

T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
SAYISAL YÖNTEMLER BİLİM DALI

YÜKSEK LİSANS TEZİ

SOSYAL MEDYA MADENCİLİĞİ VE BİR
UYGULAMA

Büşra AYAN

2501150586

TEZ DANIŞMANI
Dr. Öğr. Üyesi Mustafa CAN

İSTANBUL – 2018



T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



YÜKSEK LİSANS
TEZ ONAYI

ÖĞRENCİNİN;

Adı ve Soyadı : BÜŞRA AYAN Numarası : 2501150586
Anabilim Dalı /
Anasanat Dalı / Programı : SAYISAL YÖNTEMLER Danışmanı ☐ DR.ÖĞR.ÜYESİ MUSTAFA CAN
Tez Savunma Tarihi : 03.07.2018 Saati : 10.30
Tez Başlığı : SOSYAL MEDYA MADENCİLİĞİ VE BİR UYGULAMA.

TEZ SAVUNMA SINAVI, İÜ Lisansüstü Eğitim-Öğretim Yönetmeliği'nin 36. Maddesi uyarınca yapılmış,
sorulan sorulara alınan cevaplar sonunda adayın tezinin KABULÜNE OYBİRLİĞİ / OYÇOKLUĞUYLA karar verilmiştir.

JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
PROF.DR.TUĞBA ŞİMŞEK		Kabul
DR.ÖĞR.ÜYESİ MUSTAFA CAN		Kabul
DR.ÖĞR.ÜYESİ AYŞE DEMİRHAN		Kabul

YEDEK JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
DOÇ.DR.NİHAT TAŞ		
DOÇ.DR.CEREN ERDİN GÜNDÖĞDÜ		

ÖZ
SOSYAL MEDYA MADENCİLİĞİ VE BİR UYGULAMA
BÜŞRA AYAN

Bu çalışma, farklı sektörlerde faaliyet gösteren rakip firmaların Twitter verilerini analiz ederek, firmaların Twitter değişkenlerinin firmalara göre anlamlı farklılık ve uyum gösterip göstermediğinin tespit edilmesini, firmaların Twitter’da paylaştıkları içeriklerin kümelenmesini ve hangi içerik kümesinin en fazla etkileşime yol açtığı belirlenmesini amaçlamaktadır. Bu çalışmada ayrıca, özel bir döneme ait verilerden hareketle firmaların Twitter etkinliklerinin (performanslarının) ölçülmesi hedeflenmiştir. Bu kapsamda, Sosyal Medya Madenciliği süreci izlenerek iki bölümden oluşan bir uygulama gerçekleştirilmiştir. Birinci bölümde firmaların 2017 yılına ait Twitter verileri elde edilerek, firmaların Twitter değişkenlerinin firmalara göre değişkenlik gösterip göstermediği Ki-Kare, uyum gösterip göstermediği ise Uygunluk Analizi uygulanarak saptanmıştır. Firmaların tweetlerinin kümeleri bir dizi Metin Madenciliği ön işleme metotları yardımıyla ve hangi içerik kümelerinin daha fazla etkileşim aldığı Pareto İlkesi kapsamında ele alınarak analiz edilmiştir. İkinci bölümde ise, 2018 Şubat ayı boyunca kullanıcılar tarafından firmalar hakkında paylaşılan Twitter verileri ile firmaların bu döneme ait Twitter verileri ele alınarak, firmaların Twitter etkinliklerinin ölçülmesi gerçekleştirilmiştir. Twitter etkinlik ölçüm kriterleri arasında yer alan tweet değerini saptamak amacıyla, kullanıcılar tarafından paylaşılan tweetlerde Duygu Analizi gerçekleştirilmiş ve firmalar hakkında paylaşılan tweetlerin duygu durumları olumlu, nötr ya da olumsuz olmak üzere sınıflandırılmıştır.

Anahtar Kelimeler: sosyal medya, sosyal medya verileri, sosyal medya madenciliği, twitter madenciliği, metin madenciliği

ABSTRACT
SOCIAL MEDIA MINING AND AN APPLICATION
BÜŞRA AYAN

This study aims to determine whether Twitter variables of the firms have significant differences and correspondence with respect to the firms, to cluster the Twitter feeds that firms share on Twitter and to find out which cluster has the maximum interaction through analyzing the Twitter data of the rival firms operating in different sectors. In this study, it is also aimed to measure the effectiveness (performance) of the firms on Twitter during a specific time period. In this context, an application that consists of two parts was carried out by following the process of Social Media Mining. In the first part, the Twitter data of the firms for the year 2017 were gathered, Chi-Square and Correspondence Analysis were applied to find if there were significant differences and correspondence between Twitter variables and the firms. The clusters of the tweets that firms share on Twitter and which cluster received more interaction were analyzed with the help of a number of Text Mining pre-processing methods and the Pareto Principle, respectively. In the second part, Twitter data about the firms shared by the users and the firms' Twitter data were captured during February 2018 to measure the effectiveness of the firms on Twitter. In order to determine the tweet value among the Twitter effectiveness' measurement criteria, Sentiment Analysis was carried out on the tweets shared by the users and the sentiment of the tweets about the firms was classified as positive, neutral or negative.

Keywords: social media, social media data, social media mining, twitter mining, text mining

ÖNSÖZ

Sosyal medya madenciliği son yıllarda gelişmekte olan, disiplinler arası metotlar yardımı ile sosyal medya verilerinin analizini mümkün kılan bir alandır. Konunun güncel ve yeni şekillenmeye başlıyor oluşu ile birlikte özellikle Türkçe literatürdeki çalışmaların yetersizliği, bu tez çalışmasının sosyal medya madenciliği alanında gerçekleştirilmesinde motive edici olmuştur. Çalışmada farklı sektörlerde faaliyet gösteren rakip firmaların Twitter kullanımlarına ait çıkarımlarda bulunabilmek ve belirli bir döneme ait Twitter verilerinden faydalanılarak firmaların Twitter etkinliklerini ölçümleyebilmek amacı ile sosyal medya madenciliği adımları izlenerek bir uygulama gerçekleştirilmiştir.

Tez sürecinde yönlendirici önerileri ve yorumları ile bana yol gösteren tez danışman hocam Dr. Öğr. Üyesi Mustafa CAN'a, tez konumun belirlenmesine ve diğer süreçlere olan katkısı ile her zaman desteğini hissettiğim hocam Prof. Dr. Umman Tuğba GÜRSOY ŞİMŞEK'e, değerli fikirleri ile beni yönlendiren hocam Doç. Dr. Semen SON TURAN'a, veri çekim aşamasındaki katkılarından dolayı Sayın Adem ÖZAY'a ve Semih PARTAL'a sonsuz teşekkürlerimi sunuyorum.

Eğitim hayatım boyunca göstermiş oldukları destek ve hoşgörülerini için babam Kenan AYAN'a, annem Ayşe AYAN'a ve kardeşlerime ne kadar teşekkür etsem azdır. Tez süreci boyunca yanımda olan ve desteklerini esirgemeyerek teze katkıda bulunan arkadaşlarım Gamze EROĞLU'na, Sinem ÇABUK'a ve Elif GÜNER'e teşekkürü bir borç bilirim.

Büşra AYAN
İSTANBUL, 2018

İÇİNDEKİLER

ÖZ	iii
ABSTRACT.....	iv
ÖNSÖZ	v
İÇİNDEKİLER	vi
TABLolar LİSTESİ	x
ŞEKİLLER LİSTESİ	xii
KISALTMALAR LİSTESİ.....	xv
GİRİŞ	1

BİRİNCİ BÖLÜM

SOSYAL MEDYA VE TEMEL KAVRAMLAR

1.1. İnternetin Temelleri ve Gelişimi	3
1.2. Web Evreleri	6
1.2.1. Web 1.0.....	6
1.2.2. Web 2.0.....	8
1.2.2.1. Kullanıcıların Oluşturduğu İçerik.....	11
1.2.3. Web 3.0.....	15
1.2.4. Web 4.0 ve Ötesi.....	17
1.3. Sosyal Medya	19
1.3.1. Sosyal Medyanın Tarihçesi.....	20
1.3.2. Sosyal Medya Kavramı.....	21
1.3.3. Sosyal Medyanın Özellikleri.....	25
1.3.4. Sosyal Medyanın Temel Yapıtaşları.....	27
1.4. Yeni Bir Veri Kaynağı Olarak Sosyal Medya.....	31
1.4.1. Sosyal Medya Verileri	31

1.4.1.1.	Yapılarına Göre Sosyal Medya Verileri	32
1.4.1.2.	Özelliklerine Göre Sosyal Medya Verileri	33
1.4.1.3.	İçeriklerine Göre Sosyal Medya Verileri	36
1.4.2.	Sosyal Medya Veri Dağıtım Platformları	39
1.4.2.1.	Sosyal Ağlar	41
1.4.2.2.	İçerik Paylaşım Toplulukları	42
1.4.2.3.	İşbirlikçi Platformlar	43
1.4.2.4.	Sosyal Yayıncılık Platformları	44
1.4.2.5.	Podcast	45
1.4.2.6.	RSS	46
1.4.2.7.	Sanal Dünyalar	47

İKİNCİ BÖLÜM

SOSYAL MEDYA MADENCİLİĞİ

2.1.	Sosyal Medya Madenciliği Kavramı	48
2.2.	Sosyal Medya Madenciliğinde Karşılaşılan Zorluklar	50
2.3.	Literatür Taraması	52
2.4.	Sosyal Medya Madenciliği Süreci	65
2.4.1.	Yetkilendirme	66
2.4.2.	Veri Toplama	67
2.4.3.	Veri Ön İşleme	68
2.4.4.	Modelleme ve Analiz	69
2.4.5.	Değerlendirme	70
2.5.	Sosyal Medya Madenciliğinde Kullanılan Yöntemler	71
2.5.1.	Veri Madenciliği	71
2.5.2.	Metin Madenciliği	73
2.5.2.1.	Ön İşleme	76
2.5.2.2.	Gösterim	78
2.5.2.3.	Bilgi Keşfi	79

2.5.2.3.1. Sınıflandırma.....	79
2.5.2.3.2. Kümeleme.....	81
2.5.2.3.3. Birliklilik Analizi	82
2.5.3. Duygu Analizi.....	83

ÜÇÜNCÜ BÖLÜM

UYGULAMA

3.1. Araştırmanın Amacı	87
3.2. Araştırmanın Örnekleme.....	88
3.3. Araştırmanın Yöntemi.....	91
3.4. Sosyal Medya Madenciliği Uygulaması	93
3.4.1. Firmaların Twitter Kullanımlarına Ait Değerlendirmeler	93
3.4.1.1. Yetkilendirme	93
3.4.1.1.1. Aplikasyon Oluşturma	93
3.4.1.1.2. Yetkilendirme İşlemi	94
3.4.1.2. Veri Toplama	96
3.4.1.3. Veri Ön İşleme.....	98
3.4.1.3.1. Veri İndirgeme.....	99
3.4.1.3.2. Veri Seti	99
3.4.1.4. Modelleme ve Analiz.....	100
3.4.1.4.1. Tanımlayıcı İstatistikler	103
3.4.1.4.2. Ki-Kare Analizi.....	108
3.4.1.4.3. Uygunluk Analizi.....	112
3.4.1.4.4. İçerik Analizi	123
3.4.1.4.4.1. Tweet Analizi.....	123
3.4.1.4.4.2. Rt Analizi.....	131
3.4.1.4.4.3. Hashtag Analizi.....	134
3.4.1.5. Değerlendirme.....	136
3.4.2. Firmaların Twitter Etkinliklerinin Ölçümü.....	137

3.4.2.1. Yetkilendirme	138
3.4.2.2. Veri Toplama	138
3.4.2.3. Veri Ön İşleme	139
3.4.2.3.1. Veri İndirgeme	139
3.4.2.3.2. Veri Temizleme	139
3.4.2.3.3. Veri Seti	140
3.4.2.4. Modelleme ve Analiz	141
3.4.2.4.1. Tanımlayıcı İstatistikler	144
3.4.2.4.2. Duygu Analizi	151
3.4.2.5. Değerlendirme	154
SONUÇ	161
KAYNAKÇA	166
EKLER	184

TABLolar LİSTESİ

Tablo 1.1: Web 1.0 ve Web 2.0 Karşılaştırması	8
Tablo 1.2: Kullanıcıların Oluşturduğu İçeriği Destekleyen Etkenler	14
Tablo 1.3: Sosyal Medya Verilerinin İçeriklerine Göre Sınıflandırılması	38
Tablo 1.4: Sosyal Medya Veri Dağıtım Platformları Sınıflandırmaları	39
Tablo 2.1: Metin Madenciliği Ön İşleme Adımları	76
Tablo 2.2: Karışıklık Matrisi.....	80
Tablo 2.3: Kappa İstatistiği Değer Aralıkları ve Yorumları	85
Tablo 2.4: Krippendorff Alfa Katsayısı Değer Aralıkları ve Yorumları	86
Tablo 3.1: Twitter Terim İsimleri ve Açıklamaları	88
Tablo 3.2: Uygulamanın Örneklemini Oluşturan Sektör ve Firma Bilgileri	90
Tablo 3.3: Uygulamada İzlenen Adımlar.....	92
Tablo 3.4: Uygulamanın Birinci Bölümüne Ait Değişken Tablosu	100
Tablo 3.5: Firmalara Göre Firma Tweet Türleri Ki-Kare Analizi Sonuçları.....	109
Tablo 3.6: Firmalara Göre Firma Tweet Paylaşım Günleri Ki-Kare Analizi Sonuçları	110
Tablo 3.7: Firmalara Göre Firma Tweet Zaman Dilimleri Ki-Kare Analizi Sonuçları	111
Tablo 3.8: Uygunluk Analizi Satır ve Sütun Profillerinin Hesaplanması	113
Tablo 3.9: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri	114
Tablo 3.10: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı	115
Tablo 3.11: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri	117
Tablo 3.12: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı	118

Tablo 3.13: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri	120
Tablo 3.14: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı	121
Tablo 3.15: Kümelere Göre En Sık Kullanılan Kelimeler ve Terim Frekansları	126
Tablo 3.16: Firmalar Tarafından Paylaşılan Tweetlerin Kümeleri ve Açıklamaları	126
Tablo 3.17: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekanslarının Pareto İlkesi Kapsamında Dağılım Yüzdeleri	133
Tablo 3.18: Firmaların Twitter Kullanımlarına Ait Değerlendirmelerin Özet Bulguları	136
Tablo 3.19: 2018 Şubat Ayı Boyunca Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Frekans Dağılımları.....	140
Tablo 3.20: Uygulamanın İkinci Bölümüne Ait Değişken Tablosu	140
Tablo 3.21: Farklı Amaçlar Doğrultusunda Twitter Etkinliğinin Ölçüm Kriterleri	142
Tablo 3.22: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Kaynak Bilgisine Göre Frekans Dağılımları	151
Tablo 3.23: Araştırmacılar Tarafından Gerçekleştirilen Duygu Durum Kodlamalarının Krippendorff Alfa Katsayısıyla Hesaplanan Güvenilirlik Değerleri	152
Tablo 3.24: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Duygu Durumlarının Firmalara Göre Frekans Dağılım Yüzdeleri.....	153

ŞEKİLLER LİSTESİ

Şekil 1.1: 2006 Aralık Sayısı Time Dergisi Kapağı	12
Şekil 1.2: Web'in Gelişim Evreleri.....	18
Şekil 1.3: Sosyal Medyanın Temel Yapıtaşları.....	28
Şekil 1.4: İyi, Kötü ve Çirkin Bağlamında Sosyal Medya Verileri	34
Şekil 2.1: Sosyal Medya Madenciliği Süreci.....	66
Şekil 2.2: Veri Madenciliği Süreci.....	72
Şekil 2.3: Metin Madenciliği Süreci	75
Şekil 3.1: Twitter Veri Çekimi- Twitter Aplikasyonu Oluşturma	94
Şekil 3.2: Twitter Veri Çekimi- RapidMiner Yetkilendirme İşlemi.....	95
Şekil 3.3: Twitter Veri Çekimi- RapidMiner Yetkilendirme İzin Sayfası.....	96
Şekil 3.4: RapidMiner Firmalara Ait Veri Çekim Süreci	97
Şekil 3.5: Firma 1'e Ait RapidMiner Twitter Sorgu Sonuç Ekranı Örneği.....	97
Şekil 3.6: RapidMiner Veri İndirgeme Süreç Modeli.....	99
Şekil 3.7: Uygulamanın Birinci Bölümüne Ait Analiz Adımları	101
Şekil 3.8: Firmalar Tarafından Paylaşılan Tweetlerin Frekans Dağılımları	103
Şekil 3.9: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Frekans Dağılımları	104
Şekil 3.10: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekans Dağılımları	105
Şekil 3.11: Firmalar Tarafından Paylaşılan Hashtag İçeren Tweetlerin Frekans Dağılımları	106
Şekil 3.12: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimi Frekans Dağılımları	107
Şekil 3.13: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Frekans Dağılımları	108
Şekil 3.14: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi	116

Şekil 3.15: Firma Tweetlerinin Paylaşım Günleri ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi	119
Şekil 3.16: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi.....	122
Şekil 3.17: Tweet Analizi RapidMiner Süreç Modeli	124
Şekil 3.18: Tweet Analizi Metin Ön İşleme RapidMiner Süreç Modeli	125
Şekil 3.19: Firmalar Tarafından Paylaşılan Tweetlerin Kümelere Göre Frekans Dağılımları	128
Şekil 3.20: Firmalar Tarafından Paylaşılan Tweetlerin Kümelerinin Firmalara Göre Frekans Dağılım Yüzdeleri	129
Şekil 3.21: Firmalar Tarafından Paylaşılan Tweetlerin Kümelere Göre Frekans Dağılımları ve Almış Oldukları Rt Sayıları	131
Şekil 3.22: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekanslarının Pareto İlkesi Kapsamında Firma Tweet Kümelerine Göre Dağılım Yüzdeleri	133
Şekil 3.23: Firmalar Tarafından Paylaşılan Tweetlerin İçerdiği Tekil ve Toplam Hashtag Frekans Dağılımları	135
Şekil 3.24: Şubat Ayı Boyunca Firmalar Tarafından Paylaşılan Tweetlerin Frekans Dağılımları	144
Şekil 3.25: Şubat Ayı Boyunca Firmalar Tarafından Paylaşılan Tweetlerin Almış Oldukları Rt Frekans Dağılımları	145
Şekil 3.26: Kullanıcılar Tarafından Kozmetik Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)	146
Şekil 3.27: Kullanıcılar Tarafından Pazaryeri Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)	148
Şekil 3.28: Kullanıcılar Tarafından Elektronik Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)	149
Şekil 3.29: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Zaman Dilimlerine Göre Frekans Dağılımları	150
Şekil 3.30: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Firmalara Göre Frekans Dağılımları	154

Şekil 3.31: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Olumlu Tweetlerin Firmalara Göre Frekans Dağılımları.....	155
Şekil 3.32: Firmaların Şubat Ayı Takipçi Kazanımları	156
Şekil 3.33: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Yanıt Tweetlerinin Firmalara Göre Frekans Dağılımları.....	157
Şekil 3.34: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Toplamda Aldığı Rt Frekanslarının Firmalara Göre Dağılımları	158
Şekil 3.35: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Toplamda Aldığı Beğenilerin Firmalara Göre Frekans Dağılımları	159
Şekil 3.36: Firma Twitter Etkinliklerinin Kriterlere Göre Toplam Performans Sıralaması.....	160

KISALTMALAR LİSTESİ

API	:	Uygulama Programlama Arayüzü
CRISP-DM	:	Çapraz Endüstri Veri Madenciliği Standart Süreci
IDF	:	Ters Metin Frekansı
OAuth	:	Open Authorization
ReCal OIR	:	Reliability Calculator for Ordinal, Interval, and Ratio Data
Rt	:	Retweet
TF	:	Terim Frekansı
TF-IDF	:	Terim Frekansı- Ters Metin Frekansı
UGC	:	Kullanıcıların Oluşturduğu İçerik
VTBK	:	Veri Tabanlarında Bilgi Keşfi

GİRİŞ

Sosyal medya, bireylere özgür ve etkileşimli bir ortam sunmaktadır. Sosyal medya kullanıcıları çeşitli platformlar aracılığı ile duygu ve düşüncelerini aktarmakla beraber başka kullanıcılar tarafından destek görebilmektedir. Bilgi ve fikirler daha önce hiçbir medya ortamında olmadığı kadar hızlı yayılmakta ve hatta sanal ortamdan gerçek dünyaya taşınarak eylemlere dönüşebilmektedir. Büyük yankı uyandırabilen bu medya ortamında üretilen verilerden günümüzde birçok alanda faydalanılmaktadır.

Bireyler, kullandıkları sosyal medya platformuna göre farklılık göstermekle birlikte, fotoğraf, video, konum bilgisi vb. gibi çeşitli türlerde veriler üreterek büyük verinin oluşumunda rol oynamaktadır. Kullanıcılar tarafından üretilen bu verilerin işlenmesi, son dönemlerde farklı disiplinlerdeki araştırmacıların dikkatini çekmeye başlamıştır. Sosyal medya madenciliği olarak adlandırılan yeni bir alan gelişmekte ve müşteri ilişkileri yönetimi, halkla ilişkiler, ürün yönetimi ve reklamcılık gibi çeşitli alanlarda uygulamaları görülmektedir.

Sosyal medya platformlarından çeşitli araçlar ve arayüzler yardımı ile verinin elde edilerek analize hazırlanması ve disiplinler arası metotlar yardımı ile analiz edilerek analiz bulgularının değerlendirilmesi süreçlerinden oluşan sosyal medya madenciliği, birçok araştırmacı tarafından farklı alanlarda uygulamaların gerçekleştirilmesine olanak tanıyan bir tekniktir. Sosyal medyanın gelişmesi ve sosyal medya platformların çeşitlenmesi ile birlikte bu medya ortamında üretilen ve paylaşılan verilerinin analizi gün geçtikçe daha da önem kazanmaktadır.

Tezin birinci bölümünde öncelikle sosyal medyanın teknolojik boyutunu oluşturan İnternetin gelişimine değinilerek Web sürümleri açıklanmıştır. Sosyal medya kavramı ile birlikte sıkça anılan Web 2.0 ve kullanıcıların oluşturduğu içerik (UGC) kavramına detaylı bir şekilde yer verilmiştir. Sosyal medyanın gelişimi, sosyal medya kavramı ile ilgili çeşitli tanımlamalar, sosyal medyanın özellikleri ve temel yapıtaşları sosyal medyaya detaylı bir şekilde değinilmek amacıyla açıklanmıştır. Sosyal medya

madenciliği tekniğinin veri kaynağı olan sosyal medya verileri yapıları, özellikleri ve içerikleri göz önünde bulundurularak ele alınmıştır. Son olarak sosyal medya verilerinin üretildiği sosyal medya platformları ile ilgili çeşitli sınıflandırmalara ve bu sınıflandırmalar doğrultusunda platform açıklamalarına yer verilmiştir.

Tezin ikinci bölümünde ise sosyal medya madenciliği kavramı ve bu madencilik sürecinde karşılaşılan zorluklar açıklanmıştır. Daha sonra sosyal medya madenciliği ile ilgili yapılan çalışmalara, kapsamlı bir literatür taraması bölümünde yer verilmiştir. Uygulamada başvuru olan sosyal medya madenciliği süreç adımları açıklandıktan sonra, bu alanla ilgili yapılan çalışmalarda sıklıkla başvuru olan veri madenciliği, metin madenciliği ve duygu analizi yöntemlerine değinilmiştir.

Tezin üçüncü bölümünde çalışmanın amacı, örnekleme ve yöntemi verildikten sonra sosyal medya madenciliği süreci izlenerek iki bölümden oluşan bir uygulama gerçekleştirilmiştir. Sonuç bölümünde ise uygulamada başvuru olan analizlerin bulgularına yer verilmiştir.

BİRİNCİ BÖLÜM

SOSYAL MEDYA VE TEMEL KAVRAMLAR

İnternetin gelişimi, beraberinde birçok alanda değişime zemin hazırlamıştır. Özellikle Web 2.0 teknolojisi temelinde gelişen sosyal medya kavramı, geleneksel medyanın alışılmış tek taraflı paylaşım yapısına tezat oluşu nedeniyle insanlar tarafından kolayca benimsenmiş, bu alanda yapılan uygulamalar ve elde edilen etkileşimlerle gücü yadsınamaz bir hale gelmiştir.

Sosyal medya kavramı ile ilgili olarak literatürde araştırmacılar tarafından yapılan çeşitli tanımlamalar mevcuttur. Tanımlamalar göz önünde bulundurulduğunda, sosyal medya kavramı açıklamalarında sıkça değinilen ve bazen de bu kavramın yerine kullanılan terimlerin söz konusu olduğu görülmektedir. Bu bölümde sosyal medya kavramı ile ilgili tanımlamalara yer verilmeden önce sosyal medyanın teknolojik boyutu, Web sürümlerine değinilerek açıklanmıştır. Ardından sosyal medyanın tarihçesine, kavram tanımlarına, özelliklerine, temel yapıtaşlarına, verilerine ve verilerinin dağıtım platformlarına yer verilmiştir.

1.1 İnternetin Temelleri ve Gelişimi

Yeni iletişim teknolojilerinin temeli olan İnternetin gelişim tarihçesine bakıldığında teknolojik gelişmeler, yönetim ve operasyon ihtiyaçları, toplumsal etkiler ve ticarileşme olmak üzere dört önemli başlık göze çarpmaktadır (Leiner v.d., 2009:23):

- **Teknolojik gelişmeler:** Paket anahtarlama araştırmaları, ARPANET ve altyapı boyutlarının geliştirilmesine yönelik yapılan çalışmalar İnternetin teknolojik gelişim boyutunu oluşturmaktadır.
- **Yönetim ve operasyon ihtiyaçları:** Yönetim ve operasyon temelinde küresel ve karmaşık bir altyapıya ihtiyaç duyulması İnternetin gelişmesinde rol oynamıştır.

- **Toplumsal etkiler:** Toplum tarafından büyük bir haberleşme ağının kurulmasına ve teknolojik gelişmelerin sağlanmasına duyulan gereksinimler toplumsal boyutu oluşturmaktadır.
- **Ticarileştirme:** Araştırma sonuçlarının geniş bir bilgi altyapısı içinde etkili bir şekilde paylaşımını sağlamak amacı ile oluşan ticari boyuttur.

Sputnik 1'in uzaya gönderilmesi ve Küba Füze Krizi temelinde tüm bu gelişmeleri tetikleyen iki öncül uluslararası olay olarak kabul edilmektedir (Green, 2010:20). Bu bağlamda ele alındığında İnternetin kökeni 1957 yılında Sovyetler Birliği'nin Sputnik 1 uydu sistemini fırlatmasına dayanmaktadır. Amerika Birleşik Devletleri ile Sovyetler Birliği arasındaki uzay yarışı (space race) bu şekilde başlamıştır (Dickey ve Lewis, 2011:2).

Sovyetler Birliği tarafından dünyanın yörüngesine yerleştirilen Sputnik'in fırlatılışı bilim ve mühendislik alanlarında gösterilen gelişimin bir simgesi olarak görülmüştür. Sputnik'in başarıyla yörüngeye oturmasını sağlayan R-7 roketinin nükleer savaş başlıklarını kıtalararası taşıyabilecek bir prototip balistik füze olması, ABD'de siyaset ve bilim camiası tarafından korkutucu bir unsur olarak değerlendirilmiştir. Sovyetler Birliği'nin bu başarısının ardından ABD, Sovyetler Birliği'ne yanıt olarak NASA'yı (Ulusal Havacılık ve Uzay Dairesi) ve ARPA'yı (İleri Araştırma Projeleri Ajansı; bugünkü adıyla DARPA) 1958 yılında kurmuştur. Bu olayı takiben gerçekleşen Küba Füze Krizi 1962'de ABD, Sovyetler Birliği ve Küba arasında meydana gelmiştir. ABD casus uçaklarının çektiği hava fotoğrafları, Sovyet Birliği'nin füze üslerini Küba'da inşa ettiğini göstermiş ve ilişkiler iyice gerilerek dünyayı nükleer savaş tehdidi altında bırakmıştır. İnternetin nükleer bir saldırıda ortadan kaldırılabilir bir komuta merkezi olmaksızın dağıtılmış bir ağ olarak geliştirilmesi fikrini bu olay tetiklemiştir (Green, 2010:20-21).

İnternetin bugünkü kullanımına çok benzeyen sisteminin ilk çatısı 1962 yılı Ağustos ayında J.C.R. Licklider'in Massachusetts Institute of Technology'de (MIT) tartışmaya açtığı "Galaktik Ağ-Galaksiler Arası Bilgisayar Ağı" kavramına kadar uzanmaktadır.

