



AKDENİZ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



Aslıhan DURSUN CENGİZCİ

OTEL İŞLETMELERİNDE KAYIP MÜŞTERİ TAHMİNLEMESİ

Turizm İşletmeciliği Ana Bilim Dalı
Doktora Tezi

Antalya, 2020



AKDENİZ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



Aslıhan DURSUN CENGİZCİ

OTEL İŞLETMELERİNDE KAYIP MÜŞTERİ TAHMİNLEMESİ

Danışman

Doç. Dr. Meltem CABER

Turizm İşletmeciliği Ana Bilim Dalı

Doktora Tezi

Antalya, 2020

T.C.
Akdeniz Üniversitesi
Sosyal Bilimler Enstitüsü Müdürlüğüne,

Aslıhan DURSUN CENGİZCİ'nin bu çalışması, jürimiz tarafından Turizm İşletmeciliği Ana Bilim Dalı Doktora Programı tezi olarak kabul edilmiştir.

Başkan	: Prof. Dr. İbrahim Sani MERT	(İmza)
Üye (Danışmanı)	: Doç. Dr. Meltem CABER	(İmza)
Üye	: Doç. Dr. Tahir ALBAYRAK	(İmza)
Üye	: Doç. Dr. Olgun KİTAPCI	(İmza)
Üye	: Dr. Gözdegül BAŞER	(İmza)

Tez Başlığı: Otel İşletmelerinde Kayıp Müşteri Tahminlemesi

Onay: Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

Tez Savunma Tarihi :16/01/2020

Mezuniyet Tarihi : 06/02/2020

(İmza)
Prof. Dr. İhsan BULUT
Müdür

AKADEMİK BEYAN

Doktora Tezi olarak sunduğum “Otel İşletmelerinde Kayıp Müşteri Tahminlemesi” adlı bu çalışmanın, akademik kural ve etik değerlere uygun bir biçimde tarafımdan yazıldığını, yararlandığım bütün eserlerin kaynakçada gösterildiğini ve çalışma içerisinde bu eserlere atıf yapıldığını belirtir; bunu şerefimle doğrularım.

İmza

Aslıhan DURSUN CENGİZCİ





T.C.
AKDENİZ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
TEZ ÇALIŞMASI ORJİNALLİK RAPORU
BEYAN BELGESİ



SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

ÖĞRENCİ BİLGİLERİ	
Adı-Soyadı	Ashhan DURSUN CENGİZCİ
Öğrenci Numarası	20155204001
Enstitü Ana Bilim Dalı	Sosyal Bilimler Enstitüsü Turizm İşletmeciliği ABD
Programı	Doktora
Programın Türü	() Tezli Yüksek Lisans (x) Doktora
Danışmanın Unvanı, Adı-Soyadı	Doç. Dr. Meltem CABER
Tez Başlığı	Otel İşletmelerinde Kayıp Müşteri Tahminlemesi
Turnitin Ödev Numarası	1244834549

Yukarıda başlığı belirtilen tez çalışmasının a) Kapak sayfası, b) Giriş, c) Ana Bölümler ve d) Sonuç kısımlarından oluşan toplam 162 sayfalık kısmına ilişkin olarak, 22/01/2020 tarihinde tarafımdan Turnitin adlı intihal tespit programından Sosyal Bilimler Enstitüsü Tez Çalışması Orjinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nda belirlenen filtrelemeler uygulanarak alınmış olan ve ekte sunulan rapora göre, tezin/dönem projesinin benzerlik oranı;

alıntılar hariç % 3

alıntılar dahil % 14'tür.

Danışman tarafından uygun olan seçenek işaretlenmelidir:

(x) Benzerlik oranları belirlenen limitleri aşmıyor ise;

Yukarıda yer alan beyanın ve ekte sunulan Tez Çalışması Orjinallik Raporu'nun doğruluğunu onaylarım.

() Benzerlik oranları belirlenen limitleri aşıyor, ancak tez/dönem projesi danışmanı intihal yapılmadığı kanısında ise;

Yukarıda yer alan beyanın ve ekte sunulan Tez Çalışması Orjinallik Raporu'nun doğruluğunu onaylar ve Uygulama Esasları'nda öngörülen yüzdelik sınırlarının aşılmasına karşın, aşağıda belirtilen gerekçe ile intihal yapılmadığı kanısında olduğumu beyan ederim.

Gerekçe:

Benzerlik taraması yukarıda verilen ölçütlerin ışığı altında tarafımda yapılmıştır. İlgili tezin orijinallik raporunun uygun olduğunu beyan ederim.

22/01/2020

(imzası)
Doç. Dr. Meltem CABER

İÇİNDEKİLER

ŞEKİLLER LİSTESİ	iii
GRAFİKLER LİSTESİ	iv
TABLOLAR LİSTESİ	v
KISALTMALAR LİSTESİ	vi
ÖZET	vii
SUMMARY	viii
TEŞEKKÜR.....	ix
GİRİŞ.....	2

BİRİNCİ BÖLÜM

OTEL İŞLETMELERİNDE MÜŞTERİ İLİŞKİLERİ YÖNETİMİ VE UYGULAMALARI

1.1. Müşteri İlişkileri Yönetimi (M.İ.Y.)	3
1.2. Veri Tabanlı Pazarlama	6
1.3. Elektronik Müşteri İlişkileri Yönetimi (E-M.İ.Y.).....	11
1.4. Dijital Pazarlama	14
1.5. Müşteri Sadakati ve Müşteriyi Elde Tutma.....	18
1.6. Otel İşletmelerinde Güncel Müşteri İlişkileri Yönetimi Uygulamaları.....	22
1.6.1. Müşteri Veri Tabanları	24
1.6.2. Sadakat Programları.....	24
1.6.3. Çevrimiçi İtibar Yönetimi Sistemleri	26
1.6.4. M.İ.Y. Yazılımları	28
1.6.5. Kurumsal Kaynak Planlama Sistemleri (K.K.P.)	30

İKİNCİ BÖLÜM

KAYIP MÜŞTERİ TAHMİNLEME YÖNTEMLERİ

2.1. Makine Öğrenimi	31
2.2. Turizm ve Otel İşletmeciliğinde Makine Öğrenimi	35
2.3. Kayıp Müşteri Kavramı.....	38
2.4. Kayıp Müşteri Tahminleme.....	40
2.5. Kayıp Müşteri Tahminleme Yöntemleri	42
2.6. Farklı Sektörlerde Kayıp Müşteriyi Tanımlamak	46
2.7. Kayıp Müşteri Tahminlemesinde Kullanılan Değişkenler	48
2.8. Kayıp Müşteri Tahminleme Çalışmalarına Yönelik Literatür	55

ÜÇÜNCÜ BÖLÜM

OTEL İŞLETMELERİNDE KAYIP MÜŞTERİ TAHMİNLEME UYGULAMASI

3.1.	Araştırmanın Amacı ve Kapsamı	61
3.2.	Araştırma Modeli	61
3.3.	Araştırmanın Yöntemi ve Bulguları	62
3.3.1.	Veri Setinin Temizlenmesi ve Veri Ön İşleme.....	62
3.3.2.	Aykırı Gözlemlerin Temizlenmesi	67
3.3.3.	Kayıp Müşterinin Tanımlanması ve Zaman Aralıkları.....	69
3.3.4.	Değişkenler	70
3.3.5.	Veri Setinin Bölünmesi : Öğrenme, Doğrulama ve Test Etme	97
3.3.6.	Değişken Seçimi	99
3.3.7.	Öğrenme Aşaması ve Çapraz Doğrulama.....	102
3.3.8.	Test Aşaması ve Model Performansı.....	107
3.3.9.	Tahminleme Aşaması ve Tahmin Sonuçları.....	109
SONUÇ		121
KAYNAKÇA.....		132
EK 1 - ORANGE ÇALIŞMA MODELİ.....		145
ÖZGEÇMİŞ.....		146

ŞEKİLLER LİSTESİ

Şekil 1.1 Veri Ambarı Örneği.....	7
Şekil 1.2 Veri Türleri ve Özellikleri.....	8
Şekil 1.3 Müşteri Yaşam Döngüsü.....	9
Şekil 1.4 Öncelik Matrisi Örneği.....	10
Şekil 1.5 Dijital Pazarlama Uygulama Teknikleri.....	15
Şekil 1.6. Otel İşletmelerinde Dijital Pazarlama Bütçesinin Kullanımı.....	17
Şekil 1.7 Yatırım / Karlılık Zıtlığı.....	19
Şekil 1.8 Bağlılık Matrisi	20
Şekil 1.9 Müşteri Sadakati Piramidi	21
Şekil 1.10 Otel İşletmelerinde Çevrimiçi İtibar Yönetimi	27
Şekil 1.11 M.İ.Y. Yazılımı Başarı Faktörleri	29
Şekil 2.1 Disiplinler Arası İlişkiler.....	31
Şekil 2.2 Posta Zarfları Üzerindeki El Yazısı Örnekleri	33
Şekil 2.3 SentiVoice Sistemi Çalışma Mantığı	36
Şekil 2.4 Destinasyon Yönetim Sistemi Çalışma Mantığı	37
Şekil 3.1 Araştırma Modeli	62
Şekil 3.2 Veri Setinden Temsili Bir Kesit	63
Şekil 3.3 Kur Kategorilerinin Düzenlenmesinde Kullanılan Kodlar.....	64
Şekil 3.4 Kur Dönüştürmede Kullanılan Kodlar	64
Şekil 3.5 Seyahat Eşlikçilerini Belirlemede Kullanılan Kodlar	65
Şekil 3.6 Otel Kategorisi Değişkenlerinin Ayırıştırılmasında Kullanılan Kodlar.....	65
Şekil 3.7 Münferit Satın Alımların Belirlenmesinde Kullanılan Kodlar.....	66
Şekil 3.8 Lokal Aykırılık Faktörü Hesaplamasında Kullanılan Kodlar	67
Şekil 3.9 Aykırı Gözlemler Saçılma Diyagramı.....	68
Şekil 3.10 Lokal Aykırılık Faktör Skorları Saçılma Diyagramı.....	68
Şekil 3.11 Lokal Aykırılık Faktör Skorları Saçılma Diyagramı II	69
Şekil 3.12 Zaman Aralıkları	69
Şekil 3.13 Öğrenme, Doğrulama ve Test Veri Setlerinin Oranı.....	97
Şekil 3.14 k -Katlamalı Çapraz Doğrulama	98
Şekil 3.15 Karışıklık Matrisi Örneği	103
Şekil 3.16 ROC Eğrisi Örneği	104
Şekil 3.17 Öncelik Karar Matrisi.....	120

GRAFİKLER LİSTESİ

Grafik 3.1 Satın Alımlar Arasında Geçen Zaman Değişkeni Dağılım Grafiği	71
Grafik 3.2 İlişki Süresi Değişkeni Dağılım Grafiği.....	72
Grafik 3.3 Son Üç Yılda Harcanan Tutar Değişkeni Dağılım Grafiği	73
Grafik 3.4 Gerçek Tutar Değişkeni Dağılım Grafiği.....	73
Grafik 3.5 Son Üç Yıldaki Gecelik Harcama Tutarı Değişkeni Dağılım Grafiği	74
Grafik 3.6 Gerçek Gecelik Harcama Tutarı Değişkeni Dağılım Grafiği.....	75
Grafik 3.7 Son Üç Yıldaki Konaklama Başına Harcama Tutarı Dağılım Grafiği.....	76
Grafik 3.8 Gerçek Konaklama Başına Harcama Tutarı Dağılım Grafiği	76
Grafik 3.9 Son Üç Yıldaki Geliş Sayısı Dağılım Grafiği	77
Grafik 3.10 Gerçek Geliş Sayısı Dağılım Grafiği	78
Grafik 3.11 Güncellik Değişkeni Dağılım Grafiği	78
Grafik 3.12 Konaklama Eşlikçileri Değişkenleri Dağılım Grafiği.....	79
Grafik 3.13 Konaklama Yapılan Otel Değişkenleri Dağılım Grafiği.....	80
Grafik 3.14 Grubun Diğer Otellerinde Konaklama Değişkeni Dağılım Grafiği	81
Grafik 3.15 Sadakat Kartı Değişkeni Dağılım Grafiği	82
Grafik 3.16 Satın Alım Yöntemi Değişkeni Dağılım Grafiği	83
Grafik 3.17 Son Üç Yıldaki Geceleme Sayısı Dağılım Grafiği	84
Grafik 3.18 Gerçek Geceleme Sayısı Dağılım Grafiği.....	84
Grafik 3.19 Son Üç Yıldaki Konaklama Başına Geceleme Sayısı Dağılım Grafiği	85
Grafik 3.20 Gerçek Konaklama Başına Geceleme Sayısı Dağılım Grafiği	86
Grafik 3.21 Gecelik Ödeme Tutarı Değişimi Dağılım Grafiği.....	87
Grafik 3.22 Ortalama Harcama Tutarı Değişimi Dağılım Grafiği	87
Grafik 3.23 Ortalama Geceleme Sayısı Değişimi Dağılım Grafiği.....	88
Grafik 3.24 Sezon Değişkeni Dağılım Grafiği	89
Grafik 3.25 Ayrılma Davranışı Bağımlı Değişkeni Dağılım Grafiği	90
Grafik 3.26 Yaş Değişkeni Dağılım Grafiği.....	90
Grafik 3.27 Milliyet Değişkenleri Dağılım Grafiği.....	91
Grafik 3.28 Pazar Değişkeni Dağılım Grafiği	92
Grafik 3.29 ROC Eğrisi Rastgele Orman (Çapraz Doğrulama)	106
Grafik 3.30 ROC Eğrisi Lojistik Regresyon (Çapraz Doğrulama)	106
Grafik 3.31 ROC Eğrisi Rastgele Orman (Test)	108
Grafik 3.32 ROC Eğrisi Lojistik Regresyon (Test).....	108

TABLOLAR LİSTESİ

Tablo 2.1 Kayıp Müşteri Tahminleme Yöntemleri ve Çalışma Alanları	44
Tablo 2.2 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Davranış Değişkenleri	49
Tablo 2.3 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Diğer Değişkenler	51
Tablo 2.4 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Algı Değişkenleri	53
Tablo 3.1 Milliyet Kategorileri	66
Tablo 3.2 Sezon Değişkeni İçin Kullanılan Kategoriler	88
Tablo 3.3 Mevcut Çalışmada Kullanılan Davranışsal Değişkenler	93
Tablo 3.4 Mevcut Çalışmada Kullanılan Demografik Değişkenler	96
Tablo 3.5 Öğrenme ve Test Etme Veri Setleri	98
Tablo 3.6 Rastgele Orman Değişken Önem Düzeyleri	100
Tablo 3.7 Lojistik Regresyon Değişken Önem Düzeyleri	101
Tablo 3.8 Çapraz Doğrulama Öğrenme Modeli Performansları	105
Tablo 3.9 Rastgele Orman Çapraz Doğrulama Karışıklık Matrisi	107
Tablo 3.10 Lojistik Regresyon Çapraz Doğrulama Karışıklık Matrisi	107
Tablo 3.11 Model Performansları	107
Tablo 3.12 Rastgele Orman Karışıklık Matrisi	109
Tablo 3.13 Lojistik Regresyon Karışıklık Matrisi	109
Tablo 3.14 Rastgele Orman Tahminleri: Nümerik Değişkenler	111
Tablo 3.15 Rastgele Orman Tahminleri: Kategorik Değişkenler	113
Tablo 3.16 Öncelik Grupları Betimsel İstatistikleri: Nümerik Değişkenler	116
Tablo 3.17 Öncelik Grupları Betimsel İstatistikleri: Kategorik Değişkenler	118

KISALTMALAR LİSTESİ

K.K.P.	Kurumsal Kaynak Planlama
E-M.İ.Y.	Elektronik Müşteri İlişkileri Yönetimi
M.İ.Y.	Müşteri İlişkileri Yönetimi
RFM	Güncellik, sıklık, tutar
vb.	Ve bunun gibi



ÖZET

Otel işletmelerinin rakipleri arasından sıyrılarak faaliyetlerini karlı bir şekilde sürdürebilmesi, müşterileriyle uzun süreli ve kalıcı ilişkiler kurarak müşterilerini elde tutabilmesine bağlıdır. Bu bağlamda hangi müşterilerin otel işletmesiyle ilişkilerini sürdürüp hangilerinin terk edeceğine yönelik doğru bir tahminleme, başka bir deyişle kayıp müşteri tahminlemesi (customer churn prediction), daha hızlı ve etkili elde tutma faaliyetlerinin uygulanmasında yol gösterici olacaktır. Gerçekleştirilen bu çalışmanın öncelikli amacı, bir otel işletmesinin müşteri veri tabanına makine öğrenimi metodları uygulayarak işletmenin gelecek dönemdeki kayıp müşterilerini tahminlemek ve bu müşterileri elde tutmaya yönelik öneriler geliştirmektir. Bu bağlamda Antalya, Belek bölgesinde faaliyet gösteren ve bünyesinde beş yıldızlı üç resort otel bulunan bir zincir otel işletmesinin devamlı müşterilerine ait veriler kullanılarak kayıp müşteri tahminlemesi uygulanmış ve Lojistik Regresyon ile Rastgele Orman algoritmaları kıyaslanmıştır. Çalışmanın bulgularına göre Rastgele Orman algortması daha yüksek performans göstermekte ve gelecek üç yıl içerisinde ayrılma ihtimali olan devamlı müşterileri %80 oranında (AUC 0,80) doğru tahminleyebilmektedir. Ayrıca çalışmada Rastgele Orman algoritması tarafından RFM (güncellik, sıklık, tutar) temelli geleneksel değişkenler daha önemli bulunurken (önem sırasıyla güncellik, satın alımlar arasında geçen zamanların ortalaması, son üç yıldaki toplam geceleme sayısı, ilişki süresi, son üç yıldaki ortalama harcama tutarındaki değişim ve toplam geliş sayısı), Lojistik Regresyon metodu tarafından önerilen en önemli değişkenler daha çok otel işletmelerine yönelik değişkenler olarak ortaya çıkmıştır (önem sırasıyla zincir otel işletmesinin diğer otellerinde de konaklamış olmak, toplam geliş sayısı, çocuklu konaklamak, çift olarak konaklamak ve Türkiye vatandaşı olmak).

Son olarak elde edilen Rastgele Orman modeli aracılığıyla müşterilerin bir sonraki dönemdeki ayrılma olasılıkları hesaplanmıştır. Müşterilerin ayrılma olasılıkları ve potansiyel kazanç getirileri göz önünde bulundurularak bir öncelik matrisi oluşturulmuş; “kararsızlar”, “tutumlular”, “alternatif arayanlar” ve “fırsatçılar” olarak adlandırılan müşteri bölümleri için önerilere yer verilmiştir.

Anahtar Kelimeler: Kayıp müşteri tahminlemesi, makine öğrenimi, müşteri ilişkileri yönetimi, rastgele orman, lojistik regresyon

SUMMARY

CUSTOMER CHURN PREDICTION IN HOTEL FIRMS

Hotel firms should be able to maintain long-term and lasting relationships with their customers to continue their activities profitably by standing out from their competitors. Thus, an accurate estimation of which customers will maintain their relations with the hotel firm and which ones will leave, in other words predicting the customer churn, will help hotel firms to carry out faster and more effective retention activities. The main purpose of this study is to estimate the future customer churn for hotel customers by applying machine learning methods and to develop recommendations for retaining prospective churners. In this context, customer churn prediction has been applied by comparing Logistic Regression and Random Forest algorithms using data belonging to the repeat customers of a chain hotel firm with three five-star resort hotels operating in Belek region, Antalya. According to the results of the study, the Random Forest algorithm has shown the best performance predicting the prospective churners that are likely to leave in the next three years 80% correctly (AUC 0.80). Besides, the Random Forest algorithm found the traditional variables based on RFM (recency, frequency, monetary) to be more important (respectively: recency, inter-purchase time, total number of overnight stays in the last three years, length of relationship, change in average expenditure amount in the last three years, total frequency). On the other hand, the most important variables suggested by the Logistic Regression method were mostly hotel industry-oriented (respectively: accommodation at the other hotels of the hotel chain, total frequency, accommodation with children, accommodation as a couple, being a Turkish citizen).

Finally, the probability of churn for the next period was calculated through the Random Forest model. A priority matrix was created by taking into consideration the churn probability and customer value. Suggestions for customer segments called “hesitant”, “economic”, “alternative seeker” and “opportunistic” were shared.

Keywords: Customer churn prediction, machine learning, customer relationship management, random forest, logistic regression

TEŞEKKÜR

Öncelikle, akademisyenliğe ilk adımlarımı sayesinde atmış olduğum değerli danışman hocam Prof. Dr. Meltem CABER'e sonsuz sabrı, desteği ve katkıları için en içten teşekkürlerimi sunarım.

Tez izleme süresince değerli görüşleri ve yönlendirmeleriyle bu tezin oluşmasında çok büyük katkıları olan değerli hocalarım Prof. Dr. Tahir ALBAYRAK ve Doç. Dr. Olgun KİTAPCI'ya ayrıca teşekkür ederim.

Değerli hocam Prof. Dr. F. Bike KOCAOĞLU'na ve çalışma arkadaşlarım Dr. Caner ÜNAL, Dr. Sezer KARASAKAL ve Dr. Oğuz DOĞAN'a sabırları ve destekleri için teşekkür ederim. Her seferinde beni motive ederek devam etmemde yardımcı oldunuz.

Bu çalışmanın gerçekleşmesinde büyük katkı sağlayan değerli Hakan AKSÖZ ve Nida KİRAZ'a en içten teşekkürlerimi sunarım. Ayrıca verilerini analiz ettiğim otel işletmesinde çalışan bütün satış pazarlama, bilgi işlem, misafir ilişkileri ve önbüro ekiplerine misafir profillerini böylesine özenli ve dikkatli bir şekilde yönetmiş olmaları nedeniyle teşekkür ederim. Sizlerin özverisi ve emeği olmadan böyle bir çalışma gerçekleştirilemezdi.

Doktora tez çalışmalarım süresince her bunaldığım da yanımda olan sevgili kardeşim Ceren DURSUN'a ve bu süreçte bana en büyük desteği sağlayan sevgili annem Sema DURSUN ve babam İsmail DURSUN'a emekleri için sonsuz teşekkürlerimi sunarım. Okulda gece geç saatlere kadar çalıştıktan sonra eve döndüğümüzde tüm ütülerimizin yapılmış, bulaşıklarımızın toplanmış olduğunu gördüğümde ocağa bıraktığın yemeği yerken senin gibi bir annem olduğu için her defasında teşekkür ettim. Bu tezde en büyük emek senindir canım annem.

Son olarak bu yolculukta beni en iyi anlayan sevgili meslektaşım, en iyi arkadaşım ve eşim Süleyman CENGİZCİ'ye hep yanımda olduğu ve yapabileceklerime olan inancı için teşekkür ederim. Çalışma odasında sırt sırta gün ağırana kadar çalışıp uyumadan denize gittiğimiz günlere selam olsun.

GİRİŞ

Günümüzde çeşitli seyahat portalları aracılığıyla otel müşterilerinin kolaylıkla ulaşabilecekleri birçok alternatifin bulunduğu yoğun rekabet ortamında, otel işletmeleri varlıklarını sürdürmeye ve kar paylarını arttırmaya çalışmaktadır. Böyle bir ortamda yeni müşteri çekmenin yaratacağı pazarlama maliyetleri göz önünde bulundurulduğunda bir defa otel müşterisi olmuş ve sunulan hizmetleri deneyimlemiş müşterileri elde tutmak yeni müşteri edinmeye kıyasla daha karlı olacaktır. Mevcut müşterileri elde tutmanın yenisini elde etmekten daha maliyetli olduğunun bilincinde olan işletmeler pazarlama çalışmalarının büyük bir kısmını müşterilerini elde tutmaya ayırmaktadır. Bu bağlamda otel işletmelerinin müşterilerini daha yakından tanımaları, istek ve ihtiyaçlarını anlamaları önem arz etmektedir. Çoğu işletme yıllarca en iyi müşterileri hakkında en ufak bir bilgiye sahip olmadan üretim yapmakta ve hizmet sunmaktadır. Oysaki her işletme için karlılığın ve başarılı pazarlama çalışmalarının birinci ve en önemli adımı müşterilerini tanımaktır (Rud, 2001: 184). Müşterilerin kişisel ihtiyaçlarını anlamak yani müşterilerini en iyi şekilde “tanımak” için işletmeler veri tabanlarından ve veri madenciliği tekniklerinden faydalanmaktadırlar. Böylece daha fazla ürün/hizmet satışı sağlanabilmekte, uzun dönemli müşteri sadakati yaratılabilmekte ve müşteri elde tutma oranları arttırılabilmektedir (Chen ve Popovich, 2003: 686).

Bunun yanı sıra otel işletmelerinin hangi müşterilerinin ayrılma davranışında bulunacağını önceden tahmin edebilmesi daha hızlı ve etkili elde tutma faaliyetlerinde bulunmada yol gösterici olacaktır. Dolayısıyla, müşteri elde tutma oranlarını arttırmanın en etkili yolu müşteri ayrılmadan önce harekete geçebilmektir. Kayıp müşteri tahminleme modelleri (customer churn prediction models), işletmelerin bu amaçlarını gerçekleştirmelerinde yardımcı olabilmektedir. Kayıp müşteri modelleri müşterilerin ayrılma olasılığını tahminlemek için kullanılan makine öğrenimi algoritmalarıdır. Bir işletmenin müşterilerini kaybetmesi karlılığı üzerinde büyük olumsuz etkilere neden olacağından, çoğu işletme kayıp müşteri modellerini sadakat programlarının temeline yerleştirmektedir (Rud, 2001: 257).

Müşteri kaybına neden olan unsurların ve gelecekte işletmeyi terk etme riski olan müşterilerin belirlenmesi; gerekli önlemlerin alınmasında, müşteri geri kazanma stratejilerinin geliştirilmesinde ve sonuç olarak otel işletmelerinin karlılığı üzerinde önemli rol oynamaktadır. Bu bağlamda, gerçekleştirilen bu çalışmanın amacı makine öğrenimi yöntemlerinden faydalanılarak bir zincir otel işletmesinin müşteri veri tabanında bulunan müşteri verilerine kayıp müşteri tahminlemesi uygulanmasıdır.

Bu amaçla:

- (1) İki farklı makine öğrenimi algoritması olan Lojistik Regresyon ve Rastgele Orman yöntemlerinin karşılaştırması yapılarak, otel işletmelerine yönelik müşteri verilerinde hangi tahmin modelinin daha başarılı olduğu tespit edilecektir.
- (2) Otel işletmelerine yönelik müşteri verileriyle gerçekleştirilen bir kayıp müşteri analizinde hangi değişkenlerin kayıp müşteri davranışını yordamada daha önemli olduğu tespit edilecektir.
- (3) Otel işletmesini bir sonraki dönem içerisinde terk etme ihtimali yüksek olan müşteriler önceden tespit edilecektir.

Daha önce turizm alanında sadece sosyal kulüp üyeleri (Maier ve Prusty, 2016:803), ticari otel müşterileri (şirketler) (De Wit, 2017: 2) ve otel rezervasyon web sitesi müşterileri (Han, 2018:1) için gerçekleştirilmiş kayıp müşteri tahminleme yönteminin bir resort otel zincirinin müşteri verilerine uyarlanmasıyla bu araştırmanın mevcut literatüre katkı sağlaması amaçlanmaktadır.

Çalışmanın birinci bölümünde otel işletmelerinde müşteri ilişkileri yönetimi ve uygulamaları, ikinci bölümünde makine öğrenimi ve kayıp müşteri analizine yönelik daha önce gerçekleştirilen çalışmalardan bahsedilecektir. Üçüncü bölümde ise araştırma yöntemi ve bulgular paylaşılacaktır. Son olarak, ulaşılan sonuçlardan bahsedilecek; otel işletmeleri ve gelecek çalışmalar için önerilere yer verilecektir.

BİRİNCİ BÖLÜM

OTEL İŞLETMELERİNDE MÜŞTERİ İLİŞKİLERİ YÖNETİMİ VE UYGULAMALARI

1.1. Müşteri İlişkileri Yönetimi (M.İ.Y.)

Son yıllardaki teknolojik gelişmeler dünyanın farklı yerlerindeki insanları birbirine bağlayamaya imkân sağlamış, dünyayı küresel bir köy haline getirmiş, dolayısıyla ekonomik çevreyi değiştirmiştir. Bu durum müşterilerin bilgiye daha hızlı ulaşabilmesine ve daha seçici olmalarına yol açarak müşteri sadakati sağlamayı daha zor hale getirmiştir. Değişen bu piyasa koşullarında işletmelerin hayatlarını sürdürebilmeleri için müşterilerini anlamaları (Srivastava vd., 2019: 38) ve diğer işletmelerden farklılaşmaları şarttır (Demir ve Kırdar, 2007: 293). Bu bağlamda, müşteri ilişkileri yönetimi - M.İ.Y. (customer relationship management - CRM), müşterilerin değişen ihtiyaçlarını anlama (Ryals ve Payne, 2001: 3) ve farklılaşma yolunda işletmelere fırsat tanımaktadır. Müşterilerle kurulan ilişkiler uzun vadede satın alma davranışlarını arttıracak ve müşteri sadakatini beraberinde getirecektir (Demir ve Kırdar, 2007: 293). Nitekim, M.İ.Y. sayesinde işletmeler mevcut müşteri verilerinden yararlanarak sundukları hizmetleri geliştirebilmekte ve karlılıklarını arttırabilmektedir (Ryals ve Payne, 2001: 3).

Müşterilerle uzun dönemli ilişkiler geliştirmeyi amaç edinen ilişkisel pazarlamanın tanımı, ilk defa Leonard L. Berry (1983:25) tarafından yapılmıştır. Berry (2002:71) bu konunun önemine şu soruyla dikkat çekmiştir: “Neden önemli miktardaki pazarlama bütçesi ve kaynak işletmemize yeni müşteriler çekip, bir defa müşterimiz olanları daha sonra görmezden gelmek için harcansın?” Çoğu işletme pazarlama amaçlı kullanılan kaynakları yeni müşteriler elde etmeye ayırmakta olup mevcut müşterileri elde tutmak için çok fazla çaba harcamamaktadır. Böyle bir pazarlama anlayışı işletmeleri kısıtlayarak kaynaklarını boşa harcamalarına neden olacaktır. Çünkü işletmeler yeni müşteriler edinerek ulaşacakları karlılığa mevcut müşterilerini ellerine tutarak da ulaşabilirler. Ayrıca, mevcut müşterileri elde tutmak için ayrılacak pazarlama bütçesi yeni müşteriler çekmek için yapılacak harcamaların çok daha altında kalmaktadır. Bu durum göz önünde bulundurulduğunda pazarlama çalışmalarında “sadece yeni müşteri” yaklaşımıyla hareket etmek, işletmeler açısından karlı olmaktan uzak bir uygulamadır (Berry, 2002: 60). Dolayısıyla ilişkisel pazarlama kısaca “müşterilerle ilişki kurmak, sürdürmek ve geliştirmek” olarak tanımlanabilmektedir (Gülmez ve Kitapçı, 2003: 81).

Başarılı bir işletme olmanın anahtarı potansiyel müşterilerin arzu ve ihtiyaçlarını tahmin edebilmeye ve proaktif davranarak uygun ürün ve hizmetleri sunabilmeye dayalıdır. Başarılı

işletmeler pazarlama stratejilerini tasarlarken müşterilerini ilk sıraya ve merkeze yerleştirirler. Çünkü müşteri davranışlarında herhangi bir değişiklik söz konusu olduğunda işletme karlılığında önceden öngörülemeyen etkiler ortaya çıkabilmektedir (Srivastava vd., 2019: 38). Bunun yanı sıra, işletmeler mevcut müşterini iyi tanırlarsa onlarla ilişkilerinde etkili ve doğru iletişimde bulunarak taleplerine hitap edebilirler. Bunun sonucunda işletmelerin müşterilerini elde tutma oranlarında ciddi bir artış sağlanabilir (Chen ve Popovich, 2003: 676). Bu bağlamda ilişkisel pazarlama, müşteri davranışlarını anlayıp yorumlayarak müşteriye verilen değeri hissettirebilmek ve müşterinin güvenini kazanmak, böylelikle karşılıklı bilgi alışverişini en üst seviyeye getirebilmeyi amaçlamaktadır. Bu amaçla ilişkisel pazarlama müşteri değerini ve müşteri memnuniyetini arttırarak müşteri sadakati sağlamayı hedeflemektedir (Gülmez ve Kitapçı, 2003: 81-82).

M.İ.Y. 90'lı yılların ortalarında ortaya çıkmış olup, satış elemanları ve bilgi işlem uzmanları tarafından geliştirilmiştir (McKim ve Hughes, 2001: 224). Diğer bir ifadeyle M.İ.Y. bilgi teknolojilerindeki gelişmelerden ve müşteri merkezli süreçlerdeki organizasyonel değişimlerden evrilmiştir (Chen ve Popovich, 2003: 672). Müşteri davranışlarını izleyen ve analiz eden M.İ.Y. teknolojileri, işletmelerin en karlı müşterilerini kolayca belirlemelerine ve pazarlama çalışmalarında daha fazla satın alma potansiyeli olan bu müşterilere odaklamalarına olanak sağlamaktadır (Chen ve Popovich, 2003: 676; McKim ve Hughes, 2001: 224).

M.İ.Y. yeni müşteriler edinmeye, daha önemlisi mevcut müşterilerle ilişkileri geliştirmeye ve sürdürmeye yardımcı olan bir işletme stratejisidir (Chen ve Popovich, 2003: 682). İlişkisel pazarlama kavramının ortaya çıkışından bu yana geçen süre içerisinde müşterileri elde tutmanın değeri, M.İ.Y.'in müşterileri elde tutma ve karlılık üzerindeki etkileri vb. konular ilgi çekmiştir. İlişkisel pazarlama uygulayan işletmeler artık pazarlama yatırımlarını sadece yeni müşteriler elde etmek için kullanmamakta, bunun yanı sıra mevcut müşterilerini elde tutmak ve onlarla olan ilişkileri geliştirmek için de kullanmaktadır (Berry, 2002: 69). M.İ.Y. ilişkisel pazarlama stratejilerinin uygulanmasında bilgi teknolojilerinin kullanılmasını ifade etmektedir (Ryals ve Payne, 2001: 3; Wilson vd., 2002: 194). Dolayısıyla, M.İ.Y. pazarlamada önemi gittikçe artan bir konu olarak öne çıkmaktadır (Ryals ve Payne, 2001: 3). Başarılı M.İ.Y. uygulayan işletmeler müşterileri tarafından sadakat ve uzun dönemli karlılıkla ödüllendirilmektedir (Chen ve Popovich, 2003: 672).

M.İ.Y. bilgi teknolojileri ve pazarlama stratejileri arasında stratejik bir köprü kurmaya çalışmaktadır. Bu çabaları gerçekleştirebilmek için müşteri odaklı bilgi-yoğun stratejilere ihtiyaç duyulmaktadır. M.İ.Y. kavramını “bilgi destekli ilişkisel pazarlama” olarak da tanımlamak mümkündür (Ryals ve Payne, 2001: 3). Bu bağlamda, hizmet sektörü bilgi

yönetimini (knowledge management) ve M.İ.Y.'i bilgi temelinde entegre etmektedir. Bilgi, M.İ.Y.'in iç boyutlarından birisidir. İlgili bilgiler müşterilere aittir, müşterilerden elde edilir ve müşterilerin memnuniyeti için kullanılırlar. İşletmeler müşteri verilerini farklı kanallardan ve “temas noktalarından” (customer touch points) elde edebilmektedir. Müşteri temas noktaları hizmetin sunumu sırasında müşteriyle etkileşime geçilen tüm anları ifade etmektedir (Srivastava vd., 2019: 41). Bir işletmenin müşteri temas noktaları interneti, e-postayı, satışı, tele-pazarlama operasyonlarını, müşteri hatlarını, kioskuları vb. içerebilmektedir. Bu müşteri temas noktaları genellikle ayrı bilgi sistemleriyle kontrol edilmektedir. M.İ.Y. bu temas noktalarını müşteri bakış açısı etrafında birbirine bağlamaktadır (Chen ve Popovich, 2003: 672 - 673). M.İ.Y. temas noktalarından elde edilen müşteri verisini analiz ederek satın alma davranışlarını tahminlemeyi (Chen ve Popovich, 2003: 676) ve davranışsal örüntüleri ortaya koyarak müşteri ilişkilerinin devamlılığını sağlamayı hedeflemektedir (Srivastava vd., 2019: 38).

M.İ.Y. müşteri merkezli bir iş modelidir. M.İ.Y., müşteri geri bildirimlerini onların bakış açısından ele alarak temel (core) iş süreçlerinin yeniden tasarlanmasını amaçlayan, süreklilik arz eden yönetsel ve operasyonel bir çabadır (Chen ve Popovich, 2003: 682). M.İ.Y., müşteri verisini aramayı, toplamayı ve depolamayı; bu veriyi doğrulayarak tüm örgüt içerisinde paylaşmayı; kişiselleştirilmiş müşteri deneyimleri yaratmak için veriyi tüm departmanlara yaymayı amaçlamaktadır (Sigala, 2006: 334). M.İ.Y. çalışmalarının müşteri memnuniyeti ve sadakati üzerinde etkili olabilmesi ve hizmet sektörünün yüksek performans gösterebilmesi ancak bilgi teknolojileri ve yazılımlarının işletmeler içerisinde tüm operasyonel ve fonksiyonel departmanlarla alt departmanları kapsayacak şekilde entegre edilmesine ve desteklenmesine bağlıdır (Srivastava vd., 2019: 40). Bu yönüyle M.İ.Y. insanların, süreçlerin ve teknolojinin bir kombinasyonudur ve işletmelerin bütüncül bir yaklaşımla müşterileri anlama çabasıdır. M.İ.Y. teknolojileri ön plan (satış, pazarlama, müşteri hizmetleri vb.) ve arka plan (finans, lojistik, insan kaynakları vb.) iş süreçlerini müşteri temas noktalarıyla birbirine bağlamaktadır (Chen ve Popovich, 2003: 672 - 673). Nitekim, başarılı bir M.İ.Y. uygulaması için müşteri verisinin kaydedilmesi ve bir veri tabanının oluşturulması önem arz etmektedir. Bu açıdan bilgi yönetimi ve veri madenciliği teknolojileri M.İ.Y.'in başlıca unsurlarındandır. M.İ.Y. teknoloji tabanlı bir strateji olup temelinde bilgi yönetimi ve veri madenciliği süreçleri bulunmaktadır (Srivastava vd., 2019: 41). Bir işletmede M.İ.Y.'nin başarılı olabilmesi için gerekli kritik başarı faktörleri aşağıdaki şekilde özetlenebilir (King ve Burgess, 2008: 423):

- İşletmede üst yönetimin tam desteği
- İşletmede M.İ.Y. stratejisi kapsamında gerekli iletişimin sağlanması

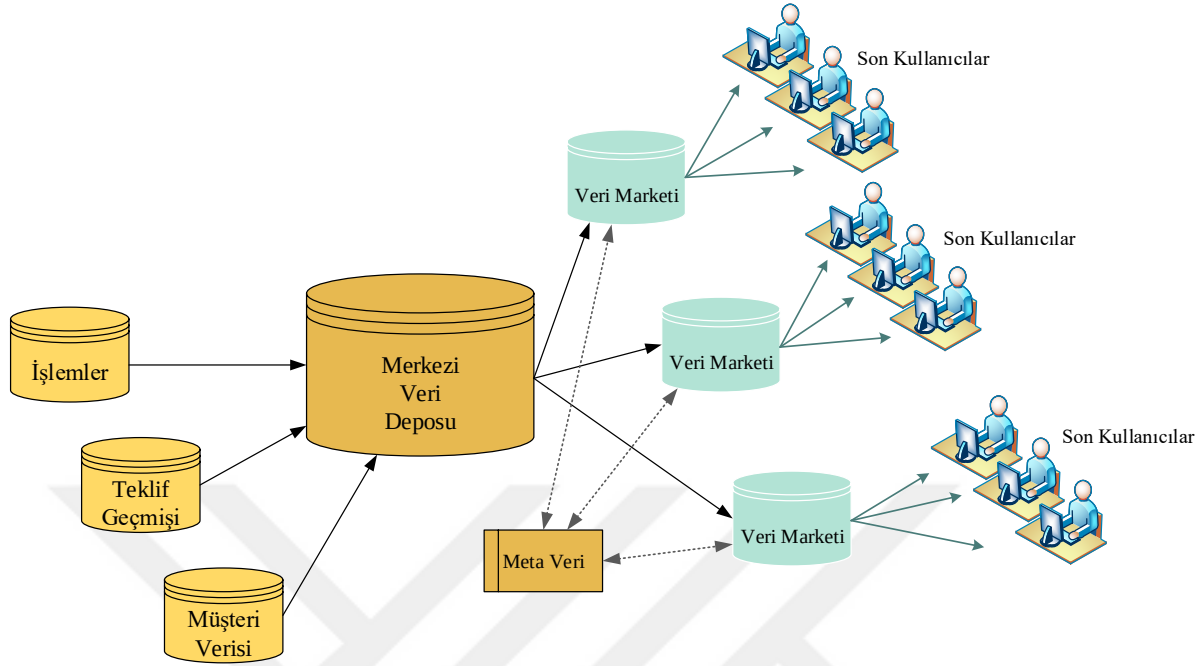
- İşletmenin bilgi yönetimi konusundaki yetkinliği
- İşletme içinde veri paylaşımı yapacak bölümlerin istekli olması
- İşletme çalışanlarının değişim süreçlerinde istekli olması
- İşletmenin teknolojik alt yapısının hazır olması
- İşletmenin kültürel değişim konusundaki yetkinliği ve müşteri odaklı olması
- İşletmenin süreç değişimleri konusundaki yetkinliği
- İşletmenin sistem entegrasyonu konusundaki yetkinliği

Sonuç olarak, M.İ.Y. müşterilerin karlılık potansiyelini arttırmaya ve müşterileri elde tutmaya odaklanmaktadır. Dolayısıyla M.İ.Y. açısından kayıp müşteri yönetimi önem arz eden bir konudur (Liu vd., 2019: 3).

1.2. Veri Tabanlı Pazarlama

Otuz yılı aşkın süredir M.İ.Y. uygulamalarında ilişkisel pazarlama, doğrudan pazarlama ve veri tabanlı pazarlama teknikleri bir arada kullanılmaktadır (Chaffey ve Ellis-Chadwick, 2019: 300). M.İ.Y. uygulamaları ile bir işletmedeki tüm müşteri verisinin kaydedildiği büyük veri ambarları (data warehouse) oluşturulmaktadır (McKim ve Hughes, 2001: 224). Veri ambarı iki ya da daha fazla veri tabanındaki bilgileri birbirine bağlayan bir yapıdır (Şekil 1.1). Veri kaynaklarını kullanarak veriyi merkezi veri deposuna alır, birleştirir, temizler ve özetleyerek veri marketlerine (data mart) dağıtır. Veri marketleri departmanlara özgü veri ambarları olup sadece son kullanıcının ihtiyaç duyacağı veriyi içermektedir. Veri üzerinde analiz yapacak kişiler ilgili veri marketindeki veriyi kullanmaktadır. Meta veri ise veri marketleri için izin görevi görmektedir (Rud, 2001:31). Müşteri verisi üç temel türde olabilmektedir: demografik veri, davranışsal veri ve psikografik yani tutumsal veri. Her veri türünün birtakım güçlü ve zayıf yönleri bulunmaktadır (Şekil 1.2). Demografik veri genel olarak kişisel özellikleri tanımlamakta ve cinsiyet, yaş, gelir, medeni durum, etnik köken, eğitim düzeyi, sahiplik vb. özellikleri içermektedir. Bu tarz veriler davranışsal verilerden farklı olarak çok sıklıkla değişiklik göstermemektedir. Dolayısıyla demografik veri sabit yani durağan bir yapıda olması sayesinde tahminleme analizlerinde sıklıkla tercih edilmektedir. Ancak müşterilerin çoğunlukla kendileriyle ilgili bu tarz bilgileri paylaşmaktan çekindiği ya da yanlış bilgi verdiği bilinmektedir. Davranışsal veri ise harcama tutarları, satın alım türü ve tarihi, ödeme tarihi ve tutarı, web sitesi üzerindeki etkinlik vb. türdeki davranışsal bilgileri içermektedir. Bu veri türü tahmin yeteneği en yüksek olan veri türüdür. Bununla birlikte elde edilmesi en maliyetli olan bilgileri kapsamaktadır. Diğer taraftan tahminleme analizlerine ilave bir boyut kazandıran psikografik veriler müşterilerin düşüncelerini, yaşam tarzlarını ve kişisel değerlerini

içermektedir. Bu türdeki veri müşteri anketleri ve odak grup çalışmalarıyla toplanabildiğinden elde edilmesi maliyetli olmaktadır (Rud, 2001: 26).



Şekil 1.1 Veri Ambarı Örneği

Kaynak: Rud, 2001: 32

Veri ambarı oluşturulurken işletmedeki bilgi işlem ve pazarlama departmanları birlikte çalışarak operasyonel veri tabanı ile pazarlama veri tabanını birbirinden ayırmalıdır. Operasyonel veri tabanı müşterinin satın alımlarına ve finansal işlemlerine yönelik bilgi içermektedir. Pazarlama veri tabanı ise mevcut müşterilerin, kayıp müşterilerin ve potansiyel müşterilerin bilgilerini içermekte olup pazarlama departmanı tarafından yönetilmektedir. Veri tabanı pazarlamasında kullanılacak standart bir veri ambarında müşterilere yönelik aşağıdaki veriler bulunabilmektedir (Clow ve Baack, 2014: 313-314):

- Müşteri isim ve adresleri
- E-posta adresleri ve işletmenin web sitesi ziyaretlerinin kaydı
- Satın alma işlemlerinin kayıtları
- Müşteri şikâyetleri, bilgi talepleri ve müşterilere verilen yanıtlar vb. müşteri etkileşimlerinin kayıtları
- Müşteri anketlerinin sonuçları
- Müşteri tercihleri
- Müşteri profilleri
- Pazarlama promosyonları

- Pazarlama kampanyaları ve müşterilerin ilgili kampanyalara yanıtları
- Demografik ve psikografik veriler
- Müşteri kümeleri (bölümlendirme) ve müşteri yaşam boyu değeri vb. veri tabanı analiz sonuçları

İşletmeler elde ettikleri veriyi analiz ederek ortaya çıkan bulguları müşterileri tanımlamaya, işletmeye çekmeye ve elde tutmaya yönelik stratejiler oluşturmak ve kullanılacak stratejilere karar vermek için kullanmaktadır. Elde edilen veri analiz edilerek daha anlamlı hale getirilebilmekte ve yorumlanarak raporlanabilmektedir. Veri analizi sayesinde gelecekteki müşteri trendleri tahmin edilebilmektedir (Srivastava vd., 2019: 38).

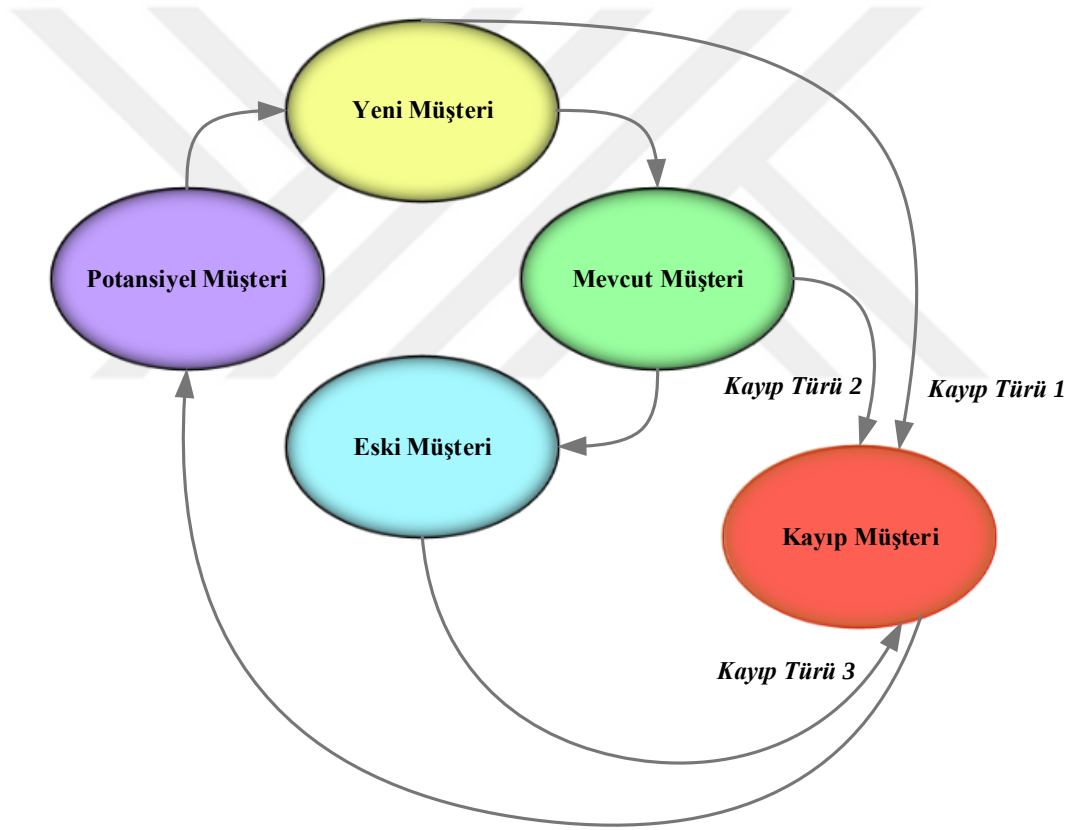
	<i>Tahmin Gücü</i>	<i>Sabitlik</i>	<i>Maliyet</i>
<i>Demografik Veri</i>	Orta	Yüksek	Düşük
<i>Davranışsal Veri</i>	Yüksek	Düşük	Yüksek
<i>Psikografik Veri</i>	Orta	Orta	Yüksek

Şekil 1.2 Veri Türleri ve Özellikleri

Kaynak: Rud, 2001:27

Veri tabanlı pazarlama müşterilerden elde edilen bilgilerin analiz edilmesiyle ulaşılan sonuçların pazarlama faaliyetlerini geliştirecek şekilde kullanılması olarak tanımlanmaktadır (Kitapçı, 2006: 2). Veri tabanlı pazarlama (database marketing) 1980'lerin ortalarında pazarlama uzmanları tarafından geliştirilmiştir. Veri tabanlı pazarlama uygulayan işletmeler pahalı olmayan, sadece etkili pazarlama çalışmaları için ihtiyaç duyulan ve veriyi saklayan veri tabanlarını kullanmaktadır. İşletmeler yapılan analizler aracılığıyla müşteri sadakatini, elde tutma oranlarını, satışları ve karlılığı arttırmayı amaçlamaktadır (McKim ve Hughes, 2001: 224). Nitekim hizmet işletmeleri iş süreçleri boyunca büyük miktarda veri toplamaktadır. Toplanan veri, veri madenciliği yardımıyla hızlı ve uygun şekilde işlenmekte ve elde edilen bilgi müşteri memnuniyetini hedefleyen karar süreçlerine katkı sağlamaktadır. Çoğu hizmet işletmesi kendi veri tabanlarını oluşturmakta ve veri madenciliği teknolojilerinden yararlanmaktadır. Bu veri tabanı yönetim sistemleri özellikle kişiselleştirilmiş hizmetlerin sunumuna olanak sağlamak; müşterilerin kişisel zevklerini, tercihlerini, hoşlarına giden ve gitmeyen konuları vb. anlamayı kolaylaştırmaktadır (Srivastava vd., 2019: 38). Veri tabanlı pazarlamanın en çarpıcı örneklerinden biri web mikro sitesi ile veri tabanının aynı sunucu içerisine yerleştirilmesi ya da aynı sunucudaymış gibi işlev görmesi için elektronik olarak birbirine bağlanmasıdır. Böylece ortak çalışan veri tabanı ve mikro site müşterilere yönelik bir iletişim aracı haline gelmektedir. Müşterilerin bilgisayarına çerezler (cookies) yerleştirilerek müşteriler web sitelerine giriş yaptığında kendilerine göre düzenlenmiş ve görmeyi arzu ettikleri konuları içeren web sayfalarıyla karşılaşmaktadırlar (Hughes, 2003: 6-7).

Hughes'a (2003: 251) göre tüm müşterilere aynı şekilde davranmak ve hepsine üst kalitede hizmet sunmayı amaçlamak doğru bir strateji değildir. Çünkü müşteriler birbirlerinden farklıdır ve işletmeler zaman içerisinde tüm müşterilerine aynı şekilde davranmamaları gerektiğini öğrenmektedir. Kimi müşteriler sadık olma eğilimindedir ve işletmelere güven duyarlar. Kimi müşteriler ise bütçelerinin büyük kısmını farklı işletmelere harcarlar. Bu tarz müşteriler kar sağlamaktan ziyade işletmelere maliyet oluşturmaktadır (Hughes, 2003: 251). M.İ.Y. uygulamaları aracılığıyla işletmeler öncelikli müşterilerinin profilini belirlemekte, hatta müşteri bazlı portreler oluşturabilmektedir (Chen ve Popovich, 2003: 676). Bu anlayışın temelini oluşturan bire-bir pazarlama (one to one marketing) yaklaşımı çoğunlukla pazarlama otomasyonu olarak adlandırılmaktadır.



Şekil 1.3 Müşteri Yaşam Döngüsü

Kaynak: Qi vd., 2009: 263

Son yıllarda işletmeler kişiselleştirilmiş e-posta gönderimlerini ve web sitesi sayfalarını oluşturmak için pazarlama otomasyonu teknolojilerinden faydalanmaktadır. Ancak bireysel düzeye taşınan ilişkileri yönetmek yüksek maliyetli bir durumdur. Bu nedenle işletmeler pazarlama otomasyonunu kişiselleştirilmiş iletişim kurmak ve müşteri teklifleri sunmak için

kullanmakta ve bunu belirli müşteri bölümlerini veya gruplarını hedef alarak gerçekleştirmektedir (Chaffey ve Ellis-Chadwick, 2019: 300).

İşletmelerin müşterilerini tanıyabilmesi ve müşteri bölümlerini oluşturabilmesi için bir müşteri tanımlama (customer identification) sürecine ihtiyaçları vardır. Bu süreç içerisinde işletmeler müşterilerini analiz etmekte, potansiyel müşteri bölümlerini ortaya çıkarmakta ve hedef müşterileri belirlemektedir. Bu süreç müşteri yaşam döngüsü yönetiminin temelidir (Şekil 1.3). Müşteri tanımlama süreci içerisinde potansiyel müşteriler belirlenebilir ve bir kısmı yeni müşterilere dönüştürülebilir. Yeni müşteriler işletmeye katıldıktan sonra süreç şu şekilde işleyecektir: *yeni müşteri* → *mevcut müşteri* → *eski müşteri*. Bu süreç içerisinde müşteri kaybı her adımda ortaya çıkabilir. Ayrılan müşteri ise (churner customer) yeniden işletmenin potansiyel müşterisi olabilir. Müşteri tanımlama sayesinde, işletmeler hedefledikleri en karlı müşteri bölümünün özelliklerini belirleyebilir, bunun yanı sıra kayıp müşteri tahminlemesi yaparak gelecekte ayrılma potansiyeli olan müşterileri ayır edebilir (Qi vd., 2009: 262- 263).



Şekil 1.4 Öncelik Matrisi Örneği

Kaynak: Qi vd., 2009: 260

Araştırmacılar regresyon modelleri sayesinde hangi müşterilerin hizmet alımını durdurmaya daha yatkın olduğunun tahminlenebileceğini keşfetmişlerdir. Bu modeller müşteri işlem geçmişi ile diğer müşteri verilerinin kombinasyonunu kullanarak müşteri kaybında öncü rol oynayan değişkenlerin belirlenmesini sağlamaktadır. Bu sayede gelişmiş veri tabanı

pazarlama teknikleri kullanılarak işletmeyi terk etme potansiyeli olan müşterilere odaklanılabilmektedir. Kayıp müşteri tahminleme ve müşteri yaşam boyu değeri analizleri hangi müşterilerin daha karlı olduğuna yönelik kapsamlı bilgiler sunmaktadır (Hughes, 2003: 11-12). Bu bilgiler müşterileri elde tutma stratejilerinde yol gösterici olan öncelik matrisinde kullanılmaktadır. Şekil 1.4'te sunulan matriste müşterilerin ayrılma olasılığı ve işlemeye kazandıracağı değer değişkenleri düşük, orta ve yüksek olması durumlarına göre ayrı bölümlere ayrılmıştır. Müşterinin tekrar gelmesi durumunda kazandıracağı değer ne kadar yüksekse, elde tutmaya yönelik teşviklerin sunulması konusunda aynı ayrılma ihtimali düzeyindeki diğer müşterilerden daha yüksek önceliğe sahiptir. Bununla birlikte kazandıracağı değer aynı düzeyde olan müşteriler arasında ayrılma ihtimali yüksek olan müşteriler daha öncelikli olacaktır. Çünkü müşterinin kazandıracağı değerle müşteriyi kaybetme ihtimali kıyaslandığında müşterinin ayrılma olasılığının yüksek olması öncelik düzeyini arttırmaktadır. Dolayısıyla öncelik matrisi üzerindeki 1 numaralı kısımda yer alan müşteriler elde tutma stratejileri için en yüksek önceliğe, 9 numaralı kısımda yer alan müşteriler ise en düşük önceliğe sahiptir (Qi vd., 2009: 260).

Veri tabanı pazarlamasının müşteri sadakati ve tekrar eden satışlar üzerindeki etkisi ancak müşteri faydası sağlandığında ortaya çıkabilir. Eğer veri tabanındaki bilgiler müşterileri memnun etmiyorsa, müşteriler işletmenin iletişim çabalarını görmezden gelebilir, sadakat kartlarını kullanmayabilir veya rakip işletmeleri tercih edebilir. Bu gerçeği göz ardı eden çoğu işletmenin veri tabanı pazarlamasında başarısızlığa uğradığı görülmektedir (Hughes, 2003: 2-3).

1.3. Elektronik Müşteri İlişkileri Yönetimi (E-M.İ.Y.)

M.İ.Y. müşteri problemlerini çözmek, kişiselleştirilmiş hizmet ve deneyimler sunmak için müşteri bilgilerinin toplanmasını, analiz edilmesini ve kullanılmasını amaçlamaktadır. İşletmelerin müşterilerle ilişki kurmak ve yönetmek için web sitelerindeki internet araçlarını kullanmaları ise elektronik müşteri ilişkileri yönetimi (E-M.İ.Y.) olarak ifade edilmektedir (Sigala, 2006: 336). Chaffey ve Ellis-Chadwick'e göre (2019: 302) E-M.İ.Y, veri tabanları, kişiselleştirilmiş mesajlar, müşteri hizmetleri, e-posta ve sosyal medya pazarlaması gibi tekniklerden yararlanılarak çevrimiçi hizmet kullanımının devamlılığını sağlamak ve müşterilerin karlılığını arttırmak için dijital iletişim teknolojilerinden yararlanılmasıdır. E-M.İ.Y. uygulamalarını diğer yönetim anlayışlarından ayıran en önemli özelliği işletmenin müşteriyi ilişki kurmada, ilişkiyi güçlendirmede ve kişisel düzeye taşımada teknolojinin gücünden faydalanmasıdır. E-M.İ.Y. uygulamaları aracılığıyla müşterilerin detaylı profilleri

oluşturularak pazar bölümlendirme yapılabilir ve kişiselleştirilmiş hizmet sunumunda bulunulabilir (Caber, 2010:1).

Günümüzde, çevrimiçi M.İ.Y. uygulamalarını genişletmek ve bu uygulamaları çevrimdışı M.İ.Y. programlarıyla bütünleştirmek mümkün hale gelmiştir. E-M.İ.Y. uygulamaları turizm ve otel işletmelerinin hali hazırda kullanmakta oldukları bilişim altyapısına kolaylıkla entegre edilebilmektedir (Caber, 2010:1). Dolayısıyla, E-M.İ.Y. uygulamaları, konaklama işletmesi müşterilerinin memnuniyet düzeyleri ve müşteri olma niyetleri üzerinde etkili olabilmektedir (Sigala, 2006: 333).

E-M.İ.Y dijital teknolojinin ve dijital verilerin M.İ.Y. çalışmalarını desteklemesi amacıyla geliştirilen stratejileri ve oluşturulan planları kapsamaktadır. Dijital pazarlama çalışmaları da E-M.İ.Y. kapsamı içindedir. E-M.İ.Y uygulamaları özetle aşağıdaki çalışmaları içermektedir (Chaffey ve Ellis-Chadwick, 2019: 302):

- Web sitesi kullanılarak ve çevrimiçi sosyal ortamlarda yer alınarak müşteri elde etmeyi sağlamak; çevrimiçi veya çevrimdışı satış iletişiminde e-posta ve web-temelli içerik kullanılarak satın almayı özendirmek
- Müşteri profilini yönetmek, e-posta gönderim listeleriyle müşteri profillerinin entegrasyonunu sağlamak ve hedef müşterileri belirlemek
- Mobil, e-posta ve sosyal medya müşteri iletişim olanaklarını üst (up-sell) ve çapraz satışı (cross-sell) destekleyecek şekilde yönetmek
- Veri madenciliği ile hedef müşteri belirleme çalışmalarını desteklemek
- Çevrimiçi ortamda kitlesel kişileştirme imkânları sunarak “bir sonraki en iyi ürünü/ hizmeti” tavsiye etmek
- Çevrimiçi müşteri hizmetleri sunmak (sık sorulan sorular, geri arama ve canlı görüşme desteği vb.)
- Çevrimiçi hizmet kalitesi yönetilerek ilk defa satın alma yapan müşterilerin tekrar satın almasına yol açacak bir deneyim sunmak
- Satın alma sürecinin ve müşteri yaşam döngüsünün bir parçası olarak farklı medya araçlarını kullanmak ve bu şekilde çok kanallı (multi-channel) bir müşteri deneyimi yaratmak

Lee-Kelley vd. (2003: 247) tarafından gerçekleştirilen bir çalışmanın bulgularına göre E-M.İ.Y. uygulamaları internet müşterilerinin sadakat düzeyini arttırmaktadır. Chen ve Chen (2004: 338) ise işletmelerde E-M.İ.Y. uygulamalarının sağladığı faydaları somut ve soyut olmak üzere iki başlık altında ele almıştır:

- **Somut Faydalar**

- İşletme gelirini ve karlılığını artırır
- İş dönüş süresini (turnaround time) hızlandırır
- İşletme içi maliyetleri azaltır
- İş gören verimliliğini artırır
- Pazarlama maliyetlerini azaltır
- Müşteri elde tutma oranlarını artırır
- Pazarlama yatırımlarının geri dönüşünü artırır

- **Soyut Faydalar**

- Müşteri memnuniyetini artırır
- Müşterilerin işletmeyle ilgili olumlu ağızdan ağıza reklam yapmasını sağlar
- Müşteriye sunulan hizmet kalitesini artırır
- İş süreçlerini geliştirir
- Daha yakın iletişim yönetimi sağlar (işletmeye ulaşmak kolaylaşır)
- Daha etkili ve derinleştirilmiş müşteri bölümlene yapabilmeyi sağlar
- Hedef müşterileri belirlemede ve müşteri profillerini oluşturmada daha net olmayı sağlar
- Müşteri taleplerini daha iyi anlamayı ve müşteri taleplerine daha doğru şekilde hitap edebilmeyi sağlar

Konaklama işletmeciliğinde, E-M.İ.Y. aracılığıyla yüksek fayda ve değer elde eden müşterilerin ilgili işletme ile ilişkilerini sürdürmeye devam ettikleri, bu sayede işletme karlılığının arttığı görülmektedir. Müşterilerin E-M.İ.Y. uygulamalarına katılımı ve bunun sürdürülebilirliği sağlanarak müşteri faydasının artırılabilmesi için otel işletmelerinin müşterileriyle olan bu ilişkilerinde neye değer verildiğini bilmesi ve elektronik hizmet kalitesi bağlamında hangi faydaların elde edilmek istendiğini anlaması gerekmektedir (Sigala, 2006: 334). Bunun yanı sıra E-M.İ.Y. uygulamalarının başarılı olabilmesi için işletmedeki bütün çalışanların ve üst yönetimin katılım ve desteğine ihtiyaç vardır (Caber, 2010:1).

