YAPAY ZEKA

Doç.Dr. Selçuk ALP
alp@yildiz.edu.tr

İÇERİK

- Yapay Zekaya Giriş
- Makine Öğrenmesi
- Uzman Sistemler
- Arama Yöntemleri
- Kümeleme Algoritmaları
- Sınıflandırma Algoritmaları
- Genetik Algoritmalar
- Yapay Sinir Ağları
- Karınca Kolonisi Optimizasyonu
- Tabu Arama
- Parçacık Sürü Optimizasyonu

KONULAR

- Kümeleme İşlemleri
- Kümeleme Tanımı
- Kümeleme Uygulamaları
- Kümeleme Yöntemleri

KÜMELEME (CLUSTERING)

- Kümeleme birbirine benzeyen veri parçalarını ayırma işlemidir ve kümeleme yöntemlerinin çoğu veri arasındaki uzaklıkları kullanır.
- Nesneleri kümelere (gruplara) ayırma
- ***Küme,*** birbirine benzeyen nesnelerden gruplardır.
  - Aynı kümedeki nesneler birbirine daha çok benzer
  - Farklı kümedeki nesneler birbirine daha az benzer

KÜMELEME (CLUSTERING)

- Danışmansız öğrenme: Hangi nesnenin hangi sınıfa ait olduğu ve sınıf sayısı belli değil
- Uygulamaları:
  - verinin dağılımını anlama
  - başka veri madenciliği uygulamaları için ön hazırlık

KÜMELEME (CLUSTERING)

Kümeleme Uygulamaları

- Örüntü tanıma
- Görüntü işleme
- Ekonomi
- Aykırılıkları belirleme
- WWW
  - Döküman kümeleme
  - Kullanıcı davranışlarını kümeleme
  - Kullanıcıları kümeleme
- Diğer veri madenciliği algoritmaları için bir önişleme adımı
- Veri azaltma – küme içindeki nesnelerin temsil edilmesi için küme merkezlerinin kullanılması

KÜMELEME (CLUSTERING)

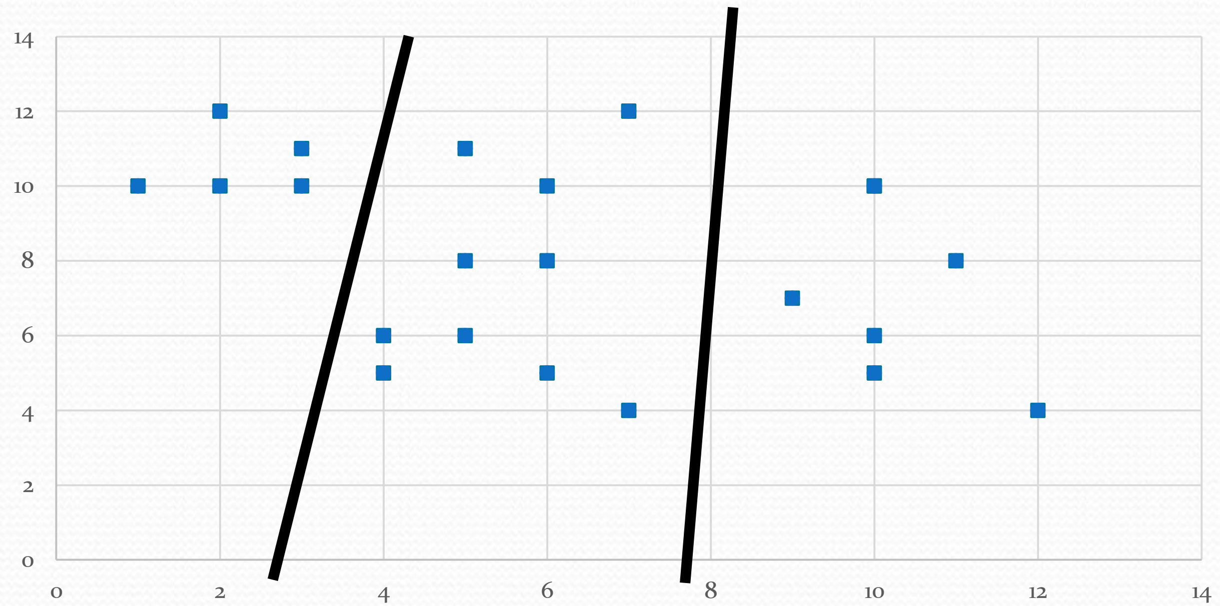
- Ölçeklenebilirlik
- Farklı tipteki niteliklerden oluşan nesneleri kümeleme
- Farklı şekillerdeki kümeleri oluşturabilme
- En az sayıda giriş parametresi gereksinimi
- Hatalı veriler ve aykırılıklardan en az etkilenme
- Model oluşturma sırasında örneklerin sırasından etkilenmeme
- Çok boyutlu veriler üzerinde çalışma
- Kullanıcıların kısıtlarını göz önünde bulundurma
- Sonucun yorumlanabilir ve anlaşılabilir olması

KÜMELEME (CLUSTERING)

