Denis Duval
denis.duval13@gmail.com

Lab session # 5
ALTEGRAD 2022

12/06/22

1 Question 1

The architecture can be naturally extended to directed or/and weighted graphs. For graph that are directed, we have to take into account that we can't always go from u to v and from v to u. So during a random walk if we're at a node u we should take a neighbour v randomly such that (u, v) is and edge.
If furthermore the graph is weighted, then the neighbour shouldn't be taken uniformly randomly but with a certain discrete distribution that models the weights (if the weight is bigger w.r.t. to the other weights, there should more chance to move to the corresponding node). For example we could say that the probability of taking edge (u, v) is $\frac{weight(u,v)}{\sum_{(u,w) edge} weight(u,w)}$.

2 Question 2

If the graph is just nodes with no edges then the architecture is a basic fully-connected neural network. In fact there is no message passing so the graph behaves like a bunch of independent features. More precisely, every node has degree 0 so the adjacency matrix A is the zero matrix and $\tilde{A}, \tilde{D}$ are the identity matrix. Hence $\hat{A}$ is the identity matrix so it can be ignored in all layers.

3 Question 3

The receptive field is the number of message passing layers. It can be proven by induction. After 0 message passing layers, the aggregate feature only depends on the node's feature. For some $i \geq 0$, let's assume that after i message passing layers for all nodes the aggregate feature associated with it only depends on nodes that are at distance at most i of it. Then if a new message passing layer is performed then a node receives information from all its neighbours, themselves depending on nodes at distance at most i. So a node will depend only the nodes that are distance at most $i + 1$.

4 Question 4

For the two graphs we have the feature matrix $X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

For the graph K_4:

- the adjency matrix is $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$. The normalized one is $\frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$.
- the product $\tilde{A}XW^0$ is $\frac{1}{15}\begin{bmatrix} -16 & 10 \\ -16 & 10 \\ -16 & 10 \\ -16 & 10 \end{bmatrix}$. The output $Z^0 = \frac{1}{3}\begin{matrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \\ 0 & 2 \end{matrix}$.
- the product $\tilde{A}Z^0W^1$ is $\frac{1}{45}\begin{bmatrix} -16 & 24 & 20 \\ -16 & 24 & 20 \\ -16 & 24 & 20 \\ -16 & 24 & 20 \end{bmatrix}$. The output $Z^1 = \begin{bmatrix} 0 & 0.53 & 0.44 \\ 0 & 0.53 & 0.44 \\ 0 & 0.53 & 0.44 \\ 0 & 0.53 & 0.44 \end{bmatrix}$

For the graph S_4:

- the adjency matrix is $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. The normalized one is $\frac{1}{\sqrt{3}}\begin{bmatrix} \frac{1}{\sqrt{3}} & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$.

- the product $\tilde{A}XW^0$ is $\begin{bmatrix} -1.65 & 1.03 \\ -0.46 & 0.29 \\ -0.46 & 0.29 \\ -0.46 & 0.29 \end{bmatrix}$. The output $Z^0 = \begin{matrix} 0 & 1.03 \\ 0 & 0.29 \\ 0 & 0.29 \\ 0 & 0.29 \end{matrix}$.
- the product $\tilde{A}Z^0W^1$ is $\begin{bmatrix} -0.34 & 0.51 & 0.42 \\ -0.24 & 0.36 & 0.30 \\ -0.24 & 0.36 & 0.30 \\ -0.24 & 0.36 & 0.30 \end{bmatrix}$. The output $Z^1 = \begin{bmatrix} 0 & 0.51 & 0.42 \\ 0 & 0.36 & 0.30 \\ 0 & 0.36 & 0.30 \\ 0 & 0.36 & 0.30 \end{bmatrix}$.

In the first graph we see that the output vector for each is the same. In the second they're all the same for the degree 1 nodes. In general, the nodes that have the same neighborhood in the graph will have the same output vector (same degree, same labels of neighbors, which themselves have the same degrees etc...).