Graph Mining CSF426
Lab 3
Time: 5 PM – 7 PM
Date: August 31, 2024
Instructor IC- Vinti Agarwal

Instructions: All questions need to be answered. Use jupyter notebook only to attempt the questions. No other IDE (VS code, Pycharm, Google collab allowed). For theoretical questions, you can type answers in the jupyter notebook itself.

Objective: The objective of this lab session is to show the practical difference between linear dimensionality reduction technique (PCA) and non-linear dimensionality reduction technique (Laplacian Eigenmaps).

Q1. Implement the approach of dimensionality reduction using Laplacian Eigenmaps by following the below mentioned steps:

a. Construct the adjacency graph by following the approach of k-nearest neighbours. (k=10). This will result in a connected graph where each node is connected to its k-nearest neighbours.
b. Compute the weight matrix using the Heat Kernel formula:

$$W_{ij} = e^{-\frac{\|X_i - X_j\|^2}{2.\sigma^2}} \quad \text{if i and j are connected, else 0}$$

$$\sigma = 1.0$$

c. Compute unnormalized Laplacian matrix using

L = D – W ,

where D is degree matrix of weight matrix W.
d. Calculate the first m eigenvectors (corresponding to first m smallest eigenvalues).
e. The obtained eigenvectors are the transformed data in reduced m dimensions.

Q2. Also, implement the approach of PCA to reduce the dimensions of data. Plot the transformed data/embeddings obtained from both PCA and Laplacian Eigenmaps.

Q3. Do the following operations:

a. Replace the unnormalized Laplacian matrix with Random Walk Laplacian matrix and Normalized Laplacian matrix.

   Random walk Laplacian matrix: $L_{rw} = D^{-1}.L$
   Normalized Laplacian matrix: $L_{sym} = D^{-1/2}.L.D^{-1/2}$.

b. Replace the heat-map weight matrix with simple 1/0 matrix and observe the results.
c. Observe the effect of changing the number of nearest neighbours to [20,30,40,50].
d. Observe the effect of changing sigma (σ) in heat-map weight matrix to [0.1,0.5,2,5]