



T.C.

ALANYA ALAADDİN KEYKUBAT ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

TURİZM İŞLETMECİLİĞİ ANABİLİM DALI

**MAKİNE ÖĞRENMESİNE DAYALI ÇEVİRİMİÇİ OTEL GÖRÜŞLERİNİN
DUYGU ANALİZİ VE KONU MODELLEME YAKLAŞIMLARIYLA
DEĞERLENDİRİLMESİ**

Doktora Tezi

Egemen Güneş TÜKENMEZ

**Danışman
Prof. Dr. Kemal KANTARCI**

ALANYA

2023

T.C.
ALANYA ALAADDİN KEYKUBAT ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

MAKİNE ÖĞRENMESİNE DAYALI ÇEVİRİMİÇİ OTEL GÖRÜŞLERİNİN
DUYGU ANALİZİ VE KONU MODELLEME YAKLAŞIMLARIYLA
DEĞERLENDİRİLMESİ

Doktora Tezi

Egemen Güneş Tükenmez
Anabilim Dalı: Turizm İşletmeciliği
Program Adı: Turizm İşletmeciliği ve Otelcilik

Danışman
Prof. Dr. Kemal KANTARCI

ALANYA
(2023)

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilemeyen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmamın Alanya Alaaddin Keykubat Üniversitesi tarafından kullanılan “bilimsel intihal tespit programıyla tarandığını ve “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm.

.....

Egemen Güneş TÜKENMEZ

TEŞEKKÜR

Uzun uğraşlar, sabır ve disiplin gerektiren doktora sürecinde, bana her türlü desteği sağlayan ve yol gösteren danışman hocam Prof. Dr. Kemal KANTARCI'ya, fikirleri ile tez sürecime büyük katkı sağlayan başta tez izleme jüri üyelerim Prof. Dr. Murat Alper BAŞARAN ve Dr. Öğr. Üyesi Serpil KOCAMAN'a, bu süreçte bana yardımcı olan tüm hocalarıma, varlıkları ile büyük mutluluk duyduğum kardeşim Alp Burteçin TÜKENMEZ, annem Serpil TÜKENMEZ ve babam Kerim TÜKENMEZ'e ve hayatımın her anında yanımda olan sevgili eşim Özge KURŞUN TÜKENMEZ'e teşekkürlerimi sunarım.

Egemen Güneş TÜKENMEZ
Alanya, 2023

ÖZET

MAKİNE ÖĞRENMESİNE DAYALI ÇEVİRİMİÇİ OTEL GÖRÜŞLERİNİN DUYGU ANALİZİ VE KONU MODELLEME YAKLAŞIMLARIYLA DEĞERLENDİRİLMESİ

Egemen Güneş TÜKENMEZ

Turizm İşletmeciliği Anabilim Dalı

Alanya Alaaddin Keykubat Üniversitesi, Lisansüstü Eğitim Enstitüsü,

Temmuz, 2023 (120 Sayfa)

Otel işletmelerinin Web 2.0 platformları üzerinden sahip olduğu kullanıcı tarafından oluşturulan içerikler, tüketici memnuniyetini ve otellerin hizmet performansını gösteren önemli göstergelerden birisidir. Bu içeriklerin yapısal ve anlamsal özelliklerini ortaya çıkartmak, otel işletmelerinin hizmet performanslarının seviyesini görmesinde etkin bir rol oynamaktadır. Böylece oteller, hizmet sağladığı tüketicilerin memnuniyetine etki eden unsurları net bir şekilde belirleyebilmektedir. Metin madenciliği alanında kullanılan makine öğrenmesi yöntemleri, büyük veri kapsamında ele alınan bu içeriklerin etkili bir şekilde analiz edilmesini sağlamaktadır.

Bu çalışmada, Türkiye'nin önemli turizm merkezlerinden birisi olan Alanya'da faaliyet gösteren otellere yapılan çevrimiçi kullanıcı içerikleri incelenmiştir. Bu içeriklerde konaklama yapan kişilerin en fazla değindiği konular ve bu konuların tüketicilerin memnuniyetine olan etkileri, konu modelleme ve duygu analizi yöntemleri ile ortaya çıkarılmıştır.

Otellerin sağladığı olanaklar, animasyon ve personel konuları, tüketicilerin memnuniyetine en fazla pozitif etki eden unsurlar olmuştur. Ön büro ve oda konuları ise negatif olarak memnuniyete etki eden konulardır. Personel konusu, tüketicilerin pozitif yorumda bulunmasına en fazla katkı sağlayan konu olurken, önbüro konusu ise tüketicilerin en fazla negatif yorumda bulunmasına sebep olmaktadır. Ayrıca araştırmada ortaya çıkan konulardan duyulan memnuniyet seviyesi, otel sınıfına ve seyahat türlerine göre değişiklik göstermektedir.

Anahtar Sözcükler: Otel İşletmeciliği, Metin Madenciliği, Makine Öğrenmesi, Duygu Analizi, Konu Modelleme

ABSTRACT

EVALUATION OF ONLINE HOTEL REVIEWS BASED ON MACHINE LEARNING WITH SENTIMENT ANALYSIS AND TOPIC MODELING APPROACHES

Egemen Güneş TÜKENMEZ

Department of Tourism Management

Graduate School of Alanya Alaaddin Keykubat University,

July, 2023

User-generated content in Web 2.0 platforms is a significant indicator of consumer satisfaction and service performance for hotel business. Revealing the structural and semantic attributes of this content plays an essential role in assessing the level of hotel service performance. By this way, hotels can precisely identify the factors influencing their consumers satisfaction. Machine learning techniques in the field of text mining effectively provides to analyze these contents within the scope of big data.

In this study, online user-generated content related to hotels operating in Alanya, one of Turkey's prominent tourism centers, has been examined. The topics most emphasized by consumers who have stayed at these hotels and the effects of these topics on consumer satisfaction have been appeared via topic modeling and sentiment analysis methods.

According to the findings, customers staying at the related hotels have identified animation, staff, and amenities as the topics that positively influence consumer satisfaction. However, front desk and room-related matters have been identified as the topics with the most negative impact. The staff topic has the most positive impact on consumers' positive comments but, the front desk topic tends to induce negative comments. In addition, the level of satisfaction with the topics revealed in the research varies according to the hotel class and travel types.

Keywords: Hotel Business, Text Mining, Machine Learning, Sentiment Analysis, Topic Modeling

İÇİNDEKİLER

İÇ KAPAK SAYFASI

ONAY SAYFASI	i
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	ii
TEŞEKKÜR	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER	vi
ŞEKİLLER LİSTESİ	viii
TABLolar LİSTESİ.....	xi
SİMGELER ve KISALTMALAR	xii
1. GİRİŞ	1
2. LİTERATÜR	3
2.1 Hizmet Kavramı.....	3
2.2 Hizmetlerin Özellikleri	4
2.3 Kalite Kavramı.....	5
2.4 Hizmet Kalitesi Çerçevesinde Geliştirilen Ölçekler	6
2.5 Konaklama İşletmelerinde Hizmet Kalitesi	8
2.6 Konaklama İşletmelerinde Çevrimiçi Hizmet Kriterleri	11
2.7 Konaklama Sektöründe Kullanıcı Tarafından Oluşturulan İçeriklerin Önemi	13
2.8 Kullanıcı Tarafından Oluşturulan İçerikler ve Metin Madenciliği	16
2.9 Duygu Analizi (Sentiment Analysis) Kavramı	17
2.9.1 Konaklama işletmelerinde duygu analizi çalışmaları	19
2.10 Konu Modelleme (Topic Modeling) Kavramı.....	21
2.10.1 Konaklama işletmelerinde konu modelleme çalışmaları	22
3. YÖNTEM	27
3.1 Araştırmanın Amacı	27
3.2 Araştırmanın Yöntemi	29
3.3 Araştırmanın Modeli	30
3.3.1 Veri setinin hazırlanması	30
3.3.2 Ön işleme (preprocessing) aşaması.....	31
3.3.3 Duygu analizi (sentiment analysis).....	33
3.3.4 Gizli dirichlet ayrımı (latent dirichlet allocation)	35

3.3.5 Kelime çantası (bag of words)	40
3.3.6 Terim ve ters belge frekansı (term and inverse document frequency)	40
3.3.7 Doküman gömme (document embedding)	43
3.3.8 Sinir ağları (neural networks)	43
3.3.9 Karmaşıklık matrisi (confusion matrix).....	45
3.3.10 Alıcı işletim karakteristiği (receiver operating characteristic)	46
3.3.11 Model doğrulama yöntemleri (model validation methods)	47
4. BULGULAR.....	48
4.1 Veri Setine Ait Bilgiler	48
4.2 Duygu Analizi Sonuçları	56
4.2.1 Duygu analizi sınıflandırılması.....	57
4.3 Konu Modelleme Sonuçları	62
4.4 Konu Modelleme Yöntemiyle Ortaya Çıkan Konular	67
4.5 Duygu ve Konu Sınıfı İlişkileri	76
4.6 Konulara Göre Duygu Sınıfı Dağılımları	77
4.7 Otel Sınıflarına Göre Konu Modelleme ve Duygu Analizi Sonuçları.....	78
4.8 Uzaklık Matrisi ve Hiyerarşik Kümeleme Analizi Sonuçları	85
4.9 Uyum Analizi Sonuçları	88
4.10. Lojistik Regresyon Analizi Sonuçları	91
5. TARTIŞMA, SONUÇ VE ÖNERİLER	96
KAYNAKLAR	103
EK: İNTİHAL RAPORU.....	118
ÖZGEÇMİŞ	119

ŞEKİLLER LİSTESİ

Şekil 3.1 VADER Formülü	34
Şekil 3.2 Gizli Dirichlet Ayrımı Çalışma Şekli	38
Şekil 3.3 Gizli Dirichlet Ayrımı Yapısı	38
Şekil 3.4 Doküman ve Kelime Matrisi Oluşumu	38
Şekil 3.5 Doküman ve Kelimelere Göre Konu Oluşumu	39
Şekil 3.6 Terim ve Ters Doküman Matrisi Formülleri	42
Şekil 3.7 MLP Sinir Ağları Modeli	45
Şekil 4.1 Araştırmada Elde Edilen Örnek Veri Seti	48
Şekil 4.2 Otel Sınıflarına Göre Puan Dağılımları	50
Şekil 4.3 Otel Sınıflarına Göre Yorum Sayıları	51
Şekil 4.4 Yıllara Göre Yorum Sayıları	52
Şekil 4.5 Seyahat Tiplerine Göre Yorum Sayıları	53
Şekil 4.6 Otel Sınıflarına Göre Puan Dağılımı	54
Şekil 4.7 Yıllara Göre Otel Sınıflarına Yapılan Puanların Dağılımları	54
Şekil 4.8 Yıllara ve Otel Sınıflarına Göre Yorumların Sayısal Dağılımı	55
Şekil 4.9 Veri Setindeki Kelime Bulutu	55
Şekil 4.10 Otel Sınıflarına Göre Duygu Analizi Sonuçları	56
Şekil 4.11 Yıllara Göre Duygu Analizi Sonuçları	56
Şekil 4.12 Karmaşıklık Matrisi Sayısal Gösterimi	58
Şekil 4.13 Karmaşıklık Matrisi Yüzdesel Gösterimi	59
Şekil 4.14 Duygu Analizi Karmaşıklık Matrisi Sayısal Gösterimi	59
Şekil 4.15 Duygu Analizi Karmaşıklık Matrisi Yüzdesel Gösterimi	60
Şekil 4.16 Negatif Duygu Sınıflarının ROC Eğrisi	60
Şekil 4.17 Nötr Duygu Sınıflarının ROC Eğrisi	61
Şekil 4.18 Pozitif Duygu Sınıflarının ROC Eğrisi	61

Şekil 4.19 Konu Modelleme Çapraşıklık Sonuçları	63
Şekil 4.20 Konu Modelleme Uyumluluk Sonuçları	63
Şekil 4.21 Olanaklar ve Otel Sınıfı Arasındaki ANOVA Analizi	65
Şekil 4.22 Olanaklar ve Duygu Türü Arasındaki ANOVA Analizi	65
Şekil 4.23 Olanaklar ve Seyahat Tipi Arasındaki ANOVA Analizi	66
Şekil 4.24 Duygu ve Otel Sınıfı Arasındaki Ki-Kare Analizi	66
Şekil 4.25 Duygu ve Seyahat Türü Arasındaki Ki-Kare Analizi	66
Şekil 4.26 Olanaklar	68
Şekil 4.27 Personel	69
Şekil 4.28 Önbüro.....	70
Şekil 4.29 Destinasyon	71
Şekil 4.30 Rezervasyon	72
Şekil 4.31 Animasyon.....	73
Şekil 4.32 Yiyecek-İçecek.....	74
Şekil 4.33 Oda	75
Şekil 4.34 Duygu Skoru ve Konular Arasındaki Korelasyon Analizi	76
Şekil 4.35 Konulara Göre Duygu Sınıflarının Dağılımı.....	78
Şekil 4.36 Olanaklar Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	79
Şekil 4.37 Personel Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	80
Şekil 4.38 Önbüro Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı.....	81
Şekil 4.39 Destinasyon Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	82
Şekil 4.40 Rezervasyon Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	83
Şekil 4.41 Animasyon Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	83
Şekil 4.42 Yiyecek-İçecek Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı....	84
Şekil 4.43 Odalar Konusunun Otel Sınıflarına Göre Duygu Türü Dağılımı	84
Şekil 4.44 Tüm Verilerin Hiyerarşik Kümeleme Analizi.....	86

Şekil 4.45 Pozitif Duygu Sınıfının Hiyerarşik Kümeleme Analizi	87
Şekil 4.46 Nötr Duygu Sınıfının Hiyerarşik Kümeleme Analizi.....	87
Şekil 4.47 Negatif Duygu Sınıfının Hiyerarşik Kümeleme Analizi	88
Şekil 4.48 Konu ve Otel Sınıfı Arasındaki Uyum Analizi	89
Şekil 4.49 Konu ve Duygu Sınıfı Arasındaki Uyum Analizi	90
Şekil 4.50 Konu ve Seyahat Tipi Arasındaki Uyum Analizi.....	91
Şekil 4.51 Negatif Duygu Sınıfının Nomogramda Gösterimi	93
Şekil 4.52 Nötr Duygu Sınıfının Nomogramda Gösterimi	94
Şekil 4.53 Pozitif Duygu Sınıfının Nomogramda Gösterimi	95

TABLÖLER LİSTESİ

Tablo 2.1 Ürünler ve Hizmetlerin Karakteristik Özellikleri	5
Tablo 2.2 Çevrimiçi Platformlarda Kullanılan Otel Hizmet Kriterleri	12
Tablo 3.1 Gizli Dirichlet Ayrımı Parametreleri	37
Tablo 3.2 BOW Doküman-Kelime Matrisi Oluşumu	41
Tablo 3.3 BOW ve Terim Frekansı Oluşumu	41
Tablo 3.4 BOW ve TF*IDF Matrisinin Oluşumu	42
Tablo 3.5 Karmaşıklık Matrisi	46
Tablo 4.1 Otel Sınıflarına Göre Veri Setindeki Bilgiler	49
Tablo 4.2 Değerlendirme Puanları ve Yorum Sayısı	49
Tablo 4.3 Otel Sınıflarına Göre Yorum Sayısı ve Yüzdeleri	50
Tablo 4.4 Yıllara Göre Yorum Sayıları ve Yüzdeleri	52
Tablo 4.5 Seyahat Tiplerine Göre Yorum Sayıları ve Yüzdeleri	53
Tablo 4.6 Sınama Seti ile Duygu Sınıflarının Sınıflama Performansı	57
Tablo 4.7 K-Katlamalı Doğrulama ile Duygu Sınıflarının Sınıflama Performansı	58
Tablo 4.8 Duygu Analizi Sınıflama Sonucu	59
Tablo 4.9 Konulara Göre Duygu Sınıflarının Dağılımı	78
Tablo 4.10 Uzaklık Matrisi	86
Tablo 4.11 Lojistik Regresyon Analizi Sonuçları	92

SİMGELER VE KISALTMALAR

Simgeler

α	Doküman ve konu dağılımını kontrol eden Dirichlet parametresi
β	Konu ve kelime dağılımını kontrol eden Dirichlet parametresi
θ	Doküman ve konu dağılım parametresi
ϕ	Konu ve kelime dağılımını sağlayan parametre

Kısaltmalar

ADAM	Adaptive Moment Estimation
AKTOB	Akdeniz Turistik Otelciler ve İşletmeciler Birliği
ALTİD	Alanya Turistik İşletmeciler Derneği
BOW	Bag of Words
CPU	Central Process Unit
CSV	Comma Seperated Values
DEM	Document Embedding
FN	False Negative
FP	False Positive
GDA	Gizli Dirichlet Ayrımı
IDF	Inverse Document Frequency
KTB	Kültür ve Turizm Bakanlığı
LDA	Latent Dirichlet Allocation
LOR	Log Odd Ratio
MLP	MultiLayer Perceptron
NLTK	Natural Language Toolkit
REGEX	Regular Expression
RELU	Recifier Linear Unit
ROC	Receiver Operating Characteristic
TF	Term Frequency
TN	True Negative
TP	True Positive
URL	Uniform Resource Loader
VADER	Valence Aware Dictionary and Sentiment Reasoner

1. GİRİŞ

Teknoloji ve bilişim alanında yaşanan gelişmeler, turizm sektöründe rekabeti arttıran ve insanların turizm sistemine dahil olmasını sağlayan en önemli etmelerin başında gelmektedir (Buhalis, 2000). Turizm sisteminin içerisinde en önemli paya sahip bileşenlerden birisi olan konaklama işletmeleri ve bu işletmelerin en büyük parçasını oluşturan otel işletmeleri, turizm sisteminde rekabette üst sıralara gelmek, hizmet kalitelerini arttırmak ve bu sayede tüketici memnuniyetini arttırabilmek için çeşitli önlemler almaktadır (Kandampully & Suhartanto, 2000). Bu sebeple tüketicilerin otel işletmelerinde geçirmiş olduğu zaman diliminde hangi sorunlarla karşılaştığını ve otel işletmelerinin sunmuş olduğu hizmetlerden hangilerinin tüketicilerin memnuniyetini etkilediğini araştırmak, bu işletmelerin hizmet kalitesini arttırmak için önemli bir unsur olmaktadır (Weng vd., 2012).

Özellikle internet kullanımının yaygınlaşmaya başlaması ve firmaların bu alanda yapmış olduğu yenilikler, tüketicilerin giderek deneyimlerini çevrimiçi platformlarda paylaşmasını hızlandırmıştır. Bu hızlanmanın en fazla yaşandığı sektörlerden birisini de konaklama sektörü oluşturmaktadır. Bir otel işletmesinde konaklama yapacak olan tüketiciler, giderek çevrimiçi internet sitelerinde konaklama yapmayı düşündüğü işletmeleri araştırmakta, bu işletmeler ile ilgili yazılan yorumları okumaktadır (Li vd., 2013). Eğer bir tüketici, konaklama yapmış olduğu bir işletmede olumsuz bir durum yaşar ise, bu durumu iade etme ya da değiştirme şansı bulunamamaktadır. Bu sebeple çevrimiçi mecralarda konaklama işletmelerinde daha önce konaklamış kişilerin paylaşmış olduğu deneyimlerin görülmesi, tüketicilerin önem verdiği bir konu olmaktadır (Mankad vd., 2016).

Konaklama sektöründe artan rekabet, doğal olarak otel işletmelerinin bu mecralarda kendi işletmeleri için belirtilen düşünceleri anlaması ve etkin bir şekilde analiz ederek işletme ile ilgili alacakları kararlara uygulaması, otel işletmelerinin geleceği açısından önemi bir durumdur. Web 2.0 araçlarının ve bu alandaki olanakların gelişmesi ile tüketiciler giderek bu alandaki mecralarda daha sık görüş bildirmektedir. Ancak kullanıcılar tarafından bu mecralarda oluşturulan içerikler, giderek büyümekte ve analiz edilmesi karmaşık bir süreç olmaktadır (Hoskisson & Wren, 2015).

Turistik destinasyonlarda bulunan otel işletmelerinin sayısı ve bu işletmelere yapılan yorumların sayısal olarak büyüklüğüne bakıldığında, bu yorumların etkili bir şekilde incelenmesi destinasyona olan imajın ortaya çıkarılmasında önemli bir husustur (Nowacki & Niezgoda, 2020).

Otel işletmelerine yapılan çevrimiçi yorumların önemi göz önüne alındığında, bu araştırmada Türkiye’de turizm sektöründe önemli paya sahip olan Alanya destinasyonundaki otellere Tripadvisor sitesi üzerinden yapılan yorumlar incelenmiştir. Analiz sürecinde makine öğrenmesi modelleri kullanılarak, konaklama yapan kişilerin memnuniyetini etki eden konular ortaya çıkarılmıştır. Elde edilen veriler, büyük veri kapsamında analizlere olanak sağlayan ve makine öğrenmesi yöntemlerinin kullandığı metin madenciliği yöntemleri ile farklı açılardan analiz edilmiştir. Araştırma kapsamında ele alınan otellere yapılan yorumlardaki konuların neler olduğu ve bu konuların tüketicilerin memnuniyetine etkisi ve konuların birbirleri ile olan ilişkileri ortaya çıkarılmıştır.

Çalışma beş bölümden oluşmaktadır. İlk bölümde hizmet kalitesi kavramına ve konaklama işletmelerinde uygulanan hizmet kalitesini ölçmek için geliştirilen ve literatürde sıkça kullanılan modeller ana hatları ile açıklanmıştır. Ayrıca otellerin hizmet kalitesini belirleyen ve çevrimiçi yorumlarda bulunan etmenlerin ne şekilde kullanıldığı belirtilmiştir. İkinci bölümde, metin madenciliği kapsamında, araştırmada kullanılan duygu analizi ve konu modelleme kavramları ele alınarak, bu alanda yapılan çalışmalar ve bu metotların büyük verinin analizinde sağlamış olduğu kolaylıklar belirtilmektedir. Üçüncü bölümde araştırmada kullanılan yöntemlerin çalışma prensipleri belirtilmiştir.

Dördüncü bölümde veri setinin genel özelliklerine ve araştırmada elde edilen bulgulara yer verilmektedir. Bu bölümde, metin setine uygulanan duygu analizi sonuçlarına ve bu analizin sınıflama performansı belirtilmiştir. Buna ek olarak, konu modelleme ile metin setinde 8 konu belirlenmiş olup, bu konuların neler olduğu gösterilmiştir. Bu konularda tüketicilerin duygu dağılımlarının ne şekilde olduğu ve otel sınıflarına göre nasıl değiştiği belirtilmiştir. Araştırmada ortaya çıkartılan konular, yorumlarda ne açıdan ele alındığı ve hangi konuların birbirleri ile ilişkili olduğu ise kümeleme ve uyum analizleri ile gösterilmiştir. Ayrıca oluşturulan konuların, hangi duygu sınıfına ne derecede etki ettiği ise lojistik regresyon analizinde gösterilmiştir. Beşinci bölümde ise, araştırmada elde edilen bulgular değerlendirilerek, öneriler getirilmiştir.

2. LİTERATÜR

Bu bölümde, yapılan tez çalışmanın kavramsal çerçevesi ve konu ile ilgili yapılan bilimsel çalışmalara yer verilmiştir. Otel işletmelerinin hizmet performansını gösteren unsurlardan birisini, bu alandaki çevrimiçi platformlardır. Bu platformlara tüketiciler tarafından yazılan değerlendirmeler, otel işletmelerinin hizmet performansının belirlenmesinde önemli göstergelerden birisini oluşturmaktadır. Metin madenciliği alanında yapılan gelişmeler ile, bu değerlendirmelerin analizleri makine öğrenmesi yöntemleri ile büyük veri analizi kapsamında etkin sonuçlar verilebilmesini sağlamaktadır. Bu sebeple hizmet ve hizmet kalitesi kavramları ile otel işletmelerinde hizmet performansının ölçülmesinde kullanılan başlıca değerlendirme ölçekleri ile beraber, kullanıcı tarafından oluşturulan içeriklerin otel performansını değerlendirmesindeki kriterler bu bölümde açıklanmıştır. Buna ek olarak, bu kriterlerin analiz edilmesinde yararlanılan makine öğrenmesi yaklaşımlarına yer verilmiştir.

2.1 Hizmet Kavramı

Hizmet kavramı ile ilgili olarak tarihsel süreç içerisinde birçok tanım geliştirilmiştir ve bu tanımlarda hizmet kavramının içermiş olduğu özellikler ve hangi unsurların hizmet olarak değerlendirilebileceği ve hizmetin temel unsurları farklı araştırmalarda belirtilmiştir. Hizmet kavramı, bir tarafın diğerine sunabileceği, esasen soyut olan ve herhangi bir şeyin mülkiyeti ile sonuçlanmayan bir eylem veya performanslar bütünüdür. (Kotler, 2011). Hizmet üretimi, direkt olarak fiziksel bir ürüne bağlı olabilir veya olmayabilir. Hizmetlere örnek olarak havacılık, konaklama, kiralama, eğitim gibi dallar örnek olarak verilebilmektedir. Bu sebeple günümüzde birçok pazar, tüketicilere ürün ve hizmetlerin oluşturmuş olduğu bütüncül ürünleri sunmaktadır. Şirketler bu üretmiş oldukları ürün ve hizmetleri sunarken, veri tabanları, kendi web siteleri gibi enformasyon alanlarından da yararlanmaktadır. (Kotler vd., 2009).

Ayrıca hizmet, mal üretiminin dışında olan ve somut bir çıktısı olmayan, soyut bir ürünü ifade etmektedir. Bu kapsamda fiziksel olarak nitelenemeyen ve hizmetin olduğu yerin, zamanın, hizmetin şeklinin ve yaratmış olduğu hissel boyutlarda fayda sağlandığı faaliyetler bütünüdür (Bayat vd., 2015). Hizmet kavramı, içermiş olduğu özellikler bakımından insanların günlük yaşamında sıkça kullandığı ve ondan vazgeçemeyeceği doyumlar olarak nitelendirilmiştir. Hizmetler, insanlara maddi

doyumların yanı sıra manevi doyumlar da sağlayan unsurların bütünüdür (Taşkın & Büyük, 2002).

Teknoloji ve inovasyon alanlarında yapılan gelişmeler, hizmet sektörünün gelişmesini sağlayan başlıca etmenler olmuştur. Yapay zeka, otomasyon, ulaşım alanlarında meydana gelen yenilikler, insanların hizmet almasını kolaylaştırmakta ve hizmet çeşitliliği sağlayarak tüketicilere ihtiyaçları doğrultusunda farklı seçenekler sunmaktadır (ResearchandMarkets, 2021). Hizmet sektörünün gelişim, ülkelerin ekonomilerine ve istihdam oranlarına da yansımaktadır. Dünya Bankası verilerine göre dünyadaki tüm istihdamın yarısı hizmet sektöründe çalışmaktadır. Türkiye ise dünya ortalamasının üzerinde olmakla birlikte, ülkedeki tüm istihdamın %56'sı hizmet sektöründe çalışmaktadır (Worldbank, 2021). Bu kapsamda hizmet kavramının ekonomik olarak değeri göz önüne alındığında, hizmet rir tüketici, firma ya da hükümetin ödemeye istekli olduğu soyut bir eylem veya kullanım olarak tanımlanmıştır (McConnel vd., 2009).

2.2 Hizmetlerin Özellikleri

Hizmetlerin genel yapısına bakıldığında, hizmetlerin mallardan farklı oldukları önemli temel yönleri bulunmaktadır. Hizmetler, standart bir halde üretilmesinin zor olduğu, elle tutulamayan ve kişi ya da gruplar tarafından yarar ve doyum meydana getiren soyut faaliyetlerin bütünüdür (Sayım & Aydın, 2011). Hizmetlerin genel olarak ürünlerden farkına bakıldığında, soyutluk, ayrılmazlık, değişkenlik, bozulabilirlik, katılım ve heterojenlik unsurları temel olarak gösterilmektedir (Parasuraman & vd., 1988). Hizmetler, ürünler gibi somut şekilde algılanan varlıklar değildir. Hizmetler ürünler gibi elle tutulamazlar (Sayım & Aydın, 2011) Hizmetlerin üretimi ve tüketimi aynı anda olduğu için hizmetlerde ayrılmazlık ilkesi bulunmaktadır. Bir hizmetin sunumu, müşteri ile hizmeti sağlayanlar arasında aynı anda gerçekleşmektedir (Çiçek & Doğan, 2009).

Hizmetler aynı zamanda hizmeti alan ve hizmeti verenin deneyimine ve bilgi birikimine bağlı bir unsur olduğu için değişkenlik özelliği göstermektedir. Hizmetler aynı zamanda ürünlerde olduğu gibi depolanamaz (Ustaahmetoğlu & Savcı, 2011). Otelde satılamayan bir oda, uçakta satılamayan bir koltuk depolanıp ileride satılmak üzere kullanılamaz. Hizmetler ayrıca kişilerin özelliklerine ve ihtiyaçlarına göre değişkenlik gösterdiği için homojen yapıda değildir. Ayrıca hizmetlerde tüketiciler ürünlere göre süreçlere daha fazla dahil olmaktadır (Türk, 2009).

Tablo 1’ de ürünler ve hizmetlerin karakteristik olarak ayrıldığı genel özellikler gösterilmektedir. Hizmetlerin ölçümü subjektiflik barındırır ve bu noktada ölçümü zor olmaktadır. Dağıtım sürecinde ise hizmetler üretildiği zamanda ve yerde tüketilmektedir. Hizmet sunumunda hizmeti birçok etmen etkilediği için ve kişiden kişiye hizmet algısı değiştiğinden ürünlerin aksine geniş bir esnekliğe sahiptir. Hizmetlerin sahip olduğu bu özelliklerden dolayı, hizmetlerdeki talep ürünlerin aksine ertelenemez ve depolanamaz. Ayrıca tüketici ile olan iletişim, hizmetlerin eş zamanlı tüketimine bağlı olarak ertelenemez. Ürünlerin üretiminde otomasyon ve makineleşme daha belirgin iken, hizmetlerde ana unsur insandır ve makineleşme hizmet sürecinde aracı olarak kullanılmaktadır. Ayrıca ürünleri satın alan tüketiciler, ürünün mülkiyetini elinde bulundurabilirken, hizmetin soyutluk özelliğinden dolayı mülkiyet tüketicilere geçemez (Gronroos, 1990; Timo & Fischer, 2005; Lopez & vd., 2018).

Tablo 2. 1 Ürünler ve hizmetlerin karakteristik özellikleri

Karakteristik Özellikler	Ürünler	Hizmetler
Ölçülebilirlik	Objektif	Subjektif
Süreç	Üretim	Teslimat
Dağıtım	Üreticiden ayrı	Üretim ile aynı
Kişisel satış	Somut	Soyut
Esneklik	Sınırlı	Geniş
Zaman Aralığı	Aydan yıla	Aynı gün içinde
Birim Özelliği	Kesin	Genel
Depolama	Olanak sağlanabilir	Olanak sağlanamaz
Talep	Ertelenebilir	Ertelenemez
Tüketici ile İletişim	Satış sırasında	Hizmet üretimi anında
Lokasyon	Her zaman tüketiciye yakın olmasına gerek yoktur	Türüne göre yerinde ya da uzak bir mekânda olabilir
Kaynak	Otomasyon veya bir insan kaynağı tarafından çalışan makinelerdir	Hizmet için makinelerin kullanımı genellikle insanların istihdamından daha az önemlidir
Sahiplik	Mülkiyet tüketicilere geçer	Mülkiyet tüketicilere geçemez

Kaynak: Gronroos, 1990; Timo & Fisher, 2005; Santana vd., 2018.

2.3 Kalite Kavramı

Kalite kavramı literatürde araştırmacılar ve örgütler tarafından farklı şekillerde tanımlanmış bir unsurdur. Amerikan Kalite Kontrol Derneği’nin yapmış olduğu kalite tanımında, bir ürünün ya da hizmetin daha önce ifade edildiği veya sürdürdüğü süreçte tüketicilerin ihtiyaçlarını karşılama kabiliyetine dayanan karakteristik özelliklerdir ve her türlü eksikliklerden arınmış bir ürün ya da hizmettir. Joseph Juran’a göre kalitenin anlamı kullanıma uygunluk, Philip Crosby’ye göre ise gerekliliklere uyum olarak nitelendirilmiştir (ASQ, 2023). Kalite aynı zamanda kullanıcı tarafında memnuniyet

veya memnuniyetsizlik uyandırabilecek bir teklifin değeriyle de ilgilidir (Tamilselvi, 2016). Deming ise kaliteyi tüketicilerin ihtiyaçlarını karşılamak olarak belirtmiştir. Ishikawa devamlı gelişme süreci, Feigenbaum ise müşteri memnuniyeti olarak kalite kavramını ele almıştır (Levis vd., 2008).

Kalite kavramının ürün bazlı yaklaşımda, ürün ile ilgili bir bileşenin varlığı veya yokluğu ile belirlenebilmektedir. Değer bazlı yaklaşımda ise kalite kavramı, kabullenebilir bir fiyat karşılığında elde edinilen performanstır. Kullanıcı bazlı kalite kavramı ise müşterilerin ihtiyaç ve tercihleridir (Thareja, 2017).

Kalite kavramı, hizmetlerin ölçülmesinde ve belirli standartlara ulaştırılmasında önemli bir basamak olmaktadır. Hizmetlerin özellikleri göz önüne alındığında, hizmet kalitesinin ölçümündeki zorluklar ve hizmetin sübjektif olması, hizmetleri alan tüketicilerin memnuniyetinin ölçülmesinde ve hizmet kalitesinin ölçülmesinde belirli ölçekler getirilmesi hassas bir durum olarak ortaya çıkmaktadır. Bu aşamada araştırmacılar, hizmet kalitesinin ölçülmesi için farklı yaklaşımlar ortaya koyarak, farklı sektörlerde uygulanmasının önünü açmışlardır.

2.4 Hizmet Kalitesi Çerçevesinde Geliştirilen Ölçekler

Hizmet kavramının sahip olduğu unsurlar göz önüne alındığında, hizmet ağırlıklı sektörlerde kalitenin ölçülmesi ve tanımlanması zor bir kavramdır. Bu sebeple, hizmet kalitesinin ölçülmesinde doğru şekilde dizayn edilen model ve yöntemlerle çalışılması, işletmeler için önem arz etmektedir (Kılıç & Eleren, 2009). Hizmetler, bir objeden ziyade performans olarak meydana gelmektedir. Çoğu hizmet, kaliteyi sağlamak için satıştan önce sayılamaz, ölçülemez, envanteri çıkarılamaz, test edilemez niteliktedir. (Parasuraman vd, 1985). Soyutluk nedeniyle, firma tüketicilerin hizmetlerini nasıl algıladıklarını ve hizmet kalitesini nasıl değerlendirdiklerini anlamakta zorlanabilmektedir (Zeithaml vd., 1981).

Hizmet kalitesi bütün olarak düşünüldüğünde; bir hizmetin mükemmelliği ve üstünlüğü olarak ifade edilebilir. Ancak hizmet kalitesi tüketicilerin algılamasına da bağlı bir durumdur. Kalitenin, tüketicilerin algısına göre ve hizmet sektörleri arasında da farklılık göstermesi nedeniyle hizmet kalitesi, işletmelerin tüketici beklentilerini karşılama veya aşma yetenekleri olarak kabul edilebilir. Bu bağlamda, işletmelerin ne sunduklarından ziyade tüketici beklentileri çerçevesinde neyi sunmaları gerektiği daha büyük önem arz etmektedir (Parasuraman vd, 1985). Ayrıca tüketicilerin hizmetleri satın aldıktan sonraki deneyimleri önemlidir. Bu durum, tüketiciler tarafından algılanan

hizmet kalitesinin ölçülmesindeki önemli bileşenlerden birisini oluşturmaktadır (Johnston & Bonoma, 1981).

Hizmet kalitesinin ölçülmesi kapsamında literatürde bununla ilgili farklı şekillerde modeller ve yaklaşımlar geliştirilmiştir. Hizmet kalitesinin ölçülmesinde oluşturulan yaklaşımlardan birisi, teknik ve fonksiyonel kalite modelidir. Bir tüketici hizmet sürecinde elde ettiği ve ürünleri teknik kalite olarak algıladığı belirtilmektedir. Örneğin bir otele gelen kişinin, bir oda ve yatak alabilmesi ya da restorana gelen birinin yemek alması sonucunda teknik kalite ortaya çıkmaktadır. Bu durumda tüketiciler almış olduğu şeyi direkt olarak ölçebilmektedir ve bu şekilde ölçümü belli olan bir durum, teknik kaliteyi göstermektedir. Ancak araştırmada teknik kalitenin tek başına memnuniyeti sağlamadığı belirtilmiştir. Tüketiciler, teknik kaliteden almış olduğu şeylerin ayrıca fonksiyonu ile de ilgilenmektedir. Örneğin bir restoran ya da otele ulaşım, oradaki kişilerin davranışları, tüketicilerin hizmete karşı bakışını etkilemektedir. Bu durum ise fonksiyonel kalite olarak adlandırılmaktadır. Tüketiciler, sadece ne aldığı ile değil, bu alım sürecinin nasıl meydana geldiğine de bakmaktadır. Bu kalite modeline ayrıca imaj konusu da etki etmektedir. İnsanlar, hizmet aldıkları yerleri ve markaları önemser. Bu yüzden kurumsal imaj, insanların beklentilerini etkileyen bir durumdur. Kurumsal imaj, tüketiciler firmaları nasıl algıladığının bir sonucu olmaktadır (Gronross, 1984).

Hizmet kalitesinin ölçülmesinde sıkça kullanılan modellerden birisi Servqual modelidir. Bu modelde, hizmet sağlayan işletmelerin, tüketici memnuniyetini ve sağlamış oldukları hizmetin kalitesini ölçmek ve anlamak için beş boyut geliştirilmiştir. Bu boyutlar güvenilirlik, güvence, somutluk, empati ve yanıt verme olarak belirtilmiştir. Güvenilirlik boyutunda hizmetlerin vaat edildiği ölçüde sunulması, güvencede ise hizmeti sağlayanların sahip olduğu bilgi ve becerinin tüketicilerde güven yaratması, cevap verme boyutunda hizmet sağlayanların tüketicilerin ihtiyaçlarına, problemlerine ya da taleplerine yanıt vermesi, somutluk boyutunda hizmetin nasıl görüldüğü, empati boyutunda ise hizmeti sağlayan kişi ya da örgütlerin tüketicileri anlayabilme kabiliyeti olarak ifade edilmektedir. Tüketicilerin ve işletmelerin bu boyutlar hakkındaki sorulara vermiş olduğu cevaplar karşılaştırılarak, meydana gelen farka göre hizmet kalitesinin ne derecede olduğu belirlenmektedir (Parasuraman vd., 1988).

Hizmet kalitesinin ölçümünde geliştirilen bir diğer model olan Servperf modeli ise, Servqual modelinin geliştirilmiş bir versiyonu olmakla birlikte, Servqual modelinin ölçtüğü beklenti ve fark yerine, tüketicilerin direkt olarak algılamış olduğu hizmet

kalitesini ölçmektedir. Bu özellikle daha çok tüketicilerin hizmetten algıladığı gerçek deneyimin ölçülmesi amaçlanmaktadır ve bu nedenle hizmet sektöründe sıklıkla kullanılan hizmet ölçüm modellerinden birisi olmuştur (Cronin vd., 1992).

İnternet kullanımının yaygınlaşması ile beraber, internet üzerindeki alışverişlerin artması, bu alanda yapılan hizmet kalitesi ölçeklerinin oluşmasını da beraberinde getirmiştir. Bunların arasındaki en önemlilerinden biri olan E-S-Qual modelidir. Bu model, sistem uygunluğu, etkinlik, güvenlik ve yerine getirme boyutlarından oluşan bir ölçektir ve internet üzerindeki hizmet kalitesinin ölçülmesini sağlamaktadır. Bilgi sistemlerinin kalitesinin ölçümü ile geliştirilen Webqual yönteminde ise, tüketicilerin bir internet sitesini değerlendirmede hangi unsurları dikkate aldığı belirlenmektedir (Loiacono, 2007). Bilgi kalitesi, sistem kalitesi ve hizmette etkileşim kalitesi adı altında toplamda 3 ana boyuttan oluşan bir ölçektir. Bilgi kalitesi boyutunda, web sitesindeki bulunan bilgilerin kalitesi, web sitesi içeriğinin kendisini güncellemesi, güvenilir bir içeriğe sahip olması gibi etmenler değerlendirilmektedir. Sistem kalitesi boyutunda ise web sitelerinin performansı dikkate alınmaktadır. Örneğin web sitesinin açılma süresi, tepki süresi, kullanıcı ve ziyaretçi arayüzü gibi etmenler değerlendirilir. Hizmette etkileşim boyutunda ise tüketiciler ile web sitesi arasındaki iletişimin boyutları değerlendirilmektedir (Parasuraman vd., 2005).

2.5 Konaklama İşletmelerinde Hizmet Kalitesi

Hizmet sektörünün önemli bir parçası olan turizm endüstrisini oluşturan önemli bileşenlerden olan konaklama işletmelerinin sağladığı hizmetin doğru bir şekilde ölçülmesi, doğal olarak tüketici memnuniyetine etki eden unsurların belirlenmesindeki önemli bir basamaktır. Turizm sisteminde yer alan konaklama işletmelerinin hizmet kalitesinin ölçülmesinde önemli modeller geliştirilmiştir. Bu modellerden birisi olan Holserv, otellerdeki hizmet kalitesinin ölçülmesinde kullanılan başlıca yöntemlerden birisini oluşturmaktadır. Servqual ölçeği baz alınarak yapılan bu çalışmada, servqual ölçeğinde bulunan boyutlara 7 tane yeni boyut eklenmiş ve diğer boyutlar otel sistemine göre uygulanmıştır (Mei vd., 1999).

HotelZot ise konaklama işletmelerindeki hizmet kalitesinin ölçümünü için geliştirilen diğer bir ölçektir ve Servqual modeli baz alınarak oluşturulmuştur. Temel amacı otel işletmelerindeki hizmet kalitesini, tüketicilerin hizmet beklentilerine ve algısına dayalı olarak ortaya koymaktır. Tolerans bölgesi adını verdikleri değerlendirmede, tüketicilerin almış olduğu hizmet beklentisi, yeterlilik ve arzulanan

şekilde algılanmasına dayalı olarak ölçmüştür. Yeterlilik ve arzulanan arasındaki bölgeye ise tolerans bölgesi adını vermişlerdir. Buna göre otel işletmelerinde konaklama yapan kişilerin almış olduğu hizmette yeterlilik ve arzulanan arasındaki fark, bir diğer anlamı ile tolerans bölgesi dar bir seviyede olduğunda, otellerde konaklama yapan kişilerin otelin kendilerine sağladığı hizmetlerde homojenlik aradığını göstermektedir. Bu aşamada arzulanan ve yeterli bulunulan hizmetler arasındaki fark az olduğunda, tüketiciler otellerden beklediği hizmet, benzer seviyede olmaktadır (Nadiri vd., 2005).

Geliştirilen diğer ölçeklerden birisi olan Lodgserv ölçeğinde, otellerde konaklayan kişilerin otelin kalitesine olan algılarının, tüketici memnuniyeti ile ilişkili bir düzeyde olduğu gözlemlenmiştir. Ayrıca bu durumun otel işletmelerinde algılanan hizmet kalitesini kapsamında tüketicilerin ağızdan ağıza iletişimini güçlendirdiği ortaya konmuştur (Getty vd., 1994).

Otel işletmelerindeki hizmet kalitesinin ölçülmesinde başta Servqual modeli baz alınarak yapılan çalışmalarda, otellerdeki hizmet kalitesini farklı açılardan ele alınmıştır (Nguyen vd., 2015). Tayland'da bulunan butik ve iş konseptli otelleri ziyaret eden turistlerle yapılan çalışmada, Servqual modeli baz alınarak anket geliştirilmiştir. Bu otellerdeki hizmet kalitesinin boyutlarını ve otel konseptlerine göre farklılıkların araştırıldığı ilgili çalışmada, hizmet kalitesinin düşük olduğu gözlemlenmiştir. Oteller, kendilerinden beklediği ölçüde hizmeti götüremediği belirtilmiş olup, butik otellerin iş konseptli otellere göre müşteri beklentilerini daha fazla yerine getirdiği görülmüştür (Boonit & Rompho, 2012).

Ürdün'de bulunan otellere yapılan çalışmada ise kullanılan Servqual modelinde, otellerde konaklama yapan kişilerin Servqual modelinin ele aldığı tüm hizmet boyutlarından daha iyi derecede hizmet beklentisinin olduğu görülmüştür. Otellerde konaklayan kişiler, en düşük olarak empati ve soyutluk konusunda hizmet kalitesi algıları meydana gelmektedir (Al-Khattab & Aldehayyat, 2011).

Sri Lanka'da beş yıldızlı otellerin hizmet kalitesinin ölçüldüğü çalışmada, otelde konaklama yapan müşterilerin çoğunluğunun otellerdeki hizmet kalitesinden olumlu düzeyde etkilendiği belirtilmiştir. Özellikle hizmet kalitesinin boyutlarından olan soyutluk, duyarlılık ve güvence konusunda konaklama yapan kişilerin memnuniyet düzeylerinin yüksek olduğu bildirilmiştir. Ancak empati ve güvenilirlik konularında ise memnuniyetin daha düşük seviyede olduğu görülmüştür (Karunaratne & Jayawardena, 2010).

Mauritius’da bulunan otellerde konaklama yapan kişilerin memnuniyetinin ölçüldüğü çalışmada ise, hizmet kalitesi boyutlarından en fazla olarak empati konusunda tüketicilerin sorun yaşadığı bildirilmiştir (Juwaheer & Ross, 2004).

Türkiye’deki otel işletmelerindeki hizmet kalitesi ile ilgili yapılan çalışmalarda, hizmet kalitesinin boyutları ve kalitenin etki etmiş olduğu farklı unsurlar ele alınmaktadır. Örneğin Adana’ da iki, üç, dört ve beş yıldızlı otellerde konaklama yapan kişiler üzerinden yapılan çalışmada, otelin hizmet kalitesi ve sunduğu imajı ile otelde konaklama yapan kişilerin tekrar satın alma davranışları arasında pozitif bir ilişki olduğu belirtilmiştir (Çabuk vd., 2013).

Giresun’ da 3 yıldızlı otellerin hizmet kalitesinin ölçüldüğü çalışmada, temizlik unsuru en fazla hizmet kalitesini etkileyen boyut olarak bildirilmiştir. Bu unsuru fiyat, personel davranışları, kat hizmetleri ve restoranlardaki hizmet durumu takip etmektedir. Otellerdeki servis durumu ve iş esnekliği kalitesi ise en az öneme sahip olan durum olarak ortaya çıkmıştır. Bunu oteldeki hizmet çeşitliliği, önbüro, güvenlik, konfor takip etmektedir (Korucuk vd., 2019).

Antalya’ da faaliyet gösteren beş yıldızlı otellerin hizmet kalitesinin ölçüldüğü çalışmada, tüketicilerin otellerin sunmuş olduğu hizmetten pozitif olarak en fazla etkilendiği konular müşteri ilişkileri, yiyecek ve içecek hizmetleri, otelin fiziki özellikleri, kat ve animasyon hizmetleri olmuştur. Ayrıca otellerin sağlamış olduğu hizmet kalitesi ile müşteri memnuniyeti arasında pozitif bir ilişki olduğu belirtilmiştir. Otelin sağlamış olduğu hizmetlerden en fazla müşteri ilişkileri, müşteri memnuniyetini etkilediği belirtilmiştir (Shengelbayeva, 2009).

Antalya’ da bir otelde konaklayan 390 müşteriye yapılan çalışmada, konaklama yapan kişilerin memnuniyetinin, otelin sunmuş olduğu hizmet kalitesinden etkilendiği olduğu belirtilmiştir. Müşteri memnuniyetini etki eden hizmet kalite boyutlarından en etkili olan ise güvenilirlik, en az etkili olan ise empati olmuştur (Torun Nalbant & Demiral, 2019).

Antalya’da beş yıldızlı otellerde konaklayan yabancı turistlerin hizmet kalitesi beklentisinin ve memnuniyetlerinin araştırıldığı çalışmada, turistlerin en fazla önem vermiş olduğu boyutun güvenilirlik boyutu olduğu belirtilmiştir. Ayrıca bu boyutun artması, hizmet kalitesinin diğer boyutlarını da arttıracak ve bu durumun genel memnuniyet düzeyinde artış göstereceği ortaya çıkarılmıştır (Sevgi, 2017).

2.6 Konaklama İşletmelerinde Çevrimiçi Hizmet Kriterleri

Web 2.0 platformlarının gelişimi ile tüketicilerin bu platformlarda belirtmiş olduğu yorumlarla otel işletmelerinin hizmet performanslarının ölçüldüğü çalışmalar hız kazanmaktadır. Otellerin hizmet performanslarının sosyal ağlar aracılığı ile ölçümü, içerisinde barındırdığı verinin büyüklüğü ve çeşitliliği kapsamında giderek önemini arttırmaktadır. Bu platformlarda bulunan Tripadvisor web sitesi, otel performansını en iyi şekilde gösteren web sitelerinden biri haline gelmiştir (Kim & Park, 2017). Bu platformlarda yazılan yorumların hakkında yapılan çalışmalar, tüketicilerin memnuniyetini hangi unsurların ne oranda belirlediği, hangi unsurların yorumlarda bir arada tutulduğu ve otellerin yapısına göre bu unsurların nasıl farklılaştığını ortaya çıkarmaktadır. Otellerin sunmuş olduğu hizmetlerdeki benzerlikler ve farklılıkların ne şekilde memnuniyeti etkilediği belirtilmektedir (Kim vd., 2016).

Otellere çevrimiçi platformlarda yapılan yorumlar göz önüne alınarak, konaklama yapan bireylerin memnuniyetine etki eden faktörleri belirlemek, otel işletmelerinin finansal performanslarının iyileşmesini, elektronik ağızdan ağıza iletişimi kullanarak daha fazla talep oluşturmasını sağlayabilmektedir (Sparks & Browning, 2011).

Otellere yapılan yorumlarda, konaklama yapan kişiler otel işletmelerinde kalış sürecini değerlendirirken, birçok farklı etmenden bahsetmektedir. Bu etmenler ise, hizmetlerin genel özellikleri göz önüne alındığında kişiden kişiye farklılık göstermektedir. Örneğin bir tüketici, oda, yiyecek içecek, rezervasyon, fiyat, personel davranışları gibi durumlardan memnuniyet duyarken, başka bir müşteri aynı otelde bu durumlardan rahatsızlık duyabilmektedir. Bu sebeple, çevrimiçi platformlarda ele alınan yorumlar bir bütün olarak ele alınmalı ve otellerin tüketicilere sunmuş olduğu hizmetler farklı açılardan araştırılmalıdır. Bu şekilde otellerin hizmet performansı etkili ve çok yönlü bir şekilde ortaya konulabilmektedir (Chittiprolu vd., 2021).

Tablo 2’ de otel işletmelerine çevrimiçi platformlarda yapılan yorumların analiz edildiği çalışmalardan bazıları ve bu çalışmalarda tüketicilerin memnuniyetini etkileyen hizmet kalitesi boyutları gösterilmektedir. Otel işletmelerindeki hizmet kalitesini etkileyen boyutlar, giderek çeşitlendiği göz önüne alındığında, araştırmaların yapıldığı ülkelere ve veri kaynaklarına göre bazı farklılıklar görünmesine rağmen, temelde benzer unsurların araştırmalarda ele alındığı görülmektedir. Bunlardan ilki otellerin fiziksel olarak durumu ve tüketicilere sağlamış olduğu olanaklardır. Bu durum, hizmet

kalitesinin ölçülmesinde önemli bir etmen olarak çalışmalarda belirtilmektedir. Ayrıca personel konusu da farklı açılardan ele alınan temel bir unsur olarak göze çarpmaktadır. Bunların yanında otelin bulunduğu konum, odaların durumu, temizlik, yiyecek ve içecek hizmetleri, değer ve fiyat konusu, rezervasyon ve ön büro konuları, otelin genel olarak sağlamış olduğu hizmetler de ilgili araştırmalarda ele alınan hizmet boyutlarını oluşturmaktadır (Paladini & Peres, 2021).

