

**ARAŞTIRMA MAKALESİ /RESEARCH ARTICLE**

**EGE BÖLGESİ'NDEKİ BİR ARAŞTIRMA VE UYGULAMA HASTANESİNİN ACİL  
HASTA VERİLERİNİN VERİ MADENCİLİĞİ İLE ANALİZ EDİLMESİ**

**Sabri ERDEM<sup>1</sup>, Güzin ÖZDAĞOĞLU<sup>1</sup>**

**ÖZ**

Veri Madenciliği büyük hacimli veri gruplarındaki ilişkiler ağının ortaya çıkarılmasında kullanılan yöntemlerden birisi olup, Veri madenciliği, sınıflandırma, gruplandırma, derecelendirme gibi farklı amaçlar için farklı teknikler kullanılmaktadır. Bu çalışma kapsamında, belirli bir dönem boyunca Ege Bölgesi'ndeki bir araştırma ve uygulama hastanesinin acil servisine başvuruda bulunan 214 bin hasta verisi ele alınarak, Veri Madenciliğinde sıklıkla kullanılan birliktelik kuralı yöntemiyle, veri setindeki gizli ancak anlamlılık içeren ilişkiler ortaya çıkarılmaya çalışılmıştır. Çalışma sonuçları, bölgesel özellikler taşıyabileceği düşünülen acil servislere hastaların başvuru nedenleri ve hasta profilleri açısından bir fikir vermekte, ayrıca acil servis bölümlerinin yeniden yapılanma çalışmalarına da farklı bir açıdan yol göstererek katkıda bulunmaktadır.

**Anahtar Kelimeler :** Acil servis, Apriori algoritması, Birliktelik kuralı, Veri madenciliği.

**ANALYZING OF EMERGENCY DATA OF A TRAINING AND RESEARCH  
HOSPITAL IN AEGEAN REGION USING DATA MINING**

**ABSTRACT**

Data mining is one of the methods that discovers the relationship networks within the high volume-data sets, and uses different techniques for different objectives such as classifying, grouping, and rating. In the scope of this study, 214 thousands of patient data is handled including patients who applied to the emergency service in a particular range of time, and significant hidden information about the characteristics of the patients are tried to be obtained. The results of the study contribute by suggesting some considerations about patient's profile to the emergency services having regional characteristics that can canalize them to the right way during improvement studies.

**Keywords:** Emergency service, Apriori algorithm, Association rule, Data mining.

<sup>1</sup> Dokuz Eylül Üniversitesi İşletme Fakültesi, İşletme Bölümü.  
sabri.erdem@deu.edu.tr  
guzin.kavrukkoca@deu.edu.tr

## 1. GİRİŞ

Gelişen teknolojinin üretim ve hizmet işletmelerindeki önemli yararlarından birisi de bilgi saklama ortamı kapasitelerinin artması ve buna bağlı olarak birim bilgi saklama maliyetinin sürekli olarak azalmasıdır. Teknolojinin bir diğer önemli yararı ise bilgi işleme hızının zamana göre geometrik hızla artması ve buna bağlı olarak, birim bilgi işleme maliyetinin oldukça düşmüş olmasıdır. Bu yeni durum bilgi sistemi kullanıcılarını, mevcut işleyişlerine ait bilgi kayıtlarına ek olarak, internet ortamındaki kendilerini ilgilendiren bilgilerin otomatik olarak kaydedilmesinden işletme içi ortamın görsel ve işitsel olarak sürekli kayıt edilip depolanmasına kadar, hemen her türlü bilgiyi depolama, dağıtma, işleme ve raporlamaya sevk etmektedir.

Biriken yoğun hacimli bilginin işleme ve analiz teknikleri de zaman içerisinde evrimleşmiştir. 1990'lı yıllarda Mega Byte seviyesindeki veriler analiz edilirken 2000'li yıllarda Giga Byte seviyesindeki verilerin analizleri ve bunun için derinlik ve hıza dayalı teknikler konu edilir olmuştur (Agrawal ve Srikant, 1994). Bu teknikler arasında İş Zekâsı, OLAP, İstatistiksel Yöntemler, Parametrik Olmayan Yöntemler ve Veri Madenciliği (VM) olarak sayılabilir.

VM kavramı olarak 1900'lü yıllarda ortaya atılmış ve özellikle sınıflandırma, kümeleme ve ilişkisel analizlerde sıklıkla başvurulan yöntemler arasında yer almıştır. Birliktelik kuralları ve apriori algoritması gibi çeşitli yöntem ve tekniklerin geliştirilmesiyle, veritabanlarından yararlı bilgilerin elde edilmesinde VM gittikçe daha yararlı ve kullanışlı bir yöntem olmaya başlamıştır (Han ve Kamber, 2006).

VM, yoğun veri birikimine sahip günümüz üretim ve hizmet işletmelerinin özellikle sepet, pazar, müşteri tercih ve eğilim, finansal işlem analizlerinde sıklıkla kullanılmaktadır (Bose ve Mahapatra, 2001).

