

DUYGU ve GÖRÜŞ MADENCİLİĞİ: SON GELİŞMELER ve 3. HAVALİMANI İÇİN BİR UYGULAMA¹

Emrah BİLGİÇ², Ahmet KOÇAK³

Geliş: 25.12.2018 / Kabul: 05.04.2019

DOI: 10.29029/busbed.502586

Öz

Öteden beri insanoğlunun karar verme süreçlerinde başkalarının görüş ve yorumlarına başvurması değişmeyen bir durumdur. Bilgiye hızlıca erişimin akıllı telefonlar, tabletler ve bilgisayarlarla yapıldığı günümüz Büyük Veri Çağı'nda özellikle internet ortamlarında her konu hakkında görüşler, bilgiler ve yorumlar paylaşılmaktadır. Bu bilgi, görüş ve yorumlardan hem kişisel hem kurumsal bazda faydalı bilgiler çıkarsamak amacıyla kullanılan Metin Madenciliği son yıllarda en çok araştırılan konulardan biri haline gelmiştir. Metin Madenciliği zamanla çeşitli alanlara ayrılmış ve Görüş madenciliği veya Duygu Analizi alanları ortaya çıkmıştır. Günümüzde insanlar sosyal medya araçlarını, blogları, forumları sadece bilgi ve düşünce paylaşmak için kullanmakla kalmayıp, internet üzerinden satın alıp kullandıkları ürün/hizmet ile alakalı bilgileri de bu ortamlar aracılığıyla değerlendirmektedir. Bu çalışma İstanbul'daki Atatürk ve Sabiha Gökçen havalimanlarını kullanan yabancı müşterilerin bir internet sitesinde yaptıkları yorumları Görüş ve Duygu Madenciliği ile analiz etmeyi ve İstanbul'da kurulmuş olan 3. Havalimanı işletmecileri için bir çıkarsama sunmayı amaçlamaktadır. Analizler sonucu insanların kalabalıktan, pahalılıktan, kabalıktan, kayıp bavullardan, zorluklardan ve uçuş rötarlarından şikâyet ettikleri anlaşılmaktadır.

Anahtar Kelimeler: Metin Madenciliği, Görüş Madenciliği, Duygu Analizi, 3. Havalimanı, İstanbul Havalimanı.

1 Bu çalışma 23. Ulusal Pazarlama Kongresi'nde bildiri olarak sunulmuştur.

2 Dr. Öğr. Üyesi, Muş Alparslan Üniversitesi, İşletme Bölümü, e.bilgic@alparslan.edu.tr, ORCID: <https://orcid.org/0000-0002-9875-2299>.

3 Arş. Gör., Bingöl Üniversitesi, İşletme Bölümü, kocakmet@gmail.com, ORCID: <https://orcid.org/0000-0003-4376-4337>.

SENTIMENT AND OPINION MINING: RECENT ISSUES AND AN APPLICATION FOR THE THIRD AIRPORT

Abstract

It is an unchanging situation that human being asks others' opinions and comments during the process of decision making. In today's Big Data Age, where the fast access to information is easily made with smartphones, tablets and computers; opinions, information and comments are shared on every topic, especially on the internet. Text Mining which is used to extract useful information from views and comments for both personal and institutional basis has become one of the most researched subjects in recent years. Text Mining has been divided into different areas over time, and the topics of Opinion Mining or Sentiment Analysis have emerged. Today, people not only use social media tools, blogs, forums to share information and ideas but also use those platforms to evaluate the products/services which they purchased over the internet. This study aims to analyse comments of foreign customers who used airports of Atatürk and Sabiha Gökçen in Istanbul. Opinion Mining and Sentiment Analysis will present an inference for the operator of the 3rd Airport in Istanbul. As a result of the analysis, it is understood that people complain about the crowd, the expensive, the rude, the lost baggage, the difficulties and the flight delays.

Keywords: *Text Mining, Opinion mining, Sentiment analysis, 3rd Airport, Istanbul Airport.*

Giriş

Günümüz internet çağında sosyal medya kullanımının yaygınlaşması ile birlikte, internet sayfalarındaki içeriklerin büyük çoğunluğunu kullanıcılar üretmektedir (Çetin ve Eryiğit, 2018). Bu içeriklere kullanıcı tarafından oluşturulan içerikler adı verilmektedir (User Generated Contents: UGC). Firmalar ve kurumlar kendilerini ilgilendirebilecek sayfalardaki yorumları, görüşleri ve şikâyetleri daha kaliteli hizmet verebilmek amacıyla takip etmek zorundadır. Fakat bu içeriklerin tamamının tek tek okunması hem maliyet hem de zaman açısından mümkün görülmemektedir. Metin Madenciliği yapısal olmayan metinlerden faydalı ve anlamlı bilgiler çıkarılmaya yarayan bir süreçtir. Doküman kümeleme, doküman sınıflandırma, metin özetleme, duygu analizi, sosyal ağ analizi, metindeki konunun tespiti, web sitelerini sınıflandırma, intihal tespiti, spam ve virüs analizi gibi birçok alanda kullanılan Metin Madenciliğinin son yıllardaki en çok uygulaması ise Duygu Analizi (DA) veya başka bir isimle Görüş Madenciliği (GM) olmuştur. GM insanların bir şeye karşı olan görüşleri, tutumları ve duyguları ile alakalı olan bilişimsel bir çalışmadır (Cambria vd., 2013; Ravi ve Ravi, 2015).

