

# Türkiye'deki Otel Konuk Yorumları ve Puanlarının Metin Madenciliği ile Analizi\*

Yılmaz AĞCA\*\*

Cemil GÜNDÜZ\*\*\*

## ÖZ

*Konukların kaldıkları oteller hakkında internette yaptıkları yorumlar ve verdikleri puanlar, tesiste kalmayı düşünen konukların tercihini etkileyen önemli faktörlerden biridir. Konuklar, bir otelde rezervasyon yapmadan önce çevrimiçi seyahat acentelerindeki veya seyahat platformlarındaki konuk yorumlarını okuyabilmekte, öyle karar verebilmektedirler.*

*Bu çalışmada konukların kaldıkları oteller için verdikleri puanlar ve yaptıkları yorumlar, metin madenciliği ile analiz edilmiştir. Bunun için bir çevrimiçi seyahat acentesinden Türkiye'deki konaklama tesisleri hakkında Türkçe yapılan yorumlar ve puanlar web madenciliği ile elde edilmiş, sonrasında metin madenciliği işlemine tabi tutulmuştur. Araştırmada 60252 Türkçe konuk yorumu ve puanı analiz edilmiştir. Buna göre Türkiye'deki konaklama tesislerinin ortalama konuk puanları 3,93'tür. En yüksek puanı Villa tipi tesisler almıştır (p=4,22; n=854). Coğrafi bölge olarak en yüksek puanı İç Anadolu bölgesi (p=4,07; n=5131), il olarak Nevşehir (p=4,53; n=2320) almıştır.*

*Araştırma kapsamında uygulanan metin madenciliği sonucunda, otel yorumlarında en çok tekrar eden tekil kelimeler puanlara göre gruplandırıldığında; konukların 1 puan verdikleri tesisleri tavsiye etmedikleri, 4 ve 5 puan verdikleri tesisleri tavsiye ettiklerine dair bulgular elde edilmiştir. Konukların düşük puan verdikleri tesislerde en çok oda, kahvaltı, su ve temizlik hakkında fikir belirttikleri belirlenmiştir. Konukların yüksek puan verdikleri tesislerde ise otelin temiz olduğu ve personelin konuga karşı ilgili olduğunu ifade eden kelimeler kullandıkları görülmüştür.*

*Araştırma sonucunda Türkiye'de konaklama tesislerine yönelik Türkçe yorumlarda genel olarak, beğenilmeme, dolayısıyla düşük puan verilmesine sebep olan faktörlerin; oda, kahvaltı, temizlik ve sıcak su problemi ile ilgili olduğu metin madenciliği sonucunda belirlenmiştir. Yüksek puan alınma neden olan faktörlerin de temizlik ve personelin ilgisi ile alakalı olduğu görülmektedir. Bu araştırmanın, konuk memnuniyeti, konuk şikâyetleri ve memnuniyetle alakalı faktörlerin bilinmesi açısından; sektör yöneticilerine, girişimcilere ve araştırmacılara katkı sağlayacağı düşünülmektedir.*

**Anahtar Kelimeler:** Metin Madenciliği, Otel Konuk Yorumları, Otel Konuk Puanları  
**JEL Sınıflandırması:** L83, C80

## Analysis of Hotel Guest Reviews and Ratings in Turkey with Text Mining

### ABSTRACT

*The comments and ratings of the guests on the internet about the hotels they stayed in are one of the important factors that affect the choice of the guests who are considering staying at the property. Before booking a hotel, guests can read guest reviews on online travel agencies or travel platforms and decide accordingly.*

\* Bu makale, 2-5 Mart 2020 tarihleri arasında Bandırma'da gerçekleştirilen 2. Uluslararası Bilimsel Araştırmalar ve Yenilikçi Çalışmalar Sempozyumu'nda "Metin Madenciliği ile Türkiye'deki Otel Yorumlarının ve Puanlarının Değerlendirilmesi" başlığı ile sözlü olarak sunulmuş bildirinin güncellenmiş ve genişletilmiş halidir.

\*\* Dr. Öğr. Üyesi, Tokat Gaziosmanpaşa Üniversitesi, Niksar Meslek Yüksekokulu, yilmaz.agca@gop.edu.tr, ORCID Bilgisi: 0000-0002-5912-0977

\*\*\* Dr. Öğr. Üyesi, Tokat Gaziosmanpaşa Üniversitesi, Niksar Meslek Yüksekokulu, cemil.gunduz@gop.edu.tr, ORCID Bilgisi: 0000-0002-0222-2497

*In this study, the ratings and comments of the guests for the hotels they stayed in were analyzed by text mining. For this, the comments and scores made in Turkish about the accommodation facilities in Turkey from an online travel agency were obtained by web mining, and then they were subjected to text mining. 60252 Turkish guest comments and scores were analyzed in the study. Accordingly, the average guest rating of accommodation facilities in Turkey is 3.93. Villa type facilities got the highest score ( $p=4.22$ ;  $n=854$ ). The Central Anatolia region ( $p=4.07$ ;  $n=5131$ ) got the highest score as a geographical region, and Nevşehir ( $p=4.53$ ;  $n=2320$ ) as a province.*

*As a result of the text mining applied within the scope of the research, when the most repeated single words in the hotel comments are grouped according to the scores; It has been found that the guests do not recommend the facilities they give 1 point, they recommend the facilities they give 4 and 5 points. It was determined that the guests mostly gave their opinions about the room, breakfast, water and cleanliness in the facilities they gave low scores. On the other hand, in the facilities where the guests gave high scores, it was observed that the hotel was clean and the staff used words expressing that they were interested in the guest.*

*As a result of the research, the factors that cause dislike and therefore a low score in Turkish comments on accommodation facilities in Turkey; It has been determined as a result of text mining that it is related to the problem of room, breakfast, cleaning and hot water. It is seen that the factors that cause high scores are also related to cleanliness and the interest of the staff. In terms of knowing the factors related to guest satisfaction, guest complaints and satisfaction; it is thought that it will contribute to sector managers, entrepreneurs and researchers.*

**Key Words:** Text Mining, Hotel Customer Reviews, Hotel Customer Rates

**JEL Classification:** L83, C80

## GİRİŞ

Günümüzde internet kullanımı dünyanın büyük bir bölümünde yaygın hale gelmiştir. İnsanlar birçok nedenden dolayı internet üzerinden alışveriş yapmaktadır. Bunlardan akla ilk gelen sağladığı *kolaylık* olmasına rağmen, bundan başka birçok neden bulunmaktadır. Yapılan araştırmalar *ücretsiz kargonun* internet üzerinden alışveriş yapılmasının güçlü nedenlerinden biri olduğunu göstermektedir (Boice, 2021). Bir diğer neden internet üzerinden alışverişlerde devasa *ürün çeşitliliğine* ulaşılabilir oluşudur (Prajapati, 2021). Bir perakende işletmesinde yer alan ürün sayısı daha sınırlı olmaktadır. İnsanların internet üzerinden alışveriş yapmalarının bir diğer sebebi daha *uygun fiyatlardır* (Chevalier, 2021). İnternet üzerinden satış işlemi yapan şirketler, perakende bir mağazaya ihtiyaç duymadıkları için daha düşük maliyetlere sahip olabilmektedirler. Bu nedenle fiyat rekabetinde bu durum avantaj sağlayabilmektedir. Günün her vaktinde ulaşılabilir oluşu, fiyatları karşılaştırma ve burada verilmeyen daha birçok neden insanları çevrimiçi alışverişe yönlendirmektedir (Henneberry, 2012; Chaffey, 2017).

