

**T.C.  
CELAL BAYAR ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI  
ÜRETİM YÖNETİMİ VE PAZARLAMA  
PROGRAMI  
YÜKSEK LİSANS TEZİ**

**VERİ MADENCİLİĞİ VE TELEKOMÜNİKASYON  
SEKTÖRÜNDE UYGULAMASI**

**TUĞBA EKİNCİ**

**DANIŞMAN:  
YRD. DOÇ. DR. MUSTAFA GERŞİL**

**MANİSA**

**2009**

**T.C.  
CELAL BAYAR ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI  
ÜRETİM YÖNETİMİ VE PAZARLAMA  
PROGRAMI  
YÜKSEK LİSANS TEZİ**

**VERİ MADENCİLİĞİ VE TELEKOMÜNİKASYON  
SEKTÖRÜNDE UYGULAMASI**

**TUĞBA EKİNCİ**

**DANIŞMAN:  
YRD. DOÇ. DR. MUSTAFA GERŞİL**

**MANİSA**

**2009**

**YÜKSEKÖĞRETİM KURULU DOKÜMANTASYON MERKEZİ**

**TEZ VERİ FORMU**

**Tez No:**

**Konu:**

**Üniv. Kodu:**

**Not: Bu bölüm merkezimiz tarafından doldurulacaktır.**

**Tezin yazarının**

**Soyadı:** EKİNCİ

**Adı:** Tuğba

**Tezin Türkçe adı:**

Veri Madenciliği ve Telekomünikasyon Sektöründe Uygulaması

**Tezin Yabancı adı:**

Data Mining and An Application In Telecommunication Sector

**Tezin yapıldığı Üniversite:** CELAL BAYAR ÜNİVERSİTESİ

**Enstitü:** SOSYAL BİLİMLER ENSTİTÜSÜ

**Yılı:** 2009

**Diğer kuruluşlar:** -

**Tezin Türü:** 1- Yüksek Lisans (X)

**Dili:** Türkçe

2- Doktora

**Sayfa sayısı:** 111

3- Tıpta uzmanlık

**Referans sayısı:** 76

4- Sanatta yeterlilik

**Tez Danışmanlarının**

**Ünvanı:** Yrd. Doç. Dr.

**Adı:** Mustafa

**Soyadı:** GERŞİL

**Türkçe anahtar kelimeler:**

1- Veri Madenciliği

2- Müşteri İlişkileri Yönetimi

3- Ayrılan Müşteri Analizi

**İngilizce anahtar kelimeler:**

1- Data Mining

2- Customer Relationship Management

3- Churn Analysis

**Tarih:**

**İmza :**

## ÖZET

Son yıllarda oldukça yaygınlaşan ve önemi artan Veri Madenciliği, büyük veri kümeleri içindeki gizli kalmış, anlamlı bilgiyi ortaya çıkarma sürecidir. Veri madenciliğinin yaygın olarak kullanıldığı önemli alanlardan biri Müşteri İlişkileri Yönetimidir. Müşteri İlişkileri Yönetimi, müşterilerin elde tutulmasına, onlarla olan ilişkilerin geliştirilmesine ve müşteri memnuniyetinin artırılmasına odaklanan bir yaklaşımdır.

Bu tezin amacı, Veri Madenciliği ve Müşteri İlişkileri Yönetimi kavramlarını tanıtmak ve bunun yanı sıra telekomünikasyon sektöründe bir veri madenciliği modeli uygulamasını gerçekleştirmektir. Çalışma, telekomünikasyon sektöründe bir firmanın ayrılma eğilimi gösteren müşterilerinin profilini belirleyerek; bu müşterilere özel pazarlama stratejileri geliştirilmesini hedeflemektedir. Ayrılacak müşteri profilini belirlemek için sınıflama teknikleri kullanılmış ve uygulamanın sonuçları sunulmuştur.

## ABSTRACT

Data mining which has wide usage and becomes more important at recent years, is the process of finding hidden and unknown patterns in huge amounts of data. One of the widely applied areas of data mining is Customer Relationship Management. Customer Relationship Management is an approach which focuses on customer retention, relationship development and increase in customer satisfaction.

The purpose of this thesis is, to introduce data mining and customer relationship management in addition to make an application of a data mining model in telecommunication side. The study aims to determine the customers profile who likely to leave the firm and develop customized campaigns and marketing strategies. Classification techniques are used for determining the churners and the application results are presented.

Yüksek Lisans tezi olarak sunduđum “Veri Madenciliđi Ve Telekomünikasyon Sektöründe Uygulaması” adlı çalışmanın, tarafımdan bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldıđını ve yararlandıđım eserlerin bibliyografyada gösterilen eserlerden oluştuđunu, bunlara atıf yapılarak yararlanmıř olduđumu belirtir ve bunu onurumla dođrularım.

30/ 05/ 2009

Tuđba EKİNCİ

TEZ SAVUNMA SINAV TUTANAĞI

Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü 18.06.2009 tarih ve 18/2 sayılı toplantısında oluşturulan jürimiz tarafından Lisans Üstü öğretim Yönetmeliği'nin 24. Maddesi gereğince Enstitümüz İşletme Anabilim Dalı Üretim Yönetimi ve Pazarlama Yüksek Lisans Programı öğrencisi Tuğba EKİNCİ'nin "Veri Madenciliği ve Telekomünikasyon Sektöründe Uygulaması" Konulu tezi incelenmiş ve aday 25/06/2009 tarihinde saat 14:00'de jüri önünde tez savunmasına alınmıştır.

Adayın kişisel çalışmaya dayanan tezini savunmasından sonra 60 dakikalık süre içinde gerek tez konusu, gerekse tezin dayanağı olan anabilim dallarından jüri üyelerine sorulan sorulara verdiği cevaplar değerlendirilerek tezin,

BAŞARILI olduğuna



OY BİRLİĞİ



DÜZELTME yapılmasına



OY ÇOKLUĞU



RED edilmesine



ile karar verilmiştir.

\* Bu halde adaya 3 ay süre verilir.

\*\* Bu halde adayın kaydı silinir.



BAŞKAN  
Yrd.Doç.Dr. Mustafa GERŞİL  
(Danışman)

ÜYE

Prof.Dr. Cengiz YILMAZ



ÜYE

Doç.Dr. Özlem DOĞAN



Evet      Hayır

\*\*\* Tez, burs, ödül veya Teşvik prog. (Tüba, Fullbright vb.) aday olabilir

Tez, mutlaka basılmalıdır

Tez, mevcut haliyle basılmalıdır

Tez, gözden geçirildikten sonra basılmalıdır.

Tez, basımı gereksizdir.

## ÖNSÖZ

Çalışmalarım boyunca bilgisi ve görüşleriyle beni yönlendiren değerli danışman hocam Yrd. Doç. Dr. Mustafa Gerşil' e katkılarından dolayı teşekkürlerimi sunarım.

Ayrıca göstermiş olduğu yakın ilgi ve ayırdığı vakitler için Yrd. Doç. Dr. U. Tuğba Şimşek Gürsoy' a, tezimin uygulama bölümüne sağladığı değerli katkılarından dolayı Dr. Serkan Aydın' a teşekkürü bir borç bilirim.

Tezimin yazım aşamasında yaratıcı zekâsını, vaktini ve desteğini benden esirgemeyen sevgili H. Mehmet Arıdal' a, özel katkısından dolayı sevgili İpek Türker' e ve destekleriyle hep yanımda olan aileme sonsuz teşekkür ederim.

Yüksek lisans eğitimim boyunca emeği geçen tüm hocalarıma ayrıca teşekkür ederim.

**Tuğba Ekinci**

İstanbul

Mayıs 2009

## İÇİNDEKİLER

	Sayfa No
TEZ VERİ FORMU .....	I
ÖZET .....	II
ABSTRACT .....	II
YEMİN METNİ .....	III
TEZ SAVUNMA SINAV TUTANAĞI .....	IV
ÖNSÖZ .....	V
İÇİNDEKİLER .....	VI
ŞEKİLLER LİSTESİ .....	X
TABLolar LİSTESİ .....	XII
GİRİŞ .....	1



# BİRİNCİ BÖLÜM

## VERİ MADENCİLİĞİ

1.1. Veri Ambarı Nedir?.....	3
1.2. Veriye Erişim Araçları .....	5
1.2.1. Yapısal Sorgulama Dili (Structured Query Language - SQL) Aracı .....	5
1.2.2. Çevrimiçi Analitik İşleme (Online Analytical Processing - OLAP) Aracı.....	6
1.2.3. Veri Madenciliği (Data Mining) Aracı .....	7
1.3. Veri Madenciliği Nedir? .....	7
1.4. Veri Madenciliğine Neden İhtiyaç Duyulmuştur? .....	10
1.5. Literatürde Veri Madenciliği.....	11
1.6. Veri Madenciliği Süreci .....	13
1.6.1. İşi Anlama .....	14
1.6.2. Veriyi Anlama.....	15
1.6.3. Veriyi Hazırlama.....	16
1.6.4. Modelleme.....	18
1.6.5. Değerlendirme.....	19
1.6.6. Gerçekleştirme .....	20
1.7. Veri Madenciliğinde Karşılaşılan Sorunlar.....	21
1.8. Veri Madenciliğinin Uygulama Alanları .....	22
1.8.1. Müşteri İlişkileri Yönetimi.....	22
1.8.2. Pazarlama Yönetimi .....	23
1.8.3. Telekomünikasyon Sektörü.....	24
1.8.4. Perakende Sektörü.....	24
1.8.5. Bankacılık ve Finans Sektörü.....	25
1.8.6. Sağlık Sektörü .....	26
1.8.7. Diğer Uygulamalar .....	27
1.9. Veri Madenciliği Modelleri.....	29
1.9.1. Tahmin Edici Modeller .....	29
1.9.1.1. Sınıflandırma.....	29
1.9.1.2. Regresyon Analizi.....	37
1.9.2. Tanımlayıcı Modeller .....	39
1.9.2.1. Kümeleme Analizi .....	39
1.9.2.2. İlişki Analizi.....	40

## İKİNCİ BÖLÜM

### MÜŞTERİ İLİŞKİLERİ YÖNETİMİ

2.1. Müşteri İlişkileri Yönetimi Nedir?.....	42
2.2. CRM' nin Gelişimi.....	43
2.3. CRM' nin Faydaları .....	45
2.4. CRM Mimarisi .....	51
2.4.1. Operasyonel CRM.....	52
2.4.2. Analitik CRM.....	52
2.4.3. İşbirliksel CRM.....	53
2.5. CRM Süreci.....	53
2.6. Türkiye'de CRM Uygulamaları .....	57
2.7. CRM' de Veri Madenciliğinin Yeri .....	60

## ÜÇÜNCÜ BÖLÜM

### UYGULAMA

3.1. İşi Anlama .....	65
3.2. Veriyi Anlama .....	66
3.3. Veriyi Hazırlama .....	69
3.4. Modelleme.....	73
3.4.1. Aktif Ev Aboneleri İçin Modelleme .....	74
3.4.2. Deaktif Ev Aboneleri İçin Modelleme .....	80
3.5. Değerlendirme.....	85
3.5.1. Aktif Ev Aboneleri İçin Değerlendirme.....	85
3.5.2. Deaktif Ev Aboneleri İçin Değerlendirme .....	90
3.6. Gerçekleştirme .....	92
SONUÇ .....	93
BİBLİYOGRAFYA .....	97
EKLER.....	105
EK-1 Orijinal Veri Tablosu.....	106

## ŞEKİLLER LİSTESİ

Şekil 1.1 Veri Ambarı Mimarisi .....	4
Şekil 1.2 İş Zekâsı Çözümleri .....	5
Şekil 1.3 Veri Madenciliğinin İlişkide Olduğu Disiplinler.....	10
Şekil 1.4 Çapraz Endüstri Veri Madenciliği Standart Süreci.....	14
Şekil 1.5 CRISP-DM İş Anlama Basamağı .....	15
Şekil 1.6 CRISP-DM Veriyi Anlama Basamağı .....	16
Şekil 1.7 CRISP-DM Veriyi Hazırlama Basamağı .....	17
Şekil 1.8 CRISP-DM Modelleme Basamağı.....	19
Şekil 1.9 CRISP-DM Değerlendirme Basamağı .....	20
Şekil 1.10 CRISP-DM Gerçekleştirme Basamağı .....	21
Şekil 1.11 Veri Madenciliğinin Uygulandığı Alanlar .....	28
Şekil 1.12 Veri Madenciliği Modelleri .....	29
Şekil 1.13 Karar Ağacı Şekli.....	30
Şekil 1.14 Örnek Bir Karar Ağacı.....	32
Şekil 1.15 Yapay Sinir Ağı Mimarisinin Temel Elemanları.....	35
Şekil 1.16 Çok Katmanlı Bir Yapay Sinir Ağı.....	35
Şekil 1.17 K-En Yakın Komşu Örneği .....	36
Şekil 2.1 Pazarlama Anlayışının Tarihsel Gelişimi .....	45
Şekil 2.2 Müşteri Bağımlılığının Kar Üzerine Etkisi.....	46
Şekil 2.3 CRM' nin Etkileri .....	49
Şekil 2.4 Müşteri İlişkileri Yönetimi Mimarisi.....	51
Şekil 2.5 CRM Evreleri.....	55
Şekil 2.6 Türk Şirketlerinde CRM Yaklaşımı.....	56
Şekil 2.7 CRM' de Verileri Anlamlı Bilgilere Dönüştürme Süreci .....	61
Şekil 3.1 Uygulama Periyodu - 1 .....	64
Şekil 3.2 Uygulama Periyodu - 2 .....	64
Şekil 3.3 SPSS Clementine Programına Ait Bir Arayüz .....	65
Şekil 3.4 Data Audit Nodu .....	68
Şekil 3.5 Veri Hazırlama Aşaması .....	72
Şekil 3.6 Aktif Ev Abonelerinin Dağılımı - 1 .....	74

Şekil 3.7 Aktif Ev Abonelerinin Dağılımı - 2.....	75
Şekil 3.8 Aktif Ev Aboneleri İçin Model Geliştirme -1.....	75
Şekil 3.9 Aktif Ev Aboneleri İçin Model Geliştirme -2.....	76
Şekil 3.10 Deaktif Abonelerin Dağılımları -1.....	80
Şekil 3.11 Deaktif Abonelerin Dağılımları -2.....	80
Şekil 3.12 Deaktif Ev Aboneleri İçin Geliştirilen Spss Clementine Akımı.....	81
Şekil 3.13 Deaktif Aboneler İçin Quest Karar Ağacı .....	82
Şekil 3.14 Olasılığa Göre İşletmeden Ayrılma Dağılımı.....	87
Şekil 3.15 Model Analizi .....	90
Şekil 3.16 Modelin Kazancı.....	90

## TABLULAR LİSTESİ

Tablo 1.1 Veri Madenciliğinin Gelişim Süreci .....	9
Tablo 1.2 Karar Ağacı İçin Örnek Veri Seti .....	31
Tablo 2.1 CRM' nin İşletmelere Rakamsal Etkileri.....	49
Tablo 3.1 Değişkenler ve Açıklamaları .....	67
Tablo 3.2 Verinin Düzenlenmeden Önceki Hali.....	69
Tablo 3.3 Verinin Düzenlendikten Sonraki Hali.....	73
Tablo 3.4 Lojistik Regresyon Sonuçları - 1 .....	77
Tablo 3.5 Lojistik Regresyon Sonuçları - 2 .....	78
Tablo 3.6 Lojistik Regresyon Sonuçları - 3 .....	79
Tablo 3.7 Anlamlılık Testi .....	85
Tablo 3.8 Model Özeti .....	86
Tablo 3.9 Sınıflandırma Oranları .....	86
Tablo 3.10 İşletmeden Ayrılan Aktif Abonelerin Özellikleri - 1.....	88
Tablo 3.11 İşletmeden Ayrılan Aktif Abonelerin Özellikleri - 2.....	89
Tablo 3.12 Aktif Abonelerin İşletmeden Ayrılma Sebepleri.....	89
Tablo 3.13 İşletmeden Ayrılan Deaktif Abonelerin Özellikleri.....	91
Tablo 3.14 Deaktif Abonelerin İşletmeden Ayrılma Sebepleri .....	91

## GİRİŞ

Coğrafi sınırların kalktığı günümüzde, müşteriler istedikleri ürünü istedikleri anda temin edebilmektedir. Çok kısa bir sürede ürünlerin benzerlerinin daha ucuz fiyata pazarlara girmeye başlamasıyla, işletmeler müşterilerinin sadakatini garanti edemez hale gelmişler ve onların kişisel tercihlerine daha fazla önem vermeye başlamışlardır.

Yoğun rekabet ortamında, işletmeler rakiplerinin bir adım önüne geçebilmek için, bilgiyi etkin ve verimli bir biçimde kullanmak zorunda olduklarının bilincindedir. Bilgiyi elde edebilmek, ilişkide olunan müşteriyi yakından tanımayı, onun her alışverişinde bıraktığı veriyi toplamayı, depolamayı ve sistematik bir şekilde kullanmayı gerektirir. Depolanan veriyi işleyerek bilgiye dönüştürmede veri madenciliği, bu bilginin, müşteri ile olan ilişkileri güçlendirmesi, müşteri memnuniyetini kâra dönüştürmesi amacıyla yapılan çalışmalarda müşteri ilişkileri yönetimi kullanılmaktadır.

İşletme, Veri Madenciliği sayesinde, müşteri davranışlarını açıklamak ve ileriye yönelik tahminler yapabilmek için modeller kurmakta, Müşteri İlişkileri Yönetimi sayesinde ise, müşterileriyle ilişkilerini güçlendirmek, müşteri kaybını azaltmak ve yeni müşteriler kazanmak için bu tahminlerden stratejiler geliştirmektedir.

Bu tez çalışmasının temel amacı, işletmenin veri ambarındaki verilerinin içerisinde saklı bulunan bilgilerin Veri Madenciliği teknikleri ile ortaya çıkarılması ve Müşteri İlişkileri Yönetimi kapsamında stratejilere dönüştürülmesidir.

Tezin birinci bölümünde Veri Ambarı kavramı ele alınmış, Veri Ambarlarında tutulan veriye, erişim yöntemleri açıklanmış, bu yöntemlerden biri olan Veri Madenciliği kavramı tüm yönleriyle ele alınmış, Veri Madenciliği Süreci, bu süreçte karşılaşılan sorunlar, Veri Madenciliğinin yaygın olarak kullanıldığı alanlar, veri analizinin yapılabilmesi için kullanılan teknikler ayrıntılarıyla verilmiştir.

Tezin ikinci bölümünde Müşteri İlişkileri Yönetimi kavramı ele alınmış; Müşteri İlişkileri Yönetimi kavramının ortaya çıkışı, faydaları, süreci ve mimarisine değinilmiş, Türkiye’de gerçekleştirilmiş olan uygulamalara yer verilmiştir.

Bu tez çalışmasının üçüncü bölümü olan uygulama bölümünde ise Türkiye’ de Telekomünikasyon sektörünün öncü firmalarından birinin Mart 2008 - Ağustos 2008 dönemleri arasındaki verileri kullanılarak, işletmeden ayrılma eğilimi gösteren müşterileri tespit edilmiştir. Uygulama çalışması SPSS Clementine programı ile yapılmış, Regresyon Analizi ve Karar Ağacı Algoritmalarından faydalanılmıştır. Uygulamanın detayları adım adım ortaya konmuş, elde edilen sonuçlar değerlendirilmiş ve müşterileri elde tutmak için öneriler geliştirilmiştir.



# 1. BÖLÜM

## VERİ MADENCİLİĞİ

Birçok işletme bir insanın ömrü boyunca okuyabileceğinden daha fazla veriyi çok kısa sürede üretmektedir. Bunun sonucu olarak veri eksikliği, enformasyon yetersizliği gibi problemler yerini günümüzde veriyle baş edebilme problemine bırakmıştır. Bu problemi bertaraf etmek için çok sayıda kuruluş veri madenciliğine başvurmaktadır.

### 1.1. Veri Ambarı Nedir?

Bir işletmenin değişik bölümleri tarafından toplanan müşteri bilgilerinin, ileride değerlendirilmek üzere arka plandaki sistemde birleştirilmesinden oluşan geniş ölçekli veri deposudur. Toplanan bilgilerin sistematik olarak bir veritabanına kaydedilmesiyle elde edilen “müşteri bilgi sistemi” veri ambarı olarak adlandırılır (Swift, 2001, s.71).

Diğer bir ifadeyle, veri ambarı karar verme sürecinde kullanılan, konu tabanlı, birleştirilmiş, zamana bağımlı, verilerin kalıcı olduğu veri topluluğudur (Agrawal ve Srikant, 1994, s.5). Veri ambarının bu dört özelliği kısaca aşağıda açıklanmıştır:

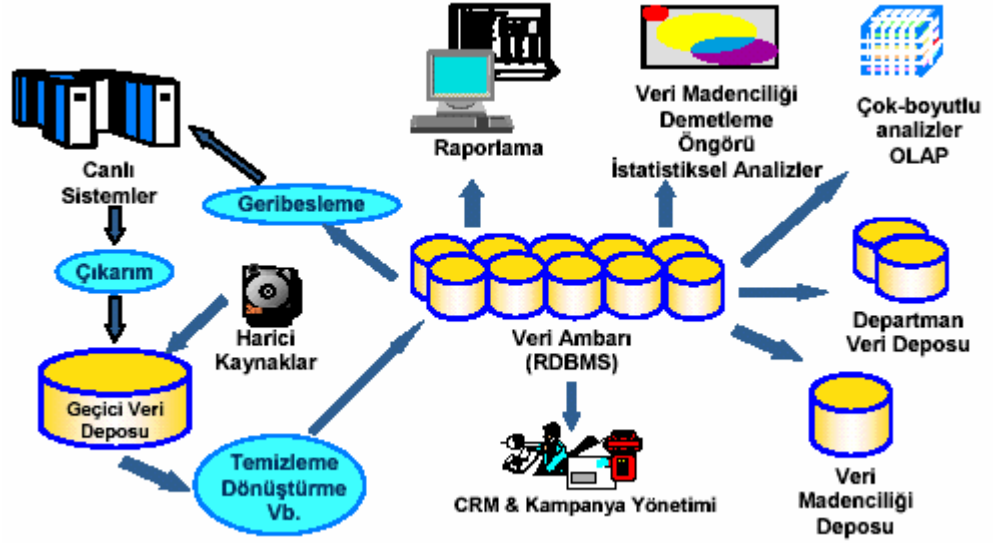
**Konu tabanlı:** Veri ambarları satış verileri, müşteri bilgileri gibi belirli bir konuya ait veriler içerir.

**Birleştirilmiş:** Veri ambarları birçok farklı kaynaktan gelen bilgilerin toplanması ile kurulur. Örneğin bir veri ambarı içinde ilişkisel veritabanları, düz metin dosyaları, işlemsel veritabanları bulunabilir.

**Zamana bağımlı:** Veri ambarlarında bilgiler periyodik aralıklarla eklenir. Veri ambarındaki her bir anahtar yapı, tarihsel olarak dizilmiş olmalıdır. Örneğin günlere göre son beş yılın satış rakamları veri ambarında bulunabilir.

**Kalıcı:** Veri ambarında veriler işlevsel veritabanlarında olduğu gibi sürekli güncellenmez. Veri ambarına eklendiği andan itibaren kalıcı olarak kaydedilir.

Şekil 1.1 Veri Ambarı Mimarisi



Kaynak: Berry M., Linoff, G., Mastering Data Mining: The Art and Science of Customer Relationship Management, John Wiley & Sons, 2000.

Şekil 1.1’ de görüldüğü gibi, veri ambarına işletmeye ait canlı sistemlerden veri girişi olabildiği gibi, işletmeye ait olmayan kaynaklardan da veri girişi mümkün olmaktadır. Canlı sistemler, günlük yapılan işleri ve işlemleri gerçekleştiren, sonuçları saklayan sistemlerdir. Bu sistemler marketlerde ya da mağazalarda stok takibi, üye borçları, satış işlemleri, ödeme kayıtları gibi bilgilerin işlendiği ve tutulduğu bilgi sistemleri olabilmektedir. Veri ambarları ve canlı sistemler arasında geri besleme ile bir ilişki kurulmuştur. Bu sayede daha iyi kararlar alınması hedeflenmiştir.

Veri ambarındaki verinin doğru ve tek tip olması için, veriler toplandıktan sonra çeşitli temizleme ve dönüştürme işlemine tabi tutulurlar. Çünkü işletmenin değişik yerlerinde aynı veriler birden fazla kez tekrar ediliyor olabilir. Normal olmayan hatalı verilerin tespit edilip temizlenmesi ve verilerin standart hale getirilmesi için dönüştürme işlemlerinin yapılması gerekmektedir.

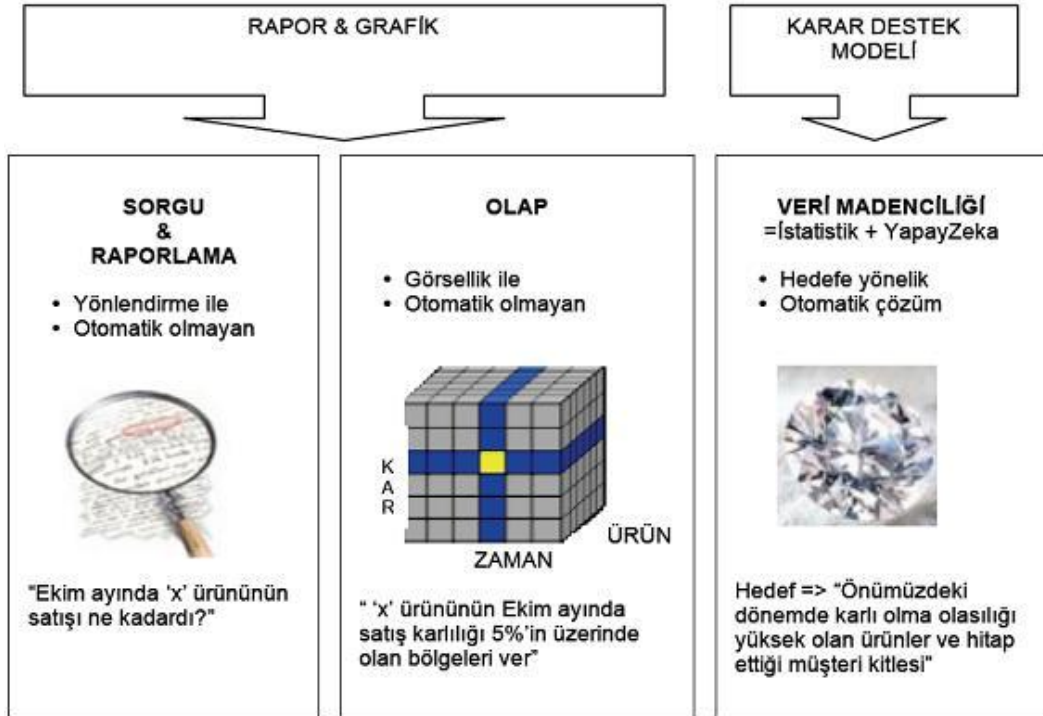
Veri Ambarı üzerinde veri madenciliği uygulamaları, çevrimiçi analitik işleme (OLAP), müşteri ilişkileri yönetimi (CRM), istatistiksel analiz ve raporlama işlemleri gerçekleştirilebilir (Han ve Kamber, 2001, s.146). Bu işlemler, ilişkisel veritabanları

üzerinde de gerçekleştirilebilmesine rağmen, veri ambarı kurulduğunda daha yüksek performans elde edilir.

## 1.2. Veriye Erişim Araçları

Veri ambarlarında veriden çıkarılabilecek dört farklı bilgi tipi vardır. Bunlar; SQL gibi bir sorgulama aracıyla kolaylıkla çıkarılabilen yüzeydeki bilgi, OLAP araçları kullanılarak analiz edilebilen çok boyutlu bilgi, veri madenciliği araçları ile çıkarılabilen gizli bilgi ve derindeki bilgidir.

Şekil 1.2 İş Zekâsı Çözümleri



Kaynak: Analitik CRM: İş Zekâsı Çözümleri (<http://www.crminturkey.org>)

### 1.2.1. Yapısal Sorgulama Dili (Structured Query Language - SQL) Aracı

SQL, ilişkisel veri tabanlarındaki bilgileri sorgulamak için standart kullanımı olan bir dildir. Standart bir dil olmasına karşılık, çeşitli veri tabanlarında SQL kullanımları arasında farklılıklar vardır. SQL komutları ile, tablolara yeni kayıt girme, var olan kayıtları sorgulama (arama ve listeleme), var olan bilgileri değiştirme ve var olan kayıtları silme işlemleri yapılabilir (<http://science.ankara.edu.tr>).

### **Sorgulama:**

SELECT [ALL|DISTINCT] { \*|alan\_adi\_listesi }

[INTO deęişken\_listesi ]

FROM tablo-ismi

[WHERE koşul]

[GROUP BY alan\_adi1 [ , alan\_adi2] ...]

[HAVING koşul]

[ORDER BY sıralama\_alanları] şeklinde genel bir ifadesi vardır. Burada;

[ ] seçimlik

{ } zorunlu

| veya anlamına gelmektedir (<http://seminer.linux.org.tr>).

Kullanıcı hazırlayacağı sorgu ile veritabanında istedięi tablodan, belirttięi koşullara uyan kayıtların tümünü ya da sadece bazı alanlarını, istedięi koşullarda gruplayarak ve/veya sıralayarak görebilir. İsterse gördüğü bu kayıtları sorgusunda “into” deyimini kullanarak veri tabanında yaratacağı yeni bir tabloya aktarabilir.

### **1.2.2. Çevrimiçi Analitik İşleme (Online Analytical Processing - OLAP) Aracı**

Veri ambarlarının gelişmesi ile beraber, verilere daha hızlı şekilde erişme ve çok boyutlu analiz ihtiyaçları, şirketleri daha farklı yapılar geliştirmeye itmiştir. Bu amaçla geliştirilen OLAP, uzman kişinin sisteme olan bakış açısını, sistemi kullanan kişilerin anlayacağı biçimde ifade ederek, ham veriden dönüştürülmüş bilgi üzerinde mümkün olan incelemelerin büyük çoğunluęuna hızlı, tutarlı, interaktif biçimde erişip analizcilere, yöneticilere ve uygulayıcılara veriden çeşitli kavrayışlar sağlayan bir yazılım teknolojisi kategorisidir (Vassiliadis ve Sellis, 1999, s.64). OLAP, verilere çok hızlı bir şekilde erişebilme ve çok boyutlu analiz ihtiyaçlarının teminine yönelik geliştirilmiştir (Berson, Smith ve Thearling, 2000, s.47).

OLAP, kullanıcının aklındaki sorulara bağılı olarak sınırlı sonuçlar döndürmekte ve sınırlı analiz imkânı sunmaktadır. Kullanıcı, bira alan kişilerin çerez de alabileceğini tahmin edebileceğinden bu fikrini OLAP ile kolaylıkla sorgulayabilirken, bira ve çocuk bezi arasındaki bağıntıya ulaşamayabilir. Bu noktada veriye erişim aracı olarak karşımıza, görünmeyen bağıntıların ve gizli bilgilerin ortaya çıkmasını sağlayan veri madenciliği analiz yöntemi çıkmaktadır.

### **1.2.3. Veri Madenciliği (Data Mining) Aracı**

İlerleyen kısımlarda daha detaylı olarak incelenecek olan veri madenciliği, yapay zekâ, istatistiksel ve matematiksel teknikler kullanarak veri ambarlarında depolanan büyük miktardaki veriden anlamlı yeni korelasyonlar, desenler ve trendler keşfetme süreci olarak tanımlanabilir. Veri madenciliğini çekici kılan en önemli özelliği geçmişe dönük modellerden ziyade geleceğe yönelik öngörülerde bulunmayı sağlayan modeller oluşturmasıdır.

### **1.3. Veri Madenciliği Nedir?**

Veri madenciliği tanımlayıcının bakış açısına, ilgi alanına ve görüşüne göre değişen tanımlara sahiptir. Literatürde bulunan veri madenciliği ile ilgili bazı tanımlar aşağıda verilmiştir (Friedman, 1997, s.1-2):

Veri madenciliği;

- Verideki anlamlı, orijinal, potansiyel olarak kullanışlı ve sonunda anlaşılabilir olan örüntülerin tanımlanmasında önemli bir süreçtir (Fayyad),
- Önceden bilinmeyen, anlaşılabilir ve etkin bilginin büyük veritabanlarından çıkarılma sürecidir (Zekulin),
- Anlamlı kuralların ve örüntülerin bulunması için geniş veri yığınları üzerine yapılan keşif ve analiz işlemleridir (Berry ve Linoff),
- Bilgi keşfi sürecinde veri içerisindeki önceden bilinmeyen ilişkileri ve örüntüleri ayırt etmek için kullanılan yöntemler kümesidir (Ferruzza),

- Daha önceden bilinmeyen ve potansiyel olarak yararlı olma durumuna sahip verinin keşfedilmesidir (Frawley),
- Veri madenciliği, veri içerisindeki faydalı örüntülerin keşfinin sürecidir (John),
- Bilinmeyen ve beklenmeyen örüntülerin büyük veritabanlarında arandığı bir karar destek sürecidir (Parsaye),
- Saklı fakat potansiyel olarak kullanışlı bilgiyi bulmak için kullanılan arama ve veri çözümleme sürecidir (Shaw).

Veri madenciliği uygulamaları, 1960'lı yıllarda IBM ve CDC gibi firmalar tarafından günün teknolojisi olan bilgisayarlar ve diskler kullanılarak, statik veri toplama şeklinde başlamıştır. Daha sonra 1980'lerde Oracle, Microsoft gibi firmalar tarafından ilişkisel veri tabanları ve SQL aracılığıyla, kayıt bazında dinamik veri toplama ile devam etmiştir. 1990'lara gelindiğinde SPSS, NCR gibi firmalar tarafından OLAP ve veri ambarları kullanılarak çok boyutlu dinamik veriler toplanmaya başlanmıştır. Günümüzde ise SPSS, SAS, IBM gelişmiş algoritmalar sayesinde proaktif olarak veri toplayabilmektedir.

Veri analizinin tarihsel gelişimi Tablo 1.1' de ayrıntılı olarak görülmektedir.

**Tablo 1.1 Veri Madenciliğinin Gelişim Süreci**

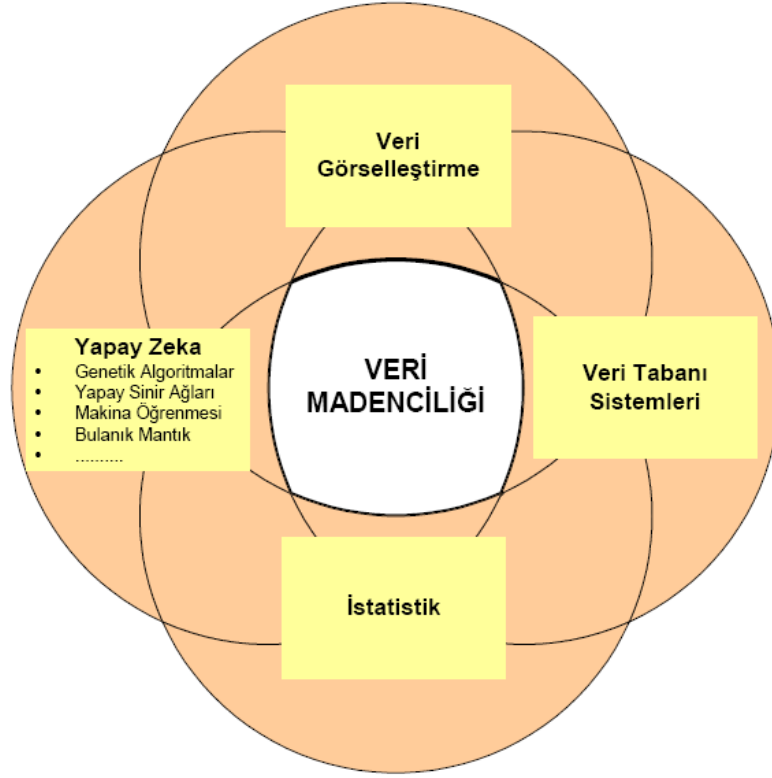
<b>ADIMLAR</b>	<b>SORUNLAR</b>	<b>TEKNOLOJİ</b>	<b>TEDARİKÇİ</b>	<b>ÖZELLİK</b>
Veri Toplama (1960)	Son 5 yılda firmanın toplam karı nedir?	Bilgisayar, Diskler	IBM, CDC	Statik veri toplama
Veri Erişimi (1980)	Geçen Mart ayında İngiltere’de gerçekleşen toplam satış miktarı nedir?	İlişkisel veri tabanları, SQL	Oracle, Sybase, Informix, IBM, Microsoft	Kayıt bazında dinamik veri toplama
Veri Ambarı ve Karar Destek Sistemleri (1990)	Geçen Mart ayında İngiltere’de gerçekleşen toplam satış miktarı nedir? Boston’da ne kadardır?	OLAP, Çok Boyutlu Veritabanları, Veri Ambarı	SPSS, Comshare, Arbor, Cognos, Microstrategy, NCR	Çok boyutlu dinamik veri toplama
Veri Madenciliği (2000)	Önümüzdeki ay Boston’da satışlar ne düzeyde gerçekleşecektir?	Çoklu Veritabanları, Gelişmiş Algoritmalar	SPSS/Clementine, Lockheed, IBM, SGI, SAS, NCR, Oracle	Proaktif veri toplama

Kaynak: Squier, L., What Is Data Mining, Data Management Association-NCR Presentation, 2001.