Licklider tarafından Galaktik Ağ, herkesin veri ve programlara herhangi bir yerden hızla erişebilmesi için küresel çapta birbirine bağlı bir bilgisayar setini ifade etmiştir. Licklider 1962 Ekim ayında Amerikan Askeri Araştırma Projesi olan DARPA'nın bilgisayar araştırma bölümünün başına geçmiş ve burada bulunduğu süre içinde araştırmacıları bilgisayar ağlarının önemi konusunda ikna etmeyi başarmıştır. MIT'den Leonard Kleinrock ise, paket anahtarlama teorisi ile ilgili ilk makaleyi Temmuz 1961'de, ilk kitabı ise 1964'te yazmış ve bilgisayar haberleşmesinde paket anahtarlama kavramının devre anahtarlama göre daha etkin bir çözüm olduğuna ulaşması ile bilgisayar ağlarına doğru önemli bir adım atılmıştır (Leiner v.d., 2009:23).

ABD Savunma Bakanlığı bünyesine bağlı ARPA'nın iki üniversitenin bilgisayarları arasında bağlantı kurarak oluşturduğu ARPANET (Amerikan Gelişmiş Savunma Araştırmaları Dairesi Ağı), 29 Ekim 1969 tarihinde kurulmuştur. Bu tarih, İnternetin ortaya çıkış tarihi olarak değerlendirilmektedir. ARPANET bilgisayarlar arasındaki ilk bağlantı olmamasına rağmen, amacı ağ oluşturmak olarak kurulan ilk bağlantı olma özelliğini taşımaktadır (Green, 2010:21). Zaman ilerledikçe bilgisayarlar hızla ARPANET'e bağlanarak host-to-host protokolü ve diğer ağ yazılımları üzerinde çalışmalar devam etmiştir. 1970'de Ağ Kontrol Protokolü (NCP) adı verilen ilk ARPANET host-to-host protokolü tamamlanmıştır (Leiner v.d., 2009:24). Bu gelişmeleri takiben Ocak 1983'te NCP'den yeni bir protokol olan, İnternet ağının altyapısını oluşturan TCP/IP'ye (Transfer Kontrol Protokolü/İnternet Protokolü) geçilmiştir. Bu geçiş tüm makinelerle uygulanması gereken bir değişiklik olduğundan dolayı uzun süreli bir planlama ile gerçekleştirilmiştir (Leiner v.d., 2009:26).

1989'da Avrupa Nükleer Araştırma Merkezi (CERN) araştırmacısı Tim Berners-Lee tarafından uluslararası protokoller sistemi (international system of protocols) fikri önerilmiş ve Web'in temelleri atılmıştır. Kullanıcılara dünyaya yayılmış, farklı türdeki birçok farklı dosyaya ait elektronik belgeleri hazırlamalarına olanak tanıyan bu dağıtılmış hipermedya istemci sunucu sistemi World Wide Web olarak adlandırmıştır (Cohen-Almagor, 2011:53). Her ne kadar Web terimi İnternet terimi yerine kullanılıyor olsa da bu iki terim birbirlerinden farklıdır. İnternet, milyonlarca

bilgisayarın küresel olarak bağlandığı, herhangi bir bilgisayarın başka bir bilgisayarla iletişim kurabileceği ağı oluşturan bir ağ şebekesi iken World Wide Web, Web sayfalarını bir tarayıcıda görüntüleyerek İnternet ortamında metin, grafik, ses, video içerebilen birbirine köprülerle (hyperlinks) bağlanan bilgiye erişmenin bir yoludur (Nath v.d., 2014:86).

Temeli İnternet üzerine kurulu olan Web, önemli gelişmeler neticesinde büyük bir bilgi platformu haline gelirken çeşitli evrelere tanık olmuştur ve gelişimi üzerine yapılan çalışmalar devam etmektedir. Web'in ilk evresi olarak adlandırılan Web 1.0, kullanıcıların içerik üretmesine olanak sağlayan Web 2.0, günümüzde çalışmaları devam eden kişiye özel sonuçlara ulaşabilmeyi hedefleyen Web 3.0, makine insan birlikteliğini amaçlayan Web 4.0 ve sonrası, bir sonraki başlıkta ele alınarak açıklanmıştır.

1.2 Web Evreleri

Web teknolojilerinin gelişmesiyle birlikte kullanıcıların sadece içeriğe erişebildiği bir dönemden, zaman ve mekân olgularından bağımsız bir şekilde hem içerik tüketicisi hem de içerik üreticisi olabildiği günümüze geçiş yaşanmıştır. Web'in gerek sanal dünyada gerekse yaşadığımız ortamda isteklerimizi tatmin edecek bir aşamaya gelmesi zaman almıştır. İnternetin gelişimi ve Web 1.0 evresinden Web 2.0 evresine geçiş ile birlikte eş zamanlı ve çift yönlü paylaşım zemin hazırlanması sonucunda sosyal medya bugünkü halini almıştır. Bu bölümde sosyal medyanın teknoloji temelini oluşturan Web 2.0 ve diğer Web sürümleri kronolojik sırayla açıklanmıştır.

1.2.1 Web 1.0

Web'in kurucusu Tim Berners-Lee tarafından “salt-okunur (read-only) Web” olarak tanımlanan Web 1.0, 1989'dan 2005'e dek sürmüştür. Temel yayıncılık ortamı olarak değerlendirilen Web 1.0 evresinde kullanıcılar görüntüledikleri içerikleri düzenleme

ve yeni içerik ekleme gibi yetkilere sahip değillerdi. Oluşturulan içerikleri görüntüleme, kullanıcıların elde etmiş oldukları tek yönlü bir katılımdı (Choudhury, 2014:8096).

Kurumsal Web sitelerinin siber alanın %90'ına tek yönlü bir şekilde hâkim oldukları dönem olarak da özetlenebilen Web 1.0 döneminde (Dwivedi v.d., 2011:978), içerik oluşumu az sayıda kişi tarafından gerçekleştirilirken, bu içeriklerin çok sayıda kişi tarafından okunabildiği bir sistem söz konusuydu. İnsanların doğrudan kaynağa giderek bilgi alabildiği Web 1.0, İnternet üzerinden erişilen birbirine bağlı hiper metin belgelerinin bir sistemi olarak faaliyet göstermekteydi (Naik ve Shivalingaiah, 2008:500).

Etkileşim açısından değerlendirildiğinde Web 1.0, statik ve tek yönlü bir yapıya sahipti. Bu evrede Web sitelerinin asıl amacı, bilgileri herkese yayınlayarak çevrimiçi bir varlık oluşturmaktı (Aghaei v.d., 2012:2). Yalnızca karşı tarafa bilgi vererek tek yönlü bir iletişim üzerine kurulu bir sistem olarak faaliyet gösteren Web 1.0'ı geleneksel kitle iletişim araçlarından ayıran özelliği çok daha fazla insana ulaşılabilir olmasıydı.

HTML, HTTP ve URI temel Web protokollerini içeren Web 1.0, statik Web sayfalardan oluşmakta ve temel üst metin işaretleme dilini (HTML) kullanmaktadır (Choudhury, 2014:8096). Web 1.0 servislerine Altavista, Yahoo ya da Netscape gibi basit giriş sayfaları, dizin hizmetleri, Web geliştirme araçları gibi basit destek araçları ve AliWeb gibi basit arama motorları örnek olarak verilebilir (Weber ve Rech, 2009:36). Bu dönemde Yahoo, Lycos ve Google gibi şirketler, bilgi arama işlevleri ve portalları geliştirmiş, Amazon ve eBay gibi şirketler de İnternette ürünlerin ve hizmetlerin satışı olan e-ticareti etkinleştirmiştir (Kambil, 2008:56).

Ardındaki temel fikir kullanıcılara bir bilgi alanı oluşturarak bilgi paylaşımı yoluyla iletişim alışverişi yapmak olan Web 1.0, işletmelerin adres, e-posta gibi iletişim bilgileri ile gazete ve dergilerdekine benzer katalog ve broşürlerini sunduğu bir

dönemdir (Hiremath ve Kenchakkanavar, 2016:707). Yapılan çalışmalarla, bireylerin İnternet ortamında katılımda bulunmadan sadece içerikleri takip ederek bilgiye ulaştığı bir evre olan Web 1.0'dan, isteyen herkesin kendi içeriklerini üretip başkalarıyla paylaşabildiği Web 2.0 evresine geçiş yaşanmıştır.

1.2.2 Web 2.0

Kaplan ve Haenlein (2010:61)'in sosyal medyanın gelişim platformu olarak tanımladığı Web 2.0, resmi olarak ilk defa 2004 yılında gerçekleştirilen bir konferansta O'Reilly ve MediaLive International arasında geçen beyin fırtınası oturumu esnasında O'Reilly Media başkan yardımcısı Dale Dougherty tarafından gündeme getirilmiştir. 2001 yılında gerçekleşen dot-com balonunun Web'in bir tür dönüm noktası olduğu konusunda fikir birliği sağlanarak Web'in geleceği masaya yatırılmıştır (O'Reilly, 2007:17). Bu beyin fırtınası oturumu sırasında Web 2.0 ile ne kastedildiğini daha iyi açıklayabilmek adına Web 1.0 ile Web 2.0 arasındaki farklar Tablo 1.1'deki gibi karşılaştırılmıştır.

Tablo 1.1: Web 1.0 ve Web 2.0 Karşılaştırması

Web 1.0	Web 2.0
Doubleclick	Google Adsense
Ofoto	Flickr
Akamai	Bittorrent
Mp3.Com	Napster
Britannica Online	Wikipedia
Kişisel Web Siteleri	Bloglar
Evite	Upcoming.Org ve Evdb
Domain İsmi Spekülasyonu	Arama Motoru Optimizasyonu
Sayfa Görme	Tıklama Başına Maliyet
Ekran Parçaları	Web Servisleri
Yayınlama	Katılım, İşbirliği
İçerik Yönetim Sistemleri	Wikiler
Dizinler (Sınıflandırma)	Etiketleme (Folksonomi)
Yapışkanlık	Sendikasyon

Kaynak: O'Reilly, 2007:18

Web 2.0 teriminin aslında yeni bir kavram olmadığı ve Web 1.0 temelinin insanları birbirine bağlamak üzerine kurulu olduğu Tim Berners-Lee tarafından dile getirilmiştir. IBM'in Web sitesinde yayınlanan bir röportajda Tim Berners-Lee, Web'in interaktif bir alana dönüşmesinin başlangıçtan beri hedeflendiğini ve Web 2.0'da Web 1.0 adına çalışan tüm insanlar tarafından üretilen standartların kullanıldığına değinmiştir (Laningham, 2006). Web 1.0'ın baştan beri temel amacı olan etkileşimli ortam, yeni gelişmelerle birlikte Web 2.0 döneminde mevcudiyet bulmuştur.

Web 2.0'ın özellikleri teknoloji, iş ve kullanıcı merkezli ele alınarak açıklanabilir. Teknoloji temelli incelendiğinde Web'in bir cihazın üzerinde yazılım düzeyinde bir platform haline geldiği görülmektedir. İş temelli incelendiğinde bilgisayar endüstrisindeki iş devrimi, İnternetin platform olarak dönüşümünün ve işletmelerin yeni platformda başarı elde edebilmek için kuralları anlamaya çalışmasının sonucu olarak doğmuştur. Kullanıcı temelli bakıldığında ise kolektif bilgi üretimi, sosyal ağ oluşturma, kullanıcı-kullanıcı arasındaki bilgi alışverişini kolaylaştıran Web uygulamaları devrine geçilmiştir ve herkes için etkileşime açık bir ortam yaratılmıştır (Choudhury, 2014:8097).

Web 2.0, bilgelik (wisdom) Web'i, insan merkezli (people-centric) Web, katılımcı (participative) Web ve okunulan-yazılan (read-write) Web gibi çeşitli şekillerde de adlandırılmaktadır. Bilgiye ulaşmanın yanı sıra katılımın da mümkün olduğu bu etkileşim Web'i iki yönlü kılmaktadır (Aghaei v.d., 2012:3).

İçerik ve uygulamaların yalnızca belirli kişiler ya da kurumlar tarafından oluşturulmadığı ve yayınlanmadığı, bunun yerine tüm kullanıcılar tarafından katılımcı ve işbirlikçi bir biçimde geliştirildiği bir platform olan Web 2.0, World Wide Web'in özel herhangi bir teknik güncellemesine tabi bulunmamakla birlikte, işlevleri için gerekli olan bir dizi temel özellik taşımaktadır. Bu özellikler arasında Web sayfalarına animasyon, etkileşim ve ses/video akışları eklemek için popüler bir yöntem olan Adobe Flash, blog girişi veya haber başlıkları gibi sık güncellenen içeriği standart bir

biçimde yayınlamak için kullanılan bir Web besleme biçimi ailesi olan RSS ve tüm sayfanın görüntüsünü etkilemeksizin Web içeriğinin güncellenmesine izin veren Web sunucularından veri alan bir teknik olan AJAX bulunmaktadır (Kaplan ve Haenlein, 2010:61).

Tüm bu özellikler beraberinde güvenlik sorunlarını da getirmektedir. Web 2.0 kullanıcısı daha fazla etkileşim ile birlikte daha az kontrole sahiptir (Choudhury, 2014:8097). Web 2.0 uygulamalarına açıkça erişilebilir olması ve bu ortamların dinamik olarak oluşturulması büyük bir güvenlik riskine neden olabilmektedir. Kötü niyetli kişilerce kişisel bilgileri ele geçirebilecek çeşitli kodların yazımı ya da içerik paylaşımı olabilmektedir (Nath v.d., 2014:87).

Web 2.0'ın neden bu denli büyük bir etkiye sahip olduğunu özetlemek adına, Anderson (2007:14-26) altı temel özellik öne sürmüştür:

- **Bireysel üretim ve kullanıcıların oluşturduğu içerik (Individual production and user generated content):** Web 2.0 ile birlikte bireyler kendi medya alanlarında bir video veya fotoğraf yükleyerek, bunları uygun anahtar kelimelerle etiketleyerek içerikleri paylaşabilmekte, blog sayfalarında kendi içeriklerini üretebilmekte, sözlüklere katkıda bulunabilmektedir. Tüm bu katılımcı yönü ile Web 2.0 kullanıcıların oluşturduğu içerik (UGC) kavramının oluşmasına zemin hazırlamıştır.
- **Kalabalığın gücünün kullanılması (Harness the power of the crowd):** Kalabalıkların bilgeliği (wisdom of crowds), kitle kaynak (crowdsourcing) ve bireysel olarak hareket eden ancak toplu sonuç üreten kişiler olarak ifade edilen folksonomiler (folksonomies) alt başlıkları ile tanımlanan kalabalığın gücünün kullanılması maddesi Web 2.0'ın sağladığı etkileşimli ortamla doğmuştur.
- **Büyük ölçekteki veriler (Data on an epic scale):** Günümüzde giderek artan miktarda veri üretilmektedir ve kullanılmaktadır. Büyük veri özellikle büyük şirketler için yönetim zorluğunu beraberinde getirirken, bu verinin işlenip

kullanılması ile şirketler, rakiplerinden sıyrılarak büyük başarı elde edebilmektedirler.

- **Katılım mimarisi (Architecture of participation):** Web 2.0'ın önemli bir ilkesi olan katılım ile birlikte hizmetler çok daha iyi bir şekilde sağlanmakta ve kullanıcıların katılımları ile daha fazla bilgi üretilerek katılım mimarisi yaratılmaktadır.
- **Ağ etkileri, güç yasaları ve uzun kuyruk kuramı (Network effects, power laws and the Long Tail):** Bir ürünü/hizmeti kullanan kişi sayısı arttıkça kullanım değerinin de artması olarak özetlenen ağ etkisi, özellikle Web 2.0 temelinde gelişen sosyal ağlarda geçerlidir. Web'in de bir güç yasası dağılımını izlediği ve uzun kuyruk kuramının geçerli olduğu görülmektedir. Örneğin Wikipedia, basılı bir ansiklopedinin içerebileceğinden on binlerce daha fazla bilgi içererek, çoğu küçük ve basılı bir ansiklopedi için önemsiz görünebilecek içeriklerin toplamı sayesinde uzun kuyruk oluşturmaktadır.
- **Açıklık (Openness):** Web'in, açık standartlarla çalışmak, açık kaynaklı yazılımları kullanmak, ücretsiz verileri kullanmak gibi açıklık temelinde çalışan güçlü bir özelliği mevcuttur. Web 2.0'ın geliştirilmesinde önemli teknolojiler, açık kaynak kodlu Firefox tarayıcısı ve deney yapılmasına izin veren genişletilebilir eklenti sistemidir.

Bu maddelerden sosyal medya kavramı ile ilgili yapılan tanımlamalarda sıklıkla kullanılan kullanıcıların oluşturduğu içerik (UGC) kavramı, Web 3.0 teriminden önce ayrı bir başlık altında ele alınarak açıklanmıştır.

1.2.2.1 Kullanıcıların Oluşturduğu İçerik

2006 Aralık sayısında Time Magazine kullanıcıların oluşturduğu içeriğe katkıda bulunan kişileri, yılın en değerli üyeleri olarak seçmiştir (Shao, 2009:8). Dergi kapağında bulunan görselde monitör ve klavyenin altında “Evet, sen. Sen Bilgi Çağı'nı kontrol ediyorsun. Senin dünyana hoş geldin” yazısı bulunmaktadır. YouTube video

oyuncu görüntüsüne yer verilerek Web 2.0 ve kullanıcıların oluşturduğu içeriğe atıfta bulunulmuştur.

Şekil 1.1: 2006 Aralık Sayısı Time Dergisi Kapağı



Kullanıcıların oluşturduğu içeriğin sosyal medya boyutunda ele alınması için belirli özellikler taşıması gerekmektedir. İktisadi İşbirliği ve Gelişme Teşkilatı'na (OECD) göre bu özellikler üç temel başlıktan oluşmaktadır. İçeriklerin herkesin erişimine açık bir Web sitesinde veya belirli bir grup insanın erişebileceği bir sosyal ağ sitesinde yayınlanması, belirli miktarda yaratıcı çaba göstermesi, profesyonel işlerin ve uygulamaların dışında üretilmiş olması gerekmektedir. Birinci koşul, e-postalarda veya anlık iletilerde değiştirilen içeriği; ikinci koşul, yalnızca mevcut içeriklerdeki tekrarlamaları (örneğin, mevcut bir gazete yazısının bir kopyasını bir kişisel blog sayfasında herhangi bir değişiklik veya yorum yapılmaksızın yayınlamak); ve üçüncü koşul, bir ticari pazar bağlamı göz önüne alınarak yaratılmış tüm içeriği hariç tutmaktadır (Kaplan ve Haenlein, 2010:61).

Bu temel üç özellik şu şekilde detaylandırılmaktadır (OECD, 2007:18):

- **Yayınlanma gerekliliği:** Teorik olarak UGC herhangi bir kullanıcı tarafından oluşturulup yayın zorunluluğu bulunmuyor olsa da, yayınlanma gerekliliğine göre içerik herkesin erişimine açık bir Web sitesinde ya da üniversite öğrencileri gibi sadece belirli bir grup insanın erişebileceği bir sosyal ağ sayfasında paylaşılmalıdır. E-posta, çift taraflı anında mesajlaşma gibi durumlar bu özellik dâhilinde UGC kapsamında değerlendirilmemektedir.
- **Yaratıcı çaba:** İçerik üretilirken kullanıcının belirli bir yenilikçi çaba harcaması gerekmektedir. Yalnızca bir televizyon şovunun bir bölümünü kopyalayıp çevrimiçi bir video sitesinde bu içeriği paylaşmak UGC olarak kabul edilmemektedir. Kişi içeriğe kendinden bir şeyler katmalıdır. Kişi kendi fotoğraflarını yüklemesi, düşüncelerini bir blog yoluyla ifade etmesi veya yeni bir müzik videosu oluşturması sonucunda ortaya çıkan bu içerikler UGC olarak değerlendirilebilir. Yaratıcı çaba sadece tek bir kişiyle değil belli bir grup ya da takım çalışması sonucunda da ortaya çıkarılabilir. Kullanıcıların birlikte Web sayfaları düzenlemeleri yaratıcı içeriğe örnek olarak gösterilebilir. Yaratıcı çabanın minimum miktarını tanımlamak zor olmakla birlikte içeriğe göre değişkenlik göstermektedir.
- **Profesyonel işlerin ve uygulamaların dışında bir oluşum:** Kullanıcıların oluşturduğu içerik genellikle profesyonel işler ve uygulamalar dışında oluşturulur ve bir kurumsal içeriğe ya da ticari pazarlama içeriğine sahip değildirler. İçerikler profesyonel olmayan kişiler tarafından kar ya da bir karşılık beklentisi olmadan oluşturulmaktadır. Bu içeriklerin oluşturulmalarındaki motivasyon unsurları; diğer katılımcılarla bağlantı kurmayı, belirli bir şöhret seviyesine veya saygınlığa erişmeyi ve kendilerini ifade edebilecekleri bir ortam sağlamayı içermektedir.

Kullanıcıların oluşturduğu içeriğin gelişim sürecine bakıldığında, 1990'lı yıllarda Yahoo ve AOL gibi portal sitelerinde bulunan bildiri tahtalarına (bulletin boards) kadar uzandığı görülmektedir. Zaman içerisinde, blogları, wikileri, resim ve video

paylaşımlarını, sosyal ağları ve diğer Web sitelerini kapsayacak şekilde gelişim gösteren kullanıcıların oluşturduğu içerikler, çeşitli platformlarda üretilmeye devam etmektedir (Shao, 2009:8).

Kullanıcıların oluşturduğu içeriğin hızlı gelişmesini ve yaygınlaşmasını sağlayan bazı etkenler söz konusudur. Bunlar teknolojik, sosyal, ekonomik ve kurumsal/yasal etkenler olarak gruplandırılabilir (OECD, 2007:28).

Tablo 1.2: Kullanıcıların Oluşturduğu İçeriği Destekleyen Etkenler

Teknolojik Etkenler	• Artan geniş bant kullanılabilirliği • Artan sabit sürücü kapasitesi ve düşük maliyetli işlemci hızları • Ses, fotoğraf ve video gibi görsel ve işitsel tüketici teknoloji araçlarının kalitesinin artması ve maliyetlerin azalması • İçerik oluşturma, dağıtma ve paylaşmayı sağlayan teknolojilerin varlığı • İçerik oluşturma, düzenleme ve dağıtmaya yardımcı olan daha basit yazılımların geliştirilmesi • Profesyonel ve profesyonel olmayan kullanıcı kaynaklı sitelerin artışı
Sosyal Etkenler	• Bilişim ve iletişim teknoloji becerileri gelişmiş, kişisel bilgilerini çevrimiçi olarak paylaşma (örneğin içerik paylaşma, tavsiye etme ve oylama) konusunda istekli ve çevrimiçi paylaşım yaparken daha az tereddüt içerisinde olan genç yaş gruplarına, bir başka deyişle dijital yerlilere doğru bir geçiş yaşanması • Kendi kendini yaratma ve ifade etme arzusu ve televizyon gibi geleneksel medya platformlarına kıyasla daha fazla etkileşim arayışı içinde olunması • Toplulukların ve işbirlikçi projelerin gelişimi • İleri yaş gruplarına da bu gibi sosyal etkenlerin yayılması ve sosyal sorumluluk, politika ve eğitim gibi toplumsal işlevlerinin yerine getirilmesi
Ekonomik Etkenler	• UGC'nin oluşturulmasını (örneğin içerik oluşturmak, düzenlemek, sağlamak gibi) sağlayan araçlara erişimin artması, düşük maliyetler ve giriş engellerinin azlığı

	• Geniş bant İnternet bağlantılarının daha düşük maliyetle kullanabilmesi • UGC'ye ve uzun kuyruk ekonomisine olan ilginin artması (mobil operatörler, telekomünikasyon hizmet sağlayıcıları, geleneksel medya yayıncıları ve arama motorları dâhil olmak üzere) • Girişim sermayesi ve diğer yatırım olanakları sayesinde ilgili girişimlerin ve UGC sitelerinin finanse edilmesi için olanakların artırılması • İçeriğin ticari anlamda değerini arttırmak için reklam ve yeni iş modellerinin çok daha elverişli hale gelmesi
Kurumsal/Yasal Etkenler	• Türemiş çalışmalar oluşturma hakkının ve yaratıcı içeriklere esnek erişim imkânı sağlayan planların artması (örneğin esnek lisanslama ve Creative Commons lisansı gibi telif hakkı planları) • Oluşturdukları içerik için kullanıcılara telif hakkı veren sözleşmelerinin artması
Kaynak: OECD, 2007:28	

Web 2.0'ın sunmuş olduğu etkileşimli ortam sonucunda kullanıcıların oluşturduğu içerik miktarı her geçen gün artarak büyük verinin devasa boyutlara ulaşmasına yol açmaktadır. Web 2.0 sonrası olarak değerlendirilen evre ise Web 3.0 olarak adlandırılmaktadır ve yapılan çalışmalar günümüzde devam etmektedir.

1.2.3 Web 3.0

Web'in gelişimine kısaca bakıldığında Web 1.0'ın tamamen bilgiye erişim ve e-ticarete ilişkin olduğu görülmektedir. Herhangi bir katılımın söz konusu olmadığı bu dönemde yalnızca Web sitelerine bağlanıp bilgiye ulaşmak mümkündü. Web 2.0 ise hayatımıza sosyal medyayı ve kullanıcı tarafından oluşturulan içeriği getirmiştir. Web 1.0'daki bu tek taraflı bilgi aktarımı yerini etkileşimli bir ortama bırakarak kullanıcılara hem okuma hem de içerik üretme fırsatı sunmuştur. Web 3.0 ise tüm bu gelişimlerin üzerinde bilgisayarların içeriğin ne anlama geldiğini anlamasını ifade etmektedir. Semantik Web olarak adlandırılan bu dönem öncesi "portakal" sözcüğü 8

adet alfa-sayısal simge olarak tanımlanırken, semantik anlamda bu sözcüğün meyve mi, sebze mi yahut bir marka mı olduğu anlaşılabilecektir (Gürsakal, 2014:48).

Semantik Web'in temelindeki amaç, kullanıcıların bilgiyi daha kolay bulmalarını, paylaşımlarını ve birleştirmelerini sağlayarak mevcut Web'e katkıda bulunarak gelişmesini sağlamaktır. Başlangıçta öngörülen Semantik Web, makinelerin karmaşık insan taleplerini anlaması ve onlara cevap vermesini sağlayan bir sistem olarak tanımlanabilmektedir (Choudhury, 2014:8098). Farklı tanımlamalar yapılan Web 3.0 için Bill Gates tarafından "yapay zekâ" yorumu yapılırken ünlü bilişim yazarı Nicholas Carr ise Web 3.0'ın insan gibi düşüneceğinin altını çizmiştir. Üçüncü nesil İnternet ağı olarak da bilinen Web 3.0 ile birlikte, temelinde birbirinden kopuk Web sayfalarının anlamsal olarak birbirine bağlanması sonucunda küresel bir veri tabanı oluşturularak içerikler insanların yanı sıra makineler tarafından da anlaşılabilecektir (Özdemir ve Karalar, 2012:102).

Web 2.0 ve Web 3.0 arasındaki en büyük fark Web 2.0'da içerik yaratıcılığı hedeflenirken Web 3.0'da bağlantılı veri setlerinin hedeflenmesidir (Aghaei v.d., 2012:6). Web 2.0'da kullanıcı bir arama motoru yardımıyla anahtar kelimeler kullanarak araştırma yaptığıında, büyük bir bilgi yığını ile karşılaşabilmekte ve bu bilgi yığını içerisinde ilgilendiğine ulaşılabilmek için zaman kaybı yaşayabilmektedir. Web 3.0 İnterneti kullanıcı odağından bilgisayarların içeriğe etkide bulunabileceği bir yapıya çevirerek, daha kısa zamanda en doğru sonuçlara ulaşmayı mümkün kılacaktır.

Bu evre akıllı telefon çağına geçilmesi ve sosyal ağların kullanımının artması ile birlikte adımları atılan Web 3.0, Bulut Bilişim (Cloud Computing), Büyük Veri (Big Data), Nesnelerin İnterneti (Internet of Things), Web 2.0'daki sosyal ağlar, güvenlik ve mevcut özellikler gibi tüm en yeni teknolojileri bir araya getiren bir birleşim olarak ele alınabilmektedir (Newman v.d., 2016:597). Web 3.0 kapsamında yapılan çalışmalar devam etmekle birlikte Web 3.0 devamında nelerin konuşulacağı da tartışılmaktadır. Web'in geleceğinin ne olacağı konusunda çeşitli fikirler üretilmekte

ve Web 3.0'ın ötesinde Web 4.0 ve Web 5.0 terimleri ile ilgili gelişmeler öngörülmektedir.