Tablo 2.2 Çevrimiçi platformlarda kullanılan otel hizmet kriterleri

Yazar ve Yıl	Veri Kaynağı	Yer	Otel Kriterleri
Xu & Li, 2016	Booking	ABD	Konum, oda, restoran, tesis, otopark, banyo, Wi-Fi, personel, sigara içilen ve oda içinde bulunan kirli hava, koridorlardaki otomatlar, yüzme havuzu, klima ve değer.
Kim & diğ., 2016	TripAdvisor	ABD	Oda, personel, otel özelliği ve görünümü, tesis ve diğer unsurlar.
Tontini & diğ., 2017	TripAdvisor	Brezilya	Konum, çalışanların tutumları, yiyecek ve içeceklerin kalitesi ve çeşitliliği, konfor, temizlik, tesisler, erişim, otel değeri, wi-fi, check-in ve check-out zamanı, oda, gürültü, yönetimin bilgisi, görünüm, ek olanaklar.
Sun & diğ., 2017	DaoDao Ctrip	Kıtalar-Arası	Hizmet, odalar ve tesisler, konum, temizlik ve değer.
Philips & diğ., 2017	Trustyou	İsviçre	Olanaklar, bina, ambiyans, odalar ve internet.
Liu & diğ., 2017	Tripadvisor	Çin	Odalar, değer, konum, temizlik ve hizmet.
Herjanto & diğ., 2017	Tripadvisor	Panama	Otel tesislerinin fiziksel özellikleri, yatak odası, otel, personel, restoranlar, konum ve fiyat faktörleri.
Fernandes & Fernandes, 2017	TripAdvisor	Portekiz	Konum, uyku kalitesi, odalar, müşteri hizmetleri, değer, temizlik, banyolar, kahvaltı, bar ve restoran, hizmet, otopark ve tesisler.
Xu, 2018	Booking	ABD	Oda, personel, konum, gürültü, wi-fi, bagaj taşıma, tesis, yatak ve banyo.
Köseoğlu & diğ., 2019	Booking	Hong Kong	Konum, oda, personel, manzara, hizmet, temizlik, kahvaltı, havuz, restoran, yemek ve diğ.
Hu & diğ., 2019	Tripadvisor	ABD	Tesis, hizmet, lokasyon, değer ve genel deneyim.
Brochado & Brochado, 2019)	Tripadvisor	Portekiz	Deneyim, otel, öğrenme, misafirlik, kamp, doğa, yiyecek, farklılıklar, eko ve yoga.
Alrawadie & Law, 2019	Tripadvisor	Türkiye	Odalar, hizmet kalitesi, otelin karakteristik özellikleri, yiyecek ve içecek.
Ying & diğ., 2020	Ctrip Tripadvisor	Çin	Oda, seyahat, yiyecek, çevre ve otel olanakları.
Sann & diğ., 2020	Tripadvisor	İngiltere	Hizmet, temizlik, oda, uyku kalitesi, lokasyon ve değer.
Rios-Martin & diğ., 2020	Tripadvisor	İspanya	Personel, servis, oda ve lokasyon
Sann & Lai, 2020	Tripadvisor	İngiltere	Rezervasyon, ön büro, kat hizmetleri, yiyecek kalitesi, güvenlik, mühendislik, pazarlama ve enformatik.
Padma & Ahn, 2020	TripAdvisor	Malezya	Otel, oda, personel, seyahat ve otel ile ilgili diğer kriterler.
Lee & diğ., 2020	Tripadvisor	ABD	Duyusal deneyim, marka, otel sınıfı, uyku,

			konum, oda, hizmet, değer ve temizlik.
Foris & diğ, 2020	Booking, Tripadvisor, Directbooking Travelminit	Romanya	Lokasyon, ulaşım, turlar, konaklama ve kısıtlar.
Doğan & diğ, 2020	Holidaycheck	Türkiye	Lokasyon, personel, yiyecek ve içecek, olanaklar ve oda.
Sann & Lai, 2020	Tripadvisor	Tayland	Personel, hizmet, oda, lokasyon, yatak, duş, uyku, kahvaltı, temizlik, olanaklar, ek hizmetler.

Kaynak: Paladini & Peres, 2021.

2.7 Konaklama Sektöründe Kullanıcı Tarafından Oluşturulan İçeriklerin Önemi

Dijital platformların teknolojinin gelişimi ile yaygın olarak kullanılmaya başlanması, insanların ürün ve hizmetler ile ilgili duygu ve düşüncelerini paylaşmasını arttırmıştır. Hizmet sektörünün önemli bir parçası olan turizm, içerisinde barındırdığı çevrimiçi platformlardaki kullanıcı deneyimleri açısından giderek zengin bir konuma gelmektedir. Bu aşamada turistler, gidecekleri destinasyonlar ve destinasyonda konaklama yapacakları yerler ile ilgili daha fazla araştırma yapmakta, daha önce konaklayan kişilerin deneyimlerine önem vermektedir (Schukert vd., 2015; Marco vd., 2018).

Konaklama işletmelerinin sunmuş olduğu hizmet kalitesi ile ilgili yapılan araştırmalar, özellikle son yirmi yılda internet kullanımının yaygınlaşması ile kullanıcı tarafından oluşturulan içerikler (User Generated Content) kullanılarak yapılan çalışmalar hız kazanmıştır (Tontini vd., 2017; Raguseo vd., 2017). Gelişen Web 2.0 teknolojilerinin kullanıcılara vermiş olduğu imkanlar, tüketicilerin bu alanda görüşlerini bildirmelerini sağlamak ve bu durum insanların tüketim davranışlarını etkilemektedir. Bu imkanlar sayesinde tüketiciler satın alma kararı verme sürecinde alternatif seçenekleri değerlendirmektedir ve seçimlerine yardımcı olacak geniş veri havuzlarına erişebilmektedir. (O’connor, 2010). Web 2.0’ da kullanıcılar yalnızca içerik üreticileri olmakla kalmayıp, aynı zamanda diğer kullanıcılarla iş birliği yapabilecek konuma gelmiştir. Bu sayede kullanıcılar aktif rol almakta ve hem tüketiciler hem de üreticiler için etkileşim yaşanabileceği büyük sanal topluluklar meydana gelmektedir (Valdivia vd., 2017).

Web 2.0 kapsamında ele alınan ürünlerden birisini ise sosyal paylaşım siteleri oluşturmaktadır. Bu siteler ile insanlar düşüncelerini ve aktivitelerini paylaşmaktadırlar ve insanlara farklı etkileşim olanakları tanıyarak, insanların birbirleri ile iletişim kurmasını ve fikirlerini paylaşım yapmasına olanak sağlamaktadır (Cooke & Buckley,

2008). Web 2.0' daki web siteleri statik değil, dinamik bir tasarıma sahip olarak kullanıcıların içerik üretme ve başından geçen deneyimlerin paylaşmasına olanak sağlar. Bu alandaki en önemli platformlardan birisini ise turizm ve seyahat platformları oluşturmaktadır. Yapılan bazı çalışmalarda web 2.0 çalışmaları, seyahat ve turizm endüstrisindeki kullanımı seyahat 2.0 (travel 2.0) olarak ele almakla beraber, seyahat 2.0 kapsamında ele alınan web siteleri seyahat 2.0 kapsamında değerlendirilen bir konsepttir (Miguens vd., 2008; Monoz vd., 2012).

Web 2.0 kapsamında ele alınan platformlardan en önemli gruplardan birisini ise, kullanıcı tarafından oluşturulan seyahat web siteleri oluşturmaktadır. Bu web siteleri, içerisinde bulundurduğu sosyal ağlar sayesinde, kullanıcıların sadece web sitesindeki ürünlerle değil, birbirleri ile de etkileşime girmesine ve ürünler hakkında incelemeler yapmasına olanak sağlanmaktadır (Amaral vd., 2014). Bu tarzda etkileşime olanak sağlayan ve kullanıcıların bu sitelerdeki turizm işletmeleri hakkında düşüncelerini paylaşmasına izin veren sitelerin başında ise Tripadvisor gelmektedir. Tripadvisor, turizm alanında web 2.0 teknolojisi baz alınarak tasarlanan ve dünyanın turizm alanında en fazla ziyaret edilen ve en popüler olan kullanıcı tarafından oluşturulan seyahat web sitesi ünvanını taşımaktadır. (Valdavia vd., 2017). Ayrıca seyahat alanındaki tüketicilerin etkileşime geçtiği dünyanın en geniş çevrimiçi ağına sahiptir (O'connor, 2010). Özellikle konaklama sektörü alanında kullanıcıları konaklama yaptığı tesislerle ilgili deneyimlerini paylaşmaya teşvik ederek turizm sektöründe sosyal etkileşimi arttırmaktadır. Tüketiciler bu gibi web sitelerini kullanarak deneyim geçirmiş oldukları konaklama işletmeleri ile ilgili görüşlerini bildirmekte ve bu işletmeleri, web sayfalarının tüketicilere sağlamış olduğu puanlama sistemini kullanarak değerlendirebilmektedir (Munar & Ooi, 2013). Tripadvisor, konaklama sektörü alanında da dünyanın en büyük seyahat sitesi olup, bünyesinde 1 milyarın üzerinde yorum barındıran ve bunların 700 milyonu otel yorumu olan, dünya çapında 500 milyona yaklaşan aylık kullanıcı sayısı, 8 milyona yaklaşan işletme sayısı ile 43 pazarda ve 22 dilde hizmet vermektedir (Tripadvisor.com, 2022).

Bu ölçekteki web sitelerinin konaklama endüstrisindeki önemi, tüketicilerin internet kullanımının ve web 2.0 teknolojilerin artması ile, tüketicilerin karar verme aşamasına etkileri artmaktadır (Tuominen, 2011). Özellikle seyahat 2.0 teknolojilerini kullanarak bu alanda faaliyette bulunan konaklama işletmelerine yapılan kullanıcı yorumları, bireylerin satın alma süreçlerini etkileyen bir unsurdur (Kantarıcı vd., 2017). Tripadvisor'ın yapmış olduğu araştırmada, tüketicilerin %80'inden fazlası web

sitelerinde bulunan yorumları seyahate çıkmadan önce okuduğu ve bu kişilerin %93'ünün otel konaklamasının seyahat deneyimi için en önemli aşama olduğu belirtilmektedir. Ayrıca tüketicilerin %77'si bir otel seçmeden önce otel yorumlarını okuduğu belirtilmektedir. Buna ek olarak, tüketicilerin %52'si ise yorum bulunmayan otellere yorum yapmamakta olduğu açıklanmıştır (Tripadvisor.com, 2019). Bu sebeple seyahat 2.0 platformlarına yapılan yorumların içeriklerini analiz etmek, konaklama işletmeleri için önemli bir durum haline gelmiştir. Bu yorumların içermiş olduğu ve konaklama işletmelerinin farklı özelliklerini değerlendirmeye imkan sağlayan olanaklar ve sayısal ölçekler, bu işletmelerde konaklama yapacak kişilerin kararlarının etkileneşinde ve otellerdeki hizmet kalitesinin belirlenmesinde önemli basamaklardan birisi olmaktadır (Barriocanal vd., 2010; Fong vd., 2017).

Kullanıcı tarafından oluşturulan seyahat web sitelerinde faaliyet gösteren konaklama işletmeleri, bu sitelerin kendilerine verdiği imkanlar ile tüketiciler bu işletmeleri farklı özelliklere göre değerlendirebilmektedir. Örneğin Tripadvisor' da tüketiciler, kendi yaşamış olduğu konaklama deneyimlerini yorumlarında ifade edebilmekte ve bu yorumların sahip olduğu sayısal ölçekler ile konaklama işletmelerini değerlendirme imkanları bulabilmektedir. Örneğin ilgili sitede tüketiciler otelin temizliğini, fiyat performansını gösteren değeri, almış olduğu hizmeti ve otelin bulunduğu lokasyonu 1'den 5'e kadar sayısal bir ölçekte değerlendirmektedir. Ayrıca otele yine 1'den 5' e kadar genel bir değerlendirme puanı verebilmektedir. Booking.com sitesinde ise bu ölçekler çalışanlar, olanaklar, temizlik, konum, fayda ve fiyat dengesi, rahatlık ve ücretsiz Wifi hizmeti olarak meydana gelmektedir ve bu ölçekler ise 1'den 10' a kadar tüketicilere sayısal olarak oteli değerlendirme imkanı vermektedir (booking.com, 2023). Benzer durum Agoda web sitesi için de geçerli olmakla beraber, bu sitede ise temizlik, konum, tesis içerikleri, hizmet ve fiyat/performans maddeleri ile misafirler konaklama işletmelerinin performansını farklı konaklama hizmetini gösteren bileşenler ile değerlendirmektedir (agoda.com, 2023).

Bu şekilde oluşturulan web sayfalarındaki kullanıcıların belirttiğı puanlar ele alınarak konaklama işletmelerinin hizmet performanslarının ölçülmesi yapılabilmektedir. Söz konusu puanların hizmet performansına etkisinin neler olduğu ve ne şekilde memnuniyeti etkilediğı, yapılan çalışmalarda farklı şekillerde ele alınmaktadır. Bu unsurların ortaya çıkarılması ile otellerin yapmış olduğu pazarlama faaliyetleri, yürütmüş olduğu tüketici ilişkileri gibi konularda otel işletmelerine öneriler getirilmektedir (Peres, 2021; Nicoli & Papadopolou, 2017; Xie vd., 2017). Oluşturulan

sayısal değerler ile yapılan araştırmalarda, otellerin sağlamış olduğu hizmet çeşitliliğinden ve bu çeşitliliğin göz önüne alınarak tüketiciler tarafından otel hizmetinin değerlendirildiği göz önüne alındığında, otellerdeki hizmet kalitesini tek bir unsurun değil, farklı unsurların etkilediği görülmektedir. Örneğin otellerin fiyat-performansına tek bir değer değil, bu değerlerin iki veya daha fazla kombinasyonları etki etmektedir ve değerlendirme yapan kişiler, bu değerleri bir bütün olarak düşünerek bu değerler otellerin hizmet kalitesini birlikte etkilemektedir (Başaran vd., 2020; Tontini vd., 2017; Ou vd., 2018).

2.8 Kullanıcı Tarafından Oluşturulan İçerikler ve Metin Madenciliği

Sayısal değerlerin konaklama hizmetlerinin performansının belirlenmesinin yanında, metin madenciliği alanında yapılan yenilikler, kullanıcıların konaklama işletmelerine seyahat 2.0 olanakları kapsamında yapmış olduğu yazılı yorumların sayısal hale getirilerek anlaşılmasını ve bu alanda yapılan analizlerin etkinliğini arttırmaktadır. (Payak vd., 2020; Kar vd., 2021) Seyahat 2.0 web sitelerinin insanların konaklama işletmelerini değerlendirmeleri için sağladığı sayısal değerlendirme yöntemleri, tüketicilerin yazmış olduğu yorumların içeriklerinin çeşitliliğine bakılarak bazen yetersiz kalabilmektedir ve bu sebeple metinler de analizlere dahil edilerek etkili sonuçlar ortaya çıkarmaktadır (Ribarsy vd., 2014). Konaklama deneyimi yaşayan birisi, otelin ve otelin bulunduğu destinasyon ile ilgili görüşlerini de bu web sitelerinde çeşitli şekillerde bildirmektedir. Ayrıca tüketicilerin bu alandaki web sitelerine yazmış olduğu yorumlarda, her bireyin otele karşı algıladığı hizmet kalitesi farklılık göstermektedir ve bu analizleri ayrıca manuel olarak analiz etmek ya da anket yoluyla elde etmek, hem zaman hem de finansal olarak maliyetli olabilmektedir (Chang vd., 2019). Örneğin bir birey için personel ile ilgili kavramlar önemli iken, bir diğeri için ise otelin sağladığı imkanlar, otele vermiş olduğu genel puanı etkileyebilmektedir. Bu sebeple metinlerin çok yönlü bir şekilde analiz edilmesi ve bu metinlerde sayısal olarak ifade edilemeyecek unsurların belirlenmesi, otellerin hizmet kalitesinin ölçülmesinde ve otellerin hangi özelliklerinin tüketicileri ne oranda etkilediğinin ortaya çıkartılması önemli bir durum olmaktadır. Bu duruma ek olarak, konaklama sektörünün faaliyet gösterdiği destinasyonlardaki imajın iyileştirilmesi bakımından da önem taşımaktadır (Chang vd., 2019; Herrero vd., 2015). Kullanıcıların yaşamış oldukları deneyimler sonrasında oluşturmuş oldukları çevrimiçi içerikler, artık bir satış asistanı olma yoluna gitmekte ve

özellikle turizm sektöründe çevrimiçi yorumların tüketiciler tarafından algılanan faydasını etkileyen faktörlerden biri olmaktadır (Ali vd., 2021).

Tüketicilerin sayısal ifadeler ile yazmış olduğu yorumlarda birçok farklı konunun ve bu konuların birçok alt konusu olduğu göz önüne alındığında, doğal dil işleme ve metin madenciliği alanında yapılan gelişmeler ve bu alanda kullanılan yöntemler, tüketicilerin ilgili platformlarda yazmış olduğu yazılı metinlerin ayrıntılı şekilde analiz edilmesini sağlamaktadır (Bahja, 2020). Bu şekilde sayısal olarak ifade edilemeyen ya da hizmet kalitesini gösteren bileşenlerin yetersiz olduğu alanlarda, yazılı ifadelerin içerikleri analiz edilerek, daha spesifik konulara ulaşılabilmekte ve bu konuları oluşturan etmenlerin ortaya çıkarılmasını sağlanabilmektedir (Xu & Lv, 2022).

Konaklama işletmelerine web 2.0 platformları üzerinden yapılan yorumlarının metin madenciliği yöntemleri ile analiz edilme yaklaşımlarından birisi ise özellikle son yıllarda kullanımı artan makine öğrenmesi ve derin öğrenme yaklaşımlarıdır. Bu yöntemler ile metinlerde gizli kalan düşünceler hesaplamalı dil bilim araçları ile ortaya daha net bir şekilde çıkabilmekte, metinlerin sınıflandırılması sağlanmakta, metinlerin kümelenmesi ve bu metinlerden ayrıntılı şekilde bilgi çıkarımı yapılması sağlanabilmektedir. Bu aşamada kullanıcıların çevrimiçi platformlarda paylaşmış olduğu içeriklerin ortaya çıkartılmasında duygu analizleri, konu sınıflandırmaları, konu modelleri gibi metin madenciliği ve doğal dil işleme kapsamında ele alınan yaklaşımlardan faydalanılmaktadır. (Nilashi vd., 2021; Bagherzadeh vd., 2021; Ren vd. 2021, Neto vd., 2020).

2.9 Duygu Analizi (Sentiment Analysis) Kavramı

Metin madenciliğinde sıklıkla uygulanan ve bu kapsamda konaklama işletmelerine yapılan çevrimiçi yorumların analizinde sıklıkla kullanılan yöntemlerden birisini duygu analizi oluşturmaktadır. Duygu analizi, bir metnin sahip olduğu duyguyu ifade eden ve bu duygunun hangi yönde olduğunu sınıflandırmaya yarayan bir yaklaşımdır. Burada metinler, pozitif, negatif ya da nötr olarak metnin alacağı duyguya göre belirli bir sınıf elde eder (Seker, 2016). Özellikle internet kullanımının artması ile tüketicilerin ürünler ve hizmetler ile ilgili olarak çevrimiçi mecralarda görüşlerini belirtmesi giderek artmaktadır. Bu görüşlerin hangi ölçüde olumlu ya da olumsuz olduğu, özellikle işletmeler tarafından yapacakları pazarlama ve yönetim stratejileri bakımından önemlidir. Tüketicilerin satın aldıkları ürün ve hizmetlere karşı almış olduğu tavrı ve bu tavrı sergileyiş biçimi, bu noktada işletmelere önemli yararlar

sağlayabilmektedir (Kaynar vd., 2016). Bunun sonucu olarak finans, politika, telekomünikasyon, yazılım mühendisliği, sosyal medya gibi alanlarda duygu analizi çalışmaları yapılmaktadır. (Najadat vd., 2018; Neri vd., 2012; Mishev vd., 2020; Obaidi vd., 2022).

Doğal dil işleme ve metin madenciliği alanında meydana gelen teknik gelişmeler, duygu analizi yöntemlerinin giderek doğru ve etkin sonuçlar vermesini sağlamış ve bu alanda veri tiplerine göre farklı yöntemler geliştirilerek, araştırmacıların bu alanda etkin çalışmalar yapmasına ve güvenilir sonuçların elde edilmesine olanak sağlamaktadır. Turizm sektörü, içerisinde birçok bileşeni olan bir sistemler bütünüdür ve bu bileşenlerinden önemli bir paya sahip olan konaklama işletmeleri alanında da duygu analizi çalışmaları artış göstermektedir (Alaei vd., 2019). Özellikle çevrimiçi platformlarının kullanımının artması ve bu alanda faaliyet gösteren konaklama işletmelerine tüketiciler tarafından yapılan yorumların analiz edilmesi hem işletmeler hem de işletmelerin bulunduğu destinasyonların imajı için önem taşımaktadır (Markopoulos vd., 2015; Gonzales vd., 2016; Jiang vd., 2021; Dohaia vd., 2018).

Duygu analizi ile yapılan çalışmalarda makine öğrenmesi metotları kullanılarak, metinlerin ya da metin setlerinin sahip olduğu duygular ortaya çıkarılabilmekte, bunlar sayısal olarak hesaplanabilmekte ve sınıflanabilmektedir (Jagdale & diğ., 2019). Böylece sınıflandırılan ve duygu yoğunluğu ve türü ortaya çıkarılan metinler, araştırmacılar tarafından çalışmalarda farklı şekillerde ele alınarak nicel ve nitel araştırmalar yapılabilmektedir (Dickinson & Hu, 2015).

Duygu analizi yapılmasının gerekçelerinden birisini de yorumların değerlendirilmesindeki sübjektifliktir. Örneğin bir tüketici, oteldeki kalış deneyimini değerlendirme sürecinde, değerlendirme kriterleri farklılık göstermektedir ve bu durum doğal olarak kişilerin puanlarına da yansımaktadır. Örneğin çevrimiçi konaklama platformlarının kullanıcılara getirmiş olduğu seçeneklerde, kişiler kendi deneyimlerini belirli bir ölçek üzerinden değerlendirme seçeneğine hizmet performansını gösteren çeşitli göstergeler aracılığı ile sahip olabilmektedir. Bir tüketici, yaşamış olduğu konaklama deneyiminden bahsederken ve otelden almış olduğu hizmeti ilgili çevrimiçi platformlarda puanlarken otel işletmesinde hangi hizmet ya da ürün yapısı kendisine göre daha önemli ise bu durum bu tüketicinin puanına ve dolayısı ile bu kişinin deneyimi bahsettiği metindeki duygu durumunu etkilemektedir (Mahadevan & Arock, 2020). Örneğin bir tüketici, oda ile ilgili kötü bir deneyim yaşamış olup, bu durumda genel değerlendirmesi 5 üzerinden 2 olabilirken, başka birisi aynı deneyimi yaşayarak 3

puan verebilmektedir. Bu sebeple tüketicilerin çevrimiçi platformlarda dile getirmiş oldukları görüşleri bir bütün olarak ele almak gerekmektedir ve tüketicilerin bu mecralardaki görüşlerinin içerikleri kapsamlı bir şekilde ele alınması önemlidir (Baccianella vd., 2009). Duygu analizi kapsamında ele alınan metotlar ile, metinler içerisindeki duygular objektif bir hale getirilmeye çalışılmaktadır.(Hogenboom vd., 2014). Bu şekilde yorumlar ile ilgili yapılan duygu analizi çalışmalarında, metinlerin içerikleri belirli bir sistem üzerinden değerlendirilmektedir (Palanisamy vd., 2013). Bu şekilde subjektifliğin önüne belirli oranda geçilmesi amaçlanmaktadır.

Metin madenciliği alanında yapılan duygu analizi çalışmaları üç farklı çeşitte yapılmaktadır. Bunlar makine öğrenmesi, kütüphane bazlı hesaplamalar (lexicon based) ve hibrit yöntemlerdir (Sham & Mohamed, 2022). Kütüphane bazlı hesaplamalarda her bir metin, metin içerisindeki kelimelerin, kütüphanedeki duygu durumuna göre almış olduğu puana göre belirli bir aralıkta puan almaktadır. Bir metindeki kelimeler ya da ifadeler, sözlüklerde negatif ya da pozitif olarak belirli bir skora ya da ağırlığa sahiptir. Bu kelime ve ifadelerin sözlüklerde almış oldukları puana göre metin setinde yer alan cümlelerin ya da belgelerin genel bir duygu puanı oluşturulmaktadır. Makine öğrenmesi modellerde ise etiketli olmayan veri setlerinde gözetimsiz, etiketli olan veri setlerinde ise gözetimli öğrenmeler gerçekleştirilerek duygu analizi cümlelere ya da dokümanlara yapılabilmektedir. Hibrit modelde ise genel olarak hem sözcük tabanlı hem de makine öğrenmesi algoritmalarının beraber kullanılması ile duygu analizi hesaplamaları ve sınıflandırmaları gerçekleştirilmektedir (Appel vd., 2016).

2.9.1 Konaklama işletmelerinde duygu analizi çalışmaları

Konaklama sektöründe duygu analizi ile ilgili yapılan çalışmalarda, farklı duygu analizi metotları kullanılmakta ve işletmelere yapılan yorumların duygu durumları, farklı açılardan ele alınarak hangi konuların tüketicilerde daha olumlu ya da olumsuz duygu yarattığı ile ilgili sonuçlar ortaya çıkarılmaktadır (Rana vd., 2016). Ayrıca konaklama işletmelerinin sahip olduğu çevrimiçi yorumları kapsamında yapılması planlanan çalışmalara yardımcı olacak şekilde hem sözlük hem de makine öğrenmesi metotları ile bu alana özgü sözlük bazlı duygu analizi çalışmaları ve bu çalışmaların makine öğrenmesi yöntemleri ile test edildiği çalışmalar da meydana gelmektedir (Bagherzade vd., 2021).

Konaklama işletmelerine yapılan çevrimiçi kullanıcı yorumlarının incelenerek duygu analizinin yapıldığı çalışmalarda ilgili işletmelerin farklı yönleri ele alınarak

analizler gerçekleştirilmektedir. Yorum yapan kişilerin geldiği destinasyonlar veya milliyeti, seyahat türleri, konaklama sınıfları ve konaklama türleri, geliş amaçları, yorumlarda vermiş oldukları otellerin hizmet performansını gösteren konulara yapmış olduğu yorumların duygu sınıflamaları, yorumların dilleri gibi seçeneklerle farklı değişkenlerin kombine edilerek duygu analizi çalışmaları gerçekleştirilmektedir. Böylece konaklama işletmelerinin pazarlama, hizmet kalitesini geliştirme gibi yapmış olduğu faaliyetlere öneriler getirmektedir (Sanz vd., 2022; Jiang vd., 2021; Matruty vd., 2023; Duan vd., 2016; Garcia vd., 2016).

Türkiye’ deki konaklama işletmelerine yapılan duygu analizi çalışmalarının sayısı az miktarda olsa da, bu alanda meydana gelen metodolojik gelişmeler ile yapılan çalışmaların sayısı son yıllarda artış göstermektedir. Antalya’da 9 ekolojik tarzda faaliyet gösteren otelin Tripadvisor sitesinden elde edilen verileri kapsamında yapılan duygu analizi çalışmasında Türkçe yorumlar kullanılmış olup, Türkçe dil tabanlı oluşturulan Zemberek kütüphanesi kullanılmıştır. Yorumlar çerçevesinde 6 adet otel kriteri belirlenmiş olup, bu kriterlerden en fazla pozitif duyguya sahip olan kriterin çevre bakımı, en fazla negatif yoruma sahip olan kriterin ise odalar kriteri olduğu ortaya çıkarılmıştır (Sezgin & Dağdelen, 2023).

Helal turizm konseptine sahip ve Antalya’da bulunan 7 adet 5 yıldızlı otelin yorumlarının duygu analizi ile yapıldığı çalışmada otellerde konaklama yapan kişilerin duygularını negatif olarak etkileyen özellikler ortaya konmuştur. Buna göre insanları en fazla rahatsız eden unsurlar otelin sessizliği ve personel ile ilgili durumlar olarak belirtilmiştir. Sessizlik konusunda insanları en çok rahatsız eden konuların yüksek ses ve müzik olduğu, personel ile ilgili en fazla rahatsız olunan konunun ise personelin iletişim ve İngilizce dil seviyesi eksikliği olarak bulunmuştur (Özen, 2019).

Duygu analizi çalışmalarında ayrıca makine öğrenmesi metotları da test edilmektedir. Örneğin Eskişehir’de ve İstanbul’da bulunan otellerin yorumlarına yapılan duygu analizinde ise makine öğrenmesi metotlarının sınıflama başarıları test edilmiştir. Araştırmadaki otellerin yorumları lojistik regresyon, K-en yakın komşu, naive bayes ve destek vektör makineleri ile sınıflandırma yapılmış olup, en iyi sınıflama başarıları 0.92 oranında lojistik regresyon vermiştir (İnan, 2019).

Booking.com ve tripadvisor.com sitesinden elde edilen toplamda 331 olumlu ve olumsuz yorum, farklı makine öğrenmesi teknikleri ile sınıflandırılarak başarıları ölçülmüştür. Çalışmada doküman terim matrisinin TF*IDF matrisine göre daha iyi

performans gösterdiği ve çalışmada kullanılan sınıflandırma metotlarından en iyisini rastgele orman yöntemi olduğu belirlenmiştir (Oğul & Ercan, 2016).

Pandemi sürecinde Twitter’da #turizm ile etiketlenen 4202 türkçe yorumun Zemberek kütüphanesi kullanılarak metinler etiketlenmiş ve bu etiketli veriler üzerine makine öğrenmesi modellerinin sınıflama performansı ölçülmüştür. Tf*Idf metodu kullanılan veri setinde, en iyi doğruluk oranını lojistik regresyon yöntemi vermiştir (Karaman & Çılan, 2020).

Tripadvisor web sitesindeki Türkçe ve İngilizce yorumlara yapılan sözlük tabanlı ve makine öğrenmesi modeli duygu analizi sınıflandırılması, sözlük tabanlı sınıflamanın İngilizce yorumlarda Türkçeye göre daha iyi çalıştığı belirtilmiştir. Makine öğrenmesi modellerinin ise sözlük tabanlı olan sınıflamada her iki dilde de daha iyi sonuç verdiği çalışmada ortaya çıkarılmıştır (Polat & Ağca, 2022).

2.10 Konu Modelleme (Topic Modeling) Kavramı

Konaklama işletmelerine yapılan yorumların konularının ortaya çıkartılması, metin madenciliği ve bu alanda kullanılan makine öğrenmesi yöntemleri ile daha hızlı ve yüksek miktarda veriyi kullanarak farklı temalar belirlenebilmektedir. Örneğin turistik destinasyonlardaki konaklama işletmelerine yapılan yorumların incelenme aşamasında, eğer konaklama işletmelerinin sayısı bu destinasyonda yüksek miktarda ise, bu yüksek miktarda olan yorumların teker teker incelenmesi ve tek tek konularının ve yorumların sınıflarının belirlenmesi, hem zaman hem de finansal açıdan büyük kayıplara neden olabilmektedir. Bu sebeple makine öğrenmesi kapsamında oluşturulan yöntemler ve bu yöntemlerde kullanılan farklı algoritmalar ile büyük verilerin daha hızlı ve etkin bir şekilde incelenmesi ve metinlerdeki konuların ortaya çıkartılmasındaki subjektifliğin belirli oranda önüne geçilmesi, araştırmacılara metin madenciliği alanında kolaylık sağlamaktadır (Ignatov & Mihalcea, 2018).

Metinlere yapılan yorumların ortaya çıkarılmasında kullanılan metin madenciliği tekniklerinden birisini ise konu modelleme oluşturmaktadır. Konu modelleme, metin seti içerisinde gizli kalan konuları ya da temaları ortaya çıkarmak için kullanılan bir yöntemdir. Elde edilen metin setleri, birçok dokumandan oluşmaktadır ve bu dokümanlar, içermiş olduğu konulara göre kategorize edilebilmekte ve bu şekilde insanlar bu konular hakkında bilgi sahibi olabilmektedir. Kümeleme yöntemleri, dokümanların gruplandırılmasında etkili bir yöntem olmasına rağmen, her zaman araştırmada istenilen sonuçları göstermeyebilmektedir. Metin setlerinde bulunan

dokümanların kümelenmesinde genellikle her bir doküman tek bir kümede gösterilmektedir. Ancak metin setleri bilindiği gibi içerisinde birçok farklı konuyu barındırmaktadır. Bu sebeple metinlerin içerisindeki dokümanlarda geçen kelimelerin oluşturduğu kümelerin ortaya çıkartılması ve her dokümanın birkaç konuya atıfta bulunması daha fazla istenilen bir durum olacaktır. Metin setinden çıkartılan konular, metin setindeki doküman koleksiyonundan otomatik olarak çıkarılmaktadır. Bu durumun ortaya çıkarılmasında kullanılan makine öğrenimi alanına konu modelleme denmektedir (Kumar & Paul, 2016).

Konu modellemede kullanılan başlıca makine öğrenmesi modelleri, gizli Dirichlet ayrımı, negatif olmayan matris ayrıştırması, gizli semantik indeksleme, hiyerarşik Dirichlet süreci, olasılıksal gizli anlam analizi, Top2vec, BERTopic gibi yöntemlerdir. Büyük miktardaki veriler bu yöntemler ile etkin bir şekilde incelenerek, tüketicilerin hangi konulara yönelik görüşlerini ifade ettiği, bu konuların kümelenmesi ve sınıflandırılması mümkün olabilmektedir (Albanese, 2022; Xu, 2018).

Anket, odak grup görüşmesi gibi geleneksel yöntemler, örneklem toplamak ve araştırmanın tahmin performansını ölçme aşamalarında daha zor olduğu için, günümüzde daha efektif yöntemlere ihtiyaç duyulmaktadır. Örneğin konaklama işletmelerinin memnuniyeti ile ilgili yapılan araştırmalarda, klasik yöntemler, bulundurduğu soru bakımından belirli limitlere sahip olmakta ve bu araştırmalar genellikle spesifik bir periyotta yapılabilmektedir. Bu durum, doğal olarak araştırmada belirli kısıtlara sebep olabilmektedir (Guo vd., 2017). Konu modelleme metotları ile, metin içeriğindeki spesifik konular ortaya çıkartılabilmektedir. Bu yöntemler, özellikle büyük miktarda verilerin ve yapılandırılmamış metinlerin analizlerinde etkili ve spesifik anlamlar çıkartabilme kabiliyetine sahip olmaktadır. Bu sayede genellenebilir sonuçlar ortaya çıkabilmektedir (Storopoli, 2019).

2.10.1 Konaklama işletmelerinde konu modelleme çalışmaları

Çevrimiçi kullanıcı içeriklerinin sıkça araştırıldığı konaklama sektörü alanında da son yıllarda konu modelleme yöntemini ile yapılan analizlerin sayısı artmaktadır. Özellikle bu alanda tüketicilerin yapmış olduğu çevrimiçi görüşler, klasik araştırma yöntemlerinden daha zengin kaynaklar içermekte ve bu kişiler geçirmiş oldukları deneyimleri daha spesifik olarak anlatma imkanına sahip olmaktadır. (Gretzel & Yoo, 2008). Konaklama işletmelerinin en büyük ve önemli kısmını oluşturan otellere sosyal paylaşım siteleri üzerinden yapılan yorumların içerikleri ve bu içeriklerin hangi

konulara değinildiği, konu modelleme kapsamında farklı makine öğrenmesi modelleri ile ortaya çıkartılabilmektedir. Ayrıca turizm sektöründeki çevrimiçi kullanıcı yorumlarının etkin bir şekilde incelenmesi, nitel ve nicel verilerin bir arada kullanılma gerekliliğini arttıran bir unsurdur. Bu sebeple, LDA gibi konu modellemesi yöntemlerin etkili bir şekilde kullanımı, zaman ve iş gücü kaybının belirli ölçüde önüne geçilmesine olanak sağlamaktadır (Jia, 2019).

Turizm sektörünün farklı bileşenleri hakkında da konu modelleme kullanılmaktadır. Atraksiyonlar, yiyecek içecek, destinasyon gibi turizmin başlıca alanlarında konu modellemesi çalışmaları yapılarak, yorumların içerikleri analiz edilmekte ve hangi konuların insanlar tarafından daha fazla dile getirildiği araştırılmaktadır. (Taecharungroj & Mathayomchan, 2019). Konaklama sektörü ile ilgili yapılan konu modellemesi çalışmalarında ise, konaklama alanında faaliyet gösteren ve web 2.0 imkanlarından yararlanan otel işletmelerine yapılan yorumlar farklı açılardan ve farklı metotlarla incelenmektedir. Bu analizlerde otellerin sınıfları, bulunduğu konum, otellere gelen tüketici tipi, zaman gibi unsurlar dikkate alınarak konu modelleme çalışmaları yapılmaktadır.

Lokasyon unsuru göz önüne alınarak yapılan çalışma Güney Kore’de gerçekleşmiş ve 104.000 yorum incelenerek konu modelleme araştırması yapılmıştır. GDA (Gizli Dirichlet Ayrımı) modelinin uygulandığı çalışmada 14 tema veri setinden çıkarılmıştır. Şehir içerisinde ve dışarısında bulunan otellerin karşılaştırılması yapıldığında, şehir içerisinde bulunan oteller profesyonel hizmet, rahatlık sağlayan otel olanaklarına ve transfer olanaklarına diğer gruba göre daha fazla değer verdiği görülmüştür (Sutherland vd., 2020). Turist tiplerine göre yapılan konu modelleme araştırmasında ise, Kamboçya’daki otellerine gelen turistlerin doğu ve batı olarak sınıflandırılmış, batı ve doğu coğrafyasından gelen turistlerin hizmet, lokasyon, oda ve destinasyon konularına önem verirken, batıdan gelen turistlerin yiyecek-içecek ve rezervasyon konularına, doğudan gelenler ise daha çok otel olanaklarına önem verdiği belirtilmiştir (Sann & Lai, 2023).

Otel işletmelerinde konaklama yapan kişileri değindiği konuların ve ayrıca bu konuların olumlu ya da olumsuz olarak algılanması, bu işletmelerin hangi konulara daha ağırlık vermesi gerektiği ile ilgili belirli bir fikir verebilmektedir. Bu aşamada tüketiciler otellerin farklı alanlarına değinerek, hangi konuların ve konuları oluşturan ifadelerin ortaya çıkarılması işletmelerin hizmet performansı açısından önemli olmaktadır. Örnek olarak çevrimiçi yorumların konu modellemesi yapıldığı bir

arařtırmada, temizlik, yiyecekler, alıřanlar ve hizmet kalitesinin durumu, otellere gelen kiřilerin en fazla deęinmiř olduęu konulardır. Bu alanlarla ilgili olarak misafirler tarafından en fazla rahatsız olunan konu ise hizmette yařanan gecikmeler olduęu belirtilmiřtir. zellikle 5 yıldıızlı otellerdeki hizmet alımı srecinde bekleme sreci, buradaki etmenlerden biri olarak grlmektedir (Afaq vd., 2022). Bu sebeplerden dolayı hem duygu analizi hem de konu modelleme alıřmalarının bir arada yrtlmesi, tketicilerin hangi konularda olumlu ya da olumsuz dřnceye sahip olmasının anlařılması adına nem tařımaktadır.

Duygu analizi ve konu modelleme analizlerinin birlikte yapıldıęı bir alıřmada, Brezilya' daki resort turizmi kapsamında ele alınan otellerin yorumları incelenmiř ve 13 adet konu belirlenmiřtir. Konular arasında oda, park imkanları ve rezervasyon en fazla negatif ynde aıklanan konular olurken, lokasyon, otel ambiyansı, alıřanlar, kahvaltı ve maliyet-fayda konuları, en fazla pozitif olarak belirtilen konular olmuřtur (Paladini & Peres, 2021). Havaalanı blgesine yakın olan zincir otellerine yapılan konu modelleme arařtırmasında ise transfer, personel, yiyecek ve iecek, temizlik, fiyat ve rezervasyon hizmetleri adı altında 7 konu bulunmuř olup, insanların memnuniyetini en fazla etki eden konunun transfer hizmetleri olduęu grlmřtr. Transfer hizmetleri, otellerin tek bařına uyguladıęı bir hizmet olmamasına raęmen, insanların otellere vermiř olduęu puana en ok etki eden unsur olduęu grlmřtr. Bu kapsamda otellerin direkt olarak kendileri saęladıęı olanaklar ile beraber evresel etkilerin de insanların konaklama deneyimine etki ettięi belirtilmiřtir. (Moro vd., 2020).

Konaklama sektrnde faaliyet gsteren  popler platformda bulunan otel verilerin incelendięi alıřmada, otellere yapılan yorumlardan temel hizmet, deęer, ekicilikler, yemek-deneyim ve ana otel rnleri olarak 5 temel kategoride GDA modeli kullanılarak konu trleri belirlenmiřtir. alıřmada insanların memnuniyet puanına gre konuların daęılımı ıkarılmıřtır ve duygu analizi sınıflaması gerekleřtirilmiřtir. Buna gre otelerde kalan misafirlerin en ok řikayet ettięi konu otelin saęladıęı temel hizmetler ve otelin saęladıęı olanaklar yani ana rnler olmaktadır. Kiřilerin en fazla beęenmiř oldukları konular ise yemek, otelde geirdikleri deneyimler ve deęer ile ilgili konular olmuřtur (Xiang vd., 2017). Benzer analizlerin yapıldıęı bir dięer alıřmada ise konaklama yapan kiřileri memnun eden en nemli zelliklerin otelin lokasyonu, olanaklarının kalitesi ile personelin saęlamıř olduęu hizmetin kalitesi ve alıřanların gler yzl bir řekilde hizmet sunmasıdır (Calheiros vd., 2017).

New York şehrinde Airbnb üzerinden faaliyet gösteren mekanlar üzerine yapılan araştırmada ise, bu mekanlar ile ilgili olarak toplamda 43 farklı konu belirlenmiştir. Belirlenen konu başlıkları, hiyerarşik kümeleme yöntemi ile dendrograma aktarılmış ve hangi konuların birbirleri ile yakın olduğu ortaya konmuştur. Araştırma sonucuna göre ortaya çıkartılan tekrar ziyaret etmek konusu ile lokasyon konusunun birlikte kullanıldığı görülmüştür. Ayrıca aile olarak destinasyonu ziyarete gelen kişiler, manzara konusunu en çok önem verdiği konu olarak belirlemişlerdir. Bu şekilde konaklama yapılan yerlerin fiziksel özellikleri ile otel yönetiminin davranışları da birbirlerine yakın konular olarak ortaya çıkarılmıştır (Sutherland & Kiatkawsin, 2020).

Duygu analizi ve konu modellemenin yapıldığı diğer bir araştırma Vietnam'ın başkentinde 405 otelin verilerine uygulanmıştır. Burada otellerin hizmetlerine pozitif olarak en fazla etki eden unsurun yiyecekler ve oda durumu olduğu belirtilmiştir. Beş yıldızlı otel gruplarında ise spesifik olarak en fazla lokasyon ve restoran konularının etkilendiği ortaya çıkarılmıştır (Tran vd., 2019).

Türkiye'deki destinasyonlarda bulunan konaklama işletmelerinin metinleri üzerinden gerçekleştirilen konu modelleme çalışmalarının sayısı son yıllarda artmakla beraber, bu alanda yapılan çalışmalar az sayıda bulunmaktadır. Örneğin Alanya'da 20 otelin sahip olduğu toplamda 2743 yoruma yapılan çalışma sonucunda, metinlerden sırasıyla çalışanlar, odalar, yiyecek içecek, işletme olanakları ve ulaşım adlı beş adet konu ortaya çıkarılmıştır. Bu konulardan en fazla olumlu olarak bahsedilen konular çalışanlar ve yiyecek içecek hizmetleridir. En fazla olumsuz olarak bahsedilen konu ise ulaşım olarak görünmektedir (Sezgin & Duman, 2023). Antalya bölgesindeki otellerin kullanıcı yorumlarının incelenmesinde ise yorumların %80'i olumlu %20'sinin ise olumsuz olduğu görülmüştü. Araştırma sonucunda Alanya'daki yorumların %68'i olumlu, %32'i olumsuz olarak görünmektedir ve destinasyonlarda en düşük düzeyde olumsuz yoruma sahip olan Antalya sınırı içerisindeki destinasyon olarak görülmüştür. Konu dağılımlarına bakıldığında ise beş adet konu belirlenmiş olup, sırasıyla en fazla deneyim, değer ve eğlence, şikayet, temel hizmetler ve yapılacak şeyler üzerine konu dağılımı görülmektedir (Büyükeke vd., 2020).

Antalya'da yer alan 164 otelin metinlerinin incelendiği ve Türkçe yorumlar üzerine yapılan çalışmada Nzemberek kütüphanesinden yararlanılmıştır. Verilere hem duygu analizi hem de Tripadvisor'da bulunan 4 adet konu değeri üzerine yapılan çalışmada, genel memnuniyet düzeyini en fazla etkileyen bu dört maddede sırasıyla hizmet memnuniyeti, oda memnuniyeti, uyku memnuniyeti ve konum memnuniyeti

olarak belirtilmiştir. Yorumlara verilen puanlarda ise en güçlü ilişkiyi sırasıyla hizmet memnuniyeti, oda memnuniyeti, uyku ve konum memnuniyeti olarak belirtilmiştir. (Tuna, 2019).

Tripadvisor web sitesinden elde edilen İngilizce ve Türkçe yorumların makine öğrenmesi modelleri ile test edildiği duygu analizi çalışmasında, İngilizce ve Türkçe sözlük tabanlı bir duygu analizi sınıflaması verilere yapılmıştır. Araştırma sonucunda İngilizce veri setindeki sınıflama başarısı Türkçe veri setine göre daha başarılı olduğu belirtilmiştir (Polat & Ağca, 2022).

7 farklı destinasyonda bulunan 142 otelin yorumlarına duygu analizi ve konu modelleme kapsamında yapılan çalışmada, otel sınıflarının arttıkça duygu skorunun da arttığı belirlenmiştir. Ayrıca ortaya konan 18 konunun oransal olarak 4 ve 5 yıldızlı otellerde en fazla pozitif yoruma sahip olan konuları temizlik, çalışanlar ve restoran ile ilgili konular olurken, oransal olarak negatif yorumların en yoğun olarak belirtildiği konu temizlik ve lokasyon konu ikilisi oluşturmaktadır. 1 ve 2 yıldızlı otellerde ise oransal olarak temizlik, çalışanlar ve restoran konularını oluşturduğu ikili gruplar olurken, negatif olarak ise manzara ve temizlik ile temizlik ve lokasyon ikilisi olmaktadır (Caner, 2019).

3. YÖNTEM

Araştırmanın amacı, araştırma sürecinde uygulanan yöntemler ve araştırma modelinin ne şekilde oluştuğu ve neleri kapsadığı bu bölümde belirtilmiştir. Elde edilen veri setinin hazırlanma süreci ve özellikleri ile kullanıcı tarafından oluşturulan içeriklerin metin madenciliği yöntemleri ile etkili bir şekilde analiz edilmesini sağlayan önışleme sürecindeki adımlar açıklanmıştır. Araştırmada duygu analizi ve konu modelleme yöntemleri, veri setine uygulanmıştır ve bu yöntemlerin nasıl uygulandığı ve içerisindeki yapısal özellikler belirtilmiştir.

3.1 Araştırmanın Amacı

Bu çalışmada Alanya’da faaliyet gösteren otel işletmelerine Tripadvisor web sitesi aracılığı ile yorum yapan kişilerin yorumlarının içerikleri analiz edilerek, bu kişilerin memnuniyetini hangi konuların ne derecede etkilediği ortaya çıkarmak, araştırmanın amacını oluşturmaktadır. Otel işletmeleri, destinasyonu ziyaret eden turistlerin turizm deneyimini etkileyen önemli unsurlardan birisidir. Bu kişilerin otel işletmelerinde yaşamış olduğu deneyimler hem destinasyona yönelik imajı, hem de destinasyonu ve otel işletmelerini tekrar ziyaret etme niyetini etkileyen bir durumdur (Marandi vd., 2023; Kladou & Mavgarani, 2015). Bu sebeple otel işletmelerine çevrimiçi platformlarda yazılan yorumların analiz edilmesi, otel işletmelerinin varlığı ve destinasyon imajı için önemli bir etmendir.

Otel işletmelerinin sağlamış olduğu hizmetlerin tüketicilerde yarattığı memnuniyet, farklı açılardan ilgili literatürdeki çalışmalarda incelenmiştir. Bu çalışmaların geneli, anket, odak grup görüşmesi, mülakat şeklinde yöntemlerle olduğu görülmekle birlikte, bu yöntemlerle yapılan analizler otel işletmelerinin hizmet kalitesinin boyutlarını belirli seçenekler kapsamınaa incelemektedir. Ayrıca bu araştırmalar ile yüksek miktarda kişiye ulaşılabilmesi finansal ve zaman olarak araştırmacılara maliyetler getirebilmektedir (Guo vd., 2017). Bu sebeple çevrimiçi platformlardaki yorum sayısının fazla olması ve bu platformları kullanan kişilerin giderek artmaya başlaması, otel işletmelerinin hizmet kalitesinin ortaya çıkarılmasında etkili olmaktadır. Ayrıca bu platformlardaki yorumlarda tüketiciler, geçirdiği konaklama deneyimlerini farklı açılardan ve farklı şekillerde ifade edebilmektedir. Bu platformlardan elde edilecek verilerin çeşitliliği göz önüne alındığında büyük veri

analizleri, çalışmalarda genellenebilir sonuçların ortaya çıkarılmasına imkan vermektedir (Rahman, 2019).

Bu sebeple bu araştırmada, metin madenciliği alanında kullanılan makine öğrenmesi yöntemleri ile otellere yapılan yorumlar analiz edilmiştir. Analiz sürecinde, makine öğrenmesi yöntemlerinden konu modelleme ve duygu analizi çalışmaları yapılmıştır. Otel işletmelerine çevrimiçi platformlarda yapılan yorumlara uygulanan konu modelleme çalışmalarının duygu analizi ile beraber uygulanması, hangi konuların tüketiciler tarafından olumlu ya da olumsuz olarak ele alındığı ve bu durumun müşteri memnuniyetini ve doğal olarak otelin sağladığı hizmet kalitesini göstermesini olanak tanıyan yaklaşımlardır (Hu vd., 2019). Bu sebeple bu araştırmada, Alanya’ daki otellerin yorumlarının büyük veri analizine olanak sağlayan konu modelleme ve duygu analizi yöntemleri beraber uygulanarak, konaklama yapan kişilerin hangi konularda pozitif ve negatif düşüncelere sahip olduğunun ortaya çıkartılması amaçlanmıştır.

Ortaya çıkartılan duygu tipleri ve konularla beraber, bu konuların duygu sınıflarına etki oranları da çalışmada belirtilmiştir. Duygu analizi ve konu modelleme çalışmalarının beraber yapıldığı çalışmalarda, hangi konuların hangi duygu türüne daha fazla etkisi olduğunun ortaya çıkarıldığı çalışmalar var olmasına rağmen, Türkiye’ de turizm alanında bu konuda yapılan çalışmaların son derece sınırlı olduğu görülmektedir.

Araştırmada bir diğer yapılan çalışma ise yorumlarda geçen konuların birliktelikleri analiz edilmiştir. Çevrimiçi platformlarda yazılan yorumlar, birden çok içeriği barındırmaktadır. Bu sebeple bu içerikler, tüketicilerin otele karşı memnuniyetini tek başına değil, bazen birlikte olarak etki ettiği görülmektedir (Başaran vd., 2020). Bu sebeple ortaya çıkan bu konuların birlikteliğinin araştırılması, otel işletmelerinin müşteri memnuniyetini arttırmak için alacakları önlemlerin ve kararların belirlenmesinde etkili olmaktadır. Bu amaçla araştırmada hiyerarşik kümeleme yöntemi ve uyum analizi ile ortaya çıkan konuların birliktelikleri değerlendirilmiştir. Ayrıca Alanya’daki otellere yapılan yorumlardaki konulara ilişkin olarak, birlikteliğin ele alındığı hiyerarşik kümeleme yöntemi ve uyum analizi yöntemi ile yapılan çalışmalara rastlanmadığından, bu yöntemlerin ortaya çıkaracağı sonuçlar Alanya’ da bulunan otel işletmelerine çok yönlü bir şekilde bilgi sağlaması amaçlanmaktadır.

