

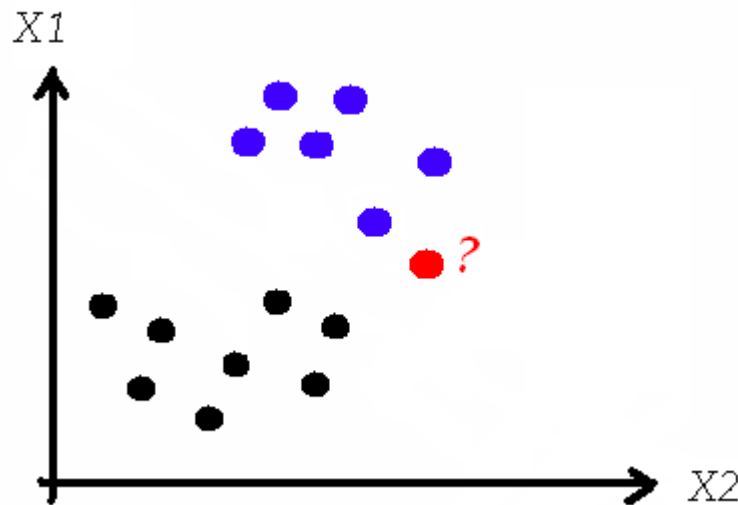
Makine Öğrenmesi

- Çok büyük miktardaki verilerin elle işlenip analiz edilmesi mümkün değildir.
 - Bu tür problemlere çözüm bulmak amacıyla makine öğrenmesi metotları geliştirilmiştir.
- Bu metotlar **geçmişteki verileri kullanarak** veriye en uygun **modeli** bulmaya çalışırlar.
- Yeni gelen verileri de bu modele göre analiz edip sonuç üretirler.

Metot türleri

- Farklı uygulamaların analizlerden farklı beklentileri olmaktadır.
- Makine öğrenmesi metotlarını bu beklentilere göre sınıflandırmak mümkündür.

Sınıflandırma



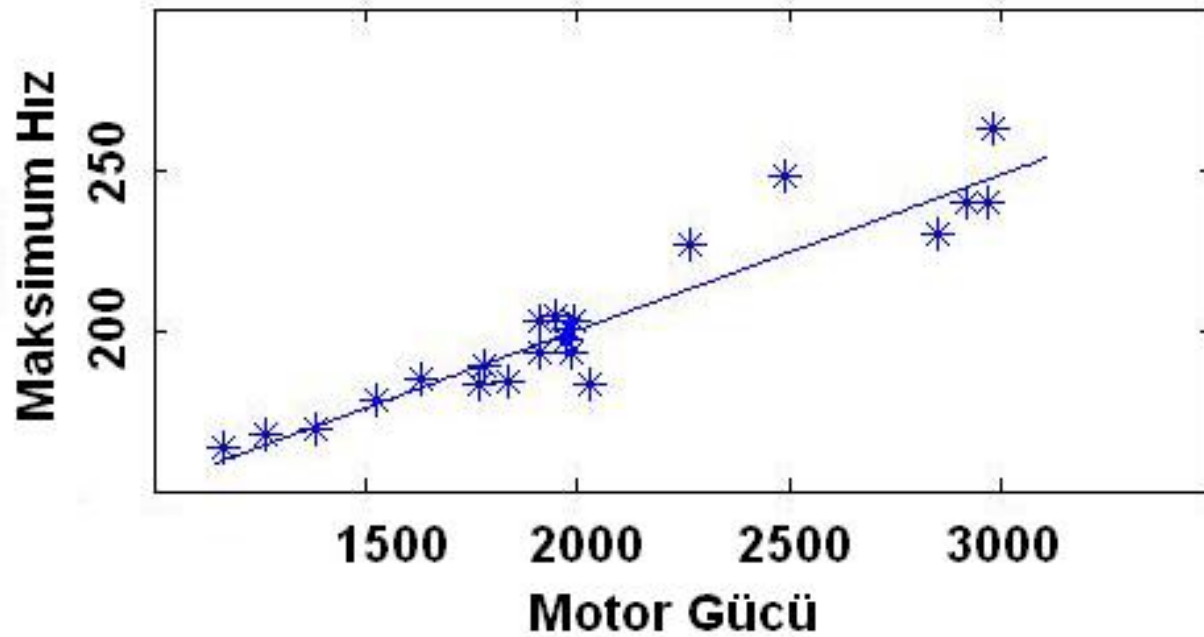
Kırmızı hangi sınıftan?

Kümeleme

- 256 rengi 16 renge nasıl indiririz?



Regresyon: Eğri Uydurma



Birliktelik Kuralları Keşfi

- Sepet analizi
 - hangi ürünler birlikte satılıyor?
- Raf düzenlemesi
 - hangi ürünler yan yana konmalı?
- Promosyonlar
 - neyin yanında ne verilmeli?