1.2.4 Web 4.0 ve Ötesi

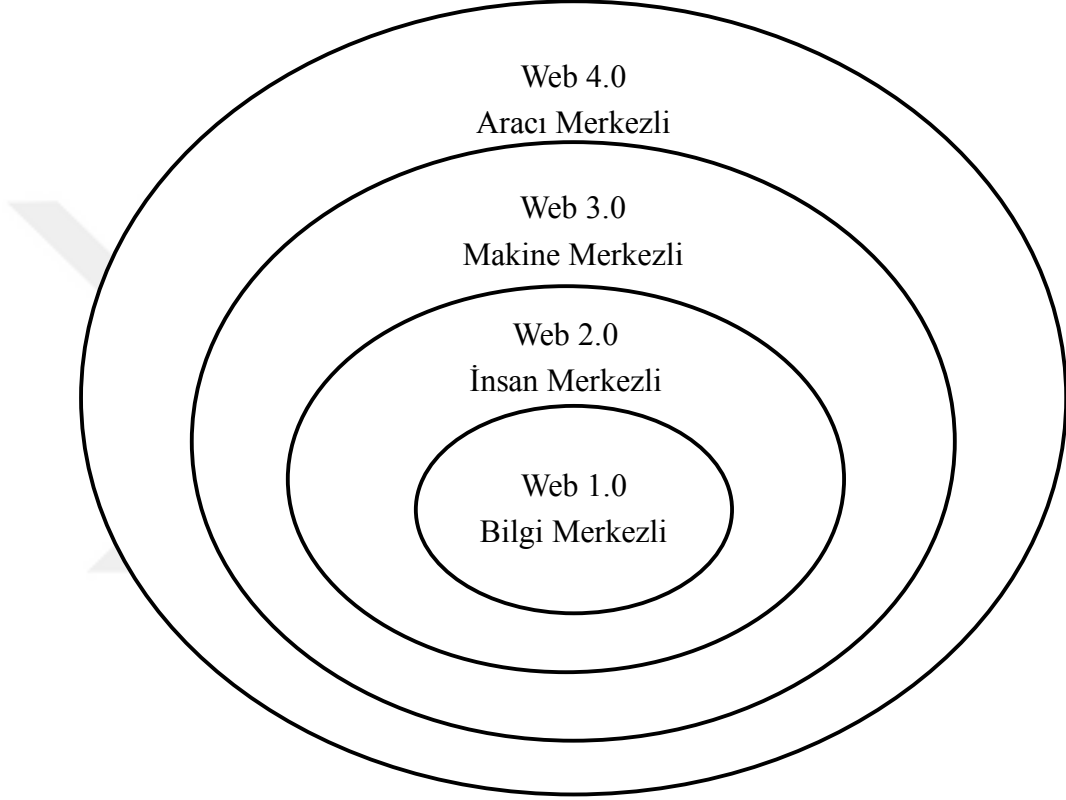
Web'in gelişiminde Web 3.0'dan sonraki evre olan, akıllı (intelligent) Web, zeki (smart) Web, işbirlikçi (collaboration) Web ve aracı merkezli (agent-centric) Web olarak çeşitli isimlendirmeleri bulunan Web 4.0 henüz fikir aşamasında olmasına karşın, temelinde tamamen sanallaştırma üzerine kurulmuş, makine ve insanın birlikteliği çalışması şeklinde özetlenebilir. Web 4.0, hem insanların hem de bilgisayarların yalnızca etkileşim kurmakla kalmayıp aynı zamanda birbirlerini anlayıp yardımcı oldukları bir evre olup, her yerde bulunan bir Web üzerinden insan ve makine akıl gücünün kullanılması olarak yorumlanabilmektedir (Murugesan, 2009:4).

Zaman içerisinde daha fazla nesnenin İnternete bağlanması sonucunda kullanıcıların oluşturduğu içerikler ve analitik hizmetlerin sağlanması başka boyutlara taşınabilecektir. Web 4.0 ile birlikte kullanıcıların etkileşimde bulunduğu sanal dünyalar ve nesneler arasındaki gerçek zamanlı entegrasyon artacaktır. Hali hazırda yer bulucu cihazlara sahip cep telefonları bizi bir mağazaya yönlendirebilmektedir (Kambil, 2008:57). Web 4.0, çevrimiçi işlevselliği fiziksel dünyaya taşıyarak önemli bir gelişmeye yol açacaktır. Örnek olarak araç anahtarının yerini bulmak için, kişinin evini Google'da aratabileceği verilebilir (Patel, 2013:416).

Değinilmesi gereken önemli bir nokta Web sürümlerinin birlikteliğidir. Web 1.0, Web 2.0, Web 3.0 ve Web 4.0 olarak adlandırılan Web evrelerinin, bir evrenin bitiminin diğer evrenin başlangıcına yol açtığı düşüncesi ile değerlendirilmesi doğru olmayacaktır. Şekil 1.2'de görüldüğü üzere bu evreler birbirleriyle ilişki içerisinde dirler. Örneğin Web 2.0 bittikten sonrası Web 3.0 gerçekleşmiştir diye okunmamalı; Web 3.0 Web 2.0 evresinin temelleri üzerinde oluşmuştur şeklinde

yorumlanmalıdır. Web sürümleri birbirlerini desteklemektedirler. Web evreleri genelde sırasıyla yaşanıyor olsa da bazı siteler ya da kullanıcılar henüz birinci evredeyken bazıları üçüncü evreyi yaşamaktadır.

Şekil 1.2: Web'in Gelişim Evreleri



Kaynak: Murugesan, 2009:3

Web 4.0 sonrası olarak değerlendirilen Web 5.0'in henüz tam olarak bir tanımı oluşmamakla birlikte bazı adlandırmalar mevcuttur. "Symbionet Web" olarak adlandırılan Web'in bu evresinde herhangi bir merkez söz konusu değildir; İnternette depolanan kişisel veriler ve bilgiler için bir Kişisel Sunucu (PS) bulunmamaktadır ve insanlar akıllı telefonlar, tabletler ve/veya kişisel robotlar gibi Akıllı İletişimci (Smart Communicator) ile birbirine bağlanmaya çalışarak 3D sanal dünyada gezinebileceklerdir (Patel, 2013:416).

Web 5.0 bir başka kaynakta “duyusal ve duygusal (sensory and emotive) Web” olarak adlandırılmıştır. Web şimdilik duyguları anlamlandıramadığı için kullanıcıların neler hissettiğini anlamak ve duyguları görselleştirmek zor olsa da, duyguları ölçebilen bazı teknolojiler mevcuttur. Örneğin www.wefeelfine.org, Web’deki duygusal ifadeleri izlemekte, bunları sınıflandırmakta ve duyguların kümelerinin frekansını ve yerini kaydetmektedir (Benito-Osorio v.d., 2013:277).

Bir diğer çalışmaya göre Web 5.0’ın, makinelerin ve insanların verileri ele alabilecekleri, etkileşim kurabilecekleri ve karar verebilecekleri biçimde işleyecekleri füzyon (fusion) Web ile ilgili olduğu değerlendirilmektedir. Web 5.0, okuma-yazma-yürütme-eş zamanlı-füzyon (read-write-execute-concurrent-fusion) Web olarak da adlandırılmaktadır (Algosaibi v.d., 2015:83).

Tüm bu gelişmeler ve değerlendirmeler ışığında Web farklı evrelere tanık olmakta ve bu bilgiye ulaşımı, etkileşimi, nesnelerin günümüzdeki konumunun her geçen gün daha da değişime uğramasına yol açmaktadır. Genel bir değerlendirme yapıldığında kullanılan teknolojilerin ve oluşturulan veri türlerinin de değişmekte olduğu göze çarpmaktadır. Web 1.0 ve sosyal medyanın temeli olan Web 2.0 yapılandırılmamış ve yarı yapılandırılmış verileri içerirken, Web 3.0 ve sonrası veri formlarının yapılandırılmış verileri içereceği öngörülmektedir (Algosaibi v.d., 2015:80).

1.3 Sosyal Medya

Sosyal medyanın gelişimine bakıldığında Web 2.0 ile ivme kazandığı aşikârdır. Ancak sosyal medyanın detaylı tarihçesi ele alındığında temellerinin atıldığı farklı platformların olduğu göze çarpmaktadır. Bu bölümde sosyal medyanın temelinin hangi yıllara ve dönüm noktalarına dayandırıldığı sosyal medyanın tarihçesi bölümünde ele alınmıştır. Sosyal medya ile ilgili yapılan çeşitli tanımlamalara ve sosyal medya ile ilgili raporlarda yer alan istatistiklere yer verilerek sosyal medyanın gücü verilerle vurgulanmıştır. Sosyal medyanın sahip olduğu özelliklere yer verilerek

bu kavram ile ilgili tanımlamalar pekiştirilmiştir. Son olarak sosyal medyanın temel yapı taşları maddeler halinde sosyal medya platformları içerisinde örnekler verilerek açıklanmıştır.

1.3.1 Sosyal Medyanın Tarihçesi

Sosyal medyaya geçiş olarak görülen çeşitli dönüm noktaları mevcuttur. 1979'da Duke Üniversitesi'nden Tom Truscott ve Jim Ellis, İnternet kullanıcılarının kamuya açık mesajlar yayınlamasını sağlayan dünya çapında bir tartışma platformu olarak kabul gören Usenet'i oluşturmuştur. Ancak bugün anladığımız sosyal medya kavramına yakın bir oluşumun başlangıcı olan Open Diary, Bruce ve Susan Abelson tarafından çevrimiçi günlük yazarlarını bir toplulukta buluşturan ilk sosyal paylaşım sitesi olarak kurulmasıyla yaklaşık 20 yıl önce başlamıştır. "Weblog" terimi ilk kez bu zamanlarda kullanılmış ve bir blog yazarının espri amaçlı "webblog" adını "we blog (blogluyoruz)" cümlesine dönüştürmesinden bir yıl sonra "blog" olarak kısaltılarak bugünkü halini almıştır. Zaman içerisinde İnternetin gelişimi ile birlikte çeşitli sosyal ağ sitelerinin oluşumu artmış ve tüm bu gelişmeler beraberinde sosyal medyanın oluşmasına katkı sağlamıştır (Kaplan ve Haenlein, 2010:60).

Bir başka kaynak sosyal medya site ve servislerinin temelini 1997 yılında kurulan Sixdegrees sosyal ağına dayandırmaktadır (Kietzmann v.d., 2011:242). Sajithra ve Patil (2013:70) ise sosyal medyanın gelişimini kronolojik bir şekilde ele almış ve sosyal medyanın temellerinin 1971'e dayandığını belirtmiştir. Bu tarihte ARPANET üzerinden Ray Tomlinson tarafından ilk e-posta gönderilmiştir. E-postanın sosyal medyanın bir parçası olarak kabul edilip edilmeyeceği konusunda tartışmalar söz konusudur, ancak basit anlamda sosyal medyanın çevrimiçi yapılan sohbetler şeklindeki tanımı dâhilinde değerlendirildiğinde, sosyal medyanın ilkel anlamda başlangıcının bu tarihlere uzandığı düşünülebilir.

Sosyal medyanın gelişimi kronolojik sırada ele alınacak olduğunda; ilk e-posta gönderimi (1971), Usenet oluşumu (1979), LISTSERV (e-posta liste yönetim programı) ile birlikte e-posta iletişiminin kullanımının artması ve anında çok sayıda kişiye ulaşmada gücünün kanıtlanması (1984), ilk İnternet Aktarmalı Sohbet (IRC) protokolünün yazılması (1988), kişisel web siteleri, tartışma ve sohbet grupları; IRC başarısından sonra web sitelerinin, tartışma ve sohbet gruplarının popülerleşmeye başlaması (1991), üyelerine arkadaşlarını ve tanıdıklarını bulmalarına yardımcı olan bir sosyal ağ sitesi olarak Classmates.com'un faaliyete geçmesi (1995), Open Diary (1998), Blogger (1999), Messenger (1999), Linden Lab (1999), Napster (1999), LiveJournal (1999), Cyworld (1999), Wikipedia (2001), Friendster (2002), Last.fm (2002), LinkedIn (2003), Myspace (2003) gibi blogların, podcastlerin ve wikilerin hayata geçmesi (1998-2004), Web 2.0 uygulamaları ve kullanıcıların oluşturduğu içerik kavramının oluşması beraberinde günümüze kadar süren gelişmeler (2005 ve sonrası) şeklinde özetlenebilir (Sajithra ve Patil, 2013:70-74).

Tüm bu farklı gelişmeler ışığında sosyal medyanın çeşitli tanımlarının olması kaçınılmazdır. Literatürde araştırmacılar tarafından yapılan çeşitli tanımlamalar mevcuttur. Bu tanımlara bir sonraki başlıkta yer verilmiştir.

1.3.2 Sosyal Medya Kavramı

Yapılan çeşitli tanımlamalar değerlendirildiğinde sosyal medyanın teknolojik ve kullanıcı temelli tanımlarının olduğu görülmektedir. Web 2.0 ve kullanıcıların oluşturduğu içerik kavramları sıkça sosyal medya ile birlikte kullanılmaktadır. Bazı araştırmacılar bu kavramlar çerçevesinde sosyal medyayı açıklamışlardır. Kaplan ve Haenlein (2010:61) tarafından sosyal medya; “Web 2.0’ın ideolojik ve teknolojik temelleri üzerine kurulmuş ve kullanıcıların oluşturduğu içeriğin üretilmesine ve paylaşılmasına izin veren bir İnternet tabanlı uygulamalar grubu” olarak tanımlanmıştır. Bir başka değerlendirmeye göre sosyal medya, Web 2.0 teknolojileri temelinde kurulu, sosyal etkileşimi, topluluk ve işbirlikçi projelerin oluşumunu

sağlayan Web siteleri olarak tanımlanmıştır (Bruns ve Bahnisch:2009:7). Genel olarak sosyal medya, kullanıcıların çevrimiçi ortamlarda bir araya gelmeleri, metin, video, görsel gibi herhangi bir içeriği üretmeleri, herhangi bir platformda daha önceden paylaşılan içerikleri değerlendirmeleri, paylaşmaları, yorum yapmaları gibi her türlü sosyal etkileşim içerisinde bulunabilmelerini sağlayan Web temelli uygulamalar ve servisler olarak tanımlanabilir (Ryan, 2016:121).

Sosyal medya ile ilgili, tek yönlü bilgi aktarımından çift yönlü bilgi paylaşımına geçilmesi üzerinde durularak yapılan tanımlamalar mevcuttur. Sosyal medya Web 1.0'dan Web 2.0'a geçiş aşamasında ortaya çıkarak bilgi paylaşımının tek yönlüden çift yönlüye geçmesini sağlamış ve “kullanıcıların ürettiği içerik”, “müşterilerin ürettiği medya” gibi kavramları ortaya çıkarmıştır (Baloğlu, 2015:1-2). Bir başka tanıma göre sosyal medya içeriklerin kolayca aktarılabildiği Web tabanlı uygulamalar kullanılarak bilgi ve görüş paylaşan, çevrimiçi olarak bir araya gelen topluluklar arasındaki uygulamalar, faaliyetler ve davranışlar olarak ifade edilmektedir (Safko ve Brake, 2009:6).

Sosyal medya, bireylerin, içerikler, siteler, ağlar gibi oluşumların ve ikisi arasındaki etkileşimlerin bir dünyası olarak düşünülmektedir (Zafarani v.d., 2014:16). Sosyal medya, tek bir kaynağa dayalı yayın mekanizmasından, kişiler ve gruplar arası gibi birden çok bir yayın mekanizmasına geçişi sağlamıştır (Evans, 2008:33). Bireylerin topluluk odaklı Web sitelerde bilgi, deneyim ve bakış açılarını istedikleri gibi paylaşabilmelerine olanak sağlayan sosyal medya, coğrafik duvarları dağıtmakta ve yeni çevrimiçi topluluklar ortaya çıkararak büyümeye devam etmektedir (Weinberg, 2009:1).

Bazı araştırmacılar ise sosyal medyanın geleneksel medyadan farklı olan özellikleri üzerinde durarak sosyal medyayı tanımlamaktadırlar. Sosyal medya Facebook, Twitter gibi geleneksel olmayan ve gazete, radyo ve televizyon gibi geleneksel medyayı içeren farklı türdeki sosyal medya sitelerinin yığını olarak görülmekle birlikte, kullanıcılara daha önce geleneksel medyada görülmemiş bir ölçekte ve hızla birbirleriyle

etkileşime geçmeleri ve ağlarını genişletebilmeleri için kolay bir yol sunmaktadır (Gundecha ve Liu, 2014:2). Bir diğer tanıma göre geleneksel medyadan oldukça farklı olan sosyal medya eşi benzeri görülmemiş bir kitle iletişim aracı olarak tanımlanmakta ve kitle kaynağın kolaylıkla yapılabildiği, farklı amaçlar için özel çıkar gruplarının olduğu doğal bir işbirliği ortamı olarak görülmektedir (Liu v.d., 2016:137). Geleneksel medyanın da sosyal medya gibi birbirinden farklı kanalları bulunsa da, sosyal medya anlık olarak değişkenlik gösterebilmesi ve katılımcı yapısı ile geleneksel medyadan farklılaşmaktadır (Evans, 2008:34).

Sosyal medya kavramının anlamını kavrayabilmek için onu oluşturan iki kelimeden de yola çıkılabilir. “Medya” kavramı genellikle yayınlar ve/veya kanallar kullanılarak reklamcılık faaliyetlerinin gerçekleştirilmesi ve fikir ve bilgilerin iletişimi anlamına gelmektedir. “Sosyal” kavramı ise, grup ya da topluluk içerisindeki bireylerin birbirleriyle olan etkileşimi anlamını taşımaktadır. Bu iki kavram birlikte ele alındığında, sosyal medya, kişilerin belirli araçlar yolu ile etkileşim kurması sonucu üretilen ve sürdürülen iletişim ve yayın platformları şeklinde tanımlanabilir (Neti, 2011:2).

Sosyal medyanın gücü ise Orsburn (2011:6-7) tarafından dikkat çekici rakamlarla açıklamıştır. Bu rakamlara göre 50 milyon kullanıcıya radyo 38 yılda, televizyon 13 yılda ve İnternet 4 yılda ulaşmıştır. Facebook ise 100 milyon kullanıcıya 9 ayda erişmiştir. İnternet tarihinin başlangıcından itibaren devam eden tüm gelişmelerle Web’in farklı evrelere tanık olmasıyla, sosyal medyanın daha farklı bir boyuta geçebileceği öngörülmektedir. Günümüzde İnternet denilince akla gelen ilk kavramlar arasında şüphesiz sosyal medya terimi de yer almaktadır.

Türkiye İstatistik Kurumu (TÜİK) verilerine göre ülkemizde İnternet kullanım amaçları arasında sosyal medya ilk sırada bulunmaktadır. TÜİK tarafından yayınlanan haber bülteninde 2016 yılının ilk üç ayında İnternet kullanan bireylerin %82,4’ü sosyal medya üzerinde profil oluşturma, mesaj gönderme veya fotoğraf vb. içerik paylaşma faaliyetlerinde bulunurken, bunu %74,5 ile paylaşım sitelerinden video izleme, %69,5

ile çevrimiçi haber, gazete ya da dergi okuma, %65,9 ile sağlıkla ilgili bilgi arama, %65,5 ile mal ve hizmetler hakkında bilgi arama ve %63,7 ile İnternet üzerinden müzik dinleme takip etmiştir (TÜİK, 2016).

We Are Social ve Hootsuite ortaklığı ile hazırlanan 2018 Ocak ayına ait İnternet, sosyal medya ve mobil kullanım istatistiklerinden oluşan kapsamlı çalışmada, Türkiye ile ilgili olan istatistikler “Digital in 2018 in Western Asia” isimli rapordan elde edilmiştir. Türkiye ile ilgili olan yıllık dijital büyüme oranları ele alındığında, Türkiye’de 2017 Ocak ayından bu yana İnternet kullanıcı sayısının %13 oranı ile 6 milyon büyüyerek 54.33 milyona ve aktif sosyal medya kullanıcısı sayısının %6 oranla 3 milyon büyüyerek 51 milyona ulaştığı görülmektedir. Kullanıcıların gün içerisinde sosyal medyada ortalama 2 saat 48 dakika vakit geçirdikleri ise bir başka dikkat çekici bulgudur. En çok kullanılan sosyal medya platformlarına bakıldığında ilk sırada YouTube gelmekte ve onu sırası ile Facebook, Instagram ve Twitter takip etmektedir (We Are Social ve Hootsuite, 2018).

Bu araştırmalar dışında, Gençlik ve Spor Bakanlığı’nın 2013 yılında yaptığı çalışma da sosyal medya kavramı ile ilgili önemli bulgular elde edilmiştir. 26 ilden 2057 gencin (15-29 yaş arası) katılımıyla gerçekleştirilen “Gençlik ve Sosyal Medya Araştırması” adlı çalışmada gençlerin sosyal medya ile ilgili tutum ve davranışları incelenmiştir. Bu araştırmanın sonuçlarına göre sosyal medyadaki en önemli konunun %37 ile özgürlük olduğu tespit edilmiştir. İnternet kullanıcısı gençlerin %96’lık gibi çok büyük bir kesiminin sosyal medya kullanıcısı olduğu göze çarpmaktadır. Bununla birlikte gençlerin %56’sı “sosyal medyanın kitleleri harekete geçirme gücü olduğu” düşüncesine sahiptir. Sosyal medyanın kullanım amacına bakıldığında ise en çok eğlence (%60) ve bilgi alma/sağlama (%59) amacıyla kullanılmaktadır. Buradan yola çıkarak, sosyal medyanın gençler açısından sadece eğlence amaçlı kullanılmadığı, aynı zamanda geleneksel medyanın alternatifi olarak da algılandığı söylenebilir. Sosyal medya gençlerin yaklaşık yarısı için aynı zamanda serbest zaman geçirme (%54), iletişim kurma (%53), gündemi takip etme/gündem yaratma (%51) ve eğitim-

öğretim-araştırma (%47) anlamına gelmektedir (T.C. Gençlik ve Spor Bakanlığı, 2013).

1.3.3 Sosyal Medyanın Özellikleri

Sosyal medya ile ilgili yapılan tanımlar ve çalışmalar neticesinde sosyal medyanın özellikleri araştırmacılar tarafından çeşitli şekillerde ifade edilmiştir. Kim v.d. (2010:219) sosyal medya sitelerinin özelliklerini kişisel profil, çevrimiçi bağlantı kurma, çevrimiçi gruplara katılma, çevrimiçi bağlantılarla iletişim kurma, kullanıcıların oluşturduğu içeriği paylaşma, fikir ve yorumda bulunma, bilgi edinme ve kullanıcıları sitede tutma başlıkları altında toplamıştır.

Mayfield (2008:5) tarafından sosyal medyanın özellikleri beş başlık altında toplanmıştır:

- **Katılım (Participation):** Sosyal medya, medya ile bireyler arasındaki duvarları kaldırmaktadır. Herkesin katkı sağlamasını ve geribildirimini teşvik etmesi sosyal medyanın önemli bir özelliğidir.
- **Açıklık (Openness):** Sosyal medya bilgi paylaşmayı ve yorum yapmayı teşvik eden özelliği ile katılıma ve geribildirime açıktır. İçeriğe erişim ve bu içeriklerden faydalanma konusunda engeller yok denecek kadar azdır.
- **Karşılıklı konuşma (Conversation):** Sosyal medya geleneksel medyadaki tek yönlü içerik aktarımının aksine iki yönlü iletişimi mümkün kılan bir özelliğe sahiptir.
- **Topluluk (Community):** Sosyal medya ortak ilgili alanlarına sahip olan kişilerin zaman ve mekân olgularının ötesinde, kolayca bir araya gelerek bir topluluk oluşturmaya olanak sağlamaktadır.
- **Bağlantılılık (Connectedness):** Çoğu sosyal medya platformu, diğer sitelere, kaynaklara ya da kişilere bağlantı verilmesine olanak tanıyarak bir ağ örmektedir.

Tang ve Liu (2010:2) tarafından ise sosyal medyanın özellikleri şu şekilde açıklanmıştır:

- **Kitle yayıncılık (Mass publication):** Bireylerin anlık bilgi paylaşımı ile birlikte kullanıcıların oluşturduğu içerik miktarı çığ gibi büyümekte ve kalabalıkların bilgeliğini (wisdom of crowds) oluşturmaktadır.
- **İşbirlikçi yazım (Collaborative writing):** Sosyal medyanın kullanıcılarına içerik üreticisi olabilme imkanı tanımaktadır. Bunun sonucunda Wikipedia platformunda olduğu gibi kaliteli içerik üretimi birçok kullanıcının katılımı ile gerçekleşmektedir.
- **Zengin kullanıcı etkileşimi (Rich user interaction):** Kullanıcıların katılımı sosyal medyanın başarısında önemli rol oynamaktadır. Kullanıcı katılımı daha fazla kullanıcı katılımını tetiklemekte ve etkileşimi arttırmaktadır.

Bir başka çalışmada ise sosyal medyanın herkes tarafından değinilen diyalog temelli olmasının dışında kontrol edilmesinin zorluğuna ve çok kolay bir şekilde bireylerin sosyal medya çatısı altında koordine olabileceğine dikkat çekilmiştir (Lee v.d., 2013:792-793):

- **Diyaloga yer vermesi (Dialogic):** Sosyal medyayı geleneksel medyadan ayıran özellik; sosyal medyanın etkileşimli ortam yapısına sahip olması, diyaloga yer vererek katılımı teşvik etmesidir. Kullanıcılar herhangi bir yönetici altında değil, bağımsız olarak ilişkiler kurmaktadır.
- **Kontrol edilememesi (Uncontrollable):** Sosyal medyanın kontrol edilebilmesi hemen hemen mümkün değildir. Sosyal medya tek bir varlığın kontrolünün ötesinde çok sayıda kullanıcı tarafından üretilen içeriklerden oluşmaktadır. Bilgi akışının çok yönlü, birbirine bağlı ve tahmin edilmesinin zor olmasından dolayı sosyal medyanın manipüle edilmesi zordur.
- **Koordinasyonsuz eylemlerin koordine etkiler yaratması (Coordinated effects of uncoordinated actions):** Koordine olmak için geleneksel

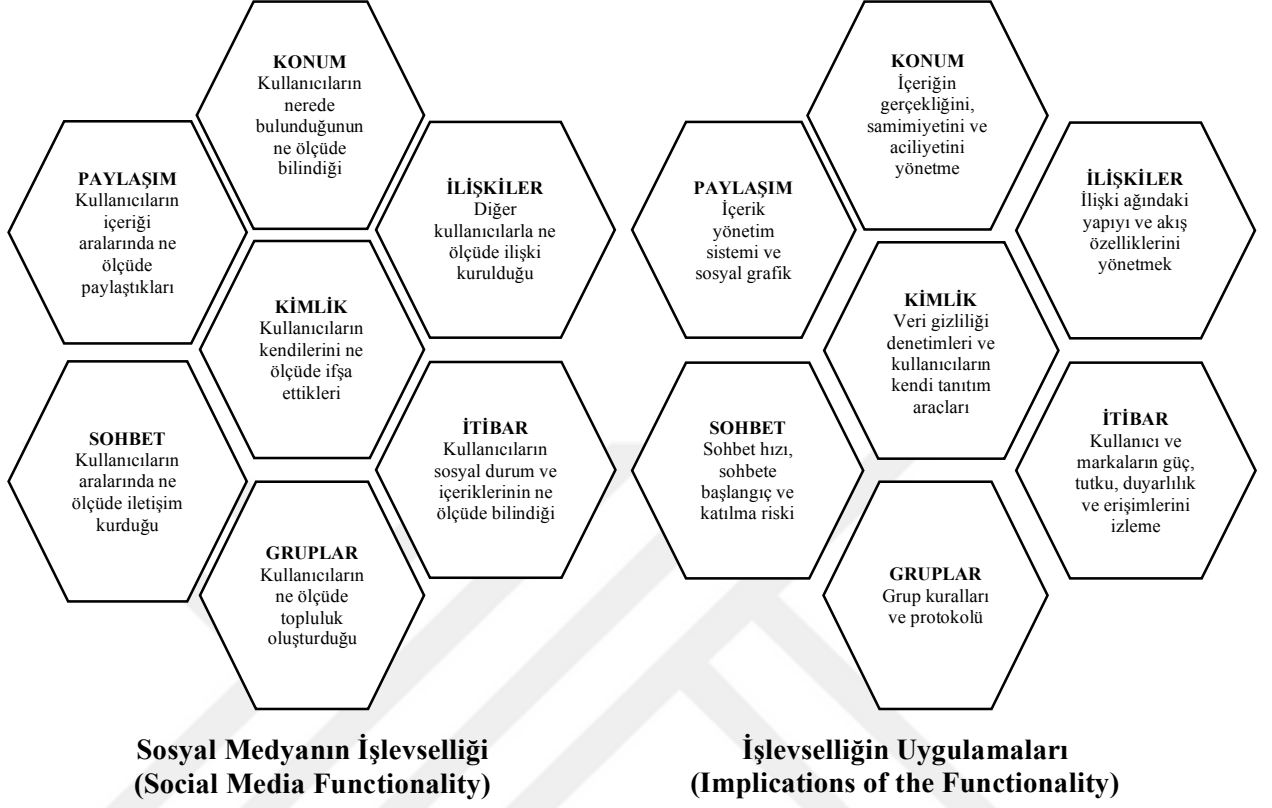
medyadaki gibi bir çaba harcamadan, dijital ortamda, tüm içerikler herkes tarafından erişilebilir, aranabilir ve izlenebilir durumdadır. Bir müşteri ürün ile ilgili olumsuz bir deneyim yaşadığında artık çok kolay bir şekilde düşüncelerini paylaşabilmenin yanında toplum tarafından onaylanma ve destek görebilmektedir.