İçerisinde bulunduğumuz Büyük Veri çağında her geçen saniyede yapısal ve yapısal olmayan üstelik çok büyük boyutlarla ifade edilebilen türde veriler çeşitli aygıtlar aracılığıyla (Nesnelerin İnterneti: IoT) toplanmakta ve kurumların depolama alanlarına saklanmaktadır. Bu çok büyük boyutlu ve karma ölçekli değişkenlerden oluşan veriyi saklamak, yönetmek ve veriden anlamlı sonuçlar çıkarsamak bahsi geçen Büyük Veri disiplini meydana getirmiştir. Bu durum ayrıca son yıllarda en çok aranan meslek olan Veri Bilimcisi mesleğini de ortaya çıkarmıştır. Büyük Veri'nin yönetimi ayrı bir çalışma alanı iken, bu verilerden çeşitli analizler sonucu faydalı bilgiler çıkarsama (Veri Madenciliği, Metin Madenciliği, Makine Öğrenmesi vb. tekniklerle) başka bir yöntem bilim olarak düşünülebilir. Bilişimsel Dilbilimi ve Veri Madenciliğinin bir alt disiplini olan DA veya GM vurgulandığı üzere dildeki görüş ve tutumlardan oluşan subjektif bilgileri tespit etmeyi ve ortaya çıkarmayı sağlayan bir dizi yöntem ve tekniklerden oluşmaktadır (Liu, 2012).

İnsanoğlu geçmişten bugüne kadar herhangi bir konuda karar verirken başkalarının görüşlerine de başvurmaya ihtiyaç duymuştur. Herhangi bir konu hakkında sunulan görüşler sadece kişinin kendisiyle alakalı karar alma süreçlerinde değil, firmaların, kurumların ürünleri/hizmetleri/politikaları hakkındaki karar alma süreçlerinde de önemlidir. Nitekim iş ortaklarının ve müşterilerin görüşleri, yorumları, duyguları ve tutumları ortaya çıkarılabildiğinde ve değerlendirilebildiğinde bu keşif iş süreçlerinde kalite artışına sebep olacaktır. Görüş, yorum, tutum gibi terimler 2000'li yıllardan sonra literatürde DA ve/veya GM başlığı altında incelenmeye başlamıştır (Piryani, 2016; Ravi ve Ravi, 2015). Bilim camiası özellikle son 15 yıldır bu konu hakkında binlerce araştırma yapmış ve ilgili alan farklı yaklaşım, teknik ve uygulamalarla geliştirilmiştir. 2018 yılında yapılmış olan bir literatür çalışmasına göre, sadece bir veri tabanında %99'u 2004 yılından sonra yayınlanmış olan 7000 adet DA/GM çalışması tespit edilmiştir (Mantyla vd., 2018).

E-ticaret web sitelerinin yanı sıra, forumlar, bloglar, sosyal ağlar, internetteki haber siteleri vb. ortamlar da insanların görüşlerini ifade edebildiği diğer platformlar olarak karşımıza çıkmaktadır. Bu platformlardan DA ve GM sayesinde sadece bir ürün/hizmet hakkındaki paylaşımlar değil, insanların ekonomik, politik olaylar hakkındaki olumlu, olumsuz düşünceleri ve kurumların, ülkelerin stratejileri hakkındaki görüşleri de ortaya çıkarılabilmektedir (Saleh vd., 2011)

Tablo 1. Duygu Analizi ve Görüş Madenciliği Çalışmalarının Dağılımı

<i>Duygu analizi</i>	68.5
<i>Görüş madenciliği</i>	29.1
<i>Duygu sınıflandırma</i>	18
<i>Görüş analizi</i>	5.6
<i>Anlamsal yönlendirme</i>	3.8
<i>Sentiwordnet</i>	2.7
<i>Görüş sınıflandırma</i>	1.4
<i>Duygu madenciliği</i>	1.3
<i>Öznellik analizi</i>	1.1
<i>Duygu</i>	1
<i>Öznellik sınıflandırma</i>	0.8

Kaynak: (Mantyla vd., 2018)

Bu çalışma, hem güncel literatürü inceleyip alanla ilgili yeni gelişmeler hakkında bilgi vermek, hem de R programlama dilinde yeni hazırlanmış olan Metin Madenciliği ve DA ve GM paketlerini güncel meselelerden biri olan İstanbul'daki 3. Havalimanı için bir uygulama ile tanıtmayı amaçlamaktadır. Çalışma hazırlanırken konu ile alakalı Türkçe çalışmaların azlığı da dikkat çekmiştir. Dolayısıyla bu çalışmanın Türkçe çalışmaları teşvik edeceği de düşünülmektedir.