Alanya’nın Türkiye turizmindeki yerine bakıldığında, Alanya bölgesindeki otellere yapılan yorumları incelemek, özellikle resort turizmin ağırlıkta olduğu bölgeler için de önemli olmaktadır. Kültür ve Turizm Bakanlığı’nın verilerine göre, 2022 yılında Türkiye’yi 51.387.513 kişi ziyaret etmiştir. Antalya’ya giriş yapan yabancı ziyaretçi

sayısı ise 12.828.472 kiři olarak belirtilmektedir. Bu kapsamda Trkiye'ye giriř yapan yabancı ziyaretilerin yaklaşık olarak drtte biri Antalya iline gelmektedir. Trkiye'de bakanlık belgeli toplam yatak sayısı 2023 nisan ayı itibarı ile 1,77 milyondur ve bu sayının 606.098'ini Antalya oluřturmaktadır. Trkiye'deki yatak kapasitesinin ise yaklaşık %35'ini Antalya ili kapsamaktadır. (KTB, 2022; AKTOB, 2023). Alanya Turistik İřletmeciler Derneęi'nin 2020 yılındaki istatistiklerine gre, Alanya'daki konaklama iřletmelerinin yatak kapasitesi 119.141 olarak belirtilmiřtir. Bu durumda Alanya, Antalya'daki yatak kapasitesinin yaklaşık olarak %20'sini oluřturmaktadır. Bu bakımdan Alanya ilesi hem Antalya hem de Trkiye de turizm sektrn oluřturan nemli paralardan biri olan konaklama sektr aısından nemli bir destinasyon konumundadır. Bu iledeki bulunan konaklama iřletmelerine yapılan evrimii yorumların analiz edilmesi, bu yeri ziyaret eden tketicilerin memnuniyetinin durumunun analiz edilmesine ve otellerin hizmet performanslarının grlmesine olanak saęlayacaktır.

Arařtırmada Mayıs 2023 tarihi itibarı ile Tripadvisor web sitesinde Alanya' da 952 otelin olduęu grlmektedir. Ancak Tripadvisor web sitesinde bu otellerin bazılarının Manavgat blgesinde olduęu grlmřtr. Ayrıca bu otel sayısının ierisinde apart, iftlik, bungalow gibi iřletmelerin de olduęu grlmektedir. Buna ek olarak bazı oteller farklı isimlerle iki ya da daha fazla otel listesinde olduęu grlmektedir. Bu sebeple, arařtırmada veri setinin kullanılacaęı her otelin konumunun Alanya ile sınırı ierisinde olmasına dikkat edilerek veri seti oluřturulmuřtur. Arařtırmada uygulanacak olan metin madencilięi modellerinin İngilizce dili ile uyumlu olması ve analizlerin gvenirlięi aısından minimum 10 ve zeri İngilizce yoruma sahip olan konaklama iřletmeleri veri setine alınmıřtır. Bu kapsamda toplamda 353 otel alıřmanın amacına uygun olarak bulunmuřtur ve bu otellerin sahip olduęu yorumlar web kazıma teknikleri ile elde edilmiřtir.

3.2 Arařtırmanın Yntemi

Arařtırmada metin madencilięi tekniklerinden metin sınıflandırılması, konu modelleme ve duygu analizi yntemleri elde edilen veri setine uygulanarak, sonular tablo ve grafikler ile gsterilmiřtir. Metin sınıflandırma iřleminin performansı ise farklı makine ęrenmesi modelleri ile test edilmiřtir. Buna ek olarak, veri setindeki elde edilen konular ve kategorik deęiřkenler arasında Anova ve Ki-Kare analizi yapılarak,

bu deęişkenlerin arasındaki farkların anlamlılıęı gösterilmiřtir. Elde edilen konular arasındaki iliřkinin grafiksel gösterimini ise, hiyerarřik kümeleme yöntemi, uyum analizleri ve uzaklık matrisi hesaplamaları ile belirtilmiřtir. Ayrıca arařtırmada ortaya çıkartılan konuların, konaklama yapan kiřilerin yorumlarının sınıfına göre hangi konunun daha fazla etkiledięini göstermek için lojistik regresyon analizi yapılmıřtır ve bu analiz sonuçları da nomogramlar vasıtası ile gösterilmiřtir. Arařtırmada ilk olarak verilerin elde edileceęi oteller belirlenmiřtir. Bu oteller otel sınıfına göre ayrı ayrı kategoriler řeklinde, her bir otelin Tripadvisor sitesindeki URL adresi, bu otellerin sahip olduęu yorum sayısı ve ingilizce yorum sayısı elde edilmiřtir. Otel sınıflamaları saęlandıktan sonra, her bir sınıfa ait otelin URL adresleri ve buradaki yorum sayısına göre otellerin her on yorumda bir deęişen URL adresleri belirlenerek veriler bu adreslerden çekilmiřtir. Verilerde konaklama yapan kiřilerin, bu otellere yapmıř oldukları tüm ingilizce yorumlar veri setine dahil edilmiřtir.

3.3 Arařtırmanın Modeli

Arařtırmada web kazıma yöntemleri ile veriler toplandıktan sonra, verilerin ayrıştırma iřlemi gerçekleřmiřtir. CVS veri türü olarak elde edilen veri seti, excel tablosu haline getirilerek analize hazırlanmıřtır. Arařtırma kapsamında ele alınacak olan farklı deęişkenlerin ve analiz türüne göre farklı excel tabloları hazırlanmıřtır ve Orange veri madencięi programı ile hazırlanan veriler analiz edilmiřtir. Veri analizinde sırasıyla ön iřleme, duygu analizi, duygu sınıflaması, eęitim ve test aşaması, konu modelleme, anova ve ki kare analizleri, uzaklık matrisi, hiyerarřik kümeleme yöntemi, uyum analizi ve regresyon analizi yapılarak bulgular elde edilmiřtir.

3.3.1 Veri setinin hazırlanması

Veriler Excel tablosunda bölümlere ayrılmıřtır. Buna göre öncelikle sitedeki verilerden elde edilen yorumlar ve yorum bařlıkları birleřtirilmiřtir. Özellikle konaklama yapan kiřiler, yařamıř oldukları deneyimi özetleyen kelime ve kelime gruplarını bařlıklarda göstermektedir ve bu durumun incelenmesi, hangi kelimelerin insanların yorumlarına etkilemesinin ortaya çıkarılması açısından gerekli bir unsurdur. Veri setindeki metin, otel sınıfı, seyahat türü, seyahat tarihi ve kiřilerin otel ile ilgili verdikleri puan excel’ de sütunlarda ayrı ayrı belirtilerek veri seti öniřlemeye uygun hale getirilmiřtir. Excel haline getirilen veri seti daha sonra Corpus denen tüm

metinlerin bulunduğu ve metin madenciliği için yapılacak olan analizlerde kullanılmasını yarayan veri türüne aktarılmıştır.

3.3.2 Önişleme (Preprocessing) aşaması

Corpus oluşturulduktan sonra ön işleme adımına geçilmiştir. Önişleme adımı, metin madenciliği ve bu kapsamda uygulanacak olan doğal dil işleme yöntemleri için en önemli adımlardan ve analizlerde uzun süre alan bölümlerden birisini oluşturmaktadır. Bu bölümde, elde edilen metinlerin hem duygu analizi hem de konu modelleme analizleri için uygun hale getirilmesi sağlanmaktadır. Özellikle önişleme sürecinin eksik ya da hatalı bir şekilde tamamlanması, analizlerin etkin bir şekilde çalışmamasına, analiz sonucunda çıkan sonuçların değerlendirilmesinin zorlaşmasına ve sonuçların güvenilirliğini olumsuz yönde etkileyebilecek durumlara sebep olabilmektedir (Hickman vd., 2022).

Her metin veri seti, kendine özgü özellikler barındırmaktadır. Otel, seyahat, destinasyon, havayolu, yiyecek içecek gibi turizmin farklı alanları ile ilgili elde edilecek metinlerin veri yapısı, bu alanda kullanılan kelimeler ve bu kelimelerin anlamsal ve yapısal özellikleri, bu sektörler ile ilgili olarak kullanılmaktadır. Elde edilen metin setinin yapısı ve ilgili olduğu sektör düşünülerek, analiz sonuçlarının etkili bir şekilde ortaya çıkartılması için ön işleme titizlikle ele alınmalı, kelime ve kelime grupları dikkatli bir şekilde incelenmeli ve sürekli olarak deneme yapılarak en iyi sonucu verebilecek önişleme aşaması uygulanmalıdır.

Önişleme sürecinde ilk olarak, metin setindeki tüm kelimeler küçük harfe dönüştürülmüştür. Bu özellik, kullanılacak olan yöntemler aynı kelimeleri yakalaması açısından önemlidir. Örneğin Otel, otel, OTEL kelimelerinde büyük-küçük harf dönüşümü yapılmadan analizlere bu şekilde sokulduğu takdirde, kullanılacak yöntemler bu 3 kelimeyi farklı olarak algılayacak ve bu durum analiz sonuçlarını etkileyecektir. Bu sebeple büyük-küçük harf dönüşümü yapılmış, daha sonra önişleme sürecindeki diğer aşamalara geçilmiştir.

Önişleme adımıdaki diğer basamak tokenization denen, metin setindeki her bir ifadeyi simgeleştirmeye yarayan işlem gerçekleşmiştir. Bu işlemde, metin setleri belirli parçalara bölünerek, analizde hangi parçaların veri setinde kullanılacağına karar verilir ve bu parçalar analize dahil edilerek işlemler yapılmaktadır. Tokenization bölümünde, metin setleri ister kelime olarak ister cümle olarak, isterse kelimeler arasındaki boşluklar hesaba katılarak ya da noktalama işaretleri ile ayrılarak parçalara

bölünebilmektedir. Bu bölünme işlemi, veri setinin özellikleri ve çalışmanın amacına göre kullanıcı tarafından belirlenir. Bu araştırmada, tokenization işlemi metin setindeki her bir kelimeye göre ayarlanmıştır. Bu noktada Regexp, Pattern: \w+ yöntemi kullanılmış olup, metin setinde bulunan ve tek olarak yazılan her bir kelime metin setinde tokenlere ayrılmıştır.

Tokenizaton işleminden sonra elde edilen metinlere filtreleme işlemi uygulanmıştır. Filtreleme işleminde, metin setindeki tüm sayısal ifadeler anlam bütünlüğünün sağlanması açısından çıkarılmıştır. Ayrıca ingilizce dilinde çok fazla kullanılan am, is, are, do, does gibi ifadeler de stopwords kapsamında veri setinden çıkarılmıştır. Bu ifadeler veri setlerinde doğal olarak çok fazla geçtiği için, metin analizlerinin sonuçlarını etkileyecektir.

Filtreleme işleminde ayrıca doküman frekans (document frequency) ile göreceli (relative) seçeneği aktif hale getirilmiştir ve sınırlılıkları %5 - %95 olarak kullanılmıştır. Bu şekilde bir kelime, corpusta bulunan yorumların %95' inden fazla ve %5' inden az geçerse, buradaki kelimeler devre dışı bırakılmıştır. Her yorumda geçen kelimeler, analiz sonuçlarını etkileyecektir ve yapılan konu modelleme çalışmasında oluşturulan her konuda bu kelimelerin geçme olasılığı yüksek olabilmekte ve bu durumda konuların farklılaşması ve değerlendirilmesi zor bir süreç olacaktır. Ayrıca yorumlarda imla hatası olan, yanlış yazılan ya da ingilizce yorumlara farklı dillerden yazılan kelimeler de bulunmaktadır. Bu sebeple bu gibi kelimelerin veri setinden çıkartılması için de göreceli filtre seçeneği minimum %5 seviyesinde tutulmuştur. Bununla birlikte, araştırmada veri setinde tek başına kullanımı spesifik bir anlamı olmayan kelimeler de veri setinden çıkarılmıştır. Örneğin stay, year, one, would, us, get, also, take, give gibi kelimeler, tek başına bir anlam nitelemediğinden anlamlı sonuçların ve oluşturulacak olan konuların anlam bütünlüğü açısından veri setinden çıkarılmıştır.

Önişleme aşamasında stemming (gövdeleme) işlemi yapılmıştır. Gövdeleme, yapım ya da çekim eki alan bir kelimeyi, kelimenin köküne indirgeyerek o kelimeyi anlamlı olarak veri setine girmesini sağlamaktadır. Örneğin "hotels" kelimesinin "hotel" şekline çevrilmesi gibi kelimelerin kök halinin veri setine konma aşamasıdır. Araştırmada yapılan gövdeleme işleminde öncelikli olarak veri setindeki tüm kelimelerin ham bir şekilde kelime sayısı ortaya çıkarılmıştır. Bu işlem, döküman frekansı sınırları içerisinde geçen kelimelere yapılmıştır. Bu kelimeler belirlendikten sonra, belirli kelimelere gövdeleme işlemi yapılarak, kök halindeki kelimeler veri setinde kullanılmıştır. Bu işlem yapılmadan önce gövdesiz bir şekilde kelimeler konu

modellemeye alınmıştır. Ancak veri setindeki bazı kelimeler bilindiği gibi farklı konuları nitelemektedir. Örneğin “clean” kelimesi, veri setinde daha çok odaları niteleyen bir kelime olarak görünmekte iken, “cleaning, cleans, cleaned” gibi kelimeler ise daha çok personeli niteleyen bir kelime grubu olarak, friendly kelimesi aynı şekilde personeli niteleyen bir kelime olarak kullanılırken, friend kelimesi ise seyahat tipini niteleyen bir şekilde kullanıldığı görülmektedir. Bilindiği gibi kelimelerin metin setleri içerisindeki kullanımı, araştırmanın yapıldığı alan ile ilgili olarak değişebilmektedir. Bu sebeple Porter Stemmer, Snowball Stemmer gibi geleneksel gövdeleme metotları yerine, veri setinde kullanılan tüm kelimelerin kullanım amacına göre bir gövdeleme işlemi gerçekleştirilmiştir. Örneğin kids, child, children kelimeleri kid kelimesine, waitress, waiters, waiter kelimeleri waiter kelimesine çevrilmiştir. Ayrıca metin setinde isim halinde bulunan tüm çoğul kelimeler de tekil hale getirilmiş ve -s,-es,-ies,-ed,-ied gibi ingilizce zamana göre çekim eki alan fiiller de kök haline getirilerek, araştırmada kök şeklindeki halleri ile kullanılmıştır.

3.3.3 Duygu analizi (Sentiment Analysis)

Araştırmadaki veri setindeki her bir yorumun sahip olduğu duyguyu ortaya çıkartılması için duygu analizi (sentiment analysis) yapılmıştır. Duygu analizi ile kişilerin her yoruma vermiş olduğu puanlardan ziyade, ortak bir sözlük ve metot kullanılarak tüm yorumların duygu analizi yapılmıştır ve bu şekilde yorumların duyguları ortak bir metoda indirgenmiştir. Araştırmada, duygu analizi yöntemlerinden VADER (Valence Aware Dictionary and Sentiment Reasoner) yöntemi kullanılmıştır. Bu yöntem, içerisinde barındırdığı pozitif ve negatif kelimeleri kullanarak, her bir yorumda hesaplama yapmakta ve yorumların duygu yoğunluğunu ve yorumların duygu tipini ortaya çıkartılmasında kullanılmaktadır. Bu şekilde metin setindeki her bir yorumun pozitif, negatif ve nötr olarak duygusal kategorisini hesaplar. Ayrıca kelimelerin metin setinde bir arada kullanılmasını da dikkate alarak kelime yoğunluğuna göre duygu skorunun ve duygu tipinin oluşmasını sağlar. VADER yöntemi, Python dili ile uyumlu kütüphanelerde bulunmaktadır ve bu yöntem ilgili kütüphanelerden indirilip veri setine adapte edilerek metin setindeki yorumların duygu analizi yapılabilir (Hutto & Gilbert, 2014). Bu çalışmada VADER yöntemi kullanılan programda Python dili ile uyumlu NLTK kütüphanesinden alınarak yapılmaktadır.

VADER yönteminde her kelimenin pozitif, negatif ve nötr olarak duygusal yoğunluğun bulunduğu kelimeler bulunmaktadır ve bir yorumun duygu analizi yapıldığında, bu kelimeler ve kelimelerin yoğunluğu dikkate alınmaktadır. Örneğin good kelimesi yanında olan very kelimesi, cümledeki duygu yoğunluğunu arttıracak bir unsur olduğundan bu durum cümlelerin pozitif bir cümle olma skorunu arttıracaktır. Ayrıca not bad gibi kullanımlar da dikkate alınarak cümlelerin duygu skoru hesaplanmasında dikkat ettiği bir unsurdur ve bu gibi durumlar doğal olarak cümlelerin pozitif bir şekilde, aksi takdirde daha negatif bir şekilde ulaşmasına sağlayacaktır. Ayrıca cümlelerde geçen kelimeler ve kelimelerin birliktelik kurallarına göre her cümlelerin bir compound(duygu skoru) skoru oluşmaktadır. Bu skor, -1 ile 1 arasında olmaktadır ve -1 ile 0.5 arasında bir değer alan cümle negatif, -0.5 ile 0.5 arasında olan bir cümle nötr, 0.5 ile 1 arasında değer alan bir cümle ise pozitif anlamda sınıflandırılmaktadır ve bu skor -1 ve 1'e yaklaştıkça duygu yoğunluğu da o derece artmış olarak kabul edilmektedir.

Önişleme yapıp cümlelerdeki kelimeler tokenize edildikten sonra yorumlardaki kelimelerin VADER algoritması tarafından matematiksel bir işlemde geçirilerek yorumların toplam skorları oluşmaktadır ve bu şekilde her bir yorum algoritmanın oluşturduğu skora göre sınıflandırılmıştır. VADER algoritmasında sözlükteki negatif ve pozitif kelimelerin kelimenin taşıdığı anlama göre bir skoru bulunmaktadır. Örneğin okay kelimesi 0.9, great kelimesi ise 3.1, horrible -2.5 değerinde sözlükte bulunmaktadır ve sözlükte yer almayan kelimeler 0 yani nötr olarak geçmektedir. Bu durumu ayrıca very, not gibi birliktelik analizi ile beraber hesaplayarak yorumların compound skoru (duygu skoru) oluşur. Algoritma, her yorumda bu kelimeleri yoğunluğuna ve sınırlılıklarına göre her kelimeyi belirler ve şekil 3.1'deki gibi her kelimenin skorunu cümle içerisinde skorları teker teker hesaplayarak ortalama oluşturur ve -1 ile 1 arasında bir değer yaratır (Ma, 2020).

$$\frac{x}{\sqrt{x^2 + \alpha}}$$

Şekil 3.1 Valence aware dictionary and sentiment reasoner formülü (Ma, 2020).

3.3.4 Gizli dirichlet ayrımı (latent dirichlet allocation)

Büyük veri kapsamında ele alınan metin madenciliği çalışmalarında makine öğrenmesi modellerinin ve bu modellerdeki sınıflama ve kümeleme metotlarının gelişimi, büyük veriden elde edilecek bilgilerin ve sonuçların çeşitliliğini ve kesinliğini hızlandırmıştır. Büyük metinlerin olduğu veri kümelerinde, metinlerin her birinin teker teker incelenip sınıflama ya da kümeleme yapılması büyük oranda zaman ve maliyet gerektiren işlemlerdir. Bu sebeple, metin setlerinin büyük miktarda olduğu veri gruplarında anlamlı sonuçlar elde edilebilecek içerik analizi yöntemleri geliştirilmektedir (Ekinci & diğ., 2020). Gizli Dirichlet Ayrımı (Latent Dirichlet Allocation), büyük miktarda metin gruplarını anlamlı bir şekilde kümeleme ve sınıflamada kullanılan ve konu modellemede kullanılan bir yöntemdir. Bir gözetimsiz makine öğrenmesi yöntemi olan GDA, metin analizlerinde son yıllarda sıklıkla kullanılan yöntemlerden biri olarak sıklıkla başvurulmaktadır. Özellikle turizm sektöründe popüler olan Tripadvisor, Booking, Agoda gibi çevrimiçi seyahat ve rezervasyon sitelerine kullanıcılar tarafından oluşturulan yorumlar, birden fazla konuyu içermektedir. Örneğin bir birey, konaklama yaptığı bir otel işletmesi hakkında söz konusu sitelerden birini kullanarak, deneyim geçirdiği otel ile ilgili düşüncelerini paylaşmaktadır. Ancak bu paylaşımda, söz konusu birey otel deneyimi ile ilgili olarak birden fazla düşüncüyü belirtmektedir. Örneğin otelin odalarını, yiyecek içecek hizmetlerini, personelin davranışlarını, otelin sunmuş olduğu olanaklar gibi farklı konuları, oluşturmuş olduğu yorumda bahsetmektedir. Buna ek olarak, tüketiciler bazen otel dışında yaşamış olduğu deneyimlerden de otel hakkında çevrimiçi platformlara yazmış olduğu yorumlarda bahsedebilmektedir. Örneğin otelin bulunduğu destinasyonda geçirdiği zaman, bu zaman içerisinde ziyaret ettiği yerler, ziyaret sürecinde ulaşım, ürün satın alma, yerel halkla ve işletme sahipleri ya da işletme çalışanları ile yaşamış olduğu deneyimlerden de yorumlarda bahsedilen unsurlardan olmaktadır. Bu süreçte, insanlar duygu ve düşüncelerini belirtirken, kelimeleri anlamlı bir şekilde bir araya getirerek bu düşüncelerini belirli bir sıraya koymakta ve geçirmiş olduğu deneyimi bu şekilde paylaşmaktadır ve bu paylaşımlarda birçok konu, yorum içerisinde farklı pozisyonlarda, gerek düzgün gerekse devrik cümleler halinde bahsedilmektedir. Bu sebeple, bu kişilerin yapmış olduğu yorumların konularının belirlenmesi dikkat gerektiren ve ön işleme süreci uzun olan bir aşama olmaktadır. Metin madenciliğinde kullanılan yöntemler ile bireylerin yorumlarda bahsettiği konuları

ele almak ve bu konular üzerine analiz yaparak, hizmet performansını hangi konuların etkilediği ortaya çıkartılması önemlidir.

Bu amaçla, metinlerde gizli kalan konuların ayrıştırılması ve bulunabilmesi için, etiketsiz olan metin verilerinde bu sebeple gözetimsiz makine öğrenmesi modelleri kullanılmaktadır. LDA modeli de bu gibi etiketi, diğer bir değişle verinin çıktısının bilinmediği metin analizlerinde sıklıkla kullanılan yöntemlerden birisi olmuştur (Chen vd., 2019).

Gizli Dirichlet Ayrımı, Bayes modeline dayanan bir modeldir. Bu modelin metin madenciliğinde kullanılmasında iki önemli faktörü göz önünde bulundurur. Metin setini dokümanlar oluşturmaktadır. Dokümanları ise kelimeler oluşturmaktadır. Her bir doküman ise, kelimelerin oluşturduğu konulardan meydana gelmektedir (Blei vd., 2003). Bu şekilde her bir konuyu kelimeler oluşturmakta, kelimeler ise dokümanları oluşturduğu için, her bir dokümanın belirli bir konu dağılımı bulunmaktadır.

GDA öncelikli olarak metin setini, doküman-kelime matrisi haline getirmektedir. Şekil 3.2’de örnek bir doküman-kelime matrisi gösterilmektedir. Örnekteki veri seti, 5 tane dokümandan oluşmaktadır. Bu doküman, bir araştırmadaki veri setinde kullanıcıların yapmış olduğu her bir yorumdur. Metin seti ise bu yorumların oluşturmuş olduğu dokümanların toplamıdır. Şekilde görüldüğü gibi bu örnekteki veri seti 5 dokümandan oluşmakta ve beş dokümanda toplamda 8 farklı kelime kullanılmıştır. İlgili matris hazırlanırken, bir dokümandaki kelimelerin Boolean veri tipine göre her bir dokümanda yer alıp alınmadığına dair matris düzenlenmektedir. Şekil 3.2 ve 3.3’te GDA çalışma şekli ve genel yapısı sunulmuştur. Şekil 3.4 ve 3.5’te ise GDA çalışma prensibinin nasıl gerçekleştirildiği ile ilgili örnek verilmiştir.

Verilen örnekte, metin seti içerisinde toplamda 5 tane doküman bulunmaktadır. Konu sayısı ise 6 olarak belirlenmiştir. Konu sayısı, kullanıcının metin setindeki veri yapısına göre değişmektedir. Şekil 3.4’teki örnekte metin setinde bulunan 5 dokümandan toplamda 8 farklı kelime oluşturulmuştur. Doküman matrisi oluşturulduktan sonra, araştırmada konu sayısı belirlenmektedir. Belirlenen konu sayısına göre, GDA yöntemi öncelikli olarak her bir konuda hangi kelimelerin yer aldığını, kelimelerin hem metin setinde hem de dokümanda geçme sıklığına göre belirler ve böylece kelime-konu matrisi oluşturur. Bunun yanında, her bir konuyu oluşturan kelimeler belirlendikten sonra, her bir dokümanın hangi konuyu hangi olasılıkla ifade ettiğini gösteren doküman-konu matrisi oluşturulur. Şekil 3.5’de

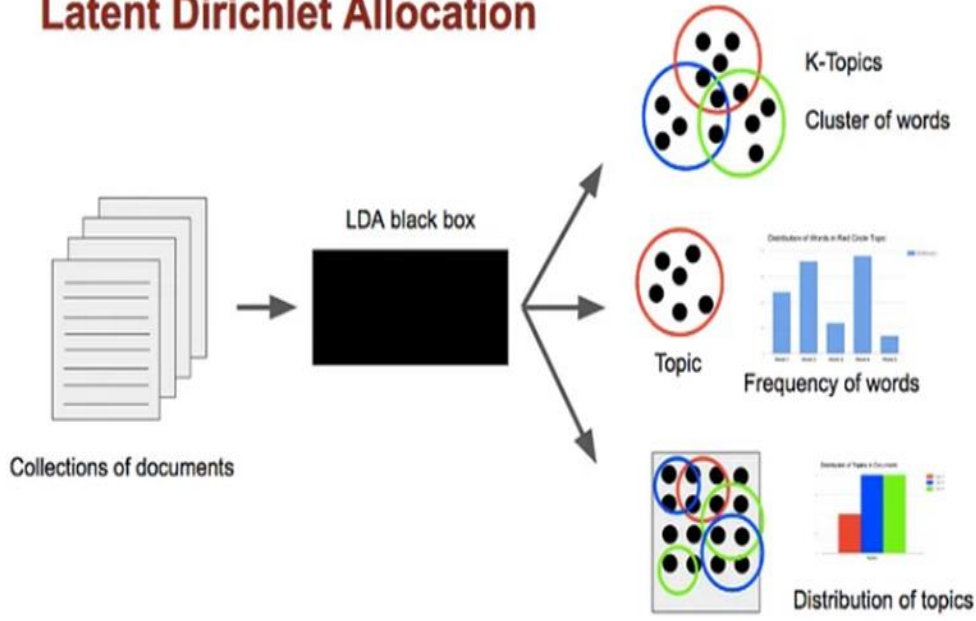
görüldüğü gibi, kelimeler hem dokümanlara hem de topiclere atanarak kümeleme işlemi şekil 3.2’ deki yapıdaki şekliyle gerçekleşmektedir

Şekil 3.3 LDA yapısını göstermektedir. Şekilde bulunan en iç kısımdaki tabaka, bir dokümandaki kelimeleri oluşturmaktadır. En dıştaki tabaka ise metin setindeki tüm dokümanları oluşturmaktadır. Buna göre öncelikli olarak LDA, bir dokümanda bulunan kelimeleri ele alır ve bu kelimelerin her birini oluşturulan konulara atar ve bu şekilde bir konuda hangi kelimelerin olasılıksal olarak dağıldığını gösterir. Bu işlem gerçekleşirken, aynı zamanda kelimelerin oluşturmuş olduğu dokümanların hangi konuları içerdiğini de hesaplamaktadır. Dolayısıyla metin setinde ilgili algoritma sürekli çalışarak, metin setindeki her kelimeyi, kelimelerin dokümanda geçmiş olduğu olasılığına göre dağıtmaktadır. Bu şekilde her bir konuda hangi kelime daha fazla geçmekte ise, o konuda o kelimelerin geçme olasılığı doğal olarak üst sıralarda olmaktadır ve ilgili konuda daha az şekilde geçen bir kelime, doğal olarak ilgili konunun alt kısımlarına yerleşmektedir. Çünkü bu kelimelerin ilgili konu kümesinde geçme olasılığı daha düşüktür. Böylece ilgili yöntem, kelimeleri ve dokümanları metin setinde sürekli olarak dağılım yapar ve metin setindeki her dokümanın ve dokümandaki kelimelerin hesaplamasını bitirdikten sonra, oluşturulan konuların kelimelerini ve kelime ağırlıklarını sunar. Bu şekilde öğrenme işlemi gerçekleştirilir ve her bir dokümanın ve ilgili konuya göre, tüm veri setinin ise oluşturulan konulara göre dağılımları ortaya çıkartılır (Seth, 2021; Lee vd., 2018; Fernandez vd., 2020; Jelodar, 2019; Maulana, 2017). Araştırmada kullanılan GDA modelindeki parametrelerin ifade edilişi Tablo 3.1’te gösterilmiştir.

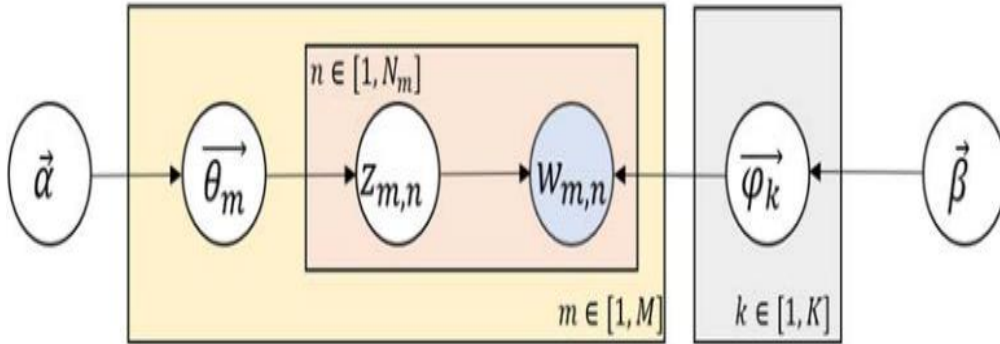
Tablo 3.1 Gizli dirichlet ayrımı parametreleri

Parametre Sembolleri	Parametreler İfadeleri
M	Metin setindeki dokümanlar
m	Metin seti içerisindeki bir doküman
N	Metin seti içerisinde ele alınan bir dokümandaki kelime sayısı
n	Metin seti içerisinde ele alınan bir dokümandaki ilgili kelime
W	Metin seti içerisinde ele alınan dokümandaki bir kelime
Z	Metin seti içerisinde ele alınan dokümandaki bir kelimenin atanmış konusu
K	Belirlenen konu sayısı
k	Belirlenen konu sayısı içerisindeki bir küme
α	Doküman ve konu dağılımını kontrol eden Dirichlet parametresi
β	Konu ve kelime dağılımını kontrol eden Dirichlet parametresi
θ	Doküman ve konu dağılım parametresi
ϕ	Konu ve kelime dağılımını sağlayan parametre
m, n	Matristeki kelimenin, konunun ve dokümanın m. sıra n. Sütundaki değeri

Latent Dirichlet Allocation



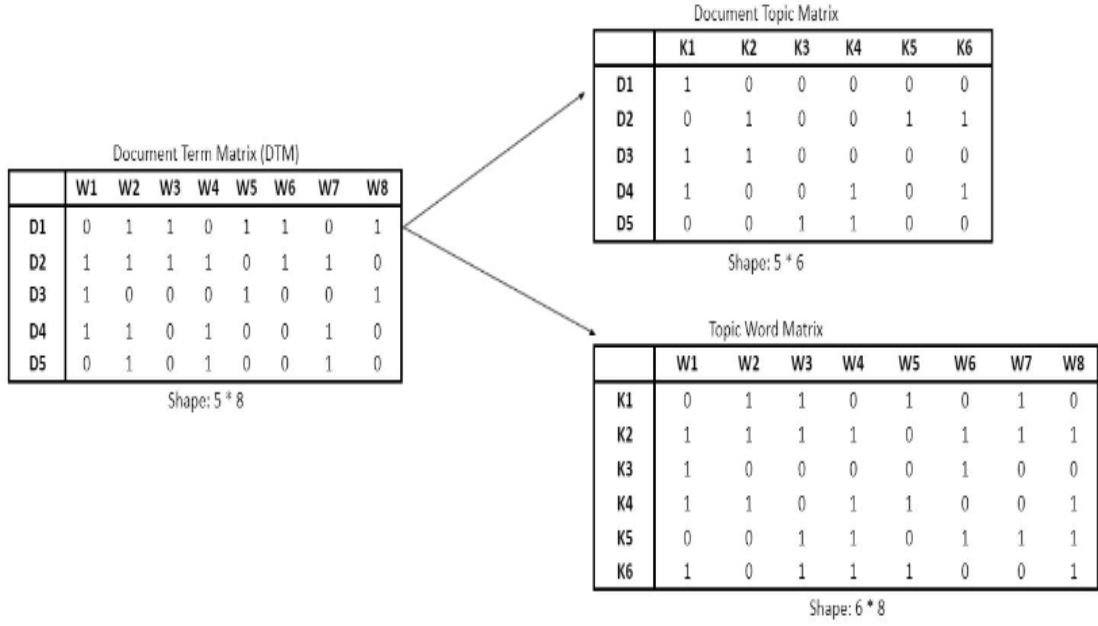
Şekil 3.2 Gizli dirichlet ayrımı çalışma şekli (Maulana, 2017).



Şekil 3.3 Gizli dirichlet ayrımı yapısı (Lee vd., 2018).

	W1	W2	W3	W4	W5	W6	W7	W8
D1	0	1	1	0	1	1	0	1
D2	1	1	1	1	0	1	1	0
D3	1	0	0	0	1	0	0	1
D4	1	1	0	1	0	0	1	0
D5	0	1	0	1	0	0	1	0

Şekil 3.4 Doküman ve kelime matrisi oluşumu (Seth, 2021).



Şekil 3.5 Doküman ve kelimelere göre konu oluşumu (Seth, 2021).

Konu sayısı belirlenirken iki önemli unsur göz önünde bulundurulmaktadır. Bunlar konu uyumluluğu (coherence) ve çapraşıklık (perplexity) sonuçlarıdır. Konu uyumluluğu, oluşturulan bir konudaki kelimelerin birbirleri ile benzerliğin ölçülmesi ile oluşturulur. Yüksek uyum skoru, konuları oluşturan kelimeler arasında uyumu göstermektedir ve seçilecek olan konu sayısında bir fikir sahibi olunmasını sağlar (Güven ve diğ, 2018). Çapraşıklık ise modelin yeni bir veri ile karşılaştığında bu veriyi tahmin edilmesinde kullanılan bir ölçüttür. Düşük çapraşıklık skoru, modelin doğru bir şekilde tahmin yürüttüğünü ve kelimeleri uygun konulara atadığını ve dokümanların doğru şekilde buna göre konulara dağılımı göstermektedir (Kapadia, 2019).

Konu modelleme kapsamında konuların dağılımının ve uyumunun gösterildiği çapraşıklık ve uyumluluk skorları, konu modelleme çalışmalarında tek başına yeterli olmayan unsurlardan birisi olabilmektedir. Örneğin konu sayısının artması, uyumluluk değerini yükseltebilmektedir ancak bu sefer çok fazla kelime her konuda olacağı için bu konuların değerlendirilmesi zor bir süreç olmaktadır. Bunun sebebi ise, konu sayısı arttıkça kelimelerin konulardaki frekans sayısındaki düşüşüdür ve bu sebeple bir kelime birden fazla konuda olma olasılığı yükselmektedir. Aynı şekilde çapraşıklık skorunun da az olması için konu sayısı düşürülebilmektedir. Ancak bu durumda konulardaki kelimelerin frekans değerleri yükselecektir ve bu konulara metin setinde en fazla geçen kelimeler gireceği için, bu sefer de konuların değerlendirilmesi ve spesifikleştirilmesi

zor bir durum olacaktır. Bu sebeplerden dolayı, konu modelleme çalışmalarında sürekli olarak deneme yapılması gerekmektedir. Bunun başlıca sebebi, hem uyumluluk ve çapraşıklık değerinin uygun bir aşamada olması, hem de konuların anlamsal bütünlüğünün yakalanması ve yorumlanabilmesi ve bunun sonucunda anlamlı bulgular elde edilebilmesidir (Kapadia, 2019; Chang vd, 2009). Konu sayısının az olması belgelerde az sayıda temanın olmasına neden olmaktadır ve araştırmada spesifik kelimelerin kaybolmasına sebep olabilmektedir. Aksi halde konu sayısının çok fazla olması durumunda, konuların değerlendirilmesi zor olabilmekte, konularda birçok benzer kelimeler çıkabilmektedir (Putri & Kusumaningrum, 2017).

3.3.5 Kelime çantası (bag of words)

Metin analizlerinde her bir cümleyi oluşturan kelime ve karakterlerin sayısal bir ifade olarak ele alınmasında ve bu şekilde vektörleştirilmesine yardımcı olan yöntemlerden birisini kelime çantası (Bag of Words (BOW)) modeli oluşturmaktadır (Zhang vd., 2010). Bu yöntem, belgelerin vektör temsiliinde belge içerisindeki her bir ögedeki kelimenin sayısı verilmektedir (Zhao & Mao, 2017). Bu modelde oluşturulan terimlerin tüm metin setinde geçme ve bu şekilde her kelime ya da kelime grubu sayısal hale getirilerek bir veri haline getirilir (Aksu ve Karaman, 2020). Böylece metin madenciliği alanında metin setlerini ve bu setlerin oluşturduğu dokümanlalar üzerinde sınıflama ve kümeleme analizleri yapılmasını sağlar (Martins vd., 2003). BOW yönteminin bu şekilde sağladığı dönüştürme ilemetinlerin makine öğrenmesi sürecine hazır hale getirilir ve metin madenciliği çalışmalarının yapılabilmesine olanak tanır. (El-Din, 2016; Yan vd., 2020).

BOW yöntemi, konu modelleme algoritmaları ile bir arada kullanıldığında daha sağlıklı sonuçlar verebilmektedir. BOW yönteminin içerisinde bulunan TF*IDF dönüşümü yapılarak TF*IDF matrisi elde edilmektedir ve bu şekilde, metin seti içerisinde çok fazla geçen kelimelerin önemi belirli ölçüde indirgenmiş olmaktadır (Zagar, 2022). Böylece her bir konuda daha spesifik kelimeler geçmekte ve bu şekilde konuların içermiş olduğu kelimelerin anlam bütünlüğü sağlanabilmektedir.

3.3.6 Terim ve ters belge frekansı (term and inverse document frequency)

Terim frekansı (Term Frequency, (TF)) ve ters belge frekansı (Inverse Term Frequency, (IDF)) metin madenciliği alanında bir kelimenin içinde bulunduğu dokümandaki ağırlığını, bir diğer ifade ile önemini gösteren ve kelimelerin ağırlıklarını

ortaya çıkartan yöntemlerdir. TF değeri ve IDF değerleri kelimelerin ayrı bir şekilde hesaplanır ve ortaya çıkan sonuç, kelimelerin doküman içerisindeki ağırlıklarını vermektedir (Sebastian, 2020).

Terim frekansının ve ters belge frekansının hesaplanmasında kullanılan formüller şekil verilmiştir. Huilgol (2020) örneği baz alınarak araştırmadaki örnek bir veri setinden BOW ve TF*IDF skoru hesaplamak için üç cümle ele alınmıştır. Cümleler aşağıdaki örnek olarak oluşturulmuştur:

- 1.cümle: This hotel was nice and staffs were helpful
- 2.cümle: This hotel was not nice and staffs were not helpful
- 3.cümle: This hotel was amazing and staffs were great

Tablo 3.2’de ilgili cümlelerin kelimeleri, her dökümana göre oluşturularak, oluşturulan metin setinde her kelimenin bulunma sıklığı belirtilmiştir.

Tablo 3.2 BOW doküman-kelime matrisi oluşumu

	This	hotel	was	nice	and	staffs	were	helpful	not	amazing	great
1.Doküman	1	1	1	1	1	1	1	1	0	0	0
2.Doküman	1	1	1	1	1	1	1	1	2	0	0
3.Doküman	1	1	1	0	1	1	1	0	0	1	1

BOW modelinin oluşturulması ise aşağıdaki şekilde gerçekleşmektedir:

1. Doküman Vektörü: [1 1 1 1 1 1 1 1 0 0 0]
2. Doküman Vektörü: [1 1 1 1 1 1 1 1 2 0 0]
3. Doküman Vektörü: [1 1 1 0 1 1 1 1 0 0 1 1]

Örnekteki üç cümlelerin TF*IDF skorlarının oluşturulması ise Tablo 3.3 ve 3.4 kapsamında verilmiştir.

Tablo 3.3 BOW ve terim frekansı oluşumu

	This	hotel	was	nice	and	staffs	were	helpful	not	amazing	great
1.Yorum	1	1	1	1	1	1	1	1	0	0	0
2.Yorum	1	1	1	1	1	1	1	1	2	0	0
3.Yorum	1	1	1	0	1	1	1	0	0	1	1

Tablo 3.4 BOW ve TF*IDF matrisinin oluşumu

	BOW1	BOW2	BOW3	TF1	TF2	TF3	IDF	LOG IDF	TF*IDF 1	TF*IDF 2	TF*IDF 3
This	1	1	1	0,125	0,1	0,125	1	0	0	0	0
hotel	1	1	1	0,125	0,1	0,125	1	0	0	0	0
was	1	1	1	0,125	0,1	0,125	1	0	0	0	0
nice	1	1	0	0,125	0,1	0	1,5	0,17	0,02	0,01	0
and	1	1	1	0,125	0,1	0,125	1	0	0	0	0
staffs	1	1	1	0,125	0,1	0,125	1	0	0	0	0
were	1	1	1	0,125	0,1	0,125	1	0	0	0	0
helpful	1	1	0	0,125	0,1	0	1,5	0,17	0,02	0,01	0
not	0	2	0	0	0,2	0	3	0,47	0	0,09	0
amazing	0	0	1	0	0	0,125	3	0,47	0	0	0,05
great	0	0	1	0	0	0,125	3	0,47	0	0	0,05

TF skoru hesaplanırken, bir kelimenin bir dokumanda geçme sıklığı, bu kelimenin bir dokumanda geçme sayısı, bu dokumandaki toplam kelime sayısına bölünmesi ile oluşmaktadır. IDF ise metin setindeki toplam doküman sayısının, kelimenin geçtiği doküman sayısına bölümünün logaritmasının alınması ile hesaplanmaktadır. TF*IDF ise hesaplanan bu iki değerin birbirleri ile çarpımı sonucunda oluşmaktadır ve bu şekilde her bir kelimenin metin içerisindeki ağırlığı tespit edilmektedir.

$$tf_{t,d} = \frac{n_{t,d}}{\text{Number of terms in the document}}$$

$$idf_t = \log \frac{\text{number of documents}}{\text{number of documents with term 't'}}$$

$$(tf_idf)_{t,d} = tf_{t,d} * idf_t$$

Şekil 3.6 Terim ve ters doküman matrisi formülleri (Huilgol, 2020).

3.3.7 Doküman gömme (document embedding)

Metinlerin sınıflandırılması, kümelenmesi ve metin setlerinden anlamlı bilgilerin çıkartılmasında kullanılan modellerden birisidir (Wang vd. 2016). DEM (Document Embedding) modeli, metin seti içerisindeki kelimelerin vektörleştirme işleminde kullanılan bir diğer modeldir. Bu aşamada, dokümandaki kelimelerin birlikteliği ve doküman içerisindeki dağılımı da göz önünde bulundurularak vektörleştirme işlemi yapılır ve bu işlemden sonra makine öğrenmesi modelleri ile sınıflandırmalar test edilmektedir (Wang vd., 2018). DEM modelinde kelimelerin sırası ve kelimelerin yapıları da dikkate alınarak metin seti içerisindeki dokümanlar vektörleştirilir. Bunu yaparken, metin setindeki her dokümanı n-gram şeklinde parçalara ayırır. Bu şekilde oluşturmuş olduğu n-gramları önceden eğitilmiş olan modellerle her dokümanı bir vektör haline getirir. Vektör haline getirildikten sonra, bu cümleler makine öğrenmesi algoritmaları ile test edilebilir aşamaya gelir. Doküman gömmede birbirlerine benzer olan kelimeler, vektörleştirme aşamasında aynı kelime ile ifade edilirler (Demsar ve diğ., 2013). Örneğin Mother, mommy, mom kelimeleri, 3 farklı kelime iken, doküman gömme modelinde her biri tek bir kelime olarak vektörleştirme aşamasında temsil edilmektedir. Bu şekilde özellikle metin sınıflandırma işlemlerinde etkili sonuçlar verebilmektedir.

Bu araştırmada doküman gömme modellerinden Fasttext yöntemi kullanılmıştır. Fasttext, açık kaynak kodlu bir kütüphane olup, içerisinde eğitilmiş birçok kelime vektörlerini barındır (fasttext.cc, 2023). Ayrıca dokümanların önceden eğitilmiş modellere göre vektörleşmesini sağlar (Grave vd., 2018; Joulin, 2016; Orange, 2023). Fasttext, içerisinde karmaşık sözcük barındıran dokümanlardaki kelimeleri anlamak için morfolojik özellikler kullanmaktadır ve bu durum vektörleştirmede kelimeleri metin setinde temsili için uygun bir hale getirmektedir. Bu şekilde doküman içerisinde özellikle yazım yanlışı olan kelimelerin temsilde kolaylık sağlar ve bu durum metin sınıflamaya yardımcı olmaktadır (Umer vd., 2023).

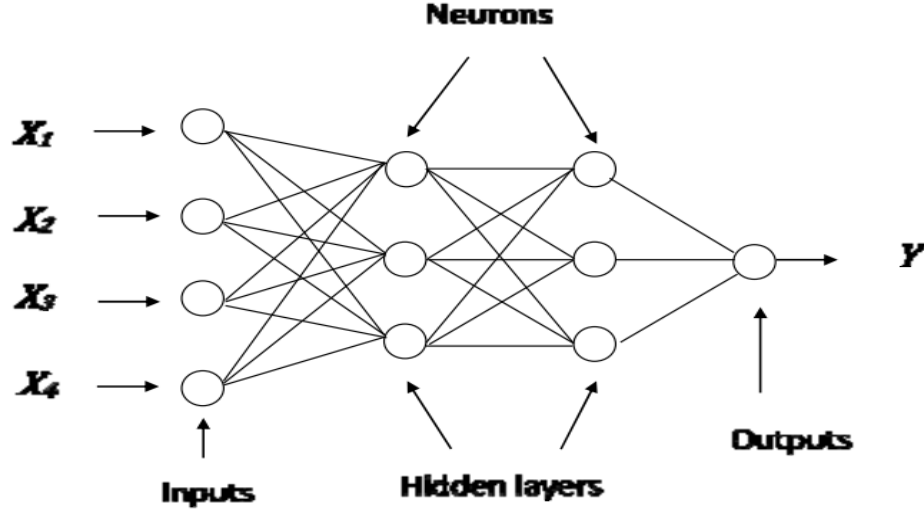
3.3.8 Sinir ağları (neural networks)

Araştırmada duygu sınıflandırması, farklı algoritmalar ile denenerek sınıflama persormansları elde edilmiştir. Ancak veri setinin büyüklüğü ve kullanılan bilgisayar CPU seviyesi dikkate alındığından, öncelikle veri setinin %25'i test edilmiş olup, bu aşamada en iyi sonucu sinir ağları model vermiştir. Bu sebeple bir sonraki aşamada veri setinin %50'si sinir ağları yöntemi ile test aşaması yapılmıştır. Sinir ağları, insanlardaki

sinir sistemi baz alınarak oluşturulan bir model olup, sinir sisteminin parçası olan nöronları ve bu nöronların arasındaki ilişkiyi sağlayan sinir ağları temel alınarak oluşturulan bir yapıdır. Bu yapı girdi, çıktı ve gizli katmanlardan meydana gelmektedir ve her katmanda nöronlar bulunmaktadır. Giriş katmanındaki nöronlar bilgiyi gizli katmanlara göndermekte ve gizli katmanlara aktarılan bilgiler, ağ yapısındaki parametreler vasıtasıyla dönüştürülüp çıktı katmanına aktarılmaktadır. Girdi ve çıktı katmanlarındaki nöron sayısı, çalışmadaki bağımlı ve bağımsız değişken sayısına göre belirlenme olup, gizli katmanlardaki nöron sayısı ve bu katmandan kaç tane oluşturulacağı ise algoritmanın performans durumuna göre araştırmacı tarafından belirlenmektedir (Atalay & Çelik, 2017; Witten, 2011).

Sinir ağlarındaki girdi ve çıktılar arasında lineer bir ilişki olmaması, bağımlı değişkenin çok fazla sayıda olması gibi durumlarda, modelin daha iyi sonuç verebilmesi için çok katmanlı algılayıcılar kullanılmaktadır (Çılgin vd., 2020). Çok katmanlı algılayıcılar (Multi-layer Perceptron), bir denetimli öğrenme algoritmasıdır ve makine öğrenmesi sürecinde bağımlı ve bağımsız değişkenler arasındaki ilişkiyi doğrusal olmayan fonksiyon şeklinde hem regresyon hem de sınıflama yapmak için öğrenmektedir (Pedregosa vd., 2011). Şekil 3.7 de örnek bir MLP sinir ağı modeli gösterilmiştir. Bu modelde görüldüğü gibi, bağımsız değişken sayısı kadar girdi katmanında, bağımlı değişken sayısı kadar da çıktı katmanında nöron kullanılmıştır. Girdi ve çıktı katmanları arasında ise 2 katman kullanılmış ve 1. gizli katmanda 3, 2. Gizli katmanda 2 adet nöron kullanılmıştır.

Katmanlara gelen bilgiler ise bazı fonksiyonlar yardımı ile bir sonraki katmanlara aktarılmaktadır. Buna kısaca aktivasyon denmektedir. Aktivasyonda bulunan fonksiyon ile bir nörona gelen bilgileri dönüştürerek bir sonraki katmanda bulunan nöronlara aktarılır. Bu aktarım sürecinin performansını daha iyi bir noktaya çıkarmak ve bilgi aktarımında hataları minimuma indirmek ve daha etkin bilgilerin nöronlara aktarılması için ise solver kullanarak, hata minimize etmeye çalışılır. Ayrıca oluşturulan sinir ağlarının aşırı uyumunun önlenmesi için Alpha= L2 parametresi kullanılmaktadır. Araştırmada solver olarak AdAm, aktivasyon olarak ise ReLU, Alpha değeri 0.00010 ve maksimum döngü 200 parametreleri kullanılmış olup, diğer parametreler default olarak Python Scikit-Learn kütüphanesinden alınmaktadır. Araştırmada 6 adet gizli katman kullanılarak, nöron sayıları sırasıyla 100, 200, 250, 200, 100 ve 50 olarak belirlenmiş olup, modelde en iyi sınıflama becerisi elde edilmiştir.



Şekil 3.7. MLP sinir ağı modeli (Marques vd., 2014).

Sinir ağlarının performansını görmek için makine öğrenmesi yöntemlerinin performanslarının gösterilmesi kapsamında en fazla kullanılan özelliklerden olan karmaşıklık matrisi ve ROC (Receiver Operating Characteristic) eğrisi bu çalışmada eğitim ve test setlerinin karşılaştırılmasında ortaya çıkan sonuçları görmek açısından kullanılmıştır.

