**BİRLİKTELİK KURALLARI ﴾֎﴿**

Bir arada gerçekleşen olayları çözümlemek için kullanılır. Örneğin bir mağazadan etek alan müşterilerin %70’inin aynı alışverişte kazakta aldığını tespit etmek.

Tanım: Olayların birlikte gerçekleşme durumlarını çözümleyen veri madenciliği yöntemlerine birliktelik kuralları(association rules) adı verilir. Bu yöntemler, birlikte olma kurallarını belirli olasılıklarla ortaya koyar. Veri madenciliğinde bir arada gerçekleşen olayları çözümlemek amacıyla özel algoritmalar geliştirilmiştir. Özellikle perakende alışverişlerde sepet analizi adı verilen bu tür çalışmalar kullanılmak üzere R paketleri geliştirilmiştir.

Pazar-sepet sepet çözümlemelerinde satılan ürünler arasındaki ilişkileri ortaya koymak için “destek” ve “güven” gibi ölçütlerden yararlanılır.

**Destek ve Güven Ölçütleri**

Birliktelik analizlerinde ögeler arasındaki ilişkileri ortaya koymak için “destek” ve “”güven” gibi ölçütten yararlanılır.

**Destek(Support)**

Bir seçenek veya seçenek kümesinin tüm gözlemler içerisindeki oranını gösterir. Yani bir seçenek veya seçenek kümesinin klasik olasılığını ifade eder. Bir diğer tanım: Bir ilişkinin tüm alışverişler içinde hangi oranda tekrarlandığını belirler.

**X seçeneği için destek**

**X ve Y seçenek kümesi için destek**

**n=gözlem sayısı**

**N(X)=X seçeneğinin görülme sayısını**

**N(X,Y)= X,Y seçenek kümesinin birlikte görülme sayısını gösterir.**

**ÖNEMLİ:** Birliktelik kurallarının oluşabilmesi için minimum destek değeri (D) nin önceden belirlenmesi gerekir. 0 ≤ D ≤ 1 aralığındaki **bu değer sıfıra(0) yaklaştıkça birlikteliklerin sayısı artar,** 1’e yaklaştığında birliktelik bulmak zorlaşacağından birlikteliklerin sayısı azalır. **Uygulamada genellikle 0.30 ya da 0.40 alınır.**

**Güven(Confidence)**

X seçeneğini seçenlerin Y seçeneğini de seçme olasılığıdır. Yani, X verilmişken Y’nin koşullu olasılığını ifade eder. Bir diğer tanım: X ürün grubunu alan müşterilerin Y ürün grubunu da alma olasılığını ortaya koyar.

Veya

şeklinde hesaplanır.

**ÖNEMLİ:** Elde edilen birlikteliklerin yorumlanmasında güven kriteri önemlidir. Güvem kriteri (G) 0≤G≤1 aralığında yer alır. **Uygulamada genelde 0.50 veya üzerinde bir değer alınır.**

**Kaldıraç(Lift)**

İstenilen bir X seçeneğinin B yerine A içinde aramasının (BᴝA ) o özellikle karşılaşma olasılığı kaç kez arttıracağını gösterir. Destek/ Güven şeklinde hesaplanır. X seçeneğine ilişkin kaldıraç değeri( K) aşağıdaki gibi hesaplanır. Lift basit bir ilişki ölçütü olarak değerlendirilebilir.

Örnek 1: Önceden verilen TV servis sağlayıcısı örneğinde 10 bireyden elde edilen veriler aşağıdaki şekildedir.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Birey | Haber | Spor | Eglence | Dizi | Sinema | Belgesel |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| 2 | 0 | 0 | 0 | 1 | 1 | 1 |
| 3 | 0 | 0 | 1 | 0 | 1 | 0 |
| 4 | 1 | 0 | 0 | 0 | 1 | 1 |
| 5 | 0 | 0 | 0 | 1 | 1 | 0 |
| 6 | 0 | 0 | 0 | 1 | 1 | 1 |
| 7 | 0 | 1 | 0 | 1 | 1 | 0 |
| 8 | 0 | 1 | 0 | 1 | 0 | 1 |
| 9 | 0 | 1 | 1 | 1 | 1 | 0 |
| 10 | 1 | 1 | 1 | 0 | 0 | 1 |

Buna göre (Dizi→ Sinema) seçenek kümesi için destek, güven ve kaldıraç değerlerini hesaplayalım.

**DESTEK**

**N(Dizi, Sinema)= 5**

**P(Dizi→ Sinema) = N(Dizi, Sinema) /n = 5/10 =0.5**

**YORUM: Dizi ve sinemanın birlikte izlenme olasılığı 0.5’tir veya müşterilerin %50’si dizi ve sinema kanallarını izlemektedir**.

