



CUMHURİYET ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü
Yönetim Bilişim Sistemleri Ana Bilim Dalı

**MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ VE
İŞ ZEKÂSI UYGULAMALARI**

Yüksek Lisans Tezi

Mehmet Ali DEVECİ

Sivas
Ocak 2018

CUMHURİYET ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü
Yönetim Bilişim Sistemleri Ana Bilim Dalı

MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ
VE İŞ ZEKÂSI UYGULAMALARI

Yüksek Lisans Tezi

Mehmet Ali DEVECİ

Tez Danışmanı
Doç. Dr. Oğuz KAYNAR

Sivas
Ocak 2018

KABUL VE ONAY

Üniversite: : Cumhuriyet Üniversitesi
Enstitü : Sosyal Bilimler Enstitüsü
Ana Bilim Dalı : Yönetim Bilişim Sistemleri
Bilim Dalı : Yönetim Bilişim Sistemleri
Tezin Başlığı : MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ VE İŞ ZEKÂSI UYGULAMALARI
Savunma Tarihi : 27/12/2017
Danışmanı : Doç. Dr. Oğuz KAYNAR

Unvanı - Adı Soyadı

İmza

Jüri Başkanı : Doç. Dr. Oğuz KAYNAR



Üye : Doç. Dr. Mehmet Ali ALAN



Üye : Doç. Dr. Metin ZONTUL



Oy Birliği

Oy Çokluğu

Mehmet Ali DEVECİ tarafından hazırlanan MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ VE İŞ ZEKÂSI UYGULAMALARI başlıklı tez, kabul edilmiştir.

..../..../.....

Prof. Dr. Ahmet ŞENGÖNÜL
Enstitü Müdürü

ETİK İLKELERE UYGUNLUK BEYANI

Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü bünyesinde hazırladığım bu Yüksek Lisans/Doktora/Sanatta Yeterlik tezinin bizzat tarafımdan ve kendi sözcüklerimle yazılmış orijinal bir çalışma olduğunu ve bu tezde;

- 1- Çeşitli yazarların çalışmalarından faydalandığımda bu çalışmaların ilgili bölümlerini doğru ve net biçimde göstererek yazarlara açık biçimde atıfta bulunduğumu;
- 2- Yazdığım metinlerin tamamı ya da sadece bir kısmı, daha önce herhangi bir yerde yayımlanmışsa bunu da açıkça ifade ederek gösterdiğimi;
- 3- Başkalarına ait alıntılanan tüm verileri (tablo, grafik, şekil vb. de dahil olmak üzere) atıflarla belirttiğimi;
- 4- Başka yazarların kendi kelimeleriyle alıntıladığım metinlerini, tırnak içerisinde veya farklı dizerek verdiğim yine başka yazarlara ait olup fakat kendi sözcüklerimle ifade ettiğim hususları da istisnasız olarak kaynak göstererek belirttiğimi,

beyan ve bu etik ilkeleri ihlal etmiş olmam halinde bütün sonuçlarına katlanacağımı kabul ederim.

09/01/2018

Mehmet Ali DEVECİ

ÖN SÖZ

Bu tez konusunu seçmemde bana yardımcı olan aynı zamanda engin bilgi birikimiyle birlikte sürekli hoşgörölü tavrıyla çalışmalarım boyunca bana yol gösteren ve hiçbir zaman desteğini esirgemeyen çok kıymetli hocam Doç. Dr. Oğuz KAYNAR'a;

Ayrıca Tezimin uygulama kısmında meslekî bilgilerinden istifade etmemi sağlayan, tez süresince katkılarını esirgemeyen değerli dostum Sayın Yasin GÖRMEZ'e;

Bugüne ulaşmamı sağlayan ve üzerimde kuşkusuz ölçülemez emeklere sahip olan değerli aileme teşekkürlerimi sunmayı bir borç bilirim.

Mehmet Ali DEVECİ

İÇİNDEKİLER

İÇİNDEKİLER	i
KISALTMALAR	v
TABLO LİSTESİ	vii
ŞEKİL LİSTESİ.....	ix
ÖZET.....	xi
ABSTRACT	xiii
1. GİRİŞ	1
2. MÜŞTERİ İLİŞKİLERİ YÖNETİMİ	7
2.1. Ürün Odaklı Yaklaşımdan Müşteri İlişkileri Yönetimine.....	7
2.1.1. Geleneksel Pazarlama Anlayışı	8
2.1.2. Modern Pazarlama Anlayışı	9
2.1.3. İşletme Odağı Olarak Müşteri	10
2.2. Müşteri ve Tüketici Kavramları	11
2.2.1. İç ve Dış Müşteri	12
2.2.2. Aktif ve Pasif Müşteri	12
2.3. Müşteri İlişkileri Yönetiminin Ortaya Çıkışı ve Gelişimi.....	12
2.4. Müşteri İlişkileri Yönetiminin Tanımı ve Önemi	13
2.5. Küreselleşme ve Müşteri İlişkileri Yönetimi	15
2.6. Müşteri İlişkileri Yönetiminin Faydaları	17
2.7. Müşteri İlişkileri Yönetiminin Unsurları	18
2.7.1.Müşteri Sadakati.....	18
2.7.1.1.Mutlak Sadakat.....	20
2.7.1.2.Gizli Sadakat	21

2.7.1.3.Sahte Sadakat	21
2.7.1.4.Sadakatsizlik.....	21
2.7.2.Müşteri Odaklılık	23
2.8. Müşteri İlişkileri Yönetimi Mimarisi	23
2.8.1.Analitik Müşteri İlişkileri Yönetimi.....	24
2.8.2.Operasyonel Müşteri İlişkileri Yönetimi.....	25
2.8.3.İşbirlikçi Müşteri İlişkileri Yönetimi	25
2.9. Müşteri İlişkileri Yönetiminin Evreleri.....	26
2.9.1.Müşteri Seçimi	26
2.9.2.Müşteri Edinme	26
2.9.3.Müşteri Koruma	27
2.9.4.Müşteri Derinleştirme	27
2.10. Müşteri İlişkileri Yönetimi Uygulama Evreleri	28
2.10.1.Müşterileri Tanımlanma	28
2.10.2.Müşterileri Farklılaştırma.....	28
2.10.3.Müşterilerle Etkileşime Girme	29
2.10.4.Kitlesele Kişiselleştirme.....	29
2.11. Müşteri İlişkileri Yönetimi Bileşenleri	29
2.11.1.İnsan	29
2.11.2.Süreç.....	30
2.11.3.Teknoloji	30
3. İŞ ZEKÂSI.....	33
3.1. İş Zekâsı Kavramları	35
3.1.1.Verii, Enformasyon ve Bilgi	35
3.1.2.OLTP ve OLAP.....	36

3.1.3.ETL	37
3.1.4.Metadata	38
3.2. İş Zekâsı Uygulama Adımları	38
3.2.1.Veri Kaynakları	38
3.2.1.1.Dahili Veriler.....	39
3.2.1.2.Harici Veriler.....	39
3.2.1.3.Kişisel Veriler	40
3.2.2.Veri Ambarları ve Veri Marketleri	40
3.2.2.1.Veri Ambarlarında Kullanılan Tablo Yapıları	41
3.2.2.1.1.Yıldız Şeması (Star Schema).....	42
3.2.2.1.2.Kar Tanesi Şeması (Snowflake Schema)	42
3.2.2.1.3.Takımyıldızı Şemalarıdır (Fact Constellations)	43
3.2.3.Veri Keşfi	44
3.2.4.Veri Madenciliği	44
3.2.5.Optimizasyon	45
3.2.6.Karar.....	45
4. VERİ MADENCİLİĞİ.....	47
4.1. Veri Madenciliğinde Kullanılan Modeller	48
4.1.1.Sınıflama ve Regresyon	48
4.1.1.1.Naive Bayes.....	49
4.1.1.2.Yapay Sinir Ağları	50
4.1.1.3.Destek Vektör Makineleri	52
4.1.1.4.Lineer Regresyon	53
4.1.2.Kümeleme	53
4.1.2.1.K-Ortalama Algoritması.....	55

4.1.2.2.Öz Örgütlenmeli Ağlar (ÖÖA).....	56
4.1.3.Öznitelik Seçim Algoritmaları	59
4.1.3.1.Ki – Kare	59
4.1.3.2.Bilgi Kazancı	60
4.1.3.3.Kazanım Oranı	60
4.1.3.4.ReliefF	61
5. MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ.....	63
5.1. RFM Analizi.....	63
5.2. Müşteri Kaybı Analizi	67
6. UYGULAMA.....	69
6.1. RFM Müşteri Segmentasyonu.....	69
6.2. Müşteri Kaybı Analizi	74
7. SONUÇ	81
KAYNAKÇA	85
ÖZ GEÇMİŞ.....	95

KISALTMALAR

MİY	: Müşteri İlişkileri Yönetimi
RFM	: Recency, Frequency, Monatery
WOM	: Word of Mouth
BPR	: Business Process Re-Engineering
CPR	: Customer Process Re-Engineering
OLTP	: Online Transactional Processing
OLAP	: Online Analitical Processing
ETL	: Extract, Transform, Loading
VM	: Veri Madenciliği
YSA	: Yapay Sinir Ağları
MLP	: Multilayer Perceptron
DVM	: Destek Vektör Makinaları
OVO	: One Versus One
OVA	: One Versus All
LR	: Lineer Regresyon
SSE	: Sum of Squared Error
ÖÖA	: Öz Örgütlenmeli Ağlar
MYBD	: Müşteri Yaşam Boyu Değeri
NB	: Naive Bayes
BK	: Bilgi kazancı
KO	: Kazanım Oranı

TABLO LİSTESİ

Tablo 1. Müşteri davranış modelleri	69
Tablo 2. Müşteri davranış modelleri (Tez çalışmasında esas alınan)	69
Tablo 3. Müşteri yaşam boyu değeri.....	70
Tablo 4. Kümeleme Yöntemleri Değerlendirme Tablosu.....	71
Tablo 5. ÖÖA kümeleme algoritmasına göre RFM gruplama.....	72
Tablo 6. K-Ortalama kümeleme algoritmasına göre RFM gruplama	73
Tablo 7. Veri seti öznitelik bilgileri	75
Tablo 8. Sınıflama sonucunda elde edilen karmaşıklık matrisleri	76
Tablo 9. Test Verisi Sonuç Değerleri.....	77
Tablo 10. Özniteliklerin Yönteme Bağlı Önem Sırası	77

ŞEKİL LİSTESİ

Şekil 1. Sadık Müşteri Neden Önemlidir (Reichheld, Teal 1996).	19
Şekil 2. Dick ve Basu'nun Sadakat Modeli (1994).....	20
Şekil 3. Müşteri İlişkileri Yönetimi Mimarisi ve İş Zekâsı	23
Şekil 4. İş Zekâsı.....	34
Şekil 5. İş Zekası Piramidi (Kaynak: Vercellis, 2009: 10)	38
Şekil 6. Yıldız tablo yapısı.....	42
Şekil 7. Kar tanesi tablo yapısı.....	43
Şekil 8. Takım yıldızı tablo yapısı	43
Şekil 9. YSA modeli	51
Şekil 10. Destek Vektör Makineleri ve Hiper Düzlem Seçimi	52
Şekil 11. K-Ortalama Algoritması (k=3)	55
Şekil 12. Öz Örgütlenmeli Ağlar (ÖÖA).....	57
Şekil 13. Bağımsız RFM Gruplama	64
Şekil 14. İç içe RFM Gruplama	65
Şekil 15. Lineer Regresyon ROC Eğrisi	78
Şekil 16. Yapay Sinir Ağları ROC Eğrisi	78
Şekil 17. Naive Bayes ROC Eğrisi	79

ÖZET

Günümüzde küreselleşmenin de etkisiyle artan rekabet işletmeler için müşterilerin ihtiyaç, talep ve beklentilerini yakından takip ederek en uygun ve en hızlı şekilde cevap verme zorunluluğu doğurmuştur. Müşteri taleplerinin öneminin artması ve işletmelerin bu talepleri takip ederek gerekli şekilde karşılık verme zorunluluğu müşteri ilişkileri yönetimi kavramını ortaya çıkarmıştır. Müşteri ilişkileri yönetimi, müşterilerle karşılıklı faydaya dayanan ilişkiler geliştirerek müşteri sadakatini kazanmayı ve bu sayede işletmeye uzun vadeli fayda sağlamayı amaçlayan bir yaklaşımdır. Bu yaklaşımın başarısı işletmenin müşteriyi doğru şekilde tanımasına bağlı olarak artmaktadır. Bu nedenle müşterilerle etkili şekilde iletişim kurmayı amaçlayan müşteri ilişkileri yönetimi uzmanları müşterilere ait her türlü bilgiye önem vermektedir. İşletmeler, teknolojinin de yardımıyla farklı kaynaklardan topladıkları verileri çeşitli veri madenciliği yöntemleriyle inceleyerek müşterilere ait daha geçerli bilgiler elde edebilmekte ve bu sayede müşterilerle daha etkili bir iletişim geliştirebilmektedir.

Bu tez çalışmasında ilk olarak bir e-ticaret sitesinden alışveriş yapan müşterilerin alışveriş kalıpları belirlenmeye ve bu kalıplara uygun pazarlama stratejileri önerilmeye çalışılmıştır. Bu amaçla müşteriler alışveriş sıklıkları, son alışveriş tarihleri ve incelenen dönem içerisindeki harcama meblağları bakımından ÖÖA ve K-means algoritmaları kullanarak 8 gruba ayrılmıştır. Oluşturulan bu 8 müşteri grubunda ÖÖA algoritması sonucunda 3, K-means algoritması sonucunda ise 4 farklı davranış modeli ortaya çıkmıştır. Çalışmada ikinci olarak bir telekomünikasyon şirketine ait müşteri verileri üzerinde yapay sinir ağları, naive-bayes, destek vektör makineleri ve lineer regresyon kullanılarak müşteri ayrılma analizi yapılmış ve söz konusu algoritmaların bu analiz için etkinliği belirlenmeye çalışılmıştır. Bu analiz için kullanılan veri seti 4667 müşteriden ve her müşteri için 19 adet işlem kaydına ait özellikler ile müşterinin terk edip terk etmediğine dair sınıf etiketlerinden oluşmaktadır. Ayrıca söz konusu özelliklerin sınıflamaya olan etkilerinin sıralanması için ki-kare, bilgi kazancı, kazanım oranı ve Relief-F olmak üzere 4 farklı metrik aynı veri seti üzerine uygulanmıştır.

Anahtar Kelimeler: RFM analizi, müşteri ayrılma analizi, müşteri ilişkileri yönetimi

ABSTRACT

At the present time, increasing competition due to influence of globalization has caused the obligation of responding to the needs, demands and expectations of the customers for the enterprises in the most appropriate and fastest way. Increasing of importance of customer requests and the obligation of enterprises to respond to these requests in the necessary manner reveals the concept of customer relationship management. Customer relationship management is an approach that aims to acquire customer loyalty by developing mutually beneficial relationships with customers and to provide long-term benefits to company. The success of this approach is increase depend on the ability of company to properly define customer's needs, characteristic and demands. For this reason, customer relationship management specialists who aim to communicate effectively with customers give importance to every kind of information belonging to customers. With the help of technology, company can obtain more valid information from customers by examining the data which collect from various sources through various data mining methods and can develop more effective communication with customers.

In this thesis, firstly, it was tried to determine the shopping patterns of the customers who shop at an e-commerce site and to suggest marketing strategies that are appropriate for the customers who included in these shopping patterns. For this purpose, customers are divided into 8 groups in terms of shopping frequency, last shopping dates and expenditure amounts in the period examined by using the self-organizing maps and K-means algorithms. In the 8 groups of customers that have been created, 3 different patterns of behavior have emerged as the result of the self-organizing maps algorithm and 4 different behavior models emerged as the result of the K-means algorithm. In the study, secondly, customer churn analysis was performed using artificial neural networks, naive-bays, support vector machines and linear regression on customer data of a telecommunication company and it has been tried to determine the effectiveness of these algorithms for this analysis. The data set used for this analysis consists of 4667 customers, 19 transaction record features for each customer and class labels indicating that the customer has churn or not-churn. In addition, four different metrics were applied to the same dataset: chi-square,

information gain, recovery rate and Relief-F for ordering the effects of classification of the features.

Keywords: RFM analysis, customer churn analysis, customer relationship management

1. GİRİŞ

Son yıllarda ortaya çıkan küreselleşme akımı ve teknolojide yaşanan gelişmeler neticesinde farklı coğrafyalarda üretilen ürün ve hizmetlere dünyanın herhangi bir yerindeki müşteriler tarafından kolayca erişilebilmektedir. Bu durum ise piyasalardaki rekabet yoğunluğunu gün geçtikçe arttırmaktadır. İşletmeler bu yeni rekabet yoğun piyasalarda üstünlük elde edebilmek için işletmenin dahili ve harici çevresine dair olabildiğince fazla veri toplamaya ve karar alma sürecinin etkinliğini arttırmak için bu verileri bilgiye dönüştürmeye çalışmaktadır. Ortaya çıkan bu yeni piyasa şartları ekonomik uygulamalarda bilgiye duyulan gereksinimi arttırmıştır ve artık bilgi klasik üretim faktörlerinin yanında yeni bir üretim faktörü sayılmaktadır.

Bilginin artan önemiyle birlikte ortaya çıkan bu yeni ekonomik ortamda, işletmelerin pazarlama anlayışında köklü değişimler meydana gelmiştir. İşletmelerin bu piyasa şartlarında başarılı olabilmesi farklılık yaratabilmesine bağlıdır. Ancak teknolojik gelişmişliğin geldiği nokta ürün veya hizmetlerdeki kalite ve fiyat odaklı farklılaşmayı taklit edilebilir hale getirmiştir. Ürün ve hizmetlerin herhangi bir yönden taklidinin kolaylaştığı günümüz piyasa şartlarında rakiplere karşı üstünlük elde etmenin en etkili yolu müşterilerle bire bir ilişkiler kurarak onları özel hissettirmektir. Bu gerçeğin farkına varan günümüz işletmelerinin pazarlama stratejilerini değiştirmesiyle ürün, üretim ve satış odaklı geleneksel pazarlama anlayışı yerini bütün üretim kararlarını müşteri odaklı olarak belirleyen bir yönetim anlayışına bırakmıştır. Modern pazarlama anlayışı olarak adlandırılan bu yaklaşımda esas unsur üretim ve satış değil müşterilerin istek ve ihtiyaçlarıdır. Bu noktada işletmenin esas amacı müşteri ile etkili bir iletişim kurarak müşteriyi tanımak, müşterinin istek ve ihtiyaçlarını anlamak, müşteriden elde edilen değeri maksimize etmek, müşteriyi elde tutmak ve en nihayetinde müşteriyi sadık hale getirmektir. İşletmelerde ortaya çıkan bu yeni anlayış müşteri ilişkileri yönetimini ortaya çıkarmıştır.

Müşteri ilişkileri yönetimi yaklaşımının sağlıklı bir şekilde uygulanabilmesi işletmelerin müşterilerini olabildiğince iyi tanımasına bağlıdır. Bilgi teknolojilerindeki gelişmeler sayesinde artık verinin kolaylıkla depolanabilir ve

analiz edilebilir olması müşteri ilişkileri yönetiminin de etkinliğini arttırmaktadır. Bu noktada müşterilere dair çeşitli kaynaklardan toplanan veriler veri madenciliği yöntemleriyle analiz edilerek faydalı bilgilere dönüştürülmekte ve bu bilgiler müşterilerle etkin bir iletişim kurulması amacıyla kullanılmaktadır. Bilgin artık yeni bir ekonomik değer olarak karşımıza çıktığı ve müşterinin artık işletmelerin odağına yerleştiği günümüz piyasa şartlarında veri madenciliği yöntemleri müşteri ilişkileri yönetiminde yoğun olarak kullanılmaktadır.

Müşteri ilişkileri yönetimi temel olarak müşteri seçimi, müşteri edinme, müşteri koruma ve müşteri derinleştirme aşamalarından oluşmaktadır. Günümüzde söz konusu bu 4 aşamada veri madenciliği yöntemleri yoğun olarak kullanılmaktadır.

Müşteri ilişkileri yönetimi süreci hedef müşteri seçimiyle başlar. İşletme için en uygun ve en karlı müşteri grubunun seçilmesi ve işletme faaliyetlerinin bu müşteri grubuna uygun hale getirilmesi hem müşteri memnuniyeti hem de işletme faydası açısından önemlidir. Woo vd. (2005) çalışmalarında bir özel bankanın kredi kartı kullanan müşterilerini talep ve ihtiyaçları, tüketim alışkanlıkları ve risk oranları bakımından sınıflamış ve işletme için en uygun müşteri grubunu bulmuştur. Bagheri ve Tarokh (2014) çalışmalarında müşterilerin demografik özelliklerinin müşteri tüketim davranışlarına ve müşteri yaşam boyu değerine etkisini k-means kümeleme algoritmasını kullanarak belirlemiştir. Khajvand vd. (2011) çalışmalarında RFM analizinde kullanılan en yakın harcama, harcama sıklığı ve harcama meblağı gibi özelliklere satın alınan ürün çeşitliliği özelliğini de ekleyerek daha iyi sonuç elde edip edemeyeceklerini görmeye çalışmışlardır. K-means algoritmasıyla verileri öncelikle son harcama, harcama sıklığı ve harcama meblağı özelliklerine göre daha sonra ise satın alınan ürün çeşitliliği özelliğini de ekleyerek 4 özellik ile kümelemiştir. Son olarak DUNN indeksini kullanarak kümelerin başarısı ölçmüş ve 4. özelliğin herhangi bir etkisinin olmadığını tespit etmişlerdir.

Müşteri ilişkileri yönetiminin ikinci aşaması müşteri edinme aşamasıdır. Bu aşama işletme belirlenen karlı müşterileri çeşitli promosyonlarla ve kampanyalarla cezbetmeye veya bu müşterilere özelleştirilmiş pazarlama stratejileri uygulayarak taleplerine en uygun şekilde cevap vermeye çalışır. Ayrıca bu aşama rekabetten dolayı kaybedilen müşterileri geri kazanmayı içerir. Hadiji vd. (2014) yaptıkları

çalışmada ücretsiz online oyun hizmeti veren bir şirketin oyunu oynamayı bırakan müşterilerini inceleyerek mevcut müşterilerin ayrılmasını öngörebilecek bir model kurmuşlardır. Coussement ve De Bock (2013) çalışmalarında karar ağaçları, genelleştirilmiş katkı modeli algoritmalarını kullanarak bir online bahis sitesi müşterilerinin en yakın işlem, işlem sıklığı, harcadığı para, iki işlem arasında geçen süre, sitede kalma süresi gibi işlemsel verilerinin yanı sıra demografik verilerini de kullanarak müşterilerin ayrılmasına işaret eden faktörleri görmeye çalışmışlardır. Cheung vd. (2003) çalışmalarında bir film sitesindeki müşterilere daha önceki tercihlerine ve kendisiyle benzerlik gösteren müşterilerin tercihlerine göre uygun film tavsiyelerinde bulunmaya çalışmışlardır.

Müşteri ilişkileri yönetiminin üçüncü aşaması müşteri koruma aşamasıdır. Bu aşamada işletmeler eldeki müşterileri bire bir pazarlama, sadakat programları ve şikâyet yönetimi gibi stratejilerle korumaya çalışır. Ayrıca müşteri ayrılma analizi müşteri edinme aşamasında müşterinin geri kazanılması için önemli olduğu gibi bu aşamada da müşteri sadakatini koruma amacıyla önem taşımaktadır. Min vd. (2002) çalışmalarında Güney Kore’de faaliyet gösteren özel bir otelin müşterilerine ait giriş-çıkış kayıtları, konaklama ücreti gibi verilerin yanı sıra müşterilerle yapılan anketleri verilerini kullanarak C5.0 karar ağaçları algoritmasıyla müşteri profillemesi için kurallar oluşturmuştur. Larivie`re ve Poel (2005) yaptıkları çalışmada Avrupa’da faaliyet gösteren ve finansal hizmetler sunan bir firma müşterilerinin demografik ve geçmiş döneme ait harcama verilerinin müşterilerin tekrar satın alma davranışı göstermelerine ve karlılıklarına olan etkisini görmeye çalışmışlardır. Larivie`re ve Poel söz konusu amaç için random forest, regression forests, lineer regresyon ve lojistik regresyon yöntemlerini kullanmışlardır. Çalışma sonucunda random forest ve regression forest yöntemlerinin lineer ve lojistik regresyondan daha iyi sonuç gösterdiğini belirlemişlerdir. Ayrıca çalışma sonucunda müşterinin karlılığı ve tekrar satın alma davranışı üzerinde, geçmiş satın alma davranışının, demografik özelliklerine nazaran daha etkili olduğunu belirlemişlerdir. Mozer vd. (2000) çalışmalarında Amerika’da faaliyet gösteren bir telekomünikasyon şirketi müşterilerine ait fatura bilgileri, müşterilerin görüşme bilgileri, çağrı merkezlerine gelen şikâyetler ve çözülme durumu, rakip müşterilerin teklifleri ve müşterilerin

demografik verileri kullanmıştır. Bu bilgiler ışığında müşterilerin tatminsizlik nedenlerini belirlemeye çalışmış ve ayrılma ihtimallerini öngörmeye çalışmışlardır.

Müşteri ilişkileri yönetiminin dördüncü ve son aşaması müşteri derinleştirme aşamasıdır. Bu aşamada işletmeler sadakatini kazandıkları müşterilere upselling ve çapraz satış stratejileri uygulayarak müşterinin toplam harcamaları içerisindeki cüzdandan paylarını arttırmaya çalışırlar. Changchien vd. (2004) çalışmalarında demografik özellikleri ve gelir düzeylerine göre gruplandıkları müşterilere kişiselleştirilmiş promosyonlar ve yan ve üst ürün tavsiyelerinde bir sistem geliştirmişlerdir. Verhoef ve Donkers (2001) sigortacılık alanında hizmet veren bir firma müşterileri üzerine yaptıkları çalışmada müşterilerin sosyo-demografik bilgilerinin yanı sıra firmanın 12 farklı sigorta paketinden hangilerini tercih ettikleri bilgisini esas alarak söz konusu müşterilerin işletme için mevcut ve potansiyel değerini hesaplamışlardır.

Bu tez çalışmasının amacı firmaların veri ambarlarındaki veriler içerisinde saklı bulunan ve firmaların karar verme süreçlerinde kullanabilecekleri enformasyonların Veri Madenciliği algoritmalarının yardımı ile ortaya çıkarılmasıdır.

Bu tez giriş bölümü dışında 4 bölümden oluşmaktadır. 2. Bölümde müşteri ilişkileri yönetiminin tarihsel gelişimi, tanımı, önemi ve faydaları gibi konular üzerinde durulmuştur. 3. Bölümde veri madenciliğini de kapsayan bir süreç olan iş zekâsı kavramından ve uygulama adımlarından bahsedilmiştir. 4. Bölümde veri madenciliği ve bu tezde kullanılacak veri madenciliği metotları anlatılmıştır. Son olarak 5. ve son bölümde ise RFM ve müşteri kaçma analizleri yapılmıştır. RFM analizi için Online Retail (“Online Retail Dataset”) isimli hazır veri seti kullanılmıştır. Bu veri seti 4322 müşterinin 2010-2011 yılları arasındaki 13 aylık süreçte tükettiği ürünlere ait bilgiler içermektedir. Bu müşterilerin tüketim alışkanlıklarının belirlenmesi için RFM analizi yöntemi kullanılmıştır. Bu analizin etkinliğinin artırılması için K-Means ve Öz Örgütlenmeli ağlar kullanılmış ve bu kümeleme analizlerinin hangisinden daha iyi sonuç alındığı tespit edilmiştir. Müşteri kaçma analizi için ise Churn in Telecom’s dataset (“Churn in Telecom’s dataset”) isimli hazır veri setinden faydalanılmıştır. Bu veri seti 4667 müşteriden ve her

müşteri için 19 adet işlem kaydına ait özellikler ile müşterinin terk edip terk etmediğine dair sınıf etiketlerinden oluşmaktadır. Sadık ya da terk eden müşterileri sınıflamada yapay sinir ağları, naive-bayes, destek vektör makineler ve lineer regresyon kullanılmıştır. Bunlardan farklı olarak veri setindeki özniteliklerin sınıflamaya olan etkilerinin sıralanması için ki-kare, bilgi kazancı, kazanım oranı ve Relief-F olmak üzere 4 farklı metrik aynı veri seti üzerine uygulanmıştır.

2. MÜŞTERİ İLİŞKİLERİ YÖNETİMİ

2.1. Ürün Odaklı Yaklaşımdan Müşteri İlişkileri Yönetimine

İşletmeler üretim fonksiyonlarını bir araya getirerek mal, hizmet veya fikir üreten kuruluşlardır. Ancak bu kuruluşlar üretim çıktılarını kar ya da herhangi başka bir fayda elde etmek amacıyla üretirler. Kar ya da başka faydalar elde etmek amacıyla kurulan işletmelerin hayatta kalabilmek için tüketicilere de uygun faydayı sağlamaları gerekmektedir. Yani bir malın sadece üretilmesi tüketiciyi cezbetmesi için yeterli değildir. Ürünün doğru zamanda, doğru miktarda, doğru yerde tüketiciye sunulması gerekmektedir. İşletmeler üretim fonksiyonuyla hammadde ve diğer girdileri müşteri ihtiyaçlarına uygun ürüne dönüştürerek şekil faydası yaratırken, pazarlama fonksiyonu ile yer, zaman ve mülkiyet faydası sağlar. Pazarlama üretimle beraber işletmenin iki önemli fonksiyonudur ve birbirini tamamlayan eylemlerdir (Erdoğan 2014: 10).