3.3.9 Karmaşıklık matrisi (confusion matrix)

Makine öğrenimi modellerinin başarısını ölçmek için çeşitli göstergeler kullanılmaktadır. Bu göstergelerin başlıca olanı karmaşıklık matrisidir. Karmaşıklık matrisi, makine öğrenmesi yöntemlerinin eğitim ve test seti olarak ayrıldıktan ve eğitim ve test işlemi bittikten sonra, yöntemlerin sınıflama becerisini gösteren, hangi verileri doğru ya da yanlış ayırdığını gösteren bir matristir. Böylece sınıflandırma yapacak olan makine öğrenmesi modelinin performansını ölçmektedir ve genellikle gözetimli öğrenme türündeki sınıflamalarda yapılmaktadır (Kumar & Paul, 2016).

Karmaşıklık matrisinde makine öğrenmesinde kullanılan yöntemin yapmış olduğu sınıflama başarılarında, matriste dört temel unsur kullanılmaktadır. Gerçek değeri doğru olup, makine öğrenmesi yöntemi bu değeri doğru olarak tahmin etmiş ise True Positive (TP), gerçek değeri doğru iken yanlış olarak sınıflandırılmış ise False Negative (FN) olarak gösterilmektedir. Buna ek olarak, gerçek değeri yanlış iken doğru olarak sınıflandırılmış ise False Positive (FP), gerçek değeri yanlış iken yanlış olarak sınıflandırılmışsa True Negative (TN) değerlerle ifade edilmektedir. Sınıflama bittikten

sonra gerçek değerler ve tahmin edilen değerler tablo 3.5’ de gösterilen şekilde bir matris halinde gösterilir.

Tablo 3.5 Karmaşıklık matrisi

Sınıflar		Gerçek Değerler	
		Pozitif	Negatif
Tahmin Edilen Değerler	Pozitif	TP	FP
	Negatif	FN	TN

Bu işlemden sonra, yöntemin sınıflandırmasının başarısının gösterilmesindeki önemli temel değerlerden bazıları duyarlılık, kesinlik, doğruluk ve F1 skorudur. Duyarlılık (recall) oranı, gerçekte pozitif olup pozitif tahmin edilen tahmin sayısının, gerçekteki tüm pozitif olan değer sayısına bölünmesi ile elde edilir. Kesinlik (precision) değeri ise, modelin doğru pozitif olarak tahmin ettiği değerlerin, tüm pozitif tahminlere bölünmesi ile oluşturulmaktadır. Doğruluk (accuracy) değeri ise, modelin tüm doğru olarak sınıflandırdığı değerlerin, toplam değerlere bölünmesi ile oluşmaktadır. F1 skoru ise, duyarlılık ve kesinlik oranlarının duyarlılık ve kesinlik toplamına bölünerek ikiyle çarpılması ile bulunmaktadır. Bu şekilde hem kesinlik hem de duyarlılık oranları birleştirilerek hesaplanır, bu şekilde özellikle dengesiz sınıflarda sıklıkla kullanılmaktadır.

$$\text{Duyarlılık (Recall)} = \frac{TP}{TP+FN} \quad (3.1)$$

$$\text{Kesinlik (Precision)} = \frac{TP}{TP+FP} \quad (3.2)$$

$$\text{Doğruluk (Accuracy)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.3)$$

$$F1 \text{ SKORU} = 2 \frac{\text{Duyarlılık} * \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}} \quad (3.4)$$

3.3.10 Alıcı işletim karakteristiği (receiver operating characteristic)

Alıcı İşletim Karakteristiği sınıflama yapılan bir modelin sınıflama performansını görsel olarak gösteren bir grafikdir. Dik koordinat sisteminde gösterilen bu eğrinin, bir ekseninde TP, diğer ekseninde ise FP oranları bulunmaktadır ve bu ikisi arasındaki sınıflama performansını koordinat sisteminde belirli eşikler göstererek

sunmaktadır. ROC eğrisinin altında kalan bölgeye ise AUC (Area Under the Curve) denmekte ve tüm koordinat sistemindeki kalan alan 1 olarak kabul edilirse, bu eğri alanı altında kalan oran modelin performansını göstermektedir. Bu değerler 1'e ne kadar yakın olursa, model o oranda iyi bir sınıflama becerisine sahip olduğu söylenebilir (Fawcett, 2006).

3.3.11 Model doğrulama yöntemleri (model validation methods)

Makine öğrenmesinin sınıflama başarılarının değerlendirilmesi için yapılan çalışmalara model doğrulama denmektedir (Tafralı, 2022). Bu aşamada eğitilen verilerin, yeni bir veri ile karşılaştığında bu verinin sınıflama performansını ölçülmesi için test seti ile karşılaştırılır. Kullanılan makine öğrenmesi yönteminin sınıflama becerisini test etmesinde farklı modeller kullanılmaktadır. Bunlardan başlıcaları sına seti (holdout) , k-katlamalı çapraz doğrulama (K-fold cross validation) ve birini dışarıda bırakma (leave-one-out) yöntemleridir. Sınama setinde verinin belirli bir bölümü eğitim seti olarak ele alınır ve eğitim seti kapsamında ele alınan veri grubu eğitilir. Eğitilme aşamasından sonra test seti olarak ayrılan veri grubu ile karşılaştırılır ve modelin performansı ölçülür. Örneğin bir veri grubunun %80'i eğitim, %20 si ise test seti şeklinde ayrılabilir ve eğitim ve test setleri karşılaştırılarak modelin başarı ölçüsü sınanır (Çolak, 2021). K-katlamalı çapraz doğrulama modelinde ise veri kümesi k kadar eşit parçaya bölünmektedir ve her parça için teker teker doğrulama işlemi gerçekleştirilmektedir. Örneğin k değeri 10 alındığında, veri kümesi 10 parçaya bölünmektedir ve her bir parça doğrulama verisi olarak kullanılmakta ve geriye kalan 9 parça eğitim verisi olarak kullanılır ve bu işlem 10 kez tekrar edilir. Modelin performansının ölçülmesinde ise her tekrar edilen işlemin ortalaması alınarak performans sonuçları ortaya çıkartılır (Öztürk, 2023). Birini dışarıda bırak modelinde ise, gözlem kümesindeki her bir gözlem her aşamada veri kümesinden ayrılır ve test edilir. Örneğin 1000 adet veri kümesi olduğu düşünülürse, veri kümesinden 1 tane örnek test seti olarak alınır ve geriye kalan 999 veri eğitilir ve bu şekilde veri setindeki her gözlem değerine test işlemi yapana kadar devam eder (Çolak, 2021).

4. BULGULAR

Bulgular bölümünde veri setine uygulanan yöntemlerden elde edilen bulgular belirtilmiştir. Veri seti oluşturulurken, otellerde konaklama yapan kişilerin otelle ilgili yazmış oldukları yorum, yorum başlığı, yorumu belirtirken otel ile ilgili 1’den 5’e kadar genel olarak otele vermiş olduğu puan, otelde kalmış olduğu tarih ve seyahat türü veri setine dahil edilmiştir. Şekil 4.1’de örnek olarak verisi çekilen bir yorum belirtilmiştir. Buna göre bu kişi 5 puan vererek ve yazmış olduğu metinde otelde geçirmiş olduğu bu deneyimde etkilendiği unsurları belirterek otelde geçirmiş olduğu deneyimden memnun olduğunu bildiren bir görüş yansıtmıştır. Bu kişi Nisan 2023’te konaklama yapmış ve çift olarak seyahat etmiştir. İlgili kişi 5 yıldızlı bir otelde konaklama yapmıştır.



Şekil 4.1 Araştırmada elde edilen örnek veri seti (Tripadvisor, 2023).

4.1 Veri Setine Ait Bilgiler

Bu bölümde araştırmada toplanan verilerin frekans değerleri ve oranları verilmektedir. Veri setindeki sunulan bilgiler kapsamında elde edilen veri seti Tablo 4.1’de otel sınıflarına göre gösterilmektedir. Buna göre çalışmada toplamda minimum 10 ve üzeri İngilizce yorumu sahip olan 353 otelin İngilizce yorumları çekilmiştir. Bu otellerin %41’i 3 yıldızlı oteller oluşturmaktadır ve veri setindeki en büyük paya sahiptir. Bunu sırasıyla 4, 5 ve 2 yıldızlı oteller izlemektedir. Tripadvisor’da 1 yıldızlı ve İngilizce yorumu 10’un üzerinde olan yorumu Mayıs 2023 tarihi itibari ile ilgili web sitesinde bulunamadığı için, tek yıldızlı oteller veri setine dahil edilememiştir. İngilizce

yorum sayısı oranına bakıldığında ise, tüm yorumların %27'sini oluşturmaktadır ve toplamda 40.691 veri araştırmada kullanılmak üzere hazırlanmıştır. İngilizce yorum sayılarının oranına bakıldığında ise her otel sınıfında dengeli bir oranın olduğu görülmektedir. Yorum sayısı ise, otellerin oda kapasiteleri göz önüne alındığında, normal olarak otel sınıfı arttıkça otellerin alabilecekleri kapasiteden dolayı yorum sayısı da artmaktadır.

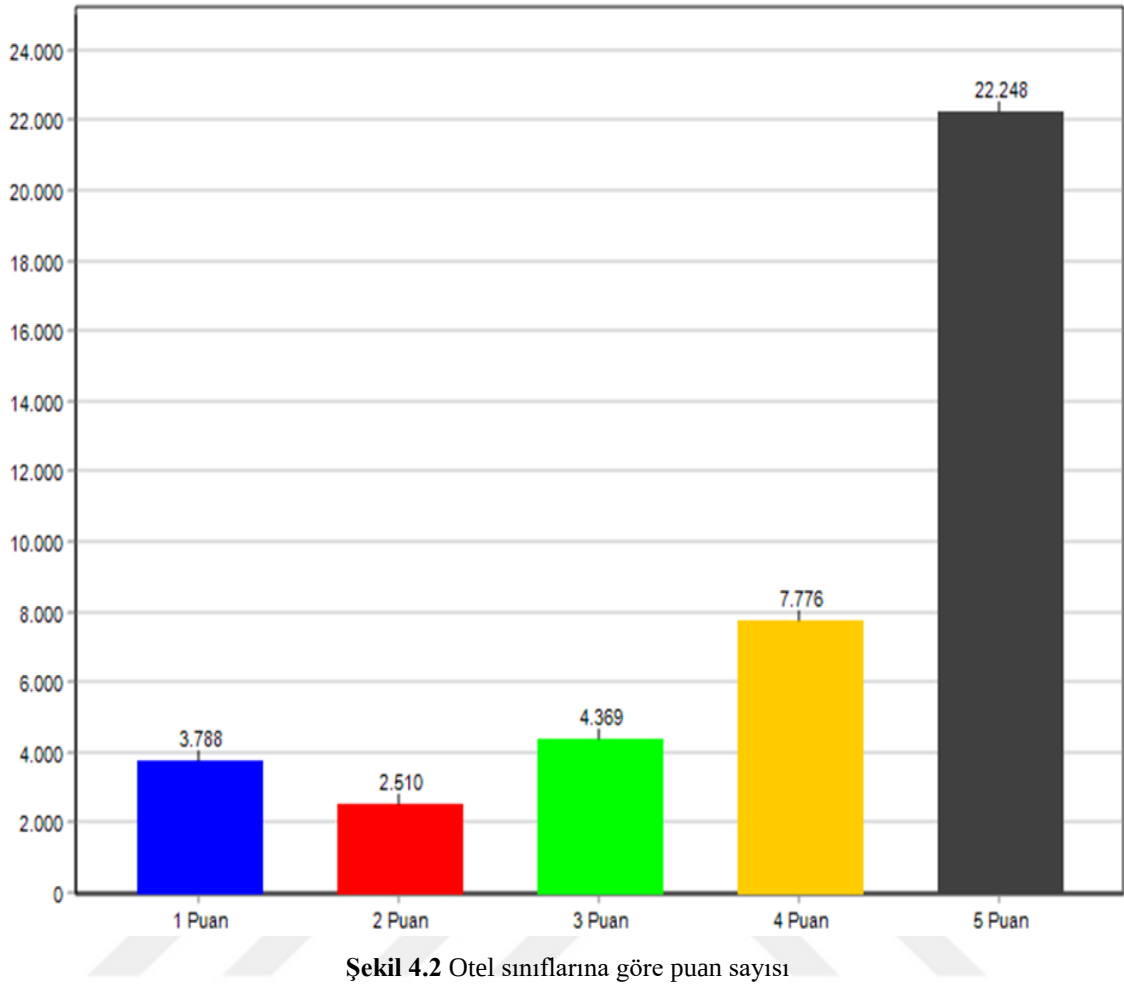
Tablo 4.1. Otel sınıflarına göre veri setindeki bilgiler

Otel Sınıfı	Otel Sayısı	Otel Yüzdesi	Tüm Yorum Sayısı	İngilizce Yorum Sayısı	İngilizce Yorum Yüzdesi
2 Yıldız	37	10	1775	505	28
3 Yıldız	143	41	19139	5433	28
4 Yıldız	97	27	52096	12281	24
5 Yıldız	76	22	87193	22472	26
Toplam	353	100	150203	40691	27

Tablo 4.2 ve şekil 4.2' de ise veri setindeki tüm otellerin, konaklama yapan kişilerin her yorumda yapmış olduğu genel değerlendirmelerin sayısal olarak gösterimi bulunmaktadır. İlgili tabloya göre konaklama yapan kişilerin yorumlarında en fazla 5 puan verdikleri görülmektedir ve bu kişilerin %54.7'si 5 puan ve %19.1'i 4 puan vererek, yaklaşık olarak % 73.8' i olumlu bir görüş bildirdiği görülmektedir. Elde edilen yorumların %9.3'ü 1 ve %6.2' si ise 2 puan vererek, olumsuz anlamda puan veren kişi oranı %15.5 olarak görülmektedir. 3 puan veren kişilerin ise tüm yorumlardaki ortalamaları %10.7 olarak görülmüştür ve 2 puan veren kişilerin oranı ile 4 puan veren kişilerin arasındaki farkın daha fazla olduğu görülmektedir. Bu durumda, bir kişi yorum yaptığında negatif bir yorumun nötr olma durumu, nötr bir yorumun pozitif olma durumuna göre daha az görüldüğü söylenebilir.

Tablo 4.2 Değerlendirme puanı ve yorum sayısı

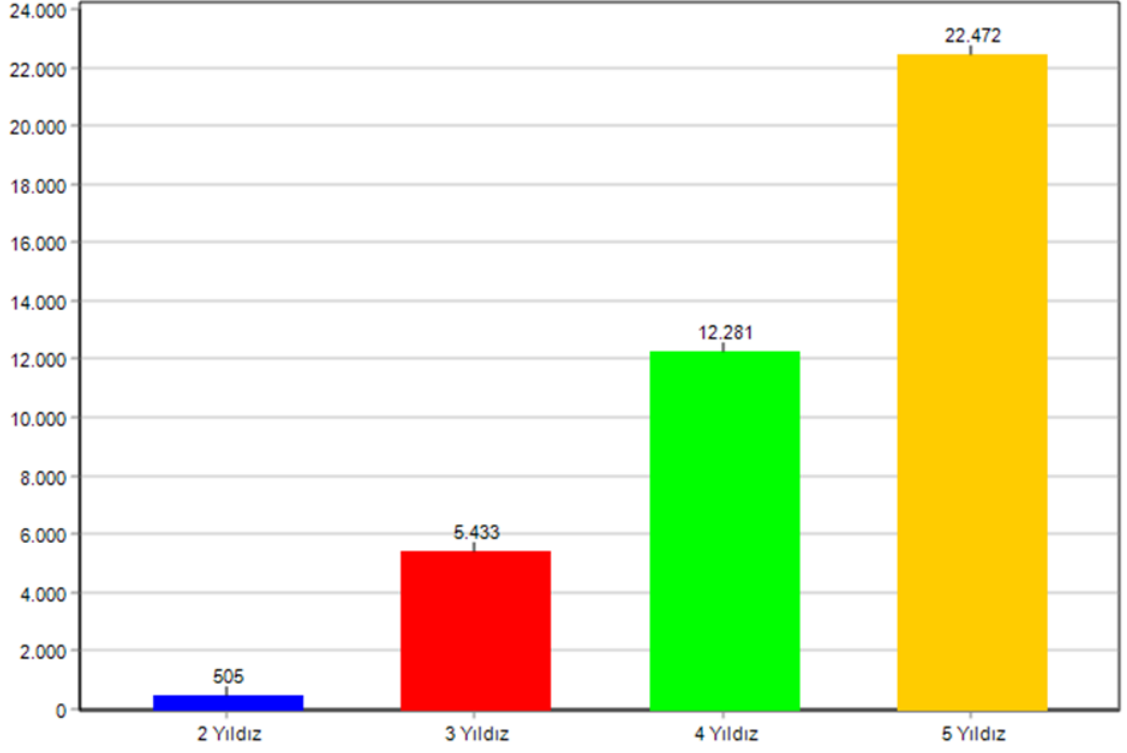
Değerlendirme Puanı	Yorum Sayısı	Yorum Yüzdesi
1 Puan	3788	9.3
2 Puan	2510	6.2
3 Puan	4369	10.7
4 Puan	7776	19.1
5 Puan	22248	54.7
Toplam	40691	100



Tablo 4.3’ ve şekil 4.3 de otel sınıflarına göre yapılan yorum sayıları gösterilmektedir. Buna göre tüm yapılan yorumların en fazla oranı 5 yıldızlı otellere yapıldığı görülmektedir. Otel sınıfı arttıkça, yorumların oranları da artmaktadır. 2 yıldızlı otellerin gerek sayısı, gerek kapasitesi ve bu otellerin genel İngilizce yorum sayısına bakıldığında, veri setinde az miktarda bulunmakta ve bu durum doğal olarak bu otellerin İngilizce yorum sayılarına da yansıdığı görülmektedir.

Tablo 4.3 Otel sınıflarına göre yorum sayısı ve yüzdeleri

Otel Sınıfı	Yorum Sayısı	Yorum Yüzdesi
2 Yıldız	505	1,24
3 Yıldız	5433	13,35
4 Yıldız	12281	30,18
5 Yıldız	22472	55,23
Toplam	40691	100

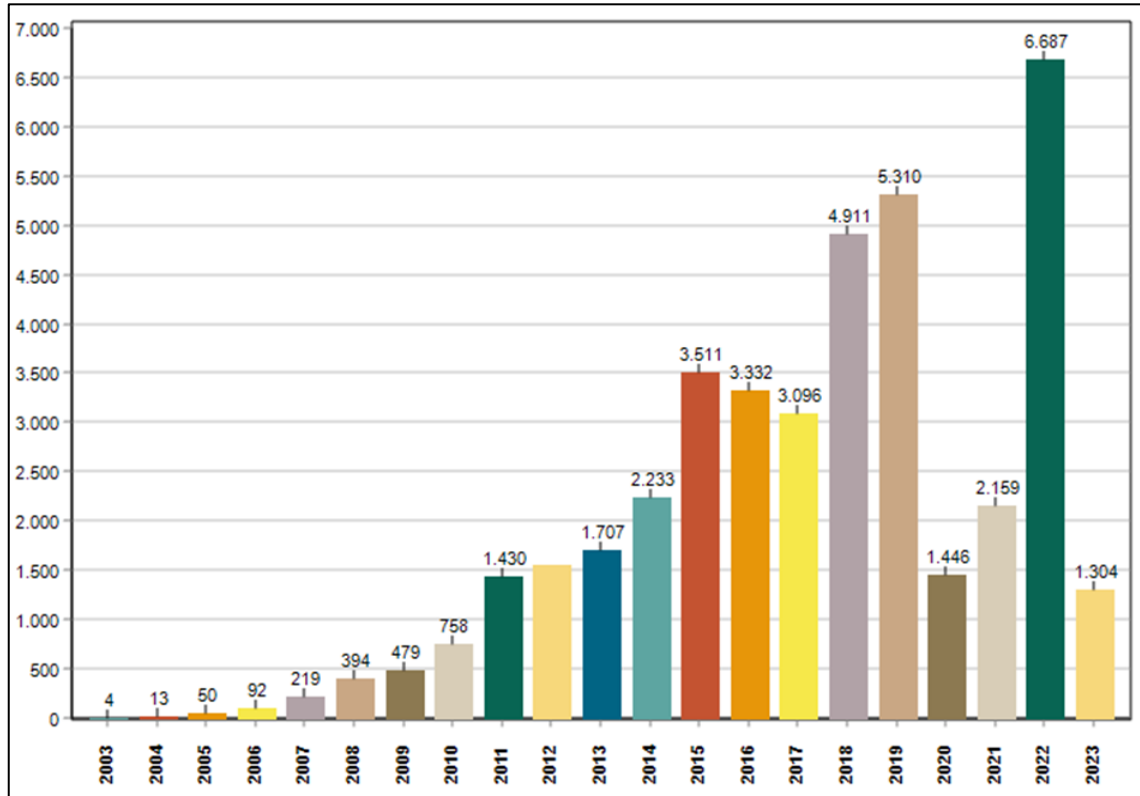


Şekil 4.3 Otel sınıflarına göre yorum sayısı

Tablo 4.4 ve Şekil 4.4 de tarihlere göre yapılan yorumlar ve bu yorumların tüm yorumlardaki oranları tarihsel olarak verilmiştir. Tripadvisor 2000 yılında kurulan bir site olduğu için, doğal olarak veri setindeki elde edilen verilerin başladığı tarih olan 2003 yılı ve bu yılı takip eden yıllarda yorum sayısının düşük olduğu görülmüştür. Yıllar içerisinde yorum sayılarında giderek artış göstermektedir. Pandemi döneminde ise bu artışta düşüş yaşandığı net olarak görülmektedir. Örneğin 2019 yılı, veri setindeki tüm yorumların %13,04'ünü oluştururken, 2020 yılında ise %3,67'sini oluşturmaktadır. Buna ek olarak 2015 yılında Rusya ile yaşanan krizin neticesinde de yorum sayısında bir önceki yıla göre 2016 ve 2017 yıllarında düşüş görülmektedir. Ancak 2022 yılında tekrardan bir artış göstermiş ve bu yıl en fazla yorumun olduğu yıl olarak karşımıza çıkmaktadır. Yıl bazındaki yorumların artışına bakıldığında, bu durum hem Tripadvisor'un hem de Alanya'daki otellerin sayısal ve buna bağlı olarak kapasite artışını ve bu otellerin ilgili siteye olan taleplerinin de artış gösterdiğini işaret ettiği söylenebilir.

Tablo 4.4 Tarihlere göre yorum sayısı ve yüzdeleri

Yıllar	Yorum Sayısı	Yorum Yüzdesi	Yıllar	Yorum Sayısı	Yorum Yüzdesi
2003	4	0,01	2015	3511	8,63
2004	13	0,03	2016	3332	8,19
2005	50	0,12	2017	3096	7,61
2006	92	0,23	2018	4911	12,07
2007	219	0,54	2019	5310	13,05
2008	394	0,97	2020	1446	3,55
2009	479	1,18	2020	1446	3,68
2010	758	1,86	2021	2159	5,31
2011	1430	3,51	2021	2159	5,34
2012	1556	3,82	2022	6687	16,43
2013	1707	4,20	2023	1304	3,20
2014	2233	5,49	Toplam	40691	100



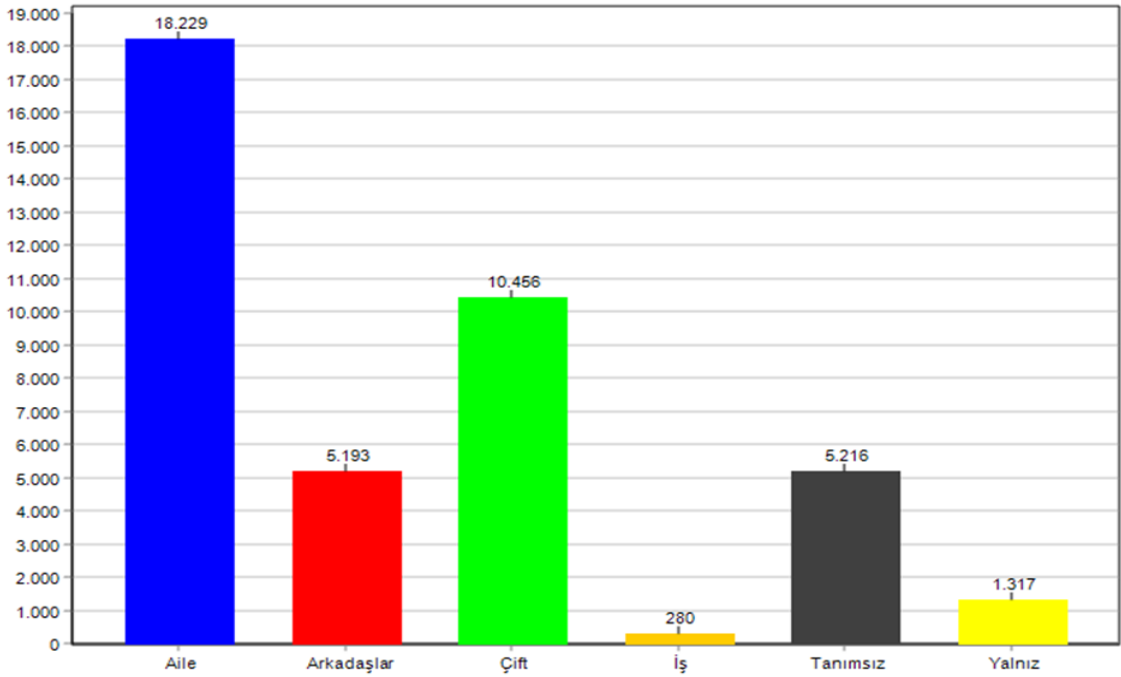
Şekil 4.4 Tarihlere göre yorum sayıları

Tablo 4.5 ve Şekil 4.5’de veri setindeki yorumda bulunan kişilerin seyahat türleri gösterilmektedir. Tripadvisor’ un vermiş olduğu seçeneklerde kişiler, yorumda bulunurken seyahat tipini de seçebilmektedirler. Ancak bu durum zorunlu bir seçenek değildir. Dolayısıyla veri setindeki bazı kişilerin seyahat türünü belirtmediği görülmektedir. Ancak veri setindeki çoğu kişi seyahat tipini belirtmiş olup, bu kişilerin en fazla oranı aile olarak görülmektedir. Genel olarak seyahat tiplerine bakıldığında ise, aile ve çift olarak seyahat eden kişilerin Alanya’ daki otelleri ağırlıklı olarak seçtiği

görülmektedir. Yalnız ve iş dolayısıyla seyahat eden kişi sayısı ise az sayıda olarak ortaya çıkmıştır.

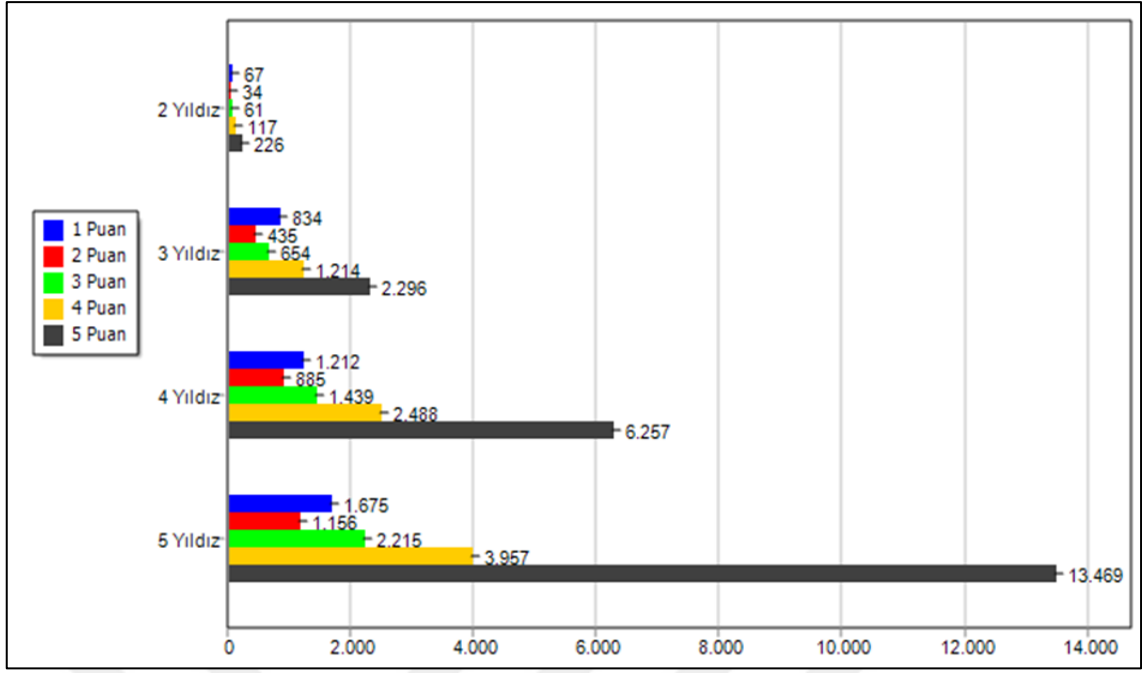
Tablo 4.5 Seyahat türlerine göre yorum sayıları ve yüzdeleri

Seyahat Tipi	Yorum Sayısı	Yorum Yüzdesi
Aile	18229	44.79
Çift	10456	25.69
Tanımsız	5216	12.81
Arkadaşlar	5193	12.76
Yalnız	1317	3.23
İş	280	0.68
Toplam	100	40691



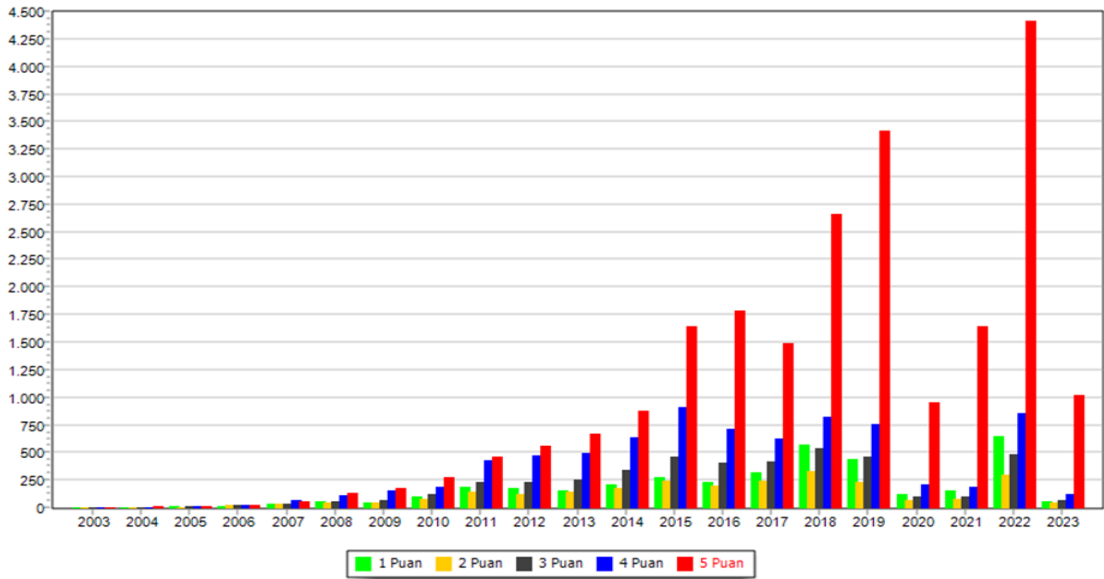
Şekil 4.5 Seyahat tiplerine göre yorum sayıları

Şekil 4.6’da ise otel sınıflarına göre konaklama yapan misafirlerin değerlendirme puanları gösterilmektedir. Puanlara bakıldığında, otel sınıfı azaldıkça, insanların vermiş oldukları değerlendirmeler arasındaki farkın açıldığı görülmektedir. Özellikle 5 yıldızlı otel grubuna verilen 5 puan, diğer otel sınıflarına ile karşılaştırıldığında aradaki farkın çok fazla olduğu görülmüştür. Bu kapsamda otel sınıfının ve doğal olarak otellerin sunmuş olduğu imkânlar, insanların otel deneyimindeki memnuniyetini etkileyen bir faktör olarak görülmektedir. Bu araştırmadaki amaçlardan birisi de bu farkın, otelin hangi unsurlarından kaynaklandığını ortaya çıkarmaktır.



Şekil 4.6 Otel sınıflarına göre puan dağılımı

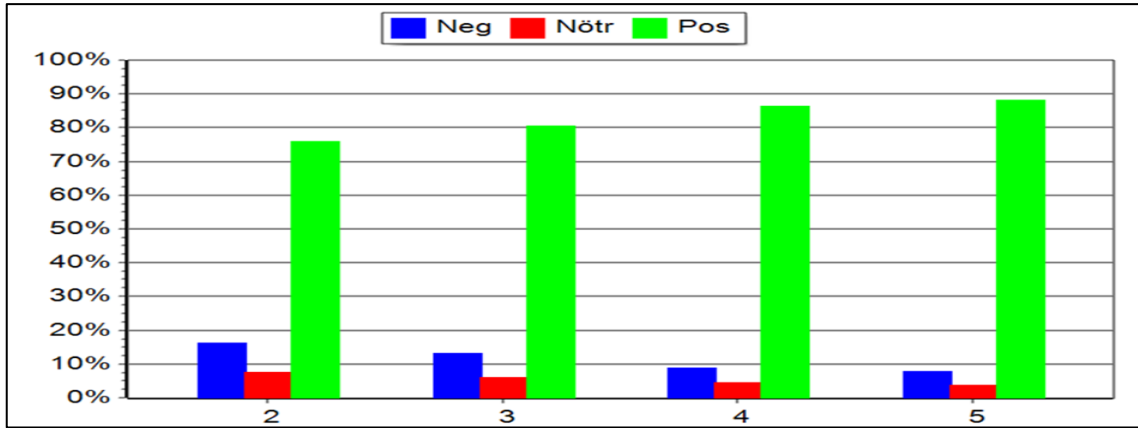
Şekil 4.7’de tarihe göre konukların vermiş oldukları puanlar görülmektedir. Yatay ekseninde yıllar, dikey ekseninde ise yorum sayısı bulunmaktadır ve her puanı bir renk temsil etmektedir. Buna göre günümüze yaklaştıkça konukların vermiş oldukları puanların arasındaki farklılık giderek artmaktadır. Özellikle 5 puan veren kişilerin diğer puan veren grupla arasındaki farka bakıldığında 2023 yılına yaklaştıkça bu farkın açıldığı görülmektedir ve beğeni oranı giderek artmaktadır.



Şekil 4.7 Yıllara göre otel sınıflarına yapılan puanların dağılımı

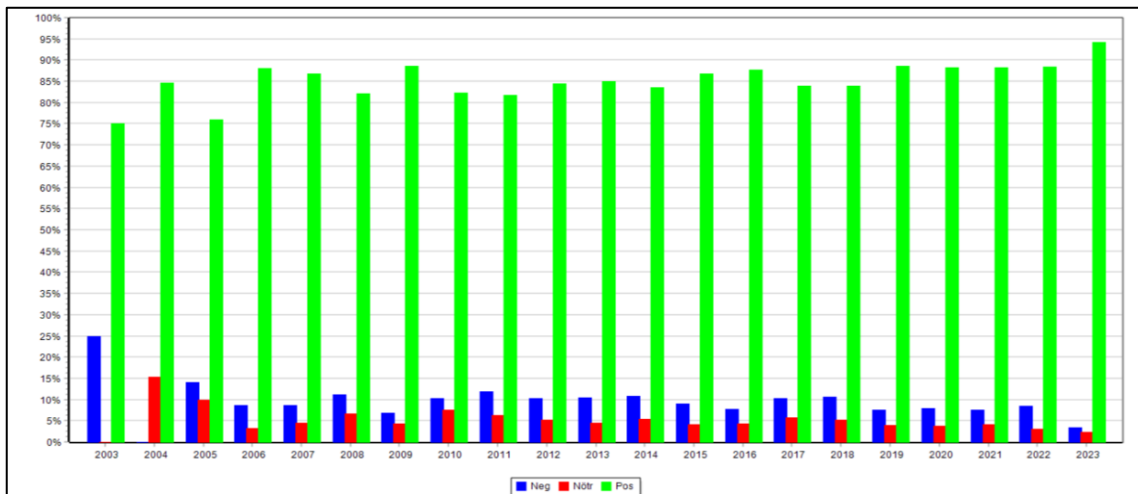
4.2. Duygu Analizi Sonuçları

Şekil 4.10’da duygu analizi sonuçlarının otel sınıflarına göre yüzdesel dağılımı verilmiştir. Buna göre otel sınıfı arttıkça, konaklama yapan kişilerin yapmış olduğu değerlendirmelerde pozitif yorum oranı, negatiflere göre artmakta olduğu gözlemlenmektedir. Otel sınıfı azaldıkça, duygu sınıflarının birbirleri ile olan yüzdesel yakınlığı artmaktadır.



Şekil 4.10 Otel sınıflarına göre duygu analizi sonuçları

Şekil 4.11’de duygu sınıflarının tüm otel sınıflarına ait verilerin tarihe göre yüzdesel dağılımı verilmektedir. Verilerin dağılımına bakıldığında, 2023 yılına yaklaşıldıkça pozitif, negatif ve nötr yorumların aralarındaki farkın giderek pozitifte doğru gittiği görülmektedir. Konaklama yapan ve deneyimlerini paylaşan kişilerin yorumlarının giderek pozitif yöne gittiği ortaya çıkmaktadır.



Şekil 4.11 Yıllara göre duygu analizi sonuçları

4.2.1 Duygu analizi sınıflandırılması

VADER yöntemi ile yapılan duygu analizi sonuçları, makine öğrenmesi modelleri ile test edilmiştir. Sınıflamada öncelikle, duygu analizi sonucunda çıkan compound skoruna göre her bir cümle pozitif, nötr ve negatif olarak gruplanmıştır ve bu gruplar hedef verisi olarak seçilmiştir. Özellik veri setinde ise her bir yorum ele alınmıştır ve bu yorumlarda ele alınan kelimeler, dokuman gömme yöntemi ile vektörleştirilmiştir ve bu şekilde sınıflamaya tabi tutulmuştur.

Veri setinin büyüklüğü ve analiz yapılan bilgisayarın işletim sisteminin gücü kapsamında, veri setinin öncelikli olarak %25' i seçilmiştir ve bu seçilen verilere eğitim ve test işlemi gerçekleştirilmiştir. Verilerin sınıflama işlemi, sinir ağları, lojistik regresyon, gradyan arttırma, stokastik gradyan iniş, rastgele orman, karar ağacı, adaptif arttırma, destek vektör makinesi, naif bayes ve k-en yakın komşu yöntemleri kullanılarak yapılmıştır. Metin setindeki pozitif, nötr ve negatif yorumların %25'i ele alınarak gerçekleştirilen ilk eğitim ve test setinde, verilerin %80'i eğitim, %20'si ise test seti için kullanılmıştır. Tablo 4.6'da model sonuçları, doğruluk değerinin büyüklüğüne göre sıralı şekilde görülmektedir. Sınıflama sonucuna göre en iyi sınıflama becerisini sinir ağları modeli oluşturmaktadır.

Tablo 4.6 Sınama seti ile duygu sınıflarının sınıflama performansı

Model	Area Under Curve	Accuracy	F1	Precision	Recall
Neural Networks	0.985	0.958	0.958	0.957	0.958
Logistic Regression	0.982	0.954	0.954	0.955	0.954
Gradient Boosting	0.987	0.953	0.949	0.948	0.953
Stochastic Gradient Descent	0.85	0.943	0.937	0.935	0.943
Random Forest	0.96	0.934	0.924	0.928	0.934
Decision Tree	0.669	0.904	0.892	0.886	0.904
Adaptive Boosting	0.8	0.894	0.898	0.902	0.894
Support Vector Machine	0.574	0.864	0.809	0.786	0.864
Naive Bayes	0.937	0.778	0.836	0.935	0.778
K- Nearest Neighbors	0.878	0.657	0.744	0.916	0.657

Sınıflama ayrıca ilgili makine öğrenmesi yöntemleri kullanılarak k-katlamalı çapraz doğrulama yöntemi ile de test edilmiştir Tablo 4.7'de gösterilmiştir. Sınıflama becerilerine göre 10-katlı çapraz doğrulama, Tablo 4.6'da gösterilen sınama setindeki değerlere göre az farkla daha iyi sonuç verdiği görülmüştür ve bu yöntemde de en iyi doğrulama sonucunu sinir ağları göstermektedir.

Tablo 4.7 K-Katlamalı doğrulama ile duygu sınıflarının sınıflama performansı

Model	Area Under Curve	Accuracy	F1	Precision	Recall
Neural Networks	0.987	0.958	0.957	0.957	0.958
Logistic Regression	0.982	0.956	0.956	0.957	0.956
Gradient Boosting	0.987	0.954	0.951	0.95	0.954
Stochastic Gradient Descent	0.852	0.945	0.938	0.936	0.945
Random Forest	0.959	0.936	0.927	0.93	0.936
Decision Tree	0.79	0.919	0.916	0.914	0.919
Adaptive Boosting	0.807	0.897	0.901	0.905	0.897
Support Vector Machine	0.571	0.864	0.806	0.777	0.864
Naive Bayes	0.939	0.783	0.839	0.935	0.783
K- Nearest Neighbors	0.863	0.68	0.756	0.904	0.68

Şekil 4.12 ve 4.13’de sinir ağırları yöntemi ile test edilen hedef verilerin sınıflama becerileri sayısal ve yüzdesel olarak görülmektedir. Görüldüğü gibi pozitif olan cümlelerin sınıflandırılmasında sinir ağırları %98.5 oranında etkili olmaktadır. Negatif yorumlarda ise %82.6 oranında başarılı bir sınıflama gerçekleştirmektedir. Nötr grubunda olan yorumların sınıflandırılmasında ise %66.3 oranında başarı sağladığı görülmektedir. Nötr duygu sınıfına sahip olan yorumlarda, yorum içerisinde hem negatif hem de pozitif duygu barındıran kelimelerin dağılımı doğal olarak pozitif ve negatiflere göre daha fazla olduğundan, nötr sınıfındaki yapılan sınıflama başarısı daha düşük düzeyde olabilmektedir. (Ray vd., 2021).

		Predicted			
		Neg	Nötr	Pos	Σ
Actual	Neg	728	112	64	904
	Nötr	102	287	65	454
	Pos	51	34	8730	8815
Σ		881	433	8859	10173

Şekil 4.12 Karmaşıklık matrisi sayısal gösterimi

		Predicted			Σ
		Neg	Nötr	Pos	
Actual	Neg	82.6 %	25.9 %	0.7 %	904
	Nötr	11.6 %	66.3 %	0.7 %	454
	Pos	5.8 %	7.9 %	98.5 %	8815
Σ		881	433	8859	10173

Şekil 4.13 Karmaşıklık matrisi yüzdesel gösterimi

Veri setinin %25'i alınarak yapılan sınıflama çalışmasında en iyi sonucu 10-katlı çapraz doğrulama yöntemi ve sinir ağları algoritması verdiği için, bu model veri setinin %50'sine de uygulanmıştır ve sonucunda doğruluk ve f1 değeri 0,93 oranında verileri doğru tespit ettiği görülmektedir. Şekil 4.14 ve 4.15'de de bu modelin karmaşıklık matrisleri oluşturulmuş ve sayısal ve yüzdesel olarak gösterilmiştir. Modelde veri sayısı arttıkça, özellikle nötr grubundaki yorumların sınıflama becerisi düştüğü, pozitif olan yorumlarda ise yaklaşık olarak aynı kaldığı, negatiflerde ise daha az oranda düştüğü görülmektedir. Tablo 4.8'de ise yapılan duygu analizinin genel sınıflama sonucu gösterilmektedir.

Tablo 4.8 Duygu analizi sınıflama sonucu

Model	Area Under Curve	Accuracy	F1	Precision	Recall
Neural Network	0.973	0.935	0.933	0.932	0.935

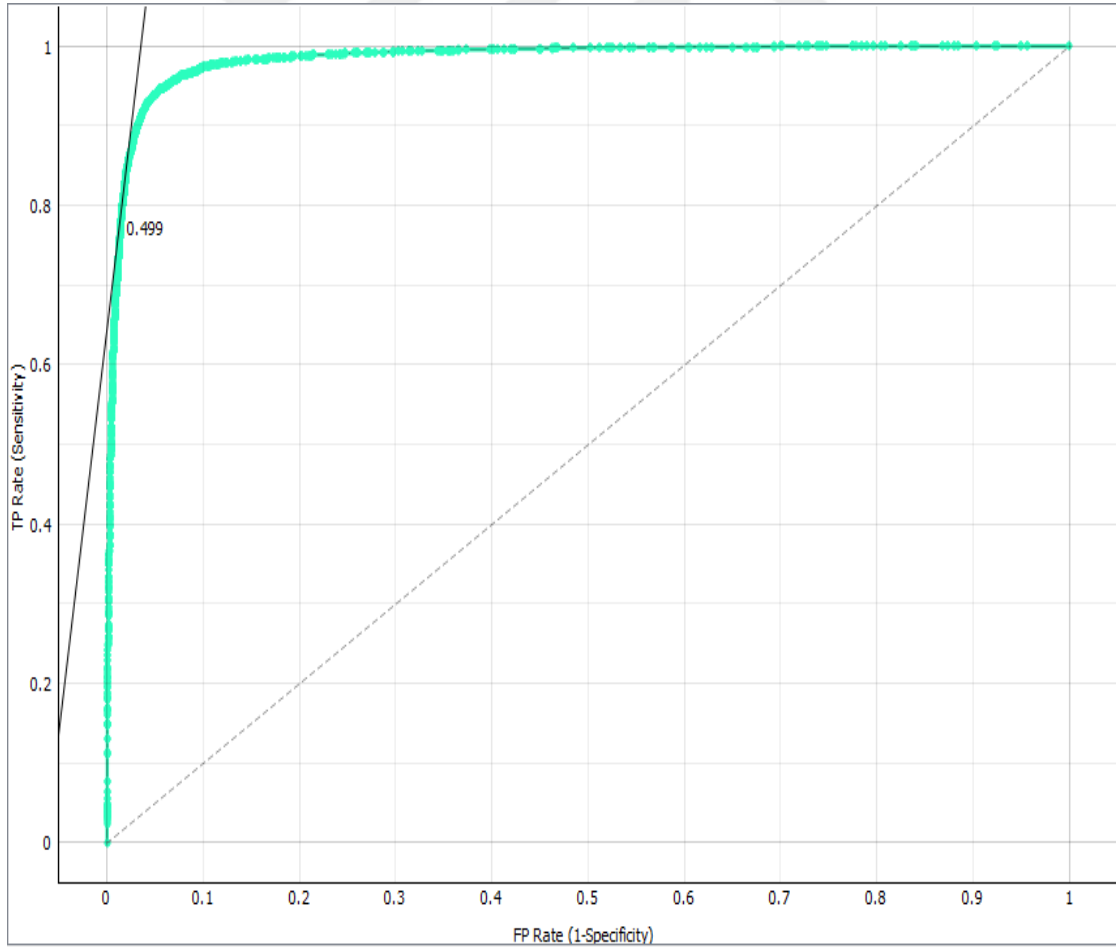
		Predicted			Σ
		Neg	Nötr	Pos	
Actual	Neg	1288	283	237	1808
	Nötr	330	385	193	908
	Pos	161	121	17347	17629
Σ		1779	789	17777	20345

Şekil 4.14 Duygu analizi karmaşık matrisi sayısal gösterimi

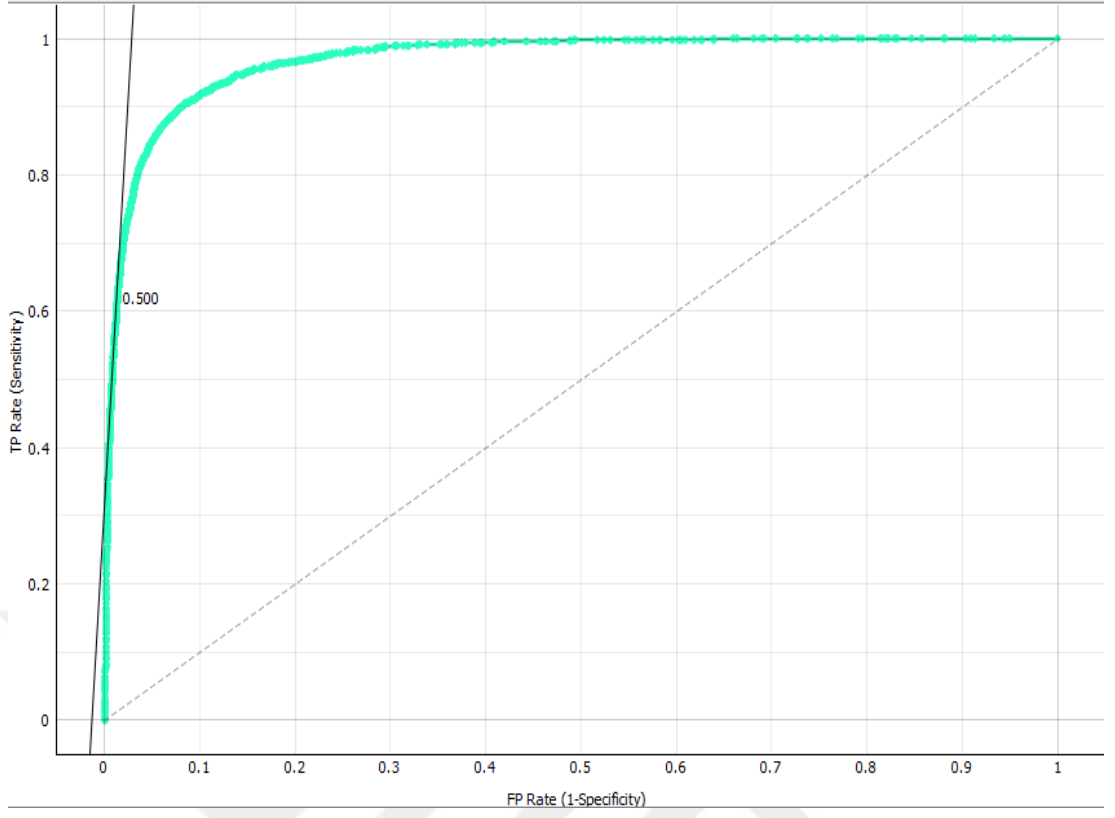
		Predicted			Σ
		Neg	Nötr	Pos	
Actual	Neg	71.2 %	15.7 %	13.1 %	1808
	Nötr	36.3 %	42.4 %	21.3 %	908
	Pos	0.9 %	0.7 %	98.4 %	17629
Σ		1779	789	17777	20345

Şekil 4.15 Duygu analizi karmaşıklık matrisi yüzdesel gösterimi

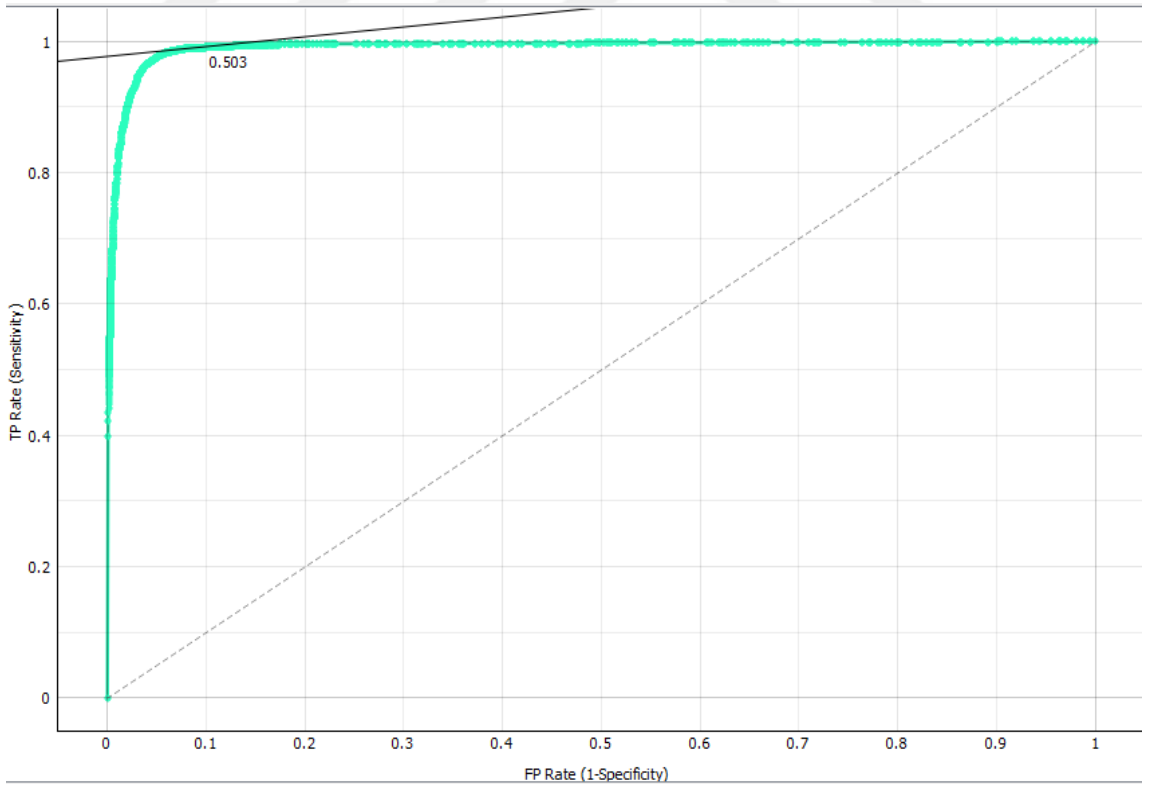
Şekil 4.16, 4.17 ve 4.18’de veri setinin %50’si kullanılarak yapılan sınıflamada, ROC eğrileri sırasıyla negatif, nötr ve pozitif duygu sınıflarının sınıflama becerisini göstermektedir. Buna göre duygu sınıflarındaki sınıflama performansı en iyi düzeyde pozitif, daha düşük düzeyde negatif, en düşük düzeyde ise nötr grubunda olmakta olduğu görülmektedir.



