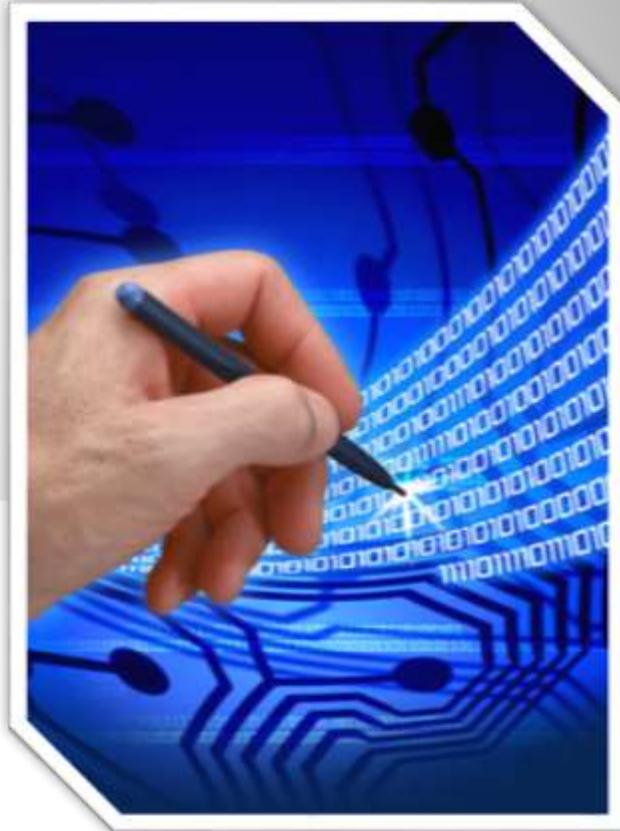
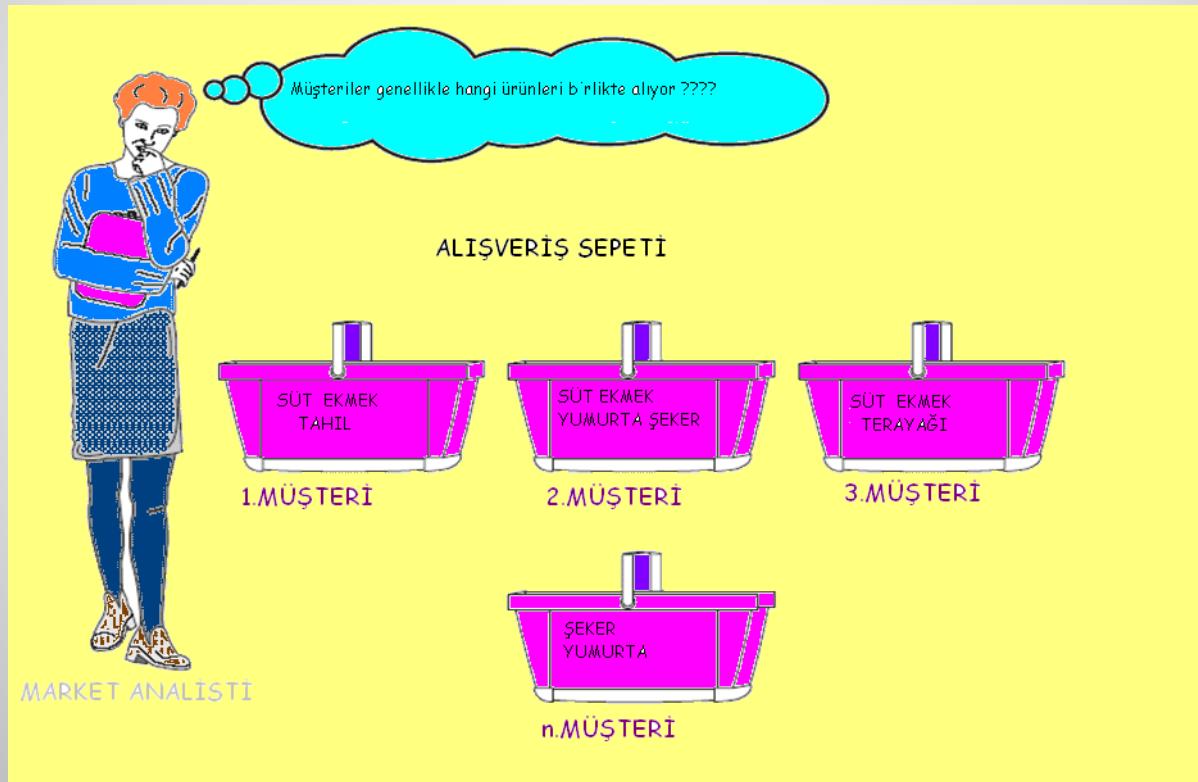