Makine Öğrenmesinin

Günlük Hayatımızdaki Uygulamaları

El yazısı / Kitap Yazısı Tanıma HCR /OCR



İşlem: Şekillerin hangi harf olduğunu tahmin etme

Kredi Taleplerini Değerlendirme

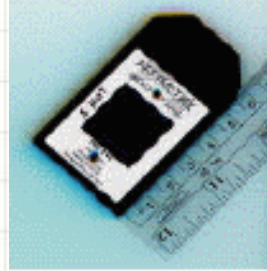
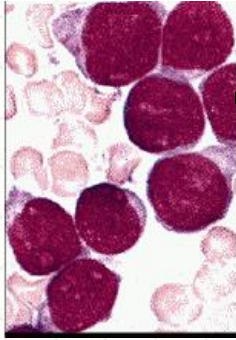
- Birisi bankadan borç ister.
- Banka borcu versin mi, vermesin mi?
- Geri ödeme nasıl olsun?

e-ticaret

- Birisi amazon.com dan bir kitap yada ürün alıyor.
 - Müşteriye alması muhtemel kitaplar önerilebilir.
 - Ama nasıl?
 - Kitapları
 - konularına
 - yazarlarına
 - birlikte satışlarına
- göre kümelemek?

Gen Mikrodizilimleri

ALL



AML

- 100 kişinin (hasta/sağlam) elimizde gen dizilimleri var.
- Bu dizilimleri analiz ederek başka birinin hasta olup olmadığını ya da hastalığının türünü öğrenebilir miyiz?
- En iyi tedaviyi önerebilir miyiz?
- Nasıl? Elimizde hangi bilgiler olmalı?

Bu adam kim? İeri alınsın mı?



Bu adam kim?

Bu adam havaalanında mı?

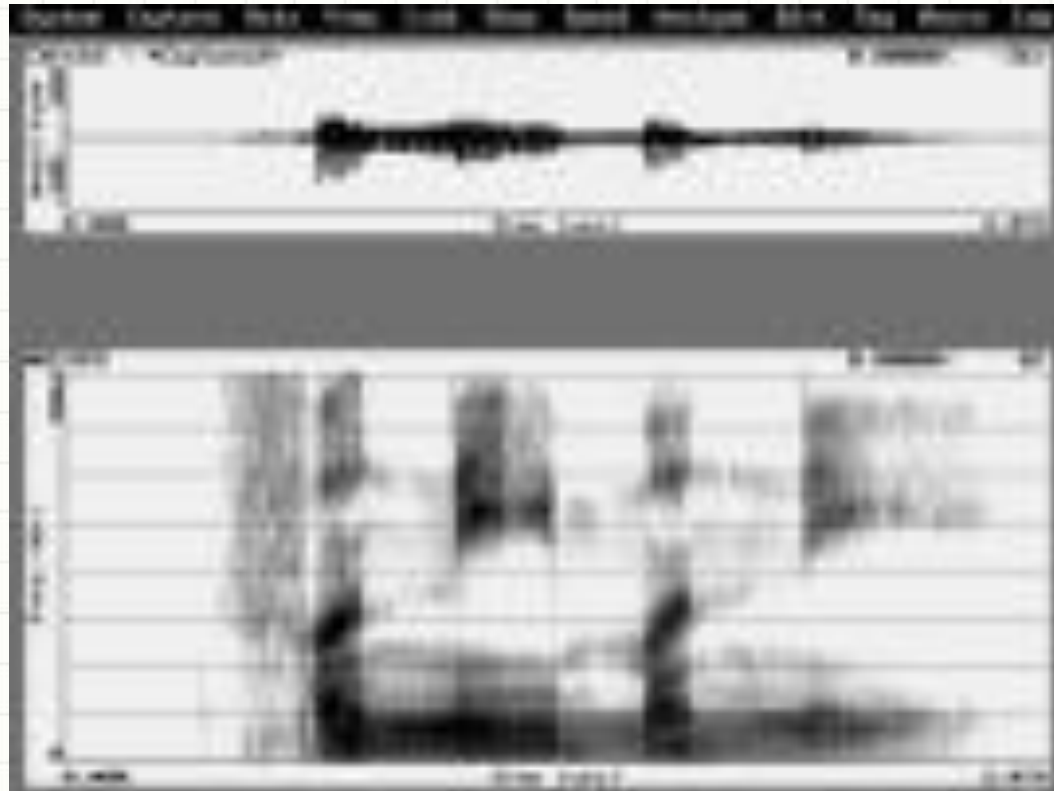


Bu parmak izi kimin?

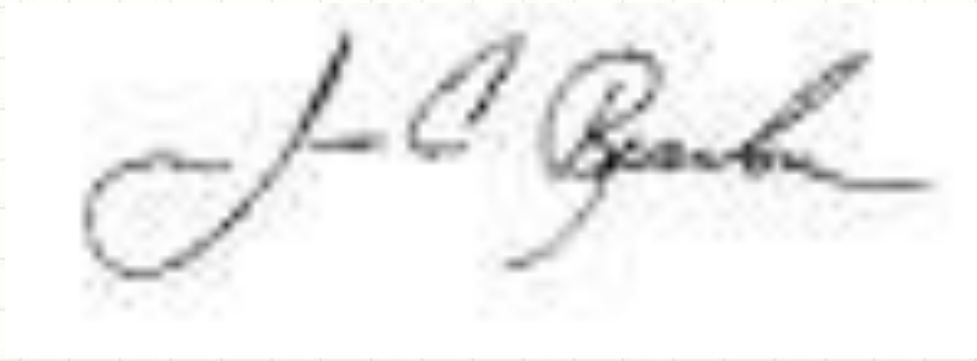
Bu adamı tutuklayalım mı?



Bu ses kimin?
Bu ses ne diyor?



Bu imza kimin? Yoksa taklit mi?



Taklit olup olmadığını nasıl anlarız?

Bu metnin konusu nedir?

