BIRCH

BIRCH is a hierarchical clustering algorithm designed to work efficiently with very large datasets.
It builds a CF Tree (Clustering Feature Tree) that summarizes the data.

- Each node stores:
- Number of points
- Linear sum
- Squared sum
- Data is inserted one by one and grouped into compact subclusters.
- Final clustering (optional) is done using another algorithm like K-Means on the subclusters.

Advantages

- Fast and Scalable
  - Works well for huge datasets (millions of points).
  - Time complexity is almost linear.
- Low Memory Usage
  - Does not store all raw points.
  - Stores only statistical summaries (CF nodes).
- Incremental Learning
  - New data can be added without re-clustering everything.
  - Useful for streaming data.
- Handles Noise
  - Can ignore outliers during CF tree construction.
- Good for Real-World Data
  - Works well for sensor data, customer data, IoT, satellite data, etc.

Disadvantages

1. **Works Best with Spherical Clusters**
    a. **Like K-Means, it assumes clusters are round-shaped.**
    b. **Fails on complex shapes (moons, spirals).**
2. **Sensitive to Threshold Parameter**
    a. **Wrong threshold → wrong number of clusters.**
3. **Not Good for Very High Dimensions**
    a. **Distance becomes less meaningful.**
4. **Order Sensitive**
    a. **The result can change if data order changes.**
5. **Less Accurate than DBSCAN for Density**
    a. **Cannot find arbitrary-shaped clusters like DBSCAN.**