Şekil 4.16 Negatif duygu sınıflarının ROC eğrisi



Şekil 4.17 Nötr duygu sınıflarının ROC eğrisi



Şekil 4.18 Pozitif duygu sınıflarının ROC eğrisi

4.3 Konu Modelleme Sonuçları

Konu modelleme analizi yapılırken, öncelikli olarak perplexity (çapraşıklık) ve coherence (uyumluluk) sonuçlarına göre ve konulara göre dağıtılan kelimelerin anlamsal bütünlüğü sağlanacak şekilde konu sayısı belirlenmektedir. Bu aşamada konu modelleme analizlerinde konuların çapraşıklık ve uyumluluk değerleri, konu sayısına göre teker teker denenerek elde edilmektedir. Bu çalışmada konu sayısı 6'dan 200'e kadar her bir konu sayısına göre çapraşıklık ve uyumluluk değerleri bulunmuştur. Çapraşıklık sonuçları Şekil 4.19'da, uyumluluk sonuçları ise Şekil 4.20'de belirtilmiştir. Çapraşıklık sonuçlarının düşük, uyumluluk sonuçlarının ise yüksek olması konu modellemede önemli bir basamağı teşkil etmesine karşın, konuların anlamsal bütünlüğünün de korunması ve değerlendirilmesinin uygun olması da önemli bir unsurdur.

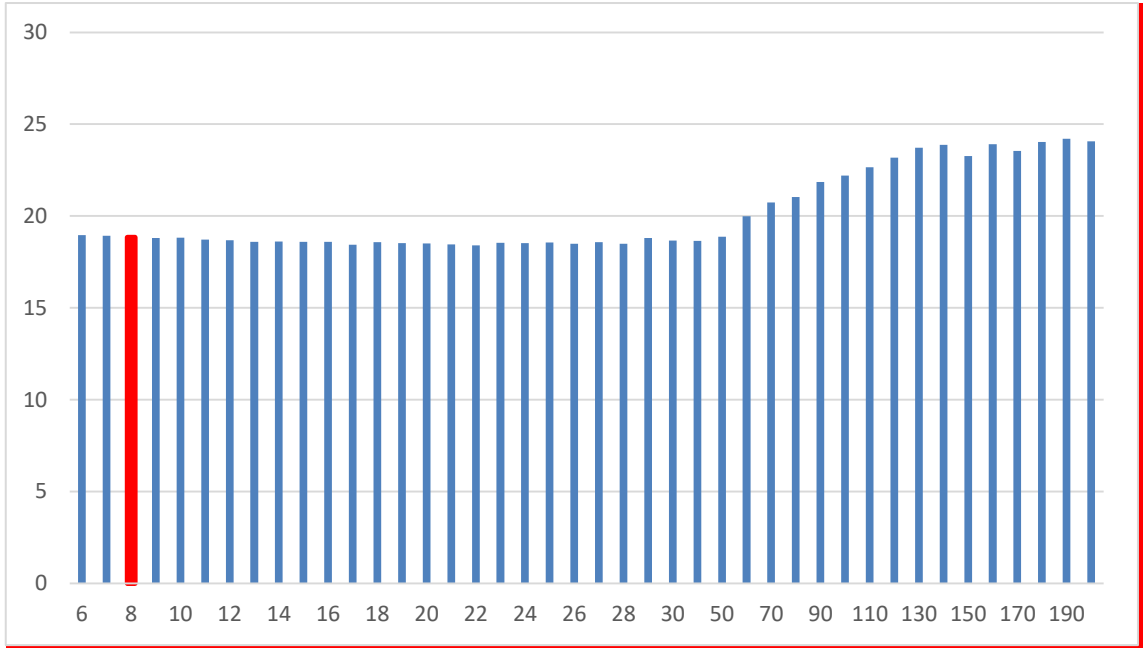
Araştırmada öncelikle konu sayısı üç dört ve beş denenmiştir ancak bu sayıdaki konularda meydana gelen sorun, veri setinde çok fazla geçen kelimelerin her konuda bulunması olmuştur. Bu sebeple konu sayısı altıdan başlanmış ve ikiyüze kadar denenmiştir. Görüldüğü gibi konu sayısı elliden fazla olduğunda, çapraşıklık değeri yükselmektedir. Aynı zamanda konu sayısı altı ile elli arasında seçildiği zaman ise çapraşıklık benzer değerleri almaktadır.

Uyumluluk değerlerine bakıldığında ise, konu sayısı 9' dan sonra arttığında uyumluluk skorunun düştüğü, 100' den sonra ise benzer seviyeye geldiği görülmüştür. Konuların değerlendirilme imkanı göz önüne alındığında ve Bu sebeple, hem çapraşıklık ve uyumluluk değeri, hem de konuların içeriklerinin değerlendirilmesi uygunluğu kapsamında araştırmada 8 tane konu seçilmiş olup, bu konular ile ilgili olarak her konunun içermiş olduğu kelimelerin özelliklerine göre konu başlıkları atanmıştır.

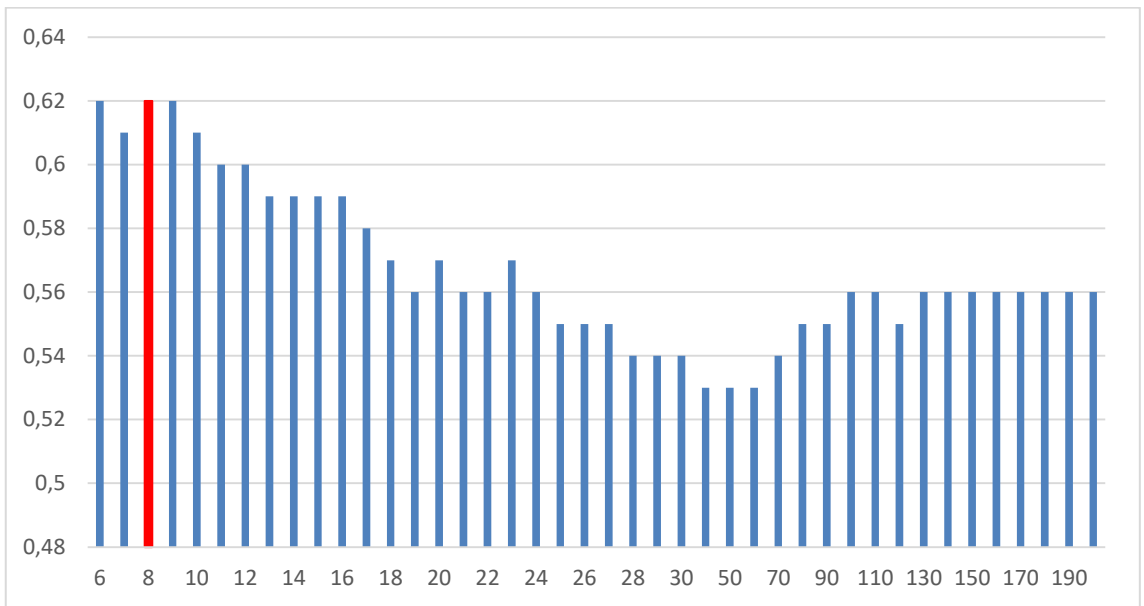
Konaklama işletmelerinde deneyimlerini sosyal mecralarda paylaşan bireyler, bilindiği gibi gerek konaklama yaptığı tesislerin gerekse tesisin bulunduğu destinasyonlarla ilgili yapmış oldukları yorumlarında birçok farklı kelimeyi aynı yorumda konumlandırmaktadır. Örneğin great kelimesi, odalar, animasyon, personel davranışı, otelin içerdiği olanaklar ve tesis özellikleri, otelin sağlamış olduğu yiyecek içecek hizmetleri gibi birçok farklı konuları betimleyen kelimeler ile kullanılmaktadır.

Bu sebeple konu modellemede oluşturulan her kümede bu kelimeler geçmektedir. Önemli olan durum, bu kelimelerin hangi kelime grupları ile bir arada

geçtiğinin ortaya konulmasıdır. Konu modellemede önemli olan unsurlardan birisi, bu gibi kelimelerin bir arada hangi özel kelimeler ile kullanıldığı ve değerlendirmenin uygun bir şekilde yapılmasıdır. bu araştırmada elde edilen konu modelleme çalışmasında, toplamda sekiz farklı konu elde edilmiştir ve bunlar kendi içeriklerine göre başlıklandırılmıştır. Başlıklandırma, her araştırmada olduğu gibi konular içerisindeki olan kelimelerin sağlamış olduğu anlam bütünlüğüne göre araştırmacılar tarafından oluşturulmaktadır.



Şekil 4.19 Konu modelleme çapraşıklık sonuçları

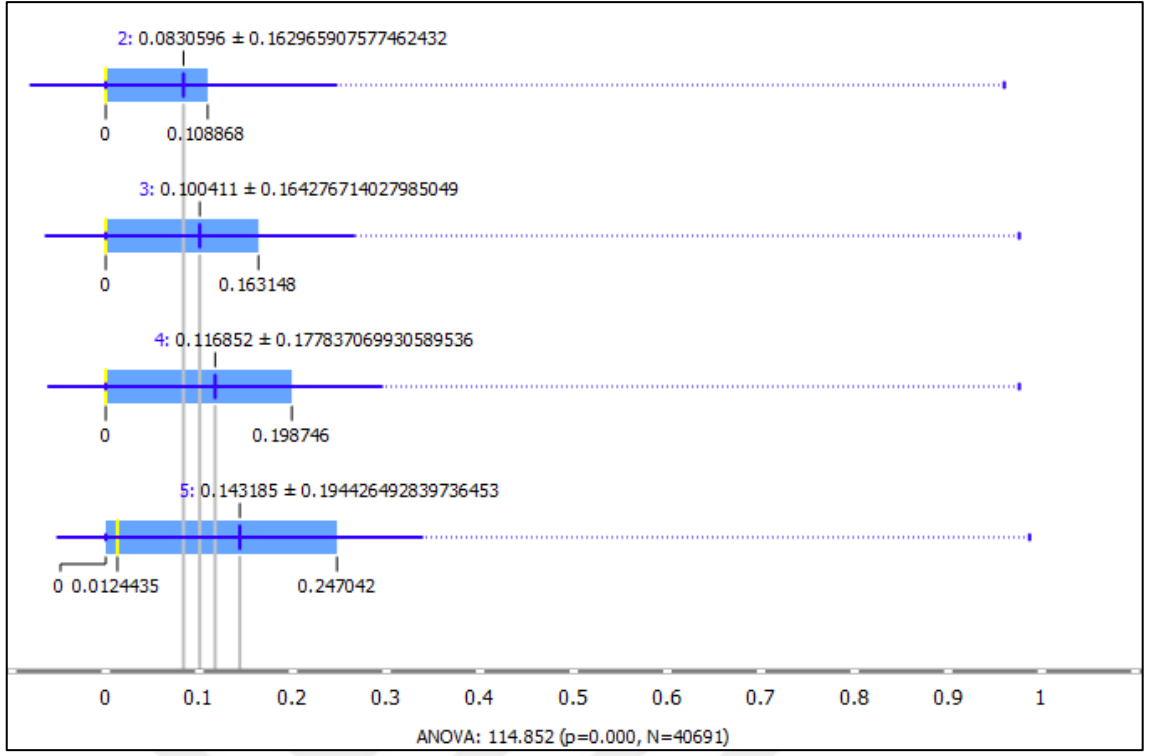


Şekil 4.20 Konu modelleme uyumluluk sonuçları

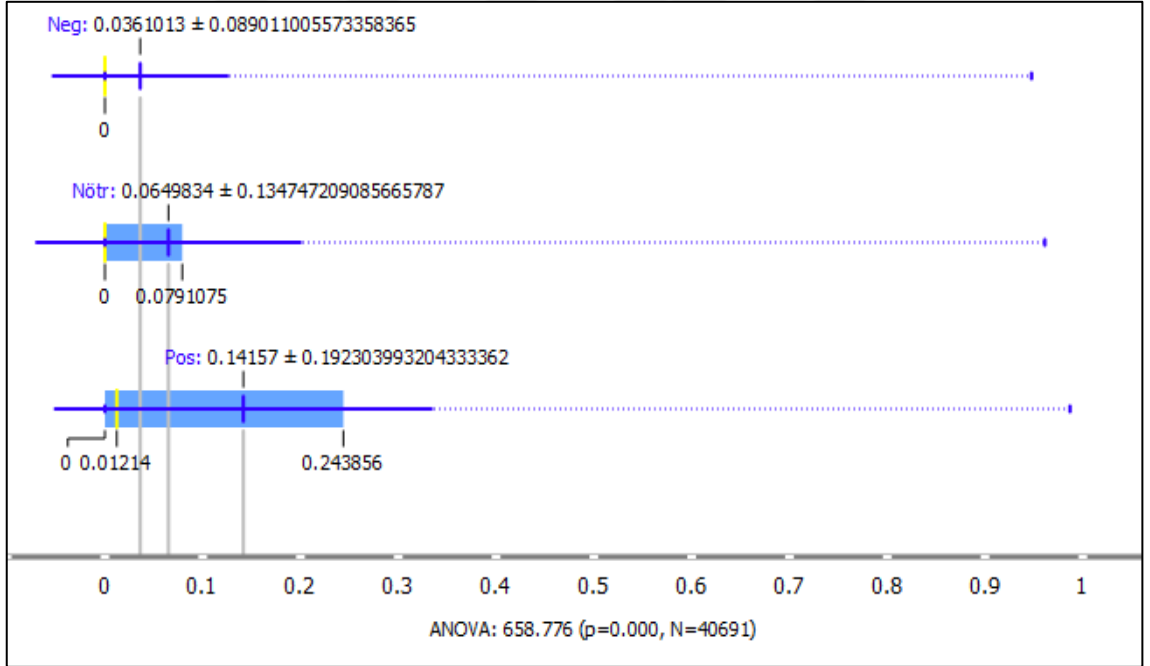
Konu modelleme çalışması sonucunda elde edilen konular, 4.26 ile 4.34 numara arasındaki şekillerde gösterilmektedir. LDA analizi sonucunda meydana çıkan konuların ilki olanaklar başlığı altında toplanmıştır. Konu modellemede ortaya çıkan konular, her bir kümeyi göstermektedir ve bu kümede birlikte kullanılan kelimeler ifade edilmektedir. Her konu kümesinde, gri renkle gösterilen çubuklar, o kelimenin metin setinde geçmiş olduğu frekansı göstermektedir. Kırmızı renk ile gösterilen ve gri çubuğun içerisinde bulunan kesim ise, o kelimenin kaç tanesinin bu kümedeki kelimeler ile beraber olduğunu göstermektedir. Örneğin şekil 4.26' ya bakıldığında, resort kelimesi, tüm veri setinde yaklaşık olarak 10.500 kez geçmekte iken, bu kelimenin yaklaşık olarak 9000 adedi 1. Konudaki bulunan kelimeler ile kullanılmaktadır. Her konuda, veri setindeki birçok kelime farklı frekans değerlerinde bulunmaktadır. Bu çalışmada, her konuda en fazla ortaya çıkan ilk 20 kelime gösterilmiştir. Her konuda, veri setindeki farklı kelimeler bulunmaktadır ancak bu kelimelerin bu konudaki kelimeler ile kullanımı, doğal olarak her konudaki kelimelere göre değişiklik göstermektedir ve bu durum da bu konularda geçen kelimelerin sıklık sırasını değiştirmektedir. Örneğin resort kelimesinin geri kalan kısmı ise diğer 7 konuda geçmektedir ancak bu kelimenin diğer konudaki kelimeler ile beraber kullanımı, 1. konuya göre daha az seviyededir. Bu sebeple konulardaki kelimeler, hem metin seti içerisindeki kullanımı hem de yorum içerisindeki kullanımına göre her konuda frekans olarak değişiklik göstermektedir. Frekans değerleri oluşturulurken, TF*IDF değerleri her kelimenin oluşturulmuş ve bu sayılara göre konu modellemesi, veri setinde çok fazla geçen ve çok az sayıda geçen kelimeleri normalize etmek için yapılmıştır.

Konu modelleme ve duygu analizi yapıldığında, değişken sınıflarının birbirleri ile farklılıklarının belirlenmesi için box plot analizi yapılmıştır. Araştırmada konular numerik bir değeri temsil ettiği için konular ve otel sınıfı, konular ve seyahat tipi, konular ve duygu sınıfları arasında anova, duygu ve otel sınıfı, seyahat tipi ve duygu türü arasında ise Ki-Kare analizi yapılmıştır. Box -plot kapsamında ortaya çıkan tüm sonuçlarda p değeri 0,0000 olarak ortaya çıkmış olup, araştırmadaki nümerik ve nominal veri türleri arasındaki ilişkilerde anlamlı farklılığın olduğu belirlenmiştir.

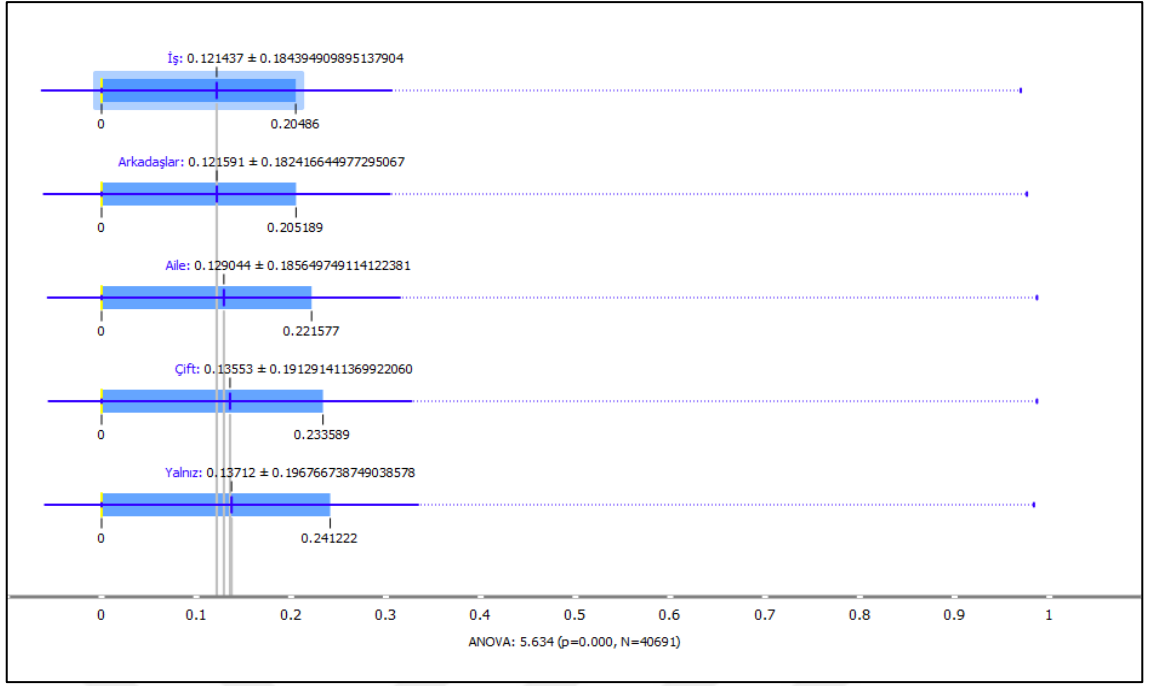
Şekil 4.21, 4.22 ve 4.23'de 1. konu grubunda örnek olarak yapılan box-plot ve ANOVA analizleri gösterilmektedir.



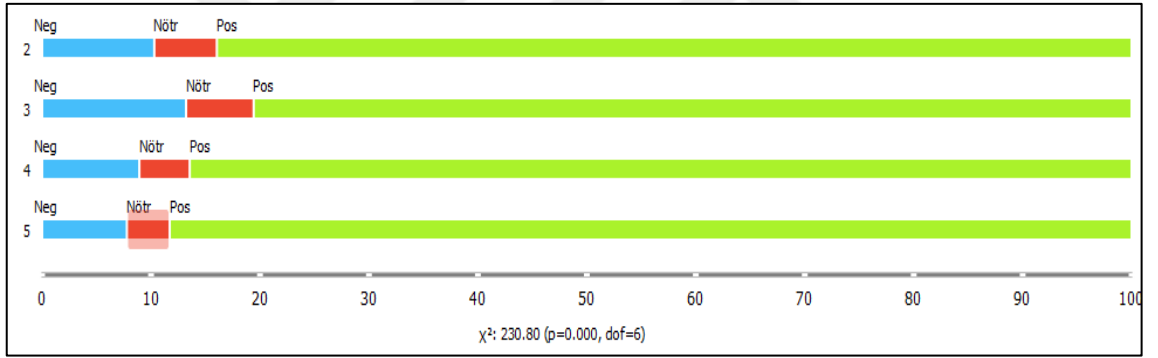
Şekil 4.21 Olanaklar ve otel sınıfı arasındaki ANOVA analizi



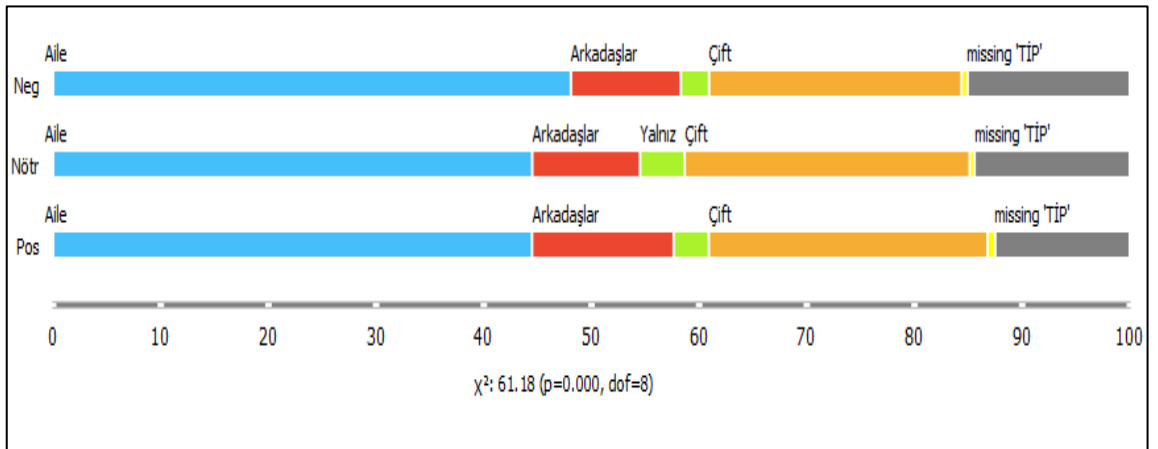
Şekil 4.22 Olanaklar ve duygu türü arasındaki ANOVA analizi



Şekil 4.23 Olanaklar ve seyahat tipi arasındaki ANOVA analizi



Şekil 4.24 Duygu ve otel sınıfı arasındaki ki-kare analizi



Şekil 4.25 Duygu ve seyahat türü arasındaki ki-kare analizi

4.4 Konu Modellemede Ortaya Çıkan Konular

Konu modelleme konularına bakıldığında, birinci konudaki kelimelere bakıldığında, facility, buffet, spa, carte, area, restaurant, standard gibi kelimeler, otelin sağlamış olduğu olanakları göstermiş olduğu ele alınarak ve diğer kelimelerin de bu kelimeler ile kullanıldığı görülerek, bu konu kümesine olanaklar adı verilmiştir ve şekil 4.26’de gösterilmiştir.

Şekil 4.27’ de oluşturulan ikinci konudaki kelimeler gösterilmiştir. Bu kelime grubuna bakıldığında, daha çok personel ve personel davranışlarını niteleyen kelimelerin olduğu görülmektedir. Bu konu grubu ise personel olarak adlandırmıştır.

Üçüncü konu ve konuyu oluşturan kelimeler ise şekil 4.28’ de gösterilmiştir. Bu konu grubuna bakıldığında ise, hem misafirleri, hem de yabancı dilleri ve ödemeler ile ilgili kelimelerin geçtiği görülmektedir. Bu gibi konuların en fazla konuşulmuş olduğu yerlerden birisi ise otellerin ön büro departmanıdır. Bu sebeple bu konu başlığı önbüro olarak belirlenmiştir.

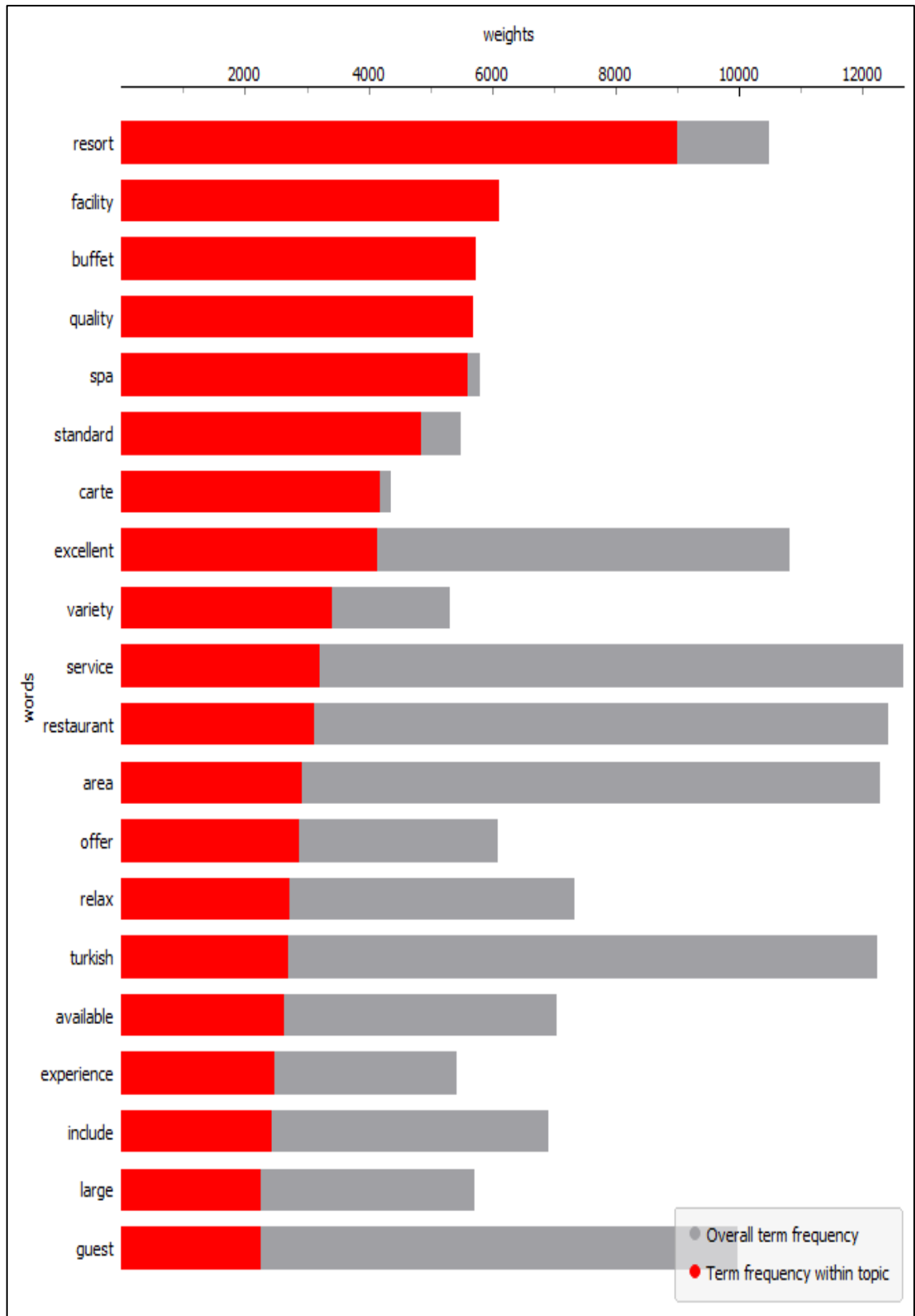
Dördüncü konuda ise alanya ve alanya destinasyonu ile ilgili kelimelerin geçtiği görülmektedir ve şekil 4.29’ da belirtilmiştir. Alanya’ nın sahip olduğu ulaşım, şehirde bulunan olanaklar, kültürel ve doğal unsurları niteleyen ve bunlar ile ilişkili olan kelimeler bu gruba toplanmıştır. Bu sebeple bu konu başlığı destinasyon olarak ele alınmıştır.

Beşinci konu şekil 4.30’da gösterilmiş olup, bu konudaki kelimelere bakıldığında daha çok kişilerin yapmış olduğu rezervasyon aşamalarını niteleyen başta transfer ile ilgili oluşturulan kelimelerin meydana geldiği görülmektedir. Bu amaçla bu gruba rezervasyon ismi verilmiştir

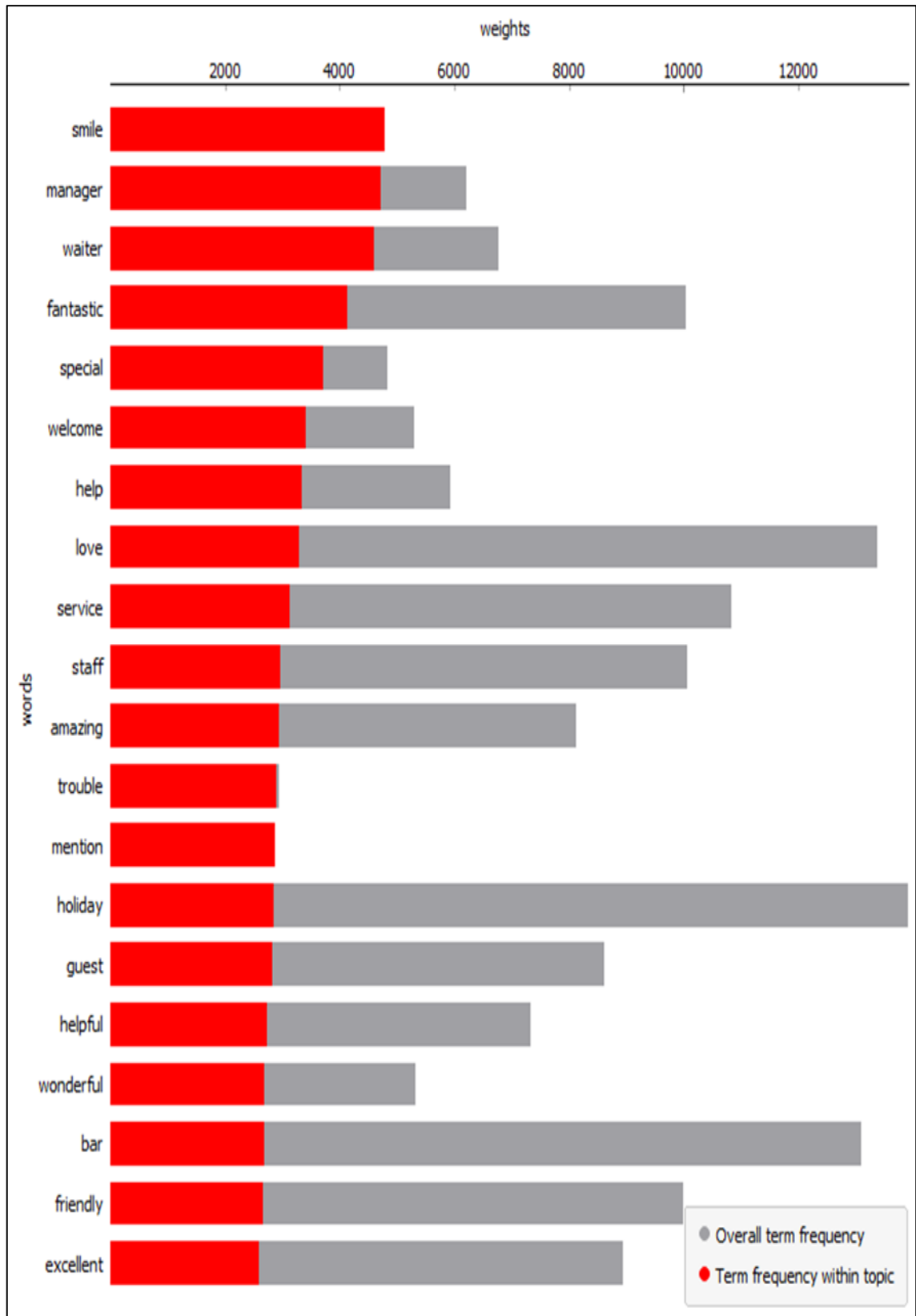
Altıncı konu ise Şekil 4.31’ de gösterilmiştir ve bu gruptaki kelimeler animasyon ve eğlence imkanlarını niteleyen kelime gruplarından ağırlıklı olarak oluştuğu görülmüştür. Bu sebeple bu konu başlığı animasyon olarak belirlenmiştir.

Yedinci konuda ise yiyecek içecek ile ilgili olan kelimelerin bir arada toplandığı görülmektedir ve bu konu başlığına da yiyecek içecek ismi verilmiştir ilgili konu şekil 4.32’de gösterilmektedir.

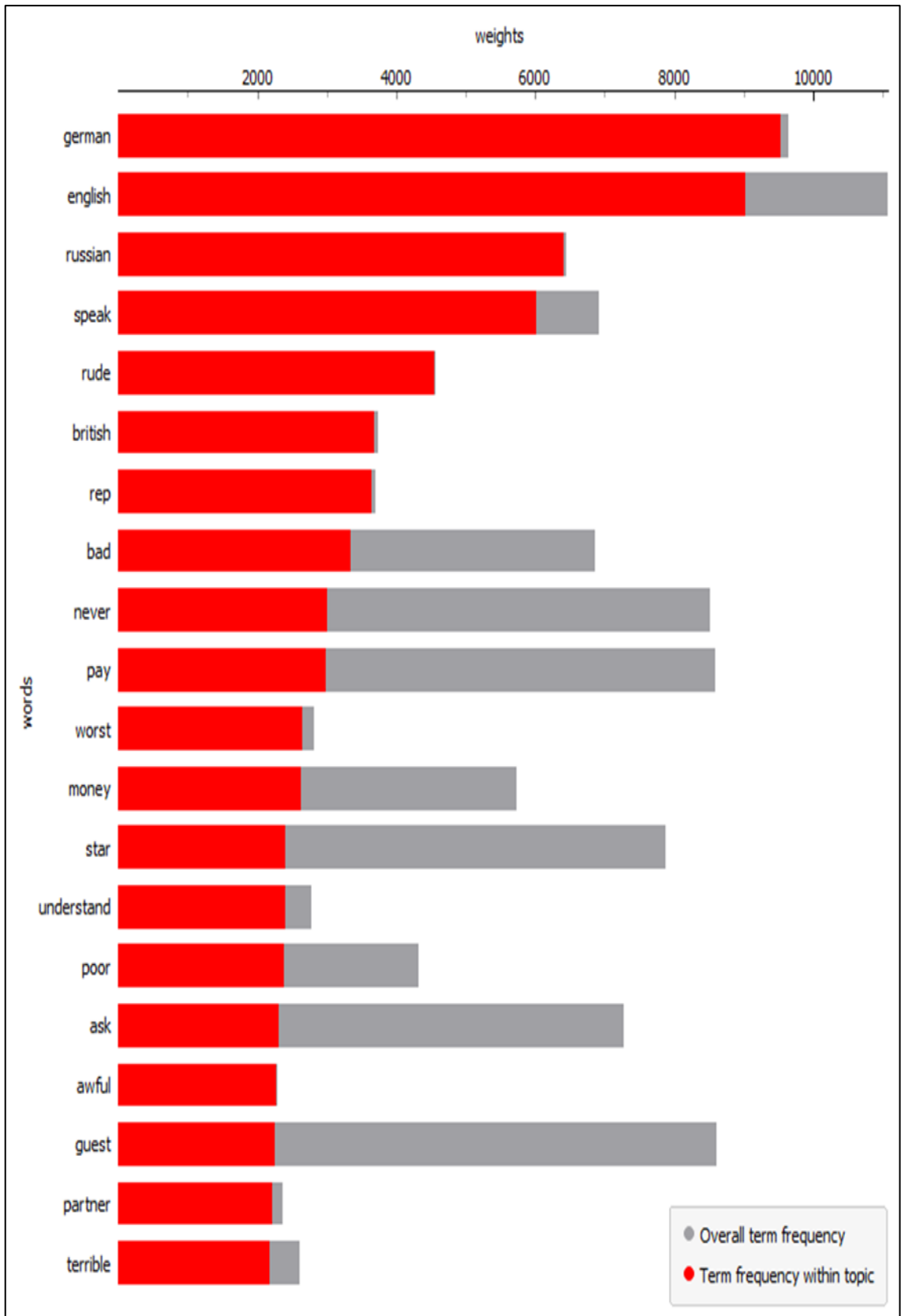
Sekizinci konu ise odalar ve odaların imkanları, odaların sahip olduğu materyaller ve bunları niteleyen kelimelerden toplandığı görülmüştür. Bu sebeple bu konu başlığında odalar adı verilmiştir ve şekil 4.33’ de gösterilmiştir.



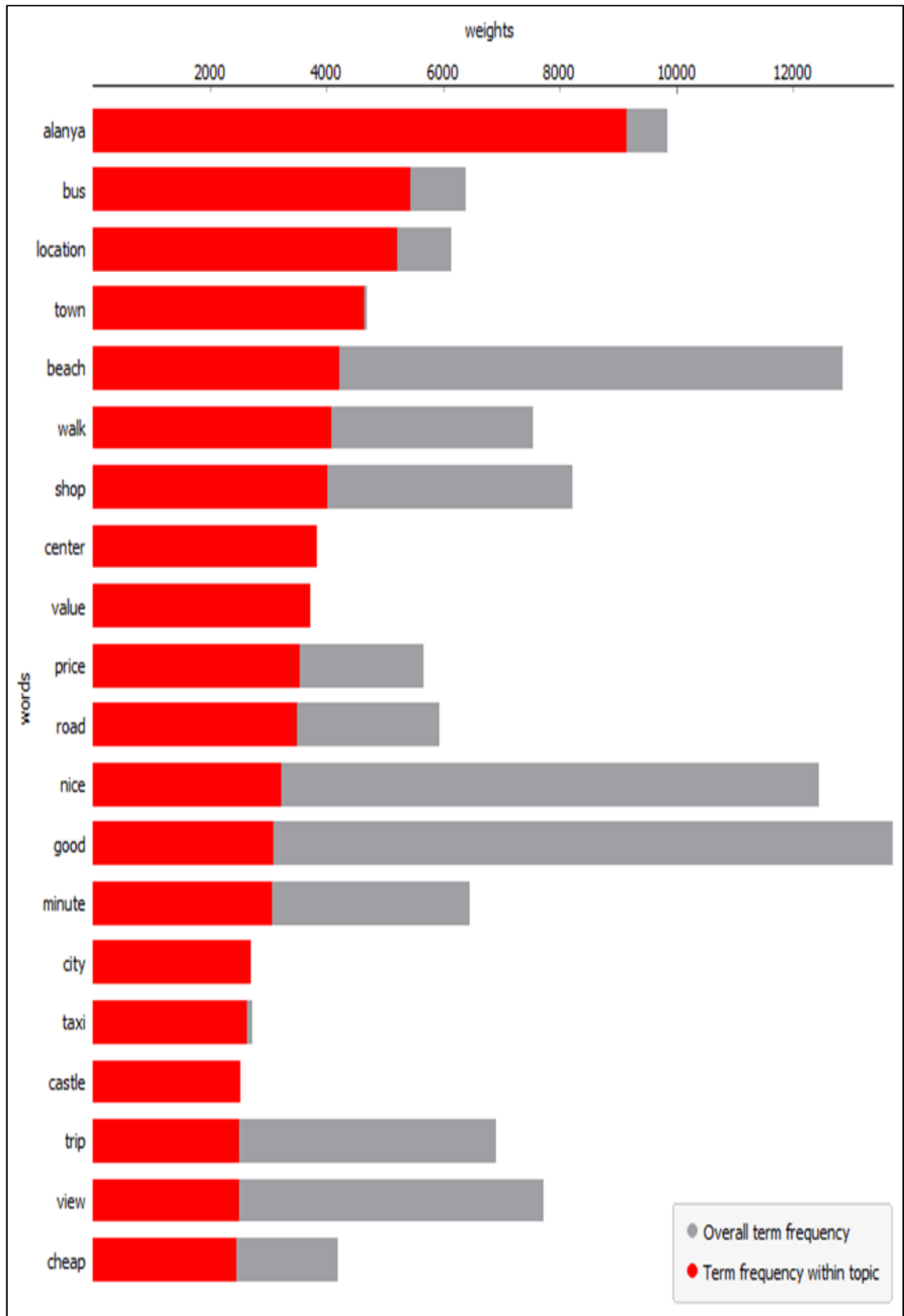
Şekil 4.26 Olanaklar



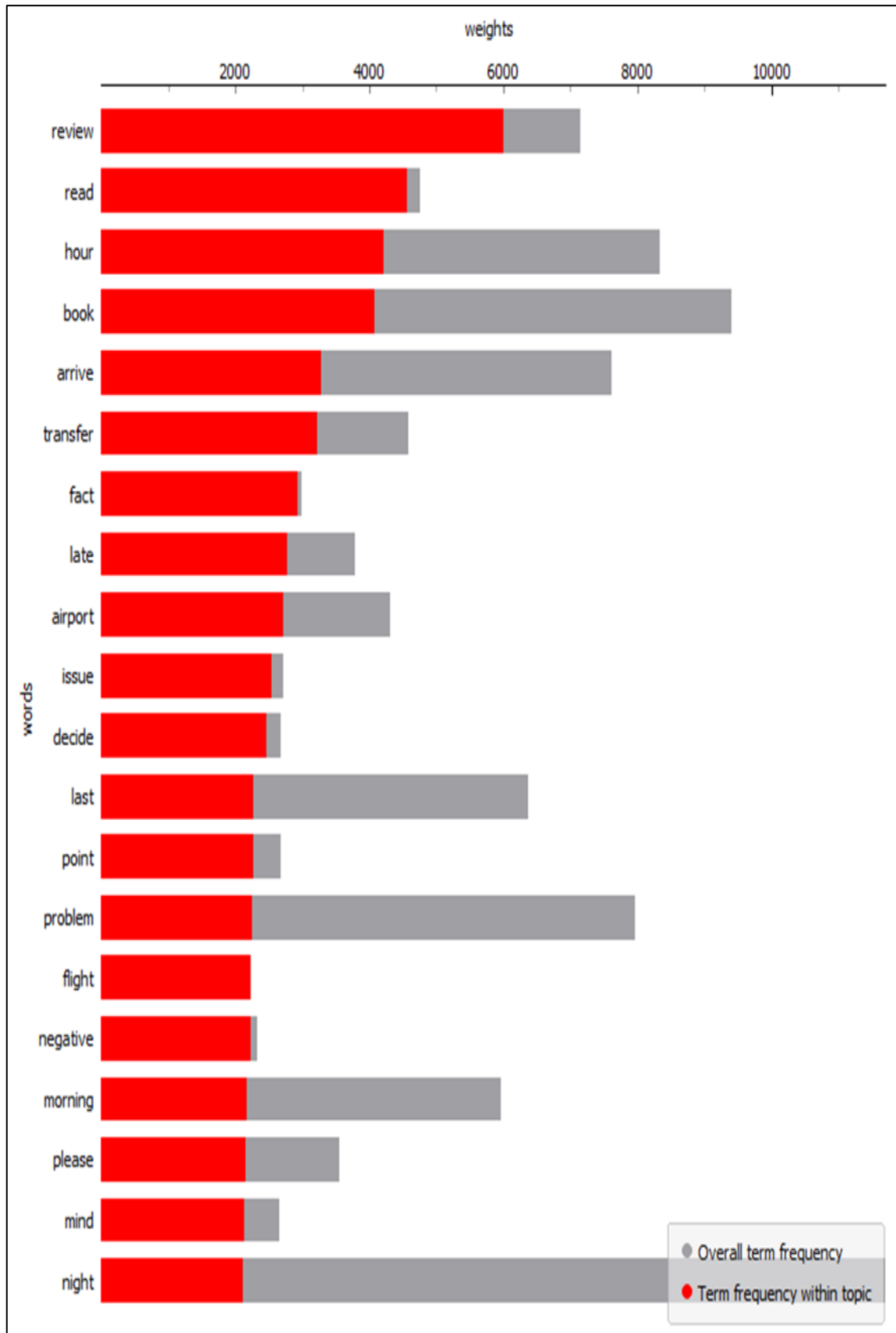
Şekil 4.27 Personel



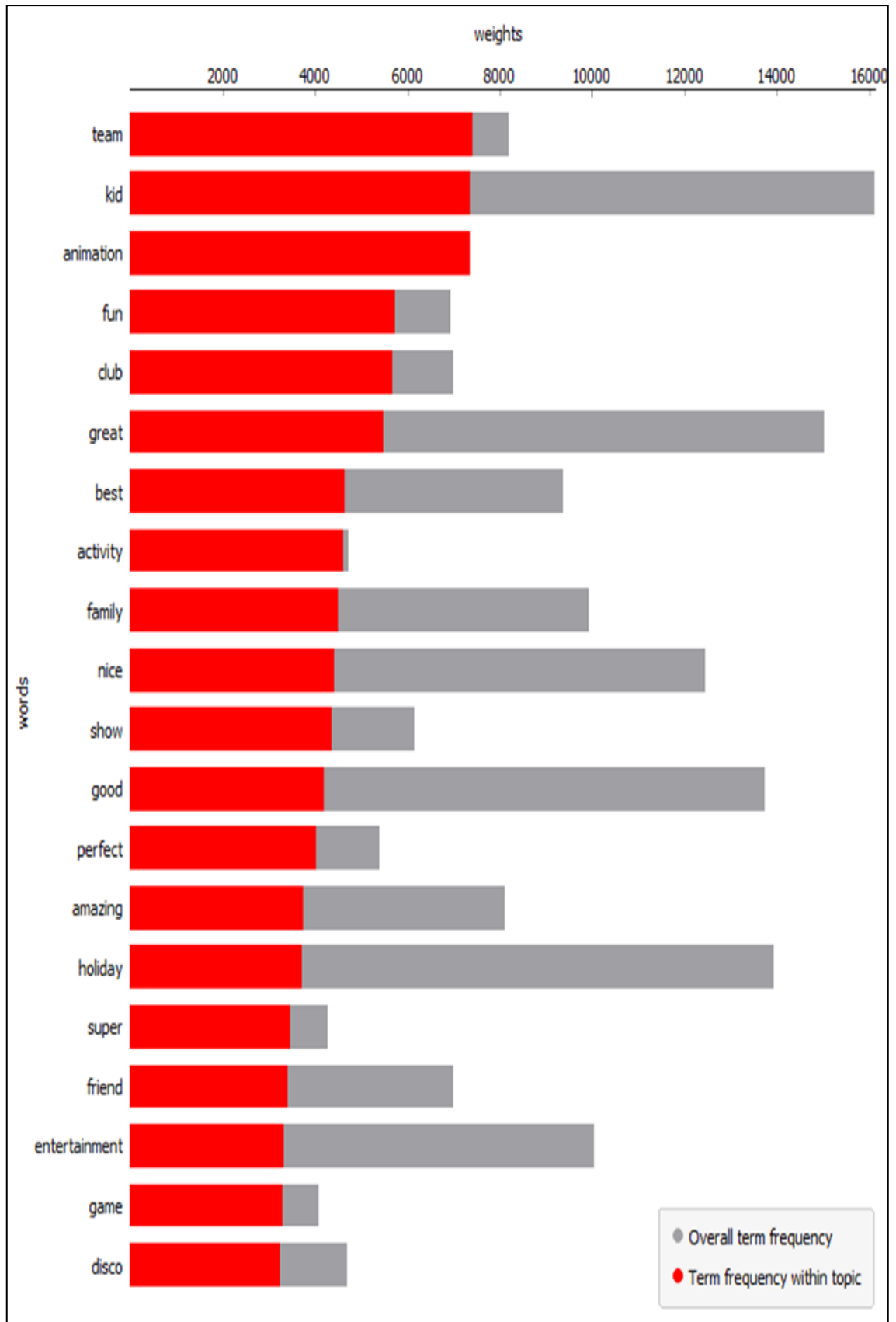
Şekil 4.28 Önbüro



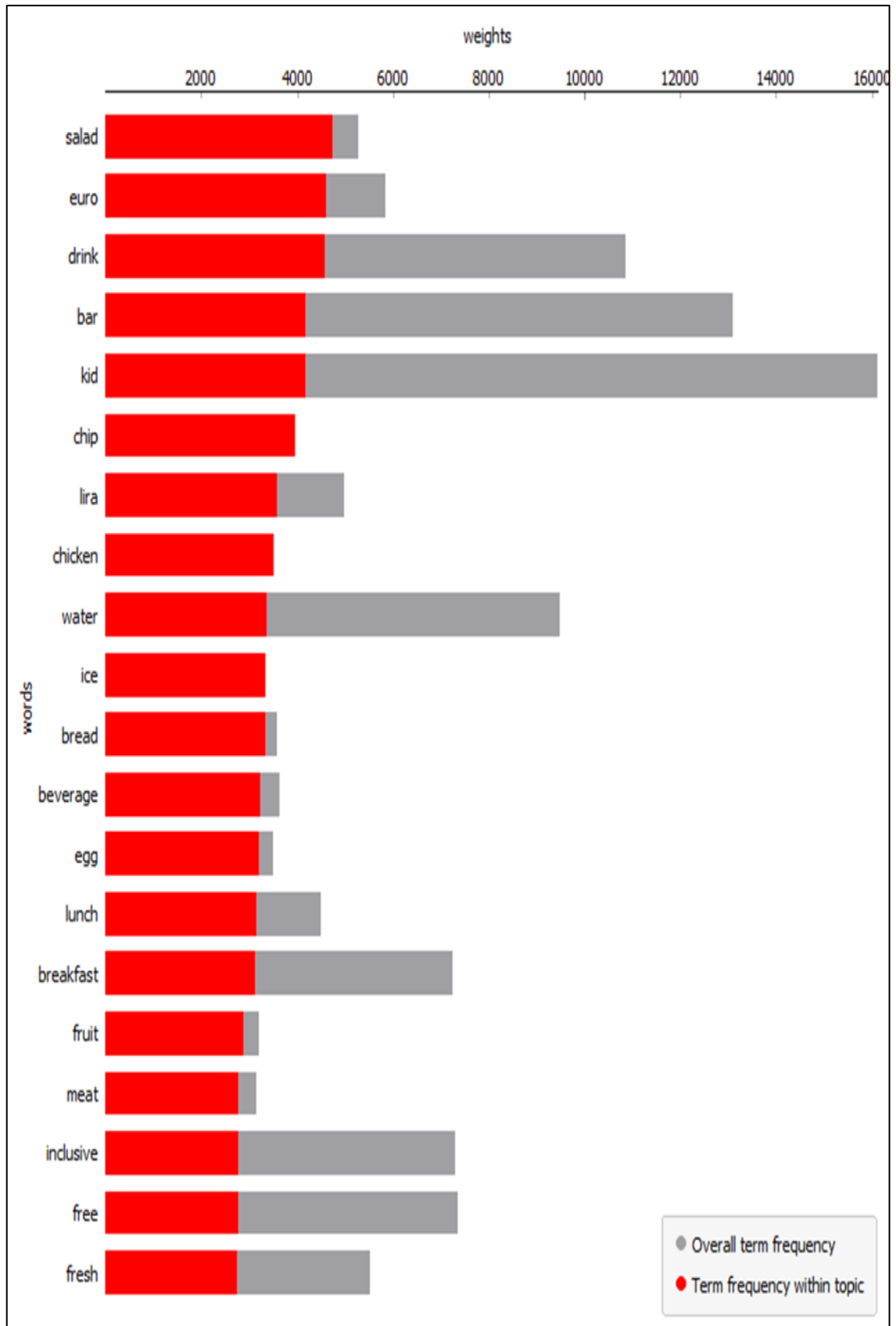
Şekil 4.29 Destinasyon



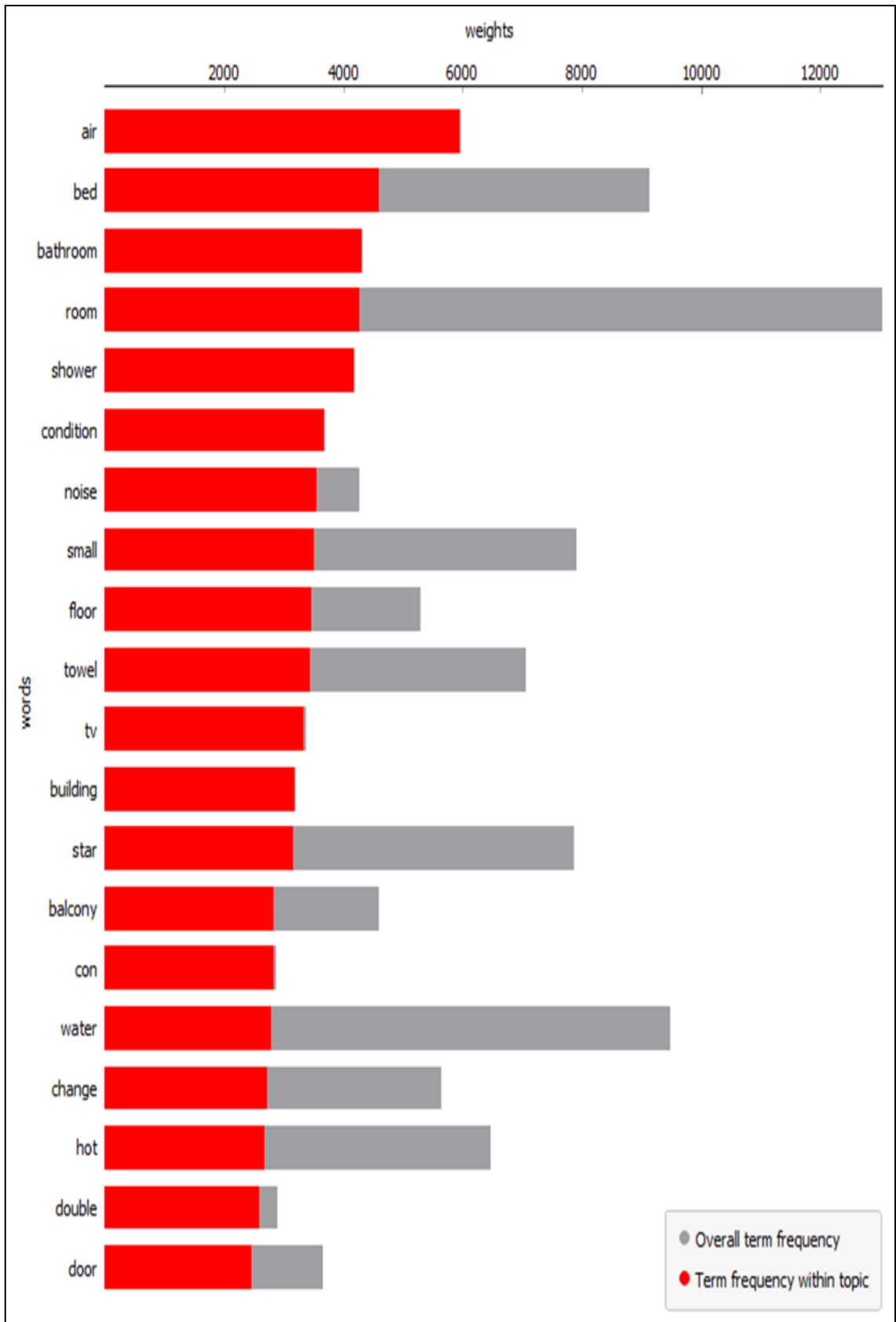
Şekil 4.30 Rezervasyon



Şekil 4.31 Animasyon



Şekil 4.32 Yiyecek-İçecek



Şekil 4.33 Oda

4.5 Duygu ve Konu Sınıfı İlişkileri

Çalışmada LDA metodu ile oluşturulan ve konularda geçen kelime içeriklerinin yapısına ve niteliğine göre adlandırılan konu başlıklarının, konaklama yapan kişilerin yazmış olduğu yorumlardaki duygu skoruna (compound score) göre Pearson korelasyon analizi yapılmıştır. Bu aşamada araştırmada ortaya konulan konular ile duygu skoru arasındaki ilişkinin yönü ve gücü ortaya çıkarılmak istenmiştir.

