



Araştırma Makalesi • Research Article

R Programlama Dili İle Pazar Sepet Analizi: Muş İl Merkezindeki Bir Süpermarkette Tüketicilerin Satın Alma Davranışlarının Tespiti Üzerine Bir Uygulama

Market Basket Analysis with R programming Language: An Application on consumer purchasing behavior of a Supermarket in Muş

Emrah Bilgiç^{a*}

^a Dr. Öğr. Üyesi, Muş Alparslan Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Güzeltepe, Muş
ORCID: 0000-0002-9875-2299

MAKALE BİLGİSİ

Makale Geçmişi:

Başvuru tarihi: 20 Eylül 2018

Düzeltilme tarihi: 05 Kasım 2018

Kabul tarihi: 23 Kasım 2018

Anahtar Kelimeler:

Birliktelik Kuralları Analizi

R Programlama Dili

Satın alma Davranışı

ÖZ

Bu araştırma, Veri Madenciliği tekniklerinden biri olan Birliktelik Kuralları Analizi'ni, Muş ilinde faaliyet gösteren bir süpermarkete ait verilere R programlama dilini kullanarak uygulamaktadır. Perakende sektöründe tüketicilerin tercihlerini, satın alma davranışlarını anlamak, en uygun üretim ve pazarlama stratejileri geliştirebilmek açısından çok önem arz etmektedir. Çalışma, araştırmaya konu olan firmaya faydalı sonuçlar sağlamakla kalmayıp, hem araştırmacıların hem de perakendecilerin sahip oldukları verileri gelişmiş algoritmalarla analiz edebilmeleri için çalışma boyunca detaylı bir şekilde kodları paylaşılan R programlama dilini kullanmaya teşvik edici niteliktedir. Ayrıca birçok çalışmadan farklı olarak ürünler arasındaki ilişkiler bulunurken çok genel ürün gruplarıyla çalışmak yerine daha ayrıntılı ürün gruplarıyla çalışılmıştır. Analizler sonucunda en güçlü satın alma davranışının; yumurta alan müşterilerin manav reyonundan da alışveriş yaptığı kuralı olduğu tespit edilmiştir.

ARTICLE INFO

Article history:

Received 20 September 2018

Received in revised form 05 November 2018

Accepted 23 November 2018

Keywords:

Association Rules Analysis

R Programming Language

Purchasing Behaviour

ABSTRACT

This study applies one of the tools of Data Mining, Association Rules Analysis by using R programming language for a grocery located in the city Mus/Turkey. Understanding consumers' preferences and behaviors is very crucial for the companies in terms of constructing optimal production and marketing strategies. In addition to providing useful information to the company, this study will also encourage researcher and retailers to use R language for analyzing data with advanced algorithms. Furthermore, while finding the relationships between the products, detailed product groups are used rather than general product categories. The results of analyzes indicate that the most powerful purchasing behavior is the rule of: the customers purchased eggs also purchased some fruit or vegetable from the greengrocery.

1. Giriş

Günümüzde bilgi teknolojilerinin hızla gelişmesi ile beraber firmalar çok daha kolay ve düşük maliyetli bir şekilde ilgilendikleri verileri elde edebilmekte, depolamakta ve analiz edebilmektedir. Örneğin, mağazalardan veya internet sitelerinden yapılan alışveriş işlemleri firmaların veri tabanlarına anında akmakta ve çeşitli paket programlar ile analiz edilebilmektedir. Günümüzde büyük çaptaki verileri (Big Data) analiz edebilmek, faydalı ve ilginç çıkarımlarda bulunabilmek amacıyla Büyük Veri Bilimi'nin

bir parçası olan Veri Madenciliği teknikleri halen yoğun bir şekilde kullanılmaktadır.

Perakende mağazalar veya e-ticaret sitelerinden alışveriş yapan müşterilerin satın alma davranışlarını belirleyebilen Birliktelik Kuralları Analizi, en çok kullanılan Veri Madenciliği tekniklerinden biri olup, birbirleri ile yüksek korelasyona sahip ürünleri tespit etmektedir. Büyük Veri çağı ile tüketicilere ait gerçek hayat verilerinden yararlanılarak önemli ve faydalı bilgiler elde edilmesine yönelik araştırmalar giderek önem kazandığından, Birliktelik Kuralları Analizi, özellikle pazarlama ile alakalı çalışmalarda

*Sorumlu yazar/Corresponding author
e-posta: dr.emrahbilgic@gmail.com

(Lee, Park ve Moon, 2013; Nengsih, 2015; Kaur ve Kang, (2016); Guo, Wang ve Li, 2017; Chiang, 2018) çok yoğun olarak karşımıza çıkmaktadır. Birliktelik Kuralları Analizi ayrıca sağlık alanındaki çalışmalarda, web sayfalarını tıklama analizleri vb. önemli Veri Madenciliği uygulamalarında da kullanılmaktadır (Aggarwal, 2015, s.93).

Bu çalışmada Muş il merkezinde faaliyet gösteren yerel bir süpermarkete ait satış verileri R programlama dili kullanılarak Birliktelik Kuralları Analizi ile analiz edilecektir. Bu sayede, müşterilerin satın alma davranışları bir başka ifade ile hangi ürünleri birlikte satın aldıkları tespit edilecek ve bu bulgularla firma kârını arttırmaya yönelik pazarlama faaliyetlerine yön verilecektir. Analizlerde kullanılan kodlar, R dilinin kullanılabilirliğini göstermek amacıyla ve hem perakendecileri hem de araştırmacıları dili kullanmaya teşvik etmek amacıyla açıklamalı bir şekilde paylaşılacaktır. İlgili literatürün güncel durumunu tespit etme amacı da taşıyan çalışma şu kısımlardan oluşacaktır: İlk bölümde güncel çalışmalara yer verildikten sonra ikinci bölümde Birliktelik Kuralları Analizi ayrıntılı bir şekilde ele alınacak, daha sonra üçüncü bölümde uygulamaya geçilecek ve dördüncü, sonuç bölümüyle de çalışma sonlandırılacaktır.

1.1. Literatür Taraması

Birliktelik Kuralları Analizi (BKA), Veri Madenciliği tekniklerinden biri olup, büyük veri setlerindeki veri noktalarının ilginç birlikteliklerini (ilişkilerini, bağımlılıklarını) keşfetmeyi amaçlamaktadır. Analize girdi sağlayan işlem (transaction) verileri bir perakende mağazasından elde edilecek her bir satın alma kaydından (bir satın almada diğer bir deyişle bir sepette tek bir ürün veya binlerce ürün yer alabilir) oluşmaktadır. Sadece perakende uygulamalarında değil, daha önce vurgulandığı gibi metin madenciliği ve web tıklama analizleri gibi farklı uygulamalarda da karşımıza çıkmaktadır (Giudici, 2005 s.90; Cios, Pedrycz, ve Swiniarski 2007, s.289; Aggarwal, 2015, s.93).