Bu mail spam mi?



Anti spam yazılımları nasıl çalışır?

Spamciler nasıl çalışıyor?

Yeni nesil spam mailler: Mesaj resimde,
Metinde ise *anti spamlardan kaçmak* için gereken kelimeler var.

Olađan dıřı bir durum var mı? Güvenlik kamerası kayıtları



Kamera kaydındaki kiři ne anlatıyor?

İřaret dili tanıma



Sonuç: İletişimin artması.

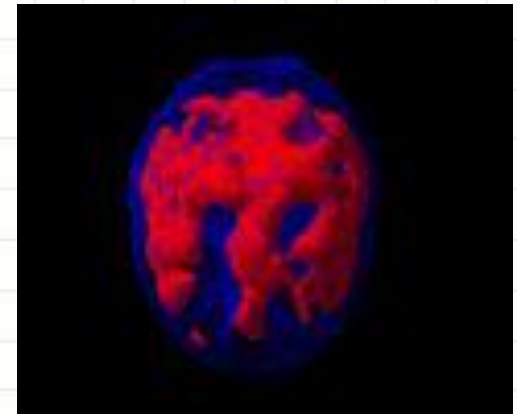
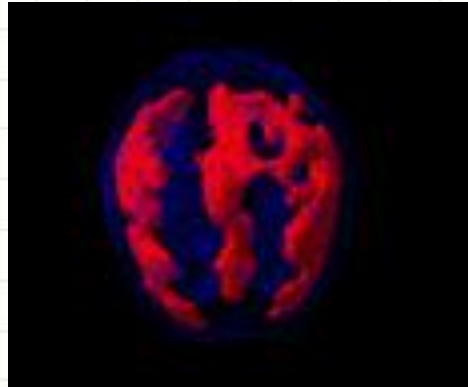
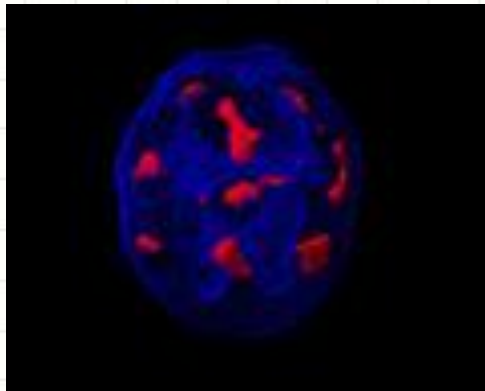
Adalet

Çin’de pilot uygulama:

- Bir şehrin mahkeme hakimleri -> bir bilgisayar programı
- Amaç: Daha adil bir dünya
 - Aynı özelliklere sahip davalarda aynı kararların alınması
- Sistemin eğitimi için neler gerekli?
 - Milyonlarca/Milyarlarca davaya ait verilerin kategorilenmesi

Beyin Aktiviteleri

- İnsanların
 - değişik şeyler düşünürken,
 - değişik duygulara sahipken,
 - problem çözerken kibeyin aktiviteleri kaydedilir.
- Amaç?



Öğrenmenin geleceği

- Öğrenme modülleri birçok uygulamaya girmiş durumda.
- Ticari olarak satılan (gerçek dünya için tasarlanmış) birçok ürün var.
- Bu tür ürün ve modüllerin sayısı giderek artmakta.
- Çözüm bekleyen ve büyük potansiyelleri olan birçok problem var.

Verilerin öğrenmede kullanımı

- Farklı **türdeki** pek çok veri işlemek gerek
 - Sayısal
 - Sayısal olmayan (kategorik, sözel, vb.)
- Veriyi sayısallaştırmak gerek

Verilerin sayısallaştırılması

- Resim:
 - Resmin her bir pikselinin değeri 0–255 arası sayılara çevrilir. Renkli resimler 3 adet, siyah beyazlar 1 adet **en*boy** büyüklüğünde matrisle ifade edilir.
- Metin:
 - Metindeki harfler, heceler ve kelimeler genelde frekanslarına göre kodlanarak sayılara çevrilir.
- Ses:
 - Genlik ve frekansın zaman içindeki değişimiyle kodlanır.

Sınıflandırma (Classification)

- **Eğitici (supervised) sınıflandırma:**
Sınıflandırma: Sınıf sayısı ve bir grup örneğin hangi sınıfa ait olduğunu bilinir
- **Eğitici (unsupervised) sınıflandırma:**
Kümeleme: Hangi nesnenin hangi sınıfa ait olduğu ve grup sayısı belirsizdir.

Sınıflandırma Tanımı

Sınıflandırmanın temel kuralları:

- Öğrenme **eğitici**lidir
- **Veri setinde** bulunan her örneğin bir dizi **özniteliği** vardır ve bu niteliklerden biri de **sınıf** bilgisidir.
- Hangi sınıfa ait olduğu bilinen nesneler (**öğrenme kümesi**: training set) ile bir model oluşturulur
- Oluşturulan model öğrenme kümesinde yer almayan nesneler (**deneme kümesi**: test set) ile denenerek başarısı ölçülür

