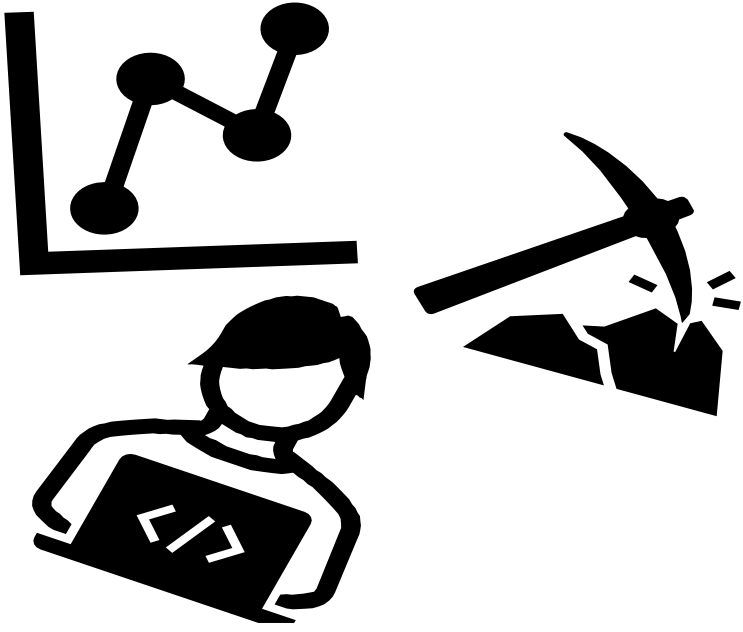




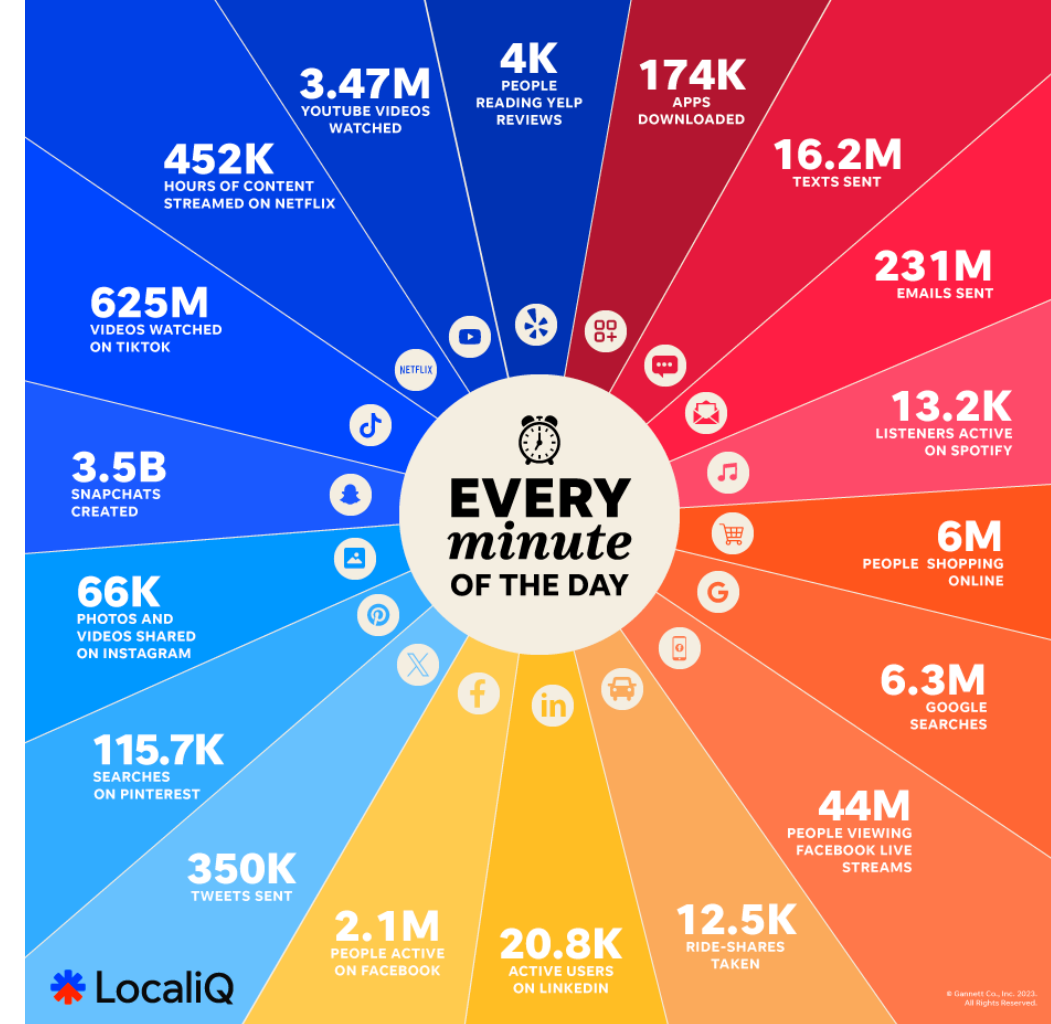
Veri Madenciliği

Birliktelik Kuralları

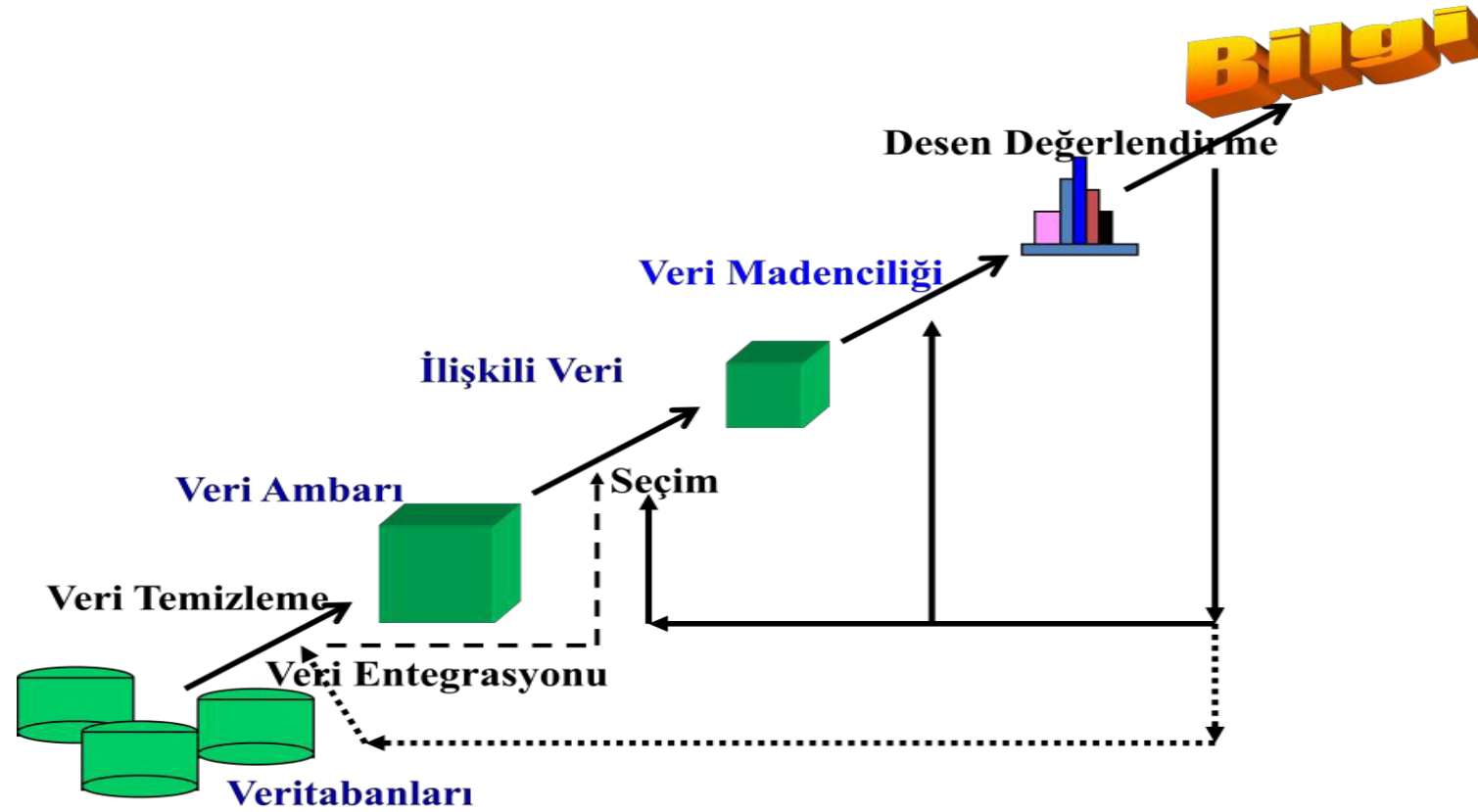
Dr. Öğr. Üyesi Murat AYDOĞAN

Veri Nedir?

- Instagram'da her gün 95 milyon fotoğraf ve video paylaşıyor. Bu dakikada 65,972'ye denk geliyor!
- Ortalama ABD kullanıcısı günde 53 dakika Instagram'da geçiriyor. Bu da yılda 297 saat demek.
- Birçok Instagram kullanıcısı sürekli yeni içerik arıyor, çünkü Instagram'ın keşfet sayfası günlük 200 milyon hesap tarafından görüntüleniyor. Bu da dakikada 138,888 yeni hesabın keşfetmesi demek.
- İnsanların %81'i ürün ve hizmet araştırmak için Instagram kullanıyor.
- İnsanların %50'si Instagram'da gördükleri bir ürünü satın almak için bir web sitesini ziyaret etmiş.
- Neredeyse her dakika yaklaşık 140,000 kullanıcı bir işletme sayfasını ziyaret ediyor.
- Instagram'da yerel pazarlama yapan işletmeler başarı görecektir, çünkü konum belirten Instagram gönderileri %79 daha fazla etkileşim alıyor.



Veri Madenciliği



Veri Madenciliği

Daha önceden bilinmeyen, geçerli ve uygulanabilir bilgilerin geniş veri tabanlarından elde edilmesi ve bu bilgilerin işletme kararları verirken kullanılmasıdır.

