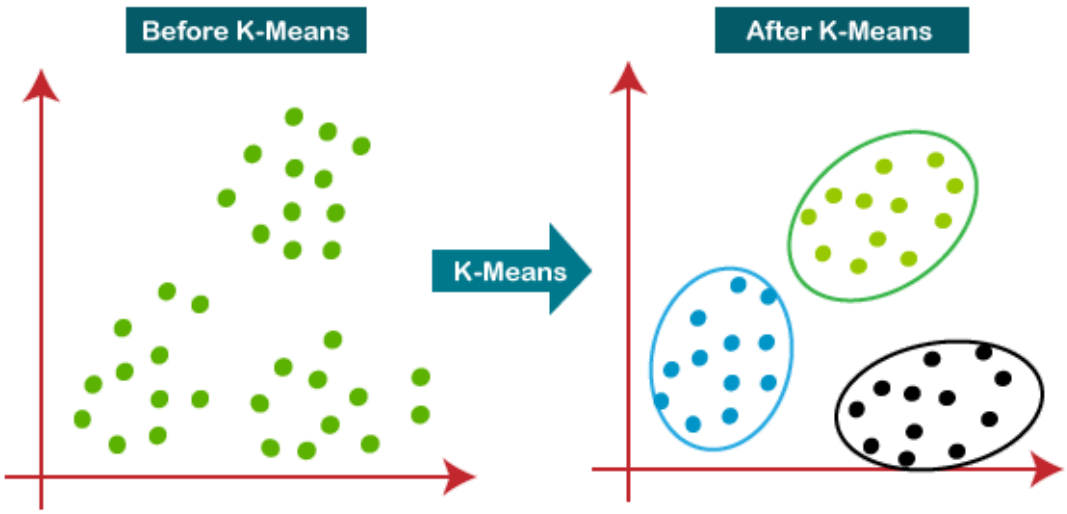

7 BEST
Data Science
Clustering
Techniques

01

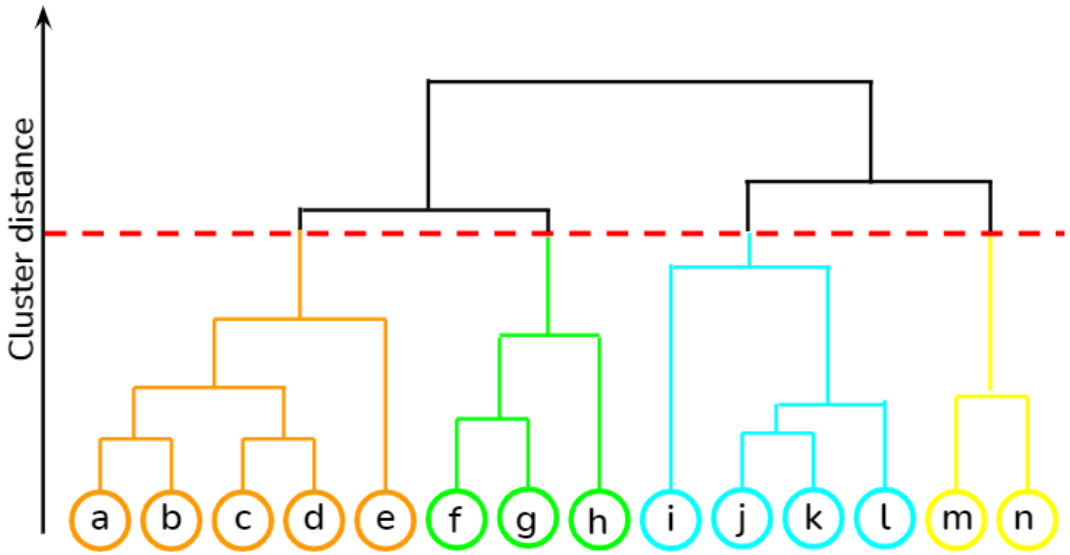
K-MEANS CLUSTERING



A centroid-based method that partitions data points into a pre-defined number of clusters (k). Simple, efficient, but sensitive to outliers and initialization.

