

## Otellere İlişkin Çevrimiçi Geribildirimlerin Makine Öğrenmesi Yöntemleriyle Duygu Analizi\*

(Sentiment Analysis of Online Feedbacks on Hotels via Machine Learning Methods)

Murat Fatih TUNA <sup>a</sup> Oğuz KAYNAR <sup>b</sup> M. Şükrü AKDOĞAN <sup>c</sup>

<sup>a</sup> Sivas Cumhuriyet Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Sivas, Türkiye, [mftuna@cumhuriyet.edu.tr](mailto:mftuna@cumhuriyet.edu.tr)

<sup>a</sup> Sivas Cumhuriyet Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Sivas, Türkiye, [okaynar@cumhuriyet.edu.tr](mailto:okaynar@cumhuriyet.edu.tr)

<sup>b</sup> Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi, Kayseri, Türkiye, [akdogans@erciyes.edu.tr](mailto:akdogans@erciyes.edu.tr)

MAKALE BİLGİSİ	ÖZET
<b>Anahtar Kelimeler:</b> Müşteri Geri Bildirimleri Çevrimiçi Rezervasyon Siteleri Otel İşletmeleri Duygu Analizi Makine Öğrenmesi  Gönderilme Tarihi 22 Nisan 2021 Revizyon Tarihi 8 Temmuz 2021 Kabul Tarihi 16 Temmuz 2021  <b>Makale Kategorisi:</b> Araştırma Makalesi	<b>Amaç</b> – Günümüzde tüketiciyi hızlı anlamının yollarından biri, onların ürün ya da hizmetler ile ilgili geri bildirimlerini hızlı ve doğru biçimde analiz etmektir. Bu anlamda geribildirimlerin içindeki duygunun bilgisayar temelli tekniklerle anlaşılması izlenecek yollardan biridir. Bu çalışmada, otel müşterilerinin geri bildirimlerinde gizli olan duyguların otelden aldıkları hizmete dair derecelmeleriyle örtüşme düzeyi araştırılmıştır. <b>Yöntem</b> – Araştırmada kullanılan yöntem makine öğrenmesi temelli duygu analizidir. Kullanılan veri seti, çevrimiçi bir rezervasyon sitesinden web kazıma yöntemiyle çekilen ve Antalya’da yer alan 164 otele ilişkin müşteri yorumlarından oluşmaktadır. Veri setindeki yorumların müşteriler tarafından yorumlara eklenen beğeni dereceleri ile uyumluluğu; Lojistik Regresyon, Rastgele Orman (RF), Karar Ağacı (CART), K-En Yakın Komşu (KNN), Destek Vektör Makinaları (SVM), Doğrusal Diskriminant Analizi (LDA) ve Naïve Bayes (NB) olmak üzere yedi farklı makine öğrenme algoritmasının kullanıldığı ikili duygu sınıflandırma yöntemi ile test edilmiştir. Test sonucunda elde edilen karmaşıklık matrislerinin oluşturulmasıyla model başarımları ölçümlenmiştir. Duygu sınıflandırmasının başarıyla yapılmasının yanı sıra elde edilen modellerin performansları karşılaştırılarak görselleştirilmiştir. <b>Bulgular</b> – Denetimli duygu sınıflandırmada kullanılan algoritmaların ortalama sınıflandırma başarımları %81,30 olarak hesaplanırken, bunların içerisinde en başarılı sonuç üreten algoritmanın Lojistik Regresyon (%87,99) olduğu anlaşılmıştır. Kullanılan yöntemler başarımları sırası ile Lojistik Regresyon (%87,99), Destek Vektör Makinaları (%86,84), Doğrusal Diskriminant Analizi (%86,24), Naïve Bayes (%82,66), Rastgele Orman (%82,00), Karar Ağaçları (%76,92) ve K-En Yakın Komşu (%63,91) şeklinde sıralanmıştır. <b>Tartışma</b> – Çalışmada uygulanan modelin, literatürle paralel olarak, otel işletmelerinin yöneticilerinin hızlı, tutarlı ve maliyet etkin pazarlama kararları almalarında bir araç olarak kullanılabileceği ve otel işletmeleri adına katma değer üretilabileceği öne sürülmüştür. Çalışmanın gerek konaklama işletmelerinin paydaşlarına gerekse bu konuda çalışma yapacak araştırmacılara destek sağlayacağı düşünülmektedir. Dahası, bu çalışmada Türkiye’nin ya da daha küresel turizm memnuniyeti yerine Antalya ilindeki otel hizmetlerinden duyulan memnuniyet incelenmiştir. Bu çalışma, Türkiye’deki gerek farklı minimal lokasyonlar gerekse daha büyük bölgesel incelemeler için benzer çalışmalar ile genişletilebilecektir. Daha ileride yapılacak çalışmalarda farklı dil kütüphaneleri kullanarak da çok dilli uygulamaların gerçekleştirilebilmesi mümkün olacaktır. Ayrıca metinsel ifadelerin konaklama işletmeleri açısından başarılı ve hızlı bir şekilde çözümlenebileceğinin kanıtlanmasının yanı sıra maliyet, zaman ve işgücü kazancının oluşacağı ön görülmüştür.

ARTICLE INFO	ABSTRACT
<b>Keywords:</b> Customer Feedbacks Online Booking Websites Hotel Businesses Sentiment Analysis Machine Learning	<b>Purpose</b> – Today, one of the ways to quickly understand the consumer is to analyse their feedbacks about the products or services of the business quickly and accurately. In this sense, understanding the sentiment in the feedback with computer-based techniques is one of the ways to be followed. Overlapping level between the sentiments hidden in the feedbacks of hotel customers and their ratings regarding the service they received from the hotel have been examined in this study. <b>Design/methodology/approach</b> – The method used in the research is machine learning-based sentiment analysis. The data set used consists of customer comments on 164 hotels in Antalya, extracted from an

\* Bu çalışma “Çevrimiçi Yorum ve Şikâyetlerin Otel İşletmeleri Üzerinden Duygu Analizi ile İncelenmesi” başlıklı doktora tezinden yararlanılarak hazırlanmıştır.

