

TELE PAZARLAMA VERİLERİNİN BİRLİKTELİK KURALLARIYLA VE CRISP-DM YÖNTEMİYLE ANALİZ EDİLMESİ*

Muhammed Bilgehan AYTAÇ*
Hasan Şakir BİLGE†

ÖZET

Veri madenciliği ile pazarlama arasındaki yoğun etkileşim günümüz işletmelerinde yaygın bir hal almıştır. Veri tabanı pazarlaması adlı yeni bir anlayış doğmuş, işletmeler müşteri bölümlene ve müşteri ilişkileri gibi konularda bu yaklaşım çerçevesinde hareket etmeye başlamışlardır. Bu çalışmada son dönemde çok hızlı bir şekilde gelişen ve ilgi odağı olan veri madenciliğinin doğrudan pazarlama üzerine örnek bir uygulaması gerçekleştirilmiştir. Doğrudan pazarlamanın bir alt kolu olarak tanımlanabilecek tele pazarlama bankalar tarafından yoğun olarak kullanılmaktadır. Ancak bankaların müşterileriyle yaptığı görüşmeler çoğunlukla müşteriler tarafından hoş karşılanmamakta hatta çoğunlukla yapılan çağrılara müşteriler hiç yanıt vermemeyi tercih etmektedirler. Tele pazarlama başarımındaki bu düşüşü arttırabilmek için veri madenciliği önemli bir araçtır. Bankalar müşterilerini daha iyi tanıdıkça onlarla daha sağlıklı ilişki kurabilecektir. Bu çalışmada da birliktelik kurallarıyla yapılan analizler sonucunda bir bankanın müşterileri hakkında aydınlatıcı bilgiler üretilmeye çalışılmıştır. Uygulama kapsamında Portekiz’de bulunan bir bankanın verileri CRISP-DM (Cross Industry Standard Process for Data Mining; Çapraz Endüstri Veri Madenciliği Standart Süreci) yöntemiyle analiz edilmiştir. Bankaların veya müşterileriyle tele pazarlama yoluyla iletişim kuran diğer firmaların müşterilerini daha iyi tanımalarına ve tekliflerini en uygun yolla belirli müşterilere iletmelerine yönelik örnek bir veri madenciliği süreci sunulmuştur.

Anahtar Kelimeler: Veri madenciliği, CRISP-DM, doğrudan pazarlama

*Bu çalışma Gazi Üniversitesi Bilişim Enstitüsü Yönetim Bilişim Sistemleri Bölümü’nde yapılan yüksek lisans tezinden üretilmiştir.

* Aksaray Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı

† Gazi Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü

ANALYZING TELEMARKETING DATASET WITH ASSOCIATION RULES AND CRISP-DM METHODOLOGY**ABSTRACT**

The intense interaction between marketing and data mining is common today's businesses. A new approach created called; database marketing. Business managers began to use this approach more widely especially in terms of customer segmentation and Customer Relationship Management (CRM). In this work, we developed a data mining application-which is rapidly growing and becoming limelight- for direct marketing. As a sub-branch of direct marketing, telemarketing is used often by banks. But today, customers do not enjoy this and commonly they prefer not to answer calls. To prevent from this decline in the popularity of telemarketing, data mining is an important tool. Banks can understand customers and their behaviors more deeply using such application. As a result, they are able to develop a stronger relationship with them. In this work, We tried to produce some beneficial information about bank customers via the results of association rules. In that context, We analyzed the dataset of a Portuguese bank using CRISP-DM (Cross Industry Standard Process for Data Mining) methodology. It offers a case data mining study for banks or other businesses which employ telemarketing for know their costumers better and convey their proposals to specific costumers in most appropriate way.

Key words: Data mining, CRISP-DM, direct marketing

1. GİRİŞ

1992 yılında yapılan bir çalışmada dünyadaki veri miktarının 20 ayda bir ikiye katlandığı ifade edilmektedir (Frawley vd., 1992) Günümüzde bu hızın hangi boyutlara ulaştığını tahmin etmek hiç de zor değildir. Veri madenciliği gelişen veri tabanı teknolojisine paralel olarak çok hızlı bir gelişme göstermiştir. Kullanılan algoritmalar, bilgisayar yazılımları, uygulama alanları çeşitlenmiştir. Hemen hemen bütün kalburüstü işletmeler sadakat kartlarıyla müşteriler hakkında veri toplamaya başlamıştır. Çevrimiçi alışverişte müşterilere alışveriş yaparken geçmişte yaptığı alışverişler analiz edilerek anında yeni teklifler sunulur hale gelmiştir. Sigorta ve finans şirketleri veri madenciliği ile müşterilerin riskliliği, dolandırıcılık tespiti (fraud detection) gibi işlemleri gerçekleştirmektedir. Veri madenciliğinin yaygın kullanım alanlarından birisini ise pazarlama teşkil etmektedir. İlgili yayınlar incelendiğinde pazarlama ile veri madenciliği etkileşimine çok sık rastlamak mümkündür (Emel ve Taşkın, 2005; 2005; 2010; Ulaş vd., 2001; Akbulut, 2006 , Ngai vd., 2009; Ulaş, 2001; Giudici ve Passerone, 2002). Temel olarak veri madenciliğinin kullanıldığı pazarlama faaliyetleri şunlardır (Çankırı vd., 2009; Olson ve Delen, 2008):

- Müşteri profili belirleme
- Hedef müşteri belirleme

- Sepet analizi
- CRM (Müşteri ilişkileri yönetimi)
- Satış tahmini

2. TEORİK ÇERÇEVE

Burada kısaca çapraz endüstri veri madenciliği standart süreci, veri madenciliği teknikleri, veri madenciliği araçlarından bahsedilecektir.