Literatürde pazarlamanın tanımıyla ilgili iki temel yaklaşım vardır: bunlardan birisi pazarlamayı bir işletme fonksiyonu olarak görürken diğeri pazarlamayı bir işletmecilik felsefesi olarak tanımlamaktadır (Erdoğan, 2014: 10). Pazarlamaya dair bu iki yaklaşım ve pazarlamayla ilgili tanımların pazarlamayı sadece tek bir yönüyle açıklaması üzerinde uzlaşmış tek bir tanım yapmayı zorlaştırmaktadır. Pazarlama ile ilgili net ve kesin bir tanım ortaya koymak mümkün olmazken farklı otoriteler tarafından yapılan tanımlar vardır. Kotler ve Keller (2012: 3)'e göre pazarlama hedef pazarları seçmeyi ve elde etmeyi, korumayı; Üstün müşteri değeri yaratarak, sunarak ve müşteriye ulaştırarak müşterilerini arttırmayı amaçlayan bir bilim ve sanattır. Amerikan Pazarlama Birliği ("AMA - American Marketing Association," 2013) pazarlamayı bireylerin ve toplumların taleplerine uygun değişimi ve arzı sağlamak üzere ürünleri üretilmesini, taşınmasını, ürünlerle ilgili tüketicilerin bilgilendirilmesini kapsayan bir kurumsal aktivite ya da süreç olarak tanımlamıştır. Nakip vd. (2012: 1)'ne göre ise "*pazarlama üretimden önce başlayıp, tüketimden sonra da devam edebilen, müşteri ihtiyaçlarını tatmin etmek ve paydaşlarına bir değer sunmak üzere ürünlerin geliştirilmesini, bu ürünlerin fiyatlandırılmasını, dağıtılmasını ve tutundurulmasını içeren ve sosyal sorumluluğu olan bir süreçtir*".

Pazarlamayla ilgili yapılan tanımların birçoğu pazarlamayı sadece kar maksatlı bir aktivite olarak tanımlarken son yıllarda kâr amacı taşımayan satış ve tutundurma faaliyetleri de birer pazarlama faaliyeti olarak tanıma dahil edilmiştir.

Pazarlama disiplininin tarihi çok eskilere dayansa da pazarlama yaklaşımı tarihin belirgin dönemlerinde tamamen değişmiştir. Pazarlama yaklaşımındaki bu değişimler belirli başlıklar altında toplandığında iki temel anlayış ortaya çıkmaktadır. Bunlar: geleneksel pazarlama anlayışı ve modern pazarlama anlayışıdır.

2.1.1. Geleneksel Pazarlama Anlayışı

Geleneksel pazarlama anlayışı piyasadaki üretici sayısının, dolayısıyla arzın, tüketici talebini karşılamada yetersiz kaldığı ülkelerde veya dönemlerde etkili olan pazarlama anlayışıdır. Üreticilerin tüketiciler karşısında avantajlı konumda olduğu bu tip piyasalarda işletmeler ne üretirsem onu satarım anlayışıyla müşteri talep ve ihtiyaçlarını ikinci plana atarak üretim ve satış süreçlerini odağına almaktadır. Geleneksel pazarlama yaklaşımı üç başlık altında incelenmektedir. Bunlar üretim yaklaşımı, ürün yaklaşımı ve satış yaklaşımıdır.

Üretim yaklaşımı sanayi devriminden sonra ortaya çıkmış ve etkisini Amerika'daki büyük buhrana kadar devam ettirmiştir. Bu dönemin en önemli özelliği talebin yüksek, arzın az olmasıdır. Talep sıkıntısı yaşamayan işletmeler için en büyük problem üretim maliyetleri ve üretim sürecinin meşakkatli olmasıdır. İşletmeler talep sıkıntısı yaşamadıkları için müşteri talep ve isteklerini ikinci plana iterek odaklarını üretim süreçlerine yoğunlaştırmışlardır. Uzmanlar arz talep dengesizliği nedeniyle bu dönemi "*ne üretirsem onu satarım*" veya "*her arz kendi talebini yaratır*" sözleriyle özetlemektedir. Her ürettiğini satabilen işletmeler için verimliliğin kaynağı üretimdir. Bu anlayış kalkınmış ülkelerde büyük buhrandan sonra terkedildiği görülse de geri kalmış ve kalkınmakta olan ülkelerde hala bu anlayışın etkileri görülmektedir. Türkiye'de bu anlayışın 24 Ocak 1980 kararlarına kadar hâkim olduğu görülmektedir. Bu kararlardan sonra bir kısım üreticilerin bu anlayışı terk etmek zorunda kaldıkları ancak yerel ölçekte iş yapan bir çok işletmenin bu anlayışı devam ettirdiği görülmüştür (Nakip, Varinli, Gülmez 2012: 13).

Üreticiler sanayi devrimi sonrasında, arzın azlığından kaynaklı rekabetsiz ortamda, üretim sürecine odaklanarak karlılıklarını arttırmaya çalışmışlardır. Ancak

zamanla pazara yeni rakiplerin girmesi ve özellikle ikinci dünya savaşı sonrası oluşan refah ortamında teşebbüsler artmıştır. Arzın talebe yaklaştığı bu dönemde pazardaki rekabet artmış ve üreticiler müşterilere daha fazla odaklanmak zorunda bırakmıştır. Müşteri taleplerinin hala değer görmediği bu dönemde işletmeler tüketicilerin sadece kalite ve ürün performansına değer verdiğini varsayarak ürünlerinde kaliteyi arttırmayı hedeflemişlerdir.

Zaman içerisinde dünyada üretimin artmaya başlaması ile birlikte rekabette giderek yoğunlaşmaya başlamıştır. İkinci dünya savaşı sonrası oluşan refah ortamıyla beraber ilk defa arz talebi geçmiş ve işletmeler satamadıkları malları stoklamak zorunda kalmışlardır. Oluşan talep fazlası ürünler işletmeleri satış süreçlerine odaklanmak zorunda bıraktı. Bu dönemde satışları arttırmak için işletmelerin reklamlara ağırlık verdikleri görülmektedir (Nakip, Varinli, Gülmez 2012: 15). Bu dönemde işletmeler müşterilerle uzun soluklu ilişkiler yerine kısa dönemli karlılıklara önem vermişlerdir. Bu nedenle dönem reklamları yanıltıcı ve aldatıcı olabilmektedir.

2.1.2. Modern Pazarlama Anlayışı

Satış yaklaşımından beklenen faydanın sağlanamaması, artan rekabet ve devletlerin müşterileri korumaya odaklı hukuksal düzenlemeleri zamanla işletmelerin ürün odaklı yaklaşımını müşteri odaklı yaklaşıma dönüştürmüştür. Devrim niteliğindeki bu değişimle beraber modern pazarlama anlayışı ortaya çıkmıştır. Durmaz'a göre (2006) modern pazarlama tüketicilerin istek ve arzularını esas alan ve onları memnun etmeyi amaçlayarak mal, hizmet ve fikir üretmek, ayrıca ürünün tanıtımı, tüketiciye ulaştırılması, satış sonrası garanti ve teknik destekle zenginleştirerek müşteri bağlılığının sağlamak yoluyla kar etmek olarak tanımlamıştır. Bu dönemde tüketici ve müşteri arasındaki fark netleşmiş ve işletmeler tüketicileri müşterileri haline getirmeyi hedeflemeye odaklanmıştır. Satış odaklı yaklaşım müşterilerle kısa dönemli ve sadece kar odaklı ilişkiyi hedeflemekle beraber değişen piyasa şartları nedeniyle müşterilerle uzun vadeli ilişkiler geliştirmek daha cazip hale gelmeye başlamıştır. Müşterilerle uzun vadeli ilişkilerin geliştirilmesi işletmelerin müşterileri daha iyi tanıması neticesinde müşterilerin taleplerine daha etkili cevap vermesini sağlarken müşterilerinde işletmeye sadakatini arttırmaktadır.

Bu sayede işletmeler sadık müşterilere sahip olmanın artlarından faydalanırken müşterilerse taleplerine etkili bir şekilde cevap alabilmektedir.

Günümüzde müşterilerin ürün tercihini sadece ürünün kalitesi ve fiyatı etkilemiyor. Üreticilerin satış sonrası hizmetleri, müşteri temsilciliğindeki ya da satış şubesindeki çalışanın tavrı, arkadaş çevresinin ya da son dönemde etkisi giderek artan sosyal mecraların ürüne veya firmaya dair yorumları tüketicilerin tüketim tercihlerine etki etmektedir. Bu noktada rekabet temelinde bir üstünlük yaratması amacıyla kullanılan bütünleşik pazarlama iletişimi yaklaşımı dayanak noktasını iki temel strateji üzerine oturtmaktadır. Bunlardan ilki imaj ve marka yönetimi bir diğeri ise müşteri ilişkileri yönetimidir (Başok Yurdakul 2003). Firmaların imaj yönetimi çalışmalarının en önemli araçlarından birisi ise hiç şüphesiz sosyal sorumluluk projeleridir.

2.1.3. İşletme Odağı Olarak Müşteri

Bilgi ve iletişim teknolojilerinde yaşanan gelişmeler tüketicilerin haber alma kaynaklarını ve iletişim kanallarını değiştirmektedir. İşletmelerin tüketicilere ulaşmak için kullandıkları kanalların değişmesi ve alternatiflerin artması geleneksel pazarlama uygulamalarının etkinliğini azaltmış ya da tamamen kullanışsız hale getirmiştir. Bu durumsa yeni pazarlama uygulamalarının geliştirilmesini zorunlu kılmıştır. Ayrıca bu yeni uygulamaların ortaya çıkmasında teknolojik gelişmelerin yanı sıra sosyolojik ve politik faktörlerinde etkisi vardır. Bu faktörlere örnek olarak; küreselleşme, serbest piyasa ekonomisine geçiş ve bu nedenle piyasada artan rekabet, teknolojik araçlar sayesinde tüketicilerin alternatiflerinin artması gösterilebilir (Özgüner 2014: 311).

Günümüz piyasa şartlarında artık işletmelerin odağında artık müşteriler vardır. Bu nedenle işletmeler için müşterilere ait her türlü bilgi büyük önem taşımakta ve müşteri ile işletme arasında karşılıklı faydaya dayalı ve sürekli ilişkilerin kurulmasında hayati önem taşımaktadır.

2.2. Müşteri ve Tüketici Kavramları

Müşteri odağının üründen müşteriye kaydığı günümüzde işletmeler için müşteri kıt kaynaktır. Bu noktada müşteri talep ve isteklerinin doğru belirlenmesi ve müşteri tatmininin artırılması müşteri kazanma ve kazanılan müşteriye elde tutmak açısından önemlidir. Bu noktada öncelikle tüketim, tüketici ve müşteri kavramlarının açıklanması yerinde olacaktır.

Tüketim en basit şekilde kişisel, ailevi ya da kurumsal ihtiyaçların karşılanması amacıyla mal veya hizmet satın alınmasıdır. Tüketici ise bu mal ve/veya hizmet alımını yapmaya istekli kişi ya da kurumdur. Bişkin'e (2011: 48) göre; satın alma, sadece günlük ihtiyaçların karşılanmasını değil, otomobil, ev eşyası veya teknolojik ürün alımından, farklı amaçlarla sigorta yaptırmaya, ev kiralamaktan, çeşitli ihtiyaçlar için kredi almaya kadar geniş bir alanı kapsar.

Müşteri ise ihtiyaç duyduğu bir ürünü ya da hizmeti belirli bir marka ya da firmadan satın alan kişi ya da kurumdur. Bu tanımlara göre her müşteri bir tüketicidir ancak her tüketici bir müşteri değildir. Tüketicinin müşteri sıfatını alabilmesi için bir firmanın ürününü tercih ediyor olması gereklidir. İşletmelerin asli amacı tüketicileri önce müşterileri haline getirmek daha sonra ise sadakatlerini kazanmaktır.

Bayraktar'dan (1994) aktaran Bişkin'e (2011: 52) göre tüketici tipleri aşağıdaki gibi incelenebilir:

- Alışkanlıklarından dolayı, sadece belirli markaları tercih eden tüketici,
- Rasyonel seçim yapma eğilimi yüksek olan tüketici,
- Fiyat odaklı olarak ürün tercihinde bulunan tüketici,
- Ani karar veren, fiziksel çekiciliğe kapılan ve markaya daha az önem veren tüketici,
- Ürün tercihlerinde duygusal davranan tüketici,
- Yeni ve henüz davranışları belirli olmayan tüketici.

Toplam kalite yönetiminin en önemli unsurlarından birisi müşteri odaklılıktır ve günümüz rekabet şartlarında işletmeler için çok büyük öneme sahiptir. Toplam

kalite yönetimi, müşteriye dış ve iç müşteri olmak üzere iki başlık altında incelemektedir(Çatı, Koçoğlu 2010: 231).

2.2.1. İç ve Dış Müşteri

Toplam kalite yönetimi müşteriye iç müşteri ve dış müşteri olmak üzere ikiye ayırmaktadır. Burada dış müşteri işletmenin nihai tüketicilerini ifade ederken İç müşteri işletmenin üretim sürecine dahil olan çalışanları ve departmanları ifade eder (Kağnıcıoğlu 2002). İç müşteri yaklaşımına göre üretim sürecinde bulunan her bir birim kendisinden önceki birimin müşterisi iken kendisinden sonraki müşterinin tedarikçisidir. Bu zincir ile her bir birimin kendi müşterisini memnun etmesi bu şekilde nihai ürünün kalitesinin artırılması ve nihai (dış) müşterinin memnuniyetinin artırılması amaçlanmaktadır.

2.2.2. Aktif ve Pasif Müşteri

Pasif müşteri işletmenin sunduğu ürün veya hizmetten sadece küçük bir kısmını tercih eden ya da yılın sadece kısıtlı bir döneminde işletmeden mal veya hizmet alımında bulunan müşteridir. Bu tür müşterilerin işletmenin gelişimine katkısı çok düşüktür. Aktif müşteri ise işletmenin sunduğu ürün ya da hizmetlerden yoğun bir şekilde faydalanan, aldığı ürün ya da hizmet neticesinde işletmeye olumlu ya da olumsuz geri bildirimlerde bulunarak işletme süreçlerinin gelişmesine katkıda bulunan müşterilerdir (Ulucan 2016: 6).

2.3. Müşteri İlişkileri Yönetiminin Ortaya Çıkışı ve Gelişimi

Pazarlama anlayışı yıllar içerisinde radikal bir değişim geçirmiştir. Başlarda pazarda talebin yüksek, arzınsa teknolojik gerilik ve pazardaki üretici azlığı sebebiyle düşük olması işletmeleri müşteriler karşısında avantajlı bir konuma taşımıştı. Üreticilerin ürettikleri ürüne müşteri bulma sorunları yoktu. Bu nedenle işletmelerde “*ne üretirsem onu satarım yeter ki üreteyim*” anlayışı hakimdi. 2. Dünya savaşı sonrasında oluşan refah ortamı iktisadi girişimleri arttırırken arz talep dengesini müşteri lehine çevirdi. Zamanla hem teknolojiye ki gelişmeler hem de pazara giren yeni üreticiler arzı arttırırken müşterilere de yeni alternatifler sunmaya başladılar. Rekabetin gün geçtikçe daha fazla hissedilmeye başlandığı bir ortamda

geleneksel pazarlama stratejileri önem kazanmaya başladı. Bunlar arasında reklam, pazar araştırması, kurum imajı, kurumsal iletişim, fiyatlandırma, promosyon gibi kolay anlaşılıp uygulanabilen kavramlar bu dönemde popülerlik kazandı (Tezcanlar 2007: 7).

Zamanla teknolojiye yaşanan gelişmeler mal ve hizmetlerin taklidini kolaylaştırmaya başladı. Bu durum tüketicilere alternatifler sunarken işletmelerin rakiplerinden farklılaşmasını zorlaştırdı. Bu durumsa müşterileri sadakatinin ve marka bağlılığının azalmasına neden oldu. Başlarda fayda sağlayan reklam, promosyon gibi stratejiler zamanla etkinliğini kaybetmeye ve yetersiz kalmaya başladı. Bu durumda işletmeler üretim odaklarını daha fazla müşteriye çevirmek zorunda kaldı. Bu dönemde 'kişiyeye özel pazarlama', 'seçkin müşteriye yönelme' ve 'müşteri odaklılık' gibi kavramlar ortaya çıkmaya başladı. İşletmeler bu şekilde üründe yapamadıkları farklılaşmayı müşterileri ile etkili iletişim kurarak yapmaya başladılar.

90'lara gelindiğinde bilgisayar teknolojilerinin yaygınlaşması ile MİY uygulamaları geliştirilmeye başlandı. Ancak bu dönemde geliştirilen programlar her bir işletme faaliyeti için farklıydı. Bu nedenle işletmeler her bir departman için farklı birer program satın almak zorunda kalıyorlardı. Bu durumsa işletme süreçleri arasındaki entegrasyonu ortadan kaldırıyor müşterilerle ilgili daha kapsamlı bir analiz yapılmasını güçleştiriyordu. 90'ların ikinci yarısında işletmelerden geri dönüşler sayesinde entegre uygulamalar geliştirilmeye başlandı.

2.4. Müşteri İlişkileri Yönetiminin Tanımı ve Önemi

İşletme kar ya da başka bir fayda elde etmek amacıyla üretim faktörlerini bir araya getiren iktisadi yapı olarak tanımlanabilir. Kar ya da fayda odaklı olmaları yönüyle işletmeler için en önemli faktör şüphesiz müşterilerdir. Kotler ve Keller (2012: 135) işletmelerin elindeki tek değer mevcut veya gelecekteki müşterilerin yaratacağı değerler olduğunu söylemektedir. İşletmeler geliştirdikleri stratejilerle diğer işletmelerden farklılaşmaya ve müşteriyi kendilerine bağlamaya çalışmaktadır. Bu noktada işletmeler rekabet üstünlüğü sağlamak amacıyla müşteri talep ve isteklerini yakından takip ederek bu talep ve isteklere uygun stratejiler geliştirme

yoluna gitmişlerdir. Bu durumsa Müşteri İlişkileri Yönetimi (MİY) kavramını ortaya çıkarmıştır.

Duran'a (2011) göre MİY'in ortaya çıkış nedenleri şöyle sıralanabilir;

- Kitlesel pazarlamanın gittikçe daha maliyetli bir yöntem haline gelmesi,
- Müşteri payının Pazar payından daha önemli hale gelmesi,
- Müşteri memnuniyeti ve müşteri sadakati kavramlarının bilinirliğinin ve öneminin artması,
- Mevcut müşterinin değerinin anlaşılması ve bu müşteriyi elde tutma çabalarına ihtiyaç duyulması,
- Bire-bir pazarlamanın önem kazanmasıyla birlikte her müşteriyle birebir ilgilenme mecburiyeti,
- Yoğun rekabet ortamı.

Müşteri ilişkileri yönetimi ile ilgili literatürde birçok tanım bulunmakla birlikte tanımlardan bazıları şunlardır; bir yönetim stratejisi olan MİY, müşterilerin istek ve beklentilerini anlamak, bunları zamanında ve doğru şekilde tespit etmek ve müşterilerin tutum ve davranışlarını etkileyecek şekilde etkileşimli iletişim kurmaktır (Özmen 2012: 57). MİY müşteri sadakatini arttırmak amacıyla müşteriler hakkındaki detaylı bilgileri kullanarak müşterilerle etkin bir ilişki geliştirme sürecidir (Kotler, Keller 2012: 135). MİY pazarlama, satış, müşteri hizmetleri ve tedarik zinciri fonksiyonlarını entegre ederek müşteri edinme, mevcut müşteriyi elde tutma ve müşteri sadakatini arttırmayı amaçlayan ve hem müşteriye hem de işletmeye fayda sağlayan bir süreçtir (Parvatiyar, Sheth 2001).

Bu tanımlar ışığında MİY'in stratejilerinin hedeflerini aşağıdaki şekilde kategorize edebiliriz:

- Müşteriyi Elde Etme,
- Mevcut Müşteriyi Elde Tutmak,
- Müşteri Sadakatini Arttırmak,
- Müşteriden Elde Edilen Karı Arttırmak.

MİY ilişkileri yönetimi müşterilerle karşılıklı faydaya dayanan ilişkiler kurmayı hedefler. Ancak esas amacı kar etmek olan işletmeler için bütün müşteriler aynı oranda karlı değildir. Budhwani'den (2002) aktaran Persson'a (2003: 37) göre; her müşteri aynı oranda değerli değildir. Hatta bazı müşteriler işletme için zararlıdır. Bu müşteriler işletmenin zaman ve enerjisini harcarken işletmeye yeterli kar sunmazlar. Pareto ilkesine göre birçok durum için sonuçların kabaca %80'i nedenlerin %20'sinden kaynaklanmaktadır. İş dünyasında ise bir şirketin yaptığı işlerin en verimli yüzde 20'si, şirket gelirlerinin yüzde 80'ini oluşturur ("80/20 Kuralı Nasıl Şirket Kurtardı" 2013). Bu ilke çerçevesinde işletmenin elde ettiği yıllık karın %80'lik kısmının %20'lik bir müşteri grubundan elde edildiği söylenebilir. İşletme karlılığı açısından büyük öneme sahip bu %20'lik kesimdeki müşterilerin doğru belirlenmesi ve bu müşterilerin talep ve isteklerine en etkin şekilde cevap verilmesi işletmenin geleceği açısından hayati öneme sahiptir. MİY müşteri memnuniyetinin yanı sıra en karlı müşterilerin belirlenerek bu müşterilere yönelik elde tutma stratejilerinin geliştirilmesini de amaçlamaktadır (Xu, Yen, Lin, Chou 2002).

Müşteri ilişkileri yönetimi bir bilgi teknolojisi sistemi değil bir yönetim felsefesidir. Ancak teknolojinin MİY süreçlerine dahil edilmesi verimliliği arttırmaktadır. Bose'e (2002) göre MİY müşteri taleplerini karşılamak amacıyla iş süreçlerinin teknolojiyle entegre edilmesini gerektirmektedir. MİY stratejisinin verimli bir şekilde uygulanması için işletme veri ambarı, web siteleri, intranet/extranet, telefonla destek sistemi, muhasebe, satış, pazarlama ve üretim gibi teknolojilerin ve iş süreçlerinin entegrasyonu gereklidir.

2.5. Küreselleşme ve Müşteri İlişkileri Yönetimi

Küreselleşme kitle iletişim araçlarının yaygınlaşmasıyla başlayan ve yeni kitle iletişim araçlarının icadı ve yaygınlaşmasıyla etkinliğini giderek arttıran bir olgudur. Küreselleşme yeni bir kavram olmamakla birlikte üzerine çok yazılıp çizilmiş bir olgu olarak karşımıza çıkmaktadır. Küreselleşmeyle ilgili çok şey söylemiş ve yazılıp çizilmiş olmasına rağmen hala irdelenecek birçok yönü vardır. Bunun en büyük nedeni küreselleşme kavramının hala devam eden ve bütün

insanlığın günlük hayatını etkilemeyen bir olgu olmasıdır. Küreselleşme etkisini birçok alanda gösterirken bizim burada irdelemek istediğimiz küreselleşmenin iş dünyasına etkisidir.

Küreselleşme ve içinde bulunduğumuz bilgi çağının en büyük sonuçlarından birisi sınırların ve mesafelerin tamamen ortadan kalkmasıdır. Bunun sonucu olarak işletmeler her ülkede ürün üretebilmekte ve satış yapabilmektedir. Hatta çok uluslu işletmeler hammaddelere ve pazara yakınlık, işçilik maliyetleri gibi faktörleri göz önünde bulundurarak üretim tesislerini dünyanın farklı noktalarına yayabilmektedir. İşletmeler bu yolla hem ülkelerdeki ekonomik ve politik riskleri kendileri için minimize etmekte hem de belirgin bir pazara bağımlı kalmadan dünyaya açılabilen yeni pazarlara nüfuz edebilmektedir. Bu şekilde rekabet üstünlüğü sağlamayı ve karlılığı artırmayı amaçlayan işletmeler için bu politikanın olumsuz yönü ise rekabet yoğunluğunun artmasıdır. İşletmelerin dünya pazarlarına açılması tüketiciler için yerel üreticilere olan bağımlılığı azaltmakta ve alternatifleri arttırmaktadır. Tüketicilerin ihtiyaçlarını karşılamaya yönelik alternatiflerinin artması ise üreticiler açısından müşteriyi elde tutmayı zorlaştırmaktadır. Bunun sonucu olarak bir dönem hâkim olan ürün odaklı anlayış yerini müşteri odaklı anlayışa bırakmıştır.

MİY'in tarihi gelişimini incelediğimiz zaman 90'ların başında ortaya çıktığını, 90'ların sonu ile 2000'lerin başında popülerliğinin giderek arttığını görürüz. Bunun en büyük nedeni küreselleşme kaynaklı rekabet artışıdır. Müşterinin talep ve beklentilerinin daha çok önem kazandığı bu dönemde, işletmeler bir yandan yeni müşteriler kazanmaya diğer yandansa eldeki müşteriyi tutmaya çalışmaktadır. Eldeki müşteriyi tutmanın maliyeti yeni müşteri kazanmaktan daha fazla olduğu için, işletmeler açısından mevcut müşterilerle iyi ilişkiler kurulması ve bu müşterilerin ihtiyaç ve taleplerindeki değişimlerin yakından takip edilmesi hayati önem taşımaktadır. Müşteriyi yakından takip etme, anlama ve iyi ilişkiler geliştirme ihtiyacı Müşteri İlişkileri Yönetimi kavramını ortaya çıkarmıştır. Zengin ve Ulama'ya (2015: 398) göre Müşteri İlişkileri Yönetimi küresel dönem pazarlama anlayışının bir sonucudur.

2.6. Müşteri İlişkileri Yönetiminin Faydaları

Rekabetin yoğunlaştığı günümüz piyasa şartlarında işletmeler için en önemli değer şüphesiz müşterilerdir. Müşterilerin elde edilmesi ve daha önemlisi müşteri sadakatinin sağlanması işletmeler için ciddi bir rekabet üstünlüğü sağlamaktadır. Müşteri ilişkileri yönetimi müşteri elde etmeyi, mevcut müşteriyi elde tutmayı, müşteri sadakati sağlamayı ve müşteri karlılığını arttırarak işletmeye avantajlar sağlamayı hedeflemektedir.

MİY'in işletmelere sağladığı faydalar aşağıda sıralanmıştır:

Müşteri Güven ve Sadakatinin Sağlanması: Müşteri ilişkileri yönetimi politikaları sayesinde müşteri talep ve istekleri yakından takip edilerek müşteri güven ve sadakati sağlanmaya çalışılmaktadır.

Müşteri Yaşam Boyu Değerinin Artması: Müşteri yaşam boyu değeri müşterinin işletmeyle ilişkide olduğu sürenin harcadığı toplam parayla çarpılması sonu bulunan bir değerdir. Müşteri ilişkileri yönetimi işletme için karlılığı yüksek müşterileri belirleyerek bu müşterilerle ilişkileri iyi tutmayı ve yaşam boyu değerini arttırmayı amaçlar.

Yüksek Müşteri Karlılığı: Tatmini artan müşterinin çapraz satış, yukarı satış ve devam eden satışlarla müşterinin firmaya aktardığı para artmaktadır. Ayrıca müşteri tavsiyeyle yeni müşterileri firmaya yönlendirmektedir.

Maliyetlerin Azalması: İşletmeler için yeni müşteri elde etmek mevcut müşterileri elde etmekten daha maliyetlidir. Müşterilerle iyi ilişkiler kurulması müşteri bağlılığını arttırırken tutundurma, reklam maliyetlerini azaltmaktadır.

Tutundurmaya Yardımcı Olma: İşletmeler için en az maliyetli ve en etkili pazarlama aracı WOM (Word of Mouth) 'dur. Türkçeye ağızdan ağıza pazarlama olarak çevrilen WOM müşterilerin memnuniyetlerini ya da memnuniyetsizliklerini çevresindekilerle paylaşarak işletmenin olumlu ya da olumsuz reklamını yapmasıdır. Tüketicilerin ürün tercihinde tanıdıklarının tavsiyelerinin rolü çok büyüktür. Tüketiciler sadakat gösterdikleri ürünü kendi çevrelerine tavsiye ederek söz konusu ürünün tutundurma faaliyetlerine olumlu etkide bulunmaktadır. Ayrıca bu durum işletmenin pazardaki imajına olumlu etki eder.

Müşteri ve Pazar Farkındalığı: MİY mevcut ve potansiyel müşterilerini tanımak için bilgi teknolojileri yoğun bir şekilde kullanır. Müşterilere verilen mağaza kartları ile hem müşterilere indirimler yapılarak promosyon faaliyetleri yürütülebilmekte hem de müşteri gruplarının tüketim alışkanlıkları izlenebilmektedir. Ayrıca çağrı merkezleri müşterilerle iletişimde kalarak tüketilen ürünlerle ilgili geri dönüşler sağlayabilmektedir.