Yapılan tanımlamalar ve özellik açıklamaları doğrultusunda sosyal medyanın, bireylere tanıdığı özgür ve etkileşimli ortam sayesinde geleneksel medyadan farklılaşarak popülarlığını gün be gün arttırdığı, bireyleri içerik paylaşmaya teşvik eden yapısı ile büyük verinin oluşumuna yol açtığı, çeşitli platformlar üzerinden bireylere birbirleri ile etkileşime geçebilme fırsatı sunduğu ve hatta çevrimiçi ortamda bir araya gelen toplulukların söylemlerini gerçek dünyaya taşımalarına ve bir davranış oluşturmalarına zemin hazırladığı söylenebilir.

1.3.4 Sosyal Medyanın Temel Yapıtaşları

Kietzmann v.d., (2011) sosyal medyayı konum, kimlik, itibar, paylaşım, sohbet, ilişkiler ve gruplar olmak üzere yedi işlevsel yapıtaşından oluşan bir petek görseli şeklinde açıklamışlardır. İşlevsellik bloğu sosyal medya kullanıcı deneyimini açıklarken, işlevselliğin uygulamaları bloğu bunların firmalara olan etkilerini incelemektedir.

Şekil 1.3: Sosyal Medyanın Temel Yapıtaşları



Kaynak: Kietzmann v.d., 2011, 243

Temel yapı taşlarını aşağıdaki gibi açıklamak mümkündür (Kietzmann v.d., 2011:243-248):

- **Kimlik (Identity):** Kimlik yapıtaşı, bir sosyal medya platformundaki kullanıcıların kimliklerini ne ölçüde açıkladıklarını göstermektedir. Kimlik bilgisi içinde ad, yaş, cinsiyet, meslek, yaşanılan yer, ilgi alanları, okuduğu okul ya da çalıştığı firma gibi kullanıcıya ait, onu tanımlayan bilgiler yer almaktadır. Bazı sosyal medya platformları için kimlik yapıtaşı önem teşkil ederken, bazı platformlarda kullanıcılar bilgilerini vermeden takma isimler ile bu platformlarda varlıklarını sürdürür. Örneğin LinkedIn platformunda kullanıcılar, ad, eğitim bilgileri, iş tecrübeleri gibi çeşitli kişisel bilgilerini

açığa vurmaktadırlar. Bu nedenle LinkedIn için kimlik yapıtaşının yapı taşları arasında daha ön planda olduğu söylenebilir.

- **Sohbet (Conversations):** Sohbet yapıtaşı kullanıcıların bir sosyal medya platformundaki diğer kullanıcılarla ne ölçüde iletişim kurduklarını göstermektedir. Çoğu sosyal medya platformunun öncelikli amacı, kullanıcıları arasında sohbet ortamı yaratmaktır. Sosyal medyada kullanıcılar yeni arkadaşlıklar edinme, farklı fikirlerdeki kişilerle tanışma imkânı bulma gibi çeşitli sebeplerle sohbet etmektedirler. Sosyal medya platformlarında sohbet düzeyleri değişkenlik göstermektedir. Örneğin Twitter platformu için sohbet unsuru kimlik unsurundan daha ön plandadır.
- **Paylaşım (Sharing):** Paylaşım yapıtaşı, kullanıcıların sosyal medya platformlarında ürettikleri veya ulaştıkları bir içeriğin paylaşılması anlamını taşır. Sosyal medya, paylaşılan resim, video, metin, ses gibi çeşitli nesneler aracılığıyla birbiriyle etkileşim kuran kişilerden oluşmaktadır. Yapılan paylaşımların sosyal medyada etkileşim kurması paylaşım yapılan platforma göre şekillenmektedir. Sosyal medya platformlarının çeşitli işlevsel hedefleri vardır. Örneğin, Instagram ve Flickr gibi platformlarda paylaşım nesnesi fotoğraf iken, LinkedIn için kariyer içerikli paylaşımlardır. Sosyal medya platformlarına bakıldığı zaman en çok paylaşım ile bağdaştırılan platform olan YouTube için paylaşım unsurunun diğer yapıtaşları arasında daha ön planda olduğu görülmektedir.
- **Konum (Presence):** Sanal ya da gerçek dünyada kullanıcıların nerede oldukları bilgisinin sunulduğu yapıtaşıdır. Örneğin Foursquare platformu için konum yapıtaşı diğer yapıtaşlarından daha önemlidir. Kullanıcılar nerede olduklarının bilgisini sosyal çevreleriyle paylaşmaktadırlar. Bazı platformlarda ise konum unsuru önemli olmamaktadır. Örneğin LinkedIn için bir kullanıcının nerede bulunduğunun bilgisi önem taşımamaktadır.
- **İlişkiler (Relationships):** İlişki yapıtaşı, sosyal medya platformlarındaki kullanıcılar arasındaki ilişkileri kapsamaktadır. Sosyal medyadaki bu ilişkiler, kullanıcıların platform üzerinden birbirleriyle olan paylaşımları, kişilerin

birbirleriyle görüşmeleri ya da sadece birbirlerinin arkadaş listelerinde yer almaları sonucu ortaya çıkmaktadır. İlişkilerin boyutu sosyal medya platformlarına göre farklılaşmaktadır. Bazı platformlarda ilişkiler oldukça resmi olmaktadır, bazı platformlar da ise bu söz konusu değildir. Örneğin LinkedIn platformu üzerinden ilişkiler oldukça resmi bir şekilde gerçekleşirken, Facebook gibi platformlarda ilişkilerde resmiyet mevcut değildir.

- **İtibar (Reputation):** İtibar yapıtaşı, kullanıcıların, kendileri de dâhil olmak üzere diğer kullanıcıların saygınlıklarını ne ölçüde tanımlayabileceklerini ifade etmektedir. İtibar genel anlamda güvenle ilişkili bir kavramdır ve yine diğer yapıtaşlarında olduğu gibi bu yapıtaşı da sosyal medya platformlarına göre farklılık göstermektedir. Örneğin; Twitter için takipçi sayısı önemliyken, LinkedIn’de başka kullanıcılardan gelen destek sayısı belirleyici olmaktadır. Bununla birlikte itibar sadece kullanıcılar için değil içerikler için de söz konusudur. Örneğin YouTube’de videoların görüntülenme sayıları ölçümlenirken, Facebook’da ise içeriklerin beğenilme sayısı değerlendirilir.
- **Gruplar (Groups):** Gruplar yapıtaşı, kullanıcıların, belirli gruplar ya da alt gruplar oluşturma boyutlarını ifade etmektedir. Bir ağ ne kadar sosyalse, o ağdaki arkadaş ya da takipçi gruplarının büyüklüğü o derece artmaktadır. Bu yapıtaşı için iki tür gruptan bahsetmek mümkündür. İlk olarak kişiler sosyal medyadaki arkadaş listelerinde yer alan kişileri yakın arkadaşlar, arkadaşlar, takipçiler ya da hayranlar gibi çeşitli şekillerde sınıflandırarak gruplar oluştururlar. İkinci olarak gerçek yaşamda dâhil olunan gruplara benzerlik göstererek herkese açık, kapalı veya gizli gruplar gibi gruplar oluşturulabilmektedir. Örneğin Facebook’da grubu yöneten, başvuranları onaylayan ve üyelerini katılmaya davet eden yöneticili gruplar bulunmaktadır.

Sosyal medyanın kullanıcılara tanıdığı etkileşimli ortam sonucunda üretilen içerikler farklı yapılara sahip olmaktadır. Örneğin Instagram’da bir fotoğraf veya video paylaşılması sonucunda çeşitli veriler üretilmektedir. Paylaşılan içeriğin beğenilmesi, tekrar paylaşılabilmesi, paylaşım yorumlarının yapılması, paylaşımın

farklı platformlarda da paylaşılabilmesi gibi çeşitli etkileşimler sonucunda aslında başta tek bir içerik paylaşımı olarak görünen kullanıcıların oluşturduğu içerik, beraberinde çok farklı türden verilerin de üretilmesine yol açabilmektedir.

1.4 Yeni Bir Veri Kaynağı Olarak Sosyal Medya

Sosyal medya verileri, sosyal medya kullanıcıları tarafından veya bu kullanıcıların etkileşimleriyle üretilen veriler olarak özetlenebilir (Danneman ve Heimann, 2014:8). Sosyal medya verileri birçok farklı sosyal medya platformundan elde edilen, bünyesinde çeşitli zorlukları barındıran, işlenmeden herhangi bir analizin yapılmasına genellikle olanak tanımayan, ancak işlendiği takdirde firmalar, sağlık kuruluşları, siyasiler, akademisyenler ve daha birçok kişiler, gruplar ve kurumlar adına önemli bulguların, örüntülerin ortaya çıkmasını sağlayan verilerdir. Bu bölümde sosyal medya verileri ve veri dağıtım platformları ele alınmıştır.

1.4.1 Sosyal Medya Verileri

Sosyal medya verileri üç başlık altında ele alınarak açıklanmıştır:

- **Yapılarına göre sosyal medya verileri:** Yapılandırılmış, yapılandırılmamış ve yarı yapılandırılmış veriler bağlamında ele alınarak değerlendirilmiştir.
- **Özelliklerine göre sosyal medya verileri:** Büyük, bağlantılı, gürültülü, heterojen, çok kaynaklı, kısmi, asimetrik gibi literatürde yer alan sosyal medya verilerinin özelliklerine göre yapılan açıklamalara yer verilmiştir.
- **İçeriklerine göre sosyal medya verileri:** Kullanıcıların oluşturduğu içerik türleri sınıflandırması gibi verilerin içerikleri ele alınarak yapılan açıklamalara yer verilmiştir.

1.4.1.1 Yapılarına Göre Sosyal Medya Verileri

Yapıları göz önünde bulundurulduğunda veriler, yapılandırılmış, yapılandırılmamış ve yarı yapılandırılmış olmak üzere üç başlık altında toplanmaktadır. Yapılandırılmış veriler (structured data) aynı tanımlanmış formata sahip, aynı düzeni izleyen veriler olarak açıklanabilmektedir. Herhangi bir format veya kurala dâhil olmayan veriler ise yapılandırılmamış veriler (unstructured data) olarak tanımlanmaktadır. Bu verilere örnek olarak metin, video, ses ve görüntüler verilebilir. Yarı yapılandırılmış veriler (semi structured data) ise içerisinde hem düzenli hem de düzensiz verileri barındırmaktadır. E-posta içeriği göz önünde bulundurulduğunda, yapılandırılmamış formata sahip olan metin bölümü dışında, yazar, konu ve özet gibi belirli bir düzene sahip bölümler de bulunmaktadır (Gürsaka, 2014:54-56). Bu açıdan ele alındığında sosyal medya verilerinin hem yapılandırılmış hem de yapılandırılmamış verileri içerdiği söylenebilir.

Sosyal medyadaki meta veriler yapılandırılmış veri olarak adlandırılmaktadır. Bu verilere kullanıcıların profillerinde yer alan bazı bilgilerle birlikte beğeni ve yorum sayısı gibi veriler örnek verilebilir. Yapılandırılmamış veriler arasında ise Facebook yorumlarından blog içeriklerine ve görsel-işitsel materyallere kadar değişen aralıkta yer alan, kullanıcı tarafından üretilen içerikler bulunmaktadır (Stieglitz v.d., 2014:91). Bir kullanıcının oluşturduğu sosyal medya profilindeki çalıştığı kurum, meslek bilgisi gibi demografik bilgileri, paylaşımlarının, beğenilerinin ve yorumlarının sayısı yapılandırılmış veri olarak tanımlanırken, bu kişinin paylaştığı metin, görsel, video gibi yapılara sahip içerikler ise yapılandırılmamış verileri içerir.

Yapılandırılmış ve yapılandırılmamış veriler olarak ele alındığında sosyal medya verileri platform temelli değişkenlik göstermekle birlikte genel bir bağlamda ele alınabilir. Yapılandırılmamış veriler genel olarak kullanıcıların oluşturduğu içeriğin tek bir ortak yapıda olmayışı olarak özetlenebilir. İletiler platform temelli değişiklik

göstermekle birlikte metin, görsel, video ve link gibi çeşitli yapıdaki verileri içerebilmektedir. Yapılandırılmış verilere ise aşağıdaki gibi örnekler verilebilir:

- Bir kullanıcının sosyal medya platformundaki ileti sayısı,
- İletilerin tarih ve zaman bilgileri,
- İletilere gelen beğeni ve yorum gibi platform temelli değişkenlik gösteren nicel bilgiler (örneğin Twitter platformu için retweet sayısı, Instagram platformu için bir videonun görüntülenme sayısı),
- Platform temelli değişkenlik göstermekle birlikte kullanıcının demografik bilgileri (örneğin LinkedIn platformunda kişi isim soy isim bilgileri, mesleği gibi bilgiler),
- Platform temelli değişkenlik göstermekle birlikte kullanıcıların platformlar arasındaki bağlantılar (örneğin Facebook platformunda kimlerin kimlerle arkadaş olduğu, Twitter platformunda kimlerin kimleri takip ettiği gibi bilgiler).

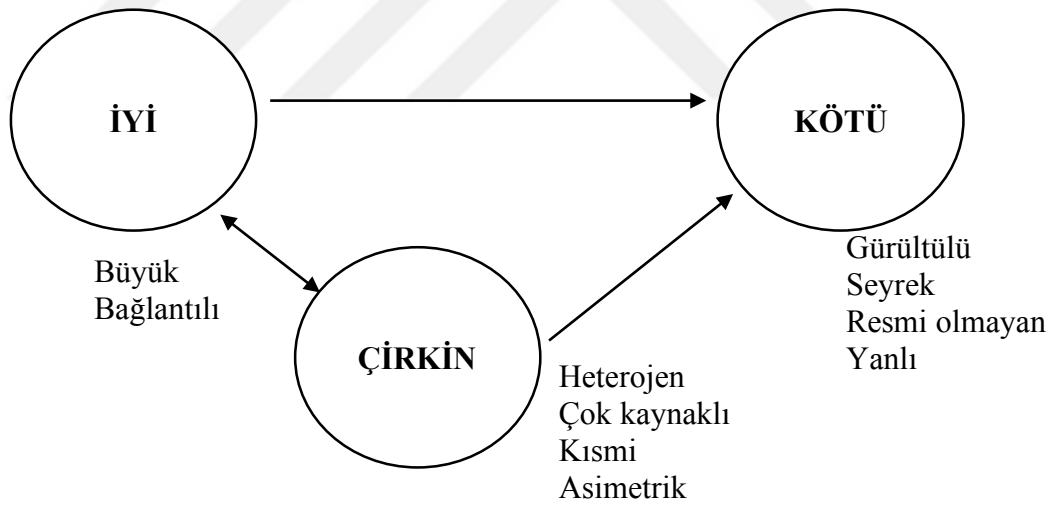
1.4.1.2 Özelliklerine Göre Sosyal Medya Verileri

Sosyal medya verileri özelliklerine göre değerlendirildiğinde, literatürde verilerin üretimi ve dağılımının çeşitli şekillerde açıklandığı görülmektedir. Sosyal medya, kullanıcıların oluşturduğu içerikler ve kullanıcılar arasındaki ilişkiler çerçevesinde değerlendirildiğinde oluşan veriler, Pareto İlkesi ile açıklanabilmektedir. Sosyal medya platformlarında, birçok kullanıcı içerik üretmekten ziyade içerik tüketicisi konumundadır. Örneğin, Twitter'da, kullanıcıların %40'ından fazlası hiç tweet atmamıştır. Kullanıcılar arasındaki ilişkilere bakıldığında, pek çok kullanıcının birkaç arkadaşı ya da bağlantısı olduğu görülmektedir. Kullanıcı bağlantıları dikkate alındığında, birkaç kullanıcının çok sayıda takipçisinin/arkadaşının olması ve çok sayıda kullanıcının ise birkaç arkadaşının bulunması durumu gözlenmektedir ki, bu Pareto İlkesi ile açıklanabilmektedir (Liu v.d., 2016:141).

Pareto İlkesine benzer şekilde 90-9-1 kuralı, sosyal medyada yer alan bireylerin %90'ının yalnızca gözlemleyerek katılımda bulunmadığını, %9'unun az miktarda katılımda bulunduğunu ve %1'inin üretilen yeni içeriklerin büyük çoğunluğunu oluşturduğunu söylemektedir. %90'lık bölümde yer alan kullanıcılar gizli meraklılar (lurkers), %9'luk bölümde yer alanlar katılımcılar (contributors) ve %1'lik bölümde yer alanlar ise yetkili kullanıcılar (superusers) olarak adlandırılmaktadır (van Mierlo, 2014:2).

Bir başka değerlendirmeye göre sosyal medya verileri büyük, bağlantılı, gürültülü ve eksik veriler olarak tanımlanmıştır (Tang v.d., 2014:20). Liu v.d. (2016) ise sosyal medya verilerini taşıdıkları özellikler bağlamında iyi, kötü ve çirkin tanımlamalarıyla açıklamışlardır.

Şekil 1.4: İyi, Kötü ve Çirkin Bağlamında Sosyal Medya Verileri



Kaynak: Liu v.d., 2016:138

Sosyal medya verilerinin “iyi” olarak tanımlanması, verilerin büyük ve bağlantılı olmasından kaynaklanmaktadır. Sosyal medya verilerinin büyük olması tek başına bir anlam ifade etmez iken, bu verilerin ilişkilendirilebilmesi birçok anlamda fayda sağlayabilmektedir (Liu v.d., 2016:138). Sosyal medyada var olan bireylerin az

miktarda içerik ürettikleri düşünülse de, sosyal medya verilerinin toplumsal özelliği ile bireylerin verileri birbirine bağlanmaktadır. Bunun sonucunda oluşan veri yığını yeni bir tür büyük veri sağlamaktadır (Tang v.d., 2014:21). Sosyal medya verileri için “Sosyal Medya Büyük Verisi (Social Media Big Data)” (Tufekci, 2014), “Büyük Sosyal Veri (Big Social Data)” (Stieglitz v.d., 2014:90) ve Sosyal Büyük Veri (Social Big Data)” (Bello-Orgaz v.d., 2016) gibi adlandırmaların literatürde yer aldığı görülmektedir.

Sosyal medya verilerinin “kötü” olarak tanımlanması ise verilerin gürültülü ve seyrek olmasından kaynaklanmaktadır. Bunun yanında verilerin konuşma dili ile yazılması ve yanlış olması da bu bağlamda değerlendirilmektedir. Verilerin gürültülü olmasının sebebi, kendilerini özgürce ifade eden kullanıcılar tarafından üretilmesidir. Kullanıcılar diledikleri içerikleri üretebilmekte ve dolayısıyla üretilen içerikler anlamsız ya da eksik ifadeler içerebilmektedir. Verilerin yetersizliği bir başka deyişle verilerin eksikliği ise sosyal ağların güç yasası dağılımı ile açıklanmaktadır. Buna göre azınlıktaki veriler zenginken, uzun kuyruktaki veriler daha az içerik ve bağlantı içermektedir (Liu v.d., 2016:138).

Sosyal medya verilerinin “çirkin” olarak tanımlanması verilerin heterojen, çok kaynaklı, kısmi ve asimetrik olmasından kaynaklanmaktadır. Kullanıcılar, kişiler, gruplar veya yayınlarla farklı bağlantılara sahip olabilmekte dolayısı ile bu verilerin heterojen bir yapıda oluşmasını sağlamaktadır. Veriler metinden, URL'lerden, ses, görüntü veya video gibi çeşitli birçok farklı yapıdan oluşabilmektedir. Verilerin kısmi olması kullanıcıların farklı sosyal medya platformlarında, yalnızca her servis için kısmi bilgi sağlamasıyla ilgilidir. Asimetrik kavramı ise bağlantılar arasındaki ilişkiler için kullanılmaktadır. Örneğin bir kullanıcı diğer bir kullanıcıyı takip ederken, takip ettiği kişi onu takip etmeyebilir (Liu v.d., 2016:138).

1.4.1.3 İeriklerine Gre Sosyal Medya Verileri

Sosyal medya verileri ierik temelli ele alındığında eřitli sınıflandırmalar ortaya ıkmaktadır. Bu veriler, ieriklerine gre řu řekilde sınıflandırılabilir (Wamba v.d., 2016:6):

- **Demografik (Demographic) Veriler:** Demografik veriler sosyal medya kullanıcılarının yař, cinsiyet, eēitim ve konum gibi paylaşılan bilgilerini iermektedir. Kullanıcıların bu bilgileri sosyal medya profillerinde kamuya aık bir řekilde grlmektedir. Demografik verilerin paylaşımı platformdan platforma temelli deēiřkenlik gsterebilmektedir.
- **rn (Product) Verileri:** rn verileri sosyal medya kullanıcılarının belirli bir markadan veya rnden bahsetmesi yoluyla retilir. Tketiciler kullandıkları bir rn ile ilgili sorunlarını ya da beklentilerini sosyal medya yolu ile paylaşabilmektedirler. Belli bir marka veya rn hakkında yapılan bu yorumlar, markanın sosyal medya sayfasında veya sosyal medya kullanıcısının kiřisel sayfasında grlebilir.
- **Psikografik (Psychographic) Veriler:** Psikografik veriler, bir rn veya bir markayla/rnle ilgili tketicilerin deēerlerini, tutumlarını, ilgi alanlarını ve yařam biimlerini bildiren verilerdir. Firmalar bu verileri analiz ettikleri takdirde tketiciler ile ilgili nemli bilgiler elde edebilmekte, tketicilerin neleri sevdiēini ve nelerden hořlanmadıēını anlayarak daha gereki bir řekilde stratejiler oluřturulabilmektedir. Bu nedenle bu verilerin anlık takip edilmesi, zellikle mřteri odaklı kuruluřlar aısından byk nem tařımaktadır.
- **Davranıřsal (Behavioral) Veriler:** Davranıřsal veriler sosyal medya platformlarındaki tketicilerin satın alma davranıřlarını temsil etmektedir. Sosyal medyadaki bireylerin davranıřsal verileri analiz eden řirketler, hedef mřterilerini tanımlayabilir ve gelecekteki satın alma eylemlerini ngrebilir.

- **Tavsiye (Referrals) Verileri:** Tavsiye verileri, sosyal medya platformlarında paylaşılan oyların ve yorumların verisidir. Bu tür veriler, sosyal medyada olumlu ya da olumsuz ağızdan ağıza iletişim yolu ile yayılan bilgileri içermektedir. Sosyal medyada olumlu bir sözlü ifade üretmek için, tüketicilerin tavsiye verilerinin analizi önem taşımaktadır.
- **Konum (Location) Verileri:** Konum verileri, tüketicilerin coğrafi konumu ile ilgili bilgilerden oluşur. Foursquare gibi sosyal medya platformlarının bazıları kullanıcıların mevcut konumlarını paylaşımları üzerine kuruludur. Bu konum verileri, şirketlere etkili stratejiler kurmak için büyük fırsatlar sağlayabilmektedir. Örneğin bir firma, müşterilerinin mevcut coğrafi konum verilerini analiz ederek, düzenleyeceği etkinliği daha iyi organize edebilir.
- **Niyet (Intention) Verileri:** Niyet verileri, kuruluşların ve şirketlerin bir ürün veya bir markayla ilgili tüketicilerin beklentilerini tahmin etmelerine ve onlarla ilgili gelecekteki faaliyetlerdeki davranışlarını saptamada yardımcı olabilecek verilerdir.

Sosyal medya verileri kullanıcıların oluşturduğu içeriklerin toplamı olarak görüldüğünden dolayı, kullanıcıların oluşturduğu içerik türleri sınıflandırması da bu veri türlerinin içeriklerini anlamak için yardımcı olmaktadır. Verileri içeriklerine göre metin, fotoğraf ve resimler, müzik ve ses, video ve film, vatandaş gazeteciliği, eğitim içeriği, mobil içerik ve sosyal içerik başlıkları altında sınıflandırmak mümkündür (OECD, 2007:34).

Tablo 1.3: Sosyal Medya Verilerinin İçeriklerine Göre Sınıflandırılması

İçerik Türü	Açıklama	Örnekler
Metin	Kullanıcıların paylaştıkları orijinal yazılar ya da yayılan diğer metin, roman ve/veya şiirler	Fanfiction.net, Quizilla.com, Writely
Fotoğraf ve resimler	Kullanıcılar tarafından çekilen ve çevrimiçi yayınlanan dijital fotoğraflar; kullanıcılar tarafından oluşturulan veya değiştirilen fotoğraflar veya resimler	Ofoto ve Flickr gibi sitelerde yayınlanan fotoğraflar; Fotoğraf blogları
Müzik ve ses	Kendi müzik/ses içeriğini kaydetmek ve/veya düzenlemek ve dijital formatta yayın yapmak, birleştirmek ve/veya dağıtmak	Remixler, müzik gruplarının Web sitelerinde veya MySpace sayfalarında evde kaydedilen müziklerin paylaşımı, podcastler
Video ve film	Mevcut içerikler, üretilen içerikler ve bunların ikisinin karışımı da dâhil olmak üzere video içeriğini kaydetmek ve/veya düzenlemek ve yayınlamak	Film fragmanı remixleri; video blogları ve videocasting; YouTube ve Google Video; Current TV
Vatandaş gazeteciliği	Vatandaşlar tarafından yapılan güncel olaylarla ilgili gazetecilik raporları (haber hikâyeleri, blog yazıları, güncel olayların fotoğraflarının veya videolarının çekilerek çevrimiçi yayınlanması)	OhmyNews, GlobalVoices ve NowPublic gibi siteler; haber değeri olan olayların fotoğrafları ve videoları; bir etkinliğin sitesindeki blog yayınları; CNN Exchange gibi işbirlikçi oluşumlar
Eğitim içeriği	Okullarda, üniversitelerde veya eğitim amacıyla oluşturulmuş içerik	H20 gibi öğretim programı paylaşım siteleri; Wikibooks, OpenCourseWare
Mobil içerik	Genellikle diğer kullanıcılara MMS, e-posta veya İnternet aracılığıyla gönderilen, cep telefonlarında veya diğer kablosuz cihazlarda yaratılan kısa mesaj, fotoğraf ve video gibi içerikler	Geleneksel medyanın erişemeyeceği doğal felaketler gibi kamu etkinliklerinin videoları ve fotoğrafları; siyasi örgütlenmede kullanılan kısa mesajlar
Sanal içerik	Çevrimiçi sanal ortam bağlamında oluşturulmuş ya da bu ortamla bütünleşmiş içerikler	Giysi, ev ve sanat eserleri de dâhil olmak üzere Second Life'da geliştirilebilir ve satılabilir çeşitli sanal ürünler

Kaynak: OECD, 2007:34

1.4.2 Sosyal Medya Veri Dağıtım Platformları

Sosyal medya platformlarının sayısı gün geçtikçe daha fazla artarken, bazı platformlar popülerliğini yitirmekte, bazıları ise yükselişe geçmektedir. Sosyal medya verilerinin üretildiği bu platformların çeşitli sınıflandırmaları mevcuttur. Bu sınıflandırmalarda hem sosyal ağlar gibi hemen hemen herkes tarafından kabul görmüş ortak başlıklar, hem de kişiler arası uygulamalar gibi azınlıktaki araştırmacılar tarafından sosyal medya platformu olarak değerlendirilen başlıklar yer almaktadır.

