CSE 454 Data Mining

Assignment 2

Deadline: 07.01.2024

a) Explain chamelon clustering algorithm. Use figures and formulas. Give the pseudo code. Explain each part in the pseudo code. (25 pts)

b) Give advantages and disadvantages of the model. Compare the algorithm with the other clustering algorithms in terms of time, perfomance and memory consumption. (10 pts)

c) What is the time complexity of it? Explain your reasons (for each step of the algorithm, I expect you to show the time complexity of that step. Just writing the final time complexity is not enough as an answer. (5pts)

d) Implement the algorithm using any programming language. Copy paste the code addition to uploading Teams. (30 pts)

e) Show clustering results for 4 randomly chosen parameter values, for each parameter. (15 pts) You must use worms dataset mentioned on the web site https://cs.joensuu.fi/sipu/datasets/

f) Show the results using another distance other than euclidean distance. (15 pts)