Şu genel amaçlar için kullanılır:

- Sınıflandırma
- Birlikte Kuralları ve İlişki Analizi
- Kümeleme

Birliktelik Kuralları ve İlişki Analizi

- Geçmiş tarihli hareketleri çözümllemek, karar destek sistemlerinde verilen kararın kalitesini artırmak için izlenen bir yaklaşımdır.
- 90'lı yılların başına değin teknik yetersizlikten dolayı, kurumlara veya müşterilere satış yapıldığı anda değil, belirli bir zaman aralığında (günlük, haftalık, aylık, yıllık) gerçekleşen satış hareketlerinin tamamına ilişkin genel veriler elektronik ortamda tutulmaktaydı.
- Barkod uygulamalarındaki gelişme ile, bir harekete ait verilerin satış hareketi oluştuğu anda toplanması ve elektronik ortama aktarılması olanaklı hale gelmiştir.
- Genellikle süpermarketlerin satış noktalarında bu tür veriler toplandığından, toplanan bu veriye market sepeti verisi adı verilmiştir.



Birliktelik Kuralları (Association Rules)

- 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.
- Bu tip birlikteliklerin keşfedilmesi, müşterilerin hangi ürünleri bir arada aldıkları bilgisini ortaya çıkarır ve market yöneticileri de bu bilgi ışığında raf düzenlerini belirleyerek satış oranlarını artırabilir ve etkili satış stratejileri geliştirebilirler.
- Market sepeti çözümlemesinin son zamanlarda çok büyük ilgi ile karşılaşmasının sebebi kullanım kolaylığı ve anlaşılabilirliğidir.
- Market sepet analizi ile birliktelik kuralları çıkarımı ilk olarak Agrawal ve çalışma arkadaşları tarafından 1993 yılında ele alınmıştır.



Mining Association Rules between Sets of Items in Large Databases

Rakesh Agrawal Tomasz Imielinski* Arun Swami

IBM Almaden Research Center
650 Harry Road, San Jose, CA 95120

Abstract

We are given a large database of customer transactions. Each transaction consists of items purchased by a customer in a visit. We present an efficient algorithm that generates all significant association rules between items in the database. The algorithm incorporates buffer management and novel estimation and pruning techniques. We also present results of applying this algorithm to sales data obtained from a large retailing company, which shows the effectiveness of the algorithm.

1 Introduction

Consider a supermarket with a large collection of items. Typical business decisions that the management of the supermarket has to make include what to put on sale, how to design coupons, how to place merchandise on shelves in order to maximize the profit, etc. Analysis of past transaction data is a commonly used approach in order to improve the quality of such decisions. Until recently, however, only global data about the cumulative sales during some time period (a day, a week, a month, etc.) was available on the computer. Progress in bar code technology has made it possible to store the so called *basket* data that stores items purchased on a per-transaction basis. Basket data type transactions do not necessarily consist of items bought together at the same point of time. It may consist of items bought by a customer over a period of time. Examples include monthly purchases by members of a book club or a music club.

*Current address: Computer Science Department, Rutgers University, New Brunswick, NJ 08803

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or special permission.

Proceedings of the 1993 ACM SIGMOD Conference
Washington DC, USA, May 1993

Several organizations have collected massive amounts of such data. These data sets are usually stored on tertiary storage and are very slowly migrating to database systems. One of the main reasons for the limited success of database systems in this area is that current database systems do not provide necessary functionality for a user interested in taking advantage of this information.

This paper introduces the problem of “mining” a large collection of basket data type transactions for association rules between sets of items with some minimum specified confidence, and presents an efficient algorithm for this purpose. An example of such an association rule is the statement that 90% of transactions that purchase bread and butter also purchase milk. The antecedent of this rule consists of bread and butter and the consequent consists of milk alone. The number 90% is the confidence factor of the rule.

The work reported in this paper could be viewed as a step towards enhancing databases with functionalities to process queries such as (we have omitted the confidence factor specification):

- Find all rules that have “Diet Coke” as consequent. These rules may help plan what the store should do to boost the sale of Diet Coke.
- Find all rules that have “bagels” in the antecedent. These rules may help determine what products may be impacted if the store discontinues selling bagels.
- Find all rules that have “sausage” in the antecedent and “mustard” in the consequent. This query can be phrased alternatively as a request for the additional items that have to be sold together with sausage in order to make it highly likely that mustard will also be sold.
- Find all the rules relating items located on shelves *A* and *B* in the store. These rules may help shelf planning by determining if the sale of items on shelf *A* is related to the sale of items on shelf *B*.