2.1. Veri Madenciliği

Veri, işlenmemiş bilgi olarak ifade edilmektedir. Veriler kullanılarak yönetime yön verecek kararlar alınmaz. Verileri işleyerek bilgi haline getirmek gerekir (Savaş vd., 2012). Bu işlem veri madenciliği olarak ifade edilir. Veri madenciliği basamakları olan bir süreçtir. Bu süreç kimi zaman otomatik kimi zaman ise yarı otomatik olarak gerçekleştirilir (Witten ve Frank, 2011;5). Tanım vermek gerekirse; veri madenciliği basit olmayan ve daha önce bilinmeyen örüntülerin veri tabanından açığa çıkarılması ve faydalı olacak bilgilerin verilerden üretilmesi olarak ifade edilebilir (Frawley vd., 1992).

Veri tabanı ise bilgisayarlarda toplanan ve yapılandırılmış şekilde bulunan veri biriktirisi. Yapılandırılmış ifadesi verilerin sistematik olarak belli bir formatta depolanmasını ifade etmektedir (Bai, 2010; 12).

2.1.1. Veri Madenciliği Teknikleri ve Araçları

Veri madenciliği alanında her geçen gün yeni yöntem ve algoritmalar geliştirilmektedir. Kullanılan teknikler geleneksel ve yeni olarak aşağıdaki gibi ikiye ayrılabilir (Koyuncugil ve Özgülbaş, 2009).

Geleneksel yöntemler:

- Regresyon
- K-en yakın komşuluk
- Kümeleme

Yeni yöntemler:

- Karar ağaçları
- Birliktelik kuralları
- Sinir ağları

Özellikle işletmelerin bulundurduğu veri tabanlarında sıklıkla uygulanan birliktelik kuralları büyük veri setlerine kolaylıkla uygulanabilmektedir. Birliktelik kuralları gözlem değerleri arasındaki ilişkiyi, veri madenciliğini yöneten kişinin tanımladığı bir başarı oranı üzerinde gerçekleştirir ve koşullu olasılık temelli değerlendirmelerle ilişkiyi özet olarak sunar (Koyuncugil ve

Özgülbaş, 2009). Pazarlama alanında özellikle sepet analizi konusunda yaygın olarak kullanılmaktadır (Giudici ve Passerone, 2002).

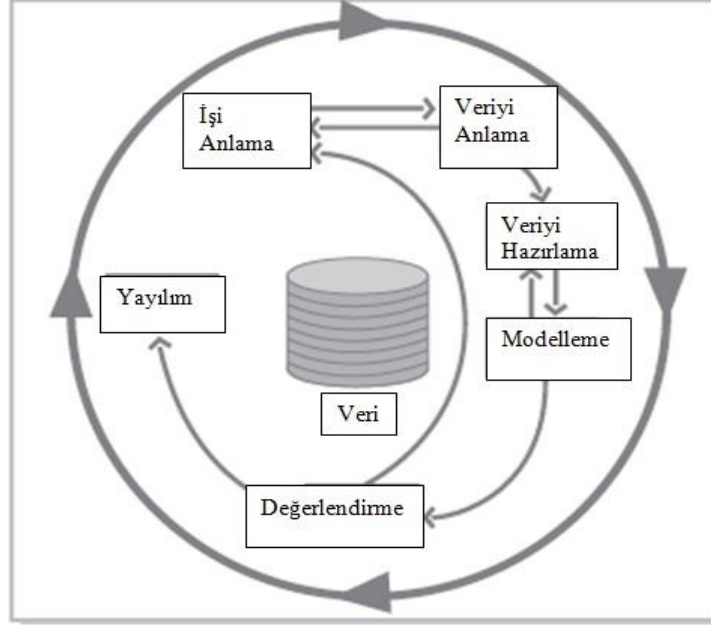
Bahsi geçen tekniklerin uygulamalarını gerçekleştirebilmek için çeşitli bilgisayar yazılımlarına ihtiyaç duyulmaktadır. Bu kapsamda çeşitli firmalar tarafından, *SPSS Clementine*, *Excel*, *SPSS*, *SAS*, *Angoss*, *KXEN*, *SQL Server*, *MATLAB ticari* ve *RapidMiner (YALE)*, *WEKA*, *R*, *C4.5*, *Orange*, *KNIME* gibi yazılımlar geliştirilmiştir.

2.1.2. Çapraz Endüstri Veri Madenciliği Standart Süreci (CRISP-DM)

Konuya ilişkin yayınlarda çeşitli süreçlere rastlamak mümkündür, ancak yaygın olarak bu döngü takip edilmektedir (Küçüksille, 2009; Wirth ve Hipp, 2000; Akbulut, 2006). Bu döngü CRISP DM olarak ifade edilmektedir. Türkçe *çapraz endüstri veri madenciliği süreci* olarak ifade edilen sürecin geliştiricileri Daimler-Chrysler (sonradan Daimler Benz AG, Almanya), NCR Sistem Mühendisliği Kopenhag (Danimarka), SPSS (İngiltere) ve OHra Verzegeringen en Bank Groep B.V (Hollanda) isimli firmalardır (Küçüksille, 2009).

Hiyerarşik olarak ifade etmek gerekirse veri madenciliği süreci aşağıdaki basamakları takip etmektedir (Chapman vd., 2000):

- İş kavrama
- Veriyi kavrama
- Veriyi hazırlama
- Modelleme
- Değerlendirme
- Yayılım



Şekil 1: Çapraz Endüstri Veri Madenciliği Süreci (Chapman vd., 2000).