- İyi kümeleme yöntemiyle elde edilen kümelerin özellikleri
  - aynı küme içindeki nesneler arası benzerlik fazla
  - farklı kümelerde bulunan nesneler arası benzerlik
- Oluşan kümelerin kalitesi seçilen benzerlik ölçütüne ve bu ölçütün gerçeklenmesine bağlı
  - Uzaklık / Benzerlik nesnelerin nitelik tipine göre değişir
    - Nesneler arası benzerlik: s(i,j)
    - Nesneler arası uzaklık: d(i,j) = 1 – s(i,j)
- İyi bir kümeleme yöntemi veri içinde gizlenmiş örüntüleri bulabilmeli
- Veriyi gruplama için uygun kümeleme kriteri bulunmalı
  - kümeleme= aynı kümedeki nesneler arası benzerliği en büyüten, farklı kümedeki nesneler arası benzerliği en küçülten fonksiyon
- Kümeleme sonucunun kalitesi seçilen kümelerin şekline ve temsil edilme yöntemine bağlı

KÜMELEME (CLUSTERING)

x1	x2
2	12
3	11
4	6
7	4
6	5
2	10
6	10
7	12
10	6
5	8
6	8
9	7
10	10
4	5
5	6
1	10
3	10
5	11
10	5
12	4
11	8
10	10



KÜMELEME (CLUSTERING)

Kümeleme Yöntemlerinde Kullanılan Uzaklıklar

- Öklid $d(i,j) = \sqrt{\sum_{k=1}^{p} \left(x_{ik} - x_{jk}\right)^2}$
- Minkowski $d(i,j) = \left|\sum_{k=1}^{p} \left(\left|x_{ik} - x_{jk}\right|^m\right)\right|^{\frac{1}{m}}$
- Manhattan $d(i,j) = \sum_{k=1}^{p} \left(\left|x_{ik} - x_{jk}\right|\right)$

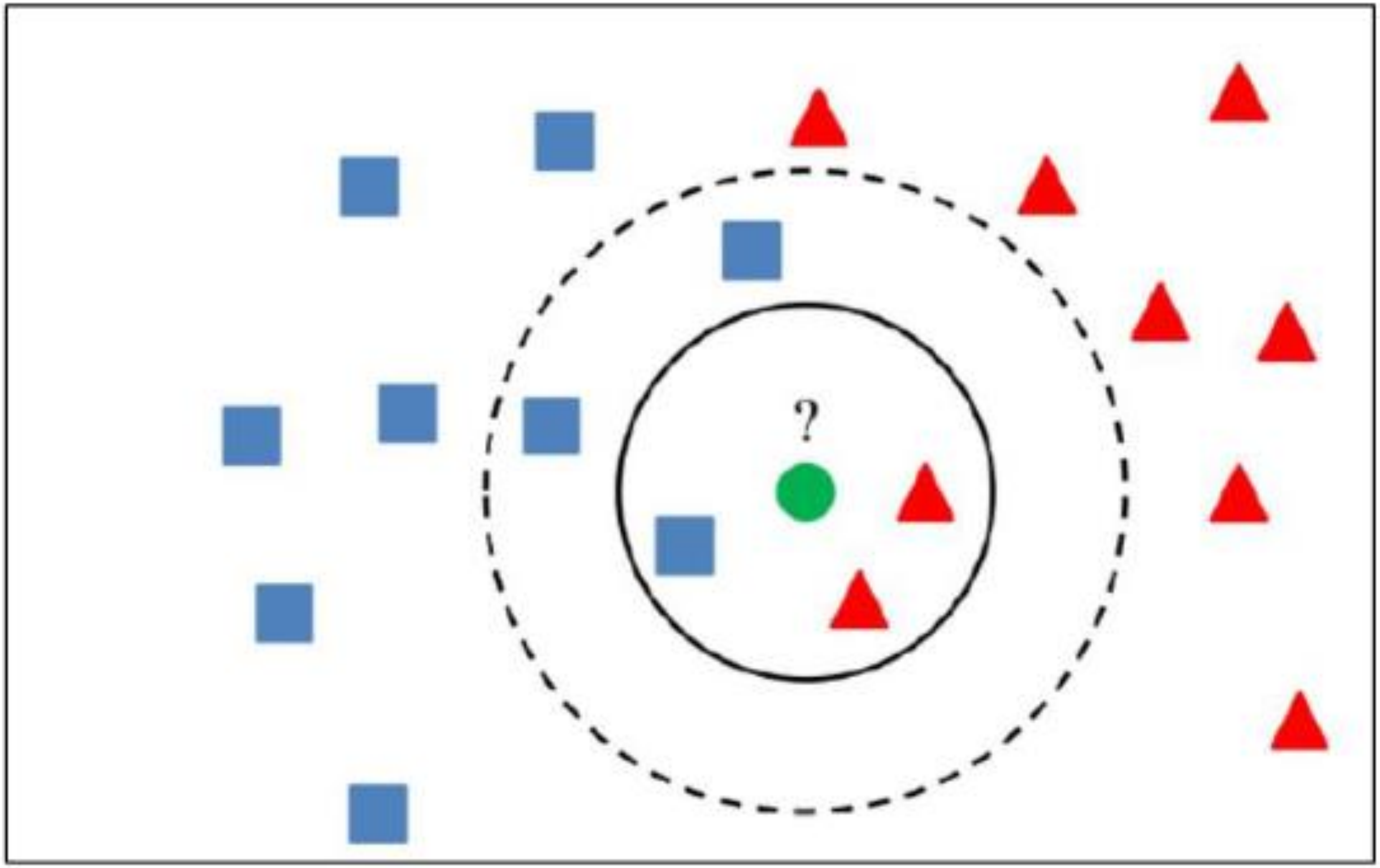
KÜMELEME YÖNTEMLERİ

- Hiyerarşik Kümeleme
  - Birleştirici Hiyerarşik Yöntemler
    - En yakın komşu algoritması
    - En uzak komşu algoritması
- Hiyerarşik Olmayan Kümeleme
  - K-Ortalamalar Yöntemi (K-Means)