Gilbert vd. (1999: 26-27) otel işletmelerinin E-M.İ.Y. uygulamalarına yönelik beş adım önermiştir:

- (1) Veri tabanı analiziyle müşteriyi tanımak (tanımla - identify): Bir otel web sitesi müşterinin ne istediğini ve ne satın aldığını ölçebilmelidir. Otel web sitesine yerleştirilen farklı tekliflere müşterilerin nasıl karşılık verdiği izlenebilmelidir. Bu ölçümler aracılığıyla müşteri profilleri oluşturulabilir.

- (2) Ürün/hizmeti geliştirmek ve daha çekici hale getirmek (geliştir - improve): Devamlı müşterilere kişiselleştirilmiş çevrimiçi anketler uygulanarak otelde sunulan hizmetlere ya da sadakat programlarına yönelik düşünceleri öğrenilebilir. Otel işletmesi bu bilgileri kullanarak müşterilerin beklentilerini karşılamayan hizmet unsurlarını geliştirebilir ve müşterilerin beklentilerini aşacak şekilde hizmet kalitesini arttırabilir.
- (3) Müşterilerin otel işletmesi hakkında bilgi edinmesini sağlamak (bilgilendir- inform): Otel işletmelerinin web siteleri stratejik bilgi merkezi olarak kullanılabilir. Otel işletmesi web sitesi aracılığıyla müşterilerini yeni ürün / hizmetler ve özel teklifler hakkında bilgilendirebilir.
- (4) Müşterileri yeni ürünler denemek, düzenli satın alım yapmak vb. için özendirmek (aklını çel - tempt): Otel işletmesi devamlı müşterilerine özel e-posta gönderimleriyle promosyon broşürleri iletebilir. Ayrıca web sitesini çapraz satışa özendirecek şekilde tasarlayabilir. Örneğin Hilton Hotel web sitesinde (www.hilton.com) bir elektronik asistan bulunmaktadır. Bu asistan müşterilere Hilton tarafından sunulan diğer hizmetleri denemelerini özendirecek kısa mesajlar göstermektedir.
- (5) Sadakat programları aracılığıyla müşterileri elde tutmak (elde tut - retain): Otel işletmesi sadık müşterilerine özel kullanıcı adı ve şifre ile giriş yapılabilen ayrı bir web sayfası tasarlayabilir ve bu web sayfasında sadece üye müşterilerine özel dergiler ve özel teklifler yayımlayabilir. Örneğin Sheraton otellerinin üyelik programı, müşterilerinin çevrimiçi etkinliklerini takip etmekte ve üye müşterilerine özel hediyelerini çevrimiçi olarak sunmaktadır.

1.4. Dijital Pazarlama

Teknolojik gelişmeler yaşanmadan önce işletmeler pazarlama faaliyetlerini telefon ve posta aracılığıyla gerçekleştirmekteydi. Müşteri siparişlerini zamanında gönderebilmek uğraş verici ve maliyetli olmaktaydı. Günümüzde, müşteriler web üzerinden siparişini ilettiğinde hemen bir “teşekkür mesajı” ile karşılaşmakta ve e-posta onayı almaktadır. Müşteri bu e-posta üzerinde bulunan bağlantıya tıklayarak dilediğinde siparişini iptal edebilmektedir. Sipariş iptal taleplerinin henüz gönderi yola çıkmadan biliniyor olmasıysa işletmeyi ek bir maliyetten kurtaracaktır. Ürün kargoya verildiğinde ürünün gönderim bilgisi müşteriye iletilmekte, müşteri ürünü aldıktan sonra ise ürünün zamanında ulaşp ulaşmadığını ya da ürünle ilgili memnuniyetini yine basit bir bağlantıya tıklayarak kolaylıkla işletmeye iletebilmektedir (Hughes, 2003: 275). Özellikle Web 4.0’ın hayatlarımıza girmesiyle müşterilerin işletmelerle olan etkileşimlerinde ve iletişim şekillerinde köklü değişiklikler yaşanmıştır. Web 4.0 ile

müşterilerin katılımı ve gücü artmış; içerik üretebildikleri sosyal medya platformlarının ve farklı iletişim araçlarının bulunduğu interaktif bir internet dünyası ortaya çıkmıştır. Günümüz dünyasının müşterileri ürün/hizmetleri araştırmak, birbirleriyle karşılaştırmak, daha önceki kullanıcıların tecrübelerine danışmak, işletmelerle ve diğer müşterilerle iletişime geçmek ve satın alım yapmak için internet kullanmaktadır. Etkili bir pazarlama programı ise işletmenin reklam ve tutundurma çalışmalarına bu yeni unsurların adapte edilmesini gerektirmektedir. Bu bağlamda dijital pazarlama; e-ticaret, internet pazarlaması ve mobil pazarlamanın bütün özelliklerini birleştiren ve dijital ayak izine sahip her türlü aracı bünyesinde barındıran bir pazarlama anlayışı olarak ortaya çıkmıştır (Clow ve Baack, 2014: 254).

Dijital pazarlama televizyon, basılı doküman, reklam panoları gibi geleneksel araçlardan farklı olarak internet ve mobil araçların kullanıldığı tanıtım faaliyetlerini ifade etmektedir. Dijital pazarlamayı tanımlamada e-pazarlama, interaktif pazarlama ve web pazarlaması kavramları da kullanılabilmektedir. Dijital pazarlama faaliyetleri elde etme, dönüştürme, ölçümleme/optimize etme ve elde tutarak geliştirme adımlarını içermekte olup (Altındal, 2013: 1147), bu adımların gerçekleştirilmesinde arama motoru pazarlaması ve optimizasyonunu, sosyal medya pazarlaması, içerik pazarlaması, e-posta pazarlaması ve uygulama geliştirme gibi yöntemlerden beslenmektedir (Şekil 1.5).



Şekil 1.5 Dijital Pazarlama Uygulama Teknikleri

Kaynak: Lanz ve Carmichael, 2015: 1

Dijital pazarlama uygulamalarında takip edilen bu adımlar aşağıdaki şekilde özetlenebilir (Altındal, 2013: 1147):

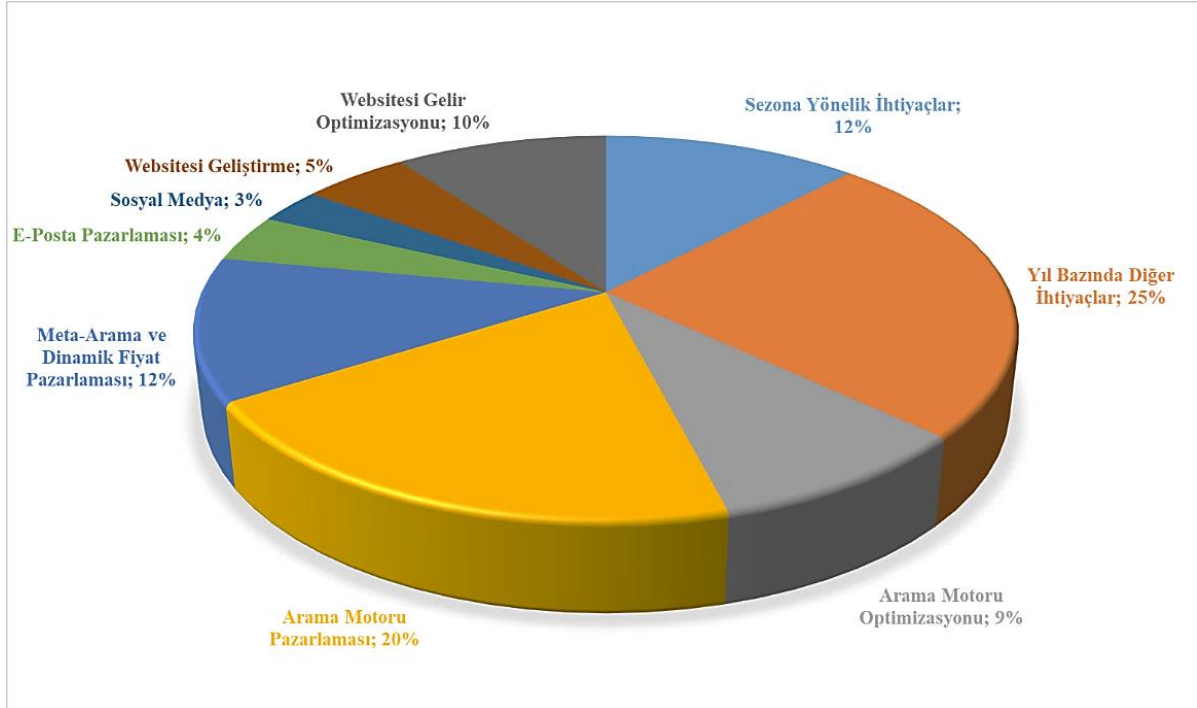
- Elde etme (Acquire): Online ortamdaki müşteriye satışın gerçekleştiği web sitesine/ alana çekmek amacıyla gerçekleştirilen faaliyetlerdir. Bu aşamada arama motoru optimizasyonu, arama motoru pazarlaması ve e-posta pazarlaması gibi yöntemler kullanılabilmektedir.
- Dönüştürme (Convert): Müşteri satışın gerçekleştiği web sitesine / alana giriş yaptıktan sonra satıcının hedefi doğrultusunda yönlendirilmesini sağlayacak faaliyetlerdir. Bu aşamada satıcının hedefi her zaman için satış yapmak olmayabilir. Bir otel web sitesi için örnek verilirse, odaların ve hizmetlerin belirli bir süre boyunca görüntülenmesi bu aşamadaki hedefler arasında yer alabilmektedir.
- Ölçümleme/ optimize etme (Measure/Optimize): Bu aşamada pazarlama uzmanları hedeflerine ne derecede ulaştıklarının ölçümlemelerini yapmaktadırlar.
- Elde tutarak geliştirme (Retain/Grow): Müşterilerle uzun dönemli ilişki kurarak bağlılıklarını sağlamak adına gerçekleştirilen faaliyetleri kapsamaktadır. Bu aşamada E-M.İ.Y. uygulamalarında bulunulabilir.

Akıllı telefon ve tabletlerin hayatımıza girmesiyle birlikte mobil pazarlama yöntemleri de ayrı bir dijital pazarlama alanı olarak ortaya çıkmıştır. Bu durum işletmelerin web sitelerini mobil araçlara uyumlu hale getirmeleri ve bu platformlara özgü içerik oluşturmalarına ihtiyaç oluşturmıştır. Özellikle facebook, twitter vb. sosyal platformların mobil uygulamalarının akıllı telefonlarımızda mevcut bulunması nedeniyle işletmeler dijital pazarlama uygulamalarına günden güne daha fazla yönelmektedir. Dolayısıyla işletmelerin bu alana yaptıkları yatırımlar ivme kazanmaya başlamıştır (Safko, 2012'den akt. Bulunmaz, 2016: 358).

Otel gelirlerinin çok küçük bir kısmının pazarlama faaliyetlerine ayrılabilirdiği göz önünde bulundurulduğunda bu bütçenin ne kadar etkili kullanıldığı otel işletmesinin başarısını doğrudan etkilemektedir. Son zamanlarda otel rezervasyonlarının yarısından fazlasının çevrimiçi platformlar üzerinden gerçekleştirildiği ve mobil uygulamalardan gerçekleştirilen sorguların bilgisayarları geçtiği bilinmektedir. Dolayısıyla günümüzün dünyası teknoloji evrenine tamamen yerleşmiş, otel işletmelerinin pazarlama departmanları kaynaklarını dijital boyuta yönlendirmiştir (Lanz ve Carmichael, 2015: 1). Otel işletmeleri çevrimiçi ortamdaki varlıklarını yönetebilmek için bir dijital pazarlama planına ihtiyaç duymaktadır. Dijital pazarlama planı kapsamında otel işletmesi müşterileriyle iletişim yönetimi stratejisi uygulayabilir. Bu strateji kapsamında müşterilerini çevrimiçi platformlarda yorum yapmaları

için cesaretlendirebilir ve bu portallarda yer alan müşteri değerlendirmelerini yanıtlayabilir. Ayrıca yorum portallarındaki bilgi ve göstergeleri kullanmak, otel web sitesi üzerine bir bağlantı ile yorum portallarında yer alan değerlendirmeleri çekmek, online seyahat acentelerinin yönetim raporlarını kullanmak ya da online itibar yönetim sistemlerinden faydalanarak yorum portallarında yer alan müşteri değerlendirmelerini izlemek ve analiz raporları almak vb. şekillerde dijital bilgiyi pazarlama çalışmalarında etkin bir şekilde kullanabilir (De Pelsmacker vd., 2018: 49). De Pelsmacker vd. (2018: 54) tarafından gerçekleştirilen bir çalışmanın sonuçları çevrimiçi yorumların izlenerek müşteri yorumlarına hızlı bir şekilde dönüş yapılması ve bir dijital pazarlama planına sahip olunmasının otellerin oda dolulukları bazında performansları üzerinde etkisi olduğunu göstermektedir.

Pazarlama uzmanları otellerin pazarlama bütçesinin %75'ini dijital pazarlama faaliyetlerine ayrılmasını önermektedir (Şekil 1.6). Elektronik ağızdan ağıza iletişim (electronic word of mouth - eWOM), “satın alınan ürünler ile kullanılan hizmetlerin özellikleri hakkında ya da bu ürün satıcıları ile hizmet sağlayıcıları hakkında tüketiciler arasında internet teknolojileri kullanılarak gerçekleştirilen enformel bilgi paylaşımlarının tümü” olarak tanımlanabilir (Litvin, 2008'den akt. De Pelsmacker vd., 2018: 47). Elektronik ağızdan ağıza iletişim birçok farklı formda bulunabileceği gibi en önemlisi çevrimiçi portallardaki yorumlardır. Bu yorumlar müşterilerin tutumları ve satın alım davranışları üzerinde etkili olmaktadır (De Pelsmacker vd., 2018: 47).



Şekil 1.6. Otel İşletmelerinde Dijital Pazarlama Bütçesinin Kullanımı

Kaynak: HeBS Digital's The Smart Hotelier's Guide to 2016 Digital Marketing Budget Planning'den akt. Lanz ve Carmichael, 2015: 3

Tripadvisor, Holidaycheck vb. çevrimiçi müşteri yorum portallarının müşterilerin otel seçimleri üzerindeki etkisi oldukça çarpıcıdır. Online portallardaki olumlu müşteri yorumları yeni müşterileri çekmede ve dolayısıyla otel gelirlerini arttırmada önem arz etmektedir. Bunun yanı sıra Tripadvisor vb. portallar müşterilerin otel özelliklerini ve fiyatlarını karşılaştırmalarına yarayan meta arama siteleridir ve aylık 350 milyonun üzerinde ziyaretçi sayısına ulaşabilmektedir.

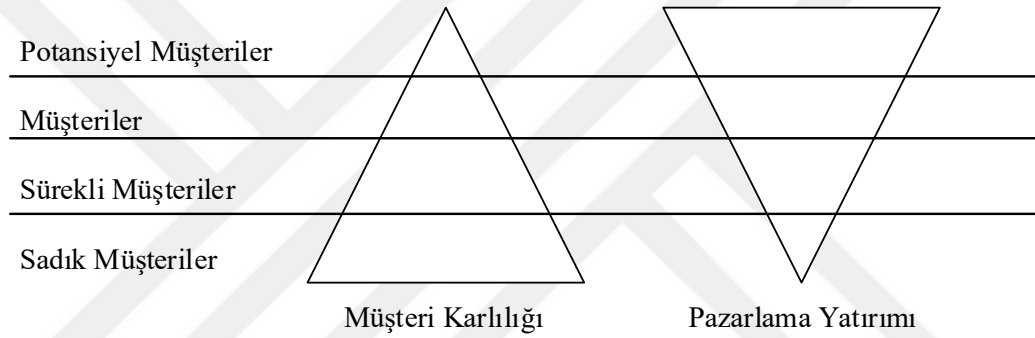
Meta arama pazarlaması satış noktalarında gerçekleştirilen tanıtım faaliyetlerini ve itibar yönetimini kapsamı nedeniyle otel işletmeleri açısından bir zorunluluktur. Bir otel işletmesinin çevrimiçi ortamlardaki varlığı müşterilerine sadece bilgi sağlamakla kalmayıp, rakipleri arasından sıyrılması ve tanınırlığının sağlaması açısından da önem arz etmektedir. Özellikle büyük bir otel markasının tanınırlığından faydalanamayacak olan bireysel otel işletmeleri için websitesi otel işletmesinin tanıtımını yapan yüzüdür. Dolayısıyla otel işletmelerinin websitelerini güncellemeleri ve geliştirmeleri müşterilerin dikkatini çekmek ve güncel bilgi sağlamak açısından önemlidir.

Arama motorları web sitesi içeriklerinin düzenli olarak geliştirilmesini talep etmektedir. Dolayısıyla otel işletmelerinde arama motoru optimizasyonuna arama motorlarının gerektirdiği standartların karşılanması ihtiyaç duyulmaktadır. E-posta pazarlaması ise otel işletmelerine potansiyel ve mevcut müşterilerine kolay bir şekilde ulaşabilmeleri konusunda yardımcı olmaktadır. Bu uygulamada önemli olan gönderilecek e-postaların bilgisayar, tablet ve mobil cihazlara uyumluluğunun sağlanmasıdır. E-posta pazarlaması ayrıca kişiselleştirilmiş dinamik içerikler oluşturma avantajına sahiptir. Web sitesini ziyaret eden kullanıcı verilerinden yararlanılarak kullanıcının o anda görüntülediği konuyla ilgili e-posta gönderimi yapılabilmekte, böylece e-posta gönderimleri kişiselleştirilebilmektedir (Lanz ve Carmichael, 2015: 2-6).

1.5. Müşteri Sadakati ve Müşteriyi Elde Tutma

1990'lı yılların başlarında yapılan araştırmalarda müşteri sadakatinin şirket karlılığının en temel gerekliliklerinden biri olduğu ortaya çıkmıştır (Reichheld vd., 2000: 134). Gerçekleştirilen çalışmalar genellikle yeni bir müşteri çekmenin müşteriyi elde tutma stratejilerine harcanan bütçeden daha maliyetli olduğunu göstermektedir (Lee-Kelley vd., 2003: 240). Çünkü sadık müşteriler daha fazla ve daha sık satın alım yapmaktadır. Sadık müşterilere hizmet sunmanın maliyeti daha düşüktür. Ayrıca işletmelerin sadık müşterileri elde tutma oranları daha yüksektir. Dolayısıyla sadık müşteriler yeni kazanılan müşterilere kıyasla daha karlıdır (Şekil 1.5) (Hughes, 2003: 250; Odabaşı, 2004: 58).

Araştırmalar müşterin ilişki kurdukları işletmelere her yıl daha fazla harcama yapma eğilimde olduklarını ortaya koymuştur. Müşteriler işletmeleri daha iyi tanıdığında ve rakip işletmelerden daha yüksek kalitede ürün/hizmet sağladığından emin olduğunda daha fazla satın alım yapma eğilimi göstermektedir. Bunun sonucunda daha uzun dönemli ilişkiler kurulmakta ve müşteriler ilerleyen zamanlarda daha özel taleplerde bulunmaya başlamaktadır. Böylelikle sadık müşterilerin karlılık oranları uzun dönemde diğer müşterilere kıyasla çok daha yüksek olmaktadır (Odabaşı, 2004: 117). Çevrimiçi ya da çevrimdışı olması farketmeksizin ağızdan ağıza iletişim oldukça kuvvetli bir pazarlama aracı olarak işletmenin olumlu ya da olumsuz tanıtımına neden olabilmektedir. Bu bağlamda sadık müşteriler aile ve arkadaşlarının yanı sıra işletmenin potansiyel müşterilerine işletmenin tanıtımını ve savunuculuğunu yapmaktadır (Kandampully vd., 2015: 381-382).



Şekil 1.7 Yatırım / Karlılık Zıtlığı

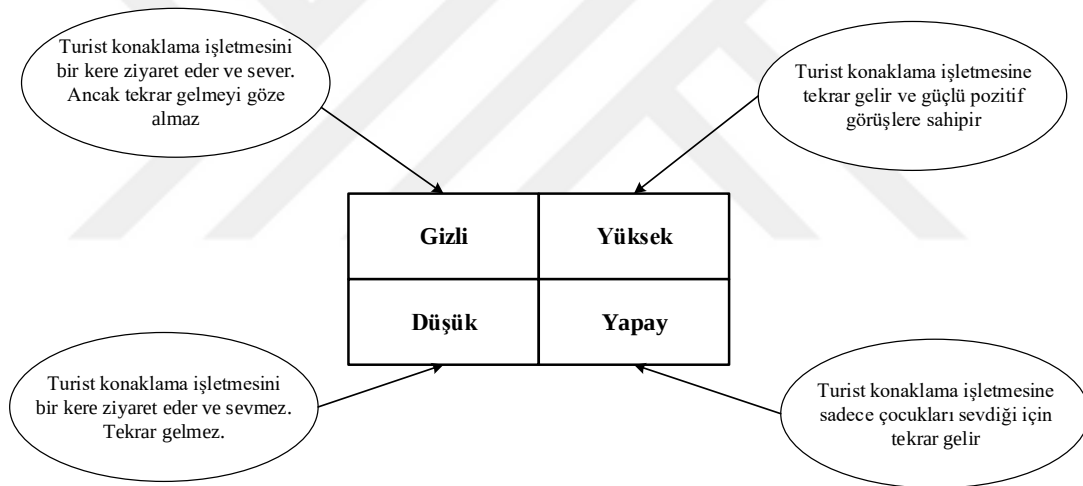
Kaynak: Odabaşı, 2004: 59

Müşteri sadakatine yönelik ilk araştırmaların daha çok ürün odaklı marka sadakatine yönelik olduğu ve bulguların genellikle memnuniyet ve kalite ile ürün sadakati arasında yüksek pozitif korelasyon gösterdiği görülmektedir. Ancak ürün sadakatine yönelik bulguların hizmet sadakatine genellenmesi mümkün olmayacağından dolayı, hizmet sektörüne odaklanmış çalışmalara ihtiyaç duyulmuştur. Kişilerarası ilişkilerin hizmet pazarlamasının en önemli unsuru olduğu ve risk algılarının hizmet sektöründe daha büyük etkileri olduğu göz önünde bulundurulduğunda müşteri sadakatinin hizmet sağlayıcıyı değiştirmenin önüne geçmede etkili olduğu ve sadakatin somut ürünlere kıyasla hizmet satın alan müşteriler arasında daha geçerli olduğu görülmektedir (Bloemer vd., 1998: 276-278). Nitekim hizmet endüstrisinde güven ve güvenilirlik gibi soyut unsurlar sadakatin oluşmasında ve sürdürülmesinde daha büyük bir rol oynamaktadır (Dick ve Basu, 1994: 108). Pazarlama uzmanları müşterilerin sadakat düzeyini arttıracak iki metod önermektedirler:

- (1) Başlarken doğru müşterileri işletmeye çekmek,
- (2) Müşteri oldukları zaman sunulan hizmette mükemmelleşmek.

Müşteriye sunulan ilgide mükemmelleşmek (excellent customer care) müşteri sadakatini geliştirme yolundaki en önemli unsurdur. Müşteri sadakatini kazanma yolunda müşteriye hizmet sunan personeller en önde gelen oyuncular. Müşteriye sunulan ilgi ve alakayı iyileştirmek için müşterilerle etkileşim halinde olan personelin bilgi ve yetkiyle güçlendirilmesi müşteri adına hareket edebilmeleri ve karar verebilmeleri adına önem arz etmektedir (Hughes, 2003: 251).

Otel işletmeleri müşterilerinin bağlılıklarını arttırmak için oldukça yüksek maliyetlere katlanmakta ancak müşterilerini hangi nedenlerden dolayı kaybettiklerini anlama yolunda yeterince çaba göstermemektedirler. Bir otel işletmesinin her sadık müşterisi için dışarda bekleyen birçok rakip otel mevcuttur. Müşteriler otel konaklamalarından memnun kalabilir ve tekrar gelme niyetlerini otel işletmesine ifade edebilirler. Ancak farklı fiyatlar ve özellikler sunan diğer otel seçenekleriyle karşılaştıklarında her zaman sadık olacaklar mıdır (Tanford vd., 2013: 806)?



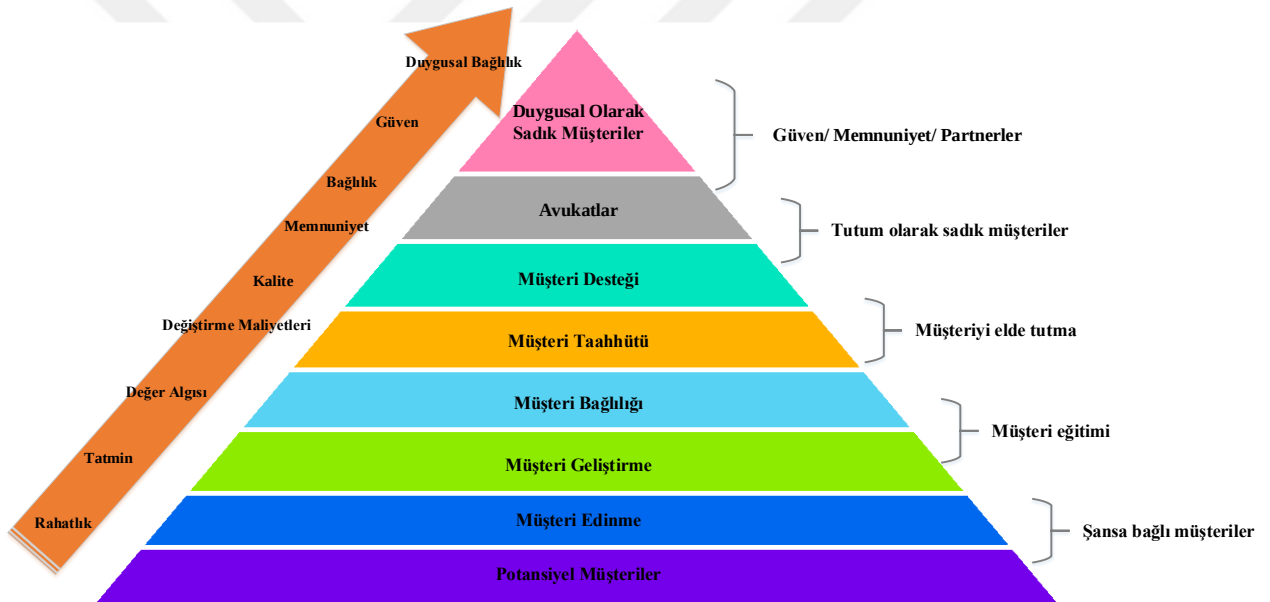
Şekil 1.8 Bağlılık Matrisi

Kaynak: Salihli Birdir, 2004: 91

Müşterilerin konaklama işletmelerine yönelik bağlılıkları bilgilerine, tutumlarına ve kullanım şekillerine bağlı olarak farklılık gösterebilmektedir (Şekil 1.6). Bu bağlamda müşterinin konaklama işletmesine yönelik bağlılığı yüksek ve düşük olabileceği gibi gizli ya da yapay bir özellik gösterebilmektedir (Salihli Birdir, 2015: 91). Otel işletmeleri hayatta kalmanın ve büyümenin anahtarının müşterilerine unutulmaz ve benzersiz deneyimler yaşatmakta saklı olduğunun bilincindedir. Dolayısıyla otel işletmelerinde müşterileri işletmeye bağlayacak kişiselleştirilmiş deneyimler tasarlamak aynı zamanda müşterilerin birer değer yaratıcısına ve marka elçisine dönüşmesine imkân sağlayarak sadakatlerini ve karlılığı arttırabilmektedir. Müşteriler güçlü yakınlık kurdukları markalara karşı daha yüksek sadakat

hissedebilmektedir. Bu nedenle hizmet işletmeleri müşteri edinmenin ötesinde müşterilerin süreçlere dahil edilmesine odaklanmaktadır.

Şekil 1.7’deki müşteri sadakati piramidinde müşteriler bir üst basamağa her geçtiğinde stratejik müşteri değeri artmaktadır. Müşterileri çekmek, dönüştürmek, dahil etmek ve bağlılıklarını sağlamak duygusal olarak sadık müşteriler yaratmanın ve marka avukatları oluşturmanın adımlarıdır. Duygusal olarak sadık müşteriler yaratmak için duygusal anlamda müşteri taahhüdü sağlamak şarttır. Bu bağlamda kişisel ilişkiler ve güven uzun dönemde müşterilerin sadakatini beslemektedir. Müşterinin hizmet sürecine dahil edilmesi durumunda müşteri sunulan hizmetlerin başarısını paylaşp hataları üstlenme eğiliminde olmakta ve işletmeye sosyal bağlılık geliştirmektedir. Marka elçileri ve avukatlar sadece kendileri için daha fazla satın alım yapmakla kalmamakta, ayrıca diğer müşterilerin davranışlarını ve satın alma niyetlerini olumlu ağızdan ağıza pazarlama aracılığıyla etkilemektedir.



Şekil 1.9 Müşteri Sadakati Piramidi

Kaynak: Kandampully vd., 2015: 381

Otel işletmelerinin müşterilerini değişen ihtiyaçlarını karşılayabilmek adına sürdürülen hizmet geliştirme süreçlerine dahil etmesi marka elçileri oluşturmada önem arz etmektedir. Örneğin Marriott otelleri “akıllıca seyahat” uygulamasıyla müşterilerinin gelecekteki seyahatlerini şekillendirmelerinde yardımlarını istemekte, modern seyahat deneyimini geliştirmek için tasarım, mutfak, teknoloji vb. konularda çığır açan fikirlerini paylaşmaya yönlendirmektedir (Kandampully vd., 2015: 383).

1.6. Otel İşletmelerinde Güncel Müşteri İlişkileri Yönetimi Uygulamaları

Pazarlama günümüzde akıllıca reklamlarla ve doğrudan e-posta gönderimleriyle çözülecek bir durum olmaktan çıkmıştır. Başarılı pazarlama uygulamaları müşterilerle etkileşimde olan herkesin katılımını gerektirmektedir. Pazarlama ve müşteri hizmetleri birbirinden ayrı düşünülemez (Hughes, 2003: 250-251). Bilgi teknolojilerindeki inovasyonlar ve rekabetteki artış nedeniyle hizmet sektöründeki pazarlama trendleri oldukça hızlı bir şekilde değişim göstermektedir. Pazarlama yaklaşımı ürün odaklı olmaktan çıkıp müşteri odaklı bir kavrama evrilmiştir. Özellikle müşterilerin misafir olarak görüldüğü hizmet endüstrisi başta olmak üzere, günümüzde müşteri odaklı stratejiler işletme felsefesinin başlıca bileşeni haline gelmiştir. İş performansı kavramı tamamen ağırlama felsefesi üzerine kuruludur. Otel işletmelerinin çabaları müşterilerin bireysel ihtiyaçlarını tatmin etmeye ve dolayısıyla kişisel ilişkiler geliştirmeye odaklanmıştır (Srivastava vd., 2019: 37).

Bilgi iletişim teknolojileri birebir deneyimlerin geliştirilmesinde ve M.İ.Y. stratejilerinin uygulanmasında katalizör bir rol oynadığından, konaklama endüstrisinde M.İ.Y. öncelikle bilgi iletişim teknoloji sağlayıcıları tarafından tanımlanmıştır. Çoğu otel işletmesi M.İ.Y. teknolojisinde yazılım firmalarının kendi yaklaşımlarına göre hareket etmelerine izin vermeleri ya da M.İ.Y. stratejilerini satın almış oldukları M.İ.Y. teknolojisine uyarlamaya çalışmaları nedeniyle etkili M.İ.Y. stratejileri uygulamada güçlük çekmektedir. Ancak M.İ.Y. bir teknoloji problemi olmaktan ziyade bir işletme problemidir ve bilgi iletişim teknolojileriyle işlevselliğin aynı potada hizalanmasını, tasarlanmasını ve işletme stratejisi ve operasyonlarıyla etkili bir şekilde koordine edilmesini gerektirir (Sigala, 2005: 392). M.İ.Y.'in başarılı bir şekilde uygulanması için müşteri bilgilerini elde etmeyi kolaylaştıracak etkin veri yönetim sistemlerine ihtiyaç duyulmaktadır. Böylece, elde edilen müşteri bilgileri müşterilerin bireysel ihtiyaçlarına yanıt verecek şekilde kişisel hizmet sunulmasında kullanılabilmekte ve müşteri memnuniyeti sağlanarak müşteriye elde tutma şansı elde edilebilmektedir. Dolayısıyla, teknoloji temelli M.İ.Y. uygulamaları özellikle otel işletmelerinin hayatta kalması ve iş hacmini büyütmesi adına zorunlu bir hal almıştır. M.İ.Y. stratejilerinin etkin bir şekilde uygulanabilmesi ve sürdürülebilmesi adına bilgi yönetimi ve veri madenciliği tekniklerinin kullanılması esastır (Srivastava vd., 2019: 37).

Küresel çapta rekabetin artışı sonucunda otel işletmelerinin faaliyetlerini devam ettirebilmeleri için müşterilerinin sürekli değişmekte olan tercihlerini ve yaşam tarzlarını anlamaları ve hizmetlerini bunlara uygun bir biçimde uyarlamaları şarttır. Otel işletmelerinde müşteriler kolay erişim, rahat odalar, hızlı giriş - çıkış işlemleri, nezaket ve uygun fiyatlandırma gibi konularda farklı önceliklere sahip olabilirler (Min vd., 2002: 274). Otel işletmelerinde

müşterileri yakından tanımak; nereden geldiklerini, ortalama ne kadar harcama yaptıklarını, ne zaman ve ne için harcamada bulunduklarını bilmek pazarlama stratejilerinin geliştirilmesinde ve karın artırılmasında işletmelere yol göstermektedir (Magnini vd., 2003: 94).

Konaklama endüstrisinde müşteri odaklı işletmeler otel yönetim sistemi (property management system - PMS), merkezi rezervasyon sistemi (central reservation system - CRS), misafir sadakat programı veri tabanlarından ve pos cihazlarından (point of sale - POS) veri madenciliği uygulamalarına fırsat tanıyacak ölçüde fazla miktarda veri toplamaktadır. Bununla birlikte internet ve web sitesi aracılığıyla oluşturulan veri miktarıyla başa çıkabilmek oldukça zorlaşmıştır. Dolayısıyla, konaklama sektörünün veri depolama ve veri madenciliği sistemlerine olan ihtiyacı her geçen gün daha fazla önem kazanmaktadır. Hizmet işletmelerinde veri madenciliği teknolojisi müşteri bilgilerini kişiselleştirilmiş hizmet sunumunda kullanılabilecek değerli çıktılara dönüştürmektedir. Böylelikle hizmet işletmeleri müşteri bilgilerini kişiselleştirilmiş ve dinamik pazarlama kararlarına dönüştürerek müşterileriyle uzun dönemli ilişkiler kurabilmekte ve uzun dönemli karlılığı sağlayabilmektedir. Veri madenciliği müşteri davranışları ve satın alım şekilleri üzerine yeni örüntülerin keşfedilmesini sağlayabilmektedir (Srivastava vd., 2019: 42).

Konaklama işletmeciliğinde M.İ.Y. uygulamalarının farklı pazarların ve müşterilerin eğilimleri doğrultusunda farklılaştırılması gerekmektedir. Bu bağlamda, müşteri verilerinin güncel ve detaylı olarak kayıt altına alınmasının önemli katkısı olacaktır. Benzer şekilde, Savoy Grubu ve Ritz-Carlton gibi otel zincirleri müşteri tercihlerinin kayıt altına alındığı kapsamlı bir veri tabanı oluşturmuş, bu bilgileri müşterilerine daha iyi hizmet sunmakta kullanmaktadırlar. Hilton otel zinciri, “Hilton Müşteri Veri Tabanı”ndan faydalanarak hedefledikleri müşteri grubuna hızlı bir şekilde ulaşabilmekte, departman yöneticileri sezonluk iş planlarını kolay bir şekilde yapabilmektedirler. Kongre ve konferans pazarına yoğunlaşan otel işletmeleri de müşteri veri tabanı kullanımı konusuna önem vermektedir. Böylece kongre katılımcılarına yönelik bilgilere ulaşabilmekte ve bu müşteri bilgilerini müşterilerine daha iyi hizmet sunmak için kullanabilmektedirler (Aksu ve Ehtiyar, 2007: 129).

Üst sınıf zincir otel gruplarından Ritz-Carlton, müşterilerinin otelde kalış süresince yaptıkları görüşmelerden elde ettikleri bilgileri kayıt altına almakta ve müşterinin bir sonraki konaklaması dünyanın herhangi bir yerindeki başka bir Ritz-Carlton’da olsa dahi, onlara kişiselleştirilmiş hizmet sunmakta kullanılmadığıdır. Hipoalerjenik yastıklar ya da fazladan havlu talebi gibi ricalar ileriki bir tarihteki konaklamada kullanılmak üzere kayıt altına alınmakta ve devamlı (repeat) müşteriler için kişiselleştirilmiş ürün ve hizmetler hazır edilmektedir (Chen ve Popovich, 2003: 676 - 677).

Otel işletmelerinde müşterilerle uzun dönemli ilişkiler geliştirmede veri tabanlarının rolü büyüktür. Kitapçı (2006: 159) tarafından gerçekleştirilen bir çalışmanın sonuçlarına göre Türkiye’de faaliyet gösteren 4 ve 5 yıldızlı otellerin %85’i müşterilerinin doğum günü ve yıl dönümü gibi özel günlerini veri tabanlarında saklamakta ve müşterilerine özel günlerinde kutlama mesajı göndermektedir.

1.6.1. Müşteri Veri Tabanları

Otel işletmeleri teknolojiye ve bilişim sistemlerindeki son gelişmelere adapte olmuş M.İ.Y. uygulamalarına veri yönetimi ve veri madenciliği süreçlerini entegre etmiştir (Srivastava vd., 2019: 38). Otel işletmeleri konaklama ve kayıt süreçlerinden müşterilere yönelik veri toplayabilmektedir. Bunun yanı sıra oteller müşterileriyle ilgili konaklama sıklığı ve harcama davranışı vb. daha detaylı ve değerli başka verilere de ulaşabilmektedirler. Bu veriler bir veri tabanı sisteminde bir araya getirilerek daha karlı müşterilerin belirlenmesi ve hedeflenmesinde kullanılacak şekilde dönüştürülebilir. Böylelikle otel işletmeleri müşterileriyle ilişki kurmaya başlayabilir (Gilbert vd., 1999: 25-26).

Konaklama endüstrisinde M.İ.Y. uygulamaları için bilgi ve iletişim teknolojilerinin önemi gereğinden fazla yüksek sesle dile getirilse de, bilgi iletişim teknolojilerinin başarılı bir M.İ.Y. uygulamasındaki desteği büyüktür (Sigala, 2005: 391). Veri depolama teknolojisi M.İ.Y.’i mümkün kılmaktadır. Çünkü bu teknoloji aracılığıyla müşteri verisi bir araya getirilip birbirleriyle ilişkilendirilerek müşteri davranışları hakkında daha derin bir anlayışa sahip olunması sağlanabilmektedir (Chen ve Popovich, 2003: 678).

Bir veri deposu (data warehouse) bir bilgi teknolojisi yönetim aracıdır. Bu araç aracılığıyla işletmenin karar vericileri tüm işletmeyi kapsayan verilerin toplandığı “müşteri veri adaları”na anında erişim sağlayabilmektedir. Bu veri depoları tüm veri tabanlarını ve insan kaynakları, satış, finans, depolama, satın alma ve pazarlama gibi operasyonel sistemleri birbirine bağlayabilmektedir. Veri depoları heterojen yapıdaki bu farklı birimlerden aldığı büyük miktardaki veriyi temizleyip işleyerek yönetilebilir hale getirir ve tüm müşteri etkileşimlerinin tarihsel kayıtlarını oluşturur (Chen ve Popovich, 2003: 677-678).

Otel işletmelerinde müşteri veri tabanlarının kullanımı müşterilerin tekrar gelmesi üzerinde etkili olmakla birlikte müşterilerin istek ve taleplerini önceden tahmin ederek müşteri sadakatini arttırmada yardımcı olmaktadır (Kitapçı, 2006: 159).

1.6.2. Sadakat Programları

Müşteri ilişkilerini yönetmenin şüphesiz en iyi yollarından biri sadakat programı uygulamaktır. Sadakat programlarının en basit uygulamaları müşterilerin yaptıkları ödemeler

sonucunda puan ya da mil kazanması ve bu birikimlerinin kendilerine ödül olarak dönmesi üzerine kurgulanmıştır. En başarılı sadakat programlarında statü seviyeleri bulunmaktadır. Müşteriler çeşitli statü seviyelerine ulaşabilmek için gerekli koşulları yerine getirmeye çalışmaktadırlar (Hughes, 2003: 220-221). Sadakat programlarının en iyi bilinen uygulamalarından biri seksenli yıllarda American Airlines tarafından sunulan sık uçan yolcu programıdır. American Airlines rakiplerine kıyasla kendi şirketleriyle uçmayı tercih eden sadık müşterilerinden daha fazla pay alabilmek amacıyla bu programı tasarlamıştır. Bu program otel işletmeleri tarafından da benimsenmiş ve uygulanmaya başlanmıştır. Günümüzde otel işletmeleri sadakat programlarından faydalanarak müşterilerine ücretsiz üst kategori odalarda konaklama, ücretsiz kahvaltı, her konaklamada bonus puan kazandırma vb. ödüller sunmakta ve marka imajlarını güçlendirerek müşterilerinin bağlılıklarını arttırmaya çalışmaktadır (Hantuli, 2016: 1) .

Otel işletmelerinde üyelik programlarında genellikle sadakat kartları kullanılmaktadır. Bu kartların iki fonksiyonu bulunmaktadır (Palmer vd., 2000: 55):

- (1) Müşterinin tekrar gelmesini sağlamak: Sadakat kartları müşterilere sunduğu teşvikler aracılığıyla müşterilerin otel işletmesine tekrar gelmesinde özendirici olmaktadır. Böylelikle müşteriler otel işletmesiyle bağ kurmakta ve sadık hale gelmektedirler.
- (2) Bilgi kaynağı olmak: Sadakat kartları aracılığıyla müşterinin konaklama ve kullanım şekilleri izlenebilmektedir. Böylelikle otel işletmeleri müşterinin konaklama sıklığı, ödediği oda ücretleri, satın alım yöntemi ve satın aldığı ekstra hizmetler hakkında veri toplayabilmektedir.

Örneğin, Marriott otel zincirinin Club Marquis sadakat programına üyelik için ücret almamaktadır. Üyelik avantajlarından faydalanabilmek için bir müşterinin beş farklı zamanda Marriott otellerinde konaklaması gerekmektedir. Sadakat programına üye olan müşteriler ise ücretsiz hizmet hattı üzerinden hızlı rezervasyon, geç kalınması halinde otomatik olarak rezervasyonun garantilenmesi, ön kayıt, lüks konaklama, çek ile ödeme ayrıcalığı, her sabah odaya ücretsiz Wall Street gazete servisi, hızlı check-out, newsletter aboneliği, kişiye özel bagaj etiketi ve üyelik kartı gibi avantajlardan yararlanabilmektedir. Club Marquis üyeliği tüm Marriott tesislerinde geçerli olmaktadır (Berry, 2002: 66).

Bütün otel işletmelerine uygun olabilecek tek bir sadakat programı örneği bulunmamaktadır. Farklı kategorilerde faaliyet gösteren otellerin kendi iç ve dış çevrelerine göre farklı sadakat programları tasarlaması gerekmektedir. Özellikle ekonomik sınıftaki otel işletmeleri üst kategori hizmet ve oda sunumu vb. teşvikleri içeren karmaşık sadakat programlarından en az fayda sağlayan işletmelerdir. Bu kategorideki otel işletmelerinin ürünleri

yüksek düzeyde standartlaştırılmış olduğundan gelişmiş hizmet sunumunda bulunmakta zorluk yaşamaktadırlar. Bu nedenle ekonomik sınıftaki otel işletmeleri daha çok oda fiyatlarında indirim ya da ücretsiz geceleme imkanı vb. ödüller sunmaktadır (Palmer, 2000: 58).

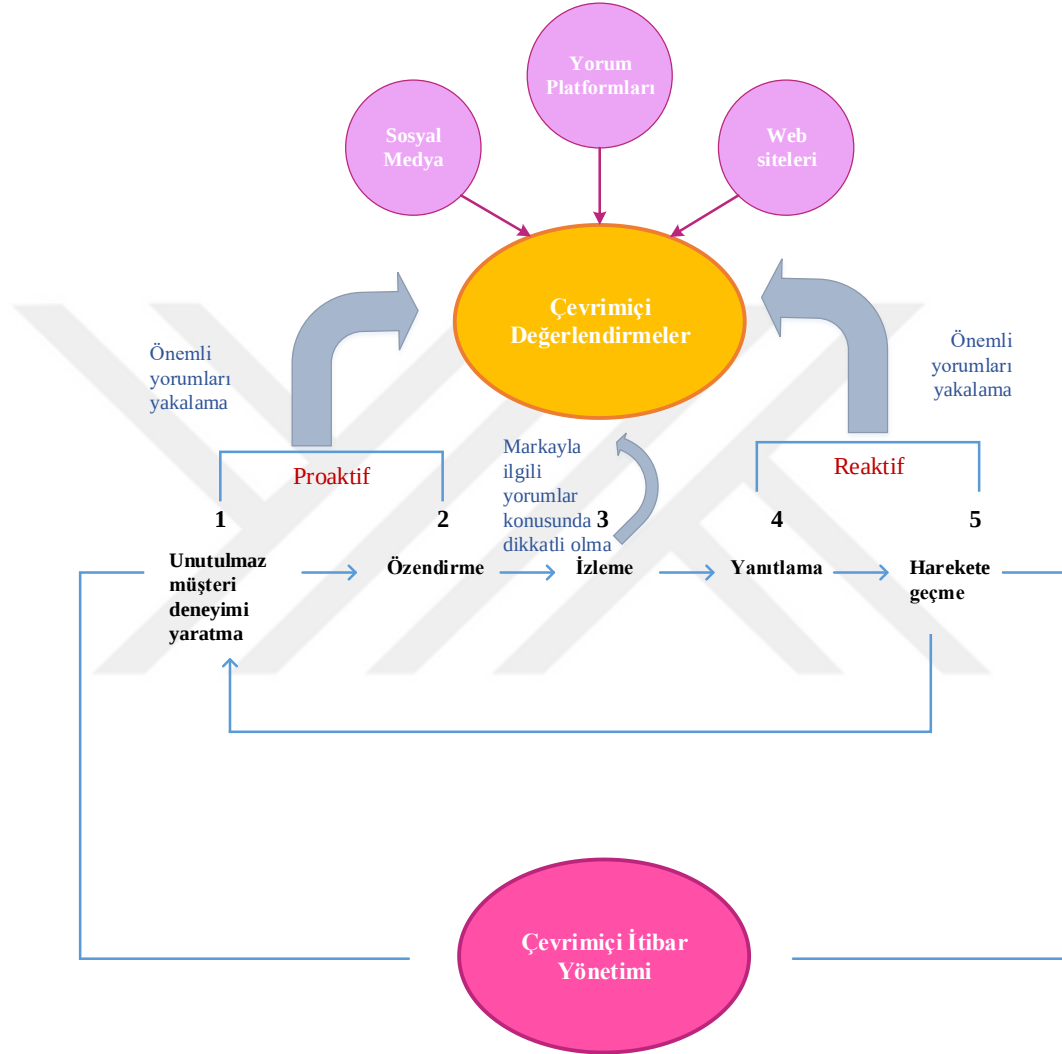
1.6.3. Çevrimiçi İtibar Yönetimi Sistemleri

Başarıyı hedefleyen otel işletmelerinin aktif bir şekilde sosyal medya üzerindeki itibarlarını yönetmeleri ve günümüz teknolojisine ayak uydurarak müşteri memnuniyeti yaratmaya çalışmaları gerekmektedir (O'Connor, 2010: 769). Sosyal medya aracılığıyla bireyler çeşitli kurumlar, markalar ve ürünler hakkında düşüncelerini paylaşabilmektedir. İşletme çalışanlardan, rakip firmalardan ve hedef müşteri kitlesinden herhangi bir bilgiyi saklamanın neredeyse imkânsız olduğunu kabul ederek işletme stratejilerini bu doğrultuda şekillendirmek gerekmektedir (Özgen, 2011: 99).

Günümüzde otel müşterileri işletmeye yönelik olumlu ve olumsuz fikirlerini internet üzerinden milyonlarla paylaşabilme imkânına sahiptir. Dolayısıyla, otel işletmelerinin müşteriler otelden ayrılmadan önce müşteri beklentilerini karşıladığından ve müşterilerini evlerine mutlu uğurladığından emin olmaları gerekir. Bu bağlamda, otel işletmelerinin proaktif bir şekilde TripAdvisor gibi sosyal iletişim sitelerindeki itibarlarını iyi bir şekilde izlemeleri ve yönetmeleri önem arz etmektedir (O'Connor, 2010: 769). Çünkü müşterilerin çevrimiçi platformlardaki değerlendirmeleri potansiyel müşteriler tarafından oldukça yararlı bir bilgi kaynağı olarak değerlendirilebilir ve turizm işletmelerinin itibarı üzerinde ciddi şekilde etkili olabilir (Horster ve Gottschalk, 2012: 229). Nitekim turistlerin %74'ünden fazlası seyahatlerini planlarken diğer turistlerin yorumlarını dikkate almaktadır. Dolayısıyla çevrimiçi yorumlar potansiyel otel müşterileri için oldukça önemli bir bilgi kaynağıdır (De Pelsmacker, 2018: 47).

Günümüzde otel müşterilerinin yorum ve değerlendirme yapabileceği çok sayıda platform bulunmakta (Tripadvisor, Holidaycheck, Trivago, Tophotels.ru, Booking.com vb.) bu platformlarda yapılan yorum ve değerlendirmeleri izlemek, yanıtlamak, gerekli önlemleri almak oldukça zorlaşmaktadır. Bu amaçla oluşturulan Reviewpro, BoomSonar ve Reputation.com gibi itibar yönetim sistemleri birçok farklı portaldaki yorumları tek ekrandan takip edebilme imkânı sunmaktadır. Bu sistemler aynı zamanda yorumların analizlerini yaparak acil önlem alınması gereken konuların tespit edilmesinde raporlar sunmakta ve olumsuz yorumlarda uyarı vererek otel işletmelerinin itibar yönetimlerinde destek sağlamaktadır (ReviewPro, erişim tarihi: 01.11.2019; Reputation.com, erişim tarihi: 01.11.2019; Ticktockboom, erişim tarihi: 01.11.2019).

Nguyen ve Coudounaris (2015: 165-166) otel işletmelerinde çevrimiçi itibar yönetimi için bir klavuz önermiştir (Şekil 1.10). Bu klavuza göre otel işletmelerinin itibarı öncelikle müşterilerin deneyimine bağlı olmaktadır. Bu nedenle otel işletmelerinin proaktif davranarak müşterilerine unutulmaz bir deneyim yaşatması müşteri duygularını etkileyerek olumlu değerlendirmelerde bulunmaları konusunda teşvik edici olacaktır.



Şekil 1.10 Otel İşletmelerinde Çevrimiçi İtibar Yönetimi

Kaynak: Nguyen ve Coudounaris, 2015: 166

Müşterilerin unutulmaz bir deneyim yaşayabilmesi ise otel işletmelerinin konaklama başladığı andan itibaren müşteriyi izleyerek yaşadığı sorunların ve olumsuzlukların çıkış yapmadan önce çözüldüğünden emin olması ve çalışanlar tarafından üstün hizmet performansı sağlanmasıyla mümkün olabilmektedir. Ayrıca otel işletmeleri proaktif davranarak otel hakkındaki düşüncelerini paylaşmaları için müşterilerini özendirmelidir. Çevrimiçi değerlendirme portallarındaki otel işletmesine yönelik yorum sayısı arttıkça rezervasyon yapma niyeti de artış göstermektedir. Çevrimiçi ortamlardaki yorumlar otel işletmesinin popülerliği ve

güvenilirliğinin bir göstergesidir. Bu amaçla otel işletmeleri müşterilerin kolaylıkla değerlendirmelerini yapabileceği basit bir sistem tasarlamalıdır. Aynı zamanda otel işletmeleri olumsuz değerlendirmelere karşı tetikte olmalı ve müşteri yorumlarını izlemelidir. Yorumların izlenmesi için önemli bulunan yorum portallarından günlük raporlamalar alınabilir. Bir sonraki aşama ise yorumların profesyonel bir şekilde kişiselleştirilerek yanıtlanmasıdır. Özellikle olumsuz yorumların doğru bir şekilde ele alınması, müşteriye değerlendirmesi için teşekkür edilmesi, müşteriden özür dilenmesi ve olumsuz durumla ilgili bir açıklama sunularak düzeltme konusunda alınacak önlemlerden bahsedilmesi önem arz etmektedir. Son aşama ise müşteri değerlendirmelerinden gelişim için gerekli konuları öğrenmek ve harekete geçmektir.

1.6.4. M.İ.Y. Yazılımları

M.İ.Y. yazılımları bir işletmenin müşteri temas noktalarında ön ofisle (satış, pazarlama, müşteri hizmetleri vb.) arka ofisi (finans, insan kaynakları vb.) birbirine bağlayan teknolojik uygulamalar olarak tanımlanabilmektedir (Chen ve Popovich, 2003: 672 - 673). Bu müşteri temas noktaları müşteri ile otel işletmesinde hizmet sunan personelin karşı kaşıya geldiği zamanları ifade etmektedir. SPA masajı ya da À la Carte restoran satışı, pazarlama bölümü tarafından yapılacak bir e-posta gönderimi ya da müşteri hizmet hattında rezervasyon yapmak için müsaitlik soran bir müşteriye sunulacak alternatifler bu temas noktalarına örnek gösterilebilir. Dolayısıyla M.İ.Y. yazılımı M.İ.Y. stratejilerinin uygulanmasını kolaylaştıran bir araç olarak da tanımlanabilir (Venturini ve Benito, 2005: 858).

Bilgi teknolojilerinin M.İ.Y.'e uyarlanması operasyonel ya da analitik formda olabilmektedir. Operasyonel M.İ.Y. satış, pazarlama ve hizmet süreçlerini otomasyonel hale getiren yazılımlar kullanmaktadır. Operasyonel M.İ.Y. yazılımları bu bağlamda satış otomasyonu (SFA – salesforce automation), kampanya yönetimi, etkinlik-bazlı pazarlama, ürün yapılandırma vb. çözümler sunmaktadır. Analitik M.İ.Y. ise SAS enstitüsü tarafından sunulan Enterprise Miner ve MicroStrategy tarafından sunulan iş zekası (Business Intelligence) gibi motorlar aracılığıyla işlev kazanmaktadır. Analitik M.İ.Y. yazılımları aşağıdaki soruların yanıtlarını bulmak amacıyla müşterilerle ilgili verileri analiz etmektedir (Ang ve Buttle, 2006: 5):

- Bu müşteriye bir sonraki teklifimiz ne olmalıdır?
- Bu müşterinin ayrılma olasılığı nedir?
- Kampanyalar için müşterileri nasıl bölümlendirebiliriz?

Otel teknolojilerine M.İ.Y. yazılımlarının entegre edilmesi sonucunda müşterilerin tercihleri ve beğenmedikleri konular tespit edilerek kişiselleştirilmiş bir konaklama deneyimi

yaşamaları sağlanabilmektedir (Sarmaniotis vd., 2013: 137). Böylece otel işletmeleri rakipleri arasında farklılaşarak pazar paylarını arttırabilirler. Bu amaçla otel işletmeleri müşterilerinin talep ve tercihlerini izleyen ve veri tabanlarında saklayan M.İ.Y. yazılımlarından faydalanmaktadır. Bir müşteri otele giriş yaptığında “üst katta King yataklı sigara içilmeyen bir oda” talep etmiş olduğu müşterinin tercihleri hakkında bilgi sağlarken aynı otel zincirinde kaçınıcı defa konakladığını ve harcamalarını da izleyebilmektedir. Müşterilere kişiselleştirilmiş deneyimler yaşatabilmek için çalışanlar tarafından elde edilen diğer özel detaylar da müşterinin profiline işlenerek gelecekteki konaklamalarında kullanılabilmektedir (Hantuli, 2016: 2).

Otellerde kullanılan M.İ.Y. yazılımları aracılığıyla müşterilerin bilgilerini veri tabanına işlemek ve gerektiğinde müşteriye kişisel hizmet sunabilmek için bu kayıtlı verileri kullanabilmek mümkün olmaktadır. Örneğin, Hyatt otellerinde konaklayan bir müşteri telefonuyla oda servisini aradığında otele henüz giriş yapmış olsa dahi telefona cevap verilirken kendisine adıyla hitap edildiğini görecektir. Çünkü günümüzde birçok otelde kullanıldığı gibi, Hyatt otellerinde de Caller ID otel içi telefon sistemi kullanılmaktadır. Bu sistem aracılığıyla Hyatt otelin müşteri veri tabanında oluşturduğu bilgiler çekilmekte aranılan departmanın ekranında müşterinin ad, soyad, oda numarası vb. bilgileri görüntülenebilmektedir. Bu uygulama aynı zamanda veri tabanlı pazarlamaya da örnek oluşturmaktadır. Müşterilerin evlerinden binlerce kilometre uzaktaki bir otelde tanınıyor ve mutluluğu için çaba harcanıyor olması kendilerini özel ve güvende hissettirmektedir (Hughes, 2003: 3).



Şekil 1.11 M.İ.Y. Yazılımı Başarı Faktörleri

Kaynak: Venturini ve Benito, 2015: 868

Padilla-Meléndez ve Garrido-Moreno (2014: 395) otel işletmelerinin M.İ.Y. yazılımlarına yatırım yaparken farklı çözümleri dikkatli bir şekilde değerlendirerek sadece

stratejik amaçları ve kaynaklarıyla uyumlu yazılımları tercih etmelerini önermiştir. M.İ.Y. yazılımı işletme performansını etkileyen bir kritik başarı faktörü olarak görülmektedir. Venturini ve Benito (2005: 870) M.İ.Y. yazılımının performans değerlendirmesinde yazılım başarı faktörlerini operasyonel fayda, müşteri yaşam döngüsüne yönelik fayda ve performans temelli fayda başlıkları altında ele almıştır. Dolayısıyla bir M.İ.Y. yazılımının işletmeye sağlayabileceği faydalar bu üç boyutun altında toplanmaktadır (Şekil 1.11).

1.6.5. Kurumsal Kaynak Planlama Sistemleri (K.K.P.)

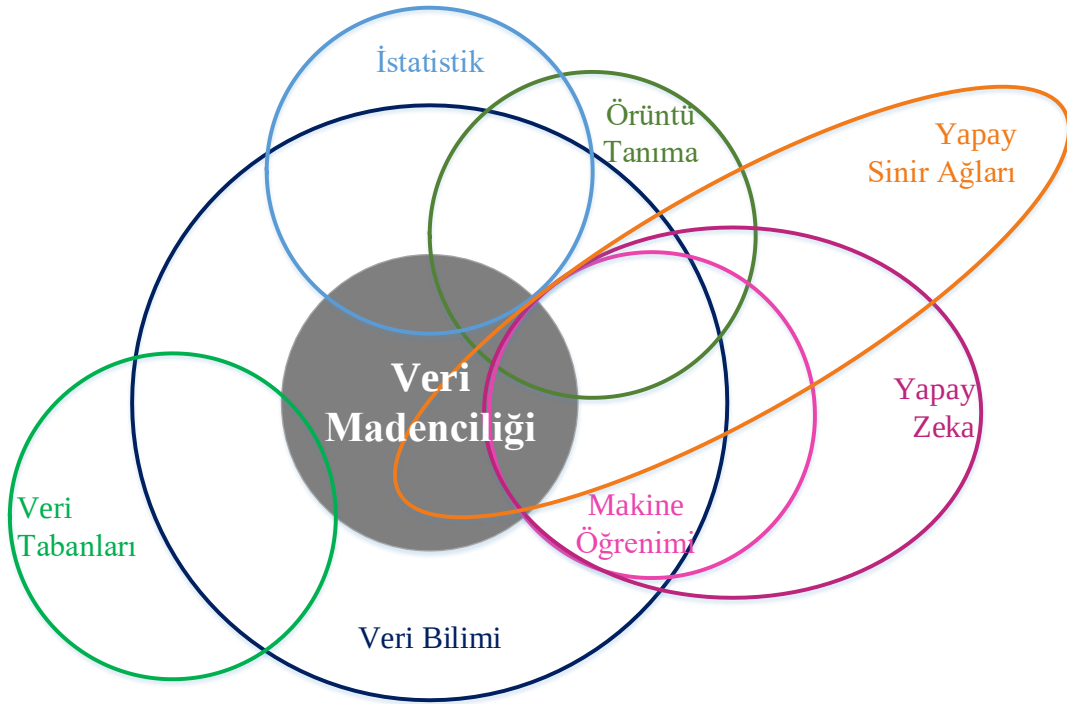
Oracle, SAP, PeopleSoft, Clarify, SAS ve Siebel gibi yazılım firmaları işletmelere M.İ.Y. uygulamaları geliştirme konusunda birbirleriyle yarışmaktadırlar. Bu firmaların çoğu aynı zamanda kurumsal kaynak planlama – K.K.P. (enterprise resource planning - ERP) sistemlerini de geliştirmektedir (Chen ve Popovich, 2003: 673). K.K.P. başarılı bir şekilde uygulandığında bir işletmenin sipariş yönetimi, üretim, insan kaynakları, finans, dağıtım vb. birimleri ile tedarikçilerini ve müşterilerini birbirine entegre bir sistemle bağlayarak veri paylaşımı sağlar (Chen ve Popovich, 2003: 678 - 679). K.K.P. sistemi M.İ.Y. uygulamalarının ön koşulu olmamasına rağmen müşterileri, tedarikçileri ve çalışanları web tabanlı erişimle M.İ.Y. sistemlerine bağlayan bir M.İ.Y. uygulaması ancak veri depoları ve/veya K.K.P. sistemleri varsa fayda sağlayabilmektedir (Solomon, 2000: 51). M.İ.Y. ve K.K.P.’nin uygulanması ve kritik başarı faktörleri benzerlik göstermektedir. İkisi de büyük ölçekli teknoloji yatırımları gerektirmektedir ve genellikle büyük yazılım firmalarından temin edilmektedir. İkisinin de başarılı olabilmesi için üst yönetimin desteğine, departmanlar arası işbirliğine, iletişime ve veri paylaşımına ihtiyaç duyulmaktadır. Farklı yönleri ise geleneksel K.K.P. finans ve insan kaynakları gibi arka ofis çalışanları tarafından kullanılırken M.İ.Y. teknolojilerinin satış ve pazarlama gibi ön ofis çalışanları tarafından kullanılmasıdır. Temel olarak M.İ.Y. gelecekte daha uygun hizmetler ve ürünler sunmak için müşterilerin istek ve tercihlerini öğrenmeyi ve müşterilerle etkileşimde bulunmayı hedeflerken K.K.P.’nin odak noktasında işletme içi rutin süreçlerin daha etkin hale getirilmesi bulunmaktadır (King ve Burgess, 2008: 422).

İKİNCİ BÖLÜM

KAYIP MÜŞTERİ TAHMİNLEME YÖNTEMLERİ

2.1. Makine Öğrenimi

İstatistik veriden öğrenme sanatıdır. Veriden öğrenme konusunda yaşanan zorluklar sonucunda istatistik biliminde köklü gelişmeler ortaya çıkmıştır. Veriden anlam çıkarma ve öğrenme sürecinde matematiksel hesaplamalar ve bilgisayar kodları hayati önem taşıdığından bu konudaki en büyük katkılar bilgisayar bilimciler tarafından sağlanmıştır (Hastie vd., 2009: xi). Makine öğrenimi veriden bilgi çıkarmayı ifade etmekte olup tahminleyici analizler ya da istatistiksel öğrenme olarak da bilinmektedir. Makine öğrenimi istatistik, yapay zeka ve bilgisayar bilimlerinin kesişiminde yer alan bir araştırma alanıdır (Müller ve Guido, 2016: 9). Diğer bir ifadeyle makine öğrenimi, yapay zekâ ve istatistik; veri madenciliği ve veri bilimini beslemektedir (Şekil 2.1). Ancak istatistik ve makine öğreniminin kesişimi bulunmamaktadır. Kimi problemlerde istatistikten kimi problemlerde ise makine öğreniminden yararlanılır. Bazı durumlarda veri ön işleme aşamasında istatistikten faydalanılırken model oluşturma aşamasında makine öğrenimi kullanılabilir. Ancak ikisi aynı anda, aynı işlem için kullanılamaz (Şeker ve Doğan, 2018: 66).