Pazar Sepet Analizi (Market Basket Analysis), BKA'nın en çok bilinen uygulaması olup, bir perakende mağazasından, müşterilerin işlemlerini (satış fişleri) analiz ederek birlikte satın alınmış ürünleri tespit etmektedir. Buradaki sepet kavramı, alışverişte kullanılan sepetleri veya alışveriş arabalarını ifade etmektedir. Pazar Sepet Analizi mağaza yerleşim planlaması, ürün çeşitlendirme, çapraz satış, raf dizaynı, katalog dizaynı, müşteri segmentasyonu, satın alma süreçleri gibi birçok önemli karar alma süreçlerinde destek sağlamaktadır. (Berry ve Linoff, 1997, s.287; Özcan ve Esnaf (2010), Çil, 2012; Leskovec, Rajaraman, ve Ullman, 2014, s.202).

BKA kullanılarak son yıllarda yapılan çalışmalar şu şekilde özetlenebilir:

Pehlivanoğlu ve Duru (2015), ortaokul öğrencilerinin sosyal ağları kullanımı hakkında yapmış oldukları çalışmada, 28 sorudan oluşan anket verilerini BKA kullanarak analiz etmişlerdir. Araştırma sonucunda "Facebook'a günde 2 saat girenler bunun kendilerini engellediğini düşünüyor ve günde 1 saat ders çalışıyorlar" gibi kurallar elde edilmiştir.

Güllüoğlu (2015), yapmış olduğu çalışmada, müşteri profillerini Apriori algoritması kullanılarak ortaya çıkan birliktelik kuralları ile ilişkilendirmiş ve en sık tercih edilen

ürün gruplarını cinsiyetlere göre ve yaş dağılımlarına göre belirlemiştir. Çalışmada İstanbul'daki bir süpermarketin 300 müşterisine ait satış verileri kullanılmıştır.

Tekin vd.(2015), Konya'daki bir süpermarketten elde ettikleri bir yıllık, toplamda 962 bin satış fişinden oluşan veri setini kullanmışlardır. Bu veri seti ile müşterilerin ürün kategorileri bazındaki satın alma davranışları Apriori algoritması ile incelenmiş ayrıca sonuçlar firmanın pazarlama faaliyetlerini desteklemek için ayrıntılı bir şekilde incelenmiştir. Örneğin et ürünlerinin ve günlük gıdaların ilişkili olduğu ürün grubu taze sebzeler iken, hazır giyim grubunun ise bisküvi grubu olduğu tespit edilmiştir.

Kokoç vd. (2016), en uygun tesis yerleşim planlamasının etkin ürün/hizmet üretimi için çok önemli olduğunu vurgulayıp, bir hastane için BKA kullanarak tesis yerleşimi için bir uygulama gerçekleştirmişlerdir. Apriori algoritmasının verdiği sonuca göre, "ambulans girişi" ve "yeşil alan" kuralı en yüksek güven değerine sahiptir.

Akgün (2016), yazılım hata tespiti amacıyla yapmış olduğu çalışmada, geliştirilen yazılım projelerinden toplanan verilerdeki hataların tespiti için, veri madenciliği tekniklerinden BKA ile yazılımcı-hata ilişkilerini içeren kuralların çıkartılması sağlamıştır.

Ertuğrul vd. (2016), birliktelik kuralları analizini, çok boyutlu ölçekleme analizi ile beraber raf dizaynı için kullanmışlardır. 750 müşterinin 16 farklı bisküvi satın alımlarını içeren veri setinde Apriori algoritması kullanılarak bisküviler için raf dizaynı yapılmıştır.

Öztürk ve Tanrısevdi (2017), pazarlama stratejileri geliştirmek amacıyla Kuşadası'nı ziyaret eden kruvaziyer gemi yolcularına ait verileri, BKA ile analiz etmişlerdir. 408 turiste ait anket verilerinin analizi sonucunda en küçük destek değeri % 30 olan ve en küçük güven değeri % 80 olan; kruvaziyer seyahate karar vermede en çok etkilenen faktör "kendi kararı" olan ve medeni durumu "evli" olan ziyaretçilerin %82 olasılıkla seyahat eşlikçileri "eşleri"dir gibi 76 adet kural belirlenmiştir.

Huseyinov ve Aytaç (2017), 9835 işleminden ve 169 ürün kategorisinden oluşan bir veri setine Apriori ve Eclat algoritmalarını uygulamışlardır. En çok satılan ürünlerin ve güven değeri %50'nin üzerindeki kuralların tespit edildiği çalışmada R programlama dili kullanıldığına vurgu yapılmış fakat paketler ve kodlar hakkında bilgi verilmemiştir.

Özçalıcı (2017), ikinci el araç piyasası için yaptığı araştırmada, bir e-ticaret sitesinden iki yüz binin üzerindeki otomobile ait 73 farklı değişkenden oluşan veri setini web kazıma yöntemi ile bir araya getirmiştir. Apriori algoritmasını kullanabilmek için değerleri tekrardan düzenlenen değişkenler ile destek değeri %10 ve güven değeri %70'in üzerinde olan 100 kural tespit edilmiştir. Çalışma sonucunda şuna benzer sonuçlar çıkarılabilir: fiyatı 30 000 TL ile 50 000 TL arasında (50 000 TL dâhil) bir araç satın alınmak istenirse, araçlarda büyük olasılıkla ABS, elektrikli ön cam, merkezi kilit, CD çalar, elektrikli ayna ve manuel vites donanımları yer alacaktır.

Sözen vd. (2017), mobilya sektörü için yapmış oldukları çalışmada anket verileri toplayarak, verilere BKA uygulamışlardır. 74 mobilya kullanıcılarına yapılan ve demografik bilgilerin yanı sıra mobilya hakkındaki soruların

(fiyat, estetik, dayanıklılık vb.) yer aldığı anketlerin BKA ile analizi sonucunda; örneğin evli kişiler mobilya satın alırken daha çok estetiğe önem vermektedir gibi kurallar elde edilmiştir.

Budak vd. (2018), bir web sitesinin ziyaretçilerinin ihtiyaçlarını ve web sitesinin tasarım ve içerik anlamında bu ihtiyaçları giderecek biçimde yenilenmesini web madenciliği ile tespit etmek amacıyla BKA uygulamasını gerçekleştirmişlerdir. Bir üniversitenin bir aylık web günlük dosyasının kullanıldığı çalışmada, veri temizleme işlemi yapıldıktan sonra elde edilen yaklaşık 39 bin satır veri (ziyaretçi IP, erişim zamanı, istek, kullanıcı adı, arana kelime gibi veriler) BKA ile analiz edilmiştir. “Yatay Geçiş” (1177kez), “Harç” (889 kez) ve “Ders Programı” (600 kez) gibi kelimelerin en çok kullanılan kelimeler olduğu ayrıca elde edilen güçlü kuralların ise şu şekilde olduğu ortaya çıkmıştır: “Yatay Geçiş Başvuruları”nı arayanların %60’ı “Yatay Geçiş” kelimesini de aratmıştır. Yine “Ders Programı” ve “Yatay Geçiş” kelimelerini arayanların %60’ı “Harç” kelimesini de aratmıştır.

Sağın ve Ayvaz (2018), bir bilgisayar firmasına ait 50 binin üzerindeki satış işlemini analiz etmiştir. 293 farklı ürünün yer aldığı veri seti 13 ana kategori ve 93 alt kategoride tekrar düzenlenmiş, ayrıca veri seti algoritmaların karşılaştırılması için iki parçaya ayrılmıştır. Apriori ve FP-Growth algoritmalarının kullanıldığı çalışmada eşik değerleri olan %40’lık güven düzeyi ve %50’lik destek değeri ile Apriori algoritmasıyla 24 kural, FP-Growth algoritması ile 2 kural tespit edilmiştir. Sonuç olarak FP-Growth algoritmasının daha hızlı çalıştığı fakat kuralları elde edebilme açısından Apriori algoritmasının daha üstün olduğu tespit edilmiştir.

