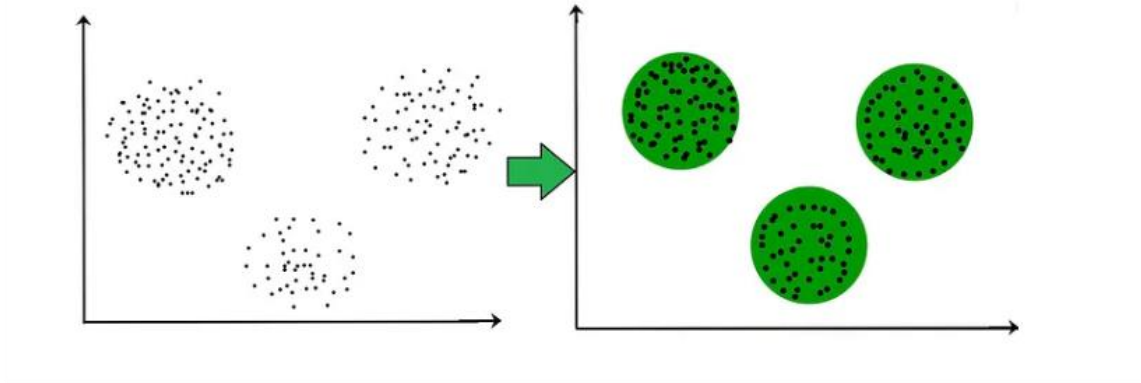


Kümeleme :

Kümeleme algoritmaları, etiketsiz verileri benzerliklerine ve farklılıklarına göre gruplara ayırmamızı sağlar. Sınıflandırma (classification) ile arasındaki temel fark burada başlar.

Kümeleme, denetimsiz bir makine öğrenmesi çeşididir. Yani kümeleme uygulayacağımız verilerde belirgin bir tanımlama bulunmaz. Verilerin benzerlik ve farklılıklarına göre onları ayırmamızı sağlar. Sınıflandırmada ise etiketli verilerle işlem yapılır.



Kümeleme yapılacak olan verinin belirgin bir tanımlaması yoktur. Şekilde veriler koordinat sistemindeki uzaklıklarına göre kümelenmiştir.

Kümelemenin Amacı:

Kümelenmenin amacı, bir grup etiketlenmemiş veride içsel gruplamayı belirlemektir.

Burada soru: neyin iyi bir kümelenme oluşturduğuna nasıl karar verileceğidir. Bir ölçüt belirlemek güçtür. Kriteri sağlaması gereken kullanıcıdır; kümelenin sonucu kullanıcının gereksinimine uymalıdır

Örneğin, iyi bir kümeleme homojen gruplar için temsilciler bularak sağlanabilir(veri azaltma/data reduction) ya da “doğal kümeler” bulma yoluna gidilebilir ve bilinmeyen özellikleri tanımlanabilir (“natural” data types), kullanışlı ve uygun gruplar bulma (“useful” data classes) veya olağandışı/istisnai veri nesneleri bulma (outlier detection) da olabilir

Kümeleme Çeşitleri:

- Hiyerarşik Kümeleme
- Gürültülü Uygulamaların Yoğunluğa Dayalı Konumsal Kümelenmesi (DBSCAN) (DBSCAN)
- K-means Kümeleme
- Ağırlık Ortalama Kaydırma Kümelemesi

- Gauss Karışım Modelleri (GMM) kullanarak Beklenti-Maksimizasyon (EM) Kümeleme