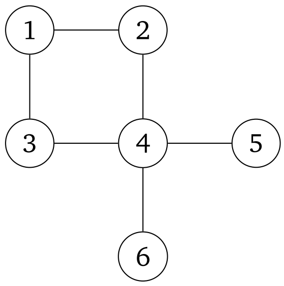
1 Question 1

The graph we consider is the following:



Since $\mathcal{N}(v_5) = \mathcal{N}(v_6) = \{v_4\}$, we know that $\mathbf{z}_5^{(1)} = \underbrace{\alpha_{5,4}^{(1)}}_{=1} \mathbf{W}^{(1)}\mathbf{z}_4^{(0)}$ and $\mathbf{z}_6^{(1)} = \underbrace{\alpha_{6,4}^{(1)}}_{=1} \mathbf{W}^{(1)}\mathbf{z}_4^{(0)}$, so $\mathbf{z}_5^{(1)} \overset{(*_1)}{=} \mathbf{z}_6^{(1)}$.

And by assumption, $\mathbf{z}_2^{(1)} \overset{(*_2)}{=} \mathbf{z}_6^{(1)}$, and $\mathbf{z}_3^{(1)} \overset{(*_3)}{=} \mathbf{z}_5^{(1)}$.
Combining $(*_1), (*_2), (*_3)$ yields $\mathbf{z}_2^{(1)} = \mathbf{z}_3^{(1)} = \mathbf{z}_5^{(1)} = \mathbf{z}_6^{(1)}$.

Then $\mathbf{z}_1^{(2)}$ is a weighted average (with weights summing up to 1) of $\mathbf{W}^{(2)}\mathbf{z}_2^{(1)}$ and $\mathbf{W}^{(2)}\mathbf{z}_3^{(1)}$, and $\mathbf{z}_4^{(2)}$ is a weighted average (with weights summing up to 1) of $\mathbf{W}^{(2)}\mathbf{z}_2^{(1)}, \mathbf{W}^{(2)}\mathbf{z}_3^{(1)}, \mathbf{W}^{(2)}\mathbf{z}_5^{(1)}$ and $\mathbf{W}^{(2)}\mathbf{z}_6^{(1)}$.
Since all these vectors are equal, and the weights sum up to 1, we have $\mathbf{z}_1^{(2)} = \mathbf{W}^{(2)}\mathbf{z}_2^{(1)}$ and $\mathbf{z}_4^{(2)} = \mathbf{W}^{(2)}\mathbf{z}_2^{(1)}$, so

$$\boxed{\mathbf{z}_1^{(2)} = \mathbf{z}_4^{(2)}}.$$

2 Question 2

If all nodes are annotated with identical features x, then for each i, $z_i^{(1)}$ is the mean of identical vectors $\mathbf{W}^{(1)}x$, and is then equal to $\mathbf{W}^{(1)}x$.
So all $z_i^{(1)}$ are equal. And it is then clear by induction that for each t, all vectors $z_i^{(t)}$ are equal.

So the message-passing layers are not useful in this case, and in the end the fully-connected layer predicts identical labels for all nodes. So obviously, the model will not achieve a high accuracy.

3 Question 3

The representations of the three graphs are given in the table below.

Readout function	$\mathbf{z}_{G_1}$	$\mathbf{z}_{G_2}$	$\mathbf{z}_{G_3}$
Sum	$[2.9, 2.3, 1.9]$	$[3.4, 1.9, 4.3]$	$[1.8, 1.2, 1.6]$
Mean	$[0.97, 0.77, 0.63]$	$[0.85, 0.48, 1.08]$	$[0.9, 0.6, 0.8]$
Max	$[2.2, 1.8, 1.5]$	$[2.2, 1.8, 1.5]$	$[2.2, 1.8, 1.5]$

We see that:

- The max function leads to identical representations for the three graphs.
- The mean function leads to different representations, but that are close to each other.
- The sum function leads to different representations, which are further away than with the mean.

Therefore, on this example, the **sum function** seems better.

4 Question 4

The adjacency matrix of C_n is $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 1 & 0 & \cdots & \cdots & 1 & 0 \end{pmatrix}$.

With all features equal to one, when computing $\tilde{\mathbf{A}}\mathbf{X}$ (in the computation of $\mathbf{Z}^{(1)}$), all coefficients are equal, so all rows of $\mathbf{Z}^{(1)}$ are equal. Then all rows of $\mathbf{Z}^{(1)}\mathbf{W}^{(2)}$ are equal, and using once again the structure of $\tilde{\mathbf{A}}$, all rows of $\mathbf{Z}^{(2)} = \tilde{\mathbf{A}}\mathbf{Z}^{(1)}\mathbf{W}^{(2)}$ are equal.

So the only difference lies in the computation of $\mathbf{z}_G$:

- If the readout function is the sum function, then $\mathbf{z}_{C_8} = 2\mathbf{z}_{C_4}$;
- If the readout function is the mean function, then $\mathbf{z}_{C_8} = \mathbf{z}_{C_4}$;
- If the readout function is the ReLU function, then $\mathbf{z}_{C_8} = \mathbf{z}_{C_4}$.

So depending on the readout function, we have $\boxed{\mathbf{z}_{G_2} = 2\,\mathbf{z}_{G_1} \text{ or } \mathbf{z}_{G_2} = \mathbf{z}_{G_1}}$.