Tablo 1.4: Sosyal Medya Veri Dağıtım Platformları Sınıflandırmaları

(Mayfield, 2008:6)	Sosyal Ağlar (Myspace, Facebook, Bebo) Bloglar Wikiler (Wikipedia) Podcastler (Apple iTunes) Forumlar İçerik Toplulukları (Flickr, Delicious, YouTube) Mikrobloglar (Twitter)
(Lietsala ve Sirkkunen, 2008:29)	İçerik Oluşturma ve Yayınlama Araçları; Bloglar, wikiler ve podcastler Sosyal Ağ Siteleri İçerik Paylaşım Siteleri (Flickr, YouTube, Delicious, Digg, Dopplr) İşbirlikçi Üretim Siteleri Sanal Dünyalar (Second Life, Habbo Hotel, World of Warcraft) Eklenti Siteleri (iLike)
(Mangold ve Faulds, 2009:358)	Sosyal Ağ Siteleri (MySpace, Facebook, Faceparty) Yaratıcı İçerik Paylaşım Siteleri: Video Paylaşım Siteleri (YouTube) Fotoğraf Paylaşım Siteleri (Flickr) Müzik Paylaşım Siteleri (Jamendo.com) Yardımcı İçerik Paylaşımı (Piczo.com) Genel Fikri Mülkiyet Paylaşım Siteleri (Creative Commons) Kullanıcı Sponsorlu Bloglar (Apple Weblog, Cnet.com) Şirket Tarafından Sponsor Olunan Web Siteleri/Bloglar (Apple.com, P&G's Vocalpoint) Şirket Tarafından Desteklenen Etkinlik/Yardım Siteleri (Dove's Campaign for Real Beauty, click2quit.com) Davet Temelli Sosyal Ağlar (ASmallWorld.net) İş Ağı Siteleri (LinkedIn)

	İşbirlikçi Web Siteler (Wikipedia) Sanal Dünyalar (Second Life) Ticaret Toplulukları (eBay, Amazon.com, Craig's List, iStockphoto, Threadless.com) Podcastler Haber Dağıtım Siteleri (Current TV) Eğitim Materyali Paylaşımı (MIT OpenCourseWare, MERLOT) Açık Kaynaklı Yazılım Toplulukları (Mozilla's spreadfirefox.com, Linux.org) Sosyal İmleme Siteleri (Digg, Delicious, Newsvine, Mixx it, Reddit)
(Safko ve Brake, 2009:449-655)	Sosyal Ağlar Yayıncılık (Bloglar, Wikiler) Resim Paylaşımı (Flickr) Müzik Paylaşımı (Podcastler) Video Paylaşımı (Vlog, YouTube) Mikrobloglar (Twitter) Canlı Yayıncılık (Justin.Tv) Sanal Dünyalar (Second Life) Oyun Siteleri (World of Warcraft) Verimlilik Uygulamaları (BitTorrent, SurveyMonkey) Toplayıcılar (Digg, FriendFeed) RSS (Atom, Google FeedBurners) Arama (Technorati) Mobil Sosyal Medya (Akıllı Telefon Uygulamaları) Kişilerarası (Skype, Viber)
(Kaplan ve Haenlein, 2010:62)	İşbirlikçi Projeler (Wikipedia) Bloglar İçerik Toplulukları (YouTube) Sosyal Ağlar (Facebook) Sanal Oyun Dünyaları (World of Warcraft) Sanal Sosyal Dünyalar (Second Life)
(Gundecha ve Liu, 2014:3)	Çevrimiçi Sosyal Ağlar (Facebook, Myspace, LinkedIn) Bloglar (Huffington Post, Business Insider, Engadget) Mikrobloglar (Twitter, Tumblr, Plurk) Wikiler (Wikipedia, Wikitravel, Wikihow) Sosyal Haber Siteleri (Digg, Slashdot, Reddit) Medya Paylaşım Siteleri (YouTube, Flickr, UstreamTV) Görüş, İnceleme ve Derecelendirme Siteleri (Epinions, Yelp, Cnet) Cevap Siteleri (Yahoo! Answers, WikiAnswers)
(Zafarani v.d., 2014:21)	Sosyal Ağlar (Facebook, LinkedIn) Mikrobloglar (Twitter) Fotoğraf Paylaşım Siteleri (Flickr, Photobucket, Picasa) Haber Toplama Siteleri (Google reader, StumbleUpon, Feedburner) Video Paylaşım Siteleri (YouTube, MetaCafe)

Canlı Yayın Siteleri (Ustream, Justin.TV)
Sanal Dünyalar (Kaneva)
Sosyal Oyunlar (World of Warcraft)
Sosyal Arama Siteleri (Google, Bing, Ask.com)
Anlık Mesajlaşma Uygulamaları (Google Talk, Skype, Yahoo! messenger)

Kullanıcılar çeşitli sınıflandırmalara tabi tutulan bu sosyal medya platformları üzerinden paylaşımlarda bulunabilmekte, diğer kullanıcılar tarafından yapılan paylaşımları ise beğenebilmek, paylaşabilmek, bu paylaşımlara yorum yapabilmek gibi kullandığı sosyal medya platformunun özelliklerine göre birbirinden farklı etkileşimlere yol açabilmektedir.

1.4.2.1 Sosyal Ağlar

Boyd ve Ellison (2007: 211) tarafından sosyal ağlar, kullanıcılara sınırları belirli bir sistemde kamuya açık ya da yarı açık profiller oluşturabilmesine, sistem içerisinde kendilerinin ve diğer kullanıcıların bağlantı listelerini görüntülemelerine olanak sağlayan Web temelli hizmetler olarak tanımlanmıştır. Sosyal ağ siteleri, çevrimiçi ortamlarda bireylerin hem yeni insanlarla tanışmalarını hem de önceden tanıdıkları kişilerle iletişim kurmalarını sağlayan Web siteleridir (Zarrella, 2009:53).

1997 yılında kurulan sosyal ağların ilk örneği olan SixDegrees.com adlı site (Boyd ve Ellison, (2007:214) ve beraberinde kurulan sitelerle birlikte kullanıcılar bir profil aracılığı ile sosyal ağlarda varlıklarını sürdürmekte, ilgi alanları ve amaçları doğrultusunda paylaşım yapabilmekte ve diğer kullanıcılarla etkileşimde bulunabilmektedir. Sosyal ağlar diğer sosyal medya platformları özelliklerinin çoğuna sahip olmasına rağmen, asıl amaçları temelinde bu platformlardan ayrılmaktadırlar. Örneğin bir sosyal ağ olan Facebook üzerinden, fotoğraf paylaşım platformu olan Instagram'daki gibi görsel paylaşımı, video içerik platformu olan YouTube'deki gibi video paylaşımı yapılabilme olmasına karşın, Facebook gibi sosyal ağların asıl amacı kullanıcılar arasında bağlantı kurmaktır.

Dünyada en çok kullanılan sosyal medya platformu olan, 2004 yılında Mark Zuckerberg tarafından kurulan Facebook en popüler sosyal ağ sitesidir (We Are Social ve Hootsuite, 2018). Temelinde bireylerin başka kişilerle iletişim kurması ve bilgi paylaşımı olan bu sosyal ağ, popülerliğini yıllar boyunca korumaktadır. Kullanıcılar oluşturdukları profillerinde fotoğraflarını, ilgi alanlarını, isim, doğum tarihi, cinsiyet gibi bilgilerini, gizli ya da açık mesajlarını ve arkadaş gruplarını sergilemektedirler. Bu platform üzerinden kullanıcılar diğer kullanıcıların profillerini gezebilmekte, yapılan paylaşımların altında bulunan “beğen”, “yorum yap” ve “paylaş” butonları ile bu paylaşımlara yorum yapabilmekte, paylaşımları beğenebilmekte ve kendileri de paylaşabilmektedir. Facebook platformu üzerinden paylaşılan herhangi bir veri metin, görsel, video ve diğer Web sitelerine bağlantılar içerebilmektedir.

Her kullanıcının kullanımına açık olan Facebook sosyal ağ sitesi gibi genel siteler bulunmakla birlikte, daha çok iş deneyimlerinin paylaşıldığı profesyonel sosyal ağ siteleri olarak tanımlanan LinkedIn gibi sosyal ağlar da mevcuttur (Kietzmann v.d., 2011:242). 2003 yılında kullanıma açılan bu sosyal ağ üzerinde kullanıcılar, iş tecrübelerini, eğitim bilgilerini, katıldıkları projeleri, yayınlarını ve yetkinlikleri gibi bilgilerini diğer kullanıcılar ile detaylı bir şekilde paylaşabilmektedirler. Kullanıcılar profesyonel sosyal ağlarını oluşturabilmekte ve özgeçmişlerine uygun kariyer fırsatlarından haberdar olabilmektedir. Bağlantılı olunan diğer kullanıcıların profilleri ziyaret edilebilir, diğer genel platformlardan farklı olarak bu ziyaret kaydedilir ve karşı tarafla bu ziyaretin bilgileri paylaşılır.

1.4.2.2 İçerik Paylaşım Toplulukları

Sosyal ağlardan farklı olarak içerik topluluklarının temel amacı, kullanıcılar arasında medya içeriğinin paylaşılmasıdır. İçerik toplulukları, BookCrossing gibi metin, Flickr gibi fotoğraf, YouTube gibi video ve Slideshare gibi PowerPoint sunuları paylaşım platformlarıdır. Bu platformlarda genellikle kullanıcıların kişisel profil sayfalarından

ziyade, topluluğa katıldıkları tarih ve paylaşılan videolar, fotoğraflar gibi temel içerikler mevcuttur (Kaplan ve Haenlein, 2010:63).

Ülkemizde en çok kullanılan sosyal medya platformu olan YouTube (We Are Social ve Hootsuite, 2018), televizyon, müzik ve film gibi geleneksel eğlence seçeneklerinin birleşimi olarak tanımlanabilmektedir. Bununla birlikte YouTube kullanıcıları, geleneksel medyadan “en çok görüntülenen” veya “en çok oy alan” video gibi çok daha fazla seçeneğe sahiptir (Shao, 2009:12). Bu platform üzerinde kullanıcılar beğenileri ile birlikte beğenmedikleri videoları belirtebilmekte, videolara yorum yapabilmekte, videoları paylaşabilmekte ve kendileri de video yükleyebilmektedir (Khan, 2017:238).

Popüler bir fotoğraf (video) paylaşım platformu olan Instagram kullanıcılara, akıllı telefonları aracılığı ile paylaşmak istedikleri resimleri ve videoları çeşitli filtreler uygulamasına ve bunları eş zamanlı başka sosyal medya platformunda paylaşmasına izin verir. Kullanıcıların resimlerini ve videolarını “caption” adı verilen bir açıklama ile tamamlamasına, açıklamalarda “#” simgesini kullanarak hashtag eklemesine, bu içerikleri paylaşmadan önce “@” simgesini kullanarak diğer kullanıcılardan söz etmelerine ya da onları etiketlemelerine izin verir. Kullanıcıların aralarında bağlantı kurması Twitter ile benzerlik göstermektedir. Kullanıcıların birbirlerini takip etmesi bağımsızdır; A kullanıcısı B kullanıcıyı takip ettiğinde B kullanıcısı A kullanıcıyı takip etmek zorunda değildir (Hu v.d., 2014:596).

1.4.2.3 İşbirlikçi Platformlar

İşbirlikçi platformlar kullanıcıların anlık ve birlikte içerik oluşturması üzerine kurulu olan, kullanıcıların oluşturduğu içeriğin ortaya çıkması şeklinde tanımlanabilen, içerisinde wikileri, sosyal haber ve imleme sitelerini bulunduran platformlar şeklinde özetlenebilir (Kaplan ve Haenlein, 2010: 62). Kaplan ve Haenlein (2014: 620) tarafından bu platformlar, wikiler, forumlar, sosyal imleme siteleri ve değerlendirme

siteleri olmak üzere dört başlıkta sınıflandırılmıştır. Bu sınıflandırmalara ek olarak Ekşi Sözlük gibi katılımcı sözlükler eklenebilir. Bir işbirlikçi hiper metin sözlüğü olarak tanımlanmakta olan Ekşi Sözlük platformu, 1999 tarihinde Sedat Kapanoğlu tarafından geliştirilmiştir. Bu platformda kayıtlı yazarlar herhangi bir konu ile ilgili, açtıkları girişler (entry) aracılığı ile içerik üretebilmektedirler (Gürel ve Yakın, 2007:204).

Adını Hawaii dilinden alan, hızlı anlamına gelen wikiler kullanıcılara, basit bir Web tarayıcısı kullanarak bir Web sayfasındaki içeriğe ekleme yapma, silme veya düzeltme olanağı sağlamaktadır (Safko ve Brake, 2009:159). Wikilerin en yaygın temsilcisi ise Wikipedia'dır. Bir diğer işbirlikçi platform olan sosyal imleme siteleri ile kullanıcılar etiketli bulutlar şeklinde düzenlenebilen Web belgelerine ait yer imlerine etiketler atayabilmektedirler. Bu şekilde istedikleri siteleri işaretleyerek diğer kullanıcıların da bunları görmesine olacak sağlarlar. 2003 yılında kurulan Delicious, en bilinen sosyal imleme sitesidir (Kaplan ve Haenlein, 2014: 619).

İşbirlikçi platformlar arasında yer alan forumlar, belirli bir konu hakkında devam etmekte olan çevrimiçi tartışmalara izin veren Web siteleridir. Kullanıcılar, forum sahibinin kurduğu kurallara ve düzenlemelere uyulduğundan emin olmak için sohbeti izleyen yöneticilerden, katkıda bulunanlardan ve üyelerden oluşan bir gruptur (Safko ve Brake, 2009:147-148). Forumlar gibi değerlendirme siteleri de kullanıcıların anlık ve birlikte sağladıkları katkılardan oluşmaktadır. İşletmeler, ürünler veya hizmetler gibi birçok konu hakkında değerlendirmelerin yayınlanabileceği web siteleri olan bu platformlara örnek olarak Epinions.com ve Amazon.com verilebilir (Saravanakumar ve SuganthaLakshmi, 2012:4448).

1.4.2.4 Sosyal Yayıncılık Platformları

Kişisel Web sayfalarının sosyal medya eşdeğeri olarak görülen bloglar üzerinde içerikler genellikle tarih bilgisine göre ters kronolojik sırayla listelenir. Genellikle

yalnızca bir kiři tarafından yönetilen bloglar, başkaları tarafından yorum eklenebilmesi ile etkileřim imkânı saęlanır. Birçok farklı versiyonu bulunmasına karřın metin tabanlı bloglar, en yaygın olarak kullanılandır (Kaplan ve Haenlein, 2010: 63).

Bir blog türü olan mikroblogger ise kiřilerin düşüncelerini belirli bir karakter sayısı sınırlandırması dâhilinde paylaşabildikleri platformlardır. Mikroblogger, kullanıcılara, faaliyetleri, görüşleri, düşünceleri doğrutusunda bilgi yayınlayıp, paylaşmalarına olanak tanımaktadır. Bloglardan farklı olarak mikroblogger, kullanıcılara sürekli olarak kısa mesajlar şeklinde içerik üretmelerine izin verir ve daha hızlı bir iletişimi teşvik eder (Java v.d., 2007:56-57).

Asıl amacı belirli bir karakter sınırlaması çerçevesinde kullanıcıların “ne yaptığını” cevaplaması temelleri üzerine kurulu bir mikroblogger hizmeti olan Twitter platformunda (Weinberg, 2009:125), takip eden ve takip edilen arasındaki ilişki karşılık gerektirmez; bir kullanıcının başka bir kullanıcıyı takip etmesi, takip edilen kullanıcının kendisini takip eden kişiyi geri takip etmesini gerektirmez. Bir kullanıcı dięer kullanıcıyı takip ederek, onun attığı “tweet” olarak adlandırılan tüm mesajlardan güncel olarak haberdar olur (Kwak v.d, 2010:591). Önceleri 140 olan karakter sınırı 2017 yılı sonlarına doğru iki katına çıkarılarak 280 olmuştur. Atılan tweetler kullanıcılar tarafından paylaşılmak istenirse “retweet” düğmesini kullanarak aynı tweet tekrar paylaşabilmektedirler. Kullanıcılar birbirlerine özel mesajlar gönderebilir, tweet içeriklerinde “#” sembolü (hashtag) kullanarak dięer kullanıcıların ve kendisinin belirli bir konuyu/kelimeyi kolayca bulmalarını sağlayabilir ve “@” işareti ile belirli bir kullanıcıdan bahsedebilir ya da kullanıcıya özel bir tweet atabilir (Gunelius, 2011:82-83).

1.4.2.5 Podcast

Adını iPod ve yayın (broadcast) birleşiminden alan podcastler, İnternette yayınlanan ve abone olunabilen ses ve medya dosyalarıdır. Tařınabilir cihazlarda ve kişisel

bilgisayarlarda yeniden oynatılabilmesi için podcastlerde genellikle besleme akışları kullanılır. Abonelik özelliği sayesinde podcastler, toplulukların oluşmasına yardımcı olmaktadır (Lietsala ve Sirkkunen, 2008:41).

Podcast hem içeriğin kendisi hem de podcasting olarak da adlandırılan besleme yöntemi anlamına gelebilmektedir. Podcast temelinde kullanıcılara, istediklerini, istedikleri yerde dinlemelerine olanak sağlayan ses ve video içeriği yaratmaktır. Ses ve video dosyaları, kullanıcı müdahalesi olmaksızın abonelere iletilebilmektedir. Bu yolla, dosyaları aramak ve indirmek için çaba sarf edilen sıkıntılı ve sıkıcı işlemi ortadan kaldırır. Kullanıcılar daha sonra istedikleri yerde bilgisayarlarında ya da taşınabilir cihazlarına indirerek podcast dinleyebilir veya izleyebilir (Jham, 2008:278-279).

1.4.2.6 RSS

Zengin Site Özeti (Rich Site Summary), RDF Site Özeti (RDF Site Summary) ve Çok Basit Besleme (Really Simple Syndication) olarak tanımlamaları bulunan RSS, kullanıcıların, abone olarak bir web sitesinden içeriği periyodik bir şekilde otomatik olarak almasına ve başkalarının abone olabileceği Web yayınları oluşturmaya izin veren bir XML kod tabanlı içerik besleyicisidir (Holvoet, 2006:32).

İnternette yeni şeyler oluşturup yayınlandığında, güncelleme isteyen herkese bir bildirim gönderilir, aboneler e-posta, mesaj veya tweet aracılığıyla otomatik olarak bilgilendirilir ve içerik otomatik olarak okuyucuya veya toplayıcı sayfasına eklenir (Safko ve Brake, 2009:345-346). Atom, FeedBurner, PingShot ve RSS 2.0 bu sitelere örnek olarak verilebilir (Safko ve Brake, 2009:616-622).

1.4.2.7 Sanal D nyalar

Sanal d nyalar  zerinde kullanıcılar, ki iselle tirilmi  avatarlar aracılı ıyla di er kullanıcılar ile etkile im i erisinde bulunurlar. Ger ek hayatın sanal ortama aktarılması ile birlikte bu platformlar  evrimi i bir    boyutlu ortam yaratırlar (Kaplan ve Haenlein, 2010: 64). Sanal d nyalar, i erisinde World of Warcraft gibi oyun d nyalarını ve Second Life gibi sosyal d nyaları barındırır. Di er sosyal medya platformlarından farklı olarak sanal d nyalar kullanıcılara deneyim, e lence ve heyecan sunmaktadır (Lietsala ve Sirkkunen, 2008:52).

Sanal oyun d nyaları ve sosyal sanal d nyalar arasında kullanıcılara tanınan  zg rl k boyutu a ısından farklılık bulunmaktadır. Sanal oyun d nyalarında kullanıcılar,  ok katılımcılı  evrimi i rol yapma oyunu (MMORPG) ba lamında uyulması gereken kurallar d hilinde hareket ederler. Sosyal sanal d nyalar ise kullanıcılara daha  zg r bir ortam sa layarak, kendi davranı larını kendilerinin se melerine imk n tanır.  rne in Second Life kullanıcılara, avatarlar aracılı ı ile y r y   yapma, sanal g ne in tadını  ıkarma ve ba ka avatarlarla konu ma gibi g nl k aktivitelerinin yanı sıra, sanal elbise ve mobilya tasarımı gibi i erik  retmelerine ve bunları sanal para kar ılı ında di er kullanıcılara satmalarına varana kadar geni  bir  zg rl k tanımaktadır (Kaplan ve Haenlein, 2010: 64).

İKİNCİ BÖLÜM

SOSYAL MEDYA MADENCİLİĞİ

2.1 Sosyal Medya Madenciliği Kavramı

Sosyal medya madenciliği, sosyal medya verilerden anlamlı örüntülerin analiz edilerek ortaya çıkarılma süreci olarak tanımlanmaktadır (Zafarani v.d., 2014:21). Bir diğer tanımlamaya göre ise sosyal medya madenciliği, sosyal medyadan elde edilen verilerin sistematik bir şekilde analiz edilmesi temelleri üzerine kuruludur. Verilerden çeşitli örüntülerin ve bilgilerin keşfi için kullanılan teknikler veri madenciliği olarak adlandırılırken, bu kavram sosyal medya verileri bağlamında değerlendirildiğinde sosyal medya madenciliği adını almaktadır (Ravindran ve Garg, 2015:4).

Yeni gelişen bir alan olan sosyal medya madenciliği, bilgisayar bilimi, veri madenciliği, makine öğrenimi, sosyal ağ analizi, ağ bilimi, sosyoloji, etnografya, istatistik, optimizasyon ve matematik gibi alanları kapsayan disiplinler arası bir alandır (Zafarani v.d., 2014:21). Sosyal medya madenciliği çeşitli alanlarda birbirinden farklı disiplinler arası metotlarla birlikte kullanılmaktadır. Literatürdeki çalışmalar incelendiğinde geniş bir uygulama alanına sahip olduğu görülmektedir. Sosyal medya madenciliği uygulamalarına bakıldığında akıllı sosyal medya izleme, itibar yönetimi, müşteri profili oluşturma, ürün yönetimi ve reklamcılık alanında yapılan çalışmalar örnek olarak verilebilir (Atzmueller, 2012:414).

Sosyal medya verilerinin üretildiği sosyal medya platformları, akademik araştırmalarda büyük miktarda veri toplamasına olanak tanıyan bir çeşit canlı laboratuvarıdır. Sadece araştırmalar için değil, pratik çözümler için de kullanılan sosyal medya verileri, kurumlar tarafından reklamcılık, halkla ilişkiler, müşteri ilişkileri yönetimi (CRM) ve iş zekâsı (BI) hedeflemeye fırsat olarak görülmektedir. Özellikle pazarlama için ek bir kanal olarak etkin bir şekilde kullanılabilmesi, birçok kurumsal firma adına önem arz etmektedir. Son zamanlarda ise siyasi kurumlar, kamuoyunu

izlemek, güncel siyasi konuları tespit etmek ve sosyal medyada kendi itibarlarını yönetmek için de sosyal medyaya eğilim göstermektedirler. Ayrıca, sağlık kurumları, sosyal medya verilerine dayanarak, zamanında müdahaleye yardımcı olacak hastalık salgınlarının teşhisi için erken uyarı sistemi oluşturabilmektedirler (Stieglitz v.d., 2014:90).

Sosyal medyanın veri kaynağı olarak kullanılması eşsiz ve yeni özellikleri beraberinde getirmektedir. Anket gibi yöntemlerle veri toplama sırasında katılımcılara, davranışları veya duyguları ile ilgili geriye dönük olarak sınırlı kapsamda sorular sorulurken veya hayali bir senaryo üzerinden ileriye yönelik davranışları sorgulanırken, sosyal medya kişisel ifadeyi ve bireyler arasındaki etkileşimi gerçek zamanlı ve geniş ölçekte gözlemleme fırsatını sunmaktadır (McCormick v.d., 2017:391).

Sosyal medya verilerinin yapılarının çeşitliliği nedeniyle, bu verilerin işlenmesinde birbirinden farklı metotlara başvurulmaktadır. Sosyal medya verilerinin analizi, metin, görsel ve video gibi içerikler ve/veya bağlantılar gibi meta veriler üzerinden hareket edilerek yapılabilmektedir. Sosyal medyada üretilen verilerin dinamik olması ile birlikte sürekli bir veri akışının gerçekleşmesi ve bu verilerin analizi ile birlikte “akan veri madenciliği” (Seker, 2014:24) kavramı literatürde kullanılmaktadır.

Sosyal medya madenciliği tüm bu tanımlamalar ve açıklamalar bağlamında, sosyal medya verileri içerisinden anlamlı bulguların, örüntülerin ve kuralların çıkarımı olarak özetlenebilir. Sosyal medya madenciliği, hazır çözümlerden ziyade, içerisinde problemler barındıran bir alandır. Sosyal medya verilerinin çeşitliliği nedeniyle, bu alan bazı zorlukları bünyesinde bulundurmaktadır. Çok büyük miktarda veri üretilmesi, verilerin oldukça dinamik ve düzensiz olması sebebiyle geleneksel veri işleme yöntemleri ile kolayca işlenemezler.

2.2 Sosyal Medya Madenciliğinde Karşılaşılan Zorluklar

Sosyal medya verileri hem yapılandırılmış hem de yapılandırılmamış içeriklere sahip olduğundan dolayı oluşan bilgi yükünden anlamlı sonuçların çıkarılması esnasında, veri madenciliği açısından önemli zorluklarla karşılaşmaktadır. Bir diğer zorluk da sosyal medya verisinin doğruluğu ve nesnelliği ile ilgilidir. Sosyal medya verileri sınırlamalara ve önyargılara maruz kalmaktadır. Bu önyargı ve sınırlamalar dikkate alınmadan gerçekleştirilen çalışmalar içerisinde yanlış yorumlama olasılığı bulundurmaktadır. Verilerin toplanması aşamasında da çeşitli zorluklarla karşılaşmaktadır. Bunlar arasında birçok sosyal medya platformunun API'ler (Uygulama Programlama Arayüzü) gibi verilere erişmek için uygun standartlaştırılmış yollar sağlamaması, API'ler sağlansa dahi çoğu zaman veri erişiminin kısıtlamalarına maruz kalınması, farklı birçok formata sahip olan verilerin büyük çoğunluğunda ön işlemeye ihtiyaç duyulması, sosyal medya platformlarında her geçen gün farklı yeniliklerin ortaya çıkması, veri toplama aşamasında ortaya çıkan gizlilik kuralının mevcut olması gibi sorunlar yer almaktadır (Stieglitz v.d., 2014:91).

Sosyal medya verilerindeki yanlışlıklar da sosyal medya madenciliği süreçlerinde sorun oluşturmaktadır. Örneğin, kötü amaçlı oluşturulmuş otomatik hesaplar, içerik kirliliği oluşturarak istatistikleri çarpıtabilmektedir. Bir çalışmada API'lerin yalnızca veri dağıtma esnasında yanlışlıklar yapmadığı bunun yanı sıra, bu verilerin dağıtıldığı mekanizmaların saldırıya uğraması sonucunda da yanlışlıkların meydana gelebileceği keşfedilmiştir. Bot gibi içerik kirleticileri de sosyal medya verileriyle çalışanlara büyük engel olmaktadır. Bot hesaplar sosyal medyada gürültüye neden olmakla birlikte istatistikleri değiştirebilmektedirler (Liu v.d., 2016:2-3).

Sosyal medya madenciliğinde karşılaşılan zorlukları şu şekilde özetlemek mümkündür (Zafarani v.d., 2014:17):

- **Büyük veri paradoksu (Big data paradox):** Sosyal medya verilerinin büyük veri olduğu aşikâr olmasına karşın her bir birey için ayrı ayrı düşünüldüğünde veriler büyük veri özelliğini kaybetmektedir. Dolayısıyla çok boyutlu ve kaynaklı bu veriler yeterli istatistiklerle birlikte etkin madencilik için kullanılmalıdır.
- **Yeterli örneklemin elde edilmesi (Obtaining sufficient samples):** Genellikle sosyal medya sitelerinde veri toplama işlemi, API aracılığıyla yapılmaktadır. Fakat bu yol ile günlük olarak sadece sınırlı miktarda veri elde edilebilmektedir. Bu ise elde edilen mevcut verilerin ne denli güvenilir ve geçerli olduğu problemini ortaya çıkarmaktadır.
- **Gürültü temizleme yanılgısı (Noise removal fallacy):** Sosyal medya verileri kullanıcılar tarafından oluşturulması sebebi ile doğası gereği içerisinde oldukça fazla miktarda gürültülü veri bulundurmaktadır. Bu gürültülü verilerin temizlenmesi aşamasında iki önemli sorun ortaya çıkmaktadır. Bu sorunlardan ilki, verilerin temizlenmesi aşamasında değerli bilgilerin de ortadan kaldırılması sonucunda oluşabilmektedir. Bir diğer sorun ise verilerin gürültülü olarak algılanıp algılanmamasının yapılacak çalışmaya göre değişkenlik göstermesidir. Tüm bu sorunlardan dolayı verilerin ön işlemeden geçirilmesi aşaması oldukça özenli bir çalışma gerektirmektedir.
- **Değerlendirme ikilemi (Evaluation dilemma):** Veri madenciliğinde bulguları değerlendirmenin standart bir prosedürü, birtakım temel gerçekliğe (ground truth) sahip olmaktır. Örneğin, bir veri seti, eğitim ve test olarak ikiye bölünmekte ve öğrenmede yalnızca eğitim verileri kullanılırken, test verileri test için temel gerçeği oluşturur. Bununla birlikte, temel gerçeklik genellikle sosyal medya madenciliğinde mevcut değildir. Güvenilir bir değerlendirme yapılmadan bulguların geçerliliği garanti edilemeyeceğinden dolayı problem oluşturmaktadır.

