

2) Given dissimilarity matrix:

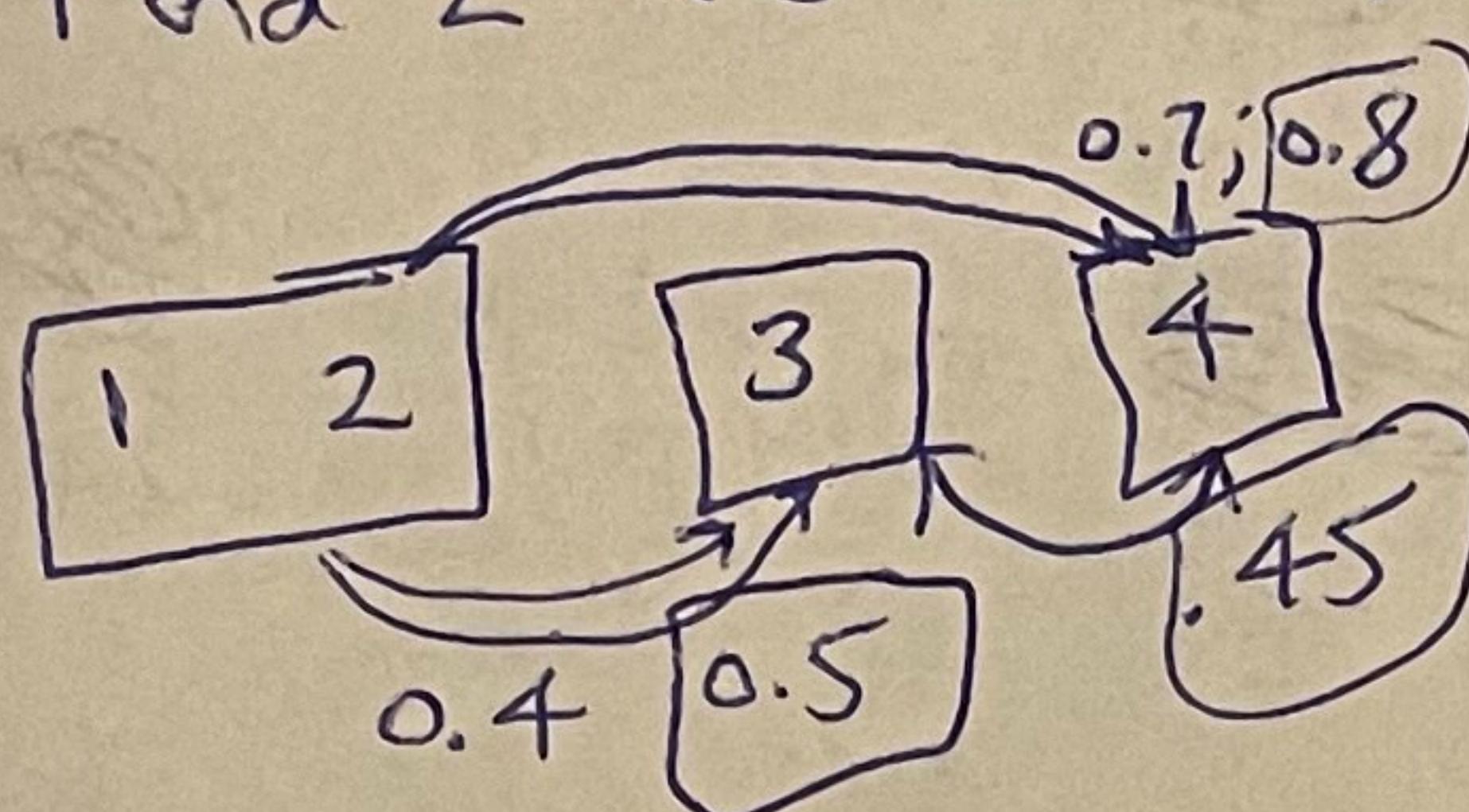
	1	0.3	0.4	0.7
2	0.3		0.5	0.8
3	0.4	0.5		0.45
4	0.7	0.8	0.45	
	1	2	3	4

