

T.C.  
BEYKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI  
ENDÜSTRİ MÜHENDİSLİĞİ BİLİM DALI

**BANKACILIK SEKTÖRÜNDE PAZARLAMA  
STRATEJİLERİNİN BELİRLENMESİNDE  
SINIFLANDIRMA VE VERİ MADENCİLİĞİ**

(Yüksek Lisans Tezi)

Tezi Hazırlayan:

**Özge UYUMAZ**

İstanbul, 2017

T.C.  
BEYKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI  
ENDÜSTRİ MÜHENDİSLİĞİ BİLİM DALI

**BANKACILIK SEKTÖRÜNDE PAZARLAMA  
STRATEJİLERİNİN BELİRLENMESİNDE  
SINIFLANDIRMA VE VERİ MADENCİLİĞİ**

(Yüksek Lisans Tezi)

Tezi Hazırlayan:

**Özge UYUMAZ**

Öğrenci No:

140792012

Danışman:

Yrd. Doç. Dr. Sabahattin Kerem AYTULUN

İstanbul, 2017

## YEMİN METNİ

Yüksek Lisans tezi olarak sunduđum “**Bankacılık Sektöründe Pazarlama Stratejilerinin Belirlenmesinde Sınıflandırma ve Veri Madenciliđi**” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere uygun şekilde tarafımdan yazıldıđını, yararlandıđım eserlerin tamamının kaynaklarda gösterildiđini ve çalışmamın içinde kullanıldıđıkları her yerde bunlara atıf yapıldıđını belirtir ve bunu onurumla dođrularım. 12/01/2017

**Aday: Özge UYUMAZ**



T.C.  
BEYKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZ SAVUNMA SINAVI SONUÇ TUTANAĞI

**Beykent Üniversitesi**  
**Fen Bilimleri Enstitüsü Müdürlüğü'ne,**

Aşağıda tez adı belirtilen yüksek lisans öğrencisi 140892012 no'lu Özge Uyumaz 'ın 12/01/2017 tarihinde yapılan tez savunma sınavı<sup>1</sup> sonucunda 60 dakika süreyle sunduğu ve savunduğu tezi hakkında<sup>2</sup> oybirliği / ~~oyçokluğu~~ ile, BAŞARILI kararı verilmiştir.

Bilgilerinize saygılarımızla arz ederiz.

---

**Anabilim Dalı** : Endüstri Mühendisliği  
**Programı** : Endüstri Mühendisliği  
**Tez Başlığı<sup>3</sup>** : Bankacılık Sektöründe Pazarlama Stratejilerinin Belirlenmesinde Sınıflandırma ve Veri Madenciliği

---

<u>Tez Sınav Jürisi</u>	<u>Öğretim Üyesi</u>
<b>Danışman</b>	: Yrd.Doç.Dr. Sabahattin Kerem Aytulun
<b>Üye</b>	: Doç.Dr. Gül Tekin Temur Arslan
<b>Üye</b>	: Doç.Dr. Oğuzhan Erdiç

İmza



<sup>1</sup> Jüri üyeleri söz konusu tezin kendilerine teslim edildiği tarihten itibaren en geç bir ay içinde toplanarak öğrenciyi tez savunma sınavına alır. Belirlenen günde yapılamayan jüri toplantısı, katılanların hazırladığı bir tutanakla enstitü yönetimine bildirilir. Bu durumda jüri en geç onbeş gün içinde toplanarak adayı tez savunma sınavına alır. Tez savunma sınav süresi en az 45 dakikadır. Yüksek lisans tez savunma sınavı, tez çalışmasının sunulması ve bunu izleyen soru-yanıt bölümlerinden oluşur ve dinleyiciye açıktır. (Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-3)

<sup>2</sup> Tez sınavının tamamlanmasından sonra jüri, tez hakkında "kabul", "düzeltme" veya "red" kararı verir. Jüri başkanı, jüri üyelerince imzalanmış sınav tutanağını, tez sınavını izleyen üç gün içinde ilgili enstitü yönetimine teslim eder. Tezi hakkında düzeltme kararı verilen öğrenci en geç üç ay içinde gerekli düzeltmeleri yaparak ve yönetmelikte belirtilen usullere uygun olarak tezini aynı jüri önünde yeniden savunur. (Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-4)

<sup>3</sup> İleride doğabilecek aksaklıkların engellenmesi için tezin başlığının yazılması gerekmektedir.

## TEŐEKKÜR

Tez alıřmam zarfında, elinden gelen tım yardımı, tavsiye ve yönlendirmelerini yapan, alıřmanın her ařamasında benden, bilgi ve hořgörösünü esirgemeyen danıřmanım Yrd. Do. Dr. Sabahattin Kerem Aytulun'a bir teőekkürü bor bilirim. Eđitim ve iř hayatımın her ařamasında beni destekleyen aileme ve tez süresince destekleri için Ahmet ađlar'a teőekkür ederim.



Adı ve Soyadı : Özge UYUMAZ  
Danışmanı : Yrd. Doç. Dr. Sabahattin Kerem AYTULUN  
Türü ve Tarihi : Yüksek Lisans, 2017  
Alanı : Endüstri Mühendisliği  
Anahtar Kelimeler : Kümeleme Analizi, Özel Bankacılık, Pazarlama Stratejileri

## ÖZ

### **BANKACILIK SEKTÖRÜNDE PAZARLAMA STRATEJİLERİNİN BELİRLENMESİNDE SINIFLANDIRMA VE VERİ MADENCİLİĞİ**

Bankacılık sektörü toplumların; ekonomik ve sosyal gündeminden etkilendiği kadar, toplumun bu konudaki faaliyetlerine de etki edebilen finansal kuruluşlardır. Günümüzde rekabet koşullarının inanılmaz derecede hızlı ve değişken olduğu bankacılık sektöründe, müşterilere yeni pazarlama yaklaşımları gerekmektedir.

Müşterinin beklentisi ve ihtiyaçlarını dikkate alınarak oluşturulan müşteri profillerine yeni pazarlama yaklaşımları ile yaklaşmak gerekmektedir. Müşterilerin finansal hareketlerinin ve tercihlerinin takip ve kaydı ile veriler oluşmuştur. Bu verilerden veri madenciliği teknikleri ile müşterinin profilini tanımak ve anlamlı sonuçlara ulaşmak önem kazanmıştır. Sonuçlar ile banka yeni pazarlama stratejileri belirleyebilir ve müşteri temsilcilerinin zamanını ve eforunu stratejiye uygun şekilde harcamasını takip edebilir.

Name and Surname : Özge UYUMAZ  
Thesis Advisor : Asst. Prof. Sabahattin Kerem AYTULUN  
Thesis Type and Date : Master, 2017  
Study Field : Industrial Engineering  
Key Words : Cluster Analysis, Private Banking, Marketing Strategy

## **ABSTRACT**

### **CLASSIFICATION AND DATA MINING TO DETERMINE MARKETING STRATEGIES AT BANKING SECTOR**

In the banking sector communities; Economic and social agenda, as well as financial institutions that can influence the society's activities in this area. Nowadays, in the banking sector where competition conditions are incredibly fast and variable, new marketing approaches are required for customers.

It is necessary to approach the customer profiles formed by taking into account the expectation and needs of the customer with new marketing approaches. The financial movements and preferences of the customers are monitored and recorded. From these data, it is important to know the data mining techniques and the profile of the client and to reach meaningful results. With the results, the bank can identify new marketing strategies and track customer representatives' time and effort to spend the effort appropriately.

# İÇİNDEKİLER

<b>ÖZ</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>İÇİNDEKİLER</b> .....	iii
<b>TABLolar LİSTESİ</b> .....	vi
<b>ŞEKİLLER LİSTESİ</b> .....	vii
<b>1. GİRİŞ</b> .....	1
1.1. Veri Madenciliği Kavramı.....	1
1.2. Veri Madenciliğinin Tarihsel Gelişimi.....	2
1.3. Veri Madenciliği Kullanım Alanları .....	4
1.4. Veri Madenciliğinin Uygulama Alanları.....	6
1.5. Veri Madenciliğinde Kullanılan Yöntem ve Modeller.....	8
1.5.1. Sınıflama ve Regresyon .....	9
1.5.1.1 Yapay Sinir Ağları.....	10
1.5.1.2 Genetik Algoritmalar .....	10
1.5.1.3 K-En Yakın Komşu .....	11
1.5.1.4 Naive – Bayes Sınıflayıcısı.....	12
1.5.1.5 Lojistik Regresyon.....	12
1.5.1.6 Karar Ağaçları ve Karar Ağaçları Algoritmaları .....	13
1.5.2. Kümeleme Analizi.....	15
1.5.2.1 Kümeleme Metodları .....	17
1.5.2.2 Hiyerarşik Yöntemler .....	18
1.5.2.3 Model Tabanlı Yöntemler .....	19
1.5.3. Birliktelik Kuralları .....	19
1.6. Gelişmiş Veri Madenciliği Teknikleri.....	19
1.6.1. Web Madenciliği .....	20
1.6.2. Metin Madenciliği .....	20
1.6.3. Multimedya Madenciliği .....	21
1.6.4. Büyük Veri .....	21
1.7. Literatür Araştırması .....	22



<b>2. VERİ MADENCİLİĞİNİN SÜRECİ</b> .....	37
2.1. Problemin Tanımlanması .....	38
2.2. Verinin Anlaşılması.....	39
2.3. Verinin Hazırlanması .....	39
2.3.1. Veri Toplama.....	40
2.3.2. Veri Değerleme .....	40
2.3.3. Verileri Bir Araya Getirme ve Temizlemesi .....	40
2.3.4. Veri Seçimi.....	41
2.3.5. Veri Dönüştürme .....	41
2.4. Modelleme.....	41
2.5. Modelin Kullanılması.....	43
2.6. Modelin Değerlendirilmesi.....	43
2.7. Modelin Uygulanması .....	43
2.8. Veri Madenciliğinde Karşılaşılan Problemler .....	44
2.8.1. Veri Tabanı Boyutu .....	45
2.8.2. Gürültülü Veri .....	45
2.8.3. Eksik Veri.....	45
2.8.4. Artık Veri .....	46
2.8.5. Dinamik Veri.....	46
<b>3. BANKACILIK VE FİNANS SEKTÖRÜNDE PAZARLAMA STRATEJİLERİ</b> .....	47
3.1. Pazarlama Kavramı .....	47
3.2. Banka Kavramı.....	49
3.2.1. Bankaların Fonksiyonları .....	49
3.2.2. Dünyada Bankacılığın Tarihsel Gelişimi .....	51
3.3. Strateji Kavramı .....	52
3.4. Pazarlama Kavramı .....	53
3.5. Bankacılıkta Stratejik Pazarlama.....	53
3.5.1. Bankacılıkta Stratejik Planlamanın Tarihsel Gelişimi .....	54
3.5.2. Bankacılıkta Stratejik Planlamayı Etkileyen Faktörler .....	56
3.5.2.1 Banka İçi Faktörler .....	56
3.5.2.2 Banka Dışı Faktörler.....	57
3.6. Banka İle Sunulan Hizmetler .....	58

<b>4. UYGULAMA</b> .....	61
4.1. Arařtırmanın Amacı ve Kapsamı .....	61
4.2. Arařtırmanın Yöntemi .....	61
4.3. Evren ve Örneklem .....	61
4.4. Verileri Anlama ve Deęişkenler .....	62
4.5. Bulgular .....	63
4.6. Bono Deęişkeninin Analiz Çıktısı .....	64
4.7. Repo Deęişkeninin Analiz Çıktısı .....	65
4.8. Sukuk Deęişkeninin Analiz Çıktısı .....	66
4.9. Faiz Geliri Deęişkeninin Analiz Çıktısı .....	67
4.10. Faiz Dıřı Gelir Deęişkeninin Analiz Çıktısı .....	68
<b>SONUÇ</b> .....	70
<b>KAYNAKLAR</b> .....	73
<b>ÖZGEÇMİŐ</b> .....	79

## TABLULAR LİSTESİ

<b>Tablo 1</b> Veri Madenciliğinin Tarihsel Süreci .....	3
<b>Tablo 2</b> Bono% özet tablosu .....	64
<b>Tablo 3</b> Repo% özet tablosu.....	66
<b>Tablo 4</b> Sukuk% özet tablosu .....	67
<b>Tablo 5</b> Faiz% (Faiz Geliri) özet tablo .....	68
<b>Tablo 6</b> Faiz Dışı% (Faiz Dışı Geliri) özet tablosu .....	69
<b>Tablo 7</b> K-Medoids küme özet tablosu .....	70
<b>Tablo 8</b> Küme özellik yorumları .....	71



## ŞEKİLLER LİSTESİ

Şekil 1	Yapay Sinir Ağlarının Katmanları.....	10
Şekil 2	Karar Ağacı Örneği.....	15
Şekil 3	Kümeleme Örneği.....	16
Şekil 4	CRISP-DM Süreci Kaynak Clay Helberg, Data Mining with Confidence, SPSS, 2002.....	38
Şekil 5	Veri Madenciliği Süreci.....	44
Şekil 6	Veri seti görüntüsü.....	63
Şekil 7	K-Medoids Algoritması Sonuçları (Merkez tipleri).....	64
Şekil 8	Bono% değişkeni scatter grafiği.....	64
Şekil 9	Repo% değişkeni scatter grafiği.....	65
Şekil 10	Sukuk% değişkeni scatter grafiği.....	66
Şekil 11	Faiz% (Faiz geliri) değişkeni scatter grafiği.....	67
Şekil 12	Faiz Dışı% (Faiz Dışı Gelir) değişkeni scatter grafiği.....	68

# 1. GİRİŞ

## 1.1. Veri Madenciliği Kavramı

Veri madenciliğinin ortaya çıkışı veri yığınlarının geniş yer kaplamasına ve büyük miktardaki verilerin yararlı bilgilere dönüştürülmesi ihtiyacına dayanmaktadır (1). Günümüzde veri depolama teknik ve sosyal açıdan çok kolay bir hale gelmiştir. İnternet, kişilere yığınla bilgi sunmakta aynı zamanda da bütün tercihlerimizi kaydetmektedir. Ticari ürünlerin tamamında barkodların kullanılması, devletlerin ve işletmelerin bilgisayarlaşması bu kayıt işlemini kolaylaştırmakta ve milyonlarca veri tabanının oluşmasına sebep olmaktadır. Oluşan bu veri tabanları güçlü ve düşük maliyetli veri tabanı sistemleriyle desteklenmektedir. Bunca veriyi işlemek ve anlamlı bir hale dönüştürmek, yeni teknikler ve araçlara duyulan ihtiyacı oluşturmuştur. Bahsedilen yeni süreç veri madenciliğinin temellerini oluşturmuştur.

Veri madenciliğini açıklamak için yapılan çeşitli tanımlar şu şekildedir:

Veri madenciliği, veri tabanları veya veri ambarlarında yer alan yığın veri içindeki gizli örüntüleri ve ilişkileri bulmak için istatistiksel algoritmaları ve yapay zekâ yöntemlerini kullanan karmaşık bir veri arama yeteneği olarak tanımlanabilir. Veri madenciliği; aynı zamanda bilgisayar bilimini, makine öğrenmesini, veri tabanı yönetimini, matematiksel algoritmaları ve istatistiği birleştiren disiplinler arası bir alandır (2). Veri madenciliği, büyük veri tabanlarından, yararlı bilgilerin otomatik, kullanıma ve analize hazır olarak çıkarılması sürecidir.

Veri madenciliği, büyük miktardaki veriden, anlamlı örüntüler ve kurallar keşfetme sürecidir (3). Veri madenciliği, istatistiksel ve matematiksel teknikler ile örüntü tanıma teknolojilerinin kullanılarak, depolama ortamlarında sıkışmış bulunan büyük miktardaki verinin elenmesi ile anlamlı yeni korelasyon, örüntü ve eğilimlerin keşfedilmesi sürecidir (4). Veri madenciliği, büyük veri depolarındaki yararlı bilginin otomatik olarak keşfedilmesi sürecidir (5).

Veri madenciliği, genellikle büyük ölçüdeki veri setlerindeki, bazı bilinmeyen veya gizli kalmış kuralların keşfine ve analizine yarayan yöntemler ve teknikler kümesidir. Kısaca veri madenciliği, veriden bilgi çıkarma sanatıdır (6). Veri

madenciliđi, tek başına yığılmıř ve ham durumdaki verinin sunamadığı bilgiyi ortaya çıkaran veri analizi sürecidir.

Veri madenciliđi, istatistik, veri tabanı teknolojisi, örüntü tanıma, makine öğrenme ve diđer alanlarla ilişkili olan bir disiplindir. Önceden tahmin edilemeyen ilişkileri bulmak için büyük veri tabanlarının ikincil analizi ile ilgilidir.

Veri madenciliđi, önceden bilinmeyen, gizli, anlamlı ve yararlı örüntülerin, büyük ölçekli veri tabanlarından otomatik biçimde elde edilmesini sađlayan, veri tabanlarındaki özbilgi keřif ve analiz sürecidir. Veri madenciliđi, büyük miktardaki veriden ilginç bilgi ya da örüntüleri çıkaran süreç veya yöntemi ifade eder.

Veri madenciliđi, büyük miktardaki veri setlerinde saklı durumda bulunan örüntü ve eğilimleri keřfetme işlemdir (7). Veri madenciliđi, yüksek kapasitelere ve yüksek verimlilik ölçümlerine ulaşmak için ihtiyaç duyulan teknolojilerin anahtar bileřenidir (8). Veri madenciliđi, veri ambarlarında yararlı olma potansiyeline sahip, aralarında beklenmedik, bilinmedik ilişkilerin olduđu verilerin keřfedilerek, hem anlaşılır hem de kullanılabilir bir biçime getirilmesine yönelik geliştirilmiř yöntemler topluluđudur (9).

Yukarıdaki tanımlardan da anlaşılabileređi gibi veri madenciliđi, geleceđe ait tahminlerin yapılabilmesi için büyük veri tabanlarındaki anlamlı, yeni ve gizli kalmıř bilgilerin keřfedilerek çeřitli tekniklerle analiz edilmesi sürecidir. Veri madenciliđi, daha büyük bir süreç olarak adlandırılan bilgi keřfi sürecinin bir bölümüdür.

## **1.2. Veri Madenciliđinin Tarihsel Geliřimi**

Bu bölümde veri madenciliđinin tarihsel geliřimi ayrıntılı olarak anlatılacaktır.

**Tablo 1** Veri Madenciliğinin Tarihsel Süreci

1950'ler	<ul style="list-style-type: none"><li>• İlk Bilgisayarlar</li></ul>
1960'lar	<ul style="list-style-type: none"><li>• Veritabanı ve verilerin depolanması</li><li>• Perseptonlar</li></ul>
1970'ler	<ul style="list-style-type: none"><li>• İlişkisel Veri Tabanı Yönetim Sistemi</li><li>• Basit Kurallara dayanan</li></ul>
1980'ler	<ul style="list-style-type: none"><li>• Büyük Miktarda Veri İçeren Veri Tabanları</li><li>• SQL Sorgu Dili</li></ul>
1990'lar	<ul style="list-style-type: none"><li>• Veri Tabanlarında Bilgi Keşfi Çalışma Grubu ve Sonuç Bildirgesi</li><li>• Veri Madenciliği İçin İlk Yazılım</li></ul>
2000'ler	<ul style="list-style-type: none"><li>• Tüm alanlar İçin Veri Madenciliği Uygulamaları</li></ul>

Veri madenciliği, kavramsal olarak 1960'lı yıllarda, bilgisayarların veri analiz problemlerini çözmek için kullanılmaya başlanmasıyla ortaya çıkmıştır. Bu dönemlerde veri taraması, veri yakalanması gibi isimler verilmiş ve bilgisayar yardımıyla gerekli sorgulama yapıldığında istenilen bilginin elde edilebileceği düşünülmüştür (10).

Günümüzde neredeyse her eve bilgisayar girmiş ve internet erişimi hemen hemen her yerden sağlanmaktadır. Disk kapasitelerinin artması, her yerden bilgiye ulaşma olasılığı, bilgisayarların çok büyük miktarlarda veri saklamasına ve daha kısa sürede işlem yapmasına olanak sağlamıştır. Geçmişten günümüze veriler her zaman yorumlanmış, bilgi elde etmek istenmiştir ve bu amaçla yeni donanımlar oluşturulmuştur. Bu sayede bilgi, geçmişten günümüze taşınır hale gelmiştir.

1950'li yıllarda ilk bilgisayarlar sayımlar için kullanılmaya başlamıştır. 1960'larda ise veri tabanı ve verilerin depolanması kavramı teknoloji dünyasında yerini almıştır. 60'ların sonunda bilim adamları basit öğrenmeli bilgisayarlar geliştirebilmişlerdir. Minsky ve Papert, günümüzde sinir ağları olarak bilinen perseptron'ların sadece çok basit olan kuralları öğrenebileceğini göstermişlerdir

(Adriaans ve Zantinge, 1997). 1970'lerde İlişkisel Veri Tabanı Yönetim Sistemleri uygulamaları kullanılmaya başlanmıştır. Bilgisayar konusunda uzman kişiler bununla beraber basit kurallara dayanan uzman sistemler geliştirmişler ve basit anlamda makine öğrenimini sağlamışlardır. 80'lerde veri tabanı yönetim sistemleri yaygınlaşmış ve bilimsel alanlarda, mühendisliklerde vb. alanlarda uygulanmaya başlanmıştır. Bu yıllarda şirketler, müşterileri, rakipleri ve ürünleri ile ilgili verilerden oluşan veri tabanları oluşturmuşlardır. Bu veri tabanlarının içerisinde çok büyük miktarlarda veri bulunmaktadır ve bunlara SQL veri tabanı sorgulama dili ya da benzeri diller kullanarak ulaşılabilir. 1990'larda artık içindeki veri miktarı katlanarak artan veri tabanlarından, faydalı ve anlamlı sonuç çıkarılabilecek bilgilerin nasıl bulunabileceği düşünölmeye başlanmıştır. Bunun üzerine konu ile ilgili çalışmalara ve yayınlara başlanmıştır (11). 1990'larda bilgisayar mühendisleri, geleneksel istatistiksel yöntemlerinin yerine algoritmik bilgisayar modülleri ile veri analizinin değeriendirilebileceğini vurgulayarak, veri madenciliği ismini kullanmışlardır. Bu yıllarda veri tabanlarındaki veri miktarları katlanarak arttığı için, büyük miktardaki veri içinden yararlı bilgilere nasıl ulaşılması gerektiği üzerinde düşünölmeye başlanmıştır ve için ilk yazılım gerçekleştirilmiştir. 2000'li yıllardan itibaren veri madenciliği sürekli gelişmiş ve geniş bir yelpazede uygulama alanı fırsatı bulmuştur. Veri madenciliğinde amaçlanan, büyük miktardaki verilerin incelenmesi olduğu için veri tabanları ile yakından ilişkilidir. Günümüzde yaygın olarak kullanılmaya başlanılan veri ambarları, günlük kullanılan veri tabanlarının birleştirilmiş ve işlenmeye daha uygun durumdaki özetini saklamayı amaçlamaktadır. Günlük veri tabanlarından istenen özet bilgi seçilerek, gerekli ön işlemeden geçtikten sonra veri ambarlarında saklanmaktadır. Hedef doğrultusunda gerekli veriler, veri ambarlarından alınarak veri madenciliği için standart bir forma çevrilmektedir. Veri ambarlarının analizi için "Online Analytic Processing (OLAP)" programları kullanılır. OLAP, çok boyutlu veri analizini sağlamaya odaklanmıştır.

### **1.3. Veri Madenciliği Kullanım Alanları**

Veri madenciliği veri ambarlarının oluşturulduğu hemen hemen her alanda uygulama alanı bulabilmektedir. Bilim ve mühendislik, sağlık ve ilaç, telekomünikasyon gibi çeşitli alanlarda da uygulama alanı bulabilmektedir. İş ve işletmeciliğe yönelik ise genel olarak üç alanda hizmet vermektedir.



- Müşteri profili belirleme: İşletme için en karlı müşterileri alt kümelere ayırma.
- Hedef Belirleme: Rakipler tarafından ele geçirilmiş karlı müşterilerin karakterlerini belirleme.
- Sepet Analizi: Müşterilerin birlikte satın aldığı ürünlerin çapraz satışa yönelik tespit edilmesi.

Bir başka kaynakta ise veri madenciliğinin kullanım alanları aşağıdaki gibidir.

- Pazarlama ve perakendecilik
- Bankacılık
- Sigortacılık
- Perakendecilik
- Borsa
- Endüstri

Son 20 yıldır Amerika Birleşik Devletleri'nde çeşitli veri madenciliği algoritmalarının gizli dinlemeden, vergi kaçakçılıklarının ortaya çıkartılmasına kadar çeşitli uygulamalarda kullanıldığı bilinmektedir. Kaynaklar incelendiğinde veri madenciliğinin en çok kullanıldığı alan olarak tıp, biyoloji ve genetik görülmektedir (11). Bankacılık ve finans sektöründe faaliyet gösteren işletmelerin müşterileri hakkında dolandırıcılık algılama, risk analizi, kredi değerliliği gibi konularında veri madenciliğini kullandıkları görülmektedir.

Sigorta şirketlerinin de hedef kitlesini tanıma, müşterilerini elde tutmada veri madenciliği çözümlerine başvurduğu yükselen bir eğilimdir. Ayrıca yine dolandırıcılık tespiti (fraud detection) kapsamında da sigorta şirketlerinin veri madenciliğine başvurdukları bilinmektedir. Veri madenciliğinin en kolay uygulanabileceği alanlardan birisi ise elektronik ticarettir. E-ticaret kapsamında faaliyet gösteren firmaların veri tabanlarını oluşturmaları, bütün faaliyetlerin elektronik ortamda gerçekleşmesinden dolayı daha kolay olmaktadır. Yine insan kaynakları yönetiminde de veri madenciliği çözümlerinin kullanıldığı görülmektedir. Muhtemel personel devirlerini tahmin etmede kullanılabilir (12).

#### 1.4. Veri Madenciliğinin Uygulama Alanları

Veri Madenciliği uygulamaları aşağıdaki gibi gruplandırılabilir. Bu bölümde ilgili gruplar ayrıntılı olarak açıklanmıştır.

- Bağıntı
- Sınıflandırma
- Regresyon
- Zaman içinde sıralı örüntüler
- Benzer zaman sıraları
- İstisnalar
- Doküman madenciliği

Bağıntı:

Amaç mallar arasındaki pozitif veya negatif korelasyonları belirlemektir. Sepet analizinde müşterilerin beraber satın aldığı malların analizi yapılır. Örneğin, “Bebek maması satın alanların %20’ si sigara da satın alır.” Bebek maması alan müşterilerin, bez de satın alacağını veya sigara satın alanların kibrit de alacağını tahmin edebiliriz ancak otomatik bir analiz bütün olasılıkları göz önüne alır ve mama ile sigara arasındaki gibi otomatik düşünülemez bağıntıları da bulmamızı sağlar.