Şekil 2.1 Disiplinler Arası İlişkiler

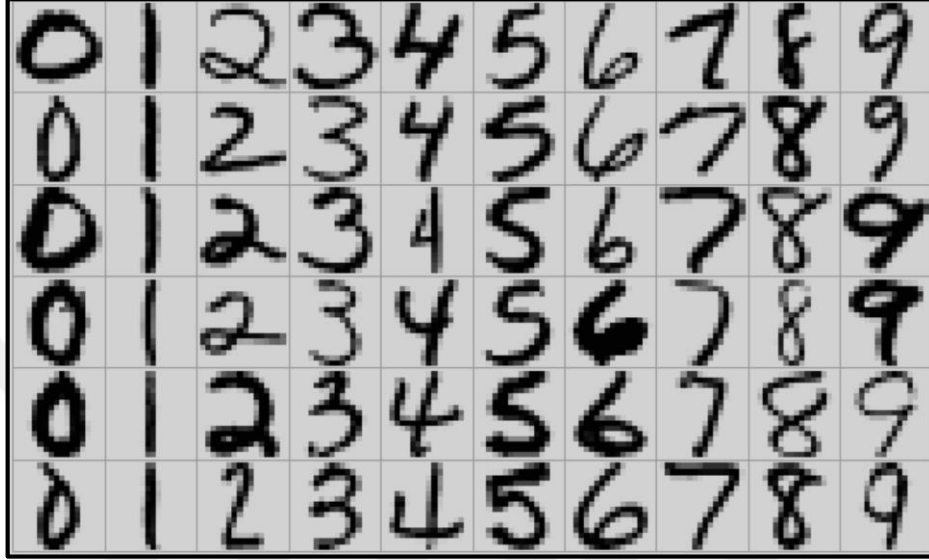
Kaynak: Şeker ve Doğan, 2018 : 65

Son yıllarda veri madenciliği, makine öğrenimi ve yapay zekâ uygulamaları gerek akademik araştırmalarda gerekse endüstriyel uygulamalarda yoğun olarak kullanılmaktadır. Günümüzde her saniye devasa ölçülerde veri üreten sosyal medya ve internet kullanımındaki yoğunluk bu sonuca yol açmıştır (Yang, 2019: 11). Hangi filmi seyredeceğimize, hangi yemeği sipariş edeceğimize ya da hangi ürünü satın alacağımıza yönelik otomatik öneriler, kişiselleştirilmiş çevrimiçi radyo, fotoğraflarımızdaki arkadaşlarımızı tanıma vb. teknikler kullanan çoğu modern websitesi ve uygulamada makine öğrenimi kullanılmaktadır. Facebook, Amazon ve Netflix gibi karmaşık websiteleri ziyaret edildiğinde göz gezdirilen hemen her noktada birden fazla makine öğrenimi modeli çalışmaktadır (Müller ve Guido, 2016: 9). Bunun yanı sıra akıllı telefonlarda bile bulunan yüz tanıma sesli komut vb. uygulamalar büyük ölçüde makine öğrenimi algoritmalarından faydalanmaktadır. Dolayısıyla yapay zekânın popülaritesi günden güne artmaktadır (Yang, 2019: 11). Ticari uygulamaların yanı sıra akademik araştırmalarda da makine öğreniminin çok büyük katkısı bulunmaktadır. Bu araştırma konuları arasında yıldızları anlamak, uzak gezegenleri bulmak, DNA zincirlerini analiz etmek ve kişiselleştirilmiş kanser tedavisi sağlamak gibi konular örnek verilebilir (Müller ve Guido, 2016: 9).

Makine öğrenimi (yapay zekâ), yeniçağın elektiriğine benzetilebilir. Elektrik ilk bulunduğunda yaşamlarımıza böylesine etki edebileceği kimsenin aklına gelmemekteydi. Makine öğreniminin de aynı potansiyele sahip olduğu görülmektedir. Çoğu insan makine öğrenimini endüstri devrimi 4.0 olarak görmekte, büyük oranda iş gücünü devre dışı bırakacağına inanmaktadır. Ancak tıpkı elektrik ve bilgisayarlı otomasyonda yaşanan gelişmelerin daha fazla iş alanı yaratarak verimliliği arttırması gibi makine öğrenimi de farklı iş kolları yaratma potansiyeline sahiptir. Çoğu insan makine öğreniminin sadece insansız araç kullanımı ve video oyunlarını insanlardan daha iyi oynamak vb. insan zekâsı seviyesinde karmaşık etkinliklere uygulanabileceğine inanmaktadır. Ancak makine öğrenimi sadece maddi olanakları yüksek kesimin değil ilgi duyan ve istekli olan herkesin yararına kullanılabilmektedir (Rebala vd., 2019: 3).

Veriden öğrenme problemleri genel olarak iki kategoride ele alınabilmektedir: denetimli öğrenme ve denetimsiz öğrenme. Denetimli öğrenmede amaç girdileri kullanarak çıktıları tahminleyebilmektir. Denetimsiz öğrenmede ise herhangi bir çıktı bulunmamakta olup amaç girdi olarak kullanılan ölçümler arasındaki ilişkileri ve desenleri betimlemektir. Modern makine öğrenimi dilinde “girdi” yordayıcı değişkenleri yani klasik istatistik diliyle bağımsız değişkenleri ifade etmektedir. “Çıktı” ise hedef değişkeni, diğer bir ifadeyle bağımlı değişkeni ifade etmektedir (Hastie vd., 2009: xi). En başarılı makine öğrenimi algoritmaları bilinen

örnekler üzerinden genelleme yaparak karar veren otomasyonlardır. Denetimli öğrenmenin alanına giren bu algoritmalarda kullanıcı algoritmaya girdileri ve karşılığında beklenen çıktıları sağlar. Algoritma ise verilen girdilerden istenen çıktıları elde etmeye yarayacak bir yol bulur. Dolayısıyla algoritma bir insanın yardımı olmadan daha önce hiç görmediği bir girdi üzerinden bir çıktı üretebilmektedir.



Şekil 2.2 Posta Zarfları Üzerindeki El Yazısı Örnekleri

Kaynak: Hastie vd., 2009:4

Örnek olarak makine öğrenimiyle bir spam e-posta sınıflandırması yapılmak istendiğinde kullanıcı algoritmaya çok sayıda e-posta (girdi) ve bu e-postaların spam olup olmadığı bilgisini (çıkıtı) sağlar. Yeni bir e-posta verildiğinde algoritma yeni e-postanın spam olup olmadığı konusunda bir tahminde bulunur. Bu şekilde girdi-çıkıtı örneklerinden öğrenen süreçlere denetimli öğrenme denmesinin nedeni bir “öğretmenin” denetiminde algoritmaya her bir örnek için çıktıları sağlayarak öğretmektir. Dolayısıyla bu algoritmalar daha kolay anlaşılır niteliktedir ve performanslarını ölçmek daha kolaydır. Denetimli öğrenme algoritmalarından aşağıdaki konularda faydalanılabilir (Müller ve Guido, 2016: 10-11; Hastie vd., 2009: 1-6):

- **Bir zarf üzerine el yazısıyla yazılmış posta kodunu tanıma:** El yazısının taranmış görüntüleri (Şekil 2.2) girdiyi, karşılığındaki rakam ise çıktıyı ifade edecektir.
- **Tıbbi görüntüleme üzerinden bir tümörün iyi huylu olup olmadığına karar verme:** Görüntüler girdiyi, tümörün iyi huylu olup olmadığı bilgisi ise çıktıyı ifade edecektir.
- **Kredi kartı işlemlerinde dolandırıcılık tespiti yapma:** Kredi kartı işlemlerinin dökümleri girdiyi, dolandırıcılık durumu ise çıktıyı ifade edecektir.

- **Kalp krizi geçiren bir hastanın tekrar kalp krizi geçirme ihtimalini tahminleme:** Demografik veriler, hastanın beslenme şekli ve klinik ölçümleri girdiyi, bu verilere göre kalp krizi geçirme durumu ise çıktıyı ifade edecektir.
- **Bir şirketin altı ay sonraki hisse bedelini tahminleme:** Şirketin performans ölçümleri ve ekonomik veriler girdiyi, bu verilere göre hisse bedelleri ise çıktıyı ifade edecektir.

Görüldüğü üzere bu örneklerde girdi ve çıktılar oldukça basit ve benzer görünse de veri toplama süreçleri tamamen farklılık gösterebilmektedir. Denetimsiz öğrenmede ise sadece girdi verisi bilinmekte çıktıyla ilgili bir veri bulunmamaktadır. Bu algoritmalar için birçok başarılı uygulama bulunmakla birlikte sonuçların anlaşılması ve algoritma performansının değerlendirilmesi güçtür. Denetimsiz öğrenme algoritmaları için aşağıdaki konular örnek gösterilebilir (Müller ve Guido, 2016:11-12):

- **Blog paylaşımlarındaki başlıkların belirlenmesi:** Büyük miktarda metin verisi bulunduğu durumlarda bu verinin özetlenmesi ve içindeki temaların tespit edilmesi istenebilir. Bu temaların ne olduğu ya da kaç tane tema bulunduğu başlangıçta bilinmiyor olabilir. Dolayısıyla bilinen bir çıktı yoktur.
- **Aynı tercihlere sahip müşteri segmentlerinin oluşturulması:** Bir müşteri kayıt verisi içerisinde hangi müşterilerin benzerlik gösterdiğini belirlemek ve benzer tercihleri olan müşteri grupları bulunup bulunmadığını anlamak istenebilir. Örneğin bir alışveriş sitesi için bu gruplar “ebeveynler”, “kitap kurtları” ya da “oyuncular” olabilir. Önceden bu grupların kimler olabileceği ya da kaç tane grup bulunabileceği bilinmediği için bilenen bir çıktı olmayacaktır.
- **Web sitesi erişimlerinde anormallik tespiti:** Kötüye kullanımı ya da yazılım hatalarını (bugs) tespit edebilmek için sıklıkla normalden farklı olan erişimleri bulmak faydalı olmaktadır. Her anormal erişim deseni birbirinden farklılık gösterebilmektedir. Dolayısıyla makine öğrenimi algoritmasını kullanacak kişinin elinde anormal davranışa örnek gösterilecek bir kayıt bulunmayabilir. Bu durumda kullanıcının sadece web trafiğini izleyebilmesi ve anormal ya da normal davranışa yönelik bir bilgisinin olmaması nedeniyle bu durum denetimsiz bir problem örneğidir.

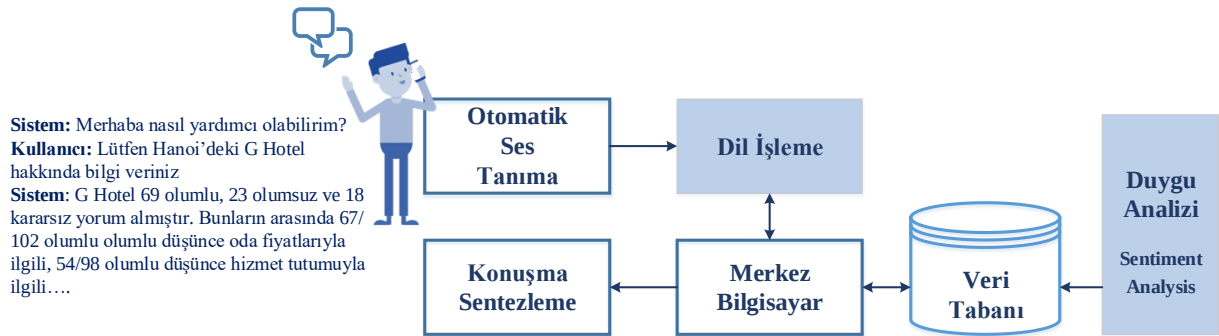
Makine öğrenimi algoritmalarında öğrenme ve test adı verilen iki aşama bulunur. Bu nedenle veri iki parçaya ayrılarak bir kısmının öğrenmede bir kısmının da öğrenilen modelin test edilmesinde kullanılması sağlanır. Veriden öğrenilen modelin daha önce görülmemiş bir yeni veri üzerinde test edilmesi algoritmanın performansının ölçülmesi açısından önem taşımaktadır (Şeker ve Doğan, 2018: 31).

2.2. Turizm ve Otel İşletmeciliğinde Makine Öğrenimi

Turizm ve otel işletmeciliği alanında gerçekleştirilen makine öğrenimi çalışmaları incelendiğinde oldukça büyük bir çoğunluğunun çevrimiçi yorum platformlarındaki yorumların analiz edilmesi konusunda yoğunlaştığı dikkat çekmektedir. Günümüzde otel müşterileri çevrimiçi müşteri yorum platformları üzerinden otellere yönelik olumlu ve olumsuz tecrübelerini ve diğer müşterilere tavsiyelerini paylaşmaktadır. Otel yorumlarına metin madenciliği uygulayarak müşteri beklentilerine yönelik daha derin bir anlayış kazanılabileceği gibi M.İ.Y. uygulamaları da bu yorumlar ışığında geliştirilebilir. Müşteri yorumlarının analiz edilmesi otel yöneticilerine hizmet kalitesini arttırmak, geliştirilmesi gereken konuları saptamak ve müşteri isteklerini anlamak için yol gösterici olmaktadır. Bu amaçla, otel müşterilerinin yorumlarının sınıflandırılmasında kullanılan duygu analizi (sentiment analysis) ve düşünce madenciliği (opinion mining) tekniklerinde (Zvarevashe ve Olugbara, 2018: 2; Raut ve Londhe, 2014: 556; García vd., 2012: 35; Shi ve Li, 2011: 950; Perikos vd., 2017: 338; Yu, 2016: 726; Akhtar, 2017: 563; Al-Smadi, 2018: 386; Markopoulos vd., 2015: 373; Ghorpade ve Ragha, 2012: 1; Zheng ve Ye, 20019: 335; Zhao vd., 2015: 23), faydalı yorumların tespit edilmesinde (O'Mahony ve Smyth, 2009: 305), yorumlar aracılığıyla otel hizmet kalitesinin değerlendirilmesinde (Duan vd., 2013: 3119; Wei vd., 2015: 72; Duan vd., 2016: 282; Phillips vd., 2015: 130), seyahat amacını tahminlemede ve Tripadvisor'da bulunan yorumların güvenilirliğini tespit etmede (Xiang vd., 2017: 625), anormal yorumları tespit edip silerek online yorum portallarındaki otel sıralamalarının güvenilirliğini arttırmada (Hsieh vd., 2014:1) ve online yorumlarda kullanılan fotoğrafların yorumun faydalı bulunması üzerindeki etkilerini değerlendirmede (Ma, 2018: 120) araştırmacılar tarafından makine öğrenimi algoritmaları kullanılmıştır.

Bunun yanı sıra Nilashi vd. (2017: 357) Tripadvisor kullanıcılarına ve Arruza vd. (2016: 1) Expedia kullanıcılarına yönelik geliştirdikleri tavsiye sistemleri (recommender system) için makine öğrenimi algoritmaları kullanmıştır. Bu sistemler diğer kullanıcıların işlemlerinden öğrenerek kullanıcının ilgi duyabileceği konuları tespit etmekte ve otel, restoran, görülecek yerler vb. önerilerde bulunmaktadır. Ayrıca Bjørkelund vd. (2012: 229) makine öğrenimi algoritmalarından faydalananak Tripadvisor yorumlarına duygu analizi ve düşünce madenciliği uygulamış ve ulaşılan sonuçları Google Maps üzerinde görüntüleyerek kullanıcıların konaklayabilecekleri en iyi otelleri kolayca tespit edebilmelerini sağlamıştır. Tran ve Phan (2015: 115) ise makine öğrenimi algoritmaları kullandıkları çalışmalarında telefon aracılığıyla otel yorumlarını sorgulama sistemi (SentiVoice) oluşturmuştur (Şekil 2.3). Bu sistem telefon ile otel bilgisi almak isteyen kullanıcıların taleplerini ses tanıma yöntemiyle otomatik olarak

tespit edebilmektedir. Ayrıca otel yorumlarına uygulanan duygu analizi ve düşünce madenciliği sonucunda elde edilen yorumları kullanıcıya sesli yanıtla bildirmektedir.



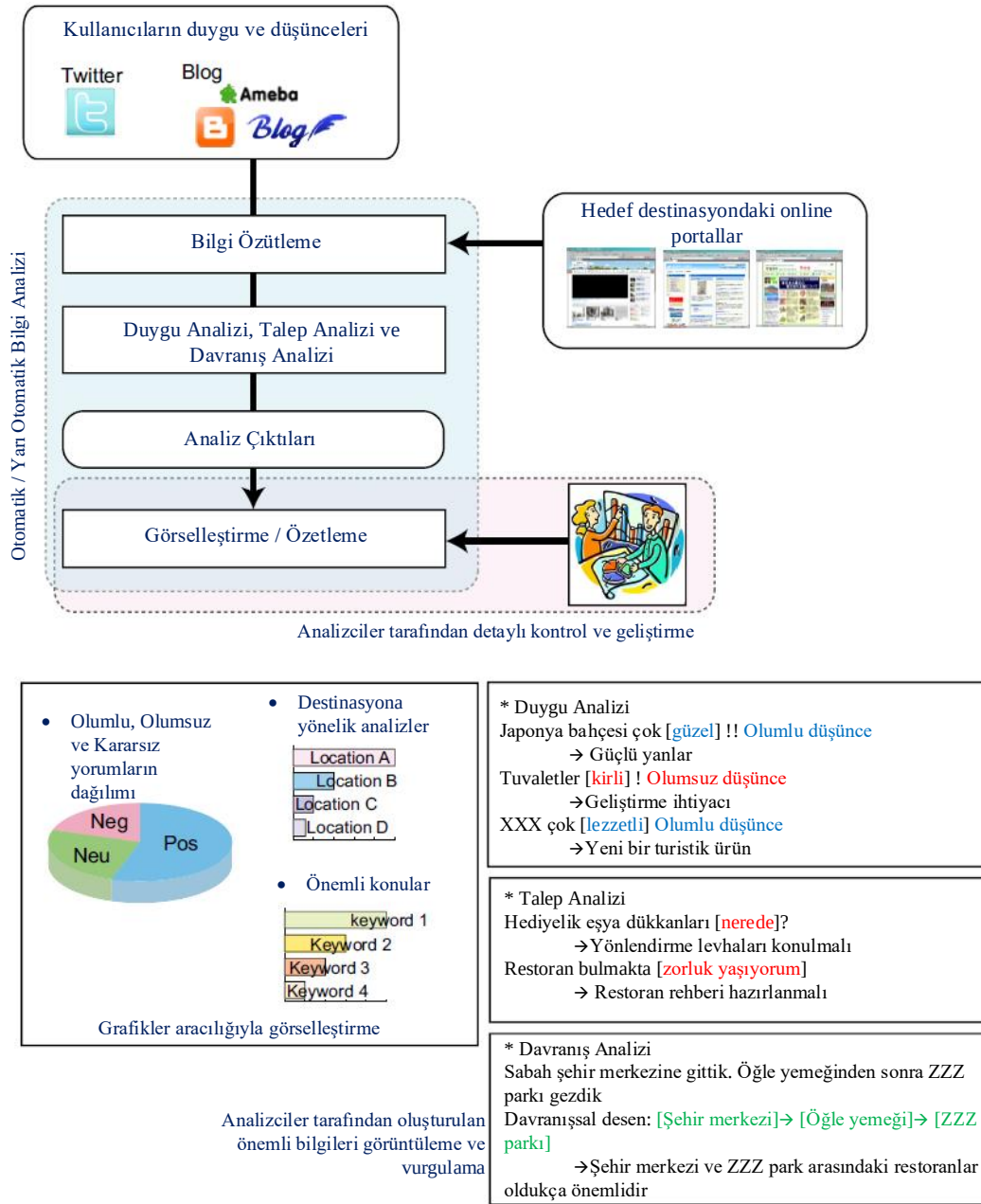
Şekil 2.3 SentiVoice Sistemi Çalışma Mantığı

Kaynak: Tran ve Phan, 2015: 115-116

Corazza vd. (2014: 45) tarafından denetimli makine öğrenimi algoritmalarından faydalanılarak bir çevrimiçi otel rezervasyon sistemi oluşturulmuştur. Çalışmada algoritmanın öğrenmesinde daha önceki müşterilerin rezervasyon bilgileri kullanılmıştır. Oluşturulan sistem müşterinin talep ettiği oda tipi müsaitse talep edilen hizmet için onay yanıtı göndermekte, müsait olamayan talepler için ise daha önceki ya da sonraki tarih alternatiflerini önermektedir. Müşteri yorumlarına duygu analizi uygulayan bir diğer çalışma da Shimada vd. (2011: 61) tarafından gerçekleştirilmiştir. Makine öğrenimi algoritmalarının kullanıldığı çalışmada Twitter üzerindeki yorumlar analiz edilerek turistik bir şehir için bir turizm bilgi sistemi oluşturulmuştur (Şekil 2.4). Bu bilgi sistemi aracılığıyla destinasyon yöneticileri güçlü ve zayıf yönlerini hızlı bir şekilde analiz edebilmekte ve turistlerin taleplerine yönelik geliştirme çalışmalarında bulunabilmektedir.

Makine öğrenimi algoritmaları turizm ve otelcilik alanında sadece otel yorumlarının analiz edilmesiyle sınırlı kalmayıp farklı araştırmalarda da kullanılmaktadır. Örneğin Kim (2011: 441) tarafından gerçekleştirilen bir çalışmada otellerin gelecekte iflas edip etmeyeceğine yönelik tahminleme çalışmasında makine öğrenimi tekniklerinden faydalanmıştır. Bu çalışmada girdi olarak otellerin finansal verileri kullanılmıştır. Ayrıca pazar profilini oluşturmada (Emel vd., 2007: 334), turizm talebi modellemesi ve tahminlemesinde (Cankurt ve Subaşı, 2015: 42; Claveria vd., 2016: 428; Kamel vd., 2008:1; Cankurt ve Subaşı, 2016: 3388; Akın, 2015: 64; Chen ve Wang, 2007: 215), otel talebi tahminlemesinde (Yang, 2014: 433; Candela Garza, 2017: 3), android işlemcili otel arama uygulamalarında (Lawrence ve Sankaranarayanan, 2012: 10), otel konumunun değerlendirilmesinde (Yang vd., 2015: 14) ve gelir yönetimi uygulamalarında (Enache, 2019: 1) makine öğrenimi algoritmalarından

faaydalanılmıřtır. Bunların haricinde birok alıřmada makine ğrenimi kullanılarak otel rezervasyon iptalleri tahminlenmiřtir (Antonio vd., 2017: 25; Antonio vd., 2019: 1; Morales ve Wang, 2010: 1). Ayrıca evrimii rezervasyon web sitesi müşterilerine (Han, 2018:1), sosyal klüp üyeleri (Maier ve Prusty, 2016:803) ve ticari otel müşterilerine (řirketler) (De Wit, 2017: 2) denetimli makine ğrenimi algoritmaları aracılıėıyla kayıp müşteri tahminlemesi uygulanmıřtır.



Şekil 2.4 Destinasyon Yönetim Sistemi alıřma Mantıėı

Kaynak: Shimada vd., 2011: 62

2.3. Kayıp Müşteri Kavramı

Günümüzde işletmeler sahip oldukları müşterilerin en değerli varlıkları olduğunun bilincindedir. Ancak, müşterilerin tercih edebileceği çok sayıda işletmenin bulunması, yeni müşteriler kazanmayı daha zor ve maliyetli bir hale getirmektedir (Van den Poel ve Larivière, 2004: 196-197). Bu nedenle işletmeler müşteri memnuniyetini sağlamanın yanısıra onları elde tutabilmenin yollarını da aramaktadırlar (Aksu ve Ehtiyar, 2007: 145). İşletmeler potansiyel risk taşıyan yeni müşteriler aramaktansa sahip oldukları müşterilerin ihtiyaçlarına odaklanarak uzun dönemli ilişkiler geliştirmeyi deneyebilirler. Uzun dönemli ilişkiler geliştirilen müşteriler daha fazla satın alım yapmakta ve eğer memnun kalırlarsa işletmenin olumlu yönde tanıtımını yapmaya istekli olmaktadır. Ayrıca uzun dönemde bu tür sadık müşterilere hizmet sağlamanın maliyeti daha düşük olmaktadır. Çünkü sadık müşteriler rakiplerin pazarlama faaliyetlerine daha az duyarlı olma eğilimindedirler (Van den Poel ve Larivière, 2004: 196-197).

Rekabetin yüksek olduğu dinamik bir pazar ortamında müşteriler aldıkları hizmetten memnun kalmadıklarında, rakip firmaların uyguladığı stratejilere ilgi duyduklarında, yeni ürün/hizmetler deneyimlemek istediklerinde vb. diğer nedenlerle satın alımlarını farklı bir işletmeden yapmayı tercih edebilirler (Vafeiadis vd., 2015: 1). Müşterinin bir hizmet sağlayıcıdan diğerine geçmesi olarak da ifade edilebilen bu durum literatürde kayıp müşteri (churner) olarak adlandırılmaktadır (Hung vd., 2006: 515). Kimi müşteriler işletmeyle bağlarını keserek işletmeden tamamen ayrılma davranışı gösterirlerken, kimileri satın alımlarını birden fazla rakip firma arasında bölüştürebilirler. Bu durum işletmelerin faaliyet gösterdiği sektörler göre farklılık gösterebilir. Özellikle sözleşmelerin söz konusu olmadığı durumlarda (perakende sektörü gibi) ayrılma davranışı daha çok kısmi kayıp (partial churn) şeklinde olmaktadır (Buckinx ve Van den Poel, 2005: 253; Miguéis vd., 2012a: 11250). Perakende sektöründe müşteriler farklı mağzalardan alışveriş yapabilir, aşamalı olarak işletmeyle ilişkilerini azaltabilir ve sonunda tam bir terk ediş söz konusu olabilir (Buckinx and Van den Poel, 2005: 255; Burez and Van den Poel, 2009: 4631). Bankacılık ve sigortacılık sektöründe müşterinin ayrılma davranışı hesap kapatma (Kumar ve Ravi, 2008: 9; Larivière ve Van den Poel, 2005: 475; Morik ve Kpcke, 2004: 325), iletişim sektöründe ise operatör değiştirme (Hwang vd., 2004: 181; Hung vd., 2006: 517), kontratın yenilenmemesi ya da iptal edilmesi (Tamaddoni Jahromi vd., 2010: 590) olarak ortaya çıkabilir.

Sözleşme ya da üyelik sistemiyle çalışan işletmelerin müşterileri kimi zaman memnun olmasalar da değiştirme maliyetleri nedeniyle aynı işletmeyle çalışmaya devam edebilirler. Bu değiştirme maliyetleri arasında toplanılan üyelik puanları ve üyelik avantajları sayılabilir. Üye

müşteriler bu haklarının yok olmasını istemedikleri için memnun olmadıkları işletmelerden ayrılma davranışında bulunmayabilirler (Kim ve Moon, 2012: 11720). Nitekim müşteri ayrılma davranışının bağlayıcı bir sözleşme ya da üyelik sisteminin bulunması ve bulunmaması durumunda farklılık gösterdiği görülmektedir. Sözleşme ya da üyelik sisteminin olduğu spor klübü üyeliği ya da cep telefonu operatörü hizmetleri gibi ilişkiler söz konusu olduğunda müşteriler hizmeti kullanmak için bir ödemede bulunmaktadır. Sözleşme bitiminde, sözleşmenin iptal edilmesi durumunda ya da üyelik yenilenmediğinde işletmeler müşterinin ayrıldığının farkına varmaktadır. Ancak ilişkiler bir sözleşmeye bağlı değilse müşteriler işletmeye ayrıldıklarını çok nadiren bildirmektedir (Suh ve Alhaery, 2016: 68; Zorn vd., 2008: 1). Çünkü sözleşme kullanılmayan ilişkilerde müşterilerin ayrılma niyetlerine yönelik işletmeyi bilgilendirme zorunluluğu yoktur ve genellikle müşterinin değiştirme maliyeti çok düşüktür (Buckinx ve Van den Poel, 2005: 253; Miguéis vd., 2012a: 11250). Böyle durumlarda işletmeler müşterinin ne zaman ayrıldığını gözlemleyememekte (Zorn vd., 2008: 1), kayıp müşterileri tanımlamakta ve bu davranışa etki eden değişkenleri belirlemekte zorluk çekmektedirler (Tamaddoni Jahromi vd., 2010: 590; Buckinx ve Van den Poel, 2005: 253). Bu şekilde çalışan sektörlerde kayıp müşteri durumu yaygın, ancak farkedilmesi güç bir “hayalet” olgudur (Tamaddoni Jahromi vd., 2010: 588). Bu nedenle işletmeler genellikle aktif olan - aktif olmayan (active -inactive) müşteri şeklinde bir tanımlamaya gitmeyi tercih etmektedir (Zorn vd., 2008: 1).

İşletmelerden ayrılan müşteriler iki ana başlık altında ele alınabilir: gönüllü kayıplar ve zorunlu (gönülsüz) kayıplar. Zorunlu kayıp, müşterilerin kendi tercihlerinden bağımsız olarak ortaya çıkan ve genellikle çevresel koşullardan etkilenen bir durum olarak tanımlanabilir. Müşterinin farklı bir ülkeye taşınması, sağlık durumu, yaşlılık, ölüm (Şeker, 2016: 27) ya da işletme tarafından ilişkinin sonlandırılması durumunda zorunlu kayıp ortaya çıkar. Zorunlu kayıp davranışında bulunan müşterileri tanımlamak kolaydır. Çünkü bu müşteriler genellikle aldıkları hizmeti kötüye kullanmaları, ödeme yapmamaları vb. nedenlerle işletme tarafından ilişkileri bitirilmiş olan ve tanımlanmış müşterilerdir. Gönüllü kayıpları tanımlamak daha güçtür. Bu tür ayrılma davranışı müşterilerin bilinçli bir şekilde işletmeden aldığı hizmetleri sonlandırması durumunda ortaya çıkar. Gönüllü kayıp davranışı iki alt kategoriye ayrılabilir; bunlar tesadüfi (incidental) ve kasıtlı (deliberate) ayrılmadır. Tesadüfi kayıp müşteri davranışı müşterinin sunulan hizmeti almasını engelleyen durumlar oluştuğunda ortaya çıkmaktadır. Örneğin; müşterinin finansal durumundaki değişiklikler nedeniyle satın alma gücünü yitirmesi gibi. Kasıtlı kayıp müşteri davranışı ise çoğu yöneticinin çözüm bulmaya çalıştığı ve tesadüfi kayıp müşteri davranışına kıyasla daha sık rastlanan bir durumdur. Bu türdeki kayıp müşteri

davranışında müşterinin bilinçli şekilde rakip bir işletmeyle çalışmaya karar vermesi durumu söz konusudur. Kasıtlı kayıp müşteri davranışına sebep olan nedenler arasında rakip işletmenin mevcut işletmede bulunmayan teknolojik olanaklara sahip olduğunun farkedilmesi, mevcut işletmede kalite ve hizmet hatalarıyla karşılaşılması ya da rakip işletmenin aynı ürün/ hizmeti daha uygun bir fiyata sunması gibi durumlar sayılabilir (Kim ve Yoon, 2004: 757; Hadden vd., 2006: 524).

2.4. Kayıp Müşteri Tahminleme

Müşterileri elde tutmak M.İ.Y.'nin başlıca amaçları arasında kabul edilir. Bu bağlamda hangi müşterilerin şirketle ilişkilerini sürdürüp hangilerinin ayrılacağına yönelik doğru bir tahminleme, başka bir deyişle kayıp müşteri tahminlemesi, etkili elde tutma kampanyaları düzenlemeyi arzu eden şirketler için büyük önem taşımaktadır. Şirketlerin kayıp müşterilere yönelik endişeleri elde tutulan müşterilerden sağlanan faydalardan ileri gelmektedir (Miguéis vd., 2012b: 338). Nitekim kayıp müşteri nedeniyle düşen satışlar sonucunda, işletmeler fırsat maliyetiyle başa çıkmak durumunda kalmakta ve yeni müşteri çekme çabalarına başvurmaktadırlar (Van den Poel ve Larivière, 2004: 198). Yeni müşterileri elde etmenin maliyetleri göz önünde bulundurulduğunda, sadık müşterileri elde tutmak ve memnun etmek daha makul bir çaba olacaktır. Müşteri kayıp oranlarını azaltmak için, işletmelerin müşteri ilişkileri yönetimine ağırlık vermesi ve proaktif davranması gerekmektedir (Zorn vd., 2008: 1). Bir işletmenin sadık müşterilerinin kayıp oranını (churn ratio) azaltması, diğer bir deyişle müşteri elde tutma oranlarında bir artış sağlaması, karlılık üzerinde etkili olacaktır (Zorn vd., 2008: 1; Van den Poel ve Larivière, 2004: 198; Verbeke vd., 2011: 2354; Ballings ve Van den Poel, 2012: 13517).

İşletmelerin kayıp müşterilerle nasıl başa çıkabileceklerini öğrenmeleri şarttır (Hung vd., 2006: 515). Bir işletme ayrılan müşterilerini belirleyebilse dahi, artık bu müşterilerini kaybetmiştir. Ayrılma potansiyeli olan müşterileri tahminlemek bu kaybın önüne geçmede işletmeye bir şans verecektir (Zorn vd., 2008: 1). Nitekim işletmeler müşterilerin satın alma tercihlerine yönelik kapsamlı verileri analiz etmede ve kayıp müşterileri tahminlemede veri madenciliği teknolojilerinden yararlanmaktadır (Wang vd., 2009: 8071). Kayıp müşteri tahminleme modellerinin amacı erken ayrılma sinyallerini tanımlamak ve gönüllü bir şekilde işletmeden ayrılma ihtimali yüksek olan müşterileri belirlemektir (Vafeiadis vd., 2015: 1; Verbeke vd., 2011: 2354; Miguéis vd., 2012a: 11250). Tahminin kesinliği, anlaşılabilirliği ve doğruluğunun ispatlanabilirliği bir kayıp müşteri tahminleme modelinin üç ana usurunu oluşturur. Tahminin kesinliği, elde tutmaya yönelik bir pazarlama kampanyasında gelecekte

işletmeyle ilişkilerini kesecek müşterilerin doğru bir şekilde hedeflenmesini sağlarken, anlaşılabilir karar setleri kayıp müşterilerin oluşmasında etkili olan asıl unsurların tespit edilmesinde ve ulaşılan bilgiyle uyumlu etkili elde tutma stratejilerinin geliştirilmesinde yol gösterici olacaktır (Verbeke vd., 2011: 2354).

İşletmeye kar sağlayan müşterileri elde tutmak için işletmenin gerçekleştirdiği faaliyetler kayıp müşteri yönetimi olarak ifade edilebilir. Kayıp müşteri yönetimi bir işletme için en önemli müşterilerin garantiye alınması sürecidir (Hung vd., 2006: 515). Müşterilerini elde tutmak isteyen bir işletmenin hangi müşterilerinin ayrılma riski olduğunu tahmin edebilmesi ve müşteri elde tutma çabalarını bu müşterilerine yönlendirmesi gerekmektedir (Wei ve Chiu, 2012: 103). Bu bağlamda, analitik çerçeveden bakıldığında, müşteri ilişkileri yönetimi altında yürütülen kayıp müşteri yönetimi faaliyetlerinin iki temel amacı vardır. Bunlar, ayrılacak olan müşterileri tahminlemek ve bu müşterileri elde tutmada hangi eylemlerin daha etkili olduğunu belirlemektir (Hung vd., 2006: 515; Ballings ve Van den Poel, 2012: 13517).

İşletmelerin elde ettikleri karın yüzde seksenini mevcut müşterilerinin yüzde yirmisinden elde ettiği (80/20 kuralı) bilinmektedir (Wang vd., 2009: 8071). Dolayısıyla işletmeler müşterilerini karlılık durumuna göre bölümlendirmekte ve elde tutma stratejilerini sadece bu kar sağlayan müşteri grubuna uygulamaktadır. Bununla birlikte, tüm müşteri tabanına ayrılma ihtimaline göre bir değer (skor) verilebilir ve kar getiren ve ayrılma olasılığı yüksek olan müşteri grubuna yönelik elde tutma stratejileri uygulanabilir (Hung vd., 2006 : 516). Örneğin, Van den Poel ve Larivière (2004: 215) kayıp müşteri tahminlemesinde tüm hesaplarını kapatan müşterilerin kayıp müşteri olarak belirlenebileceği gibi, tek bir hesapta az miktarda parası bulunan “uyuyan” müşterilerin de bu kategoriye alınabileceğini ya da tam bir ayrılmadansa kısmi olarak ayrılmış müşterilerin hedef alınabileceğini önermişlerdir. Bununla birlikte tüm ürünler yerine sadece yüksek gelir getiren ürünler için de kayıp müşteri tahminlemesi yapılabilir.

Kayıp müşteri tahminlemesinde hesaplanan ayrılma olasılığı bir müşterinin kesin olarak işletmeyle ilişkisini kesme arzusunu ya da istekliliğini ifade etmeyebilir. Hesaplanan oran müşterinin geçmişte ayrılan müşterilerle demografik özellikleri ya da hizmet kullanım davranışı gibi değişkenler açısından ne kadar benzeştiğinin bir göstergesidir. Örneğin, yakın zamanda sözleşme süresi dolacak olan müşteriler kendilerini bağlayıcı herhangi bir sözleşme olmadığından geçmişte ayrılma davranışı gösteren çoğu müşteri gibi potansiyel kayıp müşteri olarak görülür. Ancak, bu müşterilerinin birçoğunun doğru pazarlama kanalı aracılığıyla doğru zamanda uygun promosyonlarla karşılaşması halinde sözleşmelerini yenilemeleri ve işletmeyle olan ilişkilerini sürdürmeleri olasıdır. Bu nedenle pazarlama yöneticilerinin sadece müşteri

ayrılma olasılığını değil bununla birlikte elde tutmaya yönelik sunulan promosyonları kabul etme olasılığını da göz önünde bulundurmaları önem arz etmektedir (Kim ve Moon, 2012: 11718; Kumar ve Shah, 2004: 328).

Kim ve Moon'a göre (2012: 11719) kayıp müşteri analizlerinde ideal olan uygulama pazarlama yöneticilerinin uygun elde tutma teklifleri sunulmadığında ayrılma olasılığı en yüksek olan müşterileri göz önünde bulundurarak hareket etmeleri ve planlarını bu doğrultuda gerçekleştirmeleri şeklindedir. Her müşteri için ayrılma olasılığı hesaplandığında, pazarlama yöneticileri ayrılma olasılığını göz önünde bulundurarak en üstteki % X oranda müşteriyi seçip elde tutmaya yönelik pazarlama çalışmaları gerçekleştirebilirler. Sadık müşterilerin hepsi yüksek gelir bırakmayabilir ve yüksek gelir bırakan müşterilerin hepsi sadık olma eğilimi göstermeyebilir. Bu nedenle karlılığı ve sadakati eş zamanlı bir şekilde yönetmek gerekir. Kısıtlı pazarlama bütçesini promosyon teklifleriyle ilgilenmeyecek müşterilere harcamak yanlış olacaktır. Öncelikli olarak pazarlama teklifleriyle ilgilenen ve yüksek gelir bırakacak müşteriler hedeflenmelidir.

2.5. Kayıp Müşteri Tahminleme Yöntemleri

Kayıp müşteri analizine yönelik literatür incelendiğinde farklı sektörlerde gerçekleştirilmiş çalışmalar betimleyici ve tahminleyici olmak üzere iki kategoride ele alınabilir. Betimleyici çalışmaların amacı kayıp müşteriye neden olan faktörleri belirlemek iken (Ahn vd., 2006: 552; Kim ve Yoon, 2004: 751), tahminleyici çalışmalar müşteriler ayrılma davranışı göstermeden önce kimlerin ayrılacağını tahminlemeye yönelik modeller kurmayı hedeflemektedir (Burez ve Van den Poel, 2009: 4626; Coussement ve Van den Poel, 2008: 164; Hung vd., 2006: 515; Wei ve Chiu, 2002: 103).

Kayıp müşteri tahminleme çalışmalarında en çok sınıflandırma yöntemlerine başvurulduğu görülmekte olup (Miguéis vd., 2012b: 338), yapay sinir ağları (artificial neural network), genetik algoritma (genetic algorithm), karar ağaçları (decision tree), karar vektör makinesi (support vector machine) (Han, 2018:1), sağkalım analizi (survival analysis) (Lai ve Zeng, 2014: 372), Rastgele Orman algoritması (Random Forests) (Breiman, 2001: 29), güçlendirme algoritması AdaBoost (Vafeiadis vd., 2015: 3), ki-kare otomatik etkileşim belirleme yöntemi ve güçlendirilmiş versiyonu (CHAID ve E-CHAID - Exhaustive - Chi Square Automatic Interaction Detector) (Suh ve Alhaery, 2016: 67), lojistik regresyon (Han, 2018 :1), cox regresyon (Zorn vd., 2008: 3), AntMiner+ Karınca Kolonisi Optimizasyonu (Ant Colony Optimization), aktif öğrenme tabanlı yaklaşım ALBA (Active Learning Based Approach) (Verbeke vd., 2011: 2354) gibi birçok farklı yöntemden faydalanılmaktadır.

Kayıp müşteri tahminleme çalışmalarına yönelik literatür incelendiğinde sınıflandırma yöntemlerinden karar ağaçları ve lojistik regresyonun en çok kullanılan metodlar olarak öne çıktığı görülmektedir (Tablo 2.1). Karar ağaçları, Quinlan (1992) tarafından önerilmiş bir veri madenciliği yöntemi olup (akt. Miguéis vd., 2012b: 340) sınıflandırma ve tahminlemede kullanılır (Wang vd., 2009: 8072). Karar ağaçları hedef değişkenle ilişkili grupları birbirinden ayırmada en güçlü değişkenleri kullanarak veriyi alt gruplara böler (Suh ve Alhaery, 2016: 67). Analizde oluşturulan ağaç grafiği karar ağacını temsil eder. Öncelikle ağacın kök düğümüne bir veri giriş yapar ve bunu takiben çocuk düğümün nereye gideceğine karar verilir. Bu süreç verinin yaprak düğüme ulaşmasına kadar tekrar eder (Wang vd., 2009: 8072). Diğer bir değişle, belirli bir sonuca ulaşıldığında düğüm bölünmeyi bırakır ve yaprağa dönüşür. Yapraklar genellikle en çok gözlemin bulunduğu sınıfın /kategorinin adını alır. Kök düğümden yapraklara kadar ulaşan her yol bir kuralı betimlemektedir. Belirli bir kuralın bütün şartlarını karşılayan bir veri kaydı (örneğin bir müşterinin otel işletmesinden yaptığı satın alımlara yönelik girdiler) sonuçtaki yaprağın adına göre sınıflandırılmaktadır (örneğin kayıp müşteri ya da aktif müşteri) (Avon, 2016: 18).

En çok kullanılan karar ağaçları arasında C4.5 (Quinlan, 1996: 725), sınıflandırma ve karar ağaçları (CART- Classification and Regression Tree) (Breiman vd., 1984), ve ki-kare otomatik etkileşim belirleme (CHAID) (Kass, 1980) yöntemleri sayılabilir. Farklı algoritmalar kullanılmasına rağmen, karar ağacı yöntemlerinin amaçları aynıdır. Karar ağaçlarından ulaşılan sonuçlar kolay anlaşılabilir nitelikte olup daha az hesaplama kullanılarak veriler sınıflandırılabilir. Ayrıca karar ağaçları nümerik ve kategorik verilerde kullanıma uygundur (akt. Wang vd., 2009: 8072). Bununla birlikte karar ağacı algoritmaları verideki ufak değişikliklerden fazlasıyla etkilenecek ağaç yapısının ve kuralları tanımlanan ilişkilerin tutarsızlık göstermesine neden olabilmektedir. Bu durum karar ağaçlarında öğrenme aşamasında veriyi ezberleme (overfitting) sorununa yol açmaktadır (Avon, 2016: 18). Bu problemin oluşması durumunda katkı sağlamayan düğümlerin modelden çıkarılması sürecini ifade eden budama işlemi yapılması gerekebilmektedir (Quinlan ve Rivest, 1989: 230-231). Karar ağaçlarında yaşanan bu problemlerin aşılması için Rastgele Orman (Random Forest) algoritması geliştirilmiştir. Rastgele orman algoritması bir topluluk (ensemble) öğrenme algoritmasıdır. Topluluk (kollektif olarak da anılır) algoritmalarında daha güçlü bir model oluşturabilmek için farklı sınıflandırma algoritmaları bir araya getirilir. Rastgele orman algoritması da birçok karar ağacının topluluğundan oluşan bir öğrenme algoritması olup her bir sınıflandırıcı için değişkenler arasından rastgele olarak bir yordayıcı alt seti seçer.

Tablo 2.1 Kayıp Müşteri Tahminleme Yöntemleri ve Çalışma Alanları

Yazar, Yıl	Uygulama Alanı	Kullanılan Algoritmalar										
		Karar Ağaçları	Lojistik Regresyon	Destek Vektör Makinesi	Cox Regresyon	Rastgele Orman	Yapay Sinir Ağları	Naive Bayes	ADA Boost	Sağkalım Analizi	Ant Miner+	ALBA
Zorn vd., 2008	DVD kiralama firması				x							
Lim vd., 2017	Üyelik kartı şirketi	x										
Vebreke vd., 2011	İletişim										x	x
Coussement ve Van den Poel, 2008	Gazete yayınevi		x	x		x						
Miguéis vd., 2012a	Market		x									
Miguéis vd., 2012b	Market		x			x						
Van den Poel ve Larivière, 2004	Bankacılık ve Sigorta				x							
Coussement vd., 2017	İletişim		x									
Huang vd., 2010	İletişim	x		x								
Wang vd., 2009	Kablosuz ağ şirketleri	x										
Vafeiadis vd., 2015	İletişim	x	x	x			x	x				
Hadden vd., 2008	Altyapı hizmetleri						x					
Maier ve Prusty, 2016	Sosyal klüpler	x										
De Wit, 2017	Otel		x	x					x			
Han, 2018	Otel rezervasyon web siteleri		x			x						
Tamaddoni Jahromi vd., 2010	İletişim (kontratsız)	x										
Suh ve Alhaery, 2016	Kumarhane	x										
Lai ve Zeng, 2014	Dijital kütüphaneler									x		

Diğer bir ifadeyle Rastgele Orman algoritması her bir ağacı veri setinin farklı bir örneğinden ön yükleme olarak (bootstrap) oluşturur ve budamaya gerek duyulmaz. Ayrıca bir veri kaydının sınıflandırılmasında Rastgele Orman algoritması tüm karar ağaçlarını bir oylama

sistemi kullanarak birleştirir. Her ağaç, veri kaydını yerleştireceği sınıf için bir oylama yapar ve sonuç için en çok oylanan sınıf tercih edilir (Avon, 2016: 18; Miguéis vd., 2012b: 345 -346; Breiman, 2001: 29). Bu yöntem ezberleme sorununa karşı dayanıklı olup birçok sınıflandırma yöntemine kıyasla çok daha yüksek performans göstermektedir (Breiman, 2001: 29).

Lojistik regresyon, bağımlı değişkenin ikili (dichotomy) olması durumunda başvurulmuş bir regresyon analizidir (müşteri ayrıldı-ayrılmadı vb.). Lojistik regresyon veri madenciliği teknikleri arasında sayılan bir sınıflandırma metodudur. İstatiksel teoriye dayalı bilgisayar tabanlı bir tanımlama tekniği olması nedeniyle birçok avantajı bulunmaktadır. Lojistik regresyon analizi çıktıları diğer sınıflandırma algoritmalarına kıyasla daha fazla bilgi içerir. Her regresyon yaklaşımında olduğu gibi, her bir yordayıcı değişkenle sonuç değişkeni arasındaki ilişkiyi açıklar. Lojistik regresyon sadece yordayıcı değişkenin ne kadar etkili olduğuna yönelik bir ölçüm sağlamakla kalmaz (katsayı büyüklüğü –coefficient size), bununla birlikte ilişkinin yönünü de belirtir (pozitif ya da negatif) (Han, 2018: 1).

Güçlendirme algoritması (Boosting algorithm), zayıf öğrenme algoritmalarının sınıflandırma performansını arttırmak amacıyla birçok sınıflandırma modelinin karar kombinasyonlarını kullanır. Perakende firmalarda ve iletişim sektöründe kayıp müşteri tahminlemede başarılı uygulamaları bulunmaktadır. En popüler güçlendirme algoritması AdaBoost'tur (Vafeiadis vd., 2015: 3).

Sağkalım analizi (Survival Analysis) bir varlığın sağkalımını ya da spesifik bir olayla başlayan ve sonradan ortaya çıkan bir olay ile sonlanan bir süreci analiz eden bir takım istatistiksel metodlardan oluşur. Örneğin, misafir ilişkileri süresini bir varlığın (müşteri) yaşam süresi olarak düşündüğümüzde, başlangıç olayı ilişkinin başlaması ve sonlandırıcı olay müşterinin ayrılma davranışı olarak tanımlanabilir. Sağkalım analizi tıp çalışmalarında hastaların yaşam süresini analiz etmede kullanılmıştır. Son zamanlarda bu yöntemin yazılım mühendisliği, bankacılık sektörü ve kayıp müşteri analizi gibi birçok farklı alanda kullanıldığı görülmektedir (Lai ve Zeng, 2014: 371 - 372).

Karar destek makineleri (SVMs – Support Vector Machines), öğrenme algoritmalarının eşlik ettiği denetimli öğrenme modelleri olup veriyi analiz ederek örüntüleri ayırtmaktadır. Karar destek makineleri sınıflandırma ve regresyon analizleri için kullanılabilen yapısal risk minimizasyonu tabanlı bir öğrenme tekniğidir (Vafeiadis vd., 2015: 2).

AntMiner+ Karınca Kolonisi Optimizasyonu (Ant Colony Optimization) prensibine dayanan yüksek performanslı bir veri madenciliği tekniğidir ve final kural setinden alan bilgisini (domain knowledge) edinmeyi sağlar. Diğer taraftan, ALBA (Active Learning Based Approach) aktif öğrenme tabanlı yaklaşımında hem model tarafından oluşturulan kurallar daha

anlaşılabilir olmakta hem de modelin yüksek düzeyde kesinlik sağlayan öngörü yeteneği bulunmaktadır. Verbeke vd. (2011: 2354) tarafından yapılan bir çalışmada ALBA'nın sınıflandırma tekniklerinin öğrenmesini geliştirdiği böylece yüksek performansa sahip anlaşılabilir modeller sağladığı görülmüştür. AntMiner+ ise kesin, anlaşılır ve bunların ötesinde doğruluğu kanıtlanabilir modeller sağlayabilmektedir.

2.6. Farklı Sektörlerde Kayıp Müşteriyi Tanımlamak

Kayıp müşteri tahminlemede en zorlayıcı konu kayıp müşteri tanımlamasını yapmaktır. Özellikle sözleşmesiz çalışan bir işletmede bu analiz gerçekleştirileceği zaman göz önünde bulundurulması gereken birden çok faktör ortaya çıkmaktadır. Bu bağlamda işletmenin bir müşteriyi kayıp olarak tanımlayabilmesi için kaç ay ya da yıl geçmesi gerektiğini belirlemesi gerekebilir. Farklı bir yaklaşımla, temel alınan süre içerisinde müşterinin ortalama harcamaları başlangıçtaki harcamasıyla kıyaslanabilir. Ancak başlangıçtaki harcamaları çok düşük olan bir müşteri için bu hesaplamanın yapılması mantıksız olabilir (Rud, 2001: 259). Sözleşme kullanılmayan sektörlerde ayrılan müşteriyi belirlemede belirli bir süre içerisinde müşterinin etkin olmama durumu, satın alma sıklığı ve/veya müşterinin işlemleri arasında geçen süre aralığı ya da analiz zamanı ve son satınalım arasında geçen süre aralığı gibi göstergelerden yararlanılır (Suh ve Alhaery, 2016: 68). Örneğin kumar endüstrisinde müşterilerin ayrılma zamanını tahminlemede genellikle müşterinin aktif olmadığı 90 günlük bir periyod kullanılır. Ancak daha uzun süreli ziyaret döngüsü bulunan bir müşteri (örneğin 180 gün) bu durumda aktif olmasına rağmen yanlışlıkla kayıp müşteri olarak tanımlanabilir. Eğer 90 günlük periyod temel alınarak bu müşteriye bir takım geri kazanım teklifleri sunulursa, bu promosyonu değerlendirerek gelecek olan müşteri işletmeye ekstra maliyetli olacaktır. Çünkü aslında aktif olan bu müşteri herhangi bir promosyon teklifi olmasa da işletmeye geri dönecektir. Sonuç olarak bu yanlış tanımlamalar işletmeye boşa harcanmış pazarlama bütçesi olarak yansıyacaktır (Suh ve Alhaery, 2016: 65). Dolayısıyla bir işletme kayıp müşteri tanımlamasını yaparken işletme amaçlarıyla uyumlu ve harekete geçmeyi kolaylaştıracak bir hesaplama yöntemi bulmalıdır (Rud, 2001: 259).

Literatürde kayıp müşteri analizinin iletişim (Vafeiadis vd., 2015; Qureshi vd., 2013) başta olmak üzere bir çok farklı sektörde gerçekleştirildiği görülmektedir (Bakınız: Tablo 2.1) Bunların arasında bankacılık (Prasad ve Madhavi, 2012: 96), sigortacılık (Günther vd., 2014: 58), kumar (Kawale vd., 2009: 423; Suh ve Alhaery, 2016: 63), konaklama sektörü (Maier ve Prusty, 2016: 797; De Wit, 2017: 2; Han, 2018:1) ve marketler (Miguéis vd., 2012b: 337) sayılabilir. Bankacılık ve sigortacılık sektörlerinde ayrılan müşterileri tanımlamada genellikle

hesap kapatma durumu temel alınmaktadır (Kumar ve Ravi, 2008: 9; Larivière ve Van den Poel, 2005: 475; Morik ve Kpcke, 2004: 325). İletişim sektöründe ise ayrılma genellikle operator değiştirme olarak ortaya çıkmaktadır (Hwang vd., 2004: 181; Hung vd., 2006: 517; Wei ve Chiu, 2012:103). Perakende sektöründe ayrılma davranışının genellikle kısmi kayıp olarak ortaya çıktığı görülmektedir. Bu durum herhangi bir üyelik sistemi ya da sözleşme bulunmamasından dolayı müşterilerin işletmeyi ayrılma durumuyla ilgili bilgilendirmek zorunda olmayışı ve farklı işletmelerden alışveriş yaparak satın alımlarını farklı işletmeler arasında bölüştürmesinden kaynaklanmaktadır. Bu şekilde çalışılan sektörlerde müşteriler aşamalı olarak işletmeyle ilişkilerini azaltabilir ve sonuç olarak tam bir ayrılma söz konusu olabilir (Buckinx and Van den Poel, 2005: 255; Burez and Van den Poel, 2009: 4631).

Sözleşme bulunmayan sektörlerde kayıp müşteri tanımlaması yapmak karmaşık bir durumdur. Örneğin, Tamaddoni Jahromi ve arkadaşları (2010: 592), sözleşmesiz çalışan ön ödemeli cep telefonu şirketlerine yönelik gerçekleştirdikleri çalışmada kayıp müşteri tanımlamasını yapabilmek adına model oluşturma sürecini iki aşamaya bölerek müşteri verilerini dört kümeye ayrılmış ve her bir küme için tahmin dönemlerini hesaplamıştır (yöntem için bakınız: Wei ve Chiu, 2002). Benzer şekilde, kumarhanelerde kayıp müşterilerin tanımlanmasında genellikle müşterinin aktif olmadığı süre temel alınmaktadır. Bu tanımlama işletmeden işletmeye farklılık gösterebilmekle birlikte genellikle 90 günlük bir periyodun kullanıldığı görülmektedir (Suh ve Alhaery, 2016: 65).

Maier ve Prusty (2016: 802) tarafından sosyal klüplerde gerçekleştirilen bir çalışmada kayıp müşteriyi tanımlamada üyelik iptali bilgisi temel alınmıştır. Konaklama sektöründe gerçekleştirilen bir diğer çalışma ise Han (2018) tarafından otel rezervasyon web sitelerinde (Airbnb vb.) uygulanmıştır. Bu çalışmada sözleşme kullanılmaması nedeniyle ilk üç aydaki toplam talep miktarına kıyasla son üç aydaki talepleri %80 oranında azalan müşteriler kayıp müşteri olarak tanımlanmıştır. Benzer şekilde De Wit (2017: 13-14) tarafından otel işletmelerinde kontratsız çalışan ticari müşterilere yönelik gerçekleştirilen bir çalışmada kayıp müşterinin tanımlanmasında, önceki etkinliğin devam eden süreçte devam ettirilmemesi kural olarak kullanılmış ve ölçümleme ve tahmin periyodlarının belirlenmesinde Tamaddoni Jahromi ve arkadaşlarının (2010: 592) yöntemi uyarlanmıştır. Lai ve Zeng (2014: 374) tarafından dijital kütüphanelerde gerçekleştirilen kayıp müşteri tahminleme çalışmasında ise bir yıldan uzun süre boyunca kütüphanenin veri tabanından kaynak indirmeyen müşteriler ayrılan müşteri olan tanımlanmıştır.

2.7. Kayıp Müşteri Tahminlemesinde Kullanılan Değişkenler

Kayıp müşteri tahminlemesi çalışmalarında genellikle müşteri demografik bilgileri ve geçmiş davranışlarına yönelik veriler kullanılmaktadır (Suh ve Alhaery, 2016: 68). Kayıp müşteri tahminleme çalışmalarında kullanılmış davranış değişkenleri Tablo 2.2’de, demografik ve diğer değişkenler ise Tablo 2.3’te sunulmuştur. Ayrıca, Van den Poel ve Larivière (2004: 208) tarafından müşteri elde tutma çalışmalarına yönelik yapılan literatür taramasında kayıp müşteri davranışını tahminlemede kullanılmış algı değişkenleri özetlenmiştir. Bu çalışmada kullanılan kayıp müşteri analizi için sadece veri madenciliği ve makine öğrenimi teknikleriyle sınırlı kalınmamış, müşteri ayrılma davranışıyla değişkenler arasındaki ilişkiyi ele alan anket çalışmaları da ilave edilmiştir (Bakınız Tablo 2.4).

Girdi değişkenleriyle ilgili olarak, Coussement ve Van den Poel (2008:164) tarafından gerçekleştirilen bir çalışmada, yapılandırılmamış metinsel bilginin geleneksel kayıp müşteri tahminleme modellerine eklenmesi sonucunda modelin tahmin gücünün arttığı keşfedilmiştir. Yazarlar çalışmalarında müşteri hizmetlerine gönderilmiş e-posta içeriklerine metin madenciliği uygulamış ve müşteri/şirket etkileşimi değişkenleri, üyelikle ilgili değişkenler, üyelik yenilemeye özgü değişkenler ile sosyo-demografik verilerden faydanmışlardır. Benzer şekilde, Huang ve arkadaşları (2010) kayıp müşteri tahminleme çalışmalarında şikâyet bilgisi, sözleşme bilgileri, hata raporları gibi farklı değişkenlerin kullanılmasının tahmin performansını arttıracaklarını önermiştir. Ayrıca, tahminlemede kullanılan yordayıcıların PCA, (principal components analysis – temel birleşenler analizi), SOM (self-organizing maps / kendini düzenleyen haritalar) ve genetik algoritma metodu gibi teknikler yardımıyla azaltılması önerilmektedir.

Zhang vd. (2012: 102) tarafından gerçekleştirilen kayıp müşteri analizinin bulgularına göre bazı kayıp müşteriler sadece geleneksel yordayıcı değişkenler tarafından tespit edilebilirken, bazı müşteriler kişiler arası etkileşimlerden daha fazla etkilenmekte ve sadece bu etkileşimlerine yönelik değişkenler tarafından doğru tahminlenebilmektedir. Dolayısıyla, gerçekleştirilen kayıp müşteri analizlerinde müşteri davranışlarına yönelik daha kapsamlı değişkenlerin geleneksel değişkenlere ilave edilmesi tahmin modelinin performansı açısından önem arz etmektedir.