Rakiplerden Farklılaşma: Teknolojik gelişmeler sayesinde yeterli sermayeye sahip her işletme herhangi bir ürünü yüksek kaliteyle üretebilmektedir. Bu da pazardaki firmaların ürün bazında farklılaşmasını imkânsız hale getirmektedir. MİY sayesinde işletmeler ürün üzerinden yapamadıkları farklılaşma faaliyetlerini müşterilerle karşılıklı olumlu ilişkiler kurarak yapmaktadır.

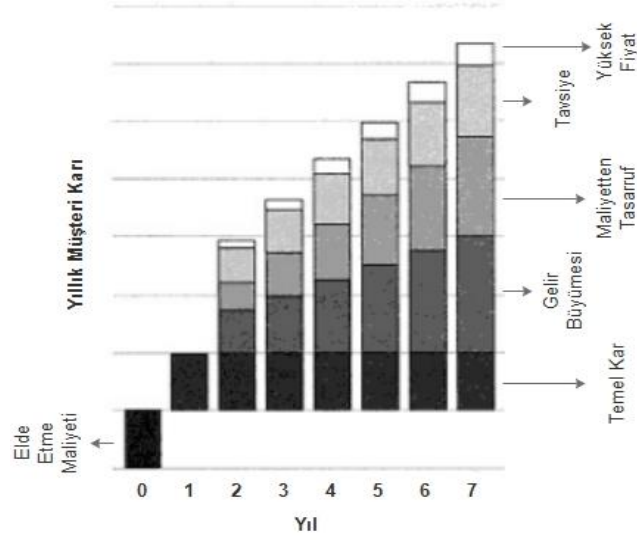
2.7. Müşteri İlişkileri Yönetiminin Unsurları

2.7.1. Müşteri Sadakati

Müşteri ilişkileri yönetiminin nihai hedeflerinden birisi müşteri sadakatini sağlamaktır. Müşteri sadakati ile ilgili birçok tanım bulunmakla birlikte müşteri sadakati Oyman'a (2002) göre: bir markaya ya da işletmeye duyulan bağlılık olarak tanımlanabilir. Dick ve Basu'ya (1994) göre: tekrar eden şekilde uzun bir dönem aralığında tekrar eden harcama yapma taahhüdüdür.

Müşteri sadakati üzerine yapılan çalışmalar incelendiğinde araştırmacıların genel olarak iki nokta üzerinde durdukları görülmektedir. Bunlardan birincisi sadakatin davranışsal yönü, ikincisi sadakatin tutumsal yönüdür (Çatı, Koçoğlu 2010: 239). Davranışsal yaklaşım müşteri sadakatini satın alma miktarı, satın alma sıklığı veya satın alma ihtimali gibi davranışsal ölçümler kullanarak değerlendirilir (Chang, Chen 2007). Tepeci'ye (1999) göre: davranışsal yaklaşımın problemi tekrar eden tüketim davranışlarının tüketicinin ürüne dair psikolojik taahhüdünü ifade etmemektedir. Örneğin bir müşterinin bir marketi seçmesinin nedeni marketin konumunun evine yakın olması olabilir bu durumda evin yakınına yeni bir market açılması durumunda müşteri eğer yeni marketin fiyat, kalite gibi değerleri daha iyi olursa orayı tercih edebilir. Tutumsal yaklaşım ise duygusal ve psikolojik bağlılığın sadakatin en net göstergesi olduğunu savunmaktadır (Bowen, Chen, 2001). Bu

yaklaşımına göre müşterinin alışveriş sıklığı çok düşük olsa bile işletme hakkında olumlu konuşması, çevresine tavsiye etmesi, başkalarını işletmeden alışveriş yapmaya ikna etmesi işletmeler için çok önemlidir. Çünkü insanların bir malı tüketmeye ikna olmalarında çevrelerindeki insanların tavsiyelerinin reklamlardan daha etkili olduğu kabul edilmiş bir gerçektir.



Şekil 1. Sadık Müşteri Neden Önemlidir (Reichheld, Teal 1996).

Reichheld ve Teal sadık müşterinin avantajlarını şöyle açıklamıştır (1996):

Elde Etme Maliyeti: İşletmeler için yeni müşteriyi elde etmek mevcut müşteriyi elde tutmaktan daha maliyetlidir. Yeni müşteriyi elde etme sürecinde işletmeler reklam, kişiye özel promosyon, satış ekibi oluşturma gibi maliyetler altına girmektedir.

Temel Kar: İşletmeyle müşteri olarak adlandırılmayacak kadar az etkileşimde bulunsa bile işletmeden mal veya hizmet satın alan tüketici gruplarıdır. Bu tüketici gruplarının işletmeye sadakatlerinin olduğunu söylemek mümkün değildir. İşletme tarafından ne kadar süre işletmeye bağlı tutulabilirse o kadar uzun süre temel düzeyde kar bırakırlar.

Gelir Büyümesi: Bir müşteri işletmeyle etkileşimde kaldığı süreyle doğru orantılı olarak daha sık alışveriş yapacak ve işletmenin bütün faaliyetlerini daha iyi öğrenerek diğer ihtiyaçları içinde işletmeyi seçmeye başlayacaktır.

Maliyetten Tasarruf: Zaman içerisinde müşteri işletmeyi, işletme de müşteriyi daha iyi tanır. Bu sayede müşteriler işletmeden sağlayamayacağı hizmetler talep etmeyi terk ederler. Ayrıca müşteri işletmeyi ve işletmenin ürünlerini tanıdıkça tavsiye ve bilgi için çalışanlara bağlılıkları azalır. İşletmelerse müşterilerini daha iyi tanıyarak profilleri hakkında daha iyi bilgiye sahip olur. Bu sayede işletmeler bir ürünü ya da hizmeti uygun profildeki müşteri grubuna yönelterek maliyetlerini azaltır.

Tavsiye: Sadık müşterilerin işletmeye sundukları üçüncü önemli fayda ise işletmeden memnun ve işletme ürün veya hizmetlerinden tatmin olmuş müşterilerin bu ürün ve hizmetleri çevresindekilere tavsiye etmeleridir.

Yüksek Fiyat: Eski müşteriler yenilere nazaran fiyat değişimlerine daha az duyarlıdır. Ayrıca yeni müşteriler kazanmak için başlatılan indirim ve kampanyalar kar oranını düşürmektedir.

		Tekrar Eden Satın Alma Davranışı	
		Yüksek	Düşük
Nispi Tutum	Yüksek	Mutlak Sadakat	Gizli Sadakat
	Düşük	Sahte Sadakat	Sadakatsizlik

Şekil 2. Dick ve Basu'nun Sadakat Modeli (1994).

2.7.1.1. Mutlak Sadakat

Mutlak sadakat gösteren müşteriler hem nispi tutumları hem de tekrar eden satın alma davranışı yüksek olan müşterilerdir. Bir müşterinin mutlak sadakatının sağlanması işletmenin pazarlama politikalarından kaynaklanabileceği gibi müşterilerin özelliklerinden de kaynaklanabilir. McGoldrick ve Andre'ye (1997) göre mutlak sadakati sağlayan müşteri özellikleri aşağıdaki gibidir;

- **Tembellik:** Bu müşteri farklı firmalar tercih ederek kendini yormak istemez. Ayrıca tasarruf etmek gibi bir amacı da yoktur.
- **Alışkanlık:** Müşteri her hafta alışveriş için aynı işletmeyi seçer ve bu alışkanlığını değiştirme gereği duymaz.

- **Uygunluk:** Müşteri için firma konumu, park imkânı, ürün çeşitliliği, açık bulunduğu saatler bakımından uygun olabilir.
- **Zamandan Tasarruf:** Müşteri çapraz alışveriş yapamayacak kadar yoğun olduğu için her ihtiyacını karşılayabileceği bir süper marketi tercih ediyor olabilir.
- **Tam Eğlence:** Müşteri sadakat gösterdiği işletmeden alışveriş yaparken eğleniyordur ve başka bir işletmeyi gezmeye gerek görmüyordur.

2.7.1.2. Gizli Sadakat

Bu sadakat düzeyindeki müşterilerin nispi tutumları yüksekken tekrar satın alma davranışları düşüktür. Yani müşteriler ürün ya da firmaya duygusal ya da psikolojik yakınlık duymakta firmayı ya da markayı beğenerek çevresindekilere tavsiye etmekte ancak tekrar eden bir alışveriş davranışı göstermemektedir.

2.7.1.3. Sahte Sadakat

Bu sadakat düzeyindeki müşteriler yüksek oranda tekrar eden alışveriş alışkanlığı gösterebilir de firmaya ya da markaya karşı nispi tutumları düşüktür. Yani müşteri alışveriş tercihinde söz konusu firma ya da markayı seçerken bu firma ya da markaya duygusal veya psikolojik bir yakınlık duymamaktadır. Bu tip müşteri gruplarının firma ya da markayı tercih neden marka memnuniyeti değil müşterinin alternatifsizliğidir.

2.7.1.4. Sadakatsizlik

Sadakatsiz müşteriler nispi tutumlarının yanı sıra tekrar satın alma davranışları da düşük olan müşterilerdir. Bu noktada sadakatsiz müşteriler için belli bir ürün ya da hizmete bağlılıkları olmayan ve tekrar satın alma amacı gütmeyen müşterilerdir denilebilir. Dick ve Basu'ya (1994) göre müşterilerin düşük nispi tutumlarının iki nedeni olabilir. Bu nedenlerden ilki pazara yeni giren bir marka ya da firma için uygun tutundurma faaliyetleri yerine getirilmemiş olabilir. Ürün müşteriye tam olarak tanıtılmamış ve diğer ürünlerden farklılıkları ve avantajları tüketiciye tam olarak açıklanamamış olabilir. İkinci neden ise pazardaki rakip firmaların ürün ve hizmetleri birbirine çok benziyor olabilir. Bu şartlar altında

pazarlama yöneticisi yüksek yüksek nispi tutum elde edemeyebilir ama agresif satış promosyonları ve tercih edilebilir ürünlere daha fazla raf alanı açarak sahte sadakat sağlayabilir.

Müşteri sadakatsizliği işletmenin yanlış pazarlama politikalarından kaynaklanabileceği gibi müşterilerin özelliklerinden de kaynaklanabilir. McGoldrick Ve Andre'ye (1997) göre sadakatsizliğe neden olan müşteri özellikleri aşağıdaki gibidir;

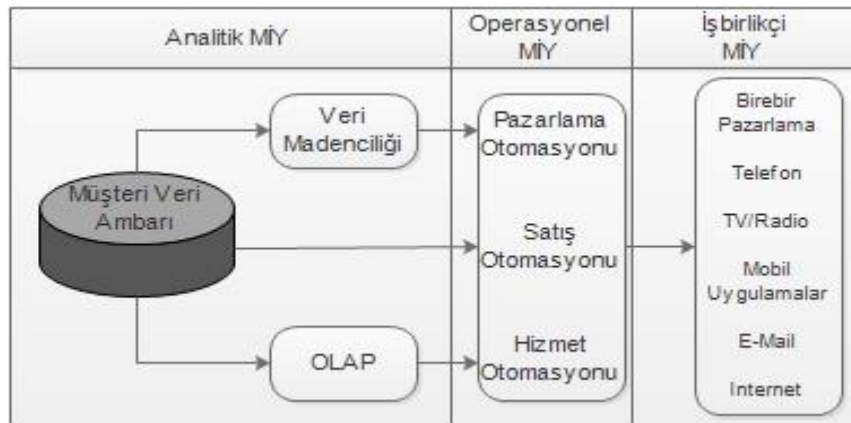
- **Akıllılık:** Müşteri farklı firmaları gezerek daha iyi bir fiyat, aynı fiyata daha yüksek kalite ve aynı fiyata daha fazla ürün gibi fırsatlar yakalayabileceğini düşünüyor olabilir.
- **Adanmışlık:** Müşteri farklı mağazalar gezerek diğer fırsatları görmeyi olması gereken tüketici davranışı olarak algılıyor olabilir. Kişi çapraz alışverişi bir görev gibi düşüyor olabilir.
- **Çoklu Eğlence:** Müşteri alışveriş faaliyetinden hoşlandığı için birden fazla mağazayı gezerek eğlencesini arttırmayı hedefliyor olabilir.
- **Para Tasarrufu:** Müşteri geliri düşük olduğu için daha harcama yapacağı alandaki tercihlerine dikkat ederek tasarruf yapmayı bu sayede sosyal hayatına, tatil planlarına, kıyafetlerine vs. daha fazla bütçe ayırmayı hedefliyor olabilir.
- **Merak:** Müşteri yeni gördüğü ve hakkında hiçbir şey bilmediği bir mağazanın formatını, ürünlerini ve markalarını merak ettiği için yeni firmaları ziyaret ediyor olabilir.
- **Zaman Yeterliliği:** Müşteri alışveriş için geniş bir zaman ayırdığı için kar etmek amacıyla farklı ürünleri indirimlerini uygun gördüğü firmalardan yapabilmektedir.
- **Zaman Öldürme:** Müşterinin gün içerisinde çok fazla boş vakti olabilir. Çapraz alışveriş sayesinde müşteri kendisine bir aktivite yaratmış olur.

2.7.2. Müşteri Odaklılık

Geçmişten günümüze artan teknolojik imkanlar ve küreselleşmeyle beraber rekabetin geldiği nokta işletmelerin nazarında müşterinin konumunu değiştirmiştir. Başlarda pazardaki talebin yüksek olması nedeniyle işletmeler talebi karşılayacak düzeyde üretimin yolunu aramaktaydı. Ancak zaman içinde pazardaki üretici sayısındaki artış ve teknolojik gelişmelerle üretimin daha kolay hale gelmesi odağı üretim sürecinden müşteriye çevirdi. Artık işletmeler için hayati soru üretilen ürüne müşteri bulmak ve dahası müşterinin memnuniyetini sağlayarak sadakatlerini kazanmaktır. Günümüzde teknolojik imkanlar sayesinde müşteriler geçmişe nazaran daha bilgili ve bilinçlidir ve seçimlerini tam olarak ne istediklerini bilerek yapmaktadırlar. Ayrıca, yeni müşteri edinme maliyetinin, mevcut müşteriye elde tutmaktan daha fazla olması, işletmeler için Müşteri İlişkileri Yönetimini (MİY) zorunlu kılmaktadır (Kağncıoğlu 2002).

2.8. Müşteri İlişkileri Yönetimi Mimarisi

MİY sistemleri müşteri odaklı sistemler olmakla birlikte uygulama alanlarına göre sınıflandırılabilirler. Bazı MİY uygulamaları müşterilerle birebir etkileşimi olan ön büro süreçlerinde kullanılırken bazıları ise idari süreçlerde karar destek sistemi olarak kullanılmaktadır. MİY sistemleri operasyonel müşteri ilişkileri yönetimi, analitik müşteri ilişkileri yönetimi ve işbirlikçi müşteri ilişkileri yönetimi olmak üzere üç başlık altında toplanabilir.



Şekil 3. Müşteri İlişkileri Yönetimi Mimarisi ve İş Zekâsı

2.8.1. Analitik Müşteri İlişkileri Yönetimi

Odağın müşteri olduğu günümüz piyasa şartlarında müşterilere uygun ürün ve hizmet sunulabilmesi için müşterilerle ilgili faydalı bilgilerin elde edilmesi zorunludur. Müşterilere dair doğru bilgiler elde etmenin yolu ise müşterinin tüketim tarihinin ve modelinin belirlenmesi gerekmektedir. Bu noktada gelişen teknoloji firmalara büyük kolaylık sağlamaktadır. Mevcut veya potansiyel müşterilere ait veriler çeşitli kaynaklardan toplanarak işlenmekte ve doğru müşteriye, doğru yöntemle yaklaşmayı sağlayacak stratejiler üretmek için kolaylık sağlamaktadır.

Analitik MİY operasyonel süreçlerden elde edilen verileri toplayıp, depolayıp, analiz ederek müşterilere dair stratejik bilgiler üretmeyi amaçlar. Analitik MİY müşterilerin geçmiş verilerinden bilgi üreterek geleceğe dair öngörüler geliştirmeyi amaçlar (Kart 2013: 13). Analitik MİY müşteri verilerinin incelenmesiyle doğru müşteriye doğru mesajı ulaştırmayı amaçlayan bir anlayıştır. Ayrıca veri madenciliği modellerini kullanarak müşteri değerini belirleye ve müşteri davranışını anlamaya çalışır (Abu Ellaban 2013: 30).

Aşağıda örnek Analitik MİY uygulamaları verilmiştir (Kelly 2000):

- **Satış Analizi:** Satış analizi yöntemiyle satışlara ait veriler farklı açılardan incelenerek ürün ya da ürün gruplarının bölgelere ve sezonlara göre satış rakamları, ürün ya da ürün gruplarını tercih eden müşterilerin demografik özellikleri belirlenebilir, ayrıca çapraz satış ve üst satış raporları oluşturulabilir. Bu sayede karar vericilerin karar vermedeki etkinliği arttırılabilir.
- **Müşteri Profil Analizi:** İşletmelerin müşteri memnuniyetini arttırmak için müşterileri farklı kriterler bakımından gruplara ayırması ve bu gruplar içerisinde kendisi için en uygun grubu seçmesi önemlidir. Müşterilerin doğru şekilde gruba bilmesi içinse müşteri profillerinin demografik ya da satış alışkanlıkları verileri üzerinden doğru analizi önemlidir.
- **Kampanya Analizi:** İşletmeler müşterilerin ilgisini çekmek amacıyla çeşitli dönemlerde kampanya ve promosyonlar geliştirirler. Bu noktada hedef müşteri ya da grubuna uygun kampanyaların belirlenmesi ve uygulanan

kampanyalara müşteri ya da müşterilerin olumlu karşılık verme oranlarının takip edilmesi önem kazanmaktadır.

- **Sadakat Analizi:** İşletmeler için karlılığın anahtarlarından birisi müşteri sadakatidir. Müşteri sadakat analizi sayesinde müşterilerin alışveriş sıklığı, son alışveriş tarihi gibi veriler incelenerek müşteriler sadık ya da sadık olmayan olarak sınıflandırılmaktadır. Bu sayede müşterilerin sadakatlerine etki eden faktörler incelenmekte ve yapıcı çözümler geliştirilebilmektedir.
- **Müşteri İrtibat Analizi:** Müşteriler işletmelerle tavsiyeler, şikayetler ve övgüler gibi birçok geri dönüşlerle etkileşime girebilir. Müşterilerin işletme faaliyetlerine dair olumlu ya da olumsuz geri dönüşleri müşteri ilişkilerinin sürdürülmesi ve hatta iyileştirilmesi açısından büyük önem taşımaktadır. Müşteri irtibat analiziyle müşterilerden elde edilen geri dönüşler incelenerek müşterilerle ilişkiler daha uzun süreye yayılabilir hatta daha da iyileştirilebilir.
- **Karlılık Analizi:** Karlılık analizi müşteri, ürün, şube vs. gibi yönleriyle karlılığı analiz ederken karlılığa etki eden faktörlerin belirlenmesinde de kullanılmaktadır.

2.8.2. Operasyonel Müşteri İlişkileri Yönetimi

Operasyonel MİY çözümleri müşteriyle birebir temasın sağlandığı ön ofis sürecinde kullanılan uygulamalardır. Müşteri verileri mail, web sitesi, çağrı merkezi, satış, pazarlama ve servis departmanı gibi müşteriyle birebir temas eden birimlerden toplanır. Daha sonrasında bu veriler müşteri veri tabanında toplanarak organize edilir ve müşteriyle ilişkili birimlerin kullanımına sunulur. Bu sistem çalışanlara müşterilerle ilgili anlık bilgilere ulaşma imkanı sağlar (Habul, Pilav-Velić, Kremić 2012). Operasyonel MİY'nin birçok uygulaması olmakla birlikte en sık kullanılanlar; pazarlama otomasyonu, satış otomasyonu ve hizmet otomasyonudur.

2.8.3. İşbirlikçi Müşteri İlişkileri Yönetimi

İşbirlikçi MİY üretim sürecine dahil olan tedarikçi, iş ortakları ve müşteriler arasında tam bir iş birliğini hedefleyen ve bu sayede müşteri taleplerine hızlı bir şekilde cevap vererek müşteri tatminini arttırmayı amaçlayan bir MİY

uygulamasıdır. İşbirlikçi MİY'nin etkin bir şekilde yürütülebilmesi için operasyonel MİY ile analitik MİY'nin birlikte koordineli bir şekilde yürütülmesi gerekmektedir.

2.9. Müşteri İlişkileri Yönetiminin Evreleri

Müşteri ilişkileri yönetimi süreci dört evrede gerçekleşmektedir. Bunlar; Müşteri seçimi, müşteri edinme, müşteri koruma ve müşteri derinleştirmedir.

2.9.1. Müşteri Seçimi

Bu aşamada esas amaç en karlı müşteri kim sorusuna cevap bulmaktır. Pazardaki her tüketici ihtiyaç ve talepleri bakımından birbirinden farklılık göstermektedir. Ancak işletmeler için her bir tüketicinin talebini karşılamak mümkün değildir. Bu nedenle işletmeler tüketicileri talep ve ihtiyaçları bakımından olabildiğince homojen gruplara ayırma yoluna gitmektedir. Bu sayede işletmeler kendileri için en karlı ve işletme imkanları dahilinde taleplerine en kolay cevap verebileceği müşteri grubunu hedef pazar olarak seçmektedir. Müşteri seçimi aşamasında şu çalışmalar yapılmaktadır (Aktepe, Baş, Tolon 2009: 67);

- Hedef Pazar Seçimi,
- Pazar Bölümlendirme,
- Ürün Konumlandırma,
- Kampanya Planları,
- Marka ve Müşteri Planlamaları,
- Yeni Ürün Lansmanları.

2.9.2. Müşteri Edinme

Müşteri ilişkileri yönetiminin ikinci aşaması müşteri edinmedir. Müşteri seçimi aşamasında belirlenen müşteri ya da müşteri grubunun dikkati işletmenin ürün ve hizmetlerine çekilerek alışveriş eylemine geçirilmeleri hedeflenmektedir. Müşteriyi tanıma, müşteriyle bağlantı kurma ve bağlantıyı sürdürmenin yollarının arandığı bu aşamada “müşterilere en etkili yoldan satış nasıl yapılabilir” sorusunun yanıtı aranır (Yıldırım, Panayırca 2016). Bu aşamada şu çalışmalar yapılmaktadır;

- İhtiyaç analizlerinin yapılması,
- Satış teklifi ve ön talep oluşturma,
- Satışın gerçekleştirilmesi.

2.9.3. Müşteri Koruma

Müşteri ilişkileri yönetiminin üçüncü aşaması müşteri koruma aşamasıdır. Bu aşamada işletme “kazanılmış müşteriyi ne kadar süre elimizde tutabiliriz” sorusuna cevap aramaktadır. Yani bu aşamada asıl hedef müşteri sadakatini kazanmaktır diyebiliriz. Bu aşamada yürütülen çalışmalar;

- Sipariş Yönetimi,
- Teslim,
- Taleplerin Organizasyonu,
- Problem Yönetimi, Refleks Sistemi.

2.9.4. Müşteri Derinleştirme

Müşteri derinleştirme Müşteri ilişkileri yönetiminin dördüncü ve son aşamasıdır. Bu aşamada müşteri sadakatinin ve karlılığının devam ettirilerek müşteri cüzdandan firmaya harcanan miktarın artırılması amaçlanmaktadır. Müşteri derinleştirme aşamasında yürütülen çalışmalar;

- Çapraz Satış Kampanyaları,
- Müşteri İhtiyaç Analizleri.

Pazarlamanın önemli fonksiyonlarından birisi hiç şüphesiz talep yaratmadır. Birden fazla tür ürün üreten işletmeler bir ürünü alan müşterileri diğer ürünleri de satarak müşteri derinleştirme yapabilir. Bunun yanı sıra müşteri ihtiyaç analizleri aracılığıyla gelecekte oluşabilecek müşteri ihtiyaçlarını önceden belirleyerek bu ihtiyaçlara rakiplerinden erken cevap verebilir.

2.10. Müşteri İlişkileri Yönetimi Uygulama Evreleri

Müşteri ilişkileri yönetiminin uygulama evrelerini aşağıdaki 4 başlık altında toplamak mümkündür:

- Müşterileri Tanımlama,
- Müşterileri Farklılaştırma,
- Müşterilerle Etkileşime Girme,
- Müşterilere Bire Bir Hizmet Sunulması.

2.10.1. Müşterileri Tanımlama

Müşteri ilişkileri yönetimi disiplininin temelinde müşteri yer almaktadır. Müşteri ilişkileri yönetimi müşterilere ihtiyaçlarına ve beklentilerine göre muamele yapmayı gerektirmektedir. Bu nedenle müşteri ilişkileri yönetimine göre her müşteriye uygun mal ve hizmet sunabilmek için öncelikle müşteriyi tanımak gerekmektedir. Müşterilerin tanımlanmasında yaş, cinsiyet, meslek vb. gibi demografik değişkenlerin yanı sıra alışkanlıklar, beklentiler, yaşam biçimleri gibi davranışsal değişkenlerde incelenmektedir (Çatı, Koçoğlu 2010: 249).

2.10.2. Müşterileri Farklılaştırma

İşletmeler kâr amacıyla kurulan organizasyonlardır. Karlarını maksimize etmeye çalışan işletmeler için her müşteri aynı değerdedir. Ayrıca söz konusu müşterilerin talep ve ihtiyaçları da farklılık göstermektedir. Müşteri farklılaştırma, müşteriler arasındaki talep ve değer farklılığından doğan bir zorunluluktur. İşletmeler için müşterileri değer ve talepleri bakımından sınıflandırmak, işletmenin elindeki kısıtlı kaynaklarını talepleri en efektif şekilde karşılanabilecek aynı zamanda en fazla kar sağlayacak müşteri grubuna yönlendirmesini sağlamaktadır.

Müşteriler işletme için iki bakımdan farklılaşmaktadır. Birincisi, her müşterinin işletmeye sağladığı değer farklıdır. İkincisi, her müşterinin işletmeden beklentileri farklıdır. Bu nedenle müşteriler sınıflandırılırken bu iki kriterin esas alınması gerekir (Çiçek 2005: 83).

2.10.3. Müşterilerle Etkileşime Girme

Müşteri ilişkileri yönetiminin üçüncü aşamasında işletmeler müşterilerle etkileşime girebilecekleri kanallar üzerinden müşterilerden geri bildirimler almaya çalışmaktadır. Bu aşamada müşterilerle satış ziyaretleri, pazarlama faaliyetleri, telefon, internet şubesi, çağrı merkezi, e-posta, müşteri hizmetlerine gelen şikayetler, fatura gönderme yoluyla iletişim kurulmaktadır. Bu etkileşim sayesinde müşterilerle çift taraflı bir ilişki kurulmakta bu sayede müşterilerle ilgili daha fazla bilgi elde edilebilmektedir (Yeler 2006: 12). Müşterilerle kurulan bu etkileşim sayesinde işletmeler müşterilerin ürüne, firmaya ya da firma çalışanına dair şikayetleri ve iş süreçlerine dair aksaklıları hızlıca tespit edebilmektedir. Ayrıca müşterinin talep ve ihtiyaçlarındaki değişimleri daha kolay takip edebilmektedir.

2.10.4. Kitlesele Kişiselleştirme

Müşteri ilişkileri yönetiminin son evresi kitlesele kişiselleştirme evresidir. Pazardaki artan rekabet müşteri taleplerinin önemini arttırmıştır. Ancak işletmelerin her bir müşterinin talebine cevap vermesinin imkansızlığı kitlesele kişiselleştirme kavramını ortaya çıkarmıştır. Kitlesele kişiselleştirme ile işletmeler üretim süreçlerini maliyetleri arttırmadan özel taleplere cevap verebilecek şekilde geliştirmişlerdir. Kitlesele pazarlama farklı müşteriye farklı şekilde muamele etmeyi gerektirmektedir (Peppers, Rogers 2000: 114).

2.11. Müşteri İlişkileri Yönetimi Bileşenleri

Müşteri ilişkileri yönetimi uygulamalarına geçmeye karar veren ve bu uygulamaları etkili bir şekilde kullanmak isteyen işletmelerin yatırımlarını ve çalışmalarını insan, süreç ve teknoloji alanlarına yoğunlaştırmaları gerekmektedir. Müşteri ilişkileri yönetimi projelerinin başarısı bu üç alanın doğru şekilde birbiriyle entegreli yürütülmesine ve bu alanlarda yapılan yatırımların dengeli dengeli dağıtılmasına bağlıdır (Murat 2005: 21).

2.11.1. İnsan

İşletmelerde değişim kavramı kültürel ve organizasyonel olarak iki başlık altında ele alınabilir. Kültürel değişimden kasıt çalışanların yeni düşünce iş yapış

biçimini anlamaları ve benimsemeleridir(Uzun 2004: 27). Organizasyonel değişimden ise işletmenin organizasyonel yapısının müşteri ilişkileri yönetimi anlayışına uygun olarak yeniden dizayn edilmesi anlaşılabilir.

