BENYAMINA Elyas
elyas.benyamina@ens-paris-saclay.fr

Lab session # 6
ALTEGRAD 2023

11/28/23

1 Question 1

Let's consider an Erdos-Renyi random graph $G(n,p)$. Let's consider one node n_i. n_i is connected to another node n_j with probability p. As n_i can be connected to $n-1$ nodes, the average degree (number of nodes it is connected to) is $(n-1)p$. As all nodes plays the same role, the average degree of each node is $(n-1)p$.
For $n = 25$ and $p = 0.4$, we have an expected number of degree of a node of $24 \cdot 0.4 = 9.6$.
For $n = 25$ and $p = 0.2$, we have an expected number of degree of a node of $24 \cdot 0.2 = 4.8$.

2 Question 2

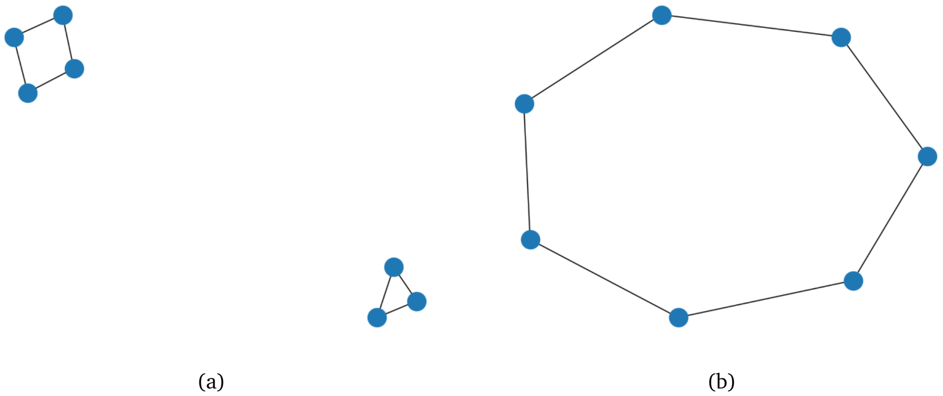
Using a non-linear function such as mean or sum can be a good solution to compare graphs of different sizes in an accurate way. Indeed, as G_1, G_2 and G_3 have different number of nodes, if we were using a linear layer as a readout function, we would have needed to change artificially (by adding an arbitrary value at the end of the vectors for example) the size of the vectors we are working with.
Also, it's better for training, and for use, as we can change the order of the graphs and get the same results.
Using sum or mean allows to capture global information from the graphs, by taking into account information from all the nodes.

3 Question 3

If we use mean readout, we obtain same output for all the cycle graphs. It's then better to use sum readout.
Using mean or sum neighbor aggregation does not seem to affect the results.
We can explain it as the input are all equals to one and therefore we have all the rows of Z^1 which are the same if we look at the operations made by the GNN model (using a sum or mean aggregation produces similar results as all the rows will be equal). Then, if we use average readout, as all the rows are equal, we'll have same rows no matter the value of n, contrary to sum readout, where the value of n changes the output.

4 Question 4

We can use graphs similar to those shown in figure 2. Let's consider the following graphs G_1 and G_2:


(a)

(b)

Figure 1: Graphs G_1 and G_2)

These two graphs are not isomorphic (it is clear as they don't share the same degree distribution), and the GNN model "sum operator-sum agreggation" cannot distinguish them (they have the same output).