Diğer sektörlerde olduğu gibi sağlık sektöründe de yoğun ve çok karakteristikli veritabanlarının oluşması, bu verilerin analiz edilmesi ihtiyacı, bu alanda da veri madenciliğinin en çok tercih edilen teknikler arasında yer almasına neden olmuştur. VM'nin sağlık sektöründeki uygulamaları daha çok biyomedikal (Lucas, 2004), hemşirelik verileri (Berger ve Berger, 2004), klinik araştırma sonuçları, planlama, kalite geliştirme alanlarında olup (Babic, 1999) acil servis gibi farklı tipte pek çok hastanın hizmet gördüğü bir alanda çok fazla olmamıştır. Oysaki acil servis gibi heterojen ve hızlı hizmetin verildiği

bir birimde doğru veriler ve doğru analizler ile daha etkin bir çalışma ortamının sağlanabileceği ve böylelikle hizmet kalitesinin artırılacağı unutulmamalıdır. Bu amaçla, Ege bölgesinde faaliyet gösteren bir Araştırma ve Uygulama Hastanesinin yoğun veri birikiminin bulunduğu, aylar bazında oluşan kayıt sayısının on binlerle ölçüldüğü acil servisinde tutulan hasta kayıtları, sıklıkla tercih edilen veri madenciliği tekniği olan birliktelik kurallarının çıkarılması yöntemiyle analiz edilerek, hastaların demografik profilleri, tanı tanımları ve acil serviste kalma sürelerini içeren anlamlı kurallar elde edilmiştir.

Çalışmanın sonraki bölümünde VM kavramı üzerinde durularak sağlık sektöründe yapılan VM çalışmalarından söz edilmiş ve uygulamada kullanılan VM tekniği açıklanmıştır. Son bölümde ise uygulama sonuçları ve yorumları sunulmuştur.

## 2. MATERYAL VE YÖNTEM

### 2.1 Veri Madenciliği

VM, belli bir yapıda düzenlenmiş veri grupları içerisinde, önceden belirlenmiş bir algoritma göre arama yaparak, karar vericilere referans olabilecek gizli bilgileri ortaya çıkarmak amacıyla geliştirilmiş tümleşik veri arama aracıdır (Frawley vd., 1991).

VM Sürecinde verinin bilgiye nasıl dönüştürüleceğini açıklayan yöntemler "VM Yöntemleri" olarak adlandırılmaktadır. VM Sürecinde işlemler, hedeflerine ve kavramsal farklılıklarına göre sınıflama, kümeleme (Kamimura vd., 2000), tahminleyici modelleme, veri görselleştirme, değişim ve sapma tespiti analizi, birliktelik kuralları (Giudici & Passerone, 2002) olarak ifade edilirken (Wei & Dong, 2000), kullandıkları teknikler ve algoritmalar bakımından da yapay sinir ağları, karar ağaçları, genetik algoritmalar, istatistik teknikler, kural çıkarımı, bulanık kümeler ve duruma dayalı nedensellik olarak sınıflandırılmaktadır.

Sınıflama, belli bir kritere göre verilerin ayrıştırılması; kümeleme, öngörülecek alanların belirlenmesi ve birbirlerine benzeyen verilerin altkümelere ayrılması; tahminleyici modelleme, veritabanındaki bazı alanların diğer alanlara bağlı olarak tahmin edilmesi; veri görselleştirme, verilerin grafiklerle sunulması; değişim ve sapma tespiti, verilerde görülen genel yapıya uymayan kural dışı davranışların ve özelliklerin tespit edilmesi; birliktelik kuralları, birbiriyle ilişkili olan değişkenlerin ortaya çıkarılması ve aralarındaki bağlantının büyüklüğünün belirlenmesi şeklinde kısaca açıklanabilir.

Birliktelik kurallarının uygulandığı çalışmalar incelendiğinde, yayınlanan çalışmaların çoğunun (Agrawal & Srikant, 1994) tarafından önerilen Apriori algoritmasına dayandığı belirtilmektedir (Chang, 2007). Veri madenciliğinde kullanılan birliktelik kuralları içerisinde arama algoritması olarak kullanılmak üzere genellikle yapay zekâ tekniklerinden ve sezgisel yöntemlerden yararlanılmaktadır.

Ürettiği faydalı bilgiler sayesinde günümüzde çok yaygın olarak kullanılan VM, özellikle bankacılık, finans, pazarlama (tüketici davranışları), imalat ve sağlık gibi, büyük ölçekte verinin bulunduğu alanlarda daha yararlı olmaktadır. Veri türü ve sayısının yüksek olduğu sağlık sektöründe VM genellikle klinik ve hastane içerisinde bilgi kalitesinin yükseltilmesi, tanı ve teşhis için ilgili görevlilere referans oluşturması, sağlık kuruluşundaki kaynakların optimum kullanımı gibi hedefler üzerinde yoğunlaşmaktadır. Sağlık veritabanında tanı analizleri (Breault vd., 2002; Kaur & Wasan, 2006), sağlık sigortasında politika geliştirme (Chae vd., 2001), sağlık ile ilgili periyodik olayların incelenmesi (<http://datamining.csiro.au>), sağlık süreçlerinde karar verme (Abidi, 2001), hemşireler için temel bilgilerin çıkarılması (Lee & Abbott, 2003), kliniklerde görevin kötüye kullanıldığı olayların belirlenmesi (Yang & Hwang, 2006), VM'nin sağlık alanındaki uygulamaları arasında yer almaktadır.

Veri madenciliğinin kullanım alanları, kümeleme, sınıflandırma, birliktelik kuralları, sıralama, tahminleme problemleri olarak gruplandırılmaktadır. Bunların arasında, sınıflandırma ve kümeleme algoritmaları ile birliktelik kuralları ise araştırmacıların yoğun olarak üzerinde çalıştığı konulardır.

Bu çalışmada Agrawal ve Srikant (1994) tarafından geliştirilen, "Apriori Algoritması" olarak adlandırılan birliktelik kuralı tekniğinin kullanıldığı sağlık sektöründen bir uygulamaya yer verilmiştir.