İnsanlar yapıları gereği değişime direnç gösterirler. Ancak hızla değişen dünyada oluşan yeni şartlara adapte olunabilmesi ve oluşan yeni talep ve beklentilere hızlı cevap verilebilmesi için işletmelerin bünyesinde yenilikçi bir kadro barındırması zorunludur. Özellikle müşteri ilişkileri yönetimi gibi işletmede organizasyonel ve kültürel değişim gerektiren bir yaklaşımın etkili yürütülebilmesi için insan faktörü çok büyük önem taşımaktadır.

2.11.2. Süreç

Üretim işletmelerin de 1995 ile 2000 yılları arasında adından sıkça söz ettiren BPR (Business Process Re-Engineering) yaklaşımı bütün iş süreçlerinin tekrar değerlendirilerek yeniden yapılandırılmasını amaçlamaktaydı. Ancak maalesef başarılı olamadı. BPR'nin başarısızlığının en önemli nedeni vizyon ve stratejiden yoksun olmasıydı. Süreçleri gözden geçirecekti ancak ne için gözden geçireceğine dair belirgin bir amacı yoktu. Müşteri ilişkileri yaklaşımının ortaya çıkışıyla beraber BPR müşteriye değer yaratma amacını yükledi ve CPR (Customer Process Re-Engineering) adını aldı (Uzun 2004: 28).

Müşteri ilişkileri yönetiminin etkili bir şekilde yürütülmesi için en önemli yöntemlerden birisi süreçlerin yeniden düzenlenmesidir. Bu noktada müşterilerden geri dönüşler alınarak değişime ihtiyacı olan süreçler belirlenir. Bu süreçler çalışanlarla değerlendirilerek yeniden tasarlanır. Daha sonrasında yeniden yapılanmış süreçler müşterilerde dahil olmak üzere ilgililere duyurulur (Çiçek 2005: 59).

2.11.3. Teknoloji

Müşteri ilişkileri yönetimi yaklaşımına geçişi hedef edinmiş bir işletmede süreçler hedefler doğrultusunda tasarlandıktan ve bu yaklaşıma dair yöneticilerin ve çalışanların desteği alındıktan sonra teknolojinin gerekliliği kendiliğinden hissedilecektir (Murat 2005: 23). Her ne kadar müşteri ilişkileri yönetimi sadece teknolojiden ve yazılımlardan ibaret değilse de teknolojinin buradaki rolü büyüktür.

Müşteri ilişkileri yönetimi çerçevesinde, yaşam biçimleri ve satın alma davranışları ile ilgili müşterilerden elde edilen veriler anlamlı bilgiye dönüştürülür. Müşterilerle ilgili elde edilen bu anlamlı bilgiler ise müşterilerle ilişkiler geliştirmek için kullanılır. İşte bu süreçte karar alıcılara yardımcı olabilecek teknolojiler mevcuttur. Bu teknolojiler sayesinde müşterilerin geçmiş davranışları, bu davranışları etkileyen faktörler ışığında gelecek davranışlarına dair öngörüler geliştirilebilir. Bu öngörüler ise pazarlama, satış, hizmet ve üretim stratejilerinin belirlenmesinde büyük faydalar sağlar (Çiçek 2005: 60).

Müşteri ilişkileri yönetiminin uygulanma sürecinde karşılaşılan hatalardan biri bu yaklaşımın sadece teknolojiden ibaret olduğu yani bir yazılım satın alınarak etkili bir şekilde uygulanabileceği düşüncesidir. Teknoloji müşteri ilişkileri yönetiminin çok önemli bir bileşeni olmakla beraber tek başına yeterli değildir. İşletmeler sadece bir yazılım satın alarak müşteri ilişkileri yönetimini doğru oluşturamazlar. İşletmenin etkili bir müşteri ilişkileri yönetimi oluşturabilmesi için teknolojiyi, insan ve süreçle dengeli bir şekilde koordine etmesi gerekir.

3. İŞ ZEKÂSI

Rekabetin giderek daha belirginleştiği günümüz piyasa şartlarında bilgi artık stratejik olarak önemli bir kaynak haline geldi (Habul & Pilav-Velić, 2010). İşletmeler günlük faaliyetleri içerisinde sürekli olarak veri üretirler. İşletmelerin iç ve dış süreçlerinde üretilen bu verilerin depolanması, analiz edilmesi ve işletmelerin karar alma sürecinde ciddi stratejik avantaj sağlayacak bilgiye dönüştürülmesi işletmeye rekabet üstünlüğü sağlaması bakımından çok önemlidir. İşte iş zekâsı bu noktada karşılaşılan ve işletmeye bilgi üretme noktasında rehberlik eden bir teknolojidir.

İş zekâsı terimini ilk olarak 1989 Gartner Group bünyesinde analist ve araştırmacı olarak çalışan Howard Dresner tarafından kullanılmıştır. Dresner, iş zekâsını, gerçeğe dayalı karar destek sistemlerini kullanarak, iş süreçlerinde karar verme becerisini geliştiren bütün metot ve fikirleri kapsayan bir şemsiye olarak tanımlamaktadır (Ateş 2008: 43).

Literatürde iş zekasıyla ilgili birçok tanım bulunmakla birlikte birkaçı aşağıda verilmiştir:

Moss ve Ateş'in (2003 : 10) tanımına göre; İş zekası ne bir ürün ne de bir sistemdir. İş zekâsı veri tabanlarını karar destek uygulamalarıyla birleştirilerek iş çevrelerinin kullanımına sunan bir mimari ve bütünleşik bir süreçtir.

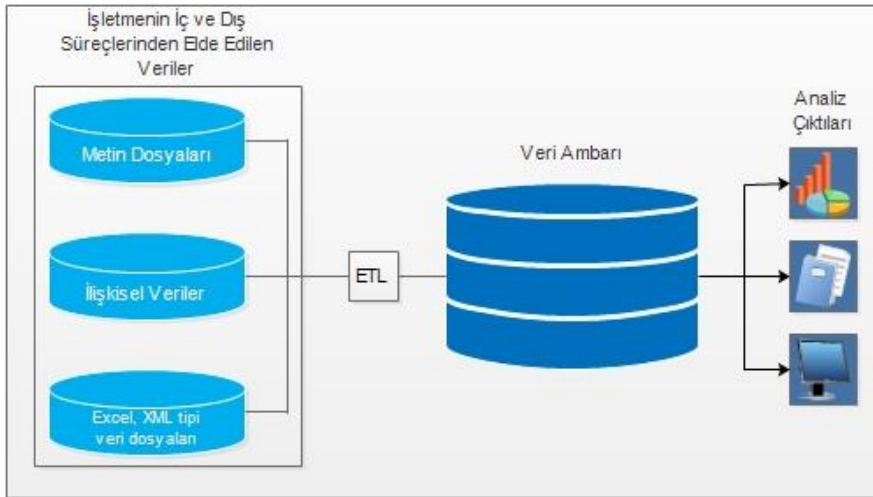
Diğer bir tanıma göre; İş zekası işletmelere rekabet üstünlüğü sağlamak amacıyla bilgi teknolojilerini kullanan ve işletmenin iç ve dış çevresinden ve kararlarından doğan riskleri belirlemeye çalışan bir konsepttir (Habul, Pilav-Velić 2010).

Agiu vd. (2014) tanımına göre; iş zekası teknik ve pratik olarak verinin toplanması işlenmesi ve analiz edilmesi için bir araçtır.

Bu tanımlar ışığında iş zekâsı işletme fonksiyonlarıyla bilişim teknolojilerini birleştirerek işletmenin farklı departmanlarından toplanan verileri depolayan, analiz ederek raporlayan ve sonuç olarak karar vericilere işletmenin iç ve dış çevresiyle ilgili faydalı bilgiler sunarak işletmeye rekabetsel üstünlük sunmayı amaçlayan bir sistemdir.

İş zekasının önemini anlamak için iş zekâsı konseptinin hangi faydaları sağladığını bilmek gerekmektedir. Aşağıda iş zekasının firmalara sağladığı bazı faydalar verilmiştir (Agiu, Mateescu, Muntean 2014).

- İş zekâsı sistemi efektif risk yönetimine imkân sağlar: iş zekâsı sistemi sayesinde işletmenin stratejik hedefleri ve finansal kayıplarına dair riskleri belirlemek kolaylaşacaktır.
- İş zekâsı çözümleri, standart ilişkisel verileri, metin dosyalarını ve Microsoft excel'den XML veri dosyalarına kadar her türlü veri türünü veri ambarlarında veya operasyonel sistemlerde entegre bir şekilde saklar. Bu şekilde bu veriler iş zekâsı uygulamalarında kullanılmak üzere daima hazır vaziyette bulunurlar.
- İş zekâsı sisteminin uygulatma maliyeti çok düşüktür, donanım ekipmanlarına yapılacak yatırım çok yüksek olmamakla birlikte yeni personele sistemin öğretilmesi çok uzun sürmez. Yapılan küçük ölçekli harcamalar ise ilk ay içinde sistem sayesinde karşılanmış olur.
- Şirket içindeki güç oyunlarının etkisini azaltır.
- Karar odaklı problemleri azaltır.



Şekil 4. İş Zekâsı

İş zekasının pratik sonuçları için aşağıdaki uygulamalar örnek olarak gösterilebilir (Tarimer 2011):

- Karlılık analizleri,
- Maliyet analizleri,
- Kaynak sonuç ilişkisi (satışlarla pazarlama giderleri arasındaki ilişki),
- Sektörel değerlendirmeler, (Toplam pazar büyümesi içinde işletmenin durumu, personel bazında satış ortalamasının sektörde ve kurumdaki durumu),
- Anlamlı modellerin kurulması, (satış rakamlarına etki eden değişkenler ile birlikte sayısallaştırılması),
- Bütçeleme, (verilerden orta vadeli nakit bütçesi oluşturulması).

3.1. İş Zekâsı Kavramları

İş zekâsı projelerinde ve ilgili kaynaklarda sıklıkla karşımıza 3 kavram çıkmaktadır. İş zekâsı kavramının ve uygulama süreçlerinin doğru anlaşılması için bu kavramları açıklık getirmek uygun olacaktır. Bu kavramlar; OLTP (On Line Transactional Processing), OLAP (On Line Analytical Processing), ETL (Extract, Transform, Load) ve Metaveri (Metadata) şeklinde sıralanabilir.

3.1.1. Veri, Enformasyon ve Bilgi

21. yüzyılda insanlar kullandıkları elektronik cihazlar, kredi kartları, internet üzerinden yaptıkları işlemler, sosyal medya kanalları üzerinden sürekli veriler üretirler. Bilginin öneminin giderek arttığı günümüz şartlarında günlük olarak üretilen bu veri yığınları karar alıcıların ilgisini çekmektedir. Ancak veriler ham haldeyken karar alma sürecinde hiçbir önem taşımamaktadır. Bu nedenle öncelikle bu veri yığınlarının uygun yöntemlerle işlenerek önce enformasyona sonrada bilgiye dönüştürülmesi gerekmektedir. Bu noktada veri, enformasyon ve bilgi arasındaki farklara değinmek yerinde olacaktır.

Veri (Data): Veriler gözlem, araştırma, deney, ölçüm ve sayım gibi değişik yöntemlerle elde edilen sayılar, rakamlar, sözcükler, metinler, resimler, olaylar vb. biçiminde temsil edilen ham gerçekliklerdir. Veriler sayısal değer ifade etmesi

durumunda nicel ve sayısal olmayan değer (nominal ve ordinal) ifade etmesi durumunda nicel veri olarak adlandırılır. (Özen 2016).

Enformasyon (Information): Enformasyon, verinin özetleme, hesaplama, sınıflandırma, gruplandırma ve analizler aracılığıyla düzenlenerek anlamlı hâle dönüştürülmesidir. Yani veriler, düzenli ve anlamlı ilişkiler çerçevesinde bir araya getirildiğinde kullanıcılar için anlamlı hale dönüşmektedir.

Bilgi (Knowledge): Yılmaz'ın (2009) tanımına göre veri ve enformasyon dışardan elde edildiği halde bilgi beyinde ortaya çıkar. Yani bilgi kişinin sahip olduğu tecrübeler, değerler, enformasyon ile yeni edindiği enformasyon ve tecrübelerin belirgin bir örüntü içerisine değerlendirilmesiyle ortaya çıkan sonuçtur. Bu bakımdan bilgi sübjektiftir ve bilgi objektif hale geldiği zaman enformasyon seviyesine iner.

3.1.2. OLTP ve OLAP

OLTP (Online Transactional Processing) ve OLAP (Online Analytical Processing) kavramları ilişkisel veri tabanları ve veri ambarları konularında sıklıkla karşılaşılan kavramlardır. OLTP tipi sistemler veri girmeye, güncellemeye ve silmeye imkân sağlayan, ayrıca sorgularla çıktılar aldığımız veri tabanı sistemleridir. Günlük hayatta kullanılan birçok veri tabanı OLTP tarzı işlem yapan veri tabanlarıdır. Günümüzde işletmelerde yaygın olarak kullanılan ERP sistemleri OLTP sistemler üzerinde işlem yapmaktadır. Bunun en önemli nedeni OLTP tarzı sistemlerin gün içerisindeki yoğun girdi-çıkıtı ve güncelleme işlemini kolayca gerçekleştirebiliyor olmasıdır. Ancak OLTP sistemleri karmaşık sorgular için uygun değildir. Bu nedenle OLAP sistemleri geliştirilmiştir (Akça 2012). OLAP sisteminin daha iyi anlaşılabilmesi için bilgi teknolojileriyle uğraşan herkesin yakından bildiği OLTP ile karşılaştırılarak anlatılması yerinde olacaktır.

- OLTP işlem yoğunluğu yüksek işletme için operasyonel süreçlerin kullanımı için uygun bir sistemken OLAP karmaşık analizler için daha uygundur.

- OLTP ancak düşük boyutlu ve karmaşık ilişkilerin olmadığı veriler üzerinde işlem yapabilirken OLAP büyük boyutlu ve karmaşık ilişkili veriler üzerinde de rahatça işlem yapabilmektedir.
- OLTP ofis çalışanları ve bilgi teknolojileri çalışanları tarafından kullanılırken OLAP analistler tarafından tercih edilmektedir. Bu nedenle OLTP çok sayıda kişi tarafından detaylı veri girişinde, OLAP ise az sayıda insan tarafından özet veriler oluşturulup analitik işlemlerde kullanılır. Analitik işlemlerden kasıt OLTP tipi sistemlerden ETL işlemiyle elde edilen sorunsuz verilerin OLAP'a yüklenerek belirli kriterlere göre analiz edilmesidir.
- OLTP sistemler OLAP'dan çok daha fazla sayıda detay barındırılmaktadır.

OLAP ileriki konularda anlatacağımız veri ambarları üzerine kurulmakla birlikte veri ambarlarından tamamen farklı bir konudur. Burada üzerinde durulması gereken hususlardan birisi OLAP'ın yukarıda bahsedilen avantajları çerçevesinde veri ambarları üzerinde kullanılması fayda sağlayacaktır. Ancak veri ambarlarında OLAP kullanılması zorunlu değildir.

3.1.3. ETL

ETL iş zekâsı projelerinin en önemli işlem adımı olmakla beraber aynı zamanda maliyet ve iş gücü bakımından işletmeye en fazla yük getiren süreçtir. ETL işlemi çıkarma, dönüştürme ve yükleme (ETL: Extract-Transform-Load) olmak üzere üç adımdan oluşmaktadır. Veri ambarı ve veri marketi oluşturma sürecinin büyük kısmını ETL süreci oluşturmaktadır. Bazı projelerin %80'ini ETL oluşturmaktadır. Bunun nedenleri arasında verinin tanımlanmasının ve çıkarılmasının zorluğu, dönüştürme sırasında kullanılan metotların karmaşıklığı, verilerdeki tutarsızlıklar ve kirlilik gösterilebilir.

Çıkarma: Verinin dahili veya harici veri kaynaklarından çıkarılma sürecidir.

Dönüştürme: ETL içerisinde en fazla vakit alan süreçtir. Farklı veri kaynaklarından gelen verilerin veri ambarı modelimize göre tekrar düzenlendiği gürültüden arındırılarak birbiriyle ve veri ambarıyla uyumlu hale dönüştürülme işlemidir.

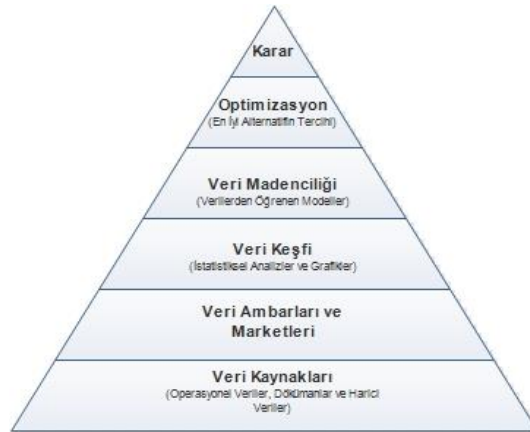
Yükleme: Kaynak sistemden gelen verilerin hedef veri ambarına ya da veri marketine yüklenme aşamasıdır.

3.1.4. Metadata

Veri hakkında veri şeklinde tanımlanan metadata; veri ambarlarının en önemli bileşenlerinden birisidir ve veri ambarını tanımlayan verilerdir. Metadata veri ambarının yapısı, ilgili isimler, veri tanımlamaları, son kullanıcı bilgileri, önceden yapılmış sorgular, kullanıcı yetkileri ve tanımları vb. bilgileri taşır. Ayrıca veri analiz sürecinde kullanılan yöntemlere dair bilgiler içerir. Örneğin; müşteri yaşam değerinin hesaplanmasında hangi formülün kullanıldığı veya müşteri sadakatinin belirlenmesinde hangi yaklaşımın esas alındığı bilgileri yine metadata altında saklanır.

3.2. İş Zekâsı Uygulama Adımları

Vercellis'e (2009: 11) göre iş zekâsı sisteminin 6 temel bileşeni vardır. Bunlar; veri kaynakları, veri ambarları ve marketleri, veri keşfi, veri madenciliği, optimizasyon ve karardır. Söz konusu adımlar iş zekâsı piramidi şeklinde aşağıda basitçe açıklanmıştır.



Şekil 5. İş Zekâsı Piramidi (Kaynak: Vercellis, 2009: 10)

3.2.1. Veri Kaynakları

Veri kaynakları iş zekâsı uygulamaları için işletmeye dair verileri barındıran kaynaklardır. Genel olarak operasyonel sistemlerden elde edilen veriler, ilişkiyel veri

tabanları, çeşitli dosya türleri (excel, notpad vs) veya xml formatta olabilir. Bu veriler e-postalardan, ERP sistemlerinden, günümüzde sosyal medya kanallarından ya da pazar araştırması yapan firmalardan elde edilebildiği gibi iş zekâsı projesinin özelliğine göre barkod okuyuculardan, otomatik kapı algılayıcılarından, GPS kayıtlarından da elde edilebilir. İş zekâsı projelerinin başlangıç noktası kabul edilen bu aşamada çeşitli kaynaklardan elde edilen veriler arasındaki uyumsuzluklar ortadan kaldırılarak bütünleştirilir.

Vercellis'e (2009: 45) göre verileri elde edilen kaynaklara göre dahili veriler, harici veriler ve kişisel veriler olarak gruplamak mümkündür.

3.2.1.1. Dahili Veriler

Dahili veriler genel olarak işletmenin yönetim, muhasebe, üretim, insan kaynakları, lojistik gibi operasyonel süreçler sırasında ortaya çıkan verilerdir. Bu veriler işletme bilgi sisteminin bel kemiğini oluşturur.

Dahili verilerin elde edildiği kaynakları üç başlık altında toplamak mümkündür.

- Geri Ofis Verileri: Bu veriler genel olarak siparişler, faturalar, depo durumu, üretim ve lojistik gibi kaynaklardan elde edilir.
- İleri Ofis Verileri: Çağrı merkezi aktiviteleri, müşteri etkileşimleri ve kampanya analizleri sonucunda elde edilen verilerdir.
- Web Tabanlı Sistemler: E-ticaret verileri, site ziyaret verileri, müşteriler tarafından gönüllü olarak doldurulan formlardan elde edilen veriler.

3.2.1.2. Harici Veriler

Harici veri kaynakları işletme dışı kaynaklardan elde edilen verilerdir. Örnek olarak pazar payları, belirgin sektörlerle ilgili trend tahminleri bunların yanı sıra ekonomik ve finansal göstergeler sayılabilir. Ayrıca farklı firmalar tarafından yapılan müşteri anketleri veya müşterilerin belirgin bir marka ya da ürüne dair fikirleri de harici veriler arasında gösterilebilir.

3.2.1.3. Kişisel Veriler

Veriler arasındaki ilişkileri daha anlamlı kılmak ya da saklı kalmış örüntüler varsa ortaya çıkarmak için iş zekâsı uzmanları iç ve dış müşterilere ait çeşitli verilere de ihtiyaç duymaktadır. Bu veriler sayesinde müşterinin satın alma davranışı üzerindeki demografik ve kültürel etkileri görebilmekte, ayrıca gelir düzeyinin ürün tercihlerine etkisi gibi anlamlı sonuçlara ulaşabilmektedir.

3.2.2. Veri Ambarları ve Veri Marketleri

Veri ambarlarıyla ilgili birçok tanım bulunmakla birlikte Tezcanlar'a (2007: 33) göre veri ambarları yönetimin karar süreçlerini desteklemek için kullanılan, konuya yönelik, entegre edilmiş, zamana bağlı ve kalıcı veri topluluğudur.

Reinschmidt ve Francoise'e (2000) göre ise veri ambarları, veri kaynaklarından toplanan verilerin ETL sürecinden geçirildikten sonra analiz amacıyla tekrar depolandıkları özelleştirilmiş veri tabanlarıdır. Veri ambarları iş zekâsı uygulamalarının başlangıç aşamasıdır ve yönetime iş süreçleri ile ilgili bilgilere erişim ve analiz imkânı sunan bir teknolojidir.

Han ve diğerlerine (2012:10) göre veri ambarı veri çıkarma, veri temizleme, verileri entegre etme, veri dönüştürme ve veri yükleme sonrası oluşan özelleştirilmiş bir veri tabanı modelidir.

Han ve diğerlerine (2012: 11) göre veri ambarlarının 4 temel özelliği vardır. Bunlar konu odaklı olmaları, entegre olmaları, zamana bağlı olmaları ve kalıcı olmalarıdır.

- Konu odaklılık: veri ambarları müşteri, tedarikçiler, üretim vb. gibi bir ana konu üzerine inşa edilir.
- Veriler arasında entegrasyon: veri ambarlarına yüklenecek veriler genelde farklı kaynaklardan elde edilirler. Bu nedenle bu veriler arasında uyumsuzluk olması olasıdır. Veri temizleme ve veri entegrasyon teknikleriyle bu veriler birbiriyle uyumlu hale getirilir.
- Zamana bağlı olma: Veri ambarlarında saklanan veriler hangi kaynaktan elde edilmiş olursa olsun zaman etiketiyle saklanır.

- Kalıcı olma: Herhangi bir hata olmadığı sürece veri ambarına yüklenen veriler üzerinde değişiklik yapılmaz.

Günümüzde farklı iş süreçlerine ait verilerin farklı veri tabanlarında saklanması ve veri tabanlarının çok büyük hacimlerde veriler içermesi analiz sürecinde birtakım zorluklar çıkarmaktadır. Veri ambarları bu zorlukların üstesinden gelmek amacıyla tasarlanmıştır. Farklı veri kaynaklarından toplanan veriler arasındaki uyumsuzluklar giderilerek veri ambarlarında saklanır. Bu sayede veriler farklı analiz yöntemlerinin uygulanması amacıyla hazır halde tutulur.

Veri marketleri ise veri ambarlarına kıyasla içerisinde daha kısıtlı veriler barındırır. Veri marketleri işletmenin belirgin bir departmanına ait özetlenmiş ve sadece o departmanın ihtiyaçlarına yönelik verileri kapsar. Veri marketlerinin birleştirilmesiyle veri ambarları oluşturulur ve bu veri ambarları kurumsal veri ambarları olarak adlandırılır (Tezcanlar 2007: 37).

3.2.2.1. Veri Ambarlarında Kullanılan Tablo Yapıları

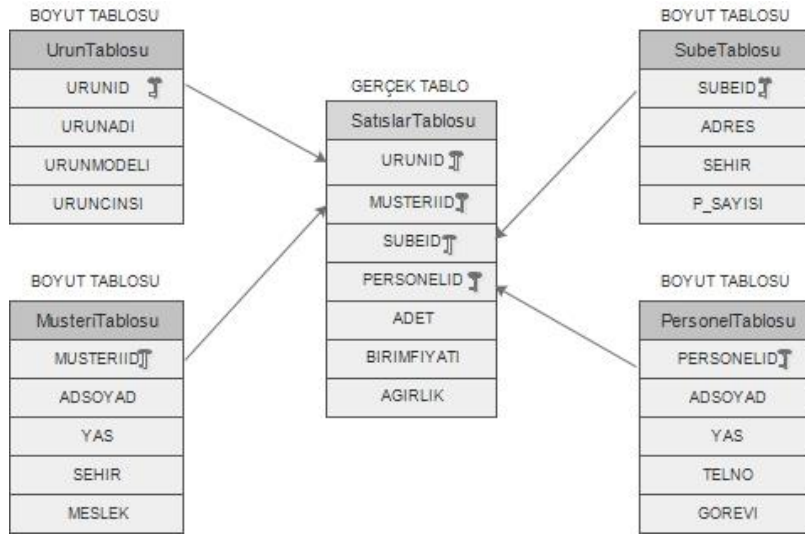
Veri ambarları belirgin bir iş probleminin çözümü amacıyla inşa edilirler. Bu noktada ilk yapılması gereken problemin tanımlanmasıdır. Daha sonra problemin çözümü için hangi tip verilere ihtiyaç duyulacağını ve verilerin saklanacağı veri tabanı yapısının nasıl oluşturulacağını belirlemek gerekir.

Problem odaklı olarak hazırlanan veri ambarlarında probleme dair veriler gerçek veriler olarak adlandırılırken, gerçek tablolardaki verilerle alakalı yan veriler boyut verilere adı verilir. Örneğin satış işleminin takip edildiği bir veri tabanı modelinde satışın tarihi, satılan malın adı, satış fiyatı gibi bilgiler satış tablosu altında saklanır. Bu tabloya gerçek tablo adı verilir. Satış işleminin gerçekleştiği şube bilgileri, satışı yapan personel bilgileri, satılan mala dair bilgiler vs. ayrı tablolar altında toplanır ve bu tablolar altındaki bilgiler gerçek tablo altında yabancı anahtarla temsil edilir.

Veri ambarları inşa edilirken 3 veri tabanı modeli kullanılabilir. Bunlar Yıldız (star schema), Kar Tanesi (snowflake schema) ve Takımyıldız Şemalarıdır (Fact Constellations).

3.2.2.1.1. Yıldız Şeması (Star Schema)

Veri ambarlarının tasarımında en çok kullanılan tasarım yıldız şemasıdır (Han, Kamber, Pei 2012: 140). Yıldız şeması modelinde gerçek tablolar merkezde yer alırken boyut tabloları gerçek tablolara bağlanır. Bu şekliyle bir yıldız benzediği için aynı isimle anılır. Aşağıdaki şekilde bir yıldız şeması örneklendirilmiştir. Ortadaki tablo gerçek tablo olarak adlandırılır ve iş probleminin esasını oluşturan tablodur. Gerçek tabloya bağlı diğer tablolar ise boyut tablosu olarak adlandırılır. Gerçek tablo içerisindeki kenarlarında anahtar sembolü bulunan değerler yabancı anahtar olarak adlandırılır ve boyut tabloları temsil eder ve bağlı bulunduğu boyut tablosuyla gerçek tablo arasındaki bağlantıyı sağlar.



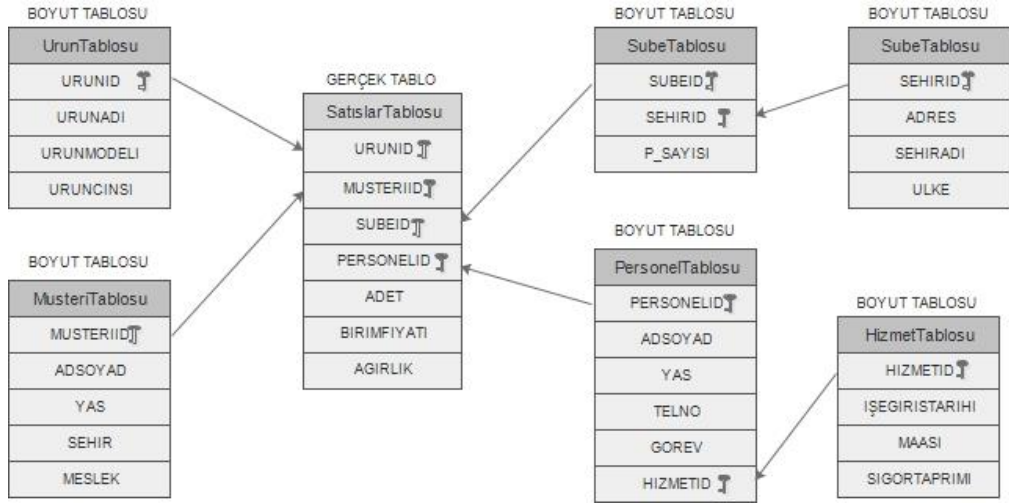
Şekil 6. Yıldız tablo yapısı

3.2.2.1.2. Kar Tanesi Şeması (Snowflake Schema)

Yıldız şeması kar tanesi şemasının bir çeşididir. Yıldız şemasından farklı olarak kar tanesi şemasında boyut tabloları altında yeni bir boyut tablosu oluşturulabilir. Bu şekilde ortaya çıkan görüntü bir kar tanesini andırmaktadır.