Örnek Veri Kümesi

Table 3.1 • The Credit Card Promotion Database

Income Range	Life Insurance Promotion	Credit Card Insurance	Sex	Age
40–50K	No	No	Male	45
30–40K	Yes	No	Female	40
40–50K	No	No	Male	42
30–40K	Yes	Yes	Male	43
50–60K	Yes	No	Female	38
20–30K	No	No	Female	55
30–40K	Yes	Yes	Male	35
20–30K	No	No	Male	27
30–40K	No	No	Male	43
30–40K	Yes	No	Female	41
40–50K	Yes	No	Female	43
20–30K	Yes	No	Male	29
50–60K	Yes	No	Female	39
40–50K	No	No	Male	55
20–30K	Yes	Yes	Female	19

Örnekler
(samples)

Örnek Veri Kümesi

Öznitelikler
(features)

The Credit Card Promotion Database

Income
Range

40–50K
30–40K
40–50K
30–40K
50–60K
20–30K
30–40K
20–30K
30–40K
30–40K
40–50K
20–30K
50–60K
40–50K
20–30K

Life Insurance
Promotion

No
Yes
No
Yes
Yes
No
Yes
No
No
Yes
Yes
Yes
Yes
No
Yes

Credit Card
Insurance

No
No
No
Yes
No
No
Yes
No
No
No
No
No
No
No
Yes

Sex

Male
Female
Male
Male
Female
Female
Male
Male
Male
Female
Female
Male
Female
Male
Female

Age

45
40
42
43
38
55
35
27
43
41
43
29
39
55
19

Sınıflandırma Yöntemleri:

- Karar Ağaçları (Decision Trees)
- Örnek Tabanlı Yöntemler (Instance Based Methods): k en yakın komşu (k nearest neighbor)
- Bayes Sınıflandırıcı (Bayes Classifier)
- Yapay Sinir Ağları (Artificial Neural Networks)
- Genetik Algoritmalar (Genetic Algorithms)
- ...

Karar Ağaçları

Karar ağaçları eğitici öğrenme için çok yaygın bir yöntemdir. Algoritmanın adımları:

1. T öğrenme kümesini oluştur
2. T kümesindeki örnekleri en iyi ayıran özniteliği belirle
3. Seçilen öznitelik ile ağacın bir düğümünü oluştur ve bu düğümden çocuk düğümleri veya ağacın yapraklarını oluştur. Çocuk düğümlere ait alt veri kümesinin örneklerini belirle
4. 3. adımda yaratılan her alt veri kümesi için
 - Örneklerin hepsi aynı sınıfa aitse
 - Örnekleri bölecek nitelik kalmamışsa
 - Kalan özniteliklerin değerini taşıyan örnek yoksa işlemi sonlandır.

Diğer durumda alt veri kümesini ayırmak için 2. adımdan devam et.

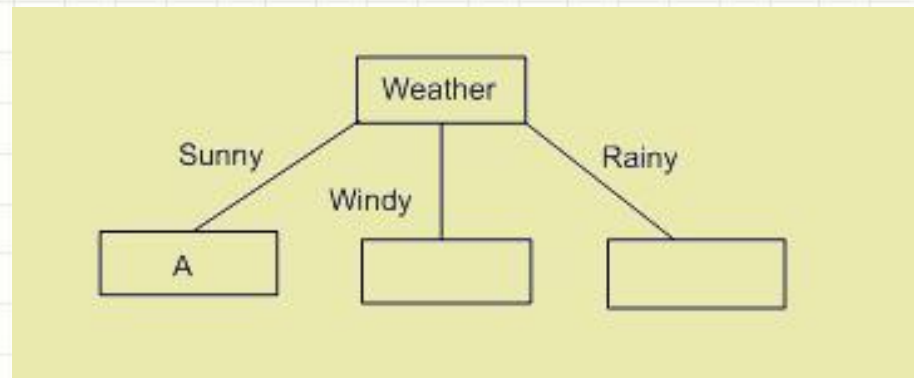
Karar Ağaçları: Haftasonu örneği

1. Adım: Veri setinden T öğrenme kümesi oluşturulur.

Weekend (Example)	Weather	Parents	Money	Decision (Category)
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay in
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

Karar Ağaçları: Haftasonu örneği

2. Adım: Veri setindeki en ayırt edici nitelik belirlenir ve ağacın kökü olarak alınır.



3. Adım: Ağacın çocuk düğümü olan A düğümüne ait alt veri kümesi belirlenir.

Weekend (Example)	Weather	Parents	Money	Decision (Category)
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W10	Sunny	No	Rich	Tennis

Karar Ağaçları: En ayırt edici nitelik nasıl bulunur?

- En ayırt edici niteliği belirlemek için her öz nitelik için **bilgi kazancı** ölçülür.
- Bilgi Kazancı ölçümünde **Entropy** kullanılır.
- **Entropy** rastgeleliği, belirsizliği ve beklenmeyen durumun ortaya çıkma olasılığını gösterir.

Karar Ağaçları

Bilgi Kazancı: Entropy

The information entropy of a discrete random variable X , that can take on possible values $\{x_1 \dots x_n\}$ is

$$\begin{aligned} H(X) = E(I(X)) &= \sum_{i=1}^n p(x_i) \log_2 (1/p(x_i)) \\ &= - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \end{aligned}$$

where

$I(X)$ is the information content or self-information of X , which is itself a random variable; and $p(x_i) = \Pr(X=x_i)$ is the probability mass function of X .