EN YAKIN KOMŞU ALGORİTMASI

- En yakın komşu yöntemine «tek bağlantı kümeleme yöntemi» adı da verilmektedir. Başlangıçta tüm gözlem değerleri birer küme olarak değerlendirilir. Adım adım bu kümeler birleştirilerek yeni kümeler elde edilir.
- Bu yöntemde öncelikle gözlemler arasındaki uzaklıklar belirlenir. Öklid uzaklık bağıntısı kullanılabilir.
- Uzaklıklar göz önüne Min d(i,j) seçilir. Söz konusu uzaklıkla ilgili satırlar birleştirilerek yeni bir küme elde edilir. Bu duruma göre uzaklıkların yeniden hesaplanması gerekir.
- Tek bir gözlemden oluşan kümeler arasındaki uzaklıkları doğrudan hesaplayabiliriz. Ancak birden fazla gözlem değerine sahip olan iki küme arasındaki uzaklığın belirlenmesi gerektiğinde farklı bir yol izlenir. İki kümenin içerdiği gözlemler arasında "birbirine en yakın olanların uzaklığı» iki kümenin birbirine olan uzaklığı" olarak kabul edilir.

EN YAKIN KOMŞU ALGORİTMASI



EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.

- Aşağıdaki tabloda verilen beş gözlem değeri, en yakın komşu algoritması ile kümelenmek isteniyor.

Gözlemler	X_1	X_2
1	4	2
2	6	4
3	5	1
4	10	6
5	11	8

- Adım1. Öncelikle uzaklık tablosu oluşturulur. Her bir gözlemin birbiriyle arasındaki öklid uzaklığı hesaplanır.

EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.

$d(1,2) = \sqrt{(4-6)^2 + (2-4)^2}$=2,83

$d(1,3) = \sqrt{(4-5)^2 + (2-1)^2}$=1,41

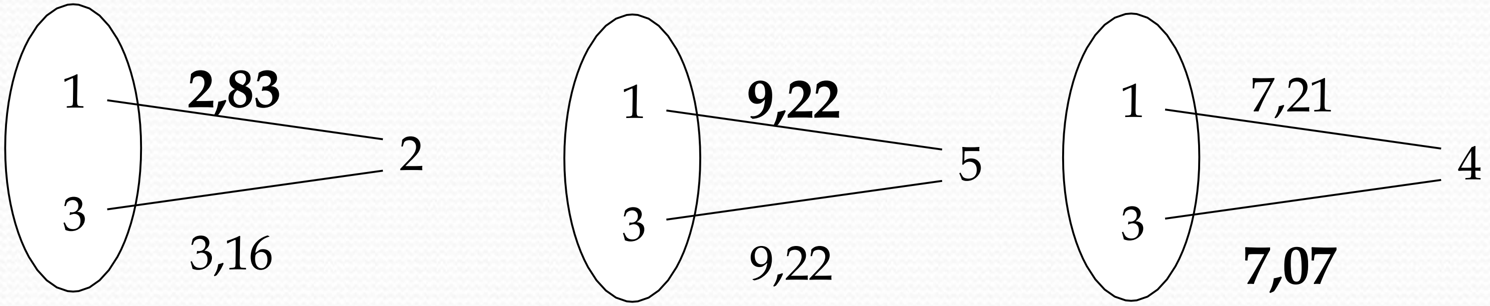
$d(1,4) = \sqrt{(4-10)^2 + (2-6)^2}$=7,21

Gözlemler	1	2	3	4	5
1					
2	2,83				
3	1,41	3,16			
4	7,21	4,47	7,07		
5	9,22	6,40	9,22	2,24	

EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.

- Adım 2. Uzaklıklar tablosunda Min d(i,j) değerinin 1,41 olduğu görülmektedir. İlgili gözlemler 1 ve 3 gözlemleridir. Bu iki değer birleştirilerek (1,3) kümesi elde edilir. Sonrasında bu kümeye göre uzaklıklar matrisi yeniden incelenir.



EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.

