1 Question 1

Embedding

$$x_i \longrightarrow \frac{x_i}{10} = y_i$$

Fully Connected

$$y_i \longrightarrow \frac{1}{2}\ln(y_i+1) - \frac{1}{2}\ln(1-y_i) \cdot z_i$$

Tanh Activation

$$y_i \longrightarrow \tanh(y_i)$$

Fully Connected

$$y_i \longrightarrow 10 \cdot y_i$$

2 Question 2

Let $W = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then:
- For X_1:

$$\phi([1.2, -0.7]) = \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1.2 \\ -0.7 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 1.2 \\ 1.7 \end{bmatrix},$$

$$\phi([-0.8, 0.5]) = \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -0.8 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} -0.3 \\ 0.5 \end{bmatrix}.$$

Sum:

$$s_1 = \phi([1.2, -0.7]) + \phi([-0.8, 0.5]) = \begin{bmatrix} 1.2 \\ 1.7 \end{bmatrix} + \begin{bmatrix} -0.3 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 2.2 \end{bmatrix}.$$

- For X_2:

$$\phi([0.2, -0.3]) = \begin{bmatrix} 0.2 \\ 1.3 \end{bmatrix}, \quad \phi([0.2, 0.1]) = \begin{bmatrix} 0.2 \\ 0.9 \end{bmatrix}.$$

Sum:

$$s_2 = \phi([0.2, -0.3]) + \phi([0.2, 0.1]) = \begin{bmatrix} 0.2 \\ 1.3 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0.9 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 2.2 \end{bmatrix}.$$

Let $\rho(s) = Ws + b$ with

$$W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then:

$$\rho(s_1) = s_1 = \begin{bmatrix} 0.9 \\ 2.2 \end{bmatrix}, \quad \rho(s_2) = s_2 = \begin{bmatrix} 0.4 \\ 2.2 \end{bmatrix}.$$

Since $\rho(s_1) \neq \rho(s_2)$, the two sets are embedded into distinct vectors, as required.

3 Task 7

As we can see on the plot, DeepSets perfectly managed to learn tu sum representation of the sets, while the LSTM lacks generalisation and already struggle on small cardinality.

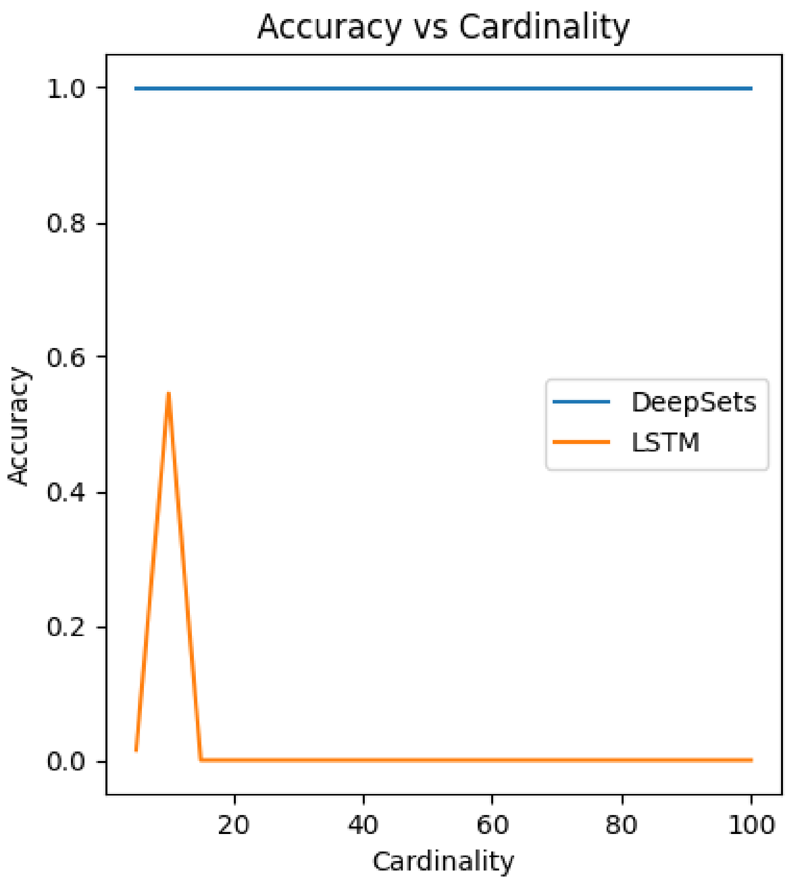


Figure 1: Accuracy against maximum cardinal.