Tablo 2.2 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Davranış Değişkenleri

Davranış Değişkenleri	Burez ve Van den Poel, 2008	Coussement ve De Bock, 2013	De Wit, 2017	Hadden vd., 2008	Han, 2018	Huang vd., 2010	Maier ve Prusty, 2016	Suh ve Alhaery, 2016	Tamaddoni Jahromi vd., 2010	Van den Poel ve Larivière, 2004	Wang vd., 2009	Zorn vd., 2008
Tutar (monetary)	✓	✓	✓					✓	✓	✓		✓
İlişki süresi (ilk gelişten itibaren geçen süre)	✓	✓			✓			✓	✓			
Güncellik (recency)		✓	✓						✓		✓	
Satın alımlar arasında geçen zaman		✓	✓							✓		
Sıklık (frequency)		✓	✓					✓				
Ödüller/ Bonus / Promosyon	✓	✓						✓				
Rezervasyon yapılan oda sayısı			✓		✓							
Şikâyet sayısı				✓	✓							
Üyelik süresi	✓						✓					
Web sitesini ziyaret sayısı												✓
Hizmet taleplerinin sayısı						✓						
İlk hizmet alınan tarih									✓			
İptal sayısı	✓											
Kalış süresi			✓									
Kişi sayısı			✓									

Davranış Değişkenleri	Burez ve Van den Poel, 2008	Coussement ve De Bock, 2013	De Wit, 2017	Hadden vd., 2008	Han, 2018	Huang vd., 2010	Maier ve Prusty, 2016	Suh ve Alhaery, 2016	Tamaddoni Jahromi vd., 2010	Van den Poel ve Larivière, 2004	Wang vd., 2009	Zorn vd., 2008
Katılım gerçekleştirilen etkinlikler							✓					
Müşterinin hizmet aldığı toplam gün sayısı									✓			
Müşterinin iki hizmet alımı arasındaki en uzun süre									✓			
Müşterinin iki hizmet kullanımı arasında geçen ortalama süre									✓			
Müşterinin son hizmet kullanımından itibaren geçen süre											✓	
Müşterinin toplam hizmet kullanımı sayısı											✓	
Müşterinin ürünü kiraladığı süre												✓
Müşterinin üyelik planını değiştirme sayısı												✓
Otelin konumu			✓									
Ödeme şekli					✓							
Rezervasyonu yapan kişi (şirket, acente ya da misafir) -Booker			✓									
Son üç ayda rezervasyon yapılan odaların ortalama ücreti					✓							
Şikâyet türü				✓								
Tamirat randevusu sayısı				✓								
Üyeliğin biteceği ay	✓											

Tablo 2.3 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Diğer Değişkenler

<i>Değişken Grupları</i>	Bolton, 1998	Burez ve Van den Poel, 2008	Coussement ve De Bock, 2013	Gustafsson, Johnson ve Ross, 2005	Hadden vd., 2008	Han, 2018	Huang vd., 2010	Maier ve Prusty, 2016	Suh and Alhaery, 2016	Van den Poel ve Larivière, 2004	Zorn vd., 2008
<i>Algısal değişkenler</i>											
Genel memnuniyet	✓			✓							✓
Yorum sayısı											✓
<i>Demografik değişkenler</i>											
Geldiği yer (bölge/şehir)		✓	✓			✓	✓	✓	✓		
Cinsiyet			✓				✓	✓	✓	✓	
Yaş		✓						✓	✓	✓	
Müşterinin şirket olup olmaması		✓									
İş yerinin konumu								✓			
Sosyal statü										✓	
Bağış alma durumu							✓				
Dil			✓								
Eğitim seviyesi										✓	

<i>Değişken Grupları</i>	Bolton, 1998	Burez ve Van den Poel, 2008	Coussement ve De Bock, 2013	Gustafsson, Johnson ve Ross, 2005	Hadden vd., 2008	Han, 2018	Huang vd., 2010	Maier ve Prusty, 2016	Suh and Alhaery, 2016	Van den Poel ve Larivière, 2004	Zorn vd., 2008
<i>İletişimsel değişkenler</i>											
Müşterinin teknik problemler vb. için işletmeyi arama sayısı		✓			✓	✓					
Son doldurulan anketten itibaren geçen süre		✓									
Müşterinin telefon numarasının olup olmadığı		✓									
Kaçırılan randevu sayısı					✓						
İşletmenin söz verip zamanında tamamlayamadığı işlerin sayısı					✓						
Müşterinin doldurduğu anket sayısı		✓									
<i>Makroçevre değişkenleri</i>											
Ülkenin refahlık indeksi										✓	
Şirket birleşmeleri										✓	
<i>Teknik değişkenler</i>											
Tamirat süresi					✓						
Sorunun çözülme süresi					✓						

Tablo 2.4 Kayıp Müşteri Tahminleme Çalışmalarında Kullanılan Algı Değişkenleri

<i>Algı Değişkenleri</i>	<i>Yazar, Yıl</i>
Genel memnuniyet	Bloemer vd., 1998 Keaveney ve Parthasarathy, 2001 Lemon vd., 2002 Mittal ve Kamakura, 2001 Mols, 1998 Nguyen ve LeBlanc, 1998 Paulin vd., 1998 Varki ve Colgate, 2001
Hizmet kalitesi	Athanassopoulos, 2000 Bloemer vd., 1998 Clark, 1997 Colgate ve Danaher, 2000 Ganesh vd., 2000 Jones vd., 2000 Jones vd., 2002 Keaveney, 1995 Levesque ve McDougall, 1996 Mittal ve Lasser, 1998 Nguyen ve LeBlanc, 1998 Paulin vd., 1998 Varki ve Colgate, 2001 Zeithaml vd., 1996
Müşteri ilişkilerinin performansı	Athanassopoulos, 2000 Bolton, 1998 Colgate ve Danaher, 2000 Ganesan, 1994 Ganesh vd., 2000 Jones vd., 2002 Keaveney, 1995 Levesque ve McDougall, 1996 Paulin vd., 1998
Hizmet hatası deneyimi	Maxham, 2001 Zeithaml vd., 1996

<i>Algı Değişkenleri</i>	<i>Yazar, Yıl</i>
Problem çözümünden duyulan memnuniyet	Conlon ve Murray, 1996 Keaveney, 1995 Levesque ve McDougall, 1996 Maxham, 2001 Zeithaml vd., 1996
Fiyat/maliyet algısı	Athanassopoulos, 2000 Colgate ve Danaher, 2000 Ganesh vd., 2000 Keaveney, 1995 Varki ve Colgate, 2001
Davranışsal niyet	Bolton vd., 2000 Ganesh vd., 2000 Lemon vd., 2002 Mols, 1998
Algılanan değiştirme maliyetleri	Ganesan, 1994 Jones vd., 2000 Jones vd., 2002 Mols, 1998
İşletmeye kıyasla rakiplerin performansı	Bolton vd., 2000 Keaveney, 1995 Levesque ve McDougall, 1996
İmaj/ İtibar	Bloemer vd., 1998 Ganesan, 1994 Nguyen ve LeBlanc, 1998
Konumsal kolaylık	Athanassopoulos, 2000 Ganesh vd., 2000 Keaveney, 1995
Öngörülen pişmanlık	Lemon vd., 2002
Etik sorunlar	Keaveney, 1995
Müşterilerin işletmeye bağımlı olduklarına yönelik algıları	Ganesan, 1994
Müşterilerin işletmenin kendilerine bağımlı olduklarına yönelik algıları	Ganesan, 1994
Mağazalar arasındaki algısal uzaklık	Popkowski vd., 2000
Güven	Ganesan, 1994

Kaynak: Van den Poel ve Larivière, 2004: 208-209

2.8. Kayıp Müşteri Tahminleme Çalışmalarına Yönelik Literatür

Coussement ve Van den Poel (2008: 165-171) kayıp müşteri tahminlemesi için kullanılan bir karar destek sisteminin performansını optimize etme konusunda gerçekleştirdikleri çalışmada, yapılandırılmamış metinsel bilginin geleneksel kayıp müşteri tahminleme modellerine eklenmesi sonucunda modelin tahmin gücünün arttığını keşfetmiştir. Yazarlar çalışmalarında Belçika’da faaliyet gösteren büyük çaplı bir gazete yayınevinin üye bilgilerini kullanmış ve müşteri hattına gönderilmiş müşteri e-postalarının içeriklerine metin madenciliği uygulamıştır. Yapılan çalışmada 30 aylık müşteri bilgisi (90.000 müşteri) içeren müşteri/şirket etkileşimi değişkenleri, üyelikle ilgili değişkenler, üyelik yenilemeye özgü değişkenler ile sosyo-demografik veriler kullanılmıştır.

Zorn ve arkadaşları (2008: 3-5) tarafından Avustralya’da bulunan bir DVD kiralama firmasının müşteri verilerine (10.156 mevcut ve eski müşteri) cox regresyon metodu uygulanarak gerçekleştirilen bir çalışmada kayıp müşterileri en anlamlı şekilde yordayan değişkenin “müşterinin üyelik planını değiştirme sayısı” olduğu bulunmuştur. Çalışmada kullanılan diğer değişkenler satınalma davranışı (firma websitesini ziyaret sayısı, yorum yapılan başlık sayısı, müşterinin DVD’yi kiraladığı süre), aylık üyelik ücreti, puanlama, müşterinin üyelik planının değiştirme sayısıdır.

Lim ve arkadaşları (2017: 76-87) tarafından Malezya’da bulunan bir sadakat kartı firmasının müşterileriyle (alışveriş merkezi, zincir eczane işletmesi, petrol şirketi) gerçekleştirilen bir çalışmada, karar ağacı ve regresyon analizinden yararlanılarak kayıp müşteri tahminlemesi yapılmış ve firmalara müşterilerini elde tutmaya yönelik faydalı önerilerde bulunulmuştur. Yazarlar gerçekleştirdikleri kayıp müşteri analizinde iki yıllık üyelik verilerinden yararlanmış ve analiz sonucunda 4 ila 8 yıllık veri kullanılmasının tahminleme için daha geçerli sonuçlara ulaştıracağını önermişlerdir.

Verbeke ve arkadaşları (2011: 2354, 2358-2359) kayıp müşteri tahmin modellemesinde veri madenciliği kullanımına yönelik bir literatür taraması gerçekleştirmişlerdir. Literatür taraması sonucunda anlaşılabilirlik ve sezgiyle elde etme (intuitiveness) konularına çok az yer verildiği görülmüş ve bu boşluğu doldurmak üzere kayıp müşteri tahminleme modellerine iki yeni veri madenciliği tekniği adapte edilmiştir. Bu bağlamda, geleneksel kural çıkarım (rule induction) tekniklerinden C4.5 ve RIPPER kıyaslanmıştır. Çalışmada bir telefon operatörünün 5000 gözlemi içeren veri seti kullanılmıştır. Veri setindeki kayıp müşterileri belirlemede en etkili olan değişkenler sırasıyla gün içerisindeki kullanım (dakika/ay), günlük kullanım ücreti (US doları/ay), müşteri hizmetlerini arama sayısı, uluslararası plan üyeliği, ses mesajı plan üyeliği, ses mesajı sayısı, uluslararası arama ücreti (US doları/ ay), uluslararası kullanım

(dakika/ay), gece kullanımı (dakika/ay), gece kullanım ücreti (US doları/ay) ve uluslararası arama sayısıdır. Çalışmanın sonuçlarına göre hem AntMiner+ hem de ALBA anlaşılır ve kesin sınıflandırma kural setleri üretmiştir.

Miguéis ve arkadaşları (2012a:11252-11255) tarafından, Avrupa menşeli bir perakende market şirketinin veri deposundan çekilen müşteri verileri (75.000 yeni müşteri) üzerinde lojistik regresyon sınıflandırma metodu kullanılarak gerçekleştirilen bir çalışmada kısmi kayıp (partial churn) tahminlemesi yapılmıştır. Çalışmada ilk satın alınan ürünün başarısı güven unsuru olarak kabul edilmiş ve keşif modelini oluşturmada ilk satın alınan ürün kategorisi ile en son tarihli müşteri-işletme etkileşimlerinin ileri ve geri kronolojik sıralamaları kullanılmıştır. Çalışmanın sonuçlarına göre oluşturulan iki modelin de standart RFM modelinden daha üstün performansa sahip olduğu görülmektedir.

Van den Poel ve Larivière (2004: 196, 214-215), Avrupa menşeli bankacılık ve sigortacılık hizmetleri sunan bir finansal hizmet şirketinin veri ambarında bulunan müşteri verilerine (47.157 müşteri) cox orantılı risk modeli (cox proportional hazard model) uyguladıkları kayıp müşteri analizi çalışmasında, müşteriye elde tutmadaki en önemli faktörlerin demografik özellikler, çevresel değişimler ve müşterilerle kurulan interaktif ve sürekli ilişkiler olarak ortaya çıktığı; bununla birlikte satın alınan ürün sayısı ile iki ürün satın alımı arasında geçen zaman değişkenlerinin ayrılma davranışı üzerindeki etkisinin zayıf olduğu görülmüştür. Çalışmanın bulgularına göre müşterilerin kritik ayrılma kararları iki dönemde ortaya çıkmaktadır. Birinci dönem kişinin ilk müşteri oluşunu takip eden ilk yıllardır. İkinci dönem ise, yirmi yıllık müşterisi olma sonrasındaki yılları kapsamaktadır.

Coussement ve arkadaşları (2017: 27, 31-32), Avrupa menşeli bir mobil iletişim firmasının müşteri verilerine (%4,52 kayıp oranıyla 30.104 müşteri) kayıp müşteri tahminleme analizi uygulamıştır. Çalışmada müşteri davranışları (giden arama dakikası vb.), müşteri-işletme etkileşimi değişkenleri (destek merkeziyle görüşme sayısı vb.), üyelikle ilgili değişkenler (tarife planının türü vb.) ve müşteri demografik bilgilerini (cinsiyet vb.) içeren 156'sı kategorik, 800'ü sürekli olmak üzere toplamda 956 değişken kullanılmıştır. Çalışmanın bulgularına göre lojistik regresyon tekniğinin yüksek performanslı bir kayıp müşteri analiz metodu olarak ortaya çıktığı ve kullanılan veri hazırlama tekniklerinin kayıp müşteri tahminleme performansını etkilediği görülmüştür.

Huang ve arkadaşları (2010: 3657), kayıp müşteri tahminleme oranlarını yükseltmek amacıyla bir sabit hat iletişim şirketinin müşteri verilerine (47.391 müşteri / yaklaşık 11.000 kayıp) kayıp müşteri tahminlemesi uygulamış ve yeni bir değişken seti ile üç yeni girdi pencere tekniği (window technique) önermiştir. Çalışmada kayıp müşteri tahminlemede karar ağaçları,

çok katmanlı algılayıcı sinir ağları (multilayer perceptron neural networks) ve destek vektör makinesi (support vector machine) kullanılmıştır. Kullanılan değişkenler; demografik veriler, hesap bilgisi, toplu arama detayları, hat bilgisi, hizmet talepleri, fatura ve ödeme bilgisi, bağış bilgisi ve Henley segmentasyonudur (müşteriler ile potansiyel müşterileri iki farklı gruba ayırarak özelliklerine, ihtiyaçlarına ve ticari değerlerine göre seviyelendiren bir algoritmadır). Çalışmanın bulgularına göre önerilen yeni ve kapsamlı değişken seti kayıp müşteri tahminlemede etkili olmuştur.

Wang ve arkadaşları (2009: 8071, 8073-8074), bir kablosuz ağ şirketinin müşteri verilerine (4000 müşteri, 60.000 işlem verisi) üç aylık bir dönem boyunca karar ağacı tekniği uygulayarak kayıp müşterilere yönelik tavsiyeler geliştiren bir sistem önermişlerdir. Çalışmada ilk 9 haftalık süreç öğrenme verisi olarak, son ay ise test verisi olarak kullanılmıştır. Değişken seçimi yapılırken, ham veride yer alan doğrusal korelasyona sahip olabilecek (örneğin alışveriş sayısı ve toplam tutar vb.) veriler istenmeyen sonuçlar oluşturabileceğinden dolayı temizlenmiş ve son kullanımdan itibaren geçen gün sayısı, kullanıcının toplam bağlantı sayısı ve müşterinin ayrılıp ayrılmadığı bilgisi kullanılmıştır. Çalışmada kayıp müşterileri önlemeye yönelik yeni pazarlama ve promosyon stratejileri geliştirmede fayda sağlayan bir tavsiye sistemi (recommender system) geliştirilmiştir.

Vafeiadis ve arkadaşları (2015: 1, 8), bir iletişim şirketinin müşteri verileri üzerinde modern makine öğrenimi metodları kıyaslanarak (yapay sinir ağları, destek vektör makineleri, karar ağaçları, naïve bayes sınıflandırma teoremi, lojistik regresyon) kayıp müşteri tahminlemesi yapılmıştır. Çalışmanın ikinci aşamasında modelin performansını arttırmak adına güçlendirme algoritması (boosting algorithm) uygulanmıştır. En etkili parametre kombinasyonunu belirlemek adına her bir metod ve parametre için Monte Carlo benzetim serileri kullanılmıştır. Çalışmanın bulgularına göre güçlendirme algoritması kullanılan modeller kullanılmayanlara kıyasla çok daha üstündür. Bununla birlikte, en iyi sınıflandırma modeli yaklaşık %97 doğruluk oranı ve %84'ün üzerinde F-skoruyla AdaBoost kullanılan POLY destek vektör makinesi (Polynomial kernel, SMV-POLY) olarak ortaya çıkmıştır.

Hadden ve arkadaşları (2008: 809-812), altyapı hizmetleri sunan bir şirketin müşteri verilerine (öğrenme verisinde %50'si kayıp 202 müşteri, doğrulama verisinde %30'u kayıp 700 müşteri) sinir ağları, regresyon analizi ve regresyon ağaçları tekniklerini kıyaslayarak kayıp müşteri tahminlemesi yapmışlardır. Çalışmada şirketin şikâyet ya da tamirin sonuçlandırılmasına yönelik verileri, şikâyet verisi ve tamir verisi değişken seti olarak kullanılmıştır. Çalışmanın sonuçlarına göre kayıp müşterileri en iyi tahminleyen yöntem Bayesian kullanılan sinir ağları tekniğiyle, ayrılmayan müşterileri tahminlemede en doğru

sonuçları doğrusal regresyon analizi vermiştir. Genel olarak en yüksek performanslı yöntemin regresyon ağaçları, en zayıf yöntemin ise sinir ağları olduğu görülmüştür.

Maier ve Prusty (2016: 803-816), sosyal klüplerde karar ağaçları ve lojistik regresyon algoritmalarından yararlanarak gerçekleştirdikleri kayıp müşteri tahminleme çalışmasında yedi yıllık bir müşteri verisi kullanılmıştır. Veri setinde aktif olan ve olmayan tüm üyelerin bilgileri yer almıştır. Üyelerin demografik bilgileri yaş, cinsiyet, yaşadığı evin konumu, çalıştığı ofisin konumu, üyelik süresi (tenure) ve üyelerin gelir getiren ve gelir beklenmeyen (non revenue generating) etkinlikleri kullanımlarına yönelik verileri kapsamaktadır. 12 aydan daha kısa süredir üye olanlar ile gönülsüz ayrılanlar (ölüm durumu vb.) analize dahil edilmemiştir. Veri seti 1.898 aktif üye (2013 yılı itibariyle aktif) ile 1.337 (2006 ve 2013 yılları arasında üyeliğini iptal eden) kayıp üye olmak üzere toplam 3.235 kişiyi içermektedir. Kayıp üyeler toplam veri setinin %42'sini oluşturmaktadır. Çalışmada orijinal veri seti rastgele iki bağımsız veri setine ayrılmış olup; bu veri setlerinin yaklaşık %70'i öğrenme verisi (%57 aktif, %43 kayıp), yaklaşık %30'u ise test verisi (%61 aktif, %39 kayıp) olarak kullanılmıştır. Çalışmanın bulgularına göre gelir beklenmeyen sosyal klüp etkinlikleri müşterinin çıkışını takip eden 120 gün içerisindeki üyelik iptalini en iyi yordayan değişkenlerdir. Maier ve Prusty'e (2016: 797) göre çalışmanın bulguları sosyal klüplerdeki ilişkisel pazarlama uygulamalarıyla ilişkili müşteri davranışlarına yönelik yeni bir anlayış sağlamak ve sosyal klüp müdürlerinin müşteriyi elde tutma stratejilerini iyileştirmelerine yönelik pratik çözümler sağlamaktadır.

Ballings ve Van den Poel (2012: 13517), kayıp müşteri tahminlemesi için kaç yıllık müşteri verisinin yeterli olacağını belirlemek üzere zaman aralığı optimizasyonu (time window optimization) uygulamışlardır. Kayıp müşteri tahminleme modeli performansındaki gelişmeyi gözlemek amacıyla lojistik regresyon ve sınıflandırma ağaçları kullanılarak müşteri verisi bir yıldan on altı yıla kadar uzatılmıştır. Çalışmanın bulgularına göre beşinci yıldan itibaren tahminleme performansında sadece marjinal bir artış gözlemlenmektedir. Yazarlar, araştırmada verileri kullanılan şirketin verilerinin %69'unu kullanmadan da başarılı bir tahminleme yapabileceğini önermektedir.

Han (2018: 2-3), otel rezervasyon web sitelerinde (Airbnb vb.) lojistik regresyon ve rastgele orman yöntemlerini kullanarak kayıp müşteri tahminleme analizi uygulamıştır. Çalışma kapsamında bir otel rezervasyon websitesinin merkezi veri tabanından 12 aylık veri alınmıştır. İlk üç aydaki toplam talep miktarına kıyasla son üç aydaki talepleri %80 oranında azalan müşteriler ayrılan müşteri olarak belirlenmiştir. Oluşturulan karar ağacına göre ayrılan müşterileri yordamada grup müşterilerin “ödeme şekli” en önemli değişken olarak ortaya çıkmış olup, “ilk üç aydaki toplam talep sayısı” ve “üyelik kartı sayısı” bunu takip etmektedir.

Uygulanan yöntemler arasında rastgele orman modeli lojistik regresyona kıyasla daha yüksek AUC (Area Under Curve) değeri hesaplanması nedeniyle tercih edilmiş olup; modelin performans oranı %78.9 olarak hesaplanmıştır.

De Wit (2017: 19-54), otel işletmelerinde kontratsız çalışan ticari müşterilere (B2B – şirketler) lojistik regresyon, AdaBoost ve destek vektör makinesi metodlarını kullanarak kayıp müşteri tahminleme analizi uygulamıştır. Çalışmada müşteriler iş amaçlı ve kontrat kullanılmayan etkinliklerde bulunan şirketler olarak tanımlanmıştır. Çalışmada kayıp müşterinin tanımlanmasında, önceki etkinliğin devam eden süreçte devam ettirilmemesi kural olarak kullanılmış ve kalibrasyon ve tahmin periyodlarının belirlenmesinde Tamaddoni Jahromi ve arkadaşlarının (2010: 592) yöntemi uyarlanmıştır. 2013-2016 yılları arasında toplanan 2830 adet B2B şirket verisi kullanılmıştır. Çalışmada kullanılan değişkenler güncellik, sıklık, iki satın alım arasında geçen ortalama süre, toplam harcama tutarı, rezervasyon türü (otel, konferans ya da ikisi birlikte), oda sayısı, kişi sayısı, kalış süresi, rezervasyonu yapan kişi (şirket, acente ya da misafir), ziyaret edilen otel konumu sayısı verilerini içermektedir. Frekansı bir olan şirketlerle, banket, düğün gibi konularla ilgili veriler analize dahil edilmemiştir. Öğrenme ve test seti için eşit kayıp/aktif müşteri oranları kullanılmış olup, kayıp oranı yaklaşık %41’dir. Öğrenme seti için verinin %70’i, test seti içinse %30’u kullanılmıştır. Çalışmanın bulgularına göre en önemli yordayıcı değişken “güncellik” olarak ortaya çıkmıştır. Bunu takiben gözlem süresi, iki satın alım arasında geçen ortalama süre, ziyaret edilen otel konumu sayıları ve sıklık değişkenleri de önemli yordayıcılar olarak bulunmuştur.

Tamaddoni Jahromi ve arkadaşları (2010: 590-596), sözleşmesiz çalışan ön ödemeli cep telefonu şirketlerine yönelik iki aşamalı bilgisayar destekli kümeleme ve sınıflandırma modelleri kullanarak kayıp müşteri tahminleme analizi uygulamışlardır. Çalışmada 1 Kasım 2007-30 Nisan 2008 tarihleri arasında 34.504 kullanıcının arama kaydı verisi kullanılmıştır. Kullanılan veri arama tarihi, arama zamanı, arama süresi, arama türü ve ücret bilgilerini içermektedir. Sözleşmesiz çalışılmasından dolayı kayıp müşteri tanımlamasını yapabilmek için model oluşturma süreci iki aşamaya bölünmüştür; kümeleme ve sınıflama. Kayıp müşteri tanımlamasını yapmada kullanılan kümeleme aşamasında arama oranı, müşterinin iki araması arasında geçen ortalama süre, en son arama yapılan tarih, ilk arama yapılan tarih, müşterinin aktif olduğu süre, bir kişinin iki araması arasında geçen en uzun süre, müşterinin arama yaptığı ya da arama aldığı toplam gün sayısı, müşterinin çalışma periyodu dahilinde gelen ve giden ayrı olmak üzere toplam arama sayıları, müşterinin tüm hizmetler için toplamda yaptığı ödeme tutarı, müşterinin çalışma periyodu dahilinde gelen ve giden ayrı olmak üzere toplam arama süreleri değişken olarak belirlenmiştir. Sonrasında ise RFM temelli bu 12 değişken kullanılarak

müşteri verileri dört kümeye ayrılmış ve tahmin periyodları hesaplanmıştır (bakınız: Wei ve Chiu, 2002). Kayıp müşteri tahminlemesinde sınıflama tekniklerinden Karar Ağacı (CART) kullanılmıştır. Çalışmanın bulgularına göre, gelen arama frekansındaki değişimler, müşterinin kullanım dakikası, gelen ve giden aramaların başlangıç frekansı aktif ve kayıp müşterileri belirlemede en önemli değişkenlerdir.

Suh ve Alhaery (2016: 71-76) tarafından Amerika Birleşik Devletleri'nde yasal kabul edilen online kumarhanelerin oyun verileri kullanılarak bir karar ağacı algoritması olan E-CHAID (Exhaustive - Chi Square Automatic Interaction Detection) yöntemiyle 2639 oyuncu verisine kayıp müşteri tahminlemesi uygulanmıştır. Çalışmada kayıp müşterileri tahminlemede sıklık, tutar, müşterisi olma süresi, bonuslar/ödülleri, müşterinin son oyunundan itibaren geçen süre (Recency), oyunlar arasında geçen ortalama süre (inter playday average) ve standart sapması (inter playday SD) gibi davranışsal verilerle; yaş, cinsiyet ve konum gibi demografik değişkenler kullanılmıştır. Ayrılan müşteriyi tanımlamada bu kuraldan yararlanılmış olup; “Recency > Inter Playday Avg + 2 x Inter Playday SD”, kuralı sağlayan müşteriler kayıp müşteri olarak kodlanmıştır. Bu kural sonucunda ulaşılan kayıp müşteri oranı %32.4'tür. Ancak araştırmada dengeli örnekleme tekniğinden faydalı olarak bu oran %49.9 olacak şekilde örneklem seçilmiştir. Çalışmanın bulgularına göre, E-CHAID metodu kullanılan karar ağacı modeli %80 oranında doğru bir sınıflandırma yapmıştır ve RFM temelli değişkenler (oyun oynanan gün sayısı, müşteri olma süresi, ilk ve son oynama günü arasında geçen süre, tutar vb.) kayıp müşterileri tahminlemede önemli yordayıcılardır. Çalışmada en önemli yordayıcılar, sırasıyla, ilk ve son oynama günü arasında geçen süre, bonus ödül tutarı, slot oyunları oynama gün sayısı, ilk üyelikten veri çevrilen toplam slot sayısı olarak bulunmuştur.

Lai ve Zeng (2014 : 376-381) tarafından Çin'de dijital kütüphanelerde sağkalım analizi kullanılarak 8054 müşterinin ayrılma davranışı analiz edilmiş ve müşteri davranışları temel alınarak müşteri sınıflandırması yapılmıştır (kilit müşteriler, sadık müşteriler, yaygın müşteriler). Çalışmada bir yıldan uzun süre boyunca kütüphanenin veri tabanından kaynak indirmeyen müşteriler kayıp müşteri olan tanımlanmıştır. Çalışmanın bulgularına göre müşteri ayrılma oranı oldukça yüksek olmakla birlikte (%65.7 kayıp) ayrılma çoğunlukla ilk üç ayda gerçekleşmektedir.

ÜÇÜNCÜ BÖLÜM

OTEL İŞLETMELERİNDE KAYIP MÜŞTERİ TAHMİNLEME UYGULAMASI

3.1. Araştırmanın Amacı ve Kapsamı

Bu araştırmanın öncelikli amacı makine öğrenimi metodlarından faydalanılarak bir zincir otel işletmesinin müşteri veri tabanında bulunan müşteri verilerine kayıp müşteri tahminlemesi uygulanmasıdır. Bu amaçla:

- (1) İki farklı makine öğrenimi algoritması olan Lojistik Regresyon ve Rastgele Orman metodlarının karşılaştırması yapılarak, otel işletmelerine yönelik müşteri verilerinde hangi yöntemin daha başarılı olduğu tespit edilecektir.
- (2) Otel işletmelerine yönelik müşteri verileriyle gerçekleştirilen bir kayıp müşteri tahminlemesinde hangi değişkenlerin kayıp müşteri davranışını yordamada daha önemli olduğu tespit edilecektir.
- (3) Otel işletmesini bir sonraki dönemde terk etme ihtimali yüksek olan müşteriler önceden tespit edilecektir.

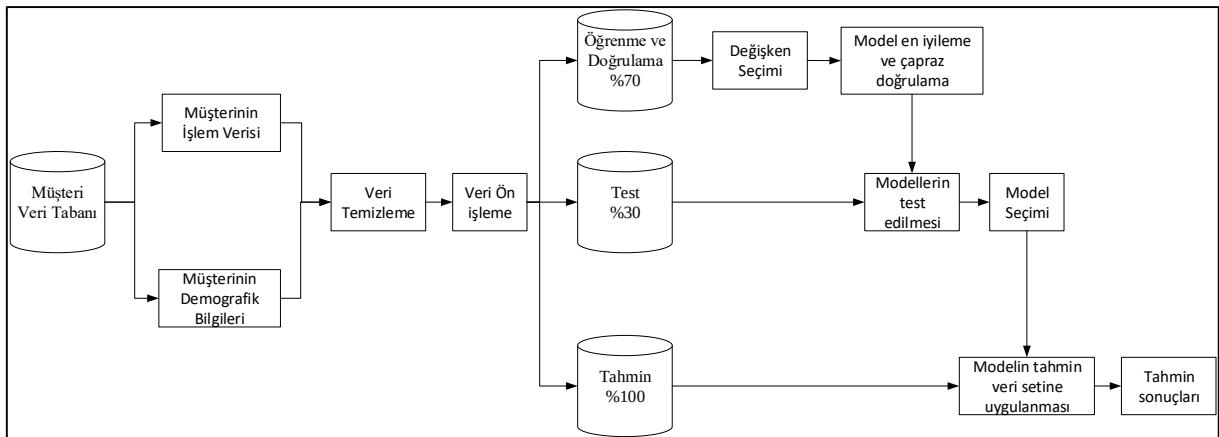
Dolayısıyla, analiz bulgularından faydalanılarak potansiyel kayıp müşteriler arasında bir önceliklendirme yapılabilir ve bu müşterileri elde tutmaya yönelik çözüm önerilerinde bulunulabilir. Ayrıca daha önce turizm alanında sadece sosyal kulüp üyeleri, B2B otel müşterileri ve otel rezervasyon websitesi kullanıcıları için gerçekleştirilmiş kayıp müşteri tahminleme yönteminin bir resort otel zincirinin müşteri verilerine uyarlanması yapılarak literatüre katkı sağlanması amaçlanmaktadır.

Bu çalışma Antalya, Belek bölgesinde faaliyet gösteren ve bünyesinde beş yıldızlı üç resort otel bulunan bir zincir otel işletmesinin en az 10 gece konaklama yapan, en az 2 defa gelen (repeat) ve çalışmada belirlenen dönemlerde aktif olan müşterilerini kapsamaktadır.

3.2. Araştırma Modeli

Bu çalışmada bir zincir otel işletmesinin müşteri veri tabanında bulunan müşteri verileri kullanılarak makine öğrenimi yöntemlerinden sınıflandırma problemlerinde yararlanılan iki metodun karşılaştırılması ve tespit edilen en iyi model aracılığıyla kayıp müşteri tahminlemesi yapılması amaçlanmaktadır. Araştırma modeli Şekil 3.1’de sunulmuştur.

Şekil 3.1’de gösterildiği üzere, araştırma kapsamında öncelikle müşteri veri tabanından çekilen müşteri işlem verileri ve demografik bilgileri literatür taramasından elde edilen bilgiler ışığında temizlenmiş ve ön işlemesi yapılmıştır. Veriyi temizleme ve analize hazır hale getirme işlemleri için Jupyter Lab (Python programlama dili) kullanılmıştır.



Şekil 3.1 Araştırma Modeli

Analize hazır hale getirilen veriler WEKA programı aracılığıyla %70'e %30 oranında ayrılarak öğrenme ve test veri setleri oluşturulmuştur. Öğrenme aşamasına başlanmadan önce değişken seçimi yapılmıştır. Değişken seçimi yapılırken WEKA programından faydalanılmıştır. Öğrenme aşamasında ise iki farklı öğrenme algoritması çapraz doğrulama metoduyla en iyileme ayarları yapılarak doğruluk oranları karşılaştırılmıştır. Kayıp müşteri analizi alan yazınında öğrenme algoritmaları iki grupta toplanmaktadır. Bunlar karar ağaçları, destek vektör makinası ve lojistik regresyon gibi temel (tekil) öğrenme modelleri ve rastgele orman ve adaboost gibi topluluk (ensemble) öğrenme modelleridir (Coussement ve De Bock, 2013:1630). Bu çalışmada topluluk öğrenmesi modellerinden Rastgele Orman ve temel öğrenme modellerinden Lojistik Regresyon kullanılmıştır. Yöntemlerin uygulanmasında bir veri madenciliği programı olan Orange'dan faydalanılmıştır. Öğrenen modeller test veri seti üzerinde denenmiş ve doğruluk oranları karşılaştırılmıştır. Test aşaması sonucunda en başarılı bulunan model tahmin veri setine uygulanarak, 2019-2021 periyodunda ayrılma davranışında bulunması muhtemel olan müşteriler tespit edilmiştir.

Orange programı ve Jupyter Lab alanına “Anaconda Navigator” aracılığıyla ücretsiz olarak ulaşılabilir. WEKA veri analizi proramı ise Waikato üniversitesi web sitesi üzerinden ücretsiz olarak kurulabilir.

3.3. Araştırmanın Yöntemi ve Bulguları

3.3.1. Veri Setinin Temizlenmesi ve Veri Ön İşleme

Araştırmada kullanılan veri seti zincir otel işletmesinin bünyesinde bulunan üç otelde (A Hotel, B Hotel ve C Hotel olarak anılacaktır) 2018 yılına kadar konaklama yapmış müşterilerin demografik verilerini ve işlem verilerini kapsamaktadır. Bu veri seti zincir otel işletmesinin Bilgi İşlem Departmanı tarafından hazırlanmış olup; müşteri id numaraları ve otel id numaralarının karşısında müşteri kartında tutulan tüm müşteri işlem verilerini içermektedir.

Bilgi İşlem Departmanı tarafından hazırlanan veri setinde 791.289 satır veri bulunmaktadır. Bu veri setinden bir temsili kesit Şekil 3.2’de örnek olarak sunulmuştur. Bu veri seti içerisinde aynı müşteri id numarası karşılığında ilgili müşterinin tüm işlem verisi bulunabildiği gibi, oteller tarafından aynı müşteriye açılan farklı kartlar olması nedeniyle hem farklı oteller hem de aynı otel için birden fazla karta sahip müşteriler bulunabilmektedir.

[illegible]

Şekil 3.2 Veri Setinden Temsili Bir Kesit

Bu nedenle öncelikle müşteri hareketleri aynı id numarası karşısında birleştirilmiş ve son konaklamadaki otel id numarası ile varsa diğer konaklamalardaki farklı otel id numaraları tek satıra alınmıştır. Bu veri seti içerisindeki değişkenler sırasıyla müşteri id numarası, otel id numarası, oda tipi, oda numarası, otel adı, giriş tarihi, çıkış tarihi, milliyet, ödeyen/eşlik eden misafir olma durumu, paket türü, acente/münferit rezervasyon bilgisi, gecelik ödeme tutarı, para birimi, yetişkin sayısı, ücretsiz çocuk sayısı, ödeyen çocuk sayısı ve bebek sayısı verileridir. Ayrı bir veri setinde ise müşteri id numaraları karşısında müşterinin yaşı, milliyeti, geldiği pazar ve sadakat kartı bilgileri bulunmaktadır. İlk veri setine müşteri id numaraları aracılığıyla bu bilgiler bağlanmıştır.

Veri setinin temizlenmesinde Jupyter Lab alanından ve Python programlama dilinden yararlanılmıştır. Pandas kütüphanesi aracılığıyla veriler ayrı sütunlara bölünerek değişken isimleri eklenmiş ve ilgili işlemler aynı müşteri id numarası altında birleştirilmiştir. Bu birleştirme işleminden sonra en eski konaklama tarihi, en güncel konaklama tarihi, güncellik, geceleme sayıları, dolar bazında harcama tutarı, seyahat eşlikçileri, son konaklama sezonu, grubun konaklanan diğer otelleri, konaklama yapılan son otel, son konaklamada acente aracısı kullanılması ya da münferit konaklanması, iki konaklama arasında geçen günlerin ortalaması ve toplam geliş sayıları değişkenleri her müşteri için Pandas ve Numpy kütüphaneleri kullanılarak hesaplanmıştır.

Harcama verilerinin hesaplanmasında tüm harcamalar ilgili kur üzerinden dolar cinsine dönüştürülmüştür. Kur kategorilerindeki kirli verilerin temizlenmesinde Şekil 3.3'deki Pandas kodlarından faydalanılmıştır. Dolar bazında harcama tutarı hesaplaması için gözlem periyodunun (Bakınız: 3.3.3. Kayıp Müşterinin Tanımlanması ve Zaman Aralığı) son dönemi

olan 2015 yılı ortalama dolar kurları esas alınmıştır. Tahmin veri seti içinse 2018 yılı ortalama dolar kurları kullanılarak aynı işlem tekrarlanmıştır. Kur çevirme işleminde kullanılan Pandas kodları Şekil 3.4’de sunulmuştur.

```
df[ListeCurr]= df[ListeCurr].replace(regex=['SEUR', '1EUR', '1EU'], value= 'EUR')
df[ListeCurr]= df[ListeCurr].replace(regex=['1USD', '1US'], value= 'USD')
df[ListeCurr]= df[ListeCurr].replace(regex=['GBP1'], value= 'GBP')
df[ListeCurr]= df[ListeCurr].replace(regex=['YTL'], value= 'TL')
df[ListeCurr]= df[ListeCurr].replace(regex=['1CHF'], value= 'CHF')
```

Şekil 3.3 Kur Kategorilerinin Düzenlenmesinde Kullanılan Kodlar

```
def EurToUsd (EUR):
    return EUR * 1.11 # 2015 ortalama kuru

def GbpToUsd (GBP):
    return GBP * 1.53 # 2015 ortalama kuru

def TLToUsd (TL):
    return TL * 0.37 # 2015 ortalama kuru

def CHFToUsd (CHF):
    return CHF * 1.04 # 2015 ortalama kuru

def CurrEX (Kur, Tutar):
    if Kur == 'EUR':
        return EurToUsd(Tutar)
    elif Kur == 'TL':
        return TLToUsd(Tutar)
    elif Kur == 'GBP':
        return GbpToUsd(Tutar)
    elif Kur == 'CHF':
        return CHFToUsd(Tutar)
    elif Kur == 'USD':
        return Tutar
    else:
        pass

df['YeniTutar1'] = df.apply(lambda x: CurrEX(x['Kur1'],x['Tutar1']),axis=1)
```

Şekil 3.4 Kur Dönüştürmede Kullanılan Kodlar

Seyahat eşlikçileri değişkeni için çocuklu, bebekli, bebekli ve çocuklu, tek, çift, yetişkin grubu halinde yapılan konaklamalar Şekil 3.5’de sunulan kodlar aracılığıyla ayrıştırılmıştır. Müşterinin grubun diğer otellerinde konaklayıp konaklamadığı bilgisi için öncelikle her bir oteldeki konaklama sayıları hesaplanmış ve A Hotel, B Hotel, C Hotel, A ve C Hotel, B ve C Hotel, B ve A Hotel ile All Hotels kategorileri Şekil 3.6’daki kodlar aracılığıyla ayrıştırılmıştır. Kodlarda geçen “L” değişkeni son konaklama yapılan otel id numarasını temsil etmektedir.

```
def Accomp (adl, chl, pchl, inf):
    if adl== 1 and chl== 0 and pchl == 0 and inf== 0:
        return 'tek'
    elif adl== 2 and chl== 0 and pchl == 0 and inf== 0:
        return 'cift'
    elif adl> 0 and chl> 0 and pchl>= 0 and inf== 0:
        return 'cocuklu'
    elif adl> 0 and chl>= 0 and pchl> 0 and inf== 0:
        return 'cocuklu'
    elif adl> 0 and chl== 0 and pchl== 0 and inf> 0:
        return 'bebekli'
    elif adl> 0 and chl>= 0 and pchl> 0 and inf> 0:
        return 'bebekli ve cocuklu'
    elif adl> 0 and chl> 0 and pchl>= 0 and inf> 0:
        return 'bebekli ve cocuklu'
    elif adl> 2 and chl== 0 and pchl== 0 and inf== 0:
        return 'yetiskin grubu'
    else:
        pass

df['Accomp'] = df.apply(lambda x: Accomp(x['adl1'],x['chl1'],x['pch1'],x['inf1'] ),axis=1)
```

Şekil 3.5 Seyahat Eşlikçilerini Belirlemede Kullanılan Kodlar

```
def OtherHotel (L, C, A, B):
    if L> 0 and C> 0 and A> 0 and B> 0:
        return 'All Hotels'
    elif L==1 and C> 0 and A>= 0 and B> 0:
        return 'All Hotels'
    elif L==2 and C> 0 and A> 0 and B>= 0:
        return 'All Hotels'
    elif L==3 and C>= 0 and A> 0 and B> 0:
        return 'All Hotels'

    elif L==1 and C== 0 and A>= 0 and B== 0:
        return 'A'
    elif L==2 and C== 0 and A== 0 and B>= 0:
        return 'B'
    elif L==3 and C>= 0 and A== 0 and B== 0:
        return 'C'

    elif L==1 and C> 0 and A>= 0 and B== 0:
        return 'A ve C'
    elif L==3 and C>= 0 and A> 0 and B== 0:
        return 'A ve C'

    elif L==2 and C> 0 and A== 0 and B>= 0:
        return 'B ve C'
    elif L==3 and C>= 0 and A== 0 and B> 0:
        return 'B ve C'

    elif L==1 and C== 0 and A>= 0 and B> 0:
        return 'B ve A'
    elif L==2 and C== 0 and A> 0 and B>= 0:
        return 'B ve A'

    else:
        pass

# HotelID 1 == A, 2 == B, 3 == C

<<<
df['DigerOtel'] = df.apply(lambda x: OtherHotel(x['L'],x['C'],x['A'],x['B'] ),axis=1)
```

Şekil 3.6 Otel Kategorisi Değişkenlerinin Ayrıştırılmasında Kullanılan Kodlar

Müşterilerin milliyet değişkeni için veri setinde 129 farklı kategori bulunmaktadır. Bu nedenle en çok ziyaret eden milliyet kategorileri kullanılmış, haricinde kalan milliyetler “Diğer” kategorisi altında toplanmıştır. Bu milliyet kategorileri oluşturulurken hem öğrenme ve test veri setindeki hem de tahmin veri setindeki en çok ziyaret eden milliyet kategorileri birlikte değerlendirilmiştir (Tablo 3.1).

Tablo 3.1 Milliyet Kategorileri

Milliyet Değişkeni	Frekans	Kümülatif Yüzde
Türkiye	15094	26,67%
Rusya Federasyonu	10527	45,27%
Almanya	9222	61,56%
İngiltere	2540	66,05%
Romanya	2310	70,13%
Fransa	1715	73,16%
Çekya	1653	76,08%
Polonya	1617	78,94%
Avusturya	1281	81,21%
Ukrayna	936	82,86%
Kazakistan	899	84,45%
İsviçre	895	86,03%
Slovakya	789	87,42%
Hollanda	735	88,72%
Belçika	733	90,02%
Macaristan	452	90,82%
Moldova	405	91,53%
İran	359	92,17%
İsveç	307	92,71%
Lüksemburg	295	93,23%
...	...	...
Toplam	56596	100%

Metin verisi olarak bulunan farklı acente kategorileri ve münferit konaklama durumu için öncelikle müşterinin son konaklamasında tercih ettiği yöntem belirlenmiştir. Acente kategorileri kirli veri içermesi ve aynı kategoriyi ifade eden çok fazla farklı sayıda içerik bulunması nedeniyle sadece “Acente” olarak kodlanmış, münferit durumunu belirlemedeyse Şekil 3.7’deki Pandas kodundan yararlanılmıştır.

```
df.loc[df['2015Agency'].str.contains('MUNF'), 'Agency'] = 'INDIVIDUAL'
df.loc[df['2015Agency'].str.contains('MÜNF'), 'Agency'] = 'INDIVIDUAL'
df.loc[df['2015Agency'].str.contains('INDIVID'), 'Agency'] = 'INDIVIDUAL'
```

Şekil 3.7 Münferit Satın Alımların Belirlenmesinde Kullanılan Kodlar

Gürültüyü engellemek ve sadık müşterilere odaklanmak adına çalışmaya konu olan otel işletmesinin sadık müşteri programındaki şartlar göz önünde bulundurularak grup otellerinin birinde en az 2 defa konaklama yapan, toplamda en az 10 gece konaklayan ve 2015 yılında aktif müşteri olan (en güncel check-out tarihi $\geq$ 01.01.2013) müşteri grubu filtrelenmiştir. Ayrıca ücretsiz konaklama yaptığı belirlenen acente çalışanı, tur rehberi, house-use kullanım yapan otel personeli vb. kayıtlar veri setinden temizlenmiştir. Bu işlemler sonrasında öğrenme ve test için kullanılacak veri setinde 9740 müşteri kalmıştır. Tahmin veri setinde ise grup otellerinden birinde en az 2 defa konaklama yapan, toplamda en az 10 gece konaklayan ve 2018 yılında aktif müşteri olan (en güncel check-out tarihi $\geq$ 01.01.2016) 12413 müşteri bulunmaktadır.

3.3.2. Aykırı Gözlemlerin Temizlenmesi

Makine öğrenmesinde aykırı gözlemler gürültü yaratarak öğrenme önünde bir engel oluşturabilir. Aykırı gözlemlerin tespiti aşaması, gürültü yaratması durumunda veri setinden silinmesi gereken anormal kayıtların araştırılmasını kapsar. Kayıp müşteri tahminlemesinde kullanılan veri setinde aykırı gözlemler farklı nedenlerden kaynaklanabilir. Bazı durumlarda müşteriler işletmenin kontrolü dışında kalan nedenlerden ötürü işletmeyi terk etmek durumunda kalabilirler.

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.neighbors import LocalOutlierFactor

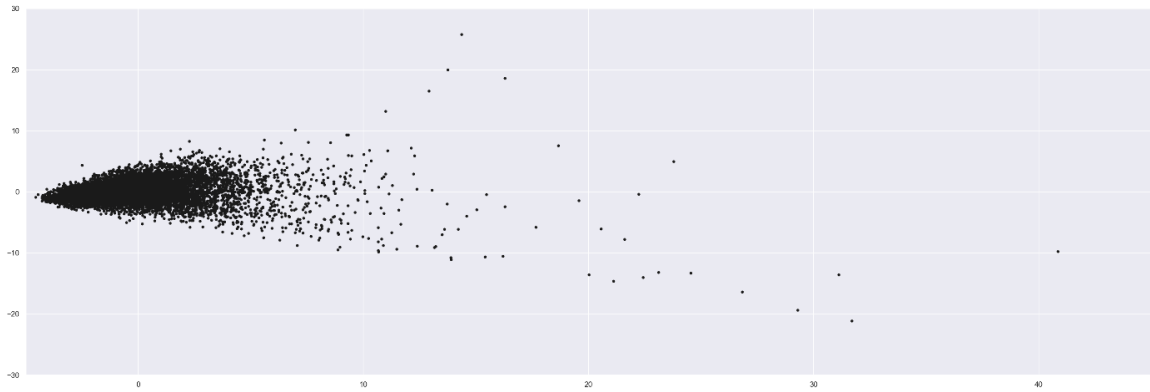
LOF = LocalOutlierFactor(n_neighbors = 20, contamination = "auto")
LOF.fit_predict(df)
output:
array([1, 1, 1, ..., 1, 1, 1])

X_score = LOF.negative_outlier_factor_
X_score[0:3]
output:
array([-0.98438006, -1.07873987, -1.05561325])