Kar tanesi ve yıldız şeması modelleri arasındaki en büyük fark, veri tekrarlarını azaltmak için kar tanecikli modelin boyut tablolarındaki verileri normalleştirilmiş biçimde tutulabilmesidir. Ayrıca bu yöntem tabloların sürekliliğini kolaylaştırırken depolama alanından da tasarruf sağlar. Ancak bu depolama tasarrufu gerçek tabloların veri büyüklüğüne kıyasla önemsizdir. Ayrıca, kar taneciği yapısı,

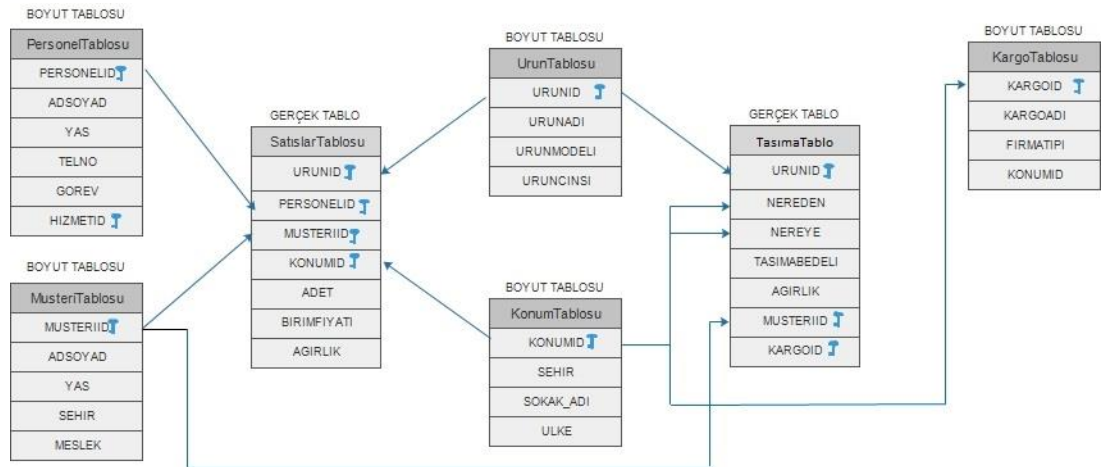
bir sorguyu yürütmek için daha fazla bağlantı gerektireceği için tarama işleminin etkinliğini azaltabilir. Sonuç olarak, sistemin performansı olumsuz etkilenebilir. Kar taneciği şeması veri tekrarlarını azaltırken, performansı olumsuz etkilediği için veri ambarı tasarımında yıldız şeması kadar popüler değildir (Han, Kamber, Pei 2012: 140). Aşağıda örnek bir kar tanesi şeması verilmiştir.



Şekil 7. Kar tanesi tablo yapısı

3.2.2.1.3. Takımyıldız Şemalarıdır (Fact Constellations)

Bu tablo türünde birden fazla gerçek tablo bulunur ve bu gerçek tablolar ortak boyut tablolarıyla birbirine bağlanır. Görüntü itibariyle birden fazla yıldız şemasının bir araya gelerek oluşturduğu yıldız kümesine benzemektedir. Bu nedenle galaksi şeması ya da takım yıldızı olarak adlandırılır.



Şekil 8. Takım yıldızı tablo yapısı

3.2.3. Veri Keşfi

İş zekâsı piramidinin 3. aşaması olan veri keşfi pasif iş zekâsı aşaması olarak da tanımlanmaktadır. Bu aşamada verilerden enformasyon çıkarmak yerine üst yönetimin hipotezini destekleyecek veriler veya örüntüler olup olmadığını incelemek için veriler üzerinde basit sorgular, raporlama analizleri ve istatistiksel metotlar kullanılır (Vercellis 2009: 113). Bilgi keşfi kavramı veri madenciliğiyle benzerlik göstermekle beraber kullandığı analiz yönteminin basitliği ve ulaştığı veya ulaşmaya çalıştığı bilgilerin yüzeyselliği bakımından farklı bir yöntemdir.

3.2.4. Veri Madenciliği

İş zekâsı piramidinin 4. Aşaması olan veri madenciliği aktif iş zekâsı olarak adlandırılır. Bu aşamada amaç veriden bilgi ve enformasyon çıkarmadır (Vercellis 2009: 77). Veri keşfinden farklı olarak bu aşama analiz öncesinde karar vericinin herhangi bir öngöründe bulunmasını gerektirmez. Veri madenciliğinin amacı karar vericiye yeni bilgiler sunmaktır. Veri madenciliği ile ilgili literatürde birçok tanım olmakla birlikte bu tanımlardan birkaçı şöyledir. Veri madenciliği konusuna 4.3. bölümde ayrıntılı olarak değinilecektir.

Agiu ve diğerlerine (2014) göre veri madenciliği gizli örüntü ve ilişkiler bularak, analitik modeller oluşturarak, sınıflama ve öngörülerini tanımlayarak elde edilen sonuçların sunulmasını kapsayan bir süreçtir. Başka bir deyişle veri madenciliği geçmiş verileri inceleyerek gelecek hakkında faydalı bilgiler üretme işlemidir.

Reinschmidt ve Francoise'e göre veri madenciliği, büyük veri yığınlarından, öncesinde bilinmeyen, geçerli ve anlaşılabilir bilgileri çıkarma ve bunları belirgin bir amaca dair önemli kararlar vermede kullanma işlemidir.

Veri madenciliğinin tanınmış uzmanlarından Piatetsky-Shapiro'ye göre veri madenciliği, verilerden daha önceden keşfedilmemiş, muhtemelen faydalı enformasyonun çıkartılması işlemi olarak tanımlamaktadır. Bu süreç kümeleme (Clustering), veri özetleme (Data Summarization), sınıflama kurallarının (Classification Rules) Öğrenilmesi, bağımlılık ağlarının (Dependency Networks)

bulunması, deęişikliklerin analizi (Analysing Changes) ve anomali tespiti (Detecting Anomaly) gibi farklı birçok teknik yaklaşımı kapsamaktadır (Akpınar 2000).

3.2.5. Optimizasyon

İş zekâsı piramidinin beşinci seviyesinde optimizasyon modelleri yer alır. Bu aşamada karar vericiler elde edilen bilgilerle en iyi seçeneęi belirlemeye çalışırlar. Karar verici eldeki problem için uygun alternatifleri belirler. Sonrasında alternatifleri deęerlendirmek ve karşılaştırmak için maliyet ya da kâr gibi bir ölçüt belirler. Bu noktada, karar verici en iyi alternatifi bu ölçütler çerçevesinde belirlemek durumundadır, yani ya en yüksek kârı ya da en düşük maliyetli alternatifi seçecektir.

3.2.6. Karar

İş zekâsı piramidinin son aşaması ise karar aşamasıdır. Bu aşamada karar alıcılar önceki aşamalardan elde edilen enformasyonlar, kendi tecrübeleri ve uzmanlık bilgileri ışığında nihai bir karar alırlar.

4. VERİ MADENCİLİĞİ

Bir iş zekâsı bileşeni olarak veri madenciliğiyle ilgili birçok tanım yapılmış olmakla birlikte Tezcanlar'a (2007: 69) göre veri madenciliği daha önceden bilinmeyen, anlamlı ve uygulanabilir bilgilerin geniş veri tabanlarından elde edilmesidir. Kalikov'a (2006: 7) göre ise VM sorgularla büyük miktarda olan veriden yararlı bilgiler, desenler ve eğilimler (genelde önceden belli olmayan) çıkarabilmektedir.

Veri madenciliğinde kullanılan modeller, tahmin edici (Predictive) ve tanımlayıcı (Descriptive) olmak üzere iki ana başlık altında incelenmektedir.

Tahmin edici modellerle, örüntüleri tanımlanmamış verilerden hareket edilerek bir örüntü modeli oluşturulur ve oluşturulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerleri tahmin edilmeye çalışılır. Örneğin bir firma şubelerinin satış verilerine ve şubelerin bulunduğu şehirlerin özelliklerini içeren bir veri setine sahip olabilir. İşletme bu verilerle satışlardaki ve taleplerdeki dönemsel değişimi görebilir ve buna göre şubelere göndereceği ürün miktarını tahmin edebilir. Bu örnekte bağımsız değişken şubelerin satışlarındaki değişim bağımlı değişken ise şubelere gönderilecek ürün miktarlarıdır. Bu modele uygun olarak kurulan model şubelere gönderilecek dönemsel ürün miktarını belirlemekte kullanılabilir. Tahmin edici modeller başlığı altında sınıflama ve regresyon modelleri kullanılmaktadır.

Tanımlayıcı modellerde ise eldeki verilerle ilgili herhangi bir sonuç bilgisi yoktur. Burada amaç bu veriler içerisindeki örüntüleri ortaya çıkarmaktır. Örneğin geliri 2000 ile 5000 arasında evli ve çocuklu bireylerin satın alma örüntüleri ile geliri 2000'in altında ama bekar bireylerin satın alma örüntülerinin benzerlik gösterdiğinin belirlenmesi tanımlayıcı modellere örnektir. Tanımlayıcı modeller kümeleme ve birliktelik kurallarıdır.

Veri madenciliği birçok işlemten oluşan bir süreçtir. Veri yığınları arasında faydalı enformasyonun elde edilmesinin yanı sıra verileri bu bilgi keşfi işlemine hazır hale getirme süreci de veri madenciliğinin bir parçasıdır. Efektif bir veri keşfi için izlenmesi gereken adımlar aşağıdaki gibidir (Savaş, Topaloğlu, Yılmaz 2012):

- Problemin tanımlanması: Veri madenciliği çalışmalarının en önemli şartlarında birisi bilgi keşfi çalışmasının hangi amaç için yapılacağı ve sonuçların başarı düzeyini ölçmede hangi ölçütlerin esas alınacağı belirlenmesidir.
- Verilerin hazırlanması: Veri madenciliği çalışmalarının ikinci aşaması olan verilerin hazırlık aşamasında veriler ön süreçlerle işlenmeye hazır hale getirilir. Tüm veri keşfi sürecinde harcanan zaman ve eforun %50 ile %80'i bu süreçte harcanmaktadır.
- Modelin kurulması ve değerlendirilmesi: Belirlenen problem için en uygun modelin kurulabilmesi, olabildiğince çok sayıda modelin denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma süreçleri en uygun model bulunana kadar tekrarlanır.
- Modelin kullanılması: Kurulan ve başarısı kabul edilen model kullanıma alınır.
- Modelin izlenmesi: Zamanla modele konu olan değişkenlerde meydana gelen değişiklikler kurulan modelin sürekli izlenmesini ve güncellenmesini zorunlu kılar.

4.1. Veri Madenciliğinde Kullanılan Modeller

4.1.1. Sınıflama ve Regresyon

Sınıflandırma ve regresyon, yeni bir kayıtın özellikleri esas alınarak, önceden tanımlanmış sınıflardan uygun olanına dahil edilmesidir. Veri madenciliğinde sınıflandırması yapılan her kayıta belirli bir sınıf etiketi atanır. Akpınar'a (2000) göre sınıflama ve regresyon esas amaç mevcut verilerden hareket ederek geleceğin tahmin edilmesidir. Sınıflama ve regresyon modelleri veri madenciliği teknikleri içerisinde en yaygın kullanıma sahip olan modellerdir. Sınıflama ve regresyon modelleri arasındaki temel fark, tahmin edilen bağımlı değişkenin kategorik veya süreklilik gösteren bir değere sahip olmasıdır. Sınıflama kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır. Ancak çok terimli lojistik regresyon (multinomial logistic regression) gibi kategorik değerlerin

de tahmin edilmesine olanak sađlayan tekniklerin geliřtirilmesiyle, her iki model giderek birbirine yaklařmakta ve bunun bir sonucu olarak aynı tekniklerden yararlanılması m¼mk¼n olmaktadır. Sınıflama ve regresyon modellerinde kullanılan bařlıca teknikler ařađıda verilmiřtir:

- Bellek Temelli Nedenleme (Memory Based Reasoning),
- Lojistik Regresyondur (Logistic Regression),
- Karar Ađađları (Decision Trees),
- K-En Yakın Komřu (K-Nearest Neighbor),
- Destek Vekt¼r Makineleri (Support Vector Machines),
- Genetik Algoritmalar (Genetic Algorithms),
- Naive-Bayes,
- Lineer Regresyon (Linear Regression).

Bu tez alıřmasında yukarıda belirtilen sınıflama regresyon modellerinde Naive-Bayes, Yapay Sinir Ađları (Artificial Neural Networks), Destek Vekt¼r Makineleri (Support Vector Machines), ve Lineer Regresyon (Linear Regression) kullanılmıřtır.

4.1.1.1. Naive Bayes

Bayes teoreminden faydalanılarak oluřturulan Naive Bayes algoritması anlařılması ve uygulanması kolay bir sınıflandırma algoritmasıdır. Olasılık esaslı alıřan bu algoritma hedef nesnenin belirli bir sınıfa ait olma olasılıđını belirtmektedir. Bayes form¼l¼ eřitlik 1’de verilmiřtir.

$$P(H|X) = \frac{P(X|H)*P(H)}{P(X)} \quad (1)$$

Eřitlikte H bir ¼zniteliđin S gibi bir sınıfa ait olma olasılıđını ifade eden hipotezdir. X giriř vekt¼r¼n¼ $P(H|X)$ ise ardıl olasılıđı temsil eder. Naive Bayes sınıflandırma algoritması ise řu řekilde alıřmaktadır;

V ’nin veri setini temsil ettiđini ve V ’deki her X ’in sınıf etiketinin tanımlı olduđunu varsayalım. X , n tane ¼znitelikten oluřan bir vekt¼rd¼r ve $X = (x_1, x_2, \dots$

x_n) şeklinde ifade edilmektedir. Ayrıca S_1, S_2, \dots, S_m ile temsil edilen m tane sınıf olduğunu varsayalım. Naive Bayes sınıflandırıcısı bir X vektörünün S_i sınıfına ait olup olmadığını bulmak için, bütün sınıflar içinde en yüksek $P(S_i|X)$ olasılığına sahip değeri bulmaya çalışır. Bu durum için Bayes teoremi eşitlik 2’de ifade edilmiştir.

$$P(S_i|X) = \frac{P(X|S_i) * P(S_i)}{P(X)} \quad (2)$$

$P(X)$ değeri tüm sınıflar için aynı olduğundan burada amaç $P(X|S_i) * P(S_i)$ ifadesini maksimize etmektir. $P(S_i)$ ifadesi, S_i sınıfındaki eleman sayısının, tüm eleman sayısına oranını ifade etmektedir. $P(X|S_i)$ ifadesi ise, X ’in n tane değer içeren bir öznitelik vektörü olduğu varsayıldığında aşağıdaki eşitlik ile hesaplanır.

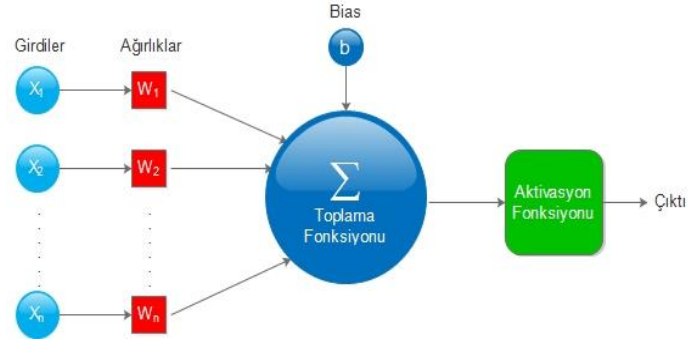
$$P(H|S_i) = \prod_{k=1}^n P(X_k|S_i) \quad (3)$$

Sonuçta, sınıflandırıcı en büyük $P(X_k|S_i)P(S_i)$ ifadesine sahip olan S_i sınıfını, X vektörünün sınıfı olarak seçer.

4.1.1.2. Yapay Sinir Ağları

İnsan sinir sisteminden esinlenilerek oluşturulan yapay sinir ağları (YSA), deneme yolu ile öğrenme yapan bir gözetimli öğrenme algoritmasıdır. YSA, veri kümeleri içerisindeki bilinmeyen ve fark edilmesi güç örüntüleri ortaya çıkarmak amacıyla kullanılır. Ancak bunun için YSA modelinin çok sayıda girdi ve girdilere ilişkin çıktı dizisiyle eğitilmesi gerekmektedir. YSA, insan beyninin fonksiyonel özelliklerine benzer şekilde, optimizasyon, öğrenme, analiz, genelleme, sınıflandırma ve ilişkilendirme gibi konularda başarılı bir şekilde uygulanmaktadır. YSA, sadece girdi ve çıktı değişkenleriyle ve arasındaki herhangi başka bir ön bilgiye ihtiyaç duymadan ve herhangi bir varsayımda bulunmadan doğrusal olmayan modellemeyi sağlayabilmektedir. Ağa, eğitim için girdi bilgileri ve bu girdilere karşılık gelen çıktı bilgileri verilmekte ve ağın girdi-çıkıtı arasındaki ilişkiyi öğrenmesi sağlanmaktadır (Yavuz, Deveci 2012).

Basit bir YSA modeli Şekil 9’da görüldüğü üzere girdi katmanındaki nöronların ağırlık ve toplama fonksiyonu üzerinden çıktı katmanına bağlanması ile oluşturulur.



Şekil 9. YSA modeli

Literatürde farklı YSA modelleri bulunmakla birlikte en sık kullanılanı çok katmanlı algılayıcı sinir ağı (MLP) modelidir. Bu model girdi katmanı, gizli katman ve çıktı katmanı olmak üzere üç katmandan oluşmaktadır. Girdi katmanı verilerin alındığı katmandır. Her bir nöron veri setindeki bir özneliği temsil ettiği için öznelik sayısı kadar girdi nöronu vardır. Çıktı katmanı ise sınıfların belirlendiği katmandır. Bu katman oluşturulmak istenen sınıf sayısı kadar nörondan oluşur. Gizli katman ise girdi katmanı ile çıktı katmanı arasında verilerin ara işleme maruz kaldığı katmandır. Gizli katman sayısı ve bir gizli katmandaki nöron sayısı tam olarak belirli olmamakla birlikte eğitimin kalitesini etkileyen önemli iki faktördür. MLP modelinde öğrenme bir katmandan sonraki katmana doğru yapıldığı için ileri beslemeli bir YSA modelidir. Kullanılan eğitim algoritması hatayı en aza indirecek şekilde ağırlıkları güncellemeyi hedefler.

$$y_i = f(\sum_{j=1}^n x_j * w_{ji}) \quad (4)$$

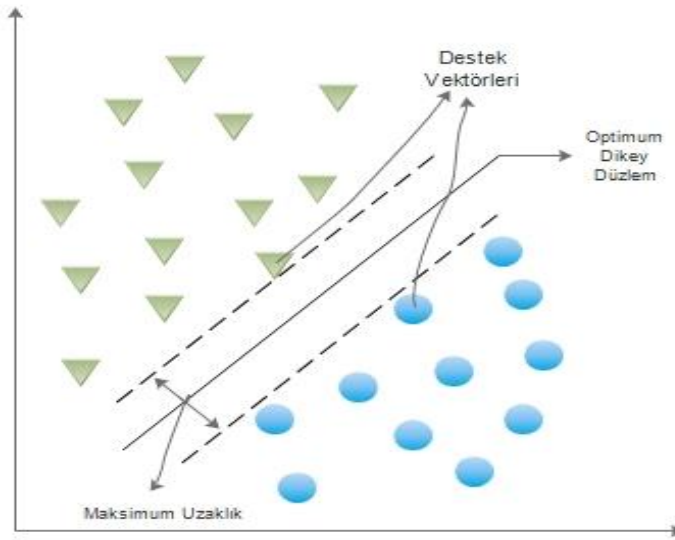
$$E(w) = \frac{1}{2} \sum_{k \in outputs} (t_k - o_k)^2 \quad (5)$$

Eşitlik 4'te MLP modelinde ileri besleme, eşitlik 5'te ise geri beslemeyi formüle edilmiştir. Eşitlik 4'de x_j mevcut katmandaki j 'inci nöronun değerini, y_j takip eden katmandaki i 'inci nörona aktarılan değeri, n mevcut katmandaki nöron sayısını w_{ji} ; mevcut katmandaki j 'inci nörondan takip eden katmandaki nörona giden ağırlığı, f ise aktivasyon fonksiyonunu (sigmoid, gauss, softmax vb.) temsil etmektedir. Eşitlik 5'de k ; veri setindeki örnek sayısını, t_k ; verilerin gerçek sınıfını, o_k ise modelin üretmiş olduğu sınıf değerini temsil etmektedir.

4.1.1.3. Destek Vektör Makineleri

İlk olarak Vapnik (2013) tarafından önerilen eğri uydurma probleminin çözümü ve sınıflama için kullanılan Destek Vektör Makineleri (DVM), istatistiksel öğrenme kuramından faydalanan bir makine öğrenmesi yöntemidir. Yaygın olarak lineer ayrışabilen verileri sınıflamada kullanılan bu yöntem, çekirdek fonksiyonları (gaussian, rbf vb.) yardımı ile lineer ayrışamayan verileri sınıflamak için kullanılabilir. DVM'deki ana amaç gözlemlenen örnekleri minimum hata ile iki sınıfa ayıran hiper düzlemi belirlemektir. Bu amaç doğrultusunda iki adet destek vektörü şekil 10'da gösterildiği gibi seçilir ve bu iki vektör arasındaki uzaklık maksimum yapılır. Son adımda ise yeni gelen örnek için eşitlik 6 kullanılarak bir sınıf değeri hesaplanır. Bu sınıf değeri 0'dan büyük olursa yeni örnek ilk sınıfa diğer durumlarda ise ikinci sınıfa atanır ve algoritma sonlandırılır.

$$y = w^t x + b \quad (6)$$



Şekil 10. Destek Vektör Makineleri ve Hiper Düzlem Seçimi

Eşitlik 6'da x ; yeni örnek için öznitelik vektörünü, w^t ; belirlenen hiper düzeleme dik olan ağırlık vektörünü, b ; sabit eşik değerini temsil etmektedir. DVM diğer makine öğrenmesi yöntemlerinden farklı olarak sadece iki sınıfa sahip verileri ayırmakta kullanılmaktadır. Daha fazla sınıflı veri setlerini ayırtmak için bire karşı bir (OVO, one versus one), ya da bire karşı hepsi (OVA, one versus all) olmak üzere iki farklı DVM tekniği geliştirilmiştir. OVO tekniğinde her bir sınıf ikilisi için,

toplamda n 'nin (toplam sınıf sayısının) ikili kombinasyonu kadar DVM modeli oluşturulur ve son adımda bu modeller istatistiksel yöntemler yardımı ile birleştirilir. OVA tekniğinde ise rastgele bir sınıf seçilerek ilk sınıf olarak ve geriye kalan diğer tüm sınıflar ikinci sınıf olarak atanarak iki sınıflı bir DVM modeli oluşturulur. Daha sonra bu işlem tüm sınıflar için ayrı ayrı uygulanarak toplamda n adet model oluşturulur. Son adımda OVO tekniğinde olduğu gibi oluşturulan modeller istatistiksel yöntemler yardımı ile birleştirilerek algoritma sonlandırılır.

4.1.1.4. Lineer Regresyon

Basit anlamı ile Lineer Regresyon (LR) bir değişkenin skorunu ikinci bir değişken kullanarak tahmin etmeyi hedeflemektedir. Burada tahmin edilmek istenen değer ölçüt değeri (criterion variable) olarak adlandırılırken diğer değişken tahmin edici değeri (predictor variable) olarak adlandırılmaktadır. Bu yöntemde ölçüt değeri bir tahmin edici değişkene bağlı olabileceği gibi, birden fazla tahmin edici değere de (değişken vektörüne) bağlı olabilmektedir. Bu amaç doğrultusunda gözlemlenen verileri en az hata ile tahmin eden lineer bir fonksiyonun parametreleri tahmin edilmektedir. Bir lineer fonksiyon eşitlik 7'deki gibi gösterilmektedir.

$$y = w_0 + x_1w_1 + x_2w_2 + \dots + x_nw_n \quad (7)$$

Bu eşitlikte y ; ölçüt değerini, n ; değişken sayısını, x_n ; n 'inci değişken için gözlemlenen değeri; w_n ise n 'inci değişken için tespit edilen ağırlık değerini temsil etmektedir. Bu yöntemdeki temel amaç eşitlik 7'de gösterilen denklem için tüm ölçüt değerlerini minimum hata ile tahmin edecek ağırlıkları tespit etmedir. Bunun için ilk olarak bir maliyet fonksiyonu belirlenir (en küçük kareler yöntemi vb.). Ardından ağırlık değerleri rastgele ya da istatistiksel yollarla belirlenir ve belirtilen maliyet fonksiyonu belirli bir eşik değerinin altına inene kadar bu ağırlık değerleri güncellenir. Yaygın olarak devamlı değerleri tahmin etmek için kullanılan LR; fonksiyon sonucunda bulunan değerlerin belirli aralıktaki eşik değerlerine bağlı olarak kategorize edilerek sınıflama yöntemi olarak da kullanılabilir.

4.1.2. Kümeleme

Kümeleme analizi en basit açıklamasıyla benzer özellik gösteren veri elemanlarının kendi aralarında gruplanması işlemidir. Hasan'a (2011: 34) göre

kümeleme, bir dizi örüntüyü (gözlem sonuçları, veri nesnelere ve özellik vektörleri) ayrık ve homojen gruplar oluşturacak şekilde gruplandırma işlemidir. Bu işlem örüntülerin benzerlik derecelerine göre kümeler ayrılmasıyla gerçekleştirilir. Kümeleme işlemi sonucunda elde edilen her başarılı (geçerli) küme içinde yer alan nesnelere arasında maksimum benzerlik ve kümeler arasında ise maksimum farklılık oluşması amaçlanır.

Boutsinas ve Gnardellis'in (2002) tanımına göre kümeleme, yaygın olarak kullanılan bir tekniktir ve amacı, bir dizi desenin ayrık ve homojen kümeler bölünmesidir. Kümeleme algoritmaları, makine öğrenmesi, sinir ağları ve istatistik gibi çeşitli alanlarda geniş bir şekilde incelenmiştir. Ayrıca, veri madenciliği, mühendislik, taksonomi, istatistiksel veri analizi ve iş uygulamaları da dahil olmak üzere birçok alanda kullanılmıştır.

Kümeleme algoritmaları özel(exclusive) ve gözetimsiz(unsupervised) bir algoritmalarlardır. Özel(exclusive) işlemlerde her nesne sadece bir kümeye ait olabilir. Özel Olmayan(nonexclusive) işlemlerde ise her nesne birden fazla kümeye dâhil edilebilmektedir. Gözetimli algoritmalarda nesnelere önceden tanımlanır ve her nesnenin bir sınıf etiketi vardır. Yeni eklenen veya sınıfı belirlenmemiş nesnelere sınıflandırılması için kullanılır. Kümeleme analizi gibi, gözetimsiz algoritmalarda ise nesnelere önceden belirlenmiş bir sınıf etiketi yoktur. Bu tip sınıflandırmalarda grupların belirlenmesi için yakınlık matrisi (proximity matrix) ile beraber nesnelere özellikleri de kullanılır

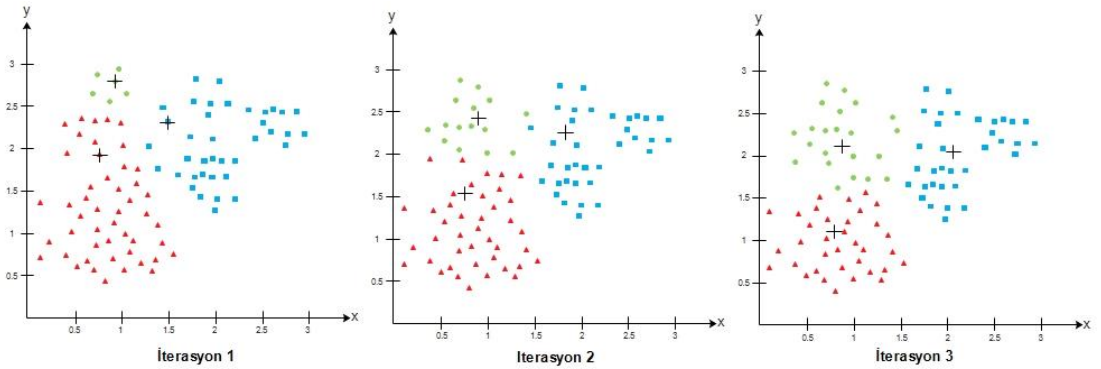
Literatürde çok sayıda kümeleme algoritması bulunmaktadır. Kümeleme algoritmasının seçiminde esas olan veri setindeki verinin tipi ve bilgi keşfinin hangi amaç için yapıldığıdır. Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir (Han, Kamber, Pei 2012: 449):

- Model tabanlı yöntemler (Model-based methods)
- Izgara tabanlı yöntemler (Grid-based methods)
- Bölme yöntemleri (Partitioning methods)
- Yoğunluk tabanlı yöntemler (Density-based methods)
- Hiyerarşik yöntemler (Hierarchical methods)

4.1.2.1. K-Ortalama Algoritması

D veri seti öklityen uzayda tanımlı n tane nesneden oluşan bir veri setidir. K-Ortalama kümeleme yöntemi D veri seti altındaki nesnelere k tane kümeye (C_1, \dots, C_k) bölmeyi amaçlamaktadır. $C_i \subset D$ ve $C_i \cap C_j = \emptyset$. K-Ortalama kümeleme algoritmasının amacı küme içindeki elemanlarının benzerliğini maksimize etmeyi amaçlarken kümeler arasındaki benzerliği minimize etmeye çalışır (Han, Kamber, Pei 2012: 451).