(a) Hierarchically cluster based on complete linkage: (maximal intercluster dissimilarity)

(i) Step 1: 1 2 3 4

↳ since 1 and 2 are closest, group them together  $\text{dissim} = 0.3$

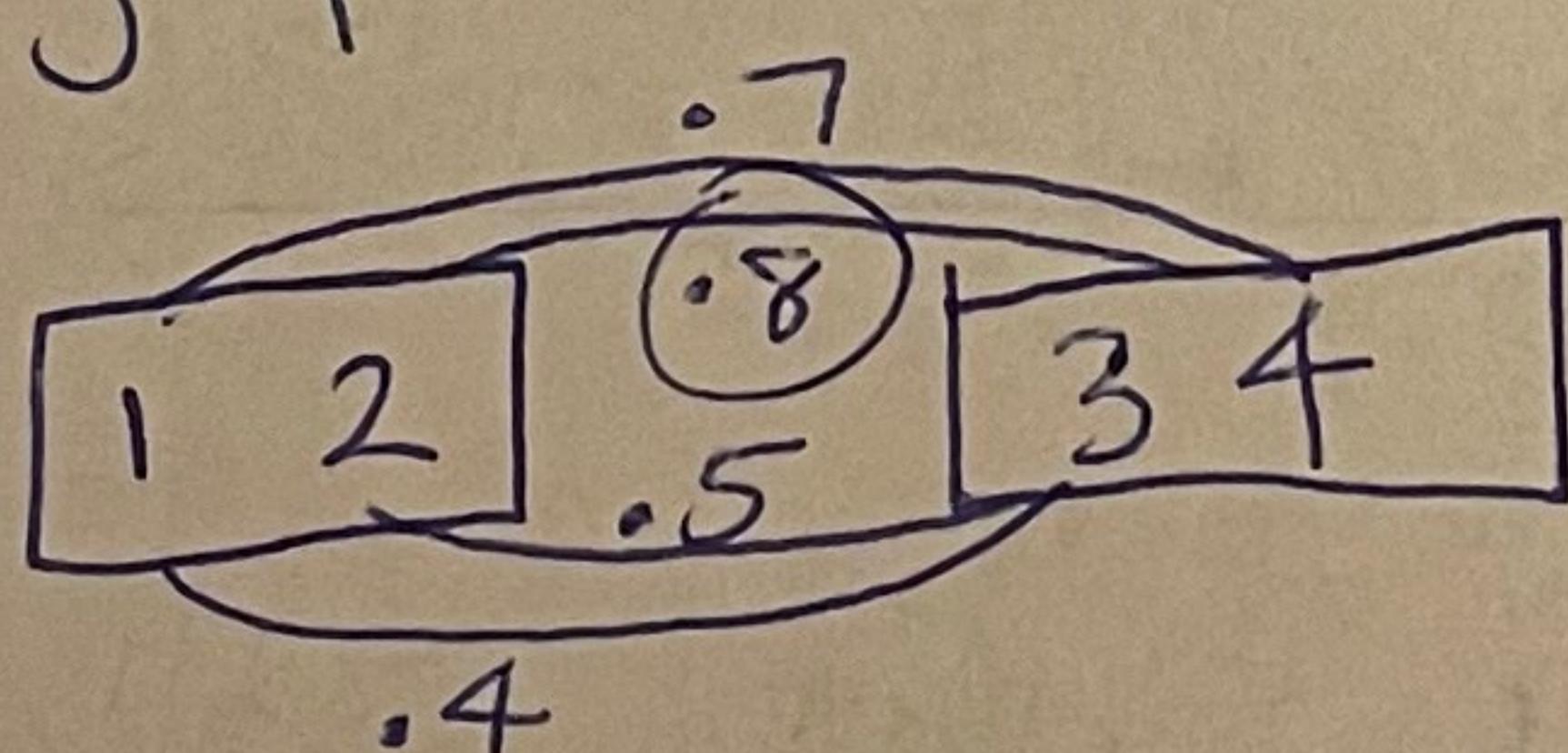
(ii) Step 2:



↳ since max diss of  $\{1, 2\}$  cluster with  $\{3\}$  is 0.5,  
 "  $\{1, 2\}$  "  $\{4\}$  is 0.8,  
 ~~$\{1, 2, 3\}$~~  "  $\{4\}$  is 0.45

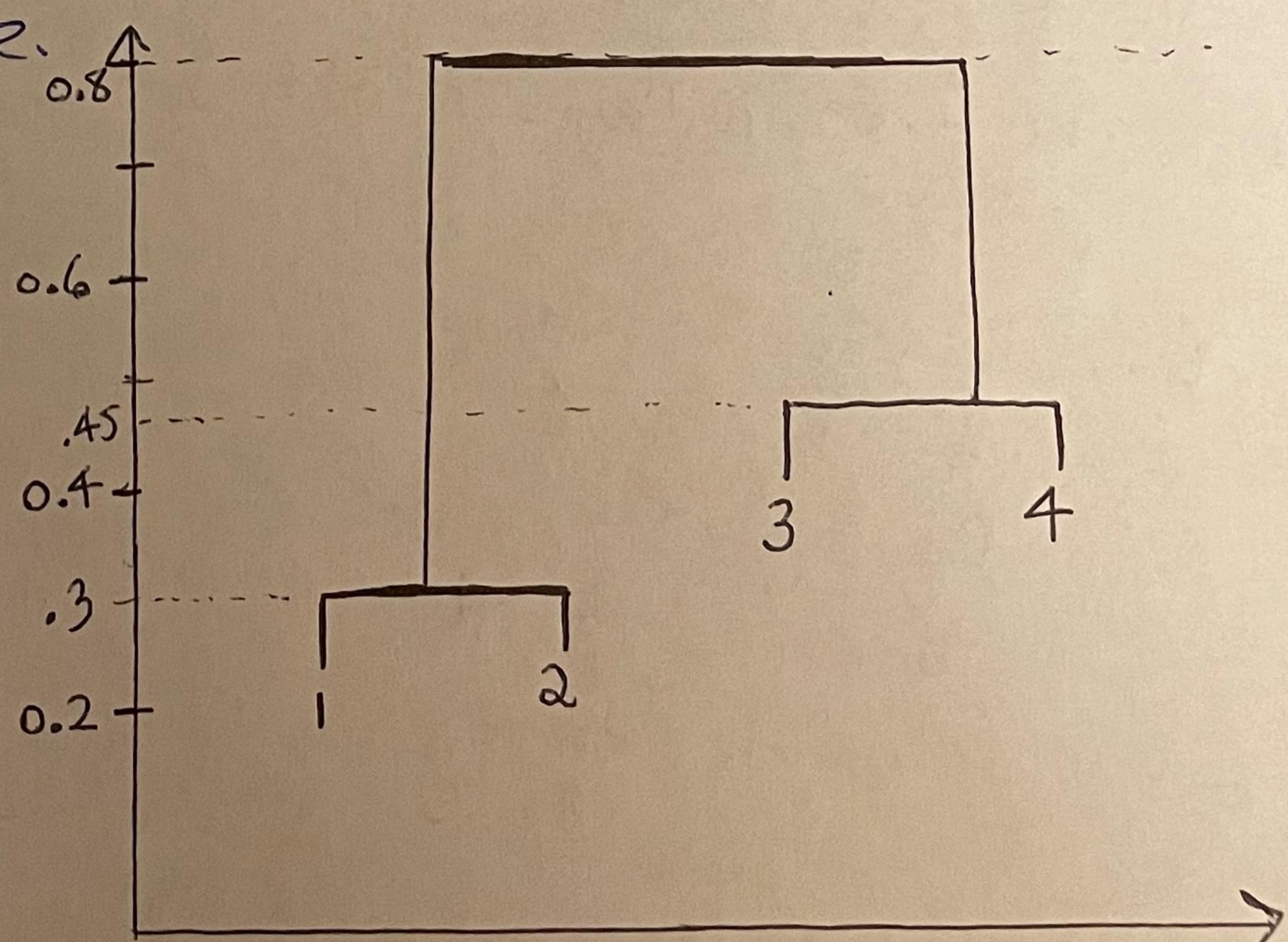
I group 3 and 4 together.