Chapman ve arkadaşları süreci şöyle izah etmişlerdir (Chapman vd., 2000):

“İlk olarak işi kavramayla süreç başlar. Bu aşama sürecin sağlıklı işlemesi için çok iyi tahlil edilmesi gereken bir basamaktır. Sürecin sonucunda ulaşılmaya arzu edilen hedefler, yönetici beklentileri açık ve net bir şekilde ifade edilmelidir. Her araştırma süreci gibi veri madenciliği de rastgele veya tesadüfi olarak ilerlememektedir. Açık iş hedefleri ve net beklentiler ortaya konmalıdır.

Daha sonra veriyi kavrama aşaması takip edilir. Elde bulundurulmuş veri tabanı ne kadar iyi ve yakından tanınıyorsa veri madenciliği süreci o kadar sağlıklı işleyecektir. Veri tabanındaki yanlış veriler, niteliklerin anlamı, kayıp veriler iyice analiz edilmelidir.

Verinin sürece uygun hale getirilmesinin yapıldığı aşama ise veriyi hazırlama aşamasıdır. Veri ayrıca kullanılacak algoritmalara ve yazılımlara hazır bir formata dönüştürülmeli, eksik veya kayıp veriler temizlenmelidir.

Modelleme ise sürecin teknik olarak gerçekleştirildiği bölümdür. Bu bölüm şu şekilde sıradüzen olarak ifade edilebilir; model tekniğini seçme, modelin sınıma tasarımını oluşturma, modelin inşası ve modelin değerlendirilmesi.

Modelleme, sürecin sözel ve metinsel olarak gerçekleştiği bölümdür. Sürecin ne derece tatminkâr olduğu özellikle değerlendirilirken gelecekte yapılacak veri madenciliği projeleri içinde bir takım kararlar alır.

Son olarak ise sürecin rapor edilmesi ve belgelendirilmesi aşaması gerçekleştirilir. Süreci bir aracı araştırma firmasının gerçekleştirdiği düşünülürse bu aşama hizmeti satın alan kullanıcı firma tarafından önemlidir.”

2.2. Doğrudan Pazarlama ve Tele Pazarlama

Genellikle pazarlamanın tutundurma çabaları ve reklamcılık faaliyetleri iki tür yaklaşımla sergilenmektedir. Bunlardan birincisi kitlesel pazarlama iken diğeri doğrudan pazarlama olmaktadır. Başarı oranı %1 gibi düşük bir rakamla ifade edilen kitlesel pazarlama yavaş yavaş etkisini kaybetmektedir (Ling ve Li, 1998). Artık işletmeler doğrudan pazarlama ile müşterileriyle birebir iletişim kurmaya daha özen göstermektedir. Doğrudan Pazarlama Birliği (Direct Marketing Association) tarafından ise doğrudan pazarlama şöyle tanımlanmıştır (Nakip vd., 2012):

“Herhangi bir yerde ölçülebilir bir tepki ya da ticari işlemi etkilemek için bir veya birden fazla reklam medyasını kullanan etkileşimli bir pazarlama sistemidir.”

Bir doğrudan pazarlama aracı olan tele pazarlama ise potansiyel müşterileri çekme, mevcut müşterilere satış yapabilme ve müşterilerin sorularına yanıt vererek veya sipariş alarak hizmet sağlama amacıyla telefonların ve çağrı merkezlerinin kullanımı olarak ifade edilebilir. Tele pazarlama genel olarak iki yönlüdür. Birincisi firmadan müşterilere hizmet ve teklif sunma amaçlı gerçekleştirilen dışa doğru tele pazarlamadır. İkincisi ise müşterilerin istek ve sorunlarını yanıtlamak suretiyle hizmet vermek veya ürün veya hizmet satın alabilmek amacıyla gerçekleştirdiği içe doğru tele pazarlamadır (Kotler ve Keller, 2012).

Tele pazarlama yöntemini kullanabilmek için işletmeler genelde bir çağrı merkezine ihtiyaç duyarlar. Bir çağrı merkezi bilgisayarlarla desteklenmiş çalışanların, içe doğru aldıkları veya dışa doğru yaptıkları telefon çağrılarının, otomatik çağrı dağıtım sistemi veya tahmine dayalı çevirme sistemleriyle kontrol edildiği ve işleme sokulduğu bir işlemler bütünüdür (Taylor ve Bain, 1999).

3. UYGULAMA

Bu kısımda ilk önce konu ile ilgili geçmişte gerçekleştirilen çalışmalar ve uygulamanın amaç ve kapsamı paylaşılacak, daha sonra uygulama Çapraz Endüstri Veri Madenciliği Standart Süreci aşamaları takip edilerek gerçekleştirilecek.

3.1. Uygulamanın Geçmişi

Veri madenciliği veya veri tabanından bilgi keşfi ilk olarak 1989 yılında tanımlanmıştır ve o günden günümüze pazarlama ile etkileşimli olarak ilerlemiştir. Bu alanda karşılaşılan bazı emsal çalışmalar burada paylaşılacaktır.