İnsanların çevrimiçi alışverişini tercih etme nedenlerinden biri de *tüketici yorumları/görüşleridir*. Bu görüşler o ürünü ilk kez alacak kişiler için yol gösterici bir özellik taşımaktadır. Bu nedenle satın alınmak istenen ürüne, o ürünü daha önce satın alanlar tarafından yapılan yorumlar, insanları çevrimiçi alışveriş konusunda motive eden önemli unsurlardan biridir (Keskinlik, Ağca ve Karaman, 2016; Cui, Lui ve Guo, 2014; Schuckert, 2015; Hu, Bose, Koh ve Liu, 2012). Bu durum sadece sanayi ürünleri için değil hizmet sektörü için de geçerlidir. Otel, restoran, kafe vb. pek çok işletme türü için tüketici yorumlarının yer aldığı web siteleri veya akıllı telefon uygulamaları mevcuttur. Birçok kullanıcı daha önce o işletmeyi ziyaret edenlerin yaptıkları yorumları okuyarak o işletme hakkında ön bilgi edinmektedir.

Bu yorumlar tüketicilere işletme hakkında bilgi vermenin yanı sıra işletmeciler için de önemli geri bildirim verisi oluşturmaktadır. Dolayısıyla işleme eleştiri aldığı şeyleri düzeltme şansı bulmaktadır. Hatta bazı sitelerde yapılan eleştiriye, işletme tarafından cevap verme olanağı da bulunmaktadır.

İnternetteki tüketici yorumları üzerine pek çok araştırma bulunmaktadır (Öğüt ve Taş, 2012; Melián-González, Bulchand-Gidumal ve López-Valcárcel, 2013; Somprasertsri ve Lalitrojwong, 2010; Trenz ve Berger, 2013; Wei ve Lu, 2013; Zhan, Loh ve Liu, 2009). Bu çalışmalar tüketici yorumlarını çeşitli analizlere tabii tutarak bilgi, örüntü çıkarmaya yönelik yapılmıştır. *Metin madenciliği* olarak adlandırılan bu yöntemler istatistik, makine öğrenmesi, web madenciliği gibi birçok aracı kullanarak metin içerisinden bilgi çıkarımı, duygu analizi, belge sınıflandırma, arama optimizasyonu gibi amaçlarla kullanılmaktadır (Miner, vd., 2012). Son yıllarda gelişen bilgisayarlar dolayısıyla artan işlem gücü, depolama kapasitelerinin artması, internetin genişlemesi ve dolayısıyla artan veri miktarı, mobil teknoloji, sosyal ağlar gibi pek çok nedenden dolayı (Ağca, 2021a) metin madenciliği uygulamaları artmıştır.

Metin madenciliği araştırmalarına diğer e-ticaret sektörleri gibi turizm sektöründe de sıklıkla rastlanmaktadır (Xu ve Li, 2016; Kuhzady ve Ghasemi, 2019; Chittiprolu, Samala ve Bellamkonda, 2021; Shen, Yang, Liu ve Li, 2021; Hananto, 2015; Lau, Lee ve Ho, 2015). Bu araştırmalar genellikle duygu analizi, veri görselleştirme, n-grams, belge sınıflandırma ve kelime sıklığı gibi yöntemleri müşteri yorumları üzerinde kullanarak, çıkarsama yapmaya odaklanmaktadır. Anlaşılacağı üzere diğer pek çok sektör gibi müşteri yorumları turizm sektörü için de oldukça önem arz etmektedir. Bu yorumlar, daha önce de bahsedildiği gibi, o tesisi tercih etme potansiyeli olan kişiler için dikkate alınan bir seçenektir (Keskinkılıç, Ağca ve Karaman, 2016). Bu çalışmada günümüzde otel tercihlerinde tüketicilerin sıklıkla faydalandığı çevrimiçi yorumlar üzerinde durulmuştur. Bu kapsamda, Türkiye'deki otellerin çevrimiçi yorumları üzerinde metin madenciliği yapılmıştır. Yapılan metin madenciliği uygulaması ile tesislerin aldıkları puanlara göre, kelime sıklıkları analiz edilmiştir. Buna ek olarak, bu çalışmada analiz edilen 60252 yorum ile Türkiye'de konukların yaptığı otel yorumları hakkında genellemeler de yapılabilecektir. Araştırmada konuklar tarafından verilen puanların tesis tiplerine, coğrafi bölgelere, illere ve şehirlere göre dağılımı hakkında çeşitli bilgiler de sunulmuştur.

Çalışmanın ilk bölümünde araştırmanın daha iyi anlaşılması için metin madenciliği hakkında çeşitli bilgiler verilmektedir. Ayrıca bu konuda daha önce yapılmış araştırmalara da yer verilmiştir. Sonrasında araştırmada kullanılan yöntem açıklanmaktadır. Bulgular kısmında ilk olarak konukların tesislere verdikleri puanlar, tesis tipine, coğrafi bölgeye, il ve şehirlere göre gruplandırılarak verilmektedir. Sonrasında verilen puanlara göre en çok tekrar eden kelimeler ve ikili kelimeler verilmektedir. Bu kelimeler kullanılarak, 1'den 5'e kadar olan puan gruplarında en çok tekrar eden kelimelerin listesi çıkarılmıştır. Böylelikle tesisin en çok eleştirilen veya en çok beğenilen özelliklerinin belirlenmesi amaçlanmaktadır.

## I. LİTERATÜR İNCELEMESİ

*Metin madenciliği*, kitaplar, web sayfaları, e-postalar, raporlar veya ürün açıklamaları gibi çeşitli belge türlerinin yapılandırılmamış metinsel verilerindeki ilginç kalıpların araştırılması veya tanımlanması yoluyla belgelerden yararlı bilgiler çıkarmaya çalışan bir dizi tekniği açıklayan bir terimdir (Truyens ve Eecke, 2014; Ağca, 2021b). Metin madenciliği, bilgi ayıklama, bilgi alma, doğal dil işleme, kümeleme ve kategorizasyon gibi amaçlarla yapılmaktadır (Weiss, Indurkha ve Zhang, 2015; Rai, 2019; Tyagi, 2021; Gaikwad, Chaugule ve Patil, 2014). Daha önce de bahsedildiği gibi gelişen bilgi ve iletişim teknolojileri metin madenciliğinin önünü açmıştır. Birçok alanda metin madenciliği yapılmaktadır. Gerek ticari işletmelerde gerekse bilimsel araştırmalarda metin madenciliği kullanılmaktadır. Özellikle ticari işletmelerde ürün ve hizmetlerin iyileştirilmesi açısından bu yorumlar/görüşler önemli değer taşımaktadır (Kim ve Srivastava, 2007; Singh, Sachdeva, Mahajan, Pande ve Sharma, 2014; Akter ve Wamba, 2016; Palese ve Usai, 2018).