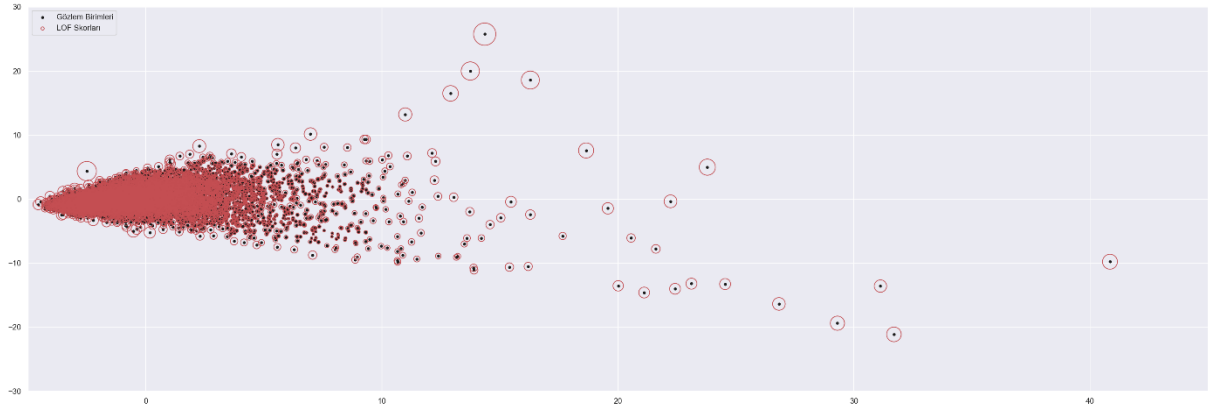
np.sort(X_score)[100:150]
output:
array([-1.91645315, -1.9107837 , -1.91014479, -1.90995094, -1.90650767,
       -1.9014946 , -1.89905056, -1.89825188, -1.89632305, -1.8924541 ,
       -1.89228199, -1.88560114, -1.88280027, -1.88203774, -1.87974186,
       -1.87912535, -1.87712575, -1.87418087, -1.8693412 , -1.8664196 ,
       -1.86626342, -1.86528023, -1.86519968, -1.86378163, -1.86289447,
       -1.86108312, -1.86049978, -1.85980628, -1.8467355 , -1.84412811,
       -1.84378492, -1.84283386, -1.84111298, -1.840717 , -1.83795954,
       -1.83498855, -1.8346316 , -1.83174434, -1.82683825, -1.82471447,
       -1.82201815, -1.82090062, -1.81922041, -1.81563837, -1.81515244,
       -1.81084084, -1.80826914, -1.80740915, -1.80645515, -1.80621061])
```

Şekil 3.8 Lokal Aykırılık Faktörü Hesaplamasında Kullanılan Kodlar

Ölüm, uzak bir ülkeye taşınmak vb. durumlar bu nedenlere örnek olarak sayılabilir. Bu kayıtların veri setinde bulunması gürültüye neden olacaktır (Avon, 2016: 44). Gürültü sorununu önlemek amacıyla öncelikle 80 yaşının üstündeki müşteriler veri setinden çıkarılmıştır. Bunun akabinde, anormallik tespiti için lokal aykırılık faktörü hesaplamasından yararlanılmıştır. Lokal aykırılık faktörü (LOF -Local outlier factor) makine öğrenimi ve sınıflandırma çalışmalarında kullanılan oldukça güçlü bir anormallik tespit yöntemidir. Bu algoritma bir gözlem noktasının komşularının yoğunluğunu temel alarak komşularından ne derecede aykırılık gösterdiğini tespit eder (Alshawabkeh vd., 2010: 104).



Şekil 3.9 Aykırı Gözlemler Saçılma Diyagramı

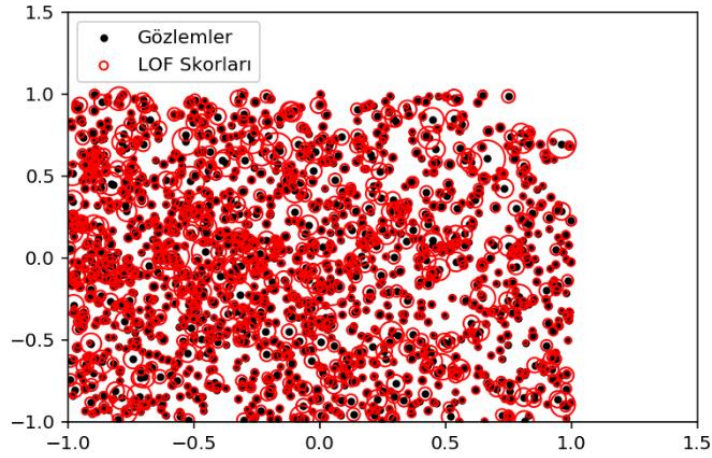


Şekil 3.10 Lokal Aykırılık Faktör Skorları Saçılma Diyagramı

Lokal aykırılık faktörü hesaplamasında Python kütüphanelerinden “scikit-learn” kullanılmıştır (Pedregosa vd., 2011:2825). Analizde ve görsellerin oluşturulmasında kullanılan Python kütüphaneleri ve lokal aykırılık faktörü hesaplamasında kullanılan kodlar Şekil 3.8’de sunulmuştur. Lokal aykırılık faktörü hesaplaması sonucunda tespit edilen aykırı gözlemler ve lokal aykırılık faktör skorları Şekil 3.9 ve Şekil 3.10’da gösterilmiştir.

Lokal aykırılık skorlarının hesaplanmasından sonra tespit edilen aykırı gözlemler veri setinden temizlenmiştir. Bu işlemin sonucunda veri setinde 9192 gözlem kalmıştır. Aykırı

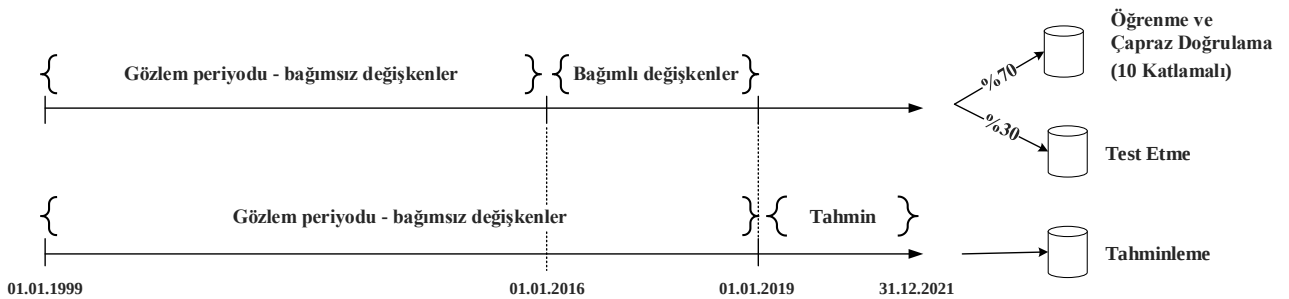
gözlemlerin temizlenmesi işleminden sonra elde edilen saçılma diyagramı Şekil 3.11’de sunulmuştur.



Şekil 3.11 Lokal Aykırılık Faktör Skorları Saçılma Diyagramı II

3.3.3. Kayıp Müşterinin Tanımlanması ve Zaman Aralıkları

Sözleşmeli çalışılmayan sektörlerde kayıp müşterinin tanımlanması oldukça güçtür. Çünkü müşterinin işletmeyle ilişkisini kestiğine yönelik bir bilgi bulunmamaktadır. Bu çalışmada kayıp müşterinin tanımlanmasında müşteri verilerinin elde edildiği otel işletmesinin müşteri sadakat programı kapsamında müşterileriyle paylaştığı kurallar temel alınmıştır. Otel işletmesi müşterilerinin toplamda 10 gecelemeyi aşması durumunda her geceleme için puan kazandırma uygulamasında bulunmaktadır. Müşteriler biriktirdikleri puanları bir sonraki konaklamalarında otel konaklama ücretinden düşerek ya da arkadaşlarına hediye göndererek kullanabilmektedir. Ancak kazanılan puanlar müşterinin otele üç yıl boyunca gelmemesi durumunda silinmektedir. Dolayısıyla, bu çalışmada müşterinin “kayıp müşteri” olarak etiketlenmesinde üç yıl kuralı kullanılmış ve zaman aralıkları üç yıllık periyodlar olarak belirlenmiştir.



Şekil 3.12 Zaman Aralıkları

Kayıp müşteri tahminlemesinde bağımsız değişkenler ve bağımlı (hedef) değişkenler için zaman aralıklarının (time window) belirlenmesi gerekir. Bağımsız değişkenler belirlenen periyod süresince gözlemlenir. Hedef değişkenine yönelik veriler, bağımsız değişkenlerin

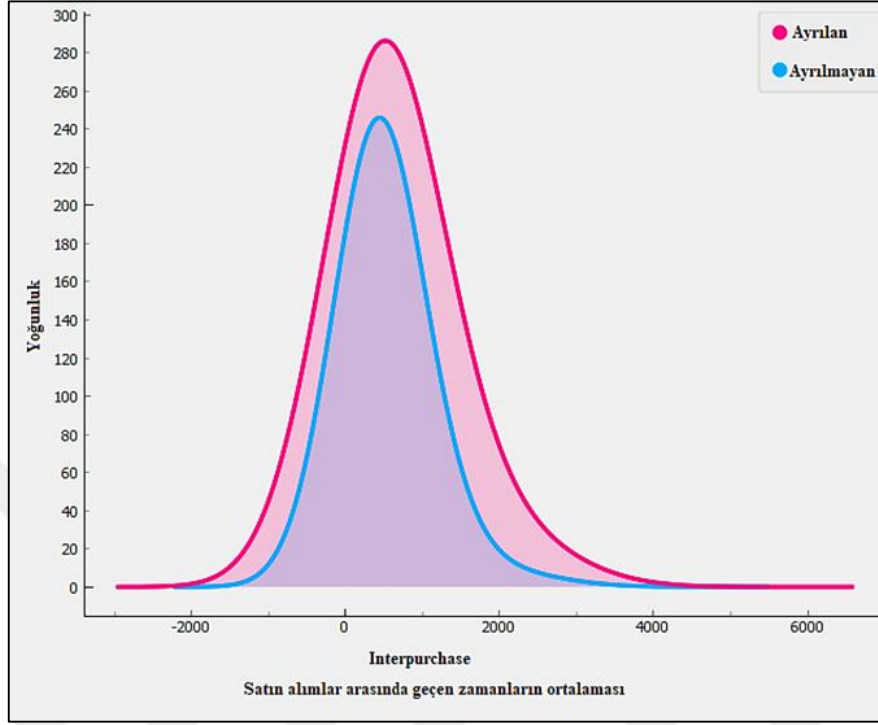
gözlemlendiği periyodu takip eden süreçte ölçülerek veri setine dahil edilir (Burez ve Van den Poel, 2008: 502; Coussement ve De Bock, 2013: 1632; De Wit, 2017: 32). Çalışmada model için oluşturulan veri seti 1999 yılından itibaren otel grubunun müşterisi olan ve 2015 yılında aktif olan 9192 müşterinin 2013 – 2015 (dahil) periyoduna kadar gerçekleştirdikleri konaklamalarından bağımsız gözlemleri ve 2016-2018 (dahil) periyodundaki hedef gözlemleri içermektedir. Tahmin için oluşturulan veri setindeyse 1999 yılından itibaren müşteri olan ve 2018 yılında aktif olan 12413 müşterinin 2016-2018 (dahil) periyoduna kadar gerçekleştirdikleri konaklamalara yönelik veriler bulunmakta ve oluşturulan model aracılığıyla 2019-2021(dahil) periyodunun hedef değişkenleri tahmin edilmeye çalışılmaktadır. Öğrenme, test etme ve tahminleme veri setleri için belirlenen zaman aralıklarının detaylı bir görseli Şekil 3.12’de sunulmuştur.

3.3.4. Değişkenler

Öğrenme aşamasında, müşterilerin geçmiş işlem verilerini gelecekteki davranışlarına bağlayan bir model oluşturulur. Bu aşamada kullanılan veri seti bağımsız girdi değişkenlerini ve bağımlı hedef değişkenini içerir. Girdi değişkenleri müşterinin belli bir süre içerisindeki davranışlarını ya da demografik bilgilerini içerebilir (Coussement ve De Bock, 2013: 1630). Çalışmada kullanılan değişkenler davranışsal ve demografik olmak üzere iki ana başlık altında toplanmaktadır. Davranışsal değişkenler arasında satın alımlar arası geçen zaman, ilişki süresi, tutar, güncellik, sıklık, konaklama süresi, konaklama eşlikçileri, otel, sadakat kartı, satınalım yöntemi, sezon ve ayrılma değişkenleri bulunmaktadır. Demografik değişkenler ise yaş, milliyet ve pazar değişkenleridir. Kullanılan davranışsal değişkenler Tablo 3.3’de, demografik değişkenler ise Tablo 3.4’de özet olarak sunulmuştur.

Müşterilerin satın alımları arasında geçen sürenin uzunluğu (interpurchase time) ayrılma davranışı üzerinde etkili olabilmektedir. Vilcassim ve Jain (1991: 35) müşterinin satın alımları arasındaki süre uzadıkça kullanılan markayı değiştirme davranışında artış olduğunu bulmuştur. Benzer şekilde, Van den Poel ve Larivière’in (2004: 213) gerçekleştirdiği çalışmanın bulgularına göre müşterinin satın alımları arasında geçen süre ortalamasındaki her bir yıllık artış ayrılma olasılığını %4,9 oranında arttırmaktadır. Gerçekleştirilen tez çalışması kapsamında bu değişkenin hesaplanmasında Van den Poel ve Larivière’in (2004: 206) yöntemi kullanılmış, müşterinin geliş tarihleri arasında geçen zamanların ortalaması alınmıştır. Kullanılan veriler nümerik olup, düşük değerler daha sık satın alımları, yüksek değerler ise daha seyrek satın alımları ifade etmektedir. Satın alımlar arasında geçen zamanların ortalaması değişkeninin aldığı değerler 11 ve 3631 arasında değişmekte olup genel ortalaması 612,2

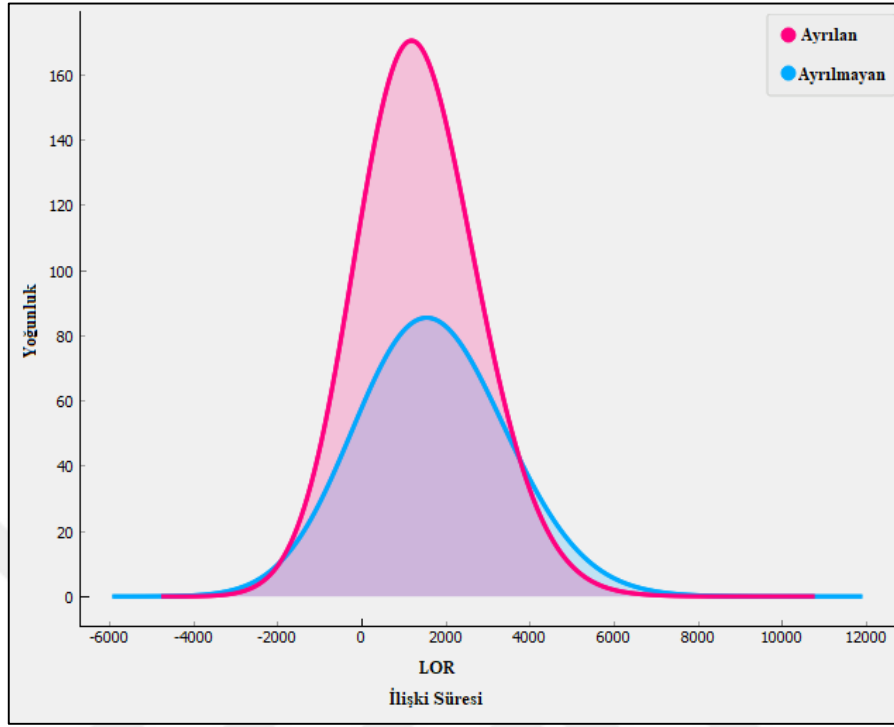
gündür ($\cong 1$ yıl 8 ay, σ (standart sapma) = 492,6). Bu değer ayrılan misafirler için ($\bar{X} = 661,3$) ayrılmayan misafirlere ($\bar{X} = 533,7$) ve genel ortalamaya kıyasla daha yüksektir. Satın alımlar arasında geçen zamanların ortalaması değişkeninin dağılım grafiği Grafik 3.1’de sunulmuştur.



Grafik 3.1 Satın Alımlar Arasında Geçen Zaman Değişkeni Dağılım Grafiği

Müşterinin işletmeyle olan ilişki süresi (LOR - length of relationship) kayıp müşterilerin en iyi yordayıcılarından biridir (Buckinx ve Van den Poel, 2005: 263; Van den Poel, 2003:19; Ballings ve Van den Poel, 2012: 13519). İlişki süresi müşterinin ilk satın alımından itibaren geçen süreyi ifade etmektedir. İlişki süresinin uzunluğu ilişkinin gücünü ifade ettiğinden, bu değer ne kadar yüksekse tekrar satın alma olasılığı da o derecede yüksek olmaktadır (Ballings ve Van den Poel, 2012: 13519). Bu çalışmada Coussement ve De Bock (2013: 1635) tarafından kullanılan metod tercih edilmiş ve müşterinin ilk konaklamasından gözlem süresinin sonuna kadar (31/12 /2015) geçen süredeki gün sayısı kullanılmıştır. Kullanılan veriler nümerik olup, müşterilerin işletmenin kaç yıllık müşterisi olduğunu ifade etmektedir. İlişki süresi değişkeninin aldığı değerler 50 ve 5963 arasında değişmekte olup genel ortalaması 1488,1 gündür ($\cong 4$ yıl, $\sigma = 915,5$). Bu değer ayrılmayan misafirler için ($\bar{X} = 1691,6$) ayrılan misafirlere ($\bar{X} = 1360,3$) ve genel ortalamaya kıyasla daha yüksektir. İlişki süresi değişkeninin dağılım grafiği Grafik 3.2’de sunulmuştur. RFM olarak anılan güncellik (R - Recency), sıklık (F - Frequency) ve tutar (M -Monetary) değişkenleri ayrılma davranışının en güçlü yordayıcıları

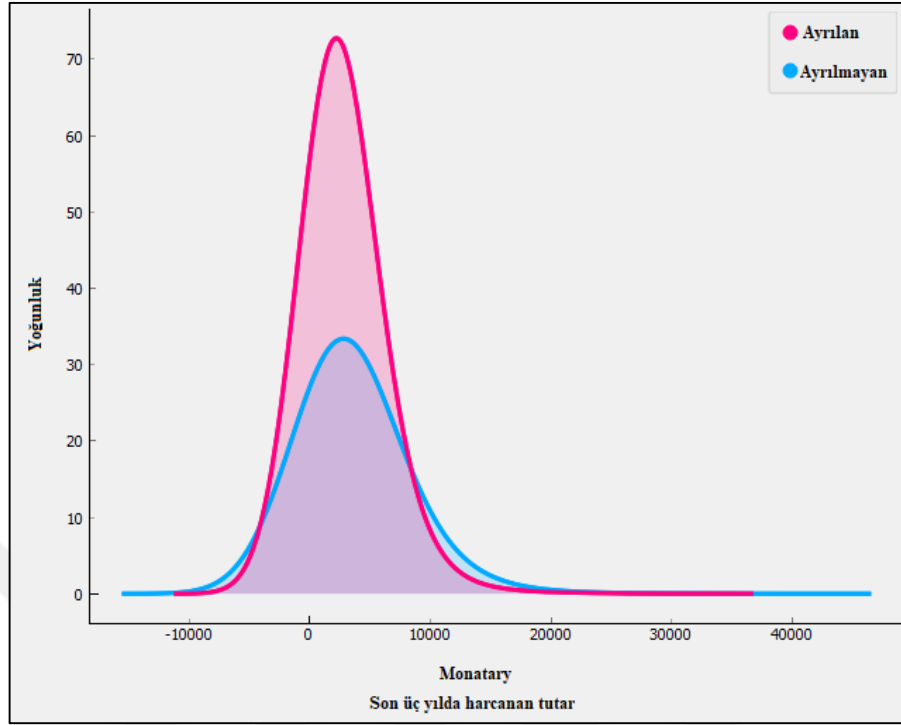
olarak öne çıkmaktadır (Suh ve Alhaery, 2016: 71; Miguéis vd., 2012a: 11253; Ballings ve Van den Poel, 2012: 13519; De Wit, 2017: 22).



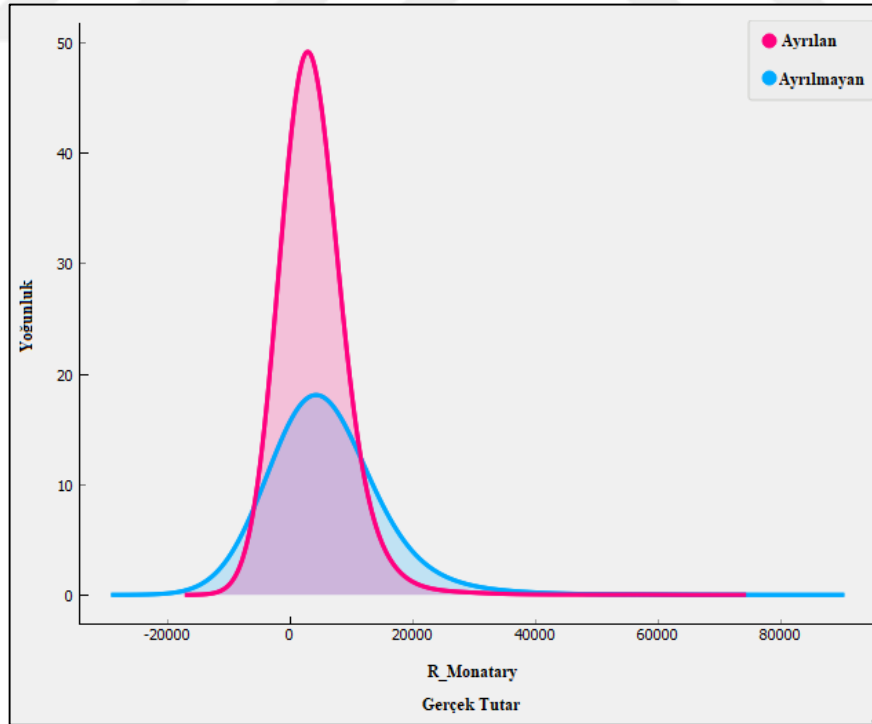
Grafik 3.2 İlişki Süresi Değişkeni Dağılım Grafiği

Müşterinin işletmeye toplam ödediği tutar değişkeni satın alma davranışını yordamada etkili olup (Buckinx ve Van den Poel, 2005: 260) bu değer yükseldikçe müşterinin tekrar satın alma olasılığı da artış göstermektedir (Ballings ve Van den Poel, 2012: 13519). Bu çalışma kapsamında tutar değişkeni için Buckinx ve Van den Poel (2005: 260) tarafından kullanılan yöntem benimsenmiştir. Buckinx ve Van den Poel (2005: 260) daha yakın zamanlar için hesaplanan değişkenlerin müşteri davranışını yordamada daha etkili olabileceğini öne sürmüş, çalışmalarına tüm satın alımlar için hesaplanan değişkenlerin yanı sıra daha güncel periyodlar için hesaplanan değişkenleri de dahil etmiştir. Bu bağlamda gözlem periyodunun tüm müşterilerin aktif olduğu 2013, 2014 ve 2015 yıllarında gerçekleştirdiği konaklamaları için ödediği toplam tutar dolar türünde hesaplanmış ve “Monetary” değişkeni olarak adlandırılmıştır. Müşterilerin ilk konaklamalarından, 31/12/2015 tarihine kadar geçen süredeki tüm konaklamaları için (ilişki süresi değişkeni göz önünde bulundurularak) bu değişken tekrar hesaplanmış ve “R_Monetary” değişkeni olarak adlandırılmıştır. Dolayısıyla “Monetary” değişkeni son üç yıllık tutarı, “R_Monetary” değişkeni ise gerçek tutarı ifade etmektedir. Son üç yılda harcanan tutar değişkeninin aldığı değerler 33,30 \$ ve 31.057,00 \$ arasında değişmekte olup genel ortalaması 3.133,12 dolardır ($\sigma = 2.732,5$). Bu değer ayrılmayan misafirler için ($\bar{X}$

= 3.621,40 \$) ayrılan misafirlere ($\bar{X} = 2.826,73$ \$) ve genel ortalama kiyasla daha yüksektir. Son üç yılda harcanan tutar değişkeninin dağılım grafiği Grafik 3.3’de sunulmuştur.



Grafik 3.3 Son Üç Yılda Harcanan Tutar Değişkeni Dağılım Grafiği

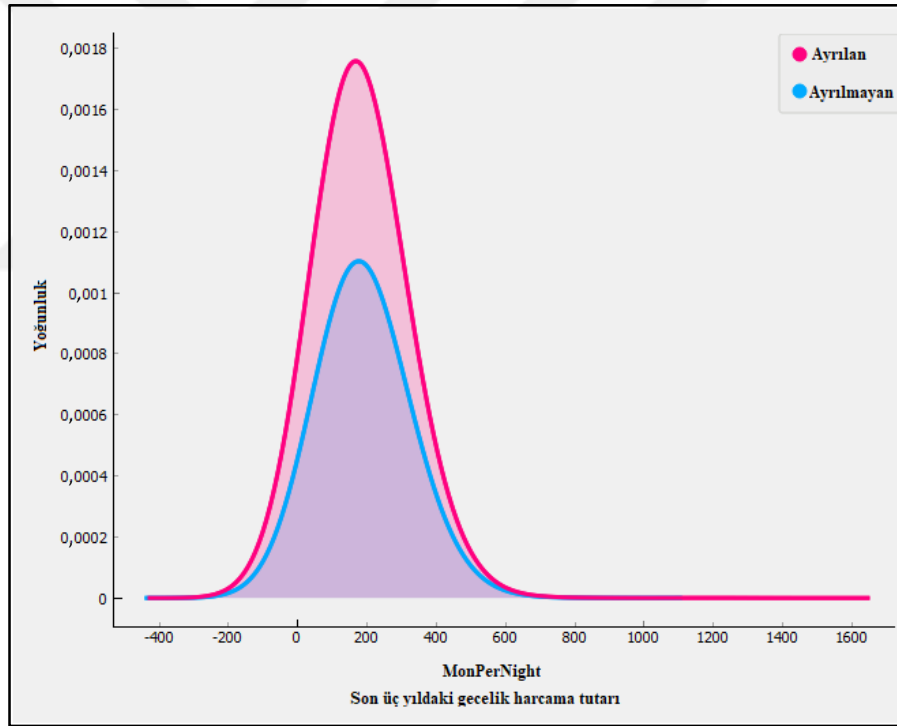


Grafik 3.4 Gerçek Tutar Değişkeni Dağılım Grafiği

Diğer taraftan, gerçek tutar değişkeninin aldığı değerler 61,05 \$ ve 61.212,25 \$ arasında değişmekte olup genel ortalaması 4.580,44 dolardır ($\sigma = 4.925,6$). Bu değer ayrılmayan

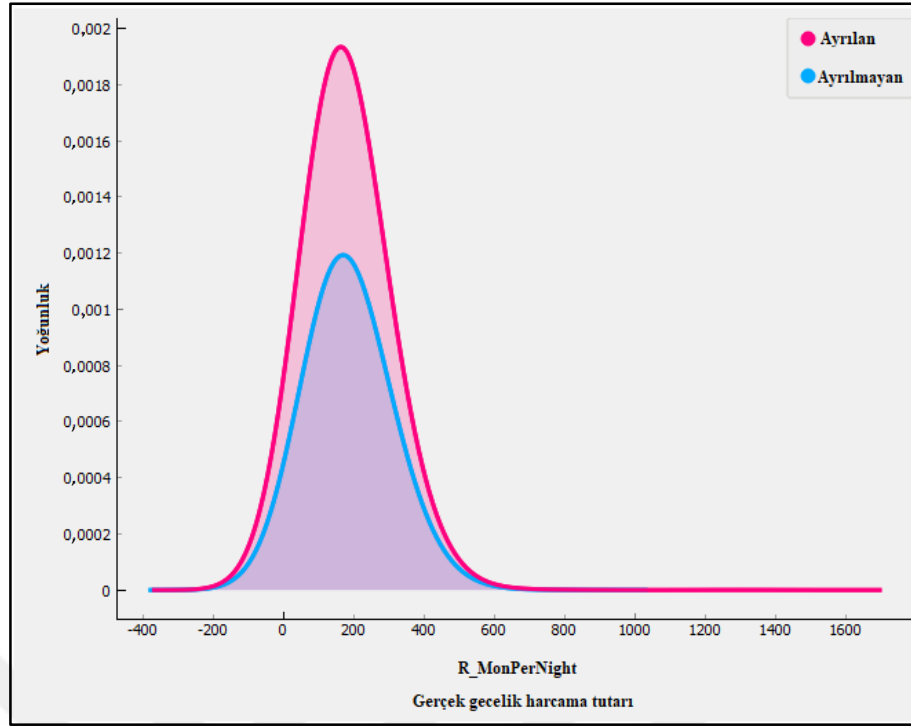
misafirler için ($\bar{X} = 5.815,04$ \$) ayrılan misafirlere ($\bar{X} = 3.805,75$ \$) ve genel ortalamaya kıyasla daha yüksektir. Gerçek tutar değişkeninin dağılım grafiği Grafik 3.4'de sunulmuştur.

Otel işletmelerinde müşterinin geldiği sezona ve konakladığı oda tipine göre ödediği tutar değişiklik gösterebilmektedir. Bu nedenle “Monetary” değişkeninin 2013, 2014 ve 2015 periyodunda gerçekleştirilen toplam geceleme sayısına bölünmesiyle hesaplanan son üç yıl için gecelik harcama tutarı (MonPerNight) ile “R_Monetary” değişkeninin müşterinin tüm gecelemleri toplamına bölünmesiyle hesaplanan gerçek gecelik harcama tutarı da (R_MonPerNight) çalışmaya dahil edilmiştir. Son üç yıl için hesaplanan gecelik harcama tutarı değişkeninin aldığı değerler 11,10 \$ ve 1.200,00 \$ arasında değişmekte olup genel ortalaması 185,81 dolardır ($\sigma = 87,1$). Bu değer ayrılmayan misafirler için ($\bar{X} = 190,79$ \$) ayrılan misafirlere ($\bar{X} = 182,69$ \$) ve genel ortalamaya kıyasla daha yüksektir. Son üç yıl için hesaplanan gecelik harcama tutarı değişkeninin dağılım grafiği Grafik 3.5'de sunulmuştur.



Grafik 3.5 Son Üç Yıldaki Gecelik Harcama Tutarı Değişkeni Dağılım Grafiği

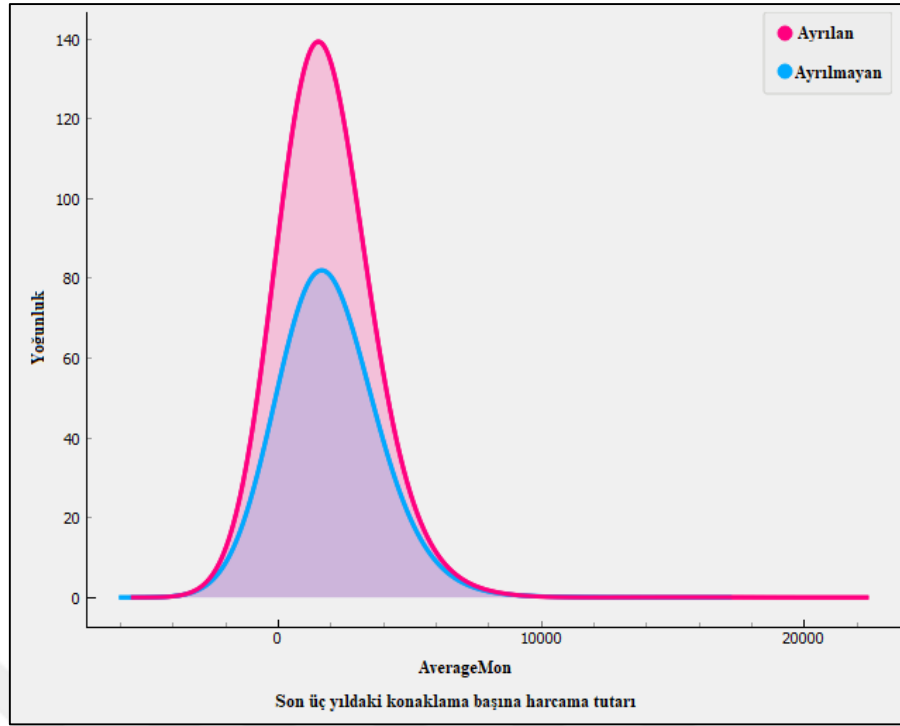
Diğer taraftan tüm konaklamalar için hesaplanan gerçek gecelik harcama tutarı değişkeninin aldığı değerler 24,79 \$ ve 1.305,51 \$ arasında değişmekte olup genel ortalaması 181,89 dolardır ($\sigma = 84,7$). Bu değer ayrılmayan misafirler için ($\bar{X} = 185,13$ \$) ayrılan misafirlere ($\bar{X} = 179,85$ \$) ve genel ortalamaya kıyasla daha yüksektir. Gerçek gecelik harcama tutarı değişkeninin dağılım grafiği Grafik 3.6'da sunulmuştur. Gecelik harcama değişkeninin yanısıra konaklama başına yapılan ödemeler konaklama süresine göre değişiklik gösterebilmektedir.



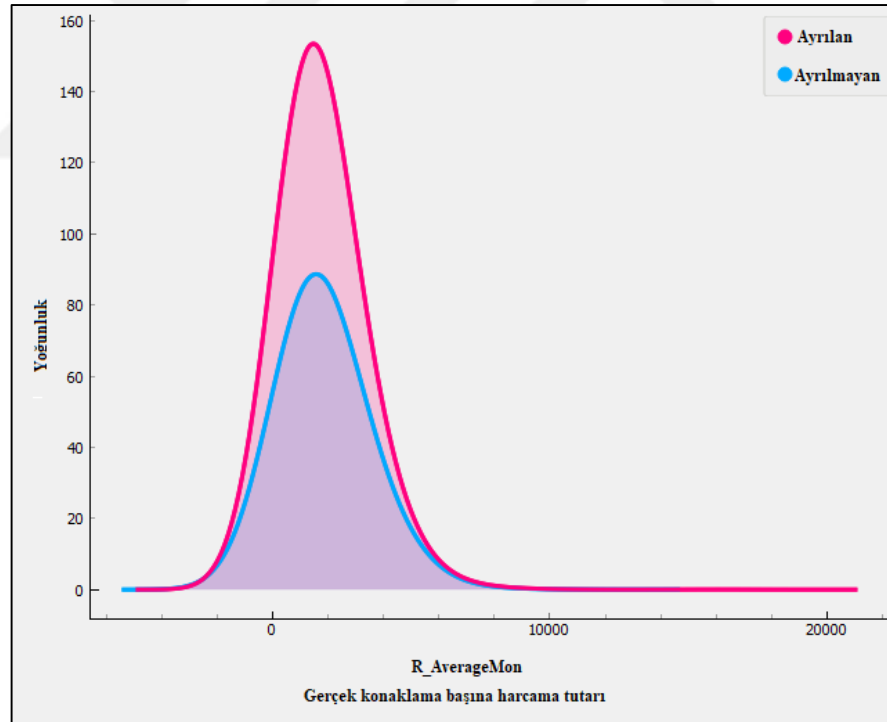
Grafik 3.6 Gerçek Gecelik Harcama Tutarı Değişkeni Dağılım Grafiği

Bu nedenle “Monetary” değişkeninin 2013, 2014 ve 2015 periyodunda gerçekleştirilen toplam geliş sayısına bölünmesiyle hesaplanan ortalama konaklama başına harcama tutarı (AverageMon) ile “R_Monetary” değişkeninin müşterinin tüm konaklamaları için toplam geliş sayısına bölünmesiyle hesaplanan tüm konaklamalar için ortalama konaklama başına harcama tutarı (R_AverageMon) değişkeni de çalışmaya dahil edilmiştir. Son üç yıl için hesaplanan konaklama başına harcama tutarı değişkeninin aldığı değerler 55,50 \$ ve 16.800,00 \$ arasında değişmekte olup genel ortalaması 1.860,82 dolardır ($\sigma = 1.231,7$). Bu değer ayrılmayan misafirler için ($\bar{X} = 1.930,50$ \$) ayrılan misafirlere ($\bar{X} = 1.817,09$ \$) ve genel ortalamaya kıyasla daha yüksektir. Son üç yıl için hesaplanan konaklama başına harcama tutarı değişkeninin dağılım grafiği Grafik 3.7’de sunulmuştur. Diğer taraftan tüm konaklamalar için hesaplanan gerçek konaklama başına harcama tutarı değişkeninin aldığı değerler 105,8 \$ ve 16.101,33 \$ arasında değişmekte olup genel ortalaması 1.816,18 dolardır ($\sigma = 1.160,54$). Bu değer ayrılmayan misafirler için ($\bar{X} = 1.863,06$ \$) ayrılan misafirlere ($\bar{X} = 1.786,76$ \$) ve genel ortalamaya kıyasla daha yüksektir. Gerçek konaklama başına harcama tutarı değişkeninin dağılım grafiği Grafik 3.8’de sunulmuştur.

RFM değişkenlerinin bir diğeri olan sıklık değişkeni müşterinin satın alım sayısını ifade etmekte olup müşterinin gelecekteki beklenen satın alım davranışını olumlu şekilde yordaması nedeniyle ayrılma davranışı için de anlamlı bir yordayıcı olarak görülmektedir (Buckinx ve Van den Poel; 2005: 260).



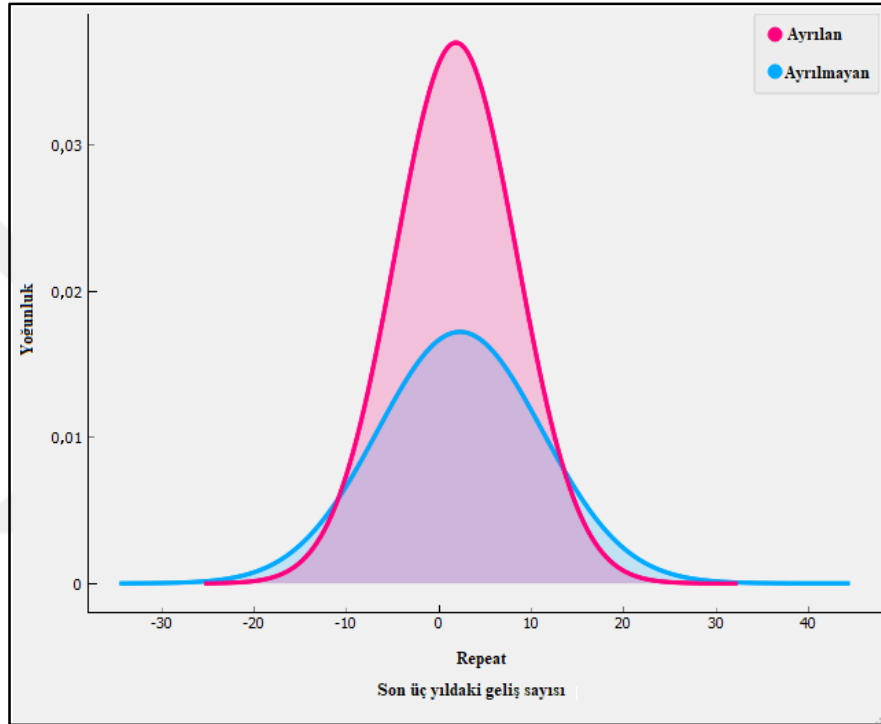
Grafik 3.7 Son Üç Yıldaki Konaklama Başına Harcama Tutarı Dağılım Grafiği



Grafik 3.8 Gerçek Konaklama Başına Harcama Tutarı Dağılım Grafiği

Çalışma kapsamında sıklık değişkeninin hesaplanmasında tutar değişkeninde olduğu gibi Buckinx ve Van den Poel'in (2005: 260) yöntemi kullanılmıştır. Bu bağlamda 2013, 2014 ve 2015 yıllarında müşterinin gerçekleştirdiği tüm konaklamaları için toplam geliş sayısı "Repeat" değişkeni olarak adlandırılmış, müşterinin ilk konaklamasından gözlem periyodunun

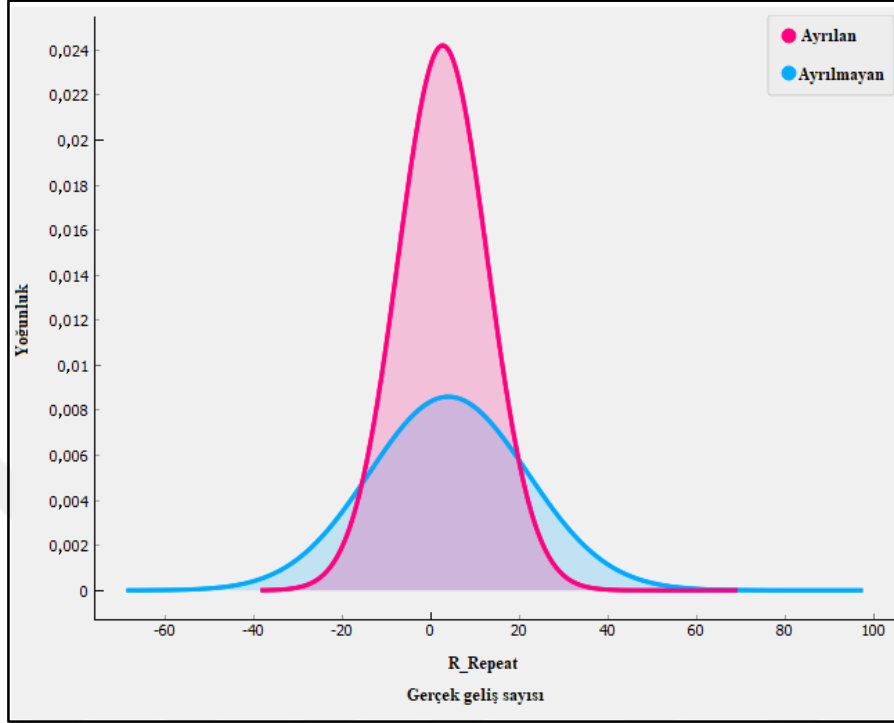
son tarihi olan 31/12/2015'e kadar geçen süredeki tüm konaklamaları için (ilişki süresi değişkeni göz önünde bulundurularak) toplam geliş sayısı "R_Repeat" değişkeni olarak adlandırılmıştır. Dolayısıyla "Repeat" değişkeni son üç yıldaki geliş sayısını, "R_Repeat" değişkeni ise gerçek geliş sayısını ifade etmektedir. Son üç yıl için hesaplanan geliş sayısı değişkeninin aldığı değerler 1 ve 9 defa arasında değişmekte olup genel ortalaması 2 defadır ($\sigma = 0,9$). Bu değer ayrılmayan misafirler için ($\bar{X} = 2,3$) ayrılan misafirlere ($\bar{X} = 1,9$) ve genel ortalamaya kıyasla daha yüksektir. Son üç yıldaki geliş sayısı değişkeninin dağılım grafiği Grafik 3.9'da sunulmuştur.



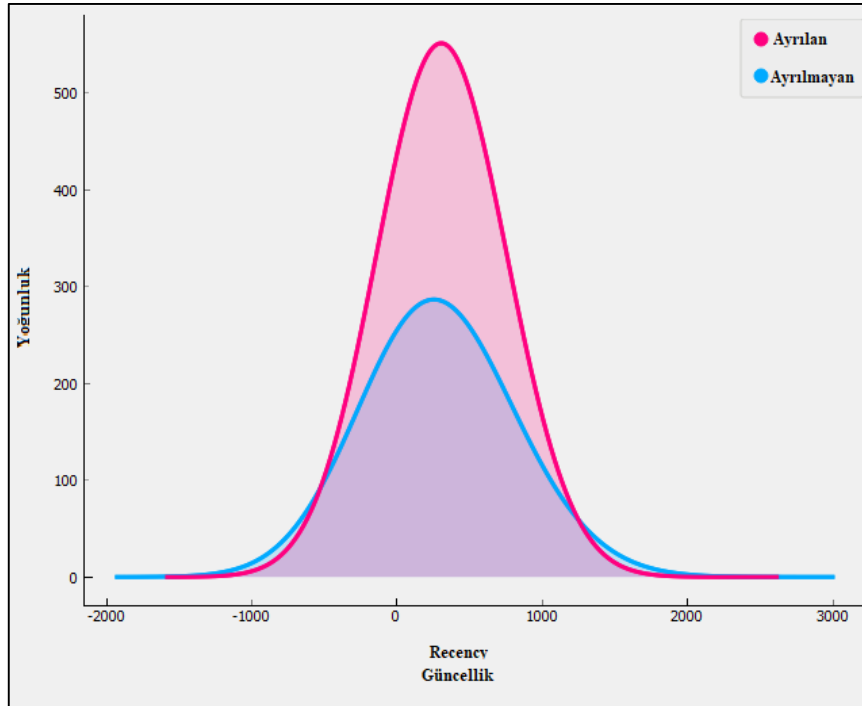
Grafik 3.9 Son Üç Yıldaki Geliş Sayısı Dağılım Grafiği

Diğer taraftan tüm konaklamalar için hesaplanan gerçek geliş sayısı değişkeninin aldığı değerler 2 ve 29 defa arasında değişmekte olup genel ortalaması 3,3 defadır ($\sigma = 2,3$). Bu değer ayrılmayan misafirler için ($\bar{X} = 4,2$) ayrılan misafirlere ($\bar{X} = 2,8$) ve genel ortalamaya kıyasla daha yüksektir. Gerçek geliş sayısı değişkeninin dağılım grafiği Grafik 3.10'da sunulmuştur. RFM değişkenlerinin sonuncusu olan güncellik değişkeni müşterinin son gelişinden gözlem periyodunun sonuna kadar geçen toplam gün sayısını ifade etmektedir. Güncel çalışmaların bulgularına göre güncellik değeri ne kadar düşükse müşterinin sadık olma olasılığı o derecede yüksek olmaktadır (Buckinx ve Van den Poel; 2005: 259). Çalışmada kullanılan güncellik değişkeni (Recency) nümerik olup, düşük değerler yakın zamandaki konaklamaları yüksek değerler ise daha eski tarihlerde gerçekleşen konaklamaları ifade etmektedir. Güncellik değişkeninin aldığı değerler 1 ve 1078 gün arasında değişmekte olup genel ortalaması 305,9

gündür ($\sigma = 210,1$). Bu değer ayrılan misafirler için ($\bar{X} = 314,4$) ayrılmayan misafirlere ($\bar{X} = 292,5$) ve genel ortalamaya kıyasla daha yüksektir. Güncellik değişkeninin dağılım grafiği Grafik 3.11’de sunulmuştur.

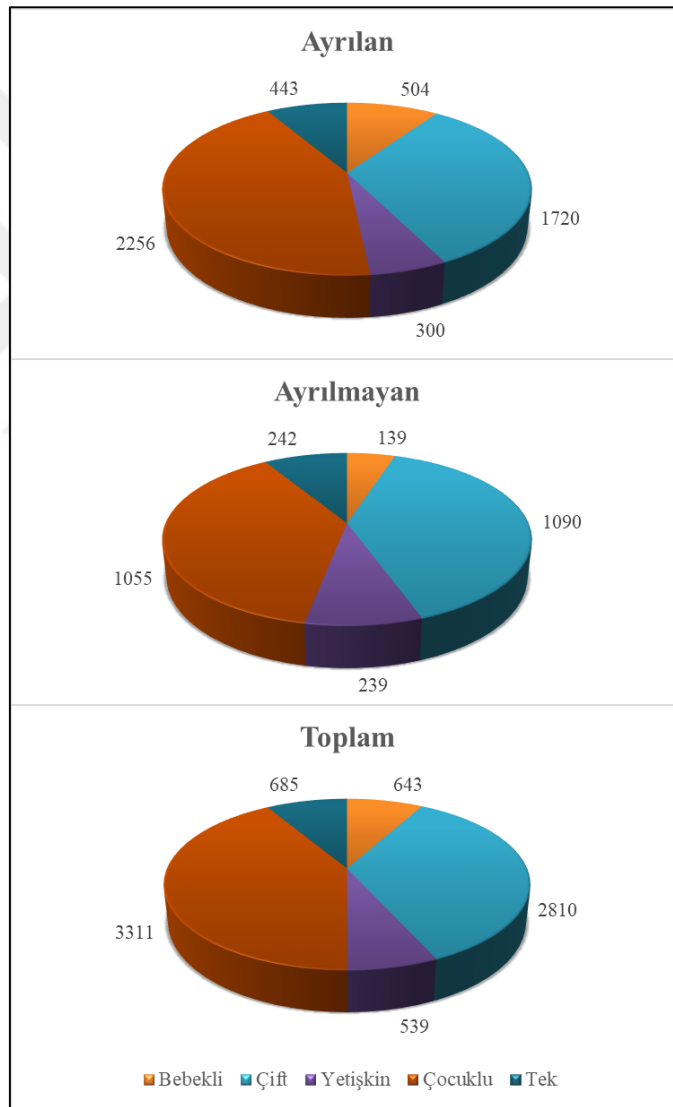


Grafik 3.10 Gerçek Geliş Sayısı Dağılım Grafiği



Grafik 3.11 Güncellik Değişkeni Dağılım Grafiği

Konaklama eşlikçileri değişkenleri oda için hesap açılan müşterinin konaklaması süresince aynı odada birlikte kaldığı diğer müşterileri ifade etmektedir. Bu değişken De Wit'in (2017: 59) çalışmasında yetişkin sayısı ve çocuk sayısı olarak kullanılmış, yetişkin sayısı değişkeni ayrılma davranışını yordamada anlamlı bulunmuştur. Bu çalışmada kullanılan konaklama eşlikçileri değişkeni için çocuklu, bebekli, yetişkin grubu halinde, tek (single) konaklama ve çift olarak konaklama durumları ikili (binary) kodlanarak ayrı değişkenler halinde çalışmaya dahil edilmiş, oda paylaşımı (room share) durumu dahil edilmemiştir. Ayrıca bebekli ve çocuklu konaklayan müşteriler için ayrı bir kategori oluşturulmamış, ikili kodlamada bu müşteriler için bebekli konaklama ve çocuklu konaklama kategorilerinin ikisi de 1 kodlanmıştır.

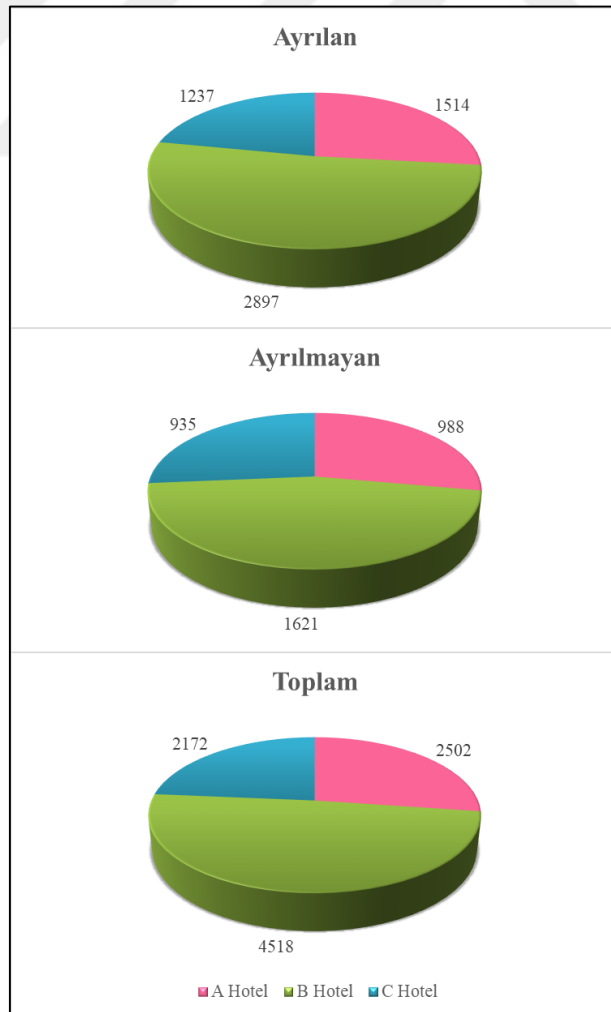


Grafik 3.12 Konaklama Eşlikçileri Değişkenleri Dağılım Grafiği

Bebekli konaklayan müşterilerin ayrılan müşteriler içerisindeki oranı %9 (f=504), ayrılmayan müşterilerde ise % 4 (f=139) olarak ortaya çıkmış olup toplam müşteriler içerisinde %7'dir (f=643). Benzer şekilde çocuklu konaklayan müşterilerin ayrılan müşteriler içerisindeki

oranı (%40, f=2256) ayrılmayan müşterilere oranla daha fazla (%30, f=139) olup toplam müşteriler içerisindeki oranı %36'dır (f=3311). Ayrıca, tek konaklayan müşterilerin ayrılan müşteriler içerisindeki oranı %8 (f=242) olmakla birlikte ayrılmayan müşteriler içerisindeki oranı daha düşük (%7, f=242), toplam müşteriler içerisindeki oranı ise %7'dir (f=685). Diğer taraftan yetişkin grubu halinde konaklayan müşterilerin ayrılmayan müşteriler içerisindeki oranı (%7, f=239), ayrılan müşterilere kıyasla (%5, f=300) daha yüksek, toplam müşteri grubu içindeki oranı %6'dır (f=539). Benzer şekilde, çift olarak konaklayan müşterilerin ayrılmayan müşteriler içerisindeki oranı %31 (f=1090), ayrılan müşterilerde ise %30 (f=1720) olarak ortaya çıkmış olup, toplam müşteriler içerisindeki oranı %31'dir (f=2810). Konaklama eşlikçileri değişkenlerinin dağılımları Grafik 3.12'de sunulmuştur.

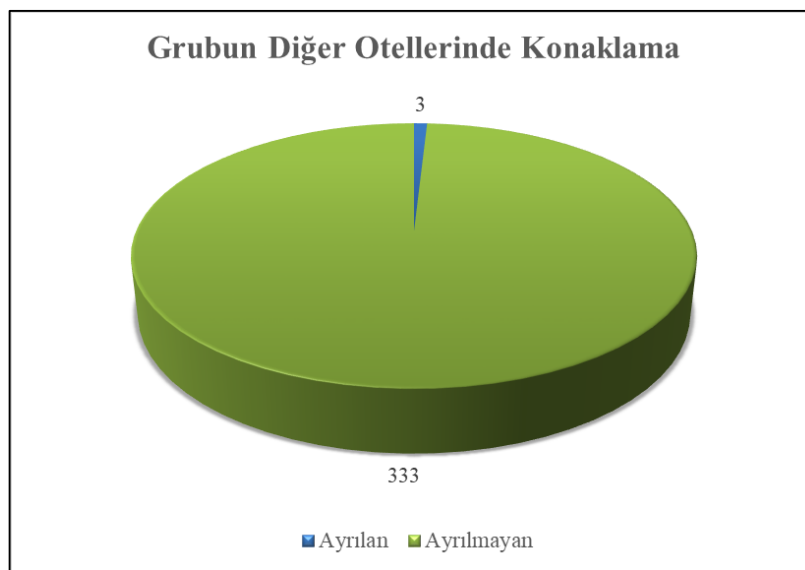
Konaklama yapılan otel değişkeni, çalışmaya konu olan otel zincirinin grup otellerini ifade etmektedir. Müşterilerin son konakladığı otel ve daha önce konaklama yaptığı otel farklılık gösterebileceği gibi, tekrar satın alım yapma kararı üzerinde de etkili olabilir. Bu nedenle son konaklama yapılan otel değişkeniyle birlikte diğer otellerdeki konaklama durumu da çalışmaya dahil edilmiştir.



Grafik 3.13 Konaklama Yapılan Otel Değişkenleri Dağılım Grafiği

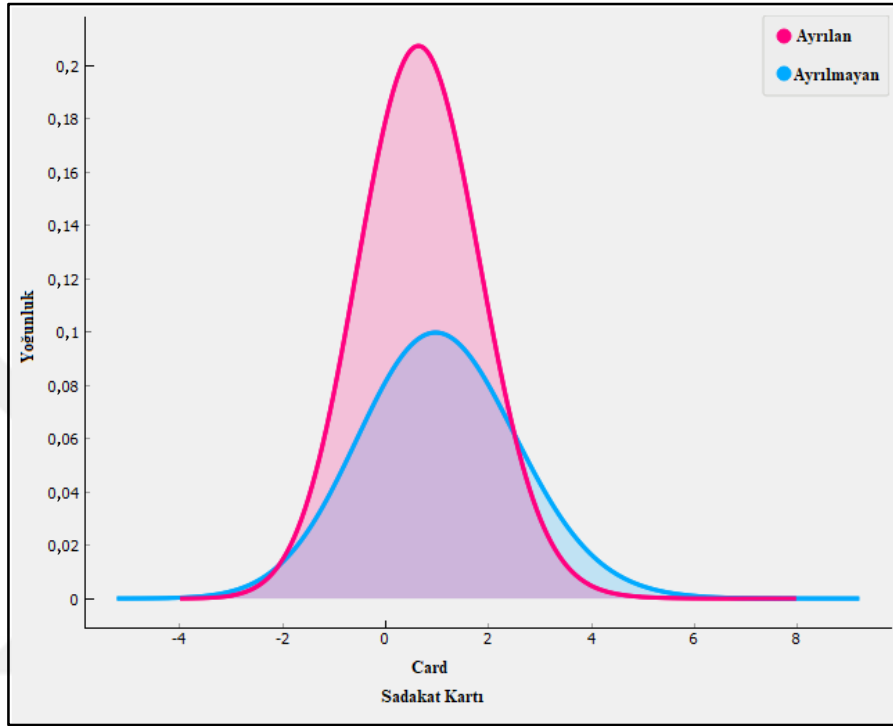
Bu bağlamda, son konaklanılan otel id kodunun A Hotel, B Hotel veya C Hotel olması durumu ile tüm konaklamaları göz önünde bulundurulduğunda grubun diğer otellerinde konaklamasının bulunup bulunmadığı durumları ikili kodlanarak ayrı değişkenler halinde çalışmaya dahil edilmiştir. Son konaklaması A Hotel’de gerçekleşen müşterilerin ayrılmayan müşteriler içerisindeki oranı (%28, f=988) ayrılan müşterilere kıyasla (%27, f=1514) daha yüksek olup toplam müşteriler içerisindeki oranı %27’dir (f=2502). Benzer şekilde son konaklaması C Hotel’de gerçekleşen müşterilerin ayrılmayan müşteriler içerisindeki oranı % 26 (f=935), ayrılan müşterilerde ise %22 (f=1237) olarak ortaya çıkmış olup, toplam müşteriler içerisindeki oranı %24’tür (2172). Diğer taraftan son konaklaması B Hotel’de gerçekleşen müşterilerin ayrılan müşteriler içerisindeki oranı %51(f=2897), ayrılmayan müşteriler arasındaki oranı ise %46 (f=1621) olarak ortaya çıkmış olup, toplam müşteriler arasındaki oranı ise %49’dur (f=4518). Son konaklanan otel değişkenlerinin dağılımları Grafik 3.13’de sunulmuştur.

Konaklama yapılan diğer otellerin dağılımları incelendiğinde, grubun diğer otellerinde konaklaması bulunan 336 müşterinin (toplam müşterilerin %4’ü), 333 tanesi ayrılma davranışında bulunmazken (ayrılmayan müşteriler içerisindeki oranı %9), sadece 3 tanesi ayrılma davranışında bulunmuştur. Ayrılma davranışında bulunan müşterilerin ikisi A ve B Hotel’de konaklama gerçekleştirmiş olup (son konaklamaları B Hotel’de), diğeri A ve C Hotelde konaklamıştır (son konaklaması C Hotel’de). Ayrılma davranışında bulunmayan 333 müşterinin ise %44’ü (f=145) A ve C Hotel’de, %24’ü (f=79) A ve B Hotel’de, %20’si (f=66) B ve C Hotel’de, %13’ü (f=43) ise grubun tüm otellerinde konaklamıştır. Grubun diğer otellerinde konaklama yapmış olma değişkeninin dağılımları Grafik 3.14’de sunulmuştur.



Grafik 3.14 Grubun Diğer Otellerinde Konaklama Değişkeni Dağılım Grafiği

Çalışmaya konu olan otel zincirinin müşterilerine uyguladığı bir sadakat programı mevcuttur. Bu sadakat programı kapsamında otel zincirinin otellerinde en az iki defa konaklama yapan ve toplam geceleme sayısı en az 20 olan müşteriler Silver müşteri; 50 ve üstü olan müşteriler Silver Plus müşteri; 120 ve üstü olan müşteriler Gold müşteri; 200 ve üstü olan müşteriler ise Platinum müşteri olarak anılmaktadır.



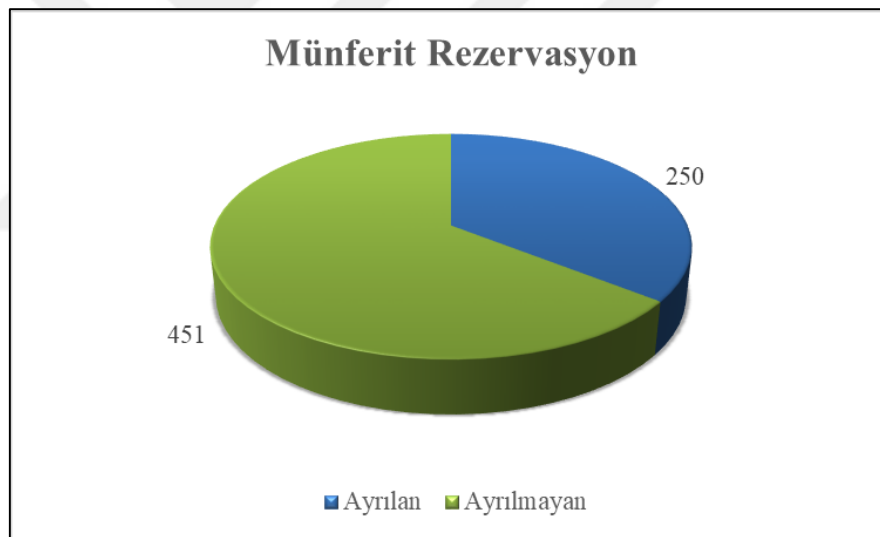
Grafik 3.15 Sadakat Kartı Değişkeni Dağılım Grafiği

Müşterilerin üyelik düzeyi yükseldikçe kendilerine sunulan imkân ve teşvikler de aynı düzede artmaktadır. Bu kapsamda üye müşterilere özel hoşgeldiniz ikramları, SPA merkezinde indirimler, banyo bukletlerinde özel sürprizler, birlikte gelen arkadaşları eşit üyelik avantajlarından faydalandırma, ücretsiz bir üst kategori odada konaklama, erken giriş – geç çıkış yapma vb. birçok imkân sunulmaktadır. Balton ve Kannan (2000: 95) tarafından gerçekleştirilen bir çalışmaya göre üye müşteriler işletmeye ödedikleri paranın karşılığını aldıklarını düşünmekte, daha kaliteli ve daha iyi hizmet aldıklarını düşünmekte ve dolayısıyla yeniden satın alma niyetleri daha yüksek olmaktadır. Bu bağlamda, müşterinin sadakat programı üyesi olup olmadığı durumu da çalışmaya dahil edilmiştir. Sadakat kartı değişkeni ordinal olup, kartı olmayan müşteriler için “0”, Silver müşteriler için “1”, Silver Plus müşteriler için “2”, Gold müşteriler için “3” ve Platinum müşteriler için “4” olarak kodlanmıştır. Müşterilerin üyelik durumları incelendiğinde %34,9’unun (f=3212) kartının bulunmadığı, %49,8’inin (f=4579) Silver müşteri olduğu, %13,1’inin (f=1202) Silver Plus müşteri olduğu, %1,6’sının (f=147) Gold müşteri olduğu, %0,6’sının (52) ise Platinum müşteri olduğu

görülmektedir. Dolayısıyla, sadakat kartı değişkeninin aldığı değerler 0 ve 4 arasında değişmekte olup genel ortalaması 0,8'dir ($\sigma = 0,8$).

Bu değer ayrılan misafirler için ($\bar{X} = 0,7$) ayrılmayan misafirlere ($\bar{X} = 1,1$) ve genel ortalamaya kıyasla daha düşüktür. Sadakat kartı değişkeninin dağılım grafiği Grafik 3.15'de sunulmuştur.

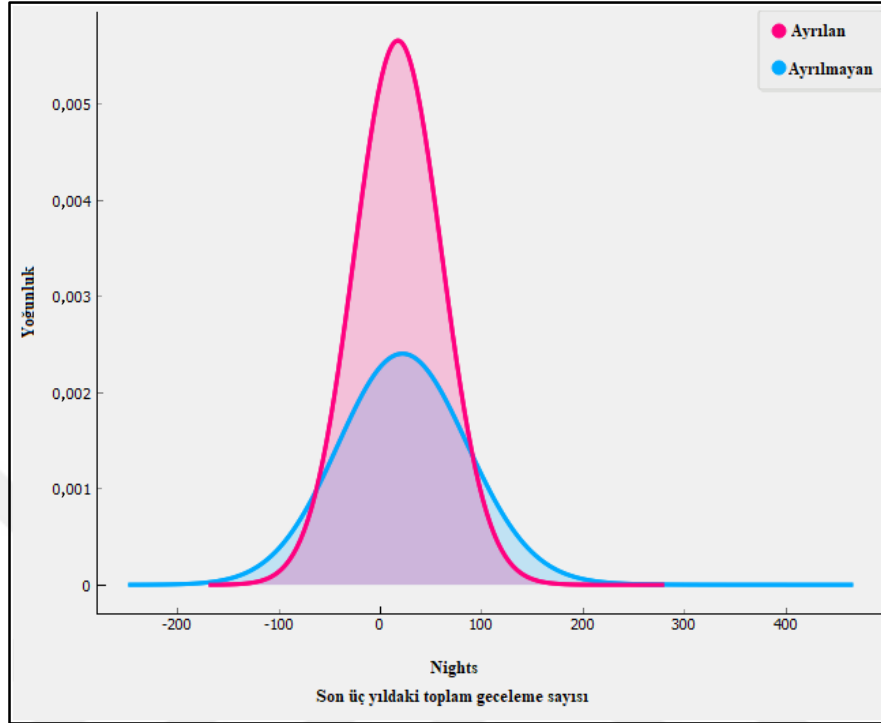
Müşteriler rezervasyonlarını acente ve tur operatörü üzerinden ya da münferit (otele direkt) olarak gerçekleştirebilir. De Wit (2017: 43) tarafından gerçekleştirilen çalışmada benzer bir değişken olan “rezervasyonları gerçekleştiren - Booker” değişkeni (şirket, müşteri ya da acente) ayrılma davranışını yordamada anlamlı bulunmuştur. Bu bağlamda, müşterinin münferit satın alım yöntemini benimsemesi durumu ikili kodlanarak çalışmaya dahil edilmiştir. Satın alım yöntemi değişkeni incelendiğinde, münferit satın alım yöntemini tercih edenlerin ayrılmayan müşteriler içerisindeki oranı %13 (f=451) olup, ayrılan müşteriler içerisindeki oranı daha düşük (%4, f=250) ve toplam müşteriler içerisindeki oranı %8'dir (f=701). Satın alım yöntemi değişkeninin dağılımları Grafik 3.16'da sunulmuştur.



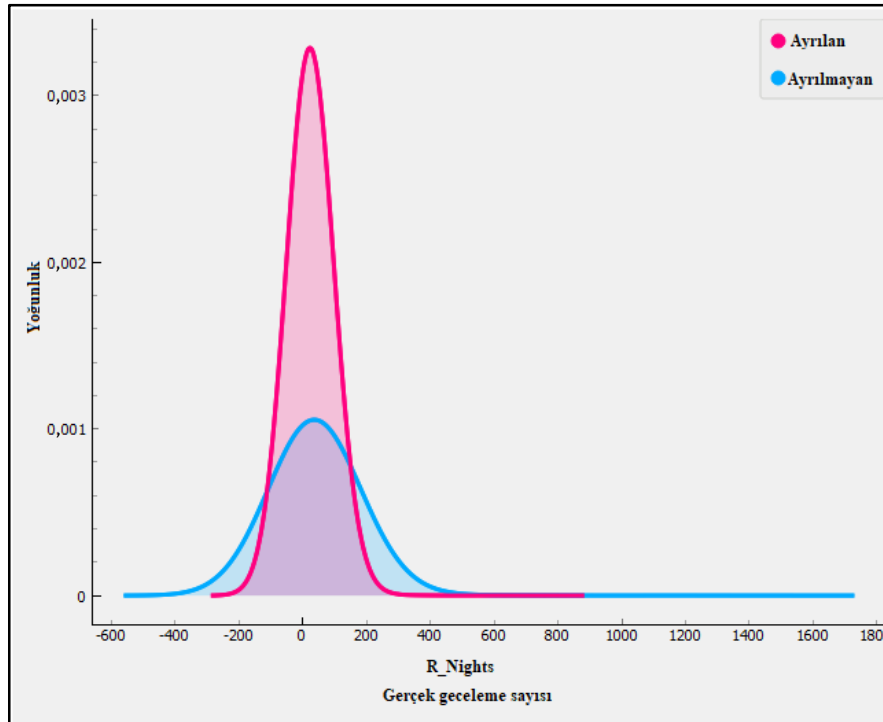
Grafik 3.16 Satın Alım Yöntemi Değişkeni Dağılım Grafiği

Otel işletmelerinde her müşteri için konaklama süreleri farklılık gösterebileceği gibi aynı müşterinin her gelişindeki konaklama süresi de değişiklik gösterebilir. De Wit (2017: 43) tarafından gerçekleştirilen çalışmada konaklama süresi değişkeni ayrılma davranışını yordamada anlamlı bulunmuştur. Dolayısıyla, müşterinin son üç yılda gerçekleştirdiği toplam geceleme sayısı (Nights) ile tüm konaklamaları için toplam geceleme sayısı (R_Nights) değişkenleri de bu çalışmaya dahil edilmiştir. Son üç yıldaki toplam geceleme sayısı değişkeninin aldığı değerler 1 ve 218 gece arasında değişmekte olup genel ortalaması 20 gecedir ($\sigma = 13,4$). Bu değer ayrılmayan misafirler için ($\bar{X} = 23,5$ gece) ayrılan misafirlere ($\bar{X} = 18,3$ gece) ve genel ortalamaya kıyasla daha yüksektir. Son üç yıl için hesaplanan toplam

geceleme sayısı değişkeninin dağılım grafiği Grafik 3.17’de sunulmuştur. Diğer taraftan tüm konaklamalar için hesaplanan gerçek geceleme sayısı değişkeninin aldığı değerler 10 ve 1163 gece arasında değişmekte olup genel ortalaması 33,2 gecedir ($\sigma = 34,8$).

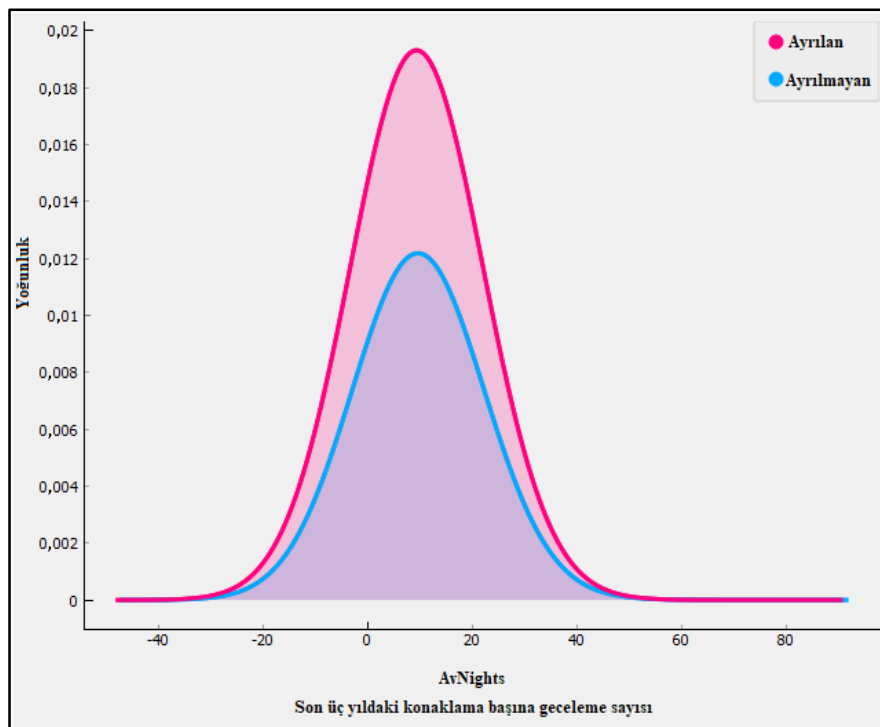


Grafik 3.17 Son Üç Yıldaki Geceleme Sayısı Dağılım Grafiği

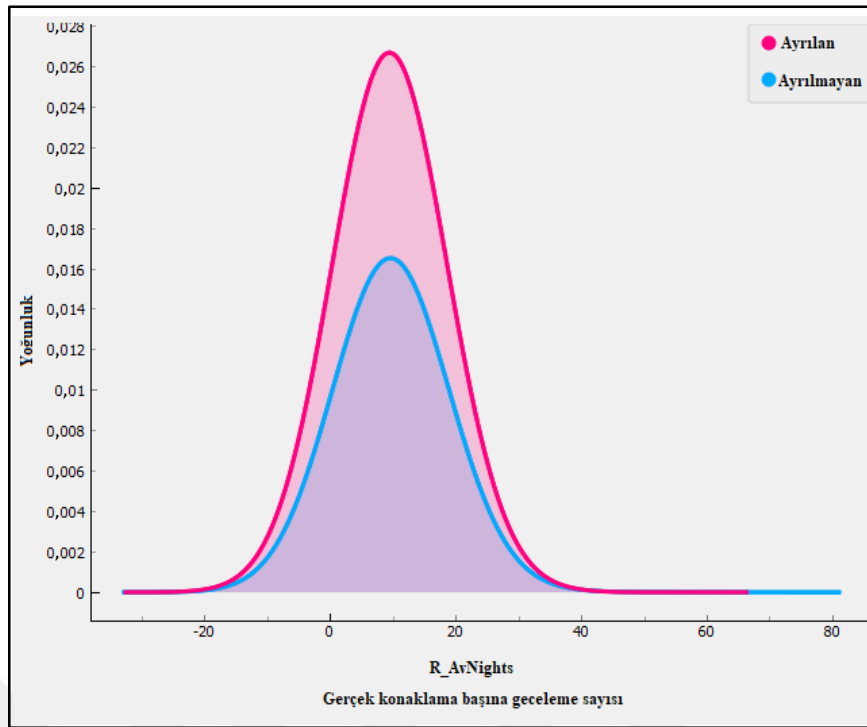


Grafik 3.18 Gerçek Geceleme Sayısı Dağılım Grafiği

Bu değer ayrılmayan misafirler için ($\bar{X} = 42,9$) ayrılan misafirlere ($\bar{X} = 27,1$) ve genel ortalamaya kıyasla daha yüksektir. Tüm konaklamalar için hesaplanan gerçek geceleme sayısı değişkeninin dağılım grafiği Grafik 3.18’de sunulmuştur. Ayrıca, müşterinin konaklama başına gerçekleştirdiği ortalama geceleme sayısının daha önemli bir yordayıcı olarak ortaya çıkma olasılığı göz önünde bulundurularak, “Nights” değişkeninin 2013, 2014 ve 2015 periyodunda gerçekleştirilen toplam geliş sayısına bölünmesiyle hesaplanan konaklama başına geceleme sayısı (AvNights) ile “R_Nights” değişkeninin müşterinin tüm konaklamaları için toplam geliş sayısına bölünmesiyle hesaplanan tüm konaklamalar için konaklama başına geceleme sayısı (R_AvNights) değişkenleri de çalışmaya dahil edilmiştir. Son üç yıldaki konaklama başına geceleme sayısı değişkeninin aldığı değerler 1 ve 43,8 gece arasında değişmekte olup genel ortalaması 9,7 gecedir ($\sigma = 3,6$). Bu değer ayrılmayan misafirler için ($\bar{X} = 9,9$ gece) ayrılan misafirlere ($\bar{X} = 9,6$ gece) ve genel ortalamaya kıyasla daha yüksektir. Son üç yıl için hesaplanan konaklama başına geceleme sayısı değişkeninin dağılım grafiği Grafik 3.19’da sunulmuştur. Diğer taraftan tüm konaklamalar için hesaplanan gerçek konaklama başına geceleme sayısı değişkeninin aldığı değerler 1,4 ve 49,9 gece arasında değişmekte olup genel ortalaması 9,7 gecedir ($\sigma = 3,4$). Bu değer ayrılmayan misafirler için ($\bar{X} = 9,8$ gece) ayrılan misafirlere ($\bar{X} = 9,6$ gece) ve genel ortalamaya kıyasla daha yüksektir. Tüm konaklamalar için hesaplanan gerçek konaklama başına geceleme sayısı değişkeninin dağılım grafiği Grafik 3.20’de sunulmuştur.

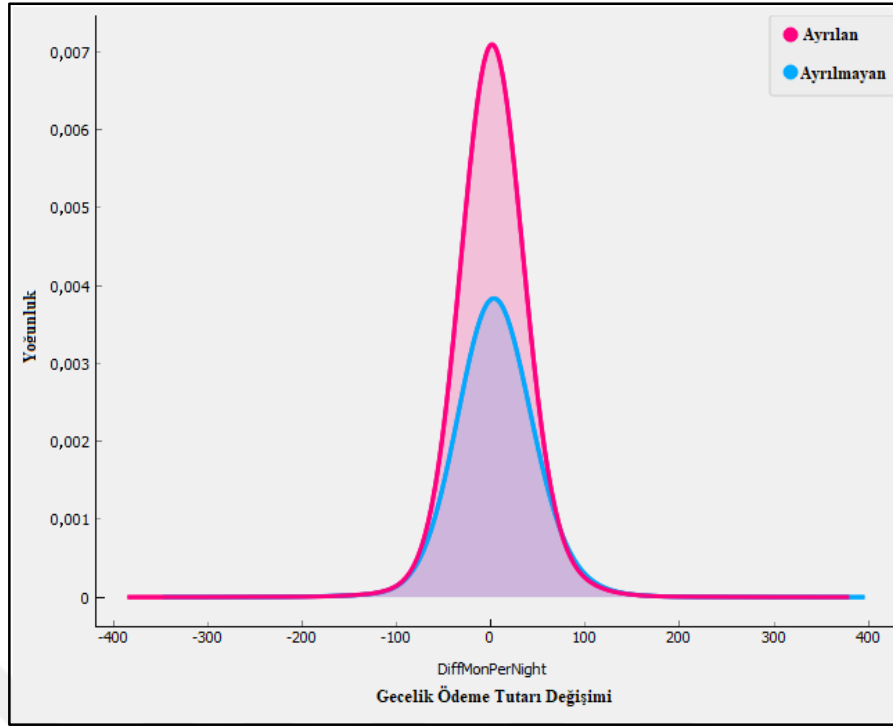


Grafik 3.19 Son Üç Yıldaki Konaklama Başına Geceleme Sayısı Dağılım Grafiği

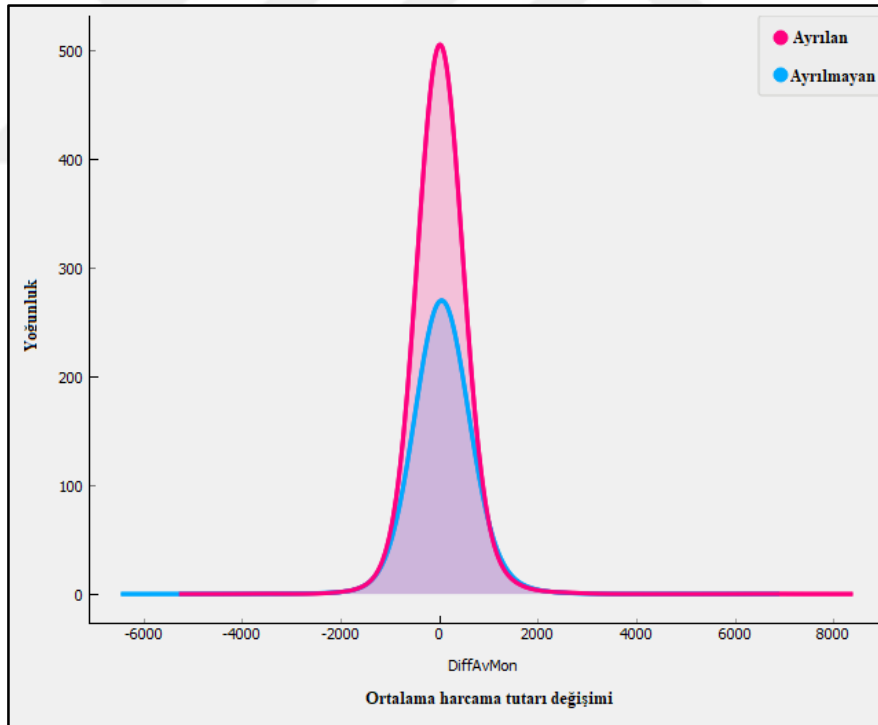


Grafik 3.20 Gerçek Konaklama Başına Geceleme Sayısı Dağılım Grafiği

Miguéis ve arkadaşları (2012b: 342) çalışmalarında müşterilerin son dönemdeki harcamalarının artış ya da azalış eğilimi göstermesi durumunun kayıp müşterileri yordamada etkili olabildiğini göstermiştir. Buradan hareketle müşterilerin son üç yıldaki ortalama geceleme, ortalama harcama ve gecelik ödeme tutarı değişkenlerinin değişimleri hesaplanarak çalışmaya dahil edilmiştir. Bu değişkenler çalışmada nümerik olarak kullanılmış olup, negatif değerler azalan eğilimleri, sıfır değeri değişim olmadığını, pozitif değerler ise artış eğilimini ifade etmektedir. Gecelik ödeme tutarı değişimi değişkeninin aldığı değerler -258,4 \$ ve 253,8 \$ arasında değişmekte olup genel ortalaması 3,9 dolardır ($\sigma = 23,2$). Bu değer ayrılmayan misafirler için ($\bar{X} = 5,7$ \$) ayrılan misafirlere ($\bar{X} = 2,8$ \$) ve genel ortalamaya kıyasla daha yüksektir. Gecelik ödeme tutarı değişimi değişkeninin dağılım grafiği Grafik 3.21’de sunulmuştur. Konaklama başına hesaplanan ortalama harcama tutarı değişimi değişkeninin aldığı değerler ise -4.408,4 \$ ve 6.625,5 \$ arasında değişmekte olup genel ortalaması 44,6 dolardır ($\sigma = 349,2$). Bu değer ayrılmayan misafirler için ($\bar{X} = 67,4$ \$) ayrılan misafirlere ($\bar{X} = 30,3$ \$) ve genel ortalamaya kıyasla daha yüksektir. Ortalama harcama tutarı değişimi değişkeninin dağılım grafiği Grafik 3.22’de sunulmuştur. Ortalama geceleme sayısı değişimi değişkeninin aldığı değerler ise -17,7 ve 15 gece arasında değişmekte olup genel ortalaması 0,01 gecedir ($\sigma = 1,2$). Bu değer ayrılmayan misafirler için ($\bar{X} = 0,1$) ayrılan misafirlere ($\bar{X} = -0,02$) ve genel ortalamaya kıyasla daha yüksektir. Ortalama geceleme sayısı değişimi değişkeninin dağılım grafiği Grafik 3.23’de sunulmuştur.

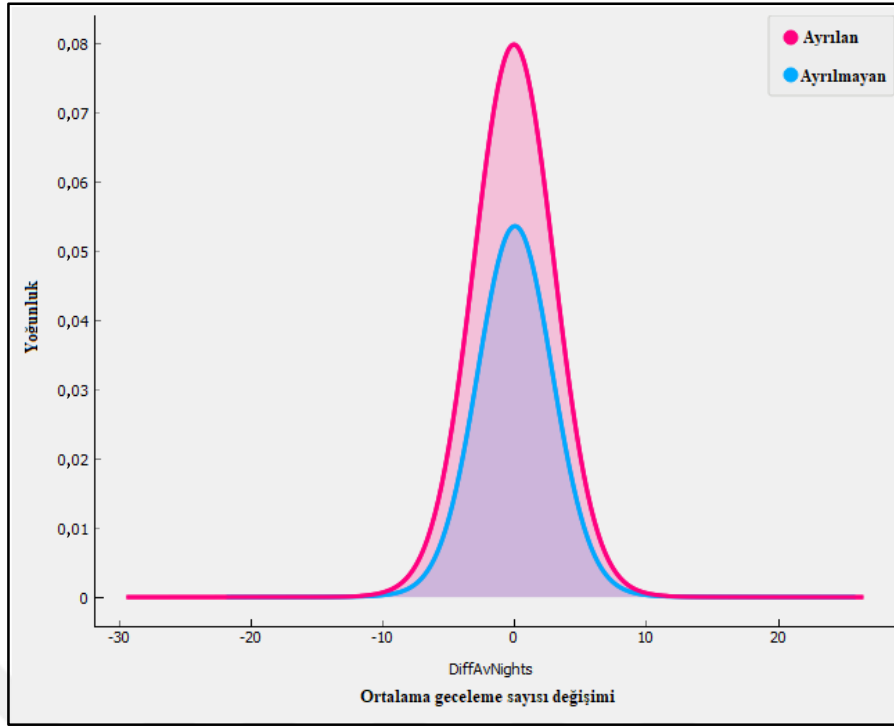


Grafik 3.21 Gecelik Ödeme Tutarı Değişimi Dağılım Grafiği



Grafik 3.22 Ortalama Harcama Tutarı Değişimi Dağılım Grafiği

Konaklama işletmelerinde konaklama yapılan periyodun müşterinin ödediği ücret, otelin doluluğu ve dolayısıyla müşterinin algıladığı hizmet kalitesi üzerinde etkisi olabilir. Burdan hareketle, zincir otel işletmesinin uyguladığı fiyat periyod kırımları ve müşteri yoğunluğu göz önünde bulundurularak son konaklamanın gerçekleştiği ay üzerinden sezon değişkeni hesaplanmıştır.



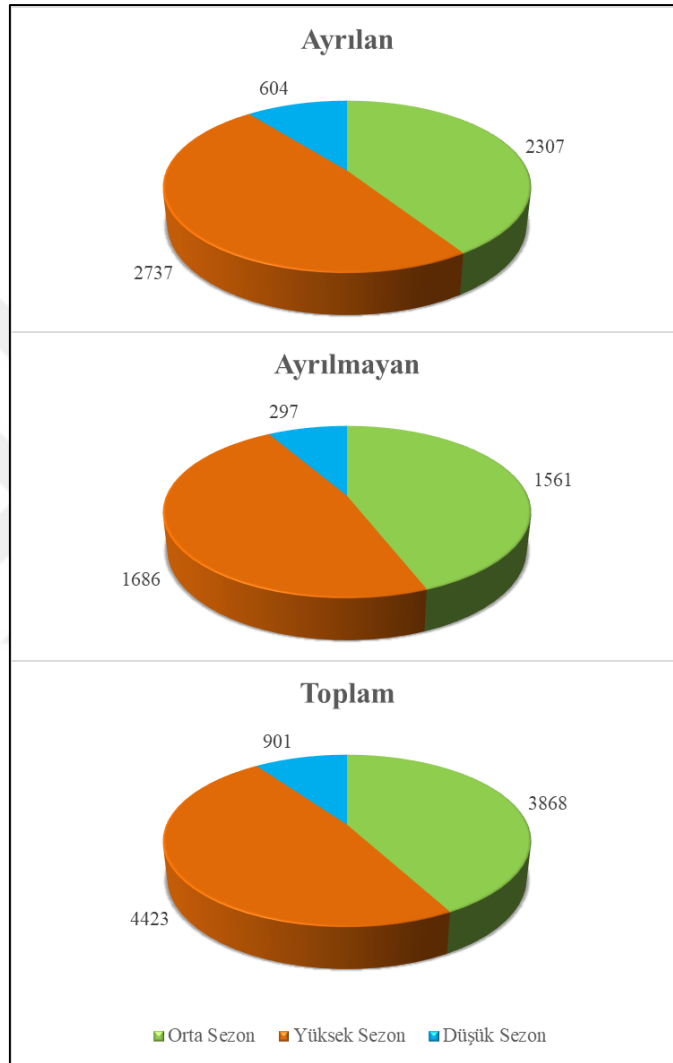
Grafik 3.23 Ortalama Geceleme Sayısı Değişimi Dağılım Grafiği

Sezon değişkeni yüksek, orta ve düşük olmak üzere üç kategori olarak belirlenmiştir (Tablo 3.2). Bu kategoriler ikili kodlanarak farklı değişkenler halinde çalışmaya dahil edilmiştir. Yüksek sezonda konaklayan müşterilerin ayrılan müşteriler içerisindeki oranı %48 ($f=2737$) olup, ayrılmayan müşteriler (%48, $f=1686$) ve toplam müşteriler içerisindeki oranıyla aynıdır (%48, $f=4423$).

Tablo 3.2 Sezon Değişkeni İçin Kullanılan Kategoriler

Ay	Sezon
Ocak	Düşük sezon
Şubat	Düşük sezon
Mart	Düşük sezon
Nisan	Orta sezon
Mayıs	Orta sezon
Haziran	Yüksek sezon
Temmuz	Yüksek sezon
Ağustos	Yüksek sezon
Eylül	Orta sezon
Ekim	Orta sezon
Kasım	Düşük sezon
Aralık	Düşük sezon

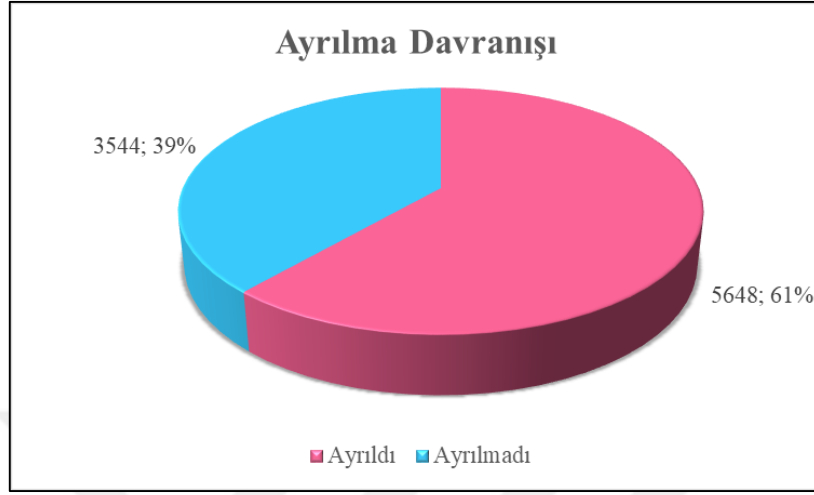
Diğer taraftan orta sezonda konaklayan müşterilerin ayrılmayan müşteriler içerisindeki oranı (%44, f=2256) ayrılan müşteriler içerisindeki oranına kıyasla daha fazla (%41, f=2307) olup toplam müşteriler içerisindeki oranı %42'dir (f=3868). Buna karşılık, düşük sezonda konaklayan müşterilerin ayrılan müşteriler içerisindeki oranı %11 (f=604) ayrılmayan müşteriler içerisindeki oranından daha yüksek (%8, f=297), toplam müşteriler içerisindeki oranı ise %10'dur (f=901). Sezon değişkeninin dağılımları Grafik 3.24'de sunulmuştur.



Grafik 3.24 Sezon Değişkeni Dağılım Grafiği

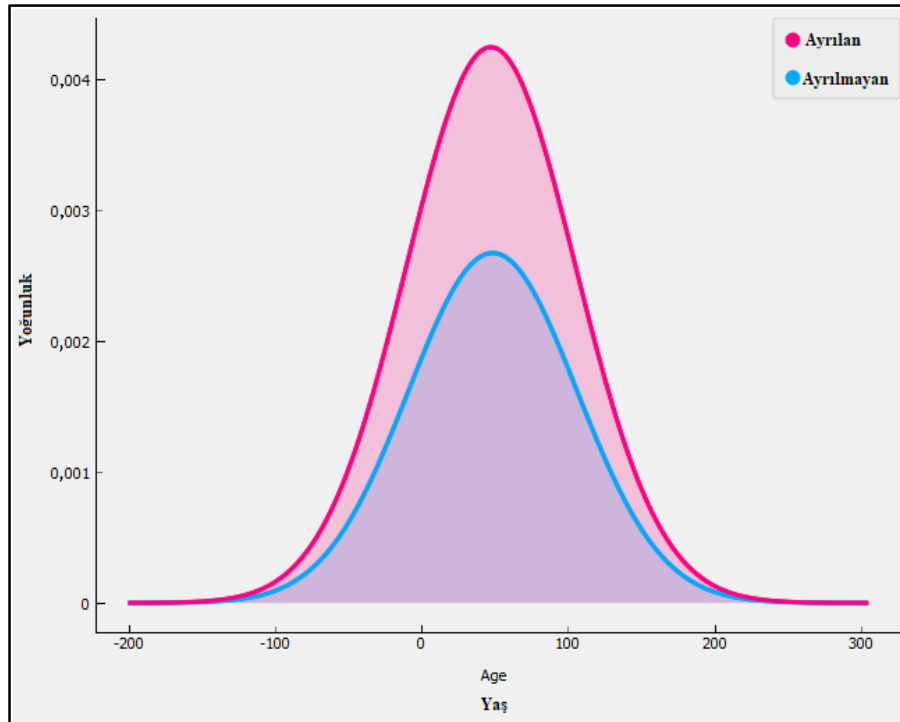
Ayrılma davranışı değişkeni çalışmanın bağımlı değişkenidir. Mevcut çalışmada ayrılma değişkeninin hesaplanmasında üç yıllık periyodlar kullanılmıştır. Dolayısıyla, ayrılma davranışı değişkeni gözlem periyodunun bitiş tarihi olan 01/01/2016 tarihinden 31/12/2018 tarihine kadar geçen süre içerisinde müşterinin zincir otel işletmesinde konaklamada bulunup bulunmadığı bilgisini içermektedir. Ayrılma davranışı değişkeni ikili kodlanmış olup, ayrılma davranışında bulunan müşteriler için “1”, ayrılmayan müşteriler için “0” değerini almaktadır (Burez ve Van den Poel, 2008: 506). Ayrılma davranışında bulunan müşterilerin toplam

müşteriler içerisindeki oranı %61 olup ($f=5648$), ayrılma davranışında bulunmayan müşterilerden (%39, $f=3544$) yüksektir. Ayrılma davranışı bağımlı değişkeninin dağılımları Grafik 3.25’de sunulmuştur.

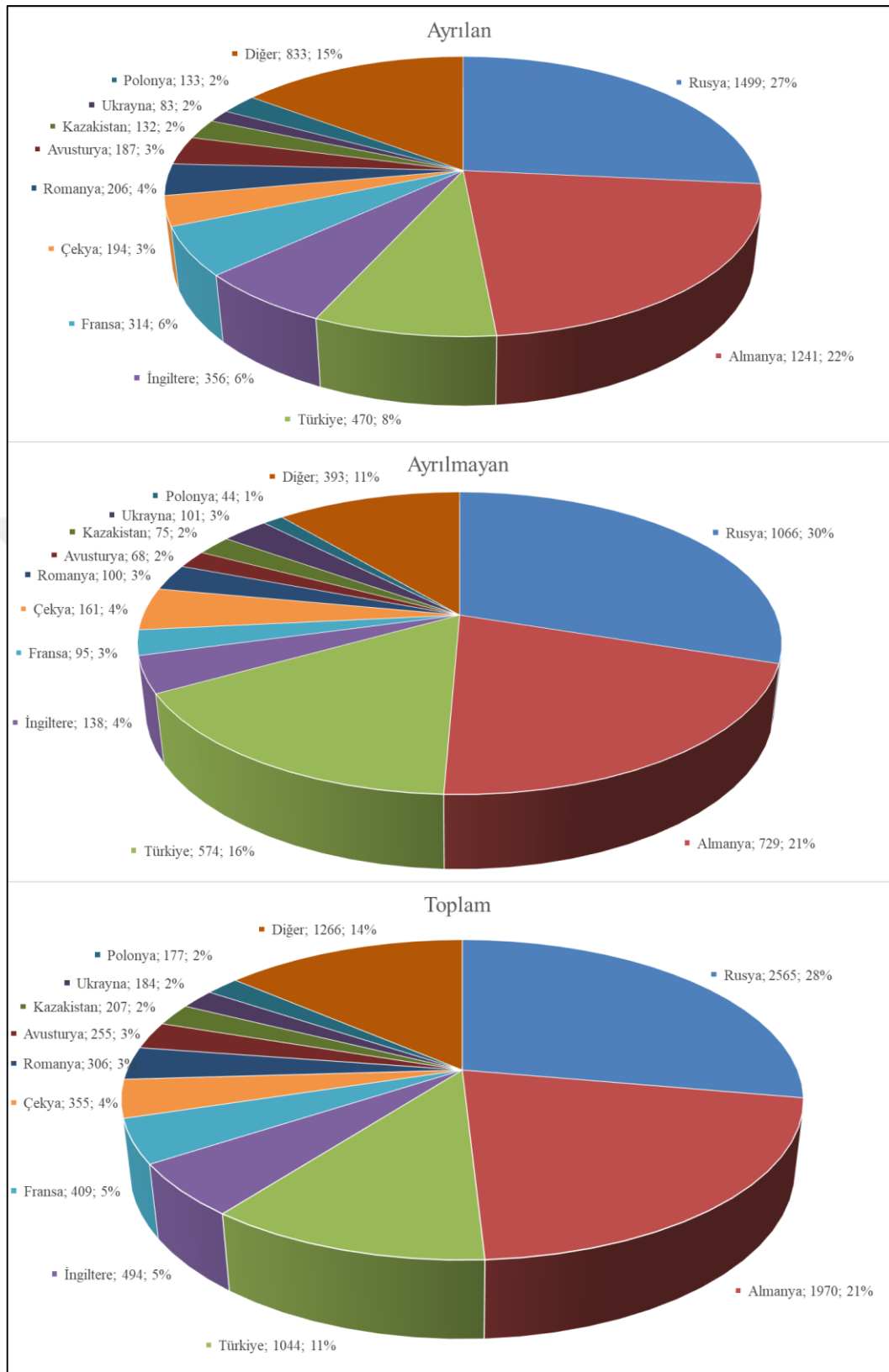


Grafik 3.25 Ayrılma Davranışı Bağımlı Değişkeni Dağılım Grafiği

Demografik yordayıcı değişkenlerden yaş değişkeni nümerik olup müşterinin gözlem periyodu sonundaki yaşını ifade etmektedir. Yaş değişkeninin aldığı değerler 20 ve 80 arasında değişmekte olup genel ortalaması 48’dir ($\sigma = 12,3$). Bu değer ayrılmayan misafirler için ($\bar{X} = 48,8$) ayrılan misafirlere ($\bar{X} = 47,6$) ve genel ortalamaya kıyasla daha yüksektir. Yaş değişkeninin dağılım grafiği Grafik 3.26’da sunulmuştur.



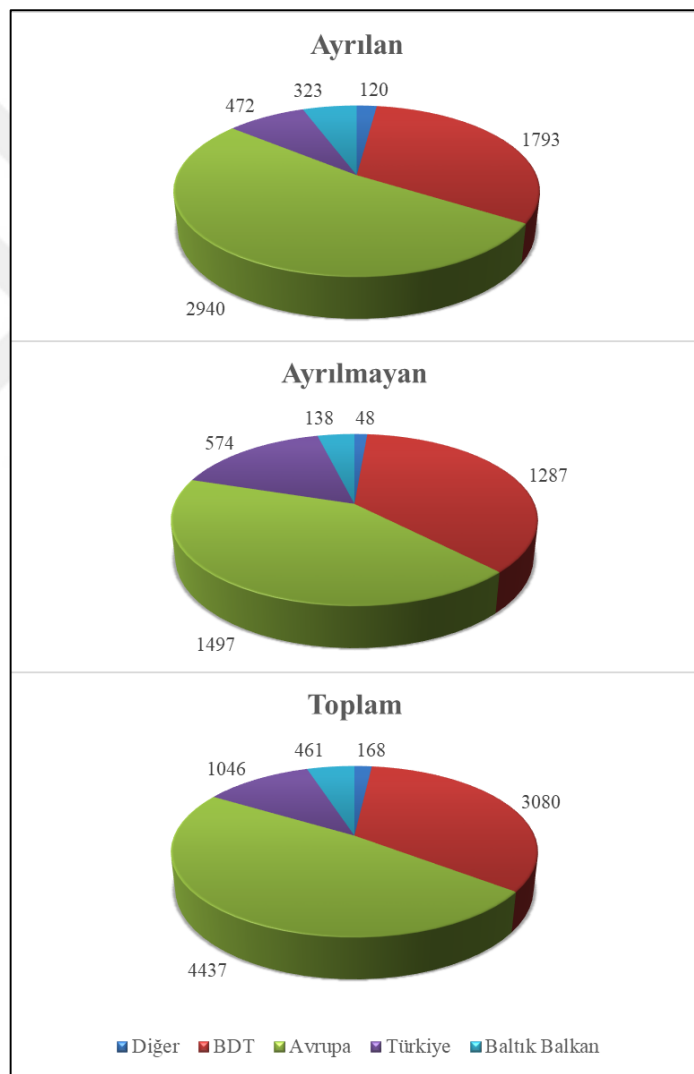
Grafik 3.26 Yaş Değişkeni Dağılım Grafiği



Grafik 3.27 Milliyet Değişkenleri Dağılım Grafiği

Milliyet değişkeni zincir otel işletmesini en çok ziyaret eden yüzde %85’lik dilimdeki ilk on bir milliyet kategorisini kapsamaktadır. Milliyet değişkenleri ikili kodlanarak çalışmaya dahil edilmiştir. Otel grubunu en çok ziyaret eden ilk üç milliyet incelendiğinde, milliyeti Rusya

olan müşterilerin ayrılmayan müşteriler içerisindeki oranı %30 (f=1066) olarak ortaya çıkmış olup, ayrılan müşteriler içerisindeki oranına kıyasla (%27, f=1499) daha yüksektir. Toplam müşteriler içerisindeki oranı ise %28'dir (f=2565). Benzer şekilde milliyeti TC olan müşterilerin ayrılmayan müşteriler içerisindeki oranı %16 (f=574), ayrılan müşteriler içerisindeki oranı ise %8 (f=470) olarak ortaya çıkmış olup, toplam müşteriler içerisindeki oranı %11'dir (1044). Diğer taraftan milliyeti Almanya olan müşterilerin ayrılan müşteriler içerisindeki oranı %22 (f=1241), ayrılmayan müşteriler arasındaki oranı ise %21 (f=729) olup, toplam müşteriler arasındaki oranı ise %21'dir (f=1970). Mevcut çalışmada kullanılan milliyet değişkenlerinin dağılımları Grafik 3.27'de sunulmuştur.



Grafik 3.28 Pazar Değişkeni Dağılım Grafiği

Son olarak müşterinin geldiği pazar bilgisi otel zincirinin satış ve pazarlama departmanı tarafından belirlenen kategoriler kullanılarak oluşturulmuş ve ikili kodlanmıştır. BDT pazarı (Bağımsız Devletler Topluluğu: Azerbaycan, Belarus, Ermenistan, Moldova, Kazakistan, Kırgızistan, Tacikistan, Özbekistan, Rusya ve Ukrayna) müşterilerinin ayrılmayan müşteriler

içerisindeki oranı %36 (f=1287), ayrılan müşteriler arasındaki oranı ise %32 (f= 1793) olarak ortaya çıkmış olup, toplam müşteriler içerisindeki oranı %34'tür (f=3080). Benzer şekilde Türkiye pazarı müşterilerinin ayrılmayan müşteriler arasındaki oranı (%16, f=574) ayrılan müşteriler arasındaki oranına (%8, f=472) kıyasla daha yüksek olup, toplam müşteriler arasındaki oranı %11'dir (f=1046). Diğer taraftan Avrupa pazarı müşterilerinin ayrılan müşteriler arasındaki oranı %52 (f=2940), ayrılmayan müşteriler arasındaki oranı ise %42 (f=1497) olup, tüm müşteriler arasındaki oranı %48'dir (f=4437). Benzer şekilde Baltık Balkan pazarı müşterilerinin ayrılan müşteriler arasındaki oranı (%6, f=323) ayrılmayan müşteriler arasındaki oranına (%4, f=138) kıyasla daha yüksek olup, tüm müşteriler arasındaki oranı %5'tir (f=461). Mevcut çalışmada kullanılan pazar değişkenlerinin dağılımları Grafik 3.28'de sunulmuştur.

Tablo 3.3 Mevcut Çalışmada Kullanılan Davranışsal Değişkenler

Değişken	Açıklama	Yazar, Yıl
Satın alımlar arası geçen zaman	▪ Interpurchase ▪ Nümerik ▪ Müşterinin geliş tarihleri arasında geçen sürelerin ortalaması	Van den Poel ve Larivière, 2004
İlişki süresi	▪ LOR ▪ Nümerik ▪ İlk konaklamadan gözlem periyodu sonuna kadar geçen gün sayısı	Coussement ve De Bock, 2013
Tutar	▪ Monetary ▪ Nümerik ▪ Son üç yıldaki konaklamalar için ödenen toplam tutar ▪ R_Monetary ▪ Nümerik ▪ Tüm konaklamalar için ödenen toplam tutar ▪ MonPerNight ▪ Nümerik ▪ Son üç yıldaki ortalama gecelik ödeme tutarı - Monetary/Nights ▪ R_MonPerNight ▪ Nümerik ▪ Tüm konaklamaların ortalama gecelik ödeme tutarı - R_Monetary/R_Nights ▪ AverageMon ▪ Nümerik ▪ Son üç yıldaki ortalama konaklama başına ödeme tutarı - Monetary/Repeat ▪ R_AverageMon ▪ Nümerik ▪ Tüm konaklamalar için ortalama konaklama başına ödeme tutarı - R_Monetary/R_Repeat	Buckinx ve Van den Poel, 2005

Değişken	Açıklama	Yazar, Yıl
Güncellik	▪ Recency ▪ Nümerik ▪ Son konaklamadan itibaren geçen gün sayısı	Coussement ve De Bock, 2013
Sıklık	▪ Repeat ▪ Nümerik ▪ Son üç yıldaki toplam geliş sayısı ▪ R_Repeat ▪ Nümerik ▪ Tüm konaklamalar için toplam geliş sayısı	Buckinx ve Van den Poel, 2005
Konaklama Eşlikçileri	▪ Acc_Chld ▪ İkili ▪ Çocuklu konaklama: evet (1), hayır (0) ▪ Acc_Sng ▪ İkili ▪ Tek konaklama: evet (1), hayır (0) ▪ Acc_Baby ▪ İkili ▪ Bebekli konaklama: evet (1), hayır (0) ▪ Acc_Adlts ▪ İkili ▪ Yetişkin grubu: evet (1), hayır (0) ▪ Acc_Couple ▪ İkili ▪ Çift konaklama: evet (1), hayır (0)	De Wit, 2017
Otel	▪ LHA ▪ İkili ▪ Son konaklama A Hotel’de: evet (1), hayır (0) ▪ LHB ▪ İkili ▪ Son konaklama B Hotel’de: evet (1), hayır (0) ▪ LHC ▪ İkili ▪ Son konaklama C Hotel’de: evet (1), hayır (0) ▪ OHID ▪ İkili ▪ Grubun diğer otellerinde de konaklama mevcut: evet (1), hayır (0)	De Wit, 2017
Sadakat Kartı	▪ Card ▪ Sıralı ▪ Kartı yok (0), Silver (1), Silver Plus (2), Gold (3), Platinum (4)	Balton ve Kannan, 2000
Satın Alım Yöntemi	▪ A_IND ▪ İkili ▪ Münferit rezervasyon: evet (1), hayır (0)	De Wit, 2017

Değişken	Açıklama	Yazar, Yıl
Konaklama süresi	- Nights - Nümerik - Son üç yıldaki toplam geceleme sayısı - R_Nights - Nümerik - Tüm konaklamalar için toplam geceleme sayısı - AvNights - Nümerik - Son üç yıldaki ortalama geceleme sayısı -Nights/Repeat - R_AvNights - Nümerik - Tüm konaklamalar için ortalama geceleme sayısı – R_Nights/R_Repeat	De Wit, 2017
Sezon	- S_Low - İkili - Düşük sezon: evet (1), hayır (0) - S_Mid - İkili - Orta sezon: evet (1), hayır (0) - S_High - İkili - Yüksek sezon: evet (1), hayır (0)	Mevcut çalışma
Eğilimler	- DiffMonPerNight - Nümerik - Gecelik ödeme tutarı değişimi: MonPerNight -R_MonPerNight - DiffAvNights - Nümerik - Ortalama geceleme sayısı değişimi: AvNights- R_AvNights - DiffAvMon - Nümerik - Ortalama harcama tutarı değişimi: AverageMon – R_AverageMon	Miguéis vd., 2012b
Ayrılma	- Churn - Nümerik - Müşteri 01/01/2016 – 31/12/2018 tarihleri arasında gelmedi: evet (1), hayır (0)	Burez ve Van den Poel, 2008

Tablo 3.4 Mevcut Çalışmada Kullanılan Demografik Değişkenler

Değişken	Açıklama	Referans
Yaş	▪ Age ▪ Nümerik ▪ Müşterinin gözlem periyodu sonundaki yaşı	Van den Poel ve Larivière, 2004; Suh ve Alhaery, 2016
