


National University of Computer and Emerging Sciences, Lahore Campus

	Course:	Information Retrieval	Course Code:	CS317
	Program:	BS(Computer Science)	Semester:	Fall 2019
	Duration:	20 Minutes	Total Marks:	10
	Paper Date:	3-Dec-19	Weight	4%
	Section:	B	Page(s):	2
	Exam:	Quiz 5	Roll No:	

Question 1:

The table below is a distance matrix for 6 objects.

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	0.84	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	0.93	0.20	0.67	0

- Apply HAC algorithm and show intermediate results. [4]
- Show the final result of hierarchical clustering with single link by drawing a dendrogram. [2 Marks]

Solution:

First Iteration: Doc A and B have smallest distance so that are closest clusters. Merge A and B.

Since it is single link similarity so similarity of doc C with new merged cluster A-B is min of the distance of (0.51 and 0.25).

	A-B	C	D	E	F
A-B	0				
C	0.25	0			
D	0.16	0.14	0		
E	0.28	0.7	0.45	0	
F	0.34	0.93	0.2	0.67	0

Second Iteration

C and D have min distance so they will be merged

	A-B	C-D	E	F
A-B	0			
C-D	0.16	0		
E	0.28	0.45	0	
F	0.34	0.2	0.67	0

Third Iteration

A-B and C-D have min distance so they will be merged

	A-B-C-D	E	F
A-B-C-D	0		
E	0.28	0	
F	0.2	0.67	0

Fourth Iteration

A-B-C-D and F have min distance so they will be merged

Question 2:

- a) What is difference between pseudo relevance feedback and relevance feedback? [2 Marks]

Name _____
Section _____

Roll No _____

b) Why do we need pseudo relevance feedback? [2 Marks]