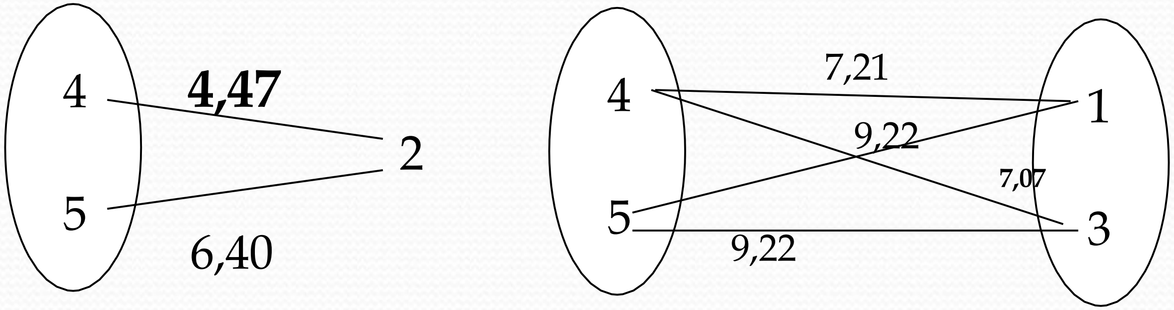
- Yeni uzaklık tablosu,

Gözlemler	(1,3)	2	4	5
(1,3)				
2	2,83			
4	7,07	4,47		
5	9,22	6,40	$\boxed{2,24}$	

- Bu tabloya bakıldığında *Min d(i,j)*=*2,24* olduğu görülür. Bu değerin 4 ve 5 gözlemleri arasındaki uzaklığı görülür. (4,5) yeni bir küme oluşturur. Bu durumda (1,3), 2 ve (4,5) kümeleri arasındaki uzaklık tablosu yeniden oluşturulur.

EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.



EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.

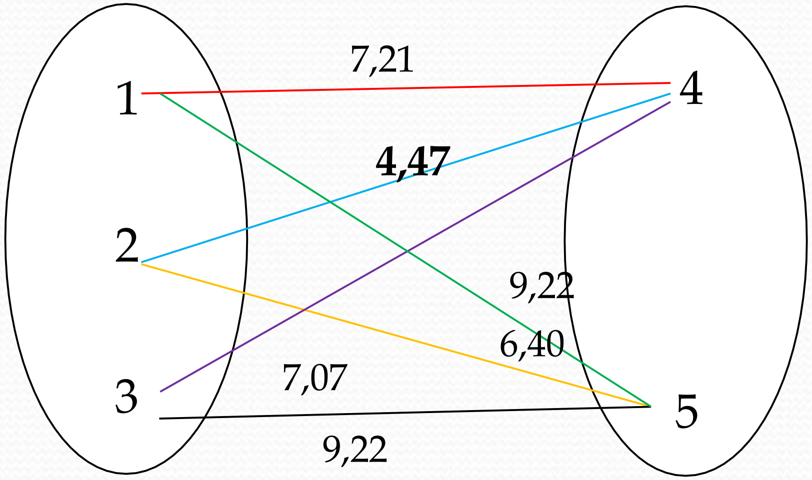
- Bu durumdaki uzaklık tablosu,

Gözlemler	(1,3)	2	(4,,5)
(1,3)			
2	2,83		
(4,5)	7,07	4,47	

- Adım 4. En son uzaklıklar tablosu incelendiğinde Min d(i,j)=2,83 olduğu görülür. O halde bu uzaklık ile ilgili olan 2 gözlemi ile (1,3) kümesi birleştirilecektir. Elde edilen (1,2,3) kümesi ile (4,5) kümesi arasındaki uzaklığı belirlemek için kümeler içindeki her bir değer eşlenir ve en küçük olan belirlenir. En küçük uzaklık 4,47 olduğuna göre iki küme arasındaki uzaklığın bu değer olduğu kabul edilir.

EN YAKIN KOMŞU ALGORİTMASI

Örnek 1.



- Adım 5. Elde edilen iki küme birleştirilerek sonuç küme bulunur. Bu küme (1,2,3,4,5) gözlemlerinden oluşan kümedir.

EN YAKIN KOMŞU ALGORİTMASI

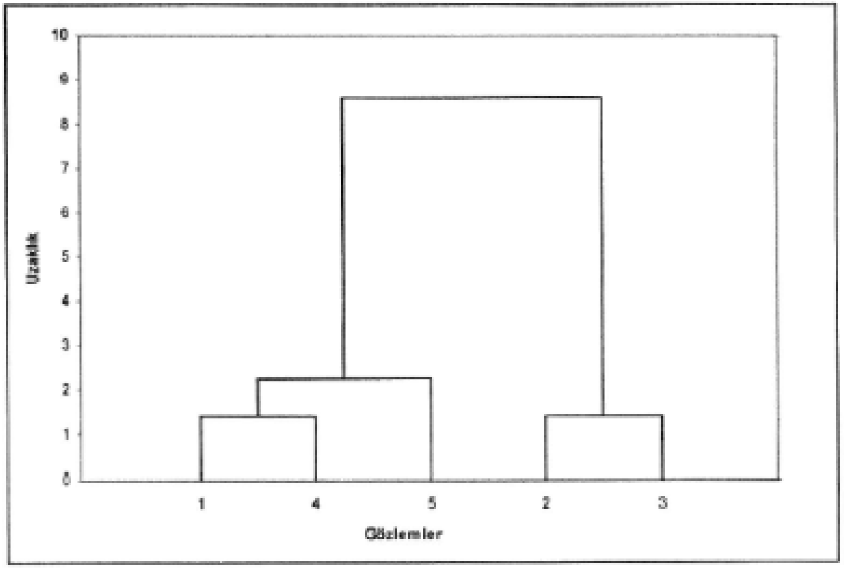
Örnek 1.

Uzaklık düzeyi göz önüne alınarak kümeler şu şekilde belirlenir.

