

Graph Mining CSF426
Lab session 16
Time: 3pm -5pm
Date: Oct 6, 2022
Instructor IC – Vinti Agarwal

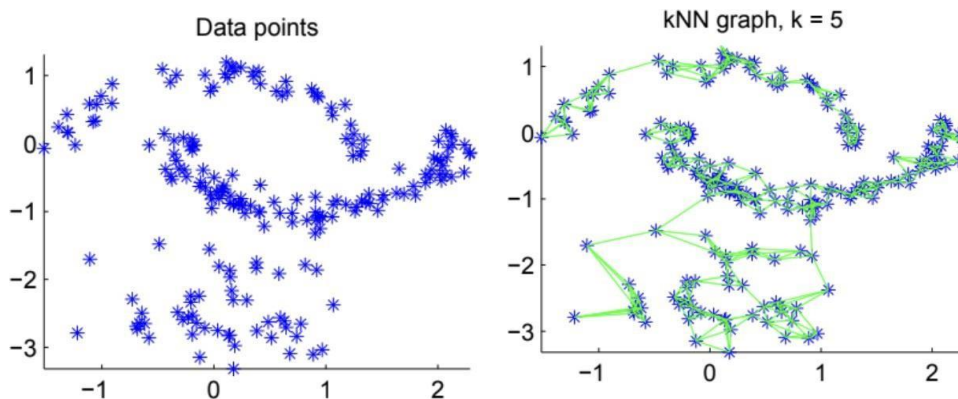
Instructions: All questions need to be answered. **You are required to write programs in jupyter notebook and submit .ipynb and pdf conversion both on canvas.** For theoretical questions, you can type answers in the jupyter notebook itself. There is no need to create a separate text file. You are free to choose any library package (unless you are explicitly asked to implement a module) in python for the implementation of the programs. Class notes support is allowed during lab sessions.

Note: The data is provided to you in the form of a jupyter file. Extra 15 minutes will be provided to submit your solutions on canvas.

[Total Marks =10]

- a) Create a similarity graph based on *k-nearest neighbor* approach (a reference paper is attached, see page no. 03) and display its variation on different values of k between 2 to 5. (An example data and corresponding 5-NN graph are shown below on some other dataset). You are free to choose your own similarity/distance metric (preferably depending on the domain the data comes from), and no general advice will be given. **[4 marks]**

Note: (For drawing graphs in 2-d plane depicting edges between points, students are allowed to seek help from TAs).



- b) Construct the weighted adjacency matrices for the k -NN similarity graph on varying values of k i.e. 2, 3, 4 and 5 and plot the sparsity level of graph (in %) vs k . **[3 + 3 =6 marks]**