Milliyet	▪ N_RUS ▪ İkili ▪ Rusya: evet (1), hayır (0) ▪ N_DEU ▪ İkili ▪ Almanya: evet (1), hayır (0) ▪ N_TC ▪ İkili ▪ Türkiye: evet (1), hayır (0) ▪ N_UK ▪ İkili ▪ İngiltere: evet (1), hayır (0) ▪ N_CZE ▪ İkili ▪ Çekya: evet (1), hayır (0) ▪ N_ROU ▪ İkili ▪ Romanya: evet (1), hayır (0) ▪ N_FR ▪ İkili ▪ Fransa: evet (1), hayır (0) ▪ N_UKR ▪ İkili ▪ Ukrayna: evet (1), hayır (0) ▪ N_KAZ ▪ İkili ▪ Kazakistan: evet (1), hayır (0) ▪ N_AUS ▪ İkili ▪ Avusturya: evet (1), hayır (0) ▪ N_POL ▪ İkili ▪ Polonya: evet (1), hayır (0) ▪ N_OTHER ▪ İkili ▪ Diğer: evet (1), hayır (0)	Coussement ve De Bock, 2013

Değişken	Açıklama	Referans
Pazar	▪ M_BDT ▪ İkili ▪ BDT pazarı: evet (1), hayır (0) ▪ M_EUROPE ▪ İkili ▪ Avrupa pazarı: evet (1), hayır (0) ▪ M_LOCAL ▪ İkili ▪ Türkiye pazarı: evet (1), hayır (0) ▪ M_BB ▪ İkili ▪ Baltık balkan pazarı: evet (1), hayır (0) ▪ M_OTHER ▪ İkili ▪ Diğer: evet (1), hayır (0)	Mevcut çalışma

3.3.5. Veri Setinin Bölünmesi : Öğrenme, Doğrulama ve Test Etme

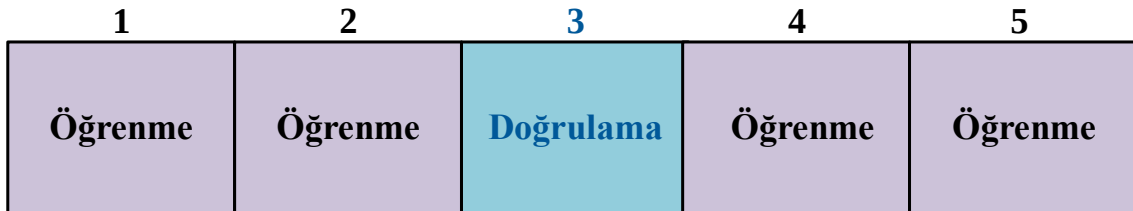
Veri sayısı fazla olduğunda en iyi uygulama veri setini rastgele olarak üç parçaya bölmektir: (1) öğrenme veri seti (training), (2) doğrulama veri seti (validating) ve (3) test veri seti (testing). Öğrenme veri seti modelin uyumu için kullanılırken, doğrulama seti hangi modelin seçileceğine karar verilirken tahmin hatasını hesaplamak için kullanılmaktadır. Test veri seti ise en son seçilen modelin genelleme hatasını (modelin daha önce görülmemiş veriler için sonuç değerlerini ne kadar doğru tahmin edebileceğinin ölçüsü) değerlendirmede kullanılır. İdeal olan test setinin saklı tutulması ve sadece veri analizinin sonunda kullanılmasıdır. Test setinin birden fazla defa kullanıldığını varsayarsak, en küçük genelleme hatasına sahip modelin seçilmesi halinde, en son seçtiğimiz modelin (final model) test hatası, gerçek hatayı eksik tahmin edecektir. Bu üç veri setindeki gözlem sayılarının nasıl seçileceğiyle ilgili genel bir kural vermek zordur, çünkü bu karar verideki sinyal gürültü oranına (signal-to-noise ratio, SNR ya da S/N) ve öğrenme setinin örneklem büyüklüğüne bağlıdır.

Yaygın kullanımda veri setinin ayrıştırılması %50 öğrenme, %25 doğrulama ve %25 test etme şeklinde olmaktadır (Şekil 3.13).



Şekil 3.13 Öğrenme, Doğrulama ve Test Veri Setlerinin Oranı

Verideki gözlem sayısı veriyi öğrenme, doğrulama ve test etme için üç parçaya bölmede yeterli değilse çapraz doğrulama (cross validation) metodundan faydalanılabilir. Yine ne kadarlık bir öğrenme verisinin yeterli olacağı konusunda kesin bir kural belirtilmesi zordur. Çünkü veriye uyum sağlayan modellerin karmaşıklığı ve sinyal gürültü oranı gibi unsurlar öğrenme verisinin büyüklüğünü etkileyebilmektedir (Hastie vd., 2009: 222). Veri setinin gözlem büyüklüğü yeterli olduğunda doğrulama için bir veri seti ayrılarak tahmin modelinin performansını değerlendirmede kullanılabilir. Ancak, genellikle veri sayısı yetersiz olmakta ve ayrı bir doğrulama seti oluşturulamamaktadır. Bu problemi aşmak adına k -katlamalı çapraz doğrulama (k -fold cross validation) metoduyla elde bulunan verinin bir kısmı modeli uydurmak için kullanılmakta ve diğer kısmı da bunu test etmede kullanılmaktadır. Bunun için veri K miktarda eşit parçalara bölünür. Örneğin $K = 5$ için, veri seti Şekil 3.14’de gösterildiği şekilde bölünecektir. Şekil 3.14’de gösterilen k . parça (şekilde 3. parça) için model, verinin geri kalan $K - 1$ adet parçasına uydurulmaktadır. Ardından uydurulmuş olan modelin (the fitted model) tahmin hatası verinin k . parçası tahmin edilerek hesaplanır. Genellikle K değeri 5 ya da 10 seçilmektedir. $K = N$ olarak seçilmesi durumu birini dışarıda bırakarak çapraz doğrulama (leave-one-out cross-validation) olarak bilinmektedir. Bu durumda $K(i) = i$ olmakta ve i . gözlem için uyumun hesaplanmasında i . gözlem hariç tutularak geri kalan verinin tamamı kullanılmaktadır (Hastie vd., 2009: 242).



Şekil 3.14 k -Katlamalı Çapraz Doğrulama

K değerinin 10 seçilmesi en yaygın görülen durumdur ve 10 katlamalı çapraz doğrulama (10-fold cross-validation) metodu olarak anılır. Bu yöntem temel tahmin algoritmasının değiştirilmesine gerek kalmadan doğrulama yapılmasına imkân verdiği için günümüzde sayısal açıdan ilgi çekici ve rağbet gören bir prosedürdür (Smola ve Vishwanathan, 2008: 56-57).

Tablo 3.5 Öğrenme ve Test Etme Veri Setleri

	Öğrenme		Test		Toplam	
Ayrılan	3953	%61	1695	%61	5648	%61
Ayrılmayan	2480	%39	1064	%39	3544	%39
Toplam	6433	%70	2759	%30	9192	%100

Mevcut çalışmada, öğrenme ve test veri setleri WEKA programının “Resample” özelliği kullanılarak bölünmüş; veri setinin %70’i öğrenme, %30’u test etme için ayrılmıştır (Tablo 3.5). Doğrulama için 10 katlamalı çapraz doğrulama yöntemi tercih edilmiştir.

3.3.6. Değişken Seçimi

Makine öğrenmesinde en iyi modelin oluşturulmasında ilk aşama doğru değişken setinin belirlenmesidir. Çoğu öğrenme modelinin başarısı yüksek ölçüde yordama gücüne sahip küçük bir değişken setinin tespit edilmesinde saklıdır. Gürültü içeren, hedef değişkeni yordamada yetersiz kalan değişkenler, model oluşturma aşamasında düşük tahmin performansına ve daha uzun süren hesaplamalara neden olabilir. Bu bağlamda, değişken seçimi aşaması veri setinde modele katkısı bulunmayan bilgilerin tespit edilmesi ve mümkün olduğunca modelden çıkarılması sürecini ifade eder (Hall ve Holmes, 2003: 1437). Hall ve Holmes (2003: 1447) değişken seçimi araçlarını kıyasladıkları çalışmalarında, yavaşlığı göz önünde bulundurulmadığı sürece model doğruluğu için en iyi değişken seçimi aracının Wrapper olduğunu bulmuştur. Wrapper değişken seçimi aracı, değişken setlerinin değerini hesaplarken hedeflenen öğrenme algoritmasını kullanır. Bununla birlikte 5 katlamalı bir çapraz doğrulamadan yararlanır. Modifiye edilmiş ileriye doğru seçme metodu kullanılarak hedef öğrenim algoritmasının performansına yapılan genel katkıya göre sıralanmış bir değişken listesi oluşturulur (Hall ve Holmes, 2003: 1440).

Mevcut çalışmada değişken setinin belirlenmesinde, Rastgele Orman için WEKA programının “Select Attributes” aracı kullanılmış; metod olarak “Wrapper (hedef algoritma: Rastgele Orman) + Greedy Stepwise” yöntemi tercih edilmiştir. Greedy (açgözlü) stepwise, değişken alt setleri arasında ileri ve geri seçme yöntemi uygulayarak değerlendirmede hiçbir değişken düşüşe neden olmayana kadar ekleme/çıkarma yapar (weka.attributeSelection.GreedyStepwise, erişim tarihi: 10.08.2019). Lojistik regresyon için ise klasik ileriye doğru seçme yöntemiyle (forward stepwise) lojistik regresyon analizi yapılmış ve analizde seçilen değişken seti kullanılmıştır.

Bu aşamadan sonra her iki değişken seti için korelasyonlar ve çoklu doğrusal bağıntı (multicollinearity) problemi olup olmadığı incelenmiştir. Rastgele Orman için belirlenen değişken setinde aralarında yüksek korelasyon bulunan (0,70 üzerinde) değişkenlerden AUC (Area Under Curve – ROC eğrisi altında kalan alan) değerleri yüksek olanlar seçilmiştir (Qi vd., 2009: 256). Çoklu doğrusal bağıntı için sıklıkla kullanılan ölçülerden biri varyans büyütme faktörü yani “VIF” (Variance Inflation Factor) değeridir. VIF değerleri her iki veri seti için de sınır değerler (VIF <10) içerisinde olup, çoklu bağıntı problemi söz konusu değildir (Ho, 2014:

297). En yüksek korelasyon değeri Rastgele Orman veri seti için 0,64 ile LOR ve R_Repeat değişkenleri arasında; Lojistik Regresyon için 0,66 ile CARD ve R_Repeat değişkenleri arasındadır. Bu işlemler sonrasında belirlenen en iyi değişken setleri Rastgele Orman için Tablo 3.6'da, Lojistik Regresyon için Tablo 3.7'de gösterilmiştir.

Tablo 3.6 Rastgele Orman Değişken Önem Düzeyleri

Değişkenler	Değişken Önem Skoru	Ayrılmayan $\bar{X}$	Ayrılan $\bar{X}$	std	VIF	AUC
Recency*	0,380	293,6	315,6	209,9	1,170	0,626
Interpurchase*	0,360	534,9	660,2	487,6	3,079	0,574
Nights*	0,340	23,4	18,3	13,4	2,004	0,610
LOR*	0,330	1698,2	1358,0	915,8	5,612	0,561
DiffAvMon*	0,300	72,38 \$	27,01 \$	348,2	1,888	0,529
R_Repeat*	0,270	4,17	2,76	2,3	5,479	0,678
LHA	0,270	0,28	0,27	0,4	1,606	0,509
DiffAvNights*	0,270	0,09	-0,04	1,2	1,866	0,584
N_RUS*	0,250	0,30	0,26	0,4	3,558	0,513
LHB*	0,250	0,45	0,51	0,5	1,722	0,522
M_EUROPE*	0,230	0,42	0,52	0,5	4,161	0,544
S_High	0,230	0,47	0,48	0,5	1,188	0,501
A_IND*	0,190	0,12	0,04	0,3	1,344	0,548
S_Low*	0,180	0,08	0,11	0,3	1,245	0,509
N_TC*	0,180	0,16	0,09	0,3	2,638	0,542
M_BB*	0,170	0,04	0,06	0,2	1,603	0,511
Acc_Baby*	0,160	0,04	0,09	0,3	1,045	0,524
N_FR*	0,150	0,02	0,06	0,2	1,101	0,508
N_POL*	0,140	0,01	0,02	0,1	1,058	0,505
OHID*	0,090	0,10	0,00	0,2	1,034	0,542

*p < 0,01

Tablo 3.6'daki bulgulara göre Rastgele Orman algoritması hesaplamalarına göre en önemli bulunan ilk üç değişken sırasıyla Recency, Interpurchase ve Nights değişkenleridir. Değişkenlerin tek başına hesaplanan AUC değerleri hesaplamasında 10 katlamalı çapraz doğrulama yönteminden faydalanılmıştır. Ayrılan ve ayrılmayan müşteri grup ortalamalarının karşılaştırılmasında t-testi kullanılmıştır. Anlamlı farklılık bulunan ortalamalara göre, ayrılma davranışı gösteren müşterilerin son geliş tarihleri daha eski ($\bar{X}$ = 315,6 gün), ortalama satın alım sıklıkları daha seyrek ($\bar{X}$ = 660,2 gün), son üç yıldaki geceleme sayıları daha düşük ($\bar{X}$ = 18,3 gece), otel zinciriyle olan ilişki süresi daha kısa ($\bar{X}$ = 1358 gün), son üç yıldaki harcama değişimlerdeki artış miktarı daha düşük ($\bar{X}$ = 27 dolar), toplam geliş sayıları daha az ($\bar{X}$ = 2,8 defa) ve son üç yıldaki ortalama geceleme sayıları azalış eğilimindedir ($\bar{X}$ = -0,04). Ayrıca ayrılan müşteriler arasında daha az sayıda Rusya vatandaşı bulunduğu ($\bar{X}$ = 0,26), son konaklamalarının B Hotel'de daha fazla gerçekleştiği ($\bar{X}$ = 0,51), Avrupa pazarından gelenlerin

daha fazla olduğu ($\bar{X} = 0,52$), münferit rezervasyon yapanların daha az olduğu ($\bar{X} = 0,04$), düşük sezonda konaklayanların daha fazla olduğu ($\bar{X} = 0,11$), daha az sayıda TC vatandaşı bulunduğu ($\bar{X} = 0,09$), Balkan Baltık pazarından gelenlerin daha fazla olduğu ($\bar{X} = 0,06$), bebekli konaklayanların daha fazla olduğu ($\bar{X} = 0,09$), daha fazla sayıda Fransız vatandaşı ($\bar{X} = 0,06$) ve Polonya vatandaşı bulunduğu ($\bar{X} = 0,02$) ve son olarak grubun diğer otellerinde konaklama gerçekleşmediği ($\bar{X} = 0,00$) görülmektedir.

Tablo 3.7 Lojistik Regresyon Değişken Önem Düzeyleri

Değişkenler	β	Standartlaştırılmış β	S.E.	Wald	odds	$ 1 - odds $	VIF
OHID*	-5,208	-0,550	0,713	53,386	0,005	0,995	1,032
R_Repeat*	-0,157	-0,199	0,025	39,399	0,854	0,146	2,995
Acc_Chld*	0,678	0,180	0,08	72,048	1,969	0,969	1,901
Acc_Couple*	0,64	0,162	0,082	61,402	1,897	0,897	1,815
N_TC*	-0,791	-0,139	0,111	50,389	0,453	0,547	1,576
Repeat*	-0,26	-0,125	0,041	39,464	0,771	0,229	1,524
Acc_Sng*	0,782	0,114	0,131	35,669	2,186	1,186	1,463
Acc_Baby*	0,78	0,111	0,132	34,85	2,182	1,182	1,086
N_FR*	0,738	0,084	0,161	21,096	2,091	1,091	1,027
Card*	-0,189	-0,079	0,066	8,219	0,828	0,172	2,749
A_IND*	-0,509	-0,074	0,125	16,563	0,601	0,399	1,361
N_UK*	0,507	0,064	0,137	13,726	1,66	0,66	1,031
S_Low*	0,334	0,055	0,109	9,324	1,397	0,397	1,186
N_CZE*	-0,517	-0,055	0,145	12,732	0,596	0,404	1,035
Acc_Adlt*	0,42	0,055	0,131	10,233	1,523	0,523	1,273
N_UKR*	-0,607	-0,047	0,194	9,811	0,545	0,455	1,013
DiffAvNights**	-0,052	-0,035	0,025	4,289	0,949	0,051	1,023
R_Monetary*	0	0,000	0	16,8	1	0	1,987

Standartlaştırılmış β katsayılarına göre sıralanmıştır. Standartlaştırılmış β hesaplanmasında $\beta\sigma/(\frac{\pi}{\sqrt{3}})$ formülü kullanılmıştır (Menard, 2011:1417).

Tablo 3.7'deki bulgulara göre lojistik regresyon değişken setindeki tüm değişkenler ayrılma davranışının anlamlı bir yordayıcısıdır (* $p < 0,01$, ** $p < 0,05$). Değişkenler arasındaki en güçlü yordayıcı (Standartlaştırılmış $\beta = -0,550$) 0,005 odds (olasılık) oranıyla OHID (grubun diğer otellerinde de konaklamış olmak) değişkeni olup; bu oran OHID değişkenindeki 1 birimlik artışın müşterinin ayrılma ihtimalini ($|1 - odds| = 0,995$) %99 oranında azalttığını göstermektedir (Ho, 2014: 392). Bu değişkeni R_Repeat (-0,199) ve Acc_Chld (0,180) değişkenleri takip etmektedir. Bulgulara göre grubun diğer otellerinde konaklamış olmak toplam geliş sayısındaki artış, TC, Ukrayna ve Çekya vatandaşı olmak, münferit rezervasyon yaptırmış olmak, son üç yıldaki geliş sayısındaki artış, üyelik düzeyindeki artış ve son üç yıldaki geceleme sayısındaki artış eğilimi ayrılma davranışı üzerinde anlamlı ve negatif yönlü bir etkiye

sahiptir. Başka bir ifadeyle, bu değişkenlerin değerlerinde artış oldukça, ayrılma olasılığı düşmektedir. Diğer taraftan, düşük sezonda konaklamış olmak, yetişkin grubu halinde konaklamak, İngiltere vatandaşı olmak, çift olarak konaklamak, çocuklu konaklamak, Fransız vatandaşı olmak, bebekli konaklamak ve tek konaklamak ayrılma davranışı üzerinde anlamlı ve pozitif yönlü bir etkiye sahiptir. Diğer bir deyişle, bu değişken değerlerinde artış gözlendikçe ayrılma olasılığı artmaktadır. Ancak, toplam harcama tutarındaki artış ya da azalışın ayrılma davranışı üzerindeki etkisi nötr olarak ortaya çıkmıştır.

3.3.7. Öğrenme Aşaması ve Çapraz Doğrulama

Öğrenme aşamasında hiç kullanılmayan bir veri üzerinde modelin performansının değerlendirilmesi, modelin gerçek gücü hakkında daha doğru bir bilgi sunabilir. Öğrenme sürecinde sıradan bir modelin yanlış bir kararla güçlü bir model olarak görülmesi riskini azaltmak amacıyla aynı işlemler birden fazla defa tekrarlanır. Bu tekrarlama işlemlerinde çapraz doğrulama yöntemi kullanılır. Tahmin modeli test verisi üzerinde denenmeden önce farklı metodlar denenerek en iyileme çalışmasının yapılması, en iyi algoritmanın oluşturulması açısından önem arz eder (Coussement ve De Bock, 2013: 1630).

Bu aşamada öğrenme algoritmalarının en iyileme (tuning) ayarlamaları yapılarak 10 katlamalı çapraz doğrulama yöntemiyle kontrolü sağlanmıştır. Rastgele Orman algoritmasında kullanılacak ağaç sayısının büyüklüğü ve ağaçlar büyürken kullanılan tesadüfi olarak seçilmiş değişken alt setindeki değişken sayısı modelin gücü üzerinde etkili olabilmektedir. Bu çalışma kapsamında Breiman (2001) tarafından önerilen yöntem benimsenmiş olup, çok sayıda ağaç kullanılmış (155 ağaç) ve değişken setinin karekökü kadar ($\sqrt{20} \cong 5$) değişken seçimi sağlanmıştır (akt. Miguéis vd., 2012b: 345). Lojistik regresyonda ise C parametresi olarak bilinen ve düzenleme gücünü ifade eden değer modelin gücü üzerinde etkili olabilmektedir (De Wit, 2016: 33). Bu çalışma kapsamında $C = 0,040$ en optimal sonucu vermiştir.

Değerlendirme kriteri olarak tahmin performansı göstergesi olan AUC değeri (ROC eğrisi altında kalan alan) ve sınıflandırma doğruluğunu belirten CA değeri (Classification Accuracy) kullanılmıştır (De Wit, 2016: 39; Miguéis vd., 2012b: 345; Burez ve Van den Poel, 2008: 500; Coussement ve Van den Poel, 2008: 167).

Sınıflandırma doğruluğu (CA) sınıflandırma modelinin doğru bir şekilde sınıflandırdığı durumların yüzdesini ifade eder. Kayıp müşteri durumunda olduğu gibi ikili bir sınıflandırma söz konusu olduğunda doğruluk oranı bir karışıklık matrisi (confusion matrix) üzerinden hesaplanabilir (Şekil 3.15).

Bu matris üzerinde doğru pozitif, doğru negatif (true positive, true negative) ve yanlış pozitif, yanlış negatif (false positive, false negative) oranları belirtilir (Maier ve Prusty, 2016: 807; Classification: Accuracy, erişim tarihi: 20.09.2019):

$$\begin{aligned} \text{Sınıflandırma Doğruluğu} &= \frac{\text{Doğru Tahminlerin Sayısı}}{\text{Tüm Tahminlerin Sayısı}} \\ &= \frac{DP + DN}{DP + YP + YN + DN} \end{aligned}$$

<i>Doğru Pozitif (DP)</i> *Gerçekte olan: Ayrıldı *Modelin tahmini: Ayrıldı	<i>Yanlış Negatif (YN)</i> *Gerçekte olan: Ayrıldı *Modelin tahmini: Ayrılmadı
<i>Yanlış Pozitif (YP)</i> *Gerçekte olan: Ayrılmadı *Modelin tahmini: Ayrıldı	<i>Doğru Negatif (DN)</i> *Gerçekte olan: Ayrılmadı *Modelin tahmini: Ayrılmadı

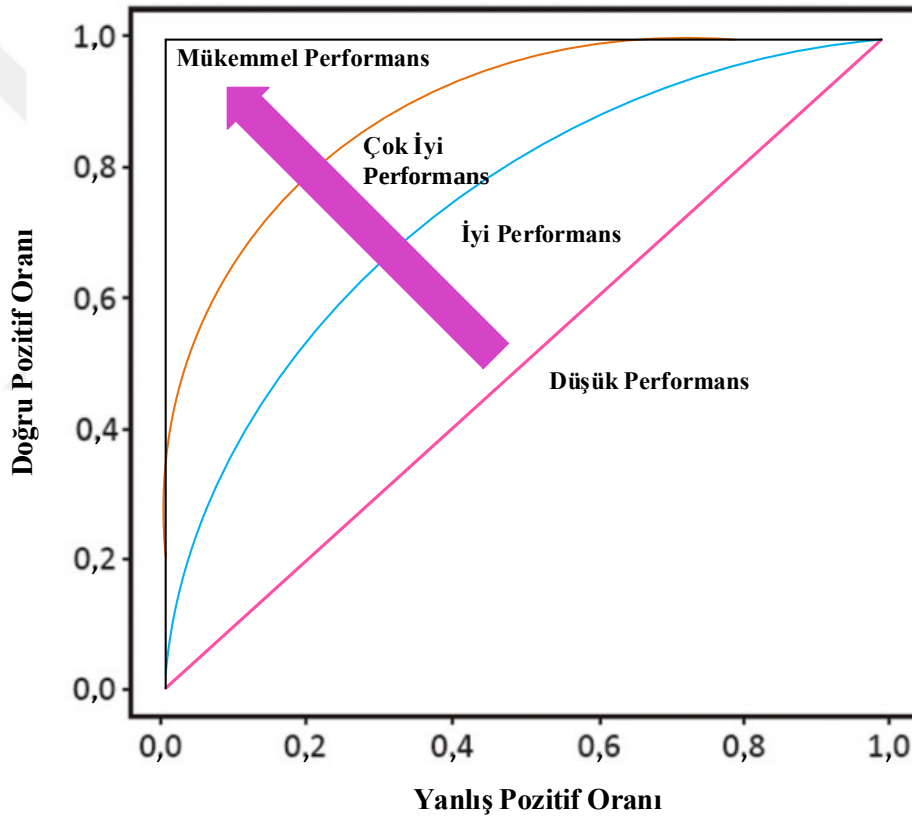
Şekil 3.15 Karışıklık Matrisi Örneği

Bu değerlerin yanısıra modelin gerçekteki pozitif durumları ne kadar doğru tahmin ettiğini gösteren duyarlılık değeri (Sensitivity ya da doğru pozitif oranı) ile negatif durumları ne kadar doğru tahmin ettiğini gösteren seçicilik değeri (Specificity ya da doğru negatif oranı) tahmin performansının önemli göstergelerindendir (Maier ve Prusty, 2016: 807; Classification: Accuracy, erişim tarihi: 20.09.2019):

$$\begin{aligned} \text{Duyarlılık} &= \frac{\text{Doğru pozitiflerin toplamı}}{\text{Gerçek pozitiflerin toplamı}} \\ &= \frac{DP}{DP + YN} \end{aligned}$$

$$\begin{aligned} \text{Seçicilik} &= \frac{\text{Doğru negatiflerin toplamı}}{\text{Gerçek negatiflerin toplamı}} \\ &= \frac{DN}{YP + DN} \end{aligned}$$

ROC (Receiver Operating Characteristic– karar vericinin etkinliği) eğrisi hedef değişkeninin iki olasılıklı, yordayıcı değişkenlerin ise sürekli olması durumunda kullanılmakta olup sınıflandırma doğruluğunu belirlemede hesaplanan duyarlılık ve seçicilik değerleri arasındaki ilişkinin grafiksel bir gösterimini sunar (Kılıç, 2013: 135). Diğer bir ifadeyle ROC, “doğru pozitif” oranının “1- doğru negatif” (yanlış pozitif) oranına karşı oluşturulan iki boyutlu bir plot grafiğidir (Şekil 3.16). İkili sınıflandırma modellerinin performansını değerlendirmede “ROC eğrisi altında kalan alan” yani AUC (area under curve) değeri kullanılır (Coussement ve Van den Poel, 2008: 167). Bu değer 0,5 ve 1 arasında değişmekte olup, 0,5 değeri iki sınıf arasındaki ayrımın tamamen şansa dayalı olduğu düşük performansı; 1 değeri ise iki sınıf arasındaki ayrımın mükemmel olduğu yüksek performansı ifade eder (Kılıç, 2013: 136).



Şekil 3.16 ROC Eğrisi Örneği

Kaynak: Gantley vd., 2018: 333; Hassouna vd., 2015: 227

Modelin tahmin başarısını değerlendirmede sınıflandırma doğruluğunun yanısıra kesinlik (presicion), duyarlılık ya da yakalama oranı (recall) ve F1 değerlerinden faydalanılmaktadır. Bilgi çıkarımı sürecinde kesinlik değeri elde edilen sonucun uygunluğunu (relevancy) gösterirken, duyarlılık değeri kaç tane uygun sonucun elde edildiğine yönelik bir ölçüm sunmaktadır.

Duyarlılık değeri olan “recall”, daha önce bahsedilen “sensitivity” ile aynı şekilde hesaplanır ve doğru pozitiflerin oranını gösterir. Diğer bir ifadeyle ayrılan müşterilerin kaç tanesinin doğru tahminlendiğinde yönelik bir oran hesaplamasıdır. Diğer taraftan kesinlik değeri model tarafından ayrılan müşteri olarak tahminlenen müşterilerin kaç tanesinin doğru tahminlendiğine yönelik bir oran hesaplamasıdır. F1 değeri ise bu iki ölçümün harmonik ortalaması olarak tanımlanmaktadır.

Bu değerler karışıklık matrisinden yararlanılarak aşağıdaki şekilde hesaplanmaktadır (Precision-Recall, erişim tarihi: 10.08.2019; Vafeiadis vd., 2015: 2) :

$$\text{Kesinlik} = \frac{\text{Doğru sınıflanan ayrılan müşteriler}}{\text{Ayrılan müşteri olarak sınıflanan tüm müşteriler}} = \frac{DP}{DP + YP}$$

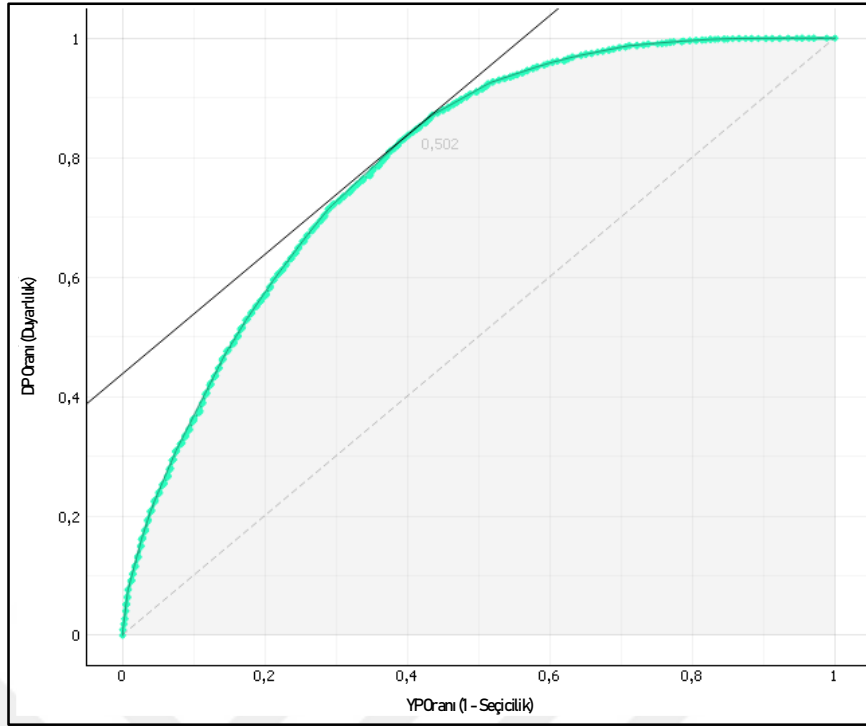
$$\text{Duyarlılık} = \frac{\text{Doğru sınıflanan ayrılan müşteriler}}{\text{Doğru sınıflanan ayrılanlar + Yanlışlıkla ayrılmayan olan sınıflanan ayrılanlar}} = \frac{DP}{DP + YN}$$

$$F1 = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$$

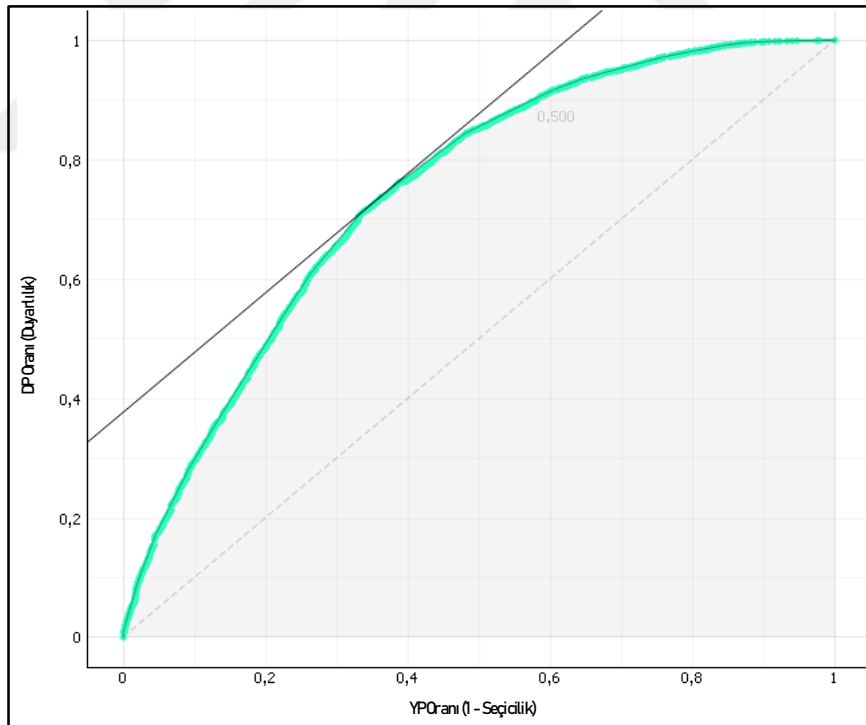
Öğrenme aşamasında 10 katlamalı çapraz doğrulama yöntemi kullanılarak performans değerleri hesaplanmış, ROC plot grafikleri ve karışıklık matrisleri oluşturulmuştur. Modellerin ayrılan ve ayrılmayan müşterileri ayırt etme gücünü ölçmek amacıyla hesaplanan AUC değerleri Tablo 3.8’de sunulmuştur. Tablo 3.8’deki değerlere göre Rastgele Orman modeli (0,79), Lojistik Regresyon modeline (0,75) kıyasla daha yüksek performans göstermektedir. Bu sonucu model başarı göstergesi olan F1 değeri ve toplam sınıflandırma doğruluğunu belirten CA değeri de desteklemektedir.

Tablo 3.8 Çapraz Doğrulama Öğrenme Modeli Performansları

Model	AUC	CA	Kesinlik	Duyarlılık	F1
<i>Rastgele Orman</i>	0,79	0,75	0,77	0,85	0,81
<i>Lojistik Regresyon</i>	0,75	0,72	0,71	0,90	0,79



Grafik 3.29 ROC Eğrisi Rastgele Orman (Çapraz Doğrulama)



Grafik 3.30 ROC Eğrisi Lojistik Regresyon (Çapraz Doğrulama)

Rastgele Orman modelinin çapraz doğrulama metoduyla hesaplanan sınıflandırma doğruluğuna yönelik veriler Tablo 3.9'daki karışıklık matrisinde sunulmuştur. Tablo 3.9'daki bulgulara göre Rastgele Orman modeli 6433 durumun 4806 tanesini doğru tahminleyerek %75 oranında sınıflandırma doğruluğuna sahiptir. Ayrılan müşteriler için modelin doğruluk oranı

%85 olup 3953 ayrılan müşterinin 3350 tanesi doğru tahmin edilmiştir. Ancak ayrılmayan müşterileri tahminleme performansı 2480 müşteride 1456 doğru tahminle %59'da kalmıştır.

Tablo 3.9 Rastgele Orman Çapraz Doğrulama Karışıklık Matrisi

Rastgele Orman		Tahmin Edilen				Σ
		Ayrılan		Ayrılmayan		
Gerçekleşen	Ayrılan	3350	% 85	603	%15	3953
	Ayrılmayan	1024	% 41	1456	%59	2480
	Σ	4347		2059		6433

Toplam Sınıflandırma Doğruluğu = %75

Lojistik regresyon modelinin çapraz doğrulama metoduyla hesaplanan sınıflandırma doğruluğuna yönelik veriler Tablo 3.10'daki karışıklık matrisinde sunulmuştur. Tablo 3.10'daki bulgulara göre Lojistik Regresyon modeli 6433 durumun 4599 tanesini doğru tahminleyerek %72 oranında sınıflandırma doğruluğuna sahiptir. Ayrılan müşteriler için modelin doğruluk oranı %90 olup 3953 ayrılan müşterinin 3544 tanesi doğru tahmin edilmiştir. Ancak ayrılmayan müşterileri tahminleme performansı 2480 müşteride 1055 doğru tahminle %43'de kalmıştır.

Tablo 3.10 Lojistik Regresyon Çapraz Doğrulama Karışıklık Matrisi

Lojistik Regresyon		Tahmin Edilen				Σ
		Ayrılan		Ayrılmayan		
Gerçekleşen	Ayrılan	3544	% 90	409	%10	3953
	Ayrılmayan	1425	% 57	1055	% 43	2480
	Σ	4969		1464		6433

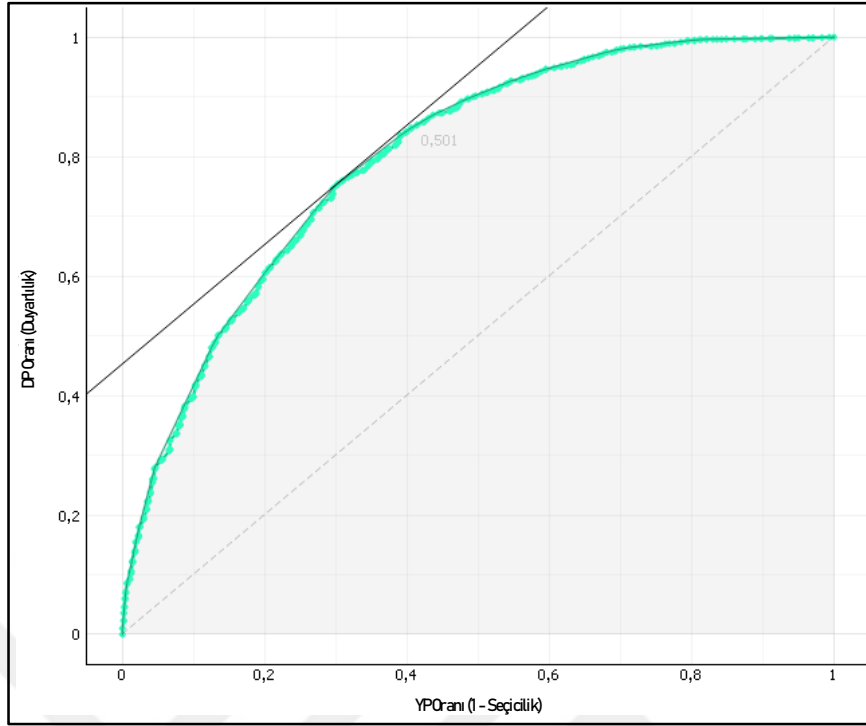
Toplam Sınıflandırma Doğruluğu = %72

3.3.8. Test Aşaması ve Model Performansı

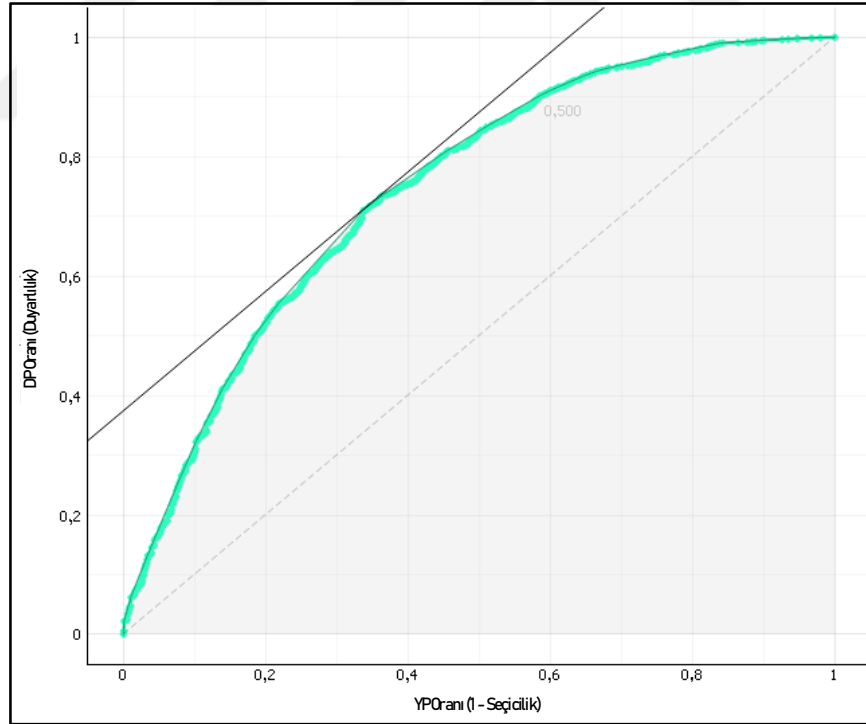
Öğrenme aşamasında çapraz doğrulama metoduyla performansları karşılaştırılarak en iyileme ayarlamaları yapılan modellerin bu aşamada daha önce gösterilmeyen bir veri seti üzerinde gerçek performansları test edilmiştir. Bu bağlamda öğrenme veri setinden öğrenen modeller test veri seti üzerinde çalıştırılmıştır. Rastgele Orman ve Lojistik Regresyon için test sonuçları ve model performansı Tablo 3.11'de sunulmuştur.

Tablo 3.11 Model Performansları

Model	AUC	CA	Kesinlik	Duyarlılık	F1
<i>Rastgele Orman</i>	0,80	0,75	0,77	0,85	0,81
<i>Lojistik Regresyon</i>	0,74	0,71	0,71	0,90	0,79



Grafik 3.31 ROC Eğrisi Rastgele Orman (Test)



Grafik 3.32 ROC Eğrisi Lojistik Regresyon (Test)

Tablo 3.11'deki değerlere göre Rastgele Orman modeli test seti üzerinde daha yüksek performans göstermiştir (0,80). Lojistik Regresyon modeli test seti üzerinde performansını düşürmüş (0,74) ve Rastgele Orman modeline kıyasla daha düşük oranda sınıflandırma

doğruluğuna ulaşmıştır. Bu sonucu model başarı göstergesi olan F1 değeri ve toplam sınıflandırma doğruluğunu belirten CA değeri de desteklemektedir.

Rastgele Orman modelinin test seti üzerinde hesaplanan sınıflandırma doğruluğuna yönelik veriler Tablo 3.12'deki karışıklık matrisinde sunulmuştur. Tablo 3.12'deki bulgulara göre Rastgele Orman modeli 2759 durumun 2069 tanesini doğru tahminleyerek %75 oranında sınıflandırma doğruluğuna sahiptir. Ayrılan müşteriler için modelin doğruluk oranı %85 olup 1695 ayrılan müşterinin 1446 tanesi doğru tahmin edilmiştir. Ancak ayrılmayan müşterileri tahminleme performansı 1064 müşteride 623 doğru tahminle %59'da kalmıştır.

Tablo 3.12 Rastgele Orman Karışıklık Matrisi

Rastgele Orman		Tahmin Edilen				Σ
Gerçekleşen		Ayrılan		Ayrılmayan		
	Ayrılan	1446	% 85	249	%15	1695
	Ayrılmayan	441	% 41	623	%59	1064
Σ		1887		872		2759

Toplam Sınıflandırma Doğruluğu = %75

Lojistik regresyon modelinin test seti üzerinde hesaplanan sınıflandırma doğruluğuna yönelik veriler Tablo 3.13'deki karışıklık matrisinde sunulmuştur. Tablo 3.13'deki bulgulara göre Lojistik Regresyon modeli 2759 durumun 1969 tanesini doğru tahminleyerek %71 oranında sınıflandırma doğruluğuna sahiptir. Ayrılan müşteriler için modelin doğruluk oranı %90 olup 1695 ayrılan müşterinin 1529 tanesi doğru tahmin edilmiştir. Ancak ayrılmayan müşterileri tahminleme performansı 1064 müşteride 440 doğru tahminle %41'de kalmıştır.

Tablo 3.13 Lojistik Regresyon Karışıklık Matrisi

Lojistik Regresyon		Tahmin Edilen				Σ
		Ayrılan		Ayrılmayan		
Gerçekleşen	Ayrılan	1529	% 90	166	%10	1695
	Ayrılmayan	624	% 59	440	% 41	1064
	Σ	2153		606		2759

Toplam Sınıflandırma Doğruluğu = %71

3.3.9. Tahminleme Aşaması ve Tahmin Sonuçları

Test aşamasının bulgularına göre 0,80 AUC değeriyle daha yüksek performansa sahip olduğu belirlenen Rastgele Orman modeli, test setinden tesadüfi olarak seçilen bir ayrılan müşteriyi %80 ihtimalle ayrılan müşteri olarak sınıflandırabilecektir (Maier ve Prusty, 2016:

812). Burdan hareketle Rastgele Orman modeli kullanılarak 1999 yılından itibaren müşteri olan ve 2018 yılında aktif olan, grubun otellerinde en az 10 gece konaklamış ve en az 2 defa gelmiş 12413 müşterinin 2016-2018 periyoduna kadar gerçekleştirdikleri konaklamalara yönelik veriler üzerinden 2019-2021 periyodunda muhtemel ayrılma davranışında bulunacak müşteriler tahminlenmiştir. Tahmin sonuçları Tablo 3.14 ve 3.15’de sunulmuştur.

Tablo 3.14’deki bulgulara göre 12413 müşterinin 7746 tanesi (%62) ayrılma davranışında bulunmaz, 4667 tanesi (%38) ayrılma davranışında bulunur şeklinde tahminlenmiştir. Bu iki müşteri grubunun RFM temelli değişkenlere yönelik ortalamaları incelendiğinde; ayrılmama olasılığı yüksek olan müşterilerde gerçek harcama tutarının ($\bar{X} = 4.684,0$ \$) ayrılma olasılığı yüksek olan müşterilere ($\bar{X} = 3.363,5$ \$) ve genel ortalamaya ($\bar{X} = 4.187,5$ \$) kıyasla daha yüksek olduğu görülmektedir. Benzer şekilde ayrılmama olasılığı yüksek olan müşterilerde gerçek geliş sayıları ($\bar{X} = 4,0$) ayrılma olasılığı yüksek olan müşterilere ($\bar{X} = 2,5$) ve genel ortalamaya ($\bar{X} = 3,5$) kıyasla daha yüksektir. Diğer taraftan güncellik değişkeninin ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 347,0$) ayrılmama olasılığı yüksek olan müşterilere ($\bar{X} = 540,4$) ve genel ortalamaya kıyasla ($\bar{X} = 467,4$) daha yakın tarihli olduğu görülmektedir. Ayrıca, gerçek harcama tutarının aksine, son üç yıldaki harcama tutarının ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 2.115,6$ \$) ayrılmama olasılığı yüksek olan müşterilere ($\bar{X} = 1.933,8$ \$) ve genel ortalamaya ($\bar{X} = 2.002,2$ \$) kıyasla daha yüksek olduğu görülmektedir. Son üç yıldaki geliş sayıları incelendiğinde ise ayrılma ihtimali yüksek olan ve ayrılmama ihtimali yüksek olan müşterilerin ortalamaları genel ortalamaya eşittir ($\bar{X} = 1,6$). Benzer şekilde gerçek gecelik ödeme tutarının ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 137,4$ \$) ayrılmama olasılığı yüksek olan müşterilere ($\bar{X} = 122,0$ \$) ve genel ortalamaya ($\bar{X} = 127,8$ \$) kıyasla daha yüksek olduğu görülmektedir. Son üç yıldaki gecelik ödeme tutarları için de aynı durum söz konusu olup ayrılma olasılığı yüksek olan müşterilerin harcamaları ($\bar{X} = 151,1$ \$) ayrılmama olasılığı yüksek olan müşterilerden ($\bar{X} = 137,1$ \$) ve genel ortalamadan ($\bar{X} = 142,3$ \$) daha yüksektir. Konaklama başına harcama tutarları incelendiğinde ise gerçek konaklama başına harcama tutarının ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 1.313,8$ \$) ayrılmama olasılığı yüksek olan müşterilere ($\bar{X} = 1.137,3$ \$) ve genel ortalamaya ($\bar{X} = 1.203,7$ \$) kıyasla daha yüksek olduğu görülmektedir. Son üç yıldaki konaklama başına harcama tutarları için de benzer şekilde ayrılma olasılığı yüksek olan müşterilerin harcamaları ($\bar{X} = 1.414,5$ \$) ayrılmama olasılığı yüksek olan müşterilerden ($\bar{X} = 1.272,5$ \$) ve genel ortalamadan ($\bar{X} = 1.325,9$ \$) daha yüksektir.

Tablo 3.14 Rastgele Orman Tahminleri: Nümerik Değişkenler

Değişkenler		Ayrılır	Ayrılmaz	$\bar{X}$	std
LOR	İlişki süresi	1534,6	2229,5	1968,3	1185,0
Recency	Güncellik	347,0	540,4	467,7	301,9
Interpurchase	Satın alımlar arası geçen zaman	836,0	744,7	779,0	697,4
R_Monetary	Gerçek harcama tutarı	3.363,5 \$	4.684,0 \$	4.187,5 \$	5768,0
Monetary	Son üç yıldaki harcama tutarı	2.115,6 \$	1.933,8 \$	2.002,2 \$	1852,9
R_MonPerNight	Gerçek gecelik ödeme tutarı	137,4 \$	122,0 \$	127,8 \$	75,4
MonPerNight	Son üç yıldaki gecelik ödeme tutarı	151,1 \$	137,1 \$	142,3 \$	77,8
R_AverageMon	Gerçek konaklama başına harcama tutarı	1.313,8 \$	1.137,3 \$	1.203,7 \$	963,0
AverageMon	Son üç yıldaki konaklama başına harcama tutarı	1.414,5 \$	1.272,5 \$	1.325,9 \$	1011,0
R_Repeat	Gerçek geliş sayısı	2,5	4,0	3,5	2,7
Repeat	Son üç yıldaki geliş sayısı	1,6	1,6	1,6	0,9
R_Nights	Gerçek geceleme sayısı	23,8	38,9	33,2	43,6
Nights	Son üç yıldaki geceleme sayısı	14,8	14,6	14,7	11,2
R_AvNights	Gerçek ortalama geceleme süresi	9,4	9,0	9,1	3,5
AvNights	Son üç yıldaki ortalama geceleme süresi	9,2	8,9	9,0	3,8
DiffAvNights	Ortalama geceleme sayısı değişimi	-0,14	0,02	-0,10	1,68
DiffMonPerNight	Gecelik ödeme tutarı değişimi	13,7 \$	15,0 \$	14,5 \$	52,0
DiffAvMon	Ortalama harcama tutarı değişimi	100,7 \$	135,2 \$	122,2 \$	569,6
Age	Yaş	45,6	47,0	46,5	13,0
Toplam		4667 (%38)	7746 (%62)		

Diğer taraftan ilişki süreleri incelendiğinde, ayrılmama ihtimali yüksek olan müşterilerin ($\cong 6$ yıl, $\bar{X} = 2229,5$ gün), ayrılma ihtimali yüksek olan müşterilere ($\cong 4$ yıl, $\bar{X} = 1534,6$ gün) ve genel ortalamaya ($\cong 5$ yıl, $\bar{X} = 1968,3$ gün) kıyasla daha uzun süredir zincir

otel işletmesinin müşterisi olduğu görülmektedir. Ayrıca satın alımlar arasında geçen süreler incelendiğinde, ayrılmama ihtimali yüksek olan müşterilerin ($\cong 2$ yıl, $\bar{X} = 744,7$ gün), ayrılma ihtimali yüksek olan müşterilere ($\cong 3$ yıl, $\bar{X} = 836,0$ gün) ve genel ortalamaya ($\cong 2$ yıl, $\bar{X} = 779,0$ gün) kıyasla daha sık satın alım yaptığı görülmektedir. Benzer şekilde gerçek geceleme sayıları ayrılmama olasılığı yüksek olan müşterilerde ($\bar{X} = 38,9$) ayrılma olasılığı yüksek olan müşterilere ($\bar{X} = 23,8$) ve genel ortalamaya ($\bar{X} = 33,2$) kıyasla daha yüksektir. Buna karşılık, son üç yıldaki geceleme sayıları incelendiğinde, ayrılma olasılığı yüksek olan müşterilerin geceleme sayılarının ($\bar{X} = 14,8$) ayrılmama olasılığı yüksek olan müşterilerden ($\bar{X} = 14,6$) ve genel ortalamadan ($\bar{X} = 14,7$) daha yüksek olduğu görülmektedir. Ayrıca ortalama geceleme süreleri incelendiğinde, gerçek ortalama geceleme sürelerinin ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 9,4$) ayrılmama olasılığı yüksek olan müşterilere ($\bar{X} = 9,0$) ve genel ortalamaya ($\bar{X} = 9,1$) kıyasla daha yüksek olduğu görülmektedir. Benzer şekilde son üç yıldaki ortalama geceleme süreleri de ayrılma olasılığı yüksek olan müşterilerde ($\bar{X} = 9,2$) ayrılmama olasılığı yüksek olan müşterilerden ($\bar{X} = 8,9$) ve genel ortalamadan ($\bar{X} = 9,0$) daha yüksektir. Buna karşılık ortalama geceleme sayısı değişimleri incelendiğinde ayrılma olasılığı yüksek olan müşteriler ($\bar{X} = -0,14$) ve genel ortalama için ($\bar{X} = -0,10$) eğilimler negatif yöndeysen, ayrılmama olasılığı yüksek olan müşterilerin ($\bar{X} = 0,02$) eğilimleri pozitif yöndedir. Ayrıca, gecelik ödeme tutarlarındaki değişimin ayrılmama olasılığı yüksek olan müşterilerde ($\bar{X} = 15,0$ \$) ayrılma olasılığı yüksek olan müşterilere ($\bar{X} = 13,7$ \$) ve genel ortalamaya ($\bar{X} = 14,5$ \$) kıyasla daha yüksek olduğu görülmektedir. Benzer şekilde ortalama harcama tutarlarındaki değişimler ayrılmama olasılığı yüksek olan müşterilerde ($\bar{X} = 135,2$ \$) ayrılma olasılığı yüksek olan müşterilere ($\bar{X} = 100,7$ \$) ve genel ortalamaya ($\bar{X} = 122,2$ \$) kıyasla daha yüksektir. Müşterilerin yaşları incelendiğinde ise ayrılmayan müşterilerin yaş ortalamasının ($\bar{X} = 47,0$) ayrılan müşterilerden ($\bar{X} = 45,6$) ve genel ortalamadan ($\bar{X} = 46,5$) daha yüksek olduğu görülmektedir.

Tablo 3.15'deki bulgular incelendiğinde, ayrılma olasılığı yüksek olan müşterilerin daha çok yüksek sezonda konaklamış olduğu (%48, $f = 2251$), buna karşılık ayrılmama olasılığı yüksek olan müşterilerin daha çok orta sezonda konakladığı görülmektedir (%45, $f = 3471$). Satın alım yöntemi olarak münferit rezervasyon yaptıran müşteriler arasında ayrılmama olasılığı yüksek olan müşterilerin oranı (%12, $f = 960$) ayrılma olasılığı yüksek olan müşterilere kıyasla (%5, $f = 214$) daha fazladır. Konaklama eşlikçilerinin dağılımları incelendiğinde iki müşteri grubunda da en çok çocuklu ve çift olarak konaklama gerçekleştirildiği görülmekle

birlikte bu kategorileri ayrılmama ihtimali yüksek olan müşterilerde tek konaklamalar (%9, f=725), ayrılma ihtimali yüksek olan müşterilerde ise bebekli konaklamalar (%8, f=369) takip etmektedir.

Tablo 3.15 Rastgele Orman Tahminleri: Kategorik Değişkenler

Değişkenler		Ayrılır	%	Ayrılmaz	%	Toplam	%
Sezon	Yüksek Sezon	2251	%48	3462	%45	5713	%46
	Orta Sezon	1964	%42	3471	%45	5435	%44
	Düşük Sezon	452	%10	813	%10	1265	%10
Satın Alım Yöntemi	Acente vd.	4453	%95	6786	%88	11239	%91
	Münferit	214	%5	960	%12	1174	%9
Konaklama Eşlikçileri	Çocuklu	2227	%48	3153	%41	5380	%43
	Çift	1294	%28	2595	%34	3889	%31
	Tek	334	%7	725	%9	1059	%9
	Yetişkin Grubu	260	%6	509	%7	769	%6
	Bebekli	369	%8	368	%5	737	%6
Son Konaklanan Otel	A Hotel	1333	%29	2536	%33	3869	%31
	B Hotel	2317	%50	3016	%39	5333	%43
	C Hotel	1017	%22	2194	%28	3211	%26
Diğer Otellerde Konaklama	Var	0	%0	4266	%55	4266	%34
	Yok	4667	%100	3480	%45	8147	%66
Milliyet	Rusya	1705	%37	2136	%28	3841	%31
	Türkiye	569	%12	1796	%23	2365	%19
	Almanya	578	%12	1183	%15	1761	%14
	Ukrayna	347	%7	427	%6	774	%6
	Çekya	248	%5	323	%4	571	%5
	Romanya	160	%3	246	%3	406	%3
	İngiltere	122	%3	217	%3	339	%3
	Fransa	97	%2	132	%2	229	%2
	Diğer	841	%18	1286	%17	2127	%17
Pazar	BDT	2238	%48	2908	%38	5146	%41
	Avrupa	1493	%32	2492	%32	3985	%32
	Türkiye	570	%12	1797	%23	2367	%19
	Baltık Balkan	254	%5	362	%5	616	%5
	Diğer	112	%2	187	%2	299	%2
Sadakat Kartı	Silver	2302	%49,3	3206	%41,4	5508	%44,4
	Silver Plus	223	%4,8	1226	%15,8	1449	%11,7
	Gold	6	%0,1	230	%3,0	236	%1,9
	Platinum	1	%0,0	114	%1,5	115	%0,9
	Kart Yok	2135	%45,7	2970	%38,3	5105	%41,1

Son konaklamanın gerçekleştirildiği otel kategorilerinde ayrılmama ihtimali yüksek olan müşterilerin A Hotel (%33, f=2536) ve C Hotel (%28, f=2194) konaklamalarındaki oranı ayrılma ihtimali yüksek olan müşterilere kıyasla daha yüksek olmasına karşın B Hotel

konaklamalarının oranı ayrılma ihtimali yüksek olan müşterilerde (%50, f=2317) daha yüksektir. Ayrıca, ayrılma ihtimali yüksek olan müşteriler arasında grubun diğer otellerinde hiç konaklama yapıldığı görülmezken, ayrılmama ihtimali yüksek olan müşterilerde bu oran %55 (f=4266) olarak ortaya çıkmıştır.

Müşterilerin milliyetleri incelendiğinde, ayrılma ihtimali yüksek olan müşteriler arasında milliyeti Rusya olanların oranı (%37, f=1705) ayrılmama ihtimali yüksek olan müşterilere (%28, f=2136) kıyasla daha yüksektir. Diğer taraftan, ayrılma ihtimali yüksek olan müşteriler arasında milliyeti Türkiye (%12, f=569) ve Almanya (%12, f=578) olanların oranının ayrılmama ihtimali yüksek olan müşterilere (sırasıyla %23, f=1796 ve %15, f=1183) kıyasla daha düşük olduğu görülmektedir. Ayrıca, BDT pazarından gelen müşterilerin ayrılma olasılığı yüksek olan müşteriler arasındaki oranı (%48, f=2238) ayrılmama ihtimali yüksek olan müşterilere kıyasla (%38, f=2908) daha yüksek olmasına karşın, Türkiye pazarından gelenler için bu durumun tersi söz konusudur (sırasıyla %12, f=570 ve %23, f=1797). Müşterilerin üyelik durumları incelendiğinde, ayrılma ihtimali yüksek olan müşterilerde Gold (%0, f=6) ve Platinum (%0, f=1) üyeliğine sahip müşteriler neredeyse görülmezken, ayrılmama ihtimali yüksek olan müşterilerde bu müşterilerin oranları daha yüksektir (sırasıyla %3, f=230 ve %1, f=114).

İşletme kaynaklarının sınırlı olduğu ve işletmelerin ayrılma potansiyeli olan tüm müşterilerine aynı öncelikte teşvik sunamayacağı bilinen bir gerçektir. Dolayısıyla çalışmaya konu olan zincir otel işletmesinin ayrılma ihtimali yüksek olan müşterilerini bir öncelik sıralamasına koyması ve teşviklerini bu müşterilerinden elde edeceği potansiyel kazanç ile kıyaslama yaparak sunması daha yerinde bir strateji olacaktır. Bu bağlamda Qi vd. (2009: 260) tarafından önerilen öncelik matrisinin bir uyarlaması yapılmıştır (Şekil 3.16). Bu matris üzerinde müşterilerin ayrılma olasılığının düşük (%75 ve altı) ve yüksek (%75 üzeri) olması durumuyla, müşterinin geri gelmesi durumunda otele kazandıracağı tahmini değer (R_AverageMont) düşük (ortalamanın altında) ve yüksek (ortalamanın üstünde) olması durumları kıyaslanmıştır.

Bu matriste öncelik düzeyleri Qi vd. (2009: 260) tarafından önerilen kriterler kullanılarak belirlenmiştir:

- Kriter 1: Aynı müşteri değer düzeyinde daha yüksek ayrılma olasılığı bulunan müşteriler daha önceliklidir.
- Kriter 2: Aynı müşteri ayrılma olasılığı düzeyinde daha yüksek müşteri değerine sahip müşteriler daha önceliklidir.

- Kriter 3: Müşterinin ayrılma olasılığının yüksek olması, müşteri değerinin yüksek olmasına kıyasla daha önemlidir.

Dolayısıyla, Şekil 3.16'da sunulan matris üzerindeki dört müşteri segmentinin öncelik sıralamaları $1>2>3>4$ şeklinde belirlenmiştir. Bu grupların betimsel analizleri Tablo 3.16 ve Tablo 3.17'de sunulmuştur.

Şekil 3.16, Tablo 3.16 ve Tablo 3.17'de sunulan bulgular birlikte değerlendirildiğinde, bu dört müşteri segmenti için aşağıdaki değerlendirmelerde bulunulabilir:

Öncelik kategorisi 1: Ayrılma ihtimali yüksek ($\bar{X} = \%80$) ve müşteri değeri yüksek ($\bar{X} = 1.946$ \$) olan müşteriler öncelik kategorisi 1 grubunda yer almaktadır. Bu müşteriler öğrenme algoritması tarafından “ayrılır” olarak kodlanan toplam müşterilerin %9,3’ünü oluşturmaktadır. Öncelik kategorisi 1 grubunda yer alan müşterilerin ortalamaları incelendiğinde, en son 1 yıl 2 ay kadar zaman önce konaklama yaptıkları ($\bar{X} = 442,7$ gün), ağırlıklı BDT (%49) ve Avrupa (%40) pazarından geldikleri, gecelik harcamalarının ($\bar{X} = 195,94$ \$) ve konaklama başına ortalama geceleme sürelerinin ($\bar{X} = 10,4$) ortalamanın üstünde, 2018 yılı sonuna kadarki toplam harcamalarının ($\bar{X} = 4.173,13$ \$) ve toplam geceklerinin ($\bar{X} = 22,2$) ise ortalamanın altında olduğu, ağırlıklı olarak yüksek sezonda (%62), A Hotel’de (%39) ve çocuklu (%59) konakladıkları, en çok Silver üyelerin bulunduğu (%57), zincir otel işletmesinin bünyesindeki otellerde yaklaşık 2 yıl 9 ay süre aralıklarla konaklama yaptıkları ($\bar{X} = 1001,5$), satın alım yöntemi olarak münferit rezervasyon tercih etmedikleri (%0), yaklaşık 4 yıldan fazla süredir otel zincirinin müşterisi oldukları ($\bar{X} = 1555,6$) ve toplam geliş sayısı ortalaması en düşük olan ($\bar{X} = 2,13$) müşteriler oldukları görülmektedir.

