

→ 1e limit depth we do pruning.  
 ↳ pre pruning (train kase se phle cut off krerge)  
 ↳ post pruning (after training whole model we cut off features)  
 Date: \_\_\_\_\_

→ generalized model or stable model wo hote hai k small change bhi krne to koi effect na ho.

## K-Means Clustering

$x_1$	$x_2$
1	1
1.5	2
3	4
5	7
3.5	5
4.5	5
3.5	4.5

$O(n \cdot k \cdot t)$   
 ↳ no. of iterations  
 ↳ no. of centroids  
 ↳ no. of data points  
 hamara point jiske close hoga  
 wo lenge  
 (1.5, 2)  $c_2$  se close hai bec dist is less  
 each point representing a cluster

① Randomly select 2 points (two centroids)  
 $c_1 = (1, 1)$ ,  $c_2 = (5, 7)$

$x_1$	$x_2$	distance $c_1$	distance $c_2$	Belong
1	1	0	7.2	$c_1$
1.5	2	1.12	6.10	$c_1$
3	4	3.61	3.61	$c_2$
5	7	3.21	0	$c_2$
3.5	5	4.72	2.5	$c_2$
4.5	5	5.31	2.06	$c_2$
3.5	4.5	4.30	2.92	$c_2$

Date: \_\_\_\_\_

C<sub>1</sub> wala ka mean,  $\left( \frac{1+1.5}{2}, \frac{1+2}{2} \right)$

C<sub>2</sub> —————  $\frac{(3+5 + 3.5 + 3.5 + 4.5 + 3.5)}{5}$ ,  $\frac{(4+7+5+5+4.5)}{5}$

pls take kote change ye job tak  
2 consecutive iterations same na aajyein

proximity  
matrix

Hierarchical clustering:

	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>5</sub>
P <sub>1</sub>	0				
P <sub>2</sub>	1.04	0			
P <sub>3</sub>	0.593	0.773	0		
P <sub>4</sub>	0.460	0.616	0.362	0	
P <sub>5</sub>	0.318	0.323	0.452	0.358	0

distances

(i) Find min distance across the matrix.

↳ called single linkage

complete linkage  
mein maximum  
distance dekhne  
hai.

→ single or complete dono se ho sakta hai



Date \_\_\_\_\_

Day \_\_\_\_\_

## "K-MEANS CLUSTERING"

- Clustering is used in unsupervised learning
- The similarity is measured by similarity function
- Shorter the distance higher similarity
- K-means is used when you have unlabeled data.
- Find groups in data, with no. of grps equal to "k"
- Data points are clustered based on feature similarity

### Hierarchical clustering

#### ▷ Agglomerative Method

→ Follows bottom up approach

↳ Single linkage → take the minimum distance

↳ Complete linkage → take the maximum distance

	BA	PI	MI	NA	RM	TO
BA	0					
PI	662	0				
MI	877	255	0			
NA	255	468	754	0		
RM	417	268	564	219	0	
TO	196	400	138	869	669	0

Page No. \_\_\_\_\_