Bilgiç ve Çakır (2018), İstanbul’da faaliyet gösteren bir süpermarket zincirine ait 175 mağazanın yaklaşık 150 bin müşterisinin 600 bin kadar işlemlerini ve mağazaların bulunduğu bölgelerin sosyo-ekonomik yapısını analiz etmişlerdir. İki farklı yaklaşımla kümelere ayrılan süpermarket mağazaları için ilk yaklaşımda sosyo-ekonomik verilerle mağaza kümelemesi yapılırken, ikinci yaklaşımda müşterilerin satın alma davranışları BKA ile tespit edilerek mağaza kümelemesi yapılmıştır. Son olarak iki yaklaşım karşılaştırılmıştır.

Bilindiği üzere Büyük Veri çağı ile hem insanların hem de firma ve kurumların hayatına Bulut Sistemler de girmiştir. Günümüzde, Veri Madenciliği, Veri Analitiği, Büyük Veri Analizi gibi isimlerle adlandırılan veri analizi konusu bulut sistemlerde de kendisine alan bulmuş ve bu yönde araştırmaların sayısı hızla artmaktadır. Dolayısıyla, büyük miktarlardaki verilerle pazar sepet analizini yüksek performanslı yapabilmek için bulut teknolojilerinden de yararlanılmaktadır. Son günlerde yapılan bir çalışmada, Yıldırım ve Birant (2018) bulut bilişimde BKA’ni ayrıntılı bir şekilde incelemişlerdir.

2. Yöntem

Bu çalışmada Veri Madenciliği tekniklerinden biri olan ve çok yaygın olarak, farklı amaçlar için kullanılan Birliktelik Kuralları Analizi uygulanacaktır. Ayrıca birliktelik kuralları analizi için Apriori algoritması tercih edilmiştir.

2.1. Birliktelik Kuralları Analizi

Birliktelik Kuralları Analizinin ayrıntılı bir şekilde incelenmesine geçmeden önce bazı tanımlamaları yapmak gereklidir. D bir süpermarkete ait satış işlemlerinin veri tabanı olsun ve bu veri tabanındaki her bir müşteriye ait satış işlemleri (satış fişi) T olsun. $I = \{I_1, I_2, \dots, I_m\}$ bir ürün kümesi ise, $T \subseteq I$ olur.

Bir birliktelik kuralı $A \Rightarrow B$ şeklinde gösterilmektedir. Kuralın solunda yer alan nesneye/ürüne sol taraf, sağındakine sağ taraf denir. Bu ifade basitçe $Sebze \Rightarrow Meyve$ olarak düşünüldüğünde, ifadenin sebze alan müşterilerin meyve de aldığı anlamı taşıdığı açıktır. Kuralın sol tarafında birden fazla ürün yer alabilir. $A \Rightarrow B$ kuralında, $A \subset I, B \subset I$ ve $A \cap B = \emptyset$ ’dir.

Bir kuralın minimum destek ve minimum güven değerini sağlaması zorunludur. Destek ve güven değerini kısaca özetlemek gerekirse, öncelikle her bir kuralın destek, güven ve kaldıraç isimli, üç değeri mevcut olduğunu belirtmek gerekir. Bunlardan destek değeri (s), D veri tabanındaki satış fişleri arasında yer alan $A \cup B$ fişlerinin oranıdır bu da hem

A ’yı hem B ’yi içeren fişlerin, bütün fişlere oranına karşılık gelmektedir. Dolayısıyla destek değeri, ilgili ürünler arasındaki bağıntının ne kadar sık olduğunu gösterir.

Diğer bir değer olan güven (c), A ’yı içeren fişlerin ne kadarının B ’yi de içerdiğini gösterir. Bu değer bir koşullu olasılık değeridir. Güven değeri kuralın kuvvetini belirtir. Bu iki değer dışında daha önce belirtildiği gibi bir de kaldıraç değeri hesaplanmaktadır.

$$\text{destek}(A \Rightarrow B) = P(A \cup B)$$

$$\text{güven}(A \Rightarrow B) = P(B|A)$$

$$\text{kaldıraç}(A \Rightarrow B) = \frac{\text{güven}(A \Rightarrow B)}{\text{destek}(B)} = \frac{\text{destek}(A \Rightarrow B)}{\text{destek}(A) \cdot \text{destek}(B)}$$

Destek, güven ve kaldıraç değerleri, bir kuralın kullanılabilirliği (gücü) ve kesinliği açısından o kuralın kalitesini ölçen değerlerdir. Daha önce belirtildiği gibi destek değeri, bir veri kümesinde, bir kuralın sol ve sağ tarafında yer alan ürünlerin birlikte kaç fişte (işlemden) yer aldığını tespit ederken güven değeri ise, kuralın sol tarafındaki ürünü içeren fişlerin kaçının aynı zamanda sağ taraftaki ürünü de içerdiğini belirtir (Cios, Pedrycz, ve Swiniarski 2007, s.290). Güven değeri bir kuralın, sağ tarafında hangi öğenin yer alacağını öngörmeye ne kadar iyi olduğunun derecesini de ölçer. Fakat sağ taraftaki öğeler çok bilindik öğeler ise kural bizim için ilgi çekici olmayabilir. Bu sebeple kuralın ilginç olup olmadığının ölçüsü olan kaldıraç değeri farklı olarak kuralın tamamını rassal olarak seçilmiş olan sağ taraf öğeleri ile karşılaştırır. Dolayısıyla, bir kuralın güven derecesi ile birlikte, kaldıraç derecesinin de göz önünde bulundurulması gerekmektedir.

Kaldıraç değerinin 1'den büyük olması kuraldaki ürünler arasındaki birlikteliğin pozitif olmasını, yani A ve B ürünlerinin beklenenden daha fazla birlikte görüldüğünü ve kuralın ilgi çekici olduğunu, 1'den küçük olması ise aralarında negatif korelasyonun olmasını ifade eder. Dolayısıyla kaldıraç değeri 1'den düşük kurallar dikkate alınmaz. Değerin tam 1'e eşit olması durumunda hiçbir korelasyonun mevcut olmadığına, bağımsızlığa işaret eder. Kaldıraç değeri ayrıca kuralın şans olarak mı ortaya çıktığını veya aksine, gerçekten beklenen ve iyi bir kural mı olduğunu belirlemektedir. Kaldıraç ismini almasının sebebi ise şu şekilde açıklanabilir: Örneğin iki ürünlü bir kuralda kaldıraç değeri 1'den büyükse bir ürünün satışı diğer ürünün satışını yükseltebilmektedir. Son olarak bu üç değer hakkında söylenmesi gereken şey bu değerlerin sınır değerlerinin analistler veya alanındaki uzman kişilerin belirlemesi gerektiğidir.

Genel olarak, birliktelik kuralı analizi iki aşamadan oluşmaktadır (Aggarwal, 2015, s.98; Han vd., 2011, s.231).