## 2.2 Apriori Algoritması

Bir VM yönteminin genel olarak, yöntemin görevi, yapısı, değerlendirme ölçeği, arama yöntemi, veri yönetimi olmak üzere beş bileşeni bulunur. Apriori algoritmasının bu bileşenler açısından incelendiğinde görevi, kural şablonlarını keşfetmek olup, değerlendirme ölçeği olarak destek ve güven parametrelerini kullanılmaktadır. En fazla kullanılan birliktelik kuralları çıkarım algoritmalarından biri olan apriori algoritması, arama yöntemi olarak Önce Genişliğine Arama (ÖGA) algoritmasını içer-

mektedir. Yapay zekâ uygulamalarında sıklıkla kullanılan bir arama tekniği (Russel & Norvig, 1995) olan ÖGA yöntemi, ağaç yapısına sahip hiyerarşideki alternatif düğümleri tarayıp kısıtlara uymayan dalları budayarak ilerler. Böylelikle tüm çözüm uzayının taranarak optimum ilişkilere veya çözümlere ulaşmasını sağlar.

Apriori algoritmasının çalışma prensibi aşağıdaki adımlarla özetlenebilir:

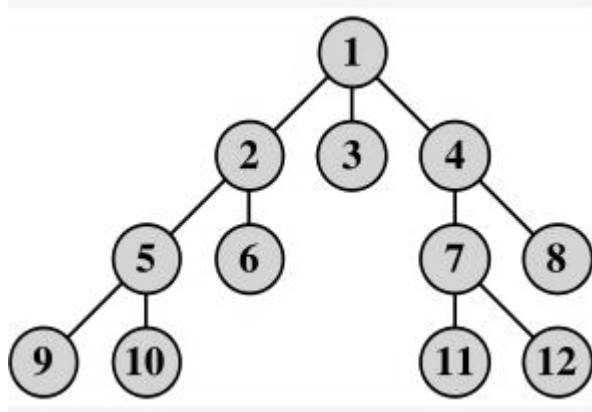
1. Ağaç üzerinde kalan en üst düzeydeki düğümleri kuyruğa ekle.
2. Kuyruğun başındaki düğümü çek.
  - a. Eğer aranan sonuç bu düğümde ise düğüm çözüm sonucu olarak açıkla ve algoritmayı bitir.
  - b. Aranan sonuç bu düğümde değil ise bu düğümü, tüm alt dallarıyla birlikte çözüm uzayından sil.
3. Eğer kuyrukta hiç alternatif düğüm kalmadıysa, aranan sonucun bulunmadığını açıkla ve algoritmayı bitir. Aksi durumda 2. adıma geri dön.

Düğüm noktalarının içinde bulunan sayılar, algoritmanın çözüm alternatifleri arasındaki arama sırasını göstermekte olup, arama biçimi olan ÖGA yönteminin çalışma prensibi Şekil 1'de gösterilmiştir.

Apriori algoritması, veri yönetimini doğrusal taramalarla sağlayıp, adım adım ve tekrarlı biçimde ilerleyerek, veri seti içerisinde sık gözlenen verileri bulmaya çalışmaktadır. Algoritma, her aşama için bir aday set belirler ve o set üzerinde sıklık analizi yapar. Bu süreç, olası tüm veri elemanları ve grupları taranmaya kadar devam eder. Bulduğu sıklıklara göre, veri seti içindeki niteliklerden bir kural ya da kurallar dizisi türetir. Algoritma tarama sırasında, en küçük destek vektörü değeri ve en küçük güven düzeyi olmak üzere iki parametre kullanır. En küçük destek değeri, bir kuralın çıkarılması için gereken en az veri sayısını kontrol ederken, en küçük güven düzeyi de çıkarılan bu kuralın öngörü gücünü kontrol eder. Bu parametrelerin değeri kullanıcı tarafından belirlenir (Chang, 2007). Örneğin, A verisi veritabanındaki bir değişkene ait değer, B verisi ise başka bir değişkene ait bir değer olmak üzere destek ve güven parametreleri aşağıdaki olasılıklarla tanımlanır:

Destek = A ve B değerlerinin birlikte bulunduğu durumların olasılığı:  $P(A \cap B)$

Güven = A değerinin bulunduğu bilindiği koşullarda B değerinin de bulunma olasılığı:  $P(B|A)$



Şekil 1. Önce Genişlik Algoritmasında Arama Sırası

Kullanıcı tarafından belirlenen minimum destek ve güven değerlerini sağlayan birlikler ile algoritmaya devam edilir. Sonuçta anlamlı bir bilgi içeren eğer-ise kurallarına ulaşılır. Apriori algoritmasının teknik çalışma şekli,  $k$ . aday veri seti  $C_k$ ; destek ve güven değeri eşik değerinin üzerinde olan  $k$ . veri seti  $L_k$  olmak üzere, aşağıdaki iki temel adımla ile tanımlanabilir

(<http://dimacs.rutgers.edu/Workshops/>):

**Adım 1 (Veri setine katılma):**  $L_{k-1}$  veri setine bir sonraki adımda eklenen verilerle birlikte  $C_k$  setinin oluşturulması.

**Adım 2 (Veri setinden eleme):** Herhangi bir  $k-1$  indisli veri setinde önceden belirlenen destek ve güven düzeyini geçemeyen veriler  $k$  indisli veri setinin elemanı olamaz.

Çalışmada acil servis sırasında tutulan kayıtlara dayanan veritabanı üzerinden apriori algoritması yürütülerek hastalara ilişkin kişisel ve demografik bilgiler, geliş tarihi ve zamanları, kalma süreleri ile tanımlar arasında çoklu kural taraması yapılmış ve anlamlı kurallara ulaşılmıştır.

Çalışmada kullanılan Apriori algoritması Borgelt tarafından (Borgelt ve Kruse, 2002) C programlama dili kullanılarak geliştirilmiştir. Programın ücretsiz sürümü, çalışma prosedürü ve kaynak kodlarıyla ilgili detaylı bilgi Borgelt'in internet erişim sayfasında (<http://www.borgelt.net/apriori.html>) sunulmaktadır.