Emel ve arkadaşları (2005) bir işletmenin satış hareketleri içeren veri tabanını kullanarak ayrıntılı bir satış analizi yapmıştır. Ulaş ve arkadaşları (2002) GİMAT A.Ş.'nin verilerini kullanarak birliktelik kurallarıyla yeni örüntüler ortaya çıkarmışlardır. Ngai ve arkadaşları (2008) veri madenciliğiyle müşteri ilişkileri yönetiminin geniş bir alan yazın taramasını ortaya koymuşlardır. Akbulut (2006) bir kozmetik firmasının müşteri bölümlemesini gerçekleştirmiştir. Timor ve Şimşek (2008) sepet analiziyle tüketici davranışını modellemişlerdir. Birant ve arkadaşları (2010) yine bir mağazalar zincirinin satışlarını birliktelik kurallarıyla incelemişlerdir. Giudici ve Passerone (2002) birliktelik kurallarının tüketici davranış araştırmalarında nasıl kullanılacağını anlatmışlardır. Ay ve Çil (2008) mağaza içi yerleştirme sisteminde veri madenciliğinin kullanılmasına yönelik bir çalışma gerçekleştirmişlerdir ve Migros A.Ş.'nin veri tabanında birliktelik kuralları ile bazı analizler gerçekleştirmişlerdir. More ve arkadaşları (2011) ise bu çalışmada da faydalandığımız ve kendilerinin oluşturdukları veri tabanı üzerinde gerçekleştirdikleri çalışmada birliktelik kuralları, nitelik seçimi ve en başarılı modelin seçimini *R Miner* adlı yazılımı kullanarak gerçekleştirmişlerdir. Bu alandaki ilgili yayınlarda ve uygulamada veri madenciliğinin, pazarlama disiplinine özellikle birliktelik kurallarıyla sıkça hizmet ettiğini görmek mümkündür.

3.2. Uygulamanın Amacı ve Kapsamı

Uygulama telepazarlama alanında faaliyet gösteren firmalara örnek bir veri madenciliği süreci gerçekleştirmek amacıyla gerçekleştirilmiştir. Alanda gitgide yaygınlaşan bir yazılım olan *WEKA* da uygulama gerçekleştirilmiş ve birliktelik kuralları yoluyla veriler analiz edilmiştir. Bu kapsamda *PredictiveApriori* algoritması tercih edilmiştir. Uygulama kapsamında kullanılan veri tabanı Portekiz'de bulunan bir bankadan ikincil veri olarak temin edilmiştir.

3.3. Uygulamanın CRISP-DM Yöntemine Göre Gerçekleştirilmesi

Uygulama *CRISP-DM* yönteminin gerektirdiği aşamalara uyularak gerçekleştirilmiş ve her bir aşamadaki işlem detaylı olarak açıklanmıştır. Tabii bu aşamada işi kavrama ve yayılım gibi basamaklar yapılan bu çalışmanın gerçek bir iş süreci olmadığından ve gerçek hayata uygulanacak iş hedefleri bulunmadığından daha çok teorik bir çerçevede gerçekleştirilmiştir.

3.3.1. İşi Kavrama

1970’li yıllarda doğan tele pazarlama günümüzde internetten (genel ağ) alışverişin yaygınlaşmasıyla giderek önemini kaybetmektedir (Chen vd., 1996). Son dönemde müşteriler istedikleri ürün veya hizmete çok hızlı bir şekilde ulaşabildikleri gibi ürün ve hizmetler hakkında bilgilere de kolayca ulaşabilmektedir. Dolayısıyla kendilerine ulaştırılan bu çağrılar artık rahatsız edici bir hal almıştır. Amerika Birleşik Devletleri’nde bu öyle bir boyuta ulaşmıştır ki, 2003 yılında *Ulusal Arama Engelleme Kayıt Merkezi* kurulmuştur ve bu merkeze 2009 yılına gelindiğinde yaklaşık 190 milyon insan kayıt yaptırmıştır. Hem rahatsız edici yapısı hem de yüksek maliyeti düşünülerek tele pazarlama yoluyla ulaşılan müşteri kitlesini azaltmak ve başarıyı arttırmak burada esas alınmıştır. Müşterilerin nitelikleri ve kampanya teklifi arasındaki bilinmeyen gizli ve bankaya ışık tutacak bilgilerin ortaya çıkarılması ve performansı arttırmaya yarayacak doğrudan pazarlama stratejileri için ipuçları oluşturmak hedef olarak belirlenmiştir.

3.3.2. Veriyi Kavrama

Uygulama kapsamında kullanılan veri tabanı *UCI Makine Öğrenmesi ve Akıllı Sistemler* isimli veri tabanı deposundan sağlanmıştır. Kaynak birçok bilimsel çalışmada kullanılan veri setlerini barındırmakta ve Amerika Birleşik Devletleri’nde bulunan *Milli Bilim Vakfı* Kurumu tarafından (*The National Science Foundation*) desteklenmektedir (Olson ve Delen, 2008). Aynı veri tabanı kullanılarak daha önce bir çalışma gerçekleştirilmiştir. Moro ve arkadaşları (2011) yılında bu veri tabanını kullanarak veriler hakkında en iyi tahmin edici modelleri tartışmış daha sonra en anlamlı nitelikleri araştırmıştır. Veri tabanındaki veriler Portekiz’deki bir bankanın doğrudan pazarlama kampanyaları hakkındadır. 17 farklı kampanya telefon görüşmeleriyle müşterilere ulaştırılmıştır. Sıklıkla bir müşteriyle teklifi ulaştırabilmek için birden fazla kez görüşüldüğü görülmektedir. Toplam 79354 telefon görüşmesi yapılmıştır. Bu görüşmeler sırasında uzun vadeli ve iyi bir faiz oranı olan mevduat önerilmiştir. 45211 örneklem ve 16 nitelik (ve bir de sonuç niteliği) barındıran veri tabanında herhangi bir kayıp veriye rastlanmamıştır. Burada kullanılan nitelikler aşağıda açıklanmıştır:

1. Yaş: Müşterilerin yaşı sayısal olarak mevcuttur.
2. Meslek: Meslekler gruplar halinde veri tabanına girilmiştir; Üst kademe yönetici, bilinmeyen, işsiz, yönetim, müstahdem, girişimci, öğrenci, mavi yakalılar, serbest meslek erbabı, emekli, teknisyen, hizmet sektöründe çalışan şeklinde seçenekler mevcuttur.
3. Evlilik durumu: Evlilik durumu da gruplar halinde bulunmaktadır; evli, boşanmış ve bekâr. Boşanmış dul olan müşteriler için de kullanılmıştır.
4. Eğitim: Bilinmeyen, orta dereceli, ilk dereceli, üçüncü dereceli olarak gruplar halinde belirtilmiştir.