1. Kullanıcı tarafından daha önce belirlenen bir destek düzeyinin (minimum destek değeri) üzerindeki bütün sık ürün kümeleri tespit edilir.
2. Sık ürün kümelerden güçlü birliktelik kuralları oluşturulur; kurallar kullanıcı tarafından belirlenen minimum bir destek ve güven değerinin üzerinde olmalıdır.

2.1.1 Apriori Algoritması

Birliktelik kuralları araştırmacılar tarafından, Apriori (Agrawal ve Srikant, 1994), FP-Growth Algorithm (Han, Pei ve Yin, 2000) ve Eclat (Zaki, Parthasarathy ve Li, 1997) gibi algoritmalar kullanılarak bulunmaktadır. Bu çalışmada da kullanılacak olan Apriori algoritması en çok kullanılan algoritma olarak karşımıza çıkmaktadır.

Agrawal'ın 1993 yılında başladığı çalışmalar 1994 yılında Srikant ile birlikte yayınladığı Apriori algoritması ile meyvesini vermiştir. Algoritma, k adet ürün kümesinin $k+1$ adet ürün kümesi için kullanıldığı, en basit kuraldan başlayıp bu kurala tek tek ürünler ekleyerek ilerleyen aşamalı bir tekniktir. Bu algoritma, verilen bir ürün kümesindeki önceden belirlenmiş bir destek değerinin üzerinde değer alan alt ürün kümesini seçerek (ilk etapta sık tekli ürünler seçilir) ve o destek değerinin altındaki diğer ürünleri ise göz ardı ederek işe başlar. İlk adımdan geçen ürünler, iki öğeli ürün kümelerini oluştururlar. Bu ürün kümelerinin hesaplanan destek değerleri de ilk başta belirlenen destek değeriyle karşılaştırılır, altında kalan ürün kümeleri yine göz ardı edilirler. Fakat göz ardı edilen bu ürün kümeleri ileride iki ürünlü kurallar için (kuralların sağ tarafları için) birer adaydırlar. Bu süreç, belirlenen destek değerine ulaşıncaya kadar sık ürünleri birleştirir, en sonunda ise artık hiçbir sık k -ürün küme bulunamaz hale gelir. Sık ürün kümelerinin tespitinden sonra ise artık kural bulma işlemine geçilir. Daha önceden belirlenmiş olan minimum bir destek değeri ve ayrıca bir güven değerinin üzerindeki birliktelik kuralları bulunur (Aggarwal, 2015, s.100; Giudici, 2005, s.93).

3. Uygulama

Bu bölümde Muş il merkezinde faaliyet gösteren yerel bir süpermarket işletmecisine ait veriler bu çalışmanın uygulama bölümüne konu olacaktır. Firmanın, gizlilik ilkeleri sebebiyle sağladığı sadece 634 müşteriye ait alışveriş işlemi analizlerde

kullanılacaktır. Her ne kadar müşteri sayısı az olsa da firmanın kıymetli müşterileri olan bu müşteriler 47 ana ürün üzerinden toplamda 16382 işlem gerçekleştirmişlerdir. Bu rakamlar sepet ortalamasının yaklaşık 25 ürün olduğunu göstermektedir. Analizlere başlamadan önce bir müşterinin aynı ürünü birden fazla alımlarını R programındaki ilgili paketin değerlendirmeye almadığını, eş (duplicate) alımları sadece tek bir alım gibi değerlendirdiğini hatırlatmakta fayda vardır. Örneğin ham veri setinde "manav" ürünleri 16382 işlemde 4424 defa yer almışken, manav reyonundan birden fazla ürün alan müşterilerin birden fazla manav kayıtları tek bir kayda düşürüldüğünde, manav ürünlerinin sayısının 465'e düştüğü, böylece 634 müşterinin yarısından fazlasının (%73'ü) manav reyonundan alışveriş yaptığı tespit edilmiştir. Aşağıdaki tablolarda, sırasıyla firmanın sağladığı ham veri setinden bir kesit (Tablo 1) ve 46ana kategoride manuel olarak kodlanmış ürünler (Tablo 2) yer almaktadır. Bu kodlamanın daha önceki yapılan birçok çalışmadan farklılık gösterdiğini söylemek mümkündür. Çünkü geçmiş çalışmalarda araştırmacılar ürünleri çok genel kategorilere (süt ürünleri, gıda, gıda dışı gibi) ayırmakta iken bu çalışmada daha detaylı bir gruplandırma, kodlama tercih edilmiştir. Çok genel kategorilerde kodlama sonucu elde edilecek birliktelik kuralları da çok genel olacaktır. Bu sebeple daha detaylı bir kodlama tercih edilmiştir ve ortaya çıkan ilişkiler daha anlamlı olacaktır. Çok daha büyük veri setleriyle analiz yapabilecek olan firmalar daha detaylı (örneğin manav reyonundaki ürünleri domates, elma, salatalık... gibi kodlamak) kodlama tercih edebilirler. Bunun sonucunda elde edilecek kurallar da çok daha detaylı olacak ve müşteri satın alma davranışlarının tespitinde fazlasıyla yol gösterici olacaktır.

Tablo 1. Firma Tarafından Sağlanan Ham Veriden Bir Kesit

Ürün	Fiş No	Barkod
005980 ÖNCÜ SALÇA DOMATES 830 GR TNK	1	8693891195843
005980 ÖNCÜ SALÇA DOMATES 830 GR TNK	1	8693891195843
005974 ÖNCÜ SALÇA BİBER TATLI 700 GR	1	8693891195942
005524 DURU NAT.SABUN 640 GR NAT.OLIVE ZEYTİN YAPRAKLI	1	8690506474195
019682 ÜRT-EKMEK SOMUN	2	2900631
019165 M-LİMON YATAK	3	2800428
019169 M-ÜZÜM BEYAZ	5	2800432
008709 KNORR ÇORBA KREMALİ TAVUK CORBASI	6	8690637018626
000024 COLGATE 100 ML ÜÇLÜ ETKİ	6	7891024132173
008702 KNORR ÇORBA EZOGELİN 65 GR	6	8690701053072

Tablo 2. 46 Ana Ürün Grubu

Manav	Kozmetik	Zeytin	Su	Baharat	Kahve
Peynir	Yumurta	Tavuk	Tuzlu bisk.	Soğuk içecek.	Çikolata
Sos	Deterjan	Sucuk	Çerez	Tatlı bis.	Tuz
Şeker	Sigara	Peynir	Yoğurt	Helva	Konserve
Fırın	Ev alet	Reçel	Margarin	Gevrek	Balık
Diş	Dondurma	Süt	Et	Kâğıt ü.	Elektronik
Çay	Şekerleme	Giyim	Meyve s.	Donmuş gıda	
Yağ	Elektrikli ev alet	Bebek	Krema	Oyuncak	