Bir diğer çalışmaya göre karşılaşılan zorluklar ölçeklenebilirlik, heterojenlik, evrim, kolektif zekâ ve değerlendirme olmak üzere beş başlık ile özetlenmiştir. Zafarani v.d.

(2014:17) tarafından değinilen değerlendirme zorluğu başlığı dışındaki maddeler aşağıdaki gibi özetlenmiştir (Tang ve Liu, 2010:7):

- **Ölçeklenebilirlik (Scability):** Sosyal medya platformlarının devasa boyutta katılımcıların, bağlantıların olduğu bir ağa sahip olmasından dolayı, geleneksel analiz teknikleri uygulandığında başarısız sonuçlar elde edilebilmektedir.
- **Heterojenlik (Heterogeneity):** Gerçekte olduğu gibi sosyal medyada da çoklu ilişkiler mevcuttur. Birçok sosyal medya platformunda kullanıcılar, etiketler ve içerik birbirleriyle iç içe geçmiş durumdadır. Bu da tek bir sosyal medya platformunda bile heterojen varlıklara yol açmaktadır. Heterojen yapıları veya etkileşimleri içeren bu farklı platformların analizi, yeni teori ve araçlar gerektirmektedir.
- **Evrin (Evolution):** Sosyal medya platformlarının dinamik olması sebebi ile zaman kavramı ayrı bir öneme sahiptir. Kullanıcılar bir içeriğe ya da platforma olan ilgilerini çok çabuk bir şekilde kaybedebilmektedirler. Bu platformlardaki dinamikleri belirlemek bu sebepten dolayı zordur.
- **Kolektif Zekâ (Collective Intelligence):** Sosyal medyada kullanıcıların oluşturduğu etiketler, yorumlar ve tavsiyeler biçimindeki kalabalıkların bilgeliğine çoğu zaman erişilebilmektedir. Sosyal bağlantı bilgisini ve kolektif zekâyı etkin bir şekilde kullanmak beraberinde çeşitli zorlukları getirmektedir.

2.3 Literatür Taraması

Sosyal medya madenciliği sürecinin açıklamasına geçmeden önce bu bölümde literatürde yapılmış çalışmalara yer verilmiştir.

af Rosenborg v.d. (2017), hazır giyim sektöründe faaliyet gösteren bir markanın sosyal medya katılımı ve finansal performansı arasındaki ilişkiyi incelemişlerdir. Çalışmada, markanın üç stil ikonu ve dört tasarımcısı karşılaştırılmış, hangilerinin üç aylık satışlarda daha fazla etkiye sahip olduğunu araştırmak adına 2012 ve 2013 yılları

arasındaki stil ikonları ve tasarımcıların yer aldığı yedi kampanyanın Facebook verileri (post, yorum ve beğenilerin sayısı) ele alınmıştır. Veri setini, tasarımcı ve stil ikonlarının kampanyalarının sosyal medya verileri ile markanın finansal verileri oluşturmuştur. Hem trend hem de mevsimsellik içeren Üstel Düzgünleştirme yöntemi kullanarak markanın satışları tahmin edilerek, stil ikon ve tasarımcı kampanyalarının etkisi ölçülmüştür. Stil ikonlarının tasarımcılardan daha fazla etki yarattığı saptanmıştır.

Ding v.d. (2017), Facebook'un "beğen" özelliğinin filmlerin gişe tarihinden önce etkin bir sosyal pazarlama aracı olarak kullanılabileceğini önermişlerdir. Çalışmada IMDb, Box Office Mojo ve Facebook'dan 2013 yılı boyunca veriler toplanmıştır. Beğenilerin gişe hasılatına etkisini incelemek amacı ile gişe öncesi beğenilerin gişe performansına etkisi kesitsel regresyon modeli kullanılarak incelenmiştir. Gişeden bir hafta önce beğenilerin sayısının %1 artması, açılış haftası gişe hasılatının yaklaşık %0,2 oranında artması ile ilişkilendirilmiştir. Gişe tarihine yaklaştığında, gişe öncesi beğenilme etkisinin daha güçlü hale geldiği gözlemlenmiş ve gişeden önceki en son beğenilerin gişe hasılatında daha etkili olduğu sonucuna ulaşılmıştır.

Oh v.d. (2017), Facebook platformu ile birlikte Twitter ve YouTube platformlarını da dâhil ederek tüketici katılım davranışları (CEB) perspektifinden sosyal medyanın filmlerin gişe performansına etkisinin olup olmadığını araştırmıştır. Bu çerçevede 106 filmlik bir veri seti oluşturulmuş, gişe gününden 1 gün öncesine ait her filmin sosyal medya profillerine ait verilerle birlikte BoxOfficeMojo'dan hafta sonu brüt gelir, yayın türü ve film türü gibi film bülten bilgileri toplanmıştır. Açılış haftası box-office brüt hasılatı bağımlı değişken, Facebook beğenileri ve etkileşimleri, YouTube görüntülenmeleri ve yorumları, Twitter takipçi sayısı ve diğer hesaplardan atılan filmle ilgili tweet sayıları bağımsız değişkenler olarak tanımlanarak OLS regresyon modeliyle analiz gerçekleştirilmiştir. Bulgular sosyal medya platformlarının filmlerin gelecek performansını olumlu yönde etkilediğini ortaya koymuştur.

İşeri v.d. (2017), Türkiye’de faaliyet gösteren üç telekomünikasyon şirketi ile ilgili konuşulmaların duygu durumlarını Twitter platformu örneklemini dâhilinde ele alarak incelemiştir. Veriler Nisan-Mayıs 2017 aralığında Twitter API desteği ile elde edilmiş, metin ön işleme aşamasından geçirilerek ünlem, vurgu gibi ifadeler veri setinden temizlenmiştir. Duygu durumlarının belirlenmesinde makine öğrenmesi sınıflandırma metoduna başvurmadan önce eğitim setini oluşturmak için her üç şirket için de elde edilen tweetler manuel olarak olumlu ve olumsuz olmak üzere iki kategoride etiketlenmiştir. İki farklı özellik çıkarımı gerçekleştirildikten sonra sınıflandırma işlemi için k-NN algoritması kullanılmıştır. Sınıflandırma yapılırken kelimeler arası uzaklıkların hesaplanmasında öklit uzaklıktan faydalanılmıştır.

Kuflik v.d. (2017), ulaşım ile ilgili 3.7 milyonluk tweet içeren veri setinden 4.000 tweet içeriğini manuel olarak kategori ve duygu durumlarını etiketleyerek, sınıflandırıcıları eğitmek ve test etmek için kullanmışlardır. Veri setinin geri kalanı daha sonra ikinci bir test adımı olarak analiz edilmiştir. Sunulan sonuçlar ulaştırma ile ilgili tweetlerin tanımlanmasında Destek Vektör Makine algoritması ikili N-gram tekniğiyle birlikte güçlü bir derecede başarı göstermiş ve ulaşım hizmetleri hakkında görüş bildiren tweetlerin belirlenmesinde de benzer başarı Naive Bayes algoritması ikili N-gram tekniğiyle kullanılarak elde edilmiştir. Genel olarak, ulaşımı ilgilendiren müşteri yorumlarının olumsuz görüşler bildirdiği gözlenmiştir.

Ayata v.d. (2017), Türkçe tweet veri setini Destek Karar Makinesi ve Rastgele Orman aracılığı ile olumlu ve olumsuz olmak üzere sınıflandırmışlardır. Bankacılık, Futbol, Telekom ve Perakende olmak üzere dört farklı sektörden oluşan veri seti öncelikle vektörel değerlere çevrilerek makinenin işleyebileceği duruma getirilmiş ve eğitim veri seti ile makine öğrenmesi gerçekleştirilmiştir. Çalışma, sektörel ve genel olmak üzere iki farklı şekilde denenmiştir. Bulgulara göre sektöre özel ayrı ayrı yapılan öğrenmelerin tüm veri seti için uygulanan öğrenmeden daha başarılı sonuçlar verdiği gözlemlenmiştir. Genel olarak Karar Destek Makinesi sınıflandırıcısının Rastgele Orman’dan daha iyi sonuçlar verdiği bulgular arasında yer almaktadır.

Yilmaz v.d. (2017), bilgi teknolojisi (IT) alanında faaliyet gösteren ilk beş firma ve kurumsal kaynak planlama (ERP) alanında faaliyet gösteren ilk beş firma örneklem dâhilinde, iki farklı alandaki firmaların Endüstri 4.0'a olan bakış açılarını karşılaştırmak amacı ile firmaların kurumsal hesaplarındaki Twitter paylaşımlarını ele almıştır. Twitter API desteği ile her firmanın kurumsal hesabına ait Twitter içerikleri elde edilmiş ve bir dizi metin ön işleme aşamasından geçirilerek analize hazırlanmıştır. Kelime frekansları elde edilmiş ve Endüstri 4.0 ile ilgili kelimeler incelenmiştir. Duygu analizi R programında çalıştırılan bir paketle gerçekleştirilmiştir. Bulgular firmaların bu kavramla ilgili çoğunlukla nötr paylaşımlar yaptığı yönündedir.

Pournarakis v.d. (2017), Ocak- Nisan 2015 tarihleri arasında bir araç paylaşım servisi ile ilgili 280.000 üzeri tweet içeriğini Twitter API aracılığı ile elde etmiş, ilk aşamada ilgili anahtar kelimelerle arama yapılarak veri setindeki gürültüleri temizledikten sonra diğer metin ön işleme aşamalarını gerçekleştirmişlerdir. Daha sonra vektör uzay modeli oluşturularak kelimeler bilgisayarın işleyebileceği bir hale getirilmiştir. İçeriklerin duygu durum sınıflandırması Gauss Kernel fonksiyonu aracılığı ile gerçekleştirilmiştir. En çok konuşulan konular Latent Dirichlet Allocation (LDA) konu modellemesi ile elde edilmiş ve Genetik Algoritma (GA) kümeleme yöntemi uygulanarak içerikler kümelere ayrılmıştır. Her kümedeki duygu durumları ayrı ayrı belirlenerek son aşamada tüketicilerin algılarına ulaşılmıştır.

Xiang v.d. (2017), ABD'de bir bölgede bulunan otellerle ilgili olarak otelin ismini, adresini, sınıfını, derece bilgilerini ve müşterilerin yaptığı tüm değerlendirmeleri içeren yaklaşık 439 bin TripAdvisor, 351 bin Expedia ve 31 bin Yelp verisini toplamışlardır. Yorumlardaki konular LDA ile çıkarılmış, duygu durumları ise belirlenen pozitif ve negatif sözcüklerden oluşturulan bir sözlük ile sınıflandırılmıştır. Platformları karşılaştırmak amacıyla, otel derecelendirmeleri ve yorumları arasındaki ilişki tahminine, konu ve duygu durumları da dâhil edilerek çoklu doğrusal regresyon uygulanmıştır. Bulgularda platform temelli birçok farklılık gözlemlenmiştir.

Ye ve Ki (2017), 2010 yılında gerçekleşen petrol sızıntısı krizini, Facebook üzerindeki kriz yanıtları ve kamunun krize tepki stratejileri konusundaki algılamalarına katkıda bulunan faktörleri içerik analizi uygulayarak araştırmıştır. Kurumun kriz iletişimi mesajları ve Facebook kullanıcılarının yorumları kurum Facebook sayfasından alınmış ve yalnızca kurumun mesajlarına doğrudan yanıt veren 16.551 içerik, dört araştırmacı tarafından içeriklerin genel bilgisine (tarih, beğeni sayısı ve ileti türü), kriz iletişim stratejilerine (bilgi verici stratejiler ve itibar yönetimi stratejileri) ve kamu yorumları ile kurum iletilerinin arasındaki uyuşmaya (uyumsuz:-1, nötr:0, uyumlu:1) göre kodlanmıştır. Bulgular, bilgi verici stratejilerin kurumun kriz yanıtlarına egemen olduğunu ve kurumun bilgi veren ve uyum stratejilerini kullandığında Facebook kullanıcıları tarafından olumlu yorumların yapıldığını ortaya koymuştur. Stratejilerin güçlendirilmesi ve üçüncü bir parti/kurum tarafından onaylanmanın beklenen etkinliği sağlayamamıştır.

Bian v.d. (2016), bireylerin “Nesnelerin İnterneti (IoT)” kavramıyla ilgili düşüncelerini araştırmak adına Twitter’den elde edilen verileri analiz etmiştir. IoT ile ilgili olan 1.57 milyar tweet Twitter API aracılığı ile toplanarak Twitter trendleri (Twitter trends) oluşturulmuştur. Kavramla ilgili bireylerin tutumunu anlayabilmek adına sözdizimi tabanlı bir yaklaşım ile duygu analizi ve gizli kavramları açığa çıkarmak için LDA’ya dayalı konu modelleme kullanılmıştır. Analiz sonuçlarına göre bireylerin IoT ile ilgili algısının çoğunlukla olumlu olduğu ve alakalı konu başlıklarının büyük veri (big data), akıllı teknoloji (smart technology), bağlantılı cihazlar (connected device), Google ve IBM gibi sektör liderleri gibi genellikle işletme ve teknolojiye odaklandığı gözlemlenmiştir. Bununla birlikte bireylerin gizlilik ve güvenlik konularında büyük kaygı duydukları sonucuna ulaşılmıştır.

Boldt v.d. (2016), bir markanın kurumsal Facebook sayfalarındaki verilerinden marka satış tahmininin mümkün olup olmadığını araştırmıştır. Araştırmanın verileri, markanın belirlenen on kurumsal Facebook sayfasındaki sosyal medya verileri (beğeni, yorum, kullanıcı ve ileti sayısı), Bloomberg Professional Service’den elde edilen 3 aylık satış verileri ve Google Trends verileri olmak üzere üç farklı veri setinden oluşturulmuştur.

Çalışmanın sonuçlarında Facebook verilerinin önemli bilgilere sahip olduğunu ortaya çıkarmıştır. Basit regresyon modellerinden bazıları yüksek öngörü doğruluğuna sahipken, çoklu regresyonların öngörme doğruluğu daha düşük çıkmıştır. Google Trends verilerinin ise istatistiksel olarak anlamlı olduğu birkaç model bulunmasına karşı tahminlerin doğruluğu, basit regresyon modellerinde olduğu kadar yüksek çıkmamıştır.

Kim v.d. (2016), Çin’de faaliyet gösteren bir kurumun sosyal medya platformu Weibo’da gelişen krize odaklanarak, krizin nasıl başarıyla yönetildiğini nitel ve nicel içerik analizi yoluyla incelemiştir. Sosyal medya krizi iletişim stratejileri ve iletişim kaynağının rolü ile ilgili çıkarımlarda bulunulmuştur. Kurum tarafından uygulanan kriz iletişim stratejileri belirlenmiş ve daha sonra kriz iletişim stratejilerinde kurumun CEO’su ve bir yöneticisi olmak üzere kullanılan iki farklı iletişim kaynağının kamunun duygu durumunda bir değişikliğe yol açıp açmadığı incelenmiştir. 1.050 yorum tabakalı tesadüfi örnekleme yolu ile seçilmiş ve olumlu, nötr ve olumsuz olmak üzere duygu durumları manuel olarak kodlanmıştır. Ki-kare test sonuçlarına bakılarak kriz iletişim stratejilerinde kurum CEO’sunun kurumun uyguladığı diğer stratejiye kıyasla daha olumlu duygulara yol açtığı sonucuna varılmıştır. Çalışma, kurumun erken kriz evresinde, mizahi yanıt stratejisini benimseyerek sosyal medyada itibar krizini yönetmede başarılı olduğunu ortaya koymuştur.

Odlum ve Yoon (2015), Twitter’ı, Ebola salgının izlenmesinde gerçek zamanlı bir bilgi yayılım gözlemlene, erken tespit ve kamusal bilgileri ve tutumları saptama yöntemi olarak ele almıştır. İçerisinde Ebola kelimesi içeren birkaç anahtar kelime kullanılarak tweetler toplanmıştır. Kamuoyu bilgilerini ve tutumları incelemek amacı ile doğal dil işleme, bilgi yayılımını gözlemlenmek için ise jeolojik görselleştirme ile zaman serisi analizi kullanılmıştır. Bulgulara göre Ebola ile ilgili toplam 42.236 tweetin paylaşıldığı ve bu tweetlerin 9.362.267.048 kişiye yayıldığı saptanmıştır. Zaman serisi analizi ile bilginin yayılma oranının kademeli olarak arttığı saptanmıştır. İçerik analizi sonucunda tartışılan dört ana konu ortaya çıkmıştır. Bunlar, risk faktörleri (bu virüsün neden oluştuğu), önlem eğitimleri (sağlık bilgisi, virüsün

semptomları, önleme yöntemleri) hastalık eğilimleri (yayılma yerleri) ve Afrika ülkeleri için dua olarak gözlemlenmiştir.

Çoban v.d. (2015), Türkçe Twitter mesajlardan oluşan bir veri setinin duygu durum analizini gerçekleştirmek üzere metin sınıflandırma yöntemlerini kullanmışlardır. Eğitim verisi için duygu durumlarının etiketlenmesinde his simgeleri olarak ifade ettikleri emoji'ler göz önünde bulundurularak, 20 bin mesaj pozitif ve negatif olarak etiketlenmiştir. Test verisi Twitter mesajları metin ön işleme aşamasından geçirilerek temizlenmiştir ve 15 bine yakın mesajın duygu durumları Destek Vektör Makineleri (SVM), Naive Bayes, Multinom Naive Bayes ve k-NN algoritmaları kullanılarak belirlenmiştir. Kelime torbası (BoW) ve N-Gram model olmak üzere iki farklı şekilde öznitelik çıkarımı yapılarak sınıflandırma sonuçlarına olan etki incelenmiştir.

Guidry v.d. (2015), Forbes 2013 verilerine göre en büyük on fast food zinciri hakkında sosyal medya platformu Instagram'da üretilen ve yayılan kriz bilgileri ile ilgili, 49'u şirketlere, 662'si kullanıcılara ait toplamda 711 Instagram içeriğini, Ekim-Aralık 2013 tarihleri arasında toplamıştır. Şirketler hakkındaki kriz bilgilerinin kullanıcılar tarafından Instagram üzerinde nasıl iletildiği, yorum ve beğeniler çerçevesinde üretilen içeriklerin etkileri, şirketler arasında kriz bilgilerine verilen yanıtlar ve Instagram katılımı açısından farklılıkların olup olmadığı araştırılmıştır. Instagram kullanıcıları tarafından üretilen içeriklerin konusu, duygu durumu (pozitif, negatif, nötr), kullanıcı bilgisi, multimedya türü (resim ya da video), şirketlerin yayınlamış oldukları içeriklerin konusu, multimedya türü ve yanıt verip vermemesi, beğeni ve yorum sıklığı şeklinde kodlamalar gerçekleştirildikten sonra SPSS programı ile veriler analiz edilmiştir. Analiz sonuçlarında şirketlerin Instagram'ı yeni yeni kullanmaya başladıkları ve herhangi bir yoruma cevap vermedikleri görülmüştür. Bulgular negatif yorumların müşteriler dışında şirket çalışanlarından da gönderildiğini ve müşterilerin servis sorunları, çalışanların ise çalışma ortamı koşullarıyla ilgili yorumlar gönderdiklerini göstermiştir. Kimliği belirsiz kişiler tarafından iletilen içeriklerin daha fazla yorumu tetiklediği gözlemlenmiştir.

Chae (2015) Şubat-Nisan 2013 tarihleri arasında tedarik zinciri ile ilgili çeşitli anahtar kelimeler aratılarak toplanan 22 bin tweet içeren veri setini, tanımlayıcı istatistikler, içerik analizi ve ağ analizinden oluşan bir analitik çerçeve önererek analiz etmiştir. Tanımlayıcı istatistikler bölümünde tweet ve hashtag sayıları, kullanıcı ve URL istatistikleri verilmiştir. İçerik analizi bölümünde en popüler kelimeler ve hashtagler frekans sayılarına göre belirlenmiş ve kümelenmiş, tweet duygu durumları SentiStrength isimli program kullanarak gerçekleştirilmiştir. Son olarak ağ analizi bölümünde kullanıcılar ele alınarak ağdaki düğüm sayısı, kenar sayısı ve derece merkeziliği hesaplanmıştır. Bulgular tedarik zinciri yönetiminde Twitter'ın potansiyeli hakkında bilgiler vermiştir.

Lassen v.d. (2014), Twitter'dan bir marka ile ilgili anahtar kelime aratarak, 2007-2013 yılları arasında yaklaşık 400 milyon tweet içeriğini Twitter API yardımı ile çekerek markanın satışlarını tahmin etmek üzere ele almıştır. Tweet içeriklerinde yorum ya da retweet gibi bir ayırım yapmadan atılan tweetler tek bir sınıf olarak ele alınmıştır. Bir diğer değişken olarak tweet içeriklerinin duygu durumları, Topsy Pro duygu analizi aracı ile otomatik olarak belirlenmiştir. 2010-2013 döneminde marka satışları ve hakkında atılan tweetler arasındaki ilişki modellenmiş ve 2007-2009 dönemi verileri oldukça gürültülü bulunduğundan dolayı araştırmacılar tarafından dâhil edilmemiştir. Çoklu doğrusal regresyon tahmin modeli geliştirilmiş, Morgan Stanley ve IDC'nin tahmin modellerine yakın bir ortalama hata ile üç aylık olmak üzere iPhone satışları başarılı bir şekilde tahmin edilmiştir.

Sert v.d. (2014), Twitter'da akademik zam konusunda kişilerin görüşlerinin zaman üzerindeki değişiminin ve gelişiminin sosyal ağ analizi çerçevesinde analiz etmişlerdir. NodeXL programının 1.000 tweet sınırlaması altında içerikler TBMM'nin gündemine göre belirlenen beş tarih için toplanmış ve sosyal ağ analizinde kullanılan görselleştirme ve sayısal ölçüler (girdi derecesi, girdi yakınlık merkeziliği, yoğunluk, ortalama derece, çap, kümeleme katsayısı) kullanılarak analiz edilmiştir. Bulgular ağ değerlerinin, yaşanan gelişmeler ışığında azalıp arttığını göstermiştir. Twitter

aracılığıyla gündemdeki gelişmeler ile ilgili halkın tepkisini ölçmenin mümkün olabildiği gözlemlenmiştir.

Ki ve Nekmat (2014), Fortune 500 şirketlerinin ne ölçüde Facebook'u kriz iletişim aracı olarak kullandığını, şirketlerin ve bireylerin göstermiş olduğu etkileşim seviyesini, kurum tarafından uygulanan kriz müdahale stratejilerini, bunlara verilen yanıtlardaki kamu duygu durumlarını, kriz türleri ile müdahale stratejilerinin ne ölçüde ilişkili olduklarını incelemiştir. 2011 yılı Fortune 500 listesinde yer alan tüm şirketlerin Ocak-Eylül 2011 tarihleri arasında Facebook iletilerinde yer alan metinler içerik analizi yöntemi ile analiz edilmiştir. 7.080 adet içeriğin duygu durumu negatif, nötr ve pozitif olmak üzere araştırmacılar tarafından manuel olarak kodlanmıştır. Bulgulara göre şirketlerden yalnızca %32'sinin kriz ortamında kamu tarafından iletilen mesajlara cevap vererek çift yönlü iletişim yarattığı görülmüştür. Katılımcıların iletilerinin büyük çoğunluğunun duygu durumlarının negatif ve nötr olduğuna ve ki-kare analizi sonuçlarına göre kamu mesajlarına cevap veren şirketlere verilen tepkilerin, cevap vermeyen şirketlere göre daha olumlu olduğuna ulaşılmıştır.

van der Meer ve Verhoeven (2013), kamunun bir kurumda yaşanan kriz hakkındaki algısının medya içeriklerinden önce neye dayandığını, medya içerikleri ile uygunluk gösterip göstermediğini ve medya bilgilendirmesinin ardından örgütsel krizin çerçevesini kişiselleştirip kişiselleştirmedeğini araştırmıştır. Çalışmada bir tesiste gerçekleşen patlama krizi ele alınmış, 38.074 Twitter mesajı ve 73 haber içeriği otomatik içerik analizi uygulanmış ve kamu ve medya çerçeveleri üç evre olarak (patlama evresi, bir sonraki evre, son evre) analiz edilmiştir. Tweet içeriklerinde ve medya metinlerinde en sık kullanılan kelime listesi derlenmiş, sözcük/oluşum matrisleri oluşturulmuş ve faktör analizi, bileşenler içindeki kelimeler arasındaki istatistiksel korelasyonun belirlenmesi için gerçekleştirilmiştir. Bulgular, ilk periyot diye tabir edilen krizin patladığı günde kamunun, panik sözcüklerini kullanarak krizi bir terör eylemi olarak algıladığını göstermiştir. İkinci periyotta ise kamunun genel kriz bilgileri ile ilgili bir çerçeve çizdiği gözlemlenmiş ve medyanın krizle ilgili paylaşımlarından sonra kamu çerçevesi ile medya çerçevesinin uyduğu

gözlemlenmiştir. Son olarak üçüncü periyotta haber ajansları “gözetleme”, “denetim”, “güvenlik” gibi kelimeler kullanırken, kamu ise krizi hükümetle ilişkilendirerek kriz çerçevesini kişiselleştirmiştir.

He v.d. (2013), üç büyük pizza zincirini ele alarak Facebook ve Twitter platformlarındaki takipçi sayısı, içerik sayısı, yorum sayısı, beğeniler ve içerik yayınlama sıklığı gibi nicel verileri toplamıştır. Daha sonra yeni bilgi ve kalıpları keşfetmek ve bu pizza zincirlerinin sosyal medyayı nasıl kullandığını daha iyi anlamak adına Facebook ve Twitter platformlarındaki içeriklere metin madenciliği metotları uygulanmıştır. Metin madenciliği sonuçlarına göre bu üç zincirin, Twitter ve Facebook platformlarını hizmetleri ve ürünleri hakkında bilgi vermekle birlikte aktif bir şekilde müşterileriyle ilgilenmek ve diyalog kurmak için de kullandıkları belirlenmiştir. Bu çalışmadan elde edilen bulgular sonucunda sosyal medyanın müşterilerle olumlu ilişkiler sürdürmede önemli bir role sahip olduğu anlaşılmıştır.

Poell ve Borra (2012), sosyal medyanın alternatif gazetecilik alanında kullanımını incelemiştir. Bu doğrultuda Kanada’da gerçekleştirilen G20 zirvesi ile ilgili sosyal medya platformları Twitter, YouTube ve Flickr’den ilgili anahtar kelime aratılarak, 11.556 tweet, 222 video ve 3.338 fotoğraf toplanıp analiz edilmiş, kullanıcıların sosyal medya raporlaması, protesto katılımı ve ortaya çıkan protesto hesapları değerlendirmiştir. Platformlar ayrı ayrı değerlendirilerek almış olduğu etkileşimler hesaplanmış ve Pareto ilkesi doğrultusunda her bir platformdaki içeriklerin yüzde kaçının en aktif kullanıcılardan ve en ilgili tweetlerden geldiği araştırılmıştır. Katılımın yalnızca küçük bir grup tarafından gerçekleştiği, büyük grupları harekete geçiremediği sonucuna ulaşılmıştır. Özellikle Flickr ve YouTube platformlarında paylaşılan içeriklerin, az sayıda kullanıcı tarafından üretildiği gözlemlenmiştir. Alternatif gazeteciliğin kitle kaynaklı olabilmesi adına en umut verici platform ise Twitter olarak değerlendirilmiştir.