(iii) Step 3:



Since max dist is 0.8, that is where these two clusters fuse.

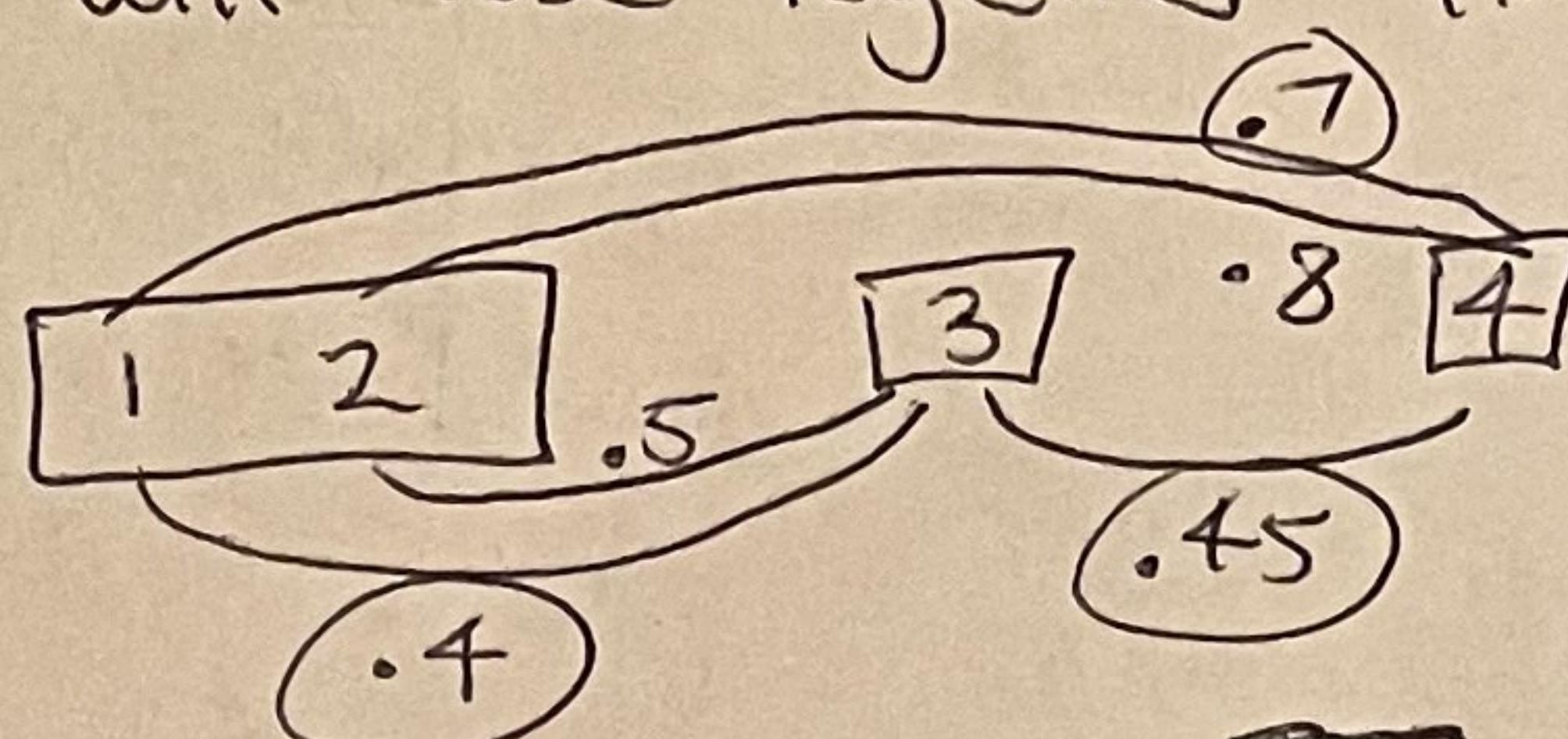
Dendrogram:



(b) Hierarchically cluster based on single linkage (=minimal intercluster dissimilarity).