### 2.3. Hastane Acil Servis Verileri

Çalışmanın uygulama verilerini, 2001-2006 yılları arasında Ege Bölgesi'ndeki bir araştırma ve uygulama hastanesinin acil servisine başvuran hastaların, acil servis ilk giriş kayıtları ve tanı bilgilerinin bulunduğu

yaklaşık 214 bin kaydı içeren veri seti oluşturulmuştur. Veri seti, "Geliş Tarihi ve Saati", "Ayrılış Tarihi ve Saati", "Hastanın Doğum Tarihi", "Cinsiyeti", "Tanı Kodu", "Tanı Açıklama", "Acil Başvuru No" ve "Hasta Protokol No" alanlarını içermektedir. Acil Başvuru No ve Hasta Protokol No alanları araştırma dışı bırakılmıştır. Bu verilerden; protokol numarası, hastanın hastaneye ilk defa gelmesiyle oluşturulan ve hastanın daha sonraki gelişlerinde de hastayı ayırt edici bilgi olarak kullanılan tekil bir kod; acil başvuru numarası ise acil servise başvuru sırasını gösteren artan özellikte tekil bir sayıdır. Dolayısıyla bu numaralar hasta profilleri hakkında kendi başına herhangi bir bilgi içermeyen tekil verilerdir. Bu nedenle veri seti üzerinde yapılan analizlerde kullanılmamıştır. Geri kalan alanlar ise kategorik hale getirilerek verinin apriori algoritması için uygun hale getirilmesi sağlanmıştır. Hastaneye özel geliştirilmiş veritabanında tutulan veriler, Ms Access (Ms Access Microsoft Firmasının tescilli markasıdır) ortamına aktarılmış ve sezgisel olarak, acil servis çalışmalarına uygun biçimde gruplandırmak ve kategoriler oluşturmak amacıyla güncelleştirme sorguları hazırlanmıştır. Veri hazırlama aşaması olarak adlandırabilecek bu işlem, aşağıdaki şekilde gerçekleştirilmiştir:

1. Geliş Tarihi, aylar itibarıyla kategorilendirilmiştir.

2. Geliş Saati, günün saat dilimlerine göre: Sabah, Öğle, Öğleden Sonra, Akşamüstü, Akşam ve Gece zaman dilimleriyle kategorilendirilmiştir.

3. Ayrılış Tarihi ve Saati, hastanın acil serviste kaldığı saat ve gün sayısını tespit etmek için kullanılmış ve sonuçlar "3 saatten az", "3 saatten fazla", "6 saatten fazla" ve "yarım günden fazla" olmak üzere kategorilendirilmiştir.

4. Hastanın Doğum Tarihi bilgisi kullanılarak hastalar "yaşlı", "orta yaşlı" ve

“genç” kategorileriyle yeniden kodlandırılmıştır.

5. Hastalara acil serviste atanan tanımlar ICD10<sup>2</sup> kodlarının en üst düzeyi temel alınarak kategorilendirilmiştir.

Gerekli veri düzeltmeleri ve kategorilendirme çalışmalarından sonra veri tablosu apriori yazılımının kabul edeceği şekilde metin dosya formatına dönüştürülmüştür. Yazılımın kabul ettiği girdi formatıyla ilgili detaylı bilgi yazarın internet erişim adresinde (<http://www.borgelt.net/apriori.html>) mevcuttur.

Veri hazırlama aşamasından sonra oluşturulan kategorik yapıdaki verilere önceden belirlenmiş destek ve güven düzeylerine bağlı olarak apriori algoritması uygulanmış ve birliktelik kuralları elde edilmiştir. Apriori algoritmasının uygulanmasındaki esas amaç, farklı değişken kombinasyonlarının, veri seti içinde gizlenmiş olan birliktelik kurallarını ortaya çıkarmasıdır. Bu kurallar arasında, önceden tahmin edilebilecek olanların olabileceği gibi tamamen gizlenmiş olanlar da olabilir.

### 3. BULGULAR

Apriori algoritmasının destek parametresi 0,1 ve güven parametresi ise 0,5 alınarak 214 bin kayıt üzerinden Tablo 1’de gösterilen 65 kural üretmiştir. Birliktelik kuralları araştırmasında destek değerinin 0.1 ve güven düzeyinin de 0.5 seçilmesinin nedeni, veri setindeki heterojenliğin, göreceli olarak, veri madenciliği uygulanan diğer alanlardaki verilerin heterojenliğine göre daha yüksek olmasıdır. Bu parametre değerleri, deneme-yanılma yoluyla, kuralların net olarak elde edilebildiği eşiklerdir ve heterojenliğe göre değerlendirildiğinde anlamlı sonuçlar ortaya koymuştur. Bu ifadeyi destekleyici, birliktelik kuralları ile ilgili bazı deneysel çalışmalarda da bu destek ve güven düzeyi değerinin kullanıldığı görülmektedir (Conen ve Leng, 2007; Aggelis, 2004). Tablodaki 1. ve 2. sütunlar, eğer-ise kurallarındaki faktörlerin koşul kategorilerini, 3.sütun koşulların birlikte yer aldığı diğer faktörü, 4. sütun ise sırasıyla destek ve güven düzeylerini belirten değerleri içermektedir.

Acil servis gibi veri heterojenliğinin üst düzeyde olduğu birim için belirlenen destek ve güven düzeylerinde apriori algoritmasının çalıştırılması için ele alınan kategorilendirilmiş veri seti için sırasıyla tekli, ikili ve üçlü kural setleri elde edilmiştir. Ortaya çıkan kurallar incelendiğinde, birliktelik ilişkilerinin özellikle

cinsiyet, kalış zamanı, acil servise geliş zamanı ve bazı tanı gruplarında yoğunlaştığı gözlemlenmiştir. Bulgular arasında acil servise başvuru tarihine ait ay bilgisine ilişkin bir kurala rastlanmamıştır. Dolayısıyla buradan acil servis birimlerinin ay ve mevsimlerden etkilenen bir kural oluşturmadığı sonucuna varılmıştır.