# *Veri Madenciliği*

BİRLİKTELİK KURALLARI  
VE  
İLİŞKİ ANALİZİ



Birliktelik kuralları, aynı işlem içinde çoğunlukla beraber görülen nesneleri içeren kurallardır. Birliktelik kurallarının kullanıldığı en tipik örnek market sepeti uygulamasıdır. Bu işlem, müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak **müşterilerin satın alma alışkanlıklarını** çözümler.



$$destek(X \rightarrow Y) = \frac{sayi(X, Y)}{n}$$

$$güven(X \rightarrow Y) = \frac{sayi(X, Y)}{sayi(X)}$$

- Örneğin 25 tane müşterinin bir defada aldığı ürün bilgilerinden yola çıkarak birliktelik kuralı şu şekilde bulunmuş olsun:

*güven(Pantolon, Kazak → Çorap)*

• Burada  $X = \{Pantolon, Kazak\}$  ve  $Y = \{\mathcal{C}orap\}$

değerleri için pantolon ve kazak alan müşterilerin bunların yanında çorap da sayı alama olasılığını ifade eder. Müşterinin bu 3 ürünü birlikte satın alma sayısı 7 ve müşteri sayısı 25 ise belirttiğimiz bu kuralın destek ölçüyü şöyle olacaktır:

$$destek(Pantolon, Kazak \rightarrow \mathcal{C}orap) = \frac{sayi(Pantolon, Kazak, \mathcal{C}orap)}{musterisayisi} = \frac{7}{25} = 0,28$$

• Eğer pantolon ve kazak alanlarının sayısının 9 ise Güven ölçütü

$$güven(Pantolon, Kazak \rightarrow \mathcal{C}orap) = \frac{sayi(Pantolon, Kazak, \mathcal{C}orap)}{sayi(Pantolon, Kazak)} = \frac{7}{9} = 0,77$$

# Apriori Algoritması

*Bu algoritma birliktelik kurallarının oluşturulmasında yararlanılan ve yaygın olarak kullanılan bir algoritmadır. Algoritma aşağıdaki adımlardan oluşmaktadır:*

- *Öncelikle destek ve güven ölçülerini karşılaştırmak için eşik değerleri belirtenir.*
- *Her bir ürün için destek sayılan hesaplanır. Eşik değeri ile karşılaştırılan destek değerlerinin içinden eşik değerinden düşük olanlar çıkarılır.*
- *Kalan ürünler ikişerli grüplənərək, grup destek sayılan hesaplanır. Tekrar eşik değerleri ile karşılaştırılan destek değerlerinden eşik değerinin altında kalanlar iptal edilir.*
- *Daha sonra üçerli, dörderli, beşerli, vb. biçimde grüplər için aynı karşılaştırma ve eleye işlemi devam etdirilir. Eşik değerlərə uyğun olduğu sürece işlemlər sürecekdir.*
- *Belirlənen ürün grubunun destek ölçülərinə bakarək birliktelik kuralları tərtibilir və bu kuralların her biri üçün güven ölçüləri belirlənir.*

## Apriori Algoritması(Örnek Çalışma)

- Öncelikle destek ve güven ölçülerini karşılaştırmak için eşik değerleri belirlenir.

Müşteri Numarası	Alınmış Çikolata Markası
MSN0101	ERDEM, ŞOKOKO, ÇİTPİT, DERYA
MSN0102	ENFES, DERYA, MAZLUM, ŞOKOKO
MSN0103	ERDEM, DERYA, ENFES
MSN0104	ŞOKOKO, ÇİTPİT, DERYA, POTPORİ,
MSN0105	ERDEM, DERYA, MAZLUM, ÇİTPİT
MSN0106	MİS, DERYA, ÇİTPİT
MSN0107	ERDEM, ZARİF, DERYA
MSN0108	MAZLUM, YURDUM, ÇİTPİT
MSN0109	YURDUM, ŞOKOKO, ÇİTPİT
MSN0110	ÇİTPİT, MİS, ERDEM, MAZLUM, DERYA

$$\text{Destek}_{\text{eşik}} = \%30$$

$$\text{Güven}_{\text{eşik}} = \%80$$

Burada eşik destek sayısı

$$0,30 * 10 = 3$$

## Apriori Algoritması(Örnek Çalışma)

- Her bir ürün için destek sayıları hesaplanır. Eşik değeri ile karşılaştırılan destek değerlerinin içinden eşik değerinden düşük olanlar çıkarılır.

