

# Örüntü Tanıma

Dr. Öğr. Üyesi Mehmet Zahid YILDIRIM

e-mail: [m.zahidyildirim@karabuk.edu.tr](mailto:m.zahidyildirim@karabuk.edu.tr)

# Bayes Teoremi

- Belirli bir kanser türünün nüfusun **%1**'ini etkilediğini varsayalım. Bu kanser için piyasada bir test var ama güvenilirliği mükemmel değil.
- Bu test kanserli kişilerin **%90**'ında pozitif sonuç veriyor yani doğru tanı koyuyor. Ancak aynı zamanda kansersiz kişilerin **%5**'inde de pozitif sonuç veriyor. Diğer bir deyişle bu kişiler de kanser olmadığı halde kanser olduğunu ortaya koyuyor.
- Endişelendiniz ve bir test yaptınız. Sonucunuz ne yazık ki pozitif çıktı.
- ***Bu durumda kanser olma olasılığınız nedir?***

# Bayes Teoremi

- Belirli bir kanser türünün nüfusun **%1**'ini etkilediğini varsayalım. Bu kanser için piyasada bir test var ama güvenilirliği mükemmel değil.
- Bu test kanserli kişilerin **%90**'ında pozitif sonuç veriyor yani doğru tanı koyuyor. Ancak aynı zamanda kansersiz kişilerin **%5**'inde de pozitif sonuç veriyor. Diğer bir deyişle bu kişiler de kanser olmadığı halde kanser olduğunu ortaya koyuyor.
- Endişelendiniz ve bir test yaptınız. Sonucunuz ne yazık ki pozitif çıktı.
- ***Bu durumda kanser olma olasılığınız nedir?***

**≈ %15**

# Bayes Teoremi

- Bayes Teoremi, 18. yüzyılda İngiliz Bakan ve matematikçi olan **Thomas Bayes** (1702 – 1761) tarafından öne sürülmüş ve özellikle geçtiğimiz son yarım yüzyılda önemi giderek daha fazla anlaşılmış bir teorem.
- Günümüzde **Bayes teoremi** fizikten kanser araştırmalarına, çevre bilimden psikolojiye kadar bir çok yerde kendine bir uygulama alanı bulmuş durumdadır.



Thomas Bayes

# Bayes Teoremi

**Kullanım Alanları: Olasılığın olduğu tüm alanlar..**

**Olasılık**, belirsizliği ifade etmek için kullanılan bir kavramdır. Bir olayın gerçekleşebilmesinin “0” ve “1” arasında değişen matematiksel değeridir.

# Bayes Teoremi

**Tanım:** Bayes Teoremi, bir olasılık değerini, bu olasılık ile ilişkili olduğunu bildiğimiz diğer olasılık değerlerini kullanarak hesaplama fikridir.

$$P(A|B) = \frac{P(A).P(B|A)}{P(B)}$$

burada,

$P(A|B)$  = B koşulu altında A'nın gerçekleşme olasılığı.

$P(B|A)$  = A koşulu altında B'nin gerçekleşme olasılığı.

$P(A)$  = A'nın olma olasılığı.

$P(B)$  = B'nin olma olasılığı.

# Bayes Teoremi

**Tanım:** Bayes Teoremi, bir olasılık değerini, bu olasılık ile ilişkili olduğunu bildiğimiz diğer olasılık değerlerini kullanarak hesaplama fikridir.

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

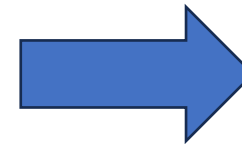
burada,

$P(A|B)$  = B koşulu altında A'nın gerçekleşme olasılığı.

$P(B|A)$  = A koşulu altında B'nin gerçekleşme olasılığı.

$P(A)$  = A'nın olma olasılığı.

$P(B)$  = B'nin olma olasılığı.