Rossi ve Magnani (2012), Twitter platformunda konu ağları (topical networks) ile takipçilerin yapısal ağı arasındaki ilişkileri değerlendirmek amacı ile popüler bir TV

şovunu ele alarak, 5-6 Ocak 2012 tarihleri arasında TV şovu süresince atılan 22.287 adet tweet üzerinden sosyal ağ analizi uygulamıştır. Oluşturulan ilk çizgede, kullanıcılar arasında en çok bahsedilen (mention) kullanıcıların TV şovu ile doğrudan ilgili hesaplar olduğu gözlemlenmiştir. Retweet zincirlerinden oluşan ikinci çizgede ise resmi hesapların yanı sıra, sıradan kullanıcıların oluşturdukları içeriklerin de ağ içerisinde baskın role sahip oldukları görülmüştür. Yeni takipçi sayısı ile tweet sayısı, cevap ağı derece merkeziliği ve retweet ağı derece merkeziliği arasındaki ilişkiye bakılmış ve tweet sayısının, yeni takipçi kazanımında çok sınırlı, retweet ve cevap ağı içerisinde merkezi bir konuma sahip olmanın ise yeni takipçi kazanımında önemli bir etkiye sahip olduğuna ulaşılmıştır.

Hui v.d. (2012), Twitter’da kriz anında yayılan içeriklerin akışlarını hakkında bilgi edinmek amacı ile bir kurumda gerçekleşen krizi ele alarak 318 kullanıcı tarafından oluşturulan 641 adet içeriği Twitter API yardımıyla elde etmişlerdir. Genel yayılım modeli (general diffusion model) çerçevesinde Twitter içeriklerinin, kurum tarafından yapılan dört bilgilendirme evresi kapsamında çizgeleri oluşturulmuştur. Bulgulara göre retweet aktivitesinin özellikle ikinci uyarı mesajı olmak üzere uyarıların iletildiği evrelerde fazlayken, tehlikenin geçtiğinin bildirildiği evrede azaldığı ve içeriklerin küçük zincirler arası etkileşime uğradığı görülmüştür. En fazla bilgi akışının kurumdan ve yerel medya kullanıcılarından olduğu bulgular arasındadır. Bir diğer bulgu ise retweet edilen içeriklerin çoğunun bir kere paylaşılmasının ardından sonlandığı ve uzun retweet zincirlerinin oluşmadığıdır.

Chatfield ve Brajawidagda (2012), Twitter platformunun bir afet uyarı sistemi olarak çalışıp çalışmadığını incelemek adına 2011 yılında Endonezya’da gerçekleşen doğal afeti ele almışlardır. Jeolojik afet uyarılarından sorumlu yerel ajansın Twitter takipçilerinin 18 Haziran - 3 Temmuz 2012 tarihleri arasında oluşturduğu 6.383 tweetten oluşan veri seti analiz edilmiştir. Sosyal ağ analizi ölçüm değerleri (düğüm sayısı, bağlantı sayısı, grafik yoğunluğu, ortalama derece, ağ çapı), yerel ajans takipçilerinin kendi aralarında ilişkilerinin olmadığını ve bazılarının yalnızca izleyici konumunda olduğunu ortaya koymuştur. Twitter’ın uyarı sistemi olarak başarılı bir

şekilde kullanabileceğinin yanı sıra, ajans takipçilerinin incelenmesi ile birlikte, bu kullanıcıların çok kısa bir sürede mesajların yayımında etkin rol oynadığı, ağda kanaat önderliği rolü üstlendiğine ulaşılmıştır.

Kumar v.d. (2011), kullanıcılarının göç hareketlerini incelemek ve göç kalıplarının neler olduğunu öğrenmek amacıyla 96.000'den fazla kullanıcı profilinden oluşan bir veri seti oluşturmuştur. Bu veri setinden, birden fazla sitede faaliyet gösteren kullanıcılar, Delicious, Digg, Flickr, Reddit, StumbleUpon, Twitter ve YouTube olmak üzere en popüler yedi sosyal medya platformuna ayrılmış ve sonucunda 17.798 kullanıcı profili veri seti olarak belirlenmiştir. Site göçü ve ilgi göçü olmak üzere iki kullanıcı göç davranışı ve belirli bir zaman diliminde aktif olup olunmadığı kıstası çerçevesinde sosyal medya platformlarındaki kullanıcı aktiviteleri belirlenmiştir. İlgi göçünde genel eğiliminin Twitter ve StumbleUpon'a yöneldiği gözlemlenmiştir. Reddit kullanıcılarının belirgin bir bölümünün Digg platformuna geçtiği görülmüştür. Kullanıcıların aynı kategorideki rakip siteler arasında dolaşmaya eğilimli olmaları sonucunda ilgi göçü gözlemlenmiştir.

Bollen v.d. (2011), Şubat-Aralık 2008 tarihleri arasında, yaklaşık 27 milyon kullanıcı tarafından üretilmiş kullanıcıların duygu durumları ile ilgili olan 10 milyona yakın tweet içeriğini kullanarak, bu içeriklerle Dow Jones Borsası Endüstri Endeksi'nde (DJIA) meydana gelen değişimlerin ilişkili olup olmadığını ve içeriklerin tahmin amaçlı kullanılıp kullanılamayacağını incelemiştir. Duygu durumlarını pozitif ve negatif olarak sınıflandıran OpinionFinder adlı araç ile daha ayrıntılı olmak üzere 6 duygu durumunu ölçen GPOMS adlı araç kullanılarak Twitter içeriklerinin duygu analizi yapılmıştır. Buna ek olarak, Yahoo! Finance aracılığı ile günlük DJIA kapanış değerleri ele alınmıştır. DJIA değerlerini GPOMS ve OpinionFinder değerleriyle ilişkilendirilen bir Granger nedensellik analizi (Granger Causality Analysis) kullanılmıştır. DJIA'yı tahminde GPOMS tarafından ölçülen "sakinlik" kavramının OpinionFinder tarafından ölçülen genel pozitif sınıflandırmasından daha başarılı sonuçlar verdiği görülmüştür. Sonuç olarak %87,6'lık bir başarı ile Dow Jones Borsası Endüstri Endeksi'nde gerçekleşen günlük iniş ve çıkışlar tahmin edilebilmiştir.

Thelwall v.d. (2011) Twitter'daki en çok konuşulan etkinliklerin aynı konuyla ilgili diğer tweetlere kıyasla pozitif ve negatif duyguların ifade gücünden dolayı oluşup oluşmadığı araştırma sorusu çerçevesinde 2010 yılına ait bir aylık süreçte 30 etkinlik ile ilgili 2,749,840 kullanıcıdan elde edilen 35 milyona yakın tweet içeriğini analiz etmiştir. Tüm iletiler SentiStrength adlı algoritma tarafından pozitif ve negatif olarak sınıflandırılmış ve daha sonra her bir konu için saatlik ortalama pozitif ve negatif duygu durum puanı hesaplanmıştır. Yüksek hacim saatleri, düşük hacim saatleri gibi tweetlerin yoğunluklarına ve saatlerine dayalı tanımlamalar yapılmış ve bu doğrultuda araştırma hipotezleri tanımlanmıştır. Wilcoxon T testi uygulanmış ve yapılan analiz sonucunda popüler olayların olumsuz duygu gücündeki artışlarla alakalı olduğuna ulaşılmıştır. Bir diğer dikkat çeken bulgu ise Oscar gibi olumlu olarak nitelendirilen birçok olay kendilerine tepki olarak artan olumsuz yorumlara sebep olmuştur.

Waters ve Jamal (2011), kâr amacı gütmeyen kuruluşların Twitter platformunu ne amaçla kullandıklarını incelemek amacıyla Chronicle of Philanthropy tarafından yayınlanan listede yer alan kuruluşlar ele alınmıştır. Bu kuruluşlar arasında Twitter'ı aktif olarak kullananlar belirlendikten sonra örneklem olarak kurumların üçte biri (n=27) rastgele seçilmiştir. Kurumların bir aylık paylaşımları dört kategori dâhilinde iki araştırmacı tarafından kodlanmıştır. Cronbach alfa değeri hesaplanarak güvenilirlik analizi gerçekleştirilmiştir. Her kategori için frekanslar hesaplanmış ve değerlendirmelerde bulunulmuştur. Bulgular, kurumların Twitter'da diyalog ve topluluk oluşturma potansiyeline rağmen tek yönlü modelleri kullandıklarını göstermiştir.

Asur ve Huberman (2010), Twitter verilerini kullanarak filmlerin gişe gelirlerini filmler gösterime girmeden tahmin etmeye çalışmışlardır. Twitter'dan veri çekme aşamasında anahtar kelime olarak film isimleri kullanılmış, 3 aylık bir zaman diliminde 24 farklı filmle ilgili 2.89 milyon tweet toplanmıştır. Her film için kritik zamanlama olarak gösterime girmeden bir hafta öncesi ve gösterime girdikten iki hafta sonrası aralığı olarak belirlenmiştir. Sosyal medya verilerinin gerçek zaman tahmininde önemli bir gösterge olabileceği vurgulanarak, film hakkında atılan

tweetlerin üretildiği oran kullanılarak gişe gelirlerinin tahmini için r^2 'si yüksek bir regresyon modeli geliştirilmiştir. Sonuçlar Hollywood Stock Exchange verileriyle karşılaştırılmıştır ve daha başarılı sonuçlar elde edildiği görülmüştür. Tahmin dışında negatif ve pozitif duyguların nasıl oluştuğu ve insanları etkilediği üzerinde de çalışılmıştır. Bu doğrultuda negatif tweetleri pozitiflerden ayırmak adına metin sınıflandırıcılar kullanılarak duygu analizi yapılmıştır.

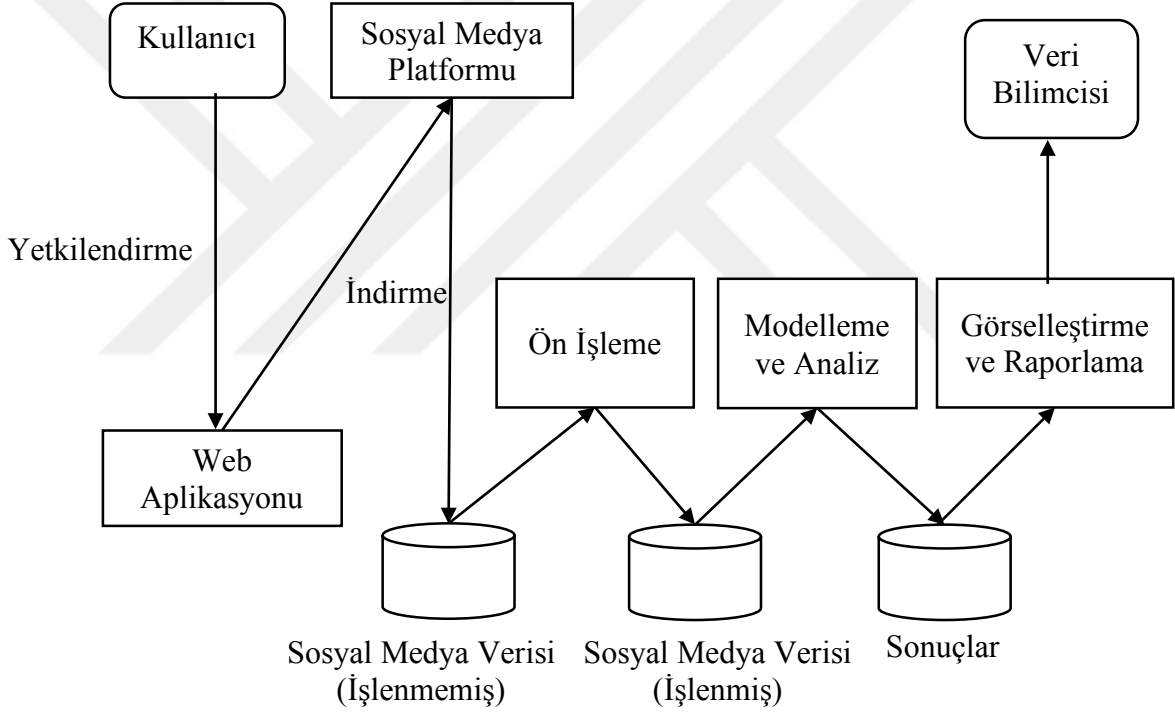
Zimbra v.d. (2009), şirketler hakkında kamuoyu görüşlerini değerlendirmek için otomatik konu çıkarımı ve duygu analizi yöntemlerini uygulayan bütünleşmiş bir çerçevenin, şirketlerin hisse senedi değerleri ile güçlü bir ilişki içinde olduğunu keşfetmişlerdir. Bu çerçeveye dayanarak, bir mağazalar zinciri ilgili bir web forumunda vaka çalışması yürütülmüş ve iletişim trafiği, içeriklerdeki duygular ve fikir liderlerinin faaliyetleri analiz edilmiştir. Veri seti Kasım 2003- Kasım 2008 tarihleri arasında 1.855 yazar tarafından 19.624 içerik ile 1.354 konu içermektedir. Web forumu verileri topladıktan sonra, maksimum entropi modeli kullanılarak bir kümeleme algoritması uygulanmıştır. Bu ön çalışmada, özellik seçimi olarak standart kelime torbası (bag-of-words) yaklaşımı kullanılmıştır. Model parametreleri, maksimum 100 iterasyona sahip 10 küme çıkarmak için belirlenmiştir. Regresyon analizi ile ilk modelde bağımsız değişken olarak yalnızca duygu durumları ele alınmış, ikinci modelde ise duygu durumları ile birlikte konu bilgileri hisse senedi değerleri tahmin edilmiştir. Konu bilgilerin eklenmesinin bir forumun duyarlılık analizini zenginleştirebileceği ve hisse senedi değeri ile temsil edilen halk algısı ile daha yakından uyumlu olduğu ortaya çıkarılmıştır.

2.4 Sosyal Medya Madenciliği Süreci

Sosyal medya madenciliği sosyal medya verilerinin elde edilmesinden, verilerden anlamlı sonuçların çıkarımına ve değerlendirilmesine kadar olan geniş bir süreci kapsamaktadır. Sosyal medya madenciliği bu bağlamda değerlendirilerek, geniş bir kapsam dâhilinde süreçler açıklanmıştır. Sosyal medya madenciliği sürecinde izlenen

adımlar Bonzanini (2016:16) tarafından yetkilendirme (authentication), veri toplama (data collection), veri temizleme ve ön işleme (data cleaning and pre-processing), modelleme ve analiz (modeling and analysis) ve sonuç sunumu (result presentation) olmak üzere beş adımda özetlenmiştir. Belirlenen amaç ve bu doğrultuda oluşturulan veri setine göre Şekil 2.1’de yer alan sosyal medya madenciliği sürecindeki adımlar değişkenlik gösterebilir.

Şekil 2.1: Sosyal Medya Madenciliği Süreci



Kaynak: Bonzanini, 2016:17

2.4.1 Yetkilendirme

Sosyal medya madenciliği sürecinin ilk adımı olan yetkilendirme, genellikle OAuth (Open Authorization) adı verilen protokol tarafından gerçekleştirilir (Bonzanini, 2016:17). Açık bir protokol olan OAuth, Web, mobil ve masaüstü uygulamalarına

basit ve standart bir yöntemle güvenli yetkilendirme imkânı sunmaktadır (Ravindran ve Garg, 2015:8). Bu protokol, kullanıcının şifre gibi hassas bilgilerini paylaşmadan, izin verilen uygulamanın kullanıcının hesap verisine erişmesine olanak tanıyan bir araçtır (Russell, 2013:13).

Yetkilendirme sürecinin adımları ve bu adımlarda izlenen yönler şu şekilde açıklanabilir (Bonzanini, 2016:17):

- **Kullanıcıdan Web uygulamasına akış:** Kullanıcı, Web uygulamasına sosyal medya platformuna erişim izni vermeyi kabul eder.
- **Web uygulamasından sosyal medya platformuna akış:** Kullanıcı, sosyal medya şifrelerini doğrudan web uygulaması ile paylaşmaz. Bu sebepten dolayı Web uygulaması, “token” ve “secret” adı verilen değerleri oluşturmak için kaynak sağlayıcısı olan sosyal medya platformu ile ilk alışverişini yapar. Bu değerler, her türlü sahtekârlık yapmaktan kaçınmada ve sorgularda kullanılır.
- **Sosyal medya platformundan kullanıcıya akış:** Kullanıcı “token” ile sosyal medya platformuna yönlendirilir. Sosyal medya platformu, Web uygulaması tarafından kullanıcının verisine erişilebilmesi için kullanıcıdan onay ister.
- **Kullanıcıdan sosyal medya platformuna akış:** Sosyal medya platformu doğasına bağlı olarak kullanıcı Web uygulamasının kullanıcı adına herhangi bir işlem yapıp yapmayacağına dair onay verir.
- **Sosyal medya platformundan Web uygulamasına akış:** Sosyal medya platformu, Web uygulaması için geçerli bir “token” yayınlar.
- **Web uygulamasından kullanıcıya akış:** “Token”, daha sonra erişimi doğrulayan kullanıcıya geri dönebilir.

2.4.2 Veri Toplama

Veri toplama adımı, yalnızca erişim izni verilen veri indirilebildiği için sosyal medya madenciliği sürecinin ilk adımı olan yetkilendirme aşamasına bağlıdır (Bonzanini,

2016:19). Hangi sosyal medya platformundan ya da platformlarından, hangi verilerin elde edileceği, belirlenen amaç doğrultusunda değişkenlik gösterebilmektedir. Kaç günlük, haftalık, aylık ya da yıllık verinin sosyal medyadan çekileceğine bağlı olarak değişkenlik göstermekle birlikte, verilerin manuel ya da otomatik bir şekilde toplanması ve bu verilerin saklanması, sosyal medya madenciliğinde önemli bir süreci oluşturmaktadır.

Birçok sosyal medya platformu verilerine API erişimi sağlamaktadır (Ravindran ve Garg, 2015:8). Bir arayüz olan API, bir uygulamanın diğer uygulama ile iletişim halinde olmasını sağlar. Bu sayede API, örneğin bir şirketin Twitter'ı kullanmak için kendi yazılım istemcilerini geliştirmesine ya da Facebook'u diğer sosyal medya servisleriyle entegre etmesine izin vererek temel sosyal medya hizmetlerinin geliştirilmesini sağlar. Araştırmacılar için ise API, sosyal medya platformlarından veri toplamak için kullanılmaktadır. Araştırmacılar yazılım komutlarıyla, analizleri için, sosyal medya platformu kullanıcıları tarafından bırakılan dijital izleri almak, depolamak ve işlemek için API arayüzünü sıklıkla kullanmaktadırlar. API tabanlı araştırmalarda verilerin otomatik bir şekilde toplanabilmesi, araştırmacılar için büyük kolaylık sağlamaktadır. Bu sayede, araştırmalarda sosyal medya verilerini toplamak, sosyal medya kullanıcılarının dijital izlerini elde etmek, anket ve mülakat gibi tekniklerden çok daha fazlasını verebilmektedir (Lomborg ve Bechmann, 2014:256).

2.4.3 Veri Ön İşleme

Veri madenciliği çalışmalarının önemli bir aşaması olan veri ön işleme sürecinde, veri temizleme (data cleaning), veri birleştirme (data integration), veri indirgeme (data reduction) ve veri dönüştürme (data transformations) teknikleri uygulanmaktadır. Bu teknikler birbirinden bağımsız değildir ve sıralı bir şekilde uygulanmaz. Veri seti ihtiyacı ve araştırmanın amacı doğrultusunda bu teknikler uygulanabilmektedir. Veri temizleme genel olarak, veri setindeki eksik verileri doldurmak, aykırı verileri saptamak, gürültüyü temizlemek ve veri setindeki tutarsızlıkları düzeltmek gibi

adımları içermektedir. Veri birleştirme aşamasında ise farklı kaynaklardan gelen veriler, uygun bir veri tabanında birleştirilir. Veri indirgeme, fazla olan değişkenlerin çıkarılması, birleştirilmesi gibi farklı yöntemlerle veri büyüklüğünün azaltılması amaçlanır. Çeşitli sebeplerden dolayı, gerek duyulduğunda normalleştirme gibi veri dönüşüm teknikleri de ön işleme aşamasında kullanılabilmektedir (Han v.d., 2011:83).

Veri ön işleme aşaması bir önceki adımda elde edilen ve depolanan ham verilerin analiz aşamasına hazırlanmasını temsil etmektedir. Bu adım herhangi bir veri madenciliği sorununun temel ve ilk adımlarından biridir. Birleştirilmiş ve temizlenmiş veri kümesinde çalıştırılan bir algoritma, herhangi bir ön işleme yöntemi uygulanmamış veri setinde çalıştırılmasından çok daha hızlı çalışmakla kalmaz, aynı zamanda da daha doğru sonuçlar üretir. Veri seti içerisindeki gürültü verileri tanımlamak ve temizlemek oldukça zor bir süreçtir. Bu süreç ne kadar başarılı gerçekleştirilirse, analiz sonuçlarındaki doğruluk oranı o denli yüksek olur (Ravindran ve Garg, 2015:14). Sosyal medya madenciliği sürecinde karşılaşılan zorluklarının açıklandığı bölümde değinilen gürültü temizleme yanılgısı, bu süreçte dikkat edilmesi gereken en önemli hususlardan birisidir.

2.4.4 Modelleme ve Analiz

Ön işlemeden geçirilen verilere, araştırmanın amacı ya da amaçları ve kurulan model doğrultusunda çeşitli analizler gerçekleştirilmektedir. Bu aşamada öncelikle bir araştırma modeli kurulur. Kurulan modelde hangi analiz yöntemlerinin uygulanacağına karar verildikten sonra gerekli adımlar izlenerek sosyal medya verileri analiz edilir. Sosyal medya madenciliğinde sıklıkla kullanılan analizlere örnek olarak metin madenciliği ve sosyal ağ analizi verilebilir (Bonzanini, 2016:20). Bu çalışmada, sosyal medya madenciliğinde kullanılan analizler ayrı bir başlık altında ele alınarak açıklanmıştır.

Sosyal medya madenciliğinde uygulanan analizler üç başlık altında toplanabilir (Ishikawa, 2015:68-69):

- **İçerik analizi temelli çalışmalar:** Bu alanda yapılan çalışmalar, kullanıcıların belli bir ürün ve hizmet ile ilgili duygularını, düşüncelerini ve şikâyetlerini analiz eder. Kullanıcıların henüz mevcut olmayan bir ürün ve hizmet hakkındaki isteklerinin ve ihtiyaçlarının araştırılması adına çeşitli analizler uygulanır.
- **Yapı analizi temelli çalışmalar:** Yapılar içerikler arası ilişkileri, kullanıcılar arası ilişkileri ve her ikisi (kullanıcılar ve içerikler) arasındaki ilişkileri içerir. Benzer içeriklerin bulunması, benzer kullanıcılardan oluşan toplulukların keşfedilmesi ve etki yaratan kullanıcıların bulunması gibi amaçlar doğrultusunda yapılar analiz edilir.
- **Geçmiş ve değişkenlik analizi temelli çalışmalar:** İçeriğin veya kullanıcıların dinamik analizine dayalı çalışmalardır. Pazar etkinliğinin ölçülmesi, belli olayların tahmini ve aralarındaki nedensel ilişkiler, gelişen eğilimler ve ihtiyaçlar analizler sonucunda belirlenebilmektedir.

2.4.5 Değerlendirme

Sonuç sunumu olarak adlandırılan bu aşama sonuçların değerlendirmesi üzerine kuruludur. Son aşama olan görselleştirme, eğilimler, anomaliler, ilişkiler gibi modelleme ve analiz aşamasının sonucunda elde edilen bulguların, soyut bir boyuttan somut bir hale dönüştürülerek kolayca anlaşılmasını sağlamaktadır (Ravindran ve Garg, 2015:19). Bir veri analizi sürecinin sonuçlarını sunmanın ve aktarmanın etkili bir yolu olan görselleştirme aşaması, farklı disiplinlerce benimsenmiştir (Bonzanini, 2016:16). Kutu ve serpilme diyagramları, kelime bulutları, karar ağaçları ve sosyal ağ analizi sürecinde kullanılan araçlar gibi çeşitli yöntemlerle veri görselleştirme yapılabilmektedir (Ravindran ve Garg, 2015:14).

2.5 Sosyal Medya Madenciliğinde Kullanılan Yöntemler

Sosyal medya madenciliği sürecinde veri, çeşitli ön işleme metotlarından geçirilerek, modelleme ve analiz aşamasına hazır hale gelir. Analiz aşamasında, araştırmanın amacı doğrultusunda değişkenlik göstermekle birlikte, sosyal medya madenciliğinde kullanılan çeşitli yöntemler söz konusudur.

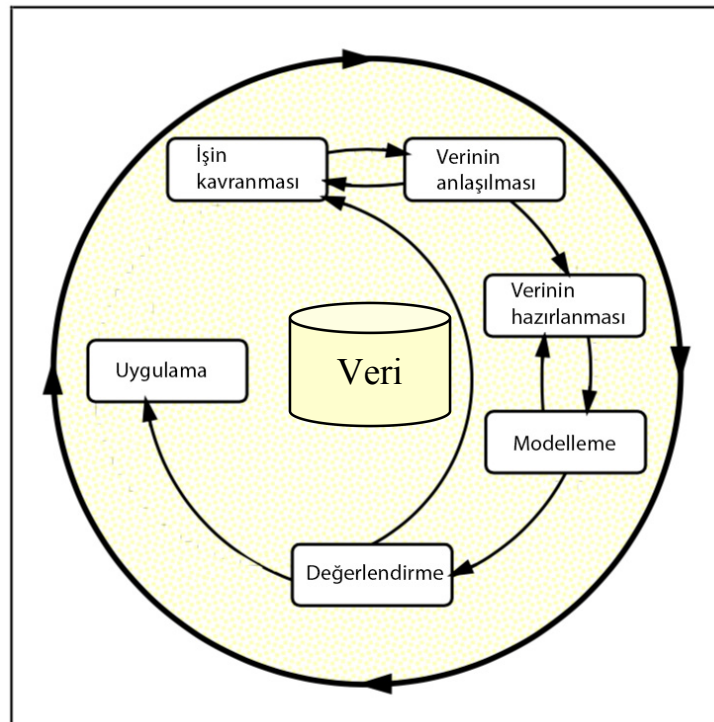
2.5.1 Veri Madenciliği

Büyük veri yığınları içerisinde anlamlı örüntülerin ve kuralların keşfedilmesini amaçlayan veri madenciliği (Hand, 2007:621), tek başına anlamsız olan veriler içerisinde bir amaç doğrultusunda bilgi elde etme işlemi olarak tanımlanabilir (Alpaydın, 2000:1).

Geçmişte veri içerisinde faydalı bilgi elde etme fikri veri madenciliği, bilgi çıkarma, bilgi keşfi, bilgi toplama gibi çeşitli isimlerle adlandırılmış olmasına karşın veri madenciliği terimi veri analistleri ve istatistikçiler tarafından kabul görmüştür. Veri madenciliği ile birlikte sıklıkla kullanılan, bilginin veri analizindeki temel amaç olduğunu vurgulayan Veri Tabanlarında Bilgi Keşfi-VTBK (Knowledge Discovery in Databases) kavramı, veriden yararlı bilginin çıkarılma sürecinin tümünü kapsamaktadır (Fayyad v.d., 1996:39). Veri madenciliği ise, VTBK sürecinin en önemli süreçleri olan modelin kurulması ve değerlendirilmesi aşamalarını oluşturmaktadır. Taşıdığı bu önem neticesinde veri madenciliği literatürde, veriden faydalı bilgi elde etme sürecinin tümü olarak kullanılmaktadır (Akpınar, 2000:2).

Veri madenciliği sürecinde izlenen aşamalar literatürde kabul görmüş standart bir sürece sahiptir. CRISP-DM olarak ifade edilen Çapraz Endüstri Veri Madenciliği Standart Süreci modeli ile Şekil 2.2'deki gibi 6 adımda özetlenmektedir.

Şekil 2.2: Veri Madenciliği Süreci



72

Veri madenciliği çalışmalarında, problemin amacı ve veri setindeki değişkenlerin özellikleri doğrultusunda çeşitli modeller uygulanabilmektedir. Uygulanan bu modeller tahmin edici (predictive) ve tanımlayıcı (descriptive) olmak üzere iki kategori altında değerlendirilebilir. Tanımlayıcı modellerde çeşitli analizler uygulanarak mevcut veriler içerisinde örüntü çıkarımı yapılır. Kümeleme (clustering), birliktelik kuralları (association rules) ve ardışık zamanlı örüntüler (sequential patterns) tanımlayıcı modellerdir (Akpınar, 2000:5). Sınıflandırma (classification), regresyon (regression) ve zaman serileri (time series) ise sonuçları (etiketleri) bilinen mevcut verilerin analiz edilerek gelecekle ilgili öngörülerin yapıldığı tahmin edici modellerdir (Weiss ve Indurkha, 1998:8).