Yazında müşteri görüşleri/yorumları (customer reviews) metin madenciliğinin önemli bir kullanım alanını oluşturmaktadır. Bu araştırmalar genellikle yorumlardan bilgi çıkarımı, özet çıkarma, olumlu/olumsuz yorumların belirlenmesi gibi amaçlarla yapılmıştır. Pek çok kişinin tatil/otel tercihlerinde konukların kaldıkları oteller hakkında yaptıkları olumlu ve olumsuz yorumlar etkili olmaktadır (Keskinkılıç, Ağca ve Karaman, 2016; Chanwisitkul, Shahgholian ve Mehandjiev, 2018; Nicholas ve Lee, 2017). Bu öneminden dolayı, konuk yorumları üzerine birçok araştırma bulunmaktadır. Örneğin Jeong ve Jeon (2008) New York'taki otellerde kalan konukların yorumları üzerine yaptıkları araştırmalarda müşteri memnuniyeti açısından *fiyatların*, *otelin konumunun* etkili olduğu sonucuna varmışlardır. Moskova'da yapılan başka bir çalışmada 5830 yorum değerlendirilmiş, konukların memnun kalmadıkları oteller için daha çok yorum yaptıkları sonucuna ulaşılmıştır. Bu çalışmada müşterilerin 5 farklı kategoride görüşleri değerlendirilmiştir. Yüksek puan verilen otellerde konukların daha çok deneyimlerinden bahsettikleri, orta puan verilen otellerde olanaklar ve konumdan, düşük puan verilen otellerde ise işlemler ve fiyattan bahsettikleri bulgusuna erişilmiştir (Han, Mankad, Gavirneni ve Verma, 2016). Tayland'da yapılan bir çalışmada Bangkok'taki bazı oteller değerlendirmeye alınmış, metin madenciliği sonucunda yüksek puan verilen yorumlarda dost canlısı ve yardımsever personelin, düşük puan verilen yorumlarda havuz ve Wifi çok etkili olduğu görülmüştür. Ayrıca bu çalışmada müşteri memnuniyetini etkileyen diğer faktörler arasında temizlik, oda ve banyo içi, uyku kalitesi ve konumun etkili olduğu belirlenmiştir (Chanwisitkul, Shahgholian ve Mehandjiev, 2018). Başka bir çalışmada Nicholas ve Lee (2017) yaptıkları metin madenciliği ile bir otelin tercih edilmesinde temizliğin, az gürültünün, uyku kalitesinin, oteldeki fiyatların ve otellerin aileye uygunluğunun etkili olduğu sonucuna varmışlardır.

Türkiye'de yapılan çalışmalara bakıldığında Yılmaz ve Aytekin (2018) TripAdvisor sitesinden en fazla yorum yapılan 25 otel ve bu otellerden gelen 9000 yorum üzerinde yürüttükleri çalışmada, otellerin konum ve erişilebilirlik kriterleri,

otellerin sunduğu yiyecek içecek konsepti, personel–hizmet kalitesi, temizlik elemanları, eğlence–animasyon programlarının otellere verilen derecelendirmelerde etkili olduğunu tespit etmişlerdir. Antalya’da otellerin de dahil olduğu başka bir çalışmada, 3588 İngilizce yorum analiz edilmiş, analiz sonucunda konukların oda, personel ve yiyecekler hakkında yorum yaptıkları sonucuna ulaşılmıştır (Atabay ve Cizel, 2020).

## II. YÖNTEM

Bu çalışmada otel, pansiyon, hostel, villa gibi konaklama tesislerine, daha önce bu tesislerde konaklamış olan konuklar tarafından yapılan *yorumlar* metin madenciliği ile analiz edilmiştir. Araştırmada, bir çevrimiçi seyahat acentası/gezi/tatil web sitesinden alınan, 60252 adet Türkçe konuk yorumu analiz edilmiştir. Verinin elde edilmesinde hedef site için araştırma sürecinde geliştirilmiş bir algoritma kullanılmıştır. Algoritmanın geliştirilmesinde Python programlama dili ve bu dilin kütüphaneleri olan Beautiful Soup ve Selenium’den faydalanılmıştır. Araştırma verisi 10/2020 tarihinde elde edilmiştir. Metin madenciliği işlemlerinde Knime Analytics Platform yazılımı kullanılmıştır. Metin madenciliği işleminde işlenen metin genellikle yarı yapılandırılmış veya yapılandırılmamış halde olmaktadır. Bu metinlerin analize hazırlanması ve gereksiz kelimelerden arındırılması için çoğu zaman bir ön işleme işlemi yapılmaktadır (Tyagi, 2021; Ağca, 2021c). Bu çalışmada veri setindeki Türkçe yorumlar seçildikten sonra, veri temizleme, tokenize ve filtreleme işlemleri uygulandıktan sonra metinler dizi haline getirilmiştir. En çok tekrar eden kelimelerden olan “*otel*” kelimesi gereksiz görülerek veri setinden çıkartılmıştır. Tek sözcüklü frekanslar alınırken “*bu, çok, için, ama, daha, da, gibi, kadar, en, o*” gibi kelimeler dikkate alınmamıştır. İki kelimeli frekanslarda bu kelimeler anlamı nitelendireceğinden temizlenmemiştir. İkili kelimelerde anlam eksikliği olacağı düşünüldüğünden temel *stopwords* işlemi bu çalışmada uygulanmamıştır. Benzer sebeplerden dolayı veri setine kelimelerin köklerine ayrılması işlemi olan *stemming* işlemi de uygulanmamıştır (Pradana ve Hayaty, 2019). N-Gram işleminde art arda gelen ikili kelimeler alınmıştır. Sadece o puan türüne özgü kelime öbeklerini bulmak için Everygram olarak adlandırılan yöntem kullanılmıştır. Bunun için bu yöntemi destekleyen NLTK kütüphanesi kullanılmıştır. Everygram işleminde önce bir yorum içerisinde birlikte yer alan kelimelerin puan türüne göre frekansları çıkartılmıştır. Sonrasında aynı puan türlerinde tekrar eden ikili kelime öbeklerinden frekansı en yüksek olan kalacak şekilde diğerleri silinmiştir. Bu işlemin amacı puana özgü kelime öbeklerinin bulunmasıdır.

## III. BULGULAR

Yapılan çalışmada toplam 60252 yorum değerlendirilmiştir. Tesis tipi olarak herhangi bir tip belirtmeyenlerle birlikte 10 tesis tipi bulunmaktadır. Tesis tipleri, yüksek puan alandan düşük puan alana sıralanmış şekilde puan ve yorum sayıları ile birlikte Tablo 1’de verilmektedir.

**Tablo 1:** Tesis Tiplerine Göre Yorum Sayıları ve Tesislerin Aldıkları Ortalama Puanlar

Sıra	Tesis Tipi	Puan	Yorum
1	Villa	4,22	854
2	Hostel	4,11	900
3	Ranch	4,09	90
4	Cottages	4,06	98
5	Apart Otel	4,04	3656
6	Kamp Alanları	3,92	952
7	Otel	3,91	32079
8	Pansiyon	3,91	18116
9	Lodge	3,91	3183
10	Belirtilmemiş	3,87	324

Tablo 1 incelendiğinde en yüksek ortalama puana Villa tipi tesislerin sahip olduğu görülmektedir. Tesis tipini belirtmeyen tesislerin en düşük ortalama puana sahip olduğu görülmektedir. Hotel, Ranch, Cottages ve Apart Otel tipinde tesislerin ortalamasının üzerinde puan aldığı görülmektedir.