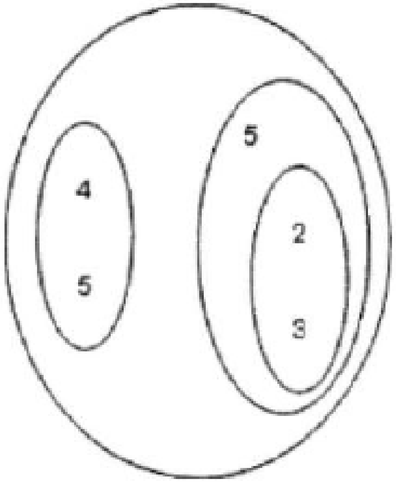
Uzaklık	Kümeler
1,41	(1, 3)
2,24	(4,5)
2,83	(1,2,3)
4,47	(1,2,3,4,5)

EN YAKIN KOMŞU ALGORİTMASI

Dendogram



Sonuç küme



EN YAKIN KOMŞU ALGORİTMASI

Örnek 2. Manhattan Uzaklığı

Aşağıdaki tabloda verilen beş gözlem değeri, en yakın komşu algoritması ile kümelenmek isteniyor.

Gözlemler	X	Y	Z
1	2	3	1
2	4	1	3
3	5	7	3
4	4	8	2
5	3	9	5

EN YAKIN KOMŞU ALGORİTMASI

Örnek 2. Manhattan Uzaklığı

Gözlemler	X	Y	Z
1	2	3	1
2	4	1	3
3	5	7	3
4	4	8	2
5	3	9	5

$$d(i,j) = \sum_{k=1}^{p} \left(\left| x_{ik} - x_{jk} \right| \right)$$

$$d(1,2) = |2-4| + |3-1| + |1-3| = 6$$

$$d(1,3) = |2-5| + |3-7| + |1-3| = 9$$

$$d(1,4) = |2-4| + |3-8| + |1-2| = 8$$

$$\dots \dots \dots$$

$$d(4,5) = |4-3| + |8-9| + |2-5| = 5$$

EN YAKIN KOMŞU ALGORİTMASI

Örnek 2. Manhattan Uzaklığı

$$d(i,j) = \sum_{k=1}^{p} \left(|x_{ik} - x_{jk}| \right)$$

Gözlemler	1	2	3	4	5
1	0.00				
2	6.00	0.00			
3	9.00	7.00	0.00		
4	8.00	8.00	3.00	0.00	
5	11.00	11.00	6.00	5.00	0.00

EN YAKIN KOMŞU ALGORİTMASI

Örnek 3. Minskowski Uzaklığı

- Aşağıdaki tabloda verilen beş gözlem değeri, en yakın komşu algoritması ile kümelenmek isteniyor.

Gözlemler	X	Y	Z
1	2	3	1
2	4	1	3
3	5	7	3
4	4	8	2
5	3	9	5

EN YAKIN KOMŞU ALGORİTMASI

Örnek 3. Minskowski Uzaklığı

Gözlemler	X	Y	Z
1	2	3	1
2	4	1	3
3	5	7	3
4	4	8	2
5	3	9	5

$$d(i,j) = \left|\sum_{k=1}^{p}\left(\left|x_{ik} - x_{jk}\right|^{m}\right)\right|^{\frac{1}{m}} \mathrm{x}$$

$$d(1,2) = [|2-4|^3 + |3-1|^3 + |1-3|^3]^{1/3} = 2.88$$

$$d(1,3) = [|2-5|^3 + |3-7|^3 + |1-3|^3]^{1/3} = 4.63$$

$$d(1,4) = [|2-4|^3 + |3-8|^3 + |1-2|^3]^{1/3} = 5,12$$

$\dots\dots\dots$

$$d(4,5) = [|4-3|^3 + |8-9|^3 + |2-5|^3]^{1/3} = 3,07$$

EN YAKIN KOMŞU ALGORİTMASI

Örnek 3. Minskowski Uzaklığı

Gözlemler	1	2	3	4	5
1	0.00				
2	2.88	0.00			
3	4.63	6.01	0.00		
4	5.12	7.01	1.44	0.00	
5	6.55	8.05	2.88	3.07	0.00

$$d(i,j) = \sum_{k=1}^{p} \left(|x_{ik} - x_{jk}|\right)$$

EN UZAK KOMŞU ALGORİTMASI

- En uzak komşu algoritmasının adımları, en yakın komşu algoritması ile benzer adımları içerir. Gözlemler arasındaki uzaklıklar hesaplanır ve minimum değerli olan birleştirilir. Sonraki küme uzaklıkları tablosu oluşturulurken en uzak mesafe kullanılır.

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

- Aşağıdaki tabloda verilen beş gözlem değeri, en uzak komşu algoritması ile kümelenmek isteniyor.

Gözlemler	X_1	X_2
1	4	2
2	6	4
3	5	1
4	10	6
5	11	8

- Adım1. Öncelikle uzaklık tablosu oluşturulur. Her bir gözlemin birbiriyle arasındaki öklid uzaklığı hesaplanır.