Merkeze dayalı kümeleme tekniği olan K-Ortalama algoritması her kümenin merkez noktasının kümeyi temsil ettiği fikrine dayanmaktadır. Bu algoritmayı uygulamak için öncelikle K değeri kullanıcı tarafından belirlenir. Sonrasında her kümenin merkez noktasını veya ortalamasını temsil etmek üzere k adet nesne rasgele seçilir. Kalan diğer değerler, kümelerin merkez noktalarına yani ortalama değerlerine olan uzaklıkları dikkate alınarak en yakın oldukları kümelere dahil edilirler. Daha sonra, her bir küme için yeni ortalama değerler hesaplanarak yeni küme merkezleri belirlenir ve tekrar nesnelere küme merkezlerine uzaklıkları incelenir. Herhangi bir değişim olmayıncaya kadar algoritma tekrar tekrar uygulanır. K-Ortalama yönteminde en yaygın olarak kullanılan uzaklık hesaplama formülü Öklit uzaklık formülüdür.



Şekil 11. K-Ortalama Algoritması (k=3)

K-Ortalama yönteminin başarısı başlangıç merkezleri olarak seçilen nesnelere bağlıdır. Başlangıç merkezlerinin belirlenmesinde birçok yaklaşım olmakla birlikte bazıları şöyledir (Hasan 2011: 37):

- k kadar rastgele nesne seçilerek küme merkezleri olarak atanır.
- Veriler rastgele k tane kümeye atanır ve küme ortalamaları alınarak ilk küme merkezleri belirlenir.
- En uç değerlere sahip veriler küme merkezleri olarak seçilir.
- Veri setinin merkezine en yakın noktalar başlangıç noktaları olarak seçilir.

K-Ortalama yönteminin başarısının değerlendirilmek için kullanılan birçok yöntem bulunmakla birlikte en yaygın kullanılan karesel hata toplamı SSE'dir (sum of squared error). Nesnelerin buldukları kümenin merkez noktalarına olan uzaklıklarının karelerinin toplamı olarak tanımlanabilen SSE değeri ne kadar düşükse kümeleme işleminin başarısı o kadar yüksektir. SSE değeri eşitlik 8'deki formülle hesaplanmaktadır (Han, Kamber, Pei 2012: 451).

$$E = \sum_{i=1}^k \sum_{p \in C_i} dist(p, m_i)^2 \quad (8)$$

p : C_i kümesindeki bir nesne,

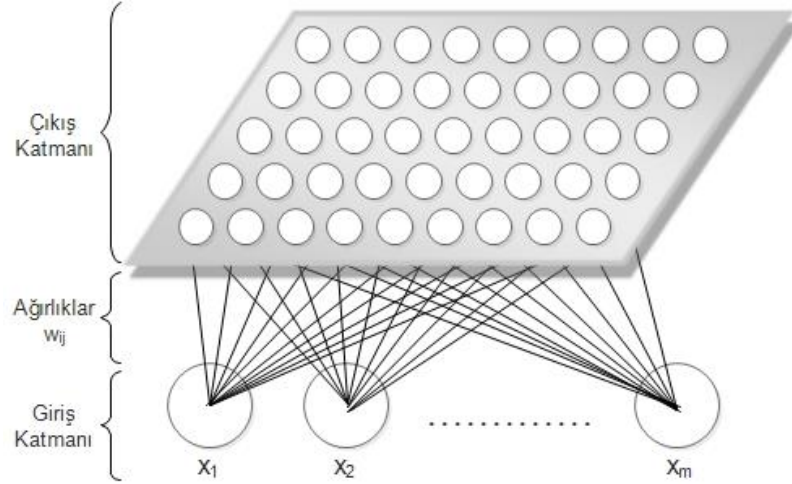
m_i : C_i kümesinin küme merkezi,

Bu formülle k tane kümenin kendi içindeki yoğunluğu incelenir.

4.1.2.2. Öz Örgütlenmeli Ağlar (ÖÖA)

Öz Örgütlenmeli Ağlar (Self-Organized Maps), 1982 yılında Teuvo Kohonen tarafından geliştirilmiştir (Kohonen 1982). Bu nedenle, Kohonen ağları olarak da bilinir. Yüksek boyutlu verilerin iki boyutlu örgü üzerine doğrusal olmayan istatistiksel yöntemlerle haritalanmasını sağlar. Şekil 11'de görüleceği üzere öz örgütlenmeli ağlar verilerin ağı sunulduğu bir giriş katmanı ile bir çıkış katmanından oluşmaktadır. Çıkış katmanında yer alan her bir nöron giriş katmanındaki nöronlara ağırlık vektörleriyle bağlıdır. Ağı giriş katmanındaki nöron sayısı veri setindeki özellik sayısı kadarken, çıkışında yer alan nöronların sayısı ise analistin isteğine bağlıdır. Çıkış katmanındaki bu nöronların farklı şekiller ve boyutlardaki dizilimlerinden farklı topolojilere sahip haritalar oluşturulur. Doğrusal, dikdörtgen, altıgen ve kübik yapıda olabilen bu dizilimlerden en çok kullanılanı dikdörtgen yapıdaki haritalardır. Nöronların farklı dizilimleri sadece topolojik komşuluğu etkiler ve bu nöronlar arasında giriş nöronlarına olan bağlantı gibi herhangi bir fiziksel

bağlantı yoktur. Topolojik komşuluk sadece ağın eğitimi sırasında yarışmayı kazanan ağırlık vektörleriyle birlikte komşu vektörlerin de ağırlıklarının değiştirilmesinde etkilidir.



Şekil 12. Öz Örgütlenmeli Ağlar (ÖÖA)

Öz Örgütlenmeli Ağlar (ÖÖA) gözetimsiz eğitim kullanan, yapısal olarak ileri beslemeli özel bir yapay sinir ağı türüdür. Gözetimli öğrenme algoritmalarında giriş ve çıkış arasındaki ilişkinin öğrenilmesi için ağa önceden öğrenilmiş giriş-çıkış çiftleri sunulur. Gözetimsiz öğrenme algoritmalarında ise bu giriş çıkış çiftlerine ihtiyaç yoktur. Ağın eğitimi için sadece giriş vektörleri yeterlidir. Ağın öğrenmesi, çıkış nöronlarına ait ağırlık vektörlerinin giriş vektörüne olan uzaklıklarına bağlı olarak yarışması esasına dayanır. Yarışmayı kazanan nöronla birlikte o nörona komşu nöronların da ağırlık değerleri değiştirilerek, ağırlık vektörlerinin girişteki verilerin oluşturduğu küme merkezlerine doğru hareketi sağlanır ve eğitim sonunda kümeleri temsil edecek referans vektörleri elde edilir. Sadece kazanan nöronun değil bu nörona komşu diğer nöronların da kazananla birlikte ağırlıklarının değiştirilmesi neticesinde birbirine benzer desenlerin uzaysal kümelenmesi sağlanmaktadır. Böylece n boyutlu bir giriş uzayında birbirine benzer kümeler, 2 boyutlu çıkış uzayında birbirlerine komşu olmaktadır.

ÖÖA algoritmasının uygulama adımları yarışmacı, işbirlikçi ve adaptasyon süreçlerinden oluşur. İlk adımda n tane çıkış nöronuna ait ağırlık vektörleri (w_j) rasgele değerler atanarak oluşturulur (Eşitlik 9).

$$w_j = w_{j1} \ w_{j2} \ w_{j3} \ \dots \ w_{jm} \quad i = 1,2,3, \dots n \quad (9)$$

İkinci aşamada giriş (x) vektörlerinde birisi rasgele seçilerek ağa sunulur (Eşitlik 10).

$$x = x_1 \ x_2 \ x_3 \ \dots \ x_m \quad (10)$$

Sonrasındaki aşama yarışma aşamasıdır. Bu aşamada her bir nörona ait ağırlık vektörü ile giriş vektörü arasındaki uzaklık değerleri (d_j) hesaplanarak karşılaştırılır (Eşitlik 11). En küçük uzaklığa sahip nöron yarışmayı kazanır ve bu nöron en uygun birim (Best matching Unit (BMU)) olarak da adlandırılır.

$$|X_i - W_j| = \sqrt{\sum_{i=1}^N (x_i - w_{ij})^2} \quad (11)$$

Algoritmanın dördüncü aşaması olan işbirlikçi süreçte sadece kazanan nöronun ağırlıkları değil aynı zamanda kazanan nörona topolojik olarak komşu olan nöronların da ağırlıkları güncellenir. Giriş vektörüne yaklaşacak şekilde ağırlıkları değiştirilen bu nöronların bir sonraki adımda kazanma olasılıkları daha da artırılmış olur. Yeni ağırlık değerleri eşitlik 12 yardımıyla hesaplanır.

$$W_j(t + 1) = W_j(t) + \theta(t)\alpha(t)(X(t) - W_j(t)) \quad (12)$$

Eşitlik 12’te $W_j(t + 1)$ bir sonraki adımdaki ağırlık vektörünü, $W_j(t)$ şu anki ağırlık vektörünü, X giriş vektörünü, (t) zamanla değişen öğrenme oranını, $\theta(t)\alpha(t)$ hem zamana hem de kazanan nöronla komşu nöronlar arasındaki uzaklığa bağlı olarak değişen topolojik komşuluk fonksiyonunu göstermektedir. Ağırlığı değişecek birimler kazanan birim ve komşuluk yarıçapı içerisindeki diğer birimlerdir. Komşu birimlerin güncellenmesi sırasında kazanan vektöre yakınlıkları göz önünde bulundurulur ve güncellenme işleminde bu yakınlık oranına da dikkat edilir. Komşuluk yarıçapını gösteren eşitlik 13’deki fonksiyon zamanla azalan bir yapıya sahiptir. Eşitlikte \check{R}_0 başlangıç yarıçapını, T_1 zaman sabitini, t ise o anki çevrimi göstermektedir.

$$\check{R} = \check{R}_0 \exp\left(-\frac{t}{T_1}\right) \quad t = 0,1,2, \dots \quad (13)$$

Eğitim sürecinde ağırlıklar sürekli güncellenerek komşuluk yarıçapı içerisinde kalan nöronların sayısı azaltılmaya ve eğitimin sonunda kazanan nöronun

yalnız bırakması amaçlanır. Böylelikle eğitimin başında çıkış uzayının geniş bir bölümünün etkileşime uğrayarak daha büyük kümelerin oluşması sağlanırken eğitim sonuna doğru çok daha küçük boyutlu ayrışmaların oluşması gerçekleştirilir.

4.1.3. Öznitelik Seçim Algoritmaları

4.1.3.1. Ki – Kare

χ^2 şeklinde de gösterilen ki-kare testi 1992’de Randy Kerber’in daha sonrasında 1995’te Huan Lui ve Rundy Setiono’nun katkılarıyla geliştirilmiştir. Bu yöntem değişkenlerin veri setini temsil edip edemediğini belirlemek amacıyla kullanılır (Kavzoğlu, Şahin, Cölkesen 2014). Ki-kare testinde H_0 ve H_1 olmak üzere iki hipotez vardır. H_0 değişkenlerin uygunsuzluğunu ifade ederken H_1 değişkenlerin veri setini temsil etmek için yetersiz olduğunu ifade eder. Ki-kare testi iki adımdan oluşan bir testtir. İlk adımda, bağlı buldukları sınıflara göre gözlemlenen değerlerin ki-kare istatistiği hesaplanır. χ^2 değeri sıfır ile pozitif sonsuz arasında değerler alabilir. Eğer bu değer sıfıra yakınsa, gözlemlenen frekans değerleri ve beklenen frekans değerleri benzer şekilde yakındır. Eğer bu değer yüksekse, gözlemlenen frekans değerleri ve beklenen frekans değerleri önemli ölçüde farklılık gösterir. Bu nedenle, ikinci aşamada χ^2 değeri Ki-kare dağılımından belirlenen bir eşik değeri ile karşılaştırılmıştır. Bu eşik değeri anlamlılık derecesine ve serbestlik derecesine göre belirlenir. Anlamlılık derecesi, boşluk modelini (null model) kullanarak eşğin üstünde bir ki-kare istatistiği elde etme ihtimalini temsil eder ve serbestlik derecesi, analiz edilen özelliklerin sayısından bir çıkararak hesaplanır. Eğer hesaplanmış değer belirlenen değerden büyükse H_1 hipotezi kabul edilir. Değilse H_0 hipotezi kabul edilir. Ki-kare istatistiğinin formülü eşitlik 14’de gösterilmektedir.

$$\chi^2 = \sum_{i=1}^n \frac{(o_i - e_i)^2}{e_i} \quad (14)$$

Bu eşitlikte, n veri setindeki öznitelik sayısını, o_i i ’inci öznitelik için gözlemlenen frekans değerini ve e_i i ’inci öznitelik için beklenen frekans değerini temsil etmektedir.

4.1.3.2. Bilgi Kazancı

Bilgi kazancı, veri setinin özniteliklere bölünmesi durumunda tahmini kaybın hesaplanması için kullanılan entropi tabanlı metriklerden biridir (Ozarkar, Patwardhan 2013). Entropi, sistemin düzensizliğini veya belirsizliğini ifade eden bir değerdir. Yüksek bir entropi değeri, sistemin yüksek oranda bilgi içerdiğini gösterir. Bilgi kazancı metriğini hesaplamak için ilk aşamada, belirli bir veri kümesinin sınıf etiketleri için entropi değeri, eşitlik 15'te formüle edildiği gibi hesaplanır.

$$E = - \sum_{i=1}^n \left(\log_2 \frac{fs(i)}{N} \right) * \frac{fs(i)}{N} \quad (15)$$

Bu eşitlikte, n veri setindeki sınıf sayısını, $fs(i)$ i 'inci sınıfın örneklem sayısını ve N toplam örneklem sayısını temsil etmektedir. Bilgi kazancı metriğini hesaplamamanın ikinci aşamasında, her öznitelik için entropi değeri hesaplanır ve bu yeni entropi değeri, her öznitelik için bilgi kazancını hesaplamak amacıyla eşitlik 15'de bulunan değerden çıkarılır. Her bir öznitelik için entropi değerinin hesaplanmasında kullanılacak formül eşitlik 16'te ve bilgi kazancı değerinin hesaplanmasında kullanılacak formül eşitlik 17'te gösterilmiştir.

$$E(i) = \sum_{k=1}^n \frac{ss(k)}{N} * \sum_{m=1}^{cs} \frac{csk(k,m)}{ss(k)} * \left(\log_2 \frac{csk(k,m)}{ss(i)} \right) \quad (16)$$

$$B(i) = E - E(i) \quad (17)$$

Eşitliklerde, $E(i)$ i 'inci öznitelik için entropi değerini, n i 'inci öznitelik altındaki farklı değerlerin sayısını, $ss(k)$ i 'inci özniteliğin k 'inci değerine ait örneklem sayısını, N veri setindeki toplam örneklem sayısını, cs veri setindeki sınıf sayısını, $csk(k,m)$ i özniteliğinin, k değişkeninin ve m sınıfına ait örneklem sayısını temsil etmektedir. Ayrıca $B(i)$ i 'inci öznitelik için bilgi kazancını ve E eşitlik 15'de hesaplanan değeri temsil etmektedir. Yukarıdaki eşitliklere göre bilgi kazancı ne kadar büyükse ilgili öznitelik o derece önemlidir.

4.1.3.3. Kazanım Oranı

Kazanım oranı, veri setinin özniteliklere bölünmesi durumunda tahmini kaybın hesaplanması için kullanılan bir diğer entropi tabanlı metriklerden biridir. Veri kümesindeki bir öznitelik çok farklı değerlere sahip olduğunda, her bir değere düşen örneklerin sayısı söz konusu öznitelik için düşüktür. Bu nedenle, bu öznitelik

için hesaplanan entropi değeri küçüldükçe ve bilgi oranı artar. Bilgi kazancı yönteminde de açıklandığı gibi, entropi değerinin büyük olması, değişkenin veri kümesini tanımlamada iyi olduğunu gösterir. Eğer bir öznitelik altında çok farklı değerler varsa, bilgi kazancı yöntemleri bu özneliği iyi bir ayırıcı olarak seçer. Sistem eğitim veri setini iyi ezberlese dahi test setini bu özneliği kullanarak doğru şekilde ayıramaz. Bu soruna bir çözüm olarak kazanım oranı, bilgi kazancını her öznitelik için bölümlene bilgileri ile normalleştirir. Eğer kazanım oranı yüksekse bu özneliğin iyi bir ayırıcı olduğu kabul edilir. Bölümlene bilgi değerinin nasıl hesaplandığı eşitlik 18’te verilmiştir. Ayrıca eşitlik 19 kazanım oranının hesaplanmasını göstermektedir.

$$S(i) = - \sum_{k=1}^n \frac{ss(k)}{N} * (\log_2 \frac{ss(k)}{N}) \quad (18)$$

$$K(i) = \frac{B(i)}{S(i)} \quad (19)$$

Yukarıdaki eşitliklerde, $S(i)$ i ’inci öznitelik için bölümlene bilgi değerini, n i ’inci öznitelik altındaki farklı değerlerin sayısını, $ss(k)$ i ’inci özneliğin k ’inci değerine ait örneklemlerin sayısını, N veri setindeki toplam örneklem sayısını, $K(i)$ i ’inci öznitelik için kazanım oranını temsil etmektedir. Ayrıca $B(i)$ eşitlik 17’te tanımlanan bilgi kazancı değerini temsil etmektedir.

4.1.3.4. ReliefF

Kira ve Rendell (1992) tarafından 1992 yılında geliştirilen Relief algoritması öznitelikleri aralarındaki ilişkiye göre ağırlıklandırılan bir öznitelik seçim yöntemidir. Bu yöntem iki sınıfı olan veri setleri için başarılı sonuç verse de çoklu sınıfa sahip veri setleri için çalışmamaktadır. Bu problemi çözmek için 1994 yılında Kononenko, çoklu sınıfı olan veri setleri içinde çalışan ReliefF algoritmasını geliştirmiştir (1994). Bu yöntemin ilk aşamasında tüm özniteliklerin ağırlıkları 0 olarak belirlenir. Daha sonra her bir adımda veri setinden rastgele bir veriyi seçer ve bu veri ile aynı sınıfa ait en yakın k (k değeri sınıf sayısının bir eksigidir) adet veri bulur ardından her bir farklı sınıfa ait en yakın veriler bulunur. Sonrasında her bir özelliğe ait ağırlıklar bu veriler kullanılarak güncellenir. Son aşamada belirlenen koşulu sağlamayan özellikler veri setinden atılarak yeni veri seti oluşturulur. Eşitlik 20’da reliefF

algoritması, Eşitlik 21’da ikili değerler için uzaklık hesaplanması, Eşitlik 22’de ise devamlı değerler için uzaklık hesaplanması formülüne edilmiştir.

$$W(x^a) = W(x^a) - \frac{\sum_{j=1}^k \text{uzaklık}(A, R_i, H_j)}{m \times k} + \frac{\sum_{C \neq \text{sınıf}(R_i) \left[\frac{P(C)}{1 - P(\text{sınıf}(R_i))} \times \sum_{j=1}^k \text{uzaklık}(A, R_i, M_j) \right]}{m \times k} \quad (20)$$

$$\text{uzaklık}(A, I_1, I_2) = \begin{cases} 0, & I_1 = I_2 \\ 1, & I_1 \neq I_2 \end{cases} \quad (21)$$

$$\text{uzaklık}(A, I_1, I_2) = |I_1 - I_2| \times \frac{1}{\max(A) - \min(A)} \quad (22)$$

Bu eşitliklerde $W(x^a)$ a ’ncı özniteliğin ağırlığını, k sınıf sayısının bir eksiğini, m döngü sayısını, R_i i ’inci döngüde seçilmiş olan veriyi, H_j seçilen veri ile aynı sınıfa ait j ’inci yakın veriyi, M_j j ’inci sınıfa ait seçilen veriye en yakın olan veriyi temsil etmektedir.

5. MÜŞTERİ İLİŞKİLERİ YÖNETİMİNDE VERİ MADENCİLİĞİ

5.1. RFM Analizi

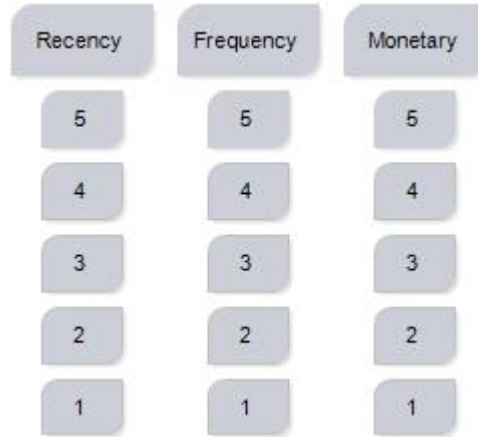
Müşteri segmentasyonunda iki temel yaklaşım vardır. Bunlardan birincisi müşterilere ait yaş, cinsiyet, memleket vb. özellikleri esas alan demografik yaklaşım, ikincisi ise müşterilerin tüketim davranışlarını esas alan davranışsal yaklaşımdır. Davranışsal yaklaşım müşterilerin tüketim alışkanlıklarını inceleyerek işletme için karlılıklarını ve gelecek dönemlerdeki taleplerini ön görmeyi mümkün kılmaktadır.

RFM analizi müşterilerin tüketim davranışlarını inceleyerek söz konusu müşterileri bu davranış şekillerine göre kümelemeyi amaçlayan bir yaklaşımdır. RFM harflerinin sıralaması rastgele olmayıp temsil ettiği kavramın önem sırasını ifade eder ve sırasıyla Recency, Frequency ve Monetary kavramlarına karşılık gelmektedir. Recency müşterinin en yakın harcama tarihi ifade eder, yakın tarihte harcama yapmış olan bir müşterinin nispeten uzak bir tarihte harcama yapmış olan bir müşteriye göre tekrar harcama yapma ihtimali daha yüksek kabul edilir (Birant 2011). Frequency incelenen zaman aralığındaki harcama sıklığını ifade eder harcama sıklığı yüksek olan müşterilerin ihtiyaçlarını karşılamada araştırmayı yapan işletmeyi yoğun olarak tercih ettiği ve işletme açısından bu müşteri grubunun değerinin yüksek kabul edilebilir. Monetary incelenen dönem içerisindeki toplam harcama miktarı anlamına gelir, işletmeler için karlılık önemli bir faktör olduğu için müşterinin işletme ürünlerine harcadığı para önem arz etmektedir.

RFM modelinin uygulanabilmesi için müşterilere ait üç temel veri gerekmektedir. Bunlar; müşterileri birbirinden ayırt etmeyi sağlayacak bir müşteri numarası, harcama yapılan tarih ve harcama miktarıdır. Bu bilgiler ışığında öncelikle her bir müşteri için en yakın işlem tarihi yani recency değeri belirlenir. Sonra her bir müşterinin işletmeye belirlenen dönem aralığında kaç kez geldiği yani frequency değeri bulunur. Son olarak yine her bir müşterinin belirlenen dönem aralığında toplam harcaması yani monetary değeri bulunur. RFM değerlerinin bulunmasının ardından bu değerler gruplandırmayı kolaylaştırmak amacıyla puanlanır. Mevcutta

Bağımsız (Independent) ve İççe (Nested) olmak üzere iki tip RFM gruplama türü bulunmaktadır.

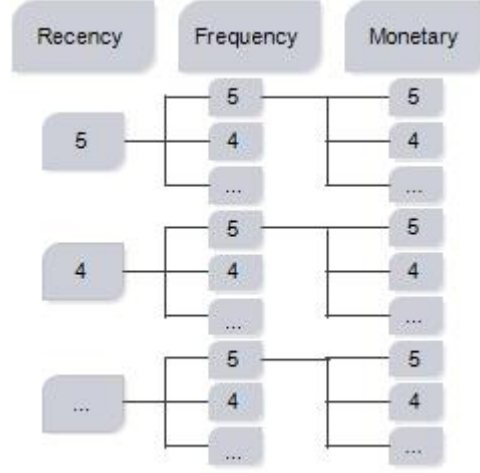
Bağımsız RFM Gruplama: Bağımsız RFM gruplamada R, F ve M faktörleri ayrı puanlanır. RFM değerlerinin bulunmasının ardından recency değerleri için artan, frequency ve monetary değerleri için azalan sırayla veriler sıralanır. Sıralanan verilerin ilk %20'lik kısmı 5, sonraki %20'lik kısmı 4, sonraki %20'lik kısım 3 olacak şekilde 5'den 1'e kadar puanlanır. Son olarak her bir R puanı 100 herbir F puanı 10 ve herbir M puanı 1 ile çarpılır. Bu şekilde her bir R, F ve M faktörünün analizdeki ağırlığı belirlenmiş olur. Oluşan RFM puanları en fazla 555 en az 111 olacak şekilde maksimum $5 \times 5 \times 5 = 125$ farklı şekilde oluşabilir. Her bir müşteri için oluşan bu RFM değerleri o müşterinin tüketim davranışına işaret eder. 555 puanlı müşteriler en iyi müşteri olup son harcamasını yakın zamanda yapmış olan, yüksek sıklıkla ve yüksek miktarda harcama yapan müşteri grubudur. Aynı şekilde 111 puanlı müşteriler işletme için en değersiz müşterilerdir. Bu müşteri grubu firmayı çok az ziyaret etmiş çok az para kazandırmış ve uzun zamandır firmayla alışveriş yapmamıştır.



Şekil 13. Bağımsız RFM Gruplama

İç içe RFM Gruplama: İççe RFM gruplamada RFM skorları verilirken öncelikle recency parametresinden başlanır. Recency parametresinin ilk %20'lik kısmı 5, ikinci %20'lik kısmı 4, sonraki %20'lik kısım 3 olacak şekilde 5'den 1'e kadar skorlanır ve 5 recency grubu elde edilir. Daha sonra her bir recency skoruna bağlı frequency parametresi %20'lik bölümlere ayrılarak 1-5 arası değerlerle

puanlanır ve 25 (5x5) frequency grubu elde edilir. Son olarak ise her bir frequency parametresine bağlı monetary parametresi 1-5 arası puanla puanlanır ve 125 (5x5x5) adet monetary grubu oluşur. Hesaplama yöntemi bağımsız RFM gruplama yöntemi göre biraz daha zor olan bu yaklaşımla RFM skorları daha fazla yayılım gösterebilir. Bu sayede bağımsız RFM gruplama yöntemiyle elde edilemeyen yeni gruplar oluşabilir (“Which RFM binning method is best for me?” 2017).



Şekil 14. İç içe RFM Gruplama

RFM analizinde karşılaşılan problem verilerin %20’lik kısımlara ayrılma sürecinde benzer değerlerin farklı gruplara düşmesi durumudur. Örneğin elimizdeki R verilerinin sırasıyla 10,10,10,9,8,8,7,6,4,3 olması durumunda verileri 5 eşit parçaya bölmemiz durumunda 2. Değer olan 10’un R skoru 5 olurken 3. Değer olan 10’un skoru 4 olmaktadır. Bu durum aynı tüketim davranışı göstermelerine rağmen 2 müşterinin farklı şekilde puanlanmasına neden olmakta ve RFM analizi bakımından tutarsızlık doğurmaktadır. Bu puanlama tabanlı RFM analizinde bunu çözmeyin birkaç yolu bulunmakla birlikte efektif değildir.

Yukarıda anlatılan RFM analizinin uygulanma adımları içerisinde puanlama aşaması aslında bir çeşit kümeleme işlemidir. Bu işlem müşterileri tüketim davranışları çerçevesinde kümelemeyi amaçlamaktadır. Bu noktada uzmanlar puanlamadan doğan sorunları çözmek amacıyla puanlama yöntemi yerine kümeleme algoritmaları kullanmayı tercih etmişlerdir.