5. Yükümlülüğü yerine getirilmeyen kredinin olup olmaması: Evet ve hayır olarak ikili terim olarak ifade edilmiştir.
6. Bakiye: Yıllık ortalama bakiye sayısal olarak ve Avro (Euro) para birimine göre ifade edilmiştir.
7. Konut kredisi: Evet veya hayır olarak ikili terim olarak ifade edilmiştir.
8. Bireysel kredi: Evet veya hayır olarak ikili terim olarak ifade edilmiştir.
9. İletişim: İletişim kurma tarzı da 3 değişkenden oluşan grup halinde ifade edilmiştir; bilinmeyen, cep telefonu, sabit telefon.
10. Gün: Ay içerisinde en son görüşülen gün sayısal olarak ifade edilmiştir.
11. Ay: Yıl içerisinde iletişim kurulan en son ay; kategorik olarak 12 aydan oluşmaktadır.
12. Süre: En son iletişim süresi sayısal olarak saniye esasına göre ifade edilmiştir.
13. Kampanya: Kampanya süresince bu müşteriyle kurulan iletişim sayısı (son görüşme de dâhil olmak üzere).
14. Geçen gün: Müşteriyle bir önceki kampanya sırasında en son ulaşıldığı günden bu yana geçen gün sayısı sayısal olarak ifade edilmiştir (-1 daha önce müşteriyle iletişime geçilmediğini ifade etmektedir).
15. Önceden kurulan iletişim sayısı: Müşteriyle bu kampanyadan önce gerçekleştirilen iletişim sayısı sayısal olarak ifade edilmiştir.
16. Önceki Çıktı: Bir önceki kampanyanın başarılı olup olmadığı 4 şekilde ifade edilmiştir; bilinmeyen, diğer, başarısız, başarılı)
17. Çıktı-Hedef değişken (y) : Önerilen mevduat teklifin kabul edilip edilmediği ikili olarak evet ve hayır şeklinde ifade edilmiştir.

3.3.3. Veriyi Hazırlama

Bu aşama genelde en zaman alıcı basamaktır (Akbulut, 2006). Çalışmada kullandığımız veri tabanı düzenli bir şekilde bulunduğundan bu aşama kolayca gerçekleştirilmiştir. Veri tabanı dosyası *Microsoft Excel* aracılığıyla bir takım işlemlere tabii tutulmuştur. Bu işlemlerin yeterli gelmediği durumlarda, veri tabanı dosyası *Microsoft Word*'de çalıştırılıp bir takım düzenlemelere tabi tutulduktan sonra tekrar *Microsoft Excel* formatında *WEKA*'da çalıştırılıp en son şekli verilmiş ve *WEKA*'ya uygun (.arff) uzantılı olarak kaydedilmiştir. Çalışma boyunca *WEKA*'nın görselleştirme imkânının yanı sıra *SPSS* programı yardımıyla bazı görselleştirmeler de kullanılmıştır. Birliktelik kurallarının kategorik verilerde çalıştığı ve sayısal verilerde çalışmadığı bilindiği için ihtiyaç duyulan yerlerde sayısal veriler *Microsoft Excel* ve *SPSS* aracılığıyla kategorik hale dönüştürülmüştür.

3.3.4. Modelleme

WEKA'nın bünyesinde barındırdığı *PredictiveApriori* algoritmasının veri tabanı üzerindeki bütün niteliklere uygulanması sonucunda elde edilen sonuçlar Tablo-1'de görüldüğü gibidir.

Tablo 1: Bütün niteliklerin PredictiveApriori ile incelenmesi

```
PredictiveApriori
=====
Best rules found:
  1. job=student y=yes 269 ==> marital=single 258   acc:(0.94665)
  2. job=entrepreneur marital=divorced 179 ==> y=no 164
acc:(0.90195)
  3. job=housemaid marital=divorced 184 ==> y=no 166   acc:(0.89301)
  4. job=unknown marital=divorced 17 ==> y=no 16   acc:(0.89172)
  5. job=unknown marital=married 203 ==> y=no 182   acc:(0.89044)
  6. job=entrepreneur marital=single 238 ==> y=no 211
acc:(0.88653)
  7. job=unknown 288 ==> y=no 254   acc:(0.88495)
  8. job=retired marital=single 108 ==> y=no 95   acc:(0.88287)
  9. job=self-employed marital=divorced 140 ==> y=no 118
acc:(0.8602)
 10. job=unemployed marital=divorced 171 ==> y=no 144   acc:(0.8574)
 11. job=housemaid marital=single 144 ==> y=no 120   acc:(0.85112)
 12. job=student marital=married 54 ==> y=no 44   acc:(0.84226)
 13. job=student marital=divorced 6 ==> y=no 5   acc:(0.82215)
 14. job=unknown y=no 254 ==> marital=married 182   acc:(0.7245)
 15. job=unknown 288 ==> marital=married 203   acc:(0.71582)
```

Algoritmanın bulduğu en iyi 15 sonuç Tablo-1'de sıralanmıştır. İlk sıradaki sonuç incelendiğinde mesleği öğrenci olan ve kampanya teklifini kabul eden müşterilerin %94 gibi yüksek bir doğruluk ile bekâr olduğu görülmektedir. İkinci sıradaki sonuç ise girişimcilik ile uğraşan ve boşanmış olan müşterilerin yüksek oran ile teklifi reddettiği görülmektedir.

