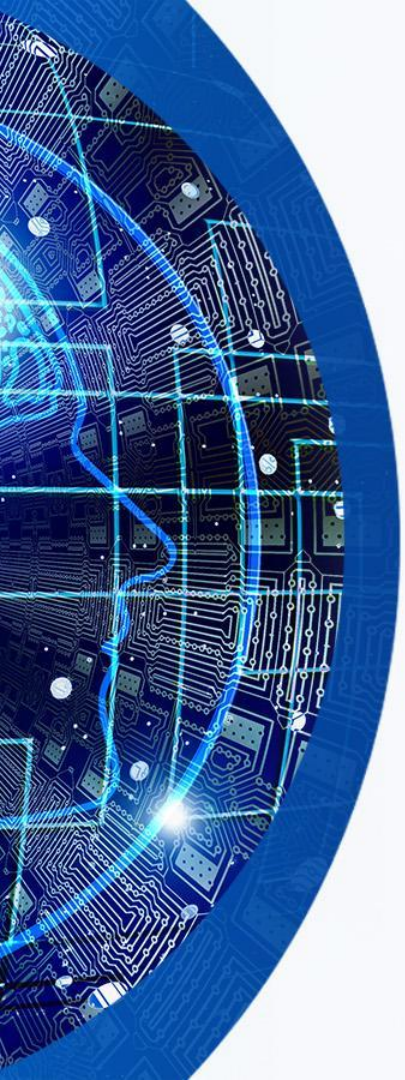
Computation of Initial Modes for K-modes Clustering Algorithm using Evidence Accumulation

Prathmesh Gaonkar (B20ME032)



Index

- K Modes algorithm
- Problem Statement
- Proposed Solution

K-modes

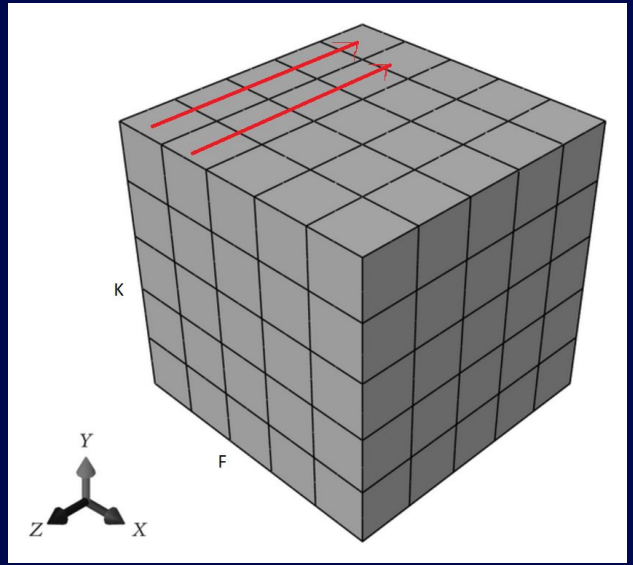
- K-modes helps in clustering categorical data.
- This algorithm uses modes instead of means.
- The algorithm uses frequency-based approach to update modes and obtain K number of clusterings.
- Firstly, we initialize K random modes for each cluster.
- Based on the simple distance metric we allocate the data points.
- Calculating the new modes of the group formed.
- Repeat the process until convergence is obtained.

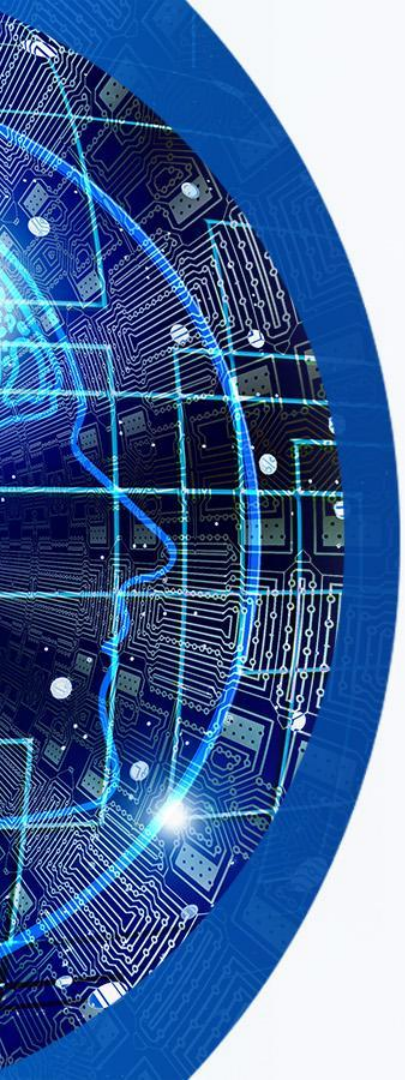
Problem-Statement

- The problem is with the random initialization of ‘K’ number of modes at the start of the algorithm.
- Because of this each time we get different clustering
- This result cannot be relied with confidence
- This paper provides a method to calculate this initial points.

Proposed Solution

- Here, the idea implemented is Evidence Accumulation.
- Iterate the process of initializing K random modes and obtain the final modes and store it in a mode pool.
- Out of each column in z axis again take mode and obtain the final initial modes.





Thank You