### Önerilen Atf/ Suggested Citation

Tuna, M.F., Kaynar, O., Akdoğan, M.Ş. (2021). Otellere İlişkin Çevrimiçi Geribildirimlerin Makine Öğrenmesi Yöntemleriyle Duygu Analizi, *İşletme Araştırmaları Dergisi*, 13 (3), 2232-2241.

Received 22 April 2021  
 Revised 8 July 2021  
 Accepted 16 July 2021

**Article Classification:**  
 Research Article

online booking site via web scraping method. Compatibility of the comments in the data set with the ratings added to the comments by the customers was tested with a binary sentiment classification via seven different machine learning algorithms including Logistic Regression (LR), Random Forest (RF), CART Decision Tree (CART), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), Naïve Bayes (NB).

**Findings** – While the average classification success of the algorithms used in supervised sentiment classification was calculated as 81.3%, it was understood that the algorithm produced the most successful results among them was Logistic Regression (87.9%). The methods used in this study were lined up as LR (%87,99), SVM (%86,84), LDA (%86,24), NB (%82,66), RF (%82,00), CART (%76,92) and KNN (%63,91).

**Discussion** – It has been suggested that the model applied in the study, in parallel with the literature, can be used as a tool for hotel managers to make fast, consistent and cost-effective marketing decisions, and that added value can be produced for hotel businesses. It is thought that the study will provide support to both the stakeholders of the accommodation businesses and the researchers who will work on this subject. Moreover, satisfaction with hotel services located in Antalya province instead of international or global tourism satisfaction were investigated in the study. This study can be extended with similar studies for both different minimal locations and larger regions in Turkey. In future studies, it will be possible to realize multilingual applications by using different language libraries. In addition, it is foreseen that the textual expressions can be successfully and quickly resolved in terms of accommodation businesses, as well as cost, time and labour savings.

## 1. Giriş

Müşterilerle karşılıklı fayda temeline dayanan ilişkilerin kurulması ve sürdürülebilmesi adına, müşterilerden gelen geri bildirimlerin hızlı ve etkili biçimde değerlendirilerek müşteriye kısa bir süre içinde dönüş yapılması gerekmektedir. Müşteri sadakatini sağlayabilmenin ilk ve en önemli koşulu, müşterilerin anlaşılması ve onlardan gelen anlık geri bildirimlerin teknolojik imkânların el verdiği en hızlı biçimde yanıtlanmasıdır (Omisakin vd., 2020:3). Geri bildirim hızının artırılması, müşterilerden elde edilecek çeşitli verilerin doğru ve istikrarlı biçimde analiz edilmesi ve bu verilerden elde edilecek bilgilerin doğru biçimde kullanılmasını gerektirmektedir. Bu süreçte verimliliğin sağlanması ise bilişim teknolojisinin etkin biçimde kullanılmasını ve standardının sürekli yükseltilmesini zorunlu hale getirecektir.

Tüketicilerden geri bildirim almanın eskiye nazaran kolaylaştığı bir gerçektir. Nitekim tüketiciler günümüzde yaygın biçimde “dijital iletişim mecraları” olarak adlandırılacak çevrimiçi birçok platform vasıtasıyla herhangi bir konudaki fikrini özgürce paylaşabilmektedir. İnsanların satın aldıkları ürün ve hizmetler hakkındaki olumlu olumsuz kişisel düşünceleri artık sadece firma ile tüketici arasında kalmamakta, sosyal medya üzerinden de paylaşılmaktadır (Güran vd., 2014:87). Paylaşılan fikirler gündemdeki bir olay ya da siyasi bir gelişme ile alakalı olabileceği gibi bir markanın ürün ya da hizmeti ile ilgili de olabilmektedir. Bunun dışında tüketiciler, bahsedilen dijital mecraları (sosyal medya, blog, tartışma forumu) kullanan ve etkileşimde bulunabildikleri diğer bireylerle yoğun bir fikir paylaşımına girişmektedirler. İşletmeler bu etkileşimi ve hızını bildiklerinden ötürü, bahsedilen mecralarda yer alacak tek bir paylaşımın bile es geçilmeden önem taşıdığını algılamıştır. Buna bağlı olarak tüketicilerin onların görüşlerini değerlendirmek ve hatta görüşün de altında yatan gizli düşüncenin ne olduğunu bilmek zorunlu hâle gelmiştir. Zaltman ve Zaltman (2008:6) tarafından da ifade edildiği üzere artık “tüketiciye gelen yüzeysel düşüncelerin altına kurcalamak” gerekmektedir. Burada ortaya çıkan tek sorun, yığın veri kümeleri içinde tıpkı samanlıkta iğne arar gibi büyük etkilere varabilen küçük etkileşim kıvılcımları aramak değildir. Yorumların nicel olmayışları, klasik analitik yöntemlerin çaresiz kalmasına da sebep olmaktadır. Çeşitli sektörden birçok işletme, tüketicilerinin düşüncelerini derinlemesine araştırmak için bilişimin sunduğu avantajlardan ve fırsatlardan yararlanmaya odaklanmıştır. Dahası bu çerçevede teknolojinin, özellikle de bilişimin sunduğu imkânlardan faydalanma eğilimi, birçok işletme için kaçınılmaz hale gelmiştir. Bu çalışmanın ikinci bölümünde çevrimiçi rezervasyon siteleri ve duygu analizi kavramları anlatılırken, üçüncü bölümde duygu analizinde gerekli olan adımlar ve kullanılan yöntemlerden bahsedilmiştir. Dördüncü bölümde ise, analiz sonuçlarına yer verilmiştir. Son bölüm olan sonuç ve tartışma bölümünde, elde edilen bulgular literatürdeki çalışmalar ile kıyaslanmıştır. Aynı zamanda başka çalışmacılar için öneriler yer almaktadır.