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

$d(1,2) = \sqrt{(4-6)^2 + (2-4)^2}$=2,83

$d(1,3) = \sqrt{(4-5)^2 + (2-1)^2}$=1,41

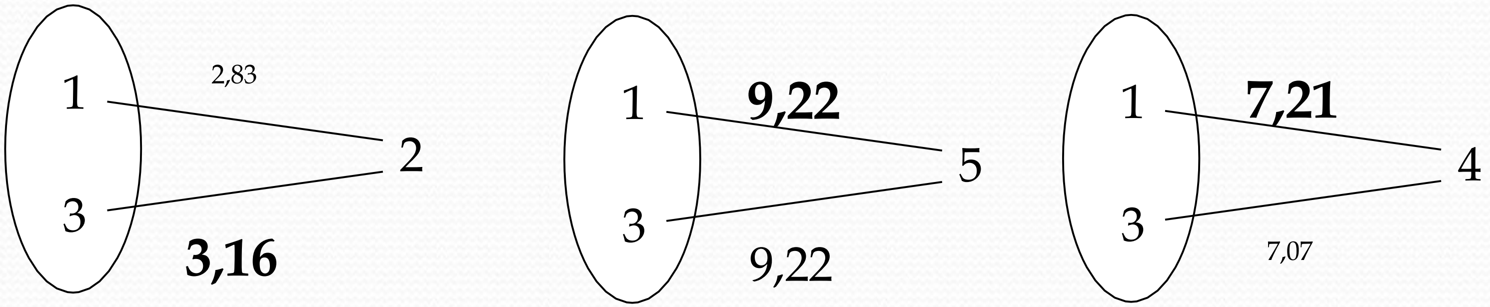
$d(1,4) = \sqrt{(4-10)^2 + (2-6)^2}$=7,21

Gözlemler	1	2	3	4	5
1					
2	2,83				
3	1,41	3,16			
4	7,21	4,47	7,07		
5	9,22	6,40	9,22	2,24	

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

- Adım 2. Uzaklıklar tablosunda Min d(i,j) değerinin 1,41 olduğu görülmektedir. İlgili gözlemler 1 ve 3 gözlemleridir. Bu iki değer birleştirilerek (1,3) kümesi elde edilir. Sonrasında bu kümeye göre uzaklıklar matrisi yeniden incelenir.



EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

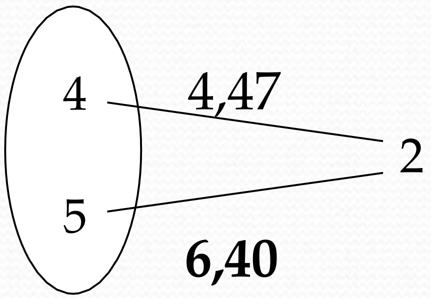
- Yeni uzaklık tablosu,

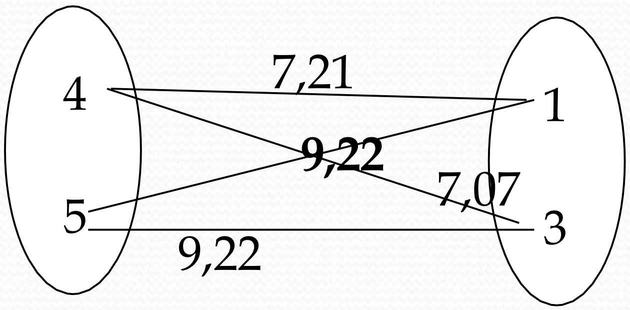
Gözlemler	(1,3)	2	4	5
(1,3)				
2	3,16			
4	7,21	4,47		
5	9,22	6,40	2,24	

- Bu tabloya bakıldığında *Min d(i,j)=2,24* olduğu görülür. Bu değerin 4 ve 5 gözlemleri arasındaki uzaklığı görülür. (4,5) yeni bir küme oluşturur. Bu durumda (1,3), 2 ve (4,5) kümeleri arasındaki uzaklık tablosu yeniden oluşturulur.

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.





EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

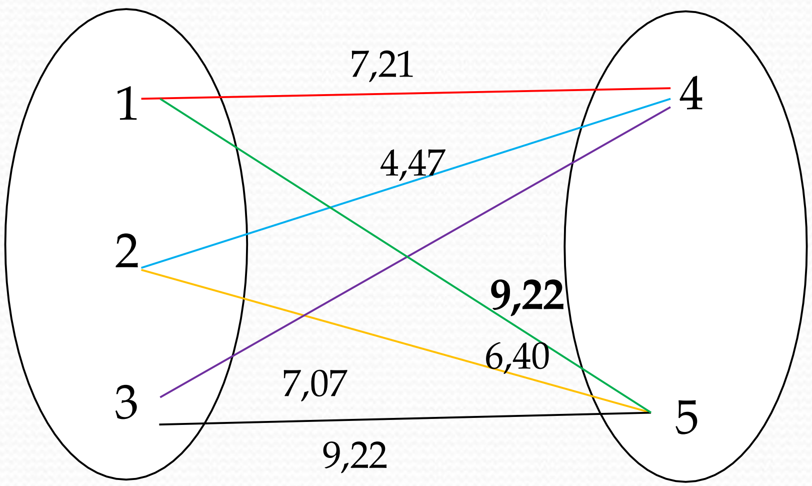
- Bu durumdaki uzaklık tablosu,

Gözlemler	(1,3)	2	(4,,5)
(1,3)			
2	2,83		
(4,5)	9,22	5,40	

- Adım 4. En son uzaklıklar tablosu incelendiğinde Min d(i,j)=2,83 olduğu görülür. O halde bu uzaklık ile ilgili olan 2 gözlemi ile (1,3) kümesi birleştirilecektir. Elde edilen (1,2,3) kümesi ile (4,5) kümesi arasındaki uzaklığı belirlemek için kümeler içindeki her bir değer eşlenir ve en küçük olan belirlenir. En küçük uzaklık 4,47 olduğuna göre iki küme arasındaki uzaklığın bu değer olduğu kabul edilir.

