Romain AMÉDÉE
romain.amedee@telecom-sudparis.eu

1 Question 1

The Erdős-Rényi random model builds a graph $G(n, p)$ by connecting the n nodes randomly with probability p. Therefore, if we consider that self-loops are not allowed, the degree of such a graph follows a binomial distribution. Therefore, the expected degree of a node is, in the general case :

$$E[degree(n, p)] = (n - 1)p$$

In our specific case we have :

- $(n = 25, p = 0.2) \implies E[degree(25, 0.2)] = 4.8$
- $(n = 25, p = 0.4) \implies E[degree(25, 0.4)] = 9.6$

2 Question 2

Fully connected layers are not commonly used as readtout functions instead of the sum or mean operation in graph level GNNs because they necessitate a fixed-size input and ouput which is not necessarily the case when dealing whith graphs (different number of nodes). Furthermore, fully connected layers are not invariant to permutations therefore the order in which the adjacency matrices of the graphs are concatenated matters for the result which is not what we want. For instance, in the exemple given in Figure 1 of the handout, the adjacency matrices are concatenated following this order : A_1, A_2 then A_3. If we were to concatenate the other way around A_3, A_2 then A_1 then the resulting vectors would be different.

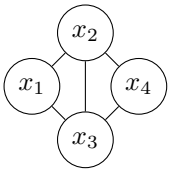
3 Question 3

We observe that :

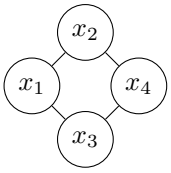
- whenever the aggregator function is set to 'mean', all the lines of the resulting matrices are identical ie. all the graphs have the same representation
- whenever the aggregator function is set to 'sum' the graph representations (lines) are different
- the function used for readout doesn't influence this behaviour

We could explain this behaviour by the fact that when we use the mean as the aggregator, the representation of a node only depends on the mean of its neighbor's representations and not on the individual values, thus resulting in identical graph representations regardless of connectivity. On the other hand, when we use the sum aggregator the representation of a node will depend on the feature vectors of neighboring nodes. This can lead to more diverse representations because even if nodes have similar neighbors, the sum operation might accentuate differences in their features.
The readout function does not have an impact on this behaviour because it is only applied after the aggregator has already computed the representation of each node so it will only affect the overall scale of the graph representations.

4 Question 4



(a)



(b)

Figure 1: Exemple of non isomorphic graphs that cannot be dinstinguished by our GNN model