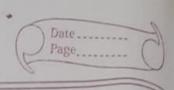


Improvised K-Mane Algorishm

- Introduction

- · clustering it a way that closeified the raw data reasonably and searche the hidden parterns that exists in dataset.
- -> Short-comings of K-moons algorithm.
 - By seeing normal K-maang algorithm, we find that algorithm has to calculate the distance of each data object from every cluster centre in each iteration.
 - Owing to which the time taken for its execution has increased manifold.
 - Comprexity of time has increased because certain data pts which remained in same cluster from start to end had to be checked for each its distance from every cluster centre after each iteration.
 - Hence me that although being quite effective was too tedious and time consuming.
 - However by experiments we have found that it is not necessary for us to calculate distance each time.



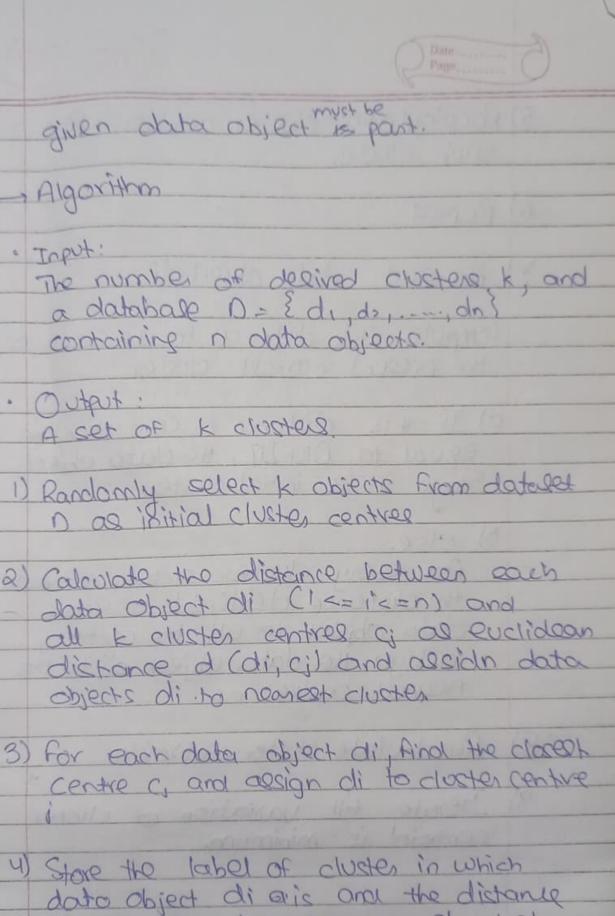
Traprovised k-means algorithm head to
Standard k-means algorithm need to
Calculate the distance of each data
of each cluster centroids every
time the iteration is conducted, which
takes up a bet of execution time
especially for large aspacity databases

- Main idea of improvised algorithmic to use 2 data structures -> 1 to store clusture label. -> 2 to store distance of all the data objects to the rangest cluster during each iteration.

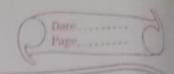
- After an iteration, we calculate distance of given data object from the cluster centroid of same cluster.

If the rew distance calculated is less than old distance then data object will be part of same cluster.

Or else distance from all cluster centroid peode to be determined in order to calculate the cluster to which



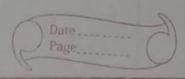
of data object of to the nearest coluster



- 5) Rocalculate cluster centraid for each cluster.
- 6) Repeat.
- 7) for each data object di

Compute it's distance to the centre of the present namest cluster

- a) If this dictance is less than an equal to Dick [i], the data object Stays in its initial cluster.
- for every cluster center, compute the distance of Codiscil of each data object to all the centere, alsign the data object di to the rearest centre, ci
- · 8) Recalculate duster centroide.
 - 9) Iterate till variotion of cluster centroid is minimum.



- -> Advantages of Optimised clustering algorithm:
- . Time complexity of execution of code decreages manifold.
 - It almost reduces to times to

 normal k-means code execution

 t- total number of k-means iteration