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.



- Adım 5. Elde edilen iki küme birleştirilerek sonuç küme bulunur. Bu küme (1,2,3,4,5) gözlemlerinden oluşan kümedir.

EN UZAK KOMŞU ALGORİTMASI

Örnek 4.

Uzaklık düzeyi göz önüne alınarak kümeler şu şekilde belirlenir.

Uzaklık	Kümeler
1,41	(1, 3)
3,16	(4,5)
2,83	(1,2,3)
4,47	(1,2,3,4,5)

K-Ortalamalar Yöntemi (K-Means)

- Bu yöntemde daha başlangıçta belli sayıdaki küme için toplam ortalama hatayı minimize etmek amaçlanır.

- N noyutlu uzayda N örnekli kümelerin verildiği kabul edilsin. Bu uzay $\{C_1, C_2, \ldots, C_k\}$ biçiminde K kümeye ayrılsın. O zaman $\sum n_k = N \ (k = 1,2, \ldots, k)$ olmak üzere C_k kümesinin ortalama vektörü M_kaşağıdaki gibi hesaplanır.

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_{ik}$$

- Burada X_k değeri C_k kümesine ait olan *i.* örnektir. C_k kümesi için kare-hata, her bir C_k örneği ile onun merkezi (centroid) arasındaki Öklid uzaklıkları toplamıdır. Bu hataya «küme içi değişme» adı da verilir.

K-Ortalamalar Yöntemi (K-Means)

- Küme içi değişmeler şu şekilde hesaplanır.

$$e_i^2 = \sum_{i=1}^{n_k} (X_{ik} - M_k)^2$$

- K kümesini içeren bütün kümeler uzayı için kare-hata içindeki değişmelerin toplamıdır. O halde söz konusu kare-hata şu şekilde hesaplanır.

$$E_k^2 = \sum_{k=1}^{K} e_k^2$$

- Kare-hata kümeleme yönteminin amacı verilen K değeri için E_k^2 değerini minimize eden K kümelerini bulmaktır. O halde k-ortalamalar algoritmasında E_k^2 değerinin bir önceki iterasyona göre azalması beklenir.

K-Means Algoritmasının Adımları

- K-Means algoritmasına başlamadan önce k küme sayısının belirlenmesi gerekir. Sonra aşağıdaki işlemler gerçekleştirilir.

1. Her bir kümenin merkezi belirlenir. Bu merkezler $M_1, M_2, \dots, M_k$ biçimindedir.
2. $e_1, e_2, \dots, e_k$küme içi değişmeler hesaplanır. Bu değişmeler E_k^2 değeri bulunur.
3. $M_k$merkez değerleri ile gözlem değerleri arasındaki uzaklıklar hesaplanır. Bir gözlem değeri hangi merkeze yakın ise o merkez ile ilgili küme içine dahil edilir.
4. Yukarıdaki 2. ve 3. adımlar kümelerde değişiklik olmayınca kadar devam ettirilir.

K-Means Algoritmasının Özellikleri

- Gerçeklemesi kolay
  - Karmaşıklığı diğer kümeleme yöntemlerine göre az
- K-Means algoritması bazı durumlarda iyi sonuç vermeyebilir.
  - Veri grupları farklı boyutlarda ise
  - Veri gruplarının yoğunlukları farklı ise
  - Veri gruplarının şekli küresel değilse
  - Veri içinde aykırılıklar varsa

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- Aşağıdaki gözlem değerleri k-ortalamalar yöntemi ile kümelenmek isteniyor.

Gözlemler	Değişken1	Değişken2
X_1	4	2
X_2	6	4
X_3	5	1
X_4	10	6
X_5	11	8

- Kümelerin sayısı başlangıçta k=2 kabul edilir. Rasgele iki küme belirlenir.

$$C_1 = \{X_1, X_2, X_4\}$$
$$C_2 = \{X_3, X_5\}$$

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

Gözlemler	Değişken1	Değişken2	Küme Üyeliği
X_1	4	2	C_1
X_2	6	4	C_1
X_3	5	1	C_2
X_4	10	6	C_1
X_5	11	8	C_2

- Adım 1. a) Belirtilen iki kümenin merkezleri şu şekilde hesaplanır.

$$M_1 = \left\{\frac{4+6+10}{3}, \frac{2+4+6}{3}\right\} = \{6.67, 4.00\} \qquad M_2 = \left\{\frac{5+11}{2}, \frac{1+8}{2}\right\} = \{8.00, 4.50\}$$

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- b) Küme içi değişmeler şu şekilde hesaplanır.

$$e_1^2 = [(4-6{,}67)^2 + (2-4.0)^2] + [(6-6{,}67)^2 + (4-4.0)^2] + [(10-6{,}67)^2 + (6-4.0)^2] = 26{,}67$$

$$e_2^2 = [(5-8)^2 + (1-4.5)^2] + [(11-8)^2 + (8-4.5)^2] = 42{,}50$$

- Bu durumda toplam kare-hata şu şekilde hesaplanır.