BEBEK BEZİ - BİRA İLİŞKİSİ

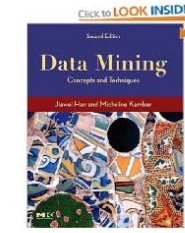
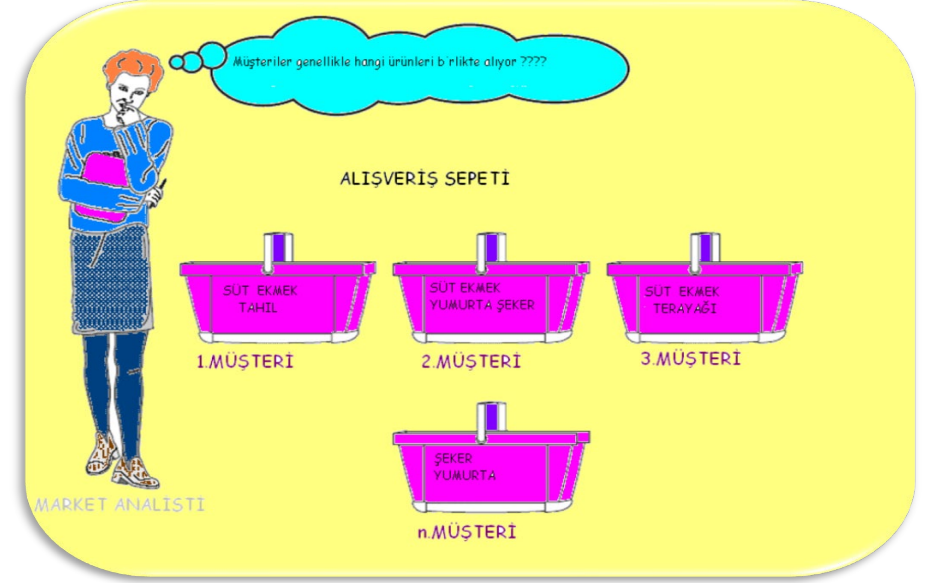
Amerika da zincir marketlerden olan Wal-Mart, müşterilerin yapmış oldukları alışveriş hareketlerini incelediklerinde Cuma akşamları bebek bezi alan erkek müşterilerin beraberinde bira aldıkları ortaya çıkmıştı.

Yapılan analizlere göre ele alınan 200.000 satın alma verisinin 4.000 tanesi bebek bezi içeriyor. 5.500 tanesi ise bira içeriyor. Bunun yanında 3.500 işlem de bebek bezi ve bira ürünlerinin ikisini de içeriyor.



Birliktelik Kuralları ve İlişki Analizi

- Market sepeti verisinde yer alan bir kayıtta, tekil olan hareket numarası, hareket tarihi ve satın alınan ürünlere ilişkin ürün kodu, miktarı, fiyatı gibi bilgiler yer almaktadır.
- Market sepet analizinde (market basket analysis) amaç, satışlar arasındaki ilişkileri bulmak ve buna bağlı kuralları çıkarmaktır.
- Bu ilişkilerin bilinmesi, şirketin kârını arttırmak için kullanılabilir. Eğer X ürününü alanların Y ürününü de çok yüksek olasılıkla aldıkları biliniyorsa ve eğer bir müşteri X ürününü alıyor ama Y ürününü almıyorsa, mevcut müşteri **“potansiyel bir Y müşterisidir”** denilebilir.
- Buna benzer veri analizleri yaparak her ürün için bir sonraki ayın satış tahminleri çıkarılabilir, birlikte satın alınan ürünler için promosyon uygulaması ve reyon dizilişleri yapılabilir, müşteriler satın aldıkları ürünlere göre gruplandırılabilir, yeni bir ürün için potansiyel müşteriler belirlenebilir.



Data Mining, Second Edition: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems) (Hardcover)
by [Jiawei Han](#) (Author), [Micheline Kamber](#) (Author), [Jian Pei](#) (Author)
★★★★★ (2 customer reviews)

RRP: £34.99

Price: **£33.24** & this item **Delivered FREE in the UK** with Super Saver Delivery. [See details and conditions](#)

You Save: **£1.75** (5%)

In stock.

Dispatched from and sold by [Amazon.co.uk](#). Gift-wrap available.

Only 3 left in stock--order soon (more on the way).

Want guaranteed delivery by Friday, September 25? Order it in the next 9 hours and 33 minutes, and choose Express delivery at checkout. [See Details](#)

24 new from £28.99 **8 used** from £29.02

Other Editions: RRP: Our Price: Other Offers:
[Hardcover](#) ~~£34.99~~ £37.99 [12 used & new](#) from £8.90

This title is part of our **Up to 35% Off 1000s of Textbooks** offer.

Frequently Bought Together



Price For All Three: £123.37

[Add all three to Basket](#)

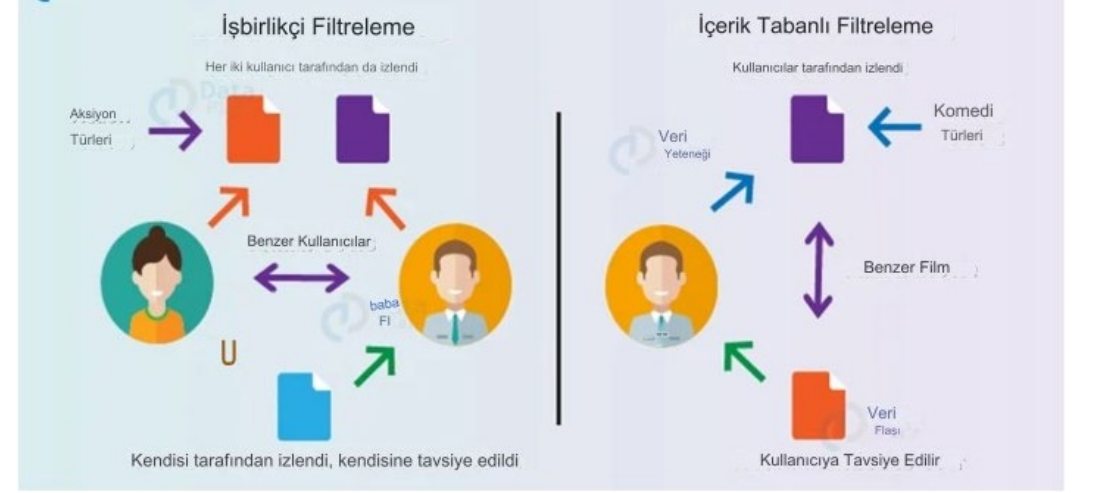
[Show availability and shipping details](#)