4 Question 3

DeepSets can serve as a submodule in a GNN, if it used as a pooling layer to aggregate node embeddings into a graph-level representation. DeepSets is a more weighted alternative to traditional pooling methods (e.g., sum, mean, max) and ensures the graph representation is invariant to node ordering.

5 Question 4

For an Erdős-Rényi random graph $G(n, p)$, the total number of possible edges in a graph with n nodes is $\binom{n}{2}$. Here, this number is 105.

$$\mathbb{E}[E] = \binom{n}{2} \cdot p$$

1. For $p = 0.2$:

$$\mathbb{E}[E] = 105 \cdot 0.2 = 21$$

2. For $p = 0.4$:

$$\mathbb{E}[E] = 105 \cdot 0.4 = 42$$

The variance of the number of edges is:

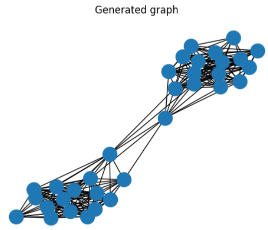
$$\text{Var}(E) = \binom{n}{2} \cdot p \cdot (1 - p)$$

1. For $p = 0.2$:

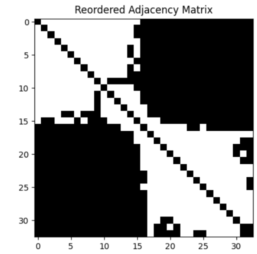
$$\text{Var}(E) = 105 \cdot 0.2 \cdot (1 - 0.2) = 105 \cdot 0.2 \cdot 0.8 = 16.8$$

2. For $p = 0.4$:

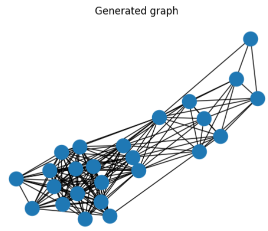
$$\text{Var}(E) = 105 \cdot 0.4 \cdot (1 - 0.4) = 105 \cdot 0.4 \cdot 0.6 = 25.2$$



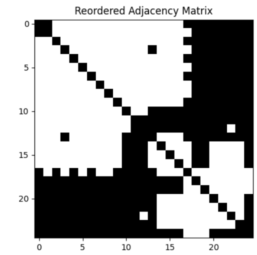
(a) Graph 1 visualization.



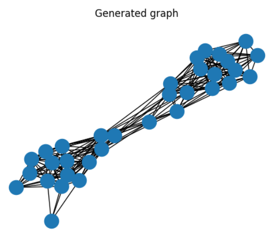
(b) Matrix associated with Graph 1.



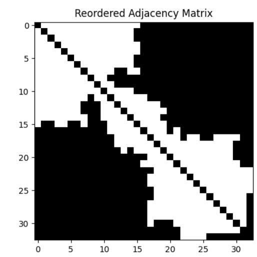
(c) Graph 2 visualization.



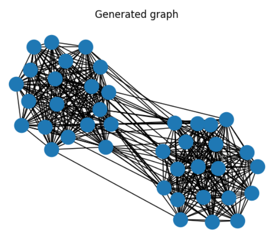
(d) Matrix associated with Graph 2.



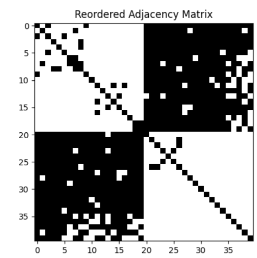
(e) Graph 3 visualization.



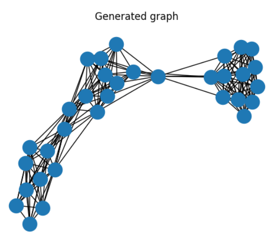
(f) Matrix associated with Graph 3.



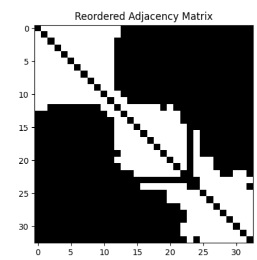
(g) Graph 4 visualization.



(h) Matrix associated with Graph 4.



(i) Graph 5 visualization.



(j) Matrix associated with Graph 5.

Figure 2: Task 11 : Graphs and their associated matrices.