Tablo 2'de ortalama puanlar ve yorum sayıları coğrafi bölgelere göre kategorize edilerek verilmektedir.

**Tablo 2:** Coğrafi Bölgelere Göre Yorum Sayıları ve Bölgelerin Ortalama Puanları

Sıra	İl	Puan	Yorum
1	İç Anadolu	4,07	5131
2	Ege	3,95	18957
3	Marmara	3,94	17948
4	Akdeniz	3,88	12021
5	Güneydoğu Anadolu	3,83	923
6	Karadeniz	3,82	4230
7	Doğu Anadolu	3,69	1042

Tablo 2'de 7 bölgeye ait ortalama puanlar ve yorum sayıları verilmiştir. Buna göre en yüksek puan alan bölgenin İç Anadolu bölgesi en düşük puan alan bölgenin Doğu Anadolu bölgesi olduğu görülmektedir.

Tablo 3'te en çok yorum alan ve en az yorum alan iller puan ve yorum sayıları ile birlikte verilmektedir. Yorum sayısı aynı zamanda talebi de ifade edebileceği için bu tablo verilmektedir.

**Tablo 3:** En Çok ve En Az Yorum Alan İllerin Yorum Sayıları ve Puanları

Sıra	İl	En Çok		En Az		
		Puan	Yorum	İl	Puan	Yorum
1	Muğla	3,98	10914	Siirt	4,00	9
2	İstanbul	3,99	10555	Kilis	3,93	14
3	Antalya	3,92	10277	Bingöl	2,50	16
4	İzmir	4,03	5297	Tunceli	3,81	16
5	Çanakkale	4,12	2509	Bayburt	4,28	18
6	Balıkesir	3,78	2341	Hakkâri	3,80	20
7	Nevşehir	4,53	2320	Şırnak	3,55	22
8	Aydın	3,62	1772	Gümüşhane	4,04	25
9	Ankara	3,59	1488	Osmaniye	4,40	25
10	Trabzon	3,71	822	Muş	3,96	26

En çok yorum alan illere bakıldığında önemli turizm destinasyonlarından Muğla, İstanbul ve Antalya'nın ilk üç sırada olduğu görülmektedir. En az yorum alan ile bakıldığında bu yerin 9 yorumla Siirt olduğu görülmektedir.

Tablo 4'te en yüksek puan alan ve en düşük puan alan iller ortalama puan ve yorum sayıları ile birlikte verilmektedir.

**Tablo 4:** En Çok ve En Az Puan Alan İllerin Yorum Sayıları ve Puanları

Sıra	En Yüksek			En Düşük		
	İl	Puan	Yorum	İl	Puan	Yorum
1	Nevşehir	4,53	2320	Bingöl	2,50	16
2	Karabük	4,43	406	İğdir	2,87	30
3	Osmaniye	4,40	25	Kırşehir	3,07	28
4	Karaman	4,30	44	Çankırı	3,08	40
5	Bayburt	4,28	18	Yozgat	3,16	32
6	Çanakkale	4,12	2509	Van	3,36	147
7	Erzincan	4,11	96	Yalova	3,45	209
8	Batman	4,05	39	Mersin	3,46	672
9	Adıyaman	4,05	86	Çorum	3,47	88
10	Gümüşhane	4,04	25	Kırklareli	3,48	134

En yüksek puan alan işe bakıldığında ilk sırada *Nevşehir* yer almaktadır. En düşük puan alan il ise *Bingöl*'dür.

Tablo 5'te en yüksek ve en düşük puan alan şehirler verilmektedir. Bu tablolarda verinin elde edildiği sitede ilçe, kasaba, köy ayrımı olmadığı için bu üç sınıfa giren yerleşim yerleri açıklamalarda *şehir* olarak adlandırılmıştır.

**Tablo 5:** En Çok ve En Az Puan Alan Şehirlerin Yorum Sayıları ve Puanları

Sıra	En Yüksek			En Düşük		
	Şehir	Puan	Yorum	Şehir	Puan	Yorum
1	Ayvalı	4,79	19	Armutlu	2,83	40
2	Adatepe	4,64	11	Başiskele	3,12	26
3	Avanos	4,59	142	Akçakoca	3,20	90
4	Beyşehir	4,50	10	Akhisar	3,26	19
5	Bozcaada	4,37	984	Aliağa	3,32	19
6	Akyaka	4,36	271	Beypazarı	3,32	75
7	Akşehir	4,33	36	Ağlasun	3,50	10
8	Bayburt	4,28	18	Boğazkale	3,50	26
9	Ayvacık	4,26	112	Amasya	3,62	185
10	Bozüyük	4,20	10	Alanya	3,63	2073

Tablo 5 incelendiğinde en yüksek puan alan iller sıralamasında ilk sırada yer alan *Nevşehir* ili burada da iki şehir ile ilk üçte yer almaktadır (*Ayvalı*, *Avanos*). *Sırası ile Adatepe*, *Bozcaada* ve *Ayvacık Çanakkale* ili sınırları içerisinde yer alırken listeye *Konya* iki şehirle girmiştir (*Beyşehir*, *Akşehir*).

En düşük puan alan şehirlere bakıldığında ilk sırada *Armutlu* yer almaktadır (*Yalova*). Bu şehri *Başiskele* (*Kocaeli*) ve *Akçakoca* (*Düzce*) takip etmektedir. Konukların tesislere verdiği puanlar bu şekildedir. Tablo 6'da ise konukların yaptığı *yorumlar* içerisinde *en çok geçen kelimeler* puanlara göre verilmektedir.

**Tablo 6:** Yorumlarda En Çok Geçen Kelimelerin Konuk Puanlarına Göre Dağılımı

Sıra	1. Puan		2. Puan		3. Puan		4. Puan		5. Puan	
	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>
1	yok	3069	yok	1050	iyi	2098	temiz	3320	güzel	11240
2	oda	2429	iyi	986	güzel	1754	güzel	3303	temiz	9617
3	var	2092	oda	979	var	1660	iyi	3188	tavsiye	8378
4	kötü	1914	var	901	kahvaltı	1642	kahvaltı	2292	iyi	7859
5	yer	1823	kahvaltı	823	temiz	1583	oda	2253	teşekkür	6754
6	hiç	1697	değil	739	alman	9	var	2167	odalar	6515
7	değil	1695	sadece	605	oda	1526	tavsiye	1665	yer	5979
8	iyi	1534	yer	599	yok	1329	personel	1590	güler	5627
9	tavsiye	1448	kötü	598	değil	1253	yakın	1547	harika	5589
10	kesinlikle	1447	hiç	548	biraz	1110	yok	1302	kahvaltı	5508
11	kahvaltı	1440	temiz	502	yer	1071	memnun	1301	memnun	5267
12	sadece	1374	su	457	personel	916	ilgili	1213	var	4886

