YOU Borachhun
borachhunyou@gmail.com

Lab session # 6
ALTEGRAD 2023

11/24/23

1 Question 1

Let $G(n, p)$ be a Erdős-Rényi random graph with n nodes and the edge probability p. Then the expected number of edges in the graph is $\binom{n}{2} \times p = \frac{n(n-1)}{2} \times p$, which corresponds to the total degree of nodes in the graph of $2 \times \frac{n(n-1)}{2} \times p = np(n-1)$. Therefore, each node in the graph has an expected number of degree of $\frac{np(n-1)}{n} = p(n-1)$.

- When $n = 25, p = 0.2$, the expected number of degree of a node equals 4.8.
- When $n = 25, p = 0.4$, the expected number of degree of a node equals 9.6.

2 Question 2

Trainable linear layers are not commonly used as readout functions in graph level GNNs for multiple reasons:

- The number of nodes in each graph are different and thus fixed-size layers cannot be used for multiple graphs.
- Optimizing the parameters of linear layers requires more computing power and time than doing the sum or mean operation, especially with large graphs.
- Graphs have unordered structures, i.e. different orders of nodes or different orders of rows in the feature matrix have the same meaning. However, linear layers have specific orders of weights and therefore feature matrices of the same graph with different row orders give different results.

3 Question 3

We can observe that the resulting vector representations of the 10 cycle graphs of different sizes are the same when the readout function uses the mean operator, for both when the neighborhood aggregation uses the sum and mean operator. In contrast, the vector representations are not the same when the sum operator is used for the readout function.
Firstly, the initial features of the nodes are set to have the same value. Moreover, each node in the cycle graphs has the same number of neighbors and thus the same number of adjacency and degree. These result with the nodes having the same vector representation after the two message passing layers, for both operators of the neighborhood aggregation. Then, when the readout function applies the mean operator over the same node representations, the graph representations are also the same. However, it is not true when the readout function uses the sum operator due to the difference of number of nodes in each graph.

4 Question 4

Figure 1: Graph G_1 (left) and G_2 (right) non-isomorphic that GNN cannot distinguish

Figure 1 shows 2 different graphs that are not isomorphic but cannot be distinguished by GNN. Similar to the previous question, with the features of the nodes initialized to the same value and each node in both graphs has the same number of adjacency, all nodes have the same vector representation after the two message passing layers. Then, when the readout function applies the sum operator over the same node representations, since both graphs have the same number of nodes, the graph representations are also the same.