02

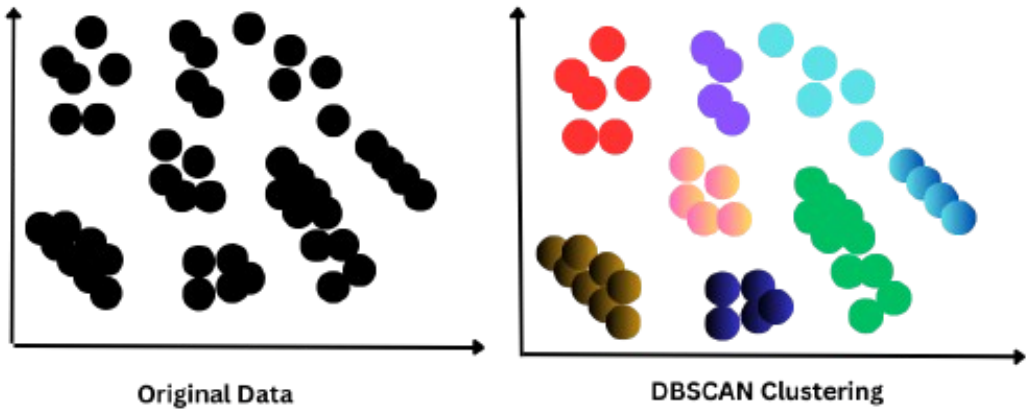
HIERARCHICAL CLUSTERING



Builds a hierarchy of clusters, either merging (agglomerative) or splitting (divisive). Good for exploratory analysis and identifying natural groupings of varying sizes.

03

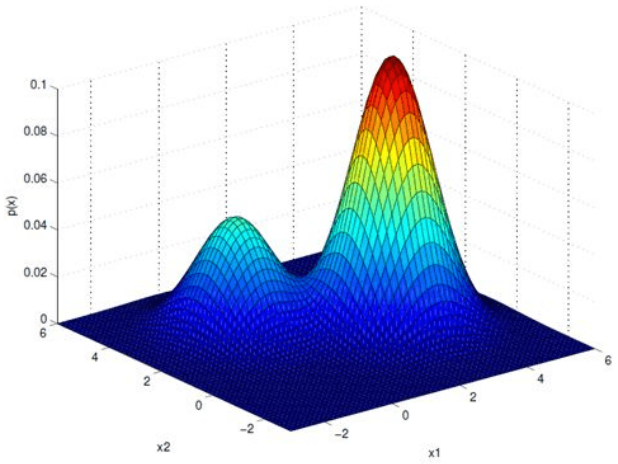
DBSCAN



Identifies clusters based on density of data points. Robust to outliers and noise, but may struggle with clusters of varying densities.