13	berbat	1287	temizlik	435	küçük	887	değil	1210	bey	4760
14	su	1279	küçük	430	yakın	882	küçük	1200	tekrar	4722
15	pis	1076	eski	414	konumu	775	tek	1178	personel	4601

Tablo 6 incelendiğinde “*yok, değil, hiç*” gibi negatif kelimelerin veya “*var, iyi*”, gibi pozitif kelimelerin hem düşük hem de yüksek puanlarda yer aldığı görülmektedir. Tahmin edileceği üzere bu kelimeler tek başlarına negatif veya pozitif anlam taşıyalar da cümle içerisinde tersi anlamda kullanılabilir. Bunun dışındaki kelimelere odaklanıldığında daha anlamlı sonuçlar ortaya çıkmaktadır. Konuklar tarafından *1 puan, 4 puan* veya *5 puan* verilen yorumlara bakıldığında ortak olarak *tavsiye* kelimesinin en çok tekrarlanan kelimelerden olduğu görülmektedir. Buradan 1 puan verilen tesislerin *tavsiye edilmediği*, 4 veya 5 puan verilen tesislerin ise aksine *tavsiye edildiği* düşünülebilir. Diğer puan türlerinde en çok tekrarlanan ilk 15 kelime listesine *tavsiye* kelimesinin girmediği görülmektedir.

Düşük puan olarak nitelendirilebilecek 1 ve 2 puan olan yorumlara bakıldığında “*oda, kahvaltı, pis, su, temizlik*” gibi kelimelerin sıklıkla tekrar edildiği görülmektedir. Buradan oda, kahvaltı, temizlik ve su gibi problemlerin veya memnuniyetsizliklerin düşük puan verilmesine neden olduğu düşünülebilir.

Yüksek puan olarak nitelendirilebilecek 4 ve 5 puan olan yorumlara bakıldığında ilk iki sırada “*temiz*” ve “*güzel*” kelimeleri görülmektedir. Ayrıca bu yorumlarda “*personel*” kelimesi de ön plana çıkmaktadır. Ayrıca yine personelle alakalı kelimeler olan “*ilgili, güler, bey*” gibi kelimeler de yüksek puan verilen yorumlarda çok tekrar etmektedir. 5 puan verilen yorumlarda diğer puan seviyelerinden farklı olarak tekrar kelimesi listeye girmiştir.

Yukarıda yorumlarda en çok tekrar eden kelimeler verilmişken, Tablo 7’de *arka arkaya en çok tekrar eden kelime çiftleri* verilmektedir.

**Tablo 7:** Yorumlarda En Çok Geçen Çift Kelimelerin Konuk Puanlarına Göre Dağılımı

Sıra	1. Puan		2. Puan		3. Puan		4. Puan		5. Puan	
	Kelime	fı	Kelime	fı	Kelime	fı	Kelime	fı	Kelime	fı
1	bu kadar	689	daha iyi	194	çok iyi	477	çok iyi	967	çok güzel	4088
2	tavsiye etmiyorum	609	çok kötü	177	güler yüzlü	388	tavsiye ederim	900	tavsiye ederim	3990
3	çok kötü	576	bu kadar	175	çok güzel	381	çok güzel	884	güler yüzlü	3988
4	daha iyi	450	çok iyi	139	daha iyi	334	güler yüzlü	845	çok iyi	2957
5	en kötü	384	çok güzel	138	genel olarak	328	genel olarak	531	çok memnun	2760
6	zorunda kaldık	288	genel olarak	127	konum olarak	252	memnun kaldık	505	çok teşekkür	2674
7	sıcak su	282	olmasın a rağmen	115	odalar temiz	236	yürüme mesafesinde	499	memnun kaldık	2225
8	pislik içinde	230	konum olarak	114	yürüme mesafesinde	228	çok yakın	451	her şey	2037
9	çok eski	219	güler yüzlü	114	fena değil	223	odalar temiz	442	çok temiz	1812
10	bile yok	205	çok küçük	109	çok yakın	220	çok memnun	384	kesinlikle tavsiye	1570



11	ertesigün	205	zorunda kaldık	100	otelin konumu	215	çok temiz	356	herkese tavsiye	1479
12	odalar çok	181	çok eski	98	idare eder	203	açık büfe	350	çok yakın	1320
13	leş gibi	180	sıcak su	98	açık büfe	185	konum olarak	328	yürüme mesafesinin de	1316
14	kötü otel	179	tavsiye etmiyorum	96	tavsiye ederim	183	gayet iyi	288	çok güzeldi	1223
15	çok küçük	169	odalar çok	94	çok küçük	182	personel çok	284	çok ilgili	1107

Tablo 7 incelendiğinde düşük puan veren konuklara ait yorumlarda “*tavsiye etmiyorum, çok kötü, en kötü*” gibi negatif kelimelerin yanı sıra yine temizlikle ilgili “*pislik içinde*” gibi kelimeler göze çarpmaktadır. Daha önce tek kelime frekanslarında yer alan “*su*” kelimesi burada da “*sıcak su*” şeklinde yer almaktadır. Bu durumda konukların sıcak su ile ilgili problemleri sıklıkla dile getirdikleri düşünülebilir. 1 ve 2 puan verilen tesislerin yorumlarında “*çok eski, çok küçük, zorunda kaldık, tavsiye etmiyorum*” ifadelerinin ortak olduğu görülmektedir. 4 ve 5 puan verilen tesislerin yorumlarında ise “*tavsiye ederim, güler yüzlü, memnun kaldık, çok temiz*” ifadelerin sıklıkla kullanıldığı görülmektedir. Ayrıca olumlu yorumlarda tesislerin konumlarıyla ilgili “*konum olarak, yürüme mesafesinde*” gibi ifadelerin de olduğu görülmektedir. Ek olarak 4 veya 5 puan verilen tesislerin yorumlarında “*güler yüzlü, çok ilgili*” gibi *çalışanlarla ilgili* ifadelerin olduğu görülmektedir.

Tablo 7’de arka arkaya kullanılan kelimeler verilmişken, Tablo 8’de ise aynı metin içerisinde birlikte tekrar eden fakat o puan türüne özgü kelimeler verilmektedir. Tablo 8 hazırlanırken öncelikle aynı puan grubunda birlikte tekrar eden kelime çiftleri bulunmuştur. Ardından bütün puan gruplarında ortak olan kelimeler tablodan silinmiştir. Bu işlem sonucunda geriye sadece belirli bir puan türünde yer alan kelime çiftleri kalmıştır. Kalan kelimelerden en çok tekrar eden, ilgili puan türüne yazılmıştır. Elde edilen sonuçlar Tablo 8’de verilmektedir.

