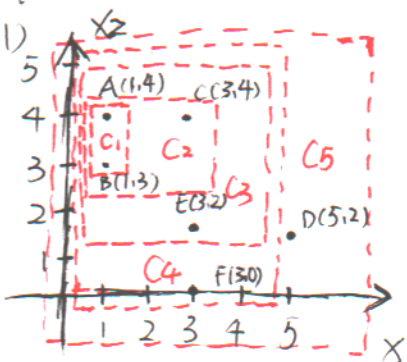
Statistical Machine Learning.

Home Work 5. 刘子 CHEN Yc3356

(1)



△ Distance Measurement Method: Manhattan Distance.
△ Notion of distance between clusters: single linkage.

(2) step one:

d(A,B) = 1 d(C,E)=2
d(A,C) = 2 d(B,C)=3
d(E,F)=2 ... ...
d(E,D)=2 I ignore other distances since they must be greater than 1.

⇒ Thus: we first cluster A and B together and we call this cluster "C1"

(3) step two:

d(C1,C)=2 I ignore other distances since, they can not be smaller than 2
d(C1,E)=3 ⇒ Thus, we cluster C1 and C together and we name this cluster "C2".

(4) step three

d(C2,E)=2 I ignore other distances since they cannot be smaller than 2.
d(C2,D)=5 ⇒ Thus, we cluster C2 and E together and we name this cluster "C3".

(5) step four:

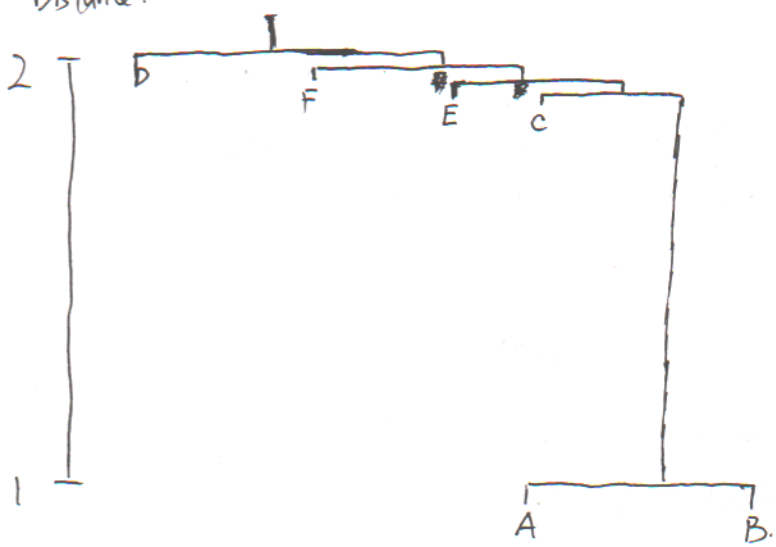
d(C3,F)=2 I ignore other distances since they cannot be smaller than 2.
d(C3,D)=2 ⇒ Here. I cluster C3 and F together and name this cluster "C4".

(6) last step:

d(C4,D)=2. Finally I cluster C4 and D together and name this cluster "C5"

cluster Dendrogram

Distance.



} these distances are all equal to two.

→ this distance is one.