3.1. En çok Satılan Ürünlerin Tespiti

Bu bölümde en çok satılan ürünler ve ürün gruplarının tespiti yapılacaktır. R programlama dilindeki *Arules* paketinde (Hahsler, Gruen, ve Hornik, 2005) bulunan *apriori* işlevi bu görevi yapabilmektedir. Aşağıdaki kod ile “Fisno” ve “urun” isimli sütunlara sahip veri seti R ortamına aktarılmış, her bir fişteki ürünler tanımlanmış ve veri setias işlevi ile matris şekline (satırlarda fiş numaraları sütunlarda ürünler ve matris değerleri 0-1 olacak şekilde)dönüştürülmüştür. *apriori* işlevi ile destek değeri %10’dan fazla olan (müşterilerin %10’undan fazlası)en çok satan ürünler tespit edilmiştir. Son olarak en çok satan ilk 20 ürün/ürün grubu ekrana yazdırılmış ve Tablo 3’de gösterilmiştir.

```
pera <- read.csv("hazir.csv",header=TRUE,sep=";")
aggpera <- split(pera$urun,pera$Fisno)
perahazir <- as(aggpera,"transactions")
peracok <-
apriori(perahazir,parameter=list(support=0.1,target="frequent"))
top20 <- sort(peracok)[1:20]
inspect(top20)
```

Tablo 3.En Çok Satılan Ürün veya Ürün Grupları

Ürün	Destek	Sayı
{manav}	0.733439	465
{yumurta}	0.705047	447
{manav,yumurta}	0.55836	354
{kozmetik}	0.430599	273
{fırın}	0.422713	268
{deterjan}	0.399054	253
{fırın,manav}	0.353312	224
{kuru gıda}	0.337539	214
{tavuk}	0.335962	213
{kozmetik,manav}	0.332808	211
{fırın,yumurta}	0.33123	210
{deterjan,manav}	0.33123	210
{şekerleme}	0.328076	208
{ev aletleri}	0.323344	205
{deterjan,yumurta}	0.321767	204
{kozmetik,yumurta}	0.320189	203
{tavuk,yumurta}	0.294953	187
{manav,tavuk}	0.290221	184
{fırın,manav,yumurta}	0.285489	181
{kuru gıda,yumurta}	0.280757	178

Müşteriler tarafından en çok satın alınan ürün veya ürün grupları incelendiğinde (destek değeri en yüksek ürünlere göre sıralama yapıldığından bazı ikili, hatta üçlü alımların destek değeri de yüksek olabilmektedir) Tablo 3 elde edilmiştir. Eğer en çok satılan tek ürünler incelenmek istenirse aşağıda gösterilen kod yazılmalıdır. Bu kod sonucunda elde edilen ürünlerin 10 tanesi Tablo 4’de gösterilmiştir.

```
tekli <- apriori(perahazir,parameter=list(support=0.1,
target="frequent",maxlen=1))
> inspect(sort(tekli))
```

Tablo 4. En Çok Satılan (Tekli) Ürün Listesi

Ürün	Destek	Fiş sayısı
{manav}	0.733439	465
{yumurta}	0.705047	447
{kozmetik}	0.430599	273
{fırın}	0.422713	268
{deterjan}	0.399054	253
{kuru gıda}	0.337539	214
{tavuk}	0.335962	213
{şekerleme}	0.328076	208
{ev aletleri}	0.323344	205
{peynir}	0.27918	177

3.2. Kuralların Tespiti

BKA hakkında Bölüm 2’de detaylı bilgi verilirken vurgulandığı gibi, sık satılan ürün veya ürün grupları birliktelik kurallarını oluştururlar. Herhangi bir kısıt koymaksızın (sol tarafta kaç ürün olduğunu belirtmeksizin)*apriori* komutu ile (*target* komutunu kullanmadan)destek değeri %10 ve güven değeri %50 tutulduğunda toplam 689 kural ile karşılaşmaktadır. Daha fazla müşteri ile yapılacak analizlerde kural sayısının binlerle, hatta on binlerle ifade edilmesi doğaldır.

689 kuralı tek tek incelemek yerine *inspect* ve *sort* komutları ile istenilen değere göre (destek, güven veya kaldırıcı) kurallar sıralanabilmektedir. Aşağıdaki komutlar ile kaldırıcı değeri en yüksek olan 20 kural Tablo 5’de verilmiştir. Bunu yapmanın ve daha fazla bilgi etmenin başka bir yolu da vardır. İlerleyen bölümde yeni kodlarla bu yol gösterilecektir.

Daha önce vurgulandığı gibi, kaldırıcı değerleri en yüksek kurallar, en ilgi çekici kurallardır. Görüldüğü üzere kaldırıcı değeri en yüksek olan kurala göre (sol tarafı iki ürünlü kural); ev aletleri reyonundan ve fırından ürün alan müşteriler peynir satın almışlardır. Firmaların uzman pazarlama ekipleri bu gibi kuralları farklı pazarlama faaliyetleri tasarlamak için kullanabilmektedir (Baralis vd., 2011; Aguinis, Forcum ve Joo, 2013).

```
perarules <-
apriori(perahazir,parameter=list(supp=0.1,conf=0.50))
inspect(sort(perarules,by="lift",20))
```

Tablo 5’in dördüncü kuralı olan, {reçel}=>{peynir} kuralını açıklamak gerekirse; %11’lik destek değeri: Reçel ile peynir alan müşterilerin oranının tüm müşteriler (analizde 634 müşteri mevcuttur) arasında %11 (yaklaşık 70 müşteri) olduğunu ifade etmektedir. %56’lık güven değeri ise: Reçel alan müşterilerin % 56’sının peynir de aldığını belirtmektedir. Ayrıca 2,02’lik kaldırıcı değeri de şu şekilde açıklanabilir: Eğer bir müşterinin peynir aldığı biliniyorsa, o

müşterinin reçel de alacağı beklentisinin (expectation) % 202 arttığını gösterir.

Tablo 5. Kaldıraç Değeri En Yüksek İlk On Kural

Kural	destek	güven	kaldıraç
{ev aletleri,fırın} => {peynir}	0.11	0.60	2.14
{deterjan,fırın,manav,yumurta}=>{peynir}	0.10	0.58	2.09
{fırın,kuru gıda}=> {peynir}	0.10	0.57	2.03
{reçel}=> {peynir}	0.11	0.56	2.02
{deterjan,evaletleri,yumurta}=>{kurugıda}	0.11	0.68	2.01
{deterjan,fırın,yumurta} => {peynir}	0.11	0.56	1.99
{deterjan,fırın,manav} => {peynir}	0.11	0.55	1.96
{süt,yumurta} => {peynir}	0.10	0.55	1.96
{deterjan,peynir} => {kuru gıda}	0.11	0.66	1.96
{deterjan,ev aletleri,manav}=> kurugıda}	0.11	0.66	1.96
{fırın,kozmetik,manav}=> {peynir}	0.11	0.54	1.95
{deterjan,ev aletleri} => {kuru gıda}	0.12	0.66	1.95
{ev aletleri,kurugıda,manav}=>{deterjan}	0.11	0.77	1.94
{sos,yumurta} => {kuru gıda}	0.10	0.65	1.93
{evaletleri,kozmetik,yumurta}=>{deterjan}	0.11	0.77	1.92
{fırın,kuru gıda,yumurta}=> {tavuk}	0.10	0.64	1.92
{evaletleri,kurugıda,yumurta}=>{deterjan}	0.11	0.76	1.91
{deterjan,fırın} => {peynir}	0.12	0.53	1.90
{ev aletleri,kozmetik} => {kuru gıda}	0.12	0.64	1.90
{yağ} => {kuru gıda}	0.10	0.64	1.89