Literatür Taraması

Boumans ve Trilling (2016) tarafından dar kapsamda üç başlıkta: hazır bir sözlükten faydalanarak sayma yöntemi, gözetimli makine öğrenmesi ve gözetimsiz makine öğrenmesi incelenen DA, Ravi ve Ravi (2015) tarafından ise altı geniş başlıkta incelenmiştir. Bu altı başlık, öznellik (subjectivity) sınıflandırma, duygu sınıflandırma, yorumun kullanışlılığı ölçümü, duygu sözlüğü (lexicon) oluşturma, görüş kelimelerini ve ürün durumunu çıkarsama ve diğer DA uygulamalarıdır. Bu çalışmada kullanılacak olan duygu sınıflandırması yönteminde hazır sözlükler yardımıyla eldeki metin sınıflandırılacaktır.

Makine Öğrenmesi tabanlı yaklaşım, sınıflandırma sırasında makine öğrenme algoritmalarını ve dilbilimsel özellikleri kullanmaktadır. Gözetimli öğrenmede algoritma eğitim veri setinden örüntüleri öğrenmektedir. Örneğin DA yapabilmek için algoritmaya daha önceden pozitif, negatif veya nötr olarak kodlanmış ifadeler öğretilir. Gözetimsiz öğrenmede ise daha önceden tanımlanmış etiketleme veya eğitim veri seti yoktur. Bunun yerine algoritma bir model kullanarak metindeki örüntüleri tespit eder. Burada araştırmacının etki edebileceği tek şey bazı parametrelerin, örneğin metnin kaç sınıfa ayrıştırılacağı gibi belirlenmesidir.

Sözlük tabanlı yaklaşım ise sınıflandırma sırasında önceden hazırlanmış duygu kavramlarından oluşan sözlüklerden yararlanmaktadır. Bu hazır sözlüklerde pozitif, negatif veya nötr duygu ifadeleri yer alır. Bazı sözlükler ise daha kapsamlı olup ayrıca *mutluluk, sevinç, kızgınlık, şaşkınlık* gibi duygu ifadelerini de içermektedir. Verilen metindeki duygu ifadeleri çeşitli algoritmalar sayesinde metin içerisinde aranıp sözlükteki duygu ifadeleri ile eşleştirilerek eldeki metin pozitif, negatif veya nötr olarak belirlenir (Eliacık ve Erdoğan 2015). Her ne kadar araştırmacılarının görüş birliğine vardığı duyguların tamamını ifade edebilecek bir sözlük olmasa da bilimsel çalışmalarda bu çalışmada da kullanılacak olunan *Bing* sözlüğü, *Laughran* sözlüğü ve *NRC* sözlüğüdür. Günümüzde Türkçe sözlükler de oluşturulmuş ve bu sözlükler geliştirilmeye devam edilmektedir (Türkçe Ulusal Derlemi, TS Corpus, STC Sözlük Türkçe Derlemi). Ayrıca, Türkçe Doğal Dil İşleme (NLP) çalışmaları da yapılmış ve yapılmaya devam edilmektedir (Eryiğit, 2014; Çetin ve Eryiğit, 2018).

Geçmişte anket yolu ile yüz yüze görüşme tekniği vb. yollar ile müşterilerinin/kullanıcılarının görüşlerini tespit eden firmalar/kuruluşlar daha önceden değinildiği gibi, insanların internet ortamlarındaki oluşturdukları içerikleri kullanarak pazarlama, ürün tasarımı, satış ve halkla ilişkiler gibi işletmeyle alakalı karar verme süreçlerini desteklemektedirler. Bu çalışmalara örnek teşkil etmeleri amacıyla aşağıdaki özetler verilmiştir.

Jin vd., (2016) ürün tasarımcılarının karar verme süreçlerini desteklemek amacıyla, müşteri yorumlarını analiz etmiştir. Bu amaçla Amazon.com sitesindeki benzer ürünlerin yorumları DA ile incelenmiştir. Çalışmanın amacı ürün tasarımı süreçlerini pazarın ihtiyaçları doğrultusunda şekillendirmektir. Birçok cep telefonu markasının binlerce yorumlarını inceleyen çalışmada müşteriler en çok telefonların ekranını, bataryasını, internet ve kablosuz ağını, uygulamalarını ve hafızasını yorumlamışlardır. Yorumlar ayrıca Bayesçi yöntem kullanılarak olumlu ve olumsuz olarak sınıflandırılmıştır.

Koçak, Polat ve Koçak (2016) Twitter kullanıcılarının havayolu firmaları hakkındaki 1879 adet yorumu olumlu, olumsuz ve nötr etiketlerle sınıflandırmıştır. Daha sonra veri setinin yüzde 25'lik kısmı eğitim veri seti olarak ayrılmış ve makineye öğretilmiştir. Yapılan SMO analizi esnasında uygulanan standart ve normalize edilmiş Kernel Polinomu sonuçlarına göre, Standart Kernel Polinomu'nun SMO algoritması içinde daha iyi bir performans gösterdiği; ancak genel olarak sınıflandırma başarısının düşük çıktığı görülmektedir. Duygu kutupları incelendiğinde '*ucuz değil*' '*hiç*' ve '*rötar*' kelimeleri öne çıkmaktadır.

