Assignment-4

Q.1 Compute the complete graph G= (V, E) of the project dataset.
Q.2. Construct MST neighborhood graph of 3-round i.e. G' =(V, E').
Q.3 Let A' be the adjacency matrix of G'. Compute the adjacency matrix A'.
Q.4 Find the degree matrix D of G', where each D(i,j)=0, i≠j and

$$D(i,i) = \sum_{j=1}^{j=n} A'(i,j)$$

Q.5 Compute the Laplacian Matrix L= D-A'.
Q.6 Compute first k eigenvectors $v_1, v_2, v_3, \ldots . v_k$ of L, where k is the number of clusters in your dataset.
Q7. Form a n*k matrix U, where each eigenvector v_i; $1 \leq i \leq k$ is stacked as a column in U. Treat each row u_i $1 \leq i \leq n$ of U as a data point.
Q.8 Apply k-means and MST based clustering algorithm to partition the points $u_{i=1,2,3,\ldots n}$ into k clusters.
Q.9 Compute the accuracy and execution time of each algorithm.

Reference:

[1] Jothi, R., Mohanty, S. K., & Ojha, A. (2016). Functional grouping of similar genes using eigenanalysis on minimum spanning tree based neighborhood graphs. *Computers in biology and medicine*, *71*, 135-148.