Aynı kurallar, destek ve güven değerine göre sıralanmak istenirse aşağıdaki kod yazılır ve Tablo 6 elde edilir.

```
perarules <-
apriori(perahazir,parameter=list(supp=0.1,conf=0.50))
inspect(sort(perarules, by=c("supp", "conf")))
```

Tablo 6. Destek Ve Güven Değeri En Yüksek İlk 15 Kural

Kural	destek	güven	kaldıraç
{yumurta}=> {manav}	0.56	0.79	1.08
{fırın} => {manav}	0.35	0.84	1.14
{kozmetik} => {manav}	0.33	0.77	1.05
{deterjan} => {manav}	0.33	0.83	1.13
{fırın} => {yumurta}	0.33	0.78	1.11
{deterjan} => {yumurta}	0.32	0.81	1.14
{kozmetik} => {yumurta}	0.32	0.74	1.05
{tavuk} => {yumurta}	0.29	0.88	1.25
{tavuk} => {manav}	0.29	0.86	1.18
{fırın,yumurta} => {manav}	0.29	0.86	1.18
{kuru gıda}=> {yumurta}	0.28	0.83	1.18
{deterjan,yumurta}=>{manav}	0.28	0.87	1.18
{kuru gıda} => {manav}	0.27	0.80	1.10
{ev aletleri} => {manav}	0.27	0.82	1.12
{şekerleme} => {manav}	0.26	0.80	1.09

Tablo 6'dan izlendiği gibi bazı kurallarda üç hatta dört ürün yer alması söz konusu olabilmektedir. Eğer araştırmacı veya firma yöneticileri sadece iki ürünlü kuralları incelemek istiyorsa aşağıdaki komutları yazmalıdır. Bu komutlar sonucunda elde edilen kuralların destek ve güven düzeyine göre en yüksek olan ilk 20'si Tablo 7'de verilmiştir.

```
perarules <-
apriori(perahazir,parameter=list(supp=0.1,conf=0.50,minlen=2,maxlen=2))
inspect(sort(perarules, by=c("supp", "conf")))
```

Tablo 7. Sadece İki Ürün İçeren Kurallar

Kural	destek	güven	kaldıraç
{yumurta} => {manav}	0.56	0.79	1.08
{fırın} => {manav}	0.35	0.84	1.14
{kozmetik} => {manav}	0.33	0.77	1.05
{deterjan} => {manav}	0.33	0.83	1.13
{fırın} => {yumurta}	0.33	0.78	1.11
{deterjan} => {yumurta}	0.32	0.81	1.14
{kozmetik} => {yumurta}	0.32	0.74	1.05
{tavuk} => {yumurta}	0.29	0.88	1.25
{tavuk} => {manav}	0.29	0.86	1.18
{kuru gıda} => {yumurta}	0.28	0.83	1.18
{kuru gıda} => {manav}	0.27	0.80	1.10
{ev aletleri} => {manav}	0.27	0.82	1.12
{şekerleme} => {manav}	0.26	0.80	1.09
{ev aletleri} => {yumurta}	0.25	0.78	1.10
{şekerleme}=> {yumurta}	0.25	0.75	1.06
{deterjan} => {kozmetik}	0.24	0.61	1.42
{peynir} => {yumurta}	0.24	0.84	1.19
{peynir} => {manav}	0.23	0.82	1.12
{deterjan} => {fırın}	0.23	0.57	1.34
{çay} => {manav}	0.22	0.84	1.15

Tablo 7'den izlendiği gibi manav, yumurta ve fırın ürünleri ile kurallarda sıkça karşılaşılmaması rastlantı değildir. Tablo 3'de bu ürünlerin firmanın en çok satan ürün grupları olduğu tespit edilmişti. En güçlü kural olarak karşımıza yumurta satın alan müşterilerin manav reyonundan da alışveriş yaptığı kuralı çıkmıştır. Firma bu tespitler ışığında daha önce de vurgulandığı gibi pazarlama stratejileri geliştirebilir.

Araştırmacılar veya firma yöneticileri kurallarda sıkça karşılaşılan bazı ürünleri analiz dışında tutmak isteyebilirler. Örnek olarak yumurta, manav ve fırın ürün gruplarının dışarıda tutulması ile geri kalan kuralları elde etmek için aşağıdaki komutlar kullanılmış, sonuçları Tablo 8'de gösterilmiştir.

```
perarules <- apriori(perahazir,
parameter=list(supp=0.1,conf=0.50),appearance=list(none=
c("yumurta","manav","fırın"),default="both"))
```

Tablo 8. Yumurta, Manav ve Fırın Dışındaki Kurallar

Kural	destek	güven	kaldıraç
{deterjan} => {kozmetik}	0.24	0.61	1.42
{kuru gıda}=> {deterjan}	0.21	0.62	1.56
{kuru gıda} => {kozmetik}	0.20	0.58	1.35
{ev aletleri} => {deterjan}	0.18	0.57	1.43
{şekerleme} => {kozmetik}	0.18	0.55	1.28
{tavuk} => {deterjan}	0.18	0.54	1.35
{ev aletleri} => {kozmetik}	0.18	0.56	1.29
{şekerleme} => {deterjan}	0.17	0.52	1.30
{tavuk} => {kozmetik}	0.17	0.51	1.18
{tavuk}=> {kuru gıda}	0.17	0.50	1.49
{çay} => {kozmetik}	0.17	0.62	1.44
{peynir} => {kozmetik}	0.16	0.58	1.35
{peynir} => {deterjan}	0.16	0.58	1.46
{ev aletleri} => {kuru gıda}	0.16	0.50	1.49
{çay} => {deterjan}	0.16	0.60	1.50
{peynir} => {kuru gıda}	0.15	0.53	1.56
{çerez} => {deterjan}	0.14	0.62	1.54
{peynir} => {ev aletleri}	0.14	0.51	1.57
{kozmetik,kurugıda}=>{deterjan}	0.14	0.70	1.76
{tatlıbisküvi}=>{kozmetik}	0.14	0.57	1.32

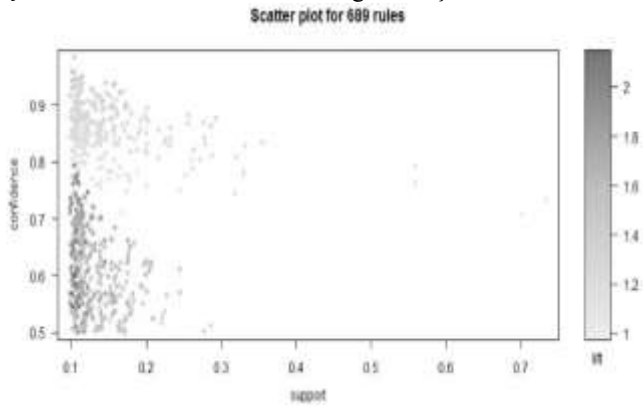
R programlama dili ile elde edilen kuralların görselleştirilmesi de söz konusudur. Örneğin *arulesViz* paketi (Hahsler ve Chelluboina, 2012) ile Kuralların Tespiti (Bölüm 3.2) bölümünde elde edilen 689 kural aşağıdaki gösterilen komutlar ile Şekil 1’de görselleştirilmiştir.

```
library(arulesViz)
plot(perarules)
```