## 2. Teorik Çerçeve

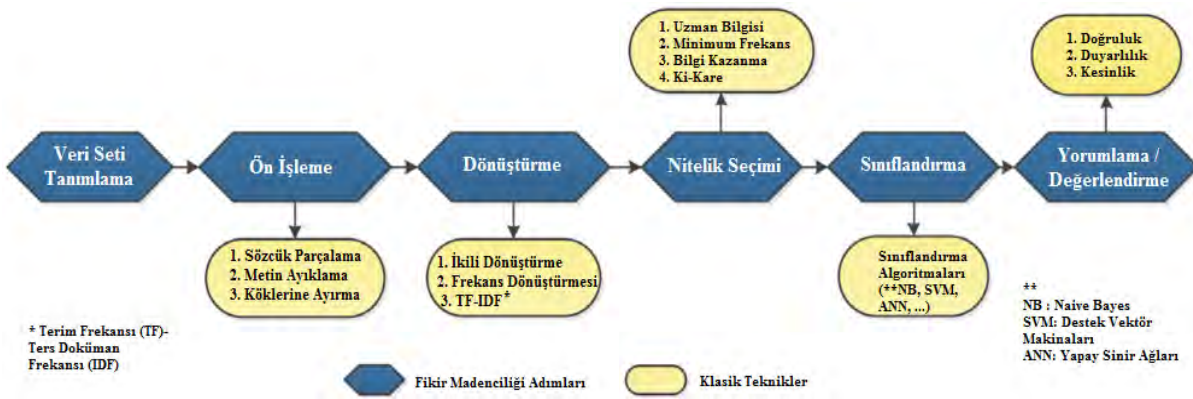
### 2.1. Çevrimiçi Rezervasyon Siteleri

Rezervasyon kolaylığı, tüm dünya çapında faaliyet gösteren rezervasyon sitelerinin faaliyete geçmesinden sonra otel işletmelerinin tercih edilmesinde daha da aranılır bir kriter hâline gelmiştir (Shuai vd., 2018:1171).

Nitekim rezervasyon siteleri, arama kriterlerinin daraltılabilmesine ve konaklama alternatiflerinin etkin biçimde karşılaştırılabilmesine olanak tanımıştır. Otel işletmeleri de bu sitelerin faaliyetlerinden iki yönde ciddi fayda sağlamışlardır. Bunlardan biri, otel işletmesi ile ilgili konaklama imkânlarını müşterilerin sorun yaşamadan öğrenebilmesidir. Bu aynı zamanda otel işletmesinin müşterilerle çevrimiçi iletişim kurması zorunlu bir birim oluşturma gereksinimini ortadan kaldırmaktadır. İkincisi ise otel müşterilerinin konaklama deneyimlerine ilişkin değerlendirmelerini doğrudan elde edilebilmesidir. Bu yönüyle rezervasyon siteleri, otel işletmeleri için bir müşteri geri bildirim platformu olarak da işlev görmektedir.

## 2.2. Duygu Analizi

Dijital mecralarda yer alan yorumların sayısı olarak çok fazla olması ve sürekli yenilerinin ekleniyor olması, aynı anda birden çok yorumu inceleyip özetleyebilecek metotlara yönelim sağlanmasına neden olmuştur. Bu metotlardan duygu analizi, metin madenciliğinin alt bileşenlerinden biri olup, bir metnin tamamı ya da bir kısmı içindeki duygunun bilgisayar temelli tekniklerle anlaşılması esasına dayanmaktadır (Han vd., 2011:512). Duygu analizinde izlenen temel adımlar şu şekilde özetlenebilir:



Şekil 1. Duygu Analizinde Kullanılan Adım ve Tekniklere Genel Bakış (Moraes vd., 2013: 623)

Duygu analizleri metinsel veriler üzerinde aşağıdaki amaçları yerine getirmek üzere kullanılmaktadır (Kaynar vd., 2016:235):

- Duygu ve görüşlerin sınıflandırılması
- Öznelliğin sınıflandırılması
- Alaycılık ve ironinin tespiti edilmesi
- Metin türü ve yazarının tespiti
- SPAM ve sahte yorumların tespiti
- Görüş ve yorumların ortaya çıkarılması
- Görüş ve yorumların özetlenmesi

Duygu analizi ile yapılan çalışmalar incelendiğinde, çeşitli dillere özgü karakteristik dil kütüphaneleri (Türkçe, Arapça, Çince, vs.) geliştirilene kadar, duygu analizi ile ilgili yapılan çalışmaların bir süre İngilizcedeki yorumlardan elde edilen veri setleri üzerinde yapıldığı göze çarpmaktadır. Literatürde dikkat çeken başka bir nokta, duygu analizi çalışmalarında hem ikili (pozitif ve negatif), hem de üçlü (pozitif, negatif ve nötr) duygu sınıflandırmalarının bulunmasıdır.

Çevrimiçi rezervasyon sitelerinden edinilen tüketici yorumlarının içerdiği duyguların, yorumlara eklenen derecelendirmelerle uyumluluğunun araştırıldığı çalışmalar bulunmaktadır. Bu çalışmalardan Wang vd. (2010) tarafından yapılan çalışmada, çevrimiçi rezervasyon sitesinde yer alan otellere ilişkin müşteri derecelemeleri, otel yorumlarına uygulanan duygu analizlerinin sonuçlarını doğrulamak için kullanılmışlardır. Kullandıkları LRR (Gizli Derecelendirme Regresyonu) modelinin en iyi sonucu (%89,60) ürettiğini ifade etmişlerdir. Raut ve Londhe (2014) tarafından yapılan çalışmada da TripAdvisor.com sitesinden elde edilen 1000 olumlu ve 1000 olumsuz otel yorum üzerinde duygu analizi gerçekleştirilmiştir. Çalışmada eğitim ve test verileri yarı yarıya ayrılmış, analizlerde ön etiketleme yapılmış hazır bir veri seti kullanılmıştır. Çalışmada

NB (Naïve Bayes), SVM (Destek Vektör Makineleri), CART (Sınıflandırma ve Regresyon Ağacı) algoritmaları ile SentiWordNet yönteminin performansları karşılaştırılmıştır. Araştırmada elde edilen doğruluk değerleri sırasıyla %88,00; %83,50; %78,40 ve %87,60 şeklinde olmuştur.