Tablo 1’e bakıldığında elde edilen kuralların genellikle yaş grupları, geliş zamanı, tanı ve cinsiyet odaklı olduğu görülmektedir. Bunun anlamı, acil servislerin karakteristik özelliklerinin belirtilen faktörlerle yakından ilişkili olduğudur. Faktör grupları açısından Tablo 1’deki sonuçlardan bir kısmı, diğer kuralların yorumlanmasında yol gösterici olmaları düşüncesiyle incelenmiştir.

İlk iki kurala bakıldığında yüksek önem düzeyinde, başvuran hastaların “kadın” olduğu görülmektedir. Genel olarak acil serviste kalma süresinin de %100 destek ve %77,8 güven düzeyinde 3 saatten az olduğu görülmüştür. Cinsiyet koşuluna bağlı ikili kurallara bakıldığında (kural aralığı 3~14), acile başvuran kişinin “kadın” olması durumunda görece olarak yüksek destek ve güven düzeylerinde “3 saatten az” birimde kaldığı, yaşının “orta yaş” kategorisinde olduğu ve başvuru saatlerinin akşam ve gece dilimlerine rastladığı söylenebilir.

Kalış zamanının “3 saatten az” olduğu durumlar göz önüne alındığında (24~26), gelen hastaların cinsiyetleri arasında “kadın” olma olasılığının daha fazla olduğu (%56,2), “orta” ve “genç” yaş grubu olma olasılığı (sırasıyla %40,9 ve %33,2) ise görece olarak diğer yaş gruplarına göre baskın olduğu görülmektedir. Gelen hastanın acil birimde 6 saatten fazla kaldığı biliniyorsa bu hastanın “kadın” olma olasılığı %12,7 destek düzeyinde, %56,7 olarak elde edilmiştir.

Acil servise başvuru saat dilimlerinin ortaya çıkardığı kurallara (30~34) göre, eğer hastanın “öğle” saat diliminde geldiği biliniyorsa bu hastanın “kadın” olması ve “3 saatten az” kalma süresine sahip olması, verilen eşik değerlerini aşarak bir kural dizisi olarak ortaya çıkmaktadır. Başvuru saat dilimlerinin “öğle”, “akşam” ve “akşam üzeri” dilimlerinde yoğunlaşması da bu kurallar içinde görülmektedir. Öğlen ve akşamüstü gelen hastalarının 3 saatten az kalma olasılıkları sırasıyla %12,7 destek düzeyinde %90,4 ve %14,7 destek düzeyinde %79,3 olarak bulunmuştur.

Tablo 1. Veri Analizi Sonucunda Ulaşılan Birliktelik Kuralları

Kural No.	1. Koşul	2. Koşul	Sonuç	Parametre değerleri (güven, destek)
1		Cinsiyet=KADIN ==>		(100,562)
2		Kalış Zaman=3 SAATTEN AZ ==>		(100,77,8)
3		Cinsiyet=KADIN ==>	Geliş Zaman=OGLEDEN SONRA	(109,564)
4		Cinsiyet=KADIN ==>	Geliş Zaman=OGLE	(12,7,56,7)
5		Cinsiyet=KADIN ==>	Kalış Zaman=6 SAATTEN FAZLA	(130,568)
6		Cinsiyet=KADIN ==>	Tani Kodu=Z	(123,568)
7		Cinsiyet=KADIN ==>	Geliş Zaman=AKSAMUZERI	(14,7,56,1)
8		Cinsiyet=KADIN ==>	Geliş Zaman=GECE	(193,545)
9		Cinsiyet=KADIN ==>	Geliş Zaman=AKSAM	(22,1,58,8)
10		Cinsiyet=KADIN ==>	Yas Grup=GENC	(332,55,2)
11		Cinsiyet=KADIN ==>	Geliş Zaman=SABAH	(203,54,5)
12		Cinsiyet=KADIN ==>	Yas Grup=ORTA YAS	(409,57,5)
13		Cinsiyet=KADIN ==>	Kalış Zaman=3 SAATTEN AZ	(77,8,55,9)
14		Cinsiyet=KADIN ==>	Yas Grup=YASLI	(25,8,55,3)
15		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=OGLE	(12,7,90,4)
16		Kalış Zaman=3 SAATTEN AZ ==>	Tani Açıklama=DIGER	(12,3,76,6)
17		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=AKSAMUZERI	(14,7,79,3)
18		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=GECE	(193,70,1)
19		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=OGLEDEN SONRA	(109,89,1)
20		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=SABAH	(203,91,7)
21		Kalış Zaman=3 SAATTEN AZ ==>	Yas Grup=GENC	(332,79,0)
22		Kalış Zaman=3 SAATTEN AZ ==>	Geliş Zaman=AKSAM	(22,1,57,9)
23		Kalış Zaman=3 SAATTEN AZ ==>	Yas Grup=YASLI	(25,8,75,3)
24		Kalış Zaman=3 SAATTEN AZ ==>	Yas Grup=ORTA YAS	(409,78,4)
25		Kalış Zaman=3 SAATTEN AZ ==>	Cinsiyet=ERKEK	(43,7,82)
26		Kalış Zaman=3 SAATTEN AZ ==>	Cinsiyet=KADIN	(56,2,77,5)
27		Kalış_Zaman=3_SAATTEN_ AZ ==>	Yas_Grup=GENC	(332,79,0)
28		Kalış_Zaman=3_SAATTEN_ AZ ==>	Yas_Grup=ORTA_YAS	(409,78,4)
29		Kalış_Zaman=6_SAATTEN_ FAZLA ==>	Cinsiyet=KADIN	(130,56,8)
30		Geliş_Zaman=OGLE ==>	Cinsiyet=KADIN	(12,7,56,7)
31		Geliş_Zaman=OGLE ==>	Kalış_Zaman=3_SAATTEN_AZ	(12,7,90,4)
32		Geliş_Zaman=AKSAM_ UZE RI ==>	Cinsiyet=KADIN	(14,7,56,1)
33		Geliş_Zaman=AKSAM_ UZE RI ==>	Kalış_Zaman=3 SAATTEN AZ	(14,7,79,3)
34		Geliş_Zaman=AKSAM ==>	Cinsiyet=KADIN	(22,1,58,8)
35		Tani_Grup=R ==>	Kalış_Zaman=3 SAATTEN AZ	(21,5,75,2)
36		Tani_Grup=R ==>	Cinsiyet=KADIN	(15,5,54,4)
37		Tani_Grup=S ==>	Cinsiyet=ERKEK	(13,8,53,9)
38		Tani_Grup=S ==>	Kalış_Zaman=3 SAATTEN AZ	(13,8,81,3)
39		Tani_Grup=J ==>	Cinsiyet=KADIN	(11,3,54,8)