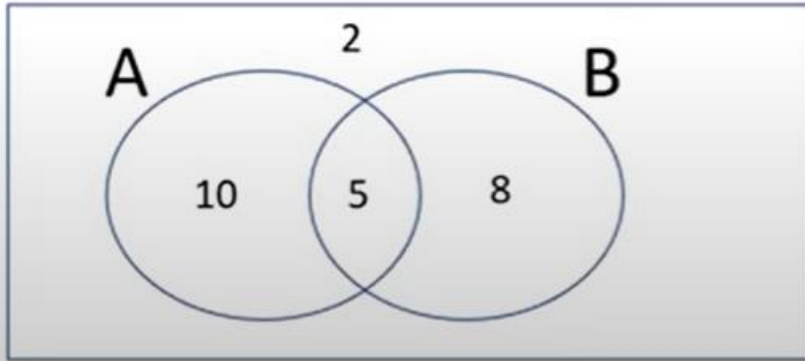


$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

# Bayes Teoremi

Örnek:



$P(A)$ : Türkçe bilme olasılığı

$P(B)$ : İngilizce bilme olasılığı

$P(A|B)$ : İngilizce bilenlerin Türkçe bilme olasılığı =  $\frac{5}{13}$

$P(B|A)$ : Türkçe bilenlerin İngilizce bilme olasılığı =  $\frac{5}{15}$



# Bayes Teoremi

**Örnek 1:** Matematik Profesörü yada şoförlük mesleklerinden birini yaptığı bilinen bir kişi var. Bu kişinin bilinen özelliği, araştırma yapmayı seviyor olması. Bu kişinin Matematik Profesörü olma olasılığı nedir?



$P(A)$ : Kişinin matematik profesörü olma ihtimali

$P(B)$ : Kişinin araştırma yapmayı sevme ihtimali

Matematik profesörlerinin ne kadarı araştırma yapmayı sever? **%90**

Şoförlerin ne kadarı araştırma yapmayı sever? **%15**

Ülkemizdeki matematik profesörü sayısı (yaklaşık) : **500**

Ülkemizdeki şoför sayısı (yaklaşık) : **700.000**

Soru: Araştırma yapmayı seven kişinin matematik profesörü olma ihtimali

$$P(A|B)=?$$

# Bayes Teoremi

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Diagram explaining the components of Bayes' Theorem:

- $P(B|A)$ : Matematik profesörlerinin araştırma yapmayı sevme ihtimali
- $P(A)$ : Kişinin matematik profesörü olma ihtimali
- $P(B)$ : Kişinin araştırma yapmayı sevme ihtimali

$P(A)$ : Kişinin matematik profesörü olma ihtimali

$P(B)$ : Kişinin araştırma yapmayı sevme ihtimali

Matematik profesörlerinin ne kadarı araştırma yapmayı sever? **%90**

Şoförlerin ne kadarı araştırma yapmayı sever? **%15**

Ülkemizdeki matematik profesörü sayısı (yaklaşık) : **500**

Ülkemizdeki şoför sayısı (yaklaşık) : **700.000**

$$P(B|A) = 0,9 \quad P(A) = \frac{500}{700500} \quad P(B) = \frac{500 * 0,9 + 700.000 * 0,15}{700500}$$

$$P(A|B) = \frac{0,9 * \frac{500}{700500}}{\frac{500 * 0,9 + 700.000 * 0,15}{700500}} \approx 0,0043$$

$$P(A|B) \approx \%0,4$$

# Bayes Teoremi

**Örnek 2:** Piknik yapmayı düşünen bir grup sabah havanın bulutlu olduğunu görüyor. Yağmurlu günlerin %50'sinin bulutlu bir günle başladığı biliniyor. Günlerin %40'ı bulutlu başlamaktadır. Bir günün yağmurlu gün olma olasılığı %10'dur. Buna göre, **bulutlu başlanan bir günde yağmur görme olasılığı** nedir?

