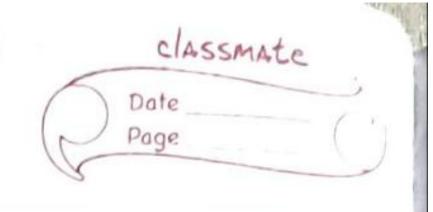


A 1, 000	Wa Mombactness or cluster cohesion: Measure
ogno ace	how close are the objects withinthe
	Same cluster. A lousex willing - clenton
	Variation is an indicator of a good
	Combactness (i.e. a good clustering). The
` '	Variation is an indicator of a good compactness (i.e. a good clustering). The different indices for evaluating the
	confactuess of clusters are base on
	distance measure such as the cluster
	between observations.
	Les ween Dobervations.
	St Separation: Measure how well-
	Sto Separated a cluster is & from
	other cluster. The indices used as
	separation neasure include:
9	distances bétween cluster centers
(A)	The fairwise minimum distances between objects in different cluders.
	between objects in afferent alusters.
	Step1: choose the number of clusters K.
	Step2: select & random Boints from
	Step2: select & Trandom Boints from the data as Centroids
	Step3: Assign all the foints to the sold
	57 ep4: Recombate the Centrold of newly
	Step4: Recombate the Centroid of newly formed elevoter.
	51ep 5: Rebeat 3 and 4.



There are essentially three stopping coitesia that can be adopted to stop the K-mean algorithm:

Centroids of newly formed clusters do not change foints remain in the Same cluster Maximum number of iterations are reached.

b) Informatices Gain = -[7] log 7 + 9 log 9 [16] 16

 $= - \left[ 0.4375 \log \left( 0.4375 \right) + 0.5625 \log \right]$   $+ 0.5625 \log \left( 0.5625 \right)$ 

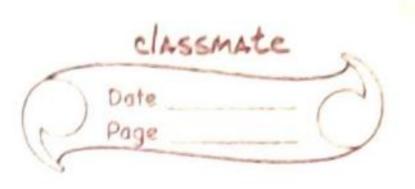
= - [0.4375 (-1.1926)+0.5625 (-0.83)]

= 0.9885

Finding splitting Attibute

3 3 3

 $E(A) = I(1) = \begin{bmatrix} 3 \log_2 \frac{3}{6} + \frac{3}{6} \log_2 \frac{3}{6} \end{bmatrix}$ 



$$=$$
  $\left[ \frac{2 \times 3}{6} \times \log_2 \frac{3}{6} \right]$ 

$$I(3) = -\left(\frac{3 \log 3}{6} + \frac{3 \log 3}{6}\right)$$

Scanned by TapScanner

Scanned by TapScanner