Şekil 4.3’de duygu analizi sonucunda her bir yorumun elde edildiği duygu skoru (compound score) ile konular arasındaki ilişki ölçülmüştür. Buna göre önbüro, oda, rezervasyon ve yiyecek-içecek ile ilgili yapılan görüşler ile duygu skoru arasında negatif bir ilişki meydana geldiği görülmektedir. En fazla negatif ilişkinin meydana geldiği alan, önbüro olarak görülmektedir. Dolayısıyla önbüro ile ilgili yorum yapan kişilerin vermiş olduğu puanlar arasında negatif bir ilişki vardır ve bir yorumda eğer önbüro ile ilgili kavramların ortaya çıkması, insanların daha çok yorumlarının negatif yönde olmasına sebep olduğu görülmüştür. Odalar ile ilgili yapılan yorumların ise en fazla ikinci negatif etkiye sahip olduğu görülmektedir. Rezervasyon ve yiyecek-içecek ile ilgili yapılan yorumların ise daha az oranda negatif olarak duygu skorunu etkilediği ortaya çıkmıştır.

Pozitif olan korelasyon yönlerine ve etkilerine bakıldığında ise, yorumların duygu skoru ile animasyon arasında en fazla pozitif ilişkinin olduğu görülmektedir. Bunu sırasıyla olanaklar, personel ve destinasyon konuları takip etmektedir. Bu konular ile ilgili yorumlarda geçen kelimeler, doğal olarak kişilerin yapmış olduğu yorumların duygu skorları ile arasında pozitif bir ilişkinin olduğu belirlenmiştir.



Şekil 4.34 Duygu skoru ve konular arasındaki korelasyon analizi

4.6 Konulara Göre Duygu Sınıfı Dağılımları

Tablo 4.9’ da duygu analizi sonucunda oluşturulan skora göre sınıflanan pozitif, nötr ve negatif yorumların, her bir konuya göre dağılım yüzdesi görülmektedir. Buna göre en pozitif olarak değerlendirilen konu animasyon olmuştur ve animasyon ile ilgil olan yorumların %71,i pozitif olarak görülmektedir. Personel ve olanaklar konusunda ise pozitif yorumlar, diğer duygu sınıflarına göre daha fazla olduğu görülmektedir. Önbüro konusu ise oluşturulan konularda en fazla negatif anlamda düşüncelere sahip olan konu olarak görülmüştür. Önbüro konusunu negatif olarak odalar takip etmektedir ve odalara yapılan yorumların %80’inden fazlası negatif ve nötr olarak ortaya çıkmıştır.

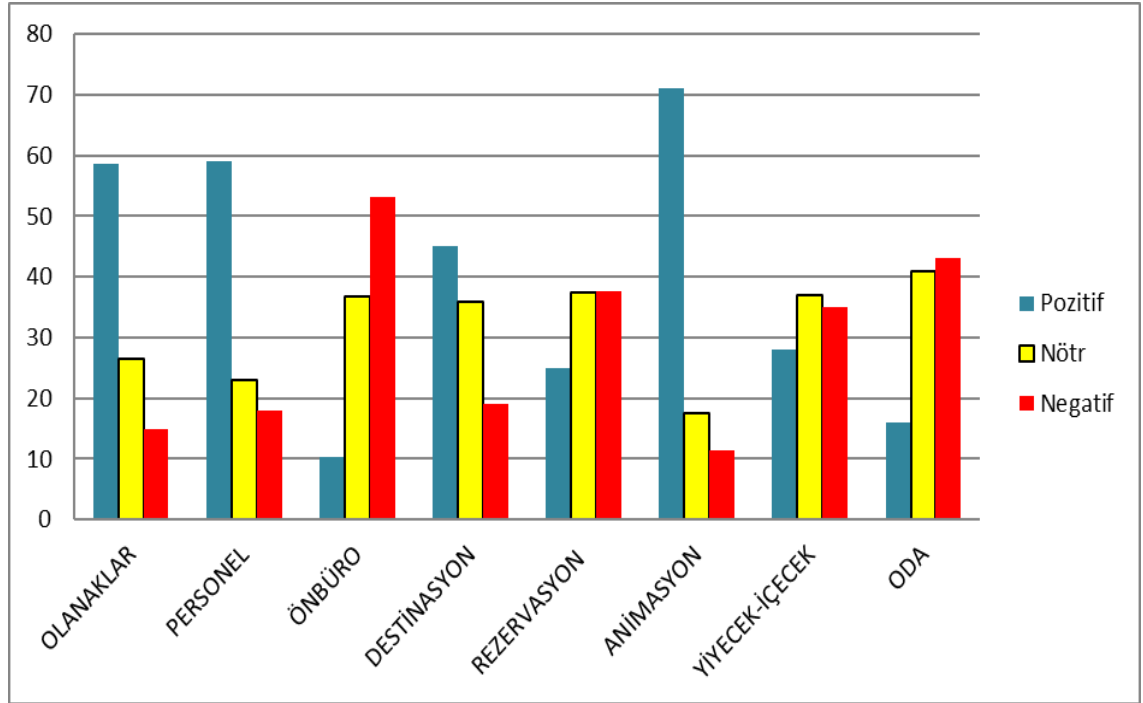
Şekil 4.35’deki grafikte ise her bir konuya ait olan duygu sınıflarının yüzdesinin grafiksel olarak dağılımı belirtilmiştir. Animasyon konusuna yapılan yorumlarda, pozitif yorumların oranı diğer duygu sınıflarına göre daha yüksek derecede olduğu görülmektedir. Bu farkı olanaklar ve personel konuları takip etmektedir ancak bu konuya yapılan yorumların duygu sınıfları arasındaki fark, animasyona göre daha az görünmekle birlikte pozitif olanlar daha yüksektir. Destinasyon konusunda ise pozitif yorumlar daha yüksek olup, nötr ve negatif yorumlara oranı daha az seviyede çıkmaktadır. Yiyecek içecek konusunun ise tüm duygu sınıfları birbirlerine yakın bir oranda ortaya çıkmıştır. Yiyecek içecek konusunda nötr duygu sınıfı diğerlerine göre yüksek iken, rezervasyon konusunda nötr ve negatif yorumların oranı, pozitiflere göre daha fazla olup ikisi eşit seviyede olduğu görülmektedir. Odalar ile ilgili olan yorumlarda ise negatifler diğer duygu sınıflarına göre yüksek olup, nötr ile arasında daha az fark bulunmaktadır. Önbüroda ise en fazla negatif yorum yüzdesine sahiptir ve negatif ve nötr olan yorumların pozitiflere göre arasındaki fark diğerlerinden daha fazla olduğu görülmektedir.

Buna göre, Alanya’ daki otellerde konaklama yapan kişilerin yapmış olduğu yorumlarda genel olarak animasyon, olanaklar, personel ve destinasyon ile ilgili olan kavramları olumlu olarak tepki verdiği görülmektedir. Bu konuların içermiş olduğu terimleri yorumlarda belirttiklerinde, kişiler genellikle olumlu olarak görüş bildirdiği görülmektedir. Ancak destinasyon konusunda pozitif duygu tipleri ile nötr ve negatifler de önemli bir yer tutmaktadır. Yiyecek içecek ve odalar konusunda ağırlıklı olarak nötr ve negatif duygu sınıflarının olduğu görülmektedir. Ancak oda konusunda yapılan yorumlarda pozitif yorumların negatif ve nötr ile arasındaki fark daha fazla olduğu

görülmektedir. Önbüro ile ilgili olan terimlerin kullanıldığı yorumların ise büyük kısmı nötr ve negatif olarak belirtilmiştir.

Tablo 4.9 Konulara göre duygu sınıflarının dağılımı

Konular	Pozitif (%)	Nötr (%)	Negatif (%)
Olanaklar	58,59	26,50	14,91
Personel	59,07	22,91	18,02
Önbüro	10,20	36,73	53,06
Destinasyon	45,05	35,90	19,05
Rezervasyon	24,92	37,38	37,69
Animasyon	71,00	17,52	11,48
Yiyecek-İçecek	28,00	36,92	35,08
Oda	16,06	40,79	43,14

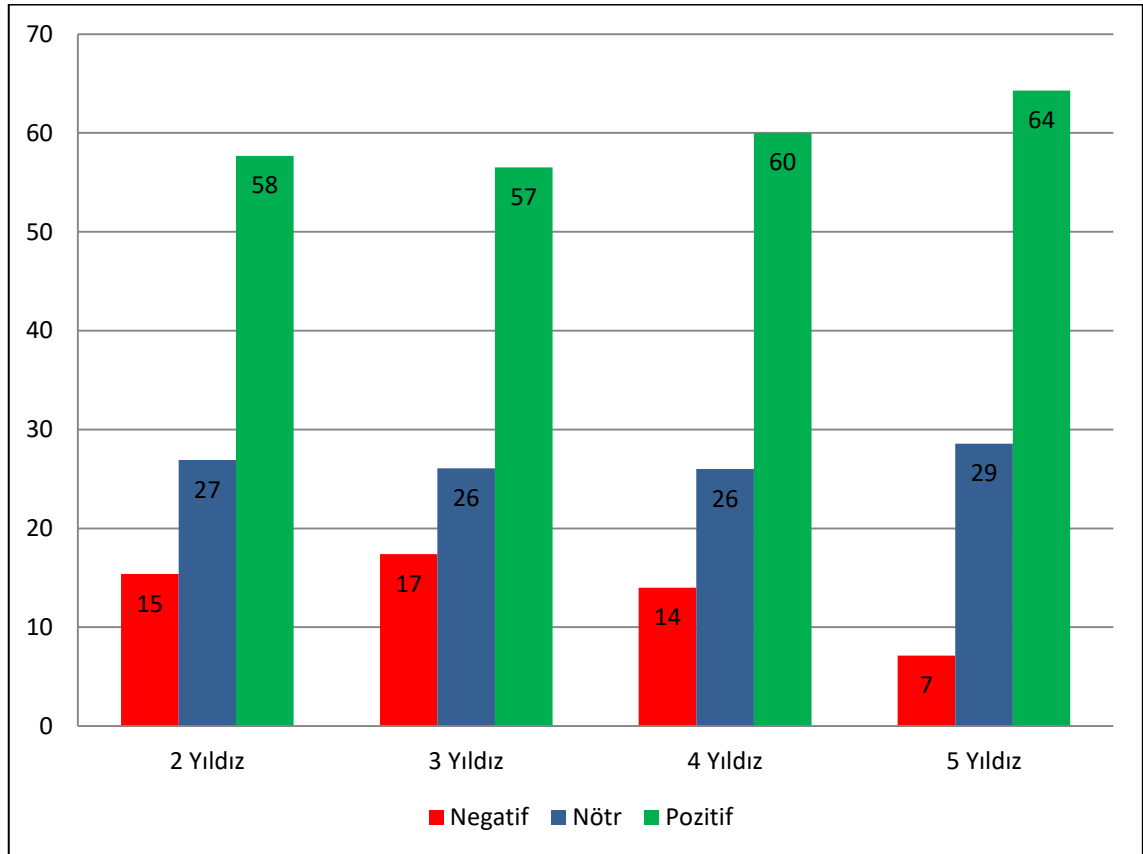


Şekil 4.35 Konulara göre duygu sınıflarının dağılımı

4.7 Otel Sınıflarına Göre Konu Modelleme ve Duygu Analizi Sonuçları

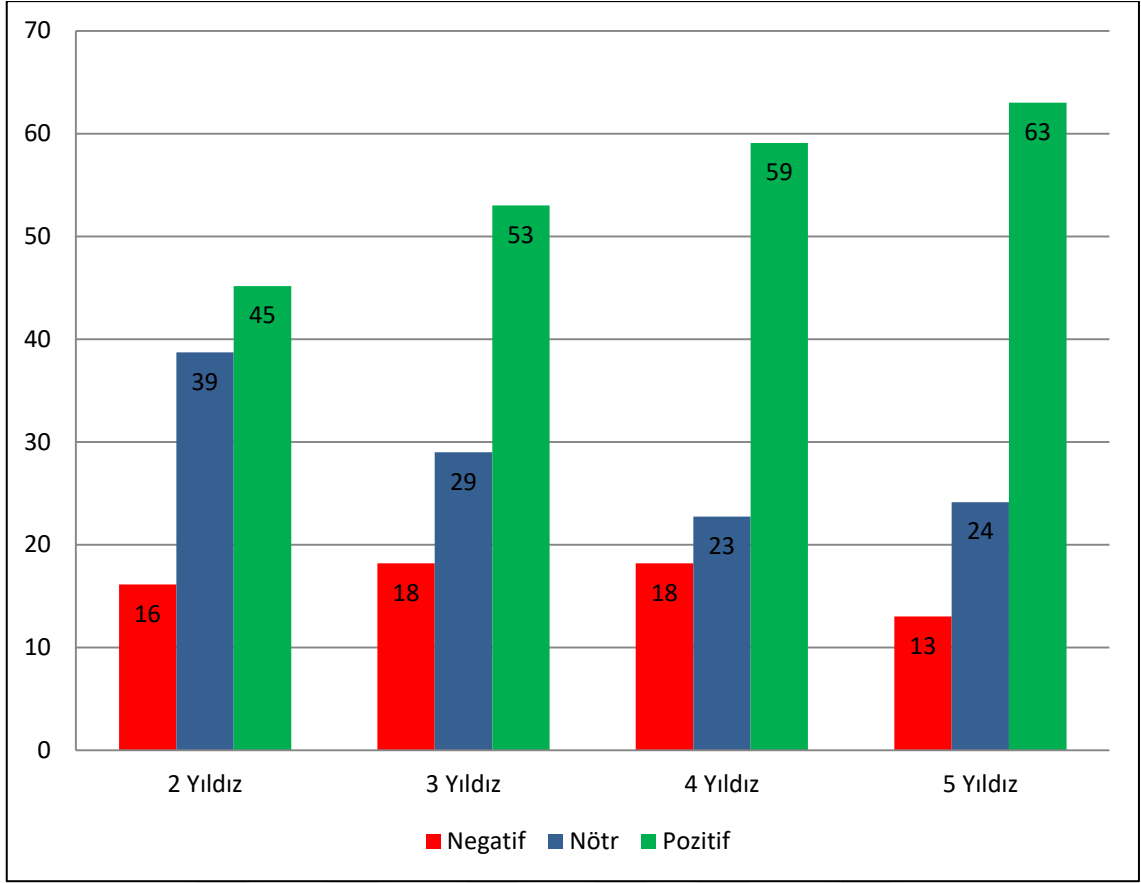
Konu modellemesinde ortaya çıkan konuların yüzdesel olarak veri seti genelinde duygu tiplerini belirlendikten sonraki aşamada otel sınıflarına göre bu konuların duygu dağılımlarının ortaya çıkartılması amaçlanmıştır. Buna göre otel sınıflarına göre hangi konuların insanlarda hangi duygu tipine yol açtığı belirlenmek istenmiştir. Bu kapsamda konu modellemede ortaya çıkan konuların her birinin otel sınıflarına göre duygu sınıfları gösterilmiştir.

Şekil 4.36’da olanaklar konusu ile ilgili yapılan yorumların otel sınıflarına göre ayrımı görülmektedir. Olanaklardaki duyguların sınıflara göre dağılımı bakıldığında, en fazla pozitif yorum oranı 5 yıldızlı otellerde görülmektedir ve otel sınıfı arttıkça pozitif yorumların diğer yorumlara göre dağılımı arttığı söylenebilir. Bu kapsamda beş yıldızlı otellerin doğal olarak sağlamış olduğu hizmetler, kişilerin olumlu olarak bu konuda değerlendirme yapmasına etkisi olmaktadır. 2 ve 3 yıldızlı otellerin sahip olduğu olanaklar hakkında yapılan yorumların duygu sınıflarına bakıldığında da, yine olumlu anlamda yorumların yapıldığı görülmektedir. Bu aşamada insanların otel sınıflarına göre otellerin sağladığı olanaklardan genel anlamda olumlu olduğu, ancak olanak seviyesinin arttıkça bu durumun insanların genel olarak beğeni seviyesini arttırmış olduğu gözlemlenmektedir.



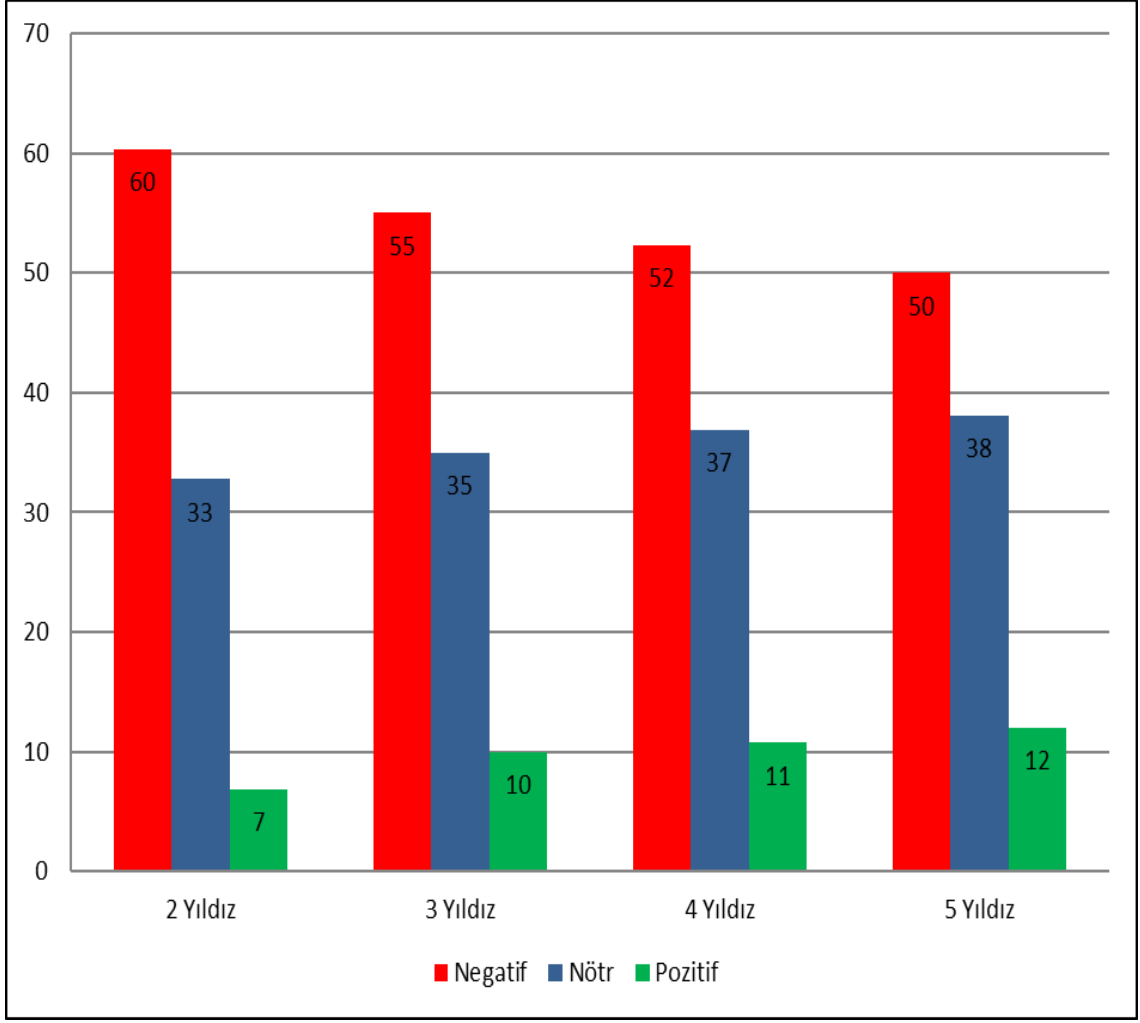
Şekil 4.36 Olanaklar konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.37’de personel konusuna ait duygu sınıflarının, otel sınıflarına göre dağılımı verilmiştir. Buna göre otel sınıfı arttıkça, personele duyulan memnuniyet artış göstermektedir.



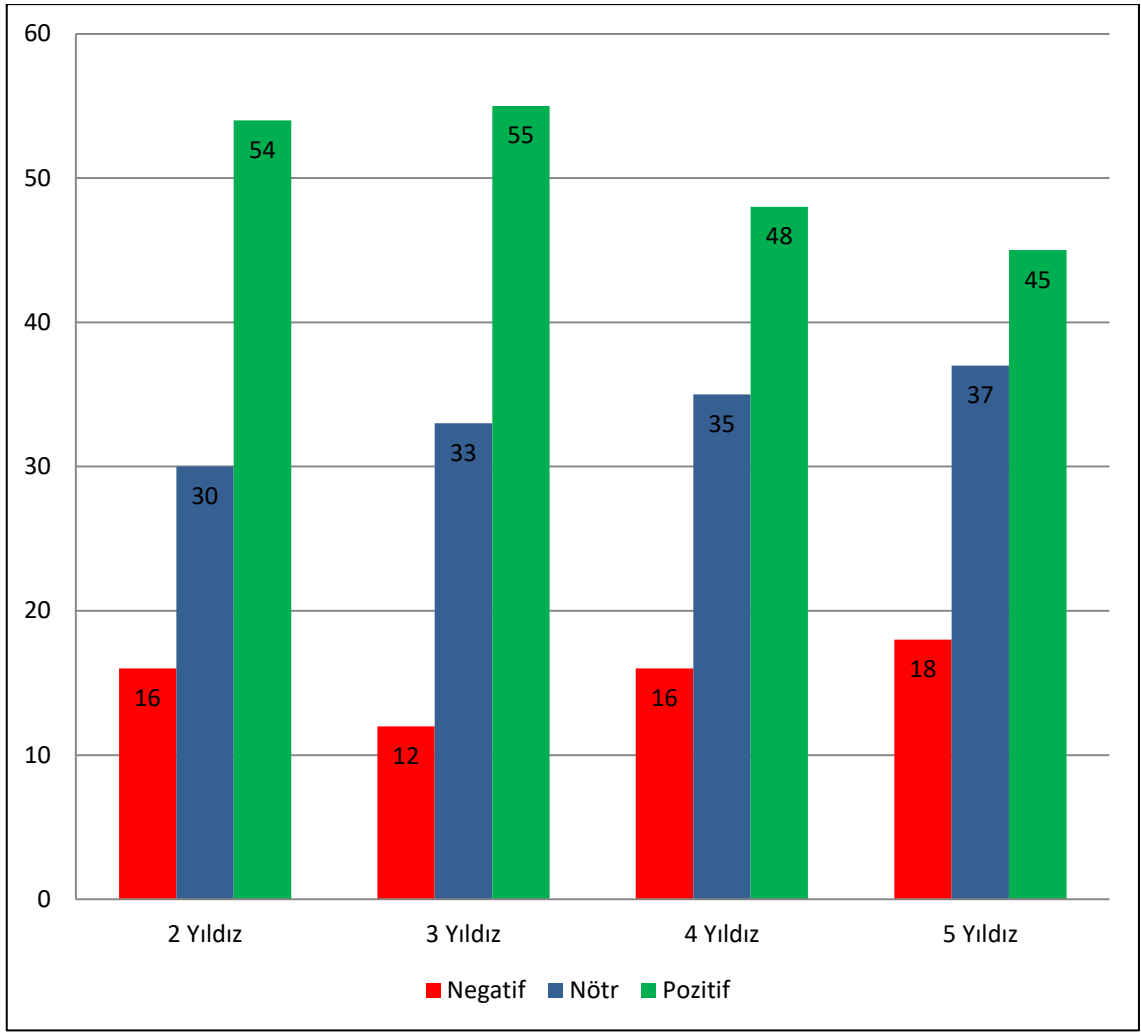
Şekil 4.37 Personel konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.38’de önbüro konusunun duygu dağılımları otel sınıflarına göre sunulmaktadır. Buna göre konular arasında negatif olarak en fazla görünen konu başlığı önbüro olarak görünmektedir. Önbüro, bilindiği gibi otel işletmelerinde konaklayan misafirlerin şikayetlerini ilk olarak sunmuş olduğu departmanlardan birisidir. Bazı otel işletmelerinde, önbüro departmanı kapsamında ayrıca misafir ilişkileri yönetimi ile ilgili departmanlar da bulunmaktadır. Bu kapsamda Alanya’ da konaklama yapan misafirlerin yapmış olduğu görüşlerde önbüro departmanı, konaklayan bireylerde negatif yönde bir intiba bıraktığı görünmektedir ve bu durum otel sınıfı azaldıkça, kişilerin negatif yaptığı yorumlar da buna paralel olarak artmakta olduğu gözlemlenmiştir. Araştırmada ortaya çıkan önbüro konusu ile ilgili olan terimlere bakıldığında, dil, davranış, ödeme gibi önbüro departmanı ile ilgili olan terimler ortaya çıkmış ve genellikle bu terimlerin kullanılmış olduğu yorumların negatif yönde bulunduğu ortaya çıkmıştır.



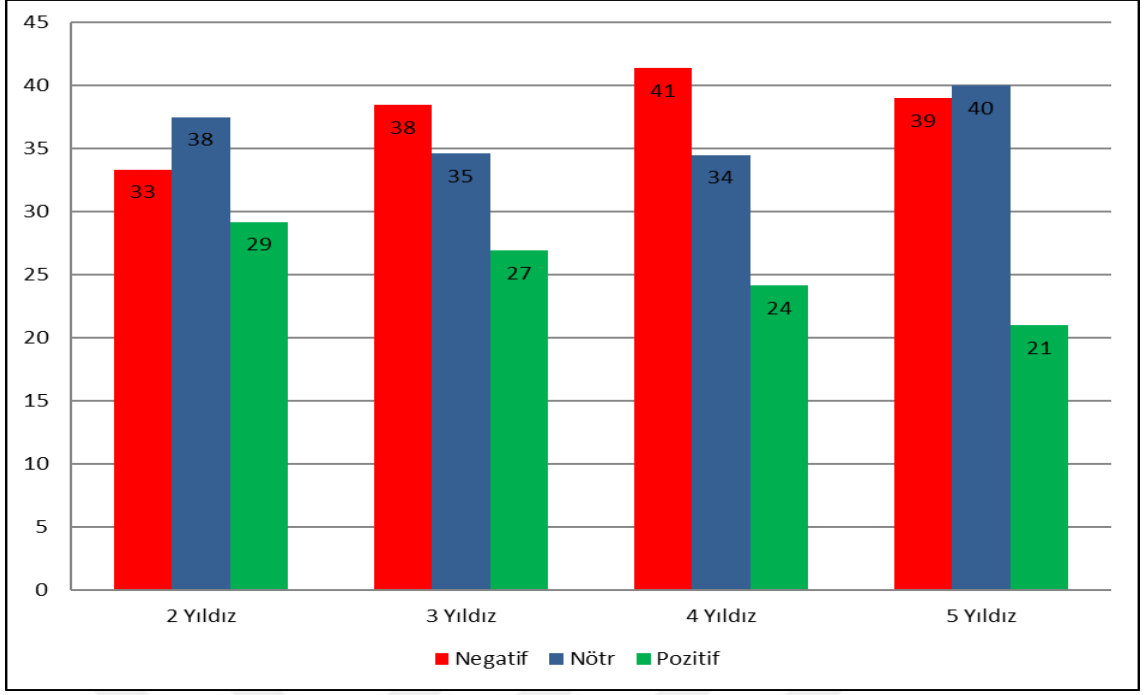
Şekil 4.38 Önbüro konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.39’da destinasyon konusunun otel sınıflarına göre yapılan duygu sınıflaması gösterilmektedir. Buna göre otel sınıfı azaldıkça destinasyon ile ilgili olarak yapılan yorumların pozitifte dönmesi artmaktadır. Bu aşamada otel sınıfı düşük olan otellerin destinasyon merkezinde olma durumları, konaklama yapan kişilerin şehir ile daha fazla entegre olması ve şehirde bulunan imkanlardan yararlanması, bu durumun meydana gelmesinin nedeni olarak gösterilebilir. Bazı 5 yıldızlı otellerin şehir merkezine uzak olması ve şehir merkezine ulaşılması gibi durumlarda meydana gelen sıkıntılar, bu otellerde konaklama yapan kişilerin destinasyon ile ilgili görüşlerini negatif olarak etkilediği söylenebilir.



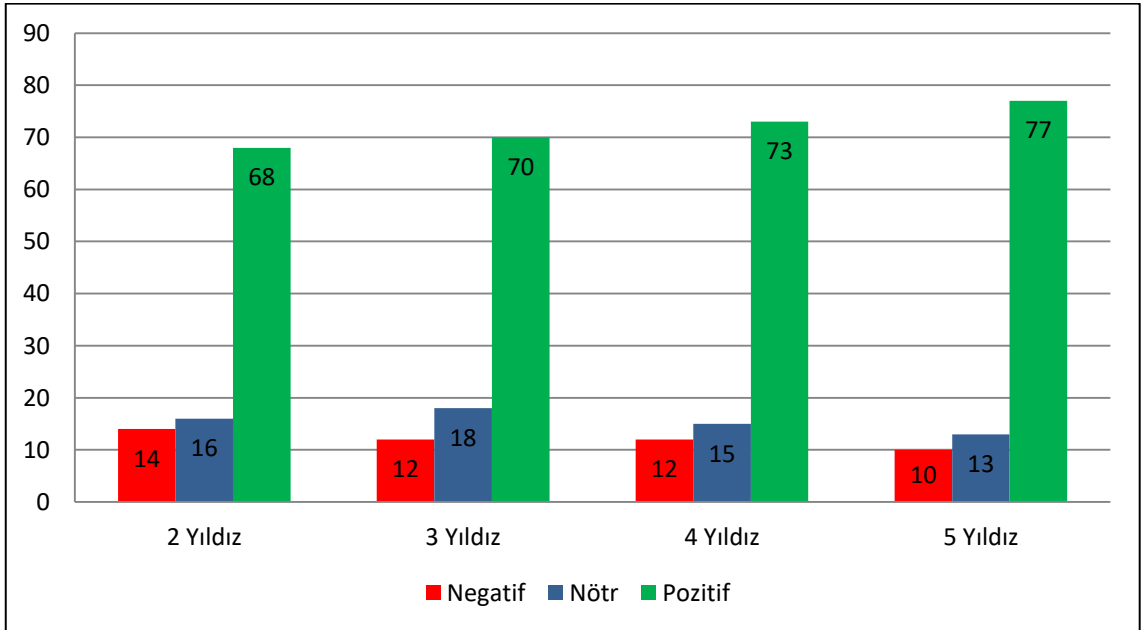
Şekil 4.39 Destinasyon konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.40’da rezervasyon konusunun otel sınıflarına göre duygu dağılımları gösterilmektedir. Araştırmada genel olarak negatif yönde olan konu gruplarından birisi rezervasyon konusunu oluşturmaktadır. Rezervasyon konusunda ilgili olan kelimelere bakıldığında transfer ve geliş ile ilgili unsurların yer aldığı görülmektedir ve bu konularda konaklama yapacak olan kişilerin yazmış olduğu yorumlarda genellikle negatif olarak yorumlar yapmış olduğu gözlemlenmektedir. Ayrıca rezervasyon konusunda otel sınıfı arttıkça, bu konuda yazılan yorumların negatif bir aşamaya geldiği belirlenmektedir. Özellikle 5 yıldızlı otellerin rezervasyon sürecinde sunmuş olduğu transfer ve otele ulaşım ile ilgili olan unsurlarda kişilerin negatif yorum yapmış olduğu belirlenmiştir.



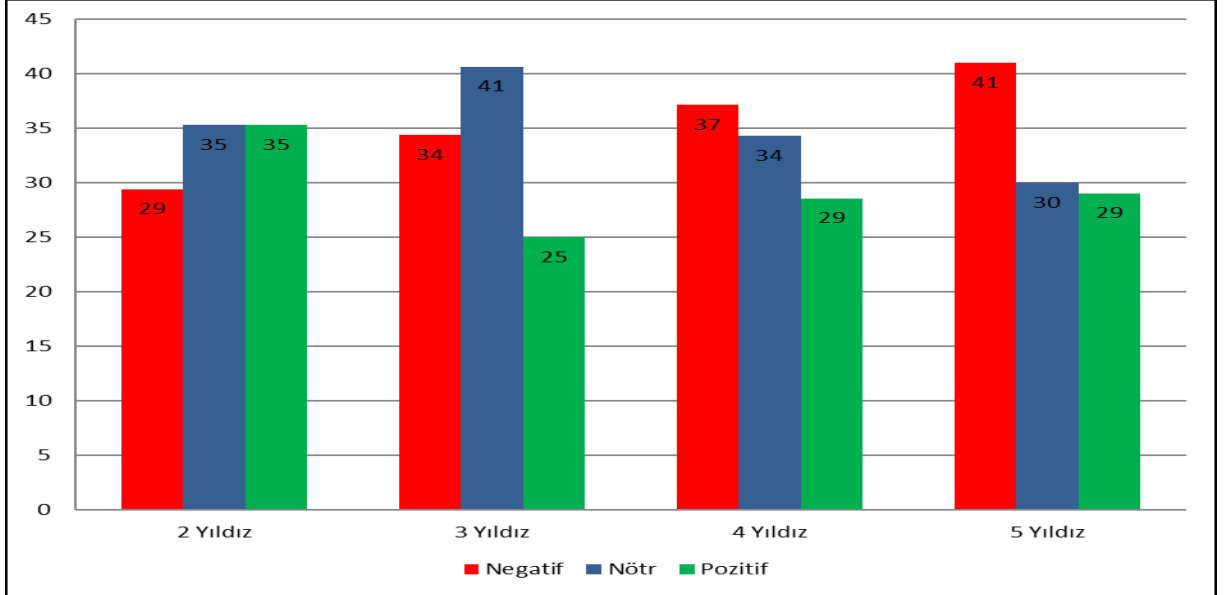
Şekil 4.40 Rezervasyon konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.41’de animasyon konusunun duygu sınıflarının otel sınıflarına göre dağılımı verilmektedir. Genel olarak bireylerin en fazla pozitif olarak yorum yapma eğiliminde olduğu alan animasyon konusu olmaktadır. Özellikle otel sınıflarının arttıkça animasyon ve diğer eğlence olanaklarına daha fazla sahip olması, bu konunun otel sınıfı arttıkça daha fazla pozitif anlamda değerlendirildiği söylenebilir.



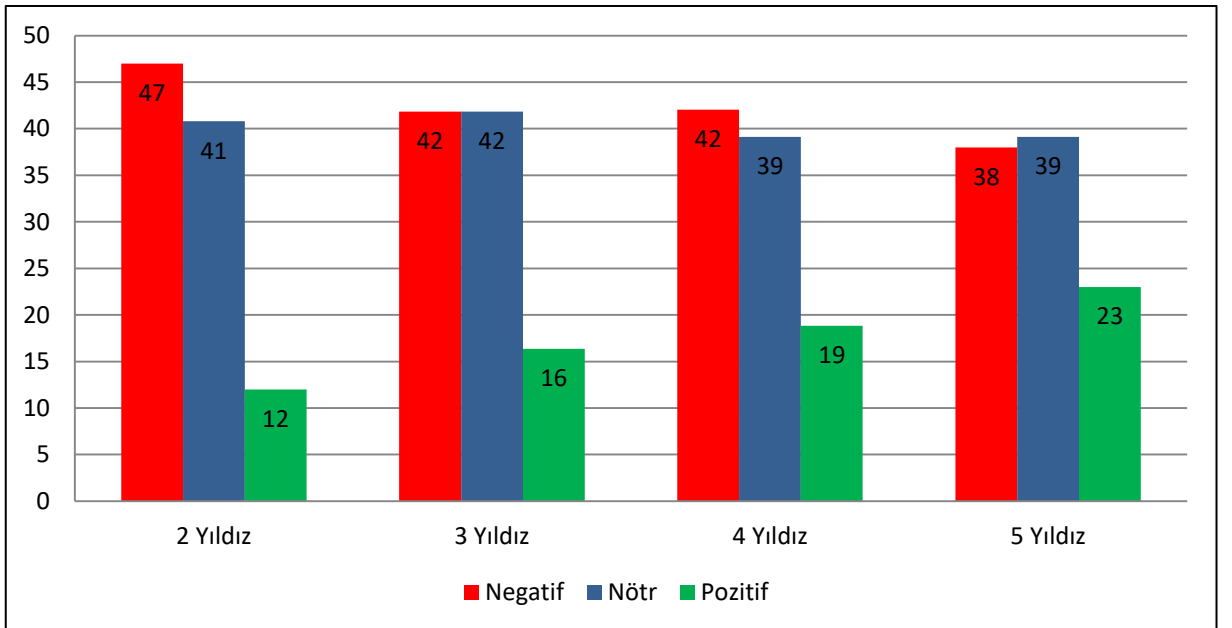
Şekil 4.41 Animasyon konusunun otel sınıflarına göre duygu türü dağılımı

Şekil 4.42’ de yiyecek içecek konu grubunun otel sınıflarına göre duygu sınıfları gösterilmiştir. Buna göre otel sınıfı arttıkça, otellerden beklenen yiyecek içecek hizmeti ile ilgili olarak negatif bir yorum yapılması eğilimi olmaktadır.



Şekil 4.42 Yiyecek-İçecek konusunun otel sınıflarına göre duygu türü

Oda ile ilgili duygu sınıfları şekil 4.43’ de gösterilmektedir. Buna göre otel sınıfı arttıkça, otellerdeki odalara yapılan yorumların içeriğinin pozitif yöne gittiği görülmesine rağmen, genel olarak odalarla ilgili genel eğilim negatif ve nötr olarak görülmektedir.



Şekil 4.43 Oda konusunun otel sınıflarına göre duygu türü dağılımı

4.8 Uzaklık Matrisi ve Hiyerarşik Kümeleme Analizi Sonuçları

Araştırmada ortaya çıkartılan konulara gözetimsiz öğrenme yöntemlerinden olan hiyerarşik kümeleme yöntemi uygulanarak hangi konuların yorumlarda bir arada bulunduğu ortaya çıkarılmak istenmiştir. Hiyerarşik kümeleme, benzerlik ve mesafe ölçütleri kullanarak, veri setindeki benzer özellik gösteren niteliklerin bir araya getirilmesi ya da bölünmesi ile oluşturulan bir kümeleme yöntemidir (Silahtaroglu, 2016; Ülgen, 2018).

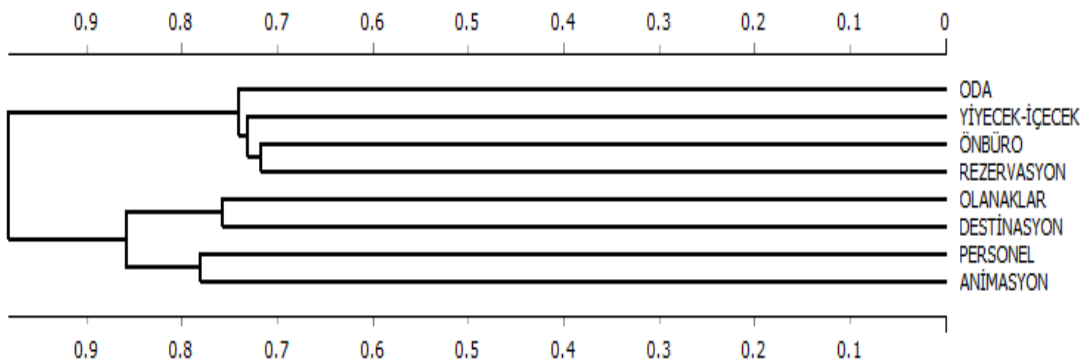
Araştırmada kullanılan kümeleme yönteminde agglomerative (birleştirici) yaklaşım kullanılmıştır. Bu yaklaşımda veri setindeki konuların hepsi ayrı bir küme olarak belirlenmekte ve benzer yapıdaki kümeler daha sonra bir araya getirilmektedir. Kümelerin uzaklıkları ise cosine (kosinüs) uzaklığı kullanılarak ve kümeler arasındaki bağlantı ise Ward yöntemi kullanılarak hesaplanmıştır. Hiyerarşik kümeleme yönteminde ele alınan niteliklerin birlikteliği gösteren görselleştirme araçlarından ise dendrogram kullanılarak, araştırmada elde edilen konuların hangilerinin yorumlarda bir arada ele alındığı gösterilmiştir. Dendrogramda bulunan veriler, veri setindeki yorumlarda ele alınma biçimine göre birbirleri ile uzaklık hesaplaması yapılarak, hangi konuların bir arada değerlendirildiği ve ne aşamada tüm konuların bir araya geldiği görülmektedir. Bu şekilde araştırmada ortaya çıkan konuların hangilerinin ilk aşamada ele alındığı ve daha sonraki aşamalarda hangi kümeler ile beraber görüldüğü gözükmemektedir.

Tablo 4.10' da konuların uzaklık matrisi hesaplamaları gösterilmektedir. Bu tabloda veri setindeki tüm yorumlar ele alınarak gösterilmiştir. Birbirleri ile en yakın ilişkide olan en yakın konu önbüro ve rezervasyon olduğu görülmektedir. Buna göre otellerde konaklayan konuklar yaptıkları yorumlarda bu iki konuyu diğer konulara göre daha fazla bir arada ele aldığı görülmektedir. Ayrıca odalar ve önbüro konularının da yine bir arada ele alınma durumunun ortaya çıktığı görülmektedir. Ayrıca ön büro ile yiyecek içecek konusunun bir arada olma durumu daha sık olarak görülmektedir. Rezervasyon ile olanaklar konusu da yine bir arada daha çok geçmiş olan konular olarak görülmektedir. Birbirlerine en uzak olarak değerlendirilen konu ise önbüro ve animasyon olduğu görülmektedir.

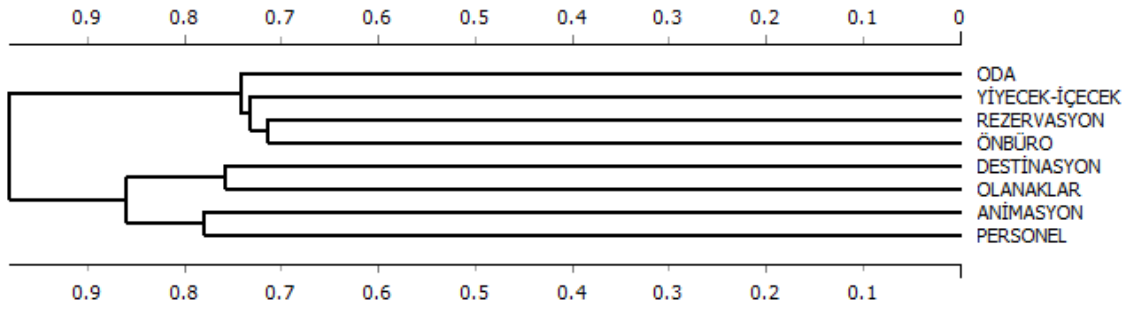
Tablo 4.10 Uzaklık matrisi

Konular	Olanaklar	Personel	Önbüro	Rezervasyon	Destinasyon	Animasyon	Oda	Yiyecek İçecek
Olanaklar		0,795	0,822	0,734	0,758	0,811	0,793	0,767
Personel	0,795		0,859	0,783	0,82	0,781	0,867	0,857
Önbüro	0,822	0,859		0,717	0,843	0,891	0,72	0,725
Rezervasyon	0,734	0,783	0,717		0,798	0,836	0,739	0,731
Destinasyon	0,758	0,82	0,843	0,798		0,836	0,787	0,816
Animasyon	0,811	0,781	0,891	0,836	0,836		0,865	0,872
Oda	0,793	0,867	0,72	0,739	0,787	0,865		0,746
Yiyecek İçecek	0,767	0,857	0,725	0,731	0,816	0,872	0,746	

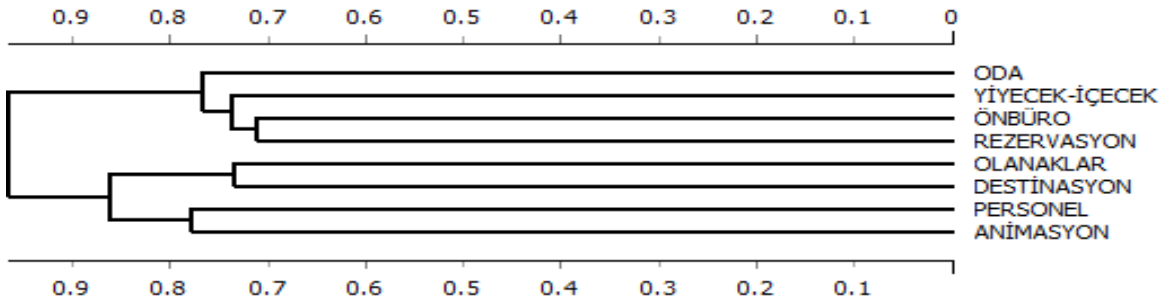
Şekil. 4.44’te hiyerarşik kümeleme analizinin dendogram grafiğinde gösterilmektedir. Eğer gözlemlenen konu kümeleri 0.85 değerinde kesilirse, veri setinde genel olarak 2 kümenin olduğu, yani konuların genel anlamda iki kümeye ayrıldığı görülmektedir. Bu kümelerin ilki oda, yiyecek ve içecek, önbüro ve rezervasyon konuları olurken, diğer küme ise olanaklar, destinasyon, personel ve animasyon olmaktadır. Bu kümeler içerisindeki konuların birbirlerine olan uzaklıklarına bakıldığında, konuların veri setindeki yorumlardaki geçme durumuna göre, birbirleri ile en fazla önbüro ve rezervasyon konusunun ele alındığı görülmektedir ve bu iki grup ilk aşamada bir küme oluşturduğu görülmüştür. Bu ikili gruba daha sonra yiyecek içecek dahil olmakta ve bu aşamayı oda konusu takip etmektedir. Diğer kümeye bakılacak olduğunda, olanaklar ve destinasyon konularının bir arada değerlendirildiği, personel ve animasyon konularının ise bir arada değerlendirildiği görülmekte, daha sonra bu iki grup bir arada yeni bir küme oluşturduğu görülmektedir.