Öncelik kategorisi 1 grubundaki müşterilerin son üç yıldaki eğilimleri incelendiğinde ortalama geceleme sayılarında ($\bar{X} = -0,32$), gecelik ödeme tutarlarında ($\bar{X} = -8,58$ \$) ve ortalama harcamalarında ($\bar{X} = -133,19$ \$) diğer kategorilere ve genel ortalamaya kıyasla ciddi bir düşüş görülmektedir. Bununla birlikte son üç yıldaki konaklama başına ortalama geceleme süreleri ($\bar{X} = 10,1$), ortalama harcamaları ($\bar{X} = 1.813,20$ \$), gecelik ödeme tutarları ($\bar{X} = 187,37$ \$) ve son üç yıldaki toplam harcamaları ($\bar{X} = 2.363,63$ \$) genel ortalamanın üstündedir. Diğer taraftan öncelik kategorisi 1 grubundaki müşterilerin son üç yıldaki toplam geceleme sayıları ($\bar{X} = 13,9$) ortalamanın altında olup, son üç yıldaki geliş sayısı ($\bar{X} = 1,31$) diğer kategorilere kıyasla en düşük ortalamaya sahiptir.

Öncelik kategorisi 2: Ayrılma ihtimali yüksek ($\bar{X} = \%81$) ve müşteri değeri düşük ($\bar{X} = 643,10$ \$) olan müşteriler öncelik kategorisi 2 grubunda yer almaktadır. Bu müşteriler öğrenme algoritması tarafından “ayrılır” olarak kodlanan toplam müşterilerin %14,3’ünü oluşturmaktadır.

Tablo 3.16 Öncelik Grupları Betimsel İstatistikleri: Nümerik Değişkenler

		Kategori 1	Kategori 2	Kategori 3	Kategori 4	$\bar{X}$
Değişkenler	N	433	668	1666	1900	
	Frekans	% 9,3	% 14,3	% 35,7	% 40,7	
	Ayrılma Olasılığı	% 80,0	% 80,6	% 63,6	% 64,2	
R_AverageMon	Kazandıracağı Değer	1.946,39 \$	643,10 \$	2.096,86 \$	718,86 \$	1.203,7 \$
LOR	İlişki süresi	1555,6	1399,6	1588,0	1530,5	1968,3
Recency	Güncellik	442,7	446,8	295,2	335,6	467,7
Interpurchase	Satın alımlar arası geçen zaman	1001,5	814,7	843,6	799,0	779,0
R_Monetary	Gerçek harcama tutarı	4.173,13 \$	1.380,72 \$	5.680,01 \$	1.844,84 \$	4.187,5 \$
Monetary	Son üç yıldaki harcama tutarı	2.363,63 \$	1.070,31 \$	3.228,67 \$	1.450,64 \$	2.002,2 \$
R_MonPerNight	Gerçek gecelik ödeme tutarı	195,94 \$	85,91 \$	196,62 \$	90,21 \$	127,8 \$
MonPerNight	Son üç yıldaki gecelik ödeme tutarı	187,37 \$	95,27 \$	200,65 \$	119,11 \$	142,3 \$
AverageMon	Son üç yıldaki konaklama başına harcama tutarı	1.813,20 \$	693,48 \$	2.123,56 \$	955,46 \$	1.325,9 \$
R_Repeat	Gerçek geliş sayısı	2,13	2,18	2,68	2,60	3,5
Repeat	Son üç yıldaki geliş sayısı	1,31	1,63	1,56	1,68	1,61
R_Nights	Gerçek geceleme sayısı	22,2	16,5	29,6	21,7	33,2
Nights	Son üç yıldaki geceleme sayısı	13,9	12,2	17,0	14,0	14,7
R_AvNights	Gerçek ortalama geceleme sayısı	10,4	7,7	10,9	8,4	9,1
AvNights	Son üç yıldaki ortalama geceleme sayısı	10,1	7,4	10,8	8,3	9,0
DiffAvNights	Ortalama geceleme sayısı değişimi	-0,32	-0,29	-0,14	-0,05	-0,10
DiffMonPerNight	Gecelik ödeme tutarı değişimi	-8,58 \$	9,36 \$	4,02 \$	28,90 \$	14,5 \$
DiffAvMon	Ortalama harcama tutarı değişimi	-133,19 \$	50,38 \$	26,70 \$	236,60 \$	122,2 \$
Age	Yaş	44,7	44,8	46,4	45,4	46,5

$\bar{X}$ = Ayrılır ve ayrılmaz olarak etiketlenmiş bütün müşterilerin genel ortalamasıdır.

Öncelik kategorisi 2 grubunda yer alan müşterilerin ortalamaları incelendiğinde, en son yaklaşık 1 yıl 3 ay kadar zaman önce konaklama yaptıkları ($\bar{X}$ = 446,8 gün), ağırlıklı Avrupa (%36) pazarından geldikleri görülmektedir. 2018 yılı sonuna kadarki toplam harcamaları ($\bar{X}$ = 1.380,72 \$), gecelik harcamaları ($\bar{X}$ = 85,91 \$), konaklama başına ortalama geceleme sürelerinin ($\bar{X}$ = 7,7), toplam geceleme süreleri ($\bar{X}$ = 16,5) ve tekrar gelişinde kazandırması beklenen değer ($\bar{X}$ = 643,10 \$) en düşük olduğu müşteri segmentidir. Ağırlıklı olarak orta sezonda

(%45), A Hotel’de (%49) ve çocuklu (%40) konakladıkları, en çok üye olmayan müşterilerin bulunduğu (%77), zincir otel işletmesinin bünyesindeki otellerde yaklaşık 2 yıl 3 ay süre aralıklarla konaklama yaptıkları ($\bar{X} = 814,7$), satın alım yöntemi olarak münferit rezervasyon tercih etmedikleri (%0) ve yaklaşık 3 yıl 10 ay kadar zamandır otel zincirinin müşterisi oldukları ($\bar{X} = 1399,6$) görülmektedir. Ayrıca en çok Türk müşterinin bulunduğu müşteri segmentidir (%27).

Öncelik kategorisi 2 grubundaki müşterilerin son üç yıldaki eğilimleri incelendiğinde ortalama geceleme sayılarında ($\bar{X} = -0,29$) düşüş görülmekle birlikte gecelik ödeme tutarlarının ($\bar{X} = -9,36$ \$) ve ortalama harcamalarının ($\bar{X} = 50,38$ \$) artış eğilimi gösterdiği görülmektedir. Bununla birlikte son üç yıldaki konaklama başına ortalama geceleme süreleri ($\bar{X} = 7,4$), ortalama harcamaları ($\bar{X} = 693,48$ \$), gecelik ödeme tutarları ($\bar{X} = 95,27$ \$), son üç yıldaki toplam harcamaları ($\bar{X} = 1.070,31$ \$) ve son üç yıldaki toplam geceleme sayıları ($\bar{X} = 12,2$) en düşük olan müşteri segmentidir. Buna karşılık öncelik kategorisi 2 grubundaki müşterilerin son üç yıldaki geliş sayıları ($\bar{X} = 1,63$) genel ortalamanın üstündedir.

Öncelik kategorisi 3: Ayrılma ihtimali düşük ($\bar{X} = \%64$) ve müşteri değeri yüksek ($\bar{X} = 2.096,86$ \$) olan müşteriler öncelik kategorisi 3 grubunda yer almaktadır. Bu müşteriler öğrenme algoritması tarafından “ayrılır” olarak kodlanan toplam müşterilerin %35,7’sini oluşturmaktadır. Öncelik kategorisi 3 grubunda yer alan müşterilerin ortalamaları incelendiğinde, en son yaklaşık 10 ay kadar zaman önce konaklama yaptıkları ($\bar{X} = 295,2$ gün) ve ağırlıklı BDT (%62) pazarından geldikleri görülmektedir. 2018 yılı sonuna kadarki toplam harcamaları ($\bar{X} = 5.680,01$ \$), gecelik harcamaları ($\bar{X} = 196,62$ \$), konaklama başına ortalama geceleme süreleri ($\bar{X} = 10,9$), toplam geceleme süreleri ($\bar{X} = 29,6$) ve tekrar gelişinde kazandırması beklenen değer ($\bar{X} = 2.096,86$ \$) en yüksek olduğu müşteri segmentidir. Toplam geliş sayıları diğer segmentlere kıyasla en yüksek olan ($\bar{X} = 2,68$) bu müşteri segmentinin ağırlıklı olarak yüksek sezonda (%62), B Hotel’de (%54) ve çocuklu (%62) konakladıkları ve Silver Plus üyelerin en fazla bulunduğu (%10) segment olduğu görülmektedir.

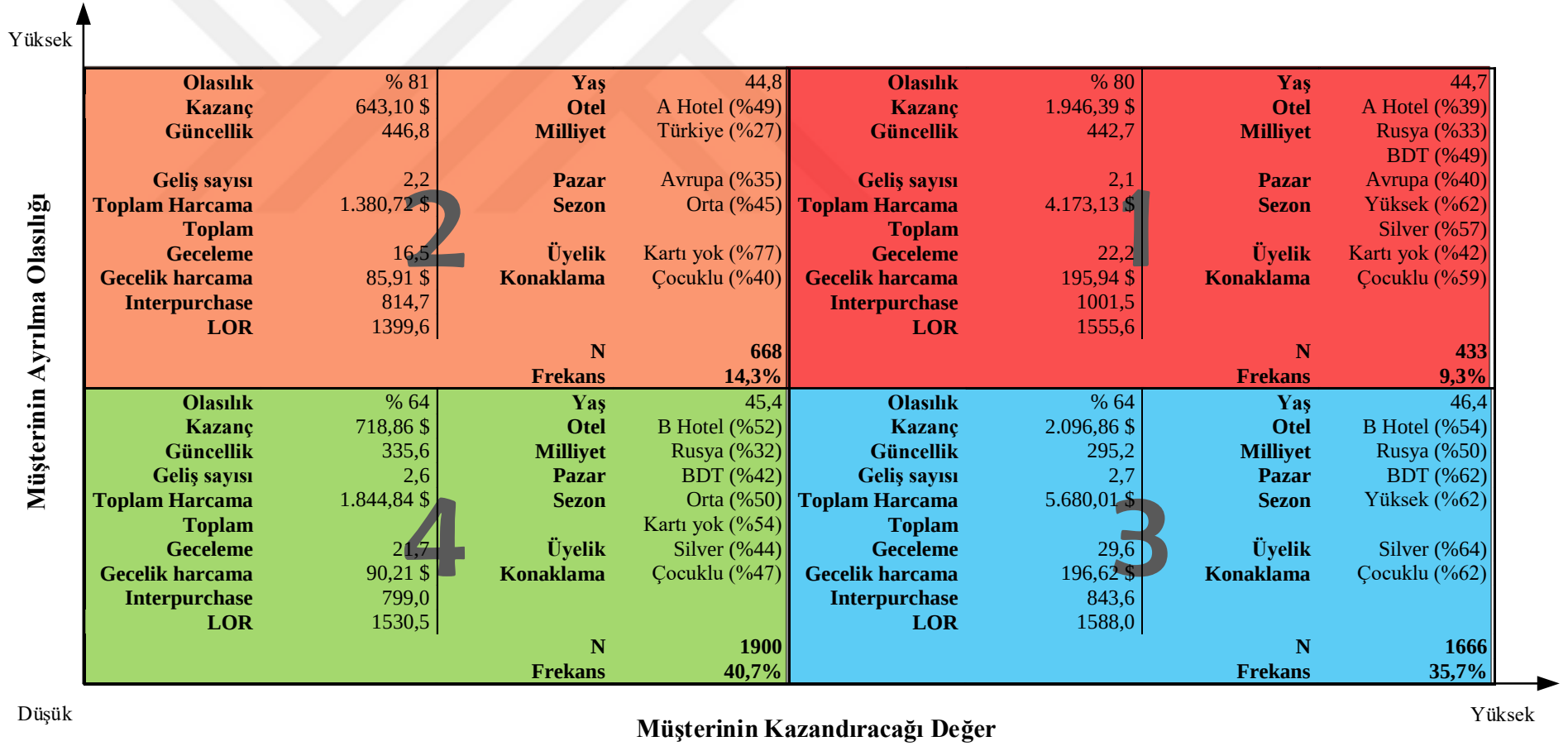
Öncelik kategorisi 3 grubundaki müşteriler zincir otel işletmesinin bünyesindeki otellerde yaklaşık 2 yıl 4 ay süre aralıklarla konaklama yaptıkları ($\bar{X} = 843,6$), ağırlıklı olarak acente vd. aracılığıyla rezervasyon yaptırdıkları (%97) ve yaklaşık 4 yıl 4 ay kadar zamandır otel zincirinin müşterisi oldukları ($\bar{X} = 1588$) görülmektedir. Öncelik kategorisi 3 grubundaki müşterilerin son üç yıldaki eğilimleri incelendiğinde ortalama geceleme sayılarında ($\bar{X} = -0,14$) düşüş görülmekle birlikte gecelik ödeme tutarlarının ($\bar{X} = 4,02$ \$) ve ortalama harcamalarının ($\bar{X} = 26,70$ \$) küçük bir artış eğilimi gösterdiği görülmektedir.

Tablo 3.17 Öncelik Grupları Betimsel İstatistikleri: Kategorik Değişkenler

		Kategori 1		Kategori 2		Kategori 3		Kategori 4	
Sezon	Yüksek Sezon	269	%62	197	%29	1032	%62	753	%40
	Orta Sezon	134	%31	299	%45	580	%35	951	%50
	Düşük Sezon	30	%7	172	%26	54	%3	196	%10
Satın Alım Yöntemi	Acente vd.	432	%100	667	%100	1608	%97	1746	%92
	Münferit	1	%0	1	%0	58	%3	154	%8
Konaklama Eşlikçileri	Çocuklu	253	%58	233	%35	995	%60	746	%39
	Çift	131	%30	191	%29	414	%25	558	%29
	Tek	10	%2	128	%19	43	%3	153	%8
	Yetişkin Grubu	28	%6	19	%3	129	%8	84	%4
	Bebekli	59	%14	51	%8	142	%9	117	%6
Son Konaklanan Otel	A Hotel	170	%39	327	%49	322	%19	514	%27
	B Hotel	160	%37	253	%38	907	%54	997	%52
	C Hotel	103	%24	88	%13	437	%26	389	%20
Diğer Otellerde Konaklama	Var	0	%0	0	%0	0	%0	0	%0
	Yok	433	%100	668	%100	1666	%100	1900	%100
Milliyet	Rusya	142	%33	121	%18	829	%50	613	%32
	Almanya	63	%15	86	%13	170	%10	259	%14
	Türkiye	6	%1	180	%27	23	%1	360	%19
	Ukrayna	40	%9	41	%6	133	%8	133	%7
	Çekya	28	%6	49	%7	77	%5	94	%5
	Romanya	19	%4	18	%3	56	%3	67	%4
	İngiltere	10	%2	19	%3	44	%3	49	%3
	Fransa	11	%3	10	%1	49	%3	27	%1
	Diğer	114	%26	144	%22	285	%17	298	%16
Pazar	BDT	214	%49	191	%29	1034	%62	799	%42
	Avrupa	175	%40	239	%36	486	%29	593	%31
	Türkiye	6	%1	180	%27	23	%1	361	%19
	Baltık Balkan	35	%8	41	%6	86	%5	92	%5
	Diğer	3	%1	17	%3	37	%2	55	%3
Sadakat Kartı	Silver	246	%57	155	%23	1071	%64	830	%44
	Kart Yok	181	%42	512	%77	425	%26	1017	%54
	Silver Plus	6	%1	1	%0	166	%10	50	%3
	Gold	0	%0	0	%0	4	%0	2	%0
	Platinum	0	%0	0	%0	0	%0	1	%0

Ayrıca son üç yıldaki konaklama başına ortalama geceleme süreleri ($\bar{X} = 10,8$), ortalama harcamaları ($\bar{X} = 2.123,56$ \$), gecelik ödeme tutarları ($\bar{X} = 200,65$ \$), son üç yıldaki toplam harcamaları ($\bar{X} = 3.228,67$ \$) ve son üç yıldaki toplam geceleme sayıları ($\bar{X} = 17,0$) en yüksek olan müşteri segmentidir. Buna karşılık öncelik kategorisi 3 grubundaki müşterilerin son üç yıldaki geliş sayıları ($\bar{X} = 1,56$) genel ortalamanın altındadır.

Öncelik kategorisi 4: Ayrılma ihtimali düşük ($\bar{X} = \%64$) ve müşteri değeri düşük ($\bar{X} = 718,86$ \$) olan müşteriler öncelik kategorisi 4 grubunda yer almaktadır. Bu segment öğrenme algoritması tarafından “ayrılır” olarak kodlanan toplam müşterilerin $\%40,7$ 'sini oluşturmaktadır. Öncelik kategorisi 4 grubunda yer alan müşterilerin ortalamaları incelendiğinde, en son yaklaşık 11 ay kadar zaman önce konaklama yaptıkları ($\bar{X} = 335,6$ gün) ve ağırlıklı BDT ($\%42$) pazarından geldikleri görülmektedir. 2018 yılı sonuna kadarki toplam harcamaları ($\bar{X} = 1.844,84$ \$), gecelik harcamaları ($\bar{X} = 90,21$ \$), konaklama başına ortalama geceleme süreleri ($\bar{X} = 8,4$) ve toplam geceleme süreleri ($\bar{X} = 21,7$) ortalamanın altındadır. Öncelik kategorisi 4 grubunda yer alan müşterilerin ağırlıklı olarak orta sezonda ($\%50$), B Hotel’de ($\%52$) ve çocuklu ($\%39$) olarak konakladığı görülmektedir. Bu müşteri segmentinde üye olmayan müşteriler ($\%54$) ve Silver üye müşteriler ($\%44$) ağırlıklı olarak bulunmaktadır. Öncelik kategorisi 4 segmenti zincir otel işletmesinin bünyesindeki otellerde yaklaşık 2 yıl 2 ay süreyle ($\bar{X} = 799,0$) en sık konaklama yapan ve diğer segmentlere kıyasla en çok münferit rezervasyon tercih eden ($\%8$) müşteri segmentidir. Bu segmentte yer alan müşterilerin yaklaşık 4 yıl 2 ay kadar zamandır otel zincirinin müşterisi oldukları ($\bar{X} = 1530,5$) görülmektedir. Öncelik kategorisi 4 grubundaki müşterilerin son üç yıldaki eğilimleri incelendiğinde gecelik ödeme tutarlarında ($\bar{X} = 28,90$ \$) ve ortalama harcamalarında ($\bar{X} = 236,60$ \$) en yüksek artışın ve ortalama geceleme sayılarında ($\bar{X} = -0,05$) en az düşüşün bulunduğu müşteri segmenti olduğu görülmektedir. Diğer taraftan son üç yıldaki konaklama başına ortalama geceleme süreleri ($\bar{X} = 8,3$), ortalama harcamaları ($\bar{X} = 955,46$ \$), gecelik ödeme tutarları ($\bar{X} = 119,11$ \$), son üç yıldaki toplam harcamaları ($\bar{X} = 1.450,64$ \$) ve son üç yıldaki toplam geceleme sayıları ($\bar{X} = 14,0$) genel ortalamanın altındadır. Buna karşılık öncelik kategorisi 4 grubu son üç yıldaki geliş sayıları ($\bar{X} = 1,68$) en yüksek olan müşteri segmentidir.



Şekil 3.17 Öncelik Karar Matrisi

SONUÇ

Bu araştırmada makine öğrenimi metodlarından faydalanılarak bir zincir otel işletmesinin devamlı (repeat) müşterileri için kayıp müşteri tahminlemesi uygulanmıştır. Çalışmaya konu olan zincir otel işletmesi bünyesinde Antalya, Belek bölgesinde faaliyet gösteren beş yıldızlı üç resort otel bulunmaktadır (çalışma kapsamında A Hotel, B Hotel ve C Hotel olarak adlandırılmıştır). Bu resort otellerin pazarlama çalışmalarında A Hotel için çiftlere hitap eden romantik otel teması, B Hotel için çocuklu ailelere hitap eden aile oteli teması ve C Hotel için daha yüksek bütçeli müşterilere hitap eden lüks konaklama teması kullanılmaktadır. Bu otellerin hepsi bünyesinde çocuk büfesi ve çocuk kulübü bulunan aile dostu oteller olup herşey dahil konseptte hizmet sunmaktadır. Bununla birlikte zincir otel otel işletmesinin sadakat programı kapsamında otel müşterilerinin kazandıkları kart avantajları ve sadakat puanları üç otelde de geçerli olmaktadır. Dolayısıyla otel işletmesinin müşterileri A Hotel, B Hotel ya da C Hotel'den herhangi birinde toplamda en az 10 gece süreyle iki defa konaklama gerçekleştirdiklerinde sadakat puanı kazanmaya başlamakta ve 20 gece konaklama sonrasında Silver üyelik sahibi olmaktadır. Burdan hareketle, bu çalışmada zincir otel işletmesinin bünyesindeki otellerde en az iki defa olmak üzere toplamda 10 gece konaklama yapan ve araştırma modeli kapsamında belirlenen periyotlarda aktif olan devamlı müşteriler için kayıp müşteri tahminlemesi yapılmıştır.

Araştırmanın öncelikli amacı daha önce farklı sektörlerde gerçekleştirilmiş ancak turizm alanında sınırlı örneklerin bulunduğu kayıp müşteri tahminleme yöntemlerinin bir zincir otel işletmesine uyarlanmasıdır. Böylelikle, tahmin modelinin bulgularından faydalanılarak potansiyel kayıp müşteriler arasında bir önceliklendirme yapılması ve bu müşterileri elde tutmaya yönelik çözüm önerilerinde bulunulması hedeflenmiştir. Bununla birlikte daha önce turizm alanında sadece sosyal kulüp üyeleri (Maier ve Prusty, 2016:803), ticari otel müşterileri (şirketler) (De Wit, 2017: 2) ve otel rezervasyon web sitesi müşterileri (Han, 2018:1) için gerçekleştirilmiş kayıp müşteri tahminleme yönteminin bir resort otel zincirinin müşteri verilerine uyarlanmasıyla bu araştırmanın mevcut literatüre katkı sağlaması amaçlanmıştır.

Bu amaçla öncelikle iki farklı makine öğrenimi algoritması olan Lojistik Regresyon ve Rastgele Orman metodlarının karşılaştırması yapılarak, otel işletmelerine yönelik müşteri verileriyle gerçekleştirilen bir kayıp müşteri tahminlemesinde hangi yöntemin daha başarılı olduğu belirlenmiştir. Model performansının değerlendirilmesinde kullanılan AUC (ROC

eğrisi altında kalan alan) değerlerine göre Rastgele Orman algoritmasının 0,80 AUC değeriyle daha yüksek performansa sahip olduğu görülmüştür. Bu bağlamda, Rastgele Orman modeli, test setinden tesadüfi olarak seçilen bir ayrılan müşteriye %80 ihtimalle doğru olarak sınıflandırabilmektedir (Maier ve Prusty, 2016: 812). Lojistik Regresyon algoritması kullanılarak oluşturulan modelin performansı ise 0,74 AUC değerinde kalmıştır.

Turizm alanında daha önce gerçekleştirilen kayıp müşteri tahminleme çalışmalarının bulguları değerlendirildiğinde, bu çalışmanın sonuçları Han (2018: 3) tarafından otel rezervasyon web sitelerinde Lojistik Regresyon ve Rastgele Orman yöntemleri kullanılarak gerçekleştirilen kayıp müşteri tahminleme analizinin sonuçlarıyla benzerlik göstermektedir. Han (2018: 3) tarafından gerçekleştirilen çalışmanın sonuçlarına göre Rastgele Orman algoritması 0,77'lik AUC değeriyle Lojistik Regresyon algoritmasından ($AUC = 0,75$) daha iyi performans göstermiştir. Benzer şekilde, Buckinx ve Van den Poel (2005: 262) tarafından kontratsız çalışılan bir hızlı tüketim ürünleri şirketinin en sadık müşterileri üzerinde Rastgele Orman, Lojistik Regresyon ve ARD Sinir Ağları (automatic relevance determination neural networks) metodları kullanılarak uygulanan kayıp müşteri tahminleme çalışmasında 0,832 AUC değeriyle en iyi performansı Rastgele Orman algoritması göstermiş olup, Lojistik Regresyon algoritması ($AUC = 0,828$) en düşük performansı göstermiştir. Ayrıca, Coussement ve De Bock (2013: 1633) tarafından kumarhane sektöründe birçok farklı öğrenme algoritmasının kıyaslandığı kayıp müşteri tahminleme çalışmasının sonuçlarına göre en yüksek performansı yine Rastgele Orman algoritması göstermiştir. Diğer taraftan, De Wit (2017: 41) tarafından otel işletmelerinde kontratsız çalışan ticari şirketlere (B2B müşteriler) lojistik regresyon, AdaBoost ve destek vektör makinesi metodları kullanılarak uygulanan kayıp müşteri tahminleme çalışmasında en iyi performans gösteren algoritma 0,75 AUC değeriyle Lojistik Regresyon olmuştur (test seti 849 müşteri). Benzer şekilde Maier ve Prusty (2016: 812) tarafından sosyal klüplerde Karar Ağacı ve Lojistik Regresyon kullanılarak gerçekleştirilen kayıp müşteri tahminleme çalışmasında Lojistik Regresyon algoritması 0,793 AUC değeriyle Karar Ağacı algoritmasından ($AUC = 0,789$) daha iyi performans göstermiştir (test seti 936 müşteri). Lojistik Regresyon algoritması Rastgele Orman algoritmasından daha hızlı çalışması ve daha yüksek performans düzeylerine ulaşması nedeniyle pazarlama yöneticileri tarafından rağbet gösterilen bir sınıflandırma tekniği olarak öne çıkabilir. Ancak, veri setinin daha zengin olduğu ve daha fazla değişkenin bulunduğu durumlarda Rastgele Orman algoritması Lojistik Regresyondan çok daha iyi performans gösterme eğilimindedir. Ayrıca, çok sayıda yordayıcı değişkenin bulunduğu durumlarda Rastgele Orman algoritmasının ezberleme sorunu (overfitting) göstermeyeceğinden emin olunurken Lojistik Regresyon söz konusu olduğunda

ezberleme sorununun oluşmayacağı garanti edilememektedir (Miguéis vd. 2012b: 348-349). Ayrıca, bir öğrenme modelini uygulamaya koymadan önce alternatif modellerin de denenmesi en iyi performans gösteren algoritmanın bulunması açısından önem arz etmektedir (Neslin vd., 2006: 209). Dolayısıyla otel işletmelerinde gerçekleştirilecek kayıp müşteri tahminleme analizlerinde Rastgele Orman algoritmasının da alternatif öğrenme modelleri arasında bulunması ve performansının test edilmesi doğru bir yaklaşım olacaktır.

Ayrıca, bu çalışmada otel işletmelerine yönelik müşteri verileriyle gerçekleştirilen bir kayıp müşteri tahminleme analizinde hangi değişkenlerin daha önemli olduğu belirlenmiştir. Bu bağlamda Rastgele Orman algoritmasının değişken önem sıralamasını bulmaya yarayan fonksiyonundan faydalanılmıştır. Rastgele Orman algoritması tarafından hesaplanan sonuçlara göre değişken seçimi sonrasında modele giren değişkenler arasında en önemli bulunan değişken güncellik (Recency) değişkenidir. Bu değişkeni önem sırasıyla satın alımlar arasında geçen zamanların ortalaması (Interpurchase), son üç yıldaki toplam geceleme sayısı (Nights), ilişki süresi (LOR), son üç yıldaki ortalama harcama tutarındaki değişim (DiffAvMon) ve toplam geliş sayısı (R_Repeat) değişkenleri izlemektedir. Bu değişkenlerden güncellik değişkeni RFM olarak anılan değişkenlerden birincisi olup, geceleme sayısı ve geliş sayısı değişkenleri Frequency - Sıklık değişkeni temelli değişkenlerdir. Harcama tutarındaki değişimi gösteren değişken ise Monetary-Tutar değişkeni temelli bir değişkendir. Bu bulgular RFM temelli geleneksel değişkenlerin en önemli değişkenler olarak ortaya çıktığı daha önce gerçekleştirilen kayıp müşteri tahminleme çalışmalarının sonuçlarıyla uyumludur (Coussement ve De Bock, 2013: 1634; De Wit, 2017: 43; Buckinx ve Van den Poel, 2005: 263).

Örneğin, Coussement ve De Bock (2013: 1634) tarafından kumarhane sektöründe gerçekleştirilen kayıp müşteri tahminleme çalışmasının sonuçlarına göre Rastgele Orman algoritması tarafından en önemli bulunan değişken “güncellik” değişkeni olmuştur, bunu güncellik temelli değişkenler, satın alımlar arasında geçen zaman değişkeni, sıklık ve tutar temelli değişkenler takip etmiştir. Ayrıca ilişki süresi değişkeni de en önemli bulunan değişkenler arasında yer almaktadır. Benzer şekilde De Wit (2017: 43) tarafından gerçekleştirilen çalışmanın bulgularına göre Lojistik Regresyon analizi sonucunda güncellik değişkeni en önemli yordayıcı değişken olarak bulunmuş, bunu sırasıyla şirket tarafından ziyaret edilen otel sayısı, ilişki süresi, satın alımlar arasında geçen zaman ve sıklık değişkenleri takip etmiştir. Ayrıca, Buckinx ve Van den Poel (2005: 263) tarafından gerçekleştirilen çalışmada Rastgele Orman algoritmasının önem düzeylerine göre RFM temelli değişkenler en önemli bulunan yordayıcı değişkenler olarak ortaya çıkmıştır. Bu değişkenler sırasıyla sıklık,

satın alımlar arasında geçen sürelerin ortalaması, tüm gelişlerdeki sıklık, ilişki süresi, son aydaki sıklık, satın alımlar arasında geçen sürelerin standart sapması ve tutar değişkenleridir.

Diğer taraftan bu çalışmada gerçekleştirilen Lojistik Regresyon analizi tarafından modele alınan değişkenler farklılık göstermiş olup hesaplanan standartlaştırılmış β katsayılarına göre en önemli bulunan değişken zincir otel işletmesinin diğer otellerinde de konaklamış olmak (OHID) değişkeni olmuştur. Bu değişkeni gerçek geliş sayısı, çocuklu konaklamak, çift olarak konaklamak ve Türkiye vatandaşı olmak değişkenleri takip etmektedir. Bu bulgular De Wit'in (2017: 43) çalışmasıyla kısmen benzerlik göstermektedir. De Wit'in (2017: 43) çalışmasında ikinci en önemli değişken olarak bulunan "şirket tarafından ziyaret edilen otel sayısı" değişkeni mevcut çalışmadaki "zincir otel işletmesinin diğer otellerinde de konaklamış olmak" değişkeniyle benzerdir. De Wit'in (2017:43) bulgularına göre şirketlerin ziyaret ettiği otel sayısı arttıkça ayrılma olasılığı azalmaktadır. Benzer şekilde, bu çalışmanın bulgularına göre müşterilerin zincir otel işletmesinin diğer otellerinde konaklamış olması müşterinin ayrılma olasılığını %99 ($|1-\text{odds}| = 0,995$) oranında düşürmektedir. Müşterilerin belirli bir markaya yönelik sadakatlerinin oluşması durumunda başka alternatifleri değerlendirme ve yeni arayışlarda bulunma davranışları azalmakta, tekrar satın alma niyetleri ise artmaktadır (Ahmad vd., 2016: 90). Dolayısıyla, bu bulgu müşterilerin marka sadakatine işaret ediyor olabilir. Diğer taraftan hızlı tüketim ürünleri alanında gerçekleştirilen bir kayıp müşteri tahminleme analizinin sonuçlarına göre RFM temelli değişkenler sadık müşterileri tahminlemede aynı markanın ürünlerini satın alma ve üyelik puanları değişkenlerinden daha başarılı yordayıcılar olarak bulunmuştur (Buckinx ve Van den Poel, 2005: 264). Bu sonuç, hizmet sektöründe gerçekleştirilen çalışmaların bulgularıyla perakende sektöründe gerçekleştirilen çalışmalar arasında farklılık oluşabileceğini göstermektedir.

Bu çalışmanın bulgularına göre gerçek geliş sayısındaki 1 birimlik artış müşterinin ayrılma olasılığını %15 oranında düşürmektedir ($|1-\text{odds}| = 0,146$). Bu sonucu destekler şekilde, daha önce gerçekleştirilen çalışmaların sonuçlarına göre daha çok satın alım yapan müşterilerin tekrar satın alma olasılıkları diğer müşterilere kıyasla daha yüksek olma eğilimindedir (Van den Poel, 2003: 8). Diğer taraftan, bu çalışmada önemli yordayıcılar olarak ortaya çıkan çocuklu konaklama durumu müşterinin ayrılma olasılığını %97 ($|1-\text{odds}| = 0,969$) oranında, çift konaklama durumu ise %90 oranında ($|1-\text{odds}| = 0,897$) arttırmaktadır. Konaklama eşlikçileri değişkenine benzer bir değişken De Wit'in (2017: 43) çalışmasında önemli bir yordayıcı olarak bulunan yetişkin sayısı değişkenidir. Çalışmanın bulgularına göre yetişkin sayısındaki artış ayrılma olasılığını arttırmaktadır. Mevcut çalışmada ise yetişkin grubu olarak konaklama durumunun ayrılma olasılığını %52 oranında ($|1-\text{odds}| = 0,523$) arttırdığı görülmüştür. Bu iki

bulgu benzerlik göstermektedir. Diğer taraftan De Wit'in (2017:43) çalışmasında önemli bir yordayıcı olarak ortaya çıkan rezervasyonu gerçekleştiren kişi/şirket değişkeni mevcut çalışmadaki satın alım yöntemi değişkeniyle benzerlik göstermektedir. Mevcut çalışmanın bulgularına göre rezervasyonun müşteri tarafından münferit olarak gerçekleştirilmesi durumu müşterinin ayrılma olasılığını %40 oranında ($|1-\text{odds}| = 0,399$) azaltırken, De Wit'in (2017:43) bulgularına göre satın alımların şirket yerine müşteri ya da acente tarafından gerçekleştirilmesi şirketin ayrılma olasılığını azaltmaktadır. Bu sonuç, satın alım yöntemi söz konusu olduğunda, resort otellerde konaklayan devamlı müşterilere uygulanan kayıp müşteri analiz bulgularıyla otellerin B2B müşterilerine yani ticari şirketlere uygulanan kayıp müşteri analiz bulguları arasında farklılık oluşabileceğini önermektedir.

Bu çalışmada lojistik regresyon analizi bulgularına göre en önemli yordayıcılardan bir diğeri de Türkiye vatandaşı olmak değişkenidir. Mevcut çalışmanın bulgularına göre müşterinin Türkiye vatandaşı olması durumunda ayrılma olasılığı %55 oranında ($|1-\text{odds}| = 0,547$) azalmaktadır. Ancak, benzer bir değişken olan müşterinin geldiği bölge değişkeni Coussement ve De Bock'un (2013:1634) çalışmasında en önemli ilk yirmi değişken arasına girememiştir.

Müşterinin son konaklamasındaki tatil sezonu ve geldiği pazar değişkenleri bu çalışmaya özgü değişkenlerdir. Rastgele Orman algoritması tarafından hesaplanan önem düzeyine göre avrupa pazarından gelmek, yüksek sezonda konaklamak, düşük sezonda konaklamak ve baltık balkan pazarından gelmek modele giren önemli değişkenler olarak ortaya çıkmıştır. Ayrıca düşük sezon değişkeninin Lojistik Regresyon bulgularına göre ayrılma olasılığını %40 oranında ($|1-\text{odds}| = 0,397$) arttırdığı görülmektedir. Diğer taraftan önceki çalışmalarda en önemli bulunan değişkenler arasında yer alan toplam harcama tutarı değişkeni (R_Monetary) Rastgele Orman modeline girememiştir. Bu durum tutar temelli değişkenlerin geceleme değişkenleriyle aralarında yüksek korelasyon bulunmasından kaynaklanıyor olabilir. Ayrıca, harcama tutarı değişkeninin Lojistik Regresyon bulgularına göre ayrılma ve ayrılmama üzerindeki etkisi nötr olarak ortaya çıkmıştır.

Müşterilerin yaşam döngüsü göz önünde bulundurularak işletmenin kayıp müşterilerinin yeni dönemdeki potansiyel müşterileri olduğu unutulmamalıdır (Qi vd., 2008: 262). Bu bağlamda bu çalışmanın en önemli çıktılarından biri de karışıklık matrisinde yanlış negatif alanına düşen kayıp müşterilerdir. Rastgele Orman algoritması tarafından öğrenme ve test aşamasında yüksek olasılık puanlarıyla ayrılmayan olarak sınıflandırılmasına rağmen gerçekte ayrılma davranışında bulunan ve dolayısıyla karışıklık matrisinde yanlış negatif alanına düşen müşterilere zincir otel işletmesi tarafından ulaşılması önem arz etmektedir. Çünkü zincir otel işletmesinin ayrılma davranışında bulunmayan müşterileri daha uzun süredir müşteri

olan, daha yüksek geliş sayısına sahip, daha çok harcama yapan ve daha sık konaklayan müşterilerdir. Ayrılmama davranışında bulunan müşterilerle aynı özellikleri göstermelerine rağmen zincir otel işletmesini terk etmiş bu müşterileri geri kazanmak işletme açısından önem arz etmektedir. Ayrıca bu müşterilerin bağlılıklarını etkileyen ve acil önlem alınması gereken hizmet hataları veya farklı faktörler söz konusu olabileceği gibi bu müşteriler ölüm, hastalık, başka ülkeye taşınma vb. nedenlerle gönülsüz ayrılma davranışında bulunmuş olabilirler. Dolayısıyla, bazı müşterilerin geleneksel değişkenler (RFM temelli) tarafından daha doğru tespit edildiği, bazı müşterilerin ise farklı değişkenler aracılığıyla daha yüksek performansla tahmin edilebildiği göz önünde bulundurulmalı, müşterilerin ayrılma davranışını yordayabilecek farklı değişkenler geleneksel değişkenlerle birlikte kullanılmalıdır (Zhang vd., 2012:102).

Çalışmanın amacı kapsamında otel işletmesini bir sonraki periyod içerisinde terk etme ihtimali yüksek olan müşterilerin önceden tespit edilmesi sağlanmıştır. Bu amaçla Rastgele Orman algoritması tarafından oluşturulan model kullanılarak 2018 yılında aktif olan, zincir otel işletmesinin bünyesindeki otellerde en az 10 gece konaklamış ve en az 2 defa gelmiş 12413 müşterinin 2016-2018 periyoduna kadar gerçekleştirdikleri konaklamalara yönelik veriler üzerinden 2019-2021 periyodunda muhtemel ayrılma davranışında bulunacak müşteriler tahminlenmiştir. Çalışmanın bulgularına göre 12413 müşterinin 7746 tanesi (%62) ayrılma davranışında bulunmaz, 4667 tanesi (%38) ayrılma davranışında bulunur şeklinde tahminlenmiştir. Müşterilerin ayrılma ve ayrılmama olasılıklarına göre değişken ortalamaları incelendiğinde alan yazınla uyumlu olarak gerçek harcama tutarının ve geliş sayılarının ayrılmama ihtimali yüksek olan müşterilerde daha yüksek olduğu görülmüştür. Bununla birlikte, ayrılma olasılığı yüksek olan müşterilerin son konaklama tarihlerinin daha güncel olduğu göze çarpmaktadır. Bu sonuç “güncellik ne kadar yakın tarihliyse müşterinin sadık olma olasılığı o derece artar” (De Wit, 2017: 43; Buckinx ve Van den Poel; 2005: 259) yönündeki genel alan yazına ters düşüyor gibi görünse de, ayrılma olasılığı yüksek olan müşterilerin ilişki sürelerinin daha kısa olduğu göz önünde bulundurularak daha yeni müşterilerden oluştuğu söylenebilir. Ayrıca ayrılma olasılığı yüksek olan müşterilerin daha yakın tarihli güncellik ortalaması kış sezonuna denk gelirken ($\cong$ Ocak ayı), ayrılmama olasılığı yüksek olan müşterilerde yaz sezonuna denk gelmektedir ($\cong$ Temmuz ayı). Çalışmanın bir diğer bulgusu ise gerçek harcama tutarının aksine, son üç yıldaki harcama tutarı temelli değişken ortalamalarının ayrılma ihtimali yüksek olan müşterilerde daha yüksek olmasıdır. Bununla birlikte son üç yıldaki geliş sayısı ortalaması her iki grupta da aynıdır. Bu durum ayrılma ihtimali yüksek olan müşterilerin gecelik ödeme tutarlarının daha yüksek olmasından

kaynaklanıyor olabilir. Zincir otel işletmesinin farklı pazarlara yönelik farklı fiyat politikaları ve bu fiyatlara göre değişen farklı hizmetleri mevcuttur. Dolayısıyla, ayrılma ihtimali yüksek olan müşterilerin ağırlıklı BDT pazarından (%48) gelmesi bu sonuca yol açmış olabilir. Bununla birlikte ayrılma ihtimali yüksek olan müşterilerin ortalama geceleme sayısı değişimleri negatif yönlü, gecelik ödeme tutarlarındaki ve ortalama harcama tutarlarındaki değişimler daha düşüktür. Diğer bir ifadeyle, ayrılma ihtimali yüksek olan müşterilerin son üç yıldaki konaklamalarında ortalama geceleme süreleri azalma eğilimi göstermiştir. Buna karşılık ayrılmama ihtimali yüksek olan müşteriler önceki konaklamalarına kıyasla daha fazla geceleme yapmışlardır. Tahminlere yönelik bulgular incelendiğinde toplam geliş sayıları daha düşük olsa da ayrılma ihtimali yüksek olan müşterilerin daha yüksek fiyatlardan konakladığı ve karlılık açısından işletme için yüksek potansiyele sahip olduğu göze çarpmaktadır.

Dolayısıyla çalışmaya konu olan zincir otel işletmesinin ayrılma ihtimali yüksek olan müşterilerini bir öncelik sıralamasına koyması ve sınırlı pazarlama bütçesini bu müşterilerinden elde edeceği potansiyel kazanç ile kıyaslama yaparak kullanması doğru bir strateji olacaktır. Bu bağlamda Qi vd. (2009: 260) tarafından önerilen öncelik matrisinin bir uyarlaması yapılmıştır. Müşterilerin ayrılma olasılığının düşük (%75 ve altı) ve yüksek (%75 üzeri) olması durumuyla, müşterinin geri gelmesi durumunda otele kazandıracağı tahmini değerin düşük (ortalamanın altında) ve yüksek (ortalamanın üstünde) olması durumları kıyaslanmış ve öncelik grupları belirlenmiştir.

Öncelik matrisindeki bulgulara göre birinci öncelik grubunda ayrılma ihtimali yüksek ve geri gelmesi durumunda kazandıracağı değer yüksek olan müşteriler yer almaktadır. Öncelik kategorisi 1 grubuna yönelik bulgular değerlendirildiğinde bu müşterilerin zincir otel işletmesinin yeni müşterileri olduğu söylenebilir. En düşük toplam geliş sayısına sahip olan bu müşteri segmentinin zincir otel işletmesiyle ilişkisi 4 yıldan fazla olmakla birlikte ziyaretleri arasında geçen süre diğer segmentlerin ve genel ortalamanın üzerindedir. Dolayısıyla bu müşteriler işletmeyi kısmi olarak terk ediyor olabilir. Diğer bir deyişle tatillerini farklı otellerde gerçekleştirerek puanlarını kaybetmemek için ya da başka nedenlerle işletmeye geri dönüyor olabilirler. Bunun haricinde öncelik kategorisi 1 grubunda yer alan müşteriler son ziyaretlerinde diğer gruplardan daha fazla harcama yapmalarına rağmen kendi ortalamalarıyla kıyaslandığında geceleme sürelerini ve dolayısıyla harcamalarını düşürmüştür. Buradan hareketle bu müşteri segmenti “kararsızlar” olarak adlandırılabilir. Potansiyel olarak işletmeye yüksek değer kazandırması beklenen bu müşteri segmentinin işletmeye henüz sadakatinin oluşmadığı görülmektedir. Zincir otel işletmesi tarafından bu müşteri segmentinin bağlılığını sağlayacak ve arttıracak önlemlerde bulunulması uygun bir strateji olabilir.

Ayrılma ihtimali yüksek ve kazandıracağı değer düşük olan müşteriler ikinci öncelik kategorisinde yer almaktadır. Ayrılma ihtimali yüksek olan Türk müşterilerin büyük çoğunluğu bu kategoride bulunmaktadır. Öncelik kategorisi 2 grubuna yönelik bulgular değerlendirildiğinde bu müşterilerin zincir otel işletmesiyle ilişki süresi en kısa olan, toplam geliş sayısı ve ödemeleri daha düşük olmasına rağmen son üç yıldaki geliş sayıları ortalamanın üstünde seyreden ve harcama eğilimleri yükselişte olan müşteri segmenti oldukları söylenebilir. Bu müşteri segmenti ağırlıklı olarak orta sezonda konaklama yapması ve kısa süreler konaklaması nedeniyle harcamaları daha düşük olsa da son üç yıldaki hareketleri izlendiğinde zincir otel işletmesiyle bağ kurmaya başlamış oldukları düşünülebilir. Dolayısıyla, bu müşteri segmentinin daha kısıtlı bütçelerle seyahat ettikleri düşünülerek muhtemel kayıpları önlemek amacıyla düşük ve orta sezonda geçerli ekonomik teşviklerin sunulması uygun bir strateji olabilir. Buradan hareketle bu müşteri segmenti “tutumlular” olarak adlandırılabilir.

Ayrılma ihtimali düşük ve müşteri değeri yüksek olan müşteriler üçüncü öncelik kategorisinde almaktadır. Öncelik kategorisi 3 grubuna yönelik bulgular değerlendirildiğinde bu müşterilerin diğer segmentlere kıyasla zincir otel işletmesiyle en uzun süreli ilişkiye, en çok geliş sayısına, en uzun geceleme süresine, en yüksek üyelik kategorisine ve en yüksek harcamalara sahip olan segment olduğu görülmektedir. Bununla birlikte bu müşteri segmentinin son üç yıldaki geliş sayısında büyük bir düşüş olduğu ve buna bağlı olarak harcama değişimlerinin diğer segmentlere kıyasla düşük bir artış seyrettiği ortaya çıkmaktadır. Öncelik kategorisi 3 segmentinin ayrılma davranışında bulunma olasılığı nispeten daha düşük olsa da, tekrar gelmesi halinde kazandıracağı değer en yüksek olan müşterilerin bu segmentte yer aldığı göz önünde bulundurulmalıdır. Müşteriler kimi zaman memnun olmasalar da değiştirme maliyetleri nedeniyle aynı işletmeyle çalışmaya devam edebilmektedir. Bu değiştirme maliyetleri arasında toplanılan üyelik puanları ve üyelik avantajları sayılabilir. Üye müşteriler bu haklarının yok olmasını istemedikleri için memnun olmadıkları işletmelerden ayrılma davranışında bulunmayabilirler (Kim ve Moon, 2012: 11720). Buradan hareketle bu müşteri segmenti “alternatif arayanlar” olarak adlandırılabilir. Bu müşteri segmentinin son üç yılda zincir otel işletmesini daha az ziyaret etmiş olması seyahatlerini başka otellere yönlendirdikleri anlamına geliyor olabilir. Bu müşteri segmentinin en yüksek üyelik puanlarına sahip müşterilerden oluşması puanları ve üyelik avantajlarını kaybetmemek adına zincir otel işletmesine tekrar gelmeleri yönünde teşvik edici olabilir. Dolayısıyla, zincir otel işletmesi bu segmentte yer alan müşterilerini kaybetmemek için yüksek sezonda geçerli olacak iki kat puan kazandırma ya da bir üst kart kategorisinin avantajlarından faydalandırma vb. müşterilerin bağlılığını geri kazandıracak stratejilerde bulunabilir.

Ayrılma ihtimali düşük ve kazandıracağı değer düşük olan müşteriler dördüncü ve son öncelik kategorisinde yer almaktadır. Öncelik kategorisi 4 grubuna yönelik bulgular değerlendirildiğinde bu müşterilerin orta sezonda gelmelerine rağmen nispeten daha yüksek harcamalarda bulunan, son üç yılda en olumlu değişimi gösteren ve geliş sayıları en yüksek olan müşteri segmenti olduğu görülmektedir. Diğer segmentlere kıyasla daha sık konaklama yapan bu segmentin zincir otel işletmesine bağlılığının son dönemde artış gösterdiği izlenmektedir. Buradan hareketle bu müşteri segmenti “fırsatçılar” olarak adlandırılabilir. Bununla birlikte bu segmentin ayrılma ihtimalinin ve kazandıracağı değer diğer segmentlere kıyasla düşük olması, zincir otel işletmesi tarafından cömert teşviklerde bulunulmasının doğru bir strateji olmayacağını düşündürmektedir. Dolayısıyla, zincir otel işletmesinin bu müşteri segmentine daha sık kendini hatırlatıcı mail gönderimlerinde bulunması, orta/düşük sezona yönelik düşük bütçeli teşvikler ya da yüksek sezona yönelik son dakika teklifleri sunarak kendine çekmeye çalışması uygun bir strateji olarak görülebilir.

Kayıp müşterileri önlemede müşterilere sunulacak teşviklere karar verilirken müşterilerin alışkanlıklarının ve ilgi alanlarının bilinmesine ihtiyaç vardır. Bu konuda işletmeler daha ileri ve derinleştirilmiş araştırmalara ihtiyaç duymaktadır (Qi vd., 2008: 262). Zincir otel işletmesinin müşteri profil sayfalarında müşterilerin ilgi alanlarına yönelik veri girişi yapılması mümkündür. Ayrıca otel işletmesinin web sitesinde müşterilerin üyelik oluşturabileceği, konaklamalarından kazandıkları puanlarını görüntüleyebilecekleri ve tercihleriyle ilgi alanlarına yönelik verileri girdiklerinde puan kazanabilecekleri bir müşteri profili alanı bulunmaktadır. Ancak çoğu müşteri tarafından bu bilgilerin henüz doldurulmadığı görülmektedir. Zincir otel işletmesinin bu eksik verileri tespit ederek müşterilerini bu bilgileri doldurmaya özendirmesi gerekmektedir. Akabinde bu veriler üzerine detaylı analizler yaparak müşterilerini daha iyi tanıyabilir ve sunacağı teşvikler için daha doğru kararlarda bulunabilir.

Sonuç olarak bu çalışmanın kayıp müşteri analizi literatürüne birçok katkısı bulunmaktadır. Öncelikle bu çalışma bir zincir otel işletmesinin müşteri verileri kullanılarak gerçekleştirilen ilk kayıp müşteri tahminleme çalışmasıdır (yazarın bilgisi dahilinde). Bu çalışmanın bulguları bir zincir otel işletmesinde Rastgele Orman algortimasının gelecek üç yıl içerisinde ayrılma olasılığı olan devamlı müşterileri %80 doğru tahminleyebileceğini göstermektedir. Dolayısıyla mevcut çalışma bu alanda gerçekleştirilecek çalışmalar için bir örnek oluşturmakta ve resort otellerin müşteri verileriyle de yüksek performanslı modeller oluşturulabileceğini önermektedir. Ayrıca çalışmada iki farklı öğrenme modelinin kıyaslaması yapılmış ve iki model için ayrı değişken setleri kullanılmıştır. Böylece hangi modelin ve hangi değişkenlerin daha yüksek performans sağladığına yönelik bulgular elde edilmiştir. Bununla

birlikte bu çalışma kontratsız çalışan işletmelerde gerçekleştirilen kayıp müşteri tahminleme analizlerine katkı sağlamaktadır. Bu alanda gerçekleştirilen çalışmalar kayıp müşteri tanımlamasının zor olması nedeniyle kısıtlıdır. Çalışma alanının resort otel müşterileri olması nedeniyle daha önce denenmemiş sezon ve pazar değişkenlerinin mevcut çalışmada öğrenme modeline girdiği görülmüştür. Ayrıca daha önce De Wit (2017:64) tarafından ticari otel müşterilerine (B2B müşteriler - şirketler) uygulanan kayıp müşteri tahminleme çalışmasında kullanılmış bir takım benzer değişkenlerin (rezervasyonu yapan, ziyaret edilen otel sayısı, otel, yetişkin sayısı) mevcut çalışmada tekrar denenmesi sağlanmış (münferit rezervasyon, grubun diğer otellerinde konaklama, son konaklanan otel, konaklama eşlikçileri) çoğu değişken için benzer bulgulara ulaşılmıştır. Son olarak mevcut çalışmada, test aşamasında elde edilen performans değerleriyle sınırlı kalan daha önceki çalışmaların ötesine geçilmiş ve gerçek veriler üzerinde tahmin modelinin uygulanması sağlanmıştır. Ayrıca kayıp müşteri tahminleme literatüründe elde edilen sonuçların nasıl kullanılabileceğine yönelik yönlendirmelerde bulunan teorik metodların (örneğin Qi vd., 2009: 260) bir uygulaması gerçekleştirilmiştir. Dolayısıyla, bu çalışmanın sonuçları bir otel işletmesinde makine öğrenimi tekniklerinden faydalanılarak kayıp müşteri tahminlemesi yapılabileceğini ve bu sayede pazarlama yöneticilerinin kararlarında etkili olacak faydalı veriler üretilebileceğini göstermektedir. Otel yöneticileri öğrenen algoritmalarından yararlanarak müşteri verilerini analiz edebilir ve pazarlama stratejilerini bu bilgiler ışığında daha doğru bir şekilde uygulayabilir.

