GALAGAIN Calvin
calvin.galagain@gmail.com


# 1 Question 1

**For directed graphs :**
- Generate directed random walks
- Change the underlying gensim Word2Vec module, so that it only considers the words to the left of the current word. That means that for each node, we look at the neighbors on each side.
- We can also set the probabilities to 0 for the random walk when the edge is not in the right direction.

**For weighted graphs :**
- We must first make a more general version of the random walk
- We create a new propability function to replace each weight if i different j : $p_{ij} = \frac{w_{ij}}{\sum w_{ij}}$

# 2 Question 2

$\hat{A} = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$
An architecture with only nodes and without edges implies that $\tilde{A} = \tilde{D}^{-\frac{1}{2}} = \mathbb{I}_n$. Which leads to $\hat{A} = \mathbb{I}_n$.

$\hat{Y} = softmax(Z^1W^2) = softmax(XW^0W^1W^2)$ is an MLP classifier with 3 layers.

# 3 Question 3

For a GCN, the receptive field is equal to the number of network layers. For each layer we multiply by $\hat{A}$ the previous hidden state which gives us that we have second order neighbours because there are 2 aggregations of neighbor values for each node (in the last layer there is no matrix A).

# 4 Question 4

**For $K_4$ :**

$$\hat{A} = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}} = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$Z^0 = ReLU(\hat{A}XW^0) = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -0.8 & 0.5 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix}$$

$$Z^1 = ReLU(\hat{A}Z^0W^1) = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0.1 & 0.3 & -0.05 \\ -0.04 & 0.6 & 0.5 \end{bmatrix} = \begin{bmatrix} 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \end{bmatrix}$$

**For $S_4$ :**

$$\hat{A} = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}} = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0.5 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 1 \end{bmatrix}$$

$$Z^0 = ReLU(\hat{A}XW^0) = \frac{1}{2} \begin{bmatrix} 0.5 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -0.8 & 0.5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & \frac{1}{2} + \frac{3}{\sqrt{2}} \\ 0 & 1 + \frac{1}{\sqrt{2}} \\ 0 & 1 + \frac{1}{\sqrt{2}} \\ 0 & 1 + \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$Z^1 = ReLU(\hat{A}Z^0W^1) = \frac{1}{2}\begin{bmatrix} 0.5 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 1 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 0 & \frac{1}{2}+\frac{3}{\sqrt{2}} \\ 0 & 1+\frac{1}{\sqrt{2}} \\ 0 & 1+\frac{1}{\sqrt{2}} \\ 0 & 1+\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0.1 & 0.3 & -0.05 \\ -0.04 & 0.6 & 0.5 \end{bmatrix} = \begin{bmatrix} 0 & 0.37 & 0.31 \\ 0 & 0.27 & 0.22 \\ 0 & 0.27 & 0.22 \\ 0 & 0.27 & 0.22 \end{bmatrix}$$

The embedding is similar when the neighborhood is close. To change each vector, it would be necessary to put a uniform distribution for each feature vector.