Graph Theory

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, $i \neq j$ 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.

Experimental Result

Silhouette Score:

Cluster Quality	K Means	MST Based Clustering
silhouette_score	0.9891455951665729	-0.06566420224954977

Time:

- ➢ Pre-processing: 2.1 s
- ➢ Graph creation: 67.4 s
- ➢ Construct MST Neighbourhood graph: 0.6 s
- ➢ Finding Eigen-vectors and U matrix: 1.1 s
- ➢ kMeans: 0.1 s
- ➢ MST based clustering: 60.8 s

Execution Time (in secs)	K Means	MST Based Clustering
Time	71.2 s	131.9 s

Observations:

Silhouette score:

K-Means > MST Based Clustering.

- ➢ Therefore it says K means has produced better cluster quality than MST-based clustering

Time:

MST Based Clustering > K Means

- ➢ K-means clustering algorithm is relatively fast compared to the MST based algorithm and also produces better clusters.