Bu araştırmanın bir takım sınırlılıkları bulunmaktadır. Öncelikle müşteri veri tabanında müşterilerin segment kodları yer almadığından toplantı katılımı, golf oyuncusu, futbol kampı vb. konaklamalarda bulunan müşteriler ayırdedilememiştir. Bu nedenle analize sadece zincir otel işletmesinin otellerinde en az iki defa konaklama yapan, en az 10 gece konaklaması bulunan ve 2015 yılında aktif olan müşteriler alınmıştır. Bunun yanı sıra zincir otel işletmesinin acente isimleri ve oda tipi için kullandığı standart bir kod bulunmadığından bu veriler analize dahil edilememiştir. Acente ve şirket isimleri metin madenciliği aracılığıyla temizlenerek standart kodlara dönüştürülebilir. Oda tipi için kullanılan kodlarda aynı ya da benzer oda tipini ifade eden birden çok kullanım bulunmaktadır. Bu kodların karşısında yer alan oda numaraları için önbüro çalışanlarıyla görüşülerek standart oda tipi kategorileri oluşturulabilir. Böylece bu değişkenler de ilerde gerçekleştirilecek analizlere dahil edilebilir. Bunun haricinde, zincir otel işletmesi müşteri veri tabanında müşterinin gecelik harcamasını ödediği para biriminde tutuyor olmasına karşın müşteri profillerinde toplam harcama tutarı verisinin tüm harcamalarının toplamını son ödediği kur üzerinden gösterdiği farkedilmiştir. Bu nedenle çalışmada tüm gecelik harcamalar her bir gözlem periyodu için periyodun son yılındaki (2015 ve 2018)

ortalama dolar kuru hesaplanarak tek para birimine dönüştürülmüştür. Zincir otel işletmesinin müşterilerinin harcamalarını kıyaslayabilmesi açısından ilgili ödemenin yapıldığı aydaki ortalama kur üzerinden müşterilerin harcamalarını tek para birimine dönüştürmesi ve toplam harcama tutarlarını bu para birimi üzerinden göstermesi daha doğru olacaktır. Ayrıca bu sayede tutar hesaplaması yapılırken ilgili periyod içerisindeki karşılaştırmalarda veri kaybı yaşanmayacaktır.

Bu araştırmanın devamında zincir otel işletmesinin satış ve pazarlama departmanı yöneticileriyle mevcut çalışmanın bulguları değerlendirilerek bir müşteri geri kazanma çalışması gerçekleştirilmesi planlanmaktadır. Bunun yanı sıra ilerde gerçekleştirilecek çalışmalarda müşterilerin şikayet, arıza ve istek verilerinin metin madenciliği tekniklerinden faydalanılarak analizlere dahil edilmesi daha yüksek performanslı tahmin modellerinin oluşturulmasında ve müşterilerin ayrılma davranışlarına yönelik daha derin bir anlayış geliştirilmesinde fayda sağlayabilir (örneğin Coussement ve Van den Poel, 2008: 164). Bununla birlikte müşteri memnuniyet anketlerinin bulguları kayıp müşteri analizlerinde kullanılacak değişkenlere ilave edilebilir (Burez ve Van den Poel, 2008: 507; Gustafsson vd., 2005: 2013). Özellikle öğrenme ve test aşamasında karışıklık matrisinin yanlış negatif alanına düşen kayıp müşterilerin ayrılma nedenlerinin keşfedilmesinde bu veriler büyük rol oynayabilir. Mevcut çalışmada öğrenme ve test veri setleri için aynı periyodlar kullanılmıştır. İlerde gerçekleştirilecek çalışmalarda model gücünün daha sonraki zamanları tahminlemede ne kadar güçlü olduğunu daha iyi anlayabilmek için test seti öğrenme setinden daha farklı bir periyodu kapsayacak şekilde hazırlanabilir (örneğin Coussement ve Van den Poel, 2008: 168).

KAYNAKÇA

- Ahmad, Z., Jun, M., Khan, I., Abdullah, M. ve Ghauri, T. A. (2016). “Examining Mediating Role of Customer Loyalty for Influence of Brand Related Attributes on Customer Repurchase Intention”. *Journal of Northeast Agricultural University (English Edition)*, 23(2): 89-96.
- Ahn, J. H., Han, S. P. ve Lee, Y. S. (2006). “Customer Churn Analysis: Churn Determinants and Mediation Effects of Partial Defection in the Korean Mobile Telecommunications Service Industry”. *Telecommunications Policy*, 30(10-11): 552-568.
- Akhtar, N., Zubair, N., Kumar, A. ve Ahmad, T. (2017). “Aspect Based Sentiment Oriented Summarization of Hotel Reviews”. *Procedia Computer Science*, 115: 563-571.
- Akın, M. (2015). “A Novel Approach to Model Selection in Tourism Demand Modeling”. *Tourism Management*, 48: 64-72.
- Aksu, A. A. ve Ehtiyar, V. R. (2007). *Turizm İşletmelerinde Çağdaş Yönetim Teknikleri*, Detay Yayıncılık, Ankara.
- Alshawabkeh, M., Jang, B. ve Kaeli, D. (2010). “Accelerating the Local Outlier Factor Algorithm on a GPU for Intrusion Detection Systems”. *Proceedings of the 3rd Workshop on General-Purpose Computation on Graphics Processing Units*, 14 Mart 2010, Pittsburg, 104-110.
- Al-Smadi, M., Qawasmeh, O., Al-Ayyoub, M., Jararweh, Y. ve Gupta, B. (2018). “Deep Recurrent Neural Network vs. Support Vector Machine for Aspect-Based Sentiment Analysis of Arabic Hotels’ Reviews”. *Journal of Computational Science*, 27: 386-393.
- Altındal, M. (2013). “Dijital Pazarlamada Marka Yönetimi ve Sosyal Medyanın Etkileri”. *Akademik Bilişim 2013 – XV. Akademik Bilişim Konferansı Bildirileri*. 23-25 Ocak 2013, Antalya, 1147-1151.
- Ang, L. ve Buttle, F. (2006). “CRM Software Applications and Business Performance”. *Journal of Database Marketing & Customer Strategy Management*, 14(1): 4-16.
- Antonio, N., de Almeida, A. ve Nunes, L. (2017). “Predicting Hotel Booking Cancellations to Decrease Uncertainty and Increase Revenue”. *Tourism & Management Studies*, 13(2): 25-39.
- Antonio, N., de Almeida, A. ve Nunes, L. (2019). “Big Data in Hotel Revenue Management: Exploring Cancellation Drivers to Gain Insights into Booking Cancellation Behavior”. *Cornell Hospitality Quarterly*, 60(4): 298–319.

- Ballings, M. ve Van den Poel, D. (2012). “Customer Event History for Churn Prediction: How Long is Long Enough?”. *Expert Systems with Applications*, 39(18): 13517-13522.
- Berry, L. L. (1983). “Relationship Marketing”. In: L. L. Berry, G. L. Shostack ve G. Upah. (Eds.): *Emerging Perspectives on Services Marketing*. American Marketing Association, Chicago, IL, 25-28.
- Berry, L. L. (2002). “Relationship Marketing of Services Perspectives from 1983 and 2000”. *Journal of Relationship Marketing*, 1(1): 59-77.
- Bjørkelund, E., Burnett, T. H. ve Nørvåg, K. (2012). “A Study of Opinion Mining and Visualization of Hotel Reviews”. *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*. 3-5 Aralık 2012, Bali, 229-238.
- Bloemer, J., De Ruyter, K. ve Peeters, P. (1998). “Investigating Drivers of Bank Loyalty: The Complex Relationship Between Image, Service Quality and Satisfaction”. *International Journal of Bank Marketing*, 16(7): 276-286.
- Bolton, R. N. (1998). “A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction”. *Marketing Science*, 17(1): 45-65.
- Breiman, L. (2001). “Random Forests”. *Machine Learning*, 45(1): 5–32.
- Buckinx, W. ve Van den Poel, D. (2005). “Customer Base Analysis: Partial Defection of Behaviourally Loyal Clients in a Non-Contractual FMCG Retail Setting”. *European Journal of Operational Research*, 164(1): 252–268.
- Bulunmaz, B. (2016). “Gelişen Teknolojiyle Birlikte Değişen Pazarlama Yöntemleri ve Dijital Pazarlama”. *TRT Akademi*, 1(2): 348-365.
- Burez, J. and Van den Poel, D. (2009). “Handling Class Imbalance in Customer Churn Prediction”. *Expert Systems with Applications*, 36: 4626–4636.
- Burez, J. ve Van den Poel, D. (2008). “Separating Financial from Commercial Customer Churn: A Modeling Step Towards Resolving the Conflict Between the Sales and Credit Department”. *Expert Systems with Applications*, 35(1-2): 497-514.
- Caber, M. (2010). *Tur Operatörü-Seyahat Acentesi Elektronik Müşteri İlişkileri Yönetiminde Web Sitesi Kalitesinin Memnuniyet ve Güven Üzerine Etkisi*. Yayımlanmamış Doktora Tezi. Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü, Antalya.
- Candela Garza, E. (2017). *Revenue Optimization for a Hotel Property with Different Market Segments: Demand Prediction, Price Selection and Capacity Allocation*. Yayımlanmamış Doktora Tezi. Massachusetts Institute of Technology, Cambridge.

- Cankurt, S. ve Subasi, A. (2015). "Developing Tourism Demand Forecasting Models Using Machine Learning Techniques with Trend, Seasonal, and Cyclic Components". *Balkan Journal of Electrical and Computer Engineering*, 3(1): 42-49.
- Cankurt, S. ve Subasi, A. (2016). "Tourism Demand Modelling and Forecasting Using Data Mining Techniques in Multivariate Time Series: A Case Study in Turkey". *Turkish Journal of Electrical Engineering & Computer Sciences*, 24(5): 3388-3404.
- Chaffey, D. ve Ellis-Chadwick, F. (2019). *Digital Marketing*. Pearson, UK.
- Chen, I. J., ve Popovich, K. (2003). "Understanding Customer Relationship Management (CRM) People, Process and Technology". *Business Process Management Journal*, 9(5): 672-688.
- Chen, K. Y.ve Wang, C. H. (2007). "Support Vector Regression with Genetic Algorithms in Forecasting Tourism Demand". *Tourism Management*, 28(1): 215-226.
- Chen, Q. ve Chen, H. M. (2004). "Exploring the Success Factors of eCRM Strategies in Practice". *Journal of Database Marketing & Customer Strategy Management*, 11(4): 333-343.
- Claveria, O., Monte, E. ve Torra, S. (2016). "Combination Forecasts of Tourism Demand with Machine Learning Models". *Applied Economics Letters*, 23(6), 428-431.
- Clow, K. ve Baack, D. (2010). *Integrated Advertising, Promotion and Marketing Communications Global Edition*. Pearson, UK.
- Conlon, D. E. ve Murray, N. M. (1996). "Customer Perceptions of Corporate Responses to Product Complaints: The Role Of Explanations". *Academy of Management Journal*, 39(4): 1040-1056.
- Corazza, M., Giovanni F., ve Francesco M. (2014). "An Artificial Neural Network-based Technique for On-Line Hotel Booking." *Procedia Economics and Finance*, 15: 45-55.
- Coussement, K. ve De Bock, K. W. (2013). "Customer Churn Prediction in the Online Gambling Industry: The Beneficial Effect of Ensemble Learning". *Journal of Business Research*, 66(9): 1629-1636.
- Coussement, K. ve Van den Poel, D. (2008). "Integrating the Voice of Customers Through Call Center Emails into a Decision Support System for Churn Prediction". *Information & Management*, 45(3): 164-174.
- Coussement, K., Lessmann, S. ve Verstraeten, G. (2017). "A Comparative Analysis of Data Preparation Algorithms for Customer Churn Prediction: A Case Study in the Telecommunication Industry". *Decision Support Systems*, 95: 27-36.

- De Pelsmacker, P., Van Tilburg, S. ve Holthof, C. (2018). "Digital Marketing Strategies, Online Reviews and Hotel Performance". *International Journal of Hospitality Management*, 72: 47-55.
- De Wit, L. (2017). *An Analysis of Non-Contractual Churn in the B2B Hotel Industry*. Yayınlanmamış Yüksek Lisans Tezi. Tilburg University Data Science: Business and Governance, Tilburg, Netherlands.
- Demir, F. O. ve Kırdar, Y. (2000). "Müşteri İlişkileri Yönetimi: CRM", *Review of Social, Economic and Business Studies*, 7(8): 293-308.
- Dick, A. S. ve Basu, K. (1994). "Customer Loyalty: Toward an Integrated Conceptual Framework". *Journal of the Academy of Marketing Science*, 22(2): 99-113.
- Duan, W., Cao, Q., Yu, Y. ve Levy, S. (2013). "Mining Online User-Generated Content: Using Sentiment Analysis Technique to Study Hotel Service Quality". *46th Hawaii International Conference on System Sciences*, 7-10 Ocak 2013, Wailea, 3119-3128.
- Duan, W., Yu, Y., Cao, Q. ve Levy, S. (2016). "Exploring the Impact of Social Media on Hotel Service Performance: A Sentimental Analysis Approach". *Cornell Hospitality Quarterly*, 57(3): 282-296.
- Emel, G. G., Taşkin, Ç. ve Akat, Ö. (2007). "Profiling a Domestic Tourism Market by Means of Association Rule Mining". *Anatolia*, 18(2): 334-342.
- Enache, M. C. (2019). "Machine Learning in Tourism Revenue Management". *Annals of the University Dunarea de Jos of Galati: Fascicle: XVII, Medicine*, 25(1): 132-136
- Ganesan, S. (1994). "Determinants of Long-Term Orientation in Buyer-Seller Relationships". *Journal of Marketing*, 58(2): 1-19.
- Ganesh, J., Arnold, M. J. ve Reynolds, K. E. (2000). "Understanding the Customer Base of Service Providers: An Examination of the Differences Between Switchers and Stayers". *Journal of Marketing*, 64(3): 65-87.
- Gantley, M., Whitehouse, H. ve Bogaard, A. (2018). "Material Correlates Analysis (MCA): An Innovative Way of Examining Questions in Archaeology Using Ethnographic Data". *Advances in Archaeological Practice*, 6(4): 328-341.
- García, A., Gaines, S. ve Linaza, M. T. (2012). "A Lexicon Based Sentiment Analysis Retrieval System for Tourism Domain". *e-Review of Tourism Research (eRTR)*, 39(10): 9166-9180.
- Ghorpade, T. ve Ragha, L. (2012). "Featured Based Sentiment Classification for Hotel Reviews Using NLP and Bayesian Classification". *2012 International Conference on*

Communication, Information & Computing Technology (ICCICT), 19-20 Ekim 2012, Mumbai, 1-5

- Gilbert, D. C., Powell-Perry, J. ve Widiyoso, S. (1999). "Approaches by Hotels to the Use of the Internet as a Relationship Marketing Tool". *Journal of Marketing Practice: Applied Marketing Science*, 5(1): 21-38.
- Gustafsson, A., Johnson, M. D. ve Roos, I. (2005). "The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention". *Journal of Marketing*, 69(4): 210-218.
- Gülmez, M. ve Kitapçı, O. (2003). "İlişki Pazarlamasının Gelişimi ve Yakın Geleceği". *CÜ İktisadi ve İdari Bilimler Dergisi*, 4(2): 81-89.
- Günther, C. C., Tvette, I. F., Aas, K., Sandnes, G. I. ve Borgan, Ø. (2014). "Modelling and Predicting Customer Churn from an Insurance Company". *Scandinavian Actuarial Journal*, 2014(1): 58-71.
- Hadden, J., Tiwari, A., Roy, R. ve Ruta, D. (2006). "Churn Prediction: Does Technology Matter". *International Journal of Intelligent Technology*, 1(2): 104-110.
- Hadden, J., Tiwari, A., Roy, R. ve Ruta, D. (2008). "Churn Prediction Using Complaints Data". *World Academy of Science, Engineering and Technology*, 10: 809-814.
- Hall, M. A. ve Holmes, G. (2003). "Benchmarking Attribute Selection Techniques for Discrete Class Data Mining". *IEEE Transactions on Knowledge and Data Engineering*, 15(6): 1437-1447.
- Han, S. (2018). "A Study on a Predictive Model of Customer Defection in a Hotel Reservation Website". *MATEC Web of Conferences* 228, 01009: 1-8.
- Hantuli, L. (2016). *An Analysis of Hotel Loyalty Programs: A Case Study of the Hilton HHonors Loyalty Program*. Yayımlanmamış Yüksek Lisans Tezi. The Collins College of Hospitality Management, California.
- Hassouna, M., Tarhini, A., Elyas, T. ve AbouTrab, M. S. (2015). "Customer Churn in Mobile Markets a Comparison of Techniques". *International Business Research*, 8(6): 224-237
- Hastie, T., Tibshirani, R. ve Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics, Stanford, California.
- Ho, R. (2013). *Handbook of Univariate and Multivariate Data Analysis with IBM SPSS*. Chapman and Hall/CRC, Boca Raton.
- Horster, E. ve Gottschalk, C. (2012). "Computer-Assisted Webnography: A New Approach to Online Reputation Management in Tourism". *Journal of Vacation Marketing*. 18(3): 229-238

- Hsieh, H. Y., Klyuev, V., Zhao, Q. ve Wu, S. H. (2014). "SVR-Based Outlier Detection and Its Application to Hotel Ranking". *2014 IEEE 6th International Conference on Awareness Science and Technology (iCAST)*, 29- 31 Ekim 2014, Paris, 1-6.
- Huang, B. Q., Kechadi, T. M., Buckley, B., Kiernan, G., Keogh, E. ve Rashid, T. (2010). "A New Feature Set with New Window Techniques for Customer Churn Prediction in Land-Line Telecommunications". *Expert Systems with Applications*, 37(5): 3657-3665.
- Hughes, A. M. (2003). *The Customer Loyalty Solution: What Works and What Doesn't in Customer Loyalty Programs*. McGraw-Hill, New York.
- Hung, S., Yen, D. C., ve Wang, H. (2006). "Applying Data Mining to Telecom Churn Management". *Expert Systems with Applications*, 31(3): 515–524.
- Hwang, H., Jung, T., ve Suh, E. (2004). "An LTV Model and Customer Segmentation Based on Customer Value: A Case Study on the Wireless Telecommunication Industry". *Expert Systems with Applications*, 26(2): 181–188.
- Kamel, N., Atiya, A. F., El Gayar, N. ve El-Shishiny, H. (2008). "Tourism Demand Foreacsting Using Machine Learning Methods". *ICGST International Journal on Artificial Intelligence and Machine Learning*, 8: 1-7.
- Kandampully, J., Zhang, T. ve Bilgihan, A. (2015). "Customer Loyalty: a Review and Future Directions with a Special Focus on the Hospitality Industry". *International Journal of Contemporary Hospitality Management*, 27(3): 379-414.
- Kawale, J., Pal, A. ve Srivastava, J. (2009). "Churn Prediction in MMORPGS: A Social Influence Based Approach". *International Conference on Computational Science and Engineering*, 29-31 Ağustos 2009, Vancouver, 423-428.
- Kılıç, S. (2013). "Klinik Karar Vermede ROC Analizi". *Journal of Mood Disorders*, 3(3): 135-40.
- Kim, H. S. ve Yoon, C. H. (2004). "Determinants of Subscriber Churn and Customer Loyalty in the Korean Mobile Telephony Market". *Telecommunications Policy*, 28(9-10): 751-765.
- Kim, S. Y. (2011). "Prediction of Hotel Bankruptcy Using Support Vector Machine, Artificial Neural Network, Logistic Regression, and Multivariate Discriminant Analysis". *The Service Industries Journal*, 31(3): 441-468.
- Kim, Y. S. ve Moon: (2012). "Measuring The Success of Retention Management Models Built on Churn Probability, Retention Probability, and Expected Yearly Revenues". *Expert Systems with Applications*, 39(14): 11718-11727.

- King, S. F. ve Burgess, T. F. (2008). "Understanding Success and Failure in Customer Relationship Management". *Industrial Marketing Management*, 37(4): 421-431.
- Kitapcı, O. (2006). *Müşteri Sadakati Yaratmak için Veri Tabanlı Pazarlama Türkiye'deki 4 ve 5 Yıldızlı Oteller Üzerine Bir Uygulama*. Yayınlanmamış Doktora Tezi. Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü, Sivas.
- Kumar, D. A. ve Ravi, V. (2008). "Predicting Credit Card Customer Churn in Banks Using Data Mining". *International Journal of Data Analysis Techniques and Strategies*, 1(1): 4-28.
- Kumar, V. ve Shah, D. (2004). "Building and Sustaining Profitable Customer Loyalty for the 21st Century". *Journal of Retailing*, 80(4): 317-329.
- Lai, Y. ve Zeng, J. (2014). "Analysis of Customer Churn Behavior in Digital Libraries". *Program*, 48(4): 370-382.
- Lanz, L. H. ve Carmichael, M. (2015). "Digital Marketing Budgets for Independent Hotels Continuously Shifting to Remain Competitive in the Online World". *Boston Hospitality Review*, 3(3): 1- 8.
- Larivière, B. ve Van den Poel, D. (2005). "Predicting Customer Retention and Profitability by Using Random Forests and Regression Forests Techniques". *Expert Systems with Applications*, 29(2): 472-484.
- Lawrence, W. ve Sankaranarayanan, S. (2012). "Smart Agent Learning Based Hotel Search System-Android Environment". *International Journal of Information Technology and Computer Science*: 4(9): 9.
- Lee-Kelley, L., Gilbert, D. ve Mannicom, R. (2003). "How E-CRM Can Enhance Customer Loyalty". *Marketing Intelligence & Planning*, 21(4): 239-248.
- Lim, T., Lee, A. S., Chia, M. P. C., Soam, W. ve Liew, K. Y. (2017). "Loyalty Card Membership Challenge: A Study on Membership Churn and Their Spending Behaviour". *Archives of Business Research*, 5(6): 68-88.
- Liu, D. R., Liao, H. Y., Chen, K. Y. ve Chiu, Y. L. (2019). "Churn Prediction and Social Neighbour Influences for Different Types of User Groups in Virtual Worlds". *Expert Systems*, 36: 1-19.
- Ma, Y., Xiang, Z., Du, Q. ve Fan, W. (2018). "Effects of User-Provided Photos on Hotel Review Helpfulness: An Analytical Approach with Deep Learning". *International Journal of Hospitality Management*, 71: 120-131.

- Magnini, V. P., Honeycutt Jr., E. D. ve Hodge, S. K. (2003). "Data Mining for Hotel Firms: Use and Limitations". *Cornell Hotel and Restaurant Administration Quarterly*, 44(2): 94-105.
- Maier, T. A. ve Prusty, S. (2016). "Managing Customer Retention in Private Clubs Using Churn Analysis: Some Empirical Findings". *Journal of Hospitality Marketing & Management*, 25(7): 797-819.
- Markopoulos, G., Mikros, G., Iliadi, A. ve Lontos, M. (2015). "Sentiment Analysis of Hotel Reviews in Greek: A Comparison of Unigram Features". V. Katsoni (Ed.), *Cultural Tourism in a Digital Era Springer Proceedings in Business and Economics*. Springer International Publishing, Athens, 373-383.
- McKim, B. ve Hughes, A. M. (2001). "How to Measure Customer Relationship Management Success". *Journal of Database Marketing & Customer Strategy Management*, 8(3): 224-231.
- Menard, S. (2011). "Standards for Standardized Logistic Regression Coefficients". *Social Forces*, 89(4): 1409-1428.
- Miguéis, V. L., Van den Poel, D., Camanho, A. S. ve e Cunha, J. F. (2012a). "Modeling Partial Customer Churn: On the Value of First Product-Category Purchase Sequences". *Expert Systems with Applications*, 39(12): 11250-11256.
- Miguéis, V. L., Van den Poel, D., Camanho, A. S. ve e Cunha, J. F. (2012b). "Predicting Partial Customer Churn Using Markov for Discrimination for Modeling First Purchase Sequences". *Advances in Data Analysis and Classification*, 6(4): 337-353.
- Min, H., Min, H., ve Emam, A. (2002). "A Data Mining Approach to Developing the Profiles of Hotel Customers". *International Journal Of Contemporary Hospitality Management*, 14(6): 274-285.
- Morales, D. R. ve Wang, J. (2010). "Forecasting Cancellation Rates for Services Booking Revenue Management Using Data Mining". *European Journal of Operational Research*, 202(2): 554-562.
- Morik, K. ve Kpcke, H. (2004). "Analysing Customer Churn in Insurance Data - A Case Study". J.F. Boulicaut, F. Esposito, F. Giannotti ve D. Pedreschi (Ed.), *Knowledge Discovery in Databases Lecture Notes in Computer Science*, Springer, Italy, 325-336.
- Müller, A. C. ve Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Sebastopol.

- Neslin, S. A., Gupta, S., Kamakura, W., Lu, J. ve Mason, C. H. (2006). "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models". *Journal of Marketing Research*, 43(2): 204-211.
- Nguyen, K. A. ve Coudounaris, D. N. (2015). "The Mechanism of Online Review Management: A Qualitative Study". *Tourism Management Perspectives*, 16: 163-175
- Nilashi, M., Bagherifard, K., Rahmani, M. ve Rafe, V. (2017). "A Recommender System for Tourism Industry Using Cluster Ensemble and Prediction Machine Learning Techniques". *Computers & Industrial Engineering*, 109: 357-368.
- O'Connor, P. (2010). "Managing a Hotel's Image on Tripadvisor". *Journal of Hospitality Marketing & Management*, 19: 754-772
- Odabaşı, Y. (2004). *Satış ve Pazarlamada Müşteri İlişkileri Yönetimi*. Sistem Yayıncılık, İstanbul.
- O'Mahony, M. P. ve Smyth, B. (2009). "Learning to Recommend Helpful Hotel Reviews". *Proceedings of the 3rd ACM Conference on Recommender Systems*, Ekim 2009, New York, 305-308.
- Özgen, E. (2011). "Halkla İlişkiler ve Yeni Medyanın İşaret Ettiği Sorunlar Üzerine Bir Tartışma". *Global Media Journal*, 1(2): 84-105
- Padilla-Meléndez, A. ve Garrido-Moreno, A. (2014). "Customer Relationship Management in Hotels: Examining Critical Success Factors". *Current Issues in Tourism*, 17(5): 387-396.
- Palmer, A., McMahon-Beattie, U. ve Beggs, R. (2000). "A Structural Analysis of Hotel Sector Loyalty Programmes". *International Journal of Contemporary Hospitality Management*, 12(1): 54-60.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., ... ve Vanderplas, J. (2011). "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research*, 12(Oct): 2825-2830.
- Perikos, I., Kovas, K., Grivokostopoulou, F. ve Hatzilygeroudis, I. (2017). "A System for Aspect-Based Opinion Mining of Hotel Reviews". *Proceedings of the 13th International Conference on Web Information Systems and Technologies - Volume 1*, 25-27 Nisan 2017, Porto, 388-394.
- Phillips, P., Zigan, K., Silva, M. M. S. ve Schegg, R. (2015). "The Interactive Effects of Online Reviews on the Determinants of Swiss Hotel Performance: A Neural Network Analysis". *Tourism Management*, 50: 130-141.
- Prasad, U. D. ve Madhavi: (2012). "Prediction of Churn Behavior of Bank Customers Using Data Mining Tools". *Business Intelligence Journal*, 5(1): 96-101.

- Qi, J., Zhang, L., Liu, Y., Li, L., Zhou, Y., Shen, Y., Liang, L. ve Li, H. (2009). “ADTreesLogit Model for Customer Churn Prediction”. *Annals of Operations Research*, 168(1): 247-265
- Quinlan, J. R. (1996). “Bagging, Boosting and C4. 5”. *AAAI/IAAI*, (1): 725-730.
- Quinlan, J. R., ve Rivest, R. L. (1989). “Inferring Decision Trees Using the Minimum Description Length Principle”. *Information and Computation*, 80(3), 227-248.
- Qureshi: A., Rehman, A. S., Qamar, A. M., Kamal, A. ve Rehman, A. (2013). “Telecommunication Subscribers' Churn Prediction Model Using Machine Learning”. *Digital Information Management (ICDIM): 2013 Eighth International Conference*, 10-12 Eylül 2013, Islamabad, 131-136.
- Raut, V. B. ve Londhe, D. D. (2014). “Opinion Mining And Summarization Of Hotel Reviews”. *2014 International Conference on Computational Intelligence and Communication Networks*, 14-16 Kasım 2014, Bhopal, 556-559.
- Rebala, G., Ravi, A. ve Churiwala, S. (2019). *An Introduction to Machine Learning*. Springer, Switzerland.
- Reichheld, F. F., Markey Jr, R. G. ve Hopton, C. (2000). “The Loyalty Effect-The Relationship Between Loyalty and Profits”. *European Business Journal*, 12(3): 134-134.
- Rud, O. P. (2001). *Data Mining Cookbook: Modeling Data for Marketing, Risk, and Customer Relationship Management*. John Wiley & Sons, New York.
- Ryals L. ve Payne A. (2001). “Customer Relationship Management in Financial Services: Towards Information-Enabled Relationship Marketing”. *Journal of Strategic Marketing*, 9(1): 3-27
- Salihli Birdir, S. (2004). *Otel İşletmelerinde Hedef Pazar Seçimi ve Konumlandırma*. Detay Yayıncılık, Ankara.
- Sarmaniotis, C., Assimakopoulos, C. ve Papaioannou, E. (2013). “Successful Implementation of CRM in Luxury Hotels: Determinants And Measurements”. *EuroMed Journal of Business*, 8(2): 134-153.
- Shi, H. X. ve Li, X. J. (2011). “A Sentiment Analysis Model for Hotel Reviews Based on Supervised Learning”. *International Conference on Machine Learning and Cybernetics*, 10 - 13 Temmuz 2011, Guilin, 950-954.
- Shimada, K., Inoue, S., Maeda, H. ve Endo, T. (2011). “Analyzing Tourism Information on Twitter for A Local City”. *First ACIS International Symposium on Software and Network Engineering*, 19-20 Aralık 2011, Seoul, 61-66.

- Sigala, M. (2005). "Integrating Customer Relationship Management in Hotel Operations: Managerial and Operational Implications". *International Journal of Hospitality Management*, 24(3): 391-413.
- Sigala, M. (2006). "E-Customer Relationship Management in the Hotel Sector: Guests' Perceptions of Perceived E-Service Quality Levels". *Turizam: Međunarodni Znanstveno-Stručni Časopis*, 54(4): 333-344.
- Smola, A. ve Vishwanathan: V. N. (2008). *Introduction to Machine Learning*. Cambridge University, London.
- Solomon, M. (2000). "Like ERP, CRM Systems Can Be A Struggle to Launch". *Computerworld*, 34(26): 51-51.
- Srivastava, S. K., Chandra, B. ve Srivastava, P. (2019). "The Impact of Knowledge Management and Data Mining on CRM in the Service Industry". Nath V., Mandal J. (Ed.) *Nanoelectronics, Circuits and Communication Systems*, 511, Springer, Singapore, 37-52.
- Suh, E. ve Alhaery, M. (2016). "Customer Retention: Reducing Online Casino Player Churn Through the Application of Predictive Modeling". *UNLV Gaming Research & Review Journal*, 20(2): 63-84.
- Şeker, E. (2016). "Müşteri Kayıp Analizi (Customer Churn Analysis)". *YBS Ansiklopedi*, 3(1): 26-29.
- Tamaddoni Jahromi, A., Sepehri, M. M., Teimourpour, B. ve Choobdar: (2010). "Modeling Customer Churn in A Non-Contractual Setting: The Case of Telecommunications Service Providers". *Journal of Strategic Marketing*, 18(7): 587-598.
- Tanford, S., Raab, C. ve Kim, Y. S. (2013). "A Model of Hotel Defection at the Purchasing Stage". *Journal of Hospitality Marketing & Management*, 22(8): 805-831.
- Tran, T. K. ve Phan, T. T. (2015). "An Upgrading Sentivoice-A System for Querying Hotel Service Reviews Via Phone". *2015 International Conference on Asian Language Processing (IALP)*, 24-25 Ekim 2015, Suzhou, 115-118.
- Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G. ve Chatzisavvas, K. C. (2015). "A Comparison of Machine Learning Techniques for Customer Churn Prediction". *Simulation Modelling Practice and Theory*, 55: 1-9.
- Van den Poel, D. (2003). "Predicting Mail-Order Repeat Buying: Which Variables Matter?". *Tijdschrift Voor Economie En Management*, 48(3): 371-404.
- Van den Poel, D. ve Larivière, B. (2004). "Customer Attrition Analysis For Financial Services Using Proportional Hazard Models". *European Journal of Operational Research*, 157(1): 196-217.

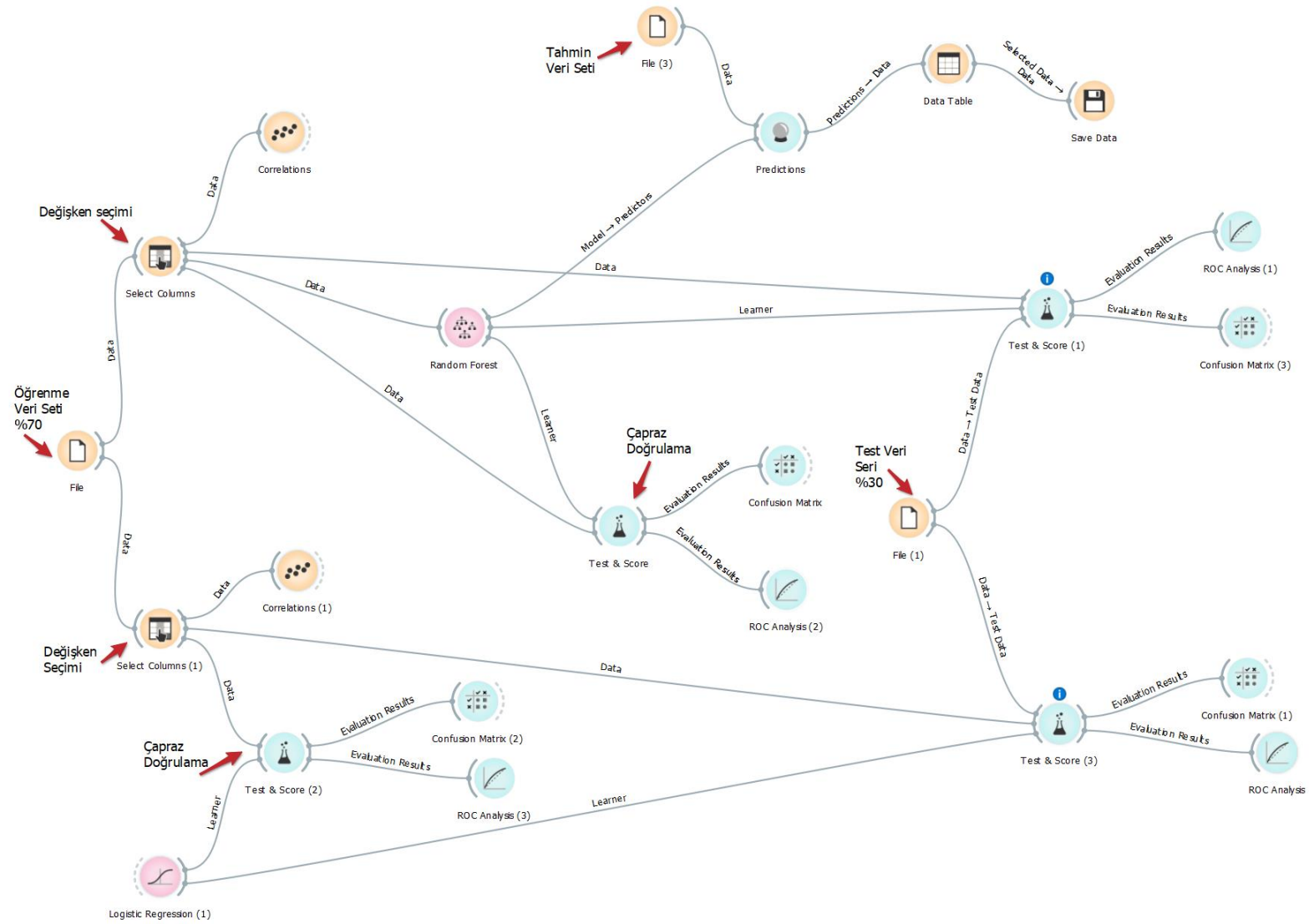
- Venturini, W. T. ve Benito, Ó. G. (2015). “CRM Software Success: A Proposed Performance Measurement Scale”. *Journal of Knowledge Management*, 19(4): 856-875.
- Verbeke, W., Martens, D., Mues, C. ve Baesens, B. (2011). “Building Comprehensible Customer Churn Prediction Models with Advanced Rule Induction Techniques”. *Expert Systems with Applications*, 38(3): 2354-2364.
- Vilcassim, N. J. ve Jain, D. C. (1991). “Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity”. *Journal of Marketing Research*, 28(1): 29-41.
- Wang, Y. F., Chiang, D. A., Hsu, M. H., Lin, C. J. ve Lin, I. L. (2009). “A Recommender System to Avoid Customer Churn: A Case Study”. *Expert Systems with Applications*, 36(4): 8071-8075.
- Wei, C. P. ve Chiu, I. T. (2002). “Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach”. *Expert Systems with Applications*, 23(2): 103-112.
- Wei, X., Luo, X., Li, Q., Zhang, J. ve Xu, Z. (2015). “Online Comment-Based Hotel Quality Automatic Assessment Using Improved Fuzzy Comprehensive Evaluation And Fuzzy Cognitive Map”. *IEEE Transactions on Fuzzy Systems*, 23(1): 72-84.
- Wilson, H., Daniel, E., ve McDonald, M., “Factors for Success in Customer Relationship Management (CRM) Systems”, *Journal of Marketing Management*, 18(1-2): (2002): 193-219.
- Xiang, Z., Du, Q., Ma, Y. ve Fan, W. (2018). “Assessing Reliability of Social Media Data: Lessons from Mining Tripadvisor Hotel Reviews”. *Information Technology & Tourism*, 18(1-4): 43-59.
- Yang, X. S. (2019). *Introduction to Algorithms for Data Mining and Machine Learning*. Academic Press, London.
- Yang, Y., Pan, B. ve Song, H. (2014). “Predicting Hotel Demand Using Destination Marketing Organization’s Web Traffic Data”. *Journal of Travel Research*, 53(4): 433-447.
- Yang, Y., Tang, J., Luo, H. ve Law, R. (2015). “Hotel Location Evaluation: A Combination of Machine Learning Tools and Web GIS”. *International Journal of Hospitality Management*, 47: 14-24.
- Yu, Y. (2016). “Aspect-based Sentiment Analysis on Hotel Reviews”. *International Journal for Scientific Research and Development*, 4(6): 726-729.
- Zhang, X., Zhu, J., Xu, S. ve Wan, Y. (2012). “Predicting Customer Churn Through Interpersonal Influence”. *Knowledge-Based Systems*, 28: 97-104.

- Zhao, Y., Dong, S. ve Yang, J. (2015). "Effect Research of Aspects Extraction for Chinese Hotel Reviews Based on Machine Learning Method". *International Journal of Smart Home*, 9(3): 23-34.
- Zheng, W. ve Ye, Q. (2009). "Sentiment Classification of Chinese Traveler Reviews by Support Vector Machine Algorithm". *Third International Symposium on Intelligent Information Technology Application*, 21-22 Kasım 2009, NanChang, 335-338.
- Zorn, S., Jarvis, W. ve Bellman, S. (2008). "Identifying Customers Likely to Churn". *ANZMAC 2008 Marketing: Shifting the Focus from Mainstream to Offbeat*, 1-3 Aralık 2008, Olympic Park, Sydney, 1-8.
- Zvarevashe, K. ve Olugbara, O. O. (2018). "A Framework for Sentiment Analysis with Opinion Mining of Hotel Reviews". *2018 Conference on Information Communications Technology and Society (ICTAS)*, 8-9 Mart 2018, Durban, 1-4.

İnternet Kaynakları

- Arruza, M., Pericich, J. ve Straka, M. (2016). "The Automated Travel Agent: Hotel Recommendations Using Machine Learning". <http://cs229.stanford.edu/proj2016spr/report/017.pdf> (erişim tarihi: 10.11.2019)
- Avon, V. (2016). "Machine Learning Techniques for Customer Churn Prediction in Banking Environments", <http://tesi.cab.unipd.it/53212/> (erişim tarihi: 10.10.2019)
- Classification: Accuracy. <https://developers.google.com/machine-learning/crash-course/classification/accuracy> (erişim tarihi: 20.09.2019)
- Precision-Recall. https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html#sphx-glr-auto-examples-model-selection-plot-precision-recall-py (erişim tarihi: 10.08.2019)
- Reputation.com. <https://www.reputation.com/industries/hospitality/> (erişim tarihi: 01.11.2019)
- ReviewPro. <https://www.reviewpro.com/products/reputation-management-for-hotels/> (erişim tarihi: 01.11.2019)
- Şeker, S.E. ve Doğan, D. (2018). *Knime ile Uçtan Uca Veri Bilimi*. http://sadiyvrenseker.com/wp-content/uploads/veribilimi_knime.pdf (erişim tarihi: 10.12.2019)
- Ticktockboom. <https://www.ttboom.com/sosyal-medya-takip-boom-sonar/> (erişim tarihi: 01.11.2019)
- Weka.attributeSelection.GreedyStepwise. <http://weka.sourceforge.net/doc.dev/weka/> (erişim tarihi: 10.08.2019)

K 1 - ORANGE ÇALIŞMA MODELİ



ÖZGEÇMİŞ

Adı ve SOYADI	Aslıhan DURSUN CENGİZCİ
Doğum Yeri - Tarihi	İstanbul – 26.11.1985
EĞİTİM DURUMU	
Mezun Olduğu Lise	Antalya Anadolu Otelcilik ve Turizm Meslek Lisesi
Lisans Diploması	Akdeniz Üniversitesi, Turizm ve Otelcilik Yüksekokulu, <i>Konaklama İşletmeciliği</i>, 2008, GPA: 3.43/4 ▪ l'Université de Perpignan Via Domitia, Institut Jacques Maillot, <i>Management du Tourisme et Hôtellerie</i>, değişim programı (2006/07)
Yüksek Lisans Diploması	Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Turizm ve Otel İşletmeciliği, 2015
Tez Konusu	Otel İşletmelerinde Müşteri Profili Oluşturmada Veri Madenciliği Tekniklerinin Kullanılması : RFM Modeli Örneği
Doktora Diploması	Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Turizm İşletmeciliği, 2020
Tez Konusu	Otel İşletmelerinde Kayıp Müşteri Tahminlemesi
Yabancı Diller	İngilizce Yökdil 2017 – 96,25 Fransızca Amerikan Kültür 2010 – C1 Sertifikası
BİLİMSEL FAALİYETLER	
Makaleler ▪ Yılmaz, Y., Ünal, C., & Dursun, A. (2019). Otel yöneticilerinin 2016 Türkiye turizm krizine ilişkin öngörülleri ve kriz yönetim uygulamaları. <i>Journal of Yaşar University (TR Index)</i>: 14(56): 468-488. https://dergipark.org.tr/en/pub/jyasar/issue/49978/513121 ▪ Caber, M., Ünal, C., Cengizci, A. D., & Güven, A. (2019). Conflict management styles of professional tour guides: A cluster analysis. <i>Tourism Management Perspectives (SSCI)</i>: 30, 89-97. https://doi.org/10.1016/j.tmp.2019.02.004 ▪ Albayrak, T., Karasakal, S., Kocabulut, Ö., & Dursun, A. (2019). Customer Loyalty Towards Travel Agency Websites: The Role of Trust and Hedonic Value. <i>Journal of Quality Assurance in Hospitality & Tourism (ESCI-Scopus)</i>: 1-28. https://doi.org/10.1080/1528008X.2019.1619497 ▪ Albayrak, T., Dursun, A., & Ünal, C. (2018). Do tourists have different motivations for online travel purchasing? A segmentation of the Russian market. <i>Journal of Vacation Marketing (SSCI)</i> doi.org/10.1177/1356766718814091	

- Yılmaz, Y., Ünal, C., & **Dursun, A.** (2018). Impacts of operational management proficiency levels of hotels on operational and marketing related decisions in time of crisis. *Ottoman: Journal of Tourism and Management Research*, 3(2): 290-305. [doi:10.26465/ojtmr.2018339511](https://doi.org/10.26465/ojtmr.2018339511)
- Karasakal, S., & **Dursun, A.** (2018). Destinasyona yönelik itici ve çekici faktörler: Bir literatür incelemesi. *Journal of Recreation and Tourism Research*, 5(1): 29-37. [ISSN:2148-5321](https://doi.org/10.26465/ojtmr.2018339511)
- Ünal, C., **Dursun, A.**, & Caber, M. (2017). A study of domestic honeymoon tourism in Turkey. *European Journal of Tourism, Hospitality and Recreation*, 8(1): 65-74. (ESCI -Scopus) doi.org/10.1515/ejthr-2017-0006
- **Dursun, A.**, & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives*, 18, 153-160. (SSCI) [doi:10.1016/j.tmp.2016.03.001](https://doi.org/10.1016/j.tmp.2016.03.001)

Ulusal ve Uluslararası Bilimsel Toplantılarda Sunulan Bildiriler

- Albayrak, T., Güzel, Ö., Caber, M., Kocabulut, Ö., **Dursun Cengizci, A.** & Güven, A. (2019). Turistlerin kalabalık algısının alışveriş davranışlarına etkisi. 20. Ulusal Turizm Kongresi, Eskişehir.
- Yılmaz Y., Ünal C. & **Dursun A.** (2017). The Effects of Distribution Channels on the Selling Prices of Hotels in Time of Crisis, ICTMM 2017: 19th International Conference on Tourism Marketing and Management, Kyoto, Japan, 16 - 17 November, 798 – 801
- Doğan O., Karasakal S., **Dursun A.** & Ünal C. (2017). Effects of Website Quality Dimensions on Repurchase Intention in Airline Industry, 2nd International Conference Tourism Dynamics and Trends, Sevilla, Spain, 26 - 29 June.
- Ünal C. & **Dursun A.** (2016). Honeymoon Tourism Market: A Study on Domestic Honeymoon Tourists Visiting Antalya, Turkey, 1st. International Conference on Tourism Dynamics and Trends, Antalya, Turkey, 04 - 07 May, 183 - 195.
- Güven A., Ünal C. & **Dursun A.** (2016). The Roles of Personality and Conflict Management Styles on the Success of Professional Tourist Guides, 1st. International Conference On Tourism Dynamics and Trends, Antalya, Turkey, 04 - 07 May, 414 - 419.
- **Dursun A.**, Ünal C. & Caber M. (2016). Sorumlu Çevresel Davranış ile Doğa-temelli Rekreatyon Tercihleri Arasındaki İlişki: Antalya'ya Gelen Alman Turistlere Yönelik Bir Araştırma, 2. Ulusal Sürdürülebilir Turizm Kongresi, Adrasan, Antalya, 28 - 30 April, 273 - 288.

Projeler

- Turist Alışverişinde Deneyim, Duygu ve Memnuniyet Arasındaki İlişkide Algılanan Kalabalığın Etkisi, TUBİTAK 1002, 117K960, *Araştırmacı*.

Ders Asistanlıkları

TRM 111 Information Technologies, 2019/20
 TRM 446 Convention and Conference Management, 2018/19
 BUSI 111 Information Technologies, 2018/19
 TRM 441 Tourism Research Methodology, 2018/19
 TRM 442 Capstone Project in Tourism Hotel Management, 2017/18
 TRM 446 Convention and Conference Management, 2017/18
 TRM 351 TQM for Tourism and Hotel Management, 2016/17

İŞ DENEYİMİ**Akademik İş Deneyimi**

Araştırma Görevlisi Turizm Fakültesi Antalya Bilim Üniversitesi 2016- devam

Profesyonel İş Deneyimi

- | | | |
|---|------------------------------|------------------------------|
| ▪ CRM Müdürü
-PRM360 ⁰ mobil uygulama
geliştirme | Papillon Hotels
& Resorts | Haziran 2014-
Ekim 2015 |
| ▪ Kalite Müdürü
-Entegre Yönetim Sistemi
(ISO 9001, 14001, 18001,
22000 & 10002) | Papillon Hotels
& Resorts | Şubat 2014 –
Haziran 2014 |
| ▪ Kalite Yönetim Sistemleri
Uzmanı
-Organizasyonel Gelişim
Projesi | Papillon Hotels
& Resorts | Haziran 2011 –
Şubat 2014 |
| ▪ Kabin Memuru | Freebird Airlines | Mart - Aralık 2010 |
| ▪ Satış ve Pazarlama Asistanı &
Turizm Koordinatörü Asistanı | Papillon Hotels
& Resorts | Haziran 2008 –
Ocak 2010 |
| ▪ Mini Club G.O.
-Animasyon | Club Med
Palmiye | Haziran - Ağustos
2006 |

E-Posta

aslihan.dursun@antalya.edu.tr