Karar Ağaçları: Entropy

- Haftasonu veri kümesindeki (T kümesi) 10 örnekten
 - 6 örnek için karar sinema
 - 2 örnek için karar tenis oynamak
 - 1 örnek için karar evde kalmak ve
 - 1 örnek için karar alışverişe gitmek olduğuna göre

- Entropy:

$$H(T) = - (6/10) \log_2(6/10) - (2/10) \log_2(2/10) - (1/10) \log_2(1/10) - (1/10) \log_2(1/10)$$

$$H(T) = 1,571$$

Karar Ağaçları: Bilgi Kazancı

- A özniteliğinin T veri kümesindeki bilgi kazancı:

$$\text{Gain}(T,A) = \text{Entropy}(T) - [\sum P(v) * \text{Entropy}(T(v))]$$

- v: A'nın muhtemel değerleri
- P(v): v'nin bulunma olasılığı

$$P(v) = |T(v)| / |T|$$

Entropy(T(v)): v'nin sınıflara göre entropisi

Karar Ağaçları: Bilgi Kazancı

- $\text{Gain}(T, \text{weather}) = ?$
 - Sunny=3 (1 Cinema, 2 Tennis) $\rightarrow P_{\text{sunny}} = 3/10$
 - Windy=4 (3 Cinema, 1 Shopping) $\rightarrow P_{\text{windy}} = 4/10$
 - Rainy=3 (2 Cinema, 1 Stay in) $\rightarrow P_{\text{rainy}} = 3/10$
 - $\text{Entropy}(T_{\text{sunny}}) = - (1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = 0,918$
 - $\text{Entropy}(T_{\text{windy}}) = - (3/4) \log_2 (3/4) - (1/4) \log_2 (1/4) = 0,811$
 - $\text{Entropy}(T_{\text{rainy}}) = - (2/3) \log_2 (2/3) - (1/3) \log_2 (1/3) = 0,918$
- $\text{Gain}(T, \text{weather}) = \text{Entropy}(T) - ((P(\text{sunny}) \text{Entropy}(T_{\text{sunny}}) + P(\text{windy}) \text{Entropy}(T_{\text{windy}}) + P(\text{rainy}) \text{Entropy}(T_{\text{rainy}}))$
 $= 1,571 - ((3/10) \text{Entropy}(T_{\text{sunny}}) + (4/10) \text{Entropy}(T_{\text{windy}}) + (3/10) \text{Entropy}(T_{\text{rainy}}))$
- $\text{Gain}(T, \text{weather}) = \mathbf{0,70}$

Karar Ağaçları: Bilgi Kazancı

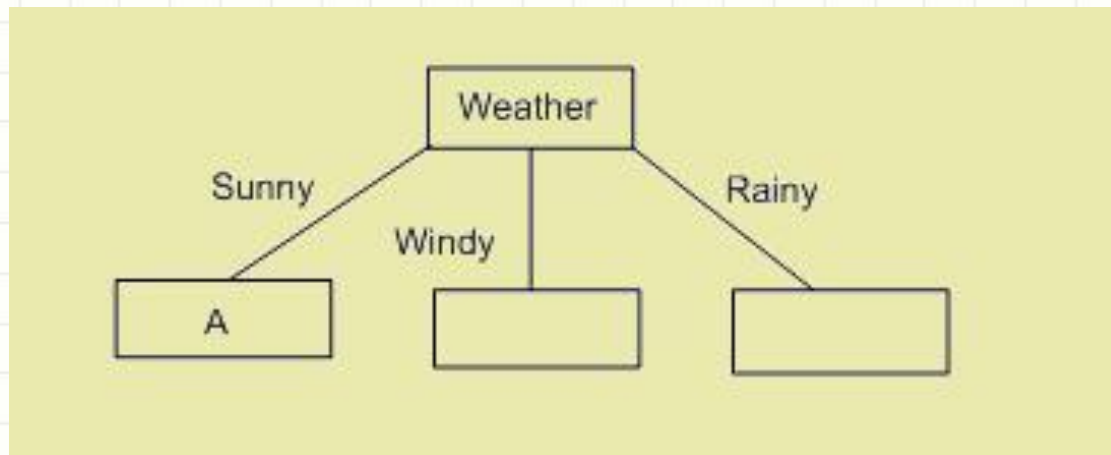
- $\text{Gain}(T, \text{parents}) = ?$
 - Yes=5 (5 Cinema)
 - No=5 (2 Tennis, 1 Cinema, 1 Shopping, 1 Stay in)
 - $\text{Entropy}(T_{\text{yes}}) = - (5/5) \log_2 (5/5) = 0$
 - $\text{Entropy}(T_{\text{no}}) = - (2/5) \log_2 (2/5) - 3(1/5) \log_2 (1/5) = 1,922$
- $\text{Gain}(T, \text{parents}) = \text{Entropy}(T) - ((P(\text{yes})\text{Entropy}(T_{\text{yes}}) + P(\text{no})\text{Entropy}(T_{\text{no}}))$
 $= 1,571 - ((5/10)\text{Entropy}(T_{\text{yes}}) + (5/10)\text{Entropy}(T_{\text{no}}))$
 $\text{Gain}(T, \text{parents}) = \mathbf{0,61}$