2015-2020 yıllarında yapılan çalışmalarda, bu çalışmada yer alan yöntemlerin kullanıldığı örnekler görmek mümkündür. Oğul ve Ercan (2016), *Booking.com* ve *TripAdvisor.com* çevrimiçi rezervasyon sitelerinden toplamış oldukları 331 adet veri üzerinde makine öğrenmesi yöntemiyle duygu analizi gerçekleştirmişlerdir. RF (Rastgele Orman), NB ve SVM (Destek Vektör Makinaları) algoritmalarının sonuçlarının karşılaştırıldığı çalışmada en yüksek performansı veren algoritmanın RF olduğu (%89,00) tespit edilmiştir. Perikos ve Hatzilygeroudis (2017), otel hizmeti alan ve çevrimiçi bir rezervasyon sitesinden elde ettikleri oluşan veri setine makine öğrenmesi yöntemiyle duygu analizi uygulamışlardır. NB, ME (maksimum entropi), SVM ve Ensemble (Topluluk) algoritmalarının kullanıldığı ve ortalama başarının %82,25 olduğu ifade edilen çalışmada, en iyi sonucu %86,00 başarı performansına sahip olan Ensemble algoritması yardımıyla elde ettiklerini duyurmuşlardır. Yin vd. (2017) Hiyerarşik Mimari (HM) ve SVM algoritmalarını kullanarak *TripAdvisor.com* ve *BeerAdvocate.com* sitelerinden elde ettikleri verilere makine öğrenmesi yöntemiyle duygu analizi uygulamışlardır. Sınıflandırmaların ortalama model başarı değerlerini HM ve SVM algoritmaları için sırasıyla %61,05 ve %84,59 şeklinde olmuştur. Al-Smadi vd., (2018), otel müşterilerine ilişkin çevrimiçi Arapça yorumları içeren hazır veri setine makine öğrenmesi yöntemiyle duygu analizi uygulamışlardır. SVM, KNN (K-En Yakın Komşuluk), Karar Ağacı, Bayes ağları ve NB algoritmalarıyla yapılan denetimli sınıflandırma testleri sonucunda %93,40 tahmin performansı ile SVM en başarılı sonuçları üreten sınıflandırıcı olarak seçilmiştir. Shuai vd. (2018) çalışmalarında Doc2Vec adı verilen yöntem yardımıyla web ortamından çektikleri 11.600 adet yorum içinden seçilen 6000 Çince otel yorumuna makine öğrenmesi temelli duygu analizi uygulamışlardır. SVM, LR (Lojistik Regresyon) ve NB algoritmalarının kullanıldığı çalışmada elde edilen doğruluk değerleri sırasıyla %84,30; %81,00 ve %74,40 şeklinde hesaplanmıştır.

2019-2020 yıllarında yapılan çalışmalar incelendiğinde; Guerreiro ve Rita (2020) çalışmalarında 661 farklı şehirdeki 47263 farklı lokantaya yapılmış olan 1.112.708 adet yorum üzerine duygu analizi uygulamışlardır. LR, CART ve RF algoritmalarının kullanıldığı çalışmada elde edilen doğruluk değerleri sırasıyla %66,86; %66,05 ve %66,83 şeklinde hesaplanmıştır. Sridharan vd. (2020) çalışmalarında duygu analizinin çözümlenmesi için kullanılan RF algoritmasının performansındaki gelişmeyi incelemişlerdir. Buldukları sonuç, veri setindeki eleman sayısının artması sonucunda RF performansının artacağı şeklindedir. Analizde kullanılan ağaç sayısının artması ile temel olarak aldıkları %85,60'lık başarı oranı %90,69'a çıkmıştır.

2021 yılında yayınlanan ve özellikle bu çalışma ile alakalı olan benzer çalışmalarda da hem kullanılan yöntemler hem de sonuçlar açısından benzerlikler gözlenmektedir. Allam vd. (2021), atılan 12548 Twitter gönderisi içinden rasgele seçilen 3000 yorum üzerine makine öğrenmesi temelli duygu analizi uygulanmıştır. Çalışmada kullanılan dil Arapça olarak seçilmiştir. SVM, LR, NB ve CNB (Complementary Naive Bayes) algoritmalarıyla yapılan analiz sonucunda ulaşılan sonuçlar sırasıyla %85,55, %65,89, %78,92 ve %83,47 olarak hesaplanmıştır. Budhi vd. (2021), duygu analizi ile berber, kafe, bar restoran vb. yerel işletmeler için yapılan Yelp yorumlarını incelemişlerdir. Çalışmalarında kullandıkları metin ifadeleri sonradan aynı kişilerce verilen 5'li likert tipi ifadeler ile doğrulayan yazarlar tarafından, bu şekilde 4,1 milyon yorum analiz edilmiştir. Birçok farklı sınıflandırma algoritması kullanılan çalışmada, bu çalışma ile alakalı olarak RF, SVM ve LR algoritmaları ortak noktadır. Elde edilen sonuçlar sırasıyla %87,95, %90,57 ve %91,16 olarak hesaplanmıştır. Kinyua vd. (2021), çalışmalarında dönemin ABD başkanı Trump'ın Twitter sayfasında paylaştığı gönderileri Dow Jones (INDU) ve S&P 500 (SPX) borsalarındaki dalgalanmalar ile kıyaslamışlardır. Çalışmada RF, CART ve LR algoritmaları toplam 11343 yorumu analiz etmek için kullanılmıştır. Twitter gönderilerinin INDU borsasındaki dalgalanma ile uygunluğu sırasıyla %98,00; %94,00 ve %99,00 ve SPX borsasındaki dalgalanma ile uygunluğu sırasıyla %92,00; %90,00 ve %91,00 olarak hesaplanmıştır. Luo (2021), farklı boyutları olan üç farklı veri seti duygu analizi çalışması yapmıştır. Analiz aşamasında SVM, NB ve LR algoritmalarının kullanıldığı çalışmada, veri setlerindeki başarı ortalaması sırasıyla %76,00; %26,00 ve %71,00 hesaplanmıştır.

Bu çalışmada, otel müşterilerinin geri bildirimlerinde gizli olan duyguların, otelden aldıkları hizmete dair derecelenmeleriyle örtüşme düzeyi araştırılmıştır. Bu amaçla, *TripAdvisor.com* sitesindeki Antalya ilinde yer alan otel işletmelerine yönelik çekilen yorumların içerdiği duyguların, müşteriler tarafından yorumlara eklenen beğeni dereceleri ile uyumluluğu test edilmiştir. Çalışmada yedi farklı makine öğrenme algoritmasının kullanıldığı ikili duygu sınıflandırma yöntemi uygulanmış ve elde edilen sonuçların başarı performansları karşılaştırılmıştır.

