Mathis EMBIT
mathis.embit@gmail.com

Lab session # 7
ALTEGRAD 2023

12/05/23

1 Question 1

LSTMs, by design, consider the sequential order of input data, making them sensitive to permutations. LSTMs might not be the best choice for tasks involving sets.

2 Question 2

- The GNN consists of two fully-connected layers knowned as "message passing layers" followed by a sum readout function and finally, by two fully-connected layers again. The DeepSets consists of a first MLP, with possibly more than 2 layers, followed by a sum, and a final MLP, with possibly more than 2 layers. Thus the main difference is that in DeepSets we are allowed to have as many hidden layers as we want in the MLP whereas the GNN is constrained to have 2 f-c layers before and after the sum.
- Considering a graph without edge, its adjacency matrix A would be 0 everywhere and $\tilde{A} = I$. That means the order relation encoded in A is no longer used by the GNN. Therefore the GNN is permutation invariant for **nodes within the same graph** and is also still invariant to permutation between graphs themselves. But the GNN is not invariant to permutation of **nodes belonging to different graphs**.

3 Question 3

1. If r = 2,
   - $P = \begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$ is a probability matrix from which homophilic graph can be sampled.
   - $P' = \begin{bmatrix} 0.05 & 0.9 \\ 0.9 & 0.05 \end{bmatrix}$ is a probability matrix from which heterophilic graph can be sampled.
2. In order to count the number of edges between different communities in the graph we need to check each possible pair of different communities and to check each possible pair of vertices (a, b), a and b being in different communities.

   Let the partition of V, $\{V_1, \dots, V_4\}$ be the 4 communities of vertices of V.

   $\forall (a, b) \in V_i \times V_j, i \neq j, \mathbb{P}((a, b) \in V) = 0.05$, therefore

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{2}\sum_{i=1}^{4}\sum_{\substack{j=1\\ j\neq i}}^{4}\sum_{a\in V_i}\sum_{b\in V_j}\mathbb{1}_{(a,b)\in V}\right] &= \frac{1}{2}\sum_{i=1}^{4}\sum_{\substack{j=1\\ j\neq i}}^{4}\sum_{a\in V_i}\sum_{b\in V_j}\mathbb{P}((a,b)\in V) \\
&= \frac{1}{2} \times 4 \times 3 \times 5 \times 5 \times 0.05 \\
&= \boxed{7.5}
\end{aligned}
$$

   We need to divide by 2 because we are treating each pair of communities twice.

4 Question 4

In a continuous setting, a loss that would be more suitable is the MSE loss function:

$$
\mathcal{L}_{\text{MSE}} = \frac{1}{n^2}\sum_{i,j=1}^{n}(A_{ij} - \hat{A}_{ij})^2
$$