**Veri Yapıları ve Algoritmalar KS-2 / Ödev-2**

Ad Soyad: Barash Serbest   
No: 5220505051

**OPTICS (Ordering Points to Identify the Clustering Structure) Algoritması**

OPTICS (Ordering Points to Identify the Clustering Structure), veri kümesindeki kümelerin yapısını tanımlamak için kullanılan bir kümeleme algoritmasıdır. Bu algoritma, veri noktalarını kümeler halinde gruplandırarak benzerlikleri belirler. OPTICS, veri setindeki yoğun bölgeleri, seyrek bölgeleri ve aykırı noktaları tanımlamak için kullanılan bir yöntemdir.

Optics algoritması, veri noktalarının bir uzayda yer aldığı ve birbirleriyle olan mesafelerinin hesaplandığı bir çizelge (reachability plot) oluşturur. Bu çizelge, veri noktalarının birbirleriyle olan ilişkisini gösterir. Optics, veri noktalarını iki önemli parametre kullanarak analiz eder: epsilon (ε) ve minPts.

Epsilon (ε), bir noktanın komşuluk yarıçapını temsil eder. Bir nokta, ε mesafesi içerisinde en az minPts sayısında komşusu varsa, bu nokta bir çekirdek nokta (core point) olarak kabul edilir. MinPts, bir noktanın çekirdek nokta olarak kabul edilmesi için gereken minimum komşu sayısını belirler.

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Şekil 1:** OPTICS algoritması Pseudo (kaba) kod Örneği

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Şekil 2:** OPTICS algoritması Pseudo kod örneği "update" fonksiyonu

Optics algoritması çalışırken şu adımları izler:

1. **Başlangıç noktası seç:** Rastgele bir nokta seçilir ve işlenmemiş olarak işaretlenir.
2. **Komşuluk arama:** Seçilen noktanın ε mesafesi içerisindeki tüm noktalar tespit edilir.
3. **Çekirdek nokta kontrolü:** Eğer seçilen noktanın ε mesafesi içerisindeki nokta sayısı minPts'ten fazlaysa, seçilen nokta bir çekirdek nokta olarak işaretlenir.
4. **Reachability mesafesi hesaplama:** Çekirdek noktalarının diğer çekirdek noktalara olan reachability mesafesi (erişilebilirlik mesafesi) hesaplanır.
5. **Optics sırasını güncelleme:** Çekirdek noktanın reachability mesafesi, çekirdek noktaların sıralandığı bir sıra (optics sırası) üzerinde güncellenir.
6. **İşaretlenmemiş nokta seçimi:** İşaretlenmemiş bir çekirdek nokta seçilir ve işlenmiş olarak işaretlenir.
7. **İşlenmiş noktaların kontrolü:** İşlenen noktaların ε mesafesi içerisindeki noktalar tespit edilir.
8. **Optics sırasının güncellenmesi:** İşlenen noktanın reachability mesafesi, mevcut optics sırası üzerinde güncellenir.
9. Tüm noktalar işlenene kadar adımlar 6-8 tekrarlanır.

Optics algoritması sonucunda oluşan reachability plot, veri noktalarının kümeleme yapısını gösterir. Bu çizelge, veri noktalarını yoğun bölgeler, seyrek bölgeler ve aykırı noktalar olarak sınıflandırır. Reachability mesafesi, noktaların birbirlerine olan benzerlik derecesini gösterir. Bu şekilde, veri kümesindeki farklı yoğunluklara sahip kümeleri ve aykırı noktaları belirlemek mümkün olur.

Optics algoritması, veri analizi, veri madenciliği, görüntü işleme, coğrafi bilgi sistemleri ve diğer birçok alanda kullanılabilir. Özellikle veri noktalarının karmaşık yapıları olan veri kümelerinde kümeleme yapmak için etkilidir. Optics, diğer kümeleme algoritmalarına göre daha esnek bir yapıya sahiptir ve veri noktalarının doğal kümeleme yapısını daha iyi yakalayabilir.

Optics algoritması, veri noktalarının reachability mesafesine dayalı olarak sıralandığı bir çizelge olan reachability plot'u kullanarak veri kümesindeki kümeleme yapılarını tanımlar. Reachability mesafesi, bir noktanın diğer noktalara olan uzaklığını temsil eder. Optics algoritması, bu mesafeleri kullanarak veri noktalarını gruplara ayırır ve yoğun bölgeleri, seyrek bölgeleri ve aykırı noktaları belirler.

Reachability plot, Optics algoritması tarafından oluşturulan bir grafik şeklinde ifade edilir. Bu grafikte, veri noktaları x ve y eksenlerinde temsil edilir. Noktaların yükseklikleri ise reachability mesafelerini gösterir. Yoğun bölgeler, reachability mesafeleri düşük olan bölgelerdir. Bu bölgelerdeki noktalar birbirine yakındır ve benzerlik gösterirler. Seyrek bölgeler ise reachability mesafeleri yüksek olan bölgelerdir. Bu bölgelerdeki noktalar arasındaki benzerlik düşüktür. Aykırı noktalar ise reachability mesafeleri belirli bir eşik değeri üzerinde olan noktalardır ve diğer noktalardan önemli ölçüde farklılık gösterirler.

Optics algoritması, veri kümesindeki doğal yapıyı daha iyi yakalayabilen bir kümeleme yöntemi olarak diğer algoritmalardan farklılık gösterir. Özellikle veri noktalarının farklı yoğunluklara ve şekillere sahip karmaşık yapılarını tanımlamada etkilidir. Optics, küme merkezi tabanlı yöntemlerden ziyade veri noktaları arasındaki ilişkileri kullanır ve daha esnek bir sınıflandırma sağlar.

