NAIM Mouna
mouna.naim@ens-paris-saclay.fr

Lab session 2
ALTEGRAD 2024

11/24/24

1 Question 1

- **First connected component: Complete graph**
  A complete graph with $n = 100$ vertices has:

$$\text{Number of edges} = \binom{n}{2} = \frac{n \cdot (n-1)}{2} = \frac{100 \cdot 99}{2} = 4950$$

  The number of triangles is:

$$\text{Number of triangles} = \binom{n}{3} = \frac{n \cdot (n-1) \cdot (n-2)}{6} = \frac{100 \cdot 99 \cdot 98}{6} = 161700$$

- **Second connected component: Complete bipartite graph ($K_{50,50}$)**
  A complete bipartite graph with two sets of 50 vertices each has:

$$\text{Number of edges} = n \cdot n = 50 \cdot 50 = 2500$$

  Since a bipartite graph does not contain any closed triangles, the number of triangles is:

$$\text{Number of triangles} = 0$$

Summary:
The total number of edges in G is:

$$4950 + 2500 = 7450$$

The total number of triangles in G is:

$$161700 + 0 = 161700$$

2 Question 2

We compute the modularity Q using the formula:

$$Q = \sum_{c=1}^{n_c} \left[\frac{l_c}{m} - \left(\frac{d_c}{2m} \right)^2 \right]$$

For the first graph (a):

$$Q_a = \underbrace{\left[\frac{6}{13} - \left(\frac{13}{2 \cdot 13} \right)^2 \right]}_{\text{Cluster 1 (Green)}} + \underbrace{\left[\frac{6}{13} - \left(\frac{13}{2 \cdot 13} \right)^2 \right]}_{\text{Cluster 2 (Blue)}} \approx 0.42$$

For the second graph (b):

$$Q_b = \underbrace{\left[\frac{2}{13} - \left(\frac{11}{2 \cdot 13} \right)^2 \right]}_{\text{Cluster 1 (Green)}} + \underbrace{\left[\frac{4}{13} - \left(\frac{15}{2 \cdot 13} \right)^2 \right]}_{\text{Cluster 2 (Blue)}} \approx -0.05$$

3 Question 3

Let's calculate the shortest path kernel for the pairs (C_4, C_4), (C_4, P_4), and (P_4, P_4).
Let:

$$\phi(P_4) = [3, 2, 0, 0]$$

$$\phi(C_4) = [4, 4, 0, 0]$$

These vectors represent the frequencies of shortest path distances for each graph. Using these vectors, we compute the kernel values as the dot product of the feature vectors.

Let us denote the kernel as k.

—

For (C_4, C_4)**:**
The kernel value is calculated as:

$$k(C_4, C_4) = \langle \phi(C_4), \phi(C_4) \rangle = 4^2 + 4^2 = 16 + 16 = 32.$$

—

For (C_4, P_4)**:**
The kernel value is calculated as:

$$k(C_4, P_4) = \langle \phi(C_4), \phi(P_4) \rangle = (4 \cdot 3) + (4 \cdot 2) = 12 + 8 = 20.$$

—

For (P_4, P_4)**:**
The kernel value is calculated as:

$$k(P_4, P_4) = \langle \phi(P_4), \phi(P_4) \rangle = 3^2 + 2^2 = 9 + 4 = 12.$$

4 Question 4

Let k denote the graphlet kernel. A kernel value $k(G, G') = 0$ means that the two graphs G and G' do not share any common substructures among the predefined graphlets (size-3 subgraphs in this case).

Example: Let us consider a triangle graph G with three nodes $(1, 2, 3)$ and a path graph G', where only adjacent nodes are connected (4-5, 5-6). These graphs do not share any graphlets of size three. This results in a kernel value of $k(G, G') = 0$.