### 3. Yöntem

#### 3.1. Verilerin Hazırlanması

Çalışmanın verileri, 1 Kasım 2017 ve 31 Mayıs 2018 tarihleri arasında *TripAdvisor.com* çevrimiçi rezervasyon sitesinden çekilen ve Antalya ili sınırları içinde yer alan toplam 164 adet otelin müşterilerine ait yorumlardan oluşmaktadır. 21.190 yorum arasından seçilen 3570 pozitif ve 3547 negatif olmak üzere toplam 7117 müşteri yorumu veri setine dahil edilmiştir. Verilerin çevrimiçi rezervasyon sitesinden çekilmesinde açık kaynak kodlu *Python* yazılımı ile yazılıma ait *Spyder* kütüphanesi ve *Anaconda* platformlarından faydalanılmıştır.

#### 3.2. Verilerin Ön İşlenmesi

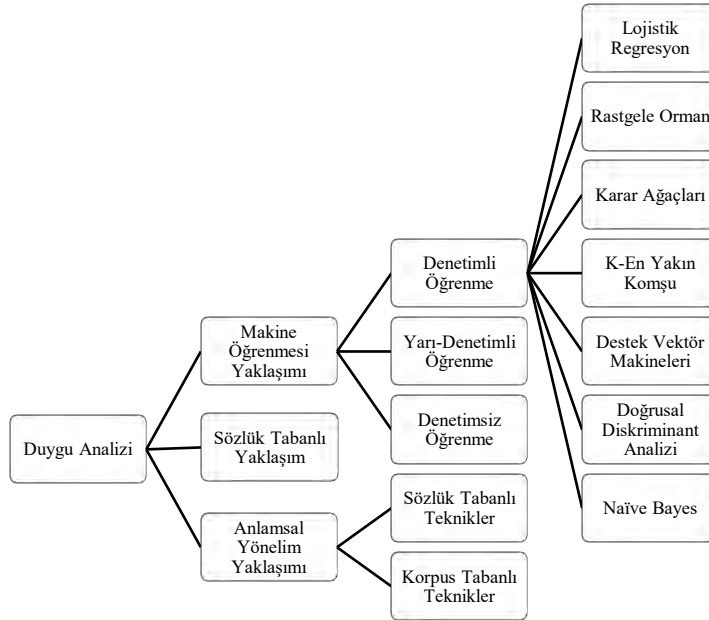
Araştırmadaki ham veriler analiz öncesinde ilk olarak etkisiz kelimelerden (bağlaç, vb.) ve noktalama işaretlerinden ayrıştırılarak yalın metin derlemi elde edilmiştir. Ardından kelimeler morfolojik köklerine ayrılmış yani 'stem' edilmiştir. *NZemberek* adı verilen Türkçe açık kaynak kodlu doğal dil işleme kütüphanesinden yararlanarak tamamlanan morfolojik kök çıkarımı sürecinin ardından n-gram yöntemi uygulanarak kelime çantası (bag-of-words) yapısı elde edilmiştir.

#### 3.3. Metin İfadelerinin Sayısallaştırılması

Ön etiketleme işlemi gerçekleştirilen yorumlardan 3 puan ve altındakiler negatif ve 4 puan ve üzerinde olanlar pozitif olarak etiketlenmiştir. Ardından verilerin vektör uzay modellemesine geçilmiştir. Metin temelli verinin sayısal veriye dönüştürülmesini sağlayan bu aşamada, yaygın bilinen Terim Frekansı-Ters Doküman Frekansı (TF-IDF) yöntemi kullanılmıştır. TF-IDF, bir kelimenin bir metindeki geçme sıklığını ve dolayısıyla metindeki göreceli önemini sayısal olarak ifade etmeyi sağlayan frekans değerine karşılık gelmektedir (Leskovec vd., 2014:14). Metin verilerinin sayısallaştırılması ile nicel hâle getirilen verilerin makine öğrenmesi ile analizine geçilmiştir.

#### 3.4. Deneylede Kullanılan Duygu Sınıflandırma Yöntemi

Araştırmadaki tüm deneyler *Python* yazılımı kullanılarak gerçekleştirilmiştir. Deneylede duygu analizi yaklaşımlarından makine öğrenmesi yaklaşımı tercih edilmiştir. Duygu analizinde kullanılan diğer yaklaşımlar ve bu araştırmada kullanılan makine öğrenmesi yaklaşımının detayları Şekil 2'de görülmektedir.



Şekil 2. Araştırmada Kullanılan Duygu Analizi Yaklaşımı ve Algoritmalar

Şekil 2'de görüldüğü üzere, araştırmada denetimli bir makine öğrenme yaklaşımı kullanılmıştır. Araştırmanın denetimli olmasının sebebi, yorumlardan elde edilen duyguların doğrulanması için müşteriler tarafından yorumlara eklenen ön etiketlenmiş derecelemelerin kullanılmasıdır. Denetimli makine öğrenmesinde, Wainer (2016:8) tarafından ikili sınıflandırmalar için önerilen 14 makine öğrenme algoritması arasından seçilen yedi sınıflandırma algoritması (Lojistik Regresyon, Random Forest, CART Sınıflandırıcı, K-En Yakın Komşu, Destek Vektör Makineleri, Doğrusal Diskriminant Analiz Sınıflandırıcısı ve Naïve Bayes) kullanılmıştır

#### 4. Bulgular

Sınıflama performansını karşılaştırmak için kullanılan metrikler; doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F ölçümü değerleridir. Bu metrikler ve matematiksel formülleri aşağıda verilmiştir:

→**Doğruluk (Accuracy):** Sınıflandırma işleminde en çok kullanılan ölçümdür. Doğru olarak sınıflandırılmış örneklerin toplam örnek sayısına oranıdır ve aşağıdaki formülle ifade edilmektedir:

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

→**Kesinlik (Precision):** Sınıflandırıcı sonucunun kesinlik derecesini verir. Pozitif olarak etiketlenen örneklerin sayısının pozitif olarak sınıflandırılmış toplam örneklere oranıdır ve aşağıdaki gibi formüle edilmektedir:

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (2)$$

→**Duyarlılık (Recall):** Gözlem ve tahmin ölçümlerinin her ikisinde de pozitif olarak etiketlenmiş örneklerin, gerçekten pozitif olan örneklerin toplam sayısına oranıdır ve aşağıdaki gibi formüle edilmektedir:

$$\text{Duyarlılık} = \frac{TP}{TP+FN} \quad (3)$$

→**F-Ölçümü:** F Ölçümü, kesinlik ve duyarlılık metrikleri kullanılarak hesaplanmaktadır. Sistemin, kesinlik veya duyarlılık yönüne doğru optimize edilmesinde kullanılmaktadır ve aşağıdaki gibi formüle edilmektedir:

$$F = \frac{2 \cdot \text{Duyarlılık} \cdot \text{Kesinlik}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4)$$

Bu çalışmada gerçekleştirilen makine öğrenmesinin sonucunda elde edilen her bir yonteme ait karmaşıklık matrisleri Tablo 1’de gösterilmektedir.

**Tablo 1.** Her Bir Yöntem İçin Elde Edilen Karmaşıklık Matrisleri

		Eğitim (%90)		Test (%10)		Tümü		
		Gözlenen		Gözlenen		Gözlenen		
		Poz.	Neg.	Poz.	Neg.	Poz.	Neg.	
LR	Tahmin	Poz.	2850	362	317	40	3167	402
		Neg.	416	2777	46	309	462	3085
RF	Tahmin	Poz.	2489	723	277	80	2766	803
		Neg.	370	2822	41	314	411	3136
CART	Tahmin	Poz.	2495	717	277	80	2772	797
		Neg.	780	2412	87	268	867	2680
KNN	Tahmin	Poz.	1838	1374	204	153	2042	1527
		Neg.	701	2491	78	277	779	2768
SVM	Tahmin	Poz.	2813	400	313	44	3125	444
		Neg.	453	2740	50	304	503	3044
LDA	Tahmin	Poz.	2812	401	312	45	3124	445
		Neg.	497	2696	55	300	552	2995
NB	Tahmin	Poz.	2894	318	322	35	3216	353
		Neg.	896	2296	100	255	996	2551

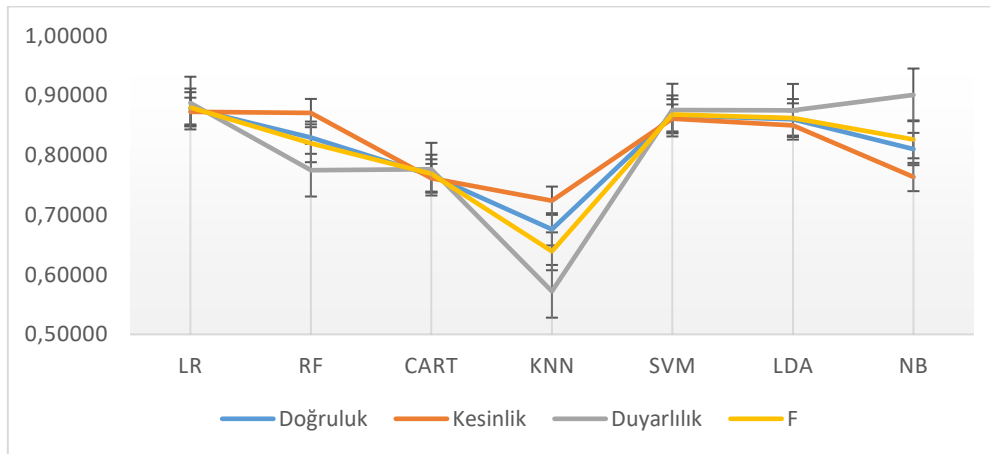
TP ve TN değerlerinin sırasıyla veri setine dâhil edilen pozitif ve negatif olan yorum sayılarına yakınlaşırken, FN ve FP değerlerinin gözlem sayısına göre daha az olduğu tablodan anlaşılmaktadır. Dolayısıyla sınıflandırıcının tahmin yeteneğinin gerçek değerler ile uyumunun arttığı gözlemlenmektedir. Yukarıdaki

açıklanan sonuç değerleri ve bu sonuç değerlerinin değerlendirme kriterleri olduğu yöntemler aşağıda Tablo 2’de verilmektedir:

**Tablo 2.** Duygu Analizi Sonucunda Elde Edilen Performans Değerleri

	YÖNTEM	Doğruluk	Kesinlik	Duyarlılık	F-Ölçümü
EĞİTİM	LR	0,87853	0,87263	0,88730	0,87990
	RF	0,82933	0,87058	0,77491	0,81996
	CART	0,76624	0,76183	0,77677	0,76923
	KNN	0,67598	0,72391	0,57223	0,63919
	SVM	0,86684	0,86130	0,87551	0,86834
	LDA	0,85982	0,84980	0,87519	0,86231
	NB	0,81043	0,76359	0,90100	0,82662
TEST	LR	0,87921	0,87328	0,88796	0,88056
	RF	0,83006	0,87107	0,77591	0,82074
	CART	0,76545	0,76099	0,77591	0,76838
	KNN	0,67556	0,72340	0,57143	0,63850
	SVM	0,86779	0,86226	0,87675	0,86944
	LDA	0,85955	0,85014	0,87395	0,86188
	NB	0,81039	0,76303	0,90196	0,82670
TÜMÜ	LR	0,87858	0,87269	0,88736	0,87997
	RF	0,82940	0,87063	0,77501	0,82004
	CART	0,76616	0,76175	0,77669	0,76915
	KNN	0,67594	0,72386	0,57215	0,63912
	SVM	0,86692	0,86136	0,87560	0,86842
	LDA	0,85989	0,84984	0,87532	0,86239
	NB	0,81043	0,76353	0,90109	0,82663

Belirtilen dört kritere göre (doğruluk, kesinlik, duyarlılık ve F Ölçümü) hesaplanan başarı oranları, tüm yöntemlerde görüldüğü üzere genel olarak %80’in üzerindedir. Dolayısıyla seçilen yöntemlerin duygu analizini hesaplamada başarılı olduğu ortaya konulmuştur. Kullanılan yöntemlerin başarı ölçümlerinin grafiği Şekil 3’te gösterilmektedir.



**Şekil 3.** Yöntemlere Göre Model Başarım Ölçümleri

Bu değişkenlere bakıldığında, her ölçüt için LR yönteminin daha yüksek performanslı sonuçlar ürettiği görülmektedir. Gerek sonuçların kesinliği gerek duyarlılığı gerekse bu ölçütler arasındaki optimizasyonun (F ölçümü) yüksekliği açısından LR yönteminin, SVM ve LDA yöntemlerine göre nispeten üstün olduğu görülmektedir.

## 5. Sonuç ve Tartışma

Bu çalışmada TripAdvisor.com sitesinden elde edilen otel müşteri yorumlarını içeren bir veri setine yedi makine öğrenme algoritmalarından (Lojistik Regresyon, Random Forest, CART Sınıflandırıcı, K-En Yakın Komşu, Destek Vektör Makineleri, Doğrusal Diskriminant Analiz Sınıflandırıcısı ve Naïve Bayes) yararlanılarak duygu analizi uygulanmıştır.

Çalışmada kullanılan bu sınıflandırıcıların performans karşılaştırması model başarımlarına göre kıyaslanmıştır. Literatür incelemesinde sunulan tüm çalışmalarda NB analizi uygulanmıştır. Bu analiz sonucunda Raut ve Londhe (2014) %88,00; Oğul ve Ercan (2016) %84,00; Perikos ve Hatzilygeroudis (2017) %85,00; Al-Smadi vd. (2018) %94,10; Shuai vd. (2018) %73,42; Allam vd. (2021) %78,10 ve Luo (2021) %26,00 başarıya ulaşmışken; bu çalışmada %81 başarıya ulaşılmıştır. Diğerlerine göre daha üstün başarı gösteren Al-Smadi vd. (2018) çalışmalarında kendi ülkelerindeki alfabeler üzerine görüntü işleme uygulamaları, diğer çalışmalarda ise engram kullanılması arasındaki farkı ortaya çıkarmıştır. Buradan yola çıkarak NB yönteminin görüntü işleme uygulamasında daha başarılı sonuçlar elde edebileceği sonucuna varılabilmektedir. Görüntü işleme ile birleştirilen bir diğer yöntem Al-Smadi vd. (2018) tarafından SVM yöntemi olduğu görülmüştür. SVM ile sonuç elde edilen Al-Smadi vd. (2018) çalışmasında başarı oranı %95,40 olarak hesaplanmış iken, engramın kullanıldığı bu çalışmada %86,84; Raut ve Londhe (2014) %83,50; Oğul ve Ercan (2016) %70,00; Perikos ve Hatzilygeroudis (2017) %85,00; Shuai vd. (2018) %79,50; Allam vd. (2021) %85,55; Budhi vd. (2021) %90,57 ve Luo (2021) %76,00 başarıya ulaşmıştır. Bu analiz sonucunda da görüntü işlemenin engram kullanmaya göre daha başarılı sonuç elde ettiği görülmektedir. Ancak "SVM analizi görüntü işlemede daha başarılıdır" yargısına tek bir çalışma ile ulaşamayacağı belirtilmelidir.

Yöntem karşılaştırmasında kullanılan bir diğer yöntem olan LR analizini kullanan Shuai vd. (2018) %79,60; Guerreiro ve Rita (2020) %66,86; Budhi vd. (2021) %90,90; Kinyua vd. (2021) %95,00 ve Luo (2021) %71,00 başarı elde etmişken; bu çalışmada ise %87,99 başarıya ulaşılmıştır. Her ne kadar birebir aynı yöntem olmasa da, temel aldığı istatistiksel yapı gereği Wang vd. (2010) kullandığı LRR analizi bu çalışmada belirtilmiştir ve regresyon temelli olan bu analiz sonucunda başarı oranı %86 olarak gözlenmiştir. CART algoritmasının kullanıldığı Raut ve Londhe (2014) %78,40; Guerreiro ve Rita (2020) %66,05; Allam vd. (2021) %65,89; Kinyua (2021) %92,00 başarı elde etmişken; bu çalışmada ise aynı teknik ile %76,92 başarı elde edilmiştir. Bir diğer sınıflandırma algoritması olan RF analizi ile yapılan çalışmalarda Oğul ve Ercan (2016) %89,00; Guerreiro ve Rita (2020) %66,83; Sridharan vd. (2020) %90,69; Budhi vd. (2021) %87,54 ve Kinyua (2021) %95,00 başarıya ulaşmışken ve bu çalışmada %82,00 başarı elde edilmiştir.

Latin alfabesi üzerine çözümlenmelerin, bu diller dışındaki alfabelerin kullanıldığı araştırmalara göre daha fazla olması yöntem açısından görüntü ve n-gram ayrımını doğurmuştur. Bu çalışmada da tercih edilen n-gram kullanımıyla kelime işleme sonuçlarının, konaklama işletmeleri açısından başarılı bir şekilde çözümlenebileceği hem literatürde incelenen araştırmacılar tarafından hem de bu çalışmada kanıtlanmıştır.

İşletmelerin gerek müşteri memnuniyet düzeyleri gerekse işletme içi süreç iyileştirmeleri hakkında daha istikrarlı kararlar alabilmeleri ve değişen piyasa koşullarına göre daha çevik hareket edebilmeleri mümkündür. Duygu analizi yardımıyla; okuma, anlama ve yorumlama adımları son derece hızlı bir şekilde aşılabilmektedir. Bahsi geçen bu üç adımın gerçekleştirilebilmesi için tahsis edilen zaman ve personel, duygu analizinin kullanılmadığı durumlarda işletme için yüksek bir maliyettir. Dolayısıyla yığın müşteri yorumlarındaki karmaşanın kişilerin öznel algılarına dayandırılarak yorumlanmasındansa makine öğrenimi yöntemiyle duygu analizinin gerçekleştirilmesi, daha doğru ve daha yansız sonuçlar elde edilmesine imkân tanımaktadır. Aynı zamanda analiz süreçlerinin kısıtlılığı ile maliyet, zaman ve işgücü kazancının elde edildiği ortadadır.