✓ **This item:** Data Mining, Second Edition: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems) by Jiawei Han

✓ **Data Mining: Practical Machine Learning Tools and Techniques, Second Edition** (The Morgan Kaufmann Series in Data Management Systems) by Ian H. Witten

Öneri Sistemleri

- Netflix dizi-film tavsiye sistemleri (2006 Netflix prize)
- Spotify şarkı tavsiye sistemleri
- YouTube video tavsiye sistemleri
- E-ticaret sitelerinde önünüze çıkan ürün önerileri



Arka planda işleyen mantığı anlamak adına bazı sorular akla geliyor:

- Beni nereden tanıyor da bu önerilerde bulunuyor?
- Ürüne tıkladığımda anında benzer ürünler geliyor, bu kadar hızlı nasıl çalışıyor bu sistem?
- Arka planda nasıl bir öneri algoritma yapısı var?

NETFLIX

YouTube

Spotify

amazon
prime video

hulu

Birliktelik Kuralları (Association Rules)

- Kuralları oluşturabilmek için destek (support) ve güven (confidence) değerlerini kullanarak, kullanıcı tarafından belirlenmiş minimum destek ve minimum güven değerlerinden yaygın birlikteliklerin belirlenmesi amaçlanmıştır.
- Market sepet analizinde, nesneler müşteriler tarafından satın alınan ürünlerdir ve bir hareket (kayıt) birçok nesneyi içinde bulunduran tek bir satın almadır.
- Birliktelik kurallarının kullanışlı olması için hem konu ile ilgili hem de anlaşılabilir olması gerekir.
- X ürünü alan bir müşterinin Y ürününü de alma durumu(birliktelik kuralı) $X \rightarrow Y$ ile gösterilir.
- Bilgisayar alan bir müşterinin aynı zamanda antivirüs yazılımı almasına ait ilişkilendirme kuralı
 - Bilgisayar $\rightarrow$ Antivirüs yazılımı [destek = %2, güven=%60]
 - **Destek:** Müşterilerin %2 si Bilgisayar ve Antivirüs yazılımını birlikte almıştır.
 - **Güven:** Bilgisayar alan müşterilerin %60'ı Antivirüs yazılımı da almıştır şeklinde ifade edilir.
- Kuralın geçerli olabilmesi için minimum destek ve güven değerlerini sağlamalıdır.

Destek Ölçütü

Destek, bir ilişkinin tüm veri seti içinde kaç kez tekrarlandığını belirler.

Kitap	DVD
Ramses	Harry Potter
Yüzüklerin Efendisi	Harry Potter
Harry Potter	Yüzüklerin Efendisi
Harry Potter	Yüzüklerin Efendisi
Harry Potter	Yüzüklerin Efendisi
Yüzüklerin Efendisi	Yüzüklerin Efendisi
Ramses	Yüzüklerin Efendisi
Yüzüklerin Efendisi	Harry Potter
Yüzüklerin Efendisi	Harry Potter
Ramses	Harry Potter

- Toplam Müşterilerin ancak **% 30** u Harry Potter kitabı ile Yüzüklerin Efendisi DVD lerini birlikte almışlardır. (Klasik Olasılık)

$$Destek(X \rightarrow Y) = \frac{Sayı(X, Y)}{N}$$

$$Destek(HP \rightarrow YE) = \frac{Sayı(HP, YE)}{10}$$

$$Destek(HP \rightarrow YE) = \frac{3}{10}$$

$$Destek(HP \rightarrow YE) = 0,30$$

Güven Ölçütü

Güven, X nesnesinin Y nesnesi ile birlikte bulunma olasılığını ifade eder.

Kitap	DVD
Ramses	Harry Potter
Yüzüklerin Efendisi	Harry Potter
Harry Potter	Yüzüklerin Efendisi
Harry Potter	Yüzüklerin Efendisi
Harry Potter	Yüzüklerin Efendisi
Yüzüklerin Efendisi	Yüzüklerin Efendisi
Harry Potter	Harry Potter
Harry Potter	Harry Potter
Yüzüklerin Efendisi	Harry Potter
Ramses	Harry Potter

- ▶ Harry Potter kitabını alan müşterilerin **%60** i Yüzüklerin Efendisi DVD sini almıştır. (Şartlı Olasılık)

$$\text{Güven}(X \rightarrow Y) = \frac{\text{Sayı}(X, Y)}{\text{Sayı}(X)}$$

$$\text{Güven}(HP \rightarrow YE) = \frac{\text{Sayı}(HP, YE)}{\text{Sayı}(HP)}$$

$$\text{Güven}(HP \rightarrow YE) = \frac{3}{5}$$

$$\text{Güven}(HP \rightarrow YE) = 0,60$$

Birliktelik Kuralı Algoritmaları

Birliktelik kurallarının belirlenmesinde kullanılabilecek çok fazla sayıda algoritma mevcuttur:

- AIS
- SETM
- APRIORI (En bilinen ve en yaygın kullanılan)
- APRIORI-TID
- APRIORI-HYBRID
- FP-GROWTH

Apriori Algoritması

Birliktelik kurallarının oluşturulmasında yararlanılan ve yaygın olarak kullanılan bir algoritmadır. Apriori algoritması aşağıdaki adımlardan oluşmaktadır:

1. Destek ve güven ölçülerini karşılaştırmak için **eşik değerleri** belirlenir.
2. Her bir ürün için **destek sayıları** hesaplanır. Eşik değeri ile karşılaştırılan destek değerlerinin içinden eşik değerinden düşük olanlar çıkarılır.
3. Kalan ürünler ikişerli gruplanarak, **grup destek sayıları** hesaplanır. Grup destek sayıları eşik değeri ile karşılaştırılır. Eşik değerinin altında kalan destek değerli gruplar listeden çıkarılır.
4. Daha sonra üçerli, dörderli, beşerli, vb. biçimde gruplar için aynı karşılaştırma ve eleme işlemi devam ettirilir. Eşik değerlere uygun olduğu sürece işlemler sürecektir.
5. Ürün grubu belirlendikten sonra birliktelik kuralları türetilir ve bu kurallarının her biri için **güven ölçüleri** belirlenir.
6. Eşik güven değerinden küçük güven değerine sahip birliktelik kuralları kural listesinden çıkarılır.

Apriori Algoritması

Kullanılacak olan veri seti tablo yada işlemsel yapıda olmalıdır.
Veri seti kategorik yapıda olmalıdır.

TID	Nesneler
100	Süt, Ekmek, Peynir
101	Yumurta, Süt
102	Ekmek, Süt
103	Peynir, Ekmek
104	Kahve, Süt
105	Ekmek, Peynir, Süt
106	Süt, Ekmek, Yumurta
107	Süt, Peynir, Ekmek
108	Yumurta, Peynir
109	Kahve

TID	Nesneler
100	Süt
100	Ekmek
100	Peynir
101	Yumurta
101	Süt
...	...
...	...
...	...
108	Yumurta
108	Peynir
109	Kahve

Apriori Örnek-1

Verilen tablodaki bilgileri kullanarak Apriori algoritması ile birliktelik kurallarını çıkarınız.

Müşteri ID (TID)	Aldığı Ürünler
5401300197	Gofret, Kola, Su, Çekirdek
5401300198	Antep Fıstığı, Çekirdek, Çikolata, Kola
5401300199	Gofret, Çekirdek, Antep Fıstığı
5401300200	Kola, Su, Çekirdek, Kahve
5401300201	Gofret, Çekirdek, Çikolata, Su
5401300202	Süt, Çekirdek, Su
5401300203	Gofret, Cips, Çekirdek
5401300204	Çikolata, Patlamış Mısır, Su
5401300205	Patlamış Mısır, Kola, Su
5401300206	Su, Süt, Gofret, Çikolata, Çekirdek

Apriori Örnek-1

Öncelikle destek ve güven eşik değerleri belirlenmelidir:

$$Destek_{min} = 0.30$$

$$Güven_{min} = 0.80$$

10 adet müşteri olduğu için destek sayısı = $10 * 0.30 = 3$ olacaktır.

Herhangi bir adımda destek sayısı 3'ten az olan değerler elimine edilecektir.

Ürün Adı	Destek
Gofret	5
Kola	4
Su	7
Çekirdek	8
Antep Fıstığı	2
Çikolata	4
Kahve	1
Süt	2
Cips	1
Patlamış Mısır	2

Müşteri ID (TID)	Aldığı Ürünler
5401300197	Gofret, Kola, Su, Çekirdek
5401300198	Antep Fıstığı, Çekirdek, Çikolata, Kola
5401300199	Gofret, Çekirdek, Antep Fıstığı
5401300200	Kola, Su, Çekirdek, Kahve
5401300201	Gofret, Çekirdek, Çikolata, Su
5401300202	Süt, Çekirdek, Su
5401300203	Gofret, Cips, Çekirdek
5401300204	Çikolata, Patlamış Mısır, Su
5401300205	Patlamış Mısır, Kola, Su
5401300206	Su, Süt, Gofret, Çikolata, Çekirdek

Apriori Örnek-1

$$Destek_{min} = 0.30$$

$$Güven_{min} = 0.80$$

Eşik destek değerine eşit yada büyük ürünler için ikili birliktelikler bulunur ve bunların destek sayıları hesaplanır:

Ürün Adı	Destek
Gofret	5
Kola	4
Su	7
Çekirdek	8
Çikolata	4

İkili	Destek	Aldığı Ürünler
Gofret, Kola	1	Gofret , Kola , Su , Çekirdek
Gofret, Su	3	Antep Fıstığı, Çekirdek , Çikolata, Kola
Gofret, Çekirdek	5	Gofret, Çekirdek, Antep Fıstığı
Gofret, Çikolata	2	Kola , Su, Çekirdek , Kahve
Kola, Su	3	Gofret , Çekirdek, Çikolata, Su
Kola, Çekirdek	3	Süt, Çekirdek, Su
Kola, Çikolata	1	Gofret, Cips, Çekirdek
Su, Çekirdek	5	Çikolata, Patlamış Mısır, Su
Su, Çikolata	3	Patlamış Mısır, Kola, Su
Çekirdek, Çikolata	3	Su , Süt, Gofret , Çikolata, Çekirdek