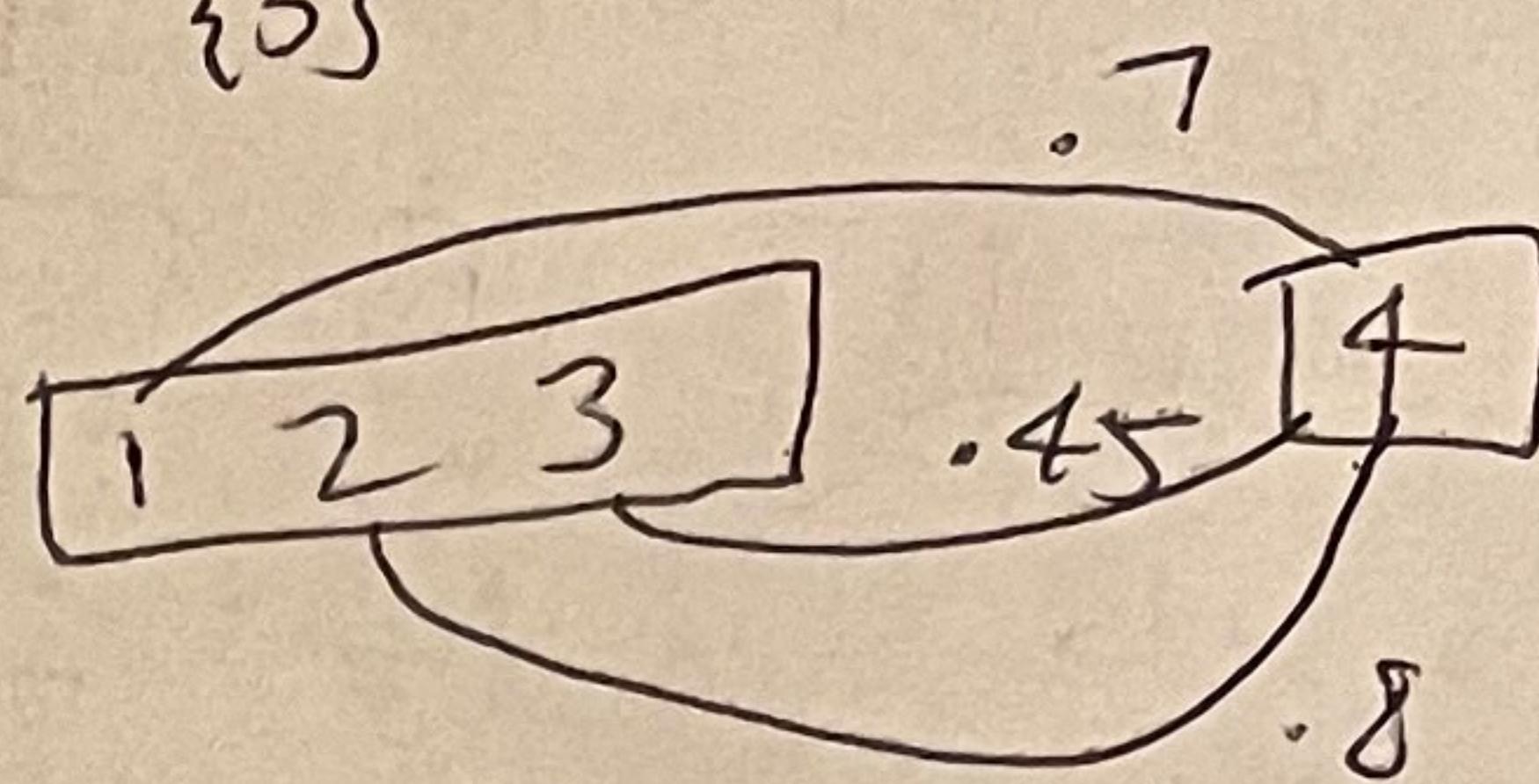
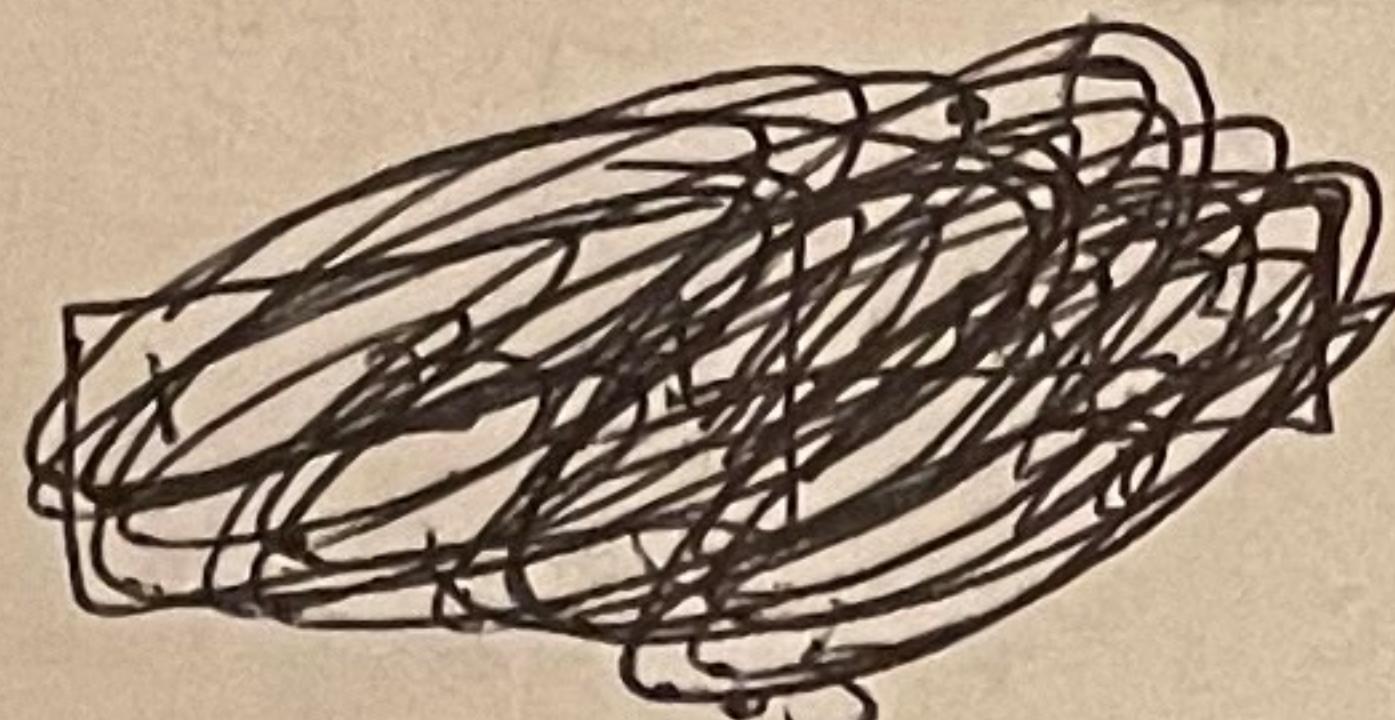
(i) Step 1: Again, as in (a), since (1) and (2) are closest together, they will fuse together first at 0.3.

