

SDOclust: Clustering with Sparse Data Observers

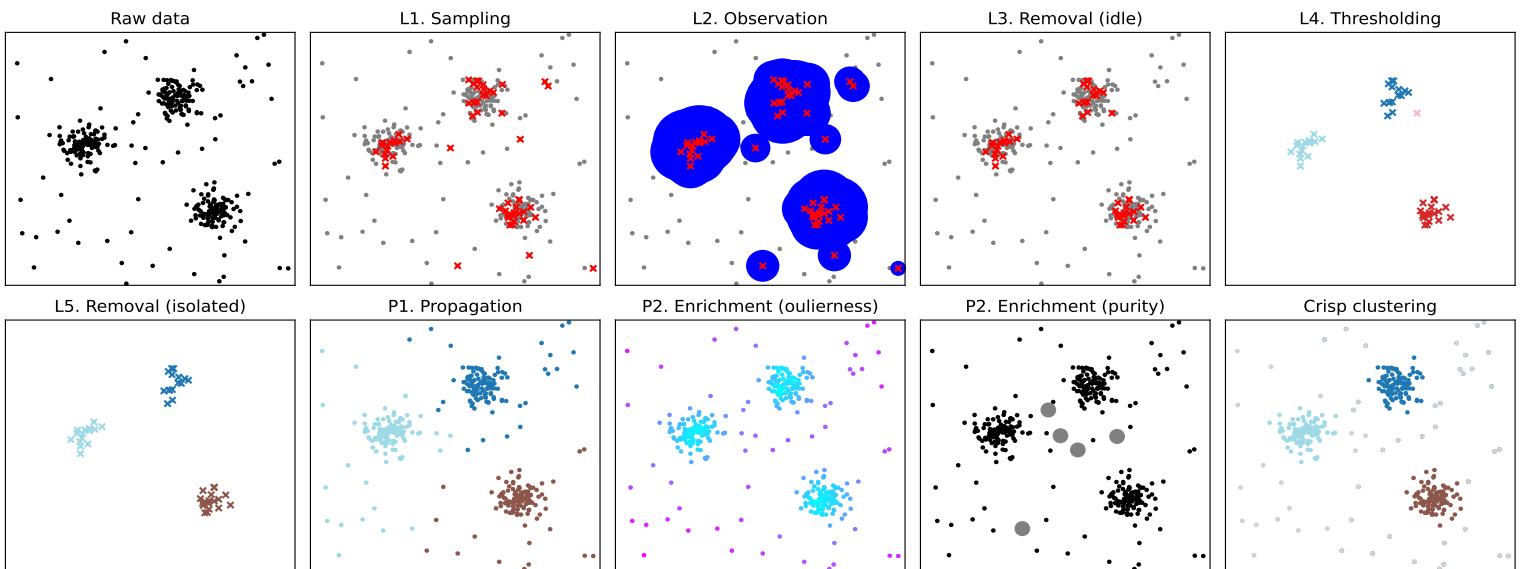
DESCRIPTION

Model-based **clustering algorithm** based on **SDO** (Sparse Data Observers) and **graph connectivity**.

MAIN FEATURES

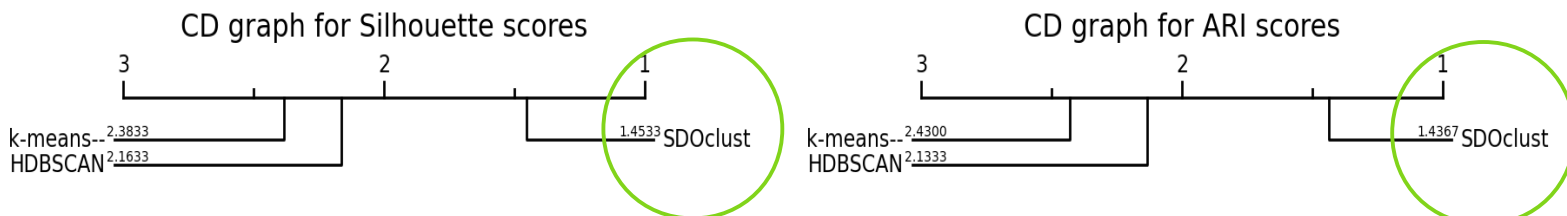
- **No parameterization** required (in most cases)
- Suitable for **non-convex** shapes
- **Noise-resistant**
- **Linear** complexity
- **High-interpretability**

PROCESS



EVALUATION

- **Comparison** of **SDOclust**, **HDBSCAN**, and **k-means--**
- **153 synthetic** datasets + **2 real** scenarios



CONCLUSIONS / INSIGHTS

- **SDOclust** shows **excellent performances** and tends to find the main partitions
- **Complementarity between HDBSCAN and SDOclust**

HDBSCAN → purity vs overclustering

SDOclust → generalization vs underclustering