

# DBSCAN

2022년 6월 3일 금요일    오후 9:35

## 1. 개념

- Density-based Spatial Clustering of Applications with Noise
- eps-neighbors
  - 한 데이터를 중심으로 epsilon 거리 이내의 데이터들을 군집
  - 너무 작으면 많은 데이터가 노이즈로 구분됨
  - 너무 크면 군집이 하나가 됨
- Minpts
  - 한 군집은 Minpts 이상의 데이터로 구성

## 2. 특징

- 데이터의 밀집 정도를 사용
- 장점
  - 군집의 갯수를 지정할 필요없음
  - 다양한 모양의 군집이 형성 될수 있고, 군집끼리 겹치지 않음
  - 하이퍼파라미터가 2개로 작음
  - 군집형태와 관계없이 클러스터링 가능
- 단점
  - 시작점에 따라 다른 모양의 군집 형성
  - eps 값에 크게 좌우됨