OPTICS Kümeleme, "Ordering Points To Identify Cluster Structure" (Küme Yapısını Belirlemek için Noktaları Sıralama) kavramının kısaltmasıdır. Bu algoritma, DBSCAN kümeleme algoritmasından esinlenerek geliştirilmiştir. DBSCAN kümeleme kavramlarına iki ek terim ekler.

OPTICS (Ordering Points To Identify the Clustering Structure), DBSCAN (Density-Based Spatial Clustering of Applications with Noise) benzeri bir yoğunluk tabanlı kümeleme algoritmasıdır, ancak değişen yoğunluklara ve şekillere sahip kümelemeleri çıkartabilir. Büyük, yüksek boyutlu veri kümelerinde farklı yoğunluklara sahip kümelemeleri tanımlamak için kullanışlıdır.

OPTICS'nin temel fikri, yoğunluk bağlantılı noktaları tanımlayarak veri kümesinin kümeleme yapısını çıkartmaktır. Algoritma, veri kümesindeki noktaların sıralı bir liste olan ulaşılabilirlik grafiğini oluşturarak verinin yoğunluk temelli bir temsilini oluşturur. Liste içindeki her nokta, diğer noktalardan bu noktaya ulaşmanın ne kadar kolay olduğunu gösteren bir ulaşılabilirlik mesafesi ile ilişkilendirilir. Benzer ulaşılabilirlik mesafelerine sahip noktaların aynı kümede olma olasılığı yüksektir.

OPTICS algoritması aşağıdaki temel adımları izler:

1. Küme minimum yoğunluğunu kontrol eden Eps parametresini belirleyin.

2. Veri kümesindeki her nokta için k-en yakın komşularına olan mesafeyi hesaplayın.

3. Rastgele bir noktadan başlayarak, veri kümesindeki her noktanın ulaşılabilirlik mesafesini, komşularının yoğunluğuna dayanarak hesaplayın.

4. Ulaşılabilirlik mesafesine göre noktaları sıralayarak ulaşılabilirlik grafiğini oluşturun.

5. Ulaşılabilirlik grafiğinden noktaları gruplayarak kümelemeleri çıkartın.

OPTICS'nin DBSCAN gibi diğer yoğunluk tabanlı kümeleme algoritmalarına göre bir avantajı, önceden belirli bir küme sayısını belirlemeyi gerektirmemesidir. Bunun yerine, verinin kümeleme yapısını çıkartır ve ulaşılabilirlik grafiğini oluşturur. Bu, kullanıcının belirli bir noktada ulaşılabilirlik grafiğini keserek küme sayısını daha esnek bir şekilde seçmesini sağlar.

Ayrıca, DBSCAN gibi diğer yoğunluk tabanlı kümeleme algoritmalarının aksine, OPTICS farklı yoğunluklara ve şekillere sahip kümelemeleri işleyebilir ve hiyerarşik yapıları tanımlayabilir.

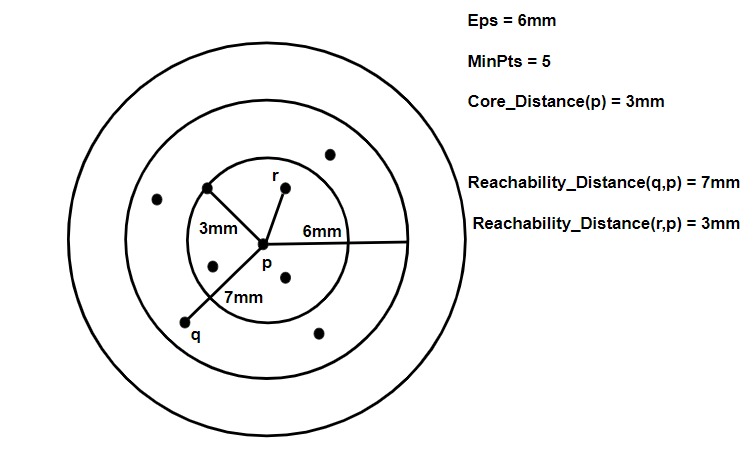
OPTICS, Python'da scikit-learn kütüphanesinde yer alan sklearn.cluster.OPTICS sınıfı kullanılarak uygulanır. Bu sınıf, minimum yoğunluk eşiği (Eps), dikkate alınacak en yakın komşu sayısı (min\_samples) ve bir ulaşılabilirlik mesafesi kesimi (xi) gibi çeşitli parametreleri alır.

Bunlar:

- Core Distance: Bir noktanın çekirdek nokta olarak sınıflandırılması için gereken minimum yarıçap değeridir. Verilen nokta bir çekirdek nokta değilse, çekirdek mesafesi tanımlanmamıştır.

- Reachability Distance: Başka bir veri noktası q'ya (Let) bağlı olarak tanımlanır. Bir nokta p ile q arasındaki ulaşılabilirlik mesafesi, p'nin çekirdek mesafesi ile p ve q arasındaki Öklidyen Mesafe'nin (veya başka bir mesafe metriğinin) maksimumu olarak tanımlanır. Ulaşılabilirlik mesafesi, q bir çekirdek nokta değilse tanımlanmaz.

Bu kümeleme tekniği, veriyi açıkça kümelere ayırmaz. Bunun yerine, Ulaşılabilirlik mesafelerinin bir görselleştirmesini üretir ve bu görselleştirmeyi kullanarak veriyi kümeleyebilir.



Kaynakça :

[ML | OPTICS Clustering Explanation - GeeksforGeeks](https://www.geeksforgeeks.org/ml-optics-clustering-explanation/)