**Şekil 4.44** Tüm verilerin hiyerarşik kümeleme analizi

Şekil 4.45, 4.46 ve 4.47’de ise araştırmadaki konular, duygu sınıflarına göre ayrılarak her bir duyguya göre oluşturulan kümeler gösterilmektedir. Şekil x de pozitif duygu sınıfına sahip olan konuların hiyerarşik kümeleri gösterilmektedir. Veri setinde pozitif yorumların, nötr ve negatif yorumlara göre çok daha fazla olmasından dolayı, pozitif yorumlar ile tüm veri setindeki yorumların ele alınarak yapılan kümeleme analizindeki sonular birbirlerine benzerlik göstermektedir. Bu durum, nötr sınıfı için de benzerlik göstermektedir. Yapılan yorumlarda negatif ya da pozitif eğilime sahip olmayan yorumların da pozitif duygu sınıfına benzer bir durum teşkil ettiği görülmektedir.



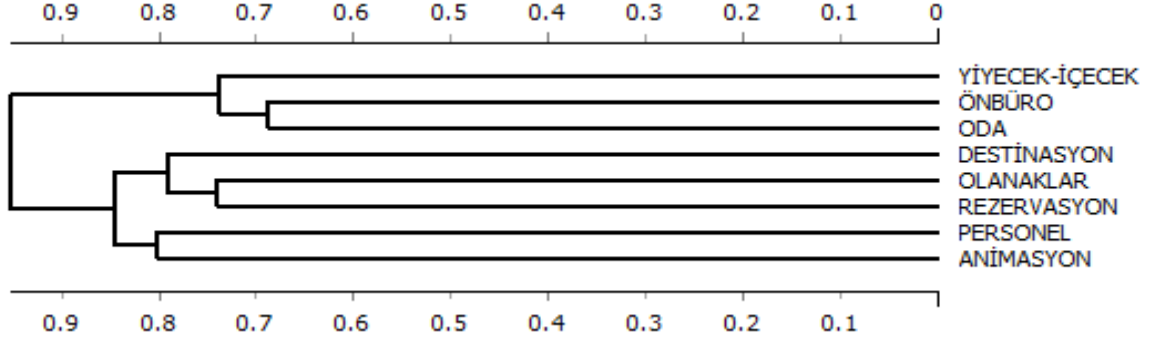
Şekil 4.45 Pozitif duygu sınıfının hiyerarşik kümeleme analizi



Şekil 4.46 Nötr duygu sınıfının hiyerarşik kümeleme analizi

Negatif duygu sınıfına ait yorumların hiyerarşik kümeleme analizi yapıldığında ise, pozitif ve nötr yorumlara göre bazı farklılıkların olduğu görülmektedir. Genel olarak kümeleme analizine bakıldığında, 0.85 değerinde kesim gerçekleşirse oluşturulacak 2 küme yiyecek-içecek, önbüro ve oda konuları olacaktır. Diğer kümedeki konular ise destinasyon, olanaklar, rezervasyon, personel ve animasyon olarak görülmektedir. Negatif yorumlarda önbüro ve oda konusu ilk önce birlikte alınmakta, daha sonra buna yiyecek içecek konusu girmektedir. Diğer grupta ise ilk olarak ele alınan ikili grup olanaklar ve rezervasyon olmaktadır. Araştırmada konu modelleme ile

oluşturulan konulara yapılan duygu analizi sonuçlarında, en fazla negatif olarak ele alınan konular önbüro ve oda konusu olarak ortaya çıkmıştır ve bu konular negatif yorumlarda da beraber ele alındığı görülmektedir.



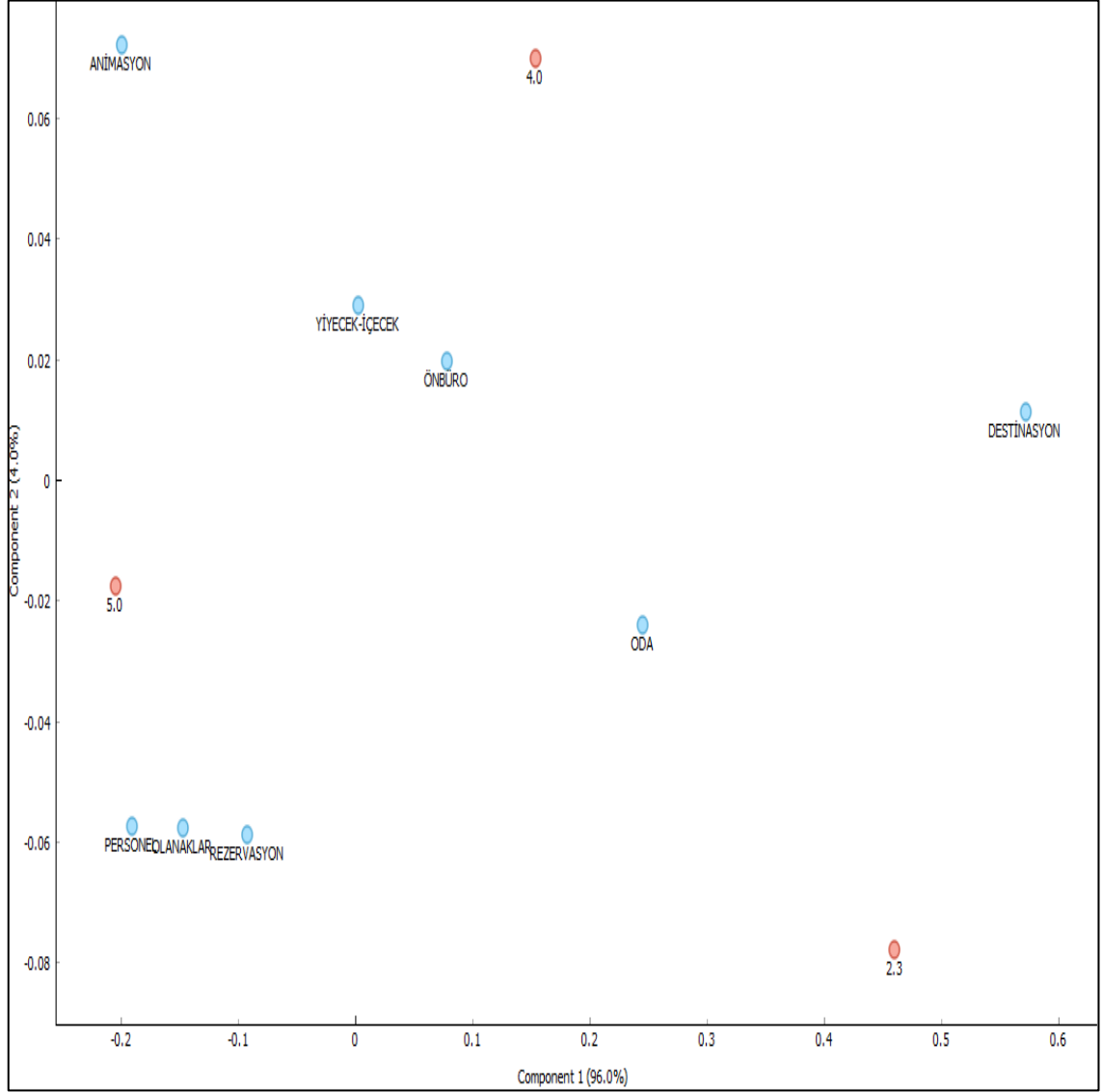
Şekil 4.47 Negatif Duygu Sınıfının Hiyerarşik Kümeleme Analizi

4.9 Uyum Analizi Sonuçları

Araştırmada otel sınıfları, konular, seyahat tipleri ve duygu sınıflarının birbirleri ile yapmış oldukları ilişkileri göstermek için uyum analizi yapılmıştır. Uyum analizi, kategorik değişkenlerin olduğu veri setlerinde bu değişkenlerin birbirleri ile ilişkisini incelemekte kullanılan yöntemlerden birisidir (Üzgören, 2007). Araştırmada uyum analizi yapılacak olan veri tipleri, konular hariç olmak üzere kategorik değişken türündedir. LDA analizi sonucu her bir yorumun konularının hesaplanması ile her yorumun uyum analiz kapsamında ele alınması için kategorik hale getirilmesi gerekmektedir. Örnek vermek gerekirse, bir yorumda bulunan olanaklar konusunun değeri %50 ve üzerinde değer alırsa, bu yorum olanaklar sınıfına atanmıştır. Bu şekilde yorumlar kategorik hale getirilerek uyum analizine hazır hale getirilmiştir. Uyum analizinde x ve y eksenlerinde inertia değeri belirtilmiş olup, iki eksenli bir harita ile tüm uyum analizleri gösterilmiştir. Araştırmada 2 ve 3 yıldızlı otel sayısı az sayıda olmasından dolayı bu kategorideki oteller birleştirilmiştir. Ayrıca seyahat tipinde iş dolayısı ile seyahat eden kişi sayısı da az miktarda olduğu için ve bazı yorumlarda seyahat tipi belirtilmediği için, seyahat tiplerinden iş ve tanımsız olanlar uyum analizi kapsamında alınmamıştır.

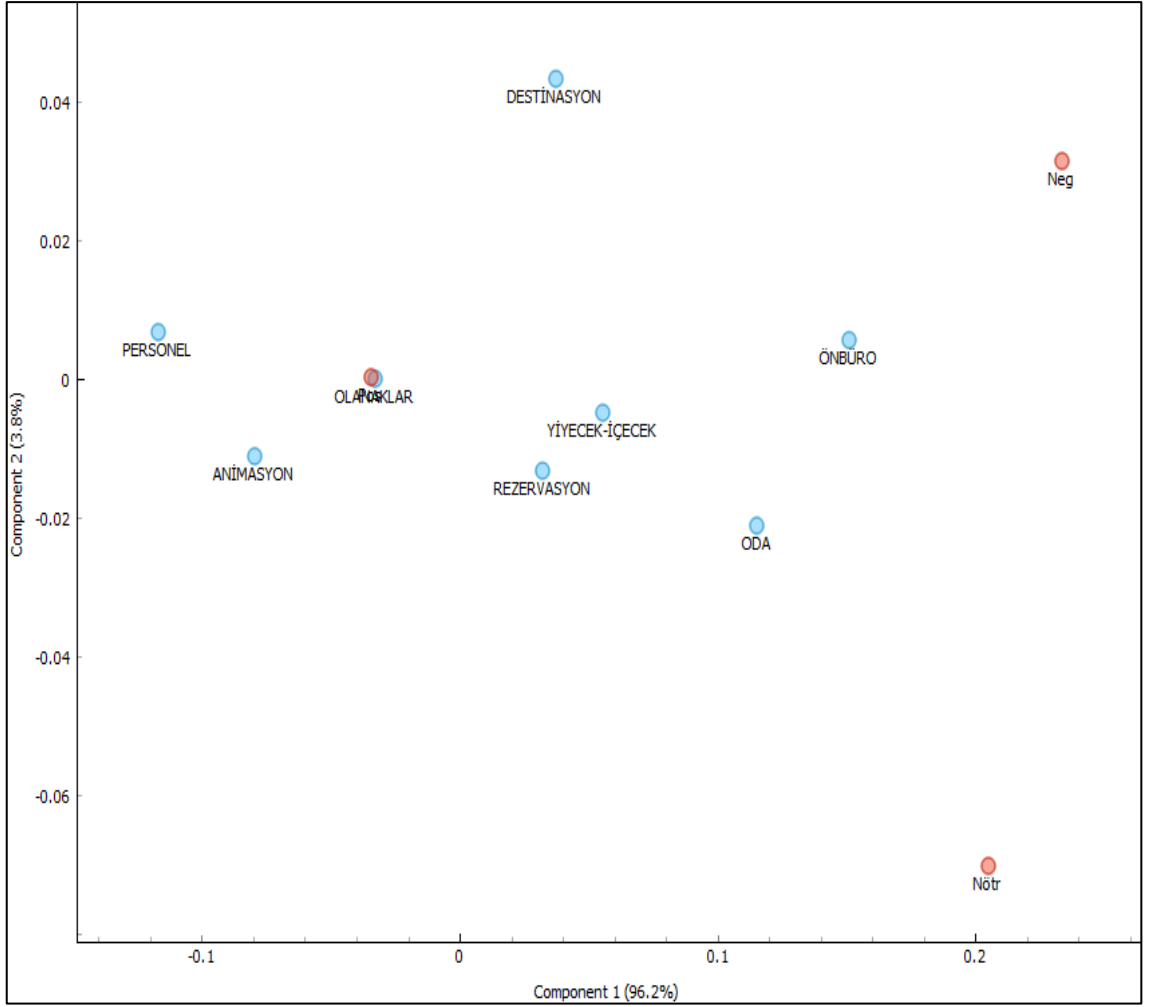
Şekil 4.48'de araştırmadaki konulara ve otel sınıfına göre uyum analizinin sonuçları gösterilmektedir. 5 yıldızlı oteller ile ilişkili olan konuların daha çok personel, olanaklar ve rezervasyon konularında olduğu görülmektedir. Yiyecek içecek, önbüro ve animasyon konusu ise daha çok 4 yıldızlı otellerle birlikte ilişkilendiği görülmektedir ancak animasyonun ilişki düzeyi daha zayıf olarak haritaya yansımaktadır. Destinasyon

ve oda konularının ise 2 ve 3 yıldızlı otellere diğer otel gruplarına göre birlikte olduğu görülmektedir.



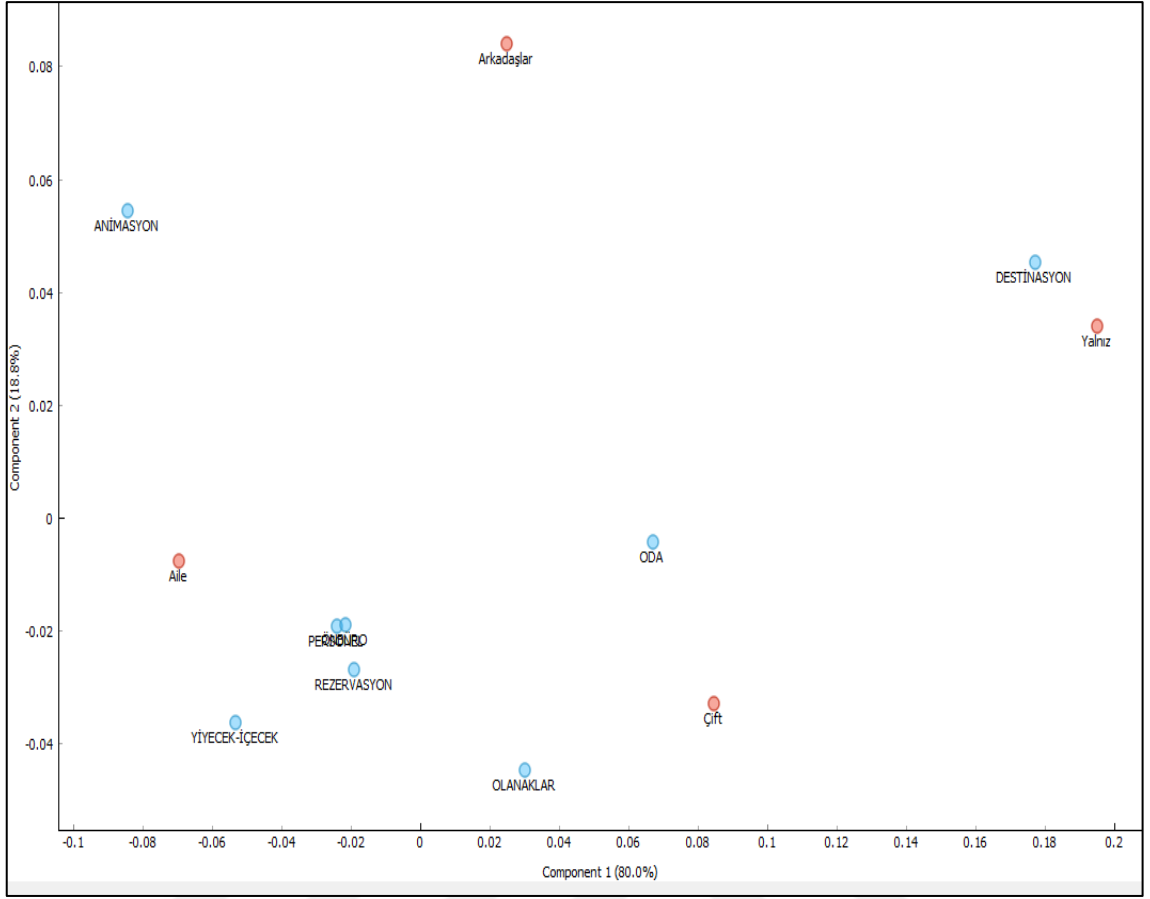
Şekil 4.48 Konu ve otel sınıfı arasındaki uyum analizi

Şekil 4.49’de ise konuların duygu sınıfı ile ilgili uyum analizi çalışması gösterilmektedir. Olanaklar konusu ile pozitif sınıf arasında çok yakın bir ilişki olduğu görülmektedir. Ayrıca pozitif sınıf ile personel, animasyon, rezervasyon ve yiyecek-içecek konuları ile ilişki görünmektedir. Yiyecek-içecek ve rezervasyon, animasyon ve personel ikilileri arasında da ayrıca ilişki görünmektedir. Tüketiciler, bu ikili konuları beraber değerlendirme eğiliminde olduğu söylenebilir. Destinasyon konusu ise daha ayrı bir noktada değerlendirildiği görülmektedir. Negatif sınıfa yakın olan konunun ise daha çok önbüro olduğu gözükmemektedir.



Şekil 4.49 Konu ve duygu sınıfı arasındaki uyum analizi

Şekil 4.50’de ise konular ve seyahat tipleri uyum analizi haritasında gösterilmektedir. Aile grubu ile personel, önbüro, rezervasyon ve yiyecek-içecek konusunun daha çok bir arada olduğu görülmekle beraber, özellikle önbüro ve personel konusunun birliktelik düzeyi yüksek görünmektedir. Özellikle aileler, bu iki konuyu beraber değerlendirme eğiliminde olduğu söylenebilir. Çifler ise daha çok oda ve otel olanakları ile ilgili konularla ilişkili olmaktadır. Yalnız seyahat edenler ise destinasyon konusu ile daha çok ilgilenmekte olduğu görülmektedir. Arkadaşları ile beraber seyahat eden kişiler ise destinasyon ve animasyon konuları ile diğer konulara göre az derecede de olsa bir ilişkide olduğu görülmektedir.



Şekil 4.50 Konu ve seyahat tipi arasındaki uyum analizi

4.10 Lojistik Regresyon Analizi Sonuçları

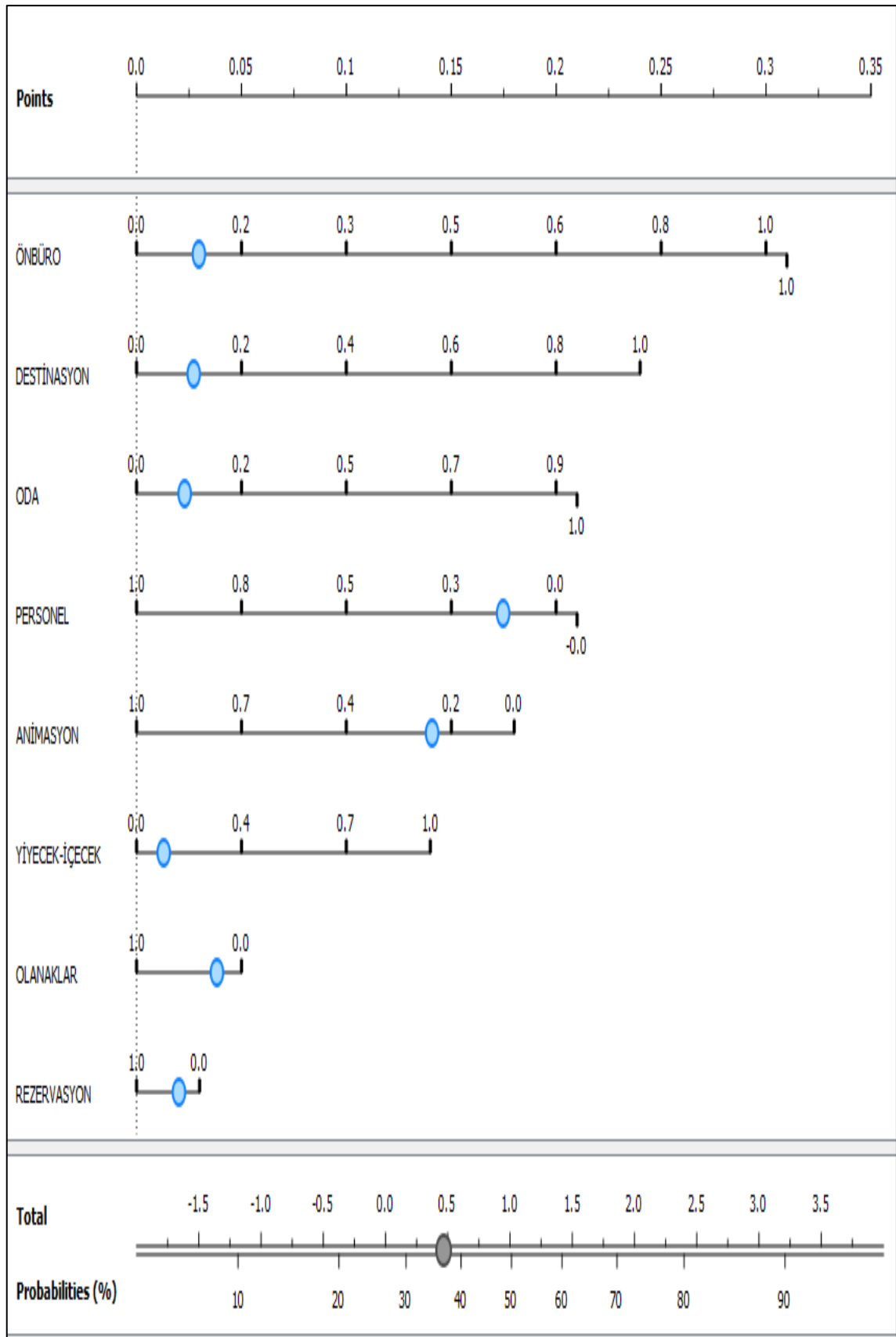
Araştırmada ortaya çıkarılan konuların hangi duygu türüne ne kadar etkilediği ile ilgili olarak lojistik regresyon analizi yapılmıştır. Lojistik regresyon analizinde, her bir yorumda bulunan konuların, duygu sınıfına etkisi araştırılmıştır. Bu araştırmada lojistik regresyon analizini görsel olarak anlatan nomogramlar her bir duygu sınıfına göre kullanılmış olup, veri setindeki hangi konunun hedef değişkeni daha fazla etkilediğini göstermekte ve bu sınıftaki olasılıksal etkileri belirtmektedir. Nomogramlardaki konularda bulunan noktalar ise, veri setinin o konudaki mevcut yerini göstermektedir. En üst sırada olan konudaki artış, mevcut duygu sınıfının olma olasılığını en fazla etkileyen konudur. En alt sıradaki ise mevcut duygu sınıfını en az derecede etkilemektedir. Şekil 4.51, 4.52 ve 4.53'te sırasıyla negatif, nötr ve pozitif sınıflara etki eden konular etki olasılığına göre büyükten küçüğe gösterilmektedir. Lojistik regresyon analizinde,

Tablo 4.11'de ise metin setinden çıkartılan konuların, duygu sınıflarına göre etkileri gösterilmektedir. İlk 3 sütunda lojistik regresyon analizinde her bir konunun

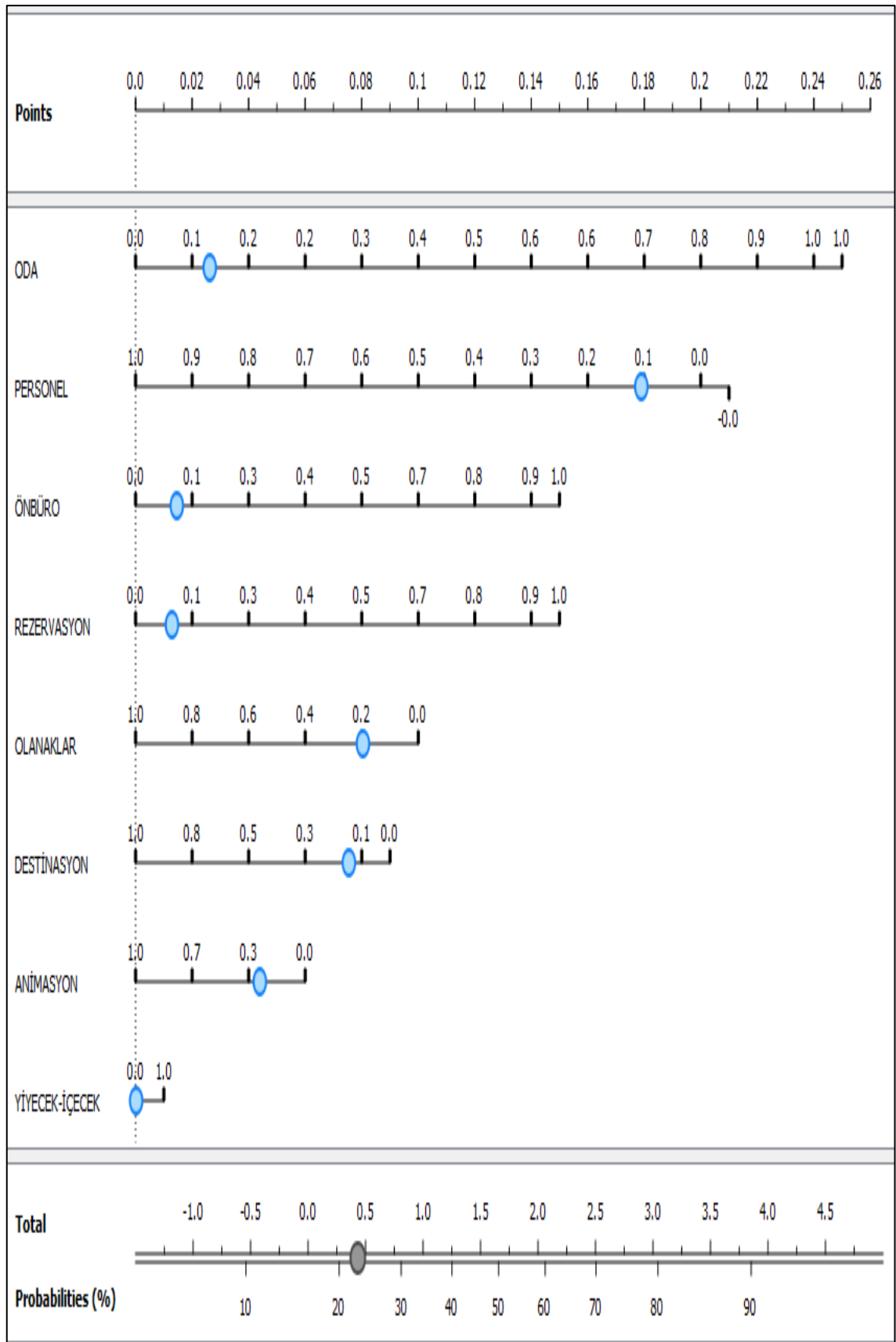
duygu sınıfına göre katsayıları ve yönleri gösterilmektedir. Son 3 sütunda ise bu katsayıların log odd ratio değerleri hesaplanmıştır. Buna göre olanaklar konusu ile ilgili 1 birimlik değişim, en fazla pozitif yorum grubuna etki ettiği görülmüştür. Konaklama işletmelerine olanaklar konusunda yorum yapıldığı takdirde, bu konu ile ilgili kullanılacak kelimeler baz alındığında bu durum kişilerin yapmış olduğu yorumların pozitif olma olasılığını 1,14 kat arttıracaktır. Personel konusuna bakıldığında ise, yine en yüksek değişim değeri pozitif sınıfta olduğu görülmektedir. Pozitif yorumdaki bir birimlik değişim, yorumların pozitif olma olasılığını 1,52 kat arttıracakları görülmektedir. Önbüro konusunda ise durum tersine işlemektedir. Önbüro konusunda bir kişi yorum yaptığı takdirde, bu yorumun negatif olma olasılığı daha yüksek görünmektedir. Benzer durum, önbüro konusundan daha az seviyede de olsa, destinasyon konusu için de geçerli olmaktadır. Rezervasyon konusunda yapılan yorumlar ise daha çok nötr duygu sınıfına etki ettiği görülmektedir. Animasyon yorumlarının artması ise bir yorumdaki duygunun pozitif olma olasılığını diğer duygu sınıflarına göre daha fazla arttırmaktadır. Yiyecek içecek konusunun geçmesi ise negatif olarak yorumların sınıflanma olasılığını daha fazla attırmaktadır. Odalar konusuna yapılan yorumların ise, yorumun pozitif olma olasılığını düşürücü bir etkiye sahip olduğu görülmektedir.

Tablo 4.11 Lojistik regresyon analizi sonuçları

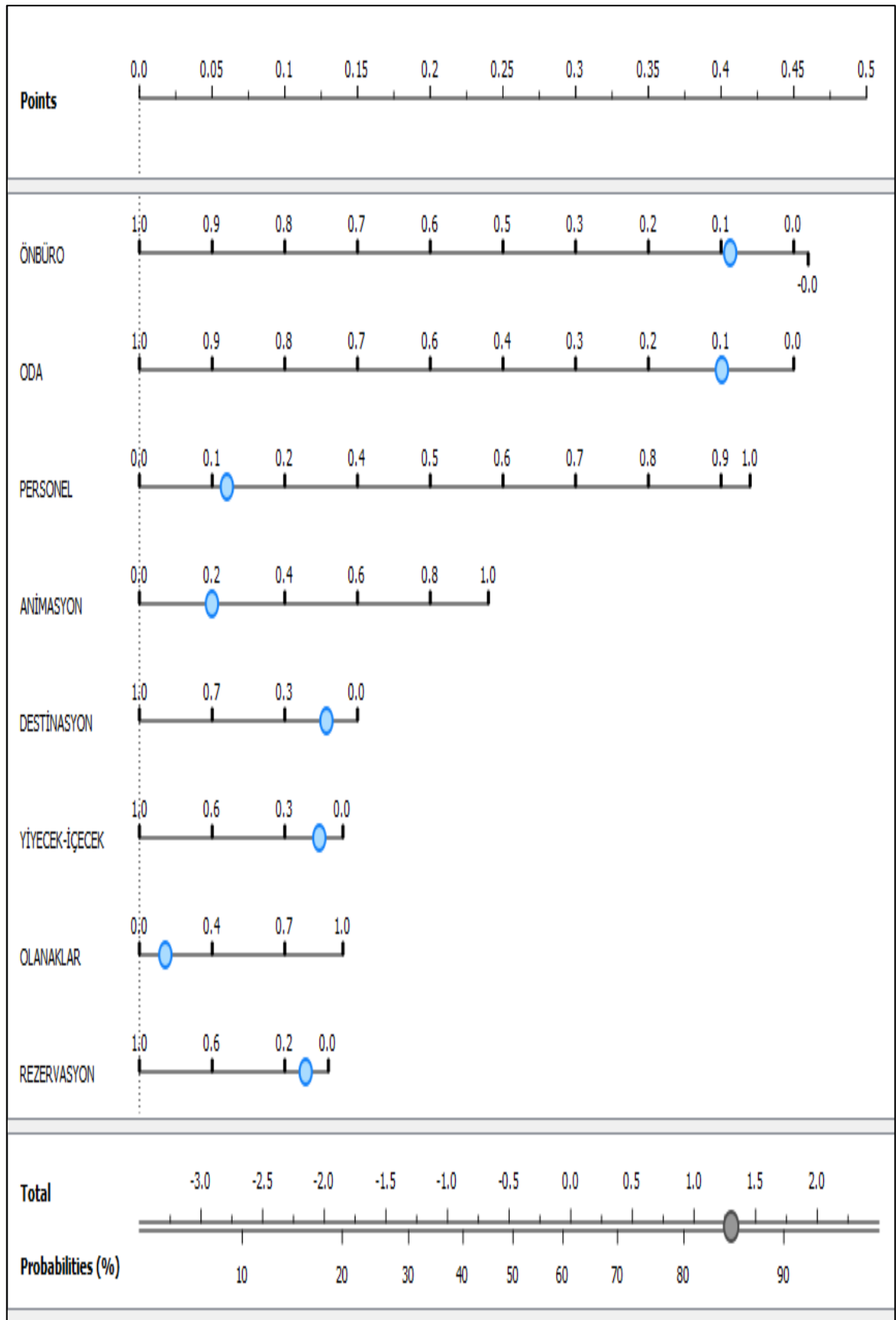
Konular/Katsayılar	Negatif	Nötr	Pozitif	LoRNegatif	LoRNötr	LoRPozitif
Kesişim Değeri	-0.56277	-1.21373	1.776496	0.56963	0.297088	5.909114
Olanaklar	-0.04462	-0.09374	0.138363	0.956359	0.910518	1.148393
Personel	-0.2072	-0.21224	0.419446	0.812853	0.80877	1.521118
Önbüro	0.305512	0.149948	-0.45546	1.35732	1.161774	0.634156
Destinasyon	0.234859	-0.08685	-0.14801	1.26473	0.916816	0.862422
Rezervasyon	-0.0224	0.149259	-0.12686	0.977851	1.160974	0.880856
Animasyon	-0.18088	-0.05645	0.237329	0.834536	0.945114	1.267858
Yiyecek-İçecek	0.13622	0.00203	-0.13825	1.145934	1.002032	0.870881
Oda	0.211326	0.241892	-0.45322	1.235316	1.273656	0.635579



Şekil 4.51 Negatif duygu sınıfının nomogramda gösterimi



Şekil 4.52 Nötr duygu sınıfının nomogramda gösterimi



Şekil 4.53 Pozitif duygu sınıfının nomogramda gösterimi

5. TARTIŞMA, SONUÇ VE ÖNERİLER

Turizm, içerisinde bulundurduğu farklı bileşenleri olan ve bu bileşenlerin birbirleri ile entegre bir şekilde faaliyet gösteren sistemler bütünüdür (Jakulin, 2017). Bu sistemi oluşturan en önemli bileşenlerden birisini de konaklama işletmeleri oluşturmaktadır. Konaklama işletmelerini oluşturan içeriklerden birisi olan otel işletmeleri, ülkelerin, bölgelerin ve destinasyonların turizm sektörü içerisinde önemli paya sahiptir (Holjevac, 2003). Turizm sistemine katılan kişilerin konaklama yapmış olduğu otel işletmelerinde geçirmiş olduğu deneyim, bu kişilerin destinasyonlara olan algılarını etkilemektedir. Bu deneyimden insanların pozitif olarak ayrılması, otel işletmelerine ve doğal olarak destinasyona tekrar ziyaret etmesindeki önemli etmenlerden birisini oluşturmaktadır (An vd., 2019).

Otel işletmelerinin sunmuş olduğu hizmetlerin tüketicilere olan etkisinin araştırılması, otel işletmelerinin geleceği açısından üzerinde durulması gereken bir konudur. Teknoloji ve internet imkanlarının gelişimi neticesinde turizm alanında artan çevrimiçi platformlar, otel işletmelerinin hizmet kalitesinin ölçülmesinde ve tüketici memnuniyetlerinin boyutlarının ortaya çıkarılmasında giderek aktif bir rol oynamaktadır (Bore vd., 2017). İnternet kullanımının giderek yaygınlaşması ile, bu platformlara olan talep de bu doğrultuda artacak ve bu platformlardaki görüşler artan rekabet koşullarında giderek ön plana gelecektir (Alrawadie vd., 2019).

Bu noktada otel işletmelerinde kalan turistlerin bu platformlarda yaptığı yorumların etkili bir şekilde incelenmesi ve turistlerin yazdıkları bu yorumlarının içeriklerinin ortaya çıkartılması için büyük veri analizi kapsamında değerlendirilen yöntemlere ihtiyaç duyulmaktadır. Çevrimiçi platformlara olan talebin artışı göz önüne alındığında, büyük veri analizlerinin sağlamış olduğu fırsatlar, otel işletmelerinde çalışan yöneticilerin alacağı kararlarda yardımcı olabilmektedir (Mousavian vd., 2023). Böylece otellerin hizmet kalitesinin arttırılması, pazarlama faaliyetleri, müşteri ilişkileri yönetimi gibi alanlarda sağlıklı kararlar vermesi sağlanabilir (Sakas vd., 2022).

Kullanıcı tarafından oluşturulan içeriklerin otel ve destinasyon yönetimlerinin karar almasına yardımcı olacak modellerden birisini ise metin madenciliği ve hesaplamalı dil bilim alanında yaşanan gelişmeler ve bu gelişmeler sonucunda ortaya çıkan yöntemlerdir. Metin madenciliği ile ilgili içerikler, başta makine öğrenmesi yöntemleri ile etkin bir şekilde analiz edilerek, otel işletmelerinin yönetsel olarak alması gereken kararlara yol gösterici nitelikte olabilmektedir (Sahebi vd., 2022).

Büyük veri ve metin madenciliği alanında gelişmeler dikkate alınarak, kullanıcı tarafından oluşturulan içerikler ve çevrimiçi seyahat platformlarının önemi göz önüne alınarak yapılan bu çalışmada, Alanya’ da bulunan otellerin Tripadvisor web sitesindeki yorumlarının içerikleri, makine öğrenmesi modelleri ile analiz edilmiştir. Özellikle metin madenciliği ve doğal dil işleme alanında kullanılan konu modelleme ve duygu analizi çalışmaları yapılarak, yorumlardaki içeriklerin başlıca hangi konuları barındırdığı, bu konuların tüketicilerin memnuniyetine ne şekilde etki ettiği ortaya çıkarılmıştır.

Alanya destinasyonu hakkında konu modelleme ve duygu analizi çalışmaları literatürde az sayıda olduğu görülmüştür. Bu çalışmanın ilgili literatüre düşünülen katkısı, konuların tüketicilerin memnuniyetine ne şekilde etki ettiği ve hangi konuların birlikte ele alındığının ortaya çıkartılmasıdır. Çevrimiçi platformlarda, tüketiciler yazmış oldukları yorumlarda birden fazla konuyu ele almaktadır. Bu konuların tüketiciler tarafından olumlu ya da olumsuz anlamda değerlendirilmesi ise, otellerin sağlamış olduğu hizmet kalitesinin performansının ölçülmesinde önemli bir basamaktır (Tuominen, 2011).

Araştırmadaki bulgulara bakıldığında, en fazla yorumun beş yıldızlı otellere yapıldığı görülmektedir. Bu durum, otellerin kapasiteleri göz önüne alındığında beklenen bir durumdur. Ancak yıllara göre otel sınıflarına gelen yorumlara bakıldığında, beş yıldızlı otellere gelen yorumların sayısı ile iki, üç ve dört yıldızlı otellere gelen yorumların farkı giderek artmakta olduğu görülmektedir. Özellikle herşey dahil konseptine sahip beş yıldızlı oteller, uyguladığı ücret politikası ve tesis içerisinde vermiş olduğu imkanlar ile daha fazla turist çekebilmektedir. Ancak bu durum, Türkiye’ ye gelen turistlerin kişi başına düşen harcamaları göz önüne alındığında, ülkeye daha az para harcayan turistlerin gelmesine ve bu durumun da ülke imajına negatif etki etmesine sebep olabilmektedir (Menekşe, 2005).

Veri setinde ayrıca görülen bir diğer husus, özellikle iş konusunda Alanya’yı az sayıda ziyaret eden kişinin bulunduğudır. Alanya’ daki otellerin konsepti genellikle aileler ve çiftler üzerine olduğu görülmektedir. Ancak iş konusunda seyahat eden turistler, ortalama olarak 1293 dolar para harcamaktadır (Hotelbusiness, 2019). Türkiye’ nin turist başı harcamasına bakıldığında ise, 2004 yılında 768 dolar iken, 2022 yılında bu sayı 589 dolara inmiştir (KTB, 2022). Alanya’ nın sahip olduğu tarihi ve doğal çekicilikleri ve otellerin sağlamış olduğu imkanlar göz önüne alındığında, iş seyahatini arttırıcı kongre, seminer, fuar, etkinlik gibi organizasyonların daha fazla gelişmesi

gerekmektedir. Antalya ve Gazipaşa havalimanlarının olması ve transfer hizmetleri sağlayan firmaların sayısı göz önüne alındığında, iş seyahati kapsamında ele alınacak organizasyonların geliştirilmesi ve bu alanda pazarlama faaliyetlerinin yapılması, daha nitelikli turistlerin gelmesini sağlayabilir.

Araştırmada ortaya çıkan sonuçlardan birisi ise, otel sınıfının arttıkça, pozitif yorum yapma eğiliminin arttığıdır. Bu kapsamda otellerin sahip olduğu özelliklerin gücü (örneğin tüketicilere sağladığı olanaklar, hizmet çeşitliliği gibi), kişilerin yapmış olduğu yorumları pozitif yönde etkilemekte olduğu söylenebilir. Dolayısıyla otellerin sahip oldukları olanakların seviyesi ve çeşitliliği, gelen kişilerin memnuniyetini arttıran bir durum olarak ortaya çıkmaktadır. Benzer durumun ortaya çıktığı çalışmalara bakıldığında, otel hizmet kalitesinin müşteri memnuniyetine en fazla etkisi olan konuların başında olanaklar gelmektedir (Calheiros vd., 2017; Philips vd., 2017; Ying vd., 2020).

Bu durum, çalışmada yapılan uyum analizi sonuçları ile de gösterilmiştir. Pozitif yorumlar ile en yakın ilişkide olan konu, olanaklar konusu olarak görülmektedir. Bu noktada otel işletmeleri, olumlu yorum alabilmeleri için otel işletmelerindeki olanakları nitelik ve nicelik olarak arttırmaları gerekmektedir. Ancak yapılan bir diğer analizde, bireylerin olumlu yorum yapmasına en fazla etki eden konunun ise personel olduğu görülmektedir. Özellikle küçük ve orta büyüklükteki işletmeler, sahip oldukları finansal ve istihdam bakımından kısa ve orta vadede oteldeki olanakların çeşitliliğini arttıramayabilir. Ancak personel konusuna yapacakları eğitimler ve personel seçimindeki dikkatli olmaları, bu işletmeleri değerlendiren tüketicilerin daha fazla olumlu yorumda bulunmasına katkı sağlayabilirler. Otel işletmelerinde personel seçimi, sürdürülebilir bir kalitenin yerleşmesinde önemli etkisi olan etmenlerden birisidir (Akyurt; 2021).

Konular ile duygu analizi skorları arasında yapılan korelasyon analizi sonuçlarına göre önbüro, oda, rezervasyon ve yiyecek-içecek arasında negatif yönlü, animasyon, olanaklar, personel ve destinasyon konularının ise pozitif yönlü bir ilişkinin olduğu saptanmıştır. Korelasyon düzeyine bakıldığında ise en yüksek negatif ilişki duygu skoru ile önbüroda olduğu görülmektedir. Önbüro, otellerde müşterilerin şikayetlerini ilk olarak ifade ettiği yer olmaktadır. Ancak yapılan bu araştırmada, önbüro ile ilgili olan düşünceler, kişilerin değerlendirmelerini negatif olarak etkileyen bir durumdur. Personel konusu, araştırmada en fazla pozitif etkiye sahipken, önbüro konusu ise en fazla negatif etkiye sahip olduğu görülmüştür. Önbüro ile ilgili kelimelere

bakıldığında, dil, davranış, ödeme gibi kelimelerin olduğu görülmektedir. Bu sebeple, otellerin önbüro personeli seçiminde daha dikkatli davranması, bu departmanda çalışan kişilerin diğer departmanlara göre daha hızlı bir şekilde eğitilerek olası negatif etkilerin indirilmesi gerekmektedir. Özellikle otellerdeki personel ile ilgili olan departmanların başında gelen önbüro, tüketicilerin memnuniyetini en fazla etkileyen unsurlardan biri oluşturmaktadır (Hartline & Jones, 1996).

Oda konusu da aynı şekilde tüketicilerin negatif yorum yapmasına neden olan konulardan birini oluşturmaktadır. Ancak oda konusundaki kelimelere bakıldığında, odalar ile ilgili olarak temizlik ile beraber bazı teknik özelliklerin de ön plana çıktığı görülmüştür. Bu sebeple odalar ile ilgili olarak otel işletmeleri, kat hizmetleri personeli ile beraber, odalar ile ilgili olan ekipmanları ve teknik servisi de güçlendirmeleri gerekmektedir. Odalar konusu, tüketicileri ön bürodan sonra en fazla negatif yorum yapmasına sebep olan ikinci konu olarak araştırmada çıkmıştır. Bu sebeple bu alanla ilgili sorunların giderilmesi, tüketicilerin olumlu yorum yapmasını sağlayacak bir etmen olarak görülmektedir.

Animasyon konusu, araştırmadaki yorumlarda sıkça geçen konuların başında olduğu görülmüştür. Alanya’ daki otellerin çoğunlukla aileler tarafından seçildiği veri setinde görülmüştür. Bu sebeple animasyon konusunda çocuk, havuz, animasyon ekibi, klüp, eğlence gibi kelimeler sıkça geçmektedir. Araştırmadaki konulardan en fazla memnuniyet duyulan konu animasyon olmuştur. Etki olarak ise personelden sonra kişilerin olumlu yorum yapmasına en çok etki eden konu animasyon olarak görülmektedir. Bu sebeple aile ve çocuk konularına otel işletmelerinin daha fazla eğilim göstermesi, bu işletmelere yapılan yorumların pozitif olma olasılığını arttıracığı söylenebilir. Ayrıca otellere yapılan yorumların duygu analizi sonucunda otel sınıfı arttıkça, pozitif duygu oranı negatiflere göre artış gösterdiği görülmektedir. Bu sebeple, özellikle iki ve üç yıldızlı oteller, personel ve personel ile ilgili olan animasyon konusuna eğilim göstererek ve bu departmanları geliştirerek, olanaklar konusuna göre daha kısa sürede tüketicilerin pozitif yönde karar vermesini sağlayabilecek stratejiler geliştirebilir. Antalya Belek’ ye bulunan otellere yapılan araştırmada, animasyon konusu, otellerde konaklama yapan kişileri en fazla memnun eden etmenlerin başında gelmektedir. (Davras & Caber, 2019).

Duygu türleri kapsamında birbirlerine en yakın gruplar yiyecek-içecek, destinasyon ve rezervasyon konularıdır. Yiyecek-içecek departmanında nötr yorumlar ile negatif yorumlar birbirlerine daha yakın durmakla birlikte, pozitif sınıf ile

aralarındaki mesafe az düzeydedir. Bu noktada oteller, yiyecek-içecek departmanı ile ilgili yapıcı kararlar alarak, bu alandaki yorumların pozitif döndürme olasılığı daha yakın görüldüğü söylenebilir. Benzer durum rezervasyon konusu için de geçerli olmaktadır. Rezervasyon konusunda ayrıca otel sınıfı arttıkça, bu konudaki düşünceler negatif yöne doğru gitmektedir. Özellikle rezervasyon konusundaki transfer, ulaşım, bekleme ile ilgili kelimelerin çıktığı görülmüştür. Bu imkanları sağlayan oteller ise genellikle dört ve beş yıldızlı olan otellerdir. Bu sebeple bu otel grupları, otel dışında sağlamış olduğu hizmetlere de önem vermesi gerekmektedir. Rezervasyon konusunda acentalar ile daha fazla iletişime geçerek, tüketicilerin yapmış olduğu rezervasyon kapsamında bekledikleri hizmeti yerine getirmeleri önemli bir durumdur.

Destinasyon konusunda yapılan yorumların duygu sınıflarında ise otel sınıfı arttıkça, destinasyon ile ilgili pozitif düşüncelerin düştüğü görülmektedir. Bu durumda, özellikle 2 ve 3 yıldızlı oteller, şehir merkezine daha yakın konumda oldukları için, destinasyonda bulunan yapıları ziyaret etme ve destinasyondaki olanaklar yararlanması, diğer gruptaki otellere göre daha fazla olduğundan dolayı, bu durum kişilerin yorumlarının duygu sınıflarına etki ettiği söylenebilir. Turistik destinasyonlarda bulunan otellerin konumunun ulaşım ağlarına ve tur imkanlarına yakın olması, alışveriş ve yiyecek içecek işletmelerine yakınlığı, turistlerin memnuniyetine olumlu düzeyde etki eden unsurlardır. Özellikle şehir tatili yapan turistlerin başlıca motivasyon kaynaklarının destinasyonların tarihi ve doğal çekiciliklerini görmektir (Algün, 2016). Bu sebeple otellerin destinasyonlardaki konumları, destinasyon hakkında daha olumlu yorum yapmaya iten sebepler olabilmektedir.

Yiyecek içecek konusu ise otel sınıfı arttıkça negatif düşünce eğilimi daha yüksek olduğu bir konu olmuştur. Bu durum, özellikle sınıf sayısı arttıkça otellerin uygulamış olduğu her şey dahil sistemin yeme içme alanında sağladığı imkanlar kapsamında sunulan yiyeceklerin ve içeceklerin kalitesi ve bunları temin edilirken geçen süre, negatif düşünceleri bu alanda arttırabilen unsurlar olabilmektedir. Sınıf sayısı düştükçe, otellerin konuklara sağladığı yiyecek içecek imkanı daha kısıtlı olmaktadır ve bu otellerin konumları şehir merkezine daha yakın bir konumda olduğundan dolayı, bu sınıflarda konaklayan kişiler yiyecek içecek ihtiyaçlarını otel dışında da sağlayabilmektedir. Bilindiği gibi otel sınıfı arttıkça yiyecek içecek beklentisi de otellerin sunmuş olduğu çeşitlilik ve özellikle her şey dahil sistem kapsamında daha fazladır. Bu durumda yiyecek içeceklerde meydana gelen kalite ya da servis sıkıntısı, kişilerin bu alanda daha fazla negatif yorumda bulunmasına sebep olabilmektedir.

Yorumlarda geen konuların birliktelik analizlerinde ortaya ıkan sonularda, rezervasyon konusunda sorun yaşıyan birisi, otele vardığında bu durumu genellikle nbüro departmanı ile konuşmaktadır. Araştırmadaki bulgular neticesinde bu iki grubun birbirlerine daha yakın olması, bu iki konuyu ilgilendiren ve tüketicilerin bu konuları yorumlarda bir arada tuttuğunu göstermektedir. Benzer yakınlık, nbüro departmanı ile odalar ve yiyecek-iecek konusunda da yaşanmaktadır. Rezervasyon konusu da aynı şekilde odalar, olanaklar ve yiyecek iecek ile daha yakın bir şekilde bir arada ele alındığı görölmektedir. Bu noktada otel tarafından kişilere verilen oda tipleri, otelin sağlamış olduđu imkanlardaki sıkıntılar ya da bu olanakların sağlanmadığının tüketiciler tarafından görülmesi, rezervasyon sonrasında yiyecek iecek imkanlarının tüketiciler tarafından beklendiğı gibi olmaması, bu duruma sebep olabilmektedir.

Araştırmadaki uyum ve kümeleme analizlerinden ıkan bir diğeri sonuç ise, negatif duyguların daha yoğun olduđu konuların birliktelikleri, pozitif olan konulara göre daha yakın konumdadır. Dolayısıyla otellere yorum yapan bireyler, yorumlarında daha ok bu konulardaki yaşanan sıkıntıları birlikte ele aldığı görölmektedir. Örneğn animasyon konusuna bakıldığında, bu konunun diğeri konularla uzaklıkları, animasyon dışındaki konularla karşılaştırıldığında daha fazla olmaktadır. Bu durum, animasyonun sahip olduđu uzaklık derecesinden az miktarda da olsa, konuların daha pozitif olarak ele alındığı konular için de geçerli olmaktadır. Özellikle negatif yorumlarda oda ve nbüro, rezervasyon ve otel olanaklarının bir arada olduđu görölmektedir. Bu sebeple oteller, tüketicilerin yapmış olduđu negatif yorumlarda öncelikli olarak bu ikili grupları bir arada alması, tüketicilerin memnuniyetini daha fazla oranda etkileyebilen bir unsur olarak söylenebilir.