40		Tani Grup=J	⇒	Kalış Zaman=3 SAATTEN AZ	(113,77.7)
41	Cinsiyet=ERKEK	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=YASLI	(11.5,75.8)
42	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Geliş Zaman=AKSAM	(13.0,58.0)
43	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Geliş Zaman=GECE	(10.5,69.1)
44	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Geliş Zaman=SABAH	(11.1,91.9)
45	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=YASLI	(14.3,75.0)
46	Cinsiyet=ERKEK	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=GENC	(14.8,79.7)
47	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=GENC	(18.4,78.4)
48	Cinsiyet=ERKEK	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=ORTA YAS	(17.3,78.5)
49	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Yas Grup=ORTA YAS	(23.5,78.3)
50	Cinsiyet=KADIN	Kalış Zaman=3 SAATTEN AZ	⇒	Tani Grup=R	(12.0,52.6)
51	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Yas Grup=YASLI	(19.5,55.0)
52	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Geliş Zaman=OGLE	(11.5,56.7)
53	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Geliş Zaman=AKSAMUZERI	(11.7,56.2)
54	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Geliş Zaman=GECE	(13.5,53.7)
55	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Geliş Zaman=SABAH	(18.6,54.6)
56	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Geliş Zaman=AKSAM	(12.8,58.9)
57	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Yas Grup=GENC	(26.3,54.8)
58	Kalış Zaman=3 SAATTEN AZ	Cinsiyet=KADIN	⇒	Yas Grup=ORTA YAS	(32.0,57.4)
59	Geliş Zaman=AKSAM UZERI	Kalış Zaman=3 SAATTEN AZ	⇒	Cinsiyet=KADIN	(11.7,56.2)
60	Geliş Zaman=GECE	Cinsiyet=KADIN	⇒	Kalış Zaman=3 SAATTEN AZ	(10.5,69.1)
61	Geliş Zaman=GECE	Kalış Zaman=3 SAATTEN AZ	⇒	Cinsiyet=KADIN	(13.5,53.7)
62	Tani Grup=R	Cinsiyet=KADIN	⇒	Kalış Zaman=3 SAATTEN AZ	(13.2,75.2)
63	Tani Grup=R	Kalış Zaman=3 SAATTEN AZ	⇒	Cinsiyet=KADIN	(16.2,61.3)
64	Tani Açıklama=AKUT Tani Grup=J	Kalış Zaman=3 SAATTEN AZ	⇒	NAZOFARINJIT	(6.0,66.5)
65	Geliş Zaman=OGLE EN SONRA	Kalış Zaman=3 SAATTEN AZ	⇒	Cinsiyet=KADIN	(9.8,56.4)

İkili kuralların son bölümü olan tanılarda (35~40) ise S tanı kodunun (yaralanma, zehirlenme, dış etmenler) daha çok “erkek” hasta başvurularında görüldüğü ve “3 saatten az” kalma süresine sahip olduğu Tablo 1’de verilen destek ve güven düzeylerinden çıkarılabilir. Belirlenen destek ve güven düzeylerinde, R (ağrı ve bayılmalar) ve J (solunum sistemi hastalıkları) tanı kodlarına ise “kadın” hastaların başvurularında rastlandığı görülmektedir. Tanı kodlarının açıklaması için ICD-10 kod listesi incelenerek, yapılan sınıflandır-

maya göre en üst başlıklar ele alınmış ve bu başlıklar Tablo 2’de gösterilmiştir.

Üçlü karakteristiklere bağlı kurallar da anlamlı ilişkileri ortaya koymaktadır. Buna göre, 41 No’lu kuraldan, cinsiyetinin “erkek” olduğu ve “3 saatten az” kaldığı bilinen bir hastanın “yaşlı” kategorisinde olma olasılığı %11,5 destek düzeyinde, %75,8 olarak elde edilmiştir. 49 No’lu kurala göre cinsiyetin “kadın” ve kalış süresinin “3 saatten daha az” olduğu biliniyorsa, bu hastanın “orta yaş” kategorisinde olma olasılığı %23,5 destek

düzeyinde %78,3dir. Aynı hasta grubu için R tanı grubunda yer alma olasılığının eşik değerlerinin üzerinde olduğu 50 No'lu kuraldan

anlaşılmaktadır. Ayrıca üç karakteristikli kural setlerinin geri kalanı incelendiğinde, daha önce açıklanan iki karakteristikli kural setlerini doğrulayan sonuçların elde edildiği görülmüştür.

