Data Mining - EX9

Deadline: Friday, Dey 14, 1403 - January 03, 2025

Question 1: Clustering

Suppose that we have the following data (one variable):

i	1	2	3	4	5	6	7	8	9	10	11	12
X	12	9	32	5	16	0	34	2	46	15	5	46

- a. Use single linkage to identify the clusters.
- b. Use complete linkage to identify the clusters.

Now suppose that we have the following data (2 variable):

X	1	1	3	3	4	4	5	5	2	3
Y	1	3	5	2	4	5	4	5	3	1

- c. Identify the cluster by applying the k-means algorithm, with $k = 2$. Try using initial cluster centers as far apart as possible.

Question 2: Neural Network

- Suppose that you need to prepare the data in Table 6.10 for a neural network algorithm.
  Define the indicatorvariables forthe occupationattribute.
- Describe thebenefits anddrawbacks ofusinglarge orsmallvalues forthe learning rate.

Question 3: Applied

- Consider the Boston dataset in Exercise EX7. Take the variable "medv" as the target variable and implement a simple neural network with a hidden layer of 3 nodes on it.
- Consider the Boston dataset. Use all its numeric variables and transform it into 3 clusters using the k-means algorithm.