Cikolata Markası	Destek Değeri
ERDEM	5
ŞOKOKO	4
ÇİTPIT	7
MAZLUM	4
ENFES	2
DERYA	8
POTPORI	1
MİS	2
YURDUM	2
ZARIF	1

Cikolata Markası	Destek Değeri
ERDEM	5
ŞOKOKO	4
ÇİTPIT	7
MAZLUM	4
DERYA	8

# Apriori Algoritması(Örnek Çalışma)

Çikolata Markası	Destek Değeri
ERDEM, SOKOKO	1
ERDEM, ÇİTIT	3
ERDEM, MAZLUM	2
ERDEM, DERYA	5
SOKOKO, ERDEM	1
SOKOKO, ÇİTIT	3
ŞOKOKO, MAZLUM	2
ŞOKOKO, DERYA	3
ÇİTIT, ERDEM	3
ÇİTIT, SOKOKO	3
ÇİTIT, MAZLUM	3
ÇİTIT, DERYA	5
MAZLUM, ERDEM	2
MAZLUM, SOKOKO	2
MAZLUM, ÇİTIT	3
MAZLUM, DERYA	3
DERYA, ERDEM	5
DERYA, ÇİTIT	5
DERYA, MAZLUM	3
DERYA, SOKOKO	3

- Kalan ürünler ikişerli grüplanarak, grup destek sayıları hesaplanır. Tekrar eşik değerleri ile karşılaştırılan destek değerlerinden eşik değerinin altında kalanlar iptal edilir.

Çikolata Markası	Destek Değeri
ERDEM, ÇİTIT	3
ERDEM, DERYA	5
SOKOKO, ÇİTIT	3
ŞOKOKO, DERYA	3
ÇİTIT, MAZLUM	3
ÇİTIT, DERYA	5
MAZLUM, DERYA	3

## Apriori Algoritması(Örnek Çalışma)

- Daha sonra üçerli,dörderli,beşerli, vb. biçimde gruplar için aynı karşılaştırma ve eleme işlemi devam ettirilir. Eşik değerlere uygun olduğu sürece işlemler sürecektir.

Çikolata Markası	Destek Değeri
ERDEM, CİTPİT, ŞOKOKO	1
ERDEM, CİTPİT, DERYA	3
ERDEM, DERYA, ŞOKOKO	2
ERDEM, DERYA, CİTPİT	3
ERDEM, DERYA, ENFES	1
ERDEM, DERYA, MAZLUM	3
ERDEM, DERYA, MİS	1
SOKOKO, CİTPİT, ERDEM	1
SOKOKO, CİTPİT, DERYA	2
SOKOKO, CİTPİT, POTPORİ	1
SOKOKO, CİTPİT, YURDUM	1
SOKOKO, DERYA, ERDEM	1
SOKOKO, DERYA, CİTPİT	2
SOKOKO, DERYA, MAZLUM	1

SOKOKO, DERYA, ENFES	1
SOKOKO, DERYA, POTPORİ	1
CİTPİT, MAZLUM, DERYA	2
CİTPİT, MAZLUM, ERDEM	2
CİTPİT, MAZLUM, YURDUM	1
CİTPİT, MAZLUM, MİS	1
CİTPİT, DERYA, ŞOKOKO	2
CİTPİT, DERYA, ERDEM	3
CİTPİT, DERYA, POTPORİ	1
CİTPİT, DERYA, MAZLUM	2
CİTPİT, DERYA, MİS	2

Çikolata Markası	Destek Değeri
ERDEM, CİTPİT, DERYA	3
ERDEM, DERYA, MAZLUM	3

## Apriori Algoritması(Örnek Çalışma)

- Belirlenen ürün grubunun destek ölçülerine bakarak birliktelik kuralları türetilir ve bu kurallarının her biri için güven ölçüleri belirlenir.

ERDEM , ÇITPIT , DERYA için birliktelik kurallarının altkümesi ;  
{ ERDEM , ÇITPIT} , { ERDEM , DERYA} , {ÇITPIT , DERYA} , {ÇITPIT} , {DERYA} , { ERDEM }