Apriori Örnek-1

Müşteri ID (TID)	Aldığı Ürünler
5401300197	Gofret, Kola, Su, Çekirdek
5401300198	Antep Fıstığı, Çekirdek, Çikolata, Kola
5401300199	Gofret, Çekirdek, Antep Fıstığı
5401300200	Kola, Su, Çekirdek, Kahve
5401300201	Gofret, Çekirdek, Çikolata, Su
5401300202	Süt, Çekirdek, Su
5401300203	Gofret, Cips, Çekirdek
5401300204	Çikolata, Patlamış Mısır, Su
5401300205	Patlamış Mısır, Kola, Su
5401300206	Su, Süt, Gofret, Çikolata, Çekirdek

$$Destek_{min} = 0.30$$

$$Güven_{min} = 0.80$$

Eşik destek değerine eşit yada büyük ikililerden yararlanarak üçlü birliktelikler bulunur ve bunların destek sayıları hesaplanır:

İkili	Destek
Gofret, Su	3
Gofret, Çekirdek	5
Kola, Su	3
Kola, Çekirdek	3
Su, Çekirdek	5
Su, Çikolata	3
Çekirdek, Çikolata	3

Üçlüler	Destek
Gofret, Su, Çekirdek	3
Gofret, Su, Kola	1
Gofret, Su, Çikolata	2
Gofret, Çekirdek, Kola	1
Gofret, Çekirdek, Çikolata	2
Kola, Su, Çekirdek	2
Kola, Su, Çikolata	0
Kola, Çekirdek, Çikolata	1
Su, Çekirdek, Çikolata	2

Apriori Örnek-1

$$Destek_{min} = 0.30$$

$$Güven_{min} = 0.80$$

{Gofret, su, çekirdek} için altkümeler dikkate alınarak birliktelik kuralları türetilir ve güven değerleri hesaplanır.

Müşteri ID (TID)	Aldığı Ürünler
5401300197	Gofret, Kola, Su, Çekirdek
5401300198	Antep Fıstığı, Çekirdek, Çikolata, Kola
5401300199	Gofret, Çekirdek, Antep Fıstığı
5401300200	Kola, Su, Çekirdek, Kahve
5401300201	Gofret, Çekirdek, Çikolata, Su
5401300202	Süt, Çekirdek, Su
5401300203	Gofret, Cips, Çekirdek
5401300204	Çikolata, Patlamış Mısır, Su
5401300205	Patlamış Mısır, Kola, Su
5401300206	Su, Süt, Gofret, Çikolata, Çekirdek

Birliktelik Kuralı	Açıklama (Şartlı Olasılık)	Güven
Gofret & Su → Çekirdek	Gofret ve Su alanların Çekirdek alma ihtimali	3/3=1,00
Gofret & Çekirdek → Su	Gofret ve Çekirdek alanların Su alma ihtimali	3/5=0,60
Su & Çekirdek → Gofret	Su ve Çekirdek alanların Gofret alma ihtimali	3/5=0,60
Gofret → Su & Çekirdek	Gofret alanların Su ve Çekirdek alma ihtimali	3/5=0,60
Su → Gofret & Çekirdek	Su alanların Gofret ve Çekirdek alma ihtimali	3/7=0,42
Çekirdek → Gofret & Su	Çekirdek Alanların gofret ve Su alma ihtimali	3/8=0,38

Apriori Örnek-1

$$Destek_{min} = 0.30$$

$$Güven_{min} = 0.80$$

Başlangıçta eşik güven değeri %80 olarak verilmişti. Güven değeri %80'e eşit yad %80'den fazla olan kurallar:

Gofret&Su → çekirdek (Güven değeri %100)

Gofret ve su alan herkes çekirdek de alacaktır (%100 ihtimalle)

Müşteri ID (TID)	Aldığı Ürünler
5401300197	Gofret, Kola, Su, Çekirdek
5401300198	Antep Fıstığı, Çekirdek, Çikolata, Kola
5401300199	Gofret, Çekirdek, Antep Fıstığı
5401300200	Kola, Su, Çekirdek, Kahve
5401300201	Gofret, Çekirdek, Çikolata, Su
5401300202	Süt, Çekirdek, Su
5401300203	Gofret, Cips, Çekirdek
5401300204	Çikolata, Patlamış Mısır, Su
5401300205	Patlamış Mısır, Kola, Su
5401300206	Su, Süt, Gofret, Çikolata, Çekirdek

Apriori Örnek-2

Tabloda verilen alışveriş bilgileri için Apriori algoritması kullanarak birliktelik kurallarını elde ediniz.

Minimum destek sayısı = 3

Minimum güven = %80

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

Apriori Örnek-2

Minimum destek sayısı = 3
Minimum güven = %80

Minimum destek değerinden daha küçük destek değerine sahip olan itemler elenir.

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

İtem	Destek Değeri
Domates	8
Çay	7
Elma	5
Şeker	4
Un	4
Ekmek	2
Üzüm	2
Makarna	2
Peynir	1
Zeytin	1

Apriori Örnek-2

Minimum destek sayısı = 3
Minimum güven = %80

Elde edilen tekli birliktelikler dikkate alınarak ikili birlikteliklerin oluşturulması

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

İtem	Destek Değeri
Domates	8
Çay	7
Elma	5
Şeker	4
Un	4

İtem set	Destek Değeri
Elma , Şeker	1
Elma , Çay	3
Elma , Un	2
Elma , Domates	5
Şeker , Çay	3
Şeker , Un	2
Şeker , Domates	3
Çay , Un	3
Çay , Domates	5
Un, domates	3

Apriori Örnek-2

Minimum destek sayısı = 3

Minimum güven = %80

İkili birlikteliklerin destek değeri eşik destek değeri ile karşılaştırılır ve eşik değerinden küçük destek değerli birliktelikler elenir.