Veri madenciliği, veritabanı sistemleri, istatistik, yapay zekâ, veri görselleştirme alanlarının bir kesişim noktası olarak doğmuş ve bu bağlamda gelişmesini sürdürmektedir.

Bu yapı Şekil 1.3’ de sembolize edilmiştir.

**Şekil 1.3 Veri Madenciliğinin İlişkide Olduğu Disiplinler**



Kaynak: Baykasoğlu, A., Veri Madenciliği ve Çimento Sektöründe Bir Uygulama, Akademik Bilişim, Gaziantep, 2005.

#### **1.4. Veri Madenciliğine Neden İhtiyaç Duyulmuştur?**

Veri madenciliği son yıllarda büyük bir kullanım hacmine ve popülerliğe ulaşmıştır. Günümüzde, veri madenciliğinin ne kadar yaygın olarak kullanıldığını işaret eden bir örnek olarak; Amerikan profesyonel basketbol ligi (NBA)'de mücadele eden 29 takımdan 16'sının IBM firmasının NBA için özel olarak hazırladığı "Advanced Scout" adlı veri madenciliği yazılımını kullanmaları verilebilir. Uygulama; basket, şut, pas, rebound, asist sayıları gibi NBA tarafından tutulan veriye dayanmaktadır. Yapılan veri madenciliği analizi sonucunda, takım koçlarının bile fark edemediği desenler bulunmuş ve bu durumlara karşı alınan önlemler neticesinde, takım çok daha iyi yerlere gelmiştir. İlk olarak, New York Nicks takımının bu başarısının ardından, diğer NBA takımlarının çoğu, aynı veri madenciliği yazılımını kullanmaya başlamışlardır (Larose, 2005, s.3).



Ürün ve hizmet sektöründe, veri madenciliği uygulamalarına duyulan ihtiyaca ilişkin bazı örnekler aşağıda sıralanmıştır (Özmen, 2001, s.3);

- Bir işletme kendi müşterisiyken rakibine giden müşterilerle ilgili analizler yaparak rakiplerini tercih eden müşterilerinin özelliklerini elde edebilir ve bundan yola çıkarak gelecek dönemlerde kaybetme olasılığı olan müşterilerin kimler olabileceği yolunda tahminlerde bulunarak onları kaybetmemek, kaybettiklerini geri kazanmak için strateji geliştirebilir.

- Ürün veya hizmette hangi özelliklerin ne derecede müşteri memnuniyetini etkilediği, hangi özelliklerinden dolayı müşterinin bunları tercih ettiği ortaya çıkarılabilir.

- En karlı mevcut müşteriler saptanarak, potansiyel müşteriler arasından en karlı olabilecekler belirlenebilir. Karlı müşteriler tespit edilerek onlara özel kampanyalar uygulanabilir. En masraflı müşteriler daha masrafsız müşteri haline dönüştürülebilir. Örneğin en çok bankacılık işlemi yapanlar ortaya çıkarılıp bunlar şube bankacılığı yerine daha masrafsız İnternet bankacılığına yönlendirilebilir.

- Bir ürün veya hizmetle ilgili bir kampanya programı oluşturmak için hedef kitlenin seçiminden başlayarak bunun hedef kitleye hangi kanallardan sunulacağı kararına kadar olan süreçte veri madenciliği kullanılabilir.

### **1.5. Literatürde Veri Madenciliği**

Veri madenciliği 1990'lı yıllarda ortaya çıkmıştır. Bir online veritabanı olan Science Direct'te 1980'den 2008 yılına kadar bir literatür taraması yapıldığında veri madenciliği ile ilgili 3492 makale olduğu görülmektedir. Tarama 2000 yılından günümüze kadar yapıldığında ise, 2849 makale bulunmaktadır. Bu da veri madenciliğinin özellikle 2000 yılından bu yana büyük bir gelişme gösterdiğinin bir kanıtıdır. Aşağıda 2004 – 2008 tarihleri arasında veri madenciliği konusunda farklı alanlarda gerçekleştirilen örnekler yer almaktadır ([www.sciencedirect.com](http://www.sciencedirect.com)).

Minami ve Dawson (2008), Japonya'daki hizmet ve perakende sektöründeki firmalarda, CRM' nin gelir ve sadakat geliştirme üzerinde doğrudan etkisi olup olmadığını araştırmışlardır. Guo ve Savickas (2008), finansal değişkenleri kullanarak döviz kurlarını tahminlemeye çalışmışlardır. Recabarren, Nussbaum ve Leiva (2008), Catolica de Chile Üniversitesinin mühendislik, eğitim, hemşirelik gibi değişik bölümlerinde okuyan birinci sınıf öğrencileri üzerinde araştırma yaparak, kültür ve internet kullanımı arasındaki ilişkiyi ortaya koymuşlardır. Caussement ve Poel (2008), geleneksel pazarlama bilgileriyle çalışan müşteri kayıplarını öngörme sistemine, çağrı merkezine gelen müşteri e-postalarını da eklemişler ve kaybetmek üzere oldukları müşterileri önceden belirleyebilmek için karar destek sistemi oluşturmuşlardır.

Abrahams ve diğerleri (2007), bir sigorta firması için en karlı müşterilerin demografik özelliklerini karar ağacı yaklaşımıyla belirlemişlerdir. Yang ve Huang (2007), TCP paketlerindeki değişimi inceleyerek saldırı tespiti için yaklaşımlar geliştirmişlerdir. Zhang, Edwards ve Harding (2007), web kullanım madenciliği teknikleriyle, online satışların kişileştirilmesini incelemişlerdir. Kusiak ve Smith (2007), veri madenciliği tekniklerini yeni ürün tasarımı ve üretim sistemi tasarımında kullanmışlardır. Bunun için ilk olarak, hızlı satın alma kararı veren, promosyonlara ilgi gösteren ya da nadir satın alma eylemi gerçekleştiren müşterilerin özelliklerini belirleyip onları gruplandırmışlardır.

Boginski ve Butensko (2006), çapraz korelasyon ve kümeleme tekniklerini kullanarak hisse senedi piyasalarının yapısal özelliklerini ortaya koymuşlardır. Jiao, Zhang ve Helande (2006), Kansai haritalama tekniği ile bir karar destek sistemi tasarlamışlardır. Jeng, Chen ve Liang (2006), genetik algoritma ile biyolojik sistemlerin kinetik parametrelerini belirlemişlerdir. Cervone, Kafatos ve Singh (2006), veri madenciliği tekniklerini kullanarak bir deprem erken uyarı sistemi geliştirmişlerdir. Kim ve Song (2006), müşteri davranışlarındaki değişimleri belirleyebilmek için değişimleri; gelişmekte olan, kötüleşmekte olan ve beklenmeyen olmak üzere üç gruba ayırıp, bunlar arasındaki benzerlik ve farklılıkları ölçerek değişimin derecesini belirlemişler ve kurdukları modeli karar ağacı analizlerinden de yararlanarak, internet üzerinden Kore'de faaliyet gösteren alışveriş merkezinde uygulamışlardır.

Facca ve Lanzi (2005), webde tutulan verileri analiz etmek için makine öğrenme algoritmalarını kullanmışlardır. Hong, Park, Jon ve Rho (2005), veri madenciliği tekniklerini kullanarak bir tedarikçi seçim modeli önermişlerdir. Huang, Chen ve Wu (2005), kümeleme tekniklerini kullanarak dağıtım merkezleri için bir sipariş yönetim sistemi geliştirmişlerdir. Liu ve Shih (2005), AHP ve birliktelik kuralı yöntemlerini birleştiren bir öneri metodolojisi geliştirerek müşterileri yaşam ömrü değerlerine göre bölümlere(segmentlere) ayırmışlardır.

Hsieh (2004), bir bankanın kredi kartı müşterileri için sinir ağları ve birliktelik kurallarını kullanarak bir davranışsal skorlama modeli yaratmış; müşterileri davranış ve karakteristiklerine göre farklı gruplara ayırarak her grubun karakteristiklerine uygun yönetim stratejileri önermek suretiyle müşteri sadakatini arttırmayı hedeflemiştir.

### **1.6. Veri Madenciliği Süreci**

Veri madenciliği çalışmalarının başarılı bir şekilde sonuçlanabilmesi için sistematik bir yaklaşım zorunludur. Veri madenciliği araçları sunan birçok satıcı, danışman firma ve araştırmacı kullanıcılara yönelik birçok süreç modelleri oluşturmuşlardır.

Örneğin, SAS SEMMA (Sample, Explore, Modify, Model, Assess) metodolojisini, NCR System Engineering (ABD-Danimarka), DaimlerChrysler AG (Almanya), SPSS Inc. (ABD) ve OHRA Verzekeringen en Bank Groep B.V (Hollanda) isimli firmaların oluşturduğu konsorsiyum ise, CRISP-DM (Cross-Industry Standard Process for Data Mining) metodolojisini geliştirmiştir.

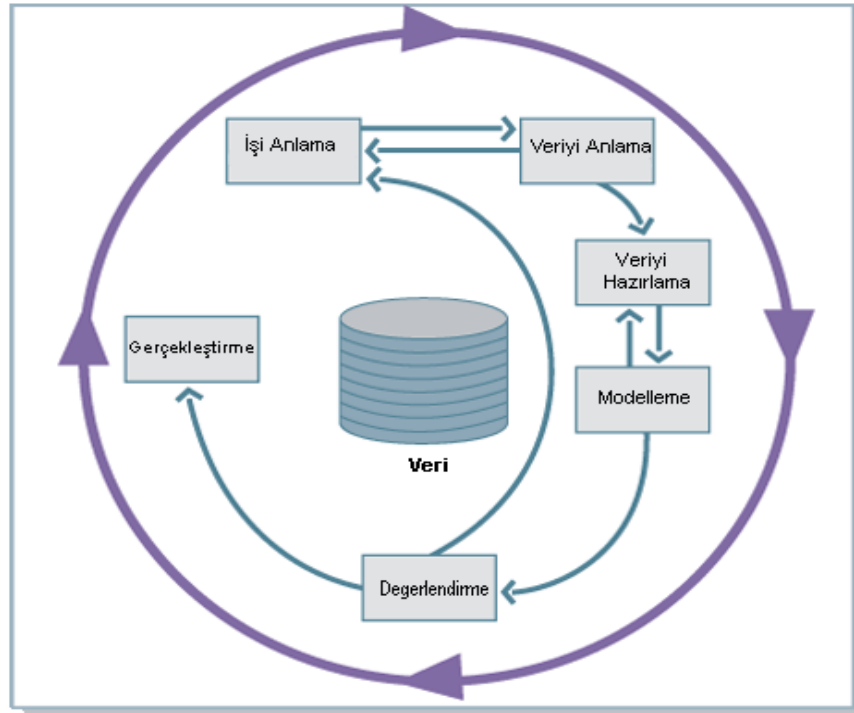
Sözcüklerin baş harfleri “Çapraz Endüstri Veri Madenciliği Standart Süreci” açılımında olan CRISP-DM, bilgi keşfi için yapılacak veri madenciliğinin temel adımlarını tanımlayan hiyerarşik bir süreçtir ([www.crisp-dm.org](http://www.crisp-dm.org)).

CRISP-DM süreci 6 aşamadan oluşmaktadır:

1. İşi Anlama
2. Veriyi Anlama
3. Veriyi Hazırlama
4. Modelleme
5. Değerlendirme
6. Gerçekleştirme

Şekil 1.4' te bu süreç görülmektedir.

**Şekil 1.4 Çapraz Endüstri Veri Madenciliği Standart Süreci**



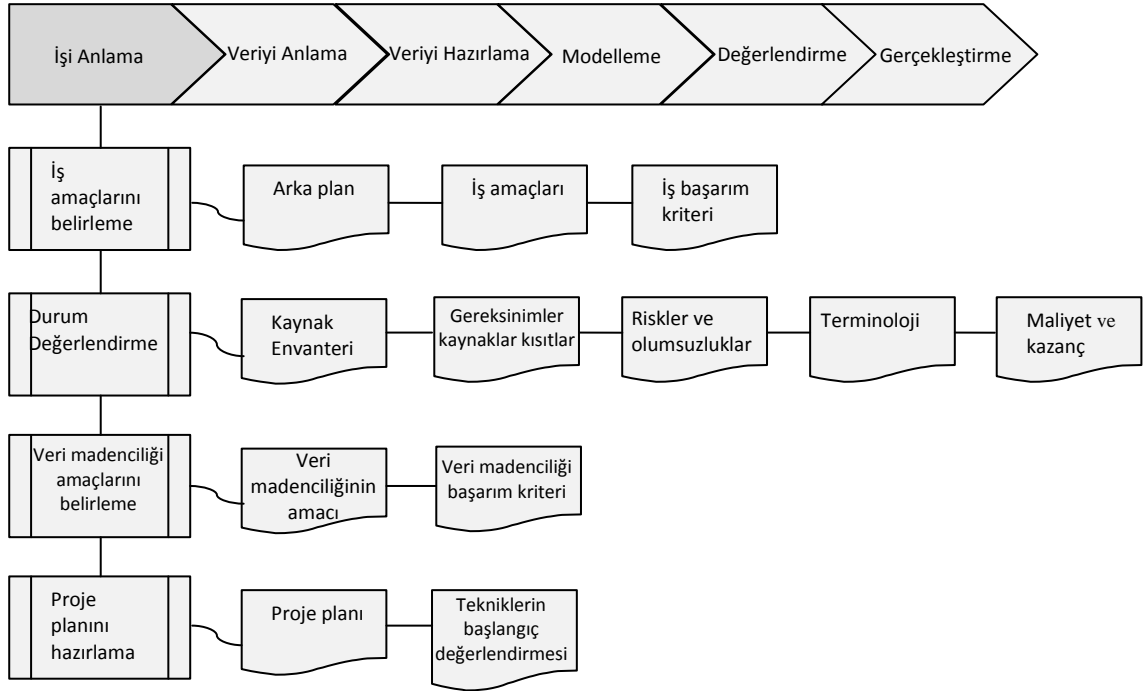
Kaynak:Process Model (www.crisp-dm.org)

### 1.6.1. İşi Anlama

CRISP-DM yöntem biliminin bu ilk adımı, projenin amaç ve gereksinimlerini, iş bakış açısı ile anlamaya ve sonrasında bu bilgiyi, veri madenciliği problemini tanımaya

ve amaçları gerçekleştirmeye yönelik bir ilk plan haline dönüştürmeye yoğunlaşmıştır (Chapman ve diğerleri, 2000, s.16).

**Şekil 1.5 CRISP-DM İşi Anlama Basamağı**

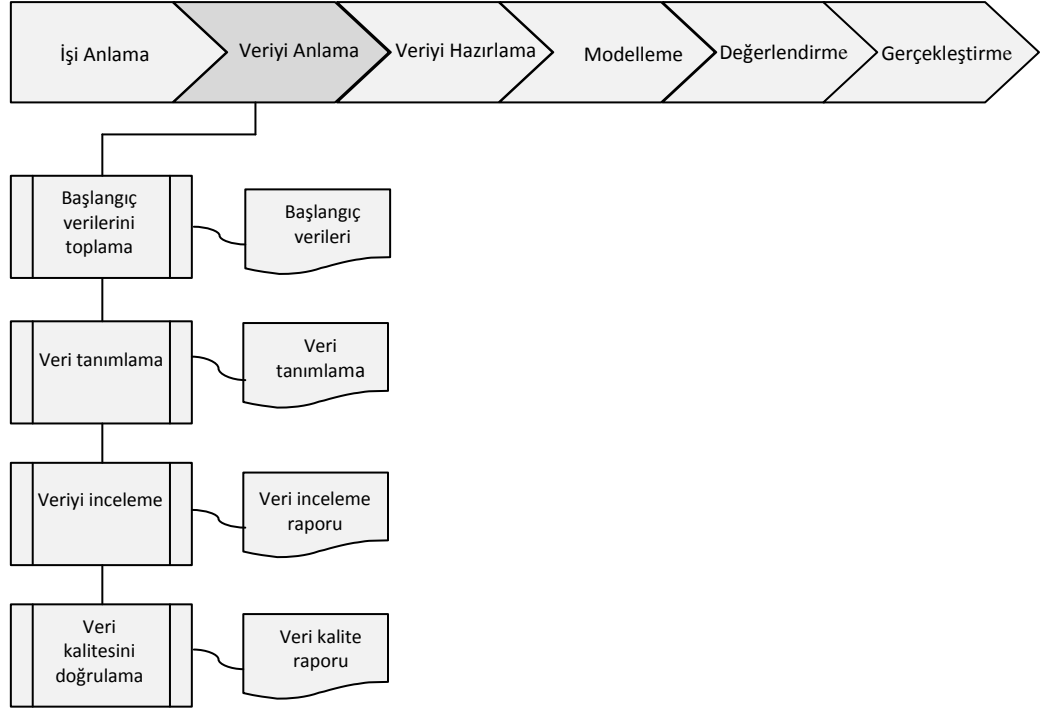


Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

### 1.6.2. Veriyi Anlama

Veriyi anlama basamağında ilk yapılması gereken işlem, verinin toplanmasıdır. Veri toplandıktan sonra, verinin daha yakından tanınması ve kalite problemlerinin belirlenebilmesi için incelenmesi gerekmektedir. Ancak veri incelenirken en önemli nokta verinin değiştirilmemesidir.

**Şekil 1.6 CRISP-DM Veriyi Anlama Basamağı**



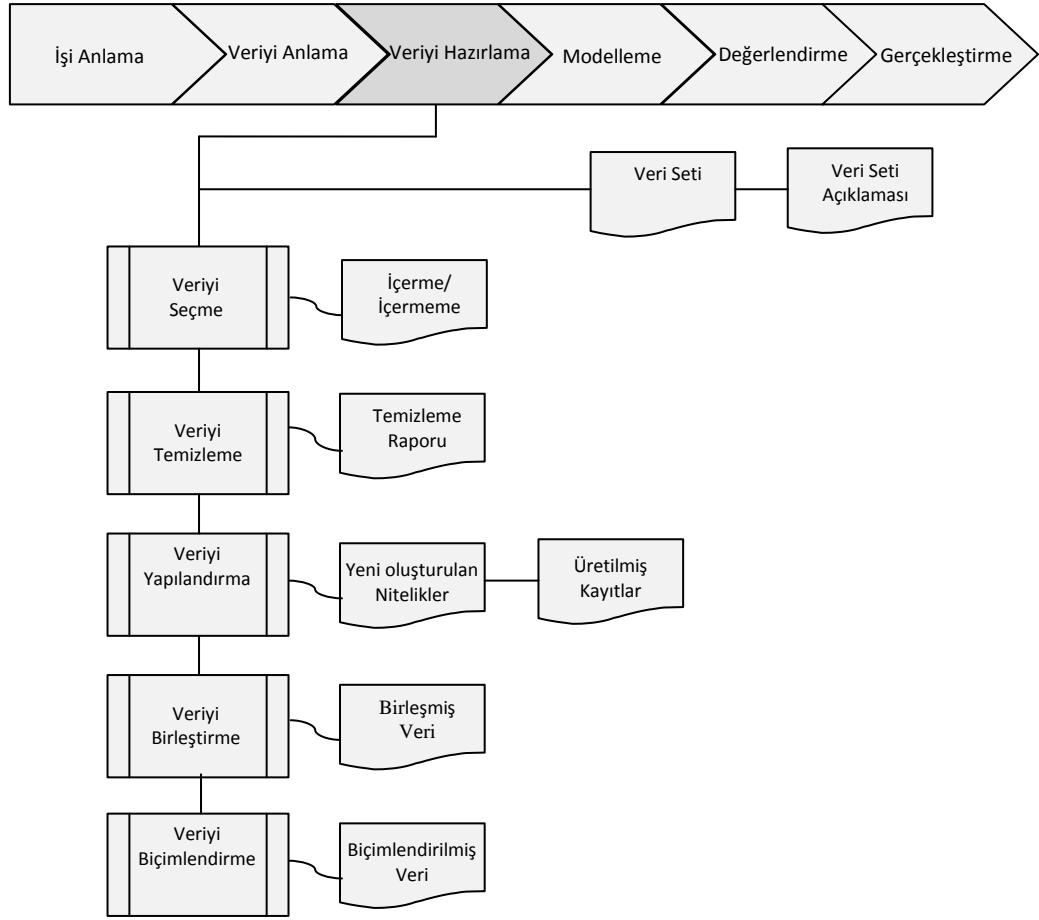
Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

### 1.6.3. Veriyi Hazırlama

Veriyi hazırlama basamağının amacı, verileri modelleme için elverişli hale getirerek, veri madenciliği uygulamalarının performansını arttırmaktır. Diğer bir deyişle, işlenmemiş verinin projede kullanılabilir duruma getirilmesi amaçlanır. Hatalı veya analizin yanlış yönleneceğine neden olabilecek veriler temizlenir. Veri farklı kaynaklardan toplanmışsa ve aralarında farklılıklar varsa gerekli dönüşümler yapılarak bu farklılıklar ortadan kaldırılır. Eksik verilerin bulunduğu kayıtlar proje için fazla enformasyon taşıyor ise silinir ya da eksik veriler çeşitli yöntemler kullanılarak tahmin edilmeye çalışılır. Bu aşama en çok iş gücü gerektiren ve toplam süreç içinde en fazla zaman alan aşamadır (Larose, 2005, s.28).

Veri hazırlama kapsamında ilk olarak, toplanmış olan veri içinden uygun olanları seçme işlemi yapılır. Veri seçimiyle hedeflenen amaç, veri madenciliği için gerekli değişkenlerin belirlenmesidir. Çünkü modelleme sırasında, sadece sonuca etki edeceği düşünülen değişkenler göz önüne alınmalıdır. Sonuca etkisinin olmadığı değerlendirilen değişkenler, veri madenciliği kapsamından çıkartılmalıdır.

**Şekil 1.7 CRISP-DM Veriyi Hazırlama Basamağı**



Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

### Veri Temizleme

Veri temizlemedeki amaç, veriyi tutarsızlık, hata, eksiklik ya da tekrarlardan arındırmaktır. Örneğin yaş değeri için eksi değer girilmesi verinin hatalı olduğunu, doğum yılı ve yaşın birbiriyle uyumlu olmaması ise veride tutarsızlıklar olduğunu gösterir.

Toplanan bilgiler eksik olabilir. Eksik veri problemlere sebep olmaktadır. Çünkü bazı veriler geçerli etki alanında sunulamaz. Örneğin; hasta veri tabanı, kırmızı kan hücreli hasta bilgilerini barındırmıyorsa, hasta veri tabanından sıtma teşhisi yapılamaz. Sorunu çözmek için bu tür veriler geliştirilir (Baykal, 2003, s.64).

### **Veri Yapılandırma**

Mevcut deęişkenlerin yeterli olmadığı durumlarda yeni deęişkenler yaratılması gerekebilir. Örneęin, uzunluk ve genişlik deęişkeni kullanılarak bunların çarpımından ortaya çıkan alan deęişkeni kullanılmak istenebilir.

### **Veri Birleştirme**

Farklı kaynaklardan toplanan verilerin birleştirilmesi gerekebilir. Örneęin, bir süper marketler zincirine ait üç tablodan birinde, tüm marketlerin genel karakteristik özellikleri, dięerinde özetlenmiş satış rakamları ve sonucunda buldukları alanın demografik özellikleri tutuluyor olsun. Üç tabloda da her market için tek kayıt bulunmaktadır. Bu tablolar yine her market için tek kayıt olacak şekilde yeni bir tabloda birleştirilebilirler.

### **Veri Biçimlendirme**

Veri, farklı kaynaklardan toplandığında, aynı veriye ait farklı biçimler görülebilmektedir. Örneęin, bir kaynakta 0/1 olarak toplanmış veri, dięer kaynakta Evet/Hayır şeklinde toplanmış olabilir. Cinsiyet deęişkeni bir kaynakta E/K iken, dięer bir kaynakta Bay/Bayan biçimine sahip olabilir.

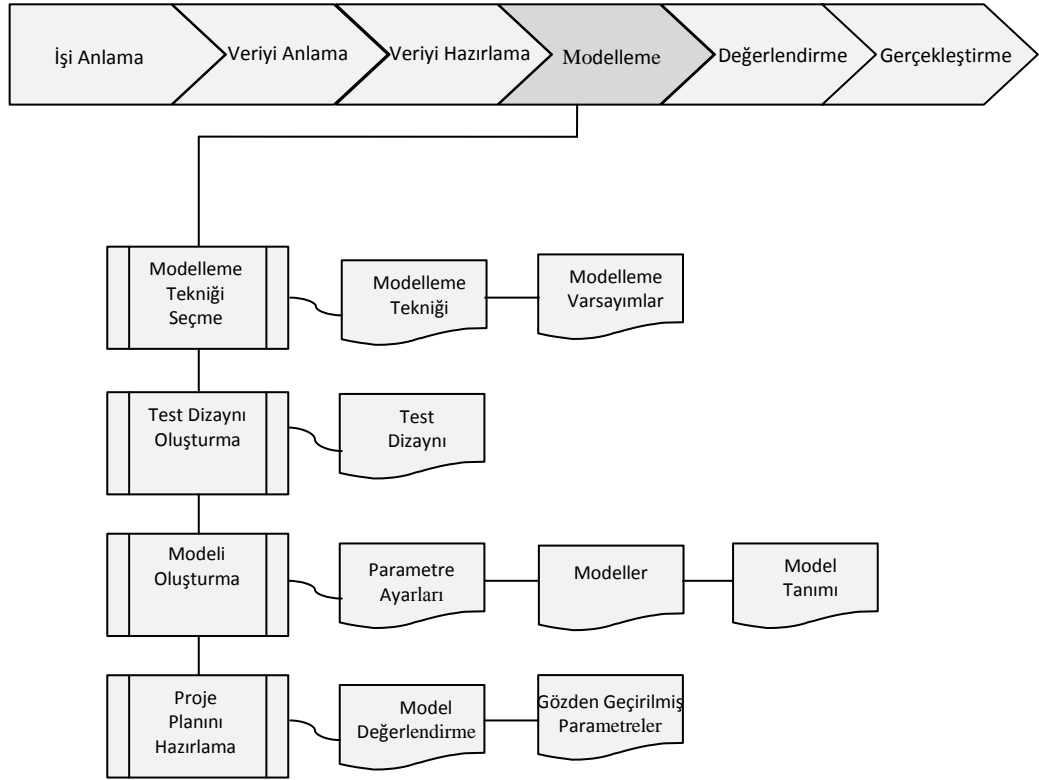
#### **1.6.4. Modelleme**

Bu aşamada, verilerden bilgi çekmek için ileri çözümleme yöntemleri kullanıldığından veri madencilięi sürecinin en gösterişli aşamasıdır. Uygun modellerin seçilip uygulanmasıyla birlikte parametreler en uygun deęişkenlere dönüştürülmektedir. Veri madencilięi, her problem tipi için farklı yöntemler içermektedir. Bazı yöntemler, veri tipi için uygun değildir ya da özel tanımlamalar gerektirmektedir. Bu nedenle gerekli olduğunda 3. aşama olan veri hazırlama aşamasına geri dönülür (Albayrak ve Yılmaz, 2009, s.37).

İlerleyen kısımlarda, karar ağaçları, yapay sinir ağları gibi birçok modelleme teknięi detaylı olarak anlatılacaktır.



**Şekil 1.8 CRISP-DM Modelleme Basamağı**

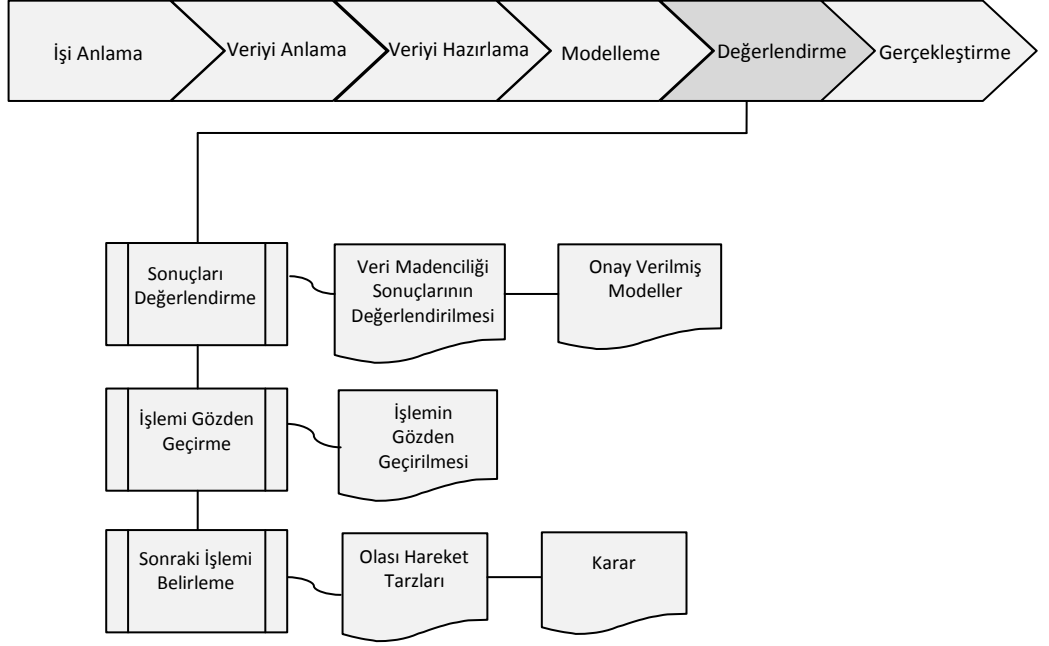


Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

### 1.6.5. Değerlendirme

Değerlendirme aşamasında, daha önce oluşturulmuş olan model, son kez tüm yönleriyle değerlendirilir, kalitesi ve etkinliği ölçülür. Modelin ilk aşamada oluşturulan proje amacına ulaşmada etkin olup olmadığı ve problemin tüm yönleri için bir çözüm sağlayıp sağlamadığı karara bağlanır. Modelin anlaşılabilirliği ve doğruluk oranı gibi konularda da model amaç için yeterli kaliteyi sağlıyorsa gerçekleştirme aşamasına geçilir (Chapman ve diğerleri, 2000, s.30).

**Şekil 1.9 CRISP-DM Değerlendirme Basamağı**



Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

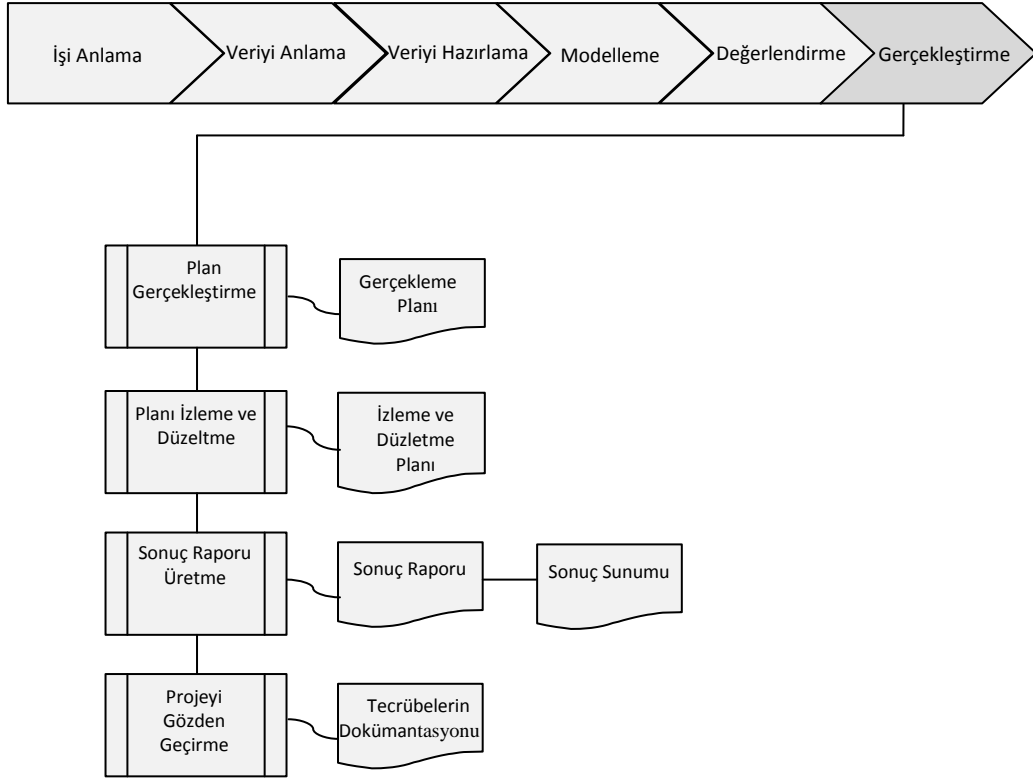
### 1.6.6. Gerçekleştirme

Bu basamak daha çok dokümantasyonun oluşmasıyla ilgilidir. Planların oluşturulduğu ve gerekirse düzeltildiği, sonuçların raporlandığı ve sunulduğu aşamadır.

Gerçekleştirme işleminde genel olarak cevaplanması gereken sorular şunlardır;

- Veri madenciliği sonuçları nasıl kullanılacak?
- Elde edilen sonuçları kimler kullanacak?
- Sonuçlar hangi sıklıkta kullanılacak?

**Şekil 1.10 CRISP-DM Gerçekleştirme Basamağı**



Kaynak: CRISP-DM ([www.crisp-dm.org](http://www.crisp-dm.org))

### **1.7. Veri Madenciliğinde Karşılaşılan Sorunlar**

Küçük veri kümelerinde hızlı ve doğru bir biçimde çalışan bir sistem, çok büyük veri tabanlarına uygulandığında tamamen farklı davranabilir. Bir veri madenciliği sistemi tutarlı veri üzerinde mükemmel çalışırken, aynı veriye gürültü eklendiğinde kayda değer bir biçimde kötüleşebilir (Akpınar, 2000, s.6).

Karşılaşılan bazı problemler şunlardır;

#### **Veritabanının Boyutu**

Veri tabanının boyutlarının veriler için yeterli olmaması durumunda, yapılan analizlerin uygulanabilirliği yoktur. Bu problemin çözümünde ise, yapılan örnekleme tekniklerinin ve örnek miktarının azaltılmasıyla bu problem aşılabilmektedir.

### **Gürültülü Veri**

Büyük veri tabanlarında pek çok niteliğin değeri yanlış olabilir. Bu hata, veri girişi sırasında yapılan insan hataları veya girilen değerlerin yanlış ölçülmesinden kaynaklanır. Veri girişi ya da veri toplanması sırasında oluşan sistem dışı hatalara gürültü adı verilir. Örneğin, ‘Yaş’ alanı için 25 girildiği halde, doğum tarihi olarak 07.02.2000 girilmesi durumunda gürültülü veri oluşur.

### **Eksik Veri**

Veride bulunan eksik veya kayıp bilgiler gözden geçirilmelidir. Örneğin bir veri tabanında yer alan bazı kişilerin ortalama aylık harcama tutarı bulunurken, bazı kayıtlarda bu bilgi eksik olabilir veya bu kayıt hiç girilmemiş olabilir. Bu durumdaki eksiklik “kayıp veriler” olarak tanımlanabilir. Bu tip kayıtları temizlemenin genel olarak iki yolu vardır. Diğer örneklere bakılarak olabilecek en uygun değerlerin verilmesi ya da boş kayıt oranı ihmal edilebilir düzeyde ise, o kayıtların tamamen silinmesi yöntemi ile bu sorun bertaraf edilebilir.