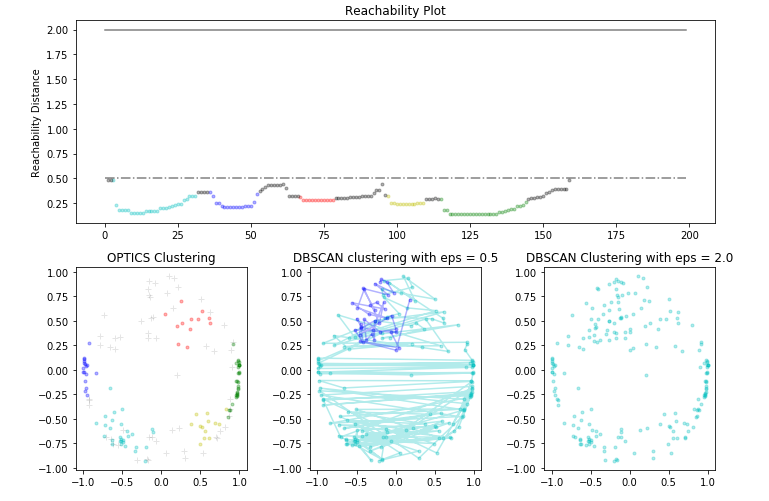
Optics algoritması, birçok uygulama alanında kullanılabilir. Örneğin, pazarlama analizinde müşteri segmentasyonu, sosyal ağ analizinde topluluk tespiti, tıp alanında hastalık sınıflandırması, coğrafi bilgi sistemlerinde nokta kümeleme gibi birçok alanda kullanım potansiyeline sahiptir. Optics algoritması, veri analizi ve veri madenciliği çalışmalarında veri noktalarının gruplara ayrılması ve örüntülerin keşfedilmesi için önemli bir araçtır.

Optics algoritması, veri noktalarının doğal kümeleme yapısını tanımlamak ve benzerlikleri belirlemek için kullanılan bir yöntemdir. Birçok farklı alanda kullanılabildiğinden dolayı, örneklerle ne için kullanıldığını daha detaylı bir şekilde anlatabilirim.

1. **Müşteri Segmentasyonu:** Bir şirket, müşteri verilerine sahipse ve bu verileri kullanarak müşterileri farklı gruplara ayırmak istiyorsa, Optics algoritması kullanılabilir. Örneğin, bir e-ticaret şirketi, müşterilerini farklı segmentlere ayırmak için müşteri satın alma geçmişi, demografik bilgiler ve davranış verilerini kullanabilir. Optics algoritması, müşterileri benzer alışveriş alışkanlıklarına, tercihlere veya demografik özelliklere göre gruplandırarak daha etkili pazarlama stratejileri geliştirmeye yardımcı olabilir.
2. **Görüntü İşleme:** Optics algoritması, görüntü işleme projelerinde de kullanılabilir. Örneğin, bir görüntü analizi projesinde, benzer özelliklere sahip piksellerin gruplandırılması gerekebilir. Optics, pikseller arasındaki benzerlikleri hesaplayarak görüntüdeki farklı nesneleri veya bölgeleri tanımlamak için kullanılabilir. Bu, nesne algılama, yüz tanıma veya görüntü sınıflandırma gibi uygulamalarda faydalı olabilir.
3. **Coğrafi Bilgi Sistemleri:** Optics algoritması, coğrafi verilerin analizinde de kullanılabilir. Örneğin, bir şehirdeki konum verilerini kullanarak farklı bölgeleri tanımlamak isteyen bir belediye, Optics algoritmasını kullanabilir. Bu sayede, benzer özelliklere sahip bölgeleri gruplandırarak şehirdeki planlama, ulaşım veya nüfus dağılımı gibi konularda daha iyi kararlar alabilir.
4. **Veri Madenciliği:** Optics algoritması, genel veri madenciliği projelerinde de kullanılabilir. Örneğin, bir finansal kuruluş, müşteri finansal verilerini analiz ederek dolandırıcılık şüphesi olan işlemleri tespit etmek isteyebilir. Optics algoritması, benzerlikleri belirleyerek potansiyel dolandırıcılık vakalarını tespit etmeye yardımcı olabilir.

Bu örnekler, Optics algoritmasının çeşitli alanlarda nasıl kullanılabileceğini göstermektedir. Temelde, veri noktalarını gruplandırarak benzerlikleri belirlemek ve veriye dayalı karar alma süreçlerini iyileştirmek için kullanılan bir algoritmadır.

C# uygulamasında uygulanan OPTICS uygulamasının temsili grafik görseli:



Şekil : OPTICS algoritması temsili grafik

**Kaynakça**

* [**https://dergipark.org.tr/tr/download/article-file/384439**](https://dergipark.org.tr/tr/download/article-file/384439)
* [**https://en.wikipedia.org/wiki/OPTICS\_algorithm**](https://en.wikipedia.org/wiki/OPTICS_algorithm)
* [**https://www.geeksforgeeks.org/ml-optics-clustering-explanation/**](https://www.geeksforgeeks.org/ml-optics-clustering-explanation/)
* [**https://www.geeksforgeeks.org/ml-optics-clustering-implementing-using-sklearn/**](https://www.geeksforgeeks.org/ml-optics-clustering-implementing-using-sklearn/)