$$E^2 = e_1^2 + e_2^2 = 26{,}67 + 42{,}50 = 69{,}17$$

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- C) M_1 ve M_2 merkezlerinden olan uzaklıkların minimum olması istendiğinden aşağıdaki hesaplamalar yapılır. Öklid uzaklık formülü kullanılarak söz konusu mesafeler hesaplanır. Örneğin (M_1, X_1) noktaları arasındaki uzaklık M_1={6.67, 4.00} ve X_1={4, 2} olduğuna göre şu şekilde hesaplanır.

$$d(M_1, X_1) = \sqrt{(6{,}67 - 4)^2 + (4 - 2)^2} = 3{,}33$$

$$d(M_2, X_1) = \sqrt{(8 - 4)^2 + (4{,}5 - 2)^2} = 4{,}72$$

Bu işlemler sonucunda $X1$ gözlem değerinin $M1$ ve $M2$ merkezlerine olan uzaklıkları göz önüne alındığında $d(M1,X1)<d(M2,X1)$ olduğu görülür. Bu durumda $M1$ merkezinin $X1$ gözlem değerine daha yakın olduğu anlaşılır. O halde $X1 \in C1$ olarak kabul edilir. Benzer biçimde tüm gözlem değerleri için tablo oluşturulur.

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

Gözlemler	M1'den Uzaklık	M2'den Uzaklık	Küme Üyeliği
X_1	$d(M_1, X_1) = 3{,}33$	$d(M_2, X_1) = 4{,}72$	C_1
X_2	$d(M_1, X_2) = 0{,}67$	$d(M_2, X_2) = 2{,}06$	C_1
X_3	$d(M_1, X_3) = 3{,}43$	$d(M_2, X_3) = 4{,}61$	C_1
X_4	$d(M_1, X_4) = 3{,}89$	$d(M_2, X_4) = 2{,}50$	C_2
X_5	$d(M_1, X_5) = 5{,}90$	$d(M_2, X_5) = 4{,}61$	C_2

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- Bu durumda yeni kümeler şu şekilde olacaktır.

$$C_1 = \{X_1, X_2, X_3\}$$
$$C_2 = \{X_4, X_5\}$$

- Adım 2. Yukarıda belirtilen iki kümenin merkezleri şu şekilde hesaplanır.

$$M_1 = \left\{\frac{4+6+5}{3}, \frac{2+4+1}{3}\right\} = \{5.00, 2.33\}$$

$$M_2 = \left\{\frac{10+11}{2}, \frac{6+8}{8}\right\} = \{10.5, 7\}$$

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- b) Küme içi değişmeler şu şekilde hesaplanır.

$$e_1^2 = [(4-5)^2 + (2-2.33)^2] + [(6-5)^2 + (4-2.33)^2] + [(5-5)^2 + (1-2{,}33)^2] = 9{,}33$$

$$e_2^2 = [(10-10{,}5)^2 + (6-7)^2] + [(11-10{,}5)^2 + (8-7)^2] = 2{,}50$$

- Bu durumda toplam kare-hata şu şekilde hesaplanır.

$$E^2 = e_1^2 + e_2^2 = 9{,}33 + 2{,}50 = 11{,}83$$

Bu değerin bir önceki iterasyonda elde edilen E^2=69,17 değerinden daha küçük olduğu görülmektedir.

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- $M1$ ve $M2$ merkezlerinden gözlem değerlerine olan uzaklıklar hesaplanır. Bunun sonucunda $d(M1,X1)<d(M2,X1)$ olduğu görülür. Bu durumda $M1$ merkezinin $X1$ gözlem değerine daha yakın olduğu anlaşılır. O halde $X1 \in C1$ olarak kabul edilir. Benzer biçimde tüm gözlem değerleri için tablo oluşturulur.

Gözlemler	M1'den Uzaklık	M2'den Uzaklık	Küme Üyeliği
X_1	$d(M_1, X_1) = 1{,}05$	$d(M_2, X_1) = 8{,}20$	C_1
X_2	$d(M_1, X_2) = 1{,}94$	$d(M_2, X_2) = 5{,}41$	C_1
X_3	$d(M_1, X_3) = 1{,}33$	$d(M_2, X_3) = 8{,}14$	C_1
X_4	$d(M_1, X_4) = 6{,}20$	$d(M_2, X_4) = 1{,}12$	C_2
X_5	$d(M_1, X_5) = 8{,}25$	$d(M_2, X_5) = 1{,}12$	C_2

K-Ortalamalar Yöntemi (K-Means)

Örnek 5.

- Bu durumda yeni kümeler şu şekilde oluşacaktır.

$$C_1 = \{X_1, X_2, X_3\}$$
$$C_2 = \{X_4, X_5\}$$

- Kümelerde önceki adıma göre herhangi bir değişme olmadığı için iterasyona son verilir.

Kaynaklar

Deperlioğlu, Ö. ve Köse, U. (2023), **PYTHON ile Yapay Zekaya Giriş**, Seçkin Yayıncılık, Ankara.

Alp, S. ve Öz. E. , **Makine Öğrenmesinde Sınıflandırma Yöntemleri ve R Uygulamaları,** Nobel Yayınları, Ankara.