**Tablo 8:** Yorumlarda Arka Arkaya En Çok Geçen Kelime Çiftlerinin Puanlara Göre Dağılımı

Sıra	1. Puan		2. Puan		3. Puan		4. Puan		5. Puan	
	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>	Kelime	f <sub>i</sub>
1	çok kötü	642	fena değil	88	konumu çok	143	çok memnun	406	görüştük üzere	969
2	tavsiye etmiyorum	622	otelin konumu	87	gayet iyi	128	teşekkür ederiz	278	çok teşekkürler	711
3	daha iyi	476	idare eder	84	personel çok	121	çok teşekkür	222	tekrar görüşmek	695
4	en kötü	398	güzel ama	82	en iyi	115	teşekkür ederim	200	her şey çok	694
5	kesinlikle tavsiye	394	temiz değildi	81	fiyatına göre	112	gayet temiz	188	gönül rahatlığıyla	667
6	tavsiye etmem	339	biraz daha	72	küçük ama	101	çok rahat	187	şiddetle tavsiye	597
7	sıcak su	334	onun dışında	58	gayet güzel	100	çok uygun	171	teşekkür ediyoruz	591

8	zorunda kaldık	308	ne yazık	57	en büyük	95	memnun kaldım	158	olmak üzere	575
9	pislik içinde	255	değil ama	56	memnun kaldık	92	güzel temiz	145	çok lezzetli	542
10	olması na rağmen	247	olarak çok	56	temiz ama	90	çalışanlar çok	144	güzel tatil	534
11	çok eski	229	yazık ki	56	odalar küçük	86	çok iyiydi	143	tavsiye ederiz	503
12	ertesi gün	221	sabah kahvaltısı	55	deniz manzaralı	85	oldukça iyi	139	çok güler	489
13	daha fazla	220	iyi değil	54	yok ama	83	çok yardımcı	138	konuda yardımcı	489
14	daha asla	216	yürüme mesafesinde	53	çok ilgili	80	aile işletmesi	126	çok keyifli	483
15	bile yok	215	çok yakın	50	cana yakın	79	yürüyüş mesafesinde	126	tatil geçirdik	474

Tablo 8 incelendiğinde 1 puandan 5 puana kadar sırası ile en çok tekrar eden kelime çiftlerinin “*çok kötü, fena değil, konumu çok, çok memnun*” ve “*görüştük üzere*” olduğu görülmektedir.

## SONUÇ

Çalışmada konaklama tesislerini ziyaret eden konukların yaptıkları çevrimiçi yorumlar metin madenciliği ile analiz edilmiştir. Araştırmada düşük veya yüksek puan alan tesislerde en çok tekrar edilen kelimeler tespit edilmiştir. Böylelikle konukların düşük puan verdikleri tesislerde nelerden şikâyet ettiklerinin veya yüksek puan verdikleri tesislerde neleri beğendiklerinin belirlenmesi amaçlanmıştır. Ayrıca konukların tesislere verdikleri puanlar; tesis tipine, coğrafi bölgeye, il ve şehre göre değerlendirilmiştir.

Araştırmada metin madenciliği ile analiz edilen yorumlarda en çok geçen kelimeler puanlara göre gruplandırılmıştır. Benzer şekilde en çok art arda geçen kelimeler de puanlara göre gruplandırılmıştır. Buna göre “*yok, değil, hiç*” gibi negatif kelimeler veya “*var, iyi*” gibi pozitif kelimeler hem olumsuz yorumlarda hem de olumlu yorumlarda yer almaktadır. Bu kelimeler tek başlarına negatif veya pozitif anlam taşıyıcılar da cümle içerisinde tersi anlamda kullanılabilir. Konuklar tarafından *1 puan, 4 puan* veya *5 puan* verilen yorumlara bakıldığında ortak olarak *tavsiye* kelimesinin en çok tekrarlanan kelimelerden olduğu görülmektedir. Buradan *1 puan* verilen tesislerin *tavsiye edilmediği*, *4* veya *5 puan* verilen tesislerin ise aksine *tavsiye edildiği* düşünülebilir. Düşük puan olarak nitelendirilebilecek *1* ve *2* puan olan yorumlara bakıldığında “*oda, kahvaltı, pis, su, temizlik*” gibi kelimelerin sıklıkla tekrar edildiği görülmektedir. Buna göre konaklama tesislerinin aldığı olumsuz yorumların tesislerin odaları, kahvaltıları, su problemleri ve temizlikleri ile ilgili olduğu düşünülebilir.

Yüksek puan olarak nitelendirilebilecek *4* ve *5* puan olan yorumlara bakıldığında ilk iki sırada “*temiz*” ve “*güzel*” kelimeleri görülmektedir. Ayrıca bu yorumlarda *personel* kelimesi de ön plana çıkmaktadır. Ayrıca yine personelle alakalı kelimeler olan “*ilgili, güler, bey*” gibi kelimeler de yüksek puan verilen

yorumlarda çok tekrar etmektedir. 5 puan verilen yorumlarda diğer puan seviyelerinden farklı olarak tekrar kelimesi listeye girmiştir. Buna göre konukların 5 puan verdikleri tesislere “*tekrar gitme*” isteği içerisinde oldukları düşünülebilir. Yazında konuk yorumları üzerine Jeong ve Jeon (2008) yaptıkları çalışmada fiyat ve otelin konumunun memnuniyet üzerinde etkili olduğu sonucuna varmışlardır. Bu araştırmada ise *konum* ve *fiyat* kelimelerine sadece 3 puan almış olan tesislerde rastlanmaktadır. Han, Mankad, Gavirneni ve Verma, 2016 yaptıkları çalışmada, 5830 yorum değerlendirilmiş, konukların memnun kalmadıkları oteller için daha çok yorum yaptıkları sonucuna ulaşılmıştır. Bu çalışmada düşük puan verilen tesisler için, *işlemler* ve *fiyattan* bahsedildiği belirtilmiştir. Bu çalışmada ise bahsedildiği gibi *fiyat* kelimesi düşük puanlı tesislerde üst sıralarda geçmemektedir.

Art arda kullanılan kelime öbeklerine bakıldığında düşük puanlı yorumlarda “*tavsiye etmiyorum, çok kötü, en kötü*” gibi negatif kelimelerin yanı sıra yine temizlikle ilgili *pislik* *içince* gibi kelimelerin geçtiği görülmektedir. Ayrıca düşük puanlı yorumlarda yorumlarında “*çok eski, çok küçük, zorunda kaldık, tavsiye etmiyorum*” ifadelerinin ortak olduğu görülmektedir. 4 ve 5 puanlı yorumlarda ise “*tavsiye ederim, güler yüzlü, memnun kaldık, çok temiz*” ifadelerin sıklıkla kullanıldığı görülmektedir.

Konuk puanlarına özgü kelime çiftlerini belirlemeye yönelik yapılan metin madenciliğinde; 1’den 5 puana kadar en çok tekrar eden kelime çiftlerinin “*çok kötü, fena değil, konumu çok, çok memnun*” ve “*görüşmek üzere*” olduğu görülmektedir. Buradan 5 puan veren konuklara özgü kelime çifti olan “*görüşmek üzere*” ‘den konukların o tesisi yeniden ziyaret etme isteğinde olduğu görülmektedir. 2 puan verilen tesislerin yorumlarında en çok tekrar eden kelime öbeği “*fena değil*” olmuştur. Buradan o tesisteki bazı şeylerin beğenildiği sonucu çıkarılabilir. Fakat yine de diğer beğenilmeye şeyler tesisin 2 puan almasına neden olmuştur. 4 ve 5 puan verilen tesislerde işletmeye veya işletme çalışanlarına “*teşekkür edildiğini*” gösteren kelimeler görülmektedir.