Karar Ağaçları: Bilgi Kazancı

- $\text{Gain}(T, \text{money}) = ?$
 - Rich=7 (3 Cinema, 2 Tennis, 1 Shopping, 1 Stay in)
 - Poor=3 (3 Cinema)
 - $\text{Entropy}(T_{\text{rich}}) = 1,842$
 - $\text{Entropy}(T_{\text{poor}}) = 0$
- $\text{Gain}(T, \text{money}) = \text{Entropy}(T) - ((P(\text{rich})\text{Entropy}(T_{\text{rich}}) + P(\text{poor})\text{Entropy}(T_{\text{poor}})))$
 $= 1,571 - ((7/10)\text{Entropy}(T_{\text{rich}}) + (3/10)\text{Entropy}(T_{\text{poor}}))$
 $\text{Gain}(T, \text{money}) = \mathbf{0,2816}$

Karar Ağaçları: Bilgi Kazancı

- $\text{Gain}(T, \text{weather}) = 0,70$
- $\text{Gain}(T, \text{parents}) = 0,61$
- $\text{Gain}(T, \text{money}) = 0,2816$
- **Weather** özelliği en büyük bilgi kazancını sağladığı için (en ayırt edici özellik olduğundan) ağacın kökünde yer alacak özellik olarak seçilir.



Karar Ağaçları: Hafta sonu örneği

3. Adım: Ağacın çocuk düğümü olan A düğümüne ait alt veri kümesi belirlenir.

Weekend (Example)	Weather	Parents	Money	Decision (Category)
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W10	Sunny	No	Rich	Tennis

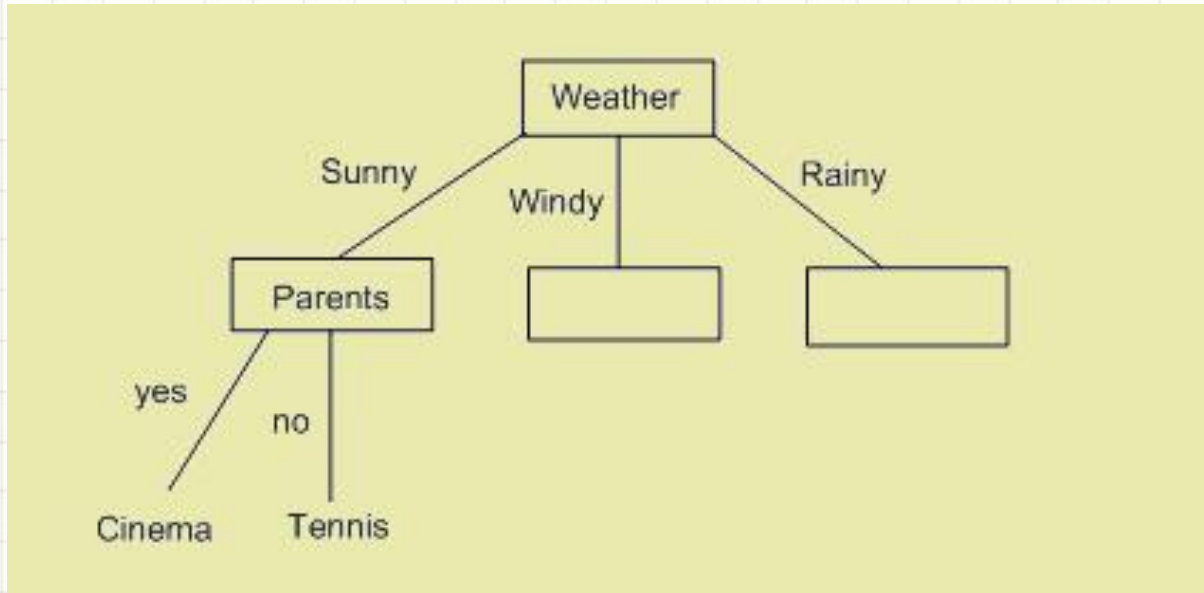
- Her alt küme için tekrar bilgi kazancı hesaplanarak en ayırt edici özellik belirlenir.

$$\begin{aligned}\text{Gain}(S_{\text{sunny}, \text{parents}}) &= 0.918 - (|S_{\text{yes}}|/|S|) * \text{Entropy}(S_{\text{yes}}) - (|S_{\text{no}}|/|S|) * \text{Entropy}(S_{\text{no}}) \\ &= 0.918 - (1/3) * 0 - (2/3) * 0 = 0.918\end{aligned}$$

$$\begin{aligned}\text{Gain}(S_{\text{sunny}, \text{money}}) &= 0.918 - (|S_{\text{rich}}|/|S|) * \text{Entropy}(S_{\text{rich}}) - (|S_{\text{poor}}|/|S|) * \text{Entropy}(S_{\text{poor}}) \\ &= 0.918 - (3/3) * 0.918 - (0/3) * 0 = 0.918 - 0.918 = 0\end{aligned}$$

Karar Ağaçları: Hafta sonu örneği

- Yeni düğüm için en ayırt edici özellik **Parents** olarak belirlenmiştir. Bu işlemler her düğüm için aşağıdaki durumlardan biri oluşuncaya kadar devam eder
 - Örneklerin hepsi aynı sınıfa ait
 - Örnekleri bölecek özellik kalmamış
 - Kalan özelliklerin değerini taşıyan örnek yok



Sınıflandırma Yöntemleri:

- Karar Ağaçları (Decision Trees)
- Örnek Tabanlı Yöntemler (Instance Based Methods): k en yakın komşu (knn)
- Bayes Sınıflandırıcı (Bayes Classifier)
- Yapay Sinir Ağları (Artificial Neural Networks)
- Genetik Algoritmalar (Genetic Algorithms)
- ...

