Compare Clustering Methods

June 23, 2020

Hongmin Li

We conduct experiment to perform six clustering methods on two dataset. The six clustering methods are listed as following.

- k-means

- k-modoids

- SC: Spectral Clustering

- HC: Hierarchical Clustering

- GMM: Gaussian-mixture-models

- Density-based spatial clustering algorithm with noise (BBSCAN)

We show information of datasets and the performance of clustering methods in following tables.

Opps, Table 4 is too wisd to show in original scale. You can manually adjust them later.

	Instance	Feature	Class
iris	150.0000	150.0000	3.0000
wine	178.0000	178.0000	3.0000

Table 1: data

	KM	KMD	HC	GMM	DBSCAN	SC
iris	85.43 ± 10.77	89.13 ± 0.65	68.00 ± 0.00	94.53 ± 9.51	66.67 ± 0.00	87.33 ± 0.00
wine	66.32 ± 6.12	67.56 ± 5.74	42.70 ± 0.00	78.37 ± 14.51	39.89 ± 0.00	52.58 ± 1.04

Table 2: acc

	KM	KMD	HC	GMM	DBSCAN	SC
iris	73.24 ± 4.96	75.41 ± 3.31	73.55 ± 0.00	88.95 ± 7.36	76.12 ± 0.00	74.98 ± 0.00
wine	42.66 ± 0.45	42.71 ± 0.78	9.14 ± 0.00	70.04 ± 14.64	0.00 ± 0.00	21.21 ± 0.06

Table 3: nmi

	KM	KMD	HC	GMM	DBSCAN	SC
iris	0.0029 ± 0.0005	0.006 ± 0.001	0.00097 ± 5e-05	0.0067 ± 0.002	0.00086 ± 9e-05	0.015 ±
wine	0.0031 ± 0.0005	0.0079 ± 0.001	0.00091 ± 5e-05	0.01 ± 0.003	0.0019 ± 0.0001	0.022 ±

Table 4: time