}, W h_{P_3}\right)\right) = \sigma\left(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.4)\right)$$
$$= \sigma\left(\mathrm{MEAN}(0.2, 0.2)\right) = 0.2$$

$$e_{P_1A_1}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_{P_1}, W h_{A_1}\right)\right) = \sigma\left(\mathrm{MEAN}(0.5 * 0.4, 0.5 * 0.2)\right)$$
$$= \sigma\left(\mathrm{MEAN}(0.2, 0.1)\right) = 0.15$$

$$e_{A_1P_2}^{\Phi} = \sigma\left(\mathrm{MEAN}\left(W h_{A_1}, W h_{P_2}\right)\right) = \sigma\left(\mathrm{MEAN}(0.5 * 0.2, 0.5 * 0.1)\right)$$
$$= \sigma\left(\mathrm{MEAN}(0.1, 0.05)\right) = 0.075$$

_ _

$$\alpha^{\Phi}_{P_1P_2} = \frac{e^{\Phi}_{P_1P_2}}{e^{\Phi}_{P_1P_2} + e^{\Phi}_{P_1P_3}} = \frac{0.125}{0.125 + 0.2} = \frac{5}{13} \approx 0.3846$$

$$\alpha^{\Phi}_{P_1P_3} = \frac{8}{13} \approx 0.6154$$

$$\alpha^{\Phi}_{P_2P_3} = 0.125$$

$$\alpha^{\Phi}_{P_1A_1} = \frac{e^{\Phi}_{P_1A_1}}{e^{\Phi}_{P_1A_1}} = 1 = \alpha^{\Phi}_{A_1P_2} = \alpha^{\Phi}_{P_1V_1} = \alpha^{\Phi}_{V_1P_3}$$

$$- -$$

PAP: $P_1A_1P_2$

$$\alpha^{\Phi}_{P_1A_1} = \frac{e^{\Phi}_{P_1A_1}}{e^{\Phi}_{P_1P_2} + e^{\Phi}_{P_1P_3}}$$

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha^{\Phi}_{im} W h_m\right) = \sigma\Big(\big(\alpha^{\Phi}_{P_1A_1} W h_{A_1}\big) + \big(\alpha^{\Phi}_{A_1P_2} W h_{P_2}\big)\Big)$$

$$= \sigma\big((1 * 0.5 * 0.2) + (1 * 0.5 * 0.1\big) = 0.015$$

PVP: $P_1V_1P_3$

$$h_i^{(k)} = \sigma\left(\sum_{m \in N(i)} \alpha^{\Phi}_{im} W h_m\right) = \sigma\Big(\big(\alpha^{\Phi}_{P_1V_1} W h_{V_1}\big) + \big(\alpha^{\Phi}_{V_1P_3} W h_{P_3}\big)\Big) =$$

$$= \sigma\big((1 * 0.5 * 0.2) + (1 * 0.5 * 0.4\big) = 0.3$$