Yıldırım ve Yüksel (2017) belirli bir tarih aralığında Borsa İstanbul'da işlem gören bir şirketin hisse senedi fiyatlarını negatif veya pozitif yönlü olarak etiketlemiştir. Daha sonra ilgili günlerde Twitter kullanıcılarının gönderdiği mesajlar üzerinden DA gerçekleştirilerek, günlük duygu polaritesi değerini elde etmişlerdir.

Duygu polaritesi değeri elde edebilmek için veri setinin bir kısmı eğitim veri setine ayrılmış ve manuel olarak pozitif-negatif şeklinde kodlanmış ve daha sonra ZeroR, J48, k-en yakın komşu yöntemi ve Naive Bayes sınıflandırma teknikleri ile bütün veri seti sınıflandırılmıştır. En iyi sınıflandırma sonucunu veren Naive Bayes yönteminden elde edilen duygu polarite değerleri ile hisse fiyatı hareketinin yönü arasında ilişki olup olmadığı analiz edilmiş ve Spearman'ın sıra ilişkisi katsayısına göre iki veri seti arasında negatif ve orta düzeyde bir ilişkiye rastlanmıştır.

Nayebi ve Abran (2017), günümüzde Apple markasının uygulama mağazasının (app-store) uygulama sahiplerine yıllık 10 milyar dolardan fazla kazanç sağladığını ve yılda 3 milyondan fazla uygulamayı piyasaya sunduğunu hatırlatarak, uygulama mağazalarının önemine dikkat çekmiş ve çalışmalarında uygulama mağazalarındaki yorumlarla alakalı yazılmış olan görüş madenciliği derlemiştir. İncelenen çalışmalar sonucunda araştırmacıların şu zorluklarla karşılaştığı gözlemlenmiştir: Yorumlarda kullanılan dilin yapısal olmayışı analizleri zorlaştırmakta, spam yorumlar analizlerin sonuçlarını saptırmakta, dolayısıyla spam yorumlara çözüm bulunması gerekmektedir. Çalışma sonucunda ise şu bulgular elde edilmiştir: Çalışmaların çoğu halen temel düzeyde olup, sınıflandırma analizi ve korelasyon analizi kullanılmaktadır, bu sebeple kullanıcı yorumlarından faydalı bilgileri ortaya çıkaracak, sadece bu alana özgü bir model geliştirilmelidir. Yorumların içerisindeki spam veya faydasız bilgiler sebebi ile yorumlar dış kaynaklardan elde edilmiş verilerle (örneğin sosyal medya, bloglar) zenginleştirilmelidir.

Singh ve Tucker (2017) internet ortamındaki ürünle alakalı yorumların müşterilerin satın alma kararlarında çok önemli olduğunu vurgulamış, bu amaçla makine öğrenmesi yaklaşımı ile müşteri yorumlarını otomatik olarak iki ana grupta sınıflandırmışlardır. Analizde kullanılacak olan veri Amazon.com sitesinden, 3 farklı model cep telefonu için yapılmış yorumlardan elde edilmiş olup, sınıflandırma işi Naif Bayes, Destekçi Vektör Makineleri ve Karar Ağaçları yöntemleri ile yapılmıştır. Başarıyla uygulanan sınıflandırma ile yorumlar, iki ana gruptan birincisi için üç (ürünün formu, fonksiyonu ve davranışı) ve diğeri için iki (hizmet ve diğer) alt gruba sınıflandırılmıştır.

Xu vd., (2017), online alışverişlerde, kullanıcıların oluşturdukları içeriklerin (UCG) önemine vurgu yaptıktan sonra, bu içeriklerin İş Zekâsı ve İş Analitiği açısından önemine de değinmiştir. Çalışmada Metin Madenciliğinin Gizli Semantik Analizi yöntemi ile otel müşterilerine ait *booking.com* adresindeki 3500 kadar yorumlarını tatmin veya tatminsizliğe yol açan faktörleri ortaya çıkarabilmek amacıyla analiz edilmiştir. Olumlu yorumlar ve olumsuz yorumlar ayrı ayrı incelenmiş olup, olumlu görüşlerin %95'ini arkadaş canlısı çalışanlar, rahat oda, iyi konum, değerli gibi terimler oluşturmaktadır. Olumsuz görüşlerin sebebi ise değersiz, rahat olmayan ve kirli oda, kaba ve yetersiz çalışanlar gibi terimlerden kaynaklanmaktadır.

Vinodhini ve Chandrasekaran (2017) online alışveriş sonrası yapılan yorumların ilgili firmanın pazarlama stratejilerini doğrudan etkilediğini belirtmiş, yapılan yorumların iki ana sınıfta (olumlu ve olumsuz) otomatik olarak gruplandırılması gerektiğini vurgulamışlardır. Bu amaçla farklı Makine Öğrenmesi sınıflayıcıları ile dengesiz (imbalanced) yapıda olan yorumlar sınıflandırılmıştır. Kullanılan veri seti *Amazon* sitesinden, fotoğraf makineleri ile ilgili yapılan yorumlardan elde edilmiş olup sınıflandırmanın performansını arttırmak amacıyla yeni bir algoritma önermişlerdir.

Ravi vd., (2017) ürünle alakalı olumsuz yorumların, ürünü almaya niyetli müşterileri satın alma kararından vazgeçirebilecek kadar güçlü olduğunu vurgulamış ve olumsuz yorumların firma tarafından hızlıca analiz edilmesi gerektiğini belirtmişlerdir. Bu amaçla, Bulanık Mantık ile DA'ni harmanlayan araştırmacılar bankalar ve sigorta şirketleri ile alakalı şikâyetleri analiz ederek, daha önceki çalışmalarda yapılan herhangi bir ürünle alakalı analizlerin aksine, bir hizmetle alakalı analizler yapmışlardır. Aguwa vd., (2017) işletmelerin karar verme süreçlerine destek olması amacıyla, bulanık mantığı kullanarak, bir otomobil firmasının müşterilerin sesi (VoC) verilerine Metin Madenciliği ve Birliktelik Kuralları Analizi uygulamış, ortaya çıkan sonuçlarla müşterilerin memnuniyet endeksi oluşturulmuştur. Yaman (2018), arkeoloji ile alakalı yapılan haberlerin içeriklerindeki yanlış yönlendirmeleri tespit etmek amacıyla negatif ve pozitif anlam ihtiva eden kelimeleri belirledikten sonra, İstanbul'la alakalı 10 yıl boyunca yayınlanmış gazete haberlerinin içerikleri DA'ne tâbi tutulmuştur.

Uygulama

Bu çalışmada daha önceden başka araştırmacılar tarafından hazırlanmış hazır duygu sözlükleri kullanılacak, yöntem olarak da eldeki dokümanda yer alan kelimelerle hazır duygu sözlüklerindeki kelimeler karşılaştırılacaktır. Bu sözlükler Bing Liu tarafından hazırlanmış Bing sözlüğü, Saif Mohammad tarafından hazırlanmış NRC sözlüğü ve Tim Loughran and Bill McDonald tarafından hazırlanan Loughran sözlüğüdür. Bu sözlükler sayesinde, yorumların yer aldığı doküman boyunca insanların genel olarak pozitif ve negatif duygu ifadeleri sınıflandırılacak ve pozitif-negatif kelime sayıları hesaplanacaktır (Pang ve Lee, 2008). Bu hesaplama ile duygu puanı elde edilecektir.

Örneklem olarak Atatürk ve Sabiha Gökçen Havalimanları hakkındaki görüşlerin ve yorumların seçilmesinin sebebi, gündemde olan 3. Havalimanı için bir çıkarsamada bulunma amacıdır. Ayrıca karakter yazımının kısıtlı olduğu Twitter yerine daha fazla görüşün, yorumun yazılabileceği Google Haritalar tercih edilmiştir. Üçüncü havalimanı işletmecisi tarafından daha fazla ortamdaki, daha fazla müşterinin yorum ve görüşleri elde edilip analiz edilmesi daha iyi sonuçların elde edilmesine yol açacaktır.

Bu doğrultuda, Atatürk Havalimanı için 2000'e yakın ve Sabiha Gökçen Havalimanı için ise yine 2000'e yakın toplamda yaklaşık 4000 yorum, İngilizce dilinde ve manuel olarak ilgili internet sitesinden alınmıştır. Verilerin Manuel olarak çekilmesinin sebebi, Google firmasının harita yorumlarına Maps JavaScript API üzerinden tam erişim için premium anahtar paketine, bir defalık kullanım için bile, yüksek bir ücret (11.000 \$) talep etmesinden kaynaklanmaktadır. Ayrıca, yorumların İngilizce alınma sebebi ise analizde kullanılacak olan programlama dilinde İngilizce sözlüklerin bulunması sebebiyledir. Daha önce vurgulandığı gibi Türkçe çalışmalar hazırlanmış ve geliştirilmeye devam edilmektedir. Yorumların yer aldığı doküman boyunca insanların genel olarak pozitif ve negatif duygu ifadeleri hazır sözlüklerle karşılaştırılıp sınıflandırılacak ve pozitif-negatif kelime sayıları hesaplanacaktır (Pang ve Lee, 2008).

R programlama dilindeki Metin Madenciliği ile alakalı son yıllarda hazırlanmış olan bazı paketler sayesinde kullanılacak olan veri (yorumların içerildiği doküman) programlama dilinin ortamına aktarılmıştır. Veri ön işleme sürecinde bütün haldeki veri seti programlama dili tarafından tek tek kelimelere ayrıştırılıp (tokenization) sıkça kullanılan anlamsız kelimeler (the, of, a, an...) veri setinden kaldırılmıştır (detaylı bilgi için: Welbers vd., 2017). Daha sonra kelimeler, eksiz haline yani kök haline çevrilmiştir (stemming). Bu işlemler için R programlama Dilindeki *tm* paketi (Feiner, 2013) kullanılmıştır. Duygu ifadelerinin tespiti için ise *tidytext* paketi (Silge ve Robinson, 2016) kullanılmıştır.

Bulgular

Daha önce vurgulandığı üzere, R programlama dili kullanılarak yapılan analizlerde şu sonuçlara ulaşılmıştır: Bing sözlüğüne göre bahsi geçen yorumlarda 2740 negatif, 2565 pozitif kelime vardır. Dolayısıyla duygu puanı -175'dir. NRC sözlüğüne göre ise, 2674 negatif, 2902 pozitif kelime vardır ve duygu puanı, 228'dir. Loughran sözlüğüne göre ise, 1486 negatif, 656 pozitif kelime vardır ve duygu puanı, -830'dur. Kullanıcı yorumlarındaki sözcüklerin her üç sözlükte de eşleştiği bazı pozitif ve negatif duygu ifade eden sözcükler incelendiğinde aşağıdaki tablo karşımıza çıkmaktadır. Bu tabloya mantıklı olan kelimeler eklenmiştir. Mantıksız olanları eklenmemiştir. Örneğin "rehber" kelimesi pozitif bir ifade olarak sözlüklerde karşımıza çıkmaktadır. Hâlbuki rehber kelimesini insanlar bir rehber ihtiyacı duyuyoruz şeklinde dile getirmişlerdir. Diğer bir örnek pozitif sözcük olarak karşımıza çıkmış bedava kelimesidir. Hâlbuki insanlar bedava internetin olmadığını (Free Wi-Fi) vurgulamaktadır. Görüldüğü üzere, pozitif ifadeler araştırmacıları yanıltabilmektedir. Ayrıca, bir cümlede her ne kadar pozitif kelime kullanılmışsa da 'değil' 'hiç' 'olmaz' gibi kelimelerle o cümlenin negatif yapılabileceği de unutulmamalıdır.

Tablo 2. Kullanıcı Yorumlarında Pozitif veya Negatif Duygu İfade Eden Sözcükler

	Bing sözlüğü	Loughran	NRC sözlüğü
P o z i t i f kelimeler	Hoş, iyi, güzel, çok iyi, modern, temiz, yeterli, çok güzel, rahat, yardımsever, süper	İyi, güzel, çok iyi, en iyi, kolay, arkadaşça, mükemmel, harika, ideal, fantastik, etkili	İyi, yemek, güzel, temiz, sanat, yardım sever, bilgi, arkadaşça, mükemmel, sevgi,
N e g a t i f kelimeler	Problem, kalabalık, ölüm, küçük, bekleme, kirli, korkunç, gecikme, pahalı, kaba, geç, zor	Problem, kötü, çok kötü, kaba, ertelenmiş, kayıp, geç, zor, karışık	Problem, kötü, ölüm, küçük, kirli, yetersiz, savaş, ertelenmiş, kayıp, pahalı, geç

Tablo 2 incelendiğinde yolcuların kalabalıktan, pahalılıktan, kabalıktan, kayıp bavullardan, zorluklardan ve uçuş rötarlarından şikâyet ettikleri anlaşılmaktadır. Ayrıca son yaşanan terör saldırısı da ölüm kelimesini yorumlarda kullandırtmıştır.

Aşağıdaki tabloda ise, NRC sözlüğü kullanılarak yorumlarda ‘keyif’ ifade eden sözcükler, ‘sinir’ ifade eden sözcükler, korku ifade eden sözcükler ve ‘güven’ ifade eden sözcükler ve kullanım sayılarına yer verilmiştir. Burada da aldatıcı eşleşmeler bulunmaktadır. Örneğin ‘court’ kelimesini mahkeme olarak algılamıştır. Hâlbuki buradaki ‘court’, ‘food court’ (yemek alanı) kelimesi ile alakalıdır. Aynı durum terminal kelimesi için de geçerlidir. Sözlükte terminal kelimesi ölüm olarak algılanmıştır.

Tablo 3. NRC Sözlüğündeki ‘Keyif’ ‘Sinir’ ‘Korku’ ‘Güven’ İfade Eden Sözcüklerle Eşleşen Müşterilerin Kullandığı Bazı Sözcükler ve Kullanım Sayıları

Keyif	Sinir	Korku	Güven	Tikslenme	Umut	Üzüntü	Sürpriz
İyi (166)	Kötü (75)	Problem (779)	Rehber (803)	Kötü (75)	Havaalanı (1068)	Problem (779)	İyi (66)
Yemek (96)	Para (33)	Kötü (75)	İyi (166)	Kirli (43)	İyi (166)	Kötü (75)	Sanat (41)
Güzel (73)	Mahkeme (23)	Ölüm (74)	Yemek (96)	Korkunç (21)	Zaman (128)	Ölüm (74)	Para (33)
Temiz (63)	Korkunç (21)	Ölüm (59)	Temiz (63)	Kaba (19)	Uzun (90)	Ölüm (terminal) (59)	Beklenti (24)
Sanat (41)	Kaba (19)	Önlemek (33)	Kapı (56)	Pahalı (16)	Bekleme (44)	Sanat (41)	Alışveriş (17)
Ödeme (34)	Pahalı (16)	Sıra (24)	Ödeme (34)	İğrenç (14)	Sanat (41)	Kalkış (25)	Ayrılık (13)
Para (33)	Tiksindirici (14)	Mahkeme (23)	Para (33)	Korkunç (14)	Ödeme (34)	Kaba (19)	Yolculuk (12)
Y.sever (32)	Korkunç (14)	Korkunç (21)	Y.sever (32)	Tuvalet (11)	Para (33)	Kayıp (17)	Umut (11)
Arkadaş (29)	Sınırlı (12)	Savaş (21)	Mükemmel (28)		Arkadaşça (29)	Pahalı (16)	Pazarlık (10)

Tablo 4. Yorumlarda En Çok Kullanılan Tek ve Çift Kelimeler

Tekli	Çiftli
<i>Yerel</i>	<i>Yerel rehber</i>
<i>Rehber</i>	<i>Bedava wifi</i>
<i>Wifi</i>	<i>Duty free</i>
<i>Bedava</i>	<i>Pasaport kontrol</i>
<i>insanlar</i>	<i>En kötü havaalanı</i>
<i>çalışanlar</i>	<i>Türk Havayolları</i>
<i>kalabalık</i>	<i>Güzel havaalanı</i>
<i>zaman</i>	<i>Büyük havaalanı</i>
<i>uçuş</i>	<i>Çok kalabalık</i>
<i>Türk</i>	<i>Transfer uçuşlar</i>

Sonuç ve Tartışma

Bilgisayarlardan sonra, tabletlerin ve akıllı telefonların hayatımıza girmesi ile birlikte, özellikle sosyal medya, blog ve forumların kullanımının hızla artması, hem müşterilerin hem firmaların/kurumların karar verme süreçlerinde bahsedilen internet ortamlarından yoğun bir şekilde faydalanmasını sağlamıştır. Dolayısıyla ortamlardaki yorumların, değerlendirmelerin, görüşlerin analizi firmalar/kurumlar tarafından zorunlu hale gelmiştir. Bu bağlamda son yıllarda akademide de bu alandaki çalışmalar artmıştır. Bilgisayar bilimciler ve istatistikçiler tarafından geliştirilen algoritmalar ve modeller sosyal bilimler alanında hızlıca uygulanmaya başlamıştır. Araştırmacılar tarafından halen, daha etkin algoritmalarla birlikte daha kapsamlı sözlükler oluşturulmaya çalışılmaktadır. Çünkü bazı cümleler her ne kadar olumlu kelimeler ihtiva etse de anlam olarak olumsuz olabilmektedir. Ayrıca araştırmacılar müşterilerin ürün ve hizmetler hakkındaki düşüncelerine dair verilerini takip eden, veri tabanında toplayan ve otomatik endeksleyerek özetleyen yazılımlar da geliştirmektedirler.

Bu çalışma için, gelecekte toplanabilecek daha fazla yorum ve görüşlerle birlikte daha detaylı analizler yapılabilir. Örneğin, günümüzde farklı alanlara özel sözlükler oluşturulmaktadır (ekonomi, finans, eğitim sözlüğü gibi) ve havalimanı işletmeciliği veya daha geniş kapsamda ulaşım işletmeciliği için oluşturulabilecek özel bir sözlükle daha anlamlı sonuçlar elde edilebilir. Ayrıca bu çalışmadaki gibi tek kelime üzerinden yapılan analizler yerine ikili kelimelerle analizler yapılabilir. Çünkü analizler sonucu, olumsuz duygu ifadelerinde karşımıza çıkan örneğin “mahkeme” kelimesi (court) aslında müşteriler tarafından bu anlamda kullanılmamış, “food court” kelimesi ayrı ayrı analize dâhil edildiği için sadece “court” kelimesi algoritma tarafından okunmuştur. Aynı durum “ölümcül” (terminal) kelimesi için de geçerlidir. Bu tip analizler için bilgisayar bilimciler tarafından yeni algoritmalar yazılmalıdır.

Son olarak, yorumların bir kısmı olumlu, olumsuz veya nötr şeklinde önceden belirlendikten sonra, Sınıflandırma Analizi kullanılarak geri kalan diğer yorumların pozitif, negatif veya nötr anlamlı olup olmadıkları belirlenebilir. Otomatiğe bağlanacak bu işlemle kurumlar daha kolay bir şekilde müşteri yorumlarını sınıflandırmış olacaklardır. Ayrıca, yapılacak bu tip analizlerde zaman kavramı da önem arz etmektedir. Çünkü daha önce vurgulandığı gibi işletmenin müşteri görüşlerini sürekli analiz etmesi ile beraber gerekli düzenlemeler yapıldıktan sonra müşterilerin yorumlarının olumlu yönde değişip değişmediği takip edilmelidir.

KAYNAKLAR

AGUWA Celestine, OLYA, Mohammad Hessam, & MONPLAISIR, Leslie (2017), Modeling of fuzzy-based voice of customer for business decision analytics, *Knowledge-Based Systems*, 125, 136-145.

BOUMANS, Jelle W, & TRILLING, Damian (2016), Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars, *Digital journalism*, 4(1), 8-23.

CAMBRIA, Erik, SCHULLER, Björn, XIA, Yunqing, & HAVASI, Catherine (2013), New avenues in opinion mining and sentiment analysis, *IEEE Intelligent Systems*, 28(2), 15-21.

ÇETİN, Fatih Samet, & ERYİĞİT, Gülşen (2018), Türkçe Hedef Tabanlı Duygu Analizi İçin Alt Görevlerin İncelenmesi–Hedef Terim, Hedef Kategori Ve Duygu Sınıfı Belirleme, *Bilişim Teknolojileri Dergisi*, 11(1), 43-56.

ELİAÇIK, Alpaslan Burak, & ERDOĞAN, Nadia (2015), Mikro Bloglardaki Finans Toplulukları için Kullanıcı Ağırlıklandırılmış Duygu Analizi Yöntemi. Paper presented at the UYMS.

ERYİĞİT, Gülşen (2014), ITU Turkish NLP web service. Paper presented at the Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics.

FEINERER, Ingo (2013), Introduction to the tm Package Text Mining in R. Accessible en ligne: <http://cran.r-project.org/web/packages/tm/vignettes/tm.pdf>.

GENC-NAYEBI, Necmiye, & ABRAN, Alain (2017), A systematic literature review: Opinion mining studies from mobile app store user reviews, *Journal of Systems and Software*, 125, 207-219.

JIN, Jian, LIU, Ying, JI, Ping, & LIU, Hongguang (2016), Understanding big consumer opinion data for market-driven product design, *International Journal of Production Research*, 54(10), 3019-3041.

KOÇAK, Bahri Baran, POLAT, Inci, & KOÇAK, Cem Burak, Twitter Kullanıcılarının Havayolu Pazarına Yönelik Duygu Kutuplarının Belirlenmesi: Bir Fikir Madenciliği Örneği, *PressAcademia Procedia*, 2(1), 684-691.

LIU, Bing (2012), Sentiment analysis and opinion mining, *Synthesis lectures on human language technologies*, 5(1), 1-167.

MÄNTYLÄ, Mika V, GRAZIOTIN, Daniel, & KUUTILA, Miikka (2018), The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16-32.

PANG, Bo, & LEE, Lillian (2008), Opinion mining and sentiment analysis, *Foundations and Trends® in Information Retrieval*, 2(1–2), 1-135.

PIRYANI, Rajesh, MADHAVI, D, & SINGH, Vivek Kumar (2017), Analytical mapping of opinion mining and sentiment analysis research during 2000–2015, *Information Processing & Management*, 53(1), 122-150.

RAVI, Kumar, & RAVI, Vadlamani (2015), A survey on opinion mining and sentiment analysis: tasks, approaches and applications, *Knowledge-Based Systems*, 89, 14-46.

RAVI, Kumar, RAVI, Vadlamani, & PRASAD, P Sree Rama Krishna (2017), Fuzzy formal concept analysis based opinion mining for CRM in financial services, *Applied Soft Computing*, 60, 786-807.

RUSHDI-SALEH, Mohammed, MARTÍN-VALDIVIA, M Teresa, UREÑA-LÓPEZ, L Alfonso, & PEREA-ORTEGA, José M. (2011), Bilingual experiments with an arabic-english corpus for opinion mining, Paper presented at the Proceedings of the International Conference Recent Advances in Natural Language Processing 2011.

SILGE, Julia, & ROBINSON, David (2016), tidytext: Text mining and analysis using tidy data principles in r, *The Journal of Open Source Software*, 1(3), 37.

SINGH, Abhinav, & TUCKER, Conrad S. (2017), A machine learning approach to product review disambiguation based on function, form and behavior classification. *Decision Support Systems*, 97, 81-91.

VINODHINI, G, & CHANDRASEKARAN, RM. (2017), A sampling based sentiment mining approach for e-commerce applications, *Information Processing & Management*, 53(1), 223-236.

WELBERS, Kasper, VAN ATTEVELDT, Wouter, & BENOIT, Kenneth (2017), Text analysis in R, *Communication Methods and Measures*, 11(4), 245-265.

XU, Xun, WANG, Xuequn, LI, Yibai, & HAGHIGHI, Mohammad (2017), Business intelligence in online customer textual reviews: Understanding consumer perceptions and influential factors, *International Journal of information management*, 37(6), 673-683.

YAMAN, Tutku Tuncalı (2018), Measuring Effects of Post-truth in Urban Archaeology: Istanbul Case, 8th International Symposium on Information Management in a Changing World, Ankara, Turkey, 2018.

YILDIRIM, Metin, & YÜKSEL, Cenk Arsun, Sosyal Medya İle Hisse Senedi Fiyatının Günlük Hareket Yönü Arasındaki İlişkinin İncelenmesi: Duygu Analizi Uygulaması, *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 33-44.

<https://std.metu.edu.tr/>

<https://tscorpus.com/>