Bu şekil kuralların destek, güven ve kaldıraç değerlerini göstermekte, kurallara ait destek ve güven değerlerinin dağılımları hakkında da bir fikir sunmaktadır. Görüldüğü gibi kurallar arasında güvenilirliği yüksek olanların destek değerleri yüzde 10 seviyelerindedir. Bayardo, Jr. ve Agrawal’a (1999) göre en ilginç kurallar, şekilde kalın noktalarla gösterilen hat boyunda yer alan kurallardır. Analizcinin amacına göre aşağıdaki komutla şekildeki destek, güven ve kaldıraç yerleri değiştirilebilir. Komut kaldıraç değerinin yerini güven değerinin alması için oluşturulmuştur. Daha ayrıntılı görselleştirmeler için Hahsler ve Chelluboina’ya (2011) başvurulabilir.

```
plot(perarules, measure = c("support", "lift"), shading = "confidence")
```

Şekil 1. Elde edilen 689 kuralın görselleştirilmesi



Daha önce *inspect* ve *sort* komutları ile incelediğimiz kuralların daha kolay ve güzel bir yöntemle de araştırılabileceği vurgulanmıştır. *Arulesviz* paketi, Xie (2016) tarafından oluşturulmuş *DT* paketini kullanarak, kuralları tıpkı bir Excel sayfasında filtreleme yapar gibi, interaktif şekilde araştırılabilmektedir. Bu görevi gerçekleştirebilmesi için programa aşağıdaki komut girilir ve Şekil 2 elde edilir.

```
> inspectDT(perarules)
```

Şekil 2. Elde edilen kuralların tablolaştırılması

id	LHS	RHS	support	confidence	lift
[1]	()	{yassimurta}	0.707	0.707	1.000
[2]	()	{deterjan}	0.733	0.733	1.000
[3]	{sepet}	{yassimurta}	0.323	0.729	1.001
[4]	{sepet}	{deterjan}	0.329	0.716	0.999
[5]	{sepet}	{yassimurta}	0.397	0.610	1.140

Firma hangi ürünlerle hangi ürünlerin birlikte satıldığını öğrenmek istediğinde (örneğin deterjan ile hangi ürünlerin birlikte satıldığını) yapması gereken basitçe LHS (sol taraf kuralı) kısmına deterjan yazmaktır. Elde edilen tablo aşağıda Şekil 3’de gösterilmiştir. Ayrıca bu interaktif tablo ile de kurallar destek, güven veya kaldıraç değerlerine göre sıralanabilmekte ve en yüksek kaldıraç değerli

kural (müşteriler ufak ev aletleri reyonu ile fırın reyonundan aldıkları ürünlerle birlikte herhangi bir peynir ürününü satın aldığı) tespit edilebilmektedir.

Şekil 3. Deterjan ürünü seçildiğinde elde edilen kurallar

id	LHS	RHS	support	confidence	lift
[102]	{deterjan}	{fırın gale}	0.239	0.726	1.557
[108]	{deterjan}	{deterjan}	0.244	0.651	1.422
[112]	{deterjan}	{sepet}	0.226	0.765	1.537
[117]	{deterjan}	{yassimurta}	0.222	0.696	1.514
[130]	{deterjan}	{deterjan}	0.216	0.610	1.231
[131]	{deterjan,yad}	{yassimurta}	0.099	0.689	1.281
[141]	{deterjan,yad}	{yassimurta}	0.096	0.610	1.201

Pakette kuralların ilişki ağını gösteren grafik de mevcuttur. Grafik temelli teknikler her bir ürünün yer aldıkları kurallardaki diğer ürünlerle olan ilişkilerini göstermektedir.

```
plot(perarules, method = "graph", engine = "htmlwidget")
```

Yukarıdaki komut ile elde edilen kurallar ağının ilk görüntüsü Şekil 4’te verilmiştir. Daha sonra kurallar ağının merkezine doğru imleç yardımıyla yaklaşım sağlanmıştır (Şekil 5). Görüldüğü gibi satış frekansları en yüksek olan ürünler (deterjan, manav, yumurta, fırın...) grafikte göz önündedir.

Şekil 4. 689 kuralın ilişki ağı

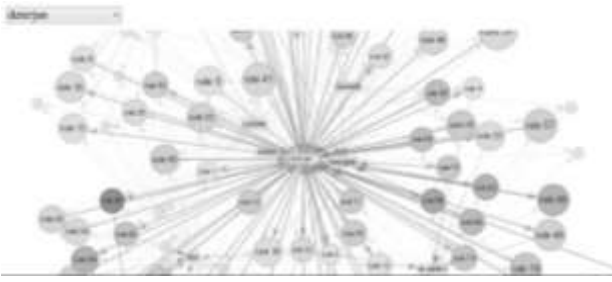


Şekil 5. İlişki ağına imleç yardımıyla yaklaşma (zoom)



Aşağıdaki son şekilde (Şekil 6) ise, Select by ID kısmında deterjan ürünü seçildiğinde karşılaşılan ağ verilmiştir.

Şekil 6. Deterjan ürününün yer aldığı kuralların ilişki ağı



4. Sonuç ve Öneriler

Bu çalışmada Muş il merkezinde faaliyet gösteren yerel bir süpermarket işletmecisinin 634 müşterisine ait alışveriş işlemleri Pazar Sepet Analizi ile incelenmiştir. Bu amaçla kullanılan Apriori Algoritması ile en çok satılan ürünlerin manav reyonundaki ürünler olduğu tespit edilmiştir. Gerçekten firmanın manav ve fırın reyonları şehirde kalitesi ile ön plana çıkmaktadır. Fırın ürünlerinin satışı ise 4. sıradadır. Ayrıca yine Apriori ile elde edilen kurallara göre manav ve fırın ürünleri en çok yumurta, kozmetik ürünleri, tavuk ve kuru gıda ürünleri ile birlikte satılmıştır.

Büyük marketler için yüzlerle ifade edilen müşterilere ait analizler yapmak elbette ki genel çıkarımlar için yol gösterici olmamaktadır. Bu çalışmada kullanılan yöntem ve analizlere ait R programlama dili kodları büyük veri setleri ile de çalışmaya uygundur. Bu doğrultuda, analizlere konu olan yerel perakendeci ile sonuçlar paylaşılacak, müşterilerin daha detaylı satın alma davranışları ve ayrıca satın alma bazlı müşteri segmentasyonu yapabilmek amacıyla, gelecekteki çalışmalar için birkaç bin müşteri veya birkaç haftalık veri seti paylaşımı için firmadan talepte bulunulacaktır.

Çok daha büyük veri setleriyle analiz yapabilecek olan bu firma ve başka perakende firmaları daha detaylı (örneğin manav reyonundaki ürünleri domates, elma, salatalık... gibi kodlamak) kodlamayı da tercih edebilirler. Bu sayede müşterilerin daha detaylı ve belirgin satın alma davranışları tespit edilebilecektir. Bu çalışmada geçmiş çalışmalardan farklı olarak detaylı ürün gruplandırması kullanılmıştır. Son olarak çalışma R programlama dilinin ilgili alanda kullanılabilirliğini gösterebilmek açısından kullanılabilir kodları detaylı bir şekilde incelemiş ve paylaşmıştır.

Kaynakça

Aggarwal, C. C. (2015). *Data Mining: The Textbook*. Springer.

