CSIT5210 Data Mining and Knowledge Discovery (Fall Semester 2021) Exercise 3 Clustering

Q1 Agglomerative Approach

There are four two-dimensional points:

x₁ (1, 2), x₂ (2, 2), x₃ (5, 2), x₄ (6, 1).

- (a) Assume that we adopt the Euclidean distance metric as the distance between any two points. Please write a matrix where its entries correspond to the pairwise distances between any two points.
- (b) Please use the agglomerative approach to group these points with distance median linkage. Draw the corresponding dendrogram for the clustering. You are required to specify the distance metric in the dendrogram.

Q2 Divisive Approach

There are the following four records with three binary attributes, namely A, B and C.

	A	В	С
1	1	1	0
2	1	1	0
3	1	0	1
4	0	0	1

Please use monothetic approach to perform hierarchical clustering over these records. If there is any tie for choosing the attributes, we choose the attributes in the order: A, B and C. Draw the dendrogram (without specifying the distance metric).