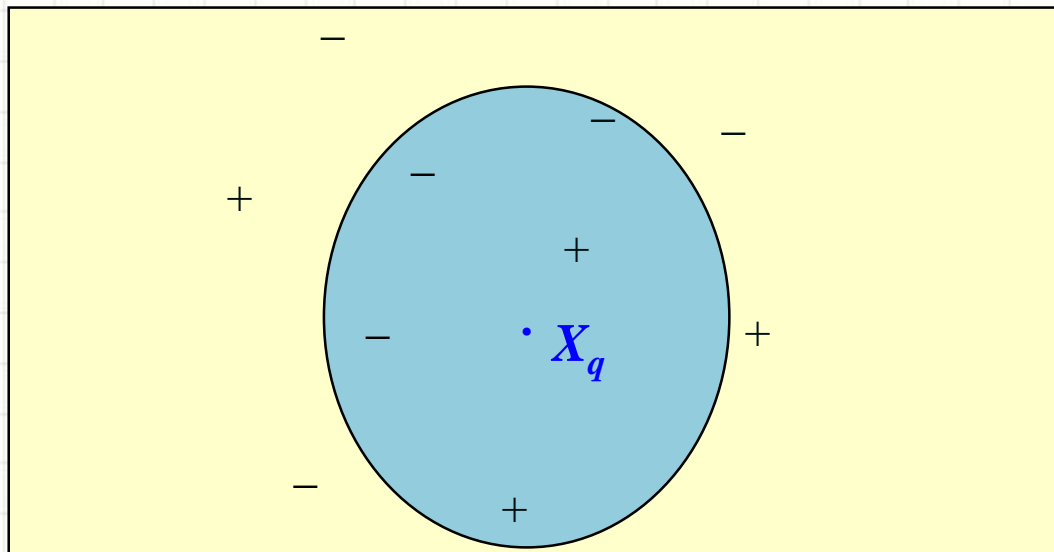
k - en yakın komşu

- Bütün örnekler n boyutlu uzayda bir nokta olarak alınır.
- Hangi sınıfa ait olduğu bilinmeyen X_q örneği için;
 - **Öklid** mesafesi kullanılarak en yakın komşular belirlenir, $\text{dist}(X_q, X_1)$
- X_q örneği, kendisine en yakın **k** örneğin (çoğunluğunun) sınıfına aittir denir.

k - en yakın komşu: Örnek

- X_q örneği ;
 - 1-en yakın komşuya göre **pozitif**
 - 5-en yakın komşuya göre **negatif**

Olarak sınıflandırılır.



Sınıflandırma Yöntemleri:

- Karar Ağaçları (Decision Trees)
- Örnek Tabanlı Yöntemler (Instance Based Methods): k en yakın komşu (knn)
- **Bayes Sınıflandırıcı (Bayes Classifier)**
- Yapay Sinir Ağları (Artificial Neural Networks)
- Genetik Algoritmalar (Genetic Algorithms)
- ...

Bayes Sınıflandırıcılar

- Bayes Sınıflayıcı: Bayes teoremine göre istatistiksel kestirim yapar.
- Bir örneğin sınıf üyelik olasılığını kestirir.
- *Naïve Bayesian sınıflandırıcı* (simple Bayesian classifier) oldukça başarılı bir sınıflayıcıdır.

Bayes Kuralı

$p(\mathbf{x} | C_j)$: Sınıf j'den bir örneğin x olma olasılığı

$P(C_j)$: Sınıf j'nin görülme olasılığı (ilk olasılık)

$p(\mathbf{x})$: Herhangi bir örneğin x olma olasılığı

$P(C_j | \mathbf{x})$: x olan bir örneğin sınıf j'den olma olasılığı (son olasılık)

$$P(C_j | \mathbf{x}) = \frac{p(\mathbf{x} | C_j)P(C_j)}{p(\mathbf{x})} = \frac{p(\mathbf{x} | C_j)P(C_j)}{\sum_k p(\mathbf{x} | C_k)P(C_k)}$$

Naïve Bayes sınıflandırıcı

- T öğrenme kümesinde bulunan her örnek n boyutlu uzayda tanımlı olsun: $\mathbf{X} = (x_1, x_2, \dots, x_n)$
- Veri kümesinde m adet sınıf bulunsun: $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_m$
- Sınıflamada son olasılığı en büyütme aranır ($\max P(\mathbf{C}_i | \mathbf{X})$)
- Bayes teoreminden türetilebilir:
$$P(\mathbf{C}_i | \mathbf{X}) = \frac{P(\mathbf{X} | \mathbf{C}_i) P(\mathbf{C}_i)}{P(\mathbf{X})}$$
- $P(\mathbf{X})$ olasılığı bütün sınıflar için sabit olduğuna göre, sadece $P(\mathbf{C}_i | \mathbf{X}) = P(\mathbf{X} | \mathbf{C}_i) P(\mathbf{C}_i)$ olasılığı için en büyük değer aranır.

Naïve Bayes sınıflandırıcı

$$P(C_i|\mathbf{X}) = P(\mathbf{X}|C_i)P(C_i)$$

Eğer bu basitleştirilmiş ifadede bütün öznitelikler bağımsız ise $P(\mathbf{X}|C_i)$ aşağıdaki şekilde yazılabilir.

$$P(\mathbf{X}|C_i) = \prod_{k=1}^n P(x_k | C_i) = P(x_1 | C_i) \times P(x_2 | C_i) \times \dots \times P(x_n | C_i)$$

Böylece hesap karmaşıklığı büyük ölçüde azaltılmış olur.

Bayes Sınıflandırıcı: Örnek

Gün	Hava Durumu	Sıcaklık Derecesi	Nem Oranı	Rüzgar	Oynama Durumu
D1	Güneşli	Sıcak	Yüksek	Zayıf	Hayır
D2	Güneşli	Sıcak	Yüksek	Şiddetli	Hayır
D3	Bulutlu	Sıcak	Yüksek	Zayıf	Evet
D4	Yağmurlu	Ilık	Yüksek	Zayıf	Evet
D5	Yağmurlu	Serin	Normal	Zayıf	Evet
D6	Yağmurlu	Serin	Normal	Şiddetli	Hayır
D7	Bulutlu	Serin	Normal	Şiddetli	Evet
D8	Güneşli	Ilık	Yüksek	Zayıf	Hayır
D9	Güneşli	Serin	Normal	Zayıf	Evet
D10	Yağmurlu	Ilık	Normal	Zayıf	Evet
D11	Güneşli	Ilık	Normal	Şiddetli	Evet
D12	Bulutlu	Ilık	Yüksek	Şiddetli	Evet
D13	Bulutlu	Sıcak	Normal	Zayıf	Evet
D14	Yağmurlu	Ilık	Yüksek	Şiddetli	Hayır

Bayes Sınıflandırıcı: Örnek

- Sınıflandırılacak **X** örneği:
 - Hava durumu = *Yağmurlu*
 - Sıcaklık = *Sıcak*
 - Nem = *Yüksek*
 - Rüzgar = *Zayıf*
- Tenis Oynanır mı ?

Bayes Sınıflandırıcı: Örnek

$$P(\text{evet} | X) = \frac{P(X | \text{evet}) P(\text{evet})}{P(X)}$$

$$P(\text{hayir} | X) = \frac{P(X | \text{hayir}) P(\text{hayir})}{P(X)}$$

- $P(X)$: bilmeye gerek yok
- Hangi olasılık yüksekse sınıf odur (**evet** ya da **hayır**)

Bayes Sınıflandırıcı: Örnek

$$P(\text{evet} | X)$$

$$= P([\text{yağmurlu} \& \text{sıcak} \& \text{yüksek} \& \text{zayıf}] | \text{evet}) \cdot P(\text{evet})$$

$$= P(\text{yağm} | \text{evet}) P(\text{sıcak} | \text{evet}) P(\text{yüksek} | \text{evet}) P(\text{zayıf} | \text{evet}) \cdot P(\text{evet})$$

$$= 3/9 \cdot 2/9 \cdot 3/9 \cdot 6/9 \cdot 9/14 = 0,010582011$$

$$P(\text{hayır} | X)$$

$$= P([\text{yağmurlu} \& \text{sıcak} \& \text{yüksek} \& \text{zayıf}] | \text{hayır}) \cdot P(\text{hayır})$$

$$= P(\text{yağm} | \text{hayır}) P(\text{sıcak} | \text{hayır}) P(\text{yüksek} | \text{hayır}) P(\text{zayıf} | \text{hayır}) \cdot P(\text{hayır})$$

$$= 2/5 \cdot 2/5 \cdot 4/5 \cdot 2/5 \cdot 5/14 = 0,018285714$$

$$P(\text{hayır} | X) > P(\text{evet} | X) \rightarrow \text{Tenis Oynan} \underline{\text{maz}}$$

Sınıflandırma Modelini Değerlendirme

- Sınıflandırma Metodu tarafından oluşturulan modelin başarısını ölçmek için
 - Doğruluk (Accuracy)
 - Hata Oranı (Error rate)
 - Specificity
 - Sensitivitygibi ölçüler kullanılır.

Sınıflandırma Modelini Değerlendirme:

- Bir M sınıflayıcısı için **doğruluk**;
 - $\text{acc}(M)$ doğru sınıflanmış örneklerin toplam örnek sayısına oranından bulunur.
- Bir M sınıflayıcısı için **hata oranı**;
 - $1 - \text{acc}(M)$ olarak hesaplanır.

Karışıklık Matrisi (Class Confusion Matrix)

	Öngörülen sınıf (Predicted Class)			
Gerçek Sınıf (Actual Class)		C ₁ (Positive)	C ₂ (Negative)	
	C ₁ (Positive)	True positive TP	False negative FN	ΣPositive
	C ₂ (Negative)	False positive FP	True negative TN	ΣNegative

sensitivity = TP / pos /* **recall**: true positive rate */

specificity = TN / neg /* **true negative rate** */

precision = TP / (TP + FP) /* **precision** */

accuracy = (TP + TN) / (pos + neg)

Örnek

Örnek	Sınıf	Sonuç
1	E	E
2	K	E
3	E	E
4	E	E
5	K	K
6	E	E
7	K	K
8	K	E
9	K	K
10	E	K

C1(**pos**)=E, C2(**neg**)=K

TP(her ikisi de E) = **4**

FP(Gerçekte E değil ama sonuç E) = **2**

FN(Gerçekte E ama sonuç E değil) = **1**

TN(her ikisi de E değil) = **3**

sensitivity = TP/pos = 4/5 = **0.8**

specificity = TN/neg = 3/5 = **0.6**

prec = TP/(TP+FP) = 4/6 = **0.67**

acc = (TP+TN)/(pos+neg) = **0.7**