Bu çalışma sadece Türkçe dilinde yazılmış olan yorumları ele almaktadır. Bu çalışmayı farklı dillerde, farklı kütüphaneler eklenerek yapmak da ilerleyen çalışmalarda mümkün olabilecektir. Böylelikle yapılan yorumların hem farklı kültürlerdeki algılanışını hem de konaklama tesislerinin uluslararası memnuniyet



düzeylerini ortaya çıkarmak mümkün olacaktır. Her ne kadar konaklama sektöründe kullanılmış olsa da farklı niteliklerdeki birçok sektörde de müşteri geri dönüşleri ile ortaya çıkan metinsel veri yığınlarının, etkili ve köklü pazarlama kararlarının alınabilmesi adına analizinin yapılabilmesi mümkündür. Gerek müşteriler gerekse işletme sahipleri açısından tespit edilen eksik noktaların iyileştirme çalışmaları ile sürece yansıtılması duygu analizinin bir getirisi olarak karşımıza çıkabilmektedir. Bu çalışma metin madenciliğinin pazarlama kapsamında kullanılması yönünde ileride yapılacak çalışmalara rehberlik edeceği düşünülmektedir.

#### Kaynakça

- Al-Smadi, M., M. Al-Ayyoub, Y. Jararweh and Qawasmeh, O. (2018). Enhancing Aspect-Based Sentiment Analysis of Arabic Hotels' Reviews Using Morphological, Syntactic and Semantic Features, *Information Processing & Management*, 56 (2), 308-319.
- Allam, A.H., Abdallah, H.M., Amer, E. and Nayel H.A: (2021). Machine Learning-Based Model for Sentiment and Sarcasm Detection, *Sixth Arabic Natural Language Processing Workshop*, Kyiv, Ukraine, 386-389.
- Budhi, G.S., Chiong, R., Pranata, I. and Hu, Z. (2021). Using Machine Learning to Predict the Sentiment of Online Reviews: A New Framework for Comparative Analysis, *Archive of Computational Methods in Engineering*, 1-24.
- Guerreiro, J. and Rita, P. (2020). How to Predict Explicit Recommendations in Online Reviews Using Text Mining and Sentiment Analysis, *Journal of Hospitality and Tourism Management*, 43, 269-272.
- Güran, A., M. Uysal ve Doğruöz, Ö. (2014). Destek Vektör Makineleri Parametre Optimizasyonunun Duygu Analizi Üzerindeki Etkisi, *DEÜ Mühendislik Fakültesi Mühendislik Bilimleri Dergisi*, 48, 86-93.
- Han, J., M. Kamber and Pei, J. (2011). *Data Mining: Concepts and Techniques (3<sup>th</sup> Ed.)*, Burlington, Morgan Kaufmann.
- Kaynar, O., Y. Görmez, M. Yıldız ve Albayrak, A. (2016). Makine Öğrenmesi Yöntemleri ile Duygu Analizi, *International Artificial Intelligence and Data Processing Symposium*, Malatya, Türkiye, 234-241.
- Kinyua, J.D., Mutigwe, C., Cushing, D.J. ve Poggi, M. (2021). An Analysis of the Impact of President Trump's Tweets on the DJIA and S&P 500 Using Machine Learning and Sentiment Analysis, *Journal of Behavioral and Experimental Finance*, 29, 1-14.
- Leskovec, J., A. Rajaraman and Ullman, J. D. (2014). *Mining of Massive Datasets*, Palo Alto, Cambridge University Press.
- Luo, X. (2021). Efficient English Text Classification Using Selected Machine Learning Techniques, *Alexandria Engineering Journal*, 60, 3401-3409.
- Moraes, R., J. F. Valiati and Gaviao Neto, W. P. (2013). Document-Level Sentiment Classification: An Empirical Comparison Between SVM and ANN, *Expert Systems with Applications*, 40 (2), 621-633.
- Oğul, B. B. ve Ercan, G. (2016). Türkçe Otel Yorumlarından Duygu Analizi, *24<sup>th</sup> Signal Processing and Communications Applications Conference (SIU)*, Zonguldak, Türkiye, 16-19 Mayıs, 497-500.
- Omisakin, O. M., Bandara, C. and Kularatne, I. (2020). Designing a Customer Feedback Service Channel Through AI to Improve Customer Satisfaction in the Supermarket Industry, *Journal of Information & Knowledge Management*, 19 (3), 1-34.
- Perikos, I. and Hatzilygeroudis, I. (2017). Aspect Based Sentiment Analysis in Social Media with Classifier Ensembles, in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, Wuhan, Çin, 24-26 May 2017, 273-278.
- Raschka, S. (2015). *Python Machine Learning*, Birmingham, Packt Publishing.
- Raut, V. B. and Londhe D. D. (2014). Opinion Mining and Summarization of Hotel Reviews, in *Sixth International Conference on Computational Intelligence and Communication Networks*, Bhopal, India, 14-16 November 2014, 556-559.

- Shuai, Q., Huang, Y., Jin, L. and Pang, L. (2018). Sentiment Analysis on Chinese Hotel Reviews with Doc2Vec and Classifiers, in *IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12-14 October 2018*, 1171-1174.
- Sridharan, K., Komarasamy, G. and Dainel Madan Raja, S. (2020). Hadoop Framework for Efficient Sentiment Classification Using Trees, *The Institution of Engineering and Technology*. 9 (5), 223-228.
- Wainer, J. (2016). Comparison of 14 Different Families of Classification Algorithms on 115 Binary Datasets, *University of Campinas / Campinas, SP*, arXiv:1606.00930.
- Wang, H., Y. Lu and Zhai, C. (2010). Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach, in *16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, USA, 24-28 July 2010*, 783-792.
- Thatware Platform (2020). Naïve Bayes, <https://thatware.co/naive-bayes/> (Erişim tarihi: 13 Aralık 2020).
- Yin, Y., Y. Song M. Zhang. (2017). Document-Level Multi-Aspect Sentiment Classification as Machine Comprehension, in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 9-11 September 2017*, 2034-2044.
- Zaltman, G. ve Zaltman, L. (2008). Pazarlama Metaforları, İstanbul, Optimist Yayım Dağıtım.