### **Dinamik Veri**

Veri tabanları dinamik olma eğilimindedir. İçerikleri her zaman ekleme, düzeltme ve silme işlemleri ile değişim halindedir. Veri tabanlarının verileri güncellemesi ya da yeni müşteri verilerini veri tabanına eklemesi, veri madenciliği sürecinde olumsuz etkilere yol açabilmektedir.

## **1.8. Veri Madenciliğinin Uygulama Alanları**

Veri madenciliği her geçen gün yeni ve farklı alanlarda kullanılmaya başlanmıştır. Günümüzde yaygın olarak kullanıldığı alanlar birkaç başlık altında toplanabilir.

### **1.8.1. Müşteri İlişkileri Yönetimi**

Yeni bir müşteri kazanmanın maliyetinin, müşteriye elde tutma maliyetinden daha yüksek olduğu ve kaybedilen bir müşteriye yeniden kazanma maliyetinin yeni müşteriler edinme maliyetinden daha fazla olduğu göz önüne alındığında, işletmeler müşteri odaklı olmak ve mevcut müşteriye ellerinde tutmak zorundadır (<http://www.spss.com.tr>).

CRM Institute tarafından yapılan bir arařtırmaya gre, yeni bir mřteri edinmek iin harcanan kaynakla, 5 eski mřterinin elde tutulması mmkn olabilmektedir. Ayrıca mřterinin iřletmeyle olan birliktelięi ilerledike, ona ayrılan elde tutma maliyeti de dřmektedir ([www.crminturkey.org](http://www.crminturkey.org)).

Hızla artan rekabet kořullarında iřletmelerin ne ıkabilmesi iin sadece finansal kořullar yeterli deęildir. Tketicilerin de her geen gn daha bilinli hale geldięi gnmzde, mřteri memnuniyeti iřletmelerin gl bir silahı haline gelmiřtir.

Mřteri memnuniyetine giden yol, iřletmelerin mřterilerini daha iyi tanınması ve mřteri gibi dřnme kapasitelerini arttırmalarından geer. Bu sayede mřteri eęilimlerini erkenden tespit etmek, řirket bařarısı iin can alıcı noktadır.

İřletmeler, mřterilerinin gemiřteki davranıřlarına, alıřkanlıklarına, demografik zelliklerine gre gelecekteki davranıřları iin tahmin modelleri oluřturmaktadırlar. Bu amala, deęiřik kanallardan gelen mřteri verileri tek bir ortamda birleřtirilip, nceden bilinmeyen anlamlı bilgilere ulařılır.

Bu bilgiler sayesinde aktif mřteriler haline gelen potansiyel mřteri adaylarının karlılıęı, ne kadar sre ile aktif mřteri olacakları ve muhtemel olarak ne zaman iřletmeyi bırakıp rakip iřletmelerden hizmet/rn alacakları tahmin edilir (Rygielski, Wang ve Yen, 2002, s.494).

Ortak zelliklerine gre, mřteriler homojen gruplara ayrılıp, her grup iin farklı uygun stratejiler de geliřtirilebilir.

### **1.8.2. Pazarlama Ynetimi**

Gnmz rekabet ortamında, iřletmelerin mřterilerine ait verilerinin anlamlı bilgi haline dnřmř řekli, satıř ve pazarlama faaliyetleri aısından kritik nem tařıyan bir unsurdur. Mřteriler hakkında normal iř srelerinde toplanan veriler, veri madencilięi araları kullanılarak, iřletmelerin mřterileri ile iliřkilerini glendirmek iin ihtiya duydukları anlamlı bilgiye dnřtirlr.

Veri madenciliği ile analizlerde kullanılacak verinin kaynağı; kredi kartı işlemleri, indirim kuponları, müşteri şikâyet aramaları, müşteri kayıtları, finans alışkanlıkları, vb. şeklinde sıralanabilir (<http://www.spss.com.tr>).

Veri madenciliği daha çok pazarlama bölümleriyle ilişkilendirilir, çünkü bu departmanlar yeni müşteri kazanmaya ve mevcut müşterileri ellerinde tutmaya odaklanırlar. Satın alma kampanyalarına kimin tepki vereceğinin tahmin edilmesi, müşterinin fiyat artışı ile değişen satın alma alışkanlıklarının belirlenmesi, yeni müşterileri çekmede hangi faktörlerin etkili olacağını tahmin edilmesi, belirli bir X ürününü alan hangi müşterilerin, o ürünle ilişkili Y ürününü de alacağını tahmin edilmesi, hangi müşterinin işletmeden ayrılma niyetinde olduğunun tahmin edilmesi pazarlama yönetiminde veri madenciliği ile yapılabilecek örneklerdendir.

### **1.8.3. Telekomünikasyon Sektörü**

Telekomünikasyon sektöründe en önemli sorun müşteri kaybıdır. Kuruluşlar hangi müşterilerini kaybedebileceklerini önceden belirleyebildikleri takdirde bu müşterilerini elde tutma amaçlı stratejiler geliştirebilir, düşük maliyetli ve etkili kampanyalar düzenleyebilirler. Kaybetme olasılığı olmayan bir müşteriye kalıcılığını sağlama amaçlı bir mesaj göndermek hem müşterinin kendisine verilmek istenen mesajın ne olduğunu algılamasını zorlaştıracak hem de maliyetleri artıracaktır.

Telekomünikasyon sektöründe bulunan firmalar ayrıntılı arama kayıtlarına sahiptirler. Bu firmalar ellerinde bulunan kayıtları kullanarak, benzer özellikler gösteren müşterileri bölümlendirip (müşteri segmentasyonu), fiyatlandırma ve promosyon stratejileri geliştirebilirler (Rygielski, Wang ve Yen, 2002, s.488).

### **1.8.4. Perakende Sektörü**

Perakende sektörü hızlı gelişen ve değişen, müşteri ile doğrudan etkileşimde bulunulan, rekabetin çok yoğun yaşandığı, müşteri satın alma davranışlarının hızlı değişkenlik gösterdiği kritik bir sektör olma özelliği taşımaktadır.

Perakende sektöründe satış terminalleri ve kodlama sistemleri sayesinde sistematik olarak veriler toplanmakta, mağaza alış veriş kartları, kredi kartları sayesinde yapılan alışverişin müşteri ile ilişkilendirilmesi sağlanmaktadır. Perakende sektöründe toplanan bu veriler yoğun rekabet ortamında avantaj elde etmede ve şirketlerin başarısında son derece önemlidir. Veri madenciliği yöntemleri sayesinde bu veriler içerisinde gizli olan bilgiler keşfedilmekte ve keşfedilen nitelikli bilgiler kurumsal rekabet ve başarı açısından kritik bir önem taşımaktadır (<http://www.spss.com.tr>).

Eskiden süpermarketlerde basit toplama makinesinden ibaret olan kasalar müşterinin o andaki satın aldığı ürünlerin toplamını hesaplamak için kullanılırdı. Günümüzde ise kasa yerine kullanılan satış noktası terminalleri sayesinde satın alma işlemlerinin bütün detayları saklanabilmektedir. Süpermarkette yapılan veri analizinde her ürün için bir sonraki ayın satış tahminlerinin çıkarılmasında, müşterilerin satın aldıkları ürünlere bağlı olarak gruplandırılmasında, yeni ürünlerin potansiyel müşterilerinin belirlenmesinde müşterilerin zaman içindeki hareketleri incelenerek tahminler yapılmasında ve önemli sonuçlar elde edilmesinde yarar sağlar. Bütün bu bilgilerin ayrıntılı olarak değerlendirilmesinde veri madenciliği devreye girmektedir (Ayanoğlu, Mert ve Giray, 2004, s.2).

Perakende sektöründe yoğunlukla, sepet analizi ile müşterilerin hangi ürünleri birlikte satın aldığı bulunur. Örneğin, “Çocuk bezi alan müşterilerin %30’u bira da satın alır.” Buradaki amaç mallar arasındaki pozitif veya negatif korelasyonları bulmaktır. Çocuk bezi alan müşterilerin mama da satın alacağı veya bira satın alanların cips de alacağı tahmin edilebilir ama, ancak otomatik bir analiz bütün olasılıkları göz önüne alır ve kolay düşünülemez, örneğin çocuk bezi ve bira arasındaki bağıntıları da bulur.

### **1.8.5. Bankacılık ve Finans Sektörü**

Bankacılık ve finans sektörü günümüzde sundukları hizmet, ürün ve servislerle bilgiye dayalı yönetime en fazla ihtiyaç duyan kuruluşlardır. Finans sektöründe en temel uygulamalar çapraz satış, risk derecelendirme, mevcut müşteriyi elde tutma, yeni müşteriler kazanma, maliyetleri azaltma, kayıp ve kaçakları engelleme, alternatif kanallar oluşturma, müşteri memnuniyetini sağlama olarak özetlenebilir. Hangi müşteri

profilinin neyi, ne zaman ve neden tercih ettiğini anlayabilen bir kuruluş hem talep yaratma, hem de doğru zamanda doğru talebi karşılama ve sunma avantajına sahip olacaktır. Mevcut müşteri kaybı, finans ve sigorta sektörlerinde en önemli problemi teşkil etmektedir. Bankalar, mevcut müşterilerden rakip bankaya geçme ihtimali olan müşterileri, profillerini ve kaybettikleri müşterilerin hangi sebepler yüzünden sistemden ayrıldıklarını tespit etmek istemektedir (<http://www.spss.com.tr>).

Bankalar veri madenciliği ile müşterilerin kredi risklerini hesaplayarak hangi müşterilerin kredi riskinin yüksek olduğunu, hangi müşterilerin geri ödemesini zamanında yapamayacağını tahmin edebilir. Kredi kartı ödemelerini aksatan, gecikmeli olarak yapan veya hiç yapmayanların özelliklerinden yola çıkarak bundan sonra aynı duruma düşebilecek muhtemel kişileri saptayabilir (Özmen, 2001, s.3).

HSBC veri madenciliği çözümleri ile yaptığı çalışmada müşteri ihtiyaçlarını ve davranışlarını tespit etmiş ve doğru müşteriye doğru önerilerle giderek pazarlama maliyetlerinde %30'luk bir azalma sağlarken, satışlarını %50 artırmıştır. (<http://www.spss.com.tr>).

Chye ve Gerry, bankacılık alanında müşteri kaybı analizi (churn analizi) ile bir uygulama gerçekleştirmişlerdir. Bu uygulama ile bir bankanın bugünkü müşterilerinin gelecek altı ayda, rakip firmalara geçip geçmeyeceği tahmin edilmeye çalışılmıştır. Churn analizi haricinde veri madenciliği, bankaların kredi başvurularını doğru bir şekilde değerlendirmelerine olanak vermektedir. Bireysel kredi başvurularında yaş, cinsiyet, meslek, iş tanımı, gelir ve giderler baz alınarak kişinin borcunu ödeyemez duruma düşmesi olasılığının tespiti yapılabilir (Chye ve Gerry, 2006, s.17).

#### **1.8.6. Sağlık Sektörü**

Doğru ve zamanında karar almanın hasta sağlığı üzerindeki etkisi tartışmasız çok önemlidir. Hastane bünyesinde toplanan operasyonel verilerden yararlanılarak, belirli bir hastalığa sahip kişilerin ortak özellikleri, tıbbi tedaviden sonra hastaların durumları, hastane maliyetleri, ölüm oranları ve salgın hastalıklar gibi konularda tahminler yapılabilir.



Dünya çapında çok sayıda başarılı uygulama örneği mevcuttur.

Kaur ve Wasan hastaların yaş ve cinsiyet gibi verilerini karar ağacı yöntemleri ile analiz ederek, hastaların göğüs kanseri olup olmadığını tahmin etmeye çalışmışlardır (Kaur ve Wasan, 2006, s.194-200).

Başka bir çalışma ise Kore Tıbbi Sigorta Kurumu (The Korea Medical Insurance Corporation) tarafından hazırlanan bir veri tabanı üzerinde yapılan yüksek tansiyon ile ilgili bir çalışmadır. İlk aşamada yüksek tansiyona sahip 9.103 kayıt üzerinde, daha sonra aynı sayıda yüksek tansiyonu olmayan kayıtlar üzerinde çalışılmıştır. Bu örnek 13.689 kayıttan oluşan öğrenme ve 4.588 kayıttan oluşan test setine bölünerek modelin eğitimi yapılmıştır. Öğrenim algoritmasında karar ağaçları algoritmalarından CHAID, C4.5, C5.0 kullanılmıştır. Bu çalışmalar sonucunda yüksek tansiyon tahmininde etkili değerler BMI, idrar proteini (urinary protein), kan glikozu, kolesterol değerleri olarak belirlenmiştir. Diyet, alınan tuz miktarı, alkol, tütün gibi değişkenlerin hiçbirinin tahminde etkili olmadığı, ayrıca grafiksel değerlerde de yalnızca yaşın etkili olduğu saptanmıştır (Kaya, Bulun ve Arslan, 2005, s.8).

San Francisco Hearth Institute; hasta sonuçlarının iyileştirilmesi, hastanın hastanede kalma süresinin azaltılması gibi amaçlarla bir çalışma başlatmış ve kurum bünyesinde toplanan verilerden hastanın geçmişine ait veriler, laboratuvar verileri, kolesterol verileri, diğer medikal verileri veri madenciliği ile bilgiye dönüştürmüştür. Yapılan analizler sonucunda geliştirilen stratejiler ile hastaların hastanede kalma süreleri azaltılmış ve yılda \$500.000 tasarruf sağlamıştır (www.spss.com.tr).

### **1.8.7. Diğer Uygulamalar**

Veri madenciliği;

- Turizm alanında, bir bölgeye turistlerin hangi sebeplerle geldiğini tespit ederek reklam kampanyası düzenleme ve bir sonraki sezonda turist sayısını arttırmada,
- Spor alanında, bir basketbol takımına alınacak oyuncunun gelecekteki performansı ve takıma sağlayacağı yararların tahmininde,

- Kütüphanecilik alanında, bir müşterinin aldığı kitabı ne zaman getireceği ve bir sonraki gelişinde hangi kitabı seçeceği konusunda tahminlerde,
- Web alanında, kullanıcıların profilini belirleyip onlara uygun ürünlerin reklam kampanyalarını hazırlamada kullanılabilir.

Şekil 1.11’ de 2008 yılı Aralık ayında KDnuggets tarafından yapılan, veri madenciliğinin sektörler bazında kullanımına ilişkin bir araştırmanın sonuçları yer almaktadır. Bu çizelgede araştırmaya katılan toplam 107 şirketin 41 adedinin müşteri ilişkileri yönetimi (CRM) alanında veri madenciliği kullandığı görülmektedir.

**Şekil 1.11 Veri Madenciliğinin Uygulandığı Alanlar**

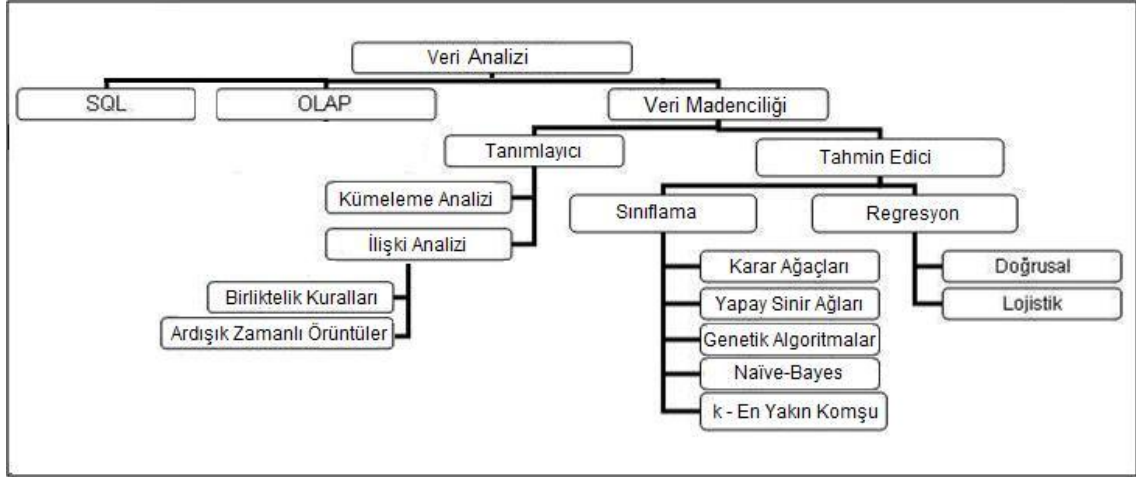


Kaynak: Data Mining Applications in 2008 (<http://www.kdnuggets.com>)

## 1.9. Veri Madenciliği Modelleri

Veri madenciliği modelleri, tahmin edici ve tanımlayıcı olmak üzere iki temel kategori altında incelenebilir. Ancak bu kategoriler birbirlerinden açık farklarla ayrılmamıştır. Anlaşılabilir oldukları ölçüde, tahmin edici modeller tanımlayıcı, tanımlayıcı modeller ise tahmin edici model olarak kullanılabilir.

Şekil 1.12 Veri Madenciliği Modelleri



### 1.9.1. Tahmin Edici Modeller

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. Örneğin bir banka önceki dönemlerde vermiş olduğu kredilere ilişkin gerekli tüm verilere sahip olabilir. Bu verilerde bağımsız değişkenler kredi alan müşterinin özellikleri, bağımlı değişken değeri ise kredinin geri ödenip ödenmediğidir. Bu verilere uygun olarak kurulan model, daha sonraki kredi taleplerinde müşteri özelliklerine göre verilecek olan kredinin geri ödenip ödenmeyeceğinin tahmininde kullanılmaktadır (Özekeş, 2003, s.67).

#### 1.9.1.1. Sınıflandırma

Sınıflandırma, yeni bir nesnenin niteliklerinin incelenerek, önceden tanımlanmış belirli sınıflar içinde hangi sınıfa ait olduğunun tespit edilmesidir. Örneğin, bir sınıflandırma modeli kredi kartı başvurularını düşük, orta ve yüksek risk grubu olarak ayırmak amacıyla kurulabilir.

Sınıflandırmaya ait diğer bir örnek aşağıda görülmektedir (Alpaydın, 2000, s.2);

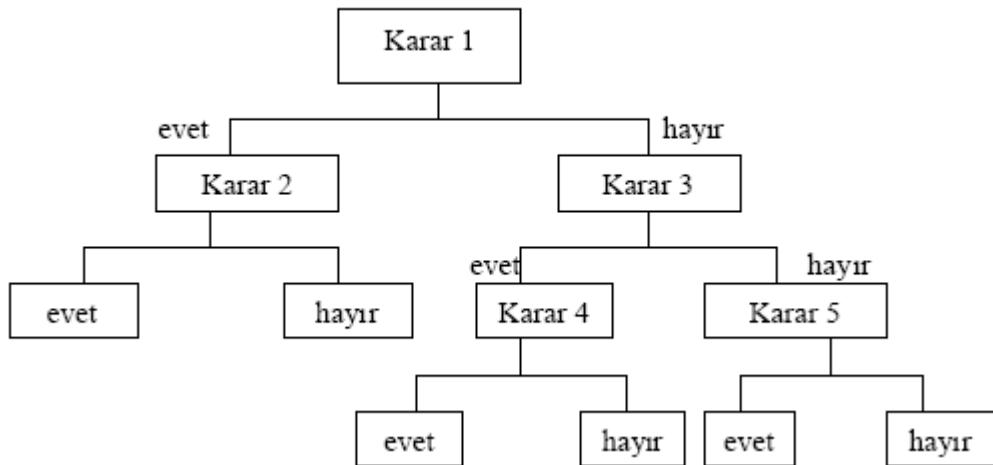
“Genç kadınlar küçük araba satın alır, yaşlı, zengin erkekler büyük, lüks araba satın alır.”

Amaç bir malın özellikleri ile müşteri özelliklerini eşlemektir. Böylece bir müşteri için ideal ürün veya bir ürün için ideal müşteri profili çıkarılabilir. Örneğin bir otomobil satıcısı şirket geçmiş müşteri hareketlerinin analizi ile yukarıdaki gibi iki kural bulursa genç kadınların okuduğu bir dergiye reklâm verirken küçük modelinin reklâmını verir.

#### 1.9.1.1.1. Karar Ağaçları

Karar ağacı yöntemini kullanarak verinin sınıflanması iki basamaklı bir işlemdir. İlk basamak öğrenme basamağıdır. Öğrenme basamağında önceden bilinen bir eğitim verisi, model oluşturmak amacı ile sınıflama algoritması tarafından çözümlenir. Öğrenilen model, sınıflama kuralları veya karar ağacı olarak gösterilir. İkinci basamak ise sınıflama basamağıdır. Sınıflama basamağında test verisi, sınıflama kurallarının veya karar ağacının doğruluğunu belirlemek amacıyla kullanılır. Eğer doğruluk kabul edilebilir oranda ise, kurallar yeni verilerin sınıflanması amacıyla kullanılır (Özekeş, 2003, s.69).

Şekil 1.13 Karar Ağacı Şekli



Kaynak: Kantardzic, M., Data Mining Concepts, Models, Methods and Algorithms, John Wiley & Sons Publishers, 2003.

Karar ağaçları, hangi demografik grupların mektupla yapılan pazarlama uygulamalarında yüksek cevaplama oranına sahip olduğunun belirlenmesi (Direct Mail), bireylerin kredi geçmişlerini kullanarak kredi kararlarının verilmesi (Credit Scoring), geçmişte işletmeye en faydalı olan bireylerin özelliklerini kullanarak işe alma süreçlerinin belirlenmesi, tıbbi gözlem verilerinden yararlanarak en etkin kararların verilmesi, hangi değişkenlerin satışları etkilediğinin belirlenmesi, üretim verilerini inceleyerek ürün hatalarına yol açan değişkenlerin belirlenmesi gibi uygulamalarda kullanılmaktadır (Akpınar, 2000, s.11).

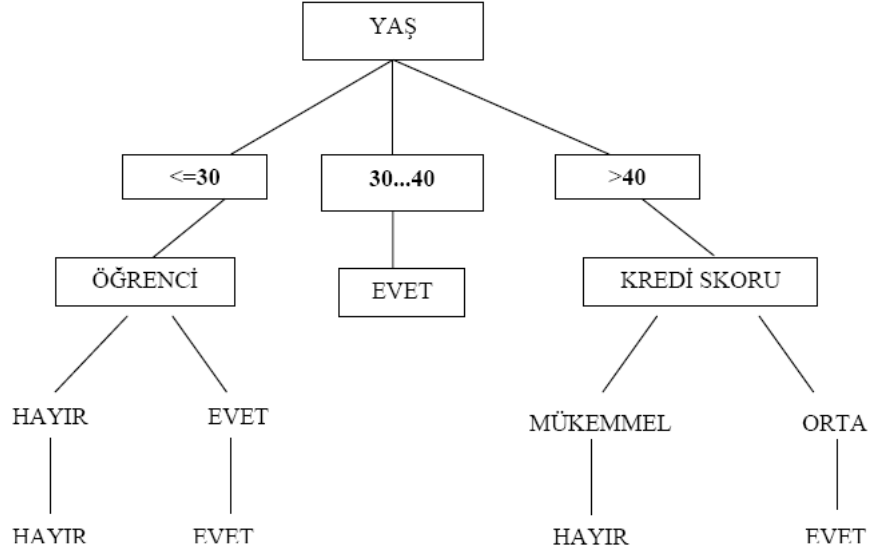
Aşağıda karar ağaçlarıyla ilgili bir örnek görülmektedir (<http://www.rhsmith.umd.edu>).

**Tablo 1.2 Karar Ağacı İçin Örnek Veri Seti**

YAŞ	GELİR	ÖĞRENCİ	KREDİ SKORU	BİLGİSAYAR ALIMI
<=30	Yüksek	Hayır	Orta	Hayır
<=30	Yüksek	Hayır	Mükemmel	Hayır
30...40	Yüksek	Hayır	Orta	Evet
>40	Orta	Hayır	Orta	Evet
>40	Düşük	Evet	Orta	Evet
>40	Düşük	Evet	Mükemmel	Hayır
31...40	Düşük	Evet	Mükemmel	Evet
<=30	Orta	Hayır	Orta	Hayır
<=30	Düşük	Evet	Orta	Evet
>40	Orta	Evet	Orta	Evet
<=30	Orta	Evet	Mükemmel	Evet
31...40	Orta	Hayır	Mükemmel	Evet
31...40	Yüksek	Evet	Orta	Evet
>40	Orta	Hayır	Mükemmel	Hayır

Tablo 1.2’ deki veri seti dikkate alınarak “bilgisayar alımı” için Şekil 1.14’ teki gibi bir karar ağacı oluşturulur.

Şekil 1.14 Örnek Bir Karar Ağacı



Karar ağaçlarına dayalı olarak geliştirilen pek çok algoritma vardır. Bu algoritmalar birbirlerinden kök, düğüm ve dallanma kriteri seçimlerinde izledikleri yol açısından ayrılırlar.

Geliştirilen bu algoritmalar içerisinde CHAID (Chi-Squared Automatic Interaction Detector; Kass, 1980), C&RT (Classification and Regression Trees; Breiman, ve diğerleri 1984), ID3 (Quinlan, 1986), Exhaustive CHAID (Biggs, de Ville ve Suen, 1991), C4.5 (Quinlan, 1993), MARS (Multivariate Adaptive Regression Splines; Friedman), QUEST (Quick, Unbiased, Efficient Statistical Tree; Loh ve Shih, 1997), C5.0 (Quinlan), SLIQ (Supervised Learning in Quest; Mehta, Aganval ve Rissanen), SPRINT (Scalable Parallelizable Induction of Decision Trees; Shafer, Agrawal ve Mehta) başlıcalarıdır.

### QUEST

QUEST (Quick, Unbiased, Efficient, Statistical Tree) 1997 yılında Loh ve Shih tarafından geliştirilmiş olan bir tekniktir. Binary (ikili) büyüyen bir ağaç algoritmasıdır. Ayrı ayrı değişken seçimi ile ilgilenir. QUEST' teki birim değişken ayırıcı, tahmini olarak tarafsız değişken seçimini gerçekleştirir. QUEST algoritmasının C&RT algoritmasına benzer yönleri vardır, ancak ağaçlar yavaş büyüyebilir ve ikili olduğu için karar ağacı çok geniş olabilir (Clementine 12.0 Algorithms Guide, 2007, s.56).

## **CHAID**

1980 yılında G.V. Kass tarafından geliştirilen CHAID algoritmasında, bağımlı değişkeni en fazla etkileyen bağımsız değişken, bağımlı değişkenin sürekli olması durumunda F testi, kategorik olması durumunda Ki Kare testi kullanılarak belirlenir. Kategorik (Nominal / Ordinal) ve sürekli değişkenler üzerinde çalışabilmesi, ağaçta her düğümü ikiden fazla alt gruba ayırabilmesi gibi nedenlerle günümüzde de tercih edilen bir algoritmadır (Akpınar, 2000, s.12).

CHAID algoritması, tahmin edici değişkenin tüm değerlerini dikkate olarak analiz yapar. Hedef değişkeni dikkate alarak istatistik olarak benzer olan değişkenleri birleştirir ve farklı olan değişkenle işlemi sürdürür. Daha sonra karar ağacının ilk dalını oluşturmak için en iyi tahmin edici değişkeni seçer. Her bir düğüm seçilen değişkenin benzer değerlerinden oluşur. Bu süreç ağaç tamamıyla büyüyene kadar tekrarlanarak devam eder.

CHAID algoritması ikili bir algoritma değildir ki bu ağaçta herhangi bir seviyede ikiden çok kategori üretmesi anlamına gelir. Bu nedenle daha geniş ağaç yaratmaya eğilimlidir. CHAID, Ki-Kare metriği vasıtasıyla, ilişki düzeyine göre farklılık rastlanan grupları ayrı ayrı sınıflamaktadır. Dolayısıyla, ağacın yaprakları, ikili değil, verideki farklı yapı sayısı kadar dallanmaktadır. ID3, C5.0 ve C&RT ikili ağaçlar türetirken, CHAID ikili olmayan çoklu ağaçlar türetir (Clementine 12.0 Algorithms Guide, 2007, s.43).

## **C&RT**

CART veya C&RT ikili ağaç olarak büyüyen bir algoritmadır. C&RT veriyi iki alt kümeye ayırır. Böylece bir sonraki adımda oluşacak olan alt küme, bir öncekinden daha homojen olacaktır. Bu süreç sonuç bulunana kadar devam eden, kendini tekrarlayan bir süreçtir. C&RT algoritması karmaşık bir algoritmadır. Sürekli ya da kategorik veriler üzerinde çalışabilir. Büyük verilerle çalışıldığında sonuç bulması uzun sürmektedir. C&RT sınıflandırma ve regresyon analizi için kullanılan bir algoritmadır (Clementine 12.0 Algorithms Guide, 2007, s.31).

## **C5.0**

C5.0 algoritması, 1993 yılında Quinlan tarafından ortaya atılmıştır. En büyük özelliği karar ağacının gereksiz dallarını tespit ederek budamasıdır. Bu algoritmanın bir diğer özelliği ise eksik veri ve sayısal değerlerle de çalışabilmesidir (Akgöbek ve Öztemel, 2006, s.5). Kayıp verileri diğer veri ve değişkenler yardımı ile tahmin ederek, daha hassas ve daha anlamlı kurallar çıkartabilen bir ağaç üretebilir. Ayrıca, hedef değişken olarak sürekli değişkenlerle çalışamaz.

### **1.9.1.1.2. Yapay Sinir Ağları**

Yapay sinir ağları (YSA) biyolojik sinir sisteminin taklit edilmesi, matematiksel olarak modellenmesi çabalarının bir sonucu olarak ortaya çıkmıştır. Yapay sinir ağları, öğrenme yolu ile yeni bilgi elde edebilmeyi sağlayan bir tekniktir.

Haykin (1999) tarafından oluşturulan tanıma göre, “Bir sinir ağı, basit işlem birimlerinden oluşan, deneysel bilgileri biriktirmeye yönelik doğal bir eğilimi olan ve bunların kullanılmasını sağlayan yoğun bir şekilde paralel dağıtılmış bir işlemcidir. Bu işlemci iki şekilde beyin ile benzerlik göstermektedir:

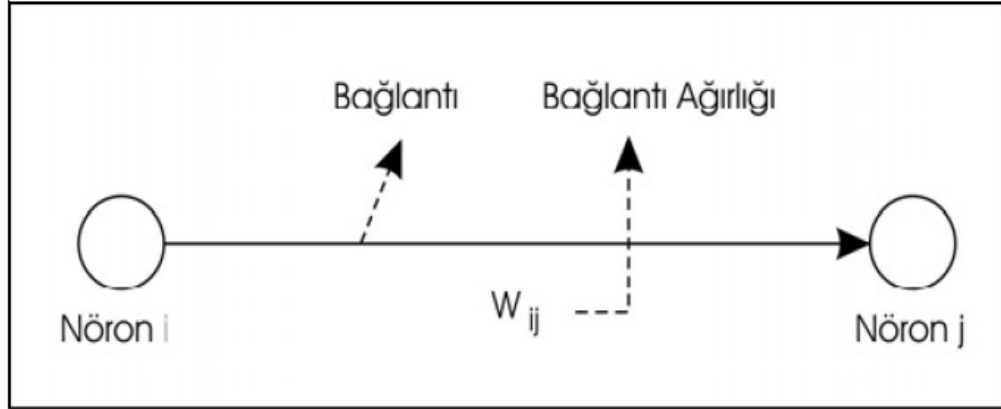
1-Bilgi, ağ tarafından bir öğrenme süreciyle çevreden elde edilir.

2-Elde edilen bilgileri biriktirmek için sinaptik ağırlıklar olarak da bilinen nöronlar arası bağlantı güçleri kullanılır.”

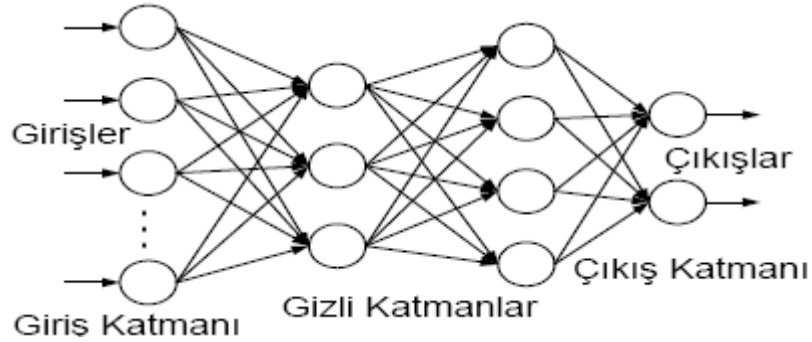
YSA’ların yapısı üç ana eleman içermektedir. Şekil 1.15’ ten de görülebileceği gibi, YSA’ların yapısını oluşturan üç ana eleman temel işlem elemanı olan nöron, girdi ve çıktı yolunu sağlayan bağlantı ve bu bağlantıların sağlamlığını gösteren bağlantı ağırlığıdır (Sattari ve diğerleri, 2007, s.339).



Şekil 1.15 Yapay Sinir Ağı Mimarisinin Temel Elemanları



Şekil 1.16 Çok Katmanlı Bir Yapay Sinir Ağı



Kaynak: Erdem, A., Uzun E., Yapay Sinir Ağları İle Türkçe Times New Roman, Arial El Yazısı Karakterlerini Tanıma, Gazi Üniversitesi, Mühendislik-Mimarlık Fakültesi Dergisi, Cilt 20, No 1, 2005

### 1.9.1.1.3. Genetik Algoritmalar

Genetik algoritma, Darwin tarafından geliştirilen “evrim teorisi” ne dayalıdır. Algoritma ilk olarak popülasyon adı verilen bir çözüm kümesi (öğrenme veri kümesi) ile başlatılır. Bir popülasyondan alınan sonuçlar bir öncekinden daha iyi olacağı beklenen yeni bir popülasyon oluşturmak için kullanılır. Evrim süreci (yeni popülasyonlar yaratma iterasyonu) tamamlandığında bağımlılık kuralları veya sınıf modelleri ortaya konmuş olur.

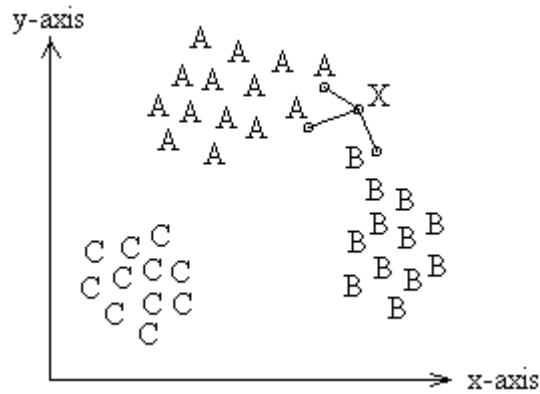
Genetik algoritmalar, veri madenciliğinde, kümeleme, tahmin, ilişki kuralları oluşturma ve müşterilerin gruplandırılması ve sınıflandırılmasında yaygın olarak kullanılmaktadır. Ayrıca işletmelerde, montaj hattı dengeleme, tesis yerleşimi ve atama gibi konularda genetik algoritmalar kullanılmaktadır (Emel ve Taşkın, 2002, s.143-144).

#### 1.9.1.1.4. K-En Yakın Komşu

Mesafeye dayalı sınıflandırma algoritmalarındandır. Sınıflandırma yapılırken eldeki verilerin birbirine olan uzaklığı veya benzerliği kullanılarak sınıflamanın gerçekleştirilmesine dayanır. Sınıflandırma yapılırken veri tabanındaki her bir kaydın diğer kayıtlarla olan uzaklığı hesaplanır (Kevin ve diğerleri, 1999, s.220). Yakınlık Öklid uzaklığı ile tanımlanır. Daha sonra, bilinmeyen örneklem, k en yakın komşu içinden en çok benzediği sınıfa atanır (Toktaş ve Demirhan, 2004, s.2).

Bir kayıt için diğer kayıtlardan sadece k adedi göz önüne alınır. Algoritmanın isminden de anlaşılacağı gibi, bu k adet kayıt, başka bir deyişle veri tabanındaki nokta, mesafesi hesaplanan noktaya diğer kayıtlara nazaran en yakın olan kayıtlardır. Bu yöntem coğrafi bilgi sistemlerinde çok kullanılır. Belirlenen bir noktaya en yakın şehir, istasyon vs. belirlenmesi aslında k-en yakın komşu algoritmasının temelini oluşturur (Kevin ve diğerleri, 1999, s.221).

Şekil 1.17 K-En Yakın Komşu Örneği



Şekil 1.17'de, sınıflandırılacak kayıt X, A sınıfına atanacaktır. Çünkü burada  $k=3$  seçilmiştir ve kayıt X'e en yakın 2 kayıt A'da, diğeri ise B sınıfında yer almaktadır.