Rabiei ve Rastegari (Rabiei & Rastegari, 2015) yaptıkları çalışmada RFM analizinde k-means kümeleme algoritmasını kullanarak müşterileri tüketim alışkanlıklarına göre 8 gruba ayırmıştır ve bu müşterilerin firma için yaşam değerini hesaplamışlardır. Son olarak müşterilerin yaşam değerleriyle bankanın yatırım tekliflerine olumlu yanıt verme eğilimleri arasındaki ilişkiyi bulmak için karar ağaçları kullanarak pozitif bir korelasyon bulmuşlardır. Ghazanfari vd. (Ghazanfari, Mohamadi, & Alizadeh, 2008) yaptıkları çalışmada RFM analizi sürecinde k-means ve fuzzy k-means algoritmalarından faydalanmışlardır. Bu iki çalışma neticesinde sonuçların kalitesini ölçmek için üç değerlendirme kriteri kullanarak k-means algoritmasının fuzzy k-meansten daha başarılı sonuçlar elde ettiğini göstermişlerdir. Sohrabi ve Khanlari (Sohrabi & Khanlari, 2007) çalışmalarında R, F ve M değerlerinin veri setindeki ortalama R, ortalama F ve ortalama M değerlerinden küçük ya da büyük olması durumuna dayanarak 8 (2x2x2=8) farklı kümeye ayrılabilceğini belirtmişlerdir. RFM değerlerini k-means kümeleme algoritmasıyla kümelere ayırarak bu kümeler için müşteri yaşam değerini hesaplamışlardır. Son olarak oluşan bu müşteri grupları için müşteri tutma stratejileri önermişlerdir.

RFM analizinde belirtilen küme sayısı veri setine göre ve analizin amacına göre değişiklik gösterebilir. Bu tezde R(FM), F(RM) ve M(RF) değerlerinin genel ortalamasının altında ya da üstünde olması durumu esas alınarak RFM verileri 8 (2x2x2) kümede incelenmiştir.

Kümeleme öncesinde F ve M değerlerinin normalize edilmesi için eşitlik 23 kullanılmıştır.

$$x' = (x - x^S)/(x^L - x^S) \quad (23)$$

R değerinin normalizasyonu için ise sayısal değeri arttıkça müşteri yaşam değerine etkisi azaldığı için eşitlik 24'deki formül kullanılmıştır.

$$x' = (x^L - x)/(x^L - x^S) \quad (24)$$

Formüllerde x normalize edilmek istenen değeri, x^L normalize edilmek istenen elemanın ait olduğu kümenin en büyük elemanını ve x^S ise normalize edilmek istenen elemanın ait olduğu kümenin en küçük elemanını ifade etmektedir.

Müşteri yaşam boyu değerinin (MYBD) hesaplanmasında eşitlik 25'deki formül kullanılmıştır:

$$MYBD = W_R * C_{Rj} + W_F * C_{Fj} + W_M * C_{Mj} \quad (\text{Liu \& Shih, 2005}) \quad (25)$$

Formülde W_R , W_F ve W_M ifadeleri ağırlıklandırılmış R, F ve M değerlerini, C_{Rj} , C_{Fj} ve C_{Mj} ifadeleri ise normalize edilmiş R, F ve M değerlerini ifade etmektedir. Kullanılacak ağırlık değerleri için Liu ve Shih'in (2005) çalışmaları esas alınmıştır. Bu çalışma doğrultusunda R, F ve M değerleri için kullanılan ağırlık değerleri $W_R=0.731$, $W_F=0.188$ ve $W_M=0.081$ şeklindedir.

5.2. Müşteri Kaybı Analizi

Günümüz rekabet yoğun piyasaları işletmeleri müşteri ilişkileri yönetimi yaklaşımını benimsemek zorunda bırakmıştır. Bu yaklaşımın esas amacı müşterilerle karşılıklı çıkara dayalı ilişkiler geliştirerek sonunda müşterinin işletmeye sadakatini sağlamaktır. Müşteri sadakatini sağlamanın işletme için birçok faydası olmakla birlikte bu faydalardan bölüm 2.7.1'de ayrıntılı olarak bahsedilmiştir.

Sadık müşteriler işletmelere pazarda önemli rekabet üstünlüğü sağlamaktadır. Ancak bu müşteriler son teknolojilere erişme imkânı, müşteri dostu çalışanlar, düşük işlem ücretleri, reklâm etkisi, coğrafi yakınlık ve çeşitli hizmet teklifleri vb. nedenlerle sürekli etkileşimde buldukları yani sadakat gösterdikleri bir firmayı terk edebilmektedirler (Farquad, Ravi, Raju 2014). İşletmeler için değerli olan bu müşterilerin ayrılması tabi ki istenmeyen bir durumdur. İşte bu noktada müşteri kaybının önüne geçilmesi için işletmenin ayrılma potansiyeli olan müşterileri ve bu müşterilerin ayrılmasına etki eden faktörleri doğru belirlemesi büyük önem taşımaktadır (Kim, Jun, Lee 2014). Müşteri kaybını etkileyen faktörlerin doğru belirlenmesi karar alıcıların işletmenin aksayan yönlerini görmesini sağlayacak ve işletmenin kendi süreçlerini iyileştirmesine ve geliştirmesine rehberlik edecektir. Ayrıca bu faktörlerin doğru tespiti ayrılma eğilimi olan müşterileri ve bu müşterilerin ayrılma eğilimlerine etki eden faktörleri belirleyerek karar alıcıların bu müşteriler için etkili elde tutma stratejileri belirlemelerine yardımcı olur (Hung, Yen, Wang 2006).

Kiřiođlu ve Topcu (2011) yaptıkları alıřmada Trkiye’deki bir telekomikasyon řirketinden elde ettikleri veriler zerinde BBF (Bayesian Belief Network) algoritmasını kullanarak mřterilerin ayrılmasına etki eden faktrleri incelemiřlerdir. Farquad vd. (2014) yaptıkları alıřmada SVM (Support Vector Machines) ile bir banka kredi kartını kullanan mřterilerin ayrılmasına etki eden faktrleri sınıflamıřlardır. Ayrıca alıřmada kullandıkları veri setinde ayrılan mřterilerle bađlılıklarını devam ettiren mřteriler arasında oransal bir denge olmamasının sınıflamayı olumsuz etkileyeceđi dřncesiyle SMOTE (Synthetic Minority Over-sampling Technique), random under-sampling, random over-sampling gibi dengeleme yntemlerini kullanmıřlardır. Hung vd. (2006) alıřmalarında Tayvan’da abonelerine kablosuz internet hizmeti veren bir firmadan elde ettikleri veriler zerinde karar ađaları kullanarak mřterilerin demografik verilerinin, fatura ve deme kalıplarının, arama detaylarının ve mřteri hizmetleriyle olan etkileřimlerinin ayrılma kararlarına etkisi zerine bir takım kurallar ıkarılmıřlardır. Lejeune (2001) alıřmasında ayrılma ynetiminde veri madenciliđi yntemlerinin nemini vurgulamıřtır. Van den Poel ve Lariviere (2004) alıřmalarında Avrupa’da bankacılık ve sigortacılık alanında faaliyet gsteren bir iřletmenin mřterilerinin ayrılmalarına etki eden faktrleri belirlemeye alıřmıřlardır. Mřterilerin tketicim davranıřı ve demografik zellikleri esas alınarak yapılan alıřmada PHM (proportional hazard modeling) yntemini uygulamıřlardır. Coussement ve Van den Poel (2008) Belika’da faaliyet gsteren gnlk bir gazetenin aboneleri zerinde gerekleřtirdikleri alıřmada abonelere ait mřteri/řirket etkileřim kayıtlarını, kayıt yenileme ile ilgili bilgileri, sosyo-demografik bilgileri ve aboneliđe dair aıklayıcı bilgileri kullanmıřtır. Bu alıřmada SVM (Support Vector Machines), RF (random forest) ve LR (logistic regression) yntemleri kullanılmıřtır.

6. UYGULAMA

6.1. RFM Müşteri Segmentasyonu

Çalışmada RFM analizi çerçevesinde müşterilerin R (belirlenen dönem içerisinde en son işlem tarihi), F (belirlenen dönem içerisinde işlem sıklığı) ve M (belirlenen dönem içerisinde harcanan topla para) özellikleri kullanılmıştır.

Çalışmada müşteriler alışveriş davranışlarına göre ÖÖA ve K-Means algoritmalarıyla kümelenmiştir. Müşterinin R, F ve M değerinin ortalama R, F ve M'den küçük ya da büyük olması durumuna göre her bir özellik için iki seçenek ve bu seçeneklerin kombinasyonu bakımından da 8 (2x2x2) davranış tipi vardır. Birant (Birant, 2011) çalışmasında bu davranış kalıplarını aşağıdaki tablo 1'deki gibi tanımlamıştır.

Tablo 1. Müşteri davranış modelleri

Kümelere	RFM Davranış Modeli	Müşteri Tipi	Kümelere	RFM Davranış Modeli	Müşteri Tipi
Küme-1	R↑F↑M↑	En İyiler	Küme-5	R↓F↑M↑	Ayrılanlar
Küme-2	R↑F↓M↑	Değerliler	Küme-6	R↓F↑M↓	Sık Gelenler
Küme-3	R↑F↑M↓	Sadıklar	Küme-7	R↓F↓M↑	Çok Harcayanlar
Küme-4	R↑F↓M↓	İlk Kez Gelenler	Küme-8	R↓F↓M↓	Belirsizler

Ancak Birant çalışmasında kümeleme için RFM skorlarını kullandığı için R değeri büyüdükçe önemi artmaktadır. Bu çalışmada ise RFM skorları yerine normalize edilmiş gerçek RFM değerleri kullanılmıştır. Bu nedenle bu tez çalışmasında RFM davranış modeli tablo 2'deki gibidir.

Tablo 2. Müşteri davranış modelleri (Tez çalışmasında esas alınan)

Kümelere	RFM Davranış Modeli	Müşteri Tipi	Kümelere	RFM Davranış Modeli	Müşteri Tipi
Küme-1	R↓F↑M↑	En İyiler	Küme-5	R↑F↑M↑	Ayrılanlar
Küme-2	R↓F↓M↑	Değerliler	Küme-6	R↑F↑M↓	Sık Gelenler
Küme-3	R↓F↑M↓	Sadıklar	Küme-7	R↑F↓M↑	Çok Harcayanlar
Küme-4	R↓F↓M↓	İlk Kez Gelenler	Küme-8	R↑F↓M↓	Belirsizler

Ayrıca hesaplanan müşteri yaşam boyu değerleri sadeleştirmek maksadıyla önem derecesini ifade edecek şekilde 1-8 aralığında skorlanmıştır. 1 en değerli 8 en değersiz müşteri sınıfını temsil etmektedir. Tablo 3'te hesaplanan MYBD'leri ve skor karşılıkları verilmiştir.

Tablo 3. Müşteri yaşam boyu değeri

	Müşteri Yaşam Boyu Değeri		Müşteri Yaşam Boyu Skoru	
	ÖÖA	K-Means	ÖÖA	K-Means
Küme-1	0,066765	0,221511	8	7
Küme-2	0,217501	0,701154	7	3
Küme-3	0,354021	0,066765	6	8
Küme-4	0,478065	0,863553	5	1
Küme-5	0,589742	0,367429	4	6
Küme-6	0,66836	0,743829	3	2
Küme-7	0,723132	0,611365	2	4
Küme-8	0,822267	0,500999	1	5

Müşteri gruplarının oluşturulmasında ÖÖA ve K-Means kümeleme algoritmaları kullanılmıştır. Bu iki algoritma neticesinde elde edilen kümeler, oluşan RFM davranış modelleri, müşterilerin kümelere ve bu davranış modellerine farklılık göstermektedir. Kümeleme algoritmalarında küme içerisindeki elemanların olabildiğince homojen ve kümelerin birbirine göre ise olabildiğince heterojen olması arzulanır. Bu nedenle uygulanan bu iki kümeleme algoritmasının başarısının değerlendirilmesinde aynı küme içerisindeki elemanların merkeze uzaklıklarının olabildiğince düşük, küme merkezleri arasındaki uzaklığında olabildiğince yüksek olması faktörleri esas alınmıştır. Bu amaçla her bir algoritma için küme elemanlarının merkeze uzaklıkları toplamı (c2etot) ve kümeler arasındaki toplam uzaklık (c2c) hesaplanmıştır. Daha sonra ise her iki algoritma da onar kez çalıştırılarak tablo 4'de gösterilen c2c ve c2etot değerleri elde edilmiştir. Elde edilen sonuçlar içerisinde ortalama c2c ve ortalama c2etot'ye en yakın olan ÖÖA algoritması için 1'inci, K-Means için ise 8'inci sonuç kullanılmıştır.

Seçilen iki sonuca göre K-Means algoritmasıyla üretilen sonuçta küme merkezleri arasındaki uzaklık nispeten daha fazla iken, ÖÖA algoritmasıyla elde

edilen sonuca göre her bir kümenin elemanlarının küme merkezine daha yakın olması daha yoğun kümeler oluştuğunu göstermektedir. Bu noktada her iki algoritma neticesinde küme merkezleri arasındaki uzaklıklar karşılaştırıldığında farkın çok küçük olması buna rağmen ÖÖA neticesinde oluşan kümeler daha yoğundur. Bu nedenle bu kümeleme işleminde ÖÖA algoritmasının daha iyi sonuç ürettiği söylenebilir.

Tablo 4. Kümeleme Yöntemleri Değerlendirme Tablosu

	SOM		K-Means	
	c2c	c2etot	c2c	c2etot
1	13,0929	9,1502	13,6583	5,7715
2	13,1026	9,2507	13,8577	16,9256
3	13,0793	9,0853	12,4642	49,0828
4	13,0261	9,1519	13,2196	7,5644
5	13,1602	9,1189	14,2448	33,1379
6	12,9042	9,2405	13,4296	15,8844
7	13,1623	9,359	13,4976	10,212
8	13,1574	8,9245	13,4452	18,4386
9	13,1811	9,0988	13,1238	19,5325
10	13,2308	8,9719	13,5868	7,1434
Ortalama	13,10969	9,13517	13,45276	18,36931

Aşağıda tablo 5 ve 6'da ÖÖA ve K-ortalama algoritmalarıyla oluşturulmuş 8 müşteri sınıfı, bu müşteri sınıflarının ortalama R, F, M değerleri, her bir müşteri sınıfının müşteri yaşam boyu değeri ve müşteri davranış modeli verilmiştir.

Tablo 5. ÖÖA kümeleme algoritmasına göre RFM gruplama

	ÖÖA					Müşteri Davranış Modeli
	MYB Skoru	Eleman Sayısı	Ortalama R	Ortalama F	Ortalama M	
Küme-1	8	268	340,15	1,21	372,95	R↑F↓M↓
Küme-2	7	336	263,29	1,27	486,89	R↑F↓M↓
Küme-3	6	365	194,03	1,77	662,63	R↑F↓M↓
Küme-4	5	378	131,10	2,23	798,26	R↑F↓M↓
Küme-5	4	716	74,40	2,56	1052,51	R↓F↓M↓
Küme-6	3	1072	34,55	2,91	1126,42	R↓F↓M↓
Küme-7	2	1163	10,02	7,08	3245,35	R↓F↑M↑
Küme-8	1	22	6,77	55,55	79464,14	R↓F↑M↑
Genel		4320	93,08	3,85	1919,17	

SOM algoritması sonucunda, oluşabilecek 8 davranış modeli içerisinde R↑F↓M↓, R↓F↓M ve R↓F↑M↑ şeklinde 3 tanesi oluşmuştur.

Küme-7 ve Küme-8'deki müşteriler RFM davranış modeli bakımından en iyi (best) müşterilerdir. Küme-8'deki müşteriler işletmenin toplam müşterilerinin %5'ni oluşturmasına rağmen toplam kazancın yaklaşık %21'i bu müşteri sınıfından sağlanmaktadır. Ayrıca bu müşteriler alışveriş sıklıklarının yüksekliği ve yakın zamana kadar işletmeyle etkileşimlerini devam ettirmiş olmaları bakımından da işletme için en değerli müşteri sınıfıdır. Müşteri yaşam boyu değerine baktığımız zamanda 8 müşteri sınıfı içerisinde birinci sırada olmaları bu durumu onaylamaktadır. Küme-7'deki müşterilere baktığımız zaman Küme-8'deki müşterilerle benzer tüketim davranışı göstermelerine rağmen müşteri yaşam boyu değeri bakımından Küme-8'den sonra gelmektedir. İşletmenin toplam müşterilerinin yaklaşık %27'sini oluşturan bu müşteri sınıfı toplam kazancın yaklaşık %45,5'unu sağlaması bakımından değerlidir.

Küme-5 ve Küme-6'daki müşteriler ilk kez gelenler (first time) olarak adlandırılan işletmeyle ilk defa etkileşimde bulunmuş müşteri sınıfıdır. Alışveriş sıklığı ve harcama miktarı bakımından benzerlik gösteren bu iki grup son etkileşim tarihi bakımından farklılık göstermektedir. Toplam müşteri sayısının yaklaşık %41'ini oluşturan bu müşteriler toplam kazancınsa yaklaşık %23,5'unu karşılamaktadır. İşletmeyle bir ya da iki kez etkileşimde bulunup tekrar ziyaret

etmeyen ancak son etkileşimi yakın zamanlı olan bu tip müşterilerin doğru pazarlama stratejileriyle işletmeyle etkileşiminin artırılma ihtimali vardır. Bu müşterilerin yaşam boyu değerlerinin tüm müşteriler içerisinde 3. ve 4. sırada olması taşıdıkları potansiyele işaret etmektedir.

Küme-1, Küme-2, Küme-3 ve Küme-4'deki müşteriler ise belirsizler (uncertain) olarak adlandırılan müşterilerdir. Bu kümeler işletmeyle ortalama 1 ya da 2 kez etkileşimde bulunmuş, yakın zamana kadar tekrar işletmeyi ziyaret etmemiş ve harcama miktarı düşük olan bu müşterilerden oluşmaktadır. İşletme müşterilerinin yaklaşık %31'ini oluşturan bu müşteri sınıfı toplam kazancın yaklaşık %10'unu sağlamaktadır.

Tablo 6. K-Ortalama kümeleme algoritmasına göre RFM gruplama

	K-Means					Müşteri Davranış Modeli
	MYB Skoru	Eleman Sayısı	Ortalama R	Ortalama F	Ortalama M	
Küme-1	7	356	261,24	1,27	482,83	R↑F↓M↓
Küme-2	3	1702	18,65	3,95	1515,45	R↓F↑M↓
Küme-3	8	268	340,15	1,21	372,95	R↑F↓M↓
Küme-4	1	11	3,36	73,73	107366,74	R↓F↑M↑
Küme-5	6	402	187,25	1,86	676,91	R↑F↓M↓
Küme-6	2	233	8,76	18,00	11254,81	R↓F↑M↑
Küme-7	4	931	63,40	2,61	1084,90	R↓F↓M↓
Küme-8	5	417	119,48	2,33	849,23	R↑F↓M↓
Genel		4320	93,08	3,85	1919,17	

K-Means algoritması sonucunda, oluşabilecek 8 davranış modeli içerisinde R↑F↓M↓, R↓F↑M↓, R↓F↑M↑ ve R↓F↓M↓ şeklinde 4 tanesi oluşmuştur.

Küme-4 ve Küme-6'daki müşteriler RFM davranışları bakımından en iyiler (best) olarak değerlendirilen müşterilerdir. Küme-4 ve Küme-6'daki müşteriler tüketim davranışı bakımından benzerlik göstermelerine rağmen Küme-4'deki müşteriler müşteri yaşam boyu değeri bakımından Küme-6'dakilerden daha değerlidirler. Küme-4'deki müşteriler işletme müşterilerinin yaklaşık %3'ünü oluştururken toplam kazancın yaklaşık %14'ü bu müşterilerden sağlanmaktadır.

Küme-6'daki müşterilerse toplam müşterilerin yaklaşık %5'ini oluştururken toplam kazancın yaklaşık %32'si bu müşteri grubundan sağlanmaktadır.

Küme-7 ilk kez gelenler (first time) olarak bilinen işletmeyle az sayıda etkileşim gerçekleştirmiş, bireysel harcamaları düşük müşterilerden oluşmaktadır. Toplam müşteri sayısının yaklaşık %21,5'ünü oluşturan bu müşteri sınıfı toplam kazancın yaklaşık %12'sini sağlamaktadır. Müşteri yaşam boyu değeri sıralamasında 4. sırada bulunan bu müşteriler uygun pazarlama stratejileriyle işletmeye bağlanabilir.

Küme-2 ise sadıklar (shopper) olarak adlandırılan bireysel harcama oranı düşük ancak yakın zamana kadar işletmeyi ziyaret etmiş ve harcama sıklığı yüksek müşterilerden oluşmaktadır. Toplam müşterilerin yaklaşık %39'unu oluşturan bu müşteriler toplam kazancın yaklaşık %31'ini sağlamaktadır. Ayrıca müşteri yaşam boyu değerine göre 3. sırada bulunan bu müşteriler işletme için mevcut durumları bakımından da potansiyelleri bakımından da değerlidirler. Uygun müşteri cezbetme ve derinleştirme stratejileri ile bu müşterilerin işletmeye sağlayacakları faydalar arttırılabilir.

Küme-1, Küme-3, Küme-5 ve Küme-8'deki müşteriler ise RFM davranışları çerçevesinde belirsiz (uncertain) olarak tanımlanan müşterilerdir. Bu kümeler işletmeyle ortalama 1 veya 2 kez etkileşimde bulunmuş, yakın zamana kadar tekrar işletmeyi ziyaret etmemiş ve harcama miktarı düşük olan bu müşterilerden oluşmaktadır. Toplam müşterilerin yaklaşık %33'ünü oluşturan bu müşteriler toplam kazancın yaklaşık %11'ini sağlamaktadır.

6.2. Müşteri Kaybı Analizi

Çalışmada, bir Telekom şirketi müşterilerinin işlem kayıtlarına ait 19 öznitelik ile yine aynı müşterilerin firmayı tercih etmekten vazgeçtiklerini veya tercih etmeye devam ettiklerini gösteren 2 farklı sınıf etiketi kullanılmıştır ("Churn in Telecom's dataset"). Bu veri setindeki öznitelikler tablo 7'de özetlenmektedir.

Tablo 7. Veri seti öznitelik bilgileri

Öznitelik No	Öznitelik Açıklama
1	Müşterinin ikamet ettiği bölge
2	Bölge alan kodu
3	Hesap kullanım süresi
4	Uluslararası plana dâhil olup olmama durumu
5	Telesekreter planına dâhil olup olmama durumu
6	Gönderilen sesli mesaj sayısı
7	Toplam gündüz kullanım dakikası
8	Toplam gündüz arama sayısı
9	Toplam gündüz ücret
10	Toplam akşam kullanım dakikası
11	Toplam akşam arama sayısı
12	Toplam akşam ücret
13	Toplam gece kullanım dakikası
14	Toplam gece arama sayısı
15	Toplam gece ücret
16	Toplam uluslararası kullanım dakikası
17	Toplam uluslararası arama sayısı
18	Toplam uluslararası ücret
19	Müşteri servisi arama sayısı

Bahsedilen veri setinde 4667 müşteriye ait kayıt bulunmaktadır. Bu kayıtların %75'lik kısmı yani 3500 adet veri eğitim verisi olarak, geriye kalan %25'lik kısmı yani 1167 adet veri ise test verisi olarak rasgele belirlenmiştir. Modellerin karşılaştırılması sırasında veri setinden doğacak üstünlüklerin engellenmesi için bölme işlemi bir kez yapılmış ve her model için aynı eğitim ve test verileri kullanılmıştır. Oluşturulan bu veri setleri üzerinde sınıflama ve öznitelik önem sırası belirleme olmak üzere iki farklı analiz yapılmıştır. Sınıflama analizleri için YSA, NB, LR ve MLP algoritmaları matlab platformu kullanılarak yapılmış, önem sırası

belirleme analizleri için ise Ki-kare, BK, KO ve reliefF metrikleri Weka platformu kullanılarak elde edilmiştir. İlk olarak sınıflama analizleri için her bir model sonucunda elde edilen karmaşıklık matris değerleri Tablo 7’de gösterilmiştir. Tabloda P, firmayı tercih etmekten vazgeçen müşteri sayısını; N, firmayı tercih etmeye devam eden müşteri sayısını; eğitim bloğu, modeli eğitmek için kullanılan veri sayılarını; test bloğu, modeli test etmek için kullanılan veri sayılarını; gözlenen bloğu, verilerin gerçek etiketlerini; tahmin bloğu ise verilerin model tarafından tahmin edilen etiketlerini temsil etmektedir. Aynı tabloda N sütunu ile N satırının kesiştiği bölgedeki sayılar TN değerlerini, N sütunu ile P satırının kesiştiği bölgedeki sayılar FN değerlerini, P sütunu ile P satırının kesiştiği bölgedeki sayılar TP değerlerini, P sütunu ile N satırının kesiştiği bölgedeki sayılar ise FP değerlerini ilgili model için göstermektedir.

Tablo 8. Sınıflama sonucunda elde edilen karmaşıklık matrisleri

			Eğitim		Test	
			Gözlenen		Gözlenen	
			N	P	N	P
Naive Bayes	Tahmin	N	1651	260	559	103
		P	171	1419	77	427
Yapay Sinir Ağları	Tahmin	N	1826	85	607	55
		P	92	1498	50	454
Logistic Regression	Tahmin	N	1696	215	588	74
		P	294	1296	107	397
Destek Vektör Makineleri	Tahmin	N	1911	0	662	0
		P	0	1590	504	0

Tablo 8’da test verisinden elde edilen veriler kullanılarak başarı oranı(A), hassasiyet(S), duyarlılık(P) ve F ölçütü(F) değerleri hesaplanmıştır. Bu hesaplama işlemi için eşitlik 26, 27, 28 ve 29’deki denklemler kullanılmıştır.

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (26),$$

$$S = \frac{TP}{TP+FN} \quad (27),$$

$$P = \frac{TP}{TP+FP} \quad (28),$$

$$F = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (29).$$

Hesaplama sonucunda elde edilen değerler tablo 9’de gösterilmiştir.

Tablo 9. Test Verisi Sonuç Değerleri

Sınıflayıcı	A	S	P	F-Ölçütü
NB	0,845	0,847	0,805	0,825
YSA	0,909	0,900	0,891	0,896
LR	0,844	0,787	0,842	0,814
DVM	0,567	0	0	0

İki model arasından ilgili ölçüt biriminde en yüksek değere sahip model en iyi model olarak değerlendirilmiş ve her bir ölçüt birimi için bu değerlendirme yapılmıştır.

Tablo 9’deki sonuçlardan da görüldüğü üzere tüm ölçüt birimlerinde YSA iyi sonuçları veren modeller olmuştur. İkinci analiz olarak eğitim veri seti üzerinde bahsedilen 4 farklı metrik değerleri hesaplanarak önem sırasına göre öznitelikler yeniden sıralanmıştır. Her bir model için önem sırasına göre sıralanan öznitelikler Tablo 10’da gösterilmiştir.

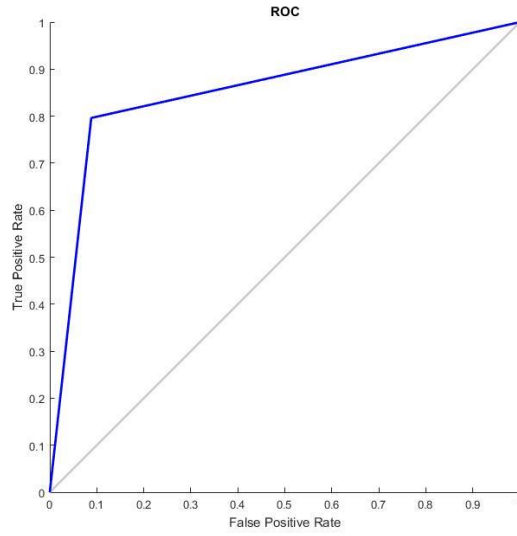
Tablo 10. Özniteliklerin Yönteme Bağlı Önem Sırası

Kullanılan	Öznitelik Sıralaması
Ki-Kare	7, 9, 19, 4, 6, 5, 17, 3, 2, 8, 10, 18, 11, 16, 15, 14, 13, 12, 1
BK	7, 9, 19, 4, 6, 5, 17, 3, 2, 8, 10, 18, 11, 16, 15, 14, 13, 12, 1
KO	7, 9, 19, 4, 6, 5, 17, 3, 2, 8, 10, 18, 11, 16, 15, 14, 13, 12, 1
ReliefF	7, 9, 17, 4, 19, 8, 12, 10, 1, 6, 16, 18, 11, 15, 13, 5, 2, 3, 14

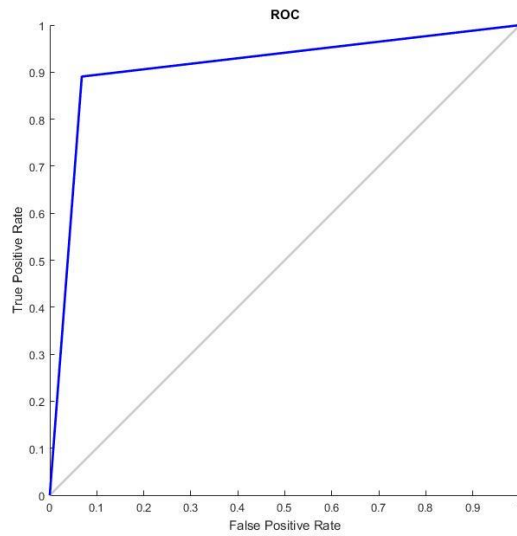
Öznitelik önem sırası analiz sonuçları incelediğinde ilk 3 algoritma birebir aynı sonuçları verirken reliefF algoritmasında bazı farklılıklar gözlenmektedir. Ancak genel hatlarıyla bu sonuçlar yorumlandığında, akşam, gece, uluslararası

kullanım ve ücret bilgileri ile kullanıcının ikamet ettiği bölgenin diğer özniteliklere oranla müşteri kaybı analizine etkisinin az olduğu anlaşılmaktadır.