Tablo 2. ICD-10 Kod Sisteminde Tanımlar

TANI GRUBU	AÇIKLAMA
A	Bazı enfeksiyöz ve paraziter hastalıklar (A00-B99)
B	
C	Neoplazmlar (C00-D48)
D	Kan ve kan yapıcı organların hastalıkları ve immün sistemin bazı bozuklukları (D50-D89)
E	Endokrin, beslenme ve metabolizma hastalıkları (E00-E90)
F	Mental ve davranışsal bozukluklar (F00-F99)
G	Sinir sistemi hastalıkları (G00-G99)
H	Göz ve adnekslerinin hastalıkları (H00-H59)
	Kulak ve mastoid çıkıntı hastalıkları (H60-H95)
I	Dolaşım sistemi hastalıkları (I00-I99)
J	Solunum sistemi hastalıkları (J00-J99)
K	Sindirim sistemi hastalıkları (K00-K93)
L	Deri ve derialtı dokunun hastalıkları (L00-L99)
M	Kas iskelet sistemi ve bağ dokusu hastalıkları (M00-M99)
N	Genitoüriner sistem hastalıkları (N00-N99)
O	Gebelik, doğum ve lohusalık (O00-O99)
P	Perinatal dönemden kaynaklanan bazı durumlar (P00-P96)
Q	Konjenital malformasyonlar, deformasyonlar ve kromozom anomalileri (Q00-Q99)
R	Semptomlar, belirtiler ve anormal klinik ve labratuvar bulguları, başka yerde sınıflanmamış (R00-R99)
S	Yaralanma, zehirlenme ve dış nedenlerin bazı diğer sonuçları (S00-T98)
T	
V	Hastalık ve ölümün dış sebepleri (V01-Y98)
W	
X	
Y	
Z	Sağlık servisleriyle temas ve sağlık durumunu etkileyen faktörler (Z00-Z99)



#### 4. SONUÇALAR

Hastane acil servisleri ani rahatsızlıklara ilk müdahaleyi gerçekleştirmek amacıyla çalışan sağlık birimleri olarak 7 gün 24 saat hizmet vermektedir. Gelen hasta profilleri incelendiğinde yaş, cinsiyet ve tanı tipleriyle, acil servis ve acil servis sonrası hastanede kalma durumları ve kalış süreleri gibi veriler arasında bir ilişkinin varlığı sorgulanarak bu hizmetlerin kalite ve etkinliğinin geliştirilmesi mümkün olmaktadır.

Birliktelik kuralları araştırmasında destek değerinin 0,1 ve güven düzeyinin de 0,5 seçilmesinin nedeni, veri setindeki heterojenliğin, göreceli olarak, veri madenciliği uygulanan diğer alanlardaki verilerin heterojenliğine göre daha yüksek olmasıdır. Bu parametre değerleri, deneme yoluyla, kuralların net olarak elde edilebildiği eşiklerdir ve heterojenliğe göre değerlendirildiğinde anlamlı sonuçlar ortaya koymuştur. Bu durum, gruplar açısından herhangi bir faktörün diğer bir faktör üzerinde etki düzeyini azaltmaktadır.

Bu çalışmada geçmişe dönük olarak sağlanabilen kapsamlı veri seti üzerinden düzenlemeler yapılmış ve veriler arasındaki ilişkilerin analizi bir veri madenciliği yöntemi olan apriori algoritması ile gerçekleştirilerek birliktelik kurallarının varlığı sorgulanmıştır. Belirli bir destek ve güven değerine göre yapılan analizler sonucunda 65 anlamlı kurula ulaşılmıştır. Elde edilen kurallardan yaş ve cinsiyet gruplarına göre acil servise başvuru zamanları ve kalış sürelerine ilişkin anlamlı bilgiler elde edilmiştir. Bu bilgiler, acil servis açısından kapasite ve eylem planlarında karar vericilere destek olabilecek özelliğindedir.

Çalışma sonuçları, bölgesel özellikler taşıyabilecekleri düşünülen acil servislere hastaların başvuru nedenleri ve hasta profilleri açısından bir fikir vermekte, ayrıca acil servis bölümlerinin yeniden yapılanma çalışmalarına da farklı bir açıdan yol göstererek katkıda bulunmaktadır.

İlgili acil servis biriminden elde edilen veri seti içerisinde sadece hastaneden çıkış bilgileri veri olarak tutulmakta ve acil müdahale yapıldıktan sonra hangi birime gönderildiğine dair bir veri grubu bulunmamaktadır. Bu kısıt daha ileri araştırmaları ve analizleri olanaksız kılmaktadır. VM yöntemlerinin başarısının veri detayına da bağlı olduğu düşünüldüğünde, karar vericilere ve çalışanlara destek olabileceği belirlenen tüm verilerin düzenli olarak kayıt altına alınması öneril-

mektedir. Aynı zamanda hastalık tanı kodlarının çok fazla olması, ana tanı gruplarının da içinde yer aldığı kuralların belirlenmesini engellemektedir. Dolayısıyla, bundan sonra acil servise gelen hastaların uluslararası standart hastalık tanı kodlarının yanında ilgili servis ya da anabilim dalı bazında da gruplandırılmış olması, gelecekteki analizler açısından daha yararlı olacaktır.

