YOU Borachhun
borachhunyou@gmail.com

Lab session # 7
ALTEGRAD 2023

12/04/23

1 Question 1

LSTM is not a permutation invariant model because the order of the model's input sequence is important. The computation at each unit in the network takes into account previous hidden states resulting from previous computations with previous elements in the sequence. Therefore, for the case of sets, inputting the same set in different orders to LSTM would get different results. Thus, this architecture should not be used on sets.

2 Question 2

The architectural difference between GNN (in Lab 6) and DeepSets is that with GNN the representation of each node in a graph is transformed from the node itself along with its neighbors (by multiplying with the adjacency matrix), while with DeepSets each element in a set is transformed only from the element itself. That said, if we consider using GNN with a graph without edges, there would be no difference between the two architecture since the representation of each node is transformed only from the node itself (multiplying with a diagonal adjacency matrix).

3 Question 3

1. For a stochastic block model graph of n total nodes with $r = 2$ blocks of equal number of nodes:
   - Let p be the probability of an edge connecting nodes in the same block, then the expected number of edges in a block is: $\binom{n/2}{2} \times p = \frac{n(n-2)}{8} p$
   - Let q be the probability of an edge connecting nodes in different blocks, then the expected number of edges connecting different blocks is: $\left(\frac{n}{2}\right)^2 \times q = \frac{n^2}{4} q$

   With the edge probability matrix $\mathbf{P} = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$:

   - For the case of homophilic graph:

   $$\begin{aligned} \frac{n(n-2)}{8} p &\gg \frac{n^2}{4} q \\ p &\gg \frac{2n}{n-2} q \end{aligned}$$

   - For the case of heterophilic graph:

   $$\begin{aligned} \frac{n(n-2)}{8} p &\ll \frac{n^2}{4} q \\ p &\ll \frac{2n}{n-2} q \end{aligned}$$

2. For a stochastic block model graph of n nodes with r blocks and edge probability matrix $\mathbf{P}$ with diagonal elements equal p and off-diagonal elements equal q:
   - The maximum total number of edges in the graph: $\binom{n}{2}$
   - The maximum number of edges within a block: $\binom{n/r}{2}$

   Therefore, the maximum number of edges between nodes in different blocks equals $\binom{n}{2} - r\binom{n/r}{2}$, and the expected number of edges between nodes in different blocks equals

   $$\left(\binom{n}{2} - r\binom{n/r}{2}\right) \times q$$

   With $n = 20, r = 4, q = 0.05$, the expectation is $\left(\binom{20}{2} - 4 \times \binom{20/4}{2}\right) \times 0.05 = 7.5$ edges.

4 Question 4

For weighted graphs, whose entries of the adjacency matrix are not constrained to 0 or 1, one suitable loss function for the reconstruction would be mean squared error (MSE):

$$\mathcal{L} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\mathbf{A}_{ij} - \hat{\mathbf{A}}_{ij} \right)^2$$

which can be used with any number in the adjacency matrix.