Araştırmada ortaya ıkan sonulardan bir diğeri, aile olarak seyahat eden kişiler, ağırlıklı olarak yiyecek iecek, rezervasyon, personel ve n büro konularından bahsetmektedir. Çiftler ise daha ok otel olanakları ve odalar ile ilgilenmektedir. Yalnız seyahat edenler ise destinasyon konusunda daha ok ilişkili olduđu görölmüştür. Arkadaşları ile beraber konakalama yapan kişiler ise, animasyon ve destinasyon konusuna daha ok ağırlık verdiğı görölmektedir. Özellikle ailelerin tercih ettiğı oteller, personel ve nbüro konusunu bir arada değerdirmeleri gerekmektedir. Ayrıca rezervasyon ve yiyecek iecek konularına daha fazla önem vermesi gerektiğı söylenebilir.

Konuların duygu sınıflarına olan etkisinin araştırıldığı lojistik regresyon analizinde, tüketiciler en fazla personel ile ilgili yorumda bulunduđu takdirde, yapılan

bir yorumun pozitif olma olasılığı daha fazla olmaktadır. Bu durumu animasyon ve olanaklar konusu takip etmektedir. Bu üç konuya otel işletmeleri önem verdiği takdirde, yorumların pozitif olma olasılığının daha fazla artacağı söylenebilir. Önbüro konusu ise en fazla negatif yorum meydana getirecek bir durum olarak ortaya çıkmıştır. Bu durumda oteller öncelikli olarak önbüro departmanına yönelmesi gerektiği ortaya çıkmaktadır. Bu şekilde bir yorumun negatif olma olasılığının önüne daha fazla geçebilir. Odalar konusu ve yiyecek içecek konuları da, ön büro konusundan sonra en fazla negatif yoruma neden olmaktadır. Bu konular hakkında otel işletmeleri, yorumların pozitif bir şekilde dönüşebilmesi için önlem almalıdır. Destinasyon konusu da ayrıca negatif yoruma sebep olan bir unsur olarak göze çarpmaktadır. Destinasyon konusu, bilindiği gibi otellerin tek başına müdahale edeceği bir konu olmamaktadır. Ancak destinasyonda turistlerin yaşamış olduğu sorunlar, otellere yapılan yorumlara da negatif yönde etki etmektedir. Bu durumun yaşanmaması için, yerel yönetimlerin destinasyonu iyileştirici faaliyetlerde bulunması önemli bir durum olarak ortaya çıkmaktadır.

KAYNAKLAR

- Afaq, A., Gaur, L., & Singh, G. (2022). A latent dirichlet allocation technique for opinion mining of online reviews of global chain hotels. *3rd International Conference on Intelligent Engineering and Management*, 201-206.
- Akdeniz Turistik Otelciler ve İşletmeciler Birliği, *Turizm İstatistikleri*, <https://aktob.org.tr/istatistik/>. (Erişim Tarihi: 15.06.2023)
- Aksu, M. Ç. & Karaman, E. (2020). FastText ve kelime çantası kelime temsil yöntemlerinin turistik mekanlar için yapılan türkçe incelemeler kullanılarak karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (20), 311-320.
- Akyurt, H. (2021). Otel işletmelerinde personel seçim kriterlerinin ahp yöntemi ile belirlenmesi: Ordu İli üzerine bir araştırma. *Journal of Global Tourism and Technology Research*, 2 (2), 59-71.
- Al Khattab, S. A., & Aldehayyat, J. S. (2011). Perceptions of Service Quality in Jordanian Hotels. *International Journal of Business and Management*, 6 (7), 226-233.
- Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on big data. *Journal of Travel Research*, 58 (2), 175-191.
- Albanese, N. C. (2022). Topic modeling with LSA, pLSA, LDA, NMF, BERTopic, Top2Vec: A comparison. <https://towardsdatascience.com/topic-modeling-with-lsa-plsa-lda-nmf-bertopic-top2vec-a-comparison-5e6ce4b1e4a5> (Erişim Tarihi: 15.06.2023).
- Algün, S. (2016). *Almanca dilini kullanan seyahat 2.0 kullanıcılarına göre otel performans analizi: Alanya alan çalışması*, (Yayınlanmamış Yüksek Lisans Tezi). Sosyal bilimler Enstitüsü, Akdeniz Üniversitesi, Antalya.
- Ali, T., Marc, B., Omar, B., Soulaïmane, K., & Larbi, S. (2021). Exploring destination's negative e-reputation using aspect based sentiment analysis approach: case of Marrakech destination on TripAdvisor. *Tourism Management Perspectives*, 40, 100892.
- Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests satisfaction from the perspective of online hotel reviewers. *International Journal of Culture Tourism and Hospitality Research*, 13 (1), 84-97.
- Amaral, F., Tiago, T., & Tiago, F. (2014). User-generated content: Tourists' profiles on Tripadvisor. *International Journal of Strategic Innovative Marketing*, 1 (3), 137-145.
- Amerikan Kalite Topluluğu, <https://asq.org/quality-resources/quality-glossary/q>. (Erişim Tarihi: 15.06.2023)
- An, S., Suh, J., & Eck, T. (2019). Examining structural relationships among service quality, perceived value, satisfaction and revisit intention for Airbnb guests. *International Journal of Tourism Sciences*, 19 (3), 145-165.
- Appel, O., Chiclana, F., Carter, J., & Fujita, H. (2016). A Hybrid approach to the sentiment analysis problem at the sentence level. *Knowledge-Based Systems*, 108, 110-124.

- Atalay, M., & Çelik, E. (2017). Büyük veri analizinde yapay zekâ ve makine öğrenmesi uygulamaları. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9 (22), 155-172.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2009). Multi-facet rating of product reviews. In *Advances in Information Retrieval: 31th European Conference on IR Research*, 31, 461-472.
- Bagherzadeh, S., Shokouhyar, S., Jahani, H., & Sigala, M. (2021). A generalizable sentiment analysis method for creating a hotel dictionary: Using big data on Tripadvisor hotel reviews. *Journal of Hospitality and Tourism Technology*, 12 (2), 210-238.
- Bahja, M. (2020). Natural Language Processing Applications in Business. Wu, M.X. & Mircea, M. (Edt). *E-Business-Higher Education and Intelligence Applications*, 53-61. IntechOpen Publishing, London.
- Başaran, M. A., Dogan, S., & Kantarci, K. (2020). On modeling of responses generated by travel 2.0 implementation: Fuzzy rule-based systems. *International Journal of Contemporary Hospitality Management*, 32 (4), 1503-1522.
- Bayat, M., Baydaş, A., & Cahit, A. T. L. I. (2015). Hizmet sektörünün kavramsal tanımı ve ulusal ekonomilerdeki önemi. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5 (9), 59-88.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal Of Machine Learning Research*, 3 (1), 993-1022.
- Boonitt, S., & Rompho, N. (2012). Measuring service quality dimensions: an empirical analysis of thai hotel industry. *International Journal of Business Administration*, 3(5), 52-63.
- Bore, I., Rutherford, C., Glasgow, S., Taheri, B., & Antony, J. (2017). A systematic literature review on eWOM in the hotel industry: Current trends and suggestions for future research. *Hospitality & Society*, 7(1), 63-85.
- Brochado, A., & Brochado, F. (2019). What makes a glamping experience great? *Journal of Hospitality and Tourism Technology*, 10 (1), 15-27.
- Buhalis, D. (2000). Trends in information technology and tourism. Gartner, W. C. & Lime, D. W. (Edt). *Trends in Outdoor Recreation, Leisure and Tourism*, 47-61. Cabi Publishing, USA.
- Büyükeke, A., Sökmen, A., & Gencer, C. (2020). Metin madenciliği ve duygu analizi yöntemleri ile sosyal medya verilerinden rekabetçi avantaj elde etme: Turizm sektöründe bir araştırma. *Journal of Tourism and Gastronomy Studies*, 8 (1), 322-335.
- Calero-Sanz, J., Orea-Giner, A., Villacé-Molinero, T., Muñoz-Mazón, A., & Fuentes-Moraleda, L. (2022). Predicting a new hotel rating system by analysing UGC content from Tripadvisor: Machine learning application to analyse service robots influence. *Procedia Computer Science*, 200, 1078-1083.
- Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. *Journal of Hospitality Marketing & Management*, 26 (7), 675-693.

- Caner, F. (2019). *Topic Modeling and Sentiment Analysis of Hotel Reviews*, (Unpublished Master Thesis). Bahçeşehir University, Graduate School of Natural and Applied Sciences, Istanbul.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., & Blei, D. (2009). Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 22. Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning techniques. In *Cognitive Informatics and Soft Computing: Proceeding of CISC 2017*, Springer Singapore, 639-647.
- Chang, Y. C., Ku, C. H., & Chen, C. H. (2019). Social Media Analytics: Extracting and Visualizing Hilton Hotel Ratings and Reviews from TripAdvisor. *International Journal of Information Management*, 48, 263-279.
- Chen, Z., Zhang, Y., Wu, C., & Ran, B. (2019). Understanding individualization driving states via latent Dirichlet allocation model. *IEEE Intelligent Transportation Systems Magazine*, 11 (2), 41-53.
- Chittiprolu, V., Samala, N., & Bellamkonda, R. S. (2021). Heritage hotels and customer experience: a text mining analysis of online reviews. *International Journal of Culture, Tourism and Hospitality Research*, 15 (2), 131-156.
- Cooke, M., & Buckley, N. (2008). Web 2.0, social networks and the future of market research. *International Journal of Market Research*, 50 (2), 267-292.
- Cronin Jr, J. J., & Taylor, S. A. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing*, 56(3), 55-68.
- Çabuk, S., Nakıboğlu, B. & Canoğlu, M. (2013). Algılanan otel imajı ve hizmet kalitesi ile tekrar satın alma niyeti arasındaki ilişkiler. *Anatolia: Turizm Araştırmaları Dergisi*, 24 (1), 96-108.
- Çılgin, C., Ünal, C., Alıcı, S., Akkol, E., & Gökşen, Y. (2020). Metin sınıflandırmada yapay sinir ağları ile bitcoin fiyatları ve sosyal medyadaki beklentilerin analizi. *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, 4 (1), 106-126.
- Çiçek, R., & Doğan, İ. C. (2009). Müşteri memnuniyetinin artırılmasında hizmet kalitesinin ölçülmesine yönelik bir araştırma: Niğde ili örneği. *Afyon Kocatepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 11 (1), 199-217.
- Çolak, U. (2021). Makine öğrenmesi Modellemeye Giriş. <https://ufukcolak.medium.com/makine-%C3%B6%C4%9Frenmesi-modellemeye-giri%C5%9F-2-405b4f3c8cfb> (Erişim Tarihi: 18.06.2023).
- Davras, Ö., & Caber, M. (2019). Analysis of hotel services by their symmetric and asymmetric effects on overall customer satisfaction: A comparison of market segments. *International Journal of Hospitality Management*, 81, 83-93.
- Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013). Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14, 2349–2353.
- Dickinson, B., & Hu, W. (2015). Sentiment analysis of investor opinions on Twitter. *Social Networking*, 4 (03), 62.

- Dohaiha, H. H., Prasad, P.W.C, Maag, A., Alsadoon, A. (2018). DeepLearning for aspect-based sentiment analysis: a comparative review, expert systems with applications. *Expert Systems with Applications*, 118, 272-299.
- Duan, W., Yu, Y., Cao, Q., & Levy, S. (2016). Exploring the impact of social media on hotel service performance: a sentimental analysis approach. *Cornell Hospitality Quarterly*, 57(3), 282-296.
- Ekinci, E., Omurca, S. İ., Kırık, E., & Taşçı, Ş. (2020). Tıp veri kümesi için gizli dirichlet ayrımı. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 22 (64), 67-80.
- El-Din, D. M. (2016). Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 7 (1). 244-252.
- Fasttext in Word Vectors for 157 Languages. <https://fasttext.cc/docs/en/crawl-vectors.html>, Erişim Tarihi: (15.06.2023).
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27 (8), 861-874.
- Fernandes, T., & Fernandes, F. (2017). Sharing dissatisfaction online: analyzing the nature and predictors of hotel guests negative reviews. *Journal of Hospitality Marketing & Management*, 27 (2), 127–150.
- Fernandez, D., Gonzalez, M., Gil, D., & Luján-Mora, S. (2020). Text mining of open-ended questions in self-assessment of university teachers: An LDA topic modeling approach. *IEEE Access*, 8, 35318-35330.
- Fong, L. H. N., Lei, S. S. I., & Law, R. (2017). Asymmetry of hotel ratings on tripadvisor: evidence from single-versus dual-valence reviews. *Journal of Hospitality Marketing & Management*, 26 (1), 67-82.
- Foris, D., Crihalmean, N., & Foris, T. (2020). Exploring the environmental practices in hospitality through book-ing websites and online tourist reviews. *Sustainability*, 12 (24), 1-18.
- García-Barriocanal, E., Sicilia, M. A., & Korfiatis, N. (2010). Exploring hotel service quality experience indicators in user-generated content: a case using tripadvisor data. *Proceedings of the 5th Mediterranean Conference of Information Systems (MCIS)*, Proceedings 33.
- García-Pablos, A., Duca, A. L., Cuadros, M., Linaza, M. T., & Marchetti, A. (2016). Correlating languages and sentiment analysis on the basis of text-based reviews. In information and communication technologies in tourism, *Proceedings of the International Conference in Bilbao*. Springer International Publishing, Spain, 565-577.
- Getty, J.M. & Thompson, K.N. (1994). Procedure for scaling perceptions of lodging quality, *Hospitality Research Journal*, 18 (2), 75-96.
- González-Rodríguez, M. R., Martínez-Torres, R., & Toral, S. (2016). Post-visit and Pre-visit tourist destination image through ewom sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management*, 28 (11), 2609-2627.

- Grave, P., Bojanowski, P., Gupta, A., Joulin, A., & Mikolov, T. (2028). Learning word vectors for 157 languages. *Proceedings of the International Conference on Language Resources and Evaluation. (LREC 2018)*. Miyazaki, Japan, 2483- 3487.
- Gretzel, U., & Yoo, K. H. (2008). Use and impact of online travel reviews. In *Information and Communication Technologies in Tourism*, Springer, Vienna, 35-46.
- Gronroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, 18 (4), 36-44.
- Gronroos C. (1990). Service management and marketing. Lexington Books: Lexington, Mass.
- Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, 59, 467-483.
- Güven, Z. A. Diri, B. & Cakaloglu, T. (2018). N-seviyeli gizli dirichlet ayırımı ile türkçe tivit duygularının sınıflandırılması. *Elektrik-Elektronik, Bilgisayar, Biyomedikal Mühendislikleri Bilimsel Toplantısı*, İstanbul, 1-5.
- Hartline, M. D., & Jones, K. C. (1996). Employee performance cues in a hotel service environment: Influence on perceived service quality, value, and word-of-mouth intentions. *Journal of business research*, 35 (3), 207-215.
- Herjanto, H., Erickson, E., & Calleja, N. F. (2017). Antecedents of business travelers' satisfaction. *Journal of Hospitality Marketing & Management*, 26 (3), 259–275.
- Herrero, A., San-Martín, H., & Hernández, J.M. (2015). How online search behavior is influenced by user-generated content on review websites and hotel interactive websites. *International Journal of Contemporary Hospitality Management*, 27, 1573–1597.
- Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text preprocessing for text mining in organizational research: Review and recommendations. *Organizational Research Methods*, 25 (1), 114-146.
- Hogenboom, A., Heerschop, B., Frasinçar, F., Kaymak, U., & Jong, F. (2014). Multi-lingual support for lexicon-based sentiment analysis guided by semantics. *Decision Support Systems*, 62, 43-53.
- Holjevac, I. A. (2003). A vision of tourism and the hotel industry in the 21st century. *International Journal of Hospitality Management*, 22 (2), 129-134.
- Hoskisson, A., & Wren, P.G. (2015). An analytical journey towards big data. *Journal of Decision Systems*, 24 (1), 87-102.
- <https://hotelbusiness.com/study-average-business-trip-cost-will-increase-in-2020/> (Erişim Tarihi: 20.06.2023).
- <https://orangedatamining.com/widget-catalog/text-mining/documentembedding/>. (Erişim Tarihi: 7.05.2023).
- <https://ir.tripadvisor.com/news-releases/news-release-details/online-reviews-remain-trusted-source-information-when-booking>. (Erişim Tarihi: 15.06.2023).
- Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management*, 72, 417-426.

- Huigol, P. (2020). Quick introduction to bag-of-words (bow) and tf-idf for creating features from text. <https://www.analyticsvidhya.com/blog/2020/02/quick-introduction-bag-of-words-bow-tf-idf/>. (Erişim Tarihi: 23.04.2023).
- Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *In Proceedings of the International AAAI Conference on Web and Social Media*, 8 (1), 216-225.
- Ignatow, Gabe & Mihalcea, Rada. (2018). *An Introduction to Text Mining: Research Design, Data Collection, and Analysis*. Sage Publication, California.
- İnan, H. E. (2019). Otel yorumlarının sınıflandırılmasında makine öğrenmesi algoritmalarının karşılaştırılması ve eskişehir otellerine yönelik yorumların duygusal analizi. *20. Ulusal ve 4. Uluslararası Turizm Kongresi*, 16-19.
- Jakulin, T. J. (2017). Systems approach to tourism: A methodology for defining complex tourism system. *Organizacija*, 50 (3), 208-215.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent dirichlet allocation (LDA) and topic modeling: Models, applications, A survey. *Multimedia Tools and Applications*, 78, 15169-15211.
- Jia, S. (2019). Understanding Cruise Tourists' Satisfaction by Analysing Their Online Ratings and Reviews. *Journal of Management and Strategy*, 10 (4), 1-6.
- Jiang, Q., Chan, C. S., Eichelberger, S., Ma, H., & Pikkemaat, B. (2021). Sentiment analysis of online destination image of hong kong held by mainland chinese tourists. *Current Issues in Tourism*, 24 (17), 2501-2522.
- Johnston, W. J., & Bonoma, T. V. (1981). The buying center: structure and interaction patterns. *Journal of Marketing*, 45(3), 143-156.
- Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv preprint*, arXiv:1607.01759.
- Juwaheer, T. D. (2004). Exploring international tourists' perceptions of hotel operations by using a modified SERVQUAL approach-a case study of Mauritius. *Managing Service Quality*, 14 (5), 350-364.
- Kandampully, J., & Suhartanto, D. (2000). Customer loyalty in the hotel industry: the role of customer satisfaction and image. *International Journal of Contemporary Hospitality Management*, 12 (6), 346-351.
- Kantarcı, K., Başaran, M. A., & Özyurt, P. M. (2017). Determining hotel performance through consumer generated travel 2.0 reviews: A case of Kyrgyzstan, *Manas Sosyal Araştırmalar Dergisi*, 6 (5), 131-145.
- Kapadia, s. (2019). Evaluate Topic Models: Latent Dirichlet Allocation, <https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0>. (Erişim Tarihi: 02.02.2023).
- Kar, A. K., Kumar, S., & Ilavarasan, P. V. (2021). Modelling the service experience encounters using user-generated content: A text mining approach. *Global Journal of Flexible Systems Management*, 22 (4), 267-288.
- Karaman, E., & Çilan, Ç. A. (2020). Sentiment analysis of tourism-related tweets during covid-19 outbreak through machine learning techniques, *Statistics & Empirical Economics Journal*, 17, 88-98.

- Karunaratne, W. M. K. K., & Jayawardena, L. N. A. C. (2010). Assessment of customer satisfaction in a five star hotel-a case study. *Tropical Agricultural Research*, 21(3), 258-265.
- Kaynar, O., Görmez, Y., Yıldız, M., & Albayrak, A. (2016). Makine öğrenmesi yöntemleri ile Duygu Analizi. In *International Artificial Intelligence and Data Processing Symposium (IDAP'16)*, 17 (18), 17-18.
- Kılıç, B., & Eleren, A. (2009). turizm sektöründe hizmet kalitesi ölçümü üzerine bir literatür araştırması. *Uluslararası Alanya İşletme Fakültesi Dergisi*, 1 (1), 91-118.
- Kim, B., Kim, S., & Heo, C. Y. (2016). Analysis of satisfiers and dissatisfiers in online hotel reviews on social media. *International Journal of Contemporary Hospitality Management*, 28 (9), 1915-1936.
- Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction: which one has more incremental predictive power in explaining hotel performance? *International Journal of Contemporary Hospitality Management*, 29 (2), 784-802.
- Kladou, S., & Mavragani, E. (2015). Assessing destination image: An online marketing approach and the case of TripAdvisor. *Journal of Destination Marketing & Management*, 4 (3), 187-193.
- Korucuk, S., Akyurt, H. & Turpcu, E. (2019). Otel işletmelerinde hizmet kalitesinin entropi yöntemi ile ölçülmesi: giresun ilindeki üç yıldızlı oteller üzerine bir araştırma. *MANAS Sosyal Araştırmalar Dergisi*, 8 (1) , 697-709.
- Kotler, P, Keller, K. L., Cunningham, P. H., & Sivaramakrishnan, S. (2009). *Marketing management. 13th ed.* Pearson Education, Canada.
- Kotler, P. (2011), "Philip Kotler's Contributions to Marketing Theory and Practice", Malhotra, N.K. (Ed.) *Review of Marketing Research: Special Issue – Marketing Legends (Review of Marketing Research, Vol. 8)*, Emerald Group Publishing Limited, Bingley, pp. 87-120.
- Köseoglu, M. A., Mehraliyev, F., Altin, M., & Okumus, F. (2019). Competitor intelligence and analysis (CIA) model and online reviews: Integrating big data text mining with network analysis for strategic analysis. *Tourism Review*, 76 (3), 529-552.
- Kumar, A., & Paul, A. (2016). *Mastering Text mining with R*, Packt Publishing Ltd. England.
- Kültür ve Turizm Bakanlığı, 2022 Yılı Sınır Giriş Çıkış İstatistikleri. <https://yigm.ktb.gov.tr/TR-9851/turizm-istatistikleri.html>, (Erişim Tarihi: 14.07.2023).
- Kültür ve Turizm Bakanlığı, Ülkemizi Ziyaret Eden Turist İstatistikleri, <https://yigm.ktb.gov.tr/Eklenti/111370,2022yilliksinirbultenixlsx.xlsx?0>. (Erişim Tarihi: 14.07.2023).
- Lee, J., Kang, J. H., Jun, S., Lim, H., Jang, D., & Park, S. (2018). Ensemble Modeling For Sustainable Technology Transfer. *Sustainability*, 10 (7), 2278.
- Lee, M., Cai, Y. J., DeFranco, A., & Lee, J. (2020). Exploring influential factors affecting guest satisfaction big data and business analytics in consumer-generated reviews. *Journal of Hospitality and Tourism Technology*, 11 (1), 137–153.

- Levis, M., Helfert, M., & Brady, M. (2008). Website design quality and form input validation: An empirical study on Irish corporate websites. *Journal of Service Science and Management*, 1, 91-100.
- Li, H., Ye, Q., & Law, R. (2013). Determinants of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pacific Journal of Tourism Research*, 18 (7), 784-802.
- Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism Management*, 59, 554–563.
- Loiacono, E.T., Watson, R.T. & D.L. Goodhue. (2007). WebQual: An instrument for consumer evaluation of web sites, *International Journal of Electronic Commerce*, 11 (3), 51-87.
- López-Santana, E., Méndez-Giraldo, G., & Molano, J. I. R. (2018). A rule-based system to scheduling and routing problem in home health care services. In *International Conference on Computational Science and Its Applications*, Cham: Springer International Publishing, 496-508.
- Ma, Y. (2020). NLP: How does NLTK.Vader Calculate Sentiment? <https://medium.com/@mystery0116/nlp-how-does-nltk-vader-calculate-sentiment-6c32d0f5046b>. (Erişim Tarihi: 8. 06. 2023).
- Mahadevan, A., & Arock, M. (2020). Integrated topic modeling and sentiment analysis: a review rating prediction approach for recommender systems. *Turkish Journal of Electrical Engineering and Computer Sciences*, 28 (1), 107-123.
- Mankad, S., Han, H. S., Goh, J., & Gavirneni, S. (2016). Understanding online hotel reviews through automated text analysis. *Service Science*, 8 (2), 124-138.
- Marandi, A., Tasavori, M., & Najmi, M. (2023). New insights into hotel customer's revisiting intentions, based on big data. *International Journal of Contemporary Hospitality Management*. <https://www.emerald.com/insight/content/doi/10.1108/IJCHM-06-2022-0719/full/html>. Erişim Adresi: 10.07.2023.
- Markopoulos, G., Mikros, G., Iliadi, A., & Lontos, M. (2015). Sentiment analysis of hotel reviews in greek: a comparison of unigram features. In *Cultural Tourism in a Digital Era: First International Conference IACuDiT*, Athens, Springer International Publishing, 373-383.
- Marques, D. B., Barradas Filho, A. O., Romariz, A. R., Viegas, I. M., Luz, D. A., Barros Filho, A. K., & Ferraudo, A. S. (2014). Recent Developments on Statistical and Neural Network Tools Focusing on Biodiesel Quality. *International Journal of Computer Science and Application*, 3(3), 97-110.
- Martins, C. A., Monard, M. C., & Matsubara, E. T. (2003). Reducing the dimensionality of bag-of-words text representation used by learning algorithms. *Proc of 3rd IASTED International Conference on Artificial Intelligence and Applications*, 228-233.
- Matrutty, J. P., Adrian, A. M., & Angdresey, A. (2023). Sentiment analysis of visitor reviews on star hotels in manado city. *Journal of Information Technology and Computer Science*, 8 (1), 21-32.

- Maulana, Z. F. (2017). Mengenal Topic Modeling. Alur Algoritma LDA. <https://toolbox.kurio.co.id/topic-modeling-696d7ba2592f>. (Erişim Tarihi: 30.06.2023).
- McConnell, C. R., Brue, S. L., & Flynn, S.M. (2009). *Economics: Principles, Problems and Policies*. 18. Edition. McGraw- Hill Irwin Publishing, Newyork.
- Mei, A. W. O., Dean, A. M., & White, C. J. (1999). Analysing service quality in the hospitality industry. *Managing Service Quality: An International Journal*, 9 (2), 136-143.
- Menekşe, R. (2005). Her şey dahil sisteminin ve sistemden faydalananlar açısından etkilerinin otel yöneticilerinin gözünden değerlendirilmesi (Marmaris örneği). *Ekonomik ve Sosyal Araştırmalar Dergisi*, 1, 97-124.
- Miguéns, J., Baggio, R., & Costa, C. (2008). Social media and tourism destinations: tripadvisor case study. *Advances in Tourism Research*, 26 (28), 1-6.
- Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., & Trajanov, D. (2020). Evaluation of sentiment analysis in finance: from lexicons to transformers. *IEEE Access*, 8, 131662-131682.
- Moro, S., Lopes, R. J., Esmerado, J., & Botelho, M. (2020). Service quality in airport hotel chains through the lens of online reviewers. *Journal of Retailing and Consumer Services*, 56, 102193.
- Mousavian, S., Miah, S. J., & Zhong, Y. (2023). A design concept of big data analytics model for managers in hospitality industries. *Personal and Ubiquitous Computing*, 1-11. <https://link.springer.com/article/10.1007/s00779-023-01714-3>. (Erişim Tarihi: 12.07.2023).
- Muñoz-Leiva, F., Hernández-Méndez, J., & Sánchez-Fernández, J. (2012). Generalising user behaviour in online travel sites through the Travel 2.0 website acceptance model. *Online Information Review*, 36 (6), 879-902.
- Nadiri, H. & Hussain, K. (2005). Diagnosing the zone of tolerance for hotel services. *Managing Service Quality*, 15 (3): 259-277.
- Najadat, H., Al-Abdi, A., & Sayaaheen, Y. (2018). Model-based sentiment analysis of customer satisfaction for the jordanian telecommunication companies. In *2018 9th International Conference on Information and Communication Systems (ICICS)*, 233-237.
- Navío-Marco, J., Ruiz-Gómez, L. M., & Sevilla-Sevilla, C. (2018). Progress in information technology and tourism management: 30 years on and 20 years after the internet-revisiting buhalis & law's landmark study about etourism. *Tourism Management*, 69, 460-470.
- Neri, F., Aliprandi, C., Capeci, F., & Cuadros, M. (2012). Sentiment analysis on social media. *2012 IEEE/ACM International Conference on Advances In Social Networks Analysis and Mining*, 919-926.
- Neto, A. P., Silva, E. A. L., Ferreira, L. V. F., & Araújo, J. F. R. (2020). Discovering the sustainable hotel brand personality on tripadvisor. *Journal of Hospitality and Tourism Technology*, 11(2), 241-254.

- Nguyen, M. Ha, N., Anh, P. & Matsui, Y. (2015). Service Quality and Customer Satisfaction: A Case Study of Hotel Industry in Vietnam. *Asian Social Science*, 11(10), 73.
- Nicoli, N., & Papadopoulou, E. (2017). TripAdvisor and reputation: a case study of the hotel Industry in Cyprus. *EuroMed Journal of Business*, 12 (3), 316-334.
- Nilashi, M., Samad, S., Ahani, A., Ahmadi, H., Alsolami, E., Mahmoud, M., & Alarood, A. A. (2021). Travellers decision making through preferences learning: a case on malaysian spa hotels in tripadvisor. *Computers & Industrial Engineering*, 158, 107348.
- Nowacki, M., & Niezgoda, A. (2020). Identifying unique features of the image of selected cities based on reviews by TripAdvisor portal users. *Scandinavian Journal of Hospitality and Tourism*, 20(5), 503-519.
- Obaidi, M., Nagel, L., Specht, A., & Klünder, J. (2022). Sentiment analysis tools in software engineering: a systematic mapping study. *Information and Software Technology*, 151, 1-14.
- O'connor, P. (2010). Managing a hotel's image on TripAdvisor. *Journal of Hospitality Marketing & Management*, 19 (7), 754-772.
- Ogul, B. B., & Ercan, G. (2016). Sentiment Classification on turkish hotel reviews. *24th Signal Processing and Communication Application Conference (SIU)*. 497-500.
- Ooi, C. S., & Munar, A. M. (2013). Digital Social Construction of a tourist site: ground zero. In *tourism social media: transformations in identity, community and culture*, Emerald Group Publishing Limited. 159-175.
- Ou, W., Huynh, V. N., & Sriboonchitta, S. (2018). Training attractive attribute classifiers based on opinion features extracted from review data. *Electronic Commerce Research and Applications*, 32, 13-22.
- Özen, İ. A. (2019). Müşteri şikâyetlerinde fikir madenciliği: antalya helal oteller üzerine bir araştırma., *2.Uluslararası Helal Turizm Kongresi*. 242-249.
- Öztürk, M. (2023). Çapraz Doğrulama Teknikleri. [https://miracozturk.com/capraz-dogrulama-teknikleri-crossvalidation/#:~:text=K%2DFold%20Cross%2DValidation%20\(genel%20performans%C4%B1%20daha%20do%C4%9Fru%20de%C4%9Ferlendirilebilir.](https://miracozturk.com/capraz-dogrulama-teknikleri-crossvalidation/#:~:text=K%2DFold%20Cross%2DValidation%20(genel%20performans%C4%B1%20daha%20do%C4%9Fru%20de%C4%9Ferlendirilebilir.) (Erişim Tarihi: 15.05.2023).
- Padma, P., & Ahn, J. (2020). Guest satisfaction & dissatisfaction in luxury hotels: an application of big data. *International Journal of Hospitality Management*, 84, 1-8.
- Paladini, P., & Peres, K. (2021). Exploring the attributes of hotel service quality in florianopolis-sc, Brazil: An analysis of tripadvisor reviews. *Cogent Business & Management*, 8 (1), 1-17.
- Palanisamy, P., Yadav, V., & Elchuri, H. (2013). Serendio: Simple and practical lexicon based approach to sentiment analysis. *Second Joint Conference on Lexical and Computational Semantics, Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, 543-548.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49 (4), 41-50.

- Parasuraman, A. B. L. L., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64 (1), 12-40.
- Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). ES-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, 7 (3), 213-233.
- Payak, A., Rai, S., Shrivastava, K., & Gulwani, R. (2020). Automatic text summarization and keyword extraction using natural language processing. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 98-103.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *The Journal of Machine Learning Research*, 12, 2825-2830.
- Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: An empirical analysis. *Journal of Travel Research*, 56 (2), 235–249.
- Polat, H. & Ağca, Y. (2022). Tripadvisor kullanıcılarının türkçe ve ingilizce yorumları kapsamında duygu analizi yöntemlerinin karşılaştırmalı analizi. *Abant Sosyal Bilimler Dergisi*, 22 (2) , 901-916.
- Putri, I. R., & Kusumaningrum, R. (2017). Latent Dirichlet Allocation (LDA) for sentiment analysis toward tourism review in Indonesia. In *Journal of Physics: Conference*, 801 (1), 012073, IOP Publishing.
- Quick Introduction to Bag-of-Words (BoW) and TF-IDF for Creating Features from Text, <https://www.analyticsvidhya.com/blog/2020/02/quick-introduction-bag-of-words-bow-tf-idf/>. (Erişim Tarihi: 23.04.2023)
- Raguseo, E., Neirotti, P., & Paolucci, E. (2017). How small hotels can drive value their way in infomediation. The case of ‘Italian hotels vs. OTAs and TripAdvisor’. *Information & Management*, 54 (6), 745-756.
- Rahman, A. (2019). Statistics-based data preprocessing methods and machine learning algorithms for big data analysis. *International Journal of Artificial Intelligence*, 17(2), 44-65.
- Rana, T. A., Cheah, Y. N., & Letchmunan, S. (2016). Topic modeling in sentiment analysis: a systematic review. *Journal of ICT Research & Applications*, 10 (1), 76-93.
- Ray, B., Garain, A., & Sarkar, R. (2021). An Ensemble-Based Hotel Recommender System Using Sentiment Analysis and Aspect Categorization of Hotel Reviews. *Applied Soft Computing*, 98, 106935.
- Ren, M., Vu, H. Q., Li, G., & Law, R. (2021). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: Implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 30 (1), 96-119.
- Researchandmarket Services Global Market Report (2021). https://www.researchandmarkets.com/reports/5781225/services-global-market-report?utm_source=GNOM&utm_medium=PressRelease&utm_code=gtsbs8&tm_ca

- mpaign=1493910+-+Global+Services+Market+Report+(2021+to+2030)+-+COVID-19+Impact+and+Recovery&utm_exec=jamu273prd. (Erişim Tarihi: 14.07.2023).
- Ribarsky, W., Wang, D. X., & Dou, W. (2014). Social media analytics for competitive advantage. *Computers & Graphics*, 38, 328-331.
- Rios-Martin, M. A., Folgado-Fernandez, J. A., Palos- Sanchez, P. R., & Castejon-Jimenez, P. (2020). The impact of the environmental quality of online feed-back and satisfaction when exploring the critical factors for luxury hotels. *Sustainability*, 12 (1), 1-24.
- Sahebi, A. G., Kordheydari, R., & Aghaei, M. (2022). A new approach in marketing research: Identifying the customer expected value through machine learning and big data analysis in the tourism industry. *Asia-Pacific Journal of Management and Technology*, 2 (3), 26-42.
- Sakas, D. P., Reklitis, D. P., Terzi, M. C., & Vassilakis, C. (2022). Multichannel digital marketing optimizations through big data analytics in the tourism and hospitality industry. *Journal of Theoretical and Applied Electronic Commerce Research*, 17 (4), 1383-1408.
- Sann, R., Lai, P.-C., & Liaw, S.-Y. (2020). Online complain-ing behavior: Does cultural background and hotel class matter? *Journal of Hospitality and Tourism Management*, 43, 80–90.
- Sann, R., & Lai, P.-C. (2020). Do expectations towards thai hospitality differ? the views of english vs Chinese speaking travelers. *International Journal of Culture Tourism and Hospitality Research*, 15 (1), 43–58.
- Sann, R., & Lai, P.-C. (2020). Understanding homophily of service failure within the hotel guest cycle: applying NLP-aspect-based sentiment analysis to the hospi-tality industry. *International Journal of Hospitality Management*, 91, 1-13.
- Sann, R. and Lai, P.-C. (2023). Topic Modeling of the Quality of Guest’s Experience Using Latent Dirichlet Allocation: Western Versus Eastern Perspectives, *Consumer Behavior in Tourism and Hospitality*, 18(1), 17-34.
- Sayım, F. & Aydın, V. (2011). Hizmet sektörü özellikleri ve sistematik olmayan risklerin sektör menkul kıymetleri ile etkileşimine dair teorik bir çalışma. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (29).
- Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: recent trends and future directions. *Journal of Travel & Tourism Marketing*, 32 (5), 608-621.
- Sebastian, A. (2020). A gentle introduction to calculating the tf-idf values. <https://towardsdatascience.com/a-gentle-introduction-to-calculating-the-tf-idf-values-9e391f8a13e5>. (Erişim Tarihi: 30.06.2023).
- Seker, S. E. (2016). Duygu Analizi (Sentimental Analysis). *YBS Ansiklopedi*, 3 (3), 21-36.
- Seth, N. (2021). Topic Modeling and Latent Dirichlet Allocation (LDA) using Gensim and Sklearn. <https://www.analyticsvidhya.com/blog/2021/06/part-2-topic-modeling-and-latent-dirichlet-allocation-lda-using-gensim-and-sklearn/>. (Erişim Tarihi: 29.06.2023).

- Sevgi, O. O. (2017). *Otel işletmelerinde yabancı turistlerin hizmet kalitesi algısının servqual yöntemiyle ölçülmesi: Antalya' da beş yıldızlı otellerde bir araştırma*. (Yayınlanmamış Yüksek Lisans Tezi). İstanbul Kültür Üniversitesi, Sosyal Bilimler Enstitüsü. İstanbul.
- Sezgin, M., & Dağdelen, Ö. (2023). Ekolojik otellerin çevrimiçi yorumlarının duygu analizi ve çok kriterli karar verme yöntemleri ile değerlendirilmesi: Antalya örneği, *Journal of Gastronomy, Hospitality and Travel*, 6(1), 131-164.
- Sezgin, M., & Duman, A. (2023). elektronik ağızdan ağıza pazarlama kapsamında konaklama işletmelerine yönelik çevrimiçi yorumların duygu analizi yöntemiyle incelenmesi: Alanya örneği. *Türk Turizm Araştırmaları Dergisi*, 7 (2), 244-265.
- Sham, N. M., & Mohamed, A. (2022). Climate change sentiment analysis using lexicon, machine learning and hybrid approaches. *Sustainability*, 14 (8), 4723.
- Shengelbayeva, A. (2009). *Konaklama işletmelerinde hizmet kalitesinin müşteri memnuniyetine etkisi: Kiriş (Antalya) örneği*. (Yayınlanmamış Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İzmir.
- Silahtaroglu, G. (2016). *Veri Madenciliği: Kavram ve Algoritmaları*. 3. Basım, Papatyabilim Yayıncılık, İstanbul.
- Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32 (6), 1310-1323.
- Storopoli, José. (2019). Topic Modeling: How and why to use in management research. *Revista Ibero-Americana de Estrategia*. 18, 316-338.
- Sun, S., Tong, K. T., & Law, R. (2017). Chinese hotel guest perception of international chain hotels under the same hotel brand in different travel destinations: the cases of intercontinental and Sheraton. *Journal of Vacation Marketing*, 23 (2), 172–188.
- Sutherland, I., & Kiatkawsin, K. (2020). Determinants of guest experience in airbnb: a topic modeling approach using LDA. *Sustainability*, 12 (8), 3402.
- Sutherland, I., Sim, Y., Lee, S. K., Byun, J., & Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichlet allocation. *Sustainability*, 12 (5), 1821.
- Taecharungroj, V., & Mathayomchan, B. (2019). Analysing Tripadvisor reviews of tourist attractions in phuket, Thailand. *Tourism Management*, 75, 550-568.
- Tafralı, S. (2022). Yapay Öğrenme: Rastgele Orman. <https://medium.com/machine-learning-t%C3%BCrkiye/yapay-%C3%B6%C4%9Frenme-rastgele-orman-e8debd886e7>, (Erişim Tarihi: 05.05.2023).
- Tamilselvi, P. (2016). Service Quality-Customer Satisfaction. *International Conference on Service Marketing – Talking The Show Abroad*, 29-31.
- Taşkın, E., & Büyük, K. (2002). Hizmet pazarlaması açısından eğitim hizmetlerinde kalite: Kütahya'daki özel dersane öğrencileri ile ilgili bir saha araştırması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 7, 203-224.
- Thareja, P. (2017). Excellence in technical education: the pathway of quality management. *OmniScience: A Multi-Disciplinary Journal*, 7 (3), 7-22.
- The World Bank (2021). Employment in services. <https://data.worldbank.org/indicator/SL.SRV.EMPL.ZS>. (Erişim Tarihi: 15.06.2023).

- Timo, N., & Fisher, R. (2005). the service–manufacturing conundrum: investigating a new approach to understanding and classifying services and manufacturing. *British Academy of Management Conference*, 1-11.
- Tontini, G., Bento, G., Milbratz, T. C., Volles, B. K., & Ferrari, D. (2017). The Critical Incident Technique (CIT) and Penalty-Reward Contrast Analysis (PRCA) applied to online reviews on TripAdvisor: Evaluation of satisfaction of hotel customers. *International Journal of Hospitality Management*, 66 (1), 106-116.
- Torun Nalbant, F. & Demiral, A. K. (2019). Otel işletmelerindeki hizmet kalitesinin müşteri memnuniyeti üzerine etkileri: antalya ilinde bir uygulama . *Süleyman Demirel Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi* , 24 (4) , 813-829.
- Tran, T., Ba, H., & Huynh, V.-N. (2019). Measuring Hotel Review Sentiment: An Aspect-Based Sentiment Analysis Approach. *Computational Color Imaging*, 393–405.
- Tuna, M.F. (2019). *Çevrimiçi Yorum ve Şikâyetlerin Otel İşletmeleri Üzerinden Duygu Analizi ile İncelenmesi*, (Yayınlanmamış Doktora Tezi). Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, Kayseri.
- Tuominen, P. (2011). The Influence of Tripadvisor Consumer-Generated Travel Reviews on Hotel Performance. *Presented at the 19th Annual Frontiers in Service Conference*, 1-11.
- Türk, D. Z. (2009). Denetim firmalarının sunduğu hizmet kalitesi müşteri tatmini ve sadakati: servperf ölçeği. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 18 (1) , 399-416.
- Umer, M., Imtiaz, Z., Ahmad, M., Nappi, M., Medaglia, C., Choi, G. S., & Mehmood, A. (2023). Impact of Convolutional Neural Network and Fasttext Embedding on Text Classification. *Multimedia Tools and Applications*, 82 (4), 5569-5585.
- Ustaahmetoğlu, E. & Savci, M. (2011). Muhasebe meslek mensuplarının sundukları hizmetin kalitesini değerlendirmesi ve örnek bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi* , (22) , 168-184.
- Uzgören, N. (2007). Uyum analizinin teorik esasları ve regresyon analiz ile benzerliğinin grafiksel boyutta karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (18).
- Ülgen, M.K. (2018). Python ile Kümeleme Algoritmaları. <https://medium.com/@k.ulgen90/python-ile-k%C3%BCmeleme-algoritmalar%C4%B1-makine-%C3%B6%C4%9Frenimi-b%C3%B6l%C3%BCm-8-8204ffa702f2>. (Erişim Tarihi: 10 Haziran 2023).
- Valdivia, A., Luzón, M. V., & Herrera, F. (2017). Sentiment Analysis in Tripadvisor. *IEEE Intelligent Systems*, 32(4), 72-77.
- Wang, J. H., Liu, T. W., Luo, X., & Wang, L. (2018). An LSTM approach to short text sentiment classification with word embeddings. *Proceedings of The 30th Conference On Computational Linguistics and Speech Processing (ROCLING 2018)*, 214-223.
- Wang, S., Tang, J., Aggarwal, C., & Liu, H. (2016). Linked document embedding for classification. *Proceedings 25th ACM International on Conference on Information and Knowledge Management*, 115-124.

- Weng, M. H., Ha, J. L., Wang, Y. C., & Tsai, C. L. (2012). A study of the relationship among service innovation, customer value and customer satisfaction: An empirical study of the hotel industry in Taiwan. *International Journal of Organizational Innovation*, 4 (3), 98-112.
- Witten I., Frank, W., Hall, M. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*, 3. Edition, Morgan Kauffman Publishing.
- www.agoda.com, (Erişim Tarihi: 12.04.2023)
- www.booking.com, (Erişim Tarihi: 12.04.2023)
- www.tripadvisor.com, (Erişim Tarihi: 12.04.2023)
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.
- Xie, K., Kwok, L., & Wang, W. (2017). Monetizing managerial responses on tripadvisor: performance implications across hotel classes. *Cornell Hospitality Quarterly*, 58 (3), 240-252.
- Xu, H., & Lv, Y. (2022). Mining and Application of Tourism Online Review Text Based on Natural Language Processing and Text Classification Technology. *Wireless Communications and Mobile Computing*, 1-13.
- Xu, J.(2018). Topic Modeling with LSA, PLSA, LDA & lda2Vec. <https://medium.com/nanonets/topic-modeling-with-lsa-psla-lda-and-lda2vec-555ff65b0b05>.(Erişim Tarihi: 12.04.2023).
- Xu, X. (2018). Does traveler satisfaction differ in various travel group compositions? evidence from online reviews. *International Journal of Contemporary Hospitality Management*, 30 (3), 1663–1685
- Xu, X., & Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. *International Journal of Hospitality Management*, 55, 57–69.
- Yan, D., Li, K., Gu, S., & Yang, L. (2020). Network-based bag-of-words model for text classification. *IEEE Access*, 8, 82641-82652.
- Ying, S., Chan, J. H., & Qi, X. (2020). Why are Chinese and North American guests satisfied or dissatisfied with hotels? an application of big data analysis. *International Journal of Contemporary Hospitality Management*, 32 (10), 3249–3269.
- Zagar, A. P. (2022). LDAvis: Visualization for LDA Topic Modelling, <https://orangedatamining.com/blog/2022/2022-03-18-ldavis/>. (Erişim Tarihi: 30.06.2023)
- Zeithaml, Valarie A. (1981). How Consumer Evaluation Processes Differ Between Goods and Services. In *Marketing of Services*, Donnelly James H., and George William R., eds. Chicago: American Marketing Association, 186–89.
- Zhang, Y., Jin, R., & Zhou, Z. H. (2010). Understanding bag-of-words model: a statistical framework. *International journal of machine learning and cybernetics*, 1, 43-52.
- Zhao, R., & Mao, K. (2017). Fuzzy Bag-of-words model for document representation. *IEEE transactions on Fuzzy Systems*, 26 (2), 794-804.

EK: İNTİHAL RAPORU

MAKİNE ÖĞRENMESİNE DAYALI ÇEVİRİMİÇİ OTEL GÖRÜŞLERİNİN DUYGU ANALİZİ VE KONU MODELLEME YAKLAŞIMLARIYLA DEĞERLENDİRİLMESİ

ORJİNALLİK RAPORU

% 4	% 3	% 1	% 1
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	acikerisim.alanya.edu.tr:8080 İnternet Kaynağı	<% 1
2	acikerisim.alanya.edu.tr İnternet Kaynağı	<% 1
3	sosyoteknik.selcuk.edu.tr İnternet Kaynağı	<% 1
4	halic.edu.tr İnternet Kaynağı	<% 1
5	acikbilim.yok.gov.tr İnternet Kaynağı	<% 1
6	www.researchgate.net İnternet Kaynağı	<% 1
7	Zekeriya Anil Guven, Banu Diri, Tolgahan Cakaloglu. "Classification of New Titles by Two Stage Latent Dirichlet Allocation", 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), 2018 Yayın	<% 1

ÖZGEÇMİŞ

Adı-Soyadı: Egemen Güneş TÜKENMEZ

Eğitim Bilgileri:

Lisans: Dokuz Eylül Üniversitesi, İşletme Fakültesi, Turizm İşletmeciliği Bölümü
(2014)

Yüksek Lisans: Dokuz Eylül Üniversitesi, İşletme Fakültesi, Sosyal Bilimler Enstitüsü,
Turizm İşletmeciliği Anabilim Dalı (2019)

Doktora: Alanya Alaaddin Keykubat Üniversitesi, Lisansüstü Eğitim Enstitüsü, Turizm
İşletmeciliği Anabilim Dalı (2023)

Mesleki Deneyim:

Resepsiyon Görevlisi: İberostar Bellis Hotel, Önbüro Departmanı, , Antalya (2010)

Stajyer: SunExpress Havayolları, Operasyon Birimi, İzmir (2013)

Araştırma Görevlisi: Haliç Üniversitesi, İşletme Fakültesi, Turizm İşletmeciliği
Bölümü (2015-2019)

Araştırma Görevlisi: Alanya Alaaddin Keykubat Üniversitesi, Turizm Fakültesi,
Turizm İşletmeciliği Bölümü (2019-Devam etmekte)

Yabancı Dil Bilgisi:

İngilizce: 91,25 (YÖKDİL)

Almanca: B1

Yayınlanan Bilimsel Faaliyetler:

Kalpaklıoğlu, N., Tükenmez, E.G. (2016).Alternatif Turizm Biçimi Olarak Gastronomi
Turizmi Pazarlaması ve Yeni Rotalar, 1. Alternatif Turizm Kongresi

Kalpaklıoğlu, N., Tükenmez,E.g.(2016). Gastronomy Tourism in Context of
Sustainability and Byzantine Cuisine in Istanbul, 8th iConEc Conference.

Yıldırım, Y., Tükenmez, E.G.,Bayram, A.(2017). Seyahat Acentalarında Sosyal Medya
Kullanımı: Twitter Örneği, II. Doğu Akdeniz Turizm Sempozyumu

Akova, O. Tükenmez, E.G., Yıldırım, Y.(2017). Relationship Between Population,
Number of Tourists,Traffic Jam and Air Pollution as Important CarryingCapacity
Components, 19th International Symposium on Environmental Pollution and its
Impact on Life in theMediterranean Region.

Kalpaklıoğlu, N.,Tükenmez, E.G.,(2018). The Use of Twitter As A Marketing Tool By
Tour Operators In Antalya, 4th International Conference Knowledge Based
Sustainable Economic Development.

Tükenmez, E.G(2021). Kadın Turisterin Karşılaştığı Seyahat Kısıtları: İstanbul Örneği,
Asead 10. Uluslararası Sosyal Bilimler Sempozyumu.

Tükenmez, E.G.(2021). Veri Zarflama Analizi ile Ülkelerin Turizm Sektörü ve Çevresel
Sürdürülebilirlik Unsurlarının Değerlendirilmesi, III. Uluslararası Sürdürülebilir
Turizm Kongresi.

- Tükenmez, E.G. (2021). Turizm Sektörü Ve Marka Etkinliklerinin Veri Zarflama Analizi Yöntemiyle Karşılaştırılması: Türkiye Ve Akdeniz Ülkeleri, 21. Ulusal Turizm Kongresi.
- Kantarcı, K., Tükenmez, E.G. , Başaran, M.A., Abdirazakov, N. (2022).Analysis of Hotel Service Quality Attributes Based on Big Data Through Online Customer Generated Reviews: Case of Almaty, Uluslararası Türk Dünyası Sosyal Bilimler Sempozyumu.
- Gül. T., Tükenmez, E.G.(2022). Turizm Öğrencilerinin Kariyer Beklentileri ve Mesleğe Bağlılıkları Üzerine Betimsel Bir Araştırma, 6th International West Asia Congress of Tourism Research
- Kantarcı, K., Tükenmez, E.G., Başaran, M.A., Abdirazakov, N. (2022). Otel Hizmet Bileşenleri İle Otel Performansları Arasındaki İlişkinin WEB 2.0 Verilerine göre analizi: Bakü Alan Araştırması, Uluslararası Türk Dünyası Sosyal Bilimler Sempozyumu.
- Tükenmez, E.G.(2022). Golf Sahalarına Yönelik Çevrimiçi Yorumların Duygu Analizi Yöntemi İle İncelenmesi: Antalya Bölgesi Örneği, Asead 10. Uluslararası Sosyal Bilimler Sempozyumu
- Ağazade, S., Tükenmez, E.G.,Uzun, M. (2023). Is the volatility of international tourism revenues affected by tourism source market structure? An empirical analysis of Turkey, *Tourism Economics* 29 (2), 291-304.
- Tükenmez, E.G. (2023). Travel Motivations and Constraints of Solo Women Travelers in Turkey.*Journal of Mediterranean Tourism Research*, 2(1), 25-44.