Sonuç olarak Türkiye’de konaklama tesislerine yapılan Türkçe yorumlarda genel olarak, beğenilmeme, dolayısıyla düşük puan verilmesine sebep olan faktörlerin; oda, kahvaltı, temizlik ve sıcak su problemi ile ilgili olduğu metin madenciliği sonucunda belirlenmiştir. Benzer olarak yüksek puan alımına neden olan faktörlerin de *temizlik* ve *personelin ilgisi* ile alakalı olduğu görülmektedir. Bu araştırmada olduğu gibi daha önce yapılan pek çok çalışmada temizliğin müşteri memnuniyeti üzerinde etkili olduğu belirlenmiştir (Dominici ve Guzzo, 2010; Xu ve Li, 2016; Davras ve Caber, 2019). Yazında benzer şekilde konaklama tesisi çalışanlarının müşteri memnuniyet üzerinde etkili olduğuna dair birçok çalışma bulunmaktadır (Albayrak ve Caber, 2015; Ekinci, Dawes ve Massey, 2008). Turizm yapısı gereği emek yoğun bir sektördür (O’Fallon ve Rutherford, 2011: 175; Kumar ve Aulia, 2021: 12). Bu çalışmadan elde edilen sonuçlar da otel çalışanlarının müşterilerin yaptıkları yorumlar üzerinde etkili olduğunu göstermektedir.

Elde edilen bulguların, otellerin hizmet kalitelerinin iyileştirilmesi açısından, konukların memnun kaldıkları veya kalmadıkları ürün veya hizmetler bağlamında, iştirakçi, yönetici ve çalışanlara kaynaklık edeceği düşünülmektedir.

Ayrıca araştırma, metin madenciliği açısından Türkçe dili üzerine yapılan kapsamlı araştırmalardan biridir. Bu açıdan bundan sonraki araştırma ve çalışmalara katkı sağlayacağı düşünülmektedir.

Araştırma sadece Türkçe dilinde yapılan yorumlar kullanılarak yapılmıştır. Başka dillerde yapılacak olan çalışmalar araştırmanın genişletilmesi açısından faydalı olacaktır. Ayrıca farklı ülke vatandaşlarının otel beklentileri ve memnuniyetleri hakkında farklar değerlendirilebilecektir. Çalışmada bölgesel olarak puan açısından bir değerlendirme yapılsa da metin madenciliği açısından bir değerlendirme yapılmamıştır. Daha sonra yapılacak çalışmalarda coğrafi bölgelerin veya illerin değerlendirilmesi, bölgelerin karakteristik özelliklerinin, beğenilen ve beğenilmeyen özelliklerinin bulunması açısından faydalı olacaktır.

### **Araştırma ve Yayın Etiği Beyanı**

Makalenin tüm süreçlerinde Yönetim ve Ekonomi Dergisi'nin araştırma ve yayın etiği ilkelerine uygun olarak hareket edilmiştir.

### **Yazarların Makaleye Katkı Oranları**

Yazarlar çalışmaya eşit oranda katkı sağlamıştır

### **Çıkar Beyanı**

Yazarın herhangi bir kişi ya da kuruluş ile çıkar çatışması yoktur.

### **KAYNAKÇA**

- Ağca, Y. (2021a). Otel Oda Fiyatlarını Açıklamada Makine Öğrenmesi Algoritmalarının Kıyaslanması. *İşletme Araştırmaları Dergisi*, 450-463
- Ağca, Y. (2021b). Alternatif Veri Elde Etme Yöntemi Web Madenciliği: Otel Oda Fiyatlarının Zamansal Analizi. *Çanakkale Onsekiz Mart Üniversitesi Yönetim Bilimleri Dergisi*, 19(42), 1013-1034
- Ağca, Y. (2021c). *R Programlama Dili ile İstatistiksel Analiz ve Veri Madenciliği*. İstanbul: Cinius Yayınları.
- Akter, S. ve Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda. *Electron Markets*, 173-194.
- Albayrak, T. ve Caber, M. (2015). Prioritisation of the hotel attributes according to their influence on satisfaction: a comparison of two techniques. *Tourism Management*, 46, 43-50.
- Atabay, L. ve Cizel, B. (2020). Comparative Content Analysis of Hotel Reviews by Mass Tourism Destination. *Journal of Tourism and Services*, 11(21), 147-166.
- Boice, M. (2021, 03 19). *The 13 Top Reasons Consumers Shop Online*. <https://www.junglescout.com/blog/reasons-consumers-shop-online/> adresinden alındı
- Chaffey, D. (2017, 11 19). *The reasons why consumers shop online instead of in stores*. <https://www.smartinsights.com/ecommerce/ecommerce-strategy/the-reasons-why-consumers-shop-online-instead-of-in-stores/> adresinden alındı
- Chanwisitkul, P., Shahgholian, A. ve Mehandjiev, N. (2018). The Reason Behind the Rating: Text Mining of Online Hotel Reviews. *2018 IEEE 20th Conference on Business Informatics (CBI)*. Vienna, Austria : IEEE.
- Chevalier, S. (2021, 07 26). *Main reasons why global consumers chose to shop online 2021*. statista.com: <https://www.statista.com/statistics/676358/reasons-online-shoppers-prefer-to-shop-online/> adresinden alındı
- Chittiprolu, V., Samala, N. ve 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.
- Cui, G., Lui, H.-K. ve Guo, X. (2014). The Effect of Online Consumer Reviews on New Product Sales. *International Journal of Electronic Commerce*, 17(1), 39-58.