Agrawal, R., & Srikant, R. (1994). Fast Algorithms For Mining Association Rules. *In Proc. 20th int. conf. very large data bases, VLDB* (Vol. 1215, 487-499).

Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using market basket analysis in management research. *Journal of Management*, 39(7), 1799-1824.

Akgün, A., & Çizel, B. (2016). Günlük Tur Programları Oluşturmada Veri Madenciliği: A Grubu Seyahat Acentası Örneği. *Turar Turizm & Araştırma Dergisi*, 6(1), 73-87.

Baralis, E., Cagliero, L., Cerquitelli, T., Garza, P., & Marchetti, M. (2011). CAS-Mine: providing personalized services in context-aware applications by means of

generalized rules. *Knowledge and information systems*, 28(2), 283-310.

Bayardo Jr, R. J., & Agrawal, R. (1999, August). Mining the most interesting rules. *In Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 145-154). ACM.

Berry, M. J. & G. Linoff. (1997). *Data Mining Techniques: For Marketing, Sales, And Customer Support*: John Wiley & Sons, Inc.

Bilgiç, E., & Çakır, Ö. (2018). Comparing Clusterings: A Store Segmentation Application. *Erusosbilder*, 32(44), 41-57

Budak, V. Ö., Kartal, E., & Gülseçen, S. (2018). Site-içi Aramalar ve Apriori Algoritması Kullanılarak Web Sitesi Ziyaretçilerinin İhtiyaç Tespitine Yönelik Bir Örnek Olay İncelemesi. *Bilişim Teknolojileri Dergisi*, 11(2), 211-222.

Chiang, W. Y. (2018). Applying Data Mining For Online Crm Marketing Strategy: An Empirical Case Of Coffee Shop İndustry in Taiwan. *British Food Journal*, 120(3), 665-675.

Cios, K. J., Pedrycz, W., Swiniarski, R. W., & Kurgan, L. A. (2007). *Data Mining: A Knowledge Discovery Approach*. Springer Science & Business Media.

Ertugrul, I., Oztas, T., Oztas, G. Z., & Ozcil, A. (2016). Shelf Layout With Integrating Data Mining And Multi-Dimensional Scaling. *European Scientific Journal*, ESJ, 12(10).

Guo, Y., Wang, M., & Li, X. (2017). Application Of An Improved Apriori Algorithm in a Mobile e-Commerce Recommendation System. *Industrial Management & Data Systems*, 117(2), 287-303.

Gulluoglu, S. S. (2015). Segmenting customers with data mining techniques. *Digital Information, Networking, and Wireless Communications (DINWC)*, 2015, 154-159.

Giudici, P. (2005). *Applied Data Mining: Statistical Methods For Business And Industry*. John Wiley & Sons.

Hahsler, M., & Chelluboina, S. (2011). Visualizing Association Rules: Introduction to the R-extension Package arulesViz. R project module, 223-238.

Hahsler, M., & Chelluboina, S. (2012). arulesViz: Visualizing Association Rules and Frequent Itemsets. R package version 0.1-5.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts And Techniques*. Elsevier.

Han, J., Pei, J., & Yin, Y. (2000). Mining Frequent Patterns Without Candidate Generation. *In ACM sigmod record* (Vol. 29, No. 2, pp. 1-12). ACM.

Huseyinov, I., & Aytaç, U. C. (2017). Identification of Association Rules İn Buying Patterns Of Customers Based on Modified Apriori and Eclat Algorithms by Using R Programming Language. *In Computer Science and Engineering (UBMK)*, 2017.

Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify The Changing Trends Of Market Data Using

- Association Rule Mining. *Procedia Computer Science*, 85, 78-85.
- Kokoç, M., Aktepe, A., Ersöz, S., & Türker, A. K. (2016). Improvement of Facility Layout By Using Data Mining Algorithms And An Application. *International Journal of Intelligent Systems and Applications in Engineering*, 4(Special Issue-1), 92-100.
- Lee, D., Park, S. H., & Moon, S. (2013). Utility-Based Association Rule Mining: A Marketing Solution For Cross-Selling. *Expert Systems with applications*, 40(7), 2715-2725.
- Leskovec, J., A. Rajaraman & J. D. Ullman. (2014). *Mining Of Massive Datasets*: Cambridge University Press.
- Michael Hahsler, Bettina Gruen and Kurt Hornik (2005), arules - A Computational Environment for Mining Association Rules and Frequent Item Sets. *Journal of Statistical Software* 14/15.
- Nengsih, W. (2015). A Comparative Study On Market Basket Analysis And Apriori Association Technique. *In Information and Communication Technology (ICICT)461-464*, IEEE.
- Özcan, T. & Ş. Esnaf. (2010). Perakende Endüstrisinde Raf Alanı Tahsis ve Mağaza Yerleşim Optimizasyonuna Bütünleşik Bir Model Önerisi. *İÜ Mühendislik Bilimleri Dergisi*. 1.1, 55-63.
- Özçalıcı, M. (2017). Predicting Second-Hand Car Sales Price Using Decision Trees and Genetic Algorithms. *Alphanumeric Journal*, 5(1), 103-114.
- Öztürk, G., & Tanrısevdi, A. (2017) Uluslararası Kruvaziyer Ziyaretçilerine Ait Özelliklerin Birliktelik Kuralı Modeli İle Analizi. *Uluslararası İktisadi ve İdari Bilimler Dergisi*, 3(1), 131-148.
- Pehlivanoglu, M. K., & Nevcihan, D. (2015). Veri Madenciliği Teknikleri Kullanılarak Ortaokul Öğrencilerinin Sosyal Ağ Kullanım Analizi: Kocaeli İli Örneği. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 3(2).
- R Core Team (2018). R: A Language And Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Sagin, A. N., & Ayvaz, B. (2018). Determination of Association Rules with Market Basket Analysis: Application in the Retail Sector. *Southeast Europe Journal of Soft Computing*, 7(1).
- Sözen, E., Bardak, T., Peker, H., & Bardak, S.(2017).Apriori Algoritması Kullanılarak Mobilya Seçimde Etkili Olan Faktörlerin Analizi. *İleri Teknoloji Bilimleri Dergisi*, 6(3), 679-684.
- Tekin, M., Y. Köse, Ö. Koyuncuoğlu & E. Tekin. (2015). The Analysis of Product Categories and Sales Relationships among Valuable Customers through Data Mining and Its Application to a National Retailer through Association Rules and Cluster Analysis. *International Interdisciplinary Business-Economics Advancement Conference*, 180.
- Yıldırım, P., & Birant, D. (2018). Bulut Bilişimde Veri Madenciliği Tekniklerinin Uygulanması: Bir literatür taraması. *Pamukkale University Journal of Engineering Sciences*, 24(2).
- Zaki, M. J., Parthasarathy, S., Ogihara, M., & Li, W. (1997, August). New Algorithms For Fast Discovery Of Association Rules. *In KDD (Vol. 97, pp. 283-286)*.
- Xie, Y., Cheng, J., Allaire, J. J., Reavis, B., Gersen, L., & Szopka, B. (2015). DT: a wrapper of the JavaScript library 'DataTables. R package version 0.1.' Available at <http://CRAN.R-project.org/package=DT> [Verified 1 March 2016].

