

Personal Reflection

What happens when algorithm parameters change?

DBSCAN

- Changing `eps` or `min_samples` can merge or split clusters or turn many points into noise.
- Small parameter shifts can radically change the structure.

K-Means

- Changing `k` directly changes the number of clusters.
- Initialization of `k` affects final cluster.

Does clustering improve or worsen with certain parameter choices?

DBSCAN

- Good parameters → captures **arbitrary shapes**, handles noise well.
- Bad parameters → everything becomes one giant cluster or most becomes noise.

K-Means

- Good `k` → clean, compact clusters.
- Bad `k` → over-segmentation or under-segmentation.
- Sensitive to scale; unscaled data can distort results.

In what situations would this algorithm fail?

DBSCAN fails when:

- Clusters have **very different densities**
- For bad `eps` or `min_samples` values.

K-Means fails when:

- Clusters are **non-spherical**.
- Clusters have **unequal sizes or densities**
- There is significant amount of noise or outliers

How do results compare on the same dataset?

Scenario	DBSCAN	K-Means
Arbitrary shapes	Good	Poor
Arbitrary density	Good	Poor
Noise/outliers	Explicitly identifies noise	Forces every point into a cluster
Simple spherical clusters	Works, but unnecessary because of large computation time	Usually very good