$$P(\text{Yağmur} | \text{Bulutlu}) = ?$$

$$P(\text{Bulutlu} | \text{Yağmur}) = 0,50$$

$$P(\text{Bulutlu}) = 0,40$$

$$P(\text{Yağmur}) = 0,10$$

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

# Naive Bayes Sınıflandırıcı

- Naive Bayes; Bayes teoremine dayanan bir sınıflandırma tekniğidir.
- Naive-Bayes sınıflandırma algoritması bir mesajın spam olup olmadığının kontrolü, bir makalenin içeriğinin hangi konuda (teknoloji, spor, politik) olduğunun sınıflandırılması veya yüz tanıma gibi alanlarda kullanılır.
- Bir sınıflandırma problemi bir çok özellikten ve bir sonuç (hedef) değişkeninden oluşur.


Weather	Car	Class
sunny	working	go-out
rainy	broken	go-out
sunny	working	go-out
sunny	working	go-out
sunny	working	go-out
rainy	broken	stay-home
rainy	broken	stay-home
sunny	working	stay-home
sunny	broken	stay-home
rainy	broken	stay-home

# Naive Bayes Sınıflandırıcı

**Örnek 1:** Hava durumuna bağlı 14 farklı veri (Gözlem) üzerinden futbol oynayıp oynanmayacağına karar verme.

Özellikler	Hedef
<b>Hava Durumu</b>	<b>Futbol Oyna</b>
Yağmurlu	Hayır
Yağmurlu	Hayır
Bulutlu	Evet
Güneşli	Evet
Güneşli	Evet
Güneşli	Hayır
Bulutlu	Evet
Yağmurlu	Hayır
Yağmurlu	Evet
Güneşli	Evet
Yağmurlu	Evet
Bulutlu	Evet
Bulutlu	Evet
Güneşli	Hayır


		Futbol Oyna			
		Evet	Hayır		
Hava Durumu	Güneşli	3	2		
	Bulutlu	4	0		
	Yağmurlu	2	3		
		9/14	5/14		



		Futbol Oyna			
		Evet	Hayır		
Hava Durumu	Güneşli	3/9	2/5	5/14	
	Bulutlu	4/9	0/5	4/14	
	Yağmurlu	2/9	3/5	5/14	
		9/14	5/14		

# Naive Bayes Sınıflandırıcı

		Futbol Oyna	
		Evet	Hayır
Hava Durumu	Güneşli	3	2
	Bulutlu	4	0
	Yağmurlu	2	3



		Futbol Oyna		
		Evet	Hayır	
Hava Durumu	Güneşli	3/9	2/5	5/14
	Bulutlu	4/9	0/5	4/14
	Yağmurlu	2/9	3/5	5/14
		9/14	5/14	

**Beklenti 1:** Güneşliken Futbol Oynar =  $P(\text{Evet} \mid \text{Güneşli}) = P(\text{Güneşli} \mid \text{Evet}) * P(\text{Evet}) / P(\text{Güneşli})$

$P(\text{Güneşli} \mid \text{Evet}) = 3/9 = 0.333$ ,  $P(\text{Güneşli}) = 5/14 = 0.357$ ,  $P(\text{Evet}) = 9/14 = 0.643$

$P(\text{Evet} \mid \text{Güneşli}) = 0.333 * 0.643 / 0.357 = 0.600$

---

**Beklenti 2:** Güneşliken Futbol Oynamaz =  $P(\text{Hayır} \mid \text{Güneşli}) = P(\text{Güneşli} \mid \text{Hayır}) * P(\text{Hayır}) / P(\text{Güneşli})$

$P(\text{Güneşli} \mid \text{Hayır}) = 2/9 = 0.222$ ,  $P(\text{Güneşli}) = 5/14 = 0.357$ ,  $P(\text{Hayır}) = 5/14 = 0.357$

$P(\text{Hayır} \mid \text{Güneşli}) = 0.222 * 0.357 / 0.357 = 0.222$

---

$P(\text{Evet} \mid \text{Güneşli}) > P(\text{Hayır} \mid \text{Güneşli}) \Rightarrow \text{Oynar}$