Sınıflandırma:

“Almanlar BMW’yi tercih eder, İsviçreliler Volvo’yu tercih eder.” Hedeflenen, malın özellikleri ile alıcının tercihlerini eşleştirmektir. Bu şekilde ürünün satışının yapılabileceği müşteri portföyünün ve müşterilerin alabileceği ürün yelpazesini önceden belirleyebilirler. Bu bilgiye ulaşan bir otomobil şirketi geçmiş müşteri hareketlerinin analizi ile yukarıdaki gibi iki kural bulursa örneğin almanların daha çok tercih ettiği restoranlara reklam verirken BMW marka otomobilinin reklamını verebilir.

Regresyon:

“Arabası olan, beş yıldan uzun süredir aynı firmada çalışan, evi olan, kullandığı kredilerinde geç ödemesi bir aydan fazla olmamış, bekar olmayan bir insanın kredi notu 790’dır” gibi bağımlı ve bağımsız değişkenler arasındaki ilişkinin

çıkarmı söz konusudur. Başvuru skorlamada bir finans kurumuna kredi için başvuran kişi ile ilgili kredi riskini notlayan örneğin bir skor hesaplanır. Bu skor kişinin özellikleri ve geçmiş kredi hareketlerine dayanılarak hesaplanır.

#### Zaman İçinde Sıralı Örüntüler:

“İlk üç taksitinden iki veya daha fazlasını geç ödemiş olan müşteriler %60 olasılıkla kanuni takibe gidiyor” gibi sonuçlar elde edilir. Davranış skoru, başvuru skorundan farklı olarak kredi almış ve taksitleri ödeyen bir kişinin sonraki taksitlerini ödeme veya geciktirme davranışını notlamayı amaçlar.

#### Benzer Zaman Sıraları:

“X şirketinin hisse fiyatları ile Y şirketinin hisse fiyatları benzer hareket ediyor” gibi zaman içindeki iki hareket serisi arasında bağıntı kurmayı amaçlar. İki malın zaman içindeki satış miktarlarını örnek verecek olursak, dondurma satışları ile kola satışları arasında pozitif, dondurma satışları ile salep satışları arasında negatif bir bağıntı beklenebilir.

#### İstisnalar ( Fark Saptanması):

Amaç önceki uygulamaların aksine kural bulmak değil, kurala uymayan istisnai hareketleri bulmaktır. Örneğin, “normalden farklı davranış gösteren müşterilerim var mı?” sorusuna cevap aranarak, olası sahtekârlıkların saptanması sağlanabilir. Visa kredi kartı için yapılan CRIS sisteminde bir yapay sinir ağı, kredi kartı hareketlerini takip ederek müşterinin normal davranışına uymayan hareketler için müşterinin bankası ile temasa geçerek müşteri onayı istenmesini sağlamaktadır.

#### Doküman Madenciliği:

Veri madenciliği teknikleri ile yazılı belgeler arasındaki ilişkileri bulmayı hedefler. Dokümanlar arasında ayrıca elle bir tasnif gerekmeden benzerlik hesaplayabilmeyi sağlar. Bu amaçla genellikle otomatik olarak çıkarılan anahtar sözcüklerin tekrar sayısı kullanılır. Doküman madenciliği, “arşivimde veya internet üzerinde bu dokümana benzer hangi dokümanlar var?” gibi sorulara cevap bulmamıza yardımcı olur. Günümüzde yaygın olarak kullanılan internet arama motorları, doküman madenciliğini kolaylaştırmıştır.

## 1.5. Veri Madenciliğinde Kullanılan Yöntem ve Modeller

Veri madenciliğinde kullanılan modeller, tahmin edici ve tanımlayıcı olmak üzere iki ana başlık altında incelenmektedir. Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır. Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılabilecek mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır (13).

Tahmin edici kurgularla örüntü tanıma işi sınıflama, regresyon ve zaman serileri yaklaşımlarını içerir. Modeller yapılacak çalışmalardaki tahminlerin amacına değişkenlik gösterir. Regresyon analizi ile çıktı niteliğini devamlı değerleri, zaman serileri ile zamanın farklılaştırıcı özellikleri ve sınıflama ile ucuz veya pahalı gibi az sayıdaki farklılaştırılmış kategoriye sahip, bir özel veri ögesi tahmini yapılmasında kullanılmalıdır. Tanımlama yapmak için sahip olunan bilgi gruplarını bulan ardışıklık, birliktelik ve kümeleme kurallarını elde etmeyi içeren birliktelik analizi ve ardışıklık keşfi davranışı kullanılır. Bu modeller fonksiyonlarına göre 3 başlık altında birleştirilebilir:

- Sınıflama (Classification) ve Regresyon (Regression)
- Kümeleme (Clustering)
- Birliktelik Kuralları (Association Rules) ve Ardışık Zamanlı Örüntüler (Sequential Patterns)

Tahmin edici (predictive) modeller sınıflama ve regresyon yöntemleridir. Tanımlayıcı (descriptive) modeller Kümeleme ve birliktelik kurallarını kullanmaktadır. Çeşitli kaynaklarda farklı betimlemere rastlamak mümkündür. Örneğin Koyuncugil ve Özgülbaş, veri madenciliği tekniklerini geleneksel teknikler ve yeni teknikler olarak ikiye ayırmışlardır. Burada en başta verilen üçlü sınıflandırma (Sınıflandırma ve regresyon, kümeleme, Birliktelik kuralları ve sıralı örüntüler) kullanılarak kısaca veri madenciliği teknikleri anlatılacaktır.

### 1.5.1. Sınıflama ve Regresyon

Özekeş ve Çamurcu'ya (2002) göre, Sınıflama ve regresyon, önemli veri sınıflarını ortaya koyan veya gelecek veri eğilimlerini tahmin eden modelleri kurabilen analiz yöntemleridir. Sınıflama kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır. Örneğin, bir sınıflama modeli banka kredi uygulamalarının güvenli veya riskli olmalarını kategorize etmek amacıyla kurulurken, regresyon modeli geliri ve mesleği verilen potansiyel müşterilerin bilgisayar ürünleri alırken yapacakları harcamaları tahmin etmek için kurulabilir. Örneğin; bir GSM operatörü belli bir gün içerisinde ve belirlenmiş bazı saatler arasında indirimli veya ücretsiz konuşma kampanyası başlatmak istediğinde, kampanyayı uygulayacağı saatler arasında en çok konuşan müşterilerini tespit etmek isteyebilir. Veri tabanında geçmişte müşterilerin konuşma verilerinin mevcut olduğunu varsayalım; o müşterilerin ne kadar süre telefonla konuştuklarını gösteren alan bağımlı değişken olacaktır. Müşterilerin diğer nitelikleri ise (yaş, cinsiyet, gelir düzeyi gibi) bağımlı değişken üzerinde etkisi araştırılan bağımsız değişken olacaktır. Söz konusu firmanın veri madenciliği uygulamaları sonucunda 24 yaşındaki bayanların o saatlerde daha fazla telefonla konuştuğunu tespit ettiğini varsayalım. Firma başlattığı bu kampanyanın tutundurma çabalarını bu kitle üzerine yoğunlaştırarak daha hızlı ve olumlu sonuçlar alacaktır. Sınıflandırma, bir veri ögesini, önceden tanımlı sınıflardan birine tasnif ederken, regresyon veri ögesini, gerçek değerli bir tahmini değişkene eşler.

Sınıflama ve regresyon modellerinde kullanılan başlıca teknikler:

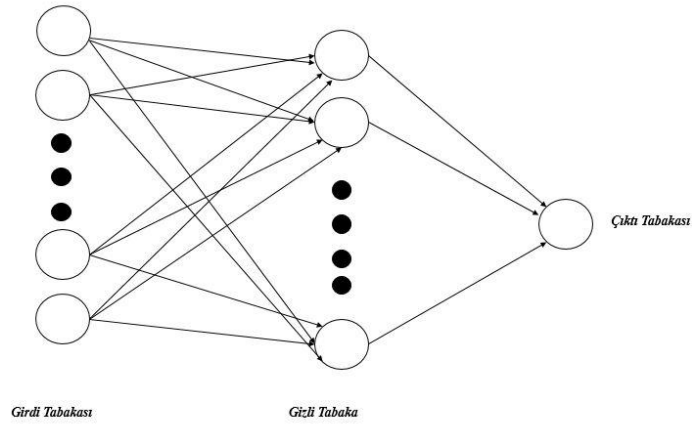
- Yapay Sinir Ağları (Artificial Neural Networks),
- Genetik Algoritmalar (Genetic Algorithms),
- K- En Yakın Komşu (K- Nearest Neighbour),
- Naive - Bayes Sınıflayıcısı,
- Lojistik Regresyon,
- Karar Ağaçları (Decision Trees)

### 1.5.1.1 Yapay Sinir Ağları

Yapay sinir ağları, insan beyninin çalışma şekline benzetilmeye çalışılarak üretilmiştir. İnsan beynindeki birçok sinir hücresinin birbirlerine değişik etki seviyeleri ile bağlanması sonucu oluşan karmaşık bir sistem olarak düşünülebilir. Çok farklı yapıda ve formlarda bulunabilen verileri hızlı bir şekilde tanımlayabilmek ve algılayabilmek için kullanılır.

Yapay sinir ağları öğrenme yeteneğine sahip, gelişmiş matematiksel yapıların hesaplanmasını içeren bir yaklaşımdır. Bu metot sinir sisteminin öğrenmesini model alan akademik araştırmaların bir sonucu olarak ortaya çıkmıştır. Sinir ağları karmaşık ve anlaşılması çok güç olan yapılardan anlam türetme becerisine sahip, dikkate değer yeteneklere sahiptir (14).

Yapay sinir ağları, insanlığın doğayı araştırma ve taklit etme çabalarının en son ürünlerinden bir tanesi olan teknolojidir. 1980'lerden itibaren yaygınlaşan ve yapay sinir ağları adı verilen programlar, basit biyolojik sinir sisteminin çalışma şeklini canlandırmak için tasarlanmışlardır. Biyolojik özelliği taşıdığı için, yani insanın beynindeki sinir ağlarının modelini oluşturabildiği için yapay sinir ağları bir insanın düşünme ve gözlemlemeye yönelik doğal yeteneklerini gerektiren problemlere çözüm üretebilmektedir.



Şekil 1 Yapay Sinir Ağlarının Katmanları (15)

### 1.5.1.2 Genetik Algoritmalar

Genetik algoritmalar, çok değişkenli fonksiyonları optimize etmeyi amaçlayan sayısal bir araçtır. Bu algoritma parametre yerine onların kodlanmış

biçimlerini kullanarak en iyiye ulaşmaya çalışır. Yapay zekânın bir uygulaması olan genetik algoritma, kısa sürede çözümleri ortaya çıkarması bakımından önemli bir tekniktir.

Genetik algoritmalar, çözümlerin kodlanması, uygunlukların hesaplanması, çoğalma, çaprazlama ve mutasyon işlemlerinin uygulanması gibi aşamaları içermektedir. Çözümlerin kodlanması aşamasında, tüm çözümlerin aynı boyutlara sahip bitler dizisi biçiminde gösterilmektedir. Popülasyondaki her üyenin uygunluk değeri hesaplanarak çoğalma aşamasına geçilmekte ve mevcut kuşaktan yeni bir popülasyon yaratılmaktadır. Mevcut gen havuzunun potansiyelini araştırmak için, bir önceki kuşaktan daha iyi nitelikler içeren yeni kromozomlar yaratmak amacıyla çaprazlama operatörü kullanılmakta ve genetik çeşitliliği korumak amacıyla mutasyon işlemi uygulanmaktadır. Tüm bu işlemlerden sonra yeni kuşak oluşturulmakta ve döngü durdurulmaktadır. Genetik algoritmalar problemlerin çözümü için evrimsel süreci bilgisayar ortamında taklit ederler. Çözüm için tek bir yapının geliştirilmesi yerine, böyle yapılardan meydana gelen bir küme oluştururlar. Problem için olası pek çok çözümü temsil eden bu küme genetik algoritma terminolojisinde nüfus adını almaktadır.

Genetik algoritmalar, çizelgeleme, tesis yerleşimi, hat dengeleme, atama ve optimizasyon problemlerinin çözümü ile finans, pazarlama ve üretim gibi alanlarda uygulanmaktadır.

### **1.5.1.3 K-En Yakın Komşu**

K En Yakın Komşu yöntemi, sınıflandırma problemini çözen denetimli öğrenme yöntemleri arasında yer almaktadır. Yöntemde; sınıflandırma yapılacak verilerin öğrenme kümesindeki normal davranış verilerine benzerlikleri hesaplanarak; en yakın olduğu düşünülen k verinin ortalamasıyla, belirlenen eşik değere göre sınıflara atamaları yapılır. Önemli olan, her bir sınıfın özelliklerinin önceden net bir şekilde belirlenmiş olmasıdır. Yöntemin performansını k-en yakın komşu sayısı, eşik değeri, benzerlik ölçümü ve öğrenme kümesindeki normal davranışların yeterli sayıda olması kriterleri etkilemektedir (16).

#### 1.5.1.4 Naive – Bayes Sınıflayıcısı

Bayes sınıflayıcıları istatistiksel sınıflayıcılardır ve bir örneğin belli bir sınıfa ait olma olasılığı gibi sınıf üyelik olasılıklarını tahmin edebilirler. Bayes sınıflaması, Bayes teoremine dayanmaktadır. VM sınıflandırma algoritmalarından olan Bayes, uygulanabilirliği ve hızlı hesaplama performansı ile araştırmacılar tarafından öne çıkan bir algoritmadır. Sınıflandırılacak olayları birbirinden bağımsız olarak ele almaktadır.

Naive Bayes algoritması sürekli veri ile çalışmadığından değişkenler kategorik hale getirilir. Naive Bayes sınıflayıcıları, belli bir sınıf için alan değerlerinin etkisinin diğer alanların değerlerinden bağımsız olduklarını varsayar. Bu varsayım sınıfların şartlı bağımsızlığı olarak adlandırılır. Bu varsayım gereken işlemleri basitleştirmek için yapılmıştır ve bu mantıkla “saf” olarak değerlendirilir.

Naive Bayes, sürekli veri ile çalışmaz. Bu nedenle sürekli değerleri içeren bağımlı ya da bağımsız değişkenler kategorik hale getirilmelidir. Naive Bayes, modelin öğrenilmesi esnasında, her çıktının öğrenme kümesinde kaç kere meydana geldiğini hesaplamaktadır. Bulunan bu değer, öncelikli olasılık olarak adlandırılır. Naive Bayes aynı zamanda her bağımsız değişken / bağımlı değişken kombinasyonunun meydana gelme sıklığını bulmaktadır. Bu sıklıklar öncelikli olasılıklarla birleştirilmek suretiyle tahminde kullanılır (17).

#### 1.5.1.5 Lojistik Regresyon

Lojistik regresyon, bağımlı değişkenin tahmini değerlerini olasılık olarak hesaplayarak olasılık kurallarına uygun sınıflama yapma imkânı veren bir istatistiksel yöntemdir. Lojistik regresyon analizinde üç temel yöntem mevcuttur (18).

- **İkili Lojistik Regresyon:** İkili cevap içeren bağımlı değişkenlerle yapılan lojistik regresyon analizidir. Bir ya da daha fazla değişken ile ikili cevap değişkeni arasındaki bağıntıyı ortaya koyar.
- **Sıralı Lojistik Regresyon:** Cevap değişkenin sıralı ölçekli olduğu durumlarda uygulanan bir yöntemdir. Sıralı ölçekli cevap değişken, en az üç kategoride gözlenen değerler içermelidir.



- **İsimsel Lojistik Regresyon:** Cevap değişkeninin isimsel ölçekli olduğu durumlarda uygulanan bir yöntemdir. Cevap değişkeninin isimsel ölçekli olduğu durumlarda uygulanan bir yöntemdir.

Lojistik Regresyon Analizinin kullanım amacı, istatistikte kullanılan diğer model yapılandırma teknikleri ile aynıdır. En az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı ile bağımsız değişkenler arasındaki ilişkiyi tanımlayabilen ve biyolojik olarak kabul edilebilir bir model kurmaktır. Lojistik regresyon modelleri, son yıllarda biyoloji, tıp, ekonomi, tarım ve veterinerlik ve taşıma sahalarında yaygın olarak kullanılmaktadır. Lojistik regresyon modellerinin yaygın bir şekilde kullanılır hale gelmesi, katsayı tahmin yöntemlerinin geliştirilmesi ve lojistik regresyon modellerinin daha ayrıntılı incelenmesine sebep olmuştur.

#### **1.5.1.6 Karar Ağaçları ve Karar Ağaçları Algoritmaları**

Karar ağaçları, sınıflandırma ve tahmin için sıkça kullanılan bir veri madenciliği yaklaşımıdır. Sinir ağları gibi diğer metodolojilerin de sınıflandırma için kullanılabilmesine rağmen karar ağaçları, kolay yorumu ve anlaşılabilirliği açısından karar vericiler için avantaj sağlamaktadır (19). Karar ağacı analizi, genellikle seçenekler üzerinde yapılan bir analiz türüdür.

Bu analizin veri madenciliğinde kullanılma sebepleri şunlardır;

- Maliyeti azdır.
- Anlaşılması ve yorumlanması kolaydır.
- Veri tabanına kolay entegre edilir.
- Güvenirliği yüksektir.

Karar ağaçları kolaylıkla sınıflama kurallarına dönüştürülebilmektedir. Bunun için algoritmaya girdi olarak verilerin belirlenen belli nitelikleri, çıktı olarak da verilerin belli bir niteliği verilir ve algoritma bu çıktı niteliğindeki değerlere ulaşmak için hangi girdi nitelik değerlerinin olması gerektiğini ağaç veri yapıları kullanarak keşfeder.

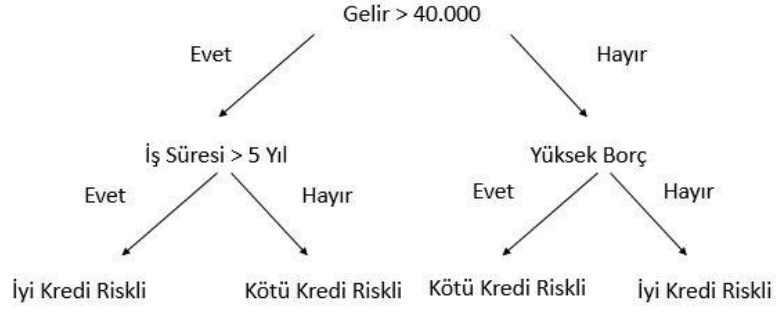
Karar ağaçlarının hedefi bağımlı değişkendeki farklılıkları maksimize edecek şekilde veriyi sıralı bir biçimde farklı gruplara ayırmaktır. Karar ağacı, adında

belirtildiği şekilde ağaç görünümünde bir tekniktir. Karar düğümleri, dallar ve yapraklardan oluşmaktadır (20). Karar ağaçlarının yapısını oluşturan unsurlar;

- **Karar Düğümü:** Veriye uygulanacak test tanımlanır. Her düğüm bir özellikteki testi gösterir.
- **Dal:** Testin sonucunu gösterir. Elde edilen her dal ile tanımlanacak sınıfın belirlenmesi amaçlanır.
- **Yaprak:** Dalın sonucunda bir sınıflandırma elde edilebiliyorsa yaprak elde edilmiş olur.

Karar ağacı tekniğini kullanarak verinin sınıflanması, öğrenme ve sınıflama olmak üzere iki basamaklı bir işlemdir. Öğrenme basamağında önceden bilinen bir eğitim verisi, model oluşturmak amacıyla sınıflama algoritması tarafından analiz edilir. Öğrenilen model, sınıflama kuralları veya karar ağacı olarak gösterilir. Sınıflama basamağında ise test verisi, sınıflama kurallarının veya karar ağacının doğruluğunu belirlemek amacıyla kullanılır. Eğer doğruluk kabul edilebilir oranda ise, kurallar yeni verilerin sınıflanması amacıyla kullanılır. Eğitim verisindeki hangi alanların, hangi sırada kullanılarak ağacın oluşturulacağı belirlenmelidir. Bu amaçla en yaygın olarak kullanılan ölçüm Entropi ölçümüdür. Entropi ölçüsü ne kadar fazla ise, o alan kullanılarak ortaya konulan sonuçlar da o oranda belirsiz ve kararsızdır. Bu nedenle karar ağacının kökünde entropi ölçüsü en az olan alanlar kullanılır.

Karar ağaçlarının en önemli avantajlarından biri EĞER-İSE (IF-THEN) yapısını kullanması ve bilgi kazanımı sunmada anlaşılabilir olmasıdır. Karar ağaçları, bir sınıf ya da değer oluşturan bir dizi kuralı gösterme yöntemidir. Örneğin, borç uygulamalarını iyi ya da kötü kredi riskine göre sınıflandırmak isteyebilirsiniz. Şekil 3' de bu problemi çözen basit bir karar ağacı örneği gösterilmektedir, aynı zamanda bu şekil bir karar ağacının tüm basit bileşenlerini de göstermektedir.



**Şekil 2** Karar Ağacı Örneği (21)

Burada; EĞER Gelir 40000 den küçük ve iş süresi 5 yıldan büyük İSE iyi kredi riskli, EĞER Gelir 40000 den küçük ve iş süresi 5 yıldan küçük İSE Kötü kredi riskli, EĞER Gelir 40000 den büyük ve yüksek borçlu İSE Kötü kredi riskli, EĞER Gelir 40000 den büyük ve yüksek borçlu değil İSE iyi kredi risklidir.

### 1.5.2. Kümeleme Analizi

Kümeleme, veri madenciliğinin temel işlemlerinden biridir. Kümeleme, veri tabanından ilginç örüntülerin keşfedildiği bir madencilik tekniğidir. Kümelemenin genel düşüncesi, veri tabanını çok sayıda kümeye ayırmak ve aynı kümeye ait verilerin mümkün olduğu kadar yakın ilişkide olmalarının sağlanmasıdır. Kümeleme basit olarak veri setlerini alt setlere parçalama işlemidir. Her alt set bir kümedir ve bir kümedeki bütün nesnelere birbirine benzerdir. Bu kümeler bir kümeleme işleminden doğmuştur. Bu bağlamda farklı kümeleme teknikleri farklı aynı veri seti üzerine uygulandığında farklı kümelemeler oluşturacaktır. Parçalama işleme bireyler tarafından değil algoritmalar tarafından gerçekleşir. Bundan dolayı kümeleme veriler arasındaki önceden bilinmeyen grupları ortaya çıkarabileceğinden kullanışlıdır.

Kümelemenin uygulama alanı çok geniştir. Kümeleme, müşteri segmentasyonu ve dolandırıcılık tespiti gibi problemlerin çözümünde yaygın biçimde kullanılır. Kümeleme uygulamalarında üç görevi yerine getirmiş oluruz (22).

- Veri setlerini kümeler içinde bölümlere ayırma,
- Kümeleme sonuçlarını doğrulama,

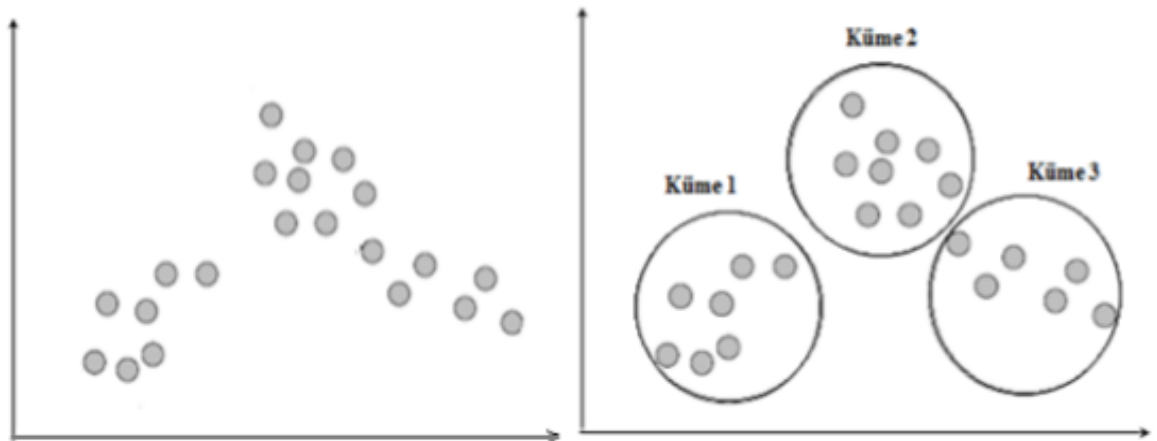
- Kümeleri yorumlama.

Kümeleme analizinin sonucunda bir kümeyi var eden veriler ya kendi içinde benzer özellikler gösterirken, diğer kümeleri var eden veriler ile benzer özellikler göstermeyeceğinden, her küme kendi içinde homojen dağılırken, kümeler arasında heterojenlik söz konusu olmaktadır. Analiz sonrası ortaya çıkan kümelerin çok boyutlu uzay gösteriminde, eğer kümeleme analizi başarılı ise küme içinde verilerin dağılımının birbirine yakınsaması, ancak birbirinden ayrı kümelerin gözle görülür şekilde birbirlerinden ayrı ve uzak görünmesi beklenir. (23).

Kümeleme analizi sonrası veriler, küme özelliklerini temsil eden bir temsilciye sahip olur. Kümeleme, bir denetimsiz öğrenme yöntemidir. Sınıflama tekniğinden farkı, sınıflama öncesinden belirlenen çıktılarına göre olduğu için veriyi denetimli bir şekilde öğretir. Kümeleme analizinin özellikleri aşağıda kısaca özetlenmiştir:

- Geçmişte belirlenmiş sınıf ve sınıf-etiketli öğrenme örnekleri ile çalışmaz.
- Kümeleme veri dağılımını anlamada fayda sağlar.

Kümeleme için basit örnek Şekil 3'de verilmiştir. Kümeleme öncesinde verilerin dağılımı ve kümeleme sonrası ortaya çıkan üç ayrı küme şeklinde gösterilmektedir.



**Şekil 3** Kümeleme Örneği

Literatürde çok sayıda kümeleme algoritması mevcuttur. Analizde kullanılması planlan algoritma seçimi, analiz amacına ve veri tipine bağlı olarak değişkenlik gösterebilir. Genel olarak başlıca kümeleme metodları, bölünmeli, hiyerarşik ve model tabanlı olarak üçe ayrılabilir.

### 1.5.2.1 Kümeleme Metodları

Bölünmeli kümeleme algoritmaları, k giriş parametresini alarak n tane nesneyi k tane kümeye böler. Bu teknikler, dendogram gibi iç içe bir kümeleme yapısı üzerinde çalışmak yerine tek-seviyeli kümeleri bulan işlemler gerçekleştirir. Tüm teknikler algoritmalar sonucu çıkan merkez noktanın ait olduğu kümeyi temsil edebilmesini amaçlar. K-ortalamar (k-means) ve k-medoids birer bölünmeli kümeleme algoritmasıdır. Veri madenciliğinde kümelemede kullanılan ve en çok bilinen uygulamalardan biri K-ortalama'dır. Öncelikle, K sayıda gözlem N gözlem içinden küme sayısına göre rastgele seçilir ve ilk kümelerin merkezi olur. İkinci olarak, kalan her bir N-K gözlem için Öklid uzaklık cinsinden en yakın küme bulunur. Her gözlem en yakın kümelere atandıktan sonra, kümenin merkezi yeniden hesaplanır. Son olarak, tüm gözlemler dağıtıldıktan sonra, gözlemler ile kümenin merkezi arasındaki Öklid uzaklık hesaplanarak en yakın kümeye atanıp atanmadığı tespit edilir. Kümelemede uygulanan K-means algoritması pek çok araştırmada kullanılmıştır.

K-means sonuçlarının değerlendirilmesinde genellikle, hata kriteri olan SSE kullanılır. Minimum SSE değere sahip olan küme en doğru sonucu verir. Verilerin ait oldukları kümeye olan uzaklık aşağıdaki formül ile ölçülür.

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} dist^2(m_i, x)$$

K-medoids algoritması, k-means yönteminin gürültülü ve aşırı verilere aşırı hassasiyeti sebebi ile 1987 yılında Kauffman ve Rousseeuw ile tasarlanmıştır. Bir grup veriyi k kümeye bölerken hedeflenen, kendi içinde maksimum derecede benzerlik sağlayıp, kümelerin birbiri arasındaki benzerliği minimum seviyede tutabilmektir. Temsilci veri, diğer verilere olan ortalama seviyedeki uzaklığı mümkün olduğunca minimuma indiren veridir. Temsilci tiplere merkez tipler (centrotype) adı verilir.

### 1.5.2.2 Hiyerarşik Yöntemler

Hiyerarşik yöntemler, kümelerin ardarda bir araya getirilme sürecidir ve bir grup, diğeriyle birleştirildikten sonra, yöntemin devamındaki basamaklarda ayrılması mümkün olmaz. Hiyerarşik yöntemin ağaç diyagramları vasıtasıyla aktarılmasına dendogram adı verilir.(24) Hiyerarşik kümeleme nesnelerin yakınlık ilişkisine göre oluşturulan kümelerden bir ağaç inşa eder. Yöntemin özellikleri aşağıdaki şekilde özetlenebilir.

- Veri tabanını belli özelliklere kümelere dendogramlar vasıtası ile böler,
- Dendogram, ayrılan yapraktan ağacın bütününe doğru yada bütünden yapraklara doğru olabilir. Ağaç kullanıcının uygun gördüğü aşamada kesilerek küme oluşumu sağlanır.
- Hiyerarşik yöntemler sıra gözetilerek yapılan işlemleri içermektedir.

Böylece başlangıçta belli hesaplamalar yapılmadan bir sonraki adıma geçilemez. Hiyerarşik olmayan yöntemlerde ise böyle bir sıralama düzenine gerek kalmadan işlemler gerçekleştirilmektedir. Buna göre hiyerarşik kümeleme analizinde kümeler küçükten büyüğe doğru belli bir hiyerarşi içerisinde oluşturulmaktadır. Bunun temelinde tek bir mutlak doğru sonucun bulunmadığı varsayımı yatmaktadır. Bu nedenle amaç az ya da fazla sayıda kümelerin elde edilmesi olmaktadır.

Hiyerarşik kümeleme, kayıtları kümelere ayırırken iki farklı yöntem izlemektedir:

- **Birleştirici (Agglomerative) Yöntem:** Bu yöntem her kaydı ayrı bir küme olarak kabul eder ve bu kümelerden(kayıtlardan) tek büyük bir küme oluşuncaya kadar kümeleri birleştirir. Yöntem başlangıçta sadece bir nesneye sahip kümeler ile başka deyişle her kaydı bir küme kabul ederek analize başlar. Daha sonra söz konusu kümeleri benzerliklerine göre birleştirir.

- **Bölünebilir (Divisive) Yöntem:** Bölünebilir yöntem birleştirici yöntemin tam tersi şekilde işler. Başlangıçta tüm kayıtlar tek büyük bir kümenin bir parçası olarak kabul edilir ve daha sonra bu küme iki veya üç alt kümeye bölünür. Her alt küme birbirine benzemeyen alt kümelere bölünerek işlem devam eder. Sonuçta oluşturulan kümeler içerisinde en iyi olanları seçilir (25).

### 1.5.2.3 Model Tabanlı Yöntemler

Model bazlı kümeleme yöntemleri, verilmiş olan bilgi ile bazı sayısal modellerin arasındaki uygunluğu en faydalı hale getirmeye çalışır. Bu yöntemler verinin olasılık dağılımlarının bir karışımından elde edildiğini varsayar. Model tabanlı kümeleme metotları iki yaklaşımdan oluşur:

**İstatiksel Yaklaşım:** Kavramsal kümeleme iki adımlı bir işlemdir: ilk olarak kümeleme yapılır sonrasında bunu tanımlama takip eder. Kavramsal kümelemedeki birçok metot, kavram veya kümelerin belirlenmesinde olasılık ölçümlerini kullanan istatistiksel yaklaşımı benimser.

### 1.5.3. Birliktelik Kuralları

Birliktelik kuralı, geçmiş kayıtlar arasındaki birliktelik eğilimin bulunması ile geleceğe uygun yeni planlar yapılmasını sağlayan çalışma metodudur. (26). Birliktelik kuralları, büyük veri kümeleri arasında birliktelik ilişkileri bulurlar. Zaman geçtikçe saklaması zorlaşan ve devamlı olarak yenisi eklenen veri sebebi ile, firmalar içerideki verileri için birliktelik kurallarını tespit etmek isterler.

Birliktelik kurallarının kullanıldığı standart vaka, market sepeti analizidir. Satın alınan mal veya hizmet sonrasında, daha sonraki alışverişleri için hangi mal ve hizmet tipine eğilimli olduğunun tespit edilmesidir. Bu yöntem ile belirlenen stratejiler doğrultusunda müşteriye daha etkin ve fazla ürün satımı gerçekleşecektir.

## 1.6. Gelişmiş Veri Madenciliği Teknikleri

Gelişen teknoloji ile birlikte son yıllarda ; kurumlar ve oluşumlar veri tabanlarında faaliyetlerine ve işlemlerine göre çok çeşitli türlerde veri depolamaktadır. Ancak bu veriler kullanılıp işlenilmediği müddetçe anlamsız bir yığın olarak sadece veri tabanlarında her geçen gün sayısı artacak şekilde saklanmaya devam eder. Bu verileri kullanılabilir hale getirecek olan yazılımların geliştirilmeye devam edilmesi, kurum ve kuruluşların verilerini anlamlı hale ve arasında gözle görülemeyen ilişkiler kurabilme isteği yeni teknikleri ortaya çıkarmıştır. Kurumlar zaman geçtikçe verilerini kullanılabılır hale uygun bilgi tipi olarak saklamaya başlamıştır. Kurumların elindeki veri işlenmemiş madene benzetilecek olursa, madeni işletmeyen içindeki anlamlara ulaşmayan kurumlar,

rakip firmaların öngörüsüne sahip olamayacakları için gelecekte rekabet etmeleri giderek zorlaşacaktır. Bu kapsamda geleceğe yönelik stratejiler belirlemenin yolu geçmiş verilerdeki ilişkiler ve örüntülerden geçmektedir, veri madenciliği kurumların mevcut durum analizleri gelecek stratejileri için cevherdir.

Veri madenciliği, veriyi bütünsel ve yapısal olarak analiz edebilir. Ancak metin ve web madenciliği yapısal özelliği bulunmayan verinin veri madenciliğinde kullanılacak hale çevrilmesinde aracı olarak kullanılır.

### **1.6.1. Web Madenciliği**

Web madenciliği yöntemleri ile yapısal olmayan web verileri yapısal veriye dönüştürülür. Web madenciliği , yapı içerik ve kullanım olarak üç kısımda incelebilir.

**Web Yapı Madenciliği:** İnternet sayfaları arasındaki, linklerdeki bağıntıyı bulmayı amaçlar.

**Web İçerik Madenciliği:** Amaç içerikte ve internetteki saklı bilgiyi keşfetmektir.

**Web Kullanım Madenciliği:** Amacı kullanıcının site ziyareti esnasında arkasında bıraktığı izlerden, erişim bilgilerinden kullanılabilecek bilgi tipi yaratmaktır.

### **1.6.2. Metin Madenciliği**

Veri her zaman standart formlarda olmaz ve farklı varyasyonlar ile yer alabilir. Veri kayıt ve içerik olarak farklı haldeyse metin madenciliği kullanılır.

Metin yazımında çeşitlilik ve denetim olmadığı için makineler metin yazımını anlayamamaktadır. Dili anlamak için yapısal olmayan öğelerden içerik oluşturmak için çeşitli metodlar vardır. Bu yöntemler, hem sorgudaki hem metindeki kelimelerin karakterlerini karşılaştıran bir temele dayanır. Bundan dolayı içeriği açıklayıcı sonuçlar elde edemez.

Metin Madenciliği uygulamaları iki ana sınıfa ayrılabilir:

**Metni özetlemek:** Yöntemde amaçlanan, metinden anlamlı anahtar kod ve içerik dosyasına atıfta bulunulabilecek bilginin çıkarımıdır.



**Metin aracılığı ile modelleme:** Metinden çıkarılan içerik ve anahtar kodlar modele girdi olarak kullanılır ve geleceğe dair varsayım yapılabilecek bilgi verir.

Metin madenciliğinin kullanım alanı çok geniştir. Bankacılık, sigortacılık sektöründe suistimali önleme alanında, müşteri segmentasyonunda, güvenlik alanında kullanılır. (28).

### **1.6.3. Multimedya Madenciliği**

Multimedya madenciliğinde çoklu ortam verileri bir arada kullanılmaktadır. Çoklu ortam kavramı; bir çok bilginin bir arada bilgisayar ortamında gösterilmesi ve sayısal olarak işlenebilmesidir. (29).

Multimedya madenciliğinin en etkin kullanılmaya başladığı alanlardan biri parmak izi işlemleridir. İnsanların el parmaklarının ucundaki parmak izi denilen kısım her insan için farklı ve bir benzeri olmayacak şekilde yaratılmıştır. Dolayısı ile her bir insan için unique(tek) olduğundan dolayı güvenli bir kodlama ve kilitleme yöntemidir. Her bir insan için parmağındaki izin içerdiği, girinti ve çıkıntılar, eğimler sisteme kaydedilir ve o kişi için eşleştirilir.

Multimedya madenciliğine eğitim alanında da geniş yer verilmektedir. Teknolojinin gelişmesi paralelinde yeni yenilikleri getirmektedir. Günümüzde artık elektronik ansiklopediler ve görsel kitaplar, ders araçları sayesinde eğitim alanında pek çok farklılıklar bulunmakta, multimedya madenciliği sayesinde sözlü anlatımlar, uygulamalar, animasyonlar ile birleştirilerek öğrencilerin öğrenim yetkinliklerini, akademik başarılarını artırmaktadır (30).

### **1.6.4. Büyük Veri**

Büyük veri kavramının geçmişi çok eskilere dayanmamaktadır. 2004 yılında Google'ın kullanmaya başladığı bir teknoloji olan büyük veri, Youtube, Facebook, Twitter gibi büyük ölçekte verisi olan firmalar tarafından da telaffuz edilmeye başlamıştır. Büyük verinin bu kadar dile getirilmesinin tek sebebi bu firmaların bu kavramı kullanması değildir. İnternet, medya ve büyük kuruluşların gözüne çarpan, çarpacak kadar büyüyen veri miktarı ve çeşitliliğidir.

İnternetin kullanımının yaygınlaşması, hayatın içindeki bir çok hizmetin kolayca internet üzerinden erişilebilir olması ve verilerin akması, çok fazla çeşitlilik,

sayı ve ayrıntıda tutulması, saklanması giderek zorlaşan alan darlığı sebebi ile yeni yöntemleri doğurmuş ve “büyük veri” bu çıkış noktasından oluşmuştur.

Büyük verinin uygulandığı alanlara ilişkin örnekler aşağıda sıralanmıştır:

***Sağlık sektöründe;*** Sağlık Kuruluşları; kayıtlı hastalarına yönelik etkin, kişiye özel hizmetler sağlayabilmek adına hasta verilerini bireysel olarak depolamaktadır. Sağlık sektörünün hızlı ve akademik gelişimi için, tuttıkları büyük veri tabanları erişime uygun durumda olmak zorundadır.

***Hükümetler;*** vatandaşları için her türlü devlet kurumu ve diğer kuruluşlardaki bilgileri tutup, işledikten sonra saklamakla yükümlüdürler.

***Finans sektöründe;*** finansal kurumlar, müşterileriyle ilgili olarak toplayıp sakladıkları bilgiler yoluyla, müşterisini tanıyıp, segmente edip, stratejilerini ona göre belirleyebilmek için büyük veri tabanlarına ihtiyaçları vardır. Bu ihtiyaçlardan doğacak şekilde, yeni internet ve mobil şube uygulamaları yazılmış olup, erişimi çok kolay olan bir çok işlem adımını online, hızlı, ulaşılabilir ve güvenli hale getirmiştir.

Sosyal medya, internet, web günlükleri, resimler, fotoğraflar, ses kayıtları, kurum verileri, GSM cihazları gibi çeşitli veriler büyük verinin kaynağı olabilir.

Büyük veri için son yıllarda en çok kullanılan kaynak sosyal medyadır. Birçok kurum sosyal medyadan, internet veya günlüklerden veri çekip, sosyal harita analizleri yapmakta, birçok farklı analizler elde edebilmektedir.

## **1.7. Literatür Araştırması**

CNN'nin yayın akışının her dakikasının aldığı ratingden, IMDB web sayfasının kaç ziyaretçi aldığına, İstanbul'a Eylül aylarında düşen yağmur miktarından, Markette en çok satılan süt markasına, Av Köpekleri Takım Yıldızı'nın parlaklık düzeyindeki değişimden, Seda otomobil satın alanların yaşlarına, pandaların doğurganlık oranından, Everest'e tırmanan dağcı sayısına kadar çevremizde yaşananların çok önemli bir kısmı sürekli olarak kaydedilmektedir. Sayılardan oluşan tüm bu kayıtlar ancak doğru şekilde bakıldığında bir anlam ifade

edilmektedir. Bu nedenle veri madenciliği en geniş anlamı ile yaşadığımız ve kaydettiğimiz olaylara anlam katmaktır.

Veri madenciliğinin uygulama alanlarını bilimsel ve iş dünyası olarak ikiye ayırmak mümkündür.

Bilimsel çalışmalarda veri madenciliği kullanımının ardında yatan sebepler; gelişmiş veri toplama yöntemleri (otomatik istasyonlar, uydu ve uzaktan algılama sistemleri, teleskop taramaları, gen çözümlenmeleri vb.) ile işlenmek üzere ham olarak çok büyük boyutlarda veri toplanması, geleneksel tekniklerin ham verileri işlemede yetersiz kalması ve hipotezler oluşturma, sınıflandırma, karar alma gibi bilimsel çalışma adımlarında bilim insanlarına destek olmasıdır.

İş dünyasında veri madenciliğinin uygulamalarının kullanılmasının temel nedeni; müşteriyi tanıyarak (müşteri gibi düşünerek) müşteri memnuniyeti sağlamak ve bu şekilde rekabet ortamında hızlı ve doğru kararları alabilmektir.

İş dünyasında her alanda veri madenciliği uygulamalarını kullanmak mümkündür ancak sundukları ürün ve servislerle bilgiye dayalı yönetime en fazla ihtiyaç duyan sektörler ve alanlar; Finans (Bankacılık, Sigortacılık), Telekomünikasyon, Pazarlama ve Perakendedir.

İş dünyasında veri madenciliği çalışmalarının büyük bölümü üç temel ihtiyacı karşılamak için kullanılır. Bu ihtiyaçlar; “Kaynakları Daha Verimli Kullanmak”, “Potansiyeli Artırmak” ve “Güvenliği Sağlamaktır”.

Aydoğan ve arkadaşları (2008), bu alanda yaptıkları çalışmalarında, bir kozmetik firmasının sattığı markalardan birinin müşteri gruplarını ve ayrılma eğilimi gösteren müşteri kesitini belirleyerek; bu müşterilere özel pazarlama stratejileri geliştirmeye yönelik bir uygulama gerçekleştirmişlerdir. Kozmetik firması bünyesindeki müşterileri elde tutmak için, ücretsiz cilt bakımı, promosyon ürün dağıtımı, indirimler uygulayarak, pazarlama stratejileri geliştirmişlerdir.. Tüm bu müşteri bölümlendirme çalışmalarını uygulayabilmek için kümeleme tekniklerinden ve ayrılacak müşteri kesitini belirlemek için ise sınıflama tekniklerinden yararlanılmıştır. Çalışmada WEKA paket programı ve Churn Analizi kullanılmıştır.

Churn Analizi, Türkiye’de ilk defa bir kozmetik firmasına uygulanarak, çalışma sonucunda bulunan bulguların mağazalarda uygulanmasında ve müşterilere ilişkin pazarlama geliştirme faaliyetlerinde, satışları ve müşterilerin ayrılma oranlarını nasıl etkilediği ve buna ilişkin yatırımın geri dönüş oranını (ROI) saptamada yol gösterebileceği düşünülmüştür (31).

Risk azaltmak, ürün veya hizmet sunumu sonrasında kayıp yaşama potansiyeli yüksek olan sektörlerde özgü çalışmalardır. En çok bankacılık ve sigortacılık sektörlerinde riskleri değerlendirmek için kullanılır. Bankalar müşterilerine kredi verdiklerinde bir finansal risk öngörürler, kurgulanan risk modelleri ile kredi alanların kredilerini geri ödeyememe ihtimallerini tahmin ederler. Krediyeye karşılık araba veya evin ipotek edildiği durumlarda risk daha düşük olmakla birlikte, kredi kartı bankalar açısından riski en yüksek kredi tipidir.

Kredi riskinin değerlendirilmesi açısından Stefansowski vd. (1997) ve Jacobson vd. (2003) bankacılık sektöründe Veri Madenciliği analizleri gerçekleştirmişlerdir. Geçmiş dönemlerde kredi çeken müşterilerin davranışlarından yola çıkarak tahmine dayalı modeller oluşturulmuştur. Böylece kimlerin kredi ödemesini geciktireceği veya ödemeyi hiç yapmayacakları oluşturulan model tarafından tahmin edilebilmiştir.

Kur ve faiz oranı tahmininde çeşitli modeller kurarak tahminlerde bulunan veri madenciliği çalışmalarından biri Vojinovic vd. (2001) tarafından gerçekleştirilmiştir. Kredi taleplerinin değerlendirilmesi ile ilgili çalışmalar Özekeş (2002), Biçer (2002), Akın (2003) ve Altıntaş (2006) tarafından gerçekleştirilmiştir. Özekeş, Uygulamanın veri setini bir Çek Bankası’nın 1993 – 1998 yılları arasındaki gerçekleştirdiği tüm işlemler oluşturur. Veri seti internetten alınmıştır. Bankanın müşterilerini, onların hesaplarını, işlemlerini, ödeme emirlerini, verilen kredileri ve kredi kartlarını açıklayan 8 tablodan oluşmaktadır. Karar ağacı tekniği kullanılan bu çalışmadaki amaç; veri setini kullanarak hangi müşterilerin aldıkları kredileri sorunsuz ödeyebilecekleri ya da hangilerinin geri ödeyemeyeceklerini tahmin etmektir. Bu sayede yeni bir kredi isteğinde bulunan müşteriye, kredinin verilmesinin uygun olup olmayacağı da tahmin edilebilir. Veri setinde yapılan incelemelerde alınan kredilerin %11’inin geri ödenmediği görülmüştür.

Sahtekârlık riski de bankalar için önemli bir konudur. Kredi kartı kaybolduğunda bankalar kaybedilme sırasında oluşan zararın bir kısmını üstlenmektedirler. Bu dönemde oluşan zararları azaltmak için sahtekârlığı tespit edici sistemler kurgulanmaktadır. Müşterilerin tipik harcama biçimlerini önceden tanımlayıp, harcama eğilimlerinde oluşan ani değişiklikleri tespit etmek ve bu doğrultuda satın alma işlemlerini onaylamayı durdurmak kullanılan yöntemlerden biridir.

Sigorta endüstrisinde risk müşterinin sigortalattığı üründe zararın oluşması ve müşterinin zararının karşılanmasını talep etmesidir. Tüm sigortacılık ürünlerinde fiyatlandırma için riskin doğru belirlenmesi gereklidir. Fiyatlar üzerinde düzenleyici kurulların etkisi nedeniyle, fiyatlandırma kârlılık açısından çok önem taşımaktadır.

Bankacılık ve sigortacılık gibi tahsilatını daha sonra yapmak üzere ürün ve hizmet sunan birçok sektör risk taşımaktadır. Telekomünikasyon şirketleri, enerji şirketleri, perakendeciler, vb. ürünün veya hizmetin ödemesini yapmayacak olan müşteri riskini taşıyan benzer sektörlerdir.

Müşterilerin, şirketin ürünlerini almaktan vazgeçerek rakip şirketin ürünlerini tercih etmeleri birçok endüstride giderek büyüyen bir sorundur. Müşterilerin bir firmadan diğer firmaya geçmesinin en önemli sebebi çoğunlukla daha iyi bir teklif almış olmalarıdır. Örneğin bankalar düşük faiz oranları ile rakiplerinin kredi kartı müşterilerinin kendi kredi kartlarını kullanmalarını sağlamaktadırlar. Kredi kartı kullanımı yaygınlaştıkça yeni kredi kartı müşterileri pazarı oldukça küçülmüştür. Bu da kredi kartı sunan bankaları yeni müşteri bulabilmek için rakiplerinin müşterilerini elde etmek zorunda bırakmaktadır. Bankalar müşterileri çekebilmek için kısa bir dönem için daha düşük kredi oranları sunmakta, oranlar normale döndüğünde müşterilerin bankada kalacaklarını düşünmektedirler. Telekomünikasyon sektöründe en önemli sorun müşteri kaybıdır. Şirketler müşterilerinin rakiplerine geçmesini engellemek için çeşitli pazarlama taktikleri uygulamakta, ürünleri sürekli yeni sunular ile çeşitlendirmektedirler.

Kuruluşlar hangi müşterilerini kaybedebileceklerini önceden belirleyebildikleri durumda, bu müşterilerini elde tutma amaçlı stratejiler geliştirebilirler. Kendi müşterisiyken rakibine giden müşterilerle ilgili analizler

yaparak rakiplerini tercih eden müşterilerinin özelliklerini elde etmek mümkündür. Bundan yola çıkarak gelecek dönemlerde kaybetme olasılığı olan müşterilerin kimler olabileceği konusunda tahminlerde bulunarak onlara özgü ürünler ve servisler geliştirebilirler. Müşteriyi elde tutmak için yapılacak maliyet her zaman için müşteriyi geri kazanmak için yapılacak maliyetten düşük olacaktır.

Timor ve Şimşek (2011) çalışmalarında veri madenciliği tekniklerinden biri olan "Birliktelik Kuralları Sepet Analizi"ni kullanarak Türkiye'de perakende sektöründe faaliyet gösteren büyük bir market zincirinin 2004 yılı Şubat, Mart, Nisan ve Mayıs aylarına ilişkin, 999 tüzel olmayan müşterilerden 627'sine ait 27.924 alışveriş kaydı bulunan verileri incelemişlerdir. Ele aldıkları verileri SPSS Clementine programı yardımıyla analiz ederek müşterilerin satın alma davranışlarını etkileyen değişkenleri karar ağaçları ile tespit etmişlerdir. Çalışmadan elde edilen bilgiler işletmelerin; ürünlere ilişkin reklam kampanya ve promosyonlarının belirlenmesinde önemli yararlar sağlamaktadır.

Timor, Ezerçe ve Şimşek (2011) perakende sektöründe yapmış oldukları bir başka çalışmada; Türkiye'nin en büyük hazır giyim perakende işletmelerinden birinin 2005 yılı Ocak, Şubat, Mart, Nisan ve Mayıs aylarına ilişkin, 5.370 müşterinin 17.164 faturasına ait 42.431 alışveriş veri kayıtlarını incelemişlerdir. Ele aldıkları verilere Birliktelik Kuralları ve Kümeleme Analizi uygulayarak müşteri profillerinin ortaya çıkarılmış, ayrıca müşterilerin satın alma davranışlarını tespit etmişlerdir. Çalışmada Kümeleme analizi ile müşterilerin demografik özelliklerine göre farklıları ortaya koyulmuştur. Çalışmadan elde edilen bilgiler işletmelerin; ürünlere ilişkin pazarlama stratejilerinin belirlenmesinde etkili olmaktadır (32).

Ürün veya hizmet sunumuna kimin yanıt vereceğini tahmin etmek maliyet düşürmek açısından önemli yöntemlerden biridir. Bir ürün veya hizmet ile ilgili bir kampanya programı oluşturmak için hedef kitlenin seçiminden başlayarak bunun hedef kitleye hangi kanallardan sunulacağı kararına kadar olan süreçte veri madenciliği kullanılabilir. Aynı grubun geçmiş davranışlarına dayandırılabilir gibi, mantıksal bir alternatif popülasyonun davranışlarına da dayandırılabilir.

- ***Kampanya şartlarını düzenlemek:*** Düzenlenecek çeşitli kampanyalarda mevcut müşteri kitlesinin seçmek ve bu müşterilerin davranış özelliklerine

yönelik yaklaşımlar geliştirmek. Bu şekilde pazarlama veya perakende kampanyalarına cevap alma oranını artırmak, müşteri ilişkileri yönetimi maliyetlerinin azaltmak hedeflenir.

- **Özel Kampanyalar düzenleme:** Potansiyel müşteriler arasından en karlı olabilecekleri belirleyerek onlara özel kampanyalar uygulamak. En masraflı müşterileri daha masrafsız müşteriler haline dönüştürmek. Örneğin en çok bankacılık işlemi yapanlar ortaya çıkarılıp bunlar şube bankacılığı yerine internet bankacılığına yönlendirmek.
- **Müşterilere özgü satış politikaları düzenlemek:** Aynı karakteristikleri (gelir düzeyi, ilgi alanları, harcama alışkanlıkları, vb.) paylaşan “model” müşteri gruplarını bulmak ve satın alma profillerine göre satış şartlarını ve fiyatları belirlemek.
- **Yeni ürün geliştirmek:** Farklı müşteri gruplarının ihtiyaç duydukları özellikleri belirtip, ihtiyaç duymadıkları özellikleri üründen çıkararak müşterilerin beklentilerini karşılayacak şekilde farklı ürünler geliştirmek.

Ay ve Çil (2010) çalışmalarında veri madenciliği tekniklerinden biri olan Birliktelik Kuralını temel alarak süpermarket yerleşim yeri düzenlemeleriyle ilgili bir uygulama geliştirmişlerdir. Çalışmada Apriori algoritması ve Çok Boyutlu Ölçekleme (ÇBÖ) yöntemlerinden yararlanılmıştır. Türkiye'nin önde gelen perakende satış firmalarından biri olan Migros Türk A.Ş.'deki MMM büyüklüğündeki bir mağazaya ait 01.01.2003 ve 31.05.2003 tarihleri arasındaki alışveriş kayıtları kullanılmış ve toplamda 186.883 adet müşteri kaydı elde edilmiştir.

Çalışmada Birliktelik Kuralından yararlanılarak elde edilen sonuçlar sayesinde marketlerde kullanımı uygun olan, müşterilerin alışverişlerini hiyerarşik bir seviyede yaptığı ızgara tip market düzeni önerisi geliştirilmiş, çok boyutlu karar verme metodlarından ÇBÖ analizi sonuçları doğrultusunda firma için yeni bir yerleşim düzeni önerilmiştir. Önerilen yerleşim düzeninin, müşterilerin davranışlarını etkilediği sonucuna ulaşılmıştır. Çalışmanın sonucunda perakendecilik sektöründe birlikte satılma eğilimi gösteren ürünler ve ürünler arasındaki ilişkiler hakkında önemli bilgiler ortaya koyulmaktadır (33).

Veri madenciliği her alanda olduğu gibi müşteri ve satış verilerinin depolanmasında, veri setleri içinde önceden bilinmeyen anlamlı olabilecek ve değerli

olabilecek verilerin elde edilmesinde, bu verilerin uygun hale getirilmesinde etkili olmaktadır.

Bu alanda Özçakır çalışmasında bir firmanın 7 tane pastane satış lokasyonu verilerini inceleyerek, bu satış verilerinden bir veri tabanı oluşturmuş ve bu verileri veri madenciliği tekniklerinin uygulanabilmesi için uygun hale getirmiştir. Satış verileri üzerinde veri madenciliği tekniklerinden Birliktelik Kuralı tekniği kullanılarak Apriori algoritmasından yararlanılmıştır. Bu algoritmayı kullanabilmek için, .net platformunda web tabanlı bir uygulama yazılımı hazırlanmış, bu uygulama yazılımı algoritmanın her kademesinde yapılan işlemleri adım adım göstermesini sağlayacak şekilde düzenlenmiştir.

Uygulanan Apriori algoritması sonucunda farklı zaman dilimi ve farklı satış lokasyonu girdi değerleri ile ilişkili olarak satış alınan ürünler olduğu görülmüştür.

Aynı ürün grubundaki ürünlerin en sık satın alınan ürünleri oluşturduğu belirlenmiştir. Bu birliktelik kuralı bilgilerinin elde edilmesi sektörde üretim ve pazarlama faaliyetlerinde etkili olmakta, ürünlerle ilgili değişikliklere gidilmesi ve satış lokasyonunda reyon farklılığının oluşturulması gibi durumları gündeme getirmektedir.

Kârlılığını artırmak için mevcut müşterilere satış yapmak, yeni müşteriler bulup onlara satış yapmaktan daha avantajlı bir yöntemdir. Bu nedenle; çapraz satış modelleri ile müşterilerin bir şirketten aldıkları ürünler dışında aynı şirketten ek ürün alma ihtimallerini, ek satış ile müşterilerin aynı ürünü tekrar veya daha çok alma olasılıklarını tahmin eden modeller oluşturulabilir.

- **Çapraz Satış:** Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi, farklı finansal göstergeler arasında gizli korelasyonların bulunması, hangi müşteri profilinin neyi, ne zaman ve neden tercih ettiğini anlayabilen modeller ile ürün satışları arasındaki bağlantı ve ilişkileri bulmak ve bu bağlantılara dayalı tahminler geliştirerek ek ürünler sunmak. Çapraz satış ile birim müşteriye yapılan satış miktarının artırılması, karsız müşteriler karlı hale getirilmesi sağlanabilir. Perakende alanında pazar sepeti analizi ile birlikte satılan ürünlerin bulunması ve buna göre stratejilerin geliştirilmesi en yaygın uygulamalardandır.



- **Ek Satış:** En iyi müşterileri veya müşteri gruplarını bulmak, bulunan bu müşteri gruplarının ihtiyaçları belirleyerek kişiselleştirilmiş ürün ve hizmetler geliştirmek, bu şekilde müşterilerin vazgeçemeyeceği ürün sunularını oluşturmak. Örneğin, yeniden sigorta poliçesi talep edecek müşterilerin tahmin edilmesi.

Firmaların temel amaçlarından biri sürekli olarak yeni müşteriler kazanmaktır. Ancak tüm müşteriler eşit ölçüde kârlı değildir. Firmalar hedef kitlelerini ilk aşamada gelir, yaş, vb. gibi bazı temel ölçütlere göre seçerler ancak seçilen tüm potansiyeller kazanılsalar bile firmaya faydalı olmayabilirler. Firma ile uzun süre çalışmayabilir, tüm ihtiyaçlarını firmadan almayabilir, sürekli olarak farklı teklifleri değerlendirerek başka firmalara geçebilir, alım sıklıkları düşük olabilir veya sadece düşük karlılığı olan ürünleri satın alıyor olabilirler. Tüm bu sebeplerle uzun vadede müşterilerden elde edilmek istenen sonuçlara ulaşamayabilir. Müşterileri elde etmek için cazip teklifler sunmak yüksek müşteri edinme maliyeti yarattığından, doğru potansiyele sahip müşteriye odaklanmak, müşteriden ömür boyu elde edilecek değeri belirlemek önemlidir.

Doğru müşteriye elde etmek için kritik yöntemlerden biri veri madenciliğidir. 3. taraflardan elde edilen başlangıç verilerinden segmentasyon ve sınıflandırmalar ile çeşitli müşteri segmentleri oluşturulabilir. Bu segmentlerden hangilerinin söz konusu ürün veya hizmeti alabileceği tespit edilir. Bu bilgiyi elde etmek için mevcut ve geçmişte alım yapmış müşterilerin alım bilgileri ve özelliklerinin yer aldığı bilgileri kullanılabilir. Alım yapma potansiyeli olan müşteriler belirlendikten sonra, hangi müşterilerin karlı olabileceği belirlenmelidir. Bu aşama da geçmiş verilere gereksinim duyar. Sınıflandırma yöntemleri ile karlı müşteriler belirlenerek ulaşılmaya çalışılır.

- **Müşterilerden ömür boyu elde edilecek getirileri belirlemek:** Müşterilerin firma ile ilişkileri boyunca yaratacakları katma değerleri tahmin etmek ve bu katma değer kategorilerine göre seçim yapmak, yaklaşım biçimleri belirlemek.

- **Kampanya şartlarını düzenlemek:** Düzenlenecek çeşitli kampanyalarda hedef müşteri kitlesinin seçmek ve bu müşterilerin davranış özelliklerine yönelik yaklaşımlar geliştirmek. Bu şekilde potansiyel müşterilere yapılan pazarlama veya perakende kampanyalarına cevap alma oranını artırmak, yeni müşteri edinme maliyetlerini azaltmak hedeflenir.

Veri madenciliği birçok alanda yaygın olarak kullanıldığı gibi sağlık sektöründe de çok etkin bir şekilde kullanılmaktadır. Sağlık sektöründe, elde bulunan büyük veri setleri içinde, klinik ve demografik bilgiler gibi öncelikli verilerin düzenlenmesi zaman ve kaynak maliyeti açısından büyük sorunlara yol açmaktadır. Özellikle sağlık sektöründe; hastalıkların tespitinde, hastanın teşhis tedavi sürecinde, hekimlerin teşhis ve tedaviye ilişkin kararlarını etkileyen verilere erişimin sağlanması aşamasında veri madenciliğinden yararlanılmakta, hastaya ilişkin tüm bilgilerin, laboratuvar sonuçlarının, röntgen, MR gibi görüntü dosyalarının tek kayıt içerisinde düzenlenmesinde veri madenciliğinin büyük önemi bulunmaktadır.

Bu konuyla ilgili olarak Koyuncuğil ve Özgürbaş (2009) çalışmalarında; sağlık sektöründe veri madenciliği konusuna bir altyapı oluşturmakta, veri madenciliği ve veri tabanı kavramları, veri ambarı oluşturma, iş zekâsı ve veri madenciliği yöntemleri konularında bilgiler ve örnekler sunmaktadır. Veri madenciliği özellikle sağlık sektöründeki öncelikli konu ve sorun alanlarına yeni bir bakış açısı kazandırmaktadır (34).

Hastane bilgi sistemleri hastalara ilişkin demografik bilgiler, klinik bilgileri hasta ve tedavi durumları, yapılan tetkikler ve diğer idari bilgileri kapsamaktadır. Bilgi sistemlerinde tıp ve sağlık alanındaki birçok verinin sayısal ortamda saklanması ve bunların düzenlenmesi noktasında klasik sorgulama yöntemleri eksik kalabilmektedir. Bu doğrultuda veri madenciliği yöntemlerinden yararlanılmaktadır. Bu sayede elde edilen bilgiler hem tıbbi araştırmalar hem de yönetim planları için önem taşımaktadır. Sağlık hizmetlerinin sunumunda ortaya çıkan hata ve suiistimal arasındaki farkın ortaya konulması, risklerin minimize edilmesi ve gerekli önlemlerin bu ayrıma göre alınması hasta güvenliği açısından oldukça önemli bir konudur. Büyük hacimli bir veri hesaba katılarak, ardışık bir süreç tasarlanması gerekmektedir. Öncelikle K-ortalamlar Kümeleme Analizi ile 'Normal Değerlerden' ayrılanlar tespit edilerek; ardından, anomali gösteren değerlerin özel bir sağlık durumu mu belirttiği, yoksa bir suiistimale mi işaret ettiğinin belirlenmesi için Karar Ağaçları ve Birliktelik Kuralları yöntemlerinin kullanımıyla elde edilen bulgular, Kümeleme Analizi ile elde edilen 'Normal Değerler' ile kıyaslanarak; anomalinin özel bir sağlık durumuna veya suiistimale işaret ettiği belirlenebilir.

Kararsız kalınan durumlarda ise uzman görüşünden faydalanılarak, analiz süreci güncellenmiş veriyle yinelenir.

Yıldırım ve diğerleri (2008), çalışmalarında dünyadaki sağlık ve tıp alanında yapılan veri madenciliğiyle ilgili çalışmalar hakkında bilgi vermiş ve Hacettepe Üniversitesi Radyoloji Bilgi Sistemleri veri tabanı üzerinde ilişkisel veri madenciliği algoritmalarını kullanarak tıbbi açıdan değerli bilgilerin ortaya çıkarılmasını amaçlamışlardır.

Çalışma sonucunda hem kaynakların daha doğru planlanması, düzenlenmesi, gelecek dönemler için planların uygulanabilmesi amaçlanmakta, hem de hastane bilgi sistemleriyle ilişkili olan radyoloji bilgi sistemi incelenerek klinik çalışmalara yön gösterilmek istenmektedir (35).

Oğuz (2009) çalışmasında metin madenciliği tekniklerini kullanarak; Akdeniz Üniversitesi Hastanesi'nin 2002-2007 yılları arasında Kulak Burun Boğaz Hastalıkları Anabilim Dalı'ndan alınan, ameliyat geçiren 600 hastaya ait hasta bilgi formunu inceleyerek yapılandırmıştır. Yapılandırılmış formata dönüştürülen formlar, hekimler için hasta ile ilgili ihtiyaç duydukları bilgilere ulaşmada ve klinik bilgilerini çıkarmada kolaylık sağlaması amacıyla analiz edilmiştir. Bu analizleri gerçekleştirebilmek için bir yazılım geliştirilmiştir. Çalışmada Birliktelik Kuralı yönteminden yararlanılmıştır.

Önce Microsoft Office Word belge formatında bulunan hasta bilgi formlarındaki veri alanları ön işlemden geçirilerek veri tablosu haline dönüştürülmüştür. Veri tablosundaki veriler sayesinde, hazırlanan metin sorgu formuyla birlikte hekimlerin hasta bilgi formlarında aradıkları özellikteki hastalara erişimleri bu ölçüde kolaylaşmıştır. Ayrıca her alana özgü oluşturulan anahtar kelime listeleriyle metin içerikleri kodlanabilmiş ve bu veriler üzerinde veri madenciliği teknikleri uygulanmıştır (36).

Dinçer ve Duru çalışmalarında veri madenciliği kümeleme tekniklerinden biri olan k-ortalamlar algoritmasını kullanarak bir yazılım geliştirmişlerdir. Bu yazılım sayesinde gırtlak kanseri ameliyat verilerinin geçmiş kayıtlarını analiz ederek, tıp doktorlarının, hasta dosyalarına ilişkin geçmiş kayıtları tek tek incelemelerine gerek duymadan tıbbi bilgilere ulaşmalarını ve elde edilen tıbbi bilgilerle geleceğe yönelik

tahminde bulunabilmelerini sağlamayı amaçlamışlardır. Bunun için Kocaeli Üniversitesi Tıp Fakültesi Hastanesi, Kulak, Burun ve Boğaz Bölümü veri tabanındaki 400 gırtlak kanseri ameliyat verisini ele almışlardır. Çalışmada gerçek veri tabanı üzerinden gidilerek, kullanılan yazılımın uygulanan etkinliği ortaya çıkarılmaya çalışılmıştır.

Çalışma sonucunda gırtlak kanseri için en etkili tedavi yönteminin, benzer hasta sonuçlarından elde edilen bilgilerle yapılan tedavi olduğu belirlenmiştir. Bu sonuç veri madenciliği yöntemlerinin sağlık sektöründe kullanılmasının tıbbi sonuçlara ulaşmada ne kadar kolaylık sağladığını göstermektedir. Ortalama yaşam süresinin artışıyla beraber, kronik hastalıkların görülme sıklığı ve buna paralel olarak getirdiği mali yük giderek artan bir seyir izlemektedir. Bu noktadan hareketle, kronik hastalıkların ortaya çıkmasını engelleyecek proaktif çözümler geliştirilmesi gerekmektedir. Her bir kronik hastalığa yönelik sosyal, ekonomik, demografik, coğrafi vb. tüm değişkenler dikkate alınarak, hastalığın ortaya çıkışında etkisi olan değişkenlerin Önemli Bileşenler Analizi, Faktör Analizi veya Lojistik Regresyon ile belirlenmesi mümkündür. Akabinde, etkisi tespit edilen değişkenlerin, etkin olduğu sınır değerler dikkate alınarak; hastalığın ortaya çıkışına işaret edebilecek risk sinyalleri geliştirilebilir.

Veri madenciliğinin bir diğer uygulama alanlarından biri finans sektörüdür. Finans sektöründe özellikle bankacılıkta veri madenciliği yaygın bir şekilde kullanılmaktadır.

Para ile ilişkili tüm alanlarda sahtekârlık (fraud) riski vardır. Sağlık, finans sektörü ve vergilendirme en çok sahtekârlık denemesine rastlanan alanlardır. Genellikle geçmiş verilere göre sahtekârlık desenleri örnekleri çıkarılabilir ve bu örneklere benzer davranışlar tespit edilerek sahtekârlık denemelerinin önüne geçilebilir. Araştırılması gereken her olay maliyet yükü getirdiğinden veya verilen hizmeti aksattığından, sahtekârlık ihtimallerini belirlerken araştırılması gereken olay sayısını minimumda tutmak esastır. (İstatistik uygulamalarındaki 1. tür ve 2. tür hata tipleri bu konuda önemlidir.) Kara para aklama, kayıp ve kaçakları engelleme, sigorta dolandırıcılıklarının tespiti, kredi kartı dolandırıcılıklarının tespiti, kaçak enerji kullananların profillerini tespit ederek olası kaçak enerji kullanıcılarını tahmin etmek, telefon görüşmelerinin (aranan yer, arama süresi, aranan zaman, vb.)

modellenmesi ve beklenen deęerlerden sapmanın olduęu konuřmaların tespiti, saęlık sigortası uygulamalarında gereksiz veya birbiriyle ilgisiz saęlık testlerinin tespiti, vergi ile ilgili yolsuzlukları ve izlerini belirlemek, aę saldırının tespit edilmesi vb. řeklinde farklı konularda veri madencilięi uygulamaları kullanılabilir.

Albayrak ve Yılmaz(2009) alıřmalarında, İMKB 100 endeksinde sanayi ve hizmet sektrlerinde yer alan 173 iřletmenin 2004–2006 yıllık finansal gstergelerini incelemiřlerdir. İncelemeler sonucunda ele alınan verilere karar aęaları teknięi CHAID algoritması uygulanmıřtır. alıřmada seilen finansal gstergelere gre sanayi ve hizmet sektrlerinde yer alan firmaları ayıran en nemli deęiřkenlerin neler olduęu belirlenmeye alıřılmıřtır. İMKB100 endeksinde faaliyet gsteren firmaları ayırmada etkili olan en nemli deęiřkenlerin karlılık, likidite, varlık kullanım etkinlięi, sermaye yapısı ve iřletme byklęu deęiřkenleri olup olmadıęını incelemiřlerdir (37).

Albayrak(2009) alıřmasında; “Trkiye’deki yerli ve yabancı ticaret bankaların finansal etkinlięe gre veri madencilięi teknikleriyle karřılařtırmıřtır. alıřmada yerli ve yabancı olarak nceden grup yelięi belirlenmiř bankalar eřitli veri madencilięi sınıflandırma teknikleri (diskriminant, lojistik regresyon ve karar aęacı modelleri) ile sınıflandırılmıřtır. alıřmada bankaların likidite, gelir-gider, karlılık ve faaliyet oranları kullanılarak karřılařtırmalar yapılmıřtır. alıřmanın sonucunda; bankaların sınıflandırmasında karar aęacı modelinin, geleneksel diskriminant ve lojistik regresyon modellerine stnlk saęlayabileceęi ve etkili bir sınıflandırma teknięi olarak kullanılabilceęi ortaya koyulmuřtur.

Ařan alıřmasında; kredi kartı kullanan banka mřterilerinin sosyo-ekonomik zelliklerinin gruplandırılmasını incelemiřtir. alıřmada kmeleme analizi teknięi ile kredi kartı kullanan banka mřterilerini sosyo-ekonomik zelliklerine gre  kmede gruplamıřtır. Bu  kmeye gre mřterilerin on adet sosyo-ekonomik deęiřkene gre farklılık gsterdięi belirlenmiřtir. alıřma sonucunda; ilgili mřterilere verilecek bireysel bankacılık hizmetlerinde ne tr mřteriyle karřılařabileceęini bilmek aısından fayda saęlandıęı, ne tr mřterilerin hangi zelliklere ve motivasyonlara sahip olduęunu bilmek aısından da yarar elde edildięi zerinde durulmuřtur (38).

Bilen ise çalışmasında; bankacılık sektöründe çalışan satış personellerinin performanslarının değerlendirilmesi ve kümeleme metodlarından k- means ile çalışanların göstermiş oldukları yıl sonu başarı notu seviyelerini dikkate alarak sınıflandırılmasını incelemiştir. Çalışmada çalışanların demografik,güncel eğitim durumu, çalıştığı pozisyon, çalıştığı lokasyona dair bilgiler ışığında veri madenciliğinde sınıflandırma yöntemleri uygulanmıştır. Uygulanan veri madenciliği teknikleri ile çalışanların performanslarına göre değerlendirilmesi yapılmış, hangi özelliklerdeki personelin hangi şubede ne oranda başarılı olduğu belirlenmiş ve bir takım kurallar oluşturulmuştur. Bu kurallar dikkate alınarak, bir personelin özelliklerine göre hangi şubelere atanabileceği ya da ataması düşünülen şubede hangi düzeyde performans gösterebileceğinin tahmin edilebileceği vurgulanmıştır.

Tosun(2006) çalışmasında; veri madenciliği yöntemlerinden faydalanarak kredi kartı müşterilerinin kaybedilme sebepleriyle ilgili sonuçlara ulaşmayı amaçlamıştır. Çalışmada müşterinin hangi nedenle yollarını ayırdığı bilgisinin yanı sıra, hangi tür müşterilerin daha fazla kaptırma olasılığını tahmin edilmeye çalışılmış, müşterilerin kaybedilme olasılıkları incelenmiştir. Çalışma sonucunda son 3 ayda hiç alışveriş yapmayan ya da hesabı hareketli olmayan müşterilerin kaybedilme olasılığının oldukça yüksek olduğu görülmüştür (39).

Çalış bankacılık sektörüyle ilgili olan çalışmasında; veri madenciliği yöntemlerinden kümeleme ve sınıflandırma ile mevcut bireysel kredi müşterilerinin analizi ve gelecekteki potansiyel müşterilerin ödeme durumlarını incelemiştir. Çalışmada veri madenciliği yazımı olarak SPSS Clementine kullanılarak, mevcut müşterileri kümeleme analizi ile değerlendirilmiş, bireysel kredi müşterilerinin ödeme durumları sınıflandırılarak, gelecekteki potansiyel müşteriler için çıkarımlar yapılmış ve müşteri portföyü oluşturulması amaçlanmıştır.

Bankacılık alanında ise veri madenciliği son derece önemli konulardan biri haline gelmiştir. Bu alanda yapılan çalışmalardan biri de Chye ve Gerry'nin banka müşterilerinin gelecek altı ayda, rakip firmaların önüne geçip geçmeyeceğini tahmin etmeye dayalı yaptıkları uygulamadır. “Churn Analizi” ile yapılan bu çalışma; müşterilerin neden başka bankalarla çalışmak istediğinin tespit edilmesi ile ilgili önemli noktalara değinmektedir. Çalışmada ayrıca Churn analizi modellemeleri sayesinde bankacılık sektöründe müşteri, ciro tahmini yapılmasında veri

madenciliğinin önemi anlatılmakta ve veri madenciliğinin diğer uygulamaları hakkında bilgiler verilmektedir.

Literatür araştırmasındaki konular veri madenciliği çalışmalarının çözüm üretebildiği örneklerdir. Verinin bulunduğu her alanda bir veri madenciliği çalışmasına gereksinim vardır.

Telekomünikasyon sektöründe web madenciliği çalışmaları söz konusudur. Dolgun ve arkadaşları, çalışmalarında bir telekomünikasyon firmasının 2070 müşterisine ait 17 değişkene sahip şirketi terk etme yapısal verisini incelemiş, veriler SPSS Clementine 12. 0 programında yer alan karar ağaçları algoritmalarından C5 kullanılarak analiz edilmiştir. Amaç şirketi terk eden müşterilere ait bir profil modeli oluşturmaktır. Bunun yanı sıra çalışmada müşterilere ait metin dosyaları kullanılarak, var olan yapısal veriye eklemeler yapılarak ikinci model, müşterilerin internet üzerinden ulaşılan web log dosyaları kullanarak üçüncü model oluşturulmaktadır. Yapısal veri ile yapısal olmayan verinin metin ve web madenciliği kullanılarak yapısal hale dönüştürülmesi incelenmiş ve bu model karşılaştırmaları detaylarıyla incelenmiştir.

Web madenciliği metodları aracılığı ile, web sitelerindeki son web teknolojileri gözden geçirilebilmekte ve elde edilen bilgiler dikkate alınarak sitedeki eksiklik, yanlışlıklar tespit edilebilmektedir. Ayrıca web sitesinin gelişimi ve etkinliği ortaya çıkarılabilmektedir. Bu durum web sitelerinin performansını artırmakta, bunun yanı sıra kullanıcılara kullanım rahatlığı sağlanmaktadır.

Özseven ve Düğenci(2011) çalışmalarında web madenciliği yöntemini kullanarak internet kullanıcılarının web sitesi ziyareti sürecince elde ettikleri erişim kayıtlarını analiz ederek, atıl durumdaki verileri bilgiye dönüştürme işlemi gerçekleştirmişlerdir. Bu işlem için Gaziosmanpaşa Üniversitesi'nin kurumsal sitesinin erişim bilgileri incelenmiş, bu kayıtları analiz edebilmek için "Log Yazılım" adlı bir yazılım geliştirmişlerdir. Yazılan kodlar web sitesi için kapsamlı istatistiksel analiz sonuçlarını ortaya çıkarmaktadır. Analiz için Birliktelik Kuralı ve Apriori Algoritması uygulanmıştır (40).

Kalikov çalışmasında e-ticaret amaçlı kurulan Pegem A yayınevinin web sitesinin veri tabanında yer alan verilere veri madenciliği teknikleri uygulayarak

analiz etmiştir. Bu çalışmaya göre Pegem A web sitesinden alınan 6 aylık alışveriş sepet tablosunun ve ona bağlı ürün tablolarının yapıları incelenmiştir. Bu veriler klasik istatistiksel tekniklerin yanı sıra Birliktelik Kuralları ve Karar Ağaçları teknikleri kullanılarak incelenmiştir. Tekniklere ilişkin uygulama sonrasında, veri tabanında bulunmakta olan sanal ürünlerin kategorilerine uygun doğru ve düzenli yerleştirilmesinde faydalı olabilecek bilgilere ulaşılmıştır.

Bu çalışma sonucunda kullanıcıların site üzerinden yapacakları alışverişlerde, hangi kitapları tercih ettikleri, ilgi alanlarına göre kitap tavsiyesinde bulunacak bir SQL sorgulaması geliştirilmeye çalışılmış, bu amaçla birliktelik kuralları tekniğinden yararlanılmıştır. Bunun yanı sıra kişilerin aylık alışveriş bilgilerine göre istatistiksel bilgiler elde edilmiştir.

Günümüzde bilgisayar sistemlerine yönelik saldırılar, bilgi güvenliğinin önemini ortaya çıkarmakta ve her geçen gün tüm sektörlerde bu konuda çeşitli koruma yöntemleri üzerinde çalışılmaktadır. Şirketler açısından bilgi güvenliğinin sağlanması önem taşımaktadır. Bu doğrultuda şirketler çeşitli koruma yöntemlerine başvurmaktadır. Bunun için de güvenlik duvarları, şifreleme sistemleri, açıklık tarayıcı araçları, nüfuz tespit sistemleri, anti-virüs yazılımları gibi birçok koruma yönteminden yararlanılmaktadır.

Sistemlerde oluşan yazılım kaynaklı açıklıklar, eksik bilgilendirme ve yetkilendirme, kullanılan şifrelerdeki zayıflıklar çoğunlukla saldırıların oluşmasına zemin hazırlamaktadır. Oluşan saldırı girişimlerinden başarılı olanlarına ise nüfuz adı verilmektedir. Bu konuda yapılan saldırı önleme çalışmalarında kullanılan nüfuz tespit sistemleri güvenlik duvarının arkasında ağ içerisinden aktiftirler.

Veri madenciliği ile yakından ilişkili olan nüfuz tespit sistemleri kullanıcı davranışları analizine dayanmaktadır. Karabağ ve diğerleri, bu konuyla ilgili yaptıkları çalışmada; kullanıcı davranışlarında veri madenciliği tabanlı nüfuz tespit sisteminde kümeleme ve sınıflandırma yöntemlerinin kullanımıyla ilgili bilgiler vermektedir. Çalışmada kullanıcılara ilişkin bir küme oluşturup ağ trafikleri incelenmiş ve saldırgan olan kullanıcılar belirlenmeye çalışılmıştır.



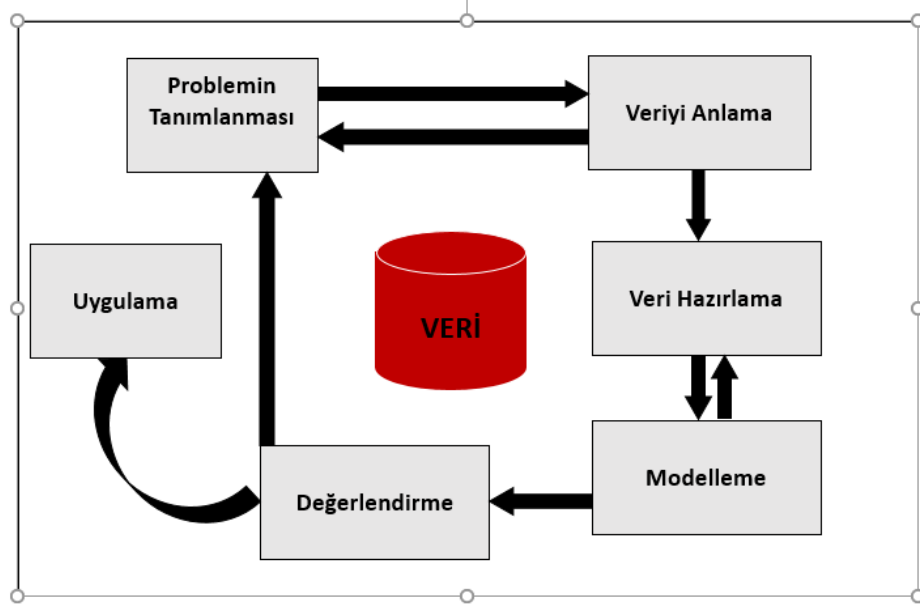
## 2 VERİ MADENCİLİĞİNİN SÜRECİ

Veri madenciliği, kendi içinde bir süreçtir. Veri madenciliğinin bir süreç olarak doğru aktarılabilmesi için, sürecin her kademesi ve akışının gözlem altında dikkatlice gözlemlenmesi gerekmektedir. Çünkü her bir akış kendinden önceki aşamanın sonucu ile şekillenir.

Oluşturulmuş olan bir süreç, veri madenciliği için standartlaştırılmış olarak getirilmiştir. Bu süreç ile ilgili olarak en yaygın kullanılan model CRISP-DM (Cross Industry Standart Process for Data Mining) modelidir (41).

Temel bir takım ortak uygulamaları içermekle birlikte, veri madenciliği süreci başlangıçta farklı uygulamacılar tarafından farklı şekilde yürütülmekteydi. Bu nedenle, uygulamada meydana gelen farklılıkları en aza indirecek bir standart süreç geliştirilmesi ihtiyacı olduğu görülmüştü. Bu ihtiyaç doğrultusunda düzenlenen çeşitli çalıştaylar ile veri madenciliği uygulamalarının gelişmesi ve tüm uygulamacılara yardımcı olacak standart bir süreç modeli oluşturulması amacıyla CRISP-DM konsorsiyumu oluşturulmuştur.

Altı adımdan oluşan CRISP-DM süreci etkileşimli ve yinelemeli bir süreçtir (41).



**Şekil 4** CRISP-DM Süreci Kaynak Clay Helberg, *Data Mining with Confidence*, SPSS, 2002

CRISP-DM, veri madenciliği ile ilgilenen kişi veya organizasyonlar tarafından en çok kullanılan yöntemdir. Bir veri madenciliği sürecinde atılması gereken adımların tümünün detayını açıklamaktadır.

1996 yılında veri madenciliği için sektörler arası standart süreç olarak tanımlanan daha sonraları CRISP-DM adıyla anılan veri madenciliği süreci aşağıdaki aşamalardan oluşmaktadır:

- Problem tanımı,
- Verinin anlaşılması,
- Verilerin hazır hale getirilmesi,
- Modelleme,
- Modelin değerlendirilmesi,
- Modelin uygulanması,

## 2.1. Problemin Tanımlanması

Yapılan çalışma ile ulaşılması planlanan amaç tanımının net olması ve çıkan sonuçların hangi ölççekler ile ölçülmesi gerektiğinin önceden belirli olması, veri madenciliğinde başarıya götürür. Bu kısıtlar ile birlikte, yanlış tahminleme ve hata

durumunda katlanılması gereken maliyet ve doğru ilerlemiş bir çalışmadan elde edilecek en yüksek fayda karşılaştırılmalı ve kararlarda göz önünde bulundurulmalıdır.

Veri madenciliği sürecini en iyi şekilde gerçekleştirebilmek için açık hedefler belirlenmelidir. Bu hedeflerin gereklilikleri açık bir şekilde belirlendikten sonra bütün bu bilgiler veri madenciliği problemine dönüştürülmeli ve hedefe yönelik giriş seviyesinde bir proje planı tasarlanmalıdır.

Temel olarak aşağıdaki faaliyetlerin gerçekleşmesi beklenmektedir (42).

- İşe yönelik hedeflerin belirlenmesi,
- Mevcut durumun değerlendirilmesi,
- Veri madenciliği hedeflerinin belirlenmesi,
- Proje planının yapılmasıdır.

## **2.2. Verinin Anlaşılması**

Bu süreç verinin toplanması ile başlar ve veriyi tanıyabilmek için şu süreçleri takip eder. Veriyi kavrama ile işi kavrama arasında çok yakın bir ilişki vardır. Veri madenciliği probleminin tasarımı ve projenin gereklilikleri, az da olsa kullanılacak veri hakkında bir anlayış gerektirmektedir (43). Bu sürecin sonunda istatistiksel özetlere bakmak gibi verinin keşfine yönelik bazı işlemler gerçekleştirilebilir. Ayrıca yine bu süreç esnasında kümeleme gibi bazı modeller veri içindeki örüntüleri kavrayabilmek için veri setine uygulanabilir.

## **2.3. Verinin Hazırlanması**

Verinin hazırlanması aşaması da veri madenciliği için en önemli aşamalardan biridir. Eksik ya da yanlış hazırlanmış bir veri model kurulduğu andan itibaren sürekli analizin en başına dönülmesine ve yeniden düzenleme gereksinimine yol açacaktır. Çalışmayı yapacak olan analistin, zaman ve maliyet kaybına neden olacaktır.

Veri ambarında toplanan verilerde hatalar, gereksiz değerler yer alabilir. Bunun yanı sıra eksik kalan değişkenler ve bu değişkenlere ilişkin aşırı değerler söz konusu olabilir.

Bu aşama belli bir sıralama olmaksızın, birden fazla kere gerçekleşebilir. Veriler için hazırlık aşaması; veri toplama, veri değerlendirme, verileri bir araya getirme ve temizleme, veri seçimi ve verileri dönüştürme etaplarından oluşmaktadır.

### **2.3.1. Veri Toplama**

Veri madenciliği modeli oluşturma sürecinde ilk olarak, tanımlanmış olan problem için ihtiyaç olunan veriler toplanır. Bahsedilen aşamada toplanan bu verilerin kaynaklarının belirlenmesi gerekmektedir. Verilerin nereden, nasıl ve hangi şartlar altında toplandığı, işletme ve organizasyon için büyük önem taşımaktadır. Çünkü bulunması gerekenden az veri kaynağının olması çalışmayı eksik bırakabileceği gibi bulunması gerekenden çok veri olması süreci uzatabilir. Bu durum ise veri kirliliğine, zaman ve emek kaybına yol açabilmektedir.

### **2.3.2. Veri Değerleme**

Çeşitli kaynaklardan bir araya getirilen veriler, verilerin birbiri ile uyumsuz halde olmasına sebep olacaktır. Uyumsuzluklar, farklı işaretler, darklı veri kodlanması, farklı ölçü birimleridir. Aynı zamanda verilerin hangi şartlar altında elde edildiği de uyumsuzluk için önemlidir. Görece iyi ve başarılı sonuç elde etmek istenilen modellerde, verilerin uyumsuzluk oranlarının düşük ve iyi durumdaki veriler olması önemlidir.

### **2.3.3. Verileri Bir Araya Getirme ve Temizlemesi**

Veri madenciliğinin bu aşamasında farklı veri kaynaklarından toplanan veride belirlenen problemler ve uyumsuzluklar çözümlenerek, veriler tek taban içinde toplanmalıdır. Yine de kolay metodlar ve gelişmiş güzel yapılan problem yok etme adımları, gelecek adımlarda daha kapsamlı problemlere sebep olabileceği göz ardı edilmemelidir. Veri madenciliğinde farklı veri kaynaklarından toplanan verilerin kullanımından kaynaklanan veri uyumsuzluklarının büyük çoğunluğu kodlama farklılıklarından, verilerin farklı zamanlara ait olmalarından, farklı ölçü birimlerine sahip olmalarından oluşur. Süreç içerisinde farklı kaynaklardan elde edilen verilerin birleştirilmesi ve temizlenmesi gerekir. Eğer mevcut verilerde eksiklikler söz konusu ise veriler zaman içinde tamamlanmaya çalışılır. Örneğin kullanılan veri tabanı küçük ise ve kayıp verilere ulaşma imkânı ve zamanı varsa, en önemlisi bu eksik verilere kesinlikle ihtiyaç duyuluyorsa, veriler elle teker teker doldurulabilir. Eksik

verilerin tamamlanması için diđer bir alternatif ise bilgilerin tahmin yöntemiyle tamamlanmasıdır. Süreç içerisinde diđer bir durum ise veri girme aşamasındaki hatalardır. Deđerlerin yanlış girilmesi veya göz ardı edilmesi sonucunda, analiz sonuçları çok farklı bir biçimde deđişebilir. Bu sebeple aykırı deđerler gözden geçirilerek atılır, analizde yapılabilecek deđişiklikler hesaplanarak, uygun ise veri tabanından silinir. Böylece bütün veriler tek bir veri tabanında düzenli bir şekilde tutulabilir.

#### **2.3.4. Veri Seçimi**

Model kurulmadan önce veri seçilmesi yapılmaktadır. Örnek olarak tahminleyici model üzerine düşünülürse, bağımlı ve bağımsız deđişkenlerin belirlenmesi ve modelde deđerlendirilecek verilerin seçilmesi demektir. Kimlik no., pozisyon kodu gibi analizle alakalı olmayan deđişkenler, kurulacak modelin etkinlik oranının azalmasına sebep olacaktır. Bu sebepten dolayı deđişkenleri modelde sokmamak gerekmektedir. Yazılımların bazıları etkinlik oranını azaltan deđişken tiplerini, kendisi otomatik olarak elemektedir. Ancak bu işlemin çalışmayı yapan araştırmacının yapması daha sağlıklı bir model çıkmasına yardımcı olacaktır.

#### **2.3.5. Veri Dönüştürme**

Bütçe tahmini geliştirmek istenen bir modelde, C/I(Cost İncome) maliyet gelir rasyosunun; direkt rasyo olarak hesaplanmış hali yerine ayrı ayrı maliyet ve gider verilerinin kullanılması gerekebilir. Çünkü ayrı ayrı verilen veri tipi, başka bir modele geçişte esneklik sağlayabilecektir.

#### **2.4. Modelleme**

Çalışmada problem netleştirildikten sonra, probleme uygun çalışan bir modelin kurulabilmesi, mümkün olduğunca çok sayıda ayrı modelin denenmesinden geçer. Dolayısı ile verinin hazırlanma ve modelin kurulma adımları, en optimal olan modele ulaşınca dek yeniden başa dönen bir akıştır. Bu aşamalarda çeşitli modelleme yöntemleri seçilir ve uygulamaya alınır Ayrıca bu tekniklerin parametreleri, teknikler için en optimal deđerlere ayarlanılır. Veri hazırlama ve modelleme arasında sıkı bir ilişki vardır (43).

Modelin kurulma aşaması, veri madenciliğinde bilgi kaynaklarından maksimum verimin alınabilmesinin sağlanması açısından çok önemlidir. İyi kurulmuş bir model, analiz sonucunda elde edilecek sonuçların kalitesini de değiştirecek, etkileyecek bir faktördür. Analiz sonucunda hangi örüntülerin bulunabileceğini tahmin edebilmesi için iyi bir veri madenciliği uygulayıcısına ihtiyaç duyulmaktadır. Çünkü model kurma süreci, analiz için kullanıcıya hangi verilerin elde hazır bulunduğunu sunar. Bu süreçte model doğru kurulmadığı takdirde, veri seti içerisinde bulunabilecek kritik ilişkilerin doğru bir şekilde sunulması ve önemli örüntüler tespit edilmesi pek mümkün değildir. Bu sebepten dolayı modelden başarılı sonuç elde etme olasılığı da azalır (44).

Model kuruluş süreci; denetimli ve denetimsiz öğrenmenin kullanılan modellere göre değişiklik teşkil eder.

Denetimli öğrenimde, sınıflar belirlenmiş bir kritere göre ayrılıp, her sınıf için çeşitli örnek verilir. denetçi tarafından ilgili sınıflar önceden belirlenen bir kritere göre ayrılarak, her sınıf için çeşitli örnekler verilir. Amaçlanan, örneklerden yola çıkılıp, her bir sınıf için özelliklerin keşfedilmesi ve kural ifadeleri ile açıklanmasıdır. Kural cümleleri, model tamamlanmış olduğunda yeni örneklerle uygulanır ve bu örnekler hangi sınıfa ait olduğu kesinleşir.

Denetimsiz öğrenme yöntemlerinde, örnek gözlemi ve örnekler arasındaki benzerlikten yola çıkarak kümelerin tanımlanması hedeflenir.

Modelde kullanılan veri tabanı çok büyük olarak karşımıza çıkabilir. Bu durumda rastlantılar bozulmayacak halde veriye örnekleme uygulanması daha sağlıklıdır. Bugünün teknolojisi ile hesaplama seçenekleri, işletim sistemleri ne kadar ileri seviyede olsa dahi, çok büyük veri tabanları üzerinde çok sayıda modelin denenmesi zaman kısıtı göz önüne alındığında, çok efektif ve mümkün olmayan bir seçenektir. Bu sebeple veri tabanının tamamına bir kaç model uygulamak yerine, tesadüfi örnekleme yapılmış bir veri tabanı dikkate alınarak, seçilen örneklem üzerinde mümkün olduğunca çok sayıda modelin uygulanması ve çalışma amacına hizmet eden güvenilir ve güçlü verinin tipine uygun en sağlıklı modelin seçilmesi daha uygundur.

## 2.5. Modelin Kullanılması

Model kurulup sonuçlar elde edildikten sonra mevcuttaki veri madenciliği kısıtlarından ortaya çıkan sonuçların yorumlanma aşamasıdır.

Bu aşamada geçerlilik, sürdürülebilirlik, fayda ve anlaşılabilirlik kriterleri dikkate alınarak üretilen sonuçlar yorumlanabilir.

Kurulan modeller direkt olarak kurum uygulamalarında kullanılabilir. Ya da ihtiyaç doğrultusunda kurumdaki diğer çalışan uygulamalar ile entegrasyonu sağlanabilir.

## 2.6. Modelin Değerlendirilmesi

Bu adım 3 aşamada gerçekleşir:

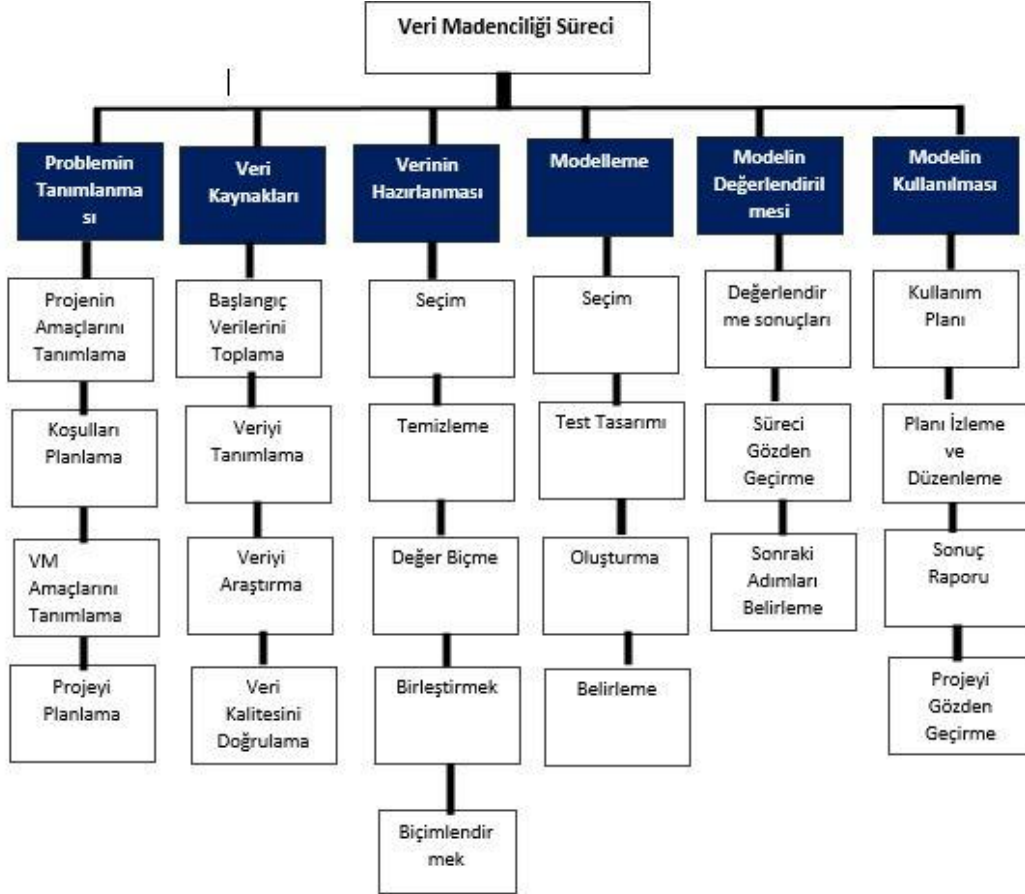
- Sonuçların değerlendirilmesi,
- Sürecin gözden geçirilmesi,
- Bir sonraki adıma karar verilmesi.

Bir veri madenciliği problemi için çeşitli pek çok yöntem kullanılmaktadır. Problem için uygun yöntem veya yöntemler belirlenerek, bunlar arasından en uygun olanlar seçilir. Model oluşturulduktan sonra uygulanacak yöntemin gereksinimlerine uygun olarak veri hazırlanma aşamasına tekrar dönülüp incelenerek gerekli görülen değişiklikler uygulanabilir (45). Modelin değerlendirilmesi aşamasında, modelin kurulup, geçerliliğine karar verilmesinden sonra, modelin iş amacına uygun olup olmadığı değerlendirilir. Sonuçlar elde edildikten sonra proje kısıtlarına ait kriterler baz alınarak oluşan sonuçlar yorumlanmaya başlanır. Bu aşamanın sonunda ise, ulaşılan veri madenciliği sonuçlarının kullanılıp kullanılmayacağı kararı verilmektedir.

## 2.7. Modelin Uygulanması

Model kurulumu ve geçerliliği onaylandıktan sonra direkt olarak ayrı bir uygulama olabileceği gibi, başka bir uygulama ile entegrasyonu sağlandıktan sonra arka tarafta çalışmaya başlayabilir. Gelecekteki şartların değişmesi olasılığı, birbiri ile

entegre ve çok sayıda kısıt ve deęişken içeren bu sistemlerin düzenli takip ve güncellenmesini gerektirmektedir. Model ile tahminlenen ve gerçekleşen arasındaki farkları gösteren analizler modelin süreklilięi ve etkinlięi konusunda bilgi verecektir. Veri madencilięi içinde gerçekleşen adımlara ilişkin olarak tablo 2 incelenirse veri madencilięi sürecinin ayrıntıları açık bir şekilde görülebilir.



**Şekil 5** Veri Madencilięi Süreci

## 2.8. Veri Madencilięinde Karşılaşılan Problemler

Boyut ve kapsam olarak büyük verilerin bulunduğu veri madencilięi çalışmalarında büyük problemler yaşanabilir. Bu kapsamda daha küçük boyutlu veri kümelerinde bazı özelliklere sahip veriler sebebi ile sorunlu çalışabilir. Çalışmanın yanlış sonuçlara sapmaması için çalışmaya başlamadan önce bu tarz problemlerin çözülmüş olması gerekmektedir.



### 2.8.1. Veri Tabanı Boyutu

Veri tabanı boyut artışı her geçen gün ivme kazanmaktadır. Veri tabanında tutulan veriler, iki boyutlu olarak yerleşir ve yayılır.

- *Yatay Boyut:* Nesnelerin özellik rakamları ile yayılır
- *Dikey Boyut:* Nesnelerdeki kayıt rakamları ile yayılır.

Günümüzde çalışan çoğu algoritma yüzler kademesindeki verilerde çalışacak şekilde yazıldığından, daha yüksek kademelerdeki verilerde uygulanabilmesi için daha fazla dikkat ve kontrol gerekir.

### 2.8.2. Gürültülü Veri

Verinin giriş yada toplanma aşamasında oluşabilecek yanlış nitelikler, verinin iletimindeki problemler, veri bazlı tutarsızlıklar ve teknolojik kısıtlar sebebi ile oluşan veri türüne gürültülü veri denir. Araştırma sonuçlarını amacından saptırma ihtimali olduğundan dolayı, ideal bir sistemin gürültülü verileri tanınması ve kapsam dışı bırakması beklenir.

### 2.8.3. Eksik Veri

Veri madenciliği yöntemlerinde iki boyuta ait nitelik ve rakamların olmaması durumunda, amaçlanan bağıntıların kurulması ve problem çözümünden oldukça uzaklaşma ihtimali bulunmaktadır. Bahsedilen eksiklik yatay ve dikey boyutta olabilecek eksikliklerdir. Yatay boyuttaki eksiklik, örnek vermek gerekirse; kişilerin çalışan bağlılığı ile cinsiyeti arasındaki bağıntıyı bulmak hedefleniyorsa, örnekleme mutlaka kişilerin cinsiyet bilgisinin olması gerekmektedir.

Dikey boyuttaki eksiklik; örnek vermek gerekirse, örneklem kümesindeki kayıtların eksik olmasıdır. Örneğin bir şirkette, kıdem ortalaması 5 ile 25 yıl arasındaki kişilerin çalışan bağlılıkları her zaman yüksek çıkıyorsa, bu bağıntıyı ortaya çıkarabilmek için örnekleme mutlaka 5-25 yıl kıdeme sahip kişi bilgilerinin bulunması gerekir.

#### **2.8.4. Artık Veri**

Problem çözümünde amaçlanan çözüme ulaşmak için oluşturulan örneklerdeki gereksiz veri tipleridir. Artık verileri tespit edip, sistem dışına çıkaran algoritmalar özellik seçimi olarak isimlendirilir.

#### **2.8.5. Dinamik Veri**

İçerik bilgisi devamlı olarak değişiklik gösteren veri tabanlarıdır. Bunlara örnek kurumsal çevrim-içi veri tabanları gösterilebilir. Sürekli değişim, veri madenciliği uygulamalarını gözardı edilemez şekilde zorlaştırmaktadır. Bu sorunlardan bazıları şunlardır:

- Analiz sonucunda ortaya konulan ilişkilendirmenin, devamkı olarak değişim gösteren verilerden hangisini temsil ettiğini bulmanın zorluğu.
- Zaman efektinin veriler ve sonuçlar üzerinde etkisinin gösterilmesinin zorluğu.

### **3. BANKACILIK VE FİNANS SEKTÖRÜNDE PAZARLAMA STRATEJİLERİ**

#### **3.1. Pazarlama Kavramı**

Pazarlama kavramı bir işletme için en önemli konulardan biridir. Pazarlama, sadece alım-satım süreci değil, bu sürecin hem öncesindeki hem de sonrasındaki faaliyetleri kapsayan geniş kapsamlı sosyal bir süreçtir. Sosyal bir süreç olması insanların istek ve ihtiyaçlarına yönelik olmasından kaynaklanır. Teknolojinin gelişmesi ve üretimin artması sonucu pazarlamanın kapsamı hizmetler ve fikirleri de içine alacak şekilde genişlemiştir (48).

Pazarlamanın çok yönlü bir süreçtir. Bu sürecin içerisinde satıcı ve müşterilerinde yer aldığı pek çok katılımcı yer almaktadır. Pazarlama kavramı uzun bir süredir araştırılmasına rağmen, dinamik yani değişken bir yapıya sahiptir. Bu dinamik sürecin baş aktörü sanayileşme sürecidir.

Pazarlama, insan istek ve ihtiyaçlarını tespit eder ve bunlara cevap arar. Üretilen mal, hizmet ve fikirler insanların ihtiyaçlarına cevap verecek düzeyde olmalıdır. Pazarlamanın amacı, müşteri tatmini sağlayarak kâr elde etmektir. Üretilen mal, hizmet ve fikirlerin başarısı, sağlayacakları müşteri tatminiyle doğru orantılıdır. Pazarlama sayesinde ihtiyaç sahibi ve bu ihtiyaca cevap verecek mal ya da hizmeti elinde bulunduran taraflar bir araya gelir ve değişim gerçekleşir. Değişimin gerçekleşebilmesi için gereken şartlar şunlardır (48).

- Değişime konu olacak mal veya hizmetin olması,
- Değişimi gerçekleştirecek tarafların olması,
- Taraflar arası uzlaşmanın olması,
- Değişimin tarafları tatmin etmesi,

Üretilen bir mal veya hizmet ne kadar iyi olursa olsun eğer tüketicinin istek ve ihtiyaçlarını karşılamıyorsa başarısız sayılır. Mal veya hizmet üretilirken dikkat edilecek bir başka nokta da mal veya hizmetin müşteri gözündeki değerinin ne olduğunun bilinmesidir. Mal veya hizmete değer biçilirken tüketicilerin satın alma

gücü göz önünde bulundurulmalıdır. Fiyat, satıcının satmaya, alıcının da ödemeye razı olduğu değerdir.

Başlangıçta ürüne odaklı olan pazarlama kavramı zamanla müşteriye odaklanmıştır ve artık işletmeler müşterilerin ihtiyaç duydukları ürün, hizmet ve pazarlanma potansiyeli bulunan şeyleri üretmeye ve müşterilerine sunmaya başlamışlardır.

Kotler'in pazarlama konusunda yaptığı üç ayrı tanım bulunmaktadır ve her tanım pazarlama konusunda belirli dönemlerde kabul gören görüşleri yansıtmaktadır. Bu tanımlara aşağıda yer verilmiştir;

Kotler (1975)'e göre pazarlama; bir firmanın belirli bir müşteri potansiyeli üzerinden belirli bir kazançta ulaşabilmek amacıyla bu müşterilerin arzularını ve gereksinimlerini karşılamak ve müşterilerin bu yolla iyi hissetmelerini sağlamak gayesiyle işletmenin sahip olduğu kaynakların öncelikle çözümlenmesi, ardından belirli bir planlamaya tabi tutulması, daha sonra örgütlü bir biçime büründürülmesi, bu işlemlerden geçirilen kaynakların kullanılmak suretiyle değerlendirilmesi ve son noktada denetlenmesi işlemine "pazarlama" adı verilmektedir (49).

Bu tanımda, müşterilerin gereksinimlerinin karşılanmasıyla tatmin edilmelerine odaklanıldığı görülmektedir.

Kotler (2000)'e göre pazarlama; müşterilerin istedikleri ve gereksinim duydukları belirli bir değer ihtiva eden ürün ve hizmetleri oluşturdukları, bu değerlerin sunumunu gerçekleştirdikleri ve başkalarıyla herhangi bir kısıtlamaya tabi olmaksızın değiştirebildikleri toplum ile alakalı bir olay olarak karşımıza çıkmaktadır.

Bu tanımda ise, pazarlama kavramının toplumsal bir kavram olduğu için bu konuda yapılan faaliyetlerin toplumu ilgilendirdiği görülmektedir.

Kotler (2007)'e göre pazarlama; belirli bir işletme tarafından kar elde etmek amacıyla hedef olarak belirlenen pazar bölümünün gereksinimlerini karşılayabilmek adına belirli bir değer ilk olarak aranması, ardından bu değer yaratılması ve son olarak ta bu değere ihtiyaç duyan pazar bölümüne sunulması şeklinde gerçekleşen bilim ve sanat etkinliğidir.

Bu tanımda önemli olan bir ihtiyacın belirlenmesi ve bu ihtiyacın karşılanması için gerekli olan süreçlerin hazırlanmasıyla müşterilerin bu ihtiyacının onlara sunulmasının sağlanmasıdır.

Pazarlama konusunda yapılan yukarıdaki tanımlardan yola çıkarak, pazarlamanın her dönemde farklı manalara geldiğini ancak anlatmak istediği noktanın temelde aynı olduğu sonucuna ulaşabiliriz.

### **3.2. Banka Kavramı**

Bu bölümde bankacılık kavramı ile ilgili çeşitli tanımlar yer almaktadır.

Banka kavramından bahsetmeden önce, banka kelimesinin kökenine inmek gerekmektedir. Banka terimi köken olarak “ masa, tezgah” anlamına gelen “banço” kelimesinden gelir. Bankacılık sektörünün ilk temsilcilerinden sayılan Lombard ailesi çalışma yerleri için, “banchum”, bu işi yapan kendileri için ise “bancheri” sıfatlarını kullanmışlardır. Bu işi de “banco” denen tahta bir masa ile varlıklarını sürdürmüşlerdir. (50).

Banka; sermaye, para ve kredi üzerine her türlü işlemi gerçekleştiren, tanzim işlemlerini yapan, özel ve tüzel kişilerin, devletin, kurum ve kuruluşların bahsedilen konulardaki her türlü ihtiyacına cevap verebilmek için çalışan bir finansal kuruluştur. Bugünün şartları gereği şimdiki bankaların faaliyet kapsamı çok çeşitli ve genişleyen bir yelpazede büyüdüğünden standart bir tanım yapabilmek çok zorlaşmıştır. Bankalar iskonto, kambiyo ve kredi işlemleri gerçekleştirir, faizi ile para alım satımı yapar, aracı rolü üstlenir, borsada aktif rol alır, kredi kart yönetimini yapar, kiralık kasalarında güvenli koruma hizmeti verir. Ülke gündeminden beslendiği gibi, ülke gündeminde önemli rol oynar.

Gerçekte yukarıda banka üzerine verilen tanımlar yeterli değildir. Günümüzde bankalar çok çeşitli hizmetler yürütmektedir. Banka kavramını daha iyi anlayabilmemiz için bankaların fonksiyonlarını incelememiz gerekmektedir.

#### **3.2.1. Bankaların Fonksiyonları**

Günümüzde rekabetin ortamının ve teknolojinin artması nedeniyle bankanın verdiği hizmette çeşitlenmeler olmuştur. Bu nedenle banka kavramını tanımlarla

anlamak yerine, fonksiyonlarını incelemek daha doğru olacaktır. Şendođdu (2006) ya g6re bu fonksiyonlar;

- eřitli yatırım araları vasıtası ile, varlıkları deęerlendirme olanaęı saęlamak,
- zel ve tuzel kiřilere , ihtiyalarına cevap verecek řekilde kredi ve finansman desteęi vermek,
- ek ve senet tahsilatları yapmak, senetleri kırdırmak
- EFT/Havale iřlemleri ile fiziki olarak para tařıma y6k6ml6l6ę6n6 hafifletmek
- Dıř ticaret ve operasyon iřlemlerine aracılık ve y6nlendirme rol6n6 6stlenmek,
- Kiralık kasa ile g6venli saklama hizmetleri vermek,
- Kredi kartı hizmeti vermek,
- Finansal danıřmanlık hizmetleri sunmak,

Genel izgileri ile bankaların fonksiyonları yukarıda bildirilmiřtir. Ancak bankaların yapısal b6y6kl6ę6 ve ekonomik izgisine g6re yukarıda verilen fonksiyonlara biilen 6nem derecesi deęiřiklik g6sterebilir. Ayrıca bankalar 6lkelerin ekonomilerine ve d6nya ekonomisine katkı saęlamaktadır. Bu katkıları ařaęıdaki gibi 6zetleyebiliriz (51).

- 6lke ekonomisinde aktif olarak kullanılmayan fonları sisteme ekerek, fonlara b6t6n olarak hareket saęlamak,
- Para politikası araları, etkin ve esnek bankacılık altyapısı gerektirdięinden, geliřtike politikanın etkinlięini y6kseltmek
- S6rd6r6len kredilendirme politikası ile gelir ve varlık daęılımını kontrol altına almak,

- Toplanılan her kaynak, etkin ve kârlı piyasalara aktararak toplum faaliyetlerinde belirleyici rol oynamak,
- Sürdüdüğü ödeme ve kredi politikası ile uluslararası ticaretin gelişmesine destek olmak,

### **3.2.2. Dünyada Bankacılığın Tarihsel Gelişimi**

Bankaların tarihçesi çok eskilere dayanmaktadır. Tarihçiler tarafından yapılan araştırmalar neticesinde, dünya tarihinin bilinen en eski finansal yapısı Mezopotamya'daki Kızıl Tapınak (MÖ 3500)'tir. (52). Tapınakların güvenilir bir yer olması, paraların burada saklanmasını doğurmuştur ve bankacılık sistemini ortaya çıkarmıştır. Devam eden yüzyıllarda ise bankacılık hizmeti tapınaklardan, ekonomik gücü elinde bulunduran ailelere doğru kaymıştır.

Hammurabi kurallarında borçların nasıl verileceği ve borçların nasıl tahsil edileceğine dair kanunlar yer almaktadır. Bu dönemde tarıma dayalı borç verme işlemi bulunmaktadır. Çiftçiler hasat dönemi öncesinde temin etmesi gereken araç ve gereçleri alır ve hasat dönemi sonrasında bu borcu faiziyle geri öderlerdi. Hammurabi kurallarında faize değinilmiş ve alınan borcun hangi faiz oranından nasıl geri ödeneceği bildirilmiştir.

Günümüz bankacılığı 1157'de Venedik Bankası'nın kurulmasıyla başlamıştır. 1408'de Genova Bankası kurulmakla birlikte, Yeniçağ ile başlayan değişiklikler, faizin geçerliliğinin kabul edilmesine neden olmuştur. Bu kabul edilme, bankacılığın gelişmesinin önündeki önemli bir engeli yok etmiştir. Sömürgelerin yayılması ve ticaretin gelişmesi bankaları da uzmanlaşmaya ve birleşmeye yöneltmiştir.

Bankacılığın tarihsel ilerlemesini duraksatan faktörlerden bir tanesi faiz yasağıdır. İnanışları gereği Hıristiyan ve Müslüman kişiler bankerlik yapamadıklarından dolayı, sektörde oluşan açığı Yahudiler doldurmuştur. İstanbul'un fethedilmesi ile Avrupa'ya doğru göç başlamış ve Yunan belgeleri Avrupa'ya giriş yapmıştır. Avrupa'da yaşanan reform sonrasında, kilisenin halk üzerindeki hakimiyetinin sona ermesi ile birlikte faiz yasağı da kalkmıştır.

19. yüzyıla değin bankalar kendi özsermayelerini üzerinden borç ile bankacılık yaparken, bu tarihten itibaren mevduat toplamak vb. hareketler ile kaynak kullandırımına başlamışlardır, ve toplum ekonomisindeki rolleri fazlası ile artmıştır.

Ticari bankacılığın gelişimi ise 1975 Petrol Krizi ile birlikte olmuştur. Petrol ihraç eden ülkelerin paralarını bankalara yatırması sonucu oluşan kaynak, bankaların kredi verme oranını arttırmıştır. Bankacılık sektöründeki bu gelişmenin teknolojik gelişme ile desteklenmesi sonucu elektronik bankacılığı ortaya çıkarmıştır.

### **3.3. Strateji Kavramı**

Sözlük manası olarak strateji, belirli bir amacı gerçekleştirmek adına hareketsel anlamda mevcut bir birlik oluşturma ve bu hareket birliğine belirli bir düzen verebilme sanatıdır. Strateji kelimesi ele alındığı zaman köklerinin Eski Yunanca olduğu net bir biçimde görülebilmektedir (53).

Yönetimsel açıdan strateji, var olan rakiplerin gerçekleştirdiği işlerin değerlendirilmesi ve devamında amaçlarına ulaşabilmek için karar birliğine varılan, dinamik, kesin sonuçlara yönelmiş ve uzun vadeli bir kararlar mekanizması şeklinde tanımlanabilir.

Chandler (1962)'ye göre strateji; yapılacak bir işin geniş vade de gereken hedefleri ile amaçlarının saptanması, ayrıca belirlenmiş olan bu amaçlara varabilmek adına ihtiyaç duyulan kaynakların sağlanması ve işin yapılış tarzının oluşturulmasıdır (54).

Daft (1998)'in bakış açısına göre strateji; bir örgütsel yapının ortaya koyduğu örgütsel hedeflerini gerçekleştirebilmesi adına yoğun rekabete sahne olan çevreyle bağlantısını gerçekleştirmesine yardımcı olan plana "strateji" adı verilmektedir (55).

James Quinn (2003)'e göre ise strateji; bir örgütsel yapı formu içerisinde gerçekleştirilmek istenen eylemleri, belirlenmiş bulunan politika ile hedefleri bir araya getirme görevini üstlenen bir plandır (54).

Aktan (2008)'in bakış açısından stratejiye odaklanmak gerekirse; ileri ki dönemlere ne şekilde varılabileceği konusunda yardımcı olan ve bütüne odaklanma davranışı gösteren planlama şeklinde belirtilebilir (56).



Stratejiler uzun vadeyi planlama aracı olduđu için işletmelere kendilerini ve rakiplerini tanıma olanağı sunar. Yukarıdaki tanımlamalardan özetle stratejiyi devam eden cümleler ile açıklanabilir. Strateji; işletmelere, sahip oldukları kaynakları en verimli şekilde kullanmalarına olanak veren, işletmeleri amaçlarına ulaştıracak misyon, vizyon, amaç ve hedefleri kapsayan, işletmenin uyumunu kolaylaştıran, sürekli yenilenmesini gerektiren ve rakipleri tanımaya fırsat veren yönetsel bir araçtır.

### **3.4. Pazarlama Kavramı**

Pazarlama çok yönlü bir süreçtir. Bu nedenle pazarlama kavramı görüş birliğine varılmadığından net bir tanım yapmak mümkün değildir. Pazarlama, sadece alım-satım süreci değil, aynı zamanda hem bu sürecin öncesini hem de sonrasını kapsayan sosyal bir süreçtir. Sosyal bir süreç olması insanların istek ve ihtiyaçlarına yönelik olmasından kaynaklanır. Pazarlama, insanın ihtiyaçlarını ve isteklerini tespit eder ve bunlara cevap arar. Pazarlamanın amacı, tatmin sağlayarak kar elde etmektir. Pazarlama sayesinde ihtiyaç sahibi ve bu ihtiyaca cevap verecek mal ya da hizmeti elinde bulunduran taraflar bir araya gelir ve değişim gerçekleşir.

### **3.5. Bankacılıkta Stratejik Pazarlama**

Bankalar da diğer işletmeler gibi uzun vadede ne yapacaklarına önceden karar vermek durumundadır. Bankanın kaynakları, hacmi, Pazar koşullarının olumlu yönde kullanılması ile yapacağı atılımlar bankanın stratejisini belirler.

Günümüzde modern pazarlama disiplini kabul etmek, finansal kuruluşlar arasındaki rekabet sebebi ile zorunlu hale gelmiştir. Pazarlama teknikleri ile, bankacılık faaliyetinin yalnızca para satışından oluşmadığı, çok çeşitli sayı ve kapsamda hizmet seçeneklerinden oluştuđu çıkarımı yapılmıştır. çok fazla hizmetler paketinden meydana geldiği anlaşılmıştır. Bu kapsamda bankalar bu hizmetlerin doğurduğu ihtiyaca yönelik hizmet ve memnuniyet sağlayan finansal kuruluşlar olarak hizmet vermeye başlamışlardır.

Bankacılık hizmetlerinde standartlaşmanın zor oluşu, hizmetlerin farklılaştırma açısından daha elverişli olduğunu gösterir. Farklılaştırma stratejisi bankalar açısından daha elverişlidir. Farklılaştırılmış pazarlama stratejisinde hedef

pazar bölümündeki tüketicilerin istek ve ihtiyaçları dikkate alınır. Müşterilerin bireysel ve kurumsal olarak sınıflandırıldığı bankacılık sektöründe bireyler, tasarruflarını değerlendirmek, havale yapmak, maaş almak gibi ihtiyaçları için banka hizmetlerini kullanırken, kurumsal müşterilerin banka hizmetlerini kullanma nedenleri de farklılık gösterir. Bu yüzden bankacılık sektöründe tüketici grupları bölümlere ayrılarak her pazar bölümü için farklı stratejiler geliştirmek gerekir.

Günümüzde hızlı gelişmeler ve değişimlere ayak uydurma isteği, sektörel yarışın etkisi ile bankalar elindeki kaynaklarını fırsata çevirme, tehdit olarak algılanabilecek nitelikleri azaltmak için strateji planlamaya önem vermeyi arttırmışlardır. Gelecekteki riskleri ortadan kaldırmak için bankalar bugünden ne yapılması gerektiğine karar verir. Bankaların stratejik planları aşağıdaki gibidir.

- Bankanın finansal ve kurum hedefleri
- Pazarlama hedefinin netleştirilmesi
- Pazarlama stratejisi
- Kampanyalar
- Bütçeler, yeni şube açılması veya şube kapatılması gibi yatırımlar
- Diğer bankaların davranışları
- Süreci değerlendirme.

### **3.5.1. Bankacılıkta Stratejik Planlamanın Tarihsel Gelişimi**

Bankalar uzun bir süre, teknolojik imkanların ve iletişim kanallarının azlığı sebebi ile rekabet olasılığı olmayan konularda faaliyette bulunmuş ve sınırlı kanunların etkisi ile birlikte 1950'lere gelinceye dek esnetilemeyen, katı bir pazarlama anlayışına sahip olmuşlardır. 1950'lerde ABD'de türeyen rakip finans kuruluşları bankaların aktif olmayan rekabet algısını yok etmiş, daha faal ve çok kapsamlı bir bakış açısına doğru yönlendirmiştir. Oluşan rekabet sonrasında bankalar modern pazarlama tekniklerine yönelmişlerdir. Yeni pazarlama yöntemleri ile

bankanın yalnızca para satışından oluşmadığı, çok çeşitli ve kapsamlı hizmetler sağlaması gerektiği kesinleşmiştir.

1960’larda bankacılık piyasasındaki en radikal değişim, kredi kartlarının hayata girmesidir. Sözkonusu yıllarda başarılı müşteri ilişkileri yönetiminin tek başına etkili olmadığı görülmüş, banka imaj ve algı yönetiminin oluşturulmasının yanısıra çeşitli ve kapsamlı reklam, satış geliştirme, pazarlama gibi paketleri de kullanmaya başlamışlardır.

Bankacılıkta asıl pazarlama anlayışı 1970’li yıllara denk gelmektedir. 1970’li yıllarda reklam faaliyetleri artmış, kredi kartları çeşitlenmeye başlamıştır. Banka personeline yönelik müşteri odaklı eğitimlerin sayısı artış göstermiş, müşteri ihtiyaçları tespiti ve bu tespitlere uyan yeni hizmetler verilmeye başlanmıştır. Bu yeniliklerin yanısıra, yeni bankacılık ürünleri, kapsamlı müşteri bağlılık çalışmaları ile desteklenmiştir.

1990’lar ve sonraki yıllarda, bankacılık sektöründe modern pazarlama disiplini kabul görmüş, yayılmış ve artış göstermiştir. Teknolojik gelişiminin artması ile diğer piyasalarda gözlemlenen benzer değişiklikleri de dikkate alarak bankacılığın önyüzü radikal değişiklik geçirmiştir. Bankalar pazarlama stratejilerinin önemini kabul etmiş, müşteriye sunulan hizmetin; müşterinin isteği doğrultusunda ve kaliteli olma yönünde ilerlemiştir.

Günümüzde bankacılık piyasasındaki strateji eğilimi, ürün odaklı olmaktan çıkıp müşteri odaklı stratejiye giderek azalmıştır. Müşteri odaklı stratejide; üründen müşteriye ulaşan süreçleri iyileştirmek yerine, müşteri ile etkin ilişki yönetimi nasıl kurulur, müşteri sadakati ve memnuniyeti nasıl sağlanabileceği dikkate alınır. Artık bankacılıkta terzi modeli, yani müşterinin ihtiyacına yönelik olarak hizmet ve ilişki yönetimi etkindir. Bu nedenle satıştan çok satış sonrası hizmetin önemi artmıştır. Günümüz şartları gibi değişken ve karmaşık müşteri profili topluluğu bulunması, yeni pazarlama yöntemlerine doğru itmiştir.

Ekonomik olarak ileri ve gelişmiş ülkelerde sektöre yönelik aktif bankacılık disiplinine karşın, Türkiye’de bankacılık 1980’lere kadar dışa kapalı, pasif çalışma sistemi içinde sürdürülmüştür. 1980’lere kadar bankalarda yalnızca reklam, satış geliştirme çabalarının kullanıldığı bilinmektedir. Bu dönem içerisinde bankalar vermiş

oldukları reklamlarda güvenilir banka imajı oluşturmaya ve bu algıyı muhafaza etmeye ağırlık vermiştir.Yani müşterilerin istek ve ihtiyaçlarını gözardı etmiştir. 1980’li yılların ortalarına doğru dışa açılımlı politikaların sonucunda yabancı bankalar piyasa içerisine girmiş ve artan rekabet koşullarında Türk Bankaları yenilenmeye zorlanmıştır.

Türk bankacılık sektöründeki hizmet kapsamının değişmesi, faiz, döviz gibi araçların eskiye oranla daha serbestleşmesi ve mevzuatın eskiye oranla daha az köşeli sınırlarının olmasıyla birlikte başlamıştır. Yaşanan günümüz teknolojisi, bankacılık sektörüne getirdiği bir çok yenilik ile birlikte bankacılığın iş yapış şekli ve süreçleri hızla değişmiştir. Ürünlerin kapsamı, çeşitliliği ve hızı çok değişmiştir.(58).

### **3.5.2. Bankacılıkta Stratejik Planlamayı Etkileyen Faktörler**

Bankacılıkta pazarlama stratejileri, banka yönetiminin kontrolü altında bulunan ve bulunmayan faktörler tarafından etkilenmektedir (59).

Banka yönetimi tarafından kontrol altında olmayan faktörleri, ulusal ve uluslararası çevre etkenleri olarak ele alabiliriz. Ulusal çevre faktörleri pazarla stratejisini direkt olarak etkilerken, ulusal olmayan çevre faktörleri, ilk olarak bankanın ulusal çevresini ve sonra bankanın kendisini etkilemektedir.

#### **3.5.2.1 Banka İçi Faktörler**

- Bankanın en iyi şekilde hizmet sunması,
- Bankanın şube sayıları ve var olan şubelerin yerleri,
- Bankanın İnsan Kaynakları politika ve kalitesi
- Bankanın işyeri ergonomisi,
- Bankanın sunmuş olduğu ürün çeşitliliği,
- Banka iç ve dış müşteri sadakat çalışmaları,

### 3.5.2.2 Banka Dışı Faktörler

Banka dışı etkenler, bankanın stratejisini en çok belirleyen etkenlerdir. Ulusal ve ulusal olmayan çevre faktörleri olarak iki ana başlık altında açılabiliriz. Bu etkenleri kısacası aşağıdaki gibi özetleyebiliriz.

#### Ulusal Çevre Faktörleri

- Nüfusa Yönelik Çevre:

Nüfusa yönelik çevre faktörü, bir işletmede pazarlama stratejilerini en çok etkileyen faktördür. Bazı değişkenler, bankaların sahip oldukları hizmetlere olan talepleri etkiler. Örneğin; ülkedeki vatandaşların sosyal aile dinamikleri, nüfustaki genç ve yaşlı kişilerin sayıca oranları vb.. Genç nüfusun hâkim olduğu ülkelerde bankaların stratejik planlaması da bu yönde olmaktadır.

- Siyasi ve Hukuksal Çevre:

Siyasi ve hukuksal çevre; doğanın, üretici ve tüketicinin korunumu, enflasyona karşı fiyat kontrolü, tekellerin oluşmasına karşı uygulanan rekabet politikası ve devlet ödenekleri ve teşvikleri gibi kapsamlı bir bölümü kapsar ve ilgilendirir. Bir ülkedeki banka, faaliyet gösterdiği lokasyondaki devletten ayrı hareket edemez ve düşünülemez. Yani bankacılık sektöründe her yapılan yeni düzenleme var olan hükümeti etkiler, hükümet tarafından alınan yeni kararlar ise bu sektörü etkiler.

- Ekonomik Çevre:

Bankacılık sektörü her daim içerisinde bulunduğu ekonomik politikalardan etkilenir ve aynı şekilde bu ekonomilerden etkilendiği zaman kendileri de ekonomik sistemi etkiler. Bankaların ekonomiye olan etkileri, kredi kullandırılması sonucunda tüketimi arttırmak, kaydı para yaratarak ekonominin likiditesini arttırmak gibi sonuçlar doğurmaktadır.

- Sosyal ve Kültürel Çevre:

Bankaların uyguladıkları pazarlama stratejileri, içerisinde bulunduğu topluma aykırı olamaz. Ayrıca sosyal ve kültürel çevre faktörlerinin, bireylerin yaşam

biçimlerinde meydana getirdiği deęişimler de bankaların pazarlama stratejisini etkiler.

- Teknoloji:

Teknolojik ilerleme, bankaların yerel piyasada rekabet gücünü koruyabilmesi açısından en önemli faktördür. Hızlı ilerleyen teknolojiyi takip etmek, bankaların pazarlama faaliyetlerinden çıkmış olup, pazarlamanın asıl unsuru haline gelmiştir.

- Rekabet:

Bankalar kendi aralarında artan bir rekabete sahip iken, aynı zamanda başka faaliyetlerini yerine getiren diğer finansal kuruluşlarla da rekabet halindedir. Dolayısıyla bu durum, bankaların pazarlama stratejilerini geliştirmiştir.

#### Ulusal Olmayan Çevre Faktörü

Ülkenin gelişmişliği, teknolojik ve sosyoekonomik durumu, dengeli gelir dağılımı, ülkenin güvenliği ve istikrarlığı sektördeki fiyatların belirlenmesindeki etmenlerdir. Fiziki ve ekonomik koşulların deęişmesi durumunda pazarlama faaliyetleri de devletten devlete farklı olur. Bu faktörler tüzel ve özel kişilerin lehine ya da aleyhine sonuçlar doğurabilir. Örneğin, bir ülkenin kredi notunun yabancı kuruluşlarca düşürülmesi, ülkeye yapılan dış yatırımları azaltacağından dolayı yabancı para dolaşımı azalır. Dolar ile ticaret yapan bir kobi, bu durumdan TL ile işlem yapan başka bir kobiden daha fazla etkilenir. (60).

### 3.6. Banka İle Sunulan Hizmetler

Bankların temel işlevi olan mevduat toplama ve borç verme işlemlerinin haricinde geliştirdikleri ürünleri de müşterilere sunarlar. Bankların ürün politikalarını belirlemesindeki asıl amaç, müşterilerin talep ve ihtiyaçlarını tatmin etmek, karlılığı arttırmak, ürünlerde alternatif yaratarak rekabet üstünlüğü sağlamak, lider banka imajı yaratmaktır.

- Mevduat:

Mevduat, vadeli mevduat ve vadesiz mevduat olarak ikiye ayrılmaktadır. Vadesiz mevduat bankalar için masrafsız mevduat türüdür. Vadeli mevduatın maliyeti, karların azalmasına neden olur.

- Kredi Hizmetleri:

Bireysel ve kurumsal müşterilerin hizmetlerini karşılamak için belirlenmiş vadelere göre geri ödemek koşulu ile kredi kullandırılmasıdır. Bireysel müşterilere göre kredi çeşitleri ihtiyaç kredisi, taşıt kredisi ve konut kredi olarak ayrılmaktadır.

- Banka Kartları:

Alışveriş kolaylığı sağlayan, nakit olma özelliğinin yanı sıra kredi özelliği olmayan borçlandırma kartlarıdır.

- ATM Hizmetleri ve POS Hizmetleri:

ATM makineleri hem müşteriler hem de bankalar için avantaj sunmaktadır. Müşteriler günün her saatinde ATM'leri kullanabildikleri için bankalar açısından maliyetlerde tasarruf sağlamaktadır.

POS makineleri, kredi kartı ve banka kartı ile yapılan alışverişlerde kullanılmaktadır. Telefon hattı üzerinden banka ile temasa geçer ve onay almasının ardından alışverişin gerçekleşmesi sağlanır.

- Telefon Bankacılığı:

Müşterilerin bankaya gitmesini engelleyerek, hem müşteri tarafından hem de banka tarafından zaman tasarrufu sağlanır.

- İnternet Bankacılığı:

İnternetin her eve ulaşması ile birlikte, bankaların üzerinde durduğu alternatif bir yöntemdir. İnternet bankacılığı kullanan müşterilerin mevduatlarına daha fazla faiz verilmesinin nedeni de budur.

- Özel Bankacılık:

Varlıkları diđer müşterilere oranla daha yüksek ve çeşitli olan müşterilerin varlıklarının, müşteriye ait olan ihtiyaç ve beklentilerine dair şekil alacak vaziyette sürdüren bankacılık segmentidir. Bankacılık segmentinden beklentileri de ürün ve yatırım araç seçeneklerinin kendilerine özel olarak şekillenmesi ve müşteri sadakati yöntemlerinin ihtiyaca özel olmasıdır.

- Çek Hizmetleri:

Bankalar, müşterilerin nakit taşıma riskini ortadan kaldırmak için uyguladığı yöntemdir.

- Kiralık Kasa Hizmetleri:

Bankalar, müşterilerin sahip olduğu değerli eşya, belge vb. değerleri saklamak için uyguladığı bir hizmettir. Bu hizmet karşılığında kasa boyutuna göre kira ödemesi alınmaktadır.

- Havale İşlemleri:

Bir bankanın şubeleri arasında hesaptan hesaba şeklinde paranın transfer edilmesidir.



## 4. UYGULAMA

### 4.1. Araştırmanın Amacı ve Kapsamı

“Bankacılık Sektöründe Pazarlama Stratejilerinin Belirlenmesinde Sınıflandırma ve Veri Madenciliği” konulu mevcut araştırmanın amacı, ülkemizde faaliyet gösteren bir özel sermayeli yatırım bankasının özel bankacılık müşterilerine yönelik pazarlama stratejisi belirlemek için, müşterileri finansallarına göre çok değişkenli bir istatistik yöntem olan Kümeleme Analizi (Cluster Analysis) sınıflandırması yapmaktır.

Çalışmanın kapsamı; ülkemizde faaliyet gösteren bir özel sermayeli yatırım bankasının 1202 adet özel bankacılık müşterisini 2016 yılındaki finansal bilgilerinden yola çıkarak kümeleme analizi gruplamaktır.

### 4.2. Araştırmanın Yöntemi

Çalışma için, veri madenciliği yazılımı olan açık kaynak kodlu “Knime“ yazılımı kullanılmıştır. Araştırma için, kümeleme yöntemlerinden küçük boyutlu data için ideal olan k-medoids yöntemi kullanılmıştır. Bu durumda çalışmanın amacı; banka için yeni müşteriler kazanmaktan ziyade, mevcut müşterileri tanımak, ihtiyaçlarını analiz edebilir duruma getirmektir. Ayrıca ihtiyaçlarına göre hem müşteri kârı ve bankanın kendi kârlılığını maximize edecek yöntemler önerebilmeyi hedeflemektir. Dolayısı ile her bir müşteri temsilcisi, şirket politikasına uygun şekilde zamanını ve eforunu harcayacaktır. Çıkacak sonuca göre, portföy yöneticileri müşteri sınıflarına özel efor ve zaman harcayıp, şirket politikasına uygun pazarlama modelleri uygulanması planlanmıştır. Günümüz şartları, bankaları stratejilerine uygun, müşteri seçmeye yöneltmiştir.

### 4.3. Evren ve Örneklem

Bankanın, CRM (Customer Relationship Management) raporlama bölümünden alınan veri, 1202 ayrı müşterinin datasını içermektedir. Veriler, özel bankacılık ürünlerine sahip müşterilerin mevcut TL karşılıklarını içermektedir.

#### 4.4. Verileri Anlama ve Değişkenler

Söz konusu bankada Özel Bankacılık segmentine ait üç ayrı ürün vardır. Ürünler; Bono, Sukuk ve Repo'dur. Tüm portföy yöneticilerine bu üç ürünü satmak üzerine hedefleri verilmiş ve yıl sonu performansları bu hedef gerçekleştirmeler üzerinden yapılması planlanmaktadır. Ancak tüm müşterilere aynı pazarlama yöntemi yaklaşıldığındaki, zaman ve kâr kaybetme oranını azaltmak için müşterilerin kümeleme yöntemi ile ifade edilmesine ihtiyaç duyulmuştur.

- **Bono:** Yatırım bankalarının borçlu adı altında, SPK üzerinden kayıt edilip kurumdan dışarıya çıkarılarak belli bir indirim üzerinden pazarlanan menkul kıymetlerdir. (64) . Genel olarak risk iştahı olmayan ve bu konuda temkinli, ancak faiz riskini alabilen, vade sonundaki kazancını en başından beri bilmek isteyen müşteri profili için uygundur. Bono ürünü vadesi ve düşük stopaj oranları sebebi ile yatırımcıların eline geçen net getiri arttırır. Vadesi sebebi ile de bankaya daha uzun süre bağlı kalır.

- **Sukuk:** Arapça "sak" kökünden gelen, sertifika anlamına gelen bir kelime olan sukuk, yatırım aracı olarak da varlığın kıymetini sabitleyerek bu vasıta ile satımıdır. Bir varlığa ait sukuğu satın alanlar, satın alınan sukuk miktarında o varlığa paydaş olurlar. Varlığa ait gelirlerde aynı oranda sukuk sahibine gider.

- **Repo :** Bir menkul kıymetin daha önceden kesin olarak netleşmiş tarihte, geri alma sözü ile satılmasıdır. Repo'nun oranı, vadesi ve getirisi her zaman net olduğu için, net kârını arttırmak isteyen bankalar ve risk almak isteyen yatırımcılar için ilk sırada tercih edilen bir ürün değildir.

1202 ayrı müşteriye ait gelen datanın özet görüntüsü aşağıdaki gibidir. Daha önce aktarıldığı gibi özel bankacılık ürünleri 3 tane olup, veride beş ayrı değişken vardır Üç ürüne ek olarak , faiz geliri ve faiz dışı gelir bilgisi de mevcuttur. Faiz gelirleri repo ve bonodan beslenmekte olup, faiz dışı gelirler ise sukuk satışından elde edilmektedir.

Müşteri ID	BONO	REPO	SUKUK	FAİZ DİŞİ GELİR	FAİZ GELİRİ	Bono%	Repo%	Sukuk%	Faiz Dış%	Faiz%
54997	0,00	12090000,00	0,00	21294,14	0,00	0,00%	99,82%	0,00%	0,18%	0,00%
55340	4999937,12	0,00	0,00	0,00	8265,75	99,83%	0,00%	0,00%	0,00%	0,17%
55530	650994,38	0,00	0,00	101,17	849741,48	43,38%	0,00%	0,00%	0,01%	56,62%
55890	7647952,98	0,00	0,00	0,00	30972,74	99,60%	0,00%	0,00%	0,00%	0,40%
55905	7168248,46	0,00	21625586,40	2957,91	1001,78	24,89%	0,00%	75,09%	0,01%	0,00%
55990	0,00	102428,00	0,00	0,00	577,22	0,00%	99,44%	0,00%	0,00%	0,56%
499998	580233,24	984,80	0,00	0,00	1549,39	99,57%	0,17%	0,00%	0,00%	0,27%
500009	194931,88	0,00	0,00	0,00	0,00	100,00%	0,00%	0,00%	0,00%	0,00%
500070	0,00	9106,92	0,00	0,00	0,00	0,00%	100,00%	0,00%	0,00%	0,00%
500094	0,00	30000,00	0,00	0,00	8,63	0,00%	99,97%	0,00%	0,00%	0,03%
500116	194894,47	0,00	0,00	111,69	443,26	99,72%	0,00%	0,00%	0,06%	0,23%
500120	2609622,24	0,00	0,00	0,00	1184,89	99,95%	0,00%	0,00%	0,00%	0,05%
501179	0,00	84400,00	0,00	0,00	0,00	0,00%	100,00%	0,00%	0,00%	0,00%
501263	2665265,81	787306,00	21801345,89	410,00	57036,36	10,53%	8,11%	86,13%	0,00%	0,23%
501286	678919,26	0,00	0,00	0,00	0,00	100,00%	0,00%	0,00%	0,00%	0,00%
501337	193606,61	177463,00	0,00	499,50	158,68	52,08%	47,74%	0,00%	0,13%	0,04%
501351	1203317,93	0,00	0,00	427,27	0,00	99,96%	0,00%	0,00%	0,04%	0,00%
501352	0,00	60153,79	0,00	0,00	30,44	0,00%	99,95%	0,00%	0,00%	0,05%
501461	465477,36	0,00	0,00	0,00	0,00	100,00%	0,00%	0,00%	0,00%	0,00%
501463	19966481,86	0,00	0,00	0,00	0,00	100,00%	0,00%	0,00%	0,00%	0,00%
501978	529031,69	0,00	0,00	0,00	0,00	100,00%	0,00%	0,00%	0,00%	0,00%
502973	2306931,84	0,00	0,00	2076,26	33310,83	98,49%	0,00%	0,00%	0,09%	1,42%
503620	0,00	14537,96	0,00	0,00	8,36	0,00%	99,94%	0,00%	0,00%	0,06%
508429	0,00	143503,39	0,00	1326,59	408,78	0,00%	98,81%	0,00%	0,91%	0,28%
512185	15442273,35	0,00	0,00	0,00	66338,31	99,57%	0,00%	0,00%	0,00%	0,43%
512186	3796304,67	0,00	0,00	0,00	16308,51	99,57%	0,00%	0,00%	0,00%	0,43%

Şekil 6 Veri seti görüntüsü

#### 4.5. Bulgular

Müşteri segmentasyonu için kümeleme analizinin gerçekleştirilebilmesi için şart olan KNIME sistemi kurulmuş ve analiz sonuçları değerlendirilerek, her bir müşteri kümesi için önerilerde bulunulmuştur.

Küme sayısı için 5 ayrı küme seçilmiştir. K- Medoids yöntemi ile çıkan sonuçları incelemeye değişken bazlı başladığındaki dağılım grafikleri aşağıda verilmiştir..

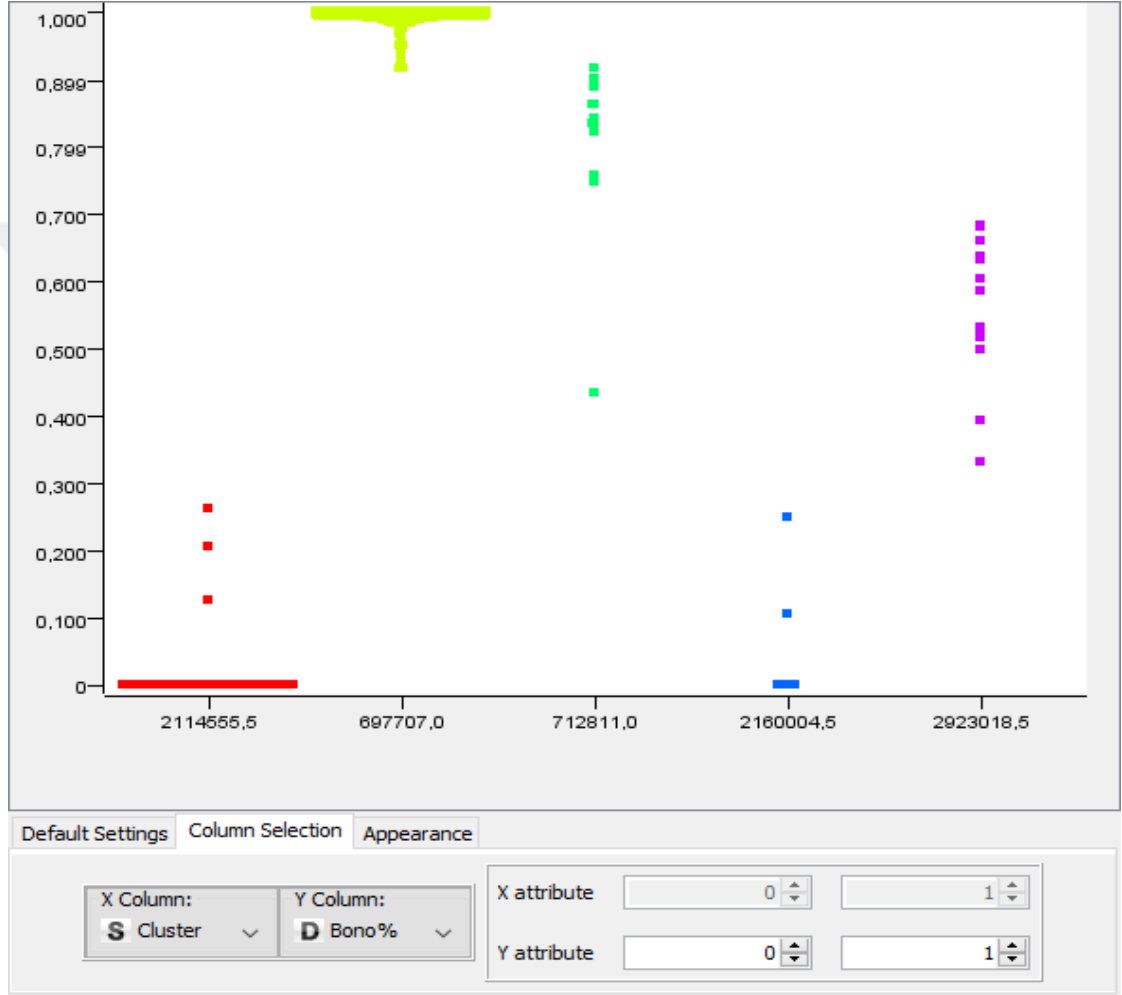
KNIME uygulamasında Color Manager yöntemi ile kümeleri renklendirilmiştir..

K-medoids sonucunda kümelerin merkezipleri (Centrotipe) aşağıdaki KNIME tablosunda görülmektedir.

Row ID	D Bono%	D Repo%	D Sukuk%	D Faiz Dışı%	D Faiz%	I partitio...
2114555.5	0	1	0	0	0	270
697707.0	0.999	0	0	0	0.001	890
2923018.5	0.586	0.414	0	0	0	14
712811.0	0.836	0.161	0	0	0.003	16
2160004.5	0	0	0.999	0	0.001	12

Şekil 7 K-Medoids Algoritması Sonuçları (Merkez tipleri)

#### 4.6. Bono Değişkeninin Analiz Çıktısı



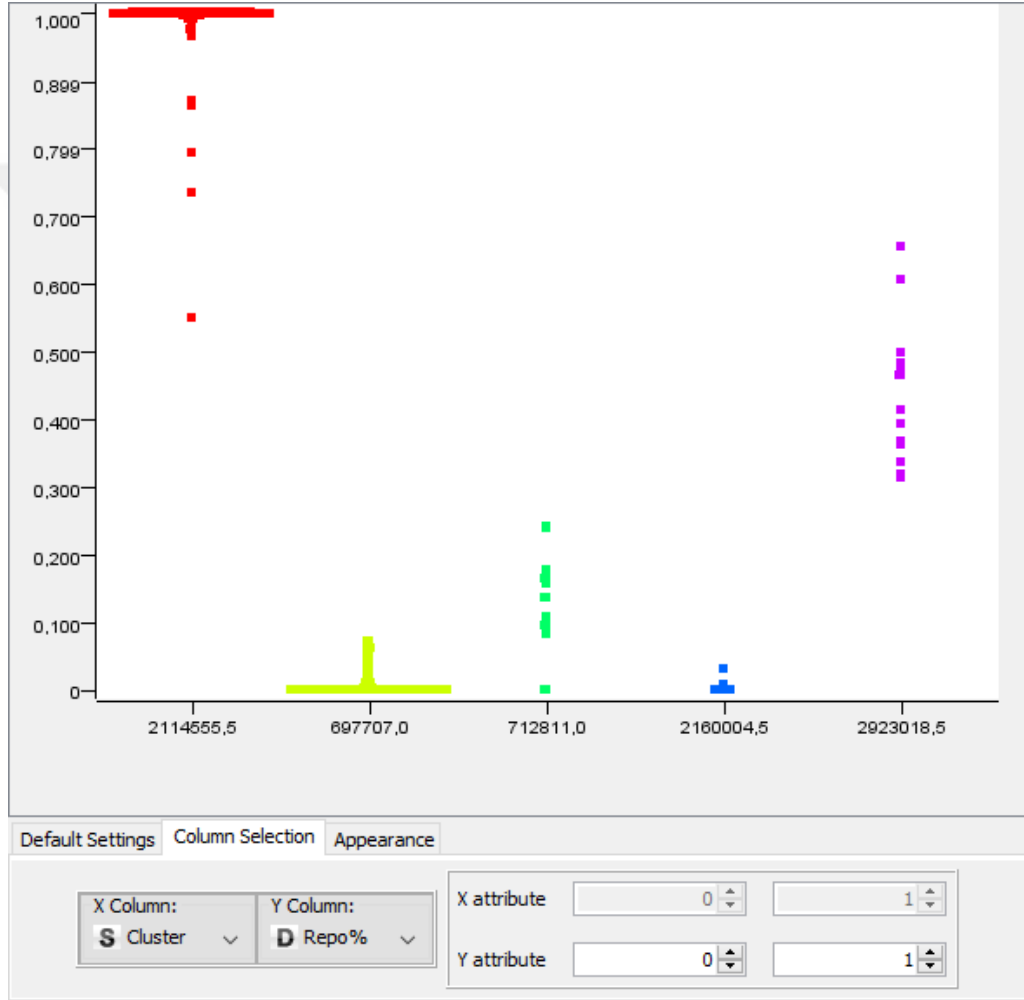
Şekil 8 Bono% değişkeni skatter grafiği

Tablo 2 Bono% özet tablosu

Değişken	Cluster				
	Kırmızı	Sarı	Yeşil	Mavi	Mor
Bono%	Yok	Yüksek	Yüksek	Yok+Düşük	Orta

Scatter Plot Dağılım grafiğinde görüldüğü üzere; bono değişkenindeki kümelerin yorumlanmasında kümelerin birbine göre konumu derecelendirilecektir. Sarı renkli kümenin, kümeler arasındaki bono değerlerinin en yüksek olan küme olduğu, ardından bir outlier farkı ile yeşil kümenin geldiği ve orta alt seviyede morların yığıldığını gözlemlenmektedir. Mavi küme ise Düşük/ Yok olarak yorumlanabilir. Kırmızı küme ise , bono değişkeninde yok olarak yorumlanacaktır.

#### 4.7. Repo Değişkeninin Analiz Çıktısı



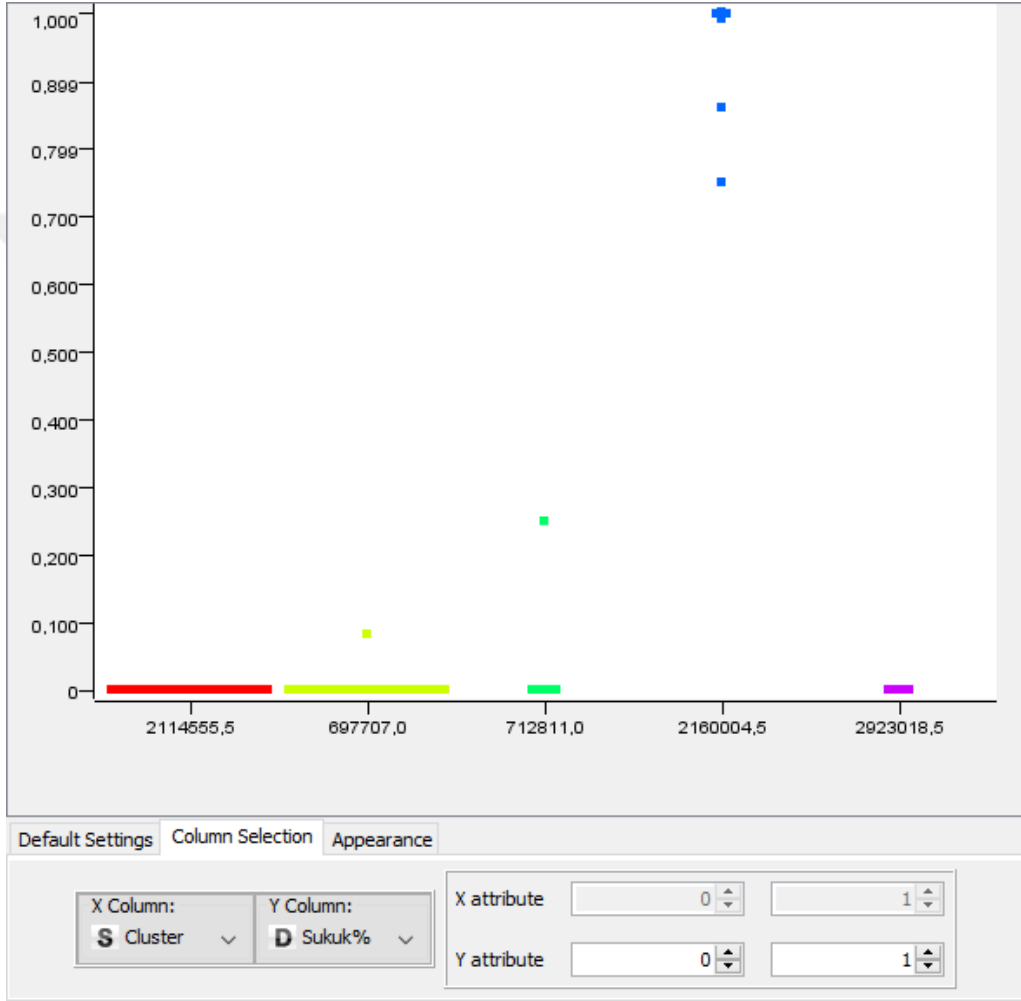
**Şekil 9** Repo% değişkeni skatter grafiği

Scatter Plot Dağılım grafiğinde görüldüğü üzere; repo değişkenindeki kümelerin yorumlanmasında kümelerin birbine göre konumu derecelendirilecektir. Kırmızı renkli kümenin, kümeler arasındaki repo değerlerinin en yüksek olan küme olduğu, ardından mor kümenin geldiği düşük seviyelerde mavi ve yeşil kümelerin dağıldığı gözlemlenmektedir. Sarı küme ise Düşük/Yok olarak yorumlanabilir.

**Tablo 3** Repo% özet tablosu

Değişken	Cluster				
	Kırmızı	Sarı	Yeşil	Mavi	Mor
Repo%	Yüksek	Yok+Düşük	Düşük	Düşük	Orta

#### 4.8. Sukuk Değişkeninin Analiz Çıktısı



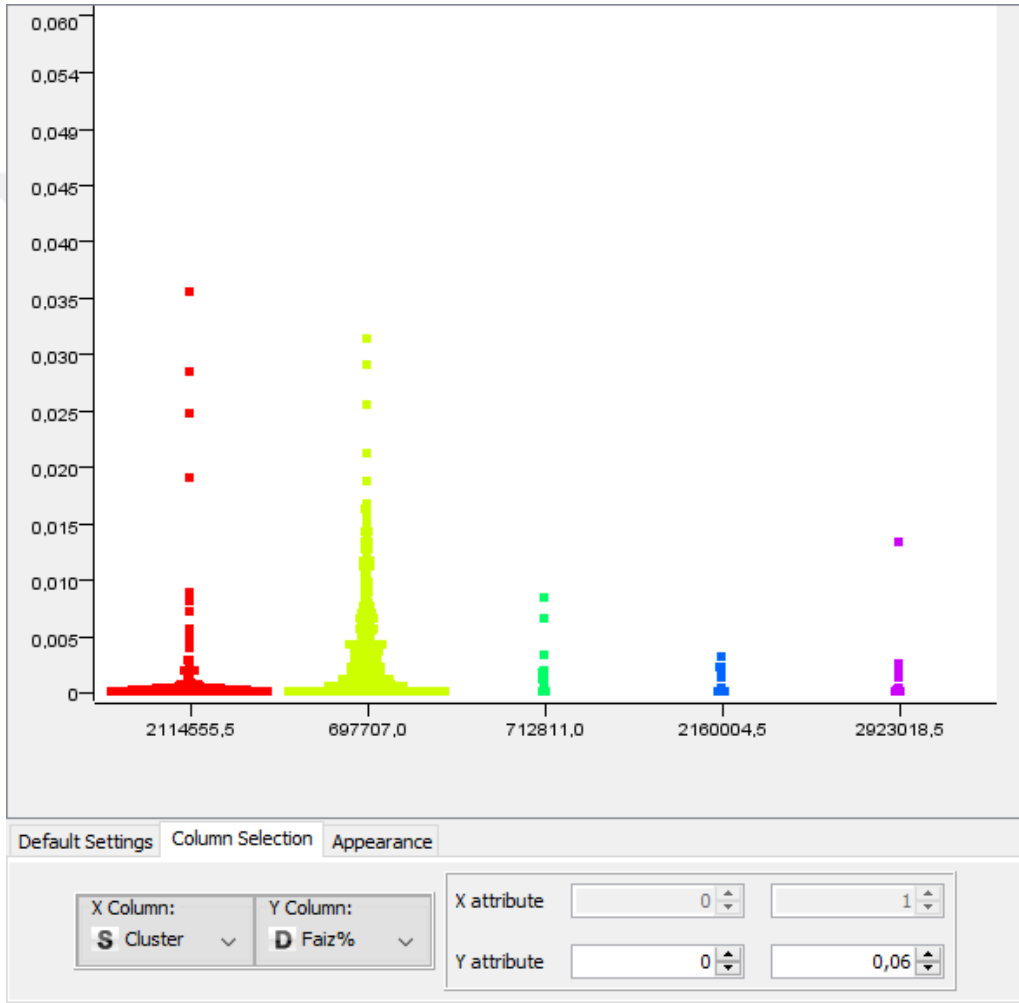
**Şekil 10** Sukuk% değişkeni skatter grafiği

Scatter Plot Dağılım grafiğinde görüldüğü üzere; sukuk değişkenindeki kümelerin yorumlanmasında kümelerin birbine göre konumu derecelendirilecektir. Mavi renkli kümede sukuk sahibi olanların neredeyse tamamının yığıldığı gözlemlenmektedir. Diğer kümeler ise neredeyse yok olarak yorumlanabilir.

**Tablo 4** Sukuk% özet tablosu

Değişken	Cluster				
	Kırmızı	Sarı	Yeşil	Mavi	Mor
Sukuk%	Yok	Yok	Yok	Büyük Kısmı Yığılmış	Yok

#### 4.9. Faiz Geliri Değişkeninin Analiz Çıktısı



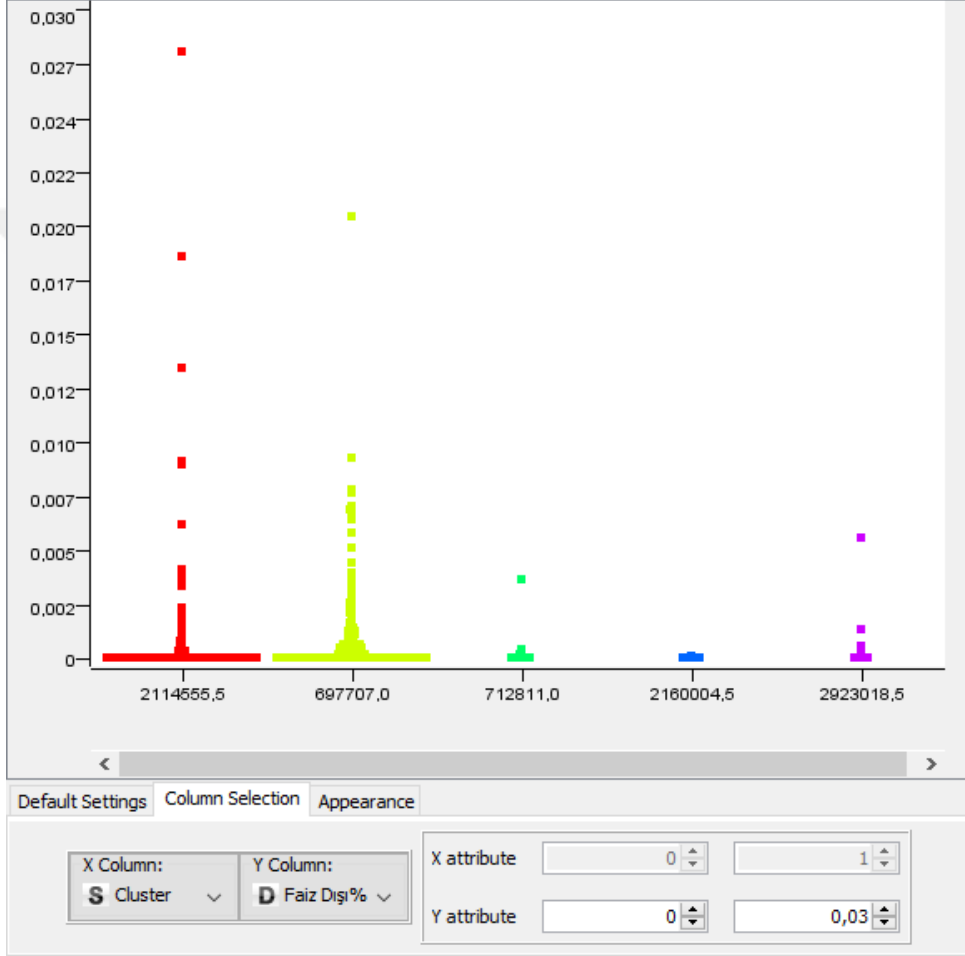
**Şekil 11** Faiz% (Faiz geliri) değişkeni skatter grafiği

Scatter Plot Dağılım grafiğinde görüldüğü üzere; faiz geliri değişkenindeki kümelerin yorumlanmasında kümelerin birbine göre konumu derecelendirilecektir. Sarı renkli kümenin, kümeler arasındaki faiz geliri değerlerinin en yüksek olan küme olduğu, ardından kırmızı kümenin geldiği düşük seviyelerde mavi, yeşil ve mor kümelerin dağıldığı gözlemlenmektedir.

**Tablo 5** Faiz% (Faiz Geliri) özet tablo

Değişken	Cluster				
	Kırmızı	Sarı	Yeşil	Mavi	Mor
Faiz Geliri	Yüksek Orta	Yüksek	Düşük	Düşük	Düşük

#### 4.10. Faiz Dışı Gelir Değişkeninin Analiz Çıktısı



**Şekil 12** Faiz Dışı% (Faiz Dışı Gelir) değişkeni scatter grafiği



Scatter Plot Dağılım grafiğinde görüldüğü üzere; faiz dışı gelir değişkenindeki kümelerin yorumlanmasında kümelerin birbine göre konumu derecelendirilecektir. Kırmızı renkli kümenin, kümeler arasındaki faiz geliri değerlerinin en yüksek olan küme olduğu, ardından kırmızı kümenin geldiği orta seviyelerde sarı kümenin dağıldığı gözlemlenmektedir. Düşük/Yok olarak mor küme, neredeyse hiç dağılmamış küme ise mavi ve yeşil kümelerdir.

**Tablo 6** Faiz Dışı% (Faiz Dışı Geliri) özet tablosu

Değişken	Cluster				
	Kırmızı	Sarı	Yeşil	Mavi	Mor
Faiz Dışı%	Yüksek Orta	Orta	Yok	Yok	Düşük/Yok

## SONUÇ

Beş ayrı kümenin, değişkenlere göre dağılımını incelediğimiz durumda bir pazarlama strateji oluşturulabilir. Yeni bir strateji için satılacak olan ürün ve değişkenlerin gerçek hayattaki durumlarını ve müşteriler açısından algısını da dikkate alınması gereklidir. Kümeleri özet bir tablo ile açıklanması gerekirse; aşağıdaki tablodan detaylı olarak inceleyebiliriz.

**Tablo 7** K-Medoids küme özet tablosu

Renk	Cluster	Kişi Sayısı	Bono%	Repo%	Sukuk%	Faiz Dışı%	Faiz%
Kırmızı	Cluster 1	270	0.0	0.9997	0.0	0.4412	0.2645
Sarı	Cluster 2	890	0.9986	0.0	0.0	0.9824	0.0013
Yeşil	Cluster 3	14	0.5863	0.4137	0.0	0.0	0.0
Mavi	Cluster 4	16	0.8359	0.1608	0.0	0.0	0.0034
Mor	Cluster 5	12	0.0	0.0	0.9986	0.0	0.0014

Tabloda müşteri profil dağılımını değişkenler açısından detaylı şekilde incelendiğinde, bankada çalışan özel bankacılık müşteri temsilcilerine yıl sonu hedeflerine ulaşması için hangi müşterilere ağırlık vermesi gerektiğine dair özet bilgiler çıkarılmıştır. Bu özet bilgiler sayesinde temsilci, portföyündeki müşterinin hangi kümeye ait olduğunu ve hangi kümedeki müşterilerine ağırlık vermesi, hangi kümedeki müşterilerine ise ağırlık vermemesi konusunda karar verebilecektir. Şirketin politikası gereği ilgili segmentte hangi stratejiyi uyguluyorlar ise, o stratejiye uygun müşteri kümesini seçerek pazarlamalarına devam edebileceklerdir. İsbetli karar sonucunda, politika için ideal müşteriye, ideal bir ürün sunabileceği için, zaman yönetimi açısından kısa ve etkili pazarlamalar doğacaktır.

Ürünlerin yapıları gereği bankaya en fazla kâr ve müşteri devamlılığı sağlayan ürün bonodur. Vadesinin uzatılabilmesi nedeniyle; doğru pazarlama stratejisi ile müşteri memnuniyeti sağlanması durumunda uzun vadeli kâr ve müşteri sadakatını beraberinde getirecektir.

Müşteri sadakati ve kâr hedeflenen bir banka olduğu için müşteri kümelerinde diğer kümelere nazaran çekici olan kümeler bonusu yüksek olan müşterilerdir.

Repo ise, güvenli bir yatırım aracı olmakla birlikte vadesi daha kısa olan ancak yine de risk alabilecek müşteri kaynağı içerdiği için kısa vadeli kâra ve potansiyele işaret eder.

Sukuk; belli bir görüşe sahip müşteri profilinin faiz geliri elde etmek istemediği için kâr paylaşımı olarak seçtiği yatırım amacı olduğundan, az getirisi olan ancak yine de belli profiller için değişmez bir yatırım aracıdır. Uzun ya da kısa vadede kâra tekabül etmez.

Kümeleri yukarıdaki bilgiler ışığında yeniden değerlendirip, özel bankacılık müşteri temsilcilerine müşterileri tanımaları için yeni adlar verilerek sunulacaktır. İstatistik bilgileri olmayabileceği, verileri doğru okumama ihtimalleri olduğundan dolayı, kümeler temsilciler için eğlenceli adlar ile isimlendirilecektir.

Yeni pazarlama stratejilerini, özel bankacılık temsilcilerine sunulması için isim önerileri tablodaki gibidir.

**Tablo 8** Küme özellik yorumları

Değişkenler	Cluster				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Kırmızı	Sarı	Yeşil	Mavi	Mor
REPO	Yüksek	Yok+Düşük	Düşük	Düşük	Orta
BONO	Yok	Yüksek	Yüksek	Yok+Düşük	Orta
SUKUK	Yok	Yok	Yok	Yığılma Yüksek Oranlı	Yok
FAİZ DIŞI GELİR	Yüksek Orta	Orta	Yok	Yok	Düşük/Yok
FAİZ GELİRİ	Yüksek Orta	Yüksek	Düşük	Düşük	Düşük
<b>KÜME ADI</b>	<b><i>Kısa Vadede Kazandıran</i></b>	<b><i>Uzun Vade Kazandıran</i></b>	<b><i>Potansiyeller</i></b>	<b><i>Kanaatkarlar</i></b>	<b><i>Kaybettirenler</i></b>

Kısa vadede kazandıran müşteri profilinin, yüksek varlıklı olmalarına karşın bankanın varlık büyüklüğünü normal bir hızda arttırırken, risk almayı sevmeden getirisi belli olan repo kullanmışlardır. Bu müşteriyi bono almaya teşvik etmek için bu müşteri tipine özel kampanyalar ile çekingenliği aşılabılır. Daha özel kampanya ve sıcak temaslar ile desteklenebilir. Özel indirimler, uçak biletleri vb hızlı ve avantajlı çözüm önerileri getirilebilir.

Uzun vadede kazandıran müşteri profili, bankanın bono politikasına en uygun müşteri tipi olmasına karşın, en kalabalık küme olma özelliğine sahiptir. Kümenin kalabalıklığı sebebi ile önemli olan müşteri bağlılığının devamının sağlanması için, kendilerini farklı ve VIP hissedecekleri özel etkinlikler sağlanabilir. Yüksek gelirli kişilerin rağbet ettiği turlar, Risk iştahı olan müşteriler oldukları için bonoyu oldukça yüksek kullanan müşterilerdir, değişiklik yaratması adına repodaki komisyon maliyetleri azaltılıp repoya da eğilim sağlanabilir.

Potansiyel müşteriler, bono alımı yapan ancak faiz geliri düşük olan kişilerdir. Alım yaptığı halde getiri olmayan müşteri tipidir. Bunun sebebi diğer bankaları tercih etmesi ya da uzak kalmış olmaları olabilir. Bu müşteriler ile yeniden iletişime geçilip eskisinden daha kuvvetli iletişim yöntemleri kullanılabilir. Bankaya yeniden gelmesini sağlayacak şekilde, kredi kartına puan yüklenebilir, kişiye özel indirimler alınabilir.

Kanaatkarlar; sözlük anlamı gibi az ile yetinebilen profili temsil eder. Sukuk ağırlıklı olduğu için müşteriye diğer ürünler faiz içerdiği için cezbetmeyebilir. Bu yüzden daha yüksek miktarlarda sukuk ürünü satışı yapabilmek için, bankayı müşteri için daha cazip ve tercih edilebilir hale getiren kampanyalar yapılabilir. Özel günlerde, bayramlarda özel hediyeler gönderilebilir.

Kaybettirenler, hem bono hem repo orta miktarda almış ancak geliri neredeyse hiç olmayan müşteri profilleridir. Kazançlarını bankaya aktarmayacak şekilde hareket ettikleri için bankaya uzun yada kısa vadede kâr sağlamayacak olup, zaman ve maliyet yönetimli stratejiye uygun olmayan müşteri tipleridir.

Son olarak çalışma, hem müşteri tarafında çalışan portföy yöneticilerinin mevcut müşterilerini daha iyi tanıması, hem de pazarlama stratejileri tarafında çalışan kişiler için yeni stratejilere temel olabilecek müşteri tanıtımı yapmayı hedeflemiştir. Uzun vadede müşteri temsilcilerinin hedeflerinin verilmesi, yeni kampanyaların planlanması konusunda temel oluşturabilmesini sağlayacaktır.

## KAYNAKLAR

- [1] Han J. , Kamber M., *Data Mining: concepts and techniques*, Morgan Kaufmann, USA, 2006.
- [2] Emel G., Taşkın Ç. , *Veri Madenciliğinde Karar ağaçları ve Bir Satış Analizi Uygulaması*, Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 2005.
- [3] Linoff G.S., Berry M.J.A., *Data Mining Techniques for Marketing, Sales and Customer Relationship Management*, Wiley, Canada, 2011.
- [4] Marakas G.M., *Decision Support Systems in the 21st Century*, Prentice Hall, USA, 2003.
- [5] Tan P.N., Steinbach M., Kumar V., *Introduction to Data Mining*, Pearson International Edition, USA, 2006.
- [6] Tuffery S., *Data Mining and Statistics for Decision Making*, Wiley, USA, 2011
- [7] Özekes S., Çamurcu A.Y., *Veri Madenciliğinde Sınıflama ve Kestirim Uygulaması*, Marmara Üniversitesi Fen bilimleri Dergisi, 2002.
- [8] Kittler r., Wang W., *The Emerging Role in Data Mining*, Solid State Tchnology, 1999.
- [9] Karacan H., Yeşilbudak M., *Kullanıcı Merkezli İnteraktif Veri Madenciliği: Bir Literatür Taraması*, Bilişim Teknolojileri Dergisi, 2010.
- [10] Köktürk F., Ankaralı H., Sümbüloğlu V., *Veri Madenciliği Yöntemlerine Genel Bakış*, Türkiye Klinikleri, 2009,
- [11] Savaş S., Topaloğlu N., Yılmaz M., *Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri*, İstanbul Ticaret Üniversitesi, Fen Bilimleri Dergisi, 2012.
- [12] Aytaç M.B., *Doğrudan Pazarlama Aracı Olarak Tele Pazarlama İçin Veri Madenciliği Çözümleneleri: Banka Müşterileri Üzerine Bir Uygulama*, Gazi Üniversitesi, Bilişim Enstitüsü, 2013.
- [13] Akpınar H., *Veritabanlarında Bilgi Keşfi ve Veri Madenciliği*, İstanbul Üniversitesi, İşletme Fakültesi Dergisi, 2000.

- [14] Çetinyokuş T., *Veri Küplerinin Bütünleşik Kullanımına Yönelik Yeni Bir OLAP Mimarisi*, Doktora Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü.2008.
- [15] Kalogirou S. A., *Applications of Artificial Neural-Networks for Energy Systems*, Applied Energy,2000.
- [16] Çalışkan S. K., ve Soğukpınar İ., *KxKNN: K-Means ve K En Yakın Komşu Yöntemleri ile Ağlarda Nüfuz Tespiti*, 2. Ağ ve Bilgi Güvenliği Sempozyumu, Girne, KKTC, 16-18 Mayıs 2008
- [17] Akbulut S., *Veri Madenciliği Teknikleri ile Bir Kozmetik Markanın Ayrılan Müşteri analizi ve Müşteri Segmentasyonu*, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2006.
- [18] Özdamar K., *Paket Programlar ile İstatiksel Veri Analizi*, Kaan Kitapevi, Eskişehir, 2004.
- [19] Chien C. F., Chen L.F, *Data Mining to Improve Personnel Selection and Enhance Human Capital : A Case Study in High-Technology Industry*, Expert Systems with Applications, 2008.
- [20] Argüden Y., Erşahin B., *Veri Madenciliği*, ARGE Danışmanlık Yayınları, İstanbul, 2008.
- [21] <http://www.twocrows.com/intro-dm.pdf> adresinden 10.06.2016 tarihinde alınmıştır.
- [22] Ching W. K., Pong M.K., *Advances in Data Mining and Modeling*, World Scientific, Hong Kong, China, 2002.
- [23] Suner A., Çelikoğlu C.C., *Toplum Tabanlı Bir Çalışmada Çoklu Uygunluk Analizi ile Sağlık Kurumu Seçimi*, Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi, 2010.
- [24] Çakmak Z., Uzgören N., Keçek G., *Kümeleme Analizi Teknikleri ile İllerin Kültürel Yapılarına Göre Sınıflandırılması ve Değişimlerin İncelenmesi*, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 2005
- [25] Berry, Michael J.A., ve Linoff, Gordon., *Mastering Data Mining, The Art and Science of Customer Relationship Management*, New York, Wiley Publishing.2000.
- [26] Özçakır F.C., Çamurcu A.Y., *Birliktelik Kuralı Yöntemi İçin Veri Madenciliği Yazılımı Tasarımı ve Uygulaması*, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 2007

- [27] Sholom M.W., Indurkha N., Zhang T., Damerau F., Text Mining: Predictive Methods for Analyzing Unstructured Information, Springer. 2004.
- [28] W. Fan, L. Wallace, S. Rich, Z. Zhang., *Tapping into the Power of Text Mining*, Communications of ACM,2006.
- [29] Özgür Yılmazel, Türkçe Metin Arama Sistemleri ve Veri tabanlarının Karşılaştırılması, 2012.
- [30] Sezgin M.E., “İkili Kodlama Kuralına Dayalı Olarak Hazırlanan Multimedya Ders Yazılımının Fen Bilgisi Öğretimindeki, Akademik Başarıya, Öğrenme Düzeylerine ve Kalıcılığına Etkisi”, Yüksek Lisans Tezi, Çukurova Üniversitesi, Sosyal Bilimler Enstitüsü, Eğitim Ana Bilim Dalı, 2002
- [31] Aydoğan E.K., Gencer C., Akbulut S., “Churn Analysis And Customer Segmentation Of A Cosmetics Brand Using Data Mining Techniques”, Sigma Mühendislik ve Fen Bilimleri Dergisi, 2008
- [32] Timor M., Ezerçe A., Şimşek U.T., “Müşteri Profili Ve Alışveriş Davranışlarını Belirlemede Kümeleme Ve Birliktelik Kuralları Analizi: Perakende Sektöründe Bir Uygulama”, Yönetim Dergisi, 2011,
- [33] Ay D., Çil İ., “Migros Türk A.Ş.’de Birliktelik Kurallarının Yerleşim Düzeni Planlamada Kullanılması”, Endüstri Mühendisliği Dergisi, 2010.
- [34] Koyuncuğil A. S., Özgülbaş N., ”Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları”, Bilişim Teknolojileri Dergisi, 2009.
- [35] Yıldırım P., Uludağ M., Görür A., “*Hastane Bilgi Sistemlerinde Veri Madenciliği*”, Akademik Bilişim 2008, Çanakkale On Sekiz Mart Üniversitesi, Çanakkale,2008.
- [36] Oğuz B., “*Metin Madenciliği Teknikleri, Kullanılarak Kulak Burun Boğaz Hasta Bilgi Formlarının Analizi*”, Akdeniz Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Ve Tıp Bilişimi Anabilim Dalı, Yüksek Lisans Tezi, 2009.
- [37] Albayrak A.S., Yılmaz Ş. K., “*Veri Madenciliği: Karar Ağacı Algoritmaları Ve İMKB Verileri Üzerine Bir Uygulama*”, Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 2009.
- [38] Aşan Z., “Kredi Kartı Kullanan Müşterilerin Sosyo Ekonomik Özelliklerinin Kümeleme Analizi İle İncelenmesi”, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 2007.

- [39] Tosun t., “*Veri Madenciliği Teknikleriyle Kredi Kartlarında Müşteri Kaybetme Analizi, Yüksek Lisans Tezi,*” İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, 2006.
- [40] Özseven T., Düğenci M., “LOG Analiz: Erişim Kayıt Dosyaları Analiz Yazılımı Ve GOP Üniversitesi Uygulaması”, *Bilişim Teknolojileri Dergisi*, 2011.
- [41] Gürsoy U., *Veri Madenciliği ve Bilgi Keşfi*, 1.Baskı, Ankara, Pegem Akademi, 2009.
- [42] Olson, D. L., ve Delen, D., “*Advanced Data Mining Techniques*”, Springer, Berlin Heidelberg, 2008.
- [43] Wirth, R., ve Hipp, J., “CRISP-DM: Towards a Standard Process Model for Data Mining” , *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, Manchester.2000
- [44] Michael J.A Berry, *Gordon Linoff, Data Mining Solutions*, U.S.A., Wiley Computer Publishing, 1998
- [45] Derya Ay, “*Veri Madenciliği ve Apriori Algoritması İle Bir Süpermarket Uygulaması*”, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Sakarya, 2009
- [46] Aydoğan E. K., & Gencer C., “ Veri Madenciliği Problemlerinde Kaba Küme Yaklaşımı Kullanılarak Sınıflandırma Amaçlı Yapılmış Olan Çalışmalar.” *Kara Harp Okulu Savunma Bilimleri Dergisi*. 2007.
- [47] Schertel S.L., “Data Mining and Its Potential Use in Textiles: A Spinning Mill” *Doktora Tezi*, North Carolina. 2002.
- [48] Mehmet Karafakıoğlu., *Pazarlama İlkeleri*, Literatür Yayıncılık. 2006.
- [49] Mustafa Halid Karaarslan, “Rekabet Üstünlüğü Sağlama Aracı Olarak Konumlandırma Stratejilerinin Stratejik Pazarlama Yönetiminde Kullanılması: Kayseri İli Tüketicilerinin Kayseri ve Ankara Mobilyaları Algılarının Ölçümü”, (Erciyes Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Yayınlanmamış Yüksek Lisans Tezi), Kayseri, 2008.
- [50] Günal, Mehmet, *Türk Bankacılık Sektörünün Sorunları ve Geleceği*, Ankara Ticaret Odası Yayınları, Ankara 2001.



- [51] Şevket Sayılğan, ‘‘Bankacılık Sisteminin Güncel Sorunları ve Çözüm Önerileri’’ İktisat İşletme ve Finans Dergisi, Yıl 18, Sayı 210, Eylül 2003.
- [52] Aydemir, Namık, ‘‘Dünden Bugüne Bankacılık’’, Yüksek Denetim Dergisi, Yıl 1, Sayı 3, Temmuz Eylül 2002.
- [53] Coşkun Can Aktan, ‘‘Strateji ve Stratejik Yönetim Kavramları’’, (Çevrimiçi) <http://www.canaktan.org/yonetim/stratejik-yonetim/strateji-kavrami.htm>, (Erişim Tarihi: 01.09.2016)
- [54] Ayşe Aktuna, ‘‘Strateji Uygulama Sürecinin Başarısını Etkileyen Faktörlerin Algılanması: Kamu Kurumları Örneği’’, (Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Yayınlanmamış Yüksek Lisans Tezi), Ankara, 2010.
- [55] Richard L. Daft, *Organization Theory and Design*, 6. Edition, Cincinnati, Ohio: South Western College Publishing, 1998
- [56] Yılmaz Demiralp, ‘‘Muhasebe Sektöründe Strateji ve İş Etiğinin Rolü Adıyaman İlinde Bulunan Muhasebe Bürolarında Uygulanması’’, (Türk Hava Kurumu Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Yayınlanmamış Yüksek Lisans Tezi), Ankara, 2014.
- [57] Remzi Altunışık & Şuayp Özdemir & Ömer Torlak, *Modern Pazarlama*, Değişim Yayınları, 2001.
- [58] [www.maximumbilgi.com](http://www.maximumbilgi.com) ( Erişim Tarihi: 02.09.2016)
- [59] Öçer, A. ve Kara, G. Bankacılık Sektöründe Pazarlama Faaliyetleri: Trabzon’da Örnek Uygulama, Pazarlama Dünyası, 2001.
- [60] ÖZALP, İnan, *Uluslararası işletmecilik*, Anadolu Üniversitesi Açık Öğretim Yayınları 2007.
- [61] BÜLBÜL, Ş., vd., ?, Propensity Skor Uygulamalarında Kümeleme Analizinin Test Amaçlı Kullanımı, [iletisim.atauni.edu.tr/eisemp/html/tammetinler/172%20.pdf](http://iletisim.atauni.edu.tr/eisemp/html/tammetinler/172%20.pdf) (28.12.2009)
- [62] SATICI, Ö., 1992, Doktor Hemşire ve Hasta Sorunlarının Çok Değişkenli Kümeleme Yöntemleri İle Araştırılması, Yayınlanmamış Doktora Tezi, Dicle Üniversitesi Sağlık Bilimler Enstitüsü, Diyarbakır, 94s.
- [63] Joseph F. Hair et. al., *Multivariate Data Analysis*, Prentice Hall, U.S.A., 1998, s. 473

- [64] ARSLAN, (2004). Spor ayakkabısı satınalma ve kullanım amaçlarına ilişkin pazar bölümlerinin oluşturulması. İstanbul: Marmara Üniversitesi İ.İ.B.F Dergisi, cilt XIX, Cilt(1), 259
- [65] <http://www.tuketicifinansman.net/2010/03/bono-nedir-bono-turleri.html#ixzz4SqYSFPjK> (Erişim Tarihi: 20.12.2016)



## ÖZGEÇMİŞ

20 Ağustos 1990 tarihi, Kayseri doğumluyum. Lise öğrenimimi Kayseri Küçükçalık Anadolu Lisesi'nde 2008 yılında tamamlamış bulunmaktayım. 2013 yılında Yıldız Teknik Üniversitesi İstatistik Bölümü'nde lisans öğrenimimi tamamladım. 2012 yılından itibaren çeşitli özel bankalarda İnsan Kaynakları alanında Ücret Yönetimi uzmanlığı üzerine çalıştım. 2014 yılında Beykent Üniversitesi Endüstri Mühendisliği Tezli Yüksek Lisans bölümünde yüksek lisans eğitimine başladım.

Mevcutta özel bir bankada Performans ve Ücret Yönetimi Uzmanı olarak görev almaktayım.

Yabancı dilim İngilizcedir.

**Özge UYUMAZ**