#### 1.9.1.1.5. Naïve-Bayes

Tüm kriterlerin sonuca olan etkilerinin olasılık olarak hesaplanması temeline dayanmaktadır. Özellikle müşterilerin belirli bir özelliğe göre sınıflandırılmasında kullanılan bir tekniktir. Belirsiz durumların olasılık teorisine dayalı olarak çıkarılması tekniğidir. Çeşitli sebeplerin aynı sonucu verdiği durumlarda, bazen sonuç bilindiği

halde, bunun hangi sebepten dolayı meydana gelmiş olabileceği bilinmeyebilir. Söz konusu sonucun hangi olasılıkla hangi sebepten kaynaklandığı araştırılmak istendiğinde Bayes teoreminden yararlanılmaktadır (Serper, 1997, s.202).

Örneğin, havanın durumuna göre tenis maçının oynanıp oynanmayacağı tahmin edilmek istensin. Sıcaklık, nem, rüzgâr bilgileri girildiğinde, sistem önceki gerçekleşmiş istatistiklerden de faydalanarak, tenis maçının oynanma ve oynanmama ihtimalini hesaplar.

Sağlık sektöründe bir kişinin tahlil sonuçlarının değerlendirilerek bir hastalığa yakalanmış olup olmama olasılığının değerlendirilmesi de Naïve-Bayes ile mümkün olmaktadır.

### **1.9.1.2. Regresyon Analizi**

Regresyon analizi, doğrusal ve lojistik olmak üzere iki kısımda incelenecektir.

#### **1.9.1.2.1. Doğrusal Regresyon Analizi**

Doğrusal regresyon analizinde amaç, bağımlı değişkeni etkileyen bağımsız değişkenlerin neler olduğunu tespit ederek, bu etkinin rakamsal değerini belirlemektir. Örneğin, gelir düzeyine göre tüketim harcamaları ya da öğrencinin devamsızlığına göre başarı durumu tahmin edilebilir.

Doğrusal regresyon analizinin kullanılabilmesi için hem açıklayıcı (bağımsız) değişkenin hem de açıklanan (bağımlı) değişkenin sürekli sayısal değişken olması gerekmektedir.

Klasik doğrusal regresyon modelleri bağımsız değişken sayısına göre basit doğrusal regresyon modeli ve çoklu regresyon modeli olmak üzere ikiye ayrılır.

Aşağıda, basit regresyon denklemi görülmektedir:

$$Y = \beta_0 + \beta_1 X$$

Y= bağımlı değişken

X= bağımsız değişken

$\alpha$ ,  $\beta_0$ ,  $\beta_1$  ise tahmin edilecek parametreleri ifade etmektedir.

#### 1.9.1.2.2. Lojistik Regresyon Analizi

Neden sonuç ilişkilerinin ortaya konulması amacıyla yapılan çoğu sosyo-ekonomik araştırmada, incelenen değişkenlerden bazıları olumlu-olumsuz, başarılı-başarısız, evet-hayır, memnun-memnun değil şeklinde iki düzeyli verilerden oluşmaktadır. Bu türde bağımlı değişkenin iki düzeyli ya da çok düzeyli kategorik verilerden oluşması durumunda; bağımlı değişken ile bağımsız değişken (ler) arasındaki neden-sonuç ilişkisinin incelenmesinde, lojistik regresyon analizi önemli bir yere sahiptir (Agresti, 2007 s.103).

Doğrusal regresyon analizinde olduğu gibi, lojistik regresyon analizinde de bazı değişken değerleri göz önüne alınarak tahmin yapılmaya çalışılır. Fakat bu iki analiz arasında üç önemli fark vardır (Bircan, 2004, s.187);

- Doğrusal regresyonda tahmin edilecek bağımlı değişken sürekli ancak lojistik regresyonda bağımlı değişken kesikli değerler alır.
- Doğrusal regresyon analizinde bağımsız değişkenin çoklu normal dağılım koşulu aranırken lojistik regresyonda böyle bir koşul aranmaz.
- Doğrusal regresyon analizinde bağımlı değişkenin değeri tahmin edilirken, lojistik regresyon analizinde ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı tahmin edilir.

Lojistik regresyon modelinin katsayıları, doğrusal regresyon modelinden farklı olarak yorumlanmaktadır. Doğrusal regresyon modelinde, her bir katsayı ait olduğu değişkenin bir birim değişmesi durumunda, bağımlı değişkenin ne kadar değişeceğini gösterirken, lojistik regresyon modelinde her bir katsayı ait olduğu değişkende meydana gelecek bir birimlik değişimin  $\ln(Y$ 'nin gerçekleşme olasılığı/ $Y$ 'nin gerçekleşmeme olasılığı) biçiminde ifade edilen olasılığın doğal logaritmasında ortaya çıkaracağı değişme şeklinde yorumlanacaktır (www.tajev.org).

Aşağıda lojistik regresyon denklemi görülmektedir:

$$\ln\left[\frac{Y_i}{1-Y_i}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

### **1.9.2. Tanımlayıcı Modeller**

Tanımlayıcı modellerde, karar vermeye rehberlik etmede kullanılabilir mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. X/Y aralığında geliri, evi ve arabası olan, ayrıca çocukları okul çağında olan aileler ile, çocuğu olmayan ve geliri X/Y aralığından düşük olan ailelerin satın alma örüntülerinin birbirlerine benzerlik gösterdiğinin belirlenmesi tanımlayıcı modellere bir örnektir (Akpınar, 2000, s.4).

30 yaş üstü evli kişiler ile 30 yaş altı bekâr kişiler üzerinde yapılan ve ödeme performanslarını gösteren bir analiz de tanımlayıcı modellere örnek olarak verilebilir.

Tanımlayıcı modeller, kümeleme ve ilişki analizi olmak üzere iki ana başlık altında incelenmektedir.

#### **1.9.2.1. Kümeleme Analizi**

Kümeleme analizi sınıflandırmanın aksine, veri kümesini önceden sınıflara ayırmaz, bunun yerine veriler dağılımlarına göre irdelenerek doğal sınıflandırmalar oluşturur. Kümeleme analizinin sınıflandırma işleminden en önemli farkı önceden belirlenmiş sınıflar ya da sınıf tanımları (etiketleri) olmamasıdır. Bu yüzden kümeleme işlemi gözetimsiz (unsupervised) veri madenciliği yöntemidir (Berkhin, 2002, s.20).

Kümeleme tekniğinde amaç üyelerinin birbirlerine çok benzediği, ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir. Kümeleme analizinde; veri tabanındaki kayıtların hangi kümelere ayrılacağı veya kümelemenin hangi değişken özelliklerine göre yapılacağı, konunun uzmanı olan bir kişi tarafından belirtilebileceği gibi, veri tabanındaki kayıtların hangi kümelere ayrılacağını geliştirilen yazılımlar da yapabilmektedir (Han ve Kamber, 2001, s.383).

Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir (Özekeş, 2003, s.71);

1. Bölme Yöntemleri (Partitioning Methods)
2. Hiyerarşik Yöntemler (Hierarchical Methods)
3. Yoğunluk Tabanlı Yöntemler (Density-based Methods)
4. Izgara Tabanlı Yöntemler (Grid-based Methods)
5. Model Tabanlı Yöntemler (Model-based Methods)

### **1.9.2.2. İlişki Analizi**

Birliktelik kuralları ve ardışık zamanlı örüntüler, pazarlama amaçlı olarak pazar sepet analizi (market basket analysis) adı altında veri madenciliğinde yaygın olarak kullanılmaktadır.

#### **1.9.2.2.1. Birliktelik Kuralları**

Birliktelik analizi, bir veri kümesindeki kayıtlar arasındaki bağlantıları arayan denetimsiz (unsupervised) veri madenciliği şeklindedir. Birliktelik analizi çoğu zaman perakende sektöründe süpermarket müşterilerinin satın alma davranışlarını ortaya koymak için kullanıldığından “pazar sepeti analizi” olarak da adlandırılır (Bland, 2002, s.45).

Birliktelik kurallarına ait örnekler aşağıda yer almaktadır:

- “Müşteriler bira satın aldıklarında %75 olasılıkla çocuk bezi de satın alırlar.”
- “Düşük yağlı peynir ve yağsız süt alan müşteriler %85 olasılıkla diyet içecek alırlar.”

#### **1.9.2.2.2. Ardışık Zamanlı Örüntüler**

Ardışık zamanlı örüntüler, birbirleri ile ilişkisi olan ancak birbirini izleyen dönemlerde gerçekleşen ilişkilerin tanımlanmasında kullanılır (Akpınar, 2000, s.6).

Ardışık zamanlı örüntüler analize ait örnekler aşağıda yer almaktadır:

- “Çadır alan müşterilerin %10’u bir ay içerisinde sırt çantası almaktadır.”
- “A hissesi %15 artarsa üç gün içinde B hissesi %60 olasılıkla artacaktır.”

## 2. BÖLÜM

### MÜŞTERİ İLİŞKİLERİ YÖNETİMİ

Yaşanan yoğun rekabet, hızla gelişen teknoloji ve bilinçlenen müşteriler, günümüzde işletmeleri değişime zorlayan başlıca faktörler arasındadır. Her geçen gün pazara yeni bir rakibin girdiği ortamda işletmeler hayatta kalabilmek için, müşterinin istek ve ihtiyaçlarını hızla anlayarak, ürün ve hizmetlerini bu yönde şekillendirmek zorundadırlar. Sadık müşteri yaratma, müşteri memnuniyetini en üst seviyeye çıkarma amacıyla yapılan çalışmalar, müşteri ilişkileri yönetimi (Customer Relationship Management) olarak karşımıza çıkmaktadır.

#### 2.1. Müşteri İlişkileri Yönetimi Nedir?

Müşteri ilişkileri yönetimi (CRM), yeni bir kavrammış gibi gözüke de aslında yeni veya günümüze özgü bir kavram değildir. Müşteri, ilişki, yönetim kavramları her zaman var olmuş fakat bunların ele alınış biçimi ve birbirleriyle ilişkileri sürekli değişen bir hal almıştır. Üretim odaklı, satış odaklı, pazarlama odaklı anlayışlardan artık müşteri odaklı yani müşteri merkezli bir anlayışa geçilmiştir. Müşteri odaklı olma anlayışı CRM kavramının ana unsurudur. Müşteriyi yapılan işlerin ortasına koyarak ürünü, hizmeti, yönetim şeklini, organizasyonu ve insan kaynağı dâhil tüm kaynakları müşteriye göre ayarlamak, düzenli ve etkin çalışmak amacıyla teknoloji, yönetim teknikleri ve çeşitli yazılımlardan yararlanmak, CRM kavramının temelini oluşturur (Bozgeyik, 2001, s.6)

CRM kavramıyla ilgili farklı gruplar, kişiler ve kurumlara ait tanımlar aşağıda verilmiştir. Buna göre CRM;

- Şirketlerin müşterileriyle olan ilişkilerini tanımlayıp geliştirmelerini sağlayan bir yönetim felsefesidir. Müşteri ile kişisel ilişkiler kurmanın ve ölçek ekonomisinin aynı potada birleşmesidir (Gartner Group),
- Müşterileri korumak, büyüme ve karlılığı sağlamak için müşterilerin ihtiyaçlarını anlamak ve bu ihtiyaçlara cevap verecek şekilde organizasyonel süreçleri değiştirmektir (Trestini),



- Gittikçe artan rekabet ortamı içerisinde işletmelerin maliyetlerini azaltıcı, gelirlerini arttırıcı, müşterilerini elinde tutmayı hedefleyen, müşteri sadakatini sağlamaya çalışan etkin bir stratejidir (Yurdakul),
- İşletmeyi müşteri odaklı hale getirerek, değerli/karlı müşteriler arasındaki farklılıkları belirleme işlemidir (Kırım),
- Müşterileri kazanma ve korumanın yanı sıra müşteri sadakatini ve karlılığını arttırmak, uygun iletişim araçlarını kullanarak müşteri davranışlarını anlamak ve etkilemek için benimsenen bir işletme felsefesidir (Swift),
- En değerli müşterileri seçmek ve yönetmek için geliştirilmiş olan bir işletme stratejisidir (www.crmguru.com).

Küreselleşmenin etkisiyle küçülen dünyamızda işletmeler için büyük bir rekabet ortamı oluşmuştur. Sayısız alternatifin bulunduğu bu ortamda, müşteriler tercihlerini kendilerinin değerini bilen ve güven veren işletmelerden yana kullanmaktadırlar.

İşletmeler bu yoğun rekabet ortamında artık geleneksel pazarlama yolu ile yeni müşteriler bulma, mevcut müşterileri elde tutma, bu müşterilere daha fazla satabilme ve müşteride sadakati sağlama gibi amaçlarını gerçekleştirmede zorluklar yaşamaktadır. Bu zorlukların üstesinden gelebilmek için birçok işletme, pazarlamada yeni açılım olan CRM' yi kullanmaya başlamıştır.

## **2.2. CRM' nin Gelişimi**

İşletmelerin zaman içindeki gelişimine paralel olarak, pazarlama anlayışlarında da gelişmeler olmuştur. Artan rekabet baskısı, çok kısa sürede taklit edilebilen ürünler, mal ve hizmet farklılaştırmasındaki zorluklar işletmeleri pazarlama alanında yeniden yapılanmaya zorlamıştır.

Günümüzün pazarlama uygulamalarına gelinceye kadar pazarlama anlayışı ve yaklaşımlarındaki değişimler ana hatlarıyla aşağıdaki gibi sıralanabilir (Tek, 1999, s.12):

Üretim anlayışı (“Ne üretirsem satarım” anlayışı),

Ürün anlayışı (“Kaliteli üretirsem satarım” anlayışı),

Satış anlayışı (“Satmasını bilirsem satarım” anlayışı),

Pazarlama anlayışı (“Tüketici ihtiyaçlarını tam karşılayarak, onu memnun edersem başarırım” anlayışı),

Sosyal pazarlama anlayışı (“Toplumun çıkarlarını koruyup tüketici ihtiyaçlarını tam karşılayarak, onu memnun edersem başarırım” anlayışı).

1800’lü yıllarda, arzın talebi karşılayamadığı, işletmelerin vasıfsız işgücü ve standart ürünlere dayanarak kitle üretiminde bulunduğu ve rekabetin üretim odaklı olduğu görülmektedir.

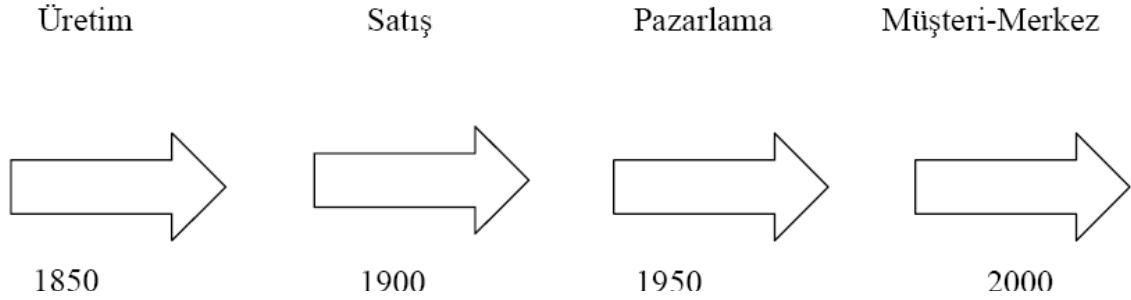
1900’lerin başlarında ise rekabetin artması ve müşterilerin farklı işletmeleri tercih etmeye başlamasıyla bir anlamda talep azalması olmaya başlamış, özellikle teknolojik gelişmeler yaşandıkça, kitle üretiminin temel özelliği olarak kabul edilen standart ürünlere olan talep daralırken, çeşitlenmiş ürünlere olan talep hacmi büyük bir artış göstermiştir (Selamoğlu, 1998, s.10). Önceden sadece ürününü üretip müşterinin satın almasını bekleyen işletmelerde, üretimin ve çeşitliliğin artmasıyla satış gücü baş göstermiş, reklâm, satış ve dağıtım kanalları ön plana çıkmıştır.

1950’lerde, işletmeler sadece satış gayesi ile yapılan çabalarla müşteriye ulaşamayacaklarını anlamışlar, müşterinin istek ve ihtiyaçlarını dikkate almaya başlamışlardır. Satış için ikna etmeye çalışmak yerine, müşterinin istediğinin gerçekte ne olduğunun kavranmaya çalışılması söz konusu olmuştur.

2000’li yıllara gelindiğinde, gerek rekabet gerekse çeşit sayısının çok olmasından dolayı müşteri odaklı bir bakış açısı şirketlere yerleşmeye başlamıştır.

İşletmeler açısından uzun dönemli müşteri bağlılığı, yeni rekabet alanında önemli bir değişimi beraberinde getirmiştir. Sanayi devrimi ile ortaya çıkan kitle üretim ve kitle pazarlama, yerini merkezde müşterinin yer aldığı yaklaşımlara bırakmıştır. Eski “tasarla-yap-sat” modelinin (ürün odaklı yaklaşım) yerini “sat-yap-yeniden tasarla” modeli (müşteri odaklı yaklaşım) almıştır. Bu bağlamda müşteri sadakatini gerçekleştirmek ve sürekli kılmak için çeşitli yöntem ve tekniklerin kullanıldığı görülmektedir. Bu yöntemlerden en önemli olanı, müşteriyi merkeze alan ve müşteri memnuniyetinin ötesine geçmeyi amaçlayan CRM yaklaşımıdır. Bu yaklaşım aynı zamanda yeni pazarlama anlayışını da ortaya koymaktadır (Rygielski, Wang ve Yen, 2002, s.488).

**Şekil 2.1 Pazarlama Anlayışının Tarihsel Gelişimi**



Kaynak: Bose, R., CRM: Key Components For Its Success, Industrial Management Data Systems, 2002

CRM’ nin yaygın olarak kullanılmasıyla, bazı yeni kavramlar ortaya çıkmıştır. Bu kavramlar temel olarak sadık müşteri, müşteri tatmini, cüzdan payı, müşteri karlılığı gibi kavramlardır. Bu kavramlara bakıldığında pazarlama satış stratejilerinin üretim satarız anlayışından müşteri odaklı bir hale geldiği görülebilmektedir. Fakat CRM sadece müşteri odaklılık anlamına gelmemektedir. Müşteri odaklılıktan ziyade müşteri merkezlilik söz konusudur. Yani ne üreteceğinden başlayıp, nasıl duyuracağına ve müşterilerle nasıl kalıcı ilişki geliştireceğine kadar geçen tüm süreç içinde çıkış noktası olarak müşteriyi temel almaktadır. Müşteri odaklılıkta ise “ne üreteceğine karar verdikten sonra, kitlesel olarak pazarlama” mantığı mevcuttur (Kırım, 2001, s.51).

### **2.3. CRM’ nin Faydaları**

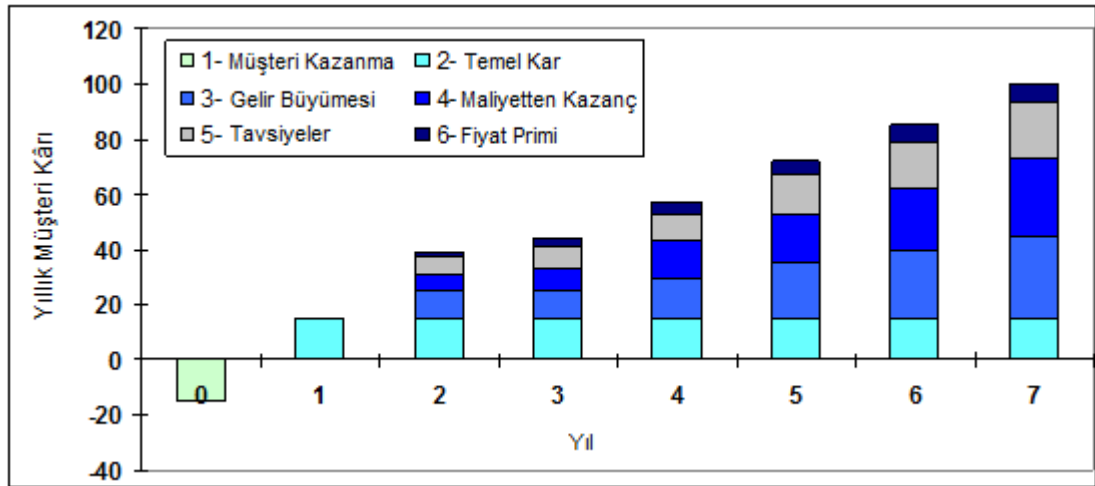
Ürünlerin birbirlerine bu kadar benzediği bir dünyada farklılığı yaratabilmenin ve rekabetçi bir üstünlük yakalayabilmenin tek yolu, müşterileri bire bir tanımak ve onlar

için bire bir üretim ve bire bir pazarlama yapmaktan geçmektedir. CRM' yi gerçekçi bir şekilde uygulamak firmaya çok önemli bir rekabet üstünlüğü kazandırmaktadır (Kırım, 2001, s.135).

CRM sayesinde, bir işletme için en kârlı ve esas olan müşteri profili belirlenerek, bu kitlenin işletmenin elinde tutulması için, müşteri ile ilgili bilgiler toplanarak, istenen değişiklikler yapılarak, ürün ya da hizmet çeşitliliği ve fiyat esnekliği sağlanmaya çalışılır. Burada amaç, işletme için müşteri merkezli bir stratejinin oluşturulmasıdır (Yereli, 2001, s.30).

İşletmelerde CRM' nin öneminin giderek artmasına paralel olarak, müşteriyle uzun dönemli kalıcı ilişkiler kurmanın ve müşteride sadakati yaratmanın da önemi giderek artmaktadır. Müşterinin işletmeyle ilişkide olduğu süre arttıkça, işletmeye getirisi de artmaktadır. Şekil 2.2' de bu durum görülmektedir.

Şekil 2.2 Müşteri Bağlılığının Kar Üzerine Etkisi



Kaynak: Swift, R. S., Accelerating Customer Relationships, Prentice Hall PTR A Pearson Education Company, Upper Saddle River, 2001

İşletmeleri CRM' ye iten birçok neden vardır. Bu nedenler aşağıdaki gibi sıralanabilir (Taşpınar, 2006, s.20):

- Yeni bir müşteriye ürün veya hizmet satmanın maliyeti, mevcut müşteriye satmaktan 6 - 7 kat daha maliyetlidir,
- Şirketler 5 senede mevcut müşterilerinin yaklaşık yarısını kaybetmektedirler,
- Mevcut müşterinin sadakatini sadece %5 artırmak, %25 - %100 arasında gelir artışı kazandırmaktadır,
- Tipik bir mutsuz müşteri, kuruma 8 ila 10 potansiyel müşteri kaybettirir,
- Zamanında ve doğru müdahale ile %85 oranında mutsuz müşteri geri kazanılabilir.

Tüm bunlar dikkate alındığında işletmeler hızla CRM uygulamalarına başlamakta ve bu sayede çok sayıda fayda elde etmektedir. CRM' nin işletmelere sağladığı faydalar aşağıdaki gibi sıralanabilir (Kırım, 2001, s.135-137):

- **Müşteri edinme maliyetlerinin düşük olması:** CRM uygulayarak pazarlama, iletişim, müşteri hizmetleri, müşteri tatmini ve müşteri takibi gibi faaliyetler için yapılan harcamalarda tasarruf sağlandığından, işletmelerin müşteri edinme maliyetleri azalmaktadır.
- **Düzenli bir iş hacmini korumak için çok sayıda müşteri edinmeye ihtiyaç duymama:** Müşterilerle uzun süreli ilişkiler amaçlandığından ve sadık müşterilerin sayısı arttığından yeni müşteri edinme ihtiyacı azalmaktadır.
- **Düşük satış maliyetleri:** Mevcut müşteriler genellikle satışa karşı daha yanıt verici olduklarından satış faaliyetlerine bağlı maliyetler azalmaktadır. Bununla birlikte, teknoloji destekli modern müşteri etkileşim kanallarının kullanılması müşterilerle ilişkileri daha etkin hale getirmekte ve pazarlama faaliyetlerinin daha düşük maliyetlerle gerçekleşmesini mümkün kılmaktadır.

- **Daha yüksek müşteri karlılığı:** Çapraz satış, satış artırma ve takip eden satışlarla müşterinin cüzdan payı arttırılmakta, tatmin düzeyi yüksek olan mevcut müşterilerin referansları ile de çok sayıda karlı müşteri işletmeye çekilmektedir.

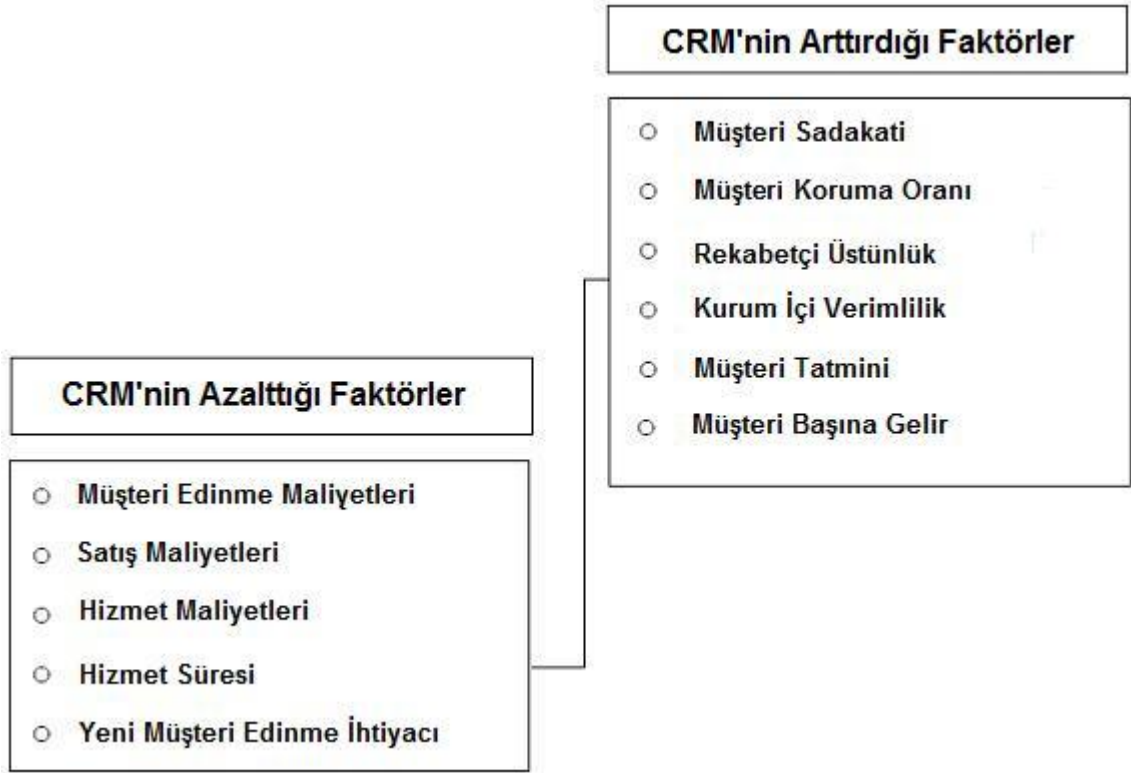
- **Müşteri koruma oranında ve müşteri sadakatinde artış:** Müşteriler kurulan uzun süreli ilişkilere bağlı olarak daha çok ve daha sık satın alma faaliyetlerinde bulunmaktadırlar. Ayrıca, müşterilerin daha fazla inisiyatif kullanmalarına imkan verilerek ilişkinin güçlendirilmesi sağlanmakta ve sonuç olarak müşteri sadakati artmaktadır.

- **Müşteri karlılığının değerlendirilebilmesi:** CRM uygulayan işletmeler, hangi müşterilerin karlı olduğunu, hangilerinin gelecekte karlı olabileceğini ve hangilerinin hiçbir zaman karlı olamayacağını öğrenebilmektedirler.

- **Tüm kurum içindeki verimi artırması:** CRM yalnızca satış, pazarlama ve müşteri hizmetlerinde değil, ilgili her birimde verimi artırma olanağı sağlar. Çünkü işletmenin tüm aktivitesi dâhili ihtiyaçlara göre değil, müşteri ihtiyaçlarına göre dizayn edilmektedir.

- **Tüm müşteri bilgilerini bütünleştirerek yeni fırsatlar sağlama:** Müşterilerle yüz yüze yapılan işlerde müşterilerle ilgili veriler toplanıp bir yerde depolanabilmektedir. Ancak, iş yapma kanalları çoğaldığında, örneğin e-mail, telefon, internet gibi kanallardan müşterilerle iş yapılmaya ve iletişime başlandığında, bu yeni veri depolarının yüz yüze iletişimden elde edilen veri depoları ile kaynaştırılması gerekmektedir. CRM ile hem geleneksel satış kanallarından elde edilen veriler hem de bu alternatif kanallardan sağlanan veriler sentezlenerek, maksimum müşteri bilgisi ve ilişkisi fırsatı sağlanabilmektedir.

Şekil 2.3 CRM' nin Etkileri



Kaynak: Agrawal, M. L., Customer Relationship Management (CRM) & Corporate Renaissance, Vol.3, No:2, 2003.

ISM CRM Başkanı Barton Goldenberg 3 yıllık dönemde CRM' nin işletmelere etkisini rakamsal olarak Tablo 2.1' deki gibi açıklamaktadır:

Tablo 2.1 CRM' nin İşletmelere Rakamsal Etkileri

Satış gelirlerinde	%10 artış
Müşteri kazanma oranında	%5 artış
Kar marjında	%1 artış
Müşteri memnuniyetinde	%3 artış
Pazarlama ve satış masraflarında	%10 azalma

Kaynak: Goldenberg, B., Customer Relationship Management :What Is It All About?, 1999

CSO Insights şirketinin 2006 yılında dünya çapında 1250' nin üzerinde şirketi kapsayan araştırmasında, 120' nin üzerinde parametre kullanılarak satış yapılarının ne

derece verimli çalıştığı değerlendirilmiştir. Şirketlerin CRM uygulamalarıyla elde ettikleri rakamsal sonuçlar şunlardır (www.csainsights.com):

- Müşterilerle % 61,9 oranında daha iyi iletişim sağlandığı,
- Müşterilerin ihtiyaçlarını % 50,3 oranında daha iyi tahmin edildiği,
- Yönetim yükünün % 41,7 oranında azaldığı,
- Gelirlerinin ise % 30,1 arttığı görülmüştür.

İyi planlanmış ve başarıya ulaşmış birçok CRM uygulaması olmasına rağmen CRM' ye yeni başlamış birçok firma bu tip sonuçlara ulaşamamaktadır. CSO Insights araştırma şirketi, 1200 CRM girişimini gözden geçirerek, firmaların projelerde CRM hedeflerine ulaşmalarını engelleyen kritik hataları belirlemiştir. Bunlar aşağıda sıralanmıştır (www.csainsights.com):

- CRM teknolojisinden çok fazla şey beklemek: CRM uygulamasını başarılı yapacak olan, teknolojinin akıllıca kullanımı ile etkili bir satış prosesinin optimizasyonudur. Eğer süreç temelde kusurlu ise teknolojinin bu sürece katkısı da az olacaktır. Dolayısıyla, araçları seçmeden önce bunlar ile ne yapılacağı net olarak bilinmesi gereklidir.

- İşletme çapında bir katılım olmaması,
- Süreç ve teknolojiye bu sistemi kullanacak olan insandan fazla önem verilmesi,
- Yöneticilerin aktif destek ve katılımının olmaması,
- Zayıf tanımlanmış stratejiler, hedefler ve amaçların varlığı,
- Yönetimin değişimi benimsemeden önce çok fazla değişimin yapılması zorunluluğu,
- Projenin mevcut işletme süreçlerine olan etkisinin dikkate alınamamış olması,
- Etkin olmayan proje ve kaynak yönetimi,
- Sistem içinde uyumlu olmayan teknolojilerin varlığı,
- Hangi aracın kullanılacağı konusunda karışıklık,

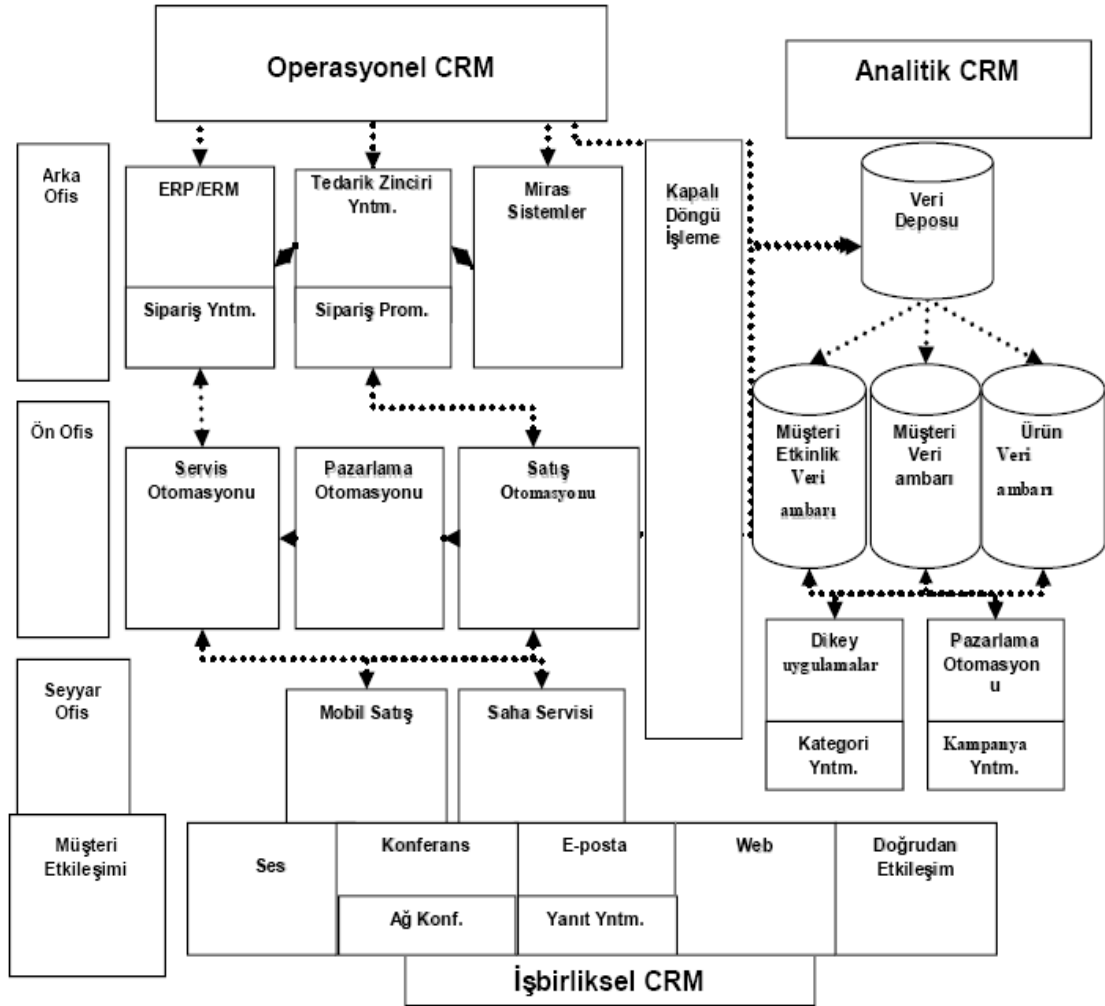


- Kirli veriler,
- İsteksiz ve direnç gösteren çalışanlara liderlik yapacak kişi eksikliği,
- Finansal kaynak yetersizliği,
- Denetimlerin eksikliği.

## 2.4. CRM Mimarisi

CRM mimarisi üç temel aşamadan oluşmaktadır. Bunlar, Operasyonel CRM, Analitik CRM ve İşbirliğine yönelik CRM' dir.

Şekil 2.4 Müşteri İlişkileri Yönetimi Mimarisi



Kaynak: Crm Nedir? (<http://www.erpcrm.com>)

### **2.4.1. Operasyonel CRM**

Operasyonel CRM, müşterilerle etkileşimi ve çalışanların süreci otomatik kullanmasını sağlamak için gerekli sistemleri kapsar. Müşteriyle doğrudan bağlantı kurmayı gerektirecek olan her faaliyet operasyonel CRM içerisinde yer almaktadır (www.sas.com).