**GÜVEN**

**N(Dizi)= 6 (Kırmızı+mavi renk)**

**P(Sinema/Dizi) = N(Dizi, Sinema) / N(Dizi)= 5/6=0.83**

**YORUM: Dizi kanallarını izleyen müşterilerin %83’ü aynı zamanda sinema kanallarını da izlemektedir.**

**P(Sinema) = 7/10=0.7( Kırmızı+yeşil)**

**P(Sinema/Dizi)= 0.83**

**K(Sinema)= P(Sinema/Dizi) / P(Sinema) =0.83 / 0.7= 1.19**

**YORUM: Sinema kanallarının izleyen müşterilere kampanya yapılacaksa, tüm müşterilerden rastgele seçim yapmak yerine dizi kanallarını izleyen müşterilerden rastgele seçim yapılırsa sinema kanalını izleyen müşteriyle karşılaşma olasılığı 1.19 katına çıkar.**

Örnek 2: Market Örneği

|  |  |
| --- | --- |
| Fiş No | Detay |
| 1 | Ekmek, Süt, Çerez, Salça |
| 2 | Biraz, Bebek Bezi, Ekmek, Yumurta |
| 3 | Bira, Kola, Bebek Bezi, Süt |
| 4 | Bira, Ekmek, Bebek Bezi, Süt, Çerez |
| 5 | Kola, Ekmek, Bebek Bezi, Süt |
| 6 | Bira, Ekmek, Bebek Bezi, Süt, Salça |
| 7 | Kola, Ekmek, Bebek Bezi, Süt |

AŞAMA 1 :Dönüştür varlık durumu 1 ile, yokluk durumu 0 ile gösterilir.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| fis no | ekmek | sut | cerez | salca | bira | bebekbezi | yumurta | kola |
| 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| 3 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| 4 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| 5 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| 6 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| 7 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |

Ekmeğin frekans değeri= 6

Sütün f.d=6

Çerezin f.d=2 gibi

Support(destek) değeri= frekans/ toplam girdi sayısı

Minimum support değerini biz belirliyoruz. **Bu örnekte %40 alınmıştır.**

AŞAMA 2: Destek(Support) değerlerini hesaplayalım

|  |  |  |
| --- | --- | --- |
| urun | frekans | support |
| ekmek | 6 | 0.857143 |
| sut | 6 | 0.857143 |
| cerez | 2 | 0.285714 |
| salca | 2 | 0.285714 |
| bira | 4 | 0.571429 |
| bebekbezi | 6 | 0.857143 |
| yumurta | 1 | 0.142857 |
| kola | 3 | 0.428571 |

Ekmeğin destek değeri 6/7=0.85

AŞAMA 3: Daha önce belirlenen minimum support değeri kullanılacak. Bu değer %40 olarak belirlenmişti. Bunun üzerinde ve eşit olan değerlerle yeni tablo oluşturulur.

|  |  |  |
| --- | --- | --- |
| urun | frekans | support |
| ekmek | 6 | 0.857143 |
| sut | 6 | 0.857143 |
| bira | 4 | 0.571429 |
| bebekbezi | 6 | 0.857143 |
| kola | 3 | 0.428571 |

AŞAMA 4: Yeni tablo oluştu. Elimizde kalan değerlerin birlikte olma durumlarını incelemek için bir tablo daha oluşturacağız. 2’li kombinasyon

|  |  |  |
| --- | --- | --- |
| urun | frekans | support |
| ekmek,sut | 5 | 0.714286 |
| ekmek,bira | 3 | 0.428571 |
| ekmek, bebekbezi | 5 | 0.714286 |
| ekmek,kola | 2 | 0.285714 |
| sut,bira | 3 | 0.428571 |
| sut,bebekbezi | 5 | 0.714286 |
| sut,kola | 3 | 0.428571 |
| bira,bebekbezi | 4 | 0.571429 |
| bira,kola | 1 | 0.142857 |
| bebekbezi,kola | 3 | 0.428571 |

Min support değeri 0.4 idi. Tabloya bak. 0.40’ a eşit ve üzerinde olanlarla yeni tablo oluştur.

AŞAMA 5:

|  |  |  |
| --- | --- | --- |
| urun | frekans | support |
| ekmek,sut | 5 | 0.714286 |
| ekmek,bira | 3 | 0.428571 |
| ekmek, bebekbezi | 5 | 0.714286 |
| sut,bira | 3 | 0.428571 |
| sut,bebekbezi | 5 | 0.714286 |
| sut,kola | 3 | 0.428571 |
| bira,bebekbezi | 4 | 0.571429 |
| bebekbezi,kola | 3 | 0.428571 |

Kombinasyonlarını yap

AŞAMA 6

|  |  |  |
| --- | --- | --- |
| urun | frekans | support |
| ekmek,sut,bira | 2 | 0.285714 |
| ekmek,sut,bebekbezi | 4 | 0.571429 |
| ekmek,bira,bebekbezi | 5 | 0.714286 |
| sut,bira,bebekbezi | 3 | 0.428571 |
| sut,bira,kola | 5 | 0.714286 |
| sut,bebekbezi,kola | 3 | 0.428571 |

O.40 eşit ve üzeri olanlar tablo yap

AŞAMA 7