Ben KABONGO B.
ben.kabongo_buzangu@ens-paris-saclay.fr

Lab session # 7
ALTEGRAD 2023

11/29/23

1 Question 1

LSTMs are not inherently permutation-invariant models. Permutation invariance means that the output of the model remains the same regardless of the order of the inputs. LSTMs, designed for sequential data, are sensitive to the order of input sequences.
For sets, which are collections of unordered elements, using LSTMs may not be the most natural choice. Sets have no specific order and LSTMs are designed to capture sequential dependencies. Although it is possible to feed an LSTM with a set, it may not fully exploit the inherent properties of sets, and model performance may be sub-optimal.

2 Question 2

2.1 Graph Neural Networks (GNNs) - DeepSets

2.1.1 Graph Neural Networks (GNNs)

- **Graph Structure**: GNNs are designed to operate on graph-structured data.
- **Node and Edge Features**: GNNs typically work with node and edge features.
- **Message Passing**:GNNs often use message-passing mechanisms to aggregate information from neighboring nodes. Each node collects information from its neighbors and updates its own representation based on this aggregated information.
- **Graph-Level Tasks**: GNNs are commonly used for graph-level tasks, where the goal is to make predictions or classifications at the level of the entire graph.

2.1.2 DeepSets

- **Set Structure**: DeepSets are designed to work with sets.
- **Permutation Invariance**: DeepSets architectures are permutation invariant.
- **Aggregation Function**: DeepSets use aggregation functions, such as summation or averaging, to process the entire set at once.
- **Set-Level Tasks**: DeepSets architectures are suitable for set-level tasks, where the goal is to make predictions or classifications at the level of the entire set.

2.2 Difference between Sets and Graphs Without Edges

A set is an unordered collection of distinct elements. The elements in a set have no inherent relationships or connections between them. Sets are defined solely by their elements, and their order is irrelevant.
A graph without edges is essentially a set of nodes. In this case, the nodes represent entities, but there are no explicit relationships (edges) between them. Each node can still have associated features, and the graph is essentially a set of nodes with certain characteristics.

3 Question 3

3.1 Edge Probability Matrices for Stochastic Block Model with r = 2

- **Homophilic Graph**:

$$P_{\text{hom}} = \begin{bmatrix} 0.8 & 0.05 \\ 0.05 & 0.8 \end{bmatrix}$$

- **Heterophilic Graph**:

$$P_{\text{het}} = \begin{bmatrix} 0.05 & 0.8 \\ 0.8 & 0.05 \end{bmatrix}$$

3.2 Expected Number of Edges in the Stochastic Block Model

Let n be the total number of nodes, r the number of blocks and n_i the number of elements in block i. If the probability that there is an edge between two nodes belonging respectively to two different blocks i and j is given by P_{ij} (here $P_{ij} = 0.05$), then the expected number of edges between two distinct blocks is given by :

$$n_i \times n_j \times P_{ij}$$

Each node of block i can be connected to a node of block j with probability P_{ij}.

- **r = 4**: Each block is assumed to contain 5 nodes. The expected number of edges is therefore $5 \times 5 \times 0.05 = 1.25$
- **r = 5**: Each block is assumed to contain 4 nodes. The expected number of edges is therefore $4 \times 4 \times 0.05 = 0.80$

4 Question 4

When dealing with weighted graphs, where the entries of the adjacency matrix can take continuous values rather than being binary, a more suitable loss function for reconstruction would take into account the continuous nature of the weights. One common approach is to use a mean squared error loss (MSE), which measures the average squared difference between the predicted and true weights.
Let's denote the predicted adjacency matrix as $\hat{\mathbf{A}}$ and the true adjacency matrix as $\mathbf{A}$ If $\mathbb{A}$ contains continuous weights, the reconstruction loss can be defined as follows:

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