#### KAYNAKLAR

- Abidi, S.S.R. (2001). Knowledge management in healthcare: Towards 'knowledge-driven' decision-support services. *International Journal of Medical Informatics* 63, 5–18.
- Aggelis, V. (2004). Association rules model of e-banking services, 5th International Conference on Data Mining, Text Mining and their Business Applications.
- Agrawal, R., Srikant, R. (1994). Fast algorithms for mining association rules. In *Proceeding of the 20th VLDB Conference Santiago*, 487-499.
- Babic A. (1999). Knowledge Discovery for Advanced Clinical data Management and Analysis, (Ed. Kokol P., Zupan B., Stare J.), *Medical Informatics Europe*, 409-414.
- Berger A. M., Berger C. R. (2004). Data Mining as a Tool for Research and Knowledge Development in Nursing, *Computers, Informatics, Nursing*. 22(3), 123-131.
- Borgelt C., Kruse R. (2002). Induction of Association Rules: Apriori Implementation, 15th Conference on Computational Statistics, Berlin, Germany.
- Bose I., Mahapatra R. K. (2001). Business Data Mining –A Machine Learning Perspective, *Information & Management*, 39(3), 211-225
- Breault, J. L., Goodall, C. R., Fos, P. J. (2002). Data mining a diabetic data warehouse. *Artificial Intelligence in Medicine* 26(1), 37–54.
- Chae, Y. M., Ho, S. H., Cho, K. W., Lee, D. H., Ji, S. H. (2001). Data mining approach to policy analysis in a health in-

- urance domain. *International Journal of Medical Information* 62(2), 103–111.
- Chang, C-L. (2007). A study of applying data mining to early Intervention for developmentally-delayed children. *Expert Systems with Applications* 33, 407–412.
- Conen, F., Leng, P. (2007). The effect of threshold values on association rule based classification accuracy, *Data & Knowledge Engineering* 60(2), 345-360.
- Frawley, W.J., Paitetsky-Shapiro, G., Matheus, C. J. (1991). Knowledge discovery in databases: an overview. Knowledge discovery in databases. California: AAAI/MIT Press, pp.1–30.
- Giudici, P., Passerone, G. (2002). Data mining of association structures to model consumer behaviour. *Computational Statistics and Data Analysis* 38(4), 533–541.
- Han J., Kamber M. (2006). Data Mining: Concepts and Techniques, Morgan Kaufman, 23-25.
- Kamimura, R. T., Bicciato, S., Shimizu, H., Alford, J., & Stephanopoulos, G. (2000). Mining of biological data II: assessing data structure and class homogeneity by cluster analysis. *Metabolic Engineering* 2(3), 228–238.
- Kaur, H., Wasan, S.K. (2006). Empirical Study on Applications of Data Mining Techniques in Healthcare. *Journal of Computer Science* 2 (2), 194-200.
- Lee, S-M., Abbott, P.A. (2003). Bayesian networks for knowledge discovery in large datasets: basics for nurse researchers. *Journal of Biomedical Informatics* 36, 389–399.
- Lucas, P., (2004), Bayesian analysis, pattern analysis, and data mining in health care, *Current Opinion in Critical Care*, 10(5), 399-403.
- Russell, S.J., Norvig, P. (1995). Artificial Intelligence: A Modern Approach, p.74-75. Prentice Hall, New Jersey.
- Wei, Z. P., Dong, H. S. (2000). Theory and practices of e-commerce. HwaTai Publications, 147–184.

Yang, W-S., Hwang, S-Y. (2006). A process-mining framework for the detection of healthcare fraud and abuse. *Expert Systems with Applications* 31, 56–68.

<http://datamining.csiro.au>, 29.04.2007

<http://dimacs.rutgers.edu/Workshops/Health/Reference/slides/madigan.ppt>; 1.05.2007

<http://sbu.saglik.gov.tr/icd10>; 15.06.2007

<http://www.borgelt.net/apriori.html>; 12.5.2008



**Sabri ERDEM**, 1972 yılında Divriği’de doğan Sabri ERDEM, ilk ve orta öğrenimini Ankara’da tamamlamış, 1996 yılında D.E.Ü. Endüstri Mühendisliği Bölümünden Mezun olduktan sonra yine aynı üniversiteden 2000 yılında Üretim Yönetimi Yüksek Lisans, 2001 yılında Bilgisayar Müh. Yüksek Lisans ve 2008 yılında Bilgisayar Müh. Doktora Derecelerini almıştır. D.E.Ü. İşletme Fakültesi Sayısal Yöntemler Anabilim Dalı’nda 1997-2004 yılları arasında araştırma görevlisi olarak çalışmış ve 2004’ten itibaren de öğretim görevlisi olarak görev yapmaktadır Evrimsel Optimizasyon, Yönetim Bilişim Sistemleri, Akıllı Sistemler, Sağlıkta Kaynak Planlama ve Etkinlik Ölçme temel çalışma ve yayın alanlarıdır.



**Güzin ÖZDAĞOĞLU**, 1978 yılında İzmir’de doğan Güzin Özdağoğlu, 2000 yılında Dokuz Eylül Üniversitesi Endüstri Mühendisliğinden mezun olup, 2003 yılında aynı bölümden master derecesini almıştır. 2003 yılında başladığı doktora öğrenimine tez aşamasında devam etmektedir. 2000 yılında Dokuz Eylül Üniversitesi İşletme Fakültesi İşletme Bölümünde başladığı Araştırma Görevliliğini halen sürdürmekte olup, araştırma alanları arasında üretim planlama, modern sezgisel yöntemler, kalite fonksiyon geçirimi ve kurumsal mimari modelleri yer almaktadır. Kendisinin bu alanlarda yayınlanmış makaleleri, bildirileri bulunmaktadır. Ayrıca, Süreç İyileştirme, ISO 9001, Norm Kadro, İş Değerleme ve Reorganizasyon projelerinde çalışmıştır. Güzin Özdağoğlu, TMMOB Makina Mühendisleri Odası İzmir Şubesinde Yönetim Kurulu üyeliğini Endüstri ve İşletme Mühendisliği Meslek Dalı Ana Komisyonu üyeliğini sürdürmektedir.