Daha sonra eğitim düzeyi ile kampanya başarısı arasındaki ilişki yine *PredictiveApriori* ile incelenmiş ve Tablo-2'de sıralanan sonuçlar elde edilmiştir.

Tablo 2: Eğitim düzeyi ile kampanya başarısı arasındaki ilişkinin PredictiveApriori ile incelenmesi

```
PredictiveApriori
=====
Best rules found:
  1. education=primary 6851 ==> y=no 6260   acc:(0.55752)
  2. education=secondary 23202 ==> y=no 20752   acc:(0.52894)
  3. education=unknown 1857 ==> y=no 1605   acc:(0.41759)
  4. education=tertiary 13301 ==> y=no 11305   acc:(0.38172)
  5. y=no 39922 ==> education=secondary 20752   acc:(0.31706)
  6. y=yes 5289 ==> education=secondary 2450   acc:(0.23449)
  7. y=yes 5289 ==> education=tertiary 1996   acc:(0.1863)
  8. y=no 39922 ==> education=tertiary 11305   acc:(0.14981)
  9. y=no 39922 ==> education=primary 6260   acc:(0.10672)
 10. education=tertiary 13301 ==> y=yes 1996   acc:(0.0982)
```

11.	education=unknown 1857 ==> y=yes 252	acc:(0.08969)
12.	y=yes 5289 ==> education=primary 591	acc:(0.0647)
13.	education=secondary 23202 ==> y=yes 2450	acc:(0.06033)
14.	education=primary 6851 ==> y=yes 591	acc:(0.04043)
15.	y=yes 5289 ==> education=unknown 252	acc:(0.03878)
16.	y=no 39922 ==> education=unknown 1605	acc:(0.02736)

Tablo-2 incelendiğinde oldukça başarısız bir birliktelik analizinin gerçekleştiği görülmektedir. Elde edilen bulgular istenilen doğruluk seviyesinde olmadığı gibi kampanyanın örneklem içerisinde büyük oranda başarısız olarak gerçekleşmesi analizin sıhhatini bozmuştur.

Takip eden kısımda konut kredisi, bireysel kredi, yükümlülüğü yerine getirilmeyen kredinin olup olmaması ve kampanya başarısı *PredictiveApriori* yöntemiyle incelenmiştir.

İlk sıradaki kurallardan biri; hem konut kredisi bulunan hem bireysel krediden faydalanan hem de kampanya teklifini kabul eden müşterilerin yasal takibe girmiş bir borcunun olmamasıdır. Bu müşteriler bankalar için birinci sınıf müşteri olarak nitelendirilebilir.

Bu üç analiz birlikte incelendiğinde kampanya teklifi niteliğinin ağırlıklı olarak “Hayır” olan örneklerin bulunduğu göze çarpmaktadır. Zira veriler gözden geçirildiğinde müşterilerin büyük çoğunluğunun kampanyayı reddettiği kolaylıkla görülebilir. “Evet” diye müşteri sayısı 5289 kişi iken, kampanyaya Hayır diye müşteri sayısı 39922’dir.

Tablo 3: Konut kredisi, bireysel kredi, yükümlülüğü yerine getirilmeyen kredinin olup olmaması ve kampanya başarısı PredictiveApriori yöntemiyle incelenmesi

PredictiveApriori		
=====		
Best rules found:		
1.	housing=yes loan=yes y=yes 265 ==> default=no 258	acc:(0.97224)
2.	housing=no loan=yes y=yes 219 ==> default=no 211	acc:(0.96263)
3.	default=yes housing=no loan=yes 168 ==> y=no 160	acc:(0.94799)
4.	default=yes housing=yes loan=yes 133 ==> y=no 126	acc:(0.94245)
5.	default=yes housing=no loan=no 212 ==> y=no 197	acc:(0.928)
6.	default=yes housing=yes y=yes 29 ==> loan=no 22	acc:(0.76272)
7.	default=yes y=yes 52 ==> loan=no 37	acc:(0.69879)
8.	housing=no loan=no y=yes 3135 ==> default=no 3120	acc:(0.62926)
9.	housing=no y=yes 3354 ==> default=no 3331	acc:(0.61872)
10.	loan=no y=yes 4805 ==> default=no 4768	acc:(0.61548)
11.	y=yes 5289 ==> default=no 5237	acc:(0.61456)
12.	housing=no loan=no 17204 ==> default=no 16992	acc:(0.61455)
13.	loan=no 37967 ==> default=no 37453	acc:(0.61455)
14.	housing=yes 25130 ==> default=no 24695	acc:(0.61454)
15.	y=no 39922 ==> default=no 39159	acc:(0.61454)
acc:(0.52663)		

Model bu şekilde kurulduğu zaman sonuçların sağlıklı olmayacağına kanaat getirilmiş ve müşteri ile görüşme süresini barındıran niteliğin kategorik olarak yeniden düzenlenmesine ve bu niteliğin başarı kıstası olarak alınmasına karar verilmiştir. Çünkü müşteri ile görüşme süresinin artması veri tabanı üzerinde son gerçekleştirilen çalışmada da ifade edildiği gibi kampanya başarısını arttırmaktadır (Moro vd., 2011). *CRISP-DM* yöntemi izlenirken 1.Şekilde de görüldüğü üzere modelleme aşamasından tekrar veri hazırlama aşamasına dönülmesi gerekebilir. Bu aşamada bu çerçevede tekrar veri hazırlama aşamasına dönülmüş ve verideki konuşma süresi niteliği kategorik olarak *SPSS* yazılımı aracılığıyla yeniden düzenlenmiştir;

0-100 saniye arası konuşmalar *çok düşük*,

100-200 saniye arası konuşmalar *düşük*,

201-300 saniye arası konuşmalar *orta*,

301-400 saniye arası konuşmalar *iyi*,

401-500 saniye arası konuşmalar *çok iyi*,

500 ve daha fazla konuşma süresi *mükemmel* olarak tanımlanmıştır. Bu işlem sonrası yapılan ilk analiz meslek, konuşma süresi ve medeni durum arasındaki ilişkinin Apriori algoritması ile incelenmesi olmuştur.

Tablo 4: Meslek, konuşma süresi ve medeni durum arasındaki ilişkinin PredictiveApriori algoritması ile incelenmesi

```

=== Run information ===

Scheme:          weka.associations.PredictiveApriori -N 100 -c -1
Relation:                               Duration-
weka.filters.unsupervised.attribute.Remove-R1,4-17
Instances:      45211
Attributes:     3
                job
                marital
                Duration_kategorik
=== Associator model (full training set) ===

PredictiveApriori
=====

Best rules found:

 7.  job=entrepreneur   Duration_kategorik=Çok   İyi   78   ==>
marital=married 63   acc:(0.75832)
 8.  job=retired   Duration_kategorik=İyi   252   ==> marital=married
194   acc:(0.75385)
 9.  job=housemaid   Duration_kategorik=Orta   187   ==>
marital=married 140   acc:(0.74076)
10.  job=retired   Duration_kategorik=Çok   İyi   150   ==>

```

```

marital=married 112    acc:(0.73856)
11.  job=unknown    Duration_kategorik=Çok    Düşü    87    ==>
marital=married 65    acc:(0.7341)
12.  job=housemaid  Duration_kategorik=İyi 110 ==> marital=married
81    acc:(0.73127)
13.  job=entrepreneur    Duration_kategorik=Orta    246    ==>
marital=married 177    acc:(0.72419)
14.  job=housemaid    Duration_kategorik=Çok    İyi    53    ==>
marital=married 39    acc:(0.71776)
15.  job=unknown 288 ==> marital=married 203    acc:(0.71365)

```

Sonuçlar arasındaki en iyi 6 adedi gereksiz tekrar barındırdığı düşünülerek elenmiştir. Ancak sonuçlar nispeten daha tatmin edici bulunmuştur. Örneğin; mesleği girişimci olan ve görüşme süresi uzun olan müşterilerin genelde evli olduğu veya evli ev hanımlarının telefonla konuşma sürelerinin genelde uzun olduğu gibi.

3.3.5. Değerlendirme

İkincil kaynaktan sağlanan veriler kullanılarak başlatılan bu veri madenciliği süreci bankaların tele pazarlama konusundaki başarımını arttırmayı hedeflemiştir. Bu doğrultuda birliktelik kurallarıyla gerçekleştirilecek analizlerin kampanya başarısının esas alınarak gerçekleştirilmesi uygun görülmüş ancak mevcut örneklemelerin arasında çok fazla kampanyayı reddeden müşterinin bulunması yeterince faydalı örüntülerin ortaya çıkarılmasını engellemiştir. Yine de yapılan ilk analiz sonucunda müşterilerin meslekleri ve medeni durumlarının kampanya başarısıyla ilişkisini gösteren yoruma açık bazı bulgular elde edilmiştir. Takip eden kısımda müşteri ile konuşma süresi niteliği kategorik olarak yeniden düzenlenmiştir ve analizler bu kapsamda daha aydınlatıcı olmaya başlamıştır.

3.3.6. Yayılım

Süreç gerekli veri düzeltme işlemlerini gerçekleştirmek üzere ve farklı algoritmaların denenmesi suretiyle devam ettirilebilir. Ayrıca konuşma süresi niteliği temel alınarak bazı sınıflandırma ve kümeleme analizlerinin gelecekte gerçekleştirilerek modelin tazelenmesi denenebilir.

4. SONUÇLAR

Veri madenciliği istatistiksel analizden farklı olarak hipotezlerin denendiği bir süreç değildir. Bu kapsamda sonuçları yorumlamak kolay olmadığı gibi süreç tekrar gerektirebilir (Koyuncugil ve Özgülbaş, 2009). Bu çalışmada da veri madenciliğinin doğrudan pazarlamayla etkileşimini göstermek

amaçlanmıştır. Çalışma kapsamında Portekiz’de bulunan bir bankanın müşterileriyle gerçekleştirdiği *tele pazarlama* sonucunda ortaya çıkan veri tabanı kullanılmıştır. Veri tabanı üzerinde *CRISP-DM* yöntemi uygulanmış ve veriler modelleme aşamasında birliktelik kurallarıyla analiz edilmiştir. İlk olarak bankaların müşterilerine sunduğu kampanyanın nasıl sonuçlandığını bildiren nitelik esas alınarak analizler yapılmış ancak burada arzu edilen sonuçlara ulaşılamamıştır. Daha önce aynı veri tabanı üzerinde yapılan çalışmada belirtildiği gibi konuşma süresinin kampanya başarısında önemli bir nitelik olduğu kabul edilmiş ve konuşma süresi kategorik olarak düzenlendikten sonra bu model kıstas alınarak yeni analizlerin yapılmasına karar verilmiştir. Örnek olarak gerçekleştirilen ilk analizde ise girişimci müşterilerin telefonla istenilen düzeyde konuştuğu ve bunların genelde bekâr olduğu gibi tatmin edici ve bankanın pazarlama stratejilerini aydınlatabilecek sonuçlar elde edilmiştir. Çalışma bir bütün olarak günümüz tele pazarlama kullanan banka ve işletmelerine örnek bir süreç ve taslak sunmaktadır.

5. KAYNAKLAR

- AKBULUT, S. (2006). *Veri Madenciliği Teknikleri İle Bir Kozmetik Markanın Ayrılan Müşteri Analizi Ve Müşteri Segmentasyonu*. Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- AY D. ve ÇİL İ. (2008), Migros Türk A.Ş. de Birliktelik Kurallarının Yerleşim Düzeni Planlamada Kullanılması, *Endüstri Mühendisliği Dergisi Özel Sayısı Cilt: 21 Sayı: 2 Sayfa: (14-29)*.
- BAI, Y. (2010) *Practical Database Programming With Visual C#.NET*. U.S.A: IEEE Press.
- CHAPMAN P., CLINTON J., KERBER R., KHABAZ T., REINARTZ T., SHEARER C., ve WIRTH R. (2000). *Step-by step Data Mining Guide*, SPSS Inc..
- CHEN M.S., HAN J., ve YU P.S. (1996), Data Mining: An Overview from a Database Perspective, *Ieee Transactions On Knowledge And Data Engijeering*, 86, (s:866-883).
- ÇANKIRI, S., KARTAL, E., YILDIRIM, K. ve GÜLSEÇEN, S . (2009). Organizasyonlarda Bilgi Yönetimi Sürecinde Veri Madenciliği Yaklaşımı, *Bilgi Çağında Varoluş: “Fırsatlar ve Tehditler” Sempozyumu*, 01-02 Ekim, İstanbul.
- EMEL, G.G. ve TAŞKIN, Ç., (2005), Pazarlama Stratejilerinin Oluşturulmasında Bir Karar Destek Aracı: Birliktelik Kuralı Madenciliği, *Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitü Dergisi*, 7(3), (s:30-59).

- EMEL, G.G. ve TAŞKIN, Ç., (2005), Veri Madenciliğinde Karar Ağaçları Ve Bir Satış Analizi Uygulaması, *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 6(2), (s:221-239).
- EMEL, G.G. ve TAŞKIN, Ç., (2010), Veri Madenciliğinde Kümeleme Yaklaşımları Ve Kohonen Ağları İle Perakendecilik Sektöründe Bir Uygulama, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15(3), (s:395-409).
- FRAWLEY, W.J., PIATETSKY-SHAPIRO ve MATHEUS C.J., (1992), Knowledge Discovery in Databases: An Overview, *AI Magazine* Vol 13, (s:58-72).
- GIUDICI P.ve PASSERONE G. (2002), Data Mining Of Association Structures to Model Consumer Behaviour, *Computational Statistics & Data Analysis*, 38, (s:533-541).
- KOTLER P. ve KELLER K.L. (2012) *Marketing Management*. Upper Saddle River, N.J.: Pearson Prentice Hall.
- KOYUNCUGİL A.S., ve ÖZGÜLBAŞ N. (2009) Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları, *Bilişim Teknolojileri Dergisi*, 2(2) (s: 21-32).
- KÜÇÜKSİLLE E., (2009). *Veri Madenciliği Süreci Kullanılarak Portföy Performansının Değerlendirilmesi Ve İMKB Hisse Senetleri Piyasasında Bir Uygulama*. Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta.
- LİNG C.X. ve Lİ C. (1998). Data Mining for Direct Marketin: Problems and Solutions. *Fourth International Conference On Knowledge Discovery And Data Mining*.(s:73-80). California, U.S.A.
- MORO S., LAUREANO R. ve CORTEZ P.(2011) Using Data Mining for Bank Direct Marketing: An Application of the CRISP-DM Methodology. *Proceedings of the European Simulation and Modelling Conference - ESM'2011*, October, 2011. (s: 117-121), Guimarães, Portugal.
- NAKİP M., VARİNLİ İ., ve GÜLMEZ M. (2012) *Güncel Pazarlama Yönetimi*, Detay Yayıncılık, Ankara.
- NGAI E.W.T., XIU L., ve CHAU D.C.K. (2009) Application of data mining techniques in customer relationship management: A literature review and classification, *Expert Systems with Applications* 36 (s:2592-2602).
- OLSON, D.L. ve DELEN, D. (2008). *Advanced Data Mining Techniques*. Berlin: Springer.

- P. TAYLOR, P. ve BAIN, P. (1999), 'An Assembly Line in the Head?: Work and Employee Relations in the Call Centre', *Industrial Relations Journal* 30(2), (s: 101-117).
- SAVAŞ S., TOPALOĞLU N., ve YILMAZ M. (2012), 'Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri', *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 21, (s: 1-23).
- TİMOR M. Ve ŞİMŞEK U.T. (2008), 'Veri Madenciliğinde Sepet Analizi İle Tüketici Davranışı Modellemesi', *Yönetim*, 59.
- ULAŞ M.A., (2001). *Market Basket Analysis For Data Mining*. Boğaziçi University Institute for Graduate Studies in Science and Engineering, İstanbul.
- ULAŞ M.A., ALPAYDIN E., SÖNMEZ N., ve KALKAN, A. (2001). 'Veri Madenciliğinde Sepet Analizi Uygulamaları', *Bilişim Zirvesi 2001*, TBD 18. Bilişim Kurultayı, 4-7 Eylül, İstanbul.
- WIRTH, R. ve HİPP, J. (2000). 'CRISP-DM: Towards a standard process model for data mining'. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, (s: 29–39). Manchester, UK.
- WITTEN, I., ve FRANK E. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. U.S.A: Morgan Kaufman