Veri madenciliği tekniklerinin sosyal medya verilerine uygulanması, son yıllarda sosyal medyanın belirgin yükselişiyle giderek daha fazla dikkat çekmektedir. Veri madenciliği, insanların bir konu hakkında sahip oldukları fikirleri daha iyi anlamak, büyük bir yığın içerisindeki grupları belirlemek, grupların zaman içindeki değişimini analiz etmek ve bir ürün/aktivite önermek gibi çeşitli amaçlarla sosyal medya ile birlikte kullanılabilir (Barbier ve Liu, 2011:327-328).

Sosyal medya verileri, veri madenciliği metotları ile analiz edilebilmektedir. Bu veriler, literatürde açıklandığı üzere, içerisinde hem yapılandırılmış hem de yapılandırılmamış verileri barındırmaktadır. İçerisinde metinleri de barındırdığından bu verilerin çeşitli algoritmalar yardımı ile analizinden önce, metin verilerinin işlenerek makinelerin analiz yapabileceği formata dönüştürülmesinin açıklanacağı metin madenciliği bölümü ele alınmıştır.

2.5.2 Metin Madenciliği

Sosyal medya platformları aracılığı ile kullanıcılar, metin temelli içerikler oluşturmaktadırlar. Bu platformlardan elde edilen verilerin işlenmesi aşamasında metin madenciliği metotlarından sıklıkla faydalanılmaktadır. Metin madenciliği, veri

madenciliği gibi, veri kaynaklarından çeşitli bilgilerin çıkarılmasını amaçlamaktadır. veri madenciliği sürecinde analiz edilen veriler yapılandırılmış verilerden oluşurken, metin madenciliği metni veri kaynağı olarak ele almaktadır.

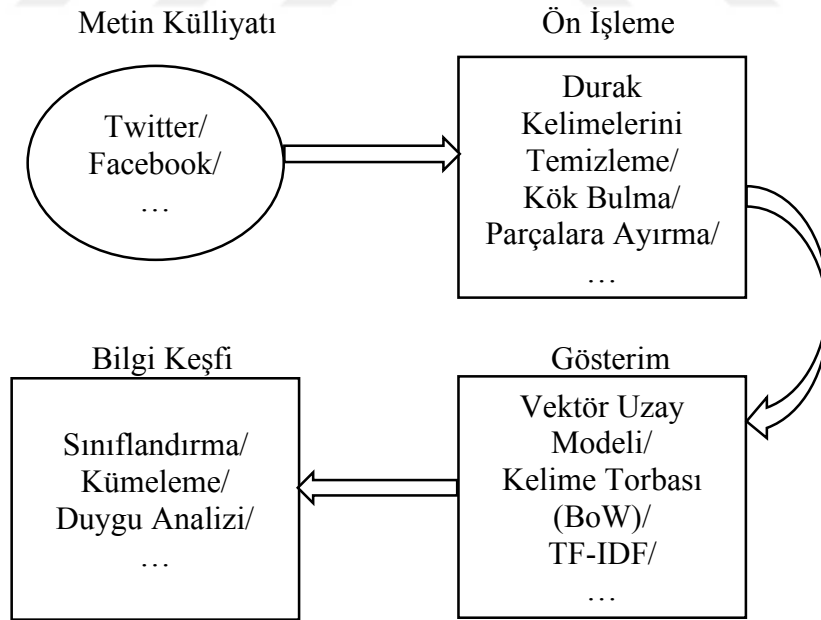
Metin madenciliği, çeşitli veri madenciliği tekniklerini kullanarak yapılandırılmamış veriden enformasyon çıkarımını amaçlamaktadır. Dokümanlardaki kelimeleri ya da kelime gruplarını analiz etme, özetleme veya bu kelimelerin istatistiklerini çıkarma gibi çok geniş uygulama alanı vardır (Gürsakal, 2014:46). Metin madenciliği yedi uygulama alanı bağlamında ele alınabilir (Miner v.d., 2012:35-38):

- **Arama ve Bilgi Getirimi (Search and Information Retrieval):** Arama ve bilgi getirimi, büyük metin veri tabanlarından anahtar kelime sorgulamalarını, metinlerin endekslenmesini, aranmasını ve elde edilmesini kapsamaktadır. Google ve Yahoo! gibi İnternet arama motorlarının yükselişi ile arama ve bilgi getiriminin bilinirliği artmıştır. E-postadan kelime işlemeye kadar neredeyse her bilgisayar uygulaması bir arama fonksiyonu içermektedir.
- **Metin Kümeleme (Document Clustering):** Bu alanda yapılan çalışmalar veri madenciliği kümeleme yöntemlerini kullanarak terimleri, paragrafları veya belgeleri gruplama ve kategorilere ayırma işlemlerinden oluşmaktadır.
- **Metin Sınıflandırması (Document Classification):** Metin sınıflandırması, etiketsiz belgelere bilinen etiketlere sahip belgelerden öğrenilen bir model kullanarak bilinen bir etiket kümesi atama işlemidir. Metin kümeleme gibi, metin sınıflandırması da veri madenciliği, istatistik ve makine öğrenme metotlarından yararlanmaktadır. Metin madenciliğinde kullanılan en önemli tekniklerden birisidir.
- **Web Madenciliği (Web Mining):** İnternette ölçek ve birbirine bağlılığa odaklanan veri ve metin madenciliği çalışmalarından oluşmaktadır. Web madenciliği, Web'deki benzersiz yapı ve büyük veri hacmi nedeniyle ayrı uygulama alanına sahiptir. Web dokümanları tipik olarak sayfalar arası köprülerle yapılandırılmış bir metin biçimindedir.

- **Bilgi Çıkarımı (Information Extraction):** Bu alanda yapılan çalışmalardaki amaç yapılandırılmamış metinden ilgili olguların ve ilişkilerin tanımlanması ve çıkarılması; yapılandırılmamış ve yarı yapılandırılmış metinlerden yapılandırılmış veri oluşturma sürecini kapsayan çalışmalardır.
- **Doğal Dil İşleme (Natural Language Processing):** Son zamanlarda doğal dil işlemenin odak noktası, istatistiksel yaklaşımları göz önünde bulundurarak metin madenciliği alanına daha da ilerlemiştir. Dil işleme ve anlama görevlerini (ör., konuşmanın bir bölümünü etiketleme) içeren çalışmalardır.
- **Kavram Çıkarımı (Concept Extraction):** Sözcük ve ifadelerin semantik olarak benzer gruplara atanma çalışmalarıdır.

Metin madenciliği aşamaları Şekil 2.3’de görüldüğü üzere üç başlık altında özetlenebilmektedir.

Şekil 2.3: Metin Madenciliği Süreci



Kaynak: Hu ve Liu, 2012:389

Metnin çeşitli kaynaklardan elde edilmesi, analize hazırlanması (ön işleme ve gösterim) ve ardından analiz edilerek bilgi keşfinin gerçekleştirilmesi, geleneksel metin madenciliği sürecidir. Bu üç aşama metin ön işleme, metnin gösterimi ve metin analizi başlıkları altında ayrı ayrı ele alınarak açıklanmıştır.

2.5.2.1 Ön İşleme

Metin ön işleme, metnin bir nevi sayısal veri şekline getirilmeye çalışıldığı aşamadır. Çalışmanın amacı doğrultusunda değişkenlik göstermekle birlikte, genellikle izlenen metin ön işleme adımları Tablo 2.1’de verilmiştir.

Tablo 2.1: Metin Madenciliği Ön İşleme Adımları

Adım	Ön İşleme Metodu	Ön İşleme Sonucu
1	Parçalara Ayırma (Tokenize)	Belgedeki her kelime veya terim ayrı bir parçaya dönüşür.
2	Durak Kelimelerini Temizleme (Stopwords Removal)	Çok yaygın kullanılan edat, bağlaç gibi kelimeleri temizlenir.
3	Filtreleme (Filtering)	Durak kelimeleri dışındaki dokümanda sıklıkla tekrar eden kelimeler çıkarılır.
4	Kök Bulma (Stemming)	Her kelime köküne indirgenir.
5	Tamlamalar (N-gram)	Kelimeleri n sayısı kadar birlikte alarak birleştirir.

Kaynak: Kotu ve Deshpande, 2014:283

Metin madenciliği ön işleme aşamalarının ilki olan parçalara ayırma (tokenize), bir metni kelimelere, sembollere veya diğer anlamlı unsurlara bölme işlemi olarak tanımlanmaktadır. Bu ön işlemin amacı metin içerisindeki kelimelerin araştırılmasıdır (Verma v.d., 2014:16). Metin bölümlenme olarak da adlandırılan bu işlem, noktalama

işaretleri ve boşluk gibi karakterler dikkate alınarak gerçekleştirilmektedir (Uysal ve Gunal, 2014:106).

Bir diğer sıklıkla kullanılan ön işleme aşaması ise dokümanda yer alan metinlerin aynı boyuta, küçük karaktere dönüştürülmesi işlemi uygulanır. Büyük ve küçük karakterdeki kelimelerin anlamları arasında bir farklılık olmadığından dolayı, analize başlamadan önce tüm karakterler küçük boyuta çevrilmektedir (Uysal ve Gunal, 2014:106).

Metnin içerisinde sıklıkla tekrar eden ve tek başlarına herhangi bir anlamı bulunmayan edat ve bağlaçlar gibi kelimeler durak kelimeleri (stopwords) olarak adlandırılmaktadır. Bu kelimelerin analize bir katkısı olmayacağı için temizlenmesi gerekmektedir. Filtreleme ise, metin içerisindeki bazı kelimelerin araştırmacının kendi isteğiyle elemesi üzerine kuruludur. Karakter sayısına ya da bir sözlüğe göre filtreleme işlemi gerçekleştirilerek metin içerisinden analize dâhil edilmek istenmeyen kelimeler çıkarılabilir (Verma v.d., 2014:17).

Dokümanda yer alan parçaları (kelimeleri) yalın hale getirerek, yapılandırılmamış formata sahip bu kelimeleri yapılandırarak, yalnızca tek bir ortak kelime dâhilinde frekans hesaplamalarının yapılabilmesi için kök bulma (stemming) işlemi gerçekleştirilir (Kotu ve Deshpande, 2014:282). Türkçe dokümanlarda kelime kökü bulmada Zemberek Kütüphanesi ve sabit ön ek algoritması (fixed-prefix algorithm) kullanılırken, İngilizce metinler için ise Porter algoritması yaygın olarak kullanılmaktadır (Uysal ve Gunal, 2014:106). Bu tez çalışmasında gerçekleştirilen uygulamanın birinci bölümünde, kelimeleri köküne indirgeme aşamasında RapidMiner aracında yer alan “Parçaları Değiştirme (Replace Tokens)” isimli operatör kullanılmıştır.

Tamlamalar (N-gram) ön işleme aşaması ise kelimelerin tekli gösterimleri yerine n sayısı kadar birlikte kelime ele alınması ile oluşmaktadır. Örneğin “sosyal medya madenciliği” kelime öbeği parçalara bölündüğünde “sosyal”, “medya”, “madenciliği”

olarak sonuç alınırken, 2’li tamlamalar yöntemi uygulandığında sonuç “sosyal medya” ve “medya madenciliği” olmaktadır.

Metin ön işleme aşamaları temel olarak bu adımlardan oluşmakla birlikte, çalışmanın amacı ve uygulanmak istenen analiz doğrultusunda değişkenlik gösterebilmektedir. Örneğin kişilerin üslubu ile ilgili çıkarımlarda bulunmak amacı ile gerçekleştirilen bir metin madenciliği çalışmasında, durak kelimelerini çıkarmama yoluna giderek kişilerin ne sıklıkla ama, fakat gibi bağlaçlar kullandığına bakılıp karşılaştırmalar yapıp üslupla ilgili çıkarımlar yapılabilir.

2.5.2.2 Gösterim

Metin ön işleme adımları gerçekleştirildikten sonra dokümanda yer alan kelimeler (terimler) ve dokümanın vektör uzay modeli (vector space model) olarak adlandırılan bir matris üzerinde gösterimi gerçekleştirilir. Bu terim-doküman matrisinde her kelimenin o içinde yer aldığı dokümana göre ağırlık değeri verilerek terim-doküman arasındaki uzaklık belirlenir (Runeson v.d., 2007:503). Literatürde çeşitli kelime ağırlıklandırma yöntemleri olsa da en yaygın olarak kullanılan yöntem TF-IDF’dir.

Açılımı Terim Frekansı – Ters Metin Frekansı (Term Frequency – Inverse Document Frequency) olan TF-IDF ağırlıklandırma algoritması, arama motorları tarafından da kullanılmaktadır. Arama motorlarında bir cümle sorgulandığında, bu cümledeki kelimeler ele alınarak öncelikle nadir sayılabilecek kelimelere yüksek bir ağırlık verilir, ardından bu nadir anahtar kelimelerin çok sayıda örneğini içeren web sayfalarına yüksek bir ağırlık verir. Böylece, yalnızca nadir bulunan anahtar kelimeleri içermeyen, bununla birlikte nadir bulunan anahtar kelimelerin çok sayıda örneğine sahip olan sayfalar, arama sonuçlarında en üst sırada yer alırlar (Kotu ve Deshpande, 2014:278).

TF-IDF hesaplaması denklem...’de verildiği üzere gerçekleştirilmektedir. Bu denkleme göre $w_{i,j}$, i teriminin j dokümanındaki ağırlığı, N külliyattaki doküman sayısı, $tf_{i,j}$ i teriminin j dokümanındaki sayısı ve df_i ise i teriminin doküman içerisindeki sayısıdır (Zhang v.d., 2011:2760).

$$w_{i,j} = tf_{i,j} \times \log \frac{N}{df_i}$$

TF-IDF iki ayrı ağırlıklandırma hesaplamasının çarpımından oluşmaktadır. TF ve IDF ayrı ayrı da çalışmalarda ağırlıklandırma yöntemi olarak kullanılmaktadır. Terim frekansı (TF) terimin dokümanda geçme sayısının dokümandaki kelime sayısına oranlanması ile bulunurken, Ters Metin Frekansı (IDF) terimin sıklığına değil, Denklem...’de görüldüğü üzere, nadirliğine odaklı hesaplanır. Örneğin; 100 dokümandan oluşan bir külliyatta “Twitter” terimi her dokümanda bulunuyorsa $\log(100/100)$ değeri 0 çıkarak bu terimin ağırlıklandırması yapılmış olacaktır.

2.5.2.3 Bilgi Keşfi

Metin ön işleme ve ağırlıklandırma aşamalarının ardından analize hazır hale gelen metin verisi, çeşitli veri madenciliği metotları ile analiz edilebilmektedir. Metnin içerisinden bilgi keşfi için uygulanan en yaygın veri madenciliği yöntemleri; kümeleme sınıflandırma, birliktelik kuralları ve duygu analizidir. Duygu analizi tezde başvurulmuş bir yöntem olduğu için ayrı bir başlıkta ele alınarak açıklanırken, bu bölümde diğer yöntemlere kısaca değinilmiştir.

2.5.2.3.1 Sınıflandırma

Sınıflandırma çalışmalarında amaç, eğitim veri setinin özelliklerine dayalı bir model oluşturarak bu modeli test verisi olarak adlandırılan yeni veri setini sınıflandırmada kullanmaktır (Ghosh v.d., 2012:224). Örneğin bir haber metnini “spor”, “politika” ve “sanat” etiketlerinden birine otomatik olarak atamak bir metin sınıflandırma

çalışmasıdır. Hangi algoritma kullanılırsa kullanılsın bu veri madenciliği tekniği, $e \in E$ (örneğin; spor, politika ve sanat) sınıfı ile etiketlenmiş olan $D = (d_1, \dots, d_n)$ dokümanlardan oluşan bir eğitim veri seti ile başlamaktadır. Bu eğitim verisi ele alınarak oluşturulan sınıflandırma modeli ile yeni bir veri setine ait olan d dokümanı için etiketleme gerçekleştirilebilmektedir (Hotho v.d., 2005:10).

$$f: D \rightarrow L \quad f(d) = e$$

Sınıflandırma modelinin performansını ölçmek için etiketlenmiş veri setinin bir bölümü rastgele ayrılarak test verisi için kullanılır. Eğitim verisi ile oluşturulan sınıflandırma modeli test verisine uygulanır. Sınıflandırma sonucunda en basit hesaplanan performans ölçümü olarak bilinen, test verisindeki doğru olarak sınıflandırılmış doküman sayısının toplam doküman sayısına oranı olan doğruluk (accuracy) elde edilir. Bununla birlikte, test verisi genellikle külliyyatın yalnızca küçük bir yüzdesini kapsadığı için doğruluk tek başına yeterli olmayabilir. Kesinlik (precision) getirilen bilginin ne kadarının istenilen bilgi ile ilgili olduğuna bakarken, duyarlılık (recall) getirilmesi gereken bilginin ne kadarının getirildiğini verir. F-ölçütü ise sınıflandırma modelinin genel performansını ölçmek için bu iki ölçütü (duyarlılık ve kesinlik) birlikte ele alan bir ölçümdür (Hotho v.d., 2005:10-11).

Doğruluk, kesinlik, duyarlılık ve F-ölçütü bir karışıklık matrisi ile ifade edilebilmektedir. Tablo 2.2’de verilen karışıklık matrisinde satırlar test verilerinin gerçek etiketlerini, sütunlar ise tahminlerini vermektedir.

Tablo 2.2: Karışıklık Matrisi

		Öngörülen Sınıf	
		Sınıf 1	Sınıf 2
Doğru Sınıf	Sınıf 1	TP(True Pozitif)	FN (False Negatif)
	Sınıf 2	FN (False Negatif)	TN (True Negatif)

Kaynak: Coşkun ve Baykal, 2011:3

Tablo 2.2’de verilen göstergeler doğrultusunda modelin başarısını ölçen göstergeler aşağıdaki gibi formüle edilmektedir (Coşkun ve Baykal, 2011:3-4):

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Kesinlik} = \frac{TP}{TP + FP}$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN}$$

$$F - \text{ölçütü} = \frac{2 \times \text{Duyarlılık} \times \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}}$$

Bu başarı ölçütleri temelinde en iyi sonucu veren, araştırmanın modeline en uygun algorithmada karar kılınarak sınıflandırma gerçekleştirilir. Karar ağaçları (decision trees), kural tabanlı algoritmalar (rule-based classifiers), destek vektör makineleri (SVM), sinir ağları (neural networks) ve istatistiksel/olasılıksal sınıflandırmalar (bayesian/generative classifiers) metin sınıflandırmada kullanılan metotlardan en yaygın olanlarıdır (Aggarwal ve Zhai, 2012: 165-166).

2.5.2.3.2 Kümeleme

Kümeleme, veri setini benzer nesnelerden oluşan gruplara ayırmaktır. Kümelere ayrılan her nesne kendi aralarında benzer olan ve diğer kümelere ait elemanlara benzemeyen nesnelerden oluşmaktadır (Ghosh v.d., 2012:230).

Kümeleme algoritmaları hiyerarşik kümeleme (hierarchical clustering), bölümlenmeli kümeleme (partitional clustering) ve spektral kümeleme (spectral clustering) olmak üzere üç ana başlık altında toplanmaktadır. Hiyerarşik kümeleme, dokümanları kademeli olarak gruplandırır. Bu kümeleme yöntemi, yukarıdan aşağıya (divisive) veya aşağıdan yukarıya (agglomerative) doğru uygulanabilir. Bu yolla veriler

benzerlik ölçütlerine göre bölünmüş veya birleştirilmiş olur. Spektral kümeleme algoritmaları, boyut küçültme için matris işlemlerini kullanır ve indirgenmiş boyutlara göre kümeler üretir. Bölümlemeli kümeleme teknikleri ise hiyerarşik kümelemeden farklı olarak, tek seferde tüm kümeleri üretir. Bu kümeleme tekniklerinden en yaygın olarak kullanılan k-ortalamlar (k-means) yöntemidir (Miner v.d., 2012: 960).

Kümeleme algoritmalarının metin verilerine uygulanması birçok alanda kullanılmaktadır. Belgelerin tutarlı kategoriler halinde hiyerarşik organizasyonu, doküman koleksiyonunun sistematik olarak taranması için çok yararlı olabilmektedir. Bunun dışında metnin içerisinde yer alan kelimelerin kümelenmesi gibi tekniklerle dokümanın genel bir özeti çıkarılabilmektedir. Bunlara ek olarak metin kümeleme algoritmaları sınıflandırma algoritmaları ile birlikte çalıştırılarak model başarılarının artırılmasına yardımcı olabilir (Aggarwal ve Zhai, 2012: 78).

2.5.2.3.3 Birliktelik Analizi

Bir veri setinde yer alan, minimum destek ve güven oranını sağlayan ilişki kurallarını bulmak, birliktelik analizinin amacıdır (Kotsiantis ve Kanellopoulos, 2006:71). $n \geq 2$ olmak üzere $S = \{i_1, i_2, \dots, i_j, \dots, i_n\}$ bir dizi öğeyi temsil etmiş olsun. $X \subset S$ ve $ij \notin X$ olmak üzere $ij \in S$ durumlarında, $X \rightarrow ij$ bir ilişki kuralı formudur. Bu kuralda X öncül, ij ise ilişki kuralının sonucudur. $X \rightarrow ij$ kuralının destek oranı, X ve ij öğelerini birlikte içeren S öğelerinin yüzdesi olarak tanımlanır. Kuralın güven oranı ise, X 'i ve ij 'yi de karşılayan öğelerin yüzdesidir (Wong v.d., 1999:120).

Örneğin, veri setinde yer alan kelimeler arasındaki birliktelikler bir algoritma (apriori, fp-growth vb.) ile analiz edilmiş ve iki kelime arasındaki birlikteliğin destek oranı %15 ve güven oranı %95 olarak bulunmuş olsun. Bu birliktelik, öncül kelimenin %95'inin sonuç kelimesi ile birlikte kullanıldığı ve bu birlikteliğin tüm kayıtların %15'ini oluşturduğu şeklinde yorumlanır.

Birliktelik kuralları, bir web arama motoru, bilgi çıkarımı (IE) ve metin arama motoru gibi metin madenciliği uygulamalarının çoğunda ana rollerden biri olarak katkıda bulunur (Manimaran ve Velmurugan, 2013:3). Metin veri tabanlarından çıkarılan ilişki kuralları, sayısal veri tabanlarından alınanlara göre oldukça farklı bir şekilde kullanılmaktadır. Kullanımdaki bu farklılık, metin veri tabanlarındaki ilişki kuralları için daha düşük minimum destek seviyesi gerektirmektedir (Holt ve Chung, 2001:182).

2.5.3 Duygu Analizi

Literatürde fikir madenciliği (opinion mining) olarak da adlandırılan duygu analizi (sentiment analysis), insanların fikirlerini, görüşlerini, değerlendirmelerini, tutumlarını ve duygularını, ürünler, hizmetler, kuruluşlar, bireyler, konular, olaylar, başlıklar ve bunların nitelikleri gibi faktörlere göre analiz eden çalışma alanıdır. 2000 yılından sonra bu alanda yapılan çalışmalar, başta sosyal medya platformlarının sağlamış olduğu ortam ile birlikte Web’de büyük boyutlarda fikir beyan eden verilerin üretilmesi sebebiyle aktif bir çalışma alanı haline gelmiştir. Duygu analizi başlarda yalnızca Doğal Dil İşleme (NLP) çalışmalarının odağı iken zamanla sosyal medya araştırmalarının merkezinde konumlanmıştır. Bu sebeple duygu analizi, yönetim bilimleri, siyaset bilimi, ekonomi ve sosyal bilimler üzerinde de önemli çalışmalarda başvuru alanı bir analizdir (Liu, 2012:1-2).

Literatür taramasında yer alan çalışmalar ele alındığında, duygu analizi çalışmaları sözlük ve makine öğrenmesi algoritmaları olmak üzere iki temelde yoğunlaşmaktadır. Özellikle Türkçe metinlerde kabul görmüş bir sözlük olmadığından, makine öğrenmesi algoritmalarının kullanılarak duygu durumu belirleme başarılarının karşılaştırıldığı çalışmalar tercih edilmektedir.

Sözlük yaklaşımında, olumlu (örneğin; güzel, harika ve iyi gibi) ve olumsuz (örneğin; kötü, yetersiz ve berbat gibi) duyguları ifade etmek için yaygın olarak kullanılan

kelimelerden oluşan görüş kelimeleri (Liu ve Zhang, 2012:423) toplandıktan sonra geliştirilerek duygu durumları belirlenebilmektedir (Can ve Alatas, 2017:94).

Makine öğrenme yaklaşımı ile duygu durumlarının belirlenmesi ise, bir sınıflandırma çalışması olarak görülmektedir. Genellikle olumlu, nötr ve olumsuz olmak üzere üç sınıflı bir eğitimli öğrenme problemi olarak ele alınmaktadır. Destek Vektör Makineleri (SVM) ve Naive Bayes gibi herhangi bir öğrenme algoritması uygulanarak duygu durumlarının belirlenmesi gerçekleştirilebilmektedir (Liu ve Zhang, 2012:422-423). Denetimli öğrenme metotları etiketli veriler üzerinden öğrenme gerçekleştirerek test verilerinin duygu durumlarını belirlerken, denetimsiz öğrenme metotları etiketli verilerin olmadığı durumlarda kullanılır (Can ve Alatas, 2017:87).

Bu iki genel başlık dışında metin verilerinin duygu durumları araştırmacılar tarafından manuel olarak da belirlenebilmektedir. Duygu durumlarının manuel olarak belirlenmesinde literatürde en az iki araştırmacının kodlamayı gerçekleştirmesi ve araştırmacılar arası güvenilirliğin belirli bir oranı sağlaması beklenmektedir. Araştırmacılar tarafından gerçekleştirilen kodlama işlemi sonucunda, araştırmacılar arasındaki güvenilirlik, puanlayıcılar arası güvenilirlik (interrater reliability) adı ile literatürde yer alan tekniklerinden faydalanılarak gerçekleştirilmektedir.

Puanlayıcılar arası güvenilirlik iki ya da daha fazla puanlayıcı tarafından yapılan kodlamalar arasındaki uyumun belirlenmesi sonucu hesaplanmaktadır. Puanlayıcılar arası güvenilirlik sonucunda elde edilen katsayı, araştırmacılar arasındaki fikir birliği hakkında bilgi vermektedir (Bilgen ve Doğan, 2017:65). Literatürde farklı güvenilirlik teknikleri bulunmaktadır. Yaygın kullanımları sebebi ile Yüzde Uyumu, Kappa İstatistiği ve Krippendorff Alfa Katsayısı sırası ile açıklanmıştır.

Yüzde Uyumu (Percent Agreement) hesaplama ve yorumlama kolaylığı sebebi ile araştırmalarda sıklıkla kullanılmaktadır. Yöntem iki kodlayıcı ile ve sadece kategorik veriler üzerine kuruludur. Bu teknik tesadüfen oluşabilecek uyumu hesaba katmaz. Bu sebeple, yüksek düzeyde hassasiyet gerektiren çalışmalarda etiketleme işlemi yapanlar

arasında uyumu kontrol etmek için çok tercih edilen bir yöntem değildir (Nili v.d., 2017:3). İki kodlayıcı için her kodlanan satır arasındaki farklar alınarak toplam uyum sayısının (farkların sıfır olduğu satır sayısının toplamı) toplam veri sayısına bölünmesi ile elde edilir.

Kappa İstatistiği (Cohen's Kappa) puanlayıcılar arası g güvenilirlikte güçlü (robust) bir istatistiktir. Korelasyon katsayılarına benzer şekilde bu istatistiğin değer aralığı, 1 mükemmel uyumu temsil etmek üzere, -1 ile +1 arasında değişmektedir. Kappa istatistiğinin negatif değerleri mümkün olmakla birlikte, Cohen bunun pratikte olası olmadığını belirtmiştir (McHugh, 2012:279).

Tablo 2.3: Kappa İstatistiği Değer Aralıkları ve Yorumları

Kappa İstatistiği	Uyumun Gücü	Verinin Güvenilirliği
0-0,20	Yok (None)	%0-4
0,21-0,39	Minimal (Minimal)	%4-15
0,40-0,59	Zayıf (Weak)	%15-35
0,60-0,79	Orta (Moderate)	%35-63
0,80-0,90	Güçlü (Strong)	%64-81
0,90-1	Çok Güçlü (Almost Perfect)	%82-100

Kaynak: McHugh, 2012:279

Tablo 2.3’de verildiği üzere, araştırmacılar arasındaki uyum