Pazarlama, satış faaliyetleri ile başlayan süreçte, siparişlerin alınması ve yerine getirilmesi, satış sonrası destek hizmeti kapsam içerisinde. Müşteri hizmetleri, sipariş yönetimi, faturalama, satış ve pazarlama otomasyonu gibi tipik iş fonksiyonlarının kapsandığı CRM çözümlerinden oluşur.

Operasyonel CRM, arka ofis (Back office), ön ofis (Front office) ve seyyar ofis (Mobile Office) unsurlarından oluşur.

### **2.4.2. Analitik CRM**

Analitik CRM, kullanıcılara ait verilerin işlenmesi, analiz ve tahminlere dönüştürülerek raporlanması işlemlerini gerçekleştirmektedir. Böylelikle CRM' nin operasyonel ve entegrasyon özellikleri üzerine analiz ve raporlama özellikleri eklenmektedir (www.microsoft.com).

Analitik CRM, veri ambarı, OLAP ve veri madenciliği uygulamaları gibi uygulamaları içerir.

Doğru kampanyalarla doğru müşterilere ulaşmak isteyen işletmelerde, Analitik CRM' den Operasyonel CRM' ye, Operasyonel CRM' den de Analitik CRM' ye verilerin dönüyor olması gerekmektedir. Çünkü yalnızca Operasyonel CRM ile siparişin alınması, yerine getirilmesi uzun vadede işletme için yeterli olmamaktadır. İşletmenin yeni çıkardığı ürünü yada hazırladığı kampanyayı müşterilerine sunmadan önce, müşterilerinin kim olduğunu bilmesi gerekmektedir. Örneğin çağrı merkezinden müşterilerine ulaşarak yeni kampanyalarla ilgili teklifler sunan bir işletmenin, “O müşteri kimdir?”, “Hobileri nelerdir?”, “Evlidir, çocuğu var mıdır?” gibi davranışsal alışkanlıklarını Analitik CRM' den öğrenmiş olması gerekmektedir.

### 2.4.3. İşbirlikli CRM

İşbirliğine yönelik CRM, aslında diğerlerinin en uygun birleşiminden oluşmaktadır. Müşteriler ile şirketler arasında tam anlamıyla bir etkileşim ve koordinasyon ağının oluşmasına imkân veren, farklı iletişim kanallarından (web, telefon, e-posta vb) gelen bilgilerin, değere dönüştürülmesini sağlayan bir süreçtir. İşbirliğine yönelik CRM çözümleri müşteri ile etkileşime imkân veren tüm fonksiyonları içermektedir. (www.microsoft.com)

Bu sayede müşteriler her an her yerde bilgi alabilme, işlem yapabilme imkânı yakalarken, işletmeler de yeni müşterilere ulaşabilmekte ve işlem hacimlerini arttırabilmektedirler.

### 2.5. CRM Süreci

Müşteri ilişkileri yönetiminde süreç, ilk olarak müşteriler hakkında olabildiğince bilgi toplamaya dayanmaktadır. Elde edilen verilerin ışığı altında müşterilerin gruplandırılması, karakteristik özelliklerinin çıkarılması ve son olarak doğru ürünün doğru müşteriye yönlendirilmesine olanak sağlar. (Kırım, 2001, s.49)

Buna göre CRM süreci, müşteri seçimi, müşteri edinme, müşteri koruma, müşteri derinleştirme olmak üzere 4 aşamadan oluşur (CRM Institute Turkey).

**Müşteri Seçimi:** CRM' nin bu evresinin ana amacı "En karlı müşteri kim?" sorusunun cevabının bulunmasıdır. Bu çerçevede şu çalışmalar yapılmaktadır:

- Hedef kitlenin belirlenmesi,
- Segmentasyon,
- Kampanya planları,
- Marka ve müşteri planlamaları,
- Yeni ürün lansmanları.

**Müşteri Edinme:** Seçilen müşterinin kazanılması için yapılan tüm çalışmalar, müşteri edinme aşaması içinde değerlendirilmektedir. Bu evrenin asıl amacı satışı gerçekleştirmektir. “Belirli bir müşteriye en etkili yoldan satış nasıl yapabiliriz?” sorusunun yanıtı bu evrenin ana amacını oluşturmaktadır. Bu kapsamda,

- İhtiyaç analizleri,
- Teklif oluşturma,
- Kapanış adımları ile ilgili çalışmalar yapılmaktadır.

İşletmeye kazandırılması gereken müşteri profiline doğru saptanması gerekmektedir. Müşteri edinme yüksek maliyete sahip olduğundan, çalışmalarına doğru hedef kitleyi saptayarak başlamak oldukça kritiktir. Örneğin; arabası olmayan bir kişiye lastik katalogu göndermek mantıklı bir yaklaşım değildir.

**Müşteri Koruma:** “Bu müşteriyi ne kadar süreyle elimizde tutabiliriz?” sorusunun yanıtının arandığı evredir. Amaç müşteriyi kuruma bağlama, onu kurumda tutabilme ve ilişkinin sürekliliğini sağlamaktır.

- Sipariş yönetimi,
- Teslim,
- Taleplerin organizasyonu ve
- Problem yönetimi müşteri koruma evresi kapsamında yapılan çalışmalardır.

Kaybedilen müşterinin yerine yeni müşteri elde etme çabaları yüksek maliyete sahip olduğundan, mevcut müşteriyi koruma işletmeler için oldukça önem taşımaktadır. Bu sebeple, mevcut müşterinin beklentileri iyi takip edilmeli ve tekrar aynı ürüne yönlendirilmesi sağlanmalıdır.

**Müşteri Derinleştirme:** Kazanılmış bir müşterinin sadakat ve kârlılığının uzun süre korunması ve müşteri harcamalarındaki payın yükseltilmesi için gereken adımları

içermektedir. Amaç sürekliliğin sağlandığı ilişkiden yeni faydalar sağlamaktır. Bu kapsamda,

- Müşteri ihtiyaç analizleri ve
- Çapraz satış kampanyaları ile ilgili çalışmalar yapılmaktadır.

Müşteri derinleştirmeye örnek olarak “hepsiburada.com” verilebilir. Kullanıcının ilgilendiği ürünler göz önüne alınarak, o ürünü tercih eden diğer müşterilerin satın aldığı ürünler de kullanıcıya önerilmektedir. Mesela; mutfak robotu alan bir kişiye, el blenderı önerilmesi söz konusu olabilir.

CRM’ nin evreleri şematik olarak Şekil 2.5’ te gösterilmiştir:

Şekil 2.5 CRM Evreleri



Kaynak: Sicimoğlu, H., Gördüğümüz Işık UFO Değil CRM, Bütünleşik Pazarlamada Marketing Türkiye, Yıl:1, Sayı: 7, Temmuz 2002.

Yeni bir müşteri edinme maliyeti, mevcut müşterileri elde tutma maliyetinin yaklaşık 5 katı, memnun olmayan müşteriyi geri kazanmanın maliyeti ise, mevcut müşteriyi elde tutmanın yaklaşık 10 katıdır (Massey, Weiss ve Holcom, 2001, s.156).

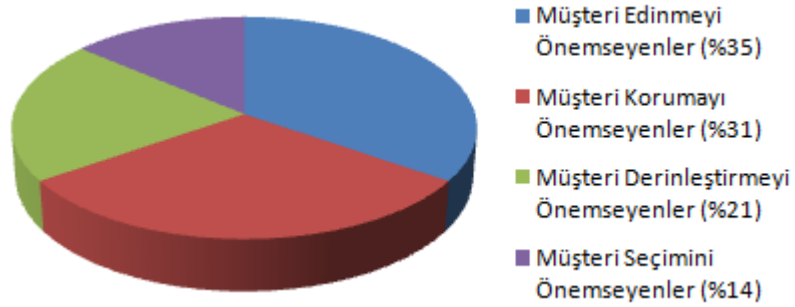
Mevcut müşterilerden sağlanan brüt gelir, yeni müşterilerden kazanılanın yaklaşık 2 katıdır. Müşteri tutmada %5 oranında bir iyileşme, kârı %25 artırmaktadır. Bu nedenle şirketler, mevcut müşterilerini elde tutmaya yönelik bir takım sadakat programları geliştirirler (Russel, 2001, s.3).

Amerika’da şirketlerin birçoğu müşterilerini koruma ve derinleştirme üzerine odaklanmakta, müşteri seçimi ve müşteri edinme adımları, özellikle oturmuş pazarlarda giderek daha az önemsenmektedir. Şirketler pazar paylarını değiştirme maliyetlerine katlanmak yerine, sadık müşteriler yaratıp müşterilerindeki cüzdan paylarını derinleştirmeyi daha ekonomik bulmaktadırlar.

Ülkemizdeki duruma bakıldığında CRM Institute Turkey’ in ortaya koyduğu sonuçlar Türk şirketlerinin daha çok müşteri edinme aşamasına önem verdiğini göstermektedir (www.crminturkey.org).

Türk şirketlerinde konu ile ilgili kişiler arasında yapılan araştırmalara göre ortaya çıkan dağılım Şekil 2.6’ da görülmektedir.

**Şekil 2.6 Türk Şirketlerinde CRM Yaklaşımı**



Bu dağılım, şirketlerin daha çok, satış kaygılarında olduğunu göstermektedir. “Müşteri Seçim” ekseninin rağbet görmemesi de şirketlerin ürünlerini konumlandırmada ve temel pazarlama ilkelerinde işi şansa bıraktıklarını göstermektedir.

## 2.6. Türkiye’de CRM Uygulamaları

Bir işletmenin başarısı, müşterilerini ne kadar iyi tanıdığına, pazardaki potansiyeli müşteriye dönüştürmede ne ölçüde başarılı olduğuna ve kazandığı müşteriye elde tutmadaki performansına bağlıdır. Başarılı işletmeleri diğer işletmelerden ayıran temel özellik, ürün ve hizmet geliştirirken bu sürece, işletme çalışanları ile müşterilerin görüşlerini dâhil etmesi durumudur. Diğer bir ifade ile ürün ve hizmet geliştirme sürecinin başlangıç noktasını tamamıyla müşteri istek ve ihtiyaçları oluşturmaktadır (Tekin ve Çiçek, 2003, s.4).

Migros’un Migros Club Kart’ı ile Bonus Card’ın altyapıları birleştirilerek, Türkiye’de ilk kez hem çok markalı hem de markaya özel sadakat kartı uygulaması olan Money Card yaratılmıştır. Müşterilere, Visa logosu taşıyan, temassız ödeme özelliğine sahip olan, aynı zamanda Migros Club Kart’ın özelliklerini taşıyan Money Card ile Migros, Tansaş, Şok, 5M, Macrocenter ve Sanal Marketten yapılan alışverişlerde ‘money’ adı verilen ve kasada para yerine geçecek puan kazanma şansı sunulmuş oldu. Böylece hem money hem de bonus puan kazanabilecek olan müşteriler, daha çok bedava alışveriş yapabileceklerdir ([www.milliyet.com.tr](http://www.milliyet.com.tr)).

Turkcell, 2002 yılından bu yana “Müşteriy(i)e Yolculuk” programı kapsamında, müşterilerini çeşitli kategorilere ayırıp, onların ihtiyaçlarına yönelik özel çözümler sunmaktadır. Örneğin gençlik paketi Gnctrkcell’ye üye olanlar şifreleriyle gittikleri sinemanın bitiminde, koltuğun altında kendilerine hediye edilen bir mp3 çalar bulabilmekte, ya da cep telefonunu değiştirdiği için numaralarını kaybeden bir first class aboneli anne, çocuğu hastalandığında Turkcell’i arayarak, doktorun numarasını nasıl bulabileceğini sorabilmektedir. Turkcell ekibi, bu abonenin verdiği ipuçlarından yararlanarak, (telefon numarasındaki birkaç rakam ya da aranan tarih gibi) telefon görüşmelerini tarayıp doktorun numarasını aboneye ulaştırmaktadır. Bir başka örnek ise, yeni doğum yapmış bir annenin, telefonundaki tüm veriler silindiği için Turkcell’i arayarak, ofis bilgisayarındaki kayıtları istemesi gösterilebilir. Turkcell ekibi, ofise eleman göndererek yedeklenen verilere ulaşıp aboneye göndermektedir. 2007 yılında Turkcell, “Müşteriy(i)e Yolculuk” adı altında yürüttüğü bu program ile dünyanın en önemli CRM ödüllerinden birinin dağıtıldığı “Gartner EMEA CRM Excellence

Awards” yarışmasında finale kalan ilk ve tek Türk şirketi olmuş ve yapılan oylama sonucu birincilik ödülüne layık görülmüştür ([www.sabah.com.tr](http://www.sabah.com.tr)).

2001 yılında, tüketici kitlesi olan kadınlarla ilişki kurmak ve bu ilişkiyi geliştirmek için “Omo Kadınlar Kulübü”nü kuran Omo, her gün çoğunluğunu ev hanımlarının oluşturduğu çok sayıda üyesine çocuk danışmanlığı, günlük yemek mönüsü, aile doktorluğu, psikolojik danışmanlık, güzellik gibi konularda hizmet vermektedir. Bir taraftan müşterileriyle sürekli iletişim halini korurken, diğer taraftan da birçok insanın ihtiyaç duyduğu pek çok konuda ücretsiz hizmet sunmaktadır.

Müşterilerine en iyi hizmeti vermeye yönelik özel bir CRM altyapısını hayata geçiren Doluca, veri tabanında kayıtlı şarap sever kitleleriyle sürekli iletişim kurarak, müşterilerinin ihtiyaç ve taleplerine hızlı cevap vermeyi amaçlamıştır. Bu amaçla şarap severlere düzenli olarak ücretsiz şarap ve yaşam dergisi Bouquet’i göndermektedir. Ayrıca gerçekleştirdiği eğitim, etkinlik ve kampanyalarını, müşteri kayıt formlarından alınan ve başka hiç bir kurum ile paylaşılmayan müşteri e-posta adreslerine göndermektedir. Tüm bu hizmetlerin yanı sıra Doluca, Bağbozumu Gezileri, şarap tadım seminerleri, şarap ve ürünler hakkındaki bilgilerin geniş ve detaylı bir şekilde aktarıldığı ve katılımcı sertifikalarının verildiği Doluca Şarap Eğitim Programı gibi projelere imza atmıştır ([www.crmpro.com.tr](http://www.crmpro.com.tr)).

Garanti Bankası, 2008 yılında çocukların tasarruf alışkanlığı kazanmasını sağlayacak, akıllı kumbara özelliği taşıyan özel ATM Minimatik’i hizmete sundu. Kart ve şifreyle çalışan, bozuk parayı tanıyan Minimatik’in, tasarruf kavramını eğlenerek öğreten bir ATM görevini üstlenmesi amaçlanmıştır. Çocuklar tıpkı yetişkinler gibi, Minimatik’e para yatırıp çekebilmekte, tasarruf etmek istedikleri miktarı tespit edebilmekte ve yaptıkları her işlemde sonra ne kadar paraları olduğunu görebilmektedirler ([www.milliyet.com.tr](http://www.milliyet.com.tr)). Minimatik, Garanti Bankasının gelecekte müşterisi olması istenilen çocuklarla küçük yaştan kurduğu bir bağ olarak değerlendirilebilir. Bankanın çocuklara yönelik ilk atılımı 2003 yılında Minibank ürünüyle başlamıştır. Yaşam boyu bankacılık fikriyle yürüttükleri çalışmalarla bir



yandan çocuklara, banka ve bankacılık ile ilgili temel bilgiler verilirken, aynı zamanda ailelere de çocuklarının geleceğini güvence altına alma fırsatı sunulmaktadır.

CRM' yi tüm restoranlarına gelen paket servis siparişlerini yönetmek için kurduğu Çağrı Merkezi'nde kullanan Burger King, bu uygulamayla restoranlarının hizmet bölgelerinin dijital haritalarını, mahalle, cadde ve sokak isimlerini bu sistemde kayıt altına almıştır. Müşteri ile iletişimin çok kısa sürdüğü fast food sektöründe, bu uygulamayı bir fırsat olarak değerlendiren şirket, böylece müşterisini daha iyi tanımak için uygun araçları sunan bir kanal da elde etmiş olmuştur. Şirket, paket servis kanalıyla cirosunu %20 artırmayı hedeflerken, müşteri bilgilerini değerlendirerek promosyonlarından reklam çalışmalarına tüm faaliyetlerini daha iyi planlayabileceğini öngörmektedir (<http://www.microsoft.com/turkiye/dynamics/basari/tabgida.msp>).

ToyotaSa, geliştirdiği CRM uygulaması ile doğrudan pazarlama altyapısı ve mobil iletişim, çağrı merkezi, internet ve bayi kanalları üzerinden yürütülen kampanyaların merkezi olarak yönetimini sağlamaktadır. Bu sayede ToyotaSa, müşterilerini daha yakından tanıyıp, müşterileriyle bire-bir iletişim kurabilmektedir. Doğrudan pazarlama prensiplerine uygun olarak kişiselleştirilmiş, hedef segmentin ihtiyaçlarına uygun kampanyalar düzenleyen ToyotaSa, müşteri memnuniyetini yüksek oranda artırabilmeyi hedeflemektedir (<http://www.veripark.com/tr/Portfolio/SuccessStories/toyotasa.asp>). ToyotaSa firması internet sitesi ile otomobil tutkunlarına beğendikleri araçları yaratma olanağı sunmaktadır. “Kendi Toyota'nı Yarat” uygulaması ile otomobil meraklıları, istedikleri Toyota modelini gövde tipi, renk ve motor seçeneğine göre seçip, aksesuarlarını ekledikten sonra aracın fiyatını öğrenebilmekte, ayrıca “Yetkili Satıcımız Sizi Arasın” bölümüyle, müşteriler belirledikleri bir ToyotaSa yetkili satıcısının, tercih ettikleri Toyota modeliyle ilgili kendilerini telefonla aramasını isteyebilmektedirler.

Trabzonspor Kulübü, maç biletlerini, kulübün dergisini ve lisanslı ürünlerini alarak, TV yayınlarını izleyerek kulüple zaten müşteri ilişkisinde bulunan taraftarlara hem gerçek müşteri saygınlığı hem de taraftar ayrıcalıkları kazandıracak bir yol olarak, kulüp üyeliği sistemini daha etkin hale getirmek için, CRM' den yararlanmaya

başlamıştır. Orta vadede 100 bin taraftarını kulüp üyesi yapmayı planlamaktadır (<http://www.microsoft.com/turkiye/dynamics/basari/ts.msp>). İlk olarak Trabzonspor ve takiben Galatasaray'da uygulamaya alınan sistemde, kulübün “müşteri” kitlesini oluşturan üye ve taraftarları, sporcuları ve kaynakları olmak üzere odaklanılan üç önemli unsur vardır. Kategorilere ayrılmış üye ve taraftarların dergi aboneliği, kombine kart alımı, sponsorluk, taraftar kartı, kulüp mağazalarından alışverişi gibi hareketlerinin kişi bazında takibi söz konusudur. Bu sayede taraftarların daha yakından tanınması ve kategorilere özel yaklaşımların geliştirilmesi mümkün olabilmektedir.

CRM altyapısı sayesinde DenizBank müşterilerine doğum günü, yeni yıl mesajları ve kampanya duyuruları gibi kişiselleştirilmiş mesajlar gönderebilirken, DenizBank müşterileri de internet bankacılığını kullanarak ya da çağrı merkezini arayarak her konuda bilgi alabilmektedir. Her müşteri için en verimli kanalın kullanıldığı bu sistem, operasyon masraflarını oldukça azaltmaktadır. Farklı kanallardan bankaya ulaşan müşterilerin yaptığı tüm işlemler şubelerdeki müşteri temsilcileri tarafından izlenebilmekte ve müşterilerin alışkanlıklarına ve tercihlerine uygun çözümler sunulabilmektedir (<http://www.veripark.com/tr/Portfolio/SuccessStories/denizbank.asp>).

Pazarlama ve halkla ilişkiler departmanındaki verimliliği arttırmak amacıyla CRM kullanmaya karar veren Delta Büro Gereçleri, 2 adet donanımlı araç olarak müşteri ziyareti esnasında müşterinin siparişi olduğu anda araçtan ürünleri direkt teslim edip anında faturasını teslim etmektedir. Aynı zamanda kırtasiyeye girildiğinde barkot entegrasyonu sayesinde kapıdaki barkodu okuyarak müşteri bilgilerine online ulaşabilmektedir. CRM uygulaması ile müşterilerini toptan, aratoptan, zincir mağaza şeklinde kategorize eden Delta, teklif sipariş yönetimi, raporlama, saha satış ve siparişlerini program üzerinden takip edebilmektedir (<http://www.comel.com.tr/basari/delta.html>).

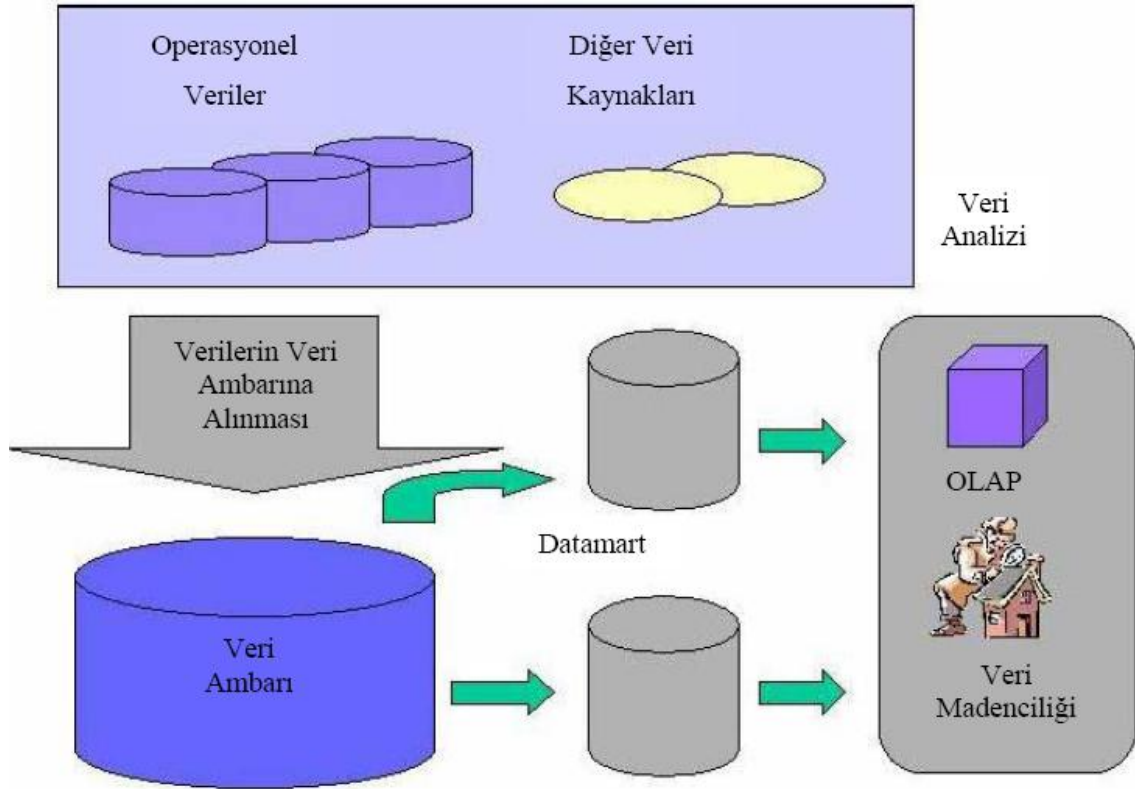
## **2.7. CRM’ de Veri Madenciliğinin Yeri**

İşletmenin müşterileriyle ilişkilerinin kurulması ve bu ilişkilerin geliştirilip sürekli hale getirilmesi gibi konularda bilişim teknolojileri kritik öneme sahiptir.

Bilişim teknolojilerinin etkin bir şekilde kullanıldığı veri madenciliği uygulamaları, işletmelerin kârlı müşteri hedefleri ve özdeşliğinin oluşturulmasında, CRM' nin kritik bir parçası olarak, önemli bir rol oynamaktadır (Insurance Systems Bulletin, 1993, s.6).

Şekil 2.7' den de görüldüğü gibi, işletme içindeki operasyonel veriler ya da diğer kaynaklardan gelen müşteri verileri veri ambarına alındıktan sonra OLAP küpleri oluşturulmak üzere ya da görünmeyen anlamlı bilgilerin keşfi için veri madenciliği uygulamalarına tabii tutulmak üzere veri analizine gönderilir.

**Şekil 2.7 CRM' de Verileri Anlamlı Bilgilere Dönüştürme Süreci**



Kaynak: Srivastava J., et. al., A Case for Analytical Customer Relationship Management, Computer Science & Engineering, University of Minnesota, Minneapolis, USA, 2002

Veri madenciliği teknikleri kullanılarak, müşterilerin yaşam tarzları, demografik özellikleri gibi işletmenin sahip olduğu müşteri verilerinden anlamlı bilgiler elde edilebilir. Örneğin, geçmişteki davranışlarından yola çıkılarak, gelecekteki davranışlarına yönelik tahmin modelleri oluşturmak mümkündür.

Firma bünyesinde geçmiş dönemlerde toplanan veriler ile; müşteri profili belirleme, kampanya yönetimi, müşteri sadakati belirleme gibi farklı amaçlı modeller geliştirilebilmektedir ve bu modeller ile istenilen karar destek amaçlı bilgilere çok hızlı bir şekilde ulaşılabilmektedir (www.spss.com.tr).

İşletmelerin müşterilerine sunduğu mağaza alışveriş kartları, alışveriş sırasında indirim sağlaması sebebiyle müşteriler tarafından düzenli kullanılırken, işletmeye, yapılan alışverişin müşteriyle ilişkilendirilebilmesi şansını doğurmaktadır. Bu sayede müşterilerin tercih ettiği markalar, fiyat aralıkları, alışveriş zamanları kısaca müşterilerin satın alma davranışları, işletme tarafından biliniyor olmaktadır. İşletmenin bu tip verilere sahip olması ve bu verilerden görünmeyen anlamlı bilgileri çıkarması rekabette kritik önem arz etmektedir.

Müşteriler tarafından hangi ürünlerin birlikte alındığının belirlenmesini sağlayan sepet analizi, o ürünlerin süpermarket içinde yakın alanlara konularak reyon düzeninin müşteri odaklı hale getirilmesini sağlamaktadır. Veri madenciliği araçları kullanılarak yapılan sepet analizinin güçlü yönü, insanların mantıksal olarak hangi ürünleri birlikte alabileceğinin düşünülmesi yerine, satış verilerine bakıp analiz ederek bu ilişkiyi veri aracılığı ile vermesidir (Baykasoğlu, Dereli ve Özen, 2004, s.2).

İşletmeler, tüm müşterilerine eşit oranda yönelmek yerine, bazı kriterlere uyan müşterilerine ürün ve hizmet sunmayı tercih etmektedirler. Bunun için, müşterileri demografik özellikleri, davranışları, ihtiyaçları ya da işletme açısından değerleri gibi ortak özelliklerine göre bölümlere ayırıp, her bölüme ayrı ayrı yaklaşım geliştirebilirler. Örneğin telekomünikasyon sektöründeki firmalar için, yüksek bölümde yer alan, işletme için değerli müşterilerin, veri madenciliği araçları kullanılarak geliştirilmiş tahmin modelinde işletmeyi terk etmesi öngörülüyorsa, bu gerçekleşmeden önce, müşterilere daha uygun tarifeye geçiş ya da bedava şehir içi dakika teklif edilebilir.

Verilerin toplanmasında, bu verilerin rekabette güçlü bir silah halini alan bilgilere dönüştürülmesinde ve bu bilgiler ışığında pazara yönelik stratejilerin belirlenmesinde veri madenciliği önemli bir araç olarak kullanılmaktadır.

### 3. BÖLÜM

#### UYGULAMA

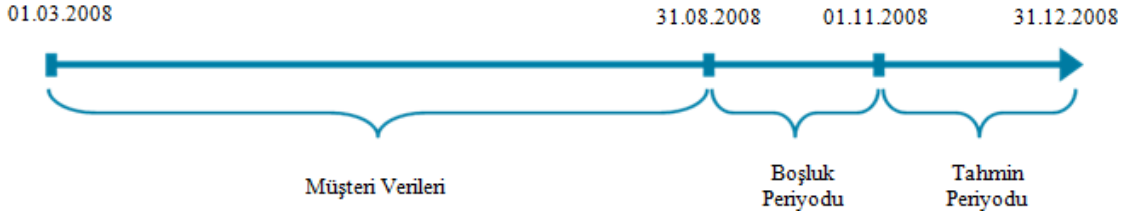
Bu uygulamada kullanılan veri seti, telekomünikasyon sektöründe faaliyet gösteren, Türkiye'nin en büyük kuruluşlarından bir tanesinin, Ege bölgesindeki İzmir, Manisa, Aydın, Muğla, Denizli, Afyon, Kütahya ve Uşak illerinden tesadüfi seçilmiş 92.217 abonesine ait Mart, Nisan, Mayıs, Haziran, Temmuz ve Ağustos 2008 kayıtlarından oluşmaktadır.

Veriler, tüzel olmayan gerçek abonelerden ev abonelerine ait olup, bu abonelerin 20.871 tanesinin Kasım-Aralık döneminde aboneliğini iptal ettirdiği, 71.346 tanesinin aboneliğinin ise devam ettiği bilinmektedir.

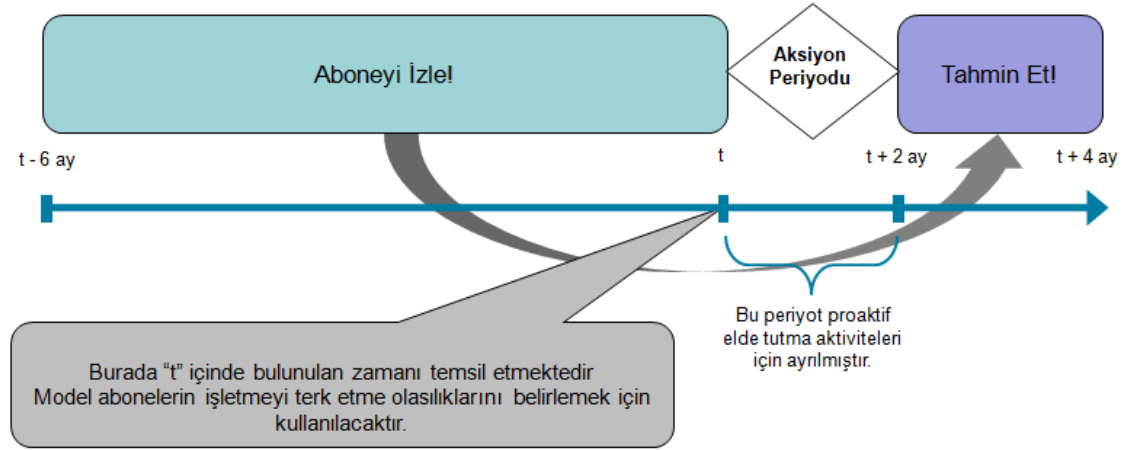
Modelde kullanılan abonelerin hepsi 01.03.2008 tarihinde işletmenin mevcut abonesi olan kişiler arasından seçilmemiştir. Bir kısmı Mart ve öncesi, bir kısmı ise Nisan, Mayıs ya da Haziran döneminde işletmenin abonesi olmuştur. Telekomünikasyon sektörü, hızla müşteri değişimi yaşayan, stabil olmayan bir sektör olduğundan, değişik zamanlarda işletmeye gelmiş abonelerle modelin geliştirilmesi daha gerçekçi olacaktır. Dikkat edilmesi gereken diğer bir nokta, Temmuz ya da Ağustos döneminde işletmenin abonesi olan kişilerin, işletmeyle ilişkide oldukları süre yeterli bulunmadığından modele alınmadığıdır.

Uygulamada, 01.03.2008 ve 31.08.2008 tarihleri arasındaki müşteri verileri kullanılarak yapılmış olan modelden sonra, 2 aylık bir boşluk periyodu bırakılması uygun görülmüştür. Burada amaç, proaktif pazarlama faaliyetleri için yeterli zamanı sağlamaktır. Proaktif pazarlama faaliyetleri, geliştirilen modelin yüksek olasılık vererek tahmin periyodunda işletmeyi terk edeceğini öngördüğü aboneler için işletme tarafından alınan aksiyonlardır. Örneğin, daha ucuza daha fazla konuşma sağlayacak tarife değişikliği, bedava şehir içi dakika, ya da belirlenen süre kadar işletmede kalınacağı taahhüt edilirse ücretsiz telefon hediye edilmesi proaktif pazarlama faaliyetlerine örnek olarak gösterilebilir.

Şekil 3.1 Uygulama Periyodu - 1



Şekil 3.2 Uygulama Periyodu - 2



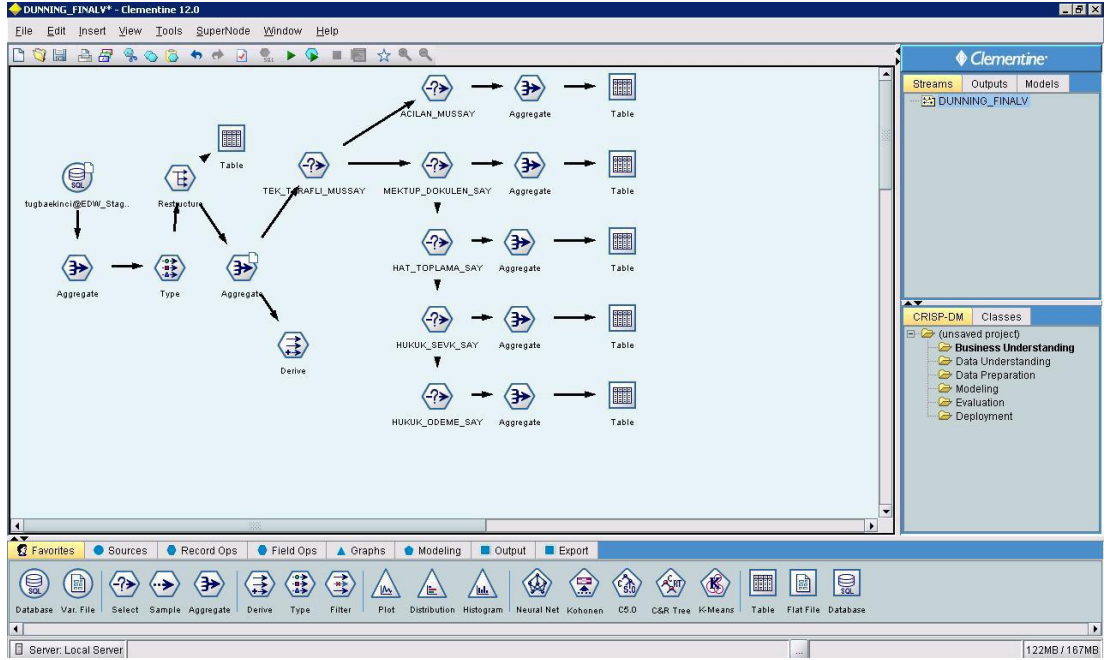
İşletmenin veri ambarından, müşterilerine ait 6 aylık verileri alınarak Microsoft MS SQL Server 2008 ile yeniden yapılandırılmış ve daha sonra bu verilerden anlamlı, bilinmeyen, daha önce fark edilememiş ve işletme için katma değer oluşturabilecek yeni enformasyon ve bilgiler elde edebilmek amacıyla, Spss Clementine programı ile veri madenciliği yöntemleri kullanılmıştır.

Spss Clementine programı, anlaşılması kolay grafikleri, güçlü fonksiyon seti, görsel arayüzü ile birçok firma ve kişi tarafından tercih edilmektedir. Veri madenciliği ve bilgi keşfi konusunda yayın ve araştırma yapan KDnuggets dergisi, Mayıs 2008'de yaptığı "En çok tercih edilen veri madenciliği çözümü" konulu araştırmasında Spss Clementine, Sas, Kxen, Oracle DM, Sql Server, Matlab gibi çözümleri geride bırakarak birinci olmuştur.

Spss Clementine kredi kartı dolandırıcılığı tespiti, müşteri davranışı analizi gibi pek çok alanda uygulanabilir.

Şekil 3.3' te Spss Clementine programına ait bir ara yüz görülmektedir.

**Şekil 3.3 SPSS Clementine Programına Ait Bir Arayüz**



Veri madenciliği için standart bir süreç söz konusudur. Bu standart süreç NCR, DaimlerChrysler, SPSS ve OHRA Bank'ın oluşturduğu konsorsiyum tarafından ortaya atılan CRISP-DM metodolojisi kapsamında incelenecektir. İlerleyen aşamalarda bu sürecin, uygulamada kullanılacak olan veri seti için nasıl işlediği açıklanmıştır.