(ii) Step 2:



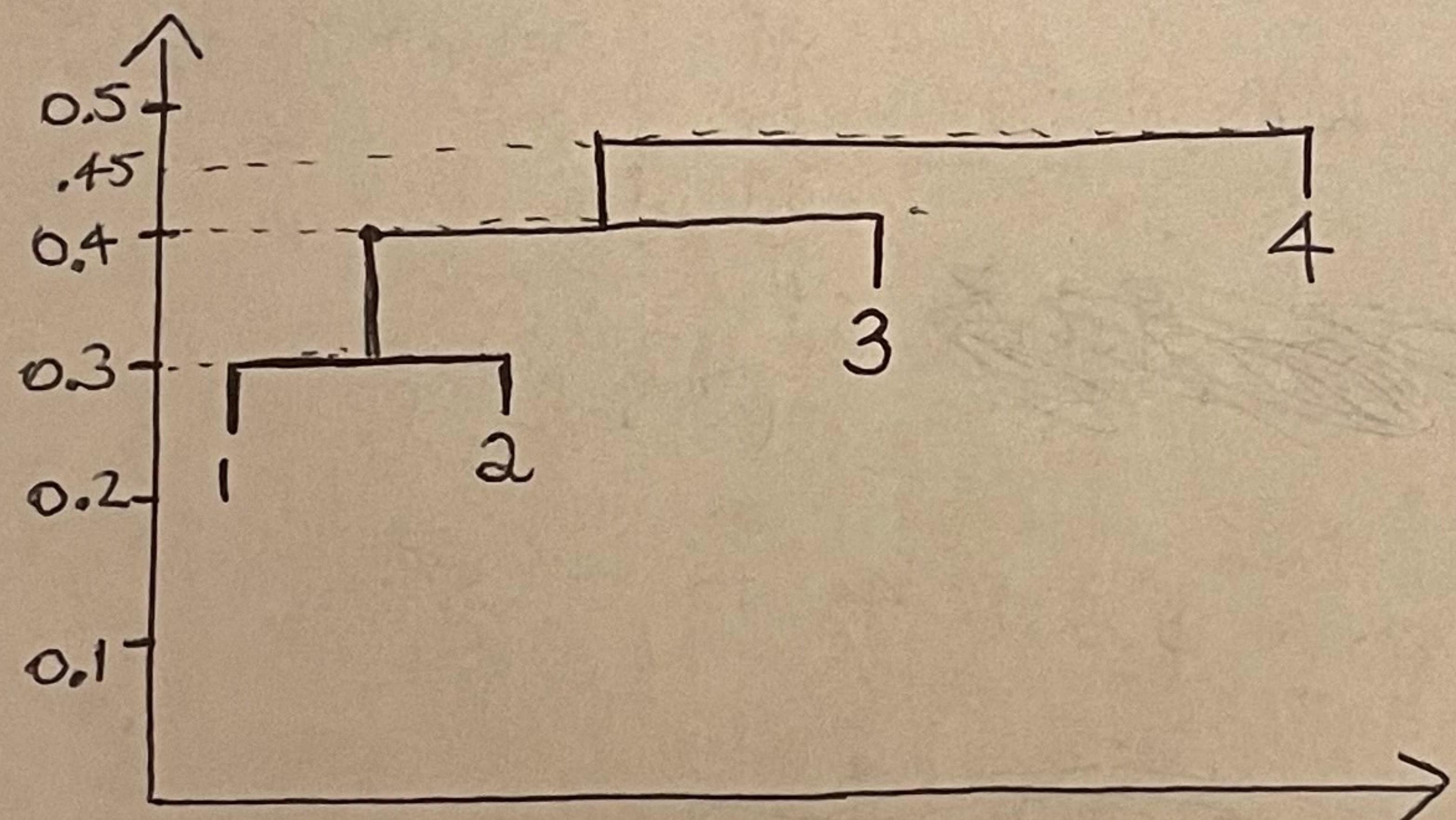
Here,  $\{1, 2\}$  will fuse with  $\{3\}$  at 0.4.

(iii) Step 3:



Here,  $\{1, 2, 3\}$  will fuse with  $\{4\}$  at 0.45.

Dendrogram:



(c.) If we cut dendrogram in (a) st. 2 clusters result, observations  $\{1, 2\}$  would be in one cluster, and obs  $\{3, 4\}$  would be in the second.

(d) In this case in (b), obs  $\{1, 2, 3\}$  are in one cluster and obs  $\{4\}$  is in the second.

(e) Draw dendrogram that is equivalent to the 1 in (a), for which 2+ leaves are repositioned, but for which meaning is the same:

