

İÇİNDEKİLER

ŞEKİL LİSTESİ	2
TABLO LİSTESİ	3
1 ÖZET	4
2 ANAHTAR KELİMELER	4
3 GİRİŞ.....	4
4 LİTERATÜR.....	5
5 METODOLOJİ.....	6
5.1 Neden Matlab?	6
5.2 Uygulama Adımları	7
5.2.1 Veri Toplama.....	7
5.2.2 Veri Ön İşleme	7
5.2.3 Verilerin Analizi	8
5.2.4 Verilerin Görselleştirilmesi	9
6 BULGULAR	10
7 SONUÇ	14
KAYNAKÇA	Error! Bookmark not defined.
EK.....	Error! Bookmark not defined.

ŞEKİL LİSTESİ

Şekil 1 Ham Veriye Ait Word Cloud	10
Şekil 2 Ön İşlem Uygulanmış Veriye Ait Kelime Bulutu.....	10
Şekil 3 Bag of N Grams’a Ait Kelime Bulutu	11
Şekil 4 LDA Modeline Ait Kelime Bulutu	12
Şekil 5 Duygu Analizi Yöntemine Ait Kelime Bulutları	14

TABLO LİSTESİ

Tablo 1 Bag Of N Grams Özellikleri	11
Tablo 2 Bag Of N Grams’a Ait Word Cloud Özellikleri	11
Tablo 3 LDA Modeli Özellikleri.....	12
Tablo 4 5x1 Tokenized Dökümanına Ait Tablo.....	13

1 ÖZET

Son zamanlarda yaşanan gelişmeler ve küreselleşme ile birlikte firmaların hayatta kalabilmesi adına önemli değişikliklere gitmesi firmalar için kaçınılmaz olmuştur. Ürün veya kar odaklı işletme modeli geçerliliğini yitirmiş yerini müşteri odaklı yapılanmalar almıştır. Günümüzün zorlu rekabet koşullarında rekabet avantajı elde edebilmek için işletmeler “müşteri kraldır” anlayışını benimsemekte ve alacağı her türlü kararda müşteri memnuniyetini esas almaktadırlar. Müşteri odaklılığın sağladığı rekabet üstünlüğünü elde etmek adına müşterilerin ne istediklerini bilmek her zamankinden daha önemli bir hale gelmiştir. Günümüz koşullarında müşteri görüşleri açısından büyük miktarda veri barındıran kaynakların başında hiç şüphesiz sanal ortamlar gelmektedir. İnsanların duygu, düşünce ve fikirlerini sanal ortamlarda ifade etmesi çok büyük miktarda veri oluşturmaktadır. Bu veriler firmalar adına önemli bir kaynak oluşturmaktadır. Bu perspektif çerçevesinde bu çalışmada müşterilerin e-alışveriş sitelerinde yapmış olduğu yorumlar ele alınacak; metin ve fikir madenciliği çalışmaları yapılacaktır.

2 ANAHTAR KELİMELER

Müşteri Odaklılık, Fikir Madenciliği, Metin Madenciliği, Duygu Analizi, N gram Counting, Sentiment Analysis, Machine L-Sentiment Analysis, Topic Model, Metin Analitiği

3 GİRİŞ

Yeni nesil internet akımı ile hızlı küresel değişimler yaşanmış web sayfaları insanların duygu ve fikirlerini gerçek zamanlı ifade ettiği mecralar haline gelmiştir. Geçmişte statik yapıda olan web sayfalarında kullanıcılar içerik tüketici konumunda iken, günümüzde web sayfaları kullanıcıların içerik oluşturduğu dinamik bir yapı haline dönüşmüştür. Bu dönüşümle herkesin yer alabildiği ağ toplumu oluşmuş ve geniş kitlelerin işbirliği ve birlikte üretimi mümkün hale gelmiştir.[1]

Sosyal ağlar müşteri ilişkileri yönetiminden satış ve pazarlamaya, insan kaynaklarından bilgi teknolojisine kadar hemen her departmanda önemli rol oynamakta ve ürünlerin pazarda tanıtımı ve ekonomiye katkı sağlamak için bilinirliği açısından önemli araçlar olarak firmaların bütünleşik yapısında yer almaktadır. [2] Bugün pek çok insan herhangi bir ürünü satın almadan önce o ürünü kullanan kişilerin e- alışveriş sitesindeki görüşlerine başvurmaktadır. Dolayısıyla e-alışveriş siteleri ürünlerle ilgili kullanıcıların yazdığı yorumlardan yola çıkarak ürün kullanıcılarında oluşan duyguların anlaşılmasında önemli ipuçları sunmaktadır. Müşteri düşüncelerini açığa çıkartmak ürünlerden hoşlanıp hoşlanmadığını tespit etmek adına duygu analizi önem arz etmektedir. Fikir madenciliği ve duygu analizinde amaç; saklı olan duygu, düşünce ve fikirleri ortaya çıkarmaktır. Duygu ve düşüncelerin açığa çıkarılıp analiz edilmesiyle firmalar müşterilerin duygu ve düşünceleri bilerek müşteri beklentilerini daha iyi karşılayabilir, böylelikle ürün gamı ve kalitesini şekillendirmek üzere kullanabilirler.[3]

Günümüz internet ortamı sayesinde müşteri fikirleri küreselleşmiş, işletme müşteri etkileşimi artmıştır. Bu yeni iletişim ortamı, Müşteri İlişkileri yönetimi açısından metin veri sağlamaktadır. Dolayısıyla işletmeler sosyal ağ ortamında ifade edilen görüşleri, müşteri odaklı üretim açısından sundukları ürün/hizmetlerin geri bildirimlerini öğrenebilmeleri açısından etkin bir analiz için gerekli veri kaynağı olarak ele almalıdırlar. Bu çalışmada, işletmeler için önemli bir veri kaynağı olan online müşteri yorumları veri havuzu olarak kullanılmış; metin

madenciliği, fikir madenciliği ve duygu analizi yöntemleriyle analiz edilmiştir. Analiz sonuçları kelime bulutları (Word cloud) ile görselleştirilmiştir.

Duygu analizi ve fikir madenciliğine ilişkin yapılan çalışmalar literatür kısmında incelenecek; sonraki bölümde kullanılan metodoloji ve bulgular ortaya konularak sonuçlar değerlendirilecektir.

4 LİTERATÜR

Fikir, duygu ve öznellik kavramları literatürde birbirinin yerine kullanılabilmektedir. Bilgisayarlarla bütünleşik çalışılması sonucu bu kavramların yerini fikir madenciliği, duygu analizi ya da öznellik analizi almıştır.[4] 1994'te Wiebe tarafından düşünce, algı ve duyguların anlaşılabilmesi adına bir algoritmanın geliştirildiği 'Hikayedeki Bakış Açısının İzlenmesi' başlıklı makale fikir madenciliği alanında ilk çalışmadır. 2003 yılında Budapeşte'de düzenlenen World Wide Web konferansında Dave ve ekibinin yaptığı çalışmayla fikir madenciliği kavramı öykü ve roman gibi düz metinler üzerinden inceleme yapılması boyutundan sıyrılmış ve web ortamında kullanılmaya başlamıştır. Bu çalışmada Dave ve arkadaşları ürünlere yönelik kullanıcı yorumlarını çeşitli sınıflandırma yöntemleriyle sınıflandırmış; özellik çıkarımı ve skorlama yaparak sonuçları karşılaştırmıştır. [5] Metinlerde duyguların tahmin edilmesi ve geleceğe yönelik çıkarımlar yapılması ve pazar duygusunun elde edilmesi için yapılan duygu analiz çalışmaları Das ve Chen (2001) ve Tong(2001) tarafından yapılmıştır. [6]

Fikir madenciliği çalışmaları üç kategoriye ayrılmaktadır; doküman seviyesi (Wilson, Wiebe & Hoffmann, 2005), cümle seviyesi (Meena & Prabhakar, 2007) ve bakış / özellik tabanlı seviyesi duygu analizi(Jo & Oh, 2011). Doküman düzeyinde fikir madenciliği çalışmasında analiz; doküman düzeyinde pozitif ve negatif olmak üzere iki kupta incelenirken, cümle düzeyinde fikir madenciliğinde bu kutupluluk cümle düzeyinde belirlenir. Pozitif ve negatif kutupluluk yorum sahibinin/yazarın düşüncesinin olumlu ya da olumsuz yansıması olarak ifade edilmektedir. Her iki düzeyde de yapılan analizler kişinin duygularını tam olarak yansıtmaz; neyi sevip sevmediğinin tam karşılığı değildir.[7]Bakış tabanlı duygu analizi yönteminde metindeki varlık ve olaylar tanımlanarak her biri için ayrı duygu çıkarımı yapılır. Bir görüş pozitif, negatif ve nötr olarak 3 sınıfta incelenir.[8] Özen 2021 yılında yaptığı çalışmasında bakış tabanlı duygu analizi yöntemi ile müşteri yorumlarını analiz etmiştir. Yapılan çalışmada "Servis çok kötü olmasına rağmen yemeği çok beğendim" cümlesinde 'servis' ve 'yemek' iki ayrı varlık olarak değerlendirilmekte ve analiz edilmektedir.[9] Ayrıca bu varlıklar için ayrı ayrı pozitif, negaif ve nötr olmak üzere fikir ve duygu çıkarımı yapılır.[10].Son zamanlarda kullanılan bu yöntem, doküman düzeyinde analiz çalışmalarında önemli bir yere sahiptir.[11]

Duygu sınıflandırma çalışması makine öğrenmesine dayalı yöntemler ve sözcük tabanlı yöntemler olmak üzere iki kategoriye ayrılmaktadır. Makine öğrenmesi yöntemi kullanılarak yapılan ilk duygu sınıflandırma çalışması 2002 yılında yapılmıştır. Pang ve arkadaşları yaptıkları çalışmada sinema filmlerinin yorumlarını ele almıştır. Yorumlar pozitif ve negatif olarak sınıflandırmışlardır. Yapılan çalışmada Kelime Torbası modeli olarak unigram kullanılmış, NB ve DVM ile sınıflandırma yapmışlardır. [12] Turney 2002 yılında otomobil ve sinema yorumları üzerine duygu analizi çalışması yapmış ve duygu analizi çalışmasında ilk defa nokta tabanlı karşılıklı bilgi skorlamasını kullanılmıştır.[13] Muller ve Collier ise NKB'yi ilk defa makine öğrenmesinde özellik seçimi olarak sinema yorumlarının duygu analizini

yapmak üzere kullanmıştır. Bu çalışmalar duygu analiz çalışmalarında NKB skarlama yöntemi kullanılmasına öncülük etmiştir.[14]

Sözcük tabanlı yöntemler ile duygu analizi çalışmalarından biri olan koşullu rastgele alanlar(KRA) yöntemi Lafferty ve arkadaşları tarafından 2001 yılında önerilmiş olan istatistiksel modelleme yöntemidir. Bu yöntem örüntü tanıma, yapay zeka, makine öğrenmesi alanlarında kullanılırken doğal dil işleme çalışmalarında da oldukça yaygın şekilde kullanılmaktadır.[15] KRA ile birlikte kullanılan kural tabanlı yaklaşım ise ilk defa 2002 yılında Turney ve arkadaşları tarafından müşteri yorumları üzerine yaptığı çalışmada kullanılmıştır.[16] 2015 yılında Maharani ve arkadaşları da ürün yorumları üzerine yaptıkları çalışmada sözcük türü örüntüleri ile ilgili algoritmalarından yararlanmışlardır.[17]

5 METODOLOJİ

Çalışmanın amacı Xiaomi 9 marka cep telefonu hakkında yapılan yorumların duygu analizinin gerçekleştirilmesi ve ürün hakkındaki fikrin açığa çıkarılmasıdır. e-alışveriş sitelerinden toplanan veriler metin tabanlı olduğu için metin madenciliği yöntemleri uygulanarak analiz edilebilecek formata dönüştürülür. Metin madenciliği çalışmalarında metinler veri kaynağı olarak kabul edilir ve amaç metin üzerinden yapılandırılmış veri elde etmektir. [18] Metin analitiği/ metin madenciliğinde büyük boyuttaki verilerin işlenerek metnin içinde yer alan ortak bilgiyi elde ederek analiz gerçekleştirilir. Metin analitiği, metinler arası benzerliği bulmak veya beklenmeyen durumları tespit etmek amacıyla kullanılan yöntemlerden oluşmaktadır. Veriyi işlemekteki amaç veriyi sınıflandırmak, kümelemek, metinden bilgi elde etmek metin üzerinde sentiment (duygu analizleri) yapmak, metni özetlemek veya elimizdeki metinden yeni metinler üretebilmektir. Bu amaçlar için uygulanan adımlar:

- Verinin toplanması
- Verinin içe aktarılması(Information Retrieval)
- Hece analizi (Lexical Analysis)
- Kelime frekans dağılımlarının hesaplanması(Word Frequency Distribution)
- Örüntü tanıma(Pattern Recognition)
- Bilgi çıkarımı(Data Mining)
- Görselleştirme(Visilution)

5.1 NEDEN MATLAB?

Veri analiz aşamalarına geçmeden önce çalışmada neden Matlab kullanılmıştır.

- Bir veya birden fazla metni HTML, Word, PDF, Excel formatlarında veriyi içeri aktarma imkanı sağlar.
- Aktarılan veriyi kolaylıkla görselleştirerek çalışmaya yön verme imkanı sağlar.
- Birçok fonksiyonu birden fazla dilde kullanılabilmektedir. Tüm fonksiyonlara İngilizce, Japonca, Almanca ve Korece başta olmak üzere tam destek sağlar.
- Matlab metin madenciliği, duygu analizi, Machine L çalışmaları için zengin araçlara sahiptir.
- Matlab'ın diğer görsellerin yanında metin analizi için önerdiği WordCloud ve Text Scotter Plot çizimleri ile veriyi analiz etmek mümkündür.

5.2 UYGULAMA ADIMLARI

5.2.1 Veri Toplama

Çalışmada e-alışveriş sitelerinden Xiaomi 9 marka cep telefonu hakkında yapılan yorumlar analiz edilmiştir. Veri çeşitliliğinin sağlanması amacıyla Amazon, Trendyol, Hepsiburada, gittigidiyor, n11, Vatan. Epey, Akakce, Cimri, çiçek sepeti olmak üzere 10 farklı online alışveriş ve karşılaştırma sitesinden 2372 müşteri yorumu kullanılarak veri havuzu oluşturulmuştur.

Veriler Excel’de toplanmıştır. Yorum sahibine ilişkin bilgiler, satıcı, kargo veya alışveriş sitesine ilişkin yorumlar çıkarılmış sadece cep telefonuna ilişkin yorumlar bırakılmıştır.

Önce Excel’e aktarılıp; salt cep telefonuna ilişkin verilere ulaşıldıktan sonra, işlenmemiş veriler Matlab’a aktarılmıştır.

5.2.2 Veri Ön İşleme

Veriyi içe aktardıktan sonra ham veri kelime bulutu olarak görselleştirilerek başlangıç analizi gerçekleştirilmiştir. Kelime bulutunda yer alan product, bought, phone, packaging, Thank gibi ürün hakkında bilgi vermeyen fakat sıkça kullanılmış olan kelimeler çıkarılmıştır. Veri havuzunda ayıklama işlemi yapıldıktan sonra veri, analiz için kullanılacak yöntemlere hazır hale getirilir. e-alışveriş sitelerinden toplanan veriler metin tabanlı olduğu için metin madenciliği yöntemleri uygulanarak analiz edilebilecek formata dönüştürülür. Uygulanan işlemler:

1.Tokenization(Tokanize the text)

Metinleri istenen özelliklere göre parçalama işlemidir. Kelime, karakter ve n gram olmak üzere 3 farklı şekilde simgeleştirme yapılabilir. İlk olarak verilere tokenization uygulanmıştır. Cümleler kelimelere parçalanmıştır.

2.Remove Stop Words

Tek başına anlam ifade etmeyen kelimeler çıkarılmıştır.

3.Normalize Words

Kelimeleri kök haline dönüştürme işlemi gerçekleştirilmiştir.

4.Noktalama işaretlerinin silinmesi

5.Büyük harfler küçük harflere dönüştürülmesi

6.Boşlukların silinmesi

7.3 harf ve 3 harften kısa kelimelerin silinmesi

8.10 harf ve 10 harften uzun kelimelerin silinmesi işlemleri uygulanmıştır.

5.2.3 Verilerin Analizi

Yorumlar N gram counting, topic model, sentiment analys, machine learning-sentiment analys olmak üzere 4 farklı yöntemle analiz edilmiştir.

5.2.3.1 N-Gram Counting

İstatistiksel doğal dil işlemede yaygın olarak kullanılmaktadır. Araştırma için kelimeler her n gram n kelimedenden oluşacak şekilde modellenmektedir.

Karakter seviyesi n-gram modelleri metin sıkıştırma yaygın olarak kullanılmaktadır.1990 yılında T.Bell ve Witten'in PPM modeli kullanılarak yaptığı çalışma buna örnek olarak verilebilir. [19] Son zamanlarda n-gram modellerinin metin madenciliği modellerinde de etkili olduğu görülmüştür. [20] N-gram modelleriyle metin kategorizasyonunu 1994'Te Cavnar ve Trenkle denemiştir.[21] 1994'te C.Apt'e ve Weiss içinde dilden bağımsız metin kategorizasyonunda N-gram modeli kullanmıştır ve çalışmada İngilizce ve Almanca olmak üzere kelime tabanlı dil kullanmışlardır.[22]

Neden N-Gram Kullanılmıştır?

N Gram modeller ardışık bir dizide bir sonraki elemanın bulunması için tahminlemede Markov modelinden yararlanan istatistiksel bir dildir. N-gram modelleri aynı amaç için kullanılan diğer algoritmalara göre ölçeklenebilir ve diğer modellere nazaran çok daha basittir.

N sayısı büyüdükçe bağlama dair daha fazla bilgi sağlamanın yanı sıra küçük n sayılar ile ufak ölçekli deney ve modellere de kolaylıkla uygulanabilir ve başarılı sonuçlar elde edilir. Ayrıca bir n gram modeli ile istatistiksel parametreler kullanarak doğal dilleri de modellemek mümkündür.

Ngram pekçok amaç için doğal dil işlemeye dair süreçlerde kullanılabilir. Özellikle telefonlarımızda mesaj yazarken kullanılan kelime tahmin özelliği ve otomatik düzeltme işleminde n gramdan yararlanılmaktadır.

N-gram modeller ayrıca imla düzeltmede de kullanılmaktadır. Sadece ASCII karakterleri kullanılarak yazılmış Türkçe bir metni ç,ı,ş,ü,ö,ğ ve i karakterleri eklemek yani ortografiye çevirmek için de n gram modellemeden yararlanılır. Örneğin elimizdeki metinde 'gol' sözcüğü geçmekte bu sözcüğün doğru yazılmış hali göl ya da gol olabilir. Algoritma, o karakterini ö'ye çevirmek ya da olduğu gibi bırakmak seçeneklerini değerlendirirken n gram modellerinden yararlanarak önceki ve sonra gelen sözcükleri değerlendirerek tahminleme yapar ve doğru kararı verir.

Bunların yanı sıra Ngram SVM, MaxEnt, Naive Bayes, modelleri gibi makine öğrenmesi modellerinde de kullanılmaktadır.

Metin analizinde metin içinde kelimenin kaç kez geçtiğini bilmek için bag of N-grams modeli kullanılmıştır. Kelime grubunun kaç kelimedenden oluşmasını istiyorsak o kelime sayısına özel algoritmalar oluşturulabilmektedir. Çalışmada ikili kelime gruplarına özel algoritma oluşturulmuştur.

5.2.3.2 Topic Model

- Konu modellerinde, bir doküman koleksiyonunda bulunup çıkartılan konular bir terim koleksiyonu olarak temsil edilirler.
- Metin dokümantasyonlarının büyük bir kısmında belli kalıpları /anlamsallığı ortaya çıkararak makine öğrenmesi ve doğal dil işlemede yer alan konuları keşfetmek için kullanılan metin madenciliği aracıdır.

Oluşturulan Bag Of N-Gram modelleriyle olasılık tabanlı bir yöntem olan gözetimsiz makine öğrenimi yöntemlerinden LDA yöntemiyle veri içerisinde temel alınan alt konu başlıkları analiz edilmiştir.

5.2.3.3 Sentiment Analysis

- Kullanıcıların sübjektif yorum ve değerlendirmelerinin taşıdığı anlam, duygu ve tutumların analiz edilmesidir.
- Fikir madenciliği olarak da bilinen duygu analizi ile belirli bir konuda sözlü ve yazılı metinlerin içerisindeki öznel dil öğeleri anlamlandırılabilir.
- Pozitif ve negatif önde gelen iki duygu sözlüğü mevcuttur.

Yapılan çalışmada veriler numerikleştirilerek pozitif ve negatif duygular açığa çıkarılır.

5.2.3.4 Machine Learning-Sentiment Analysis

Metni Machine L kullanarak sınıflandırma algortimalarının arka planında metinsel verinin sayısal veriye dönüştürülme işlemi vardır. Makine öğrenmesi algoritmalarının çalışabilmesi için metinsel veri sayısal matrislere dönüştürülmelidir. Metnin Machine L-Sentiment Analysis kullanılarak analiz edilmesi aşamaları:

- Makine öğrenmesi ve derin öğrenme için metin Word embedding kullanılarak numerikleştirilir.
- Duygu sınıflandırması yapmak için eğitilmiş sistem kullanılmakta ve sistem önceden etiketlenmiş eğitim verisi ile eğitilmektedir.
- Opinion Lexion English ile sistem eğitilmiş, pozitif ve negatif duygular sözlük üzerinden öğretilmiş ve duygu sınıflandırılması yapılır.

5.2.4 Verilerin Görselleştirilmesi

Veri madenciliğinin kollarından biri olan metin madenciliği metotları içinde kelime bulutları, metinlerde en sık kullanılan kelimeleri vurgulama ve görselleştirmeye olanak tanımaktadır. Metin verisinden metin veya etiket bulutu denebilecek kelime bulutları oluşturulabilir. Tablo olarak görselleştirmek yerine kelime bulutları kullanılmasının sebeplerini şu şekilde sıralayabiliriz:

- Metin daha hızlı analiz etmeyi sağlar.
- Kelime bulutu olarak ortaya çıkan anahtar kelimeleri görselleştirmeye olanak tanır.

Veriye ön işlem yapıldıktan sonra good, very ve price kelimelerinin ön plana çıktığı görülmektedir.

Veri bag of N grams modeliyle analiz edilmiştir. [2371x5432 double] sayım gerçekleşmiş, [1x1323] kelime bilgisi bulunmuş, Ngram uzunluğu 2 olan 5432 Ngrams ‘a ayrılmıştır. Bag of N grams özellikleri Tablo1’de, oluşturulan kelime bulutu özellikleri Tablo 2’de yer almaktadır.

COUNTS	[2371x5432 double]
VOCABULARY	[1x1323 string]
N GRAMS	[5432x2 string]
N GRAM LENGTHS	2
NUM N GRAMS	5432
NUM DOCUMENTS	2371

Tablo 1 Bag Of N Grams Özellikleri

WORD DATA	[1x5432 string]
SIZE DATA	[1x5432 double]
MAX DISPLAY WORDS	100

Tablo 2 Bag Of N Grams’a Ait Word Cloud Özellikleri

Veri Bag of N Grams modeliyle analiz edildikten sonra oluşan kelime bulutu Şekil 3’te yer almaktadır. Bag of N grams modeliyle analiz edildikten sonra Şekil 3’te yer alan kelime bulutundan da görüldüğü üzere very fast, very nice, very good kelimeleri ön plana çıkmaktadır. Ürünle ilgili en çok çok hızlı, çok hoş, çok iyi yorumlarında bulunulmuştur.



Şekil 3 Bag of N Grams’a Ait Kelime Bulutu

Gözetimsiz makine öğrenimi yöntemlerinden olan LDA yöntemiyle 10 alt başlık oluşturulmuş bunların 4’ü Şekil 4’te yer aldığı üzere kelime bulutu olarak görselleştirilmiştir. LDA modeline ait özellikler Tablo 3’te yer almaktadır.



Şekil 4 LDA Modeline Ait Kelime Bulutu

NumTopics	10
WordConcentration	1
TopicConcentration	0.2691
CorpusTopicProbabilities	[0.1695 0.1111 0.1047 0.1041 0.1018 0.1007 0.0915 0.0860 0.0809 0.0498]
DocumentTopicProbabilities	[2371×10 double]
TopicWordProbabilities	[5432×10 double]
Vocabulary	[1×5432 string]
TopicOrder	initial-fit-probability
FitInfo	[1×1 struct]

Tablo 3 LDA Modeli Özellikleri

LDA topic 1’de very well, very nice, very good, the best, the price kelimeleri ön plana çıkmaktadır. Ürünle ilgili yapılan yorumlarda LDA yöntemiyle açığa çıkan kelimeler bize ürün hakkında en iyi, çok hoş, çok iyi yorumlarının ağırlıkta olduğunu göstermektedir. 2. Alt başlıkta price perform, charge go, go long, long time kelimeleri baskın olarak görülmektedir. Bu alt başlıkta oluşan kelimeler bize ürünle ilgili fiyat performans ürünü olduğu, şarjının uzun süre gittiği yorumlarının sıklıkla yapıldığını göstermektedir. 3.alt başlıkta very nice, very good, very fast kelimeleriyle ürünün çok hızlı ve çok iyi/hoş olduğu yorumları ön plana çıkmaktadır. 4.başlıkta very satisfy, feature very, very nice, very fast, fast ship kelimeleriyle ürünün çok tatmin edici, özelliklerinin çok olduğu, çok hızlı, teslimatın hızlı yapıldığı yorumları ön plana çıkmaktadır.

Duygu analizinde önce belgeye tokenize işlemi uygulanmıştır. İlk beş token (belirteç/simge) Tablo 4’te yer almaktadır.

3 tokens	best the price
3 tokens	hip reliable eller
2 tokens	recommend nice
2 tokens	excel price
1 tokens	Success

Tablo 4 5x1 Tokenized Dökümanına Ait Tablo

VaderSentimentLexion fonksiyonu kullanılarak belirteçli belgelerin duyarlılığı değerlendirildi. İlk birkaç belgeye ait puanlama

0.6369

0

0.6486

0.4588

0.5719

şeklinde. 1'e yakın puanlar olumlu duyguları, -1'e yakın puanlar olumsuz duyguları göstermektedir. Metnin kelime bulutlarında olumlu ve olumsuz duygularla görselleştirilmiş hali Şekil 5'te yer almaktadır. Açığa çıkan pozitif duygulara ait kelime bulutunda camera, nice, good, price, days kelimelerinin daha baskın olduğu görülmektedir. Bu kelimelerden nice ve good ürüne dair iyi, güzel, hoş olduğu duyguları hakkında bilgi vermektedir. Yapılan bazı çalışmalar sıklıkla kullanılan kelimeler ve duyguların birbirinin tamamlayıcısı olduğu varsayımında bulunmaktadır. Örneğin bu kelime bulutunda olduğu üzere kamerasının iyi, fiyatının güzel olduğu varsayımlarında bulunmak gibi. Yine pozitif kelime bulutunda diğerlerine kıyasla daha baskın kelimelere göz atacak olursak quality, great, liked, quickly, fast, satisfied kelimelerinin ön plana çıktığı görülmektedir. Ürünün kaliteli, hızlı, tatmin edici, sevildiği fikir ve duyguları açığa çıkmıştır. Negatif kelime bulutunda ordered, days, price, fast, arrived kelimeleri daha baskın olarak görüntülenmiştir. Negatif kelime bulutu incelendiğinde açığa çıkan kelimelerden big kelimesi telefonun büyük olduğuna dair oluşan negatif bir duygu olmakla birlikte diğer kelimeler ürüne dair olumsuz duygular hakkında bilgi vermemektedir. Affordable, useful, beautiful gibi ürünün uygun fiyatlı, güzel, kullanışlı olduğunu ifade eden kelimeler de negatif kelime bulutunda yer almaktadır. Negatif kelime bulutundan elde edilen sonuçlardan iki farklı çıkarımda bulunabiliriz; bunlardan ilki ürün hakkında olumsuz duygu ve fikirlerin yer almadığı ve üründen tam anlamıyla memnun kalındığı, ikincisi negatif duygu analizi sonuçlarından ürünle ilgili olumsuz duygu ve düşüncelere ulaşamadığı sonuçlarına varabiliriz.



Şekil 5 Duygu Analizi Yöntemine Ait Kelime Bulutları

7 SONUÇ

XİAOMİ 9 marka cep telefonuna dair 10 farklı online alışveriş ve karşılaştırma sitesinden aktarılan 2372 müşteri yorumu, duygu analizi ve metin madenciliği yöntemleriyle analiz edilmiştir. Analiz sonuçları kelime bulutu olarak görselleştirilmiş ve ürünle ilgili fikir ve duygular açığa çıkarılmıştır. Yapılan analiz sonucu ürünün fiyat performans ürünü olduğu, hızlı, kaliteli, uygun fiyatlı, çok özellikli, çok iyi, çok hoş, kamerasının iyi olduğu, şarjının uzun süreli olduğu fikir ve duyguları açığa çıkmıştır. Bunların yanı sıra ürünün büyük bulunması dışında ürünle ilgili olumsuz duygu ve düşünce açığa çıkmamıştır.

Yapılan çalışma piyasaya sürülen ürünlerin müşteri beklentilerini karşılayıp karşılamadığı, sevilip sevilmediğinin tespiti adına önem teşkil etmektedir. Yorumlara dair analizler mevcut ürünle ilgili değerlendirmede **önemli bir yere sahip olmasının yanı sıra** yeni ürün gamı oluşturulurken ürün özellikleri ve spesifikasyonların belirlenmesi için kullanılabilir. Ürünle ilgili müşteriler için önemli olan noktaların belirlenmesi ve müşterilerin tanınması adına açığa çıkan bilgiler değerlendirilebilir. Online müşteri yorumlarından açığa çıkan fikir ve duyguları; reklam ve pazarlama stratejilerinden, ürün geliştirmeye, Müşteri ilişkileri Yönetiminden Talep yönetimine birçok üretim planlama ve kontrol sistemi bileşenlerinde ele alıp değerlendirmek mümkündür.