### 3.1. İşi Anlama

Günümüzde işletmelerin mevcut müşterilerini kaybetmesi yaşadıkları en büyük problemlerinden biridir. Rekabetin yoğun yaşandığı, her geçen gün pazarda yeni bir rakibin benzer mal ve hizmetler sunmaya başladığı sektörlerde müşterilerin rakipleri tercih etmesi çok sık görülmektedir.

Müşteriler işletmeyi terk etmeden önce, işletme hangi müşterilerin ayrılacağı bilgisini elindeki verilerden çıkarabilirse, o müşterilere özel tedbirler alabilir ve müşteri kayıp oranını azaltabilir.

Bu tez çalışmasında amaç, uygulamaya konu olan firmanın sunduğu ürün ve hizmetlerden vazgeçen müşterileri tespit etmek amacıyla ayrılan müşteri analizini

(churn analizi) gerçekleştirmektir. İşletmenin abonesi olduğu sürede en az bir defa telefonunu kullanmış (evinden arama yapmış) ‘aktif’ ev aboneleri için Lojistik Regresyon Analizi kullanılarak ve işletmenin abonesi olduğu sürede hiç telefonunu kullanmamış (evinden arama yapmamış) ‘deaktif’ ev aboneleri için Karar Ağacı tekniği kullanılarak verilerin analiz edilmesi hedeflenmiştir.

### **3.2. Veriyi Anlama**

Veriyi anlama basamağının ilk aşamasında verinin elde edilmesi gerekmektedir. Bunun için uygulamaya konu olan firmanın veri tabanındaki değişik tablolarda tutulan abone bilgileri, abonenin sahip olduğu hattın özellikleri, hattın kullanım durumu ile ilgili bilgiler ilgili tablolardan çekilmiştir.

Verinin daha yakından tanınması için incelenmesi gerekmektedir. Ancak bu inceleme sırasında unutulmaması gereken nokta, verinin değiştirilmemesidir. Bu kapsamda veri setinde hatalı, eksik ya da tekrarlı kayıtların olup olmadığı incelenmiştir.



Tablo 3.1’ de veri setini oluşturan bazı değişkenler ve açıklamaları görülmektedir.

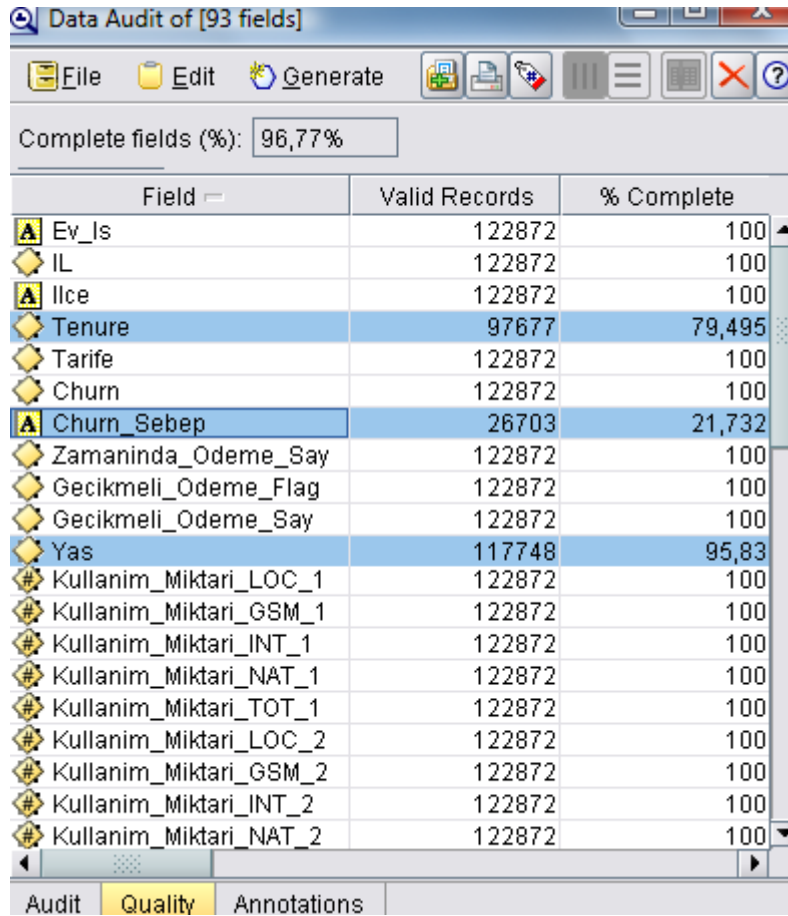
**Tablo 3.1 Değişkenler ve Açıklamaları**

Değişken	Açıklama
Ev_Is	Telefonun ev telefonu mu iş telefonu mu olduğu bilgisi
IL	Abonenin yaşadığı il
Ilce	Abonenin yaşadığı ilçe
Tenure	Abonenin işletmenin ne kadar zamandır abonesi olduğunu gösteren süre
Tarife	Aboneye ait tarife
Churn	Abonenin işletmeden ayrılıp ayrılmadığı bilgisi
Churn_Sebep	Abonenin işletmeden ayrılmışsa ayrılma sebebi
Zamaninda_Odeme_Say	Zamanında ödenen fatura sayısı
Gecikmeli_Odeme_Flag	Gecikmeli ödenen faturanın olup olmadığı bilgisi
Gecikmeli_Odeme_Say	Gecikmeli ödenen fatura sayısı
Kullanim_Miktari_LOC_N	N. aydaki şehir içi kullanım miktarı
Kullanim_Miktari_GSM_N	N. aydaki Gsm kullanım miktarı
Kullanim_Miktari_INT_N	N. aydaki uluslararası kullanım miktarı
Kullanim_Miktari_NAT_N	N. aydaki şehirlerarası kullanım miktarı
Kullanim_Miktari_TOT_N	N. aydaki toplam kullanım miktarı
Arama_Sayisi_LOC_N	N. aydaki şehir içi arama sayısı
Arama_Sayisi_GSM_N	N. aydaki Gsmi arama sayısı
Arama_Sayisi_INT_N	N. aydaki uluslararası arama sayısı
Arama_Sayisi_NAT_N	N. aydaki şehirlerarası arama sayısı
Arama_Sayisi_N	N. aydaki toplam arama sayısı (Loc+Na+Gsm+Int)
Oto_Odeme	Otomatik ödemenin olup olmadığı bilgisi
Recency_Out	6 aylık periyotta en son arama yapılan günden, 6. ayın sonuna kadar olan süre
Arama_Yapilmayan_Gun_sayisi_N	N. aydaki arama yapılmayan gün sayısı (Sadece yapılan aramalar bazında)
Adsl	Telefona bağlı Adsl olup olmadığı bilgisi
Kullanim_Ucreti_N	N. aydaki kullanım ücreti
Int	Telefonun uluslararası aramaya açık olup olmadığı bilgisi
Detay	Detaylı fatura alınıp alınmadığı bilgisi
Yas	Abonenin yaşı
Gsm	Telefonun cep telefonlarını aramaya açık olup olmadığı bilgisi
Na	Telefonun şehirlerarası aramaya açık olup olmadığı bilgisi
Out	Telefonun dış aramaya açık olup olmadığı bilgisi
Arizali_Gun	Telefonun arızalı olduğu gün sayısı

Tablo 3.1’ de ‘N’ 1., 2., 3., 4., 5. ve 6. ayları temsil etmekte olup, sonunda N bulunan değişkenlerden ilgili ayların hepsi için olmak üzere toplam 6’şar adet mevcuttur.

Değişkenler üzerinde gerekli incelemeyi yapmak için, Spss Clementine programının “Data Audit” nodunun “Quality” tabı kullanılmıştır. Data Audit noduna ait ekran görüntüsünün yer aldığı Şekil 3.4 incelendiğinde, Tenure, Churn\_Sebep ve Yas değişkenlerinin eksik değerler içerdiği görülmektedir. Churn\_Sebep değişkeni, abonelerin işletmeden ayrılma sebebini tuttuğundan tüm abonelerde bu alanın dolu olması beklenemez. Tenure değişkeni %79,5 dolu iken, Yas değişkeni %95,8 oranında doludur. Analizin yapılabilmesi için yeni yapay kayıtlar eklenerek, bu değerlerin tamamlanması ya da eksik değerlerin bulunduğu kayıtların silinmesi yoluna gidilmelidir.

Şekil 3.4 Data Audit Nodu



Field	Valid Records	% Complete
Ev_Is	122872	100
IL	122872	100
Ilce	122872	100
Tenure	97677	79,495
Tarife	122872	100
Churn	122872	100
Churn_Sebep	26703	21,732
Zamaninda_Odeme_Say	122872	100
Gecikmeli_Odeme_Flag	122872	100
Gecikmeli_Odeme_Say	122872	100
Yas	117748	95,83
Kullanim_Miktari_LOC_1	122872	100
Kullanim_Miktari_GSM_1	122872	100
Kullanim_Miktari_INT_1	122872	100
Kullanim_Miktari_NAT_1	122872	100
Kullanim_Miktari_TOT_1	122872	100
Kullanim_Miktari_LOC_2	122872	100
Kullanim_Miktari_GSM_2	122872	100
Kullanim_Miktari_INT_2	122872	100
Kullanim_Miktari_NAT_2	122872	100

Veriler incelenirken, Manisa ili için İlçe değişkeninde ‘BORLU’ ve İzmir ili için İlçe değişkeninde ‘ÇİĞLİ.’ değerlerine rastlanmıştır. Bunlar hatalı olduğundan, bu değerlere sahip kayıtların silinmesi gerekmektedir.

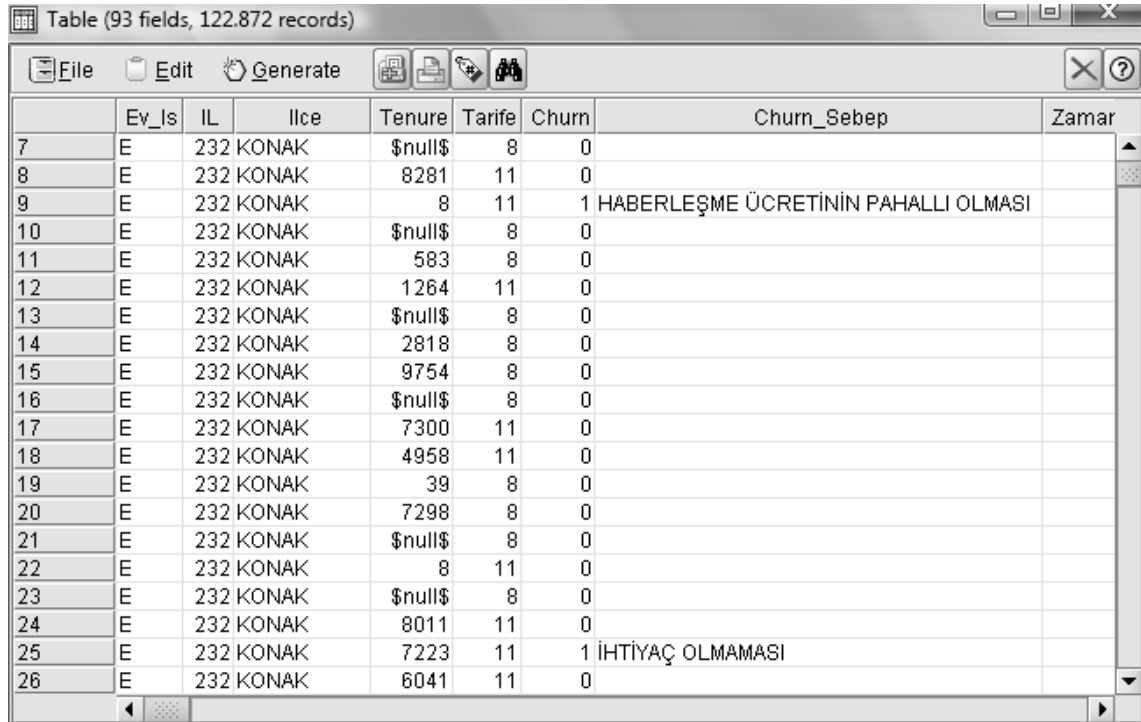
Veri tabanında tutulan, Ev\_Is değişkenine gerek olmadığı tespit edilmiştir. Yapılacak olan uygulama, tüzel olmayan gerçek abonelerden ev abonelerinin kayıtları ile gerçekleştirileceğinden, Ev\_Is değişkeni sürekli evi temsil eden ‘E’ değerini tutacaktır.

### 3.3. Veriyi Hazırlama

Veriyi hazırlama aşamasında, önceki aşamada tespit edilen aksaklıklar düzeltilmeli, veri modelleme için uygun hale getirilmelidir. Hatalı, eksik ya da tutarsız veri seti ile geliştirilen model doğru olmayacaktır.

Verinin analize hazırlanmadan önceki durumu Tablo 3.2’ de görülmektedir.

**Tablo 3.2 Verinin Düzenlenmeden Önceki Hali**



	Ev_Is	IL	İlçe	Tenure	Tarife	Churn	Churn_Sebep	Zamar
7	E	232	KONAK	\$null\$	8	0		
8	E	232	KONAK	8281	11	0		
9	E	232	KONAK	8	11	1	HABERLEŞME ÜCRETİNİN PAHALLI OLMASI	
10	E	232	KONAK	\$null\$	8	0		
11	E	232	KONAK	583	8	0		
12	E	232	KONAK	1264	11	0		
13	E	232	KONAK	\$null\$	8	0		
14	E	232	KONAK	2818	8	0		
15	E	232	KONAK	9754	8	0		
16	E	232	KONAK	\$null\$	8	0		
17	E	232	KONAK	7300	11	0		
18	E	232	KONAK	4958	11	0		
19	E	232	KONAK	39	8	0		
20	E	232	KONAK	7298	8	0		
21	E	232	KONAK	\$null\$	8	0		
22	E	232	KONAK	8	11	0		
23	E	232	KONAK	\$null\$	8	0		
24	E	232	KONAK	8011	11	0		
25	E	232	KONAK	7223	11	1	İHTİYAÇ OLMAMASI	
26	E	232	KONAK	6041	11	0		

Veriyi anlama aşamasında tespit edilen Yas değişkeninde eksik değerlerin bulunduğu kayıtlar uygulama veri setinden çıkarılmıştır. Ayrıca uç nokta olarak görülen 18'in altındaki ve 90'in üzerindeki değerlere sahip kayıtlar da veri setine alınmamıştır.

Tenure değişkeninin veriyi anlama aşamasında eksik değerlere sahip olduğu anlaşılmıştı. Ayrıca, müşterinin Tenure değeri bilinmiyorsa işletme tarafından ilgili tabloda o alana 20.000' in üzerinde herhangi bir rakam girildiği öğrenildi. Bunun için Tenure değişkeninde 20.000' in üzerinde bir değere ya da boş değere sahip kayıtların silinmesi yoluna gidildi. Ayrıca Temmuz ve Ağustos döneminde gelen müşterilerin modellemeye girmemesi için Tenure değeri 62'den küçük olan kayıtlar da silindi.

Tenure değişkeni, daha önce de belirtildiği gibi, abone ile işletme arasındaki ilişkinin süresini göstermektedir. Model geliştirilirken kolaylık olması için Tenure değişkeninden TNR isimli yeni bir değişken yaratılmıştır. TNR değişkeni Tenure' nin alacağı değerlere göre abonenin kaç aydır işletmenin abonesi olduğunu göstermekte ve 0 - 6 arasında bir değer almaktadır. Aşağıda TNR' nin alacağı değerler ve bu değerleri alabilmesi için gereken koşullar birlikte verilmiştir:

		<b><u>TNR</u></b>
Tenure < 31	⇒	0
31 ≤ Tenure < 62	⇒	1
62 ≤ Tenure < 92	⇒	2
92 ≤ Tenure < 123	⇒	3
123 ≤ Tenure < 153	⇒	4
153 ≤ Tenure < 184	⇒	5
184 ≤ Tenure	⇒	6

TNR değişkeninden TNR1, TNR2, TNR3, TNR4, TNR5 ve TNR6 değişkenleri yaratılmıştır. İlgili ayda kişinin aboneliği varsa 1, tersi durumunda ise 0 değerini almaktadırlar. Bu değişkenlerin, modelleme aşamasında abonelerin aktif ve deaktif olmalarını sınamak için kullanılması düşünülmektedir.

Aşağıda hangi durumlarda 0, hangi durumlarda 1 değerini aldıkları görülmektedir:

	<u>TNR1</u>	<u>TNR2</u>	<u>TNR3</u>	<u>TNR4</u>	<u>TNR5</u>	<u>TNR6</u>
TNR >= 6	1					
TNR < 6	0					
TNR >= 5		1				
TNR < 5		0				
TNR >= 4			1			
TNR < 4			0			
TNR >= 3				1		
TNR < 3				0		
TNR >= 2					1	
TNR < 2					0	
TNR >= 1						1
TNR < 1						0

Yeni oluşturulan TNR\_W değişkeni ise, TNR1, TNR2, TNR3, TNR4, TNR5 ve TNR6 değişkenlerinin ağırlıklı toplamlarından oluşmakta ve ileride ağırlıklı ortalama hesaplanırken kullanılması planlanmaktadır.

$$\text{TNR\_W} = \text{TNR1} * 1 + \text{TNR2} * 2 + \text{TNR3} * 3 + \text{TNR4} * 4 + \text{TNR5} * 5 + \text{TNR6} * 6$$

Abonelerin faturalarının zamanında ödenme sayısını tutan Zamanında\_Odeme\_Say, gecikmeli ödeme mevcutsa 1 değerini alan, tersi durumda 0 olan Gecikmeli\_Odeme\_Flag ve faturaların gecikmeli ödenme sayısını tutan Gecikmeli\_Odeme\_Say değişkenleri modelden çıkarılıp bunların yerine, gecikmeli ödeme sayısının, zamanında ödeme sayısına oranı olan Gecikmeli\_Zamanında\_Orani değişkeni yaratılmıştır.

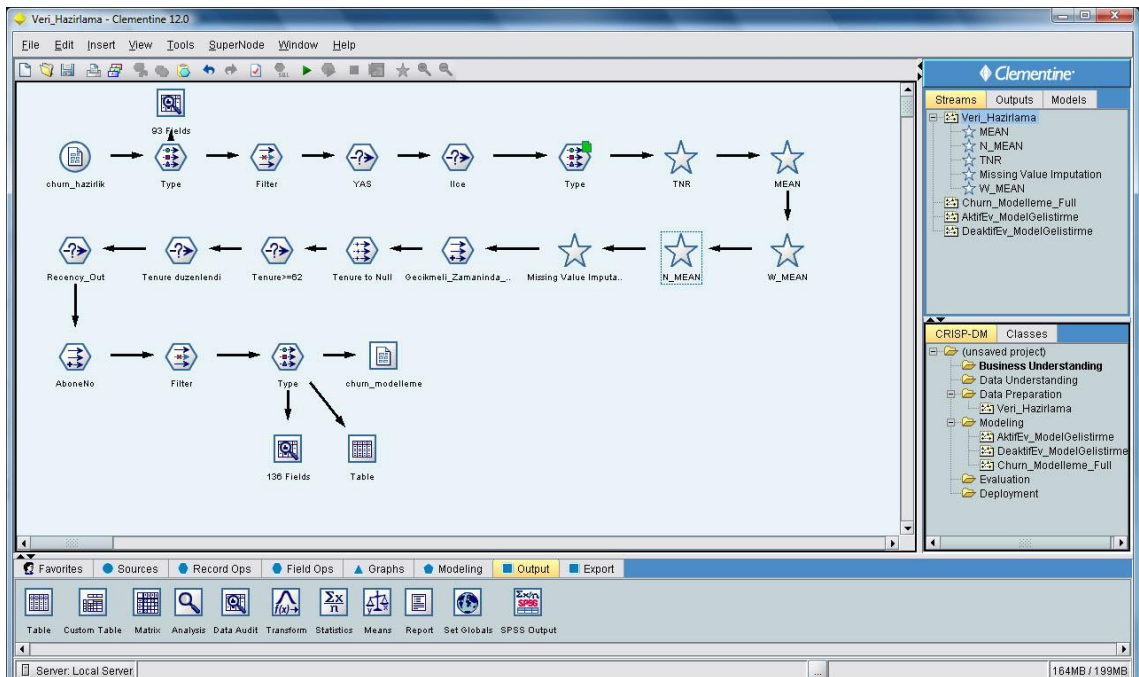
Şehir içi, şehirlerarası, uluslararası ve Gsm için ayrı ayrı olmak üzere abone olunan aylardaki kullanım miktarları ve arama sayıları ayrıca yine abone olunan aylardaki arama yapılmayan gün sayıları ve kullanım ücretleri ile ilgili değişkenler

modellemeye alınmayıp bunların yerine, ortalamaları, ağırlıklı ortalamaları ve ağırlıklı ortalamanın ortalamaya bölünmesiyle elde edilen normalize edilmiş ortalamaları oluşturulup modellemeye alınmıştır. Örneğin, abonenin ilgili aydaki şehir içi kullanım miktarını tutan Kullanim\_Miktari\_LOC\_1, Kullanim\_Miktari\_LOC\_2, Kullanim\_Miktari\_LOC\_3, Kullanim\_Miktari\_LOC\_4, Kullanim\_Miktari\_LOC\_5 ve Kullanim\_Miktari\_LOC\_6 yerine bunların ortalaması olan MEAN\_Kullanim\_Miktari\_LOC, ağırlıklı ortalaması olan W\_MEAN\_Kullanim\_Miktari\_LOC ve W\_MEAN\_Kullanim\_Miktari\_LOC'un MEAN\_Kullanim\_Miktari\_LOC'a bölünmesiyle oluşturulan N\_MEAN\_Kullanim\_Miktari\_LOC alanı yaratılmıştır.

Veriyi anlama basamağında modellemeden çıkarılması tespit edilen Ev\_Is değişkeni çıkarılmıştır. Ayrıca gizlilik gereği, müşteri numaralarını tutan MusteriKey, abone numaralarını tutan AboneKey ve telefon numaralarını tutan HizmetNo modellemeye alınmamış, her bir aboneye, o abone için tek olan ve 1'den başlayıp 1'er 1'er artan AboneNo verilmiştir.

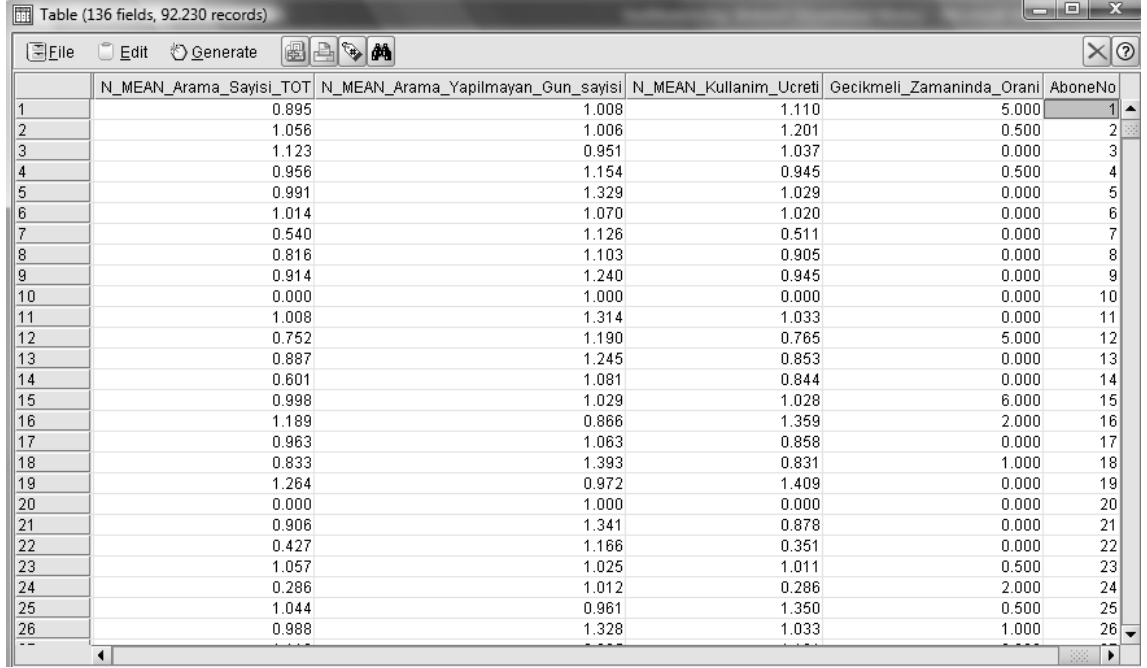
Şekil 3.5' te Spss Clementine'de verinin hazırlanması aşaması görülmektedir.

**Şekil 3.5 Veri Hazırlama Aşaması**



Verinin analize hazırlandıktan sonraki durumu Tablo 3.3’ te görülmektedir.

**Tablo 3.3 Verinin Düzenlendikten Sonraki Hali**



	N_MEAN_Arama_Sayisi_TOT	N_MEAN_Arama_Yapilmayan_Gun_sayisi	N_MEAN_Kullanim_Ucreti	Gecikmeli_Zamaninda_Orani	AboneNo
1	0.895	1.008	1.110	5.000	1
2	1.056	1.006	1.201	0.500	2
3	1.123	0.951	1.037	0.000	3
4	0.956	1.154	0.945	0.500	4
5	0.991	1.329	1.029	0.000	5
6	1.014	1.070	1.020	0.000	6
7	0.540	1.126	0.511	0.000	7
8	0.816	1.103	0.905	0.000	8
9	0.914	1.240	0.945	0.000	9
10	0.000	1.000	0.000	0.000	10
11	1.008	1.314	1.033	0.000	11
12	0.752	1.190	0.765	5.000	12
13	0.887	1.245	0.853	0.000	13
14	0.601	1.081	0.844	0.000	14
15	0.998	1.029	1.028	6.000	15
16	1.189	0.866	1.359	2.000	16
17	0.963	1.063	0.858	0.000	17
18	0.833	1.393	0.831	1.000	18
19	1.264	0.972	1.409	0.000	19
20	0.000	1.000	0.000	0.000	20
21	0.906	1.341	0.878	0.000	21
22	0.427	1.166	0.351	0.000	22
23	1.057	1.025	1.011	0.500	23
24	0.286	1.012	0.286	2.000	24
25	1.044	0.961	1.350	0.500	25
26	0.988	1.328	1.033	1.000	26

Bu çalışmada kullanılmak üzere uygulamaya konu olan firmadan alınan orijinal veri tablosunun formatı EK-1’de gösterilmiştir.

### 3.4. Modelleme

Tanımlanan problem için en uygun modelin bulunabilmesi olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir ([www.bilgiyonetimi.org](http://www.bilgiyonetimi.org)).

Bu tez çalışmasında, aboneliği boyunca en az bir arama yapmış kişiler aktif, hiç arama yapmamış kişiler deaktif kabul edilmiş ve iki grup için ayrı ayrı model geliştirilmiştir.

Aboneleri aktif ya da deaktif olarak ayırabilmek için aşağıdaki koşullara bakılmıştır:

Aktif abone olma koşulu =>  $(Kullanim\_Ucreti\_1 * TNR1 + Kullanim\_Ucreti\_2 * TNR2 + Kullanim\_Ucreti\_3 * TNR3 + Kullanim\_Ucreti\_4 * TNR4 + Kullanim\_Ucreti\_5 * TNR5 + Kullanim\_Ucreti\_6 * TNR6) > 0$

Deaktif abone olma koşulu =>  $(Kullanim\_Ucreti\_1 * TNR1 + Kullanim\_Ucreti\_2 * TNR2 + Kullanim\_Ucreti\_3 * TNR3 + Kullanim\_Ucreti\_4 * TNR4 + Kullanim\_Ucreti\_5 * TNR5 + Kullanim\_Ucreti\_6 * TNR6) = 0$

TNRN abonenin ilgili ayda işletmede aboneliği olup olmamasına göre 0 ya da 1 değerlerini tutmaktadır. 3. Ayda işletmenin abonesi olmuş bir kişi için TNR1 ve TNR2 0 olacağından, koşula 1. Ay ve 2. Ay dâhil olmayacaktır. Abone, telefonunu kullanmışsa faturasına kullanım ücreti yansıyacak ve formülde çıkan sonuç 0'dan büyük olacağından aktif abone koşulunu sağlayacaktır.



### 3.4.1. Aktif Ev Aboneleri İçin Modelleme

Aktif ev aboneleri için model geliştirilirken Lojistik Regresyon Analizi kullanılmıştır. Uygulama için hazırlanan veri setinde bağımlı değişken (churn) abonenin işletmeden ayrılması durumunda 1, işletmede kalması durumunda 0 değerini alan iki kategorili bir değişkendir. İki kategorili verilerin gösterimi için daha basit ve uygun olduğundan, uygulamada Lojistik Regresyon Analizi tercih edilmiştir.

Lojistik Regresyon Analizi birinci bölümde ayrıntılı olarak açıklanmıştır. Bu sebeple bu bölümde sadece analiz sonuçları üzerinde durulacaktır.

Modellemeye başlamadan önce aktif ev abonelerinin işletmeden ayrılma ve işletmede kalmaya devam etme durumunu gösteren dağılım incelenmiştir. Şekil 3.6' da bu dağılım görülmektedir.

Şekil 3.6 Aktif Ev Abonelerinin Dağılımı - 1

Deger	Oran	%	Sayı
0		78,22	69438
1		21,78	19330



Aktif ev abonelerinin %78,2' sinin işletmede kalmaya devam ettiği, %21,8' inin ise işletmeden ayrıldığı görülmektedir. Geliştirilecek olan modelin bu dağılımdan etkilenmemesi ve eğilimli sonuçlar vermemesi için, bu oranın eşitlenmesi gerekmektedir. Model tüm müşteriler için işletmede kalacakları yönünde tahmin yapsa da, %78,2 oranında doğru tahminde bulunmuş olacaktır. Bunu engellemek için, Spss Clementine programında bulunan 'distribution' nodu ile oran dengelenmiştir.

Şekil 3.7' de görüldüğü gibi modele girmeden önce abonelerin %50,2' si işletmede kalmaya devam etmiş, %49,8' i işletmeyi terk etmiştir.

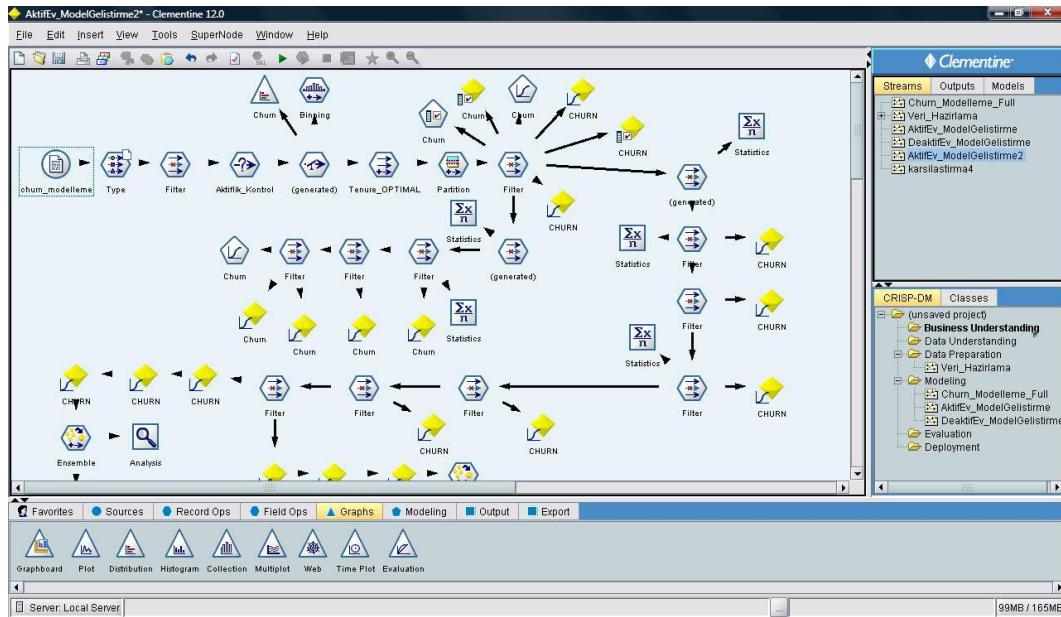
**Şekil 3.7 Aktif Ev Abonelerinin Dağılımı - 2**

Deger	Oran	%	Sayı
0		50,19	19479
1		49,81	19330

Veri setindeki farklı değişkenler ile farklı modeller denenmiş ve doğruluğu en yüksek model seçilmiştir.

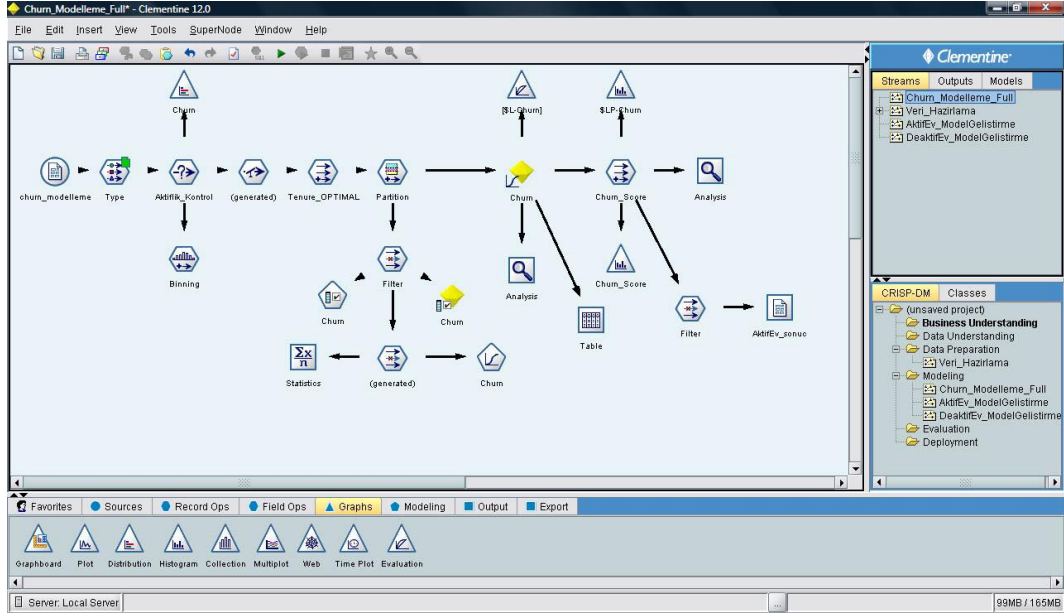
Şekil 3.8' de aktif ev aboneleri için Spss Clementine' de denenilen farklı modeller görülmektedir.

**Şekil 3.8 Aktif Ev Aboneleri İçin Model Geliştirme -1**



Şekil 3.9’ daki Spss Clementine akımında (stream), denenen farklı modellerin içinden en iyi olduğu düşünülen model ile lojistik regresyon analizi yapılmıştır.

Şekil 3.9 Aktif Ev Aboneleri İçin Model Geliştirme -2



Spss Clementine programındaki ‘Statistics’ nodu ile değişkenlerin arasındaki korelasyona bakılmış ve bağımsız değişkenler arasında yüksek korelasyon varsa, hedef değişken ile daha az ilişkisi olan değişken modellemeye alınmamıştır. Diğer bir ifadeyle, bağımlı değişken üzerinde etkili olan ancak kendi aralarında ilişkili olmayan değişkenler tercih edilmiştir.

Bir regresyon modelini başarılı kabul edebilmek için, modelde yer alan tüm bağımsız değişkenlerin anlamlı bulunması gerekmektedir. Modelde anlamlılık seviyesi 0,05 kabul edildiğinden, bu oranın üzerindeki anlamlılık seviyesine sahip değişkenler modelden çıkarılmış ve model tekrar denenmiştir.

Tablo 3.4’ te denenen bir modelin B parametreleri ile bu parametrelere ilişkin Wald istatistikleri, serbestlik dereceleri, önem seviyeleri ve Odds değerleri verilmiştir.