Birliktelik	Açıklama	Güven
ERDEM & ÇITPIT -> DERYA	ERDEM ve ÇITPIT'in bulunduğu item-sette DERYA'nın olma olasılığı	3/3=%100
ERDEM & DERYA -> ÇITPIT	ERDEM ve DERYA'nın bulunduğu item-sette ÇITPIT'in olma olasılığı	3/5=%60
ÇITPIT & DERYA -> ERDEM	ÇITPIT ve DERYA'nın bulunduğu item-sette ERDEM'in olma olasılığı	3/5=%60
ÇITPIT -> ERDEM & DERYA	ÇITPIT'in bulunduğu item-sette ERDEM ve DERYA'nın olma olasılığı	3/7 = %42
DERYA -> ERDEM & ÇITPIT	DERYA'nın bulunduğu item-sette ÇITPIT ve ERDEM'in olma olasılığı	3/8 = %38
ERDEM -> ÇITPIT & DERYA	ERDEM'in bulunduğu item-sette ÇITPIT ve DERYA'nın olma olasılığı	3/5 = %60

## Apriori Algoritması(Örnek Çalışma)

ERDEM , DERYA , MAZLUM için birliktelik kurallarının alt kümesi ;

{ERDEM , DERYA} , {ERDEM , MAZLUM} , {DERYA , MAZLUM} , {DERYA} , {MAZLUM} , {ERDEM}

Birliktelik	Açıklama	Güven
ERDEM & DERYA -> MAZLUM	ERDEM ve DERYA bulunduğu item-sette MAZLUM un olma olasılığı	3/5=%60
ERDEM & MAZLUM -> DERYA	ERDEM ve MAZLUM un bulunduğu item-sette DERYA nın olma olasılığı	2/3=%66
DERYA & MAZLUM -> ERDEM	DERYA ve MAZLUM un bulunduğu item-sette ERDEM in olma olasılığı	3/3=%100
DERYA -> ERDEM & MAZLUM	DERYAin bulunduğu item-sette ERDEM ve MAZLUM un olma olasılığı	3/4=%75
MAZLUM -> ERDEM & DERYA	MAZLUM un bulunduğu item-sette DERYA nın ve ERDEM in olma olasılığı	3/4= %75
ERDEM -> DERYA & MAZLUM	ERDEM in bulunduğu item-sette DERYA ve MAZLUM un olma olasılığı	3/5=%60

## Apriori Algoritması(Örnek Çalışma)

Birliktelik	Açıklama	Güven
ERDEM & ÇITPIT -> DERYA	ERDEM ve ÇITPIT in bulunduğu item-sette DERYA nin olma olasılığı	3/3=%100
ERDEM & DERYA -> ÇITPIT	ERDEM ve DERYA nin bulunduğu item-sette ÇITPIT in olma olasılığı	3/5=%60
ÇITPIT & DERYA -> ERDEM	ÇITPIT ve DERYA nin bulunduğu item-sette ERDEM in olma olasılığı	3/5=%60
<b>ÇITPIT -&gt; ERDEM &amp; DERYA</b>	<b>ÇITPIT &amp; DERYA -&gt; ERDEM</b>	<b>3/7 = %42</b>
Birliktelik	Açıklama	Güven
DERYA -> ERDEM & DERYA -> MAZLUM	ERDEM ve DERYA bulunduğu item-sette MAZLUM un olma olasılığı	3/5=%60
ERDEM -> DERYA & MAZLUM -> DERYA	ERDEM ve MAZLUM un bulunduğu item-sette DERYA nin olma olasılığı	2/3=%66
DERYA & MAZLUM -> ERDEM	DERYA ve MAZLUM un bulunduğu item-sette ERDEM in olma olasılığı	3/3=%100
DERYA -> ERDEM & MAZLUM	DERYAin bulunduğu item-sette ERDEM ve MAZLUM un olma olasılığı	3/4=%75
MAZLUM -> ERDEM & DERYA	MAZLUM un bulunduğu item-sette DERYA nin ve ERDEM in olma olasılığı	3/4=%75
ERDEM -> DERYA & MAZLUM	ERDEM in bulunduğu item-sette DERYA ve MAZLUM un olma olasılığı	3/5=%60

Bu iki birliktelik kuralında güven<sub>eşik</sub>=80 değeri dikkate alınarak düzenleme yapılırsa;

DERYA & MAZLUM -> ERDEM = %100

ERDEM & ÇITPIT -> DERYA = %100 ‘a göre aşağıdaki sonuçlar çıkarılır.

**DERYA ve MAZLUM un birlikte satıldığı alışverişde ERDEM in satılma olasılığı %100 ‘dür  
ERDEM in ve ÇITPIT in birlikte satıldığı alışverişte DERYA nin satılma olasılığı %100 ‘dır.**

## Kaynaklar :

- Veri Madenciliği Yöntemleri, Yalçın Özkan 06'2008
- Veri Madenciliği ,Gökhan Silahtaroğlu 06'2008
- İstanbul Ticaret Üniversitesi Dergisi Veri Madenciliği Modeller Ve Uygulama Alanları (Serhat ÖZEKES)