04

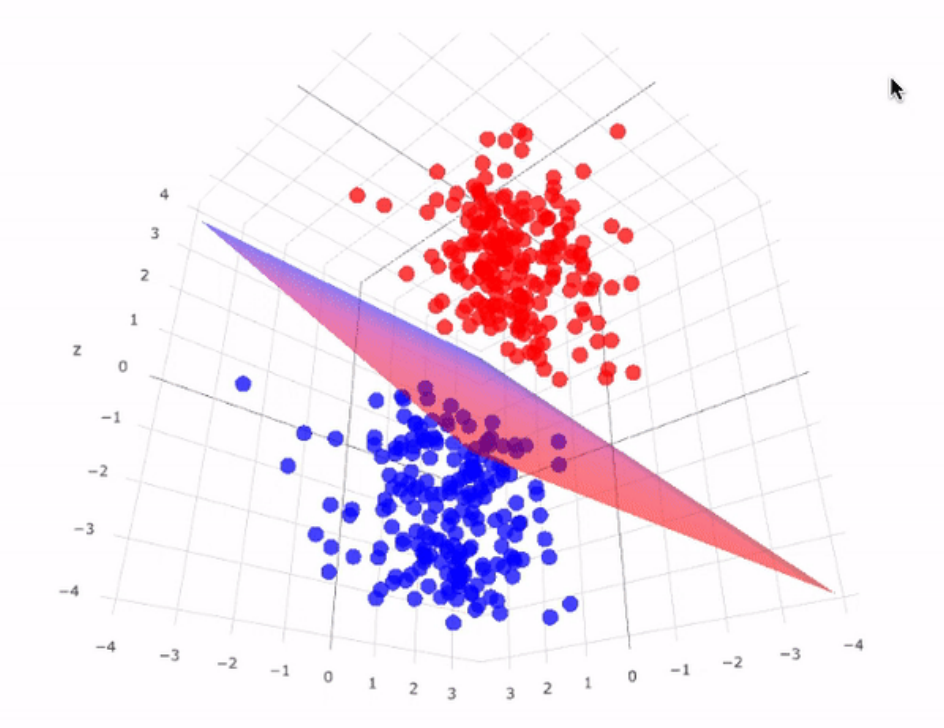
GAUSSIAN MIXTURE MODELS



Assumes data points are generated from a mixture of Gaussian distributions, where each distribution represents a cluster. Flexible for complex data shapes, but requires parameter tuning.

05

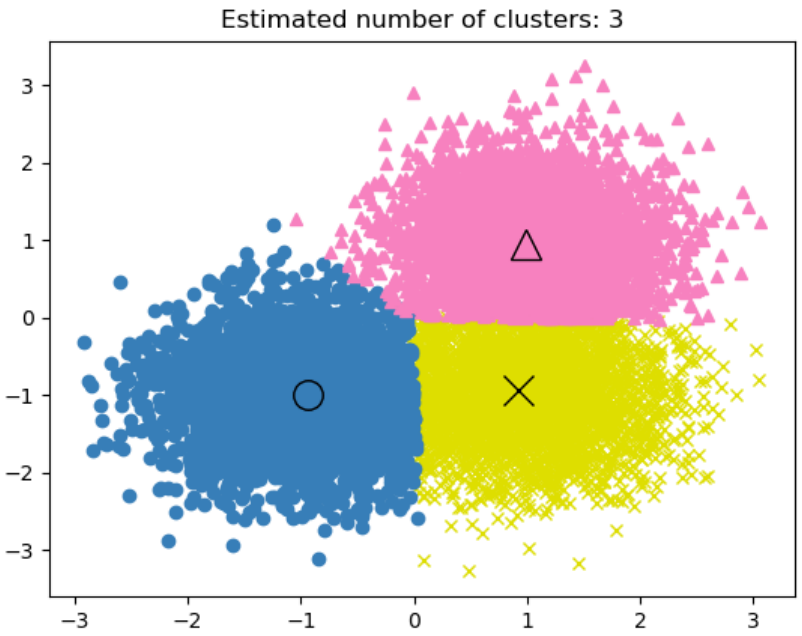
SVM CLUSTERING



Uses support vectors to define cluster boundaries. Effective for high-dimensional data and non-linear clusters, but computationally expensive.

06

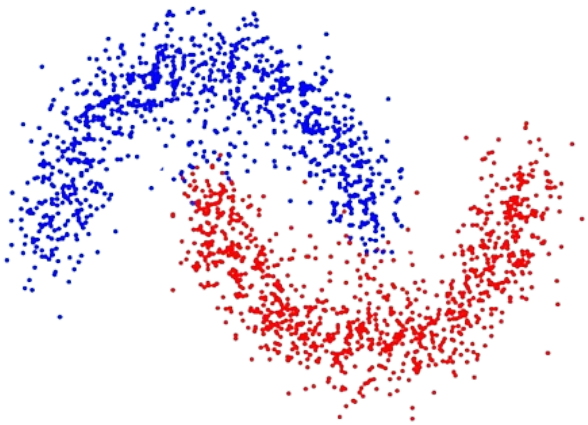
MEANSHIFT CLUSTERING



Iteratively shifts data points towards regions of higher density, ultimately converging to cluster centers. Useful for complex data shapes and identifying overlapping clusters.

07

SPECTRAL CLUSTERING



Projects data into a lower-dimensional space using dimensionality reduction techniques and then performs clustering in that space. Effective for high-dimensional data and non-linear clusters.

If you find this helpful, Repost

for AI content.

linkedin.com/in/ileonjose