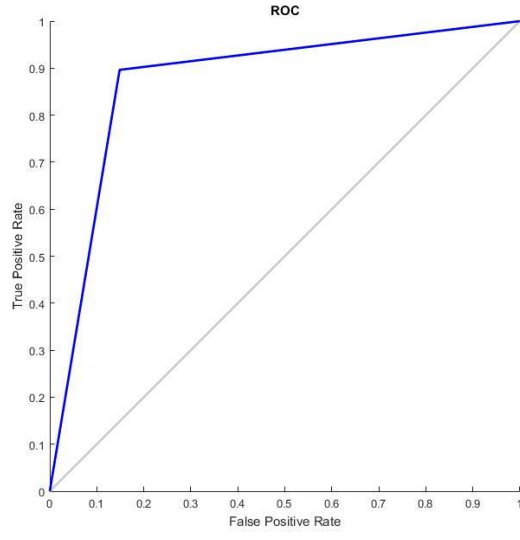
Şekil 15, 16 ve 17’de lineer regresyon, yapay sinir ağları ve naive bayes için roc eğrileri verilmiştir. DVM için ise sonuçların kötü olmasından dolayı roc eğrisi oluşturulmamıştır.



Şekil 15. Lineer Regresyon ROC Eğrisi



Şekil 16. Yapay Sinir Ağları ROC Eğrisi



Şekil 17. Naive Bayes ROC Eğrisi

7. SONUÇ

Geçtiğimiz yüzyıl içerisinde teknolojiye yaşanan gelişmeler ve bu gelişmelerinde etkisiyle ortaya çıkan küreselleşme piyasadaki rekabeti arttırırken işletmelerin odağını üründen müşteriye çevirmiştir. Müşterilerin işletme odağına yerleşmesi geleneksel işletme-müşteri ilişkisini değiştirmiştir. İşletme-müşteri ilişkisindeki bu yeni anlayışa göre işletmeler müşterilerin talep ve ihtiyaçlarını belirleyerek onlara daha iyi hizmet vermeye ve sonuç olarak sadakatlerini kazanmaya çalışmaktadır. İşletmelerin müşterilerle karşılıklı faydaya dayanan ilişkiler geliştirme ihtiyacı pazarlamanın bir alt disiplini olan müşteri ilişkileri yönetimini ortaya çıkarmıştır.

Müşteri ilişkileri yönetimi stratejilerinin etkinliği işletmenin müşterilerini doğru tanımasıyla doğru orantılıdır. Yani bilgi arttıkça MİY uygulamalarının etkinliği de artmaktadır. Son dönemde teknoloji de yaşanan gelişmeler ve teknolojinin günlük hayatın bir parçası haline gelmesi neticesinde insanlar gün içerisindeki her faaliyetlerinde dijital bir iz bırakmakta yani veri üretmektedir. Potansiyel ya da mevcut müşterilerin bıraktığı bu dijital izler birçok amaçla kullanıldığı gibi müşteri ilişkileri yönetimiyle ilgilenen uzmanlarında ilgisini çekmektedir. Çeşitli kaynaklardan toplanan bu bilgiler çeşitli veri madenciliği yöntemleri ya da istatistiki yöntemlerle incelenmektedir. Bu tez çalışmasında da amaç müşterilere ait veriler üzerinde veri madenciliği teknikleri uygulayarak faydalı bilgiler ortaya çıkarmak ve bu bilgilere uygun pazarlama stratejileri önermektir.

Bu çalışmada ilk olarak bir e-ticaret sitesi üzerinden alışveriş yapan müşteriler son alışveriş tarihleri, alışveriş sıklıkları ve harcadıkları para miktarları esas alınarak 8 kümeye ayrılmışlardır. Bu aşamada diğer çalışmalardan farklı olarak ÖÖA ve K-Means algoritmaları birlikte kullanılmış ve hangi algoritmanın daha başarılı sonu ürettiği tespit edilmeye çalışılmıştır. Kümeleme sonrasında oluşan müşteri gruplarının işletme için değerleri belirlenmiş, ayrıca söz konusu müşterilerin tüketim davranışları modellenmiştir. Tespit edilen müşteri değerleri işletmenin o müşteri grubunu elinde tutmak için ne kadarlık bir maliyeti göze almasının efektif olacağını belirlemesine yardım ederken, tüketim davranışları bu müşteri gruplarını kazanmak ve bağlılıklarını arttırmak için doğru stratejilerin geliştirilmesini sağlar.

Müşteri gruplarının belirlenmesi için ÖÖA ve K-means kümeleme algoritmaları kullanılmıştır. Kümeleme algoritmasının başarısı her bir kümenin kendi içinde homojen diğer kümelerle kıyasla ise heterojen olmasıyla ölçülür. Bu nedenle kullanılan kümeleme algoritmalarının başarısını ölçmek için kümedeki her bir elemanın merkeze uzaklıkları toplanarak minimum, kümelerin birbirlerine olan uzaklıkları da toplanarak maksimum olması hedeflenmiştir. Bu değerlendirme neticesinde ÖÖA algoritmasının daha iyi sonuç verdiği görülmüştür. ÖÖA kümeleme algoritması sonucunda en iyiler, ilk kez gelenler ve belirsizler olarak tanımlanan müşteri grupları oluşmuştur. K-means algoritması neticesinde ise ÖÖA'da oluşan kümeler ek olarak sadıklar olarak tanımlanan müşteri grubu da oluşmuştur.

Best müşteriler yakın zamana kadar işletmeyle etkileşimini devam ettiren harcama sıklığı yüksek ve harcama meblağı yüksek müşterilerdir. Bu müşteri grubu zaten işletmenin mevcut ürünlerini yoğun bir şekilde tercih etmektedir. Bu nedenle işletmenin bu müşterileri zaten tükettikleri ürünleri daha fazla tüketmeye teşvik etmek yerine ürün çeşitlendirmeye gitmesi daha yerinde olacaktır. Yani bu müşteri grubuna yönelik olarak upselling ve cross selling çalışmaları yürütmek yerine olacaktır.

First time müşteriler yakın zamanda işletmeyle etkileşimde bulunmuş ancak alışveriş sıklığı ve harcaması düşük müşterilerdir. Bu müşterilerin yakın zamanda işletmeyi tercih etmiş olmaları tekrar tercih edebileceklerine işaret kabul edilebilir. Söz konusu müşteri grubunun işletmeyi nadiren tercih etmesinin nedenlerini ve tercih ettikleri bu dönemlerde hangi faktörlerin etkili olduğunu belirlemek için elde edilen veriler yetersizdir. Ancak bu müşterilerin işletmeyi sürekli değil de düşük sıklıkta tercih etmesinin arka planında ya belirgin bir ürün çeşidi ya da dönemsel promosyonlar olduğu tahmin edilebilir. Promosyon amaçlı olarak işletmeyi tercih eden müşterilerin işletmeyle etkileşimini arttırmak için avcı ürün stratejisi uygulanabilir. Yani spesifik ürünlerde indirimler yaparak o ürünler için işletmeye gelen müşterilere başka yan ürünlerde satılabilir. Ya da belirgin bir ürün için işletmeyi tercih eden müşterilere benzer ürün tavsiyeleriyle cross selling veya up selling yapılabilir.

İşletmenin üçüncü müşteri sınıfı ise belirsiz müşterilerdir. Bu müşteriler yakın zamanda işletmeyi tercih etmemiş, alışveriş sıklıkları ve harcama meblağları düşük müşterilerdir. Bu müşteriler her ne kadar bireysel olarak işletme için en değersiz müşteri sınıfı olsa da işletmenin dönemlik kazancının %10'unu sağlaması bakımında önemlidir. Bu müşterilerin avcı ürün veya süreli promosyon stratejileriyle işletmeye bağlılıkları arttırılabilir.

İşletme müşterileri içerisindeki son grup ise shopper olarak adlandırılan harcama meblağları düşük olmasına rağmen yakın zamana kadar işletmeyle etkileşimi devam ettirmiş ve harcama sıklığı yüksek müşterilerdir. Bu müşterilerin işletme için kar oranları düşük olmasına rağmen sadakatleri yüksektir. Harcama miktarlarının ortalamasının altında olması ise alt gelir sınıfında olmaları ile açıklanabilir. Bu müşterilerin işletmeye direkt olarak sağladıkları faydalar beklentinin altında rağmen işletmeye gösterdikleri sadakat nedeniyle dolaylı faydalar sağlarlar.

Çalışmanın ilk aşamasında kullanılan veri setinde müşterilere ait demografik bilgilerin ve müşterinin gelir bilgisinin eksikliği müşteri davranışlarına etki eden çevresel, sosyal ve finansal faktörleri görmeyi engellemektedir. Ayrıca ürünlere ait dönemsel promosyon bilgilerinin eksikliği müşterinin mutlak bir sadakat mi yoksa sahte bir sadakat mi gösterdiğini belirmememizi imkânsız kılmaktadır. İleriki çalışmalarda RFM analizi sonucunda belirlenen müşteri davranış modelleri birer etiket olarak alınıp, müşterilerin demografik özelliklerine ve finansal durumlarına dair kurallar çıkarılması planlanmaktadır.

Çalışmada ikinci olarak bir telekomünikasyon şirketi müşterilerine ait veri seti kullanılarak müşterilerin ayrılma analizi naive bayes, yapay sinir ağları, lineer regresyon ve destek vektör makineleri kullanılarak yapılmış, buna ek olarak aynı veri setinin öznelikleri sınıflamaya etkisi bakımından ki-kare, bilgi kazancı, kazanım oranı ve reliefF metrikleri kullanılarak sıralanmıştır. Bu analiz sonuçları incelendiğinde tüm performans kriterlerinde en iyi sınıflama algoritmasının yapay sinir ağları olduğu, en kötü sınıflama algoritmasının da destek vektör makineleri olduğu görülmüştür. OVO ve OVA teknikleri için ayrı ayrı gamma parametresi optimizasyon işlemlerinin yapılmasına rağmen destek vektör makinesi modellerinin

veri setini ezberlemesinin önüne geçilememiştir. Bu nedenle bu veri setinin destek vektör makinesi için uygun olmadığı görülmektedir. Öznitelik sıralama için kullanılan metrik analizleri incelendiğinde tüm yöntemler için gündüz kullanım bilgisini içeren öznitelikler ile müşteri servisini arama sayısı diğer özniteliklere göre müşterileri ayrılma durumlarına göre sınıflamada daha belirleyici olmaktadır. Yine aynı analiz sonuçları incelendiğinde ki-kare, bilgi kazancı ve kazanım oranı algoritmalarının bire bir aynı reliefF algoritmasının ise bu algoritmalara benzer bir sonuç aldığı görülmüştür. Sonuçların birbirine bu kadar yakın olması, bu veri setinin öznitelik seçimi için uygun olmadığı yorumu yapılabilmektedir. Tüm bu sonuçlar düşünüldüğünde daha sonraki çalışmalarda örnek ve öznitelik sayısı daha fazla olan bir veri seti üzerinde aynı yöntemlerin denenmesi ve bu yöntemlerin analiz sonuçlarının derin öğrenme gibi daha güncel algoritma sonuçları ile karşılaştırılması düşünülmektedir.

KAYNAKÇA

- “80/20 Kuralı Nasıl Şirket Kurtardı”. (2013). Erişim Tarihi: 20 Haziran 2017.
<http://www.capital.com.tr/gelecek-trendler/8020-kurali--nasil-sirket-kurtardi-haberdetay-1254>
- Abu Ellaban, Mahmoud Ayesh (2013). *The Role of Data Mining Technology in Building Marketing and Customer Relationship Management (CRM) for Telecommunication Industry* (Doktora Tezi). İslamic University, Gaza.
- Agiu Dragoş, Mateescu Vlad ve Muntean Iulia (2014). "Business Intelligence Overview". *Database Systems Journal*. 5: 23–35.
- Akça, Mustafa. (2012). "OLTP ve OLAP Nedir?" Erişim Tarihi: 23 Ağustos 2017.
<http://mustafaakca.com/oltp-ve-olap-nedir/>
- Akpınar, Haldun (2000). "Veri Madenlerinde Bilgi Keşfi ve Veri Madenciliği". *İstanbul Üniversitesi İşletme Fakültesi Dergisi*. 29: 1–22.
- Aktepe Cemalettin, Baş Mehmet ve Tolon Metehan (2009). *Müşteri İlişkileri Yönetimi* Ankara: Detay Yayıncılık.
- "AMA - American Marketing Association" (2013). Erişim Tarihi: 16 Mayıs 2017.
<https://www.ama.org/aboutama/pages/definition-of-marketing.aspx>
- Ateş Hilal (2008). *Karar Vermede İş Zekasının Önemi: Tekstil Sektöründe Bir Araştırma* (Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi, İzmir.
- Bagheri, F. ve Tarokh, Mohammad Jafar (2014). "Customer behavior mining based on RFM model to improve the customer relationship management". *Journal of Industrial Engineering and Management Studies*. 1: 43–57.
- Başok Yurdakul, Nilay (2003). "İşletme Yönetiminde İki Stratejik Görev: İmaj-Marka Yönetimi ve Müşteri İlişkileri Yönetimi". *Manas Üniversitesi Sosyal Bilimler Dergisi*. 8: 205–211.
- Bayraktar, Meltem (1994). "Tüketici Davranışlarının Psikolojik Boyutları". *Tüketici Bülteni*. 80.

- Birant, Derya (2011). "Data Mining Using RFM Analysis. In Knowledge-Oriented Applications in Data Mining". InTech. Erişim Tarihi: 2 Mart 2017
<http://www.intechopen.com/books/knowledge-oriented-applications-in-data-mining>
- Bişkin, Ferdi (2011). *Sağlık İşletmelerinde Müşteri İlişkileri Yönetimi: Kamu ve Özel Sağlık İşletmelerinde Müşteri İlişkileri Yönetimi Uygulamaları Araştırması* (Yüksek Lisans Tezi). Selçuk, Konya.
- Bose, Ranjit (2002). "Customer relationship management: key components for IT success". *Industrial Management & Data Systems*. 102: 89–97.
- Boutsinas Basilis ve Gnardellis T. (2002). "On distributing the clustering process". *Pattern Recognition Letters*. 23: 999–1008.
- Budhwani, Karim (2002). "Once You've Found Them, Never Let Them Go". *CMA Management*. 2.
- Çatı Kahraman ve Koçoğlu Cenk Murat (2010). *Müşteri İlişkileri Yönetimi. Güncel Pazarlama Yaklaşımlarından Seçmeler*. Ankara: Detay Yayıncılık.
- Chang Yu-Heng ve Chen Fang-Yuan (2007). "Relational benefits, switching barriers and loyalty: A study of airline customers in Taiwan". *Journal of Air Transport Management*. 13: 104–109.
- Changchien S. Wesley, Lee Chin-Feng ve Hsu Yu-Jung (2004). "On-line personalized sales promotion in electronic commerce". *Expert Systems with Applications*. 27: 35–52.
- Cheung Kwok-Wai, Kwok James T., Law Martin H. ve Tsui Kwok-Ching (2003). "Mining customer product ratings for personalized marketing". *Decision Support Systems*. 35: 231–243.
- Churn in Telecom's dataset. Erişim Tarihi: 28 Kasım 2017.
<https://bigml.com/user/francisco/gallery/dataset/5163ad540c0b5e5b22000383>
- Çiçek, Ercan (2005). *Süpermarket ve Hipermarketlerde Müşteri İlişkileri Yönetimi ve Bir Uygulama* (Doktora Tezi). Selçuk, Konya.

- Coussement Kristof ve De Bock Koen W. (2013). "Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning". *Journal of Business Research*. 66: 1629–1636.
- Coussement, Kristof ve Van den Poel Dirk (2008). "Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques". *Expert Systems with Applications*. 34: 313–327.
- Dick Alan S. ve Basu Kunal (1994). "Customer loyalty: Toward an integrated conceptual framework". *Journal of the Academy of Marketing Science*. 22: 99–113.
- Duran, Mustafa. CRM Hakkında.... Erişim Tarihi: 22 Haziran 2017 <http://danismend.com/kategori/altkategori/crm-hakkinda/>
- Durmaz, Yakup (2006). "Modern Pazarlamada Tüketici Memnuniyeti ve Evrensel Tüketici Hakları". *Journal of Yaşar University*. 1: 255 – 266.
- Erdoğan, Zafer (2014). *Pazarlamanın Konusu, Kapsamı, Gelişimi ve Diğer Sosyal Bilimlerle İlişkisi. Pazarlama İlkeler ve Yönetim*. Bursa: Ekin Basım Yayın
- Farquad Mohammad Abdul Haque, Ravi Vadlamani ve Raju S. Bapi (2014). "Churn prediction using comprehensible support vector machine: An analytical CRM application". *Applied Soft Computing*. 19: 31–40.
- Ghazanfari Mehdi, Mohamadi Samira Malek ve Alizadeh Somayeh (2008). "Data-mining application for country segmentation based on the RFM model". *Int. J. Data Analysis Techniques and Strategies*. 1: 126–140.
- Habul Aida ve Pilav-Velić Amila (2010). "Business Intelligence and Customer Relationship Management". Proceedings of the ITI 2010 32nd Int. Conf. on Information Technology Interfaces, Croatia.
- Habul Aida, Pilav-Velić Amila ve Kremić Emir (2012). "Customer Relationship Management and Business Intelligence". Erişim Tarihi: 4 Temmuz 2017. <http://www.intechopen.com/books/advances-in-customer-relationship-management/customer-relationship-management-and-business-intelligence>

- Hadiji Fabian, Sifa Rafet, Drachen Anders, Thureau Christian, Kersting Kristian ve Bauckhage Christian (2014). "Predicting player churn in the wild". *2014 IEEE Conference on Computational Intelligence and Games*. 1–8.
- Han Jiawei, Kamber Micheline ve Pei Jian (2012). *Data Mining: Concepts and Techniques*. Massachusetts: Morgan Kaufmann.
- Hasan Afan (2011). *Havayolu Şirketlerinde Müşteri İlişkileri Yönetimini (CRM) Desteklemek ve Müşteri Sadakatini Değerlendirmek İçin Veri Madenciliğinin Kullanılması* (Yüksek Lisans Tezi). Yıldız Teknik Üniversitesi, İstanbul.
- Hung Shin-Yuan, Yen David C., ve Wang Hsiu-Yu (2006). "Applying data mining to telecom churn management". *Expert Systems with Applications*. 31: 515–524.
- Bowen John T. ve Chen Shiang- Lih (2001). "The relationship between customer loyalty and customer satisfaction". *International Journal of Contemporary Hospitality Management*. 13: 213–217.
- Kağnıcıoğlu Hakan (2002). "Günümüz İşletmelerinin Yaşam Anahtarı: Müşteri Odaklılık". *Ege Academic Review*. 2: 78–90.
- Kalikov Anarbek (2006). *Veri Madenciliği ve Bir E-Ticaret Uygulaması* (Yüksek Lisans Tezi). Gazi Üniversitesi, Ankara.
- Kart Özge (2013). *Decision Support System For a Customer Relationship Management Case Study* (Yüksek Lisans Tezi). Dokuz Eylül, İzmir.
- Kavzoğlu Taşkın, Şahin Emrehan Kutluğ ve Cölkesen İsmail (2014). "Heyelan Duyarlılık Analizinde Ki-Kare Testine Dayalı Faktör Seçimi". *Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*, İstanbul.
- Kelly Scott (2000). "Analytical CRM: The fusion of data and intelligence". *Interactive Marketing*. 1: 262–267.
- Khajvand Mahboubeh, Zolfaghar Kiyana, Ashoori Sarah ve Alizadeh Somayeh (2011). "Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study". *Procedia Computer Science*. 3: 57–63.

- Kim Kyoungok, Jun Chi-Hyuk ve Lee Jaewook (2014). "Improved churn prediction in telecommunication industry by analyzing a large network". *Expert Systems with Applications*. 41: 6575–6584.
- Kenji Kira ve Larry A. Rendell (1992). "A Practical Approach to Feature Selection". Proceedings of the Ninth International Workshop on Machine Learning.
- Kisioglu Pınar ve Topcu Y. Ilker (2011). "Applying Bayesian Belief Network approach to customer churn analysis: A case study on the telecom industry of Turkey". *Expert Systems with Applications*. 38: 7151–7157.
- Kohonen Teuvo (1982). "Self-organized formation of topologically correct feature maps". *Biological Cybernetics*. 43: 59–69.
- Kononenko Igor (1994). *Estimating attributes: Analysis and extensions of RELIEF*. Berlin: Heidelberg.
- Kotler Philip ve Keller Kevin Lane (2012). *Marketing Management*. New Jersey: Prentice Hall.
- Larivière Bart ve Van den Poel Dirk (2005). "Predicting customer retention and profitability by using random forests and regression forests techniques". *Expert Systems with Applications*. 29: 472–484.
- Lejeune Miguel A. P. M. (2001). "Measuring the impact of data mining on churn management". *Internet Research*. 11: 375–387.
- Liu Duen-Ren ve Shih Ya-Yueh (2005). "Integrating AHP and Data Mining for Product Recommendation Based on Customer Lifetime Value". *Information & Management*. 42: 387–400.
- McGoldrick, Peter J. ve Andre Elisabeth (1997). "Consumer misbehaviour". *Journal of Retailing and Consumer Services*. 4: 73–81.
- Min Hokey, Min Hyesung ve Emam Ahmed (2002). "A data mining approach to developing the profiles of hotel customers". *International Journal of Contemporary Hospitality Management*. 14: 274–285.

- Moss Larissa T. ve Atre, Shaku (2003). *Business Intelligence Roadmap*. Massachusetts: Addison Wesley.
- Mozer Michael C., Wolniewicz Richard H, Grimes David B., Johnson Eric ve Kaushansky Howard (2000). "Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry". *IEEE Transactions on Neural Networks*" 11: 690–696.
- Murat Eda (2005). *Müşteri İlişkileri Yönetimi ve Bir Enerji Firmasında Uygulaması* (Yüksek Lisans Tezi). Erciyes Üniversitesi, Kayseri.
- Nakip Mahir, Varinli İnci ve Gülmez Mustafa (2012). *Güncel Pazarlama Yönetimi*. Ankara: Detay Yayıncılık.
- Online Retail Dataset. Erişim Tarihi: 8 Nisan 2017,
- Oyman Mine (2002). "Müşteri Sadakati Sağlamada Sadakat Programlarının Önemi". *Kurgu Dergisi*. 19: 169-185
- Ozarkar Prajakta ve Patwardhan Manasi (2013). "Efficient Spam Classification by Appropriate Feature Selection". *Global Journal of Computer Science and Technology Software & Data Engineering*. 13: 49–57.
- Özen Üstün (2016). *Bilgi Sistemlerine Giriş: Temel Kavramlar*. Atatürk Üniversitesi Açık Öğretim Fakültesi.
- Özgüner Hicran (2014). Pazarlamada Güncel Uygulamalar. *Pazarlama İlkeler ve Yönetim*. Bursa: Ekin Basın Yayın Dağıtım.
- Özmen, Şule (2012). *Ağ Ekonomisinde Yeni Ticaret Yolu: E-Ticaret*. İstanbul: İstanbul Üniversitesi Yayınları.
- Parvatiyar Atul ve Sheth Jagdish N. (2001). "Customer Relationship Management: Emerging Practice". *Process, and Discipline*. 3: 1–34.
- Peppers Don ve Rogers Martha (2000). *The One to One Manager: Real-World Lessons in Customer Relationship*. Minnesota: Capstone Publishing Ltd.
- Persson Petra (2003). *Customer Relation Management How a CRM System Can Be Used in the Sales Process*. Lulea Universty of Technology, Stockholm.

- Rabiei Atefeh ve Rastegari Hamid (2015). "Integrating RFM and Classification for Response Modeling Based on Customer Lifetime Value". *Cumhuriyet Üniversitesi Fen Fakültesi Fen Bilimleri Dergisi*. 36: 247-253
- Reichheld Frederick F. ve Teal Thomas (1996). *The Loyalty Effect : The Hidden Force Behind Growth, Profits, and Lasting Value*. Harvard Business School Press.
- Reinschmidt Joerg ve Francoise Allison (2000). *Business Intelligence Certification Guide*. Erişim Tarihi:23 Eylül 2017 www.redbooks.ibm.com
- Savaş Serkan, Topaloğlu Nurettin ve Yılmaz Mithat (2012). "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri". *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*. 11: 1–23.
- Sohrabi Babak ve Khanlari Amir (2007). "Customer Lifetime Value (CLV) Measurement Based on RFM Model". *Iranian Accounting & Auditing Review*. 14: 7–20.
- Tarimer Güray (2011). "İş zekası uygulamaları ve işletmeler üzerine pratik etkileri". Erişim Tarihi: 31 Temmuz 2017. <https://www.dunya.com/gundem/is-zekasi-uygulamalari-ve-isletmeler-uzerine-pratik-etkileri-haberi-140317>
- Tepeci Mustafa (1999). "Increasing brand loyalty in the hospitality industry". *International Journal of Contemporary Hospitality Management*. 11: 223–230.
- Tezcanlar Pelin (2007). *Müşteri İlişkileri Yönetimi, Veri Madenciliği ve Bir Uygulama*. İstanbul Üniversitesi.
- Ulucan Nazif Can (2016). *Online Satış Platformlarında Müşteri İlişkileri Yönetiminin Performansa Etkisi Üzerinde Sosyal Medya Kullanımının Rolü: Sosyal Müşteri İlişkileri Yönetimi (Sosyal CRM)*. Haliç Üniversitesi.
- Uzun Ali Çağlar (2004). *Müşteri İlişkileri Yönetiminde Kritik Başarı Faktörleri (Yüksek Lisans Tezi)*. Muğla Üniversitesi.

- Van den Poel Dirk ve Larivière Bart (2004). "Customer attrition analysis for financial services using proportional hazard models". *European Journal of Operational Research*. 157: 196–217.
- Vapnik Vladimir (2013). "The Nature of Statistical Learning Theory". Springer Science & Business Media. Erişim Tarihi:24.06.2017. [//www.springer.com/br/book/9780387987804](http://www.springer.com/br/book/9780387987804)
- Vercellis, Carlo (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. İngiltere: Wiley.
- Verhoef Peter C. ve Donkers Bas (2001). "Predicting customer potential value an application in the insurance industry". *Decision Support Systems*. 32: 189–199.
- Which RFM binning method is best for me? Erişim Tarihi: 8 Nisan 2017 <https://communities.sas.com/t5/SAS-Communities-Library/Which-RFM-binning-method-is-best-for-me/ta-p/341327>
- Woo Ji Yong, Bae Sung Min ve Park Sang Chan (2005). "Visualization method for customer targeting using customer map". *Expert Systems with Applications*. 28: 763–772.
- Yavuz Selahattin ve Deveci Muhammet (2012). "İstatiksel Normalizasyon Tekniklerinin Yapay Sinir Ağın Performansına Etkisi". *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*. 40: 167–187.
- Yeler Mücahit (2006). *Müşteri İlişkileri Yönetimi (Crm) ve Türkiye’de Bankacılık Sektöründe Crm Uygulamalarında Eğitimin Rolü* (Yüksek Lisans Tezi). Gazi Üniversitesi.
- Yıldırım Fazlı ve Panayırıcı Uğur Cevdet (2016). *Müşteri İlişkileri Yönetimi ve Pazarlama İletişimi*. İstanbul: Papatya Yayıncılık.
- Yılmaz Malik (2009). "Enformasyon Ve Bilgi Kavramları Bağlamında Enformasyon Yönetimi ve Bilgi Yönetimi". 49: 95–118.

Xu Yurong, Yen David C., Lin Binshan ve Chou David C. (2002). "Adopting customer relationship management technology". *Industrial Management & Data Systems*. 102: 442–452.

Zengin Burhanettin ve Ulama Şevki (2015). *Müşteri İlişkileri Yönetimi (MIY). Turizm Pazarlamasında Güncel Yaklaşımlar*. İstanbul: Beta Basım Dağıtım.

ÖZ GEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı: Mehmet Ali Deveci

Uyruğu: TC

Doğum Tarihi ve Yeri: Nizip/Gaziantep

e-posta: madeveci@gmail.com

EĞİTİM

Derece	Kurum	Mezuniyet Yılı
Lisans	Cumhuriyet Üniversitesi	2012
Yüksek Lisans	Cumhuriyet Üniversitesi	2018

İŞ TECRÜBESİ

Tarih	Kurum	Görev
2013	Cumhuriyet Üniversitesi	Araştırma Görevlisi

YABANCI DİL BİLGİSİ

Yabancı Dilin Adı YDS (81,25) ÜDS (...) TOEFL (...) EILTS (...)