Kalan ikili birlikteliklerden yararlanılarak 3lü birliktelikler elde edilir.

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

İtem set	Destek Değeri
Elma , Şeker	1
Elma , Çay	3
Elma , Un	2
Elma , Domates	5
Şeker , Çay	3
Şeker , Un	2
Şeker , Domates	3
Çay , Un	3
Çay , Domates	5
Un, domates	3

Apriori Örnek-2

Minimum destek sayısı = 3

Minimum güven = %80

Üçlü birlikteliklerin oluşturulması

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

İtem set	Destek Değeri
Elma , Çay	3
Elma , Domates	5
Şeker , Çay	3
Şeker , Domates	3
Çay , Un	3
Çay , Domates	5
Un, domates	3

İtem set	Destek Değeri
Elma, çay, domates	3
Elma, çay, şeker	1
Elma, çay, un	2
Elma, domates, şeker	1
Elma, domates, un	2
Şeker, çay, domates	2
Şeker, çay, un	0
Şeker, domates, un	1
Çay, un, domates	2

Apriori Örnek-2

Minimum destek sayısı = 3
Minimum güven = %80

Dörtlü birlikteliklerin oluşturulması

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

İtem set	Destek Değeri
Elma, çay, domates, şeker	1
Elma, çay, domates, un	2
Elma, çay, domates, makarna	1

{Elma, çay, domates} üçlü birlikteliklerini kullanarak elde edilen dörtlü birlikteliklerin tamamı eşik destek sayısı altında kaldığı için en fazla üçlü birlikteliklere ulaşabildik.

Apriori Örnek-2

Minimum destek sayısı = 3
Minimum güven = %80

{elma, çay, domates} için birliktelik kuralları:

Müşteri Numarası	Aldığı Ürünler
101	Elma, Şeker, Çay, Domates
102	Ekmek, Domates, Un, Şeker
103	Elma, Domates, Ekmek
104	Şeker, Çay, Domates, Peynir,
105	Elma, Domates, Un, Çay
106	Makarna, Domates, Çay
107	Elma, Zeytin, Domates
108	Un, Üzüm, Çay
109	Üzüm, Şeker, Çay
110	Çay, Makarna, Elma, Un, Domates

Birliktelik	Açıklama	Güven
Elma & Çay -> Domates	Elma ve Çayın bulunduğu item-sette domatesin olma olasılığı	$3/3 = \%100$
Elma & Domates -> Çay	Elma ve domatesin bulunduğu item-sette çayın olma olasılığı	$3/5 = \%60$
Çay & Domates -> Elma	Çay ve Domatesin bulunduğu item-sette elmanın olma olasılığı	$3/5 = \%60$
Çay -> Elma & Domates	Çayın bulunduğu item-sette elma ve domatesin olma olasılığı	$3/7 = \%42$
Domates -> Elma & Çay	Domatesin bulunduğu item-sette çay ve elmanın olma olasılığı	$3/8 = \%38$
Elma -> Çay & Domates	Elmanın bulunduğu item-sette çay ve domatesin olma olasılığı	$3/5 = \%60$

Apriori Örnek-2

Minimum destek sayısı = 3
Minimum güven = %80

Eşik güven değerinin altında kalan kuralların elenmesi

Başlangıçta minimum güven = %80 olarak belirlenmişti.
Elde edilen kuralların güven değeri eşik değerin altında ise kural listesinden çıkartılırlar.

Elma & Çay -> Domates = %100

Elmanın ve Çayın birlikte satıldığı alışverişte domatesin satılma olasılığı **%100 'dür.**

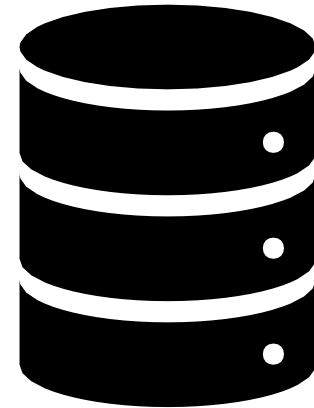
Ücretsiz veri setleri

<http://archive.ics.uci.edu/ml/index.php> (UCI Machine Learning Repository)

https://github.com/viktree/curly-octo-chainsaw/blob/master/BreadBasket_DMS.csv Market sepet analizi için veri

<https://github.com/awesomedata/awesome-public-datasets#machinelearning> Liste halinde farklı veri kaynakları

<https://www.kaggle.com/datasets>



Referanslar

- ***Veri madenciliği ders notları, Dr. Burcu Çarklı YAVUZ, SAÜ (Temel İçerik)***
- Data Mining: Concepts and Techniques, J. Han, M. Kamber, J. Pei
- Veri madenciliği Kavram ve Algoritmaları, Gökhan Silahtaroğlu, Papatya Yayıncılık
- Veri Madenciliği Yöntemleri, Yalçın Özkan, Papatya Yayıncılık
- Veri madenciliği ders notları, Doç. Dr. Nilüfer YURTAY, SAÜ
- Veri madenciliği ders notları, Dr. Tuğrul TAŞCI, SAÜ