- Davras, Ö. ve 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.
- Dominici, G. ve Guzzo, R. (2010). Customer Satisfaction in the Hotel Industry: A Case Study from Sicily. *International Journal of Marketing Studies*, 2(2), 3-12.
- Ekinci, Y., Dawes, P. ve Massey, G. (2008). Document details - An extended model of the antecedents and consequences of consumer satisfaction for hospitality services. *European Journal of Marketing*, 42(1-2), 35-68.
- Gaikwad, S. V., Chaugule, A. ve Patil, P. (2014). Text Mining Methods and Techniques. *International Journal of Computer Applications*, 85(17), 42-45.
- Han, H. J., Mankad, S., Gavirneni, N. ve Verma, R. (2016). What Guests Really Think of Your Hotel: Text Analytics of Online Customer Reviews. *Cornell Hospitality Report*, 16(2), 1-19.
- Hananto, A. (2015). Application of Text Mining to Extract Hotel Attributes and Construct Perceptual Map of Five Star Hotels from Online Review: Study of Jakarta and Singapore Five-Star Hotels. *ASEAN Marketing Journal*, 7(2), 58-80.
- Henneberry, R. (2012, 06 20). *How To Cash In On The 5 Reasons People Buy Products Online*. <https://www.crazyegg.com/blog/reasons-people-buy-products-online/> adresinden alındı
- Hu, N., Bose, I., Koh, N. S. ve Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, 52(3), 674-684.
- Jeong, M. ve Jeon, M. M. (2008). Customer Reviews of Hotel Experiences through Consumer Generated Media (CGM). *Journal of Hospitality & Leisure Marketing*, 17(1-2), 121-138.
- Keskinkılıç, M., Ağca, Y. ve Karaman, E. (2016). İnternet ve Bilgi Sistemleri Kullanımının Turizm Dağıtım Kanallarına Etkisi Üzerine Bir Uygulama. *İşletme Araştırmaları Dergisi*, 8(4), 445-472. doi:10.20491/isarder.2016.227
- Kim, Y. A. ve Srivastava, J. (2007). Impact of social influence in e-commerce decision making. *ICEC '07: Proceedings of the ninth international conference on Electronic commerce*, (s. 293-302).
- Kuhzady, S. ve Ghasemi, V. (2019). Factors Influencing Customers' Satisfaction and Dissatisfaction with Hotels: A Text-Mining Approach. *Tourism Analysis*, 24(1), 69-79.
- Kumar, S. ve Aulia, A. (2021). *The Influence of Hotel Review Towards E-Trust and Its Implication on Hotel Booking Intention*. (S. Kumar, Dü.) Indonesia.
- Lau, K.-N., Lee, K.-H. ve Ho, Y. (2015). Text Mining for the Hotel Industry. *Cornell Hotel and Restaurant Administration Quarterly*, 46(3), 344-362.
- Melián-González, S., Bulchand-Gidumal, J. ve López-Valcárcel, B. G. (2013). Online Customer Reviews of Hotels: As Participation Increases, Better Evaluation Is Obtained. *Cornell Hospitality Quarterly*, 53(3), 274-283.
- Miner, G. D., Elder, J., Fast, A., Hill, T., Nisbet, R. ve Delen, D. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data*. Oxford, UK: Academic Press.
- Nicholas, C. K.-W. ve Lee, A. S. (2017). Voice of Customers: Text Analysis of Hotel Customer Reviews. *CBDR 2017* (s. 104-111). Osaka, Japan: Association for Computing Machinery.
- O'Fallon, M. J. ve Rutherford, D. G. (2011). *Hotel Management and Operations*. New Jersey, US: John Wiley & Sons.
- Öğüt, H. ve Taş, B. K. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *The Service Industries Journal*, 32(2), 197-214.
- Palese, B. ve Usai, A. (2018). The relative importance of service quality dimensions in E-commerce experiences. *International Journal of Information Management*, 40, 132-140.
- Pradana, A. W. ve Hayaty, M. (2019). The effect of stemming and removal of stopwords on the accuracy of sentiment analysis on indonesian-language texts. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control Journal*, 4(4), 375-380.
- Prajapati, V. (2021, 01 22). *Why Do People Shop Online?* <https://www.techprevue.com/reasons-prefer-online-shopping/> adresinden alındı

- Rai, A. (2019, 06 01). *What is Text Mining: Techniques and Applications*. upgrad.com: <https://www.upgrad.com/blog/what-is-text-mining-techniques-and-applications/> adresinden alındı
- Schuckert, M. (2015). Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *Journal of Travel & Tourism Marketing*, 32(5), 608-621.
- Shen, Z., Yang, X., Liu, C. ve Li, J. (2021). Assessment of Indoor Environmental Quality in Budget Hotels Using Text-Mining Method: Case Study of Top Five Brands in China. *Sustainability*, 13(8).
- Singh, P. K., Sachdeva, A., Mahajan, D., Pande, N. ve Sharma, A. (2014). An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites. *5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*. Noida, India : IEEE.
- Somprasertsri, G. ve Lalitrojwong, P. (2010). Mining Feature-Opinion in Online Customer Reviews for Opinion Summarization. *Journal of Universal Computer Science*, 16(6), 938-955.
- Trenz, M. ve Berger, B. (2013). Analyzing Online Customer Reviews - An Interdisciplinary Literature Review And Research Agenda. *ECIS 2013 Completed Research*, (s. 83).
- Truyens, M. ve Eecke, P. V. (2014). Legal aspects of text mining. *Comput. Law Secur. Rev.*, 2182-2186.
- Tyagi, N. (2021, 05 10). *Top 7 Text Mining Techniques*. analyticssteps.com: <https://www.analyticssteps.com/blogs/top-7-text-mining-techniques> adresinden alındı
- Wei, P.-S. ve Lu, H.-P. (2013). An examination of the celebrity endorsements and online customer reviews influence female consumers' shopping behavior. *Computers in Human Behavior*, 29(1), 193-201.
- Weiss, S. M., Indurkha, N. ve Zhang, T. (2015). *Fundamentals of Predictive Text Mining* (2 b.). New York, US: Springer.
- Xu, X. ve 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.
- Zhan, J., Loh, H. T. ve Liu, Y. (2009). Gather customer concerns from online product reviews – A text summarization approach. *Expert Systems with Applications*, 36(2), 2107-2115.

## SUMMARY

This research focuses on text mining analysis in the context of the tourism sector, which is a common practice in various e-commerce industries. Specifically, the study aims to gain insights from customer reviews by employing methods such as sentiment analysis, data visualization, n-grams, document classification, and word frequency analysis. Customer comments play a crucial role in the tourism sector, as they heavily influence potential customers' decisions when selecting accommodations. Therefore, the researchers conducted text mining on hotel reviews in Turkey to identify patterns and trends based on the guests' preferences and dislikes.

The primary objective of this research is to investigate the aspects that guests pay attention to when expressing their satisfaction or dissatisfaction with hotel facilities. While individual comments can provide some understanding of guests' preferences and dislikes, this study aims to generalize these findings using a substantial dataset of 60,252 comments. Consequently, hotels that are either favored or disliked can gain valuable insights into the specific aspects that guests appreciate or find lacking. Additionally, the research seeks to uncover patterns in the comments made by guests in Turkey. Furthermore, the study provides

information regarding the distribution of guest ratings across different types of facilities, geographical regions, provinces, and cities.

The initial part of the study offers comprehensive explanations of text mining techniques to enhance readers' understanding of the research methodology. Subsequently, the research methodology employed in this study is presented. In the findings section, the guests' ratings are presented by categorizing them according to the type of facility, geographical region, province, and city. Furthermore, the study analyzes the most frequently occurring words and binary words associated with each rating score. By using these words, a list of the most commonly repeated words within each score group (ranging from 1 to 5) is compiled. Consequently, the study aims to identify the most criticized or highly praised features of hotel facilities.

This research contributes to the field of text mining in the tourism sector by employing advanced analytical techniques to extract valuable insights from a large corpus of hotel reviews. The findings enable hotel businesses to gain a comprehensive understanding of guests' preferences and dislikes. By identifying the key features that impact guest satisfaction, hotels can make informed decisions to enhance their offerings and improve customer experiences. Furthermore, the study provides valuable information about guest ratings across different dimensions, facilitating comparisons between different types of facilities, geographical regions, provinces, and cities.

In conclusion, this research utilizes text mining techniques to analyze hotel reviews in Turkey, aiming to identify the aspects that guests pay attention to when expressing their satisfaction or dissatisfaction with hotel facilities. By employing advanced text mining methods, the study provides insights into the most commonly repeated words associated with different rating scores. The findings contribute to the overall understanding of guest preferences and offer actionable information for hotels to improve their services. Additionally, the research provides valuable data on guest ratings across various dimensions, facilitating comparisons and providing a comprehensive overview of the tourism sector in Turkey.