Tablo 3.4 Lojistik Regresyon Sonuçları - 1

Churn(a)	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-4,557	,137	1108,849	1	,000			
Recency_Out	,007	,001	64,292	1	,000	1,007	1,005	1,008
MEAN_Kullanım_Miktari_NAT	-,001	,001	2,201	1	,138	,999	,998	1,000
MEAN_Arama_Sayisi_LOC	-,002	,000	22,582	1	,000	,998	,998	,999
MEAN_Arama_Sayisi_NAT	,001	,001	,405	1	,525	1,001	,999	1,003
MEAN_Arama_Sayisi_GSM	-,001	,001	2,361	1	,124	,999	,997	1,000
W_MEAN_Arama_Yapılmayan_Gun_sayisi	,031	,002	173,182	1	,000	1,032	1,027	1,037
N_MEAN_Arama_Yapılmayan_Gun_sayisi	,118	,069	2,908	1	,088	1,126	,982	1,290
N_MEAN_Kullanım_Ucreti	,105	,059	3,148	1	,076	1,111	,989	1,248
Gecikmeli_Zamanında_Orani	,081	,005	216,616	1	,000	1,084	1,072	1,096
[Oto_Odeme=0]	,023	,034	,467	1	,494	1,023	,958	1,093
[Oto_Odeme=1]	0(b)	.	.	0	.	.	.	.
[Adsl=0]	1,473	,038	1493,811	1	,000	4,362	4,048	4,701
[Adsl=1]	0(b)	.	.	0	.	.	.	.
[Int=0]	,557	,031	322,043	1	,000	1,745	1,642	1,854
[Int=1]	0(b)	.	.	0	.	.	.	.
[Detay=0]	1,229	,029	1793,141	1	,000	3,419	3,230	3,619
[Detay=1]	0(b)	.	.	0	.	.	.	.
[Gsm=0]	-,817	,049	275,506	1	,000	,442	,401	,487
[Gsm=1]	0(b)	.	.	0	.	.	.	.
[Out=0]	-,766	,186	16,875	1	,000	,465	,323	,670
[Out=1]	0(b)	.	.	0	.	.	.	.
[Tenure_OPTIMAL=1]	2,390	,109	481,568	1	,000	10,911	8,814	13,507
[Tenure_OPTIMAL=2]	2,055	,058	1274,868	1	,000	7,805	6,973	8,737
[Tenure_OPTIMAL=3]	1,935	,052	1381,241	1	,000	6,923	6,252	7,667
[Tenure_OPTIMAL=4]	1,210	,050	583,836	1	,000	3,352	3,039	3,698
[Tenure_OPTIMAL=5]	,610	,053	133,673	1	,000	1,840	1,659	2,040
[Tenure_OPTIMAL=6]	,354	,061	34,271	1	,000	1,425	1,266	1,605
[Tenure_OPTIMAL=7]	0(b)	.	.	0	.	.	.	.

Odds değeri (Exp(B)), müşterilerin işletmede kalma olasılığının, işletmeden ayrılma olasılığına oranını temsil etmektedir.

Wald istatistiği, doğrusal regresyondaki t testinin lojistik regresyondaki karşılığıdır. Lojistik regresyon denklemi ile tahmin edilen  $B_i$  katsayılarının sifıra eşit olduğu biçimindeki hipotezleri test etmektedir. Modelde Wald testinden geçemeyen bir değişken varsa, bulunan katsayının tesadüfen ortaya çıktığı, gerçekte bu değişken ile bağımlı değişken arasında bir ilişki söz edilemeyeceği anlaşılır. Wald istatistiği,

- $H_0 = B_i = 0$
- $H_1 = B_i \neq 0$  şeklindeki iki hipotezi test eder.

Tabloda görülen anlamlılık seviyeleri  $H_0$  hipotezinin reddedilmesi durumunda yapılacak hata miktarlarını göstermektedir (www.tajev.org). Örneğin,

N\_Mean\_Kullanim\_Ucreti deęişkeninin anlamlılık seviyesine bakıldığında 0,076 olduęu görülmektedir, dolayısıyla  $H_0$  hipotezi reddedilirse yapılacak hata miktarı, hata toleransı olarak belirlenen 0,05' in üstünde olduęundan kabul edilebilir bir hata miktarı olmayacaktır.

Önem seviyeleri 0,05' ten büyük olan deęişkenler çıkarılıp model tekrar çalıştırılmış ve Tablo 3.5' teki gibi sonuçlar elde edilmiştir.

**Tablo 3.5 Lojistik Regresyon Sonuçları - 2**

Churn(a)	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-4,366	,100	1918,808	1	,000			
Gecikmeli_Zamaninda_Orani	,082	,005	265,281	1	,000	1,085	1,075	1,096
MEAN_Arama_Sayisi_LOC	-,002	,000	27,811	1	,000	,998	,997	,999
N_MEAN_Arama_Yapilmayan_Gun_sayisi	,105	,063	2,786	1	,095	1,111	,982	1,257
Recency_Out	,005	,001	39,886	1	,000	1,005	1,003	1,006
W_MEAN_Arama_Yapilmayan_Gun_sayisi	,035	,002	232,740	1	,000	1,035	1,031	1,040
[Adsl=0]	1,478	,038	1517,983	1	,000	4,384	4,070	4,722
[Adsl=1]	0(b)	.	0	.	.	.	.	.
[Detay=0]	1,194	,029	1710,184	1	,000	3,300	3,118	3,492
[Detay=1]	0(b)	.	0	.	.	.	.	.
[Gsm=0]	-,848	,049	305,429	1	,000	,428	,390	,471
[Gsm=1]	0(b)	.	0	.	.	.	.	.
[Int=0]	,555	,031	321,619	1	,000	1,743	1,640	1,852
[Int=1]	0(b)	.	0	.	.	.	.	.
[Out=0]	-,704	,182	14,888	1	,000	,495	,346	,707
[Out=1]	0(b)	.	0	.	.	.	.	.
[Tenure_OPTIMAL=1]	2,443	,098	615,913	1	,000	11,506	9,487	13,954
[Tenure_OPTIMAL=2]	1,936	,057	1142,884	1	,000	6,931	6,195	7,754
[Tenure_OPTIMAL=3]	1,871	,052	1280,499	1	,000	6,495	5,862	7,196
[Tenure_OPTIMAL=4]	1,154	,050	525,774	1	,000	3,172	2,874	3,500
[Tenure_OPTIMAL=5]	,550	,053	108,043	1	,000	1,733	1,563	1,923
[Tenure_OPTIMAL=6]	,281	,061	21,562	1	,000	1,325	1,177	1,492
[Tenure_OPTIMAL=7]	0(b)	.	0	.	.	.	.	.

N\_MEAN\_Arama\_Yapilmayan\_Gun\_sayisi deęişkeni de anlamlı bulunmadığından modelden çıkarılmış ve elde edilen nihai model Tablo 3.6' da gösterilmiştir.

**Tablo 3.6 Lojistik Regresyon Sonuçları - 3**

Churn(a)	B	Std. Error	Wald	df	Sig.	Exp(B)	95,0% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-4,455	,077	3309,956	1	,000			
Recency_Out	,005	,001	51,308	1	,000	1,005	1,004	1,007
MEAN_Arama_Sayisi_LOC	-,003	,000	60,322	1	,000	,997	,997	,998
W_MEAN_Arama_Yapilmayan_Gun_sayisi	,030	,002	178,962	1	,000	1,031	1,026	1,035
Gecikmeli_Zamaninda_Orani	,078	,005	247,770	1	,000	1,082	1,071	1,092
[Adsl=0]	1,468	,038	1503,441	1	,000	4,340	4,029	4,674
[Adsl=1]	0(b)	.	.	0	.	.	.	.
[Int=0]	,778	,029	744,387	1	,000	2,177	2,059	2,303
[Int=1]	0(b)	.	.	0	.	.	.	.
[Detay=0]	1,185	,028	1745,339	1	,000	3,270	3,094	3,457
[Detay=1]	0(b)	.	.	0	.	.	.	.
[Tenure_OPTIMAL=1]	2,594	,091	818,716	1	,000	13,388	11,208	15,991
[Tenure_OPTIMAL=2]	2,085	,057	1337,213	1	,000	8,044	7,193	8,995
[Tenure_OPTIMAL=3]	1,962	,052	1442,273	1	,000	7,110	6,426	7,868
[Tenure_OPTIMAL=4]	1,186	,050	573,067	1	,000	3,274	2,971	3,608
[Tenure_OPTIMAL=5]	,630	,052	145,089	1	,000	1,877	1,695	2,080
[Tenure_OPTIMAL=6]	,357	,060	35,367	1	,000	1,429	1,270	1,608
[Tenure_OPTIMAL=7]	0(b)	.	.	0	.	.	.	.

Tablo 3.6’ da görüldüğü gibi, modele alınan tüm değişkenler anlamlı bulunmuştur. Exp (B), bağımlı değişkenin, bağımsız değişkeninin etkisi ile kaç kat daha fazla ya da % kaç oranında fazla gözlenme olasılığına sahip olduğunu belirttiğine, bir başka ifadeyle Odds oranı; incelenen iki olayın gözlenme olasılıklarından birinin diğerine oranla kaç kat daha fazla veya kaç kat daha az olarak ortaya çıkabileceğini gösterdiğine göre (Girginer ve Cankuş, 2008, s.190), tablo incelendiğinde aşağıdaki sonuçlara varılabilir:

- Recency\_Out değişkenindeki 1 birimlik artış 1,005,
- W\_Mean\_Arama\_Yapilmayan\_Gun\_Sayisi değişkenindeki 1 birimlik artış 1,031 ve
- Gecikmeli\_Zamaninda\_Oranında değişkenindeki 1 birimlik artış 1,082 birim abonelerin işletmeyi terk etme olasılığını arttıracaktır.
- Mean\_Arama\_Sayisi\_Loc değişkenindeki 1 birimlik artış 0,997 birim abonelerin işletmeyi terk etme olasılığını azaltacaktır.

Ayrıca;

- Adsl’ i olmayanların, Adsl’ i olanlara göre 4,34,



- Telefonu uluslar arası aramalara kapalı olanların, açık olanlara göre 2,177,
- Detaylı fatura almayanların, alanlara göre 3,27 ve
- Abonelik süresi 135 günün altında olanların, üzerinde olanlara göre 13,388 birim daha fazla işletmeyi terk etme riskine sahip oldukları görülmektedir.

### 3.4.2. Deaktif Ev Aboneleri İçin Modelleme

6 aylık dönem içinde, abone olduğu günden modelin geliştirildiği güne kadar ev telefonunu hiç kullanmamış aboneler deaktif abone olarak nitelendirilmektedir. Örneğin, yazlığına telefon bağlatıp, yaz dönemi dışında telefonu kullanmayan abone ya da internet bağlatabilmek için önce telefon hattı çektirmek zorunda kalan ama telefon cihazı almayan abone deaktif abonedir. Modelde, bu abonelere ait kullanım ücreti, kullanım miktarı, arama sayısı, bunların ortalaması, ağırlıklı ortalaması ya da normalize edilmiş ortalaması gibi değişkenler söz konusu değildir. Deaktif aboneler için, karar ağacı tekniği uygun görülmüş, karar ağacı algoritmalarından Quest Karar Ağacı algoritması seçilmiştir.



Şekil 3.10' da deaktif ev abonelerinin işletmede kalma ve işletmeyi terk etme durumlarını gösteren dağılım görülmektedir.

**Şekil 3.10 Deaktif Abonelerin Dağılımları -1**

Deger ▲	Oran	%	Sayı
0		55,32	1908
1		44,68	1541

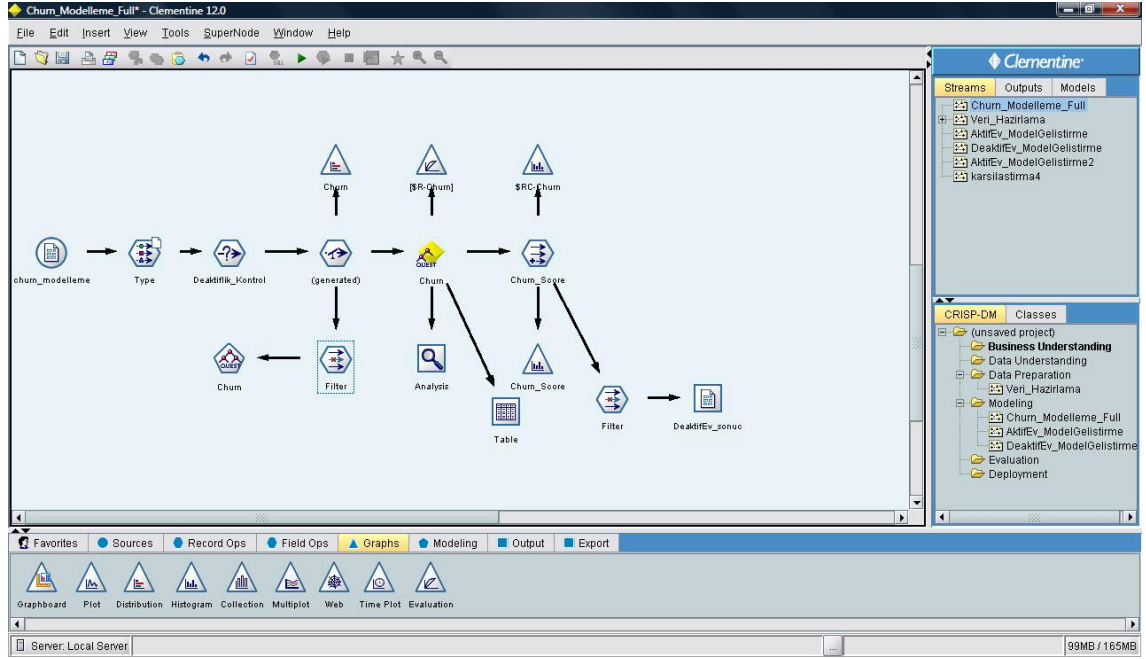
Model geliştirilmeden önce işletmede kalan ve işletmeden ayrılan oranlarını dengelemek için, Spss Clementine'nin 'distrubution' nodu kullanılmış ve Şekil 3.11' de görüldüğü gibi iki oran birbirine eşitlenmiştir.

**Şekil 3.11 Deaktif Abonelerin Dağılımları -2**

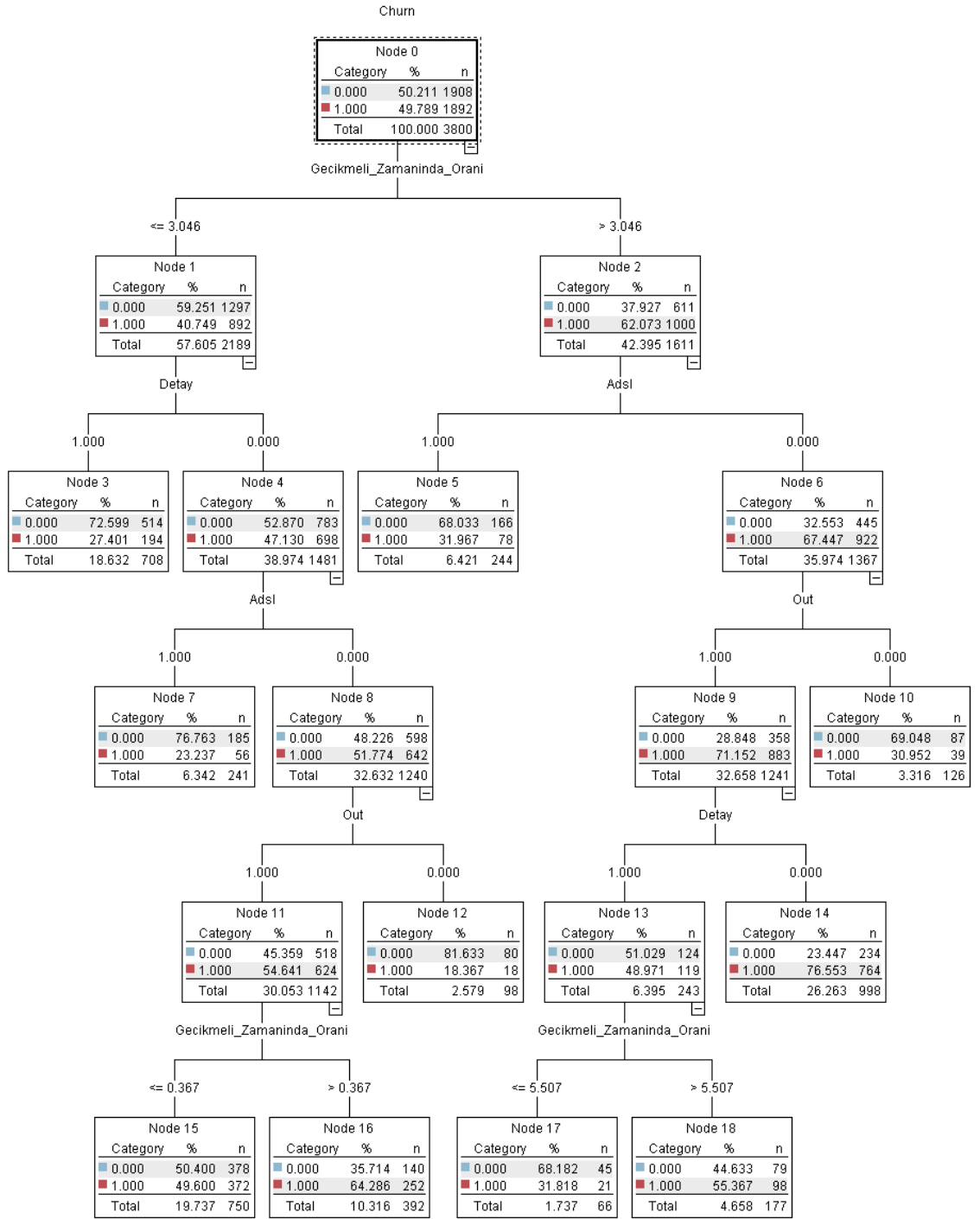
Deger ▲	Oran	%	Sayı
0		50,0	1908
1		50,0	1908

En uygun modeli oluşturabilmek için farklı değişkenler kullanılarak çok sayıda model denenmiş ve en iyisi olduğu düşünülen model ile karar ağacı oluşturulmuştur. Şekil 3.12’ de, karar ağacını oluşturmak için yapılmış Spss Clementine akımı (stream) görülmektedir.

**Şekil 3.12 Deaktif Ev Aboneleri İçin Geliştirilen Spss Clementine Akımı**



Şekil 3.13 Deaktif Aboneler İçin Quest Karar Ağacı



Şekil 3.13' teki karar ağacı incelendiğinde, abonelerin %49,79 'unun işletmeden ayrıldığı, % 50,21' inin ise işletmede kalmaya devam ettiği görülmektedir. İşletmede



kalmaya devam eden abone oranı, 3800 abonenin 1908' ine karşılık gelmektedir. İşletmeyi terk eden abone sayısı ise 1892' dir.

Abonenin gecikmeli ödenen fatura sayısının, zamanında ödenen fatura sayısına oranı arttıkça işletmeyi terk etme eğiliminde olduğu görülmektedir. Bu oranı temsil eden Gecikmeli\_Zamanında\_Orani değişkeninin ayırıcı değeri 3,046' dır.

Gecikmeli Zamanında Orani 3,046' ya küçük ya da eşit olduğunda;

Detaylı faturası olan 708 abonenin %72,6' sının işletmede kaldığı, %27,4' ünün ise işletmeyi terk ettiği, detaylı faturası yoksa, bu koşulları sağlayan 1481 abonenin %52,9' unun işletmede kaldığı, 47,1' inin ise işletmeyi terk ettiği gözlenmektedir. Bu durumda detaylı faturaya sahip abonelerin kalma eğilimi taşıdığı söylenebilir.

Detaylı faturası olmadığı halde, Adsl' i olan abonelerin %76,8' inin kaldığı, %23,2' sinin ayrıldığı, Adsl' i olmayan abonelerin ise % 48,2' sinin kaldığı, %51,8' inin ayrıldığı görülmektedir. Bu durumda Adsl sahipliğinin işletmede kalma yönünde etkisinden söz edilebilir.

Abonenin detaylı faturası ve Adsl'i yokken telefonu dış aramaya açıksa, bu koşulları sağlayan 1142 abonenin %45,36' sının işletmede kaldığı, %54,64' ünün işletmeden ayrıldığı, arama yapamayan abonelerin ise %81,63' ünün kaldığı, %18,37' sinin işletmeden ayrıldığı görülmektedir. Bu durumda telefonu dış aramaya açık olan abonelerin işletmeyi terk etme eğiliminde olduğu söylenebilir.

Detaylı faturası ve Adsl'i olmayan, telefonu dış aramalara açık olan abonelerin, gecikmeli ödedikleri fatura sayılarının zamanında ödedikleri fatura sayısına oranı 0,367' den küçük ya da eşitse, bu abonelerin %50,4' ünün işletmede kaldığı, %49,6' sının işletmeden ayrıldığı, bu oran 0,367'den büyükse, abonelerin %35,7' sinin işletmede kaldığı, %64,3' ünün ise işletmeyi terk ettiği görülmektedir.

Gecikmeli Zamanında Oranı 3.046'nın üzerinde olduğunda;

Adsl' e sahip abonelerin %68' inin işletmede kaldığı, %32' sinin işletmeden ayrıldığı, Adsl'i olmayan 1367 abonenin ise, %32,6' sının işletmede kaldığı, %67,4' ünün işletmeden ayrıldığı görülmektedir.

Adsl' i olmayan ve arama yapabilen abonelerin %28,9' unun işletmede kaldığı, %71,1' inin işletmeyi terk ettiği, Adsl'i olmayan ama telefonu aramalara kapalı olan abonelerin %69' unun işletmede kaldığı, %31' inin işletmeyi terk ettiği görülmektedir.

Adsl'i olmayan ve telefonu dış aramalara açık olan abonelerin detaylı faturası varsa, bu abonelerin %51' inin işletmede kaldığı, %49' unun işletmeden ayrıldığı, detaylı faturası yoksa, bu abonelerin %23,4' ünün işletmede kaldığı, %76,6' sının işletmeden ayrıldığı görülmektedir.

Adsl'i olmayan, telefonu dış aramalara açık olan ve detaylı fatura alan abonelerin, gecikmeli ödedikleri fatura sayılarının zamanında ödedikleri fatura sayısına oranı 5,507'den küçük ya da eşitse, bu abonelerin %68,2' sinin işletmede kaldığı, %31,8' inin işletmeden ayrıldığı, bu oran 5,507'den büyükse, abonelerin %44,6' sının işletmede kaldığı, %55,4' ünün ise işletmeyi terk ettiği görülmektedir.

Anlatılanlar özet halinde aşağıda sunulmuştur:

```
Gecikmeli_Zamaninda_Orani <= 3,046 [Mode: 0] (2.189)
  Detay = 1 [Mode: 0] ⇒ 0 (708; 0,726)
  Detay = 0 [Mode: 0] (1.481)
    Adsl = 1 [Mode: 0] ⇒ 0 (241; 0,768)
    Adsl = 0 [Mode: 1] (1.240)
      Out = 1 [Mode: 1] (1.142)
        Gecikmeli_Zamaninda_Orani <= 0,367 [Mode: 0] ⇒ 0 (750; 0,504)
        Gecikmeli_Zamaninda_Orani > 0,367 [Mode: 1] ⇒ 1 (392; 0,643)
      Out = 0 [Mode: 0] ⇒ 0 (98; 0,816)
Gecikmeli_Zamaninda_Orani > 3,046 [Mode: 1] (1.611)
  Adsl = 1 [Mode: 0] ⇒ 0 (244; 0,68)
  Adsl = 0 [Mode: 1] (1.367)
    Out = 1 [Mode: 1] (1.241)
      Detay = 1 [Mode: 0] (243)
        Gecikmeli_Zamaninda_Orani <= 5,507 [Mode: 0] ⇒ 0 (66; 0,682)
        Gecikmeli_Zamaninda_Orani > 5,507 [Mode: 1] ⇒ 1 (177; 0,554)
      Detay = 0 [Mode: 1] ⇒ 1 (998; 0,766)
    Out = 0 [Mode: 0] ⇒ 0 (126; 0,69)
```

### 3.5. Değerlendirme

Değerlendirme basamağında amaç, yapılan veri madenciliği uygulamasının sonuçlarının incelenmesi ve yapılan işlemlerin gözden geçirilmesidir. Bunun için, modelin anlamlılığı, doğruluk oranı gibi konularda yeterli kaliteyi sağlayıp sağlamadığına bakılır.

Gerçekleştirilen iki model olduğu için değerlendirme aşaması da aktif ev ve deaktif ev olmak üzere iki başlıkta incelenmiştir.

#### 3.5.1. Aktif Ev Aboneleri İçin Değerlendirme

Aktif ev aboneleri için lojistik regresyon ile geliştirilen modeli değerlendirmek için aşağıdaki tablolar kullanılmıştır:

**Tablo 3.7 Anlamlılık Testi**

		Ki-Kare	df	Sig.
Adım 1	Adım	59,550	13	,000
	Blok	59,550	13	,000
	Model	59,550	13	,000

Tablo 3.7’ de görülen anlamlılık (sig.) değeri 0,05’ ten küçük olduğundan kurulan lojistik model anlamlı bulunmuştur.

**Tablo 3.8 Model Özeti**

Adım	-2 Log likelihood	Cox & Snell R Kare	Nagelkerke R Kare
1	26,146	,532	,723

Modele ait Nagelkerke  $R^2$  değeri 0,72 olarak bulunmuştur. Bu değer, kurulan lojistik modelin kullanılan değişkenlerce açıklanma oranı göstermektedir. Ayrıca Cox ve Snell  $R^2$  değeri 0,53 olarak bulunmuştur. Açıklama oranlarının yeterince yüksek ve modelin anlamlı olduğu, ayrıca dışarıda kalan değişkenlerin çok fazla önemli olmadığı, modele alınan değişkenlerin yeterli olduğu söylenebilir.

**Tablo 3.9 Sınıflandırma Oranları**

Gözlenen	Tahmin Edilen		
	0	1	Doğruluk Yüzdesi
0	13342	6105	68,6%
1	5339	13991	72,4%
<b>Genel Yüzde</b>	<b>48,2%</b>	<b>51,8%</b>	<b>70,5%</b>

Tablo 3.9’ da verildiği gibi Recency\_Out, Mean\_Arama\_Sayisi\_Loc, W\_Mean\_Arama\_Yapılmayan\_Gun\_Sayisi, Gecikmeli\_Zamaninda\_Oraninda, Adsl, Int, Detay ve Tenure\_Optimal değişkenleri ile kurulan modelin doğru sınıflandırma oranı %70,5 olarak bulunmuştur.

Geliştirilen model, her bir abonenin işletmeden ayrılıp ayrılmayacağını tahmin etmiş, bunun için işletmeden ayrılmaları durumunda 1, kalmaları durumunda 0 değerini tutan \$L-Churn isiminde bir değişken üretmiştir. Ayrıca abonenin ayrılıp ayrılmayacağını tahmin ederken bunu belirli olasılığa bağlamış ve bu olasılığı \$LP-Churn isimli değişkende tutmuştur. Örneğin;

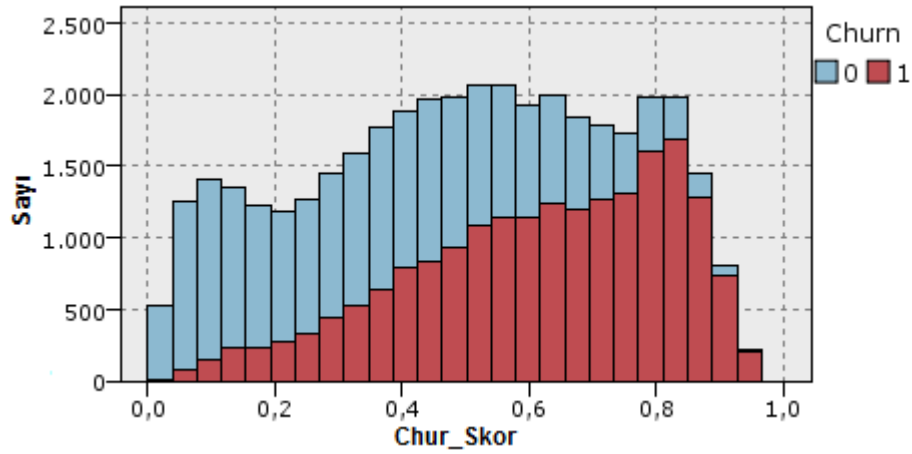
1. Abone için: \$L-Churn=0 \$LP-Churn=0,58

2. Abone için:  $\$L\text{-Churn}=1$   $\$LP\text{-Churn}=0,72$  olduğunu farz edilirse, bu durum 1. Abonenin 0,58 olasılıkla işletmede kalacağı, 2. Abonenin ise 0.72 olasılıkla işletmeden ayrılacağı yönünde yorumlanır.

Modelin aboneler için tahmin ettiği 0 ve 1 değerlerini, abonelerin gerçek davranışlarıyla kıyaslamak için Churn\_Skor isimli yeni bir değişken yaratılmıştır. Bu değişken tüm abonelerin işletmeden ayrılma olasılıklarını tutmakta olup, model abonenin işletmeden ayrılacağını tahmin ediyorsa  $\$LP\text{-Churn}$  değişkeninin değerini, abonenin işletmede kalacağını tahmin ediyorsa  $1-\$LP\text{-Churn}$  değerini almaktadır.

$$\begin{array}{l} \text{Churn_Skor} \\ \$L\text{-Churn}=1 \Rightarrow \$LP\text{-Churn} \\ \$L\text{-Churn}=0 \Rightarrow 1-\$LP\text{-Churn} \end{array}$$

Şekil 3.14 Olasılığa Göre İşletmeden Ayrılma Dağılımı



Şekil 3.14' te abonelerin modelin vermiş olduğu işletmeden ayrılma olasılığına göre dağılımı görülmektedir. Kırmızı renkli kısım işletmeden ayrılan aboneleri, mavi renkli kısım işletmede kalmaya devam eden aboneleri temsil etmektedir. Modelin verdiği olasılık arttıkça, kırmızı renkli kısım artmakta bu da işletmeden ayrılan sayısının modelin daha yüksek olasılıkla gideceğini öngördüğü kişilerde arttığı anlamını taşımaktadır.

**Tablo 3.10 İşletmeden Ayrılan Aktif Abonelerin Özellikleri - 1**

<b>Değişken</b>	<b>Değer</b>	<b>Kişi Sayısı</b>	<b>Oran</b>
<b>Adsl</b>	0: Adsl yok	18.168	94%
	1: Adsl var	1.162	6%
<b>Detay</b>	0: Detaylı fatura alınmıyor	15.744	81%
	1: Detaylı fatura alınıyor	3.586	19%
<b>Int</b>	0: Telefon uluslararası aramaya kapalı	15.492	80%
	1: Telefon uluslararası aramaya açık	3.838	20%
<b>Tenure_Optimal</b>	1: Abonelik Süresi<135 gün	7.900	41%
	2: 135 gün<=Abonelik Süresi<498 gün	4.702	24%
	3: 498 gün<=Abonelik Süresi<1795 gün	3.081	16%
	4: 1795 gün<=Abonelik Süresi<3607 gün	1.864	10%
	5: 3607 gün<=Abonelik Süresi<5499 gün	862	4%
	6: 5499 gün<=Abonelik Süresi<6710 gün	521	3%
	7: 6710 gün<=Abonelik Süresi	400	2%

Tablo 3.10’ da modelde kullanılan Adsl, Detay, Int ve Tenure\_Optimal değişkenlerinin aldığı değerlere göre işletmeden ayrılan 19.330 abonenin durumları görülmektedir. Buna göre; işletmeden ayrılan abonelerin %94’ ünün Adsl’ e sahip olmadığı, %81’ inin detaylı fatura almadığı, %80’ inin telefonunun uluslararası aramaya açık olmadığı, büyük çoğunluğu olan %41’ inin abonelik süresinin 135 günün altında olduğu görülmektedir. Geliştirilen model, Adsl’ i olmayan abonelerin, detaylı fatura almayan abonelerin, telefonu uluslararası aramaya kapalı olan abonelerin, abonelik süresi az olan abonelerin işletmeden ayrılma yönünde eğilimi olacaklarını öngörmektedir. Ortaya çıkan sonuçlar modeli destekler niteliktedir.

Tablo 3.11’ de modelde kullanılan diğer değişkenlerin, işletmeden ayrılan ve işletmede kalan aboneler ayrımında ortalama değerleri görülmektedir.

**Tablo 3.11 İşletmeden Ayrılan Aktif Abonelerin Özellikleri - 2**

Değişken	İşletmeden Ayrılan Aboneler	İşletmede Kalan Aboneler
Mean_Arama_Sayisi_Loc (dk.)	28	56
Gecikmeli_Zamaninda_Oraninda	3,2	0,9
W_Mean_Arama_Yapilmayan_Gun_Sayisi	18	10
Recency_Out (gün)	16	5

Tablo 3.11 incelendiğinde, ortalama şehir içi arama sayıları ortalamasının işletmeden ayrılan aboneler için 28, işletmede kalan aboneler için 56 gün olduğu görülmektedir. Bu durumda abonelerin şehir içi arama sayıları ortalamalarının işletmeden kalmada etkili olduğu ve geliştirilen modelin de öngördüğü gibi bu rakam arttıkça abonelerin işletmede kalmaya devam ettiği söylenebilir. Modelin kullandığı Gecikmeli\_Zamaninda\_Oraninda, W\_Mean\_Arama\_Yapilmayan\_Gun\_Sayisi ve Recency\_Out değişkenlerinin ortalamalarının işletmeden ayrılan abonelerde, işletmede kalmaya devam eden abonelerde olduğundan daha yüksek değerler aldığı görülmekte ve bu durum modelin tahminlerini desteklemektedir.

Tablo 3.12’ de aktif ev abonelerinin işletmeden ayrılma sebepleri ve bu sebepler yüzünden işletmeden ayrılan abone adetleri verilmiştir. İşletmeden ayrılan 19.330 abonenin büyük bir kısmının işletmenin sunduğu hizmetlere ihtiyaç duymaması sebebiyle işletmeden ayrıldığı görülmüştür. İşletmeyi terk etmenin diğer önemli sebebi ise haberleşme ücretinin pahalı olmasıdır.

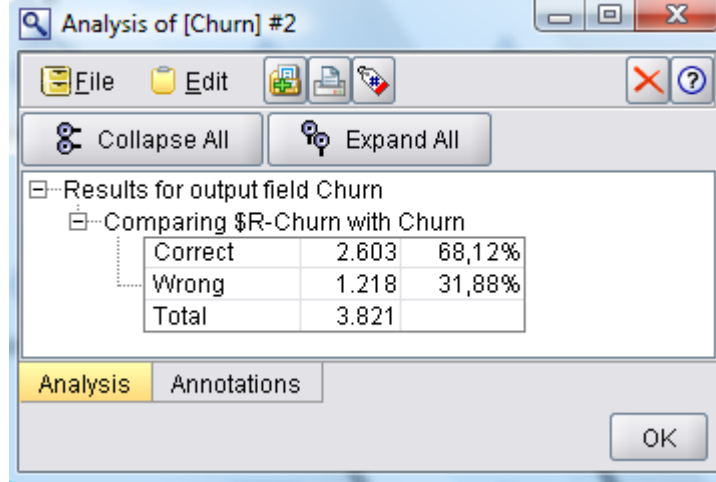
**Tablo 3.12 Aktif Abonelerin İşletmeden Ayrılma Sebepleri**

Sebepler	Sayı
Diğer Sebepler	1.441
Haberleşme ücretinin pahalı olması	3.457
Haberleşmenin Gsm ile yapılacak olması	593
Tarife paketi nedeniyle iptal	156
Taşınma(Yeni adres belli değil)	871
Hizmetlerden memnun olmama	116
Bayilerden memnun olmama	13
İhtiyaç olmaması	12.683

### 3.5.2. Deaktif Ev Aboneleri İin Deęerlendirme

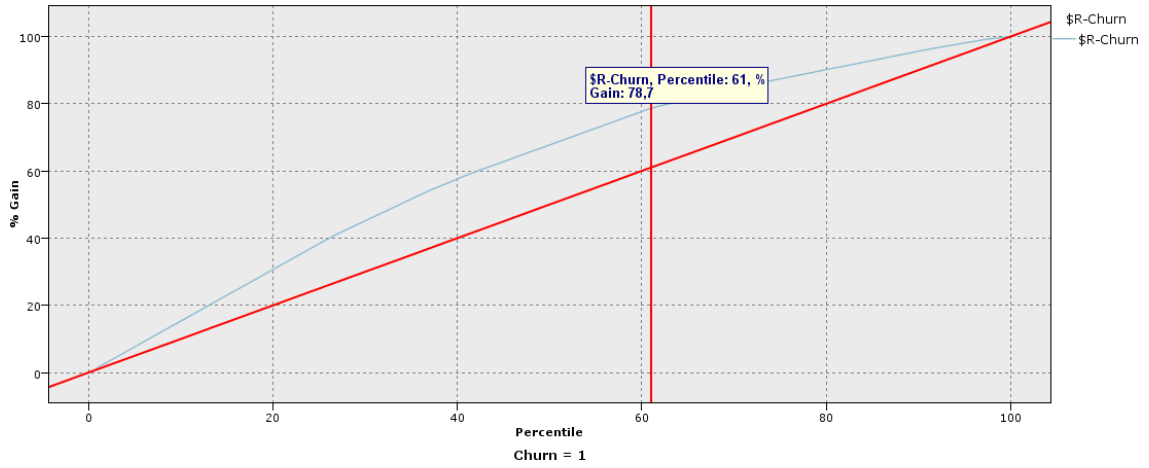
Deaktif ev aboneleri iin geliřtirilen Quest karar aęacı modelini deęerlendirmek iin modelin saęladıęı doęruluk oranına ve kazancına bakılmıřtır.