# Naive Bayes Sınıflandırıcı

- Naive bayes sınıflandırıcı veri setinde birden fazla özellik varsa basitçe bütün koşullu olasılıkların çarpımıdır.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$



$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c) \quad \text{==}$$

$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i)$

# Naive Bayes Sınıflandırıcı

## Örnek 2:

Yeni Örnek;

Outlook => Sunny,

Temp => Hot,

Humidity => Normal,

Windy => False.

$$P(y|X) = \frac{P(X|y) \cdot P(y)}{P(X)}$$

	Outlook	Temp	Humidity	Windy	Play Golf
1	Rainy	Hot	High	FALSE	No
2	Rainy	Hot	High	TRUE	No
3	Overcast	Hot	High	FALSE	Yes
4	Sunny	Mild	High	FALSE	Yes
5	Sunny	Cool	Normal	FALSE	Yes
6	Sunny	Cool	Normal	TRUE	No
7	Overcast	Cool	Normal	TRUE	Yes
8	Rainy	Mild	High	FALSE	No
9	Rainy	Cool	Normal	FALSE	Yes
10	Sunny	Mild	Normal	FALSE	Yes
11	Rainy	Mild	Normal	TRUE	Yes
12	Overcast	Mild	High	TRUE	Yes
13	Overcast	Hot	Normal	FALSE	Yes
14	Sunny	Mild	High	TRUE	No



# Naive Bayes Sınıflandırıcı

## Örnek 2:

$$P(y|X) = \frac{P(X|y) \cdot P(y)}{P(X)}$$

$x_1 = (\text{Outlook} = \text{Sunny}), x_2 = (\text{Temp} = \text{Hot}), x_3 = (\text{Humidity} = \text{Normal}), x_4 = (\text{Windy} = \text{False})$

$P(\text{Yes}) = 9/14$

$P(\text{No}) = 5/14$

$P(\text{Sunny}|\text{Yes}) = 3/9$

$P(\text{Sunny}|\text{No}) = 2/5$

$P(\text{Hot}|\text{Yes}) = 2/9$

$P(\text{Hot}|\text{No}) = 2/5$

$P(\text{Normal}|\text{Yes}) = 6/9$

$P(\text{Normal}|\text{No}) = 1/5$

$P(\text{False}|\text{Yes}) = 6/9$

$P(\text{False}|\text{No}) = 2/5$

Oranları çarpalım: "Yes".

$P(X|\text{Play Golf}=\text{Yes}) = 3/9 * 2/9 * 6/9 * 6/9 * 9/14 = 0.0211$

Oranları Çarpalım: "No".

$P(X|\text{Play Golf}=\text{No}) = 2/5 * 2/5 * 1/5 * 2/5 * 5/14 = 0.0045$

$$P(X|\text{Play Golf} = \text{Yes}) = \frac{0.0211}{0.0211 + 0.0045} = 0.8242$$

$$P(X|\text{Play Golf} = \text{No}) = \frac{0.0045}{0.0045 + 0.0211} = 0.1757$$

	Outlook	Temp	Humidity	Windy	Play Golf
1	Rainy	Hot	High	FALSE	No
2	Rainy	Hot	High	TRUE	No
3	Overcast	Hot	High	FALSE	Yes
4	Sunny	Mild	High	FALSE	Yes
5	Sunny	Cool	Normal	FALSE	Yes
6	Sunny	Cool	Normal	TRUE	No
7	Overcast	Cool	Normal	TRUE	Yes
8	Rainy	Mild	High	FALSE	No
9	Rainy	Cool	Normal	FALSE	Yes
10	Sunny	Mild	Normal	FALSE	Yes
11	Rainy	Mild	Normal	TRUE	Yes
12	Overcast	Mild	High	TRUE	Yes
13	Overcast	Hot	Normal	FALSE	Yes
14	Sunny	Mild	High	TRUE	No

# Naive Bayes Sınıflandırıcı

## Örnek 3:

TENİS OYNAMA ÖRNEĞİ					
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

# Naive Bayes Sınıflandırıcı

## Örnek 3:

$$P(\text{evet}) = 9/14$$

$$P(\text{hayır}) = 5/14$$

Hava durumu	
$P(\text{güneşli} \text{evet}) = 2/5$	$P(\text{güneşli} \text{hayır}) = 3/5$
$P(\text{bulutlu} \text{evet}) = 4/4$	$P(\text{bulutlu} \text{hayır}) = 0$
$P(\text{yağmurlu} \text{evet}) = 3/5$	$P(\text{yağmurlu} \text{hayır}) = 2/5$
Sıcaklık	
$P(\text{sıcak} \text{evet}) = 2/4$	$P(\text{sıcak} \text{hayır}) = 2/4$
$P(\text{ılık} \text{evet}) = 4/6$	$P(\text{ılık} \text{hayır}) = 2/6$
$P(\text{serin} \text{evet}) = 3/4$	$P(\text{serin} \text{hayır}) = 1/4$
Nem oranı	
$P(\text{yüksek} \text{evet}) = 3/7$	$P(\text{yüksek} \text{hayır}) = 4/7$
$P(\text{normal} \text{evet}) = 6/7$	$P(\text{normal} \text{hayır}) = 1/7$
Rüzgar	
$P(\text{şiddetli} \text{evet}) = 3/6$	$P(\text{zayıf} \text{hayır}) = 3/6$
$P(\text{zayıf} \text{evet}) = 6/8$	$P(\text{şiddetli} \text{hayır}) = 2/8$

## TENİS OYNAMA ÖRNEĞİ

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

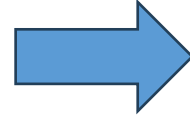
# Naive Bayes Sınıflandırıcı

## Örnek 3:

$$P(\text{evet}) = 9/14$$

$$P(\text{hayır}) = 5/14$$

Hava durumu	
$P(\text{güneşli} \text{evet}) = 2/5$	$P(\text{güneşli} \text{hayır}) = 3/5$
$P(\text{bulutlu} \text{evet}) = 4/4$	$P(\text{bulutlu} \text{hayır}) = 0$
$P(\text{yağmurlu} \text{evet}) = 3/5$	$P(\text{yağmurlu} \text{hayır}) = 2/5$
Sıcaklık	
$P(\text{sıcak} \text{evet}) = 2/4$	$P(\text{sıcak} \text{hayır}) = 2/4$
$P(\text{ılık} \text{evet}) = 4/6$	$P(\text{ılık} \text{hayır}) = 2/6$
$P(\text{serin} \text{evet}) = 3/4$	$P(\text{serin} \text{hayır}) = 1/4$
Nem oranı	
$P(\text{yüksek} \text{evet}) = 3/7$	$P(\text{yüksek} \text{hayır}) = 4/7$
$P(\text{normal} \text{evet}) = 6/7$	$P(\text{normal} \text{hayır}) = 1/7$
Rüzgar	
$P(\text{şiddetli} \text{evet}) = 3/6$	$P(\text{zayıf} \text{hayır}) = 3/6$
$P(\text{zayıf} \text{evet}) = 6/8$	$P(\text{şiddetli} \text{hayır}) = 2/8$



Yeni Örnek:  $X = \langle \text{yağmurlu, sıcak, yüksek, zayıf} \rangle$

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

$$\begin{aligned} &= P(\text{yağmurlu} | \text{evet}) \cdot P(\text{sıcak} | \text{evet}) \cdot P(\text{yüksek} | \text{evet}) \cdot P(\text{zayıf} | \text{evet}) \cdot P(\text{evet}) \\ &= 3/5 \cdot 2/4 \cdot 3/7 \cdot 6/8 \cdot 9/14 \\ &= \mathbf{0.062} \end{aligned}$$

$$P(\text{hayır} | X) = ?$$

$$\begin{aligned} &= P(\text{yağmurlu} | \text{hayır}) \cdot P(\text{sıcak} | \text{hayır}) \cdot P(\text{yüksek} | \text{hayır}) \cdot P(\text{zayıf} | \text{hayır}) \cdot P(\text{hayır}) \\ &= 2/5 \cdot 2/4 \cdot 4/7 \cdot 2/8 \cdot 5/14 \\ &= \mathbf{0.01} \end{aligned}$$

$$P(X | \text{evet}) = \frac{0.062}{0.062 + 0.01} = \mathbf{0.861}$$

$$P(X | \text{hayır}) = \frac{0.01}{0.062 + 0.01} = \mathbf{0.139}$$

$P(X | \text{evet}) > P(X | \text{hayır})$  olduğundan örnek  $X'$  in sınıfı **evet** olarak öngörülür.

# Naive Bayes Sınıflandırıcı

Ödev:

## ÖRNEK

Verilen tabloya göre Dergi Harcaması=Evet, Saat Harcaması=Evet  
Hayat Sigortası=Hayır, Kredi Kartı Sigortası=Hayır  
Cinsiyet=?

Dergi Harcaması	Kol Saati Harcaması	Hayat Sigortası	Kredi Kartı Sigortası	Cinsiyet
Evet	Hayır	Hayır	Hayır	E
Evet	Evet	Evet	Hayır	K
Hayır	Hayır	Hayır	Hayır	E
Evet	Evet	Evet	Hayır	E
Evet	Hayır	Evet	Hayır	K
Hayır	Hayır	Hayır	Hayır	K
Evet	Hayır	Evet	Evet	E
Hayır	Evet	Hayır	Hayır	E
Evet	Hayır	Hayır	Hayır	E
Evet	Evet	Evet	Hayır	K

# Naive Bayes Sınıflandırıcı

## Avantaj

- Basit ve kolay uygulanabilir.
- Az veriyle iyi işler başarabilir.
- Hızlı olmasından dolayı gerçek zamanlı sistemlerde kullanılabilir.
- Kategorik verilerde sayısal verilere göre daha iyi çalışır.

## Dezavantajları

- Tüm özelliklerin bağımsız olduğunu varsayar ancak bu gerçek hayatta nadirdir.
- Sistem dinamik olduğundan her defasında eğitim en baştan yapılır.
- Zero Probability problemi ile karşı karşıya kalabilirsiniz. Zero Probability istediğimiz örneğin veri setinde hiç bulunmaması durumudur. Bunun için en basit yöntem tüm verilere minimum değer ekleyerek (genellikle 1) bu olasılığı ortadan kaldırmaktır.



# Naive Bayes Sınıflandırıcı

## **Avantaj**

- Basit ve kolay uygulanabilir.
- Az veriyle iyi işler başarabilir.
- Hızlı olmasından dolayı gerçek zamanlı sistemlerde kullanılabilir.

## **Dezavantajları**

- Tüm özelliklerin bağımsız olduğunu varsayar ancak bu gerçek hayatta nadirdir.
- Zero Probability problemi ile karşı karşıya kalabilirsiniz. Zero Probability istediğimiz örneğin veri setinde hiç bulunmaması durumudur. Bunun için en basit yöntem tüm verilere minimum değer ekleyerek (genellikle 1) bu olasılığı ortadan kaldırmaktır.

# Naive Bayes Sınıflandırıcı

## **Uygulama Alanları**

- Gerçek Zamanlı Sistemler
- Çoklu Sınıflandırma Problemleri (Haber / E-Ticaret Kategorileri)
- Metin Sınıflandırma (Spam Filtreleme / Duygu Analizi)
- Hastalık Teşhisi
- Tavsiye Sistemleri (Recommendation System)



# Kaynaklar

Sargur Srihari (CEDAR),  
Jason Corso (SUNY at Buffalo),  
Armando Vieira (Closer),  
Luis Gustavo Martins (Catolica),  
Selim Aksoy (Bilkent)