řekil 3.15 Model Analizi



řekil 3.15 incelendięinde, modelin %68,1 oranında doęruluk payına ve %31,9 oranında yanılma payına sahip olduęu grlecektir.

řekil 3.16 Modelin Kazancı



řekil 3.16' da geliřtirilen modelin kazanç grafięi grlmektedir. Buna gre, abonelerin, modelin belirttięi en yksek ayrılma olasılıęına sahip %61' ine ulařıldıęında, gerekte iřletmeyi terk edecek abonelerin %78,7' sine ulařılmıř olunmaktadır. Bu durumda, mřteri kaybını azaltmak iin yapılan proaktif pazarlama



kampanyalarında hedef olarak %61'lik kısım belirlenirse, işletmeden ayrılacak olan abonelerin %78,7' sine ayrılmadan ulaşmak ve gerekli kampanya duyurusu ya da teklifleri yapmak mümkün olacaktır.

**Tablo 3.13 İşletmeden Ayrılan Deaktif Abonelerin Özellikleri**

Değişken	Özelliği	Kişi Sayısı	Oran
<b>Adsl</b>	0: Adsl yok	1.388	90%
	1: Adsl var	153	10%
<b>Detay</b>	0: Detaylı fatura alınmıyor	1.253	81%
	1: Detaylı fatura alınmıyor	288	19%
<b>Out</b>	0: Telefon dış aramaya kapalı	69	4%
	1: Telefon dış aramaya açık	1.472	96%

Tablo 3.13' te işletmeden ayrılan 1.541 deaktif abonenin modelin kullandığı değişkenlerin değerlerine göre sayıları verilmiştir. Buna göre, işletmeden ayrılan deaktif abonelerin %90' ının Adsl'i bulunmamakta, %81'i detaylı fatura almamakta ve %96'sının telefonu dış aramalara açık durumdadır. Ayrıca, Gecikmeli\_Zamanında\_Oranında değişkenin aldığı değerler ortalaması işletmeden ayrılan aboneler için 3,5 iken, işletmede kalmaya devam eden aboneler için 1,2'dir. Geliştirilen modele göre Adsl sahipliği, detaylı fatura sahipliği, telefonun dış aramaya kapalı olması işletmede kalma yönünde aboneler üzerinde etkilidir. Gecikmeli ödenen fatura sayısının zamanında ödenen fatura sayısına oranı arttıkça aboneler gitmeye eğilimli olmaktadır. Bu durumda modelin geliştirdiği tahminler ile ortaya çıkan sonuçların birbirine denk olduğu söylenebilir.

**Tablo 3.14 Deaktif Abonelerin İşletmeden Ayrılma Sebepleri**

Sebepler	Sayı
Diğer Sebepler	151
Haberleşme ücretinin pahalı olması	193
Haberleşmenin Gsm ile yapılacak olması	44
Tarife paketi nedeniyle iptal	11
Taşınma(Yeni adres belli değil)	65
Hizmetlerden memnun olmama	11
Bayilerden memnun olmama	2
İhtiyaç olmaması	1.431

Deaktif ev abonelerinin işletmeden ayrılma sebepleri aktif ev abonelerin ayrılma sebepleri ile benzerlik göstermektedir. Ayrılan 1.908 abonenin 1.431'inin ihtiyaç olmaması sebebiyle işletmeyi terk ettiği Tablo 3.14' te görülmektedir.

### **3.6. Gerçekleştirme**

Kurulan ve geçerliliği kabul edilen model, doğrudan bir uygulama olabileceği gibi bir başka uygulamanın alt parçası olarak da kullanılabilir. Kurulan modeller risk analizi, kredi değerlendirme, dolandırıcılık tespiti gibi işletme uygulamalarında doğrudan kullanılabilen gibi, tahmin edilen envanter düzeyleri yeniden sipariş noktasının altına düştüğünde otomatik olarak sipariş verilmesini sağlayacak bir uygulamanın içine gömülebilir ([www.bilgiyoneti.org](http://www.bilgiyoneti.org)).

Bu çalışmada elde edilen sonuçlar işletmenin pazarlama faaliyetlerinde kullanılabilir. İşletme ürün ve hizmetleriyle ilgili kampanyalar yaptığında, modellerin yüksek oranda işletmeden ayrılacağını işaret ettiği abonelere öncelik verilerek, işletmenin müşteri kaybı yavaşlatılabilir.

## SONUÇ VE ÖNERİLER

Bu tez çalışmasında, önemi gün geçtikçe artan ve yaygınlaşan veri madenciliği ve müşteri ilişkileri yönetimi ayrıntılı bir şekilde incelenmiş, telekomünikasyon sektöründe faaliyet gösteren bir firmanın verileri veri madenciliği teknikleri ile analiz edilmiştir.

Tezin amacı, veri madenciliği tekniklerinin CRISP-DM süreci ile ele alınarak, gerekli analizlerin yapılması ve veriden anlamlı bilgilerin çıkarılmasıdır. Bu kapsamda, uygulamaya konu olan firmadan ayrılmayı tercih eden müşteri davranışlarını ortaya koyan modeller geliştirilerek, CRM kapsamında müşteriyi elde tutmaya yönelik çeşitli kampanyalar ve pazarlama stratejileri önerilmiştir.

Veri seti aktif ve deaktif aboneler olmak üzere iki kısma ayrılmış ve her biri için ayrı ayrı modeller geliştirilmiştir. Uygulamada SPSS Clementine programı kullanılmıştır. Verinin temizlenmesi ve istenilen formata dönüştürülmesi işlemlerinin ardından, aktif aboneler için Lojistik Regresyon Analizi kullanılmış, deaktif aboneler için ise QUEST Karar Ağacı ile müşterinin işletmeden ayrılma davranışı modellenmiştir.

Telekomünikasyon sektörü sirkülasyonun çok fazla olduğu dinamik bir sektördür. Ayrılma eğilimi gösteren müşteri profili önceden belirlenebilirse, belirlenen hedef gruba özel kampanyalar düzenlenerek müşterilerin işletmede kalması sağlanabilir.

### **Aktif ve deaktif abone ayrımı yapmaksızın tüm abonelerde;**

- Adsl sahipliğinin abonelerin işletmede kalması yönünde etkisi olduğu gözlenmiştir. Adsl firmanın güçlü bir silahıdır. Aboneler Adsl hizmeti alabilmek için önce telefon hizmeti almaktadır. Ancak çıplak Adsl uygulamasının başlaması halinde, özellikle telefonunu kullanmayan deaktif abonelerin işletmeyi terk etme oranlarında artış olabileceği düşünülmektedir. Ayrıca telefonunu kapattıran abonelerin internet hizmetini başka firmalardan alması da mümkündür. Bunları engellemek için işletmenin hem telefon, hem de Adsl hizmetini aboneleri için cazip hale getirmesi gerekmektedir. Bunun için işletme mevcut alt yapısını kullanarak, verdiği Adsl hizmetinde ücreti değiştirmeden bant genişliğini arttırmak yolu ile iyileştirme yapabilir. Aynı fiyata daha

hızlı internet imkânı, bu hizmeti kullanmayan aboneleri kuruma çekerken, mevcut abonelerin kalmasında önemli bir etken olacaktır. Ayrıca internet üzerinden verilebilecek ücretsiz müzik, video gibi mültimedya paketleri ile abonelerin Adsl kullanımlarının arttırılması sağlanabilir. Bunların yanında, aboneler telefon hizmeti ile birlikte sahip olabilecekleri diğer hizmetler konusunda bilgilendirilmelidir. Örneğin, telefonları aracılığı ile diğer telefonlara kısa mesaj gönderebilecekleri, görüntülü haberleşme yapabilecekleri ya da uyandırma servisinden yararlanabilecekleri gibi pek çok konuda aboneler teşvik edilebilir. Bu sayede, telefonun ana amacı olan iletişimin yanı sıra sağladığı bu tür ek hizmetlerle kullanımının arttırılması mümkün olabilir.

- Detaylı fatura almayan abonelerin işletmeden ayrılma eğiliminde olduğu görülmüştür. Detaylı fatura, toplam borcun her bir kalemini ayrıntılı olarak gösterdiğinden, abonenin fatura tutarı üzerinde kontrolünün olmasını sağlar. Başka bir ifade ile, abone gelecek dönem faturaları için nasıl daha az ücret ödeyeceği ile ilgili sonuçlar çıkarabilir. Bu durumda, detaylı fatura almayan ve işletmeden gitmesi öngörülen abonelere ücretsiz detaylı fatura gönderilmesi önerilebilir.

- Abonelerin gecikmeli ödediği fatura sayısının artması, işletmeden ayrılma risklerini de arttırmaktadır. Faturaların zamanında ödenmemesinin sebebi abonenin ödeme yapmayı unutmasından ya da maddi sebeplerden kaynaklanmaktadır. Gecikmeli ödenen faturalar günlük faize uğramakta ve ödenmesi gerekenden daha fazla bir miktara ulaşmaktadır, bu da abonenin hoşnutsuzluğuna sebep olabilir. Mutsuz abone ayrılma eğilimindedir. Unutularak ödenmesi geciken faturaları engellemek için, işletme bankalarla anlaşarak abonelerin otomatik ödeme talimatı vermeleri yönünde teşvik edilmelerini sağlayabilir. Bu, bankanın yapacağı çeşitli promosyon faaliyetleriyle desteklenebilir. Maddi sebeplerden kaynaklanan gecikmeli ödemeleri önlemek için, abonelerle görüşülerek fatura oluşma ve son ödeme tarihlerinin aboneye uygun hale getirilmesi ve mevcut borçlarının taksitlendirilmesi söz konusu olabilir.

### **Aktif abonelerde;**

- Telefonu uluslar arası aramalara kapalı olanların, açık olanlara göre işletmeden ayrılmaya daha fazla eğilimli oldukları gözlenmiştir. İşletmenin yurt dışı arama tarifelerinin diğer rakiplere göre daha uygun olması, yurt dışı arama hizmetini alan abonelerin işletmede kalmaya devam etmelerini sağlamaktadır. Yurt dışı arama hizmetini farklı kanallardan alan kişilere, görsel ve yazılı medya ile ulaşılarak işletmenin fiyat avantajları aktarılabilir.

- Abonenin işletme ile ilişkide olduğu süre arttıkça işletmede kalmaya devam etme olasılığı artmaktadır. Bu durumda yeni abone olmuş kişilerin işletme ile bağı koparmalarının daha kolay olduğu söylenebilir. İşletme aboneleriyle olan ilişki süresini arttırmak için, süreye dayalı promosyon faaliyetlerinde bulunabilir. Bu faaliyetler, yeni abonelikler için, belirlenen süre boyunca abone olarak kalınacağı taahhüdü karşılığında ücretsiz telefon verilmesi şeklinde gerçekleştirilebilir.

- Abonelerin şehir içi arama sayıları arttıkça, işletmeyi terk etme olasılıklarının azaldığı görülmektedir. Bu durum işletmenin, şehir içi konuşmalarında rakiplerine göre daha fazla tercih edildiğinin göstergesidir. Arama sayısı düşük olan aboneleri şehir içi konuşmaya teşvik etmek için, yapacakları şehir içi konuşma süresi kadar ücretsiz ek konuşma süresi verilebilir ya da aynı ücrete daha fazla şehir içi konuşma süresi veren daha uygun tarifeler önerilebilir.

- Arama yapılmayan gün sayısının artması abonelerin işletmeyi terk etmesini tetiklemektedir. Bu sonuç, abonelerin işletmenin hizmetlerini kullanmadıkları için fatura ödemek istememeleriyle açıklanabilir. Abonelere telefonlarını düzenli kullanma alışkanlığı sağlamak için, günün belirli saatlerini kapsayan periyotlarında ücretsiz konuşma süresi verilebilir.

- En son arama yapılan tarihten modelin geliştirildiği tarihe kadar olan gün sayısı arttıkça, abonelerin işletmeden ayrılma eğiliminde oldukları gözlenmektedir. Diğer bir ifadeyle, aboneler daha uzun aralıklarla arama yapmaya başladıkça ayrılma eğilimleri de artmaktadır. Abonelerin daha sık arama yapmalarını sağlamak için

abonelerin profiline uygun tarife seçenekleriyle abonenin işletmede kalması sağlanabilir. Örneğin abonenin meslek grubuna yönelik indirimli tarife seçenekleri sunulabilir.

#### **Deaktif abonelerde;**

- Telefonunu aramalara kapattırılmış abonelerin, telefonu aramalara açık olan abonelere göre işletmede kalmaya daha yatkın oldukları gözlenmiştir. Deaktif aboneler arama yapmamış kişilerden oluştuğundan, abonelerin telefonunu dış aramalara kapattırmasının sebebi, abonelik hizmetini başka bir amaç için kullandıkları şeklinde yorumlanabilir. Bu sebeple telefonu dış aramalara kapalı olan aboneler işletmenin sunduğu diğer hizmetleri almaya devam ettikleri sürece, abone olarak kalmaya devam ederler. Telefonunu kullanmayan ve özellikle dış aramalara kapatmayan abonelerin işletmeden ayrılma eğilimi taşıdığı görülmektedir. Bu durumdaki abonelerin telefonunu sık kullanarak aktif abone statüsüne getirilmesi için indirimli konuşma saatleri ya da ‘30 dakika konuş, 10 dakika bizden hediye’ gibi özendirici kampanyalar düzenlenebilir. Ayrıca işletme sunduğu diğer hizmetlerle ilgili kampanyalar düzenleyip bu grubu oluşturan abonelerine ulaşabilir. Örneğin, abone Adsl hizmeti alabilmek için telefon bağlatmışsa, Adsl’ den vazgeçtiğinde telefonunu da kapattıracaktır. Bunu engellemek için okula giden çocuğu olan abonelere, internet üzerinden kullanılan bir eğitim destek aracı ücretsiz sunulabilir. Bu durumda abone çocuğunun dersleri için Adsl kullanımına devam edecek dolayısı ile telefonunu kapattırmayacaktır.

Çalışma sonucunda elde edilen bilgiler ve sunulan öneriler ile; mevcut aboneler ile ilişkilerin daha da geliştirilmesi, telefon trafiklerinin artırılması ve işletmenin sunduğu diğer hizmetlere olan talebin artırılması sağlanarak, abonelerin ayrılma oranları azaltılabilecek ve işletmenin kar marjı yükseltilebilecektir.

## BİBLİYOGRAFYA

Agrawal, M. L., Customer Relationship Management (CRM) & Corporate Renaissance, Vol.3, No:2, 2003

Agrawal, R., Srikant, R., Fast Algorithms for Mining Association Rules, Proceedings of the 20th VLDB Conference Santiago, 1994

Agresti, A., An Introduction to Categorical Data Analysis, John Wiley and Sons. Inc., 2007

Akgöbek, Ö., Öztemel E., Endüktif Öğrenme Algoritmalarının Kural Üretme Yöntemleri ve Performanslarının Karşılaştırılması, SAÜ Fen Bilimleri Enstitüsü Dergisi 10.Cilt, 1.Sayı 2006

Akpınar, H., Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, İstanbul Üniversitesi İşletme Fakültesi Dergisi, S. 1 -22, Cilt:29, Sayı: 1, 2000

Albayrak, A., Yılmaz Ş., Veri Madenciliği: Karar Ağacı Algoritmaları ve MKB Verileri Üzerine Bir Uygulama, Süleyman Demirel Üniversitesi, İktisadi ve İdari Bilimler Fakültesi Dergisi, C.14, S.1, 2009

Alpaydın, E., Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri, Bilişim Eğitim Semineri, 2000

[http://www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden\\_2k-notlar.doc](http://www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden_2k-notlar.doc)

(Son Erişim Tarihi:12.12.2008)

Ayanoğlu M., Mert K., Giray E., Perakende Sektöründe Veri Madenciliği Vazgeçilmez mi?Alternatifi CRM Mi?, YA/EM XXIV Ulusal Kongresi, Haziran 2004

Baykal, N., Veri Tabanı ve Veri Madenciliği, Tıp Bilişim Güz Okulu Sunumu, 2003

Baykasođlu A., Dereli T., Özen S. H., Müşteri Odaklı Süpermarket Dizaynı İçin Analitik Yaklaşımlar, YA/EM XXIV Ulusal Kongresi, Haziran 2004

Baykasođlu, A., Veri Madenciliđi ve Çimento Sektöründe Bir Uygulama, Akademik Bilişim, Gaziantep, 2005

Berkhin, P., Survey of Clustering Data Mining Techniques, Accrue Software Inc., San Jose, California, USA, 2002

Berry M. J. A. ve Linoff G. S., Mastering Data Mining: The Art and Science of Customer Relationship Management, John Wiley & Sons, NewYork, 2000

Berson A., Smith S., Thearling K., Building Data Mining Applications for CRM, New York: McGraw-Hill, 2000

Bircan, H. Lojistik Regresyon Analizi ve Tıp Verileri Üzerine Bir Uygulama, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2004/2

Bland C., The Discovery of Multiple Level Profile Association Rules, Doctoral Thesis, Graduate School of Computer and Information Science of University of Mississippi, Mississippi, 2002

Bose, R., CRM: Key Components For Its Success, Industrial Management Data Systems, 2002

Bozgeyik, A., CRM Niçin Önemli ve Nasıl Rekabet Avantajı Sağlar, Active Bankacılık ve Finans, Kasım-Aralık, No:21, 2001

Chapman, P. Et al., CRISP-DM 1.0 Step-by-Step Data Mining Guide, Spss Inc., 2000

Chye K. H., Gerry C. K., Data Mining and Customer Relationship Management in the Banking Industry, Singapore Management Review, Volume 24, Issue 2, 2006



Customer Service Makes Conflicting Demands On IT, Insurance Systems Bulletin, 9 (4), 1993

Emel G. G., Taşkın, Ç., Genetik Algoritmalar ve Uygulama Alanları, Uludağ Üniversitesi İİBF Dergisi, Cilt:XXI, Sayı:1, Bursa, 2002

Erdem, A., Uzun E., Yapay Sinir Ağları İle Türkçe Times New Roman, Arial El Yazısı Karakterlerini Tanıma, Gazi Üniversitesi, Mühendislik-Mimarlık Fakültesi Dergisi, Cilt 20, No 1, 13-19, 2005

Friedman, J. H., Data Mining and Statistics: What's the Connection?, Stanford University, Department Of Statistics and Stanford Linear Accelerator Center, 1997

Girginer, N., Cankuş, B., Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Estram Örneği, Celal Bayar Üniversitesi, Yönetim ve Ekonomi Dergisi, Cilt:15 Sayı:1, 2008

Goldenberg, B., Customer Relationship Management: What Is It All About, ISM, 1999 [www.mycustomer.com/lib/918](http://www.mycustomer.com/lib/918) (Son Erişim Tarihi: 10.04.2009)

Han, J., Kamber, M., Data Mining Concepts and Techniques, Morgan Kaufmann Publishers Inc., 2001

Kantardzic, M., Data Mining Concepts, Models, Methods and Algorithms, John Wiley & Sons Publishers, 2003

Kaur., H., Wasan., S., Empirical Study on Applications of Data Mining Techniques in Healthcare, Journal of Computer Science 2(2), 2006

Kaya, E., Bulun, M., Arslan, A., Tıpta Veri Ambarları Oluşturma ve Veri Madenciliği Uygulamaları, 2005 <http://ab.org.tr/ab03/program/96.html> (Son Erişim Tarihi: 10.08.2008)

Kevin B., et all., When Is Nearest Neighbour Meaningful?, 7. Uluslararası Database Theory Konferansı Bildirisi (ICDT.99), İsrail, 1999

Kırım, A., Strateji ve Birebir Pazarlama, Sistem Yayıncılık, İstanbul, 2001

Larose, D.T., Discovering Knowledge In Data: An Introduction To Data Mining, Wiley Computer Publishing, 2005

Massey P., Weiss, M., Holcom, K., Reengineering CRM: Leveraging Knowledge Assets In IBM, Decision Support Systems, 32 (2), 2001

Özekeş S., Data Mining Models and Application Areas, Istanbul Commerce University Journal of Science, No.3, 2003

Özmen, Ş., İş Hayatı Veri Madenciliği İle İstatistik Uygulamalarını Yeniden Keşfediyor, V.Ulusal Ekonometri Ve İstatistik Sempozyumu, Eylül 2001

Russel S., A Framework For CRM, California Management Review, 43 (4), 2001

Rygielski, C., Wang, J. ve Yen, D.C., Data Mining Techniques for Customer Relationship Management, Technology in Society, Volume 24, Issue 4, 2002

Sattari, M. Et. All., Yapay Sinir Ağları Yöntemi ile Savalan Sulama Rezervuarının Simülasyonu, Ankara Üniversitesi Ziraat Fakültesi Tarım Bilimleri Dergisi 13 (4), 2007

Selamoğlu A., Yönetim ve Üretim Anlayışında Değişim, Japon Modelinin Artan Etkinliği ve İnsan Unsuru, Çimento İşveren Dergisi, Ankara, Cilt:12, Sayı:6, Kasım 1998

Serper, Ö., Uygulamalı İstatistik, 1, 1. Baskı, Marmara Kitabevi, Bursa, 1997

Sicimođlu, H., Grdđmz Iřık UFO Deđil CRM, Btnleřik Pazarlamada Marketing Trkiye, Yıl:1, Sayı: 7, Temmuz 2002

Squier, L., What is Data Mining, Data Management Association-NCR Presentation, 2001

SPSS Inc., Clementine 12.0 Algorithms Guide, 2007

Srivastava J., et. al., A Case for Analytical Customer Relationship Management, Computer Science & Engineering, University of Minnesota, Minneapolis, USA, 2002

Swift, R. S., Accelerating Customer Relationships, Prentice Hall PTR A Pearson Education Company, Upper Saddle River, 2001

Tařpınar, H., Biliřim Altyapısıyla CRM, Ankara, 2006

Tek, . B., Pazarlama İlkeleri-Global Ynetimsel Yaklařım, Trkiye Uygulamaları, 8. Baskı, Beta Basım Yayım A.S., İstanbul, 1999

Tekin, M., iek, E., Deđiřim Ynetimi Srecinde Mřteri İliřkileri Ynetimi ve nemi, III. Ulusal retim Arařtırmaları Sempozyumu, Nisan 2003

Toktař, P., Demirhan M., Risk Analizinde Veri Madenciliđi Uygulamaları, YA/EM XXIV Ulusal Kongresi, Haziran 2004

Vassiliadis P., Sellis T. K., A Survey of Logical Models for OLAP Databases, ACM SIGMOD Record, Vol.28, No.4, 1999

Yereli, A. N., Mřteri İliřkileri Ynetimi (CRM) ve Gnmz Trkiye'sindeki Yeri, Celal Bayar niversitesi Ynetim ve Ekonomi Dergisi, Sayı: 1, Cilt:7, 2001

## **Diğer Kaynaklar**

“Analitik CRM: İş Zekâsı Çözümleri”

<http://www.crminturkey.org> (Son Erişim Tarihi:09.09.2008)

“SQL Veri Tabanı Fonksiyonları”

<http://science.ankara.edu.tr> (Son Erişim Tarihi: 20.03.2006)

Ülküderner, Ç., “SQL’e Giriş”

[http://seminer.linux.org.tr/wiki/index.php/Seminer\\_Notlar%C4%B1\\_2001](http://seminer.linux.org.tr/wiki/index.php/Seminer_Notlar%C4%B1_2001)

(Son Erişim Tarihi: 23.01.2008)

“Process Model”

<http://www.crisp-dm.org/Process/index.htm>

(Son Erişim Tarihi: 07.02.2009)

“Veri Madenciliğinin Yaygın Uygulama Alanları”, <http://www.spss.com.tr>

(Son Erişim: 28.08.2008)

“Data Mining Applications In 2008”

<http://www.kdnuggets.com/polls/2008/data-mining-applications.htm>

(Son Erişim Tarihi: 18.03.2009)

Berry, M., Linoff, G., “Data Mining Techniques: Decision Trees”

[http://www.rhsmith.umd.edu/faculty/bgolden/course\\_pack\\_slides.ppt](http://www.rhsmith.umd.edu/faculty/bgolden/course_pack_slides.ppt)

(Son Erişim Tarihi: 13.11.2008)

[www.sciencedirect.com](http://www.sciencedirect.com) (Son Erişim Tarihi: 25.10.2008)

Akman, S., “Regresyon Analizi”

[www.tajev.org/dosyalar/regresyon\\_sunum%20sema%20akman.ppt](http://www.tajev.org/dosyalar/regresyon_sunum%20sema%20akman.ppt) (Son Erişim Tarihi:

22.02.2009)

“Satış Hedefleri ve CRM’ nin Sağladığı Faydalar”  
www.csainsights.com (Son Erişim tarihi: 21.12.2008)

Dickie J., “CRM projelerinin Başarısız Olma Nedenleri”  
www.csainsights.com (Son Erişim Tarihi:21.12.2008)

“Crm Nedir?” <http://www.erpcrm.com> (Son Erişim Tarihi: 19.09.2008)

Akmeriç, N., “Müşteri İlişkileri Yönetimi”  
<http://www.sas.com/offices/europe/turkey/news> (Son Erişim Tarihi: 18.01.2009)

“CRM: Müşteriye Dair Her Şey”  
[http://www.microsoft.com/turkiye/dynamics/crm/crm\\_nedir.msp](http://www.microsoft.com/turkiye/dynamics/crm/crm_nedir.msp)  
(Son Erişim Tarihi:15.03.2009)

“Türkiye’nin CRM İbrelere”  
www.crminturkey.org (Son Erişim Tarihi: 17.10.2008)

Milliyet Gazetesi, “Migros ve Garanti Bankası’ndan Money Card...”  
<http://www.milliyet.com.tr/Ekonomi/SonDakika.aspx?aType=SonDakika&ArticleID=1092336> (Son Erişim Tarihi:16.05.2009)

Milliyet Gazetesi, “Garanti’den çocuklara özel ATM: Minimatik”  
<http://www.milliyet.com.tr/default.aspx?aType=HaberDetay&ArticleID=518309> (Son Erişim Tarihi:16.05.2009)

Sabah Gazetesi, “Çocuğu hasta anne için doktor buldular, CRM’de birinci oldular”  
<http://arsiv.sabah.com.tr/2007/03/10/eko105.html> (Son Erişim Tarihi:16.05.2009)

“Bir Şarap Üreticisinin CRM Yaklaşımı”  
<http://crmpro.com.tr/2006/07/bir-sarap-ureticisinin-crm-yaklami.html>  
(Son Erişim Tarihi:23.04.2009)

“CRM Uygulaması / ToyotaSA”

<http://www.veripark.com/tr/Portfolio/SuccessStories/toyotasa.asp>

(Son Erişim Tarihi: 10.02.2009)

“AçıkDeniz / DenizBank”

<http://www.veripark.com/tr/Portfolio/SuccessStories/denizbank.asp>

(Son Erişim Tarihi: 10.02.2009)

“Burger King, Microsoft CRM ile Paket Servis Hizmetini Başlatırken Müşterisini Tanıma Fırsatı da Yakaladı”

<http://www.microsoft.com/turkiye/dynamics/basari/tabgida.msp>

(Son Erişim Tarihi: 10.02.2009)

“Trabzonspor Kulübü, 100 Bin Kulüp Üyesi Hedefine Microsoft CRM ile İlerliyor”

<http://www.microsoft.com/turkiye/dynamics/basari/ts.msp>

(Son Erişim Tarihi: 10.02.2009)

“Delta CRM İle Başarıya Koşuyor”

<http://www.comel.com.tr/basari/delta.html> (Son Erişim Tarihi: 10.02.2009)

Eker H., “Veri Madenciliği Veya Bilgi Keşfi”

[http://www.bilgiyonetimi.org/cm/pages/mkl\\_gos.php?nt=538](http://www.bilgiyonetimi.org/cm/pages/mkl_gos.php?nt=538)

(Son Erişim Tarihi: 10.02.2009)

## **EKLER**

### EK-1 Orijinal Veri Tablosu Formatı

AboneNo	İL	İlce	Tenure	Tarife	Churn	Churn_Sebep	Gecikmeli_Zamaninda_Orani	Oto_Odeme	Recency_Out	Adsl	Int	Detay	Yas	Gsm	Na	Out
1	232	KONAK	3070	11	1	İhtiyaç olmaması	2	0	3	0	0	0	43	1	1	1
2	232	KONAK	1519	11	0		0,5	1	1	1	0	1	47	1	1	1
3	232	KONAK	5726	8	0		5	0	0	0	0	1	54	0	0	0
4	232	KONAK	9796	8	0		0	0	5	1	0	0	48	1	1	1
5	232	KONAK	1053	8	1	Haberleşme ücretinin pahalı olması	2	1	2	1	1	0	54	1	1	1
6	232	KONAK	686	11	1	İhtiyaç olmaması	1	0	2	0	0	0	34	1	1	1
7	232	KONAK	372	8	0		6	0	1	0	0	1	28	1	1	1
8	232	KONAK	6280	11	0		0	0	1	0	0	0	32	1	1	1
9	232	KONAK	1005	8	1	İhtiyaç olmaması	0,5	0	10	1	0	0	51	1	1	1
10	232	KONAK	848	8	0		0	0	6	0	0	0	45	1	1	1



**EK-1 (Devam)Orijinal Veri Tablosu Formatı**

Arizali_Gun	TNR_1	TNR_2	TNR_3	TNR_4	TNR_5	TNR_6	MEAN_Kullanim_Miktari_LOC	MEAN_Kullanim_Miktari_NAT	MEAN_Kullanim_Miktari_GSM	MEAN_Kullanim_Miktari_INT	MEAN_Kullanim_Miktari_TOT
0	1	1	1	1	1	1	37.371667	1.336667	12.048333	0	50.758333
0	1	1	1	1	1	1	115.545	62.008333	10.656667	2.106667	190.31333
0.49	1	1	1	1	1	1	0	0	0	0	0
0.02	1	1	1	1	1	1	673.44833	69.41	117.40667	2.68	862.945
0	1	1	1	1	1	1	141.18333	1.96	19.078333	0	162.22
0.1	1	1	1	1	1	1	112.22	17.956667	12.858333	0	143.035
0.04	1	1	1	1	1	1	137.285	2.921667	13.015	1.008333	154.23
0	1	1	1	1	1	1	235.90167	7.526667	7.41	0	250.83833
0	1	1	1	1	1	1	140.35333	46.076667	37.425	0	223.855
0	1	1	1	1	1	1	178.40833	51.598333	2.313333	0.003333	232.32333

**EK-1 (Devam)Orijinal Veri Tablosu Formatı**

MEAN_Arama_Sayisi_LOC	MEAN_Arama_Sayisi_NAT	MEAN_Arama_Sayisi_GSM	MEAN_Arama_Sayisi_INT	MEAN_Arama_Sayisi_TOT	MEAN_Arama_Yapilmayan_Gun_sayisi	MEAN_Kullanim_Ucreti	W_MEAN_Kullanim_Miktari_LOC
23.333333	0.666667	14	0	38	17.166667	7.701667	44.62
61.333333	17.166667	13.666667	1.166667	93.333333	2.666667	16.775	109.302857
0	0	0	0	0	30	0	0
406	23	72.166667	0.5	501.666667	5.166667	100.2583	640.084762
141.666667	1	17.166667	0	159.833333	1.333333	19.99667	147.050476
49.833333	5.833333	4.666667	0	60.333333	12.166667	16.92	99.347619
125.666667	1.166667	27.833333	0.166667	154.833333	4.333333	17.89667	138.942857
137.5	4.5	11.333333	0	153.5	0.5	21.15167	250.144286
48.833333	5.666667	21.166667	0	75.666667	13.166667	25.565	129.979048
79.833333	11.166667	1.833333	0.166667	93	6.833333	18.08833	155.015238

**EK-1 (Devam)Orijinal Veri Tablosu Formatı**

W_MEAN_ Kullanım_Miktari_ NAT	W_MEAN_ Kullanım_Miktari_ GSM	W_MEAN_ Kullanım_Miktari_ INT	W_MEAN_ Kullanım_Miktari_ TOT	W_MEAN_ Arama_Sayisi_ LOC	W_MEAN_ Arama_Sayisi_NAT	W_MEAN_ Arama_Sayisi_GS M
1.177143	13.10381	0	58.902381	28.714286	0.761905	16.428571
67.904286	10.349524	2.303333	189.85524	56.619048	18.238095	13.571429
0	0	0	0	0	0	0
71.002381	119.22048	4.594286	834.90191	379.38095	23.142857	68.904762
2.24	19.244762	0	168.53333	138.09524	1.142857	18.190476
11.883333	11.158571	0	122.39	43.285714	3.047619	3.714286
3.704762	13.804762	1.440476	157.89238	124.2381	1.47619	27.666667
7.301429	7.919048	0	265.36524	137.52381	4.857143	12.047619
32.62619	35.655238	0	198.26	46.047619	5.142857	19.952381
60.311429	2.254762	0.001905	217.58476	70.095238	12.619048	1.47619

**EK-1 (Devam)Orijinal Veri Tablosu Formatı**

W_MEAN_Arama_Sayisi_INT	W_MEAN_Arama_Sayisi_TOT	W_MEAN_Arama_Yapilmayan_Gun_sayisi	W_MEAN_Kullanim_Ucreti	N_MEAN_Kullanim_Miktari_LOC	N_MEAN_Kullanim_Miktari_NAT	N_MEAN_Kullanim_Miktari_GSM	N_MEAN_Kullanim_Miktari_INT	N_MEAN_Kullanim_Miktari_TOT
0	45.904762	14.380952	9.329048	1.193953	0.880656	1.087604	0	1.160447
1.380952	89.809524	2.47619	16.682381	0.945977	1.095083	0.971178	1.093354	0.997593
0	0	30	0	0	0	0	0	0
0.857143	472.28571	5.238095	97.155238	0.950459	1.022942	1.015449	1.714286	0.967503
0	157.42857	1.857143	20.23381	1.041557	1.142857	1.008723	0	1.038918
0	50.047619	14.142857	15.457619	0.885293	0.661778	0.867809	0	0.855665
0.238095	153.61905	3.904762	18.561905	1.012076	1.26803	1.060681	1.428571	1.023746
0	154.61905	0.428571	21.800476	1.060375	0.970075	1.068697	0	1.057913
0	71.142857	15.857143	23.439048	0.926085	0.708085	0.952712	0	0.885663
0.095238	84.285714	9	16.688095	0.868879	1.168864	0.974681	0.571429	0.93656

**EK-1 (Devam)Orijinal Veri Tablosu Formatı**

N_MEAN_ Arama_Sayisi_LOC	N_MEAN_ Arama_Sayisi_NAT	N_MEAN_ Arama_Sayisi_GSM	N_MEAN_ Arama_Sayisi_ INT	N_MEAN_ Arama_Sayisi_TOT	N_MEAN_ Arama_Yapilmayan_ Gun_sayisi	N_MEAN_ Kullanım_Ucreti
1.230612	1.142857	1.173469	0	1.20802	0.837725	1.211302
0.923137	1.062413	0.993031	1.183673	0.962245	0.928571	0.994479
0	0	0	0	0	1	0
0.934436	1.006211	0.9548	1.714286	0.941433	1.013825	0.969049
0.97479	1.142857	1.059639	0	0.984955	1.392857	1.011859
0.86861	0.522449	0.795918	0	0.829519	1.162427	0.913571
0.988632	1.265306	0.994012	1.428571	0.992157	0.901099	1.037171
1.000173	1.079365	1.063025	0	1.00729	0.857143	1.030674
0.942955	0.907563	0.942632	0	0.940214	1.20434	0.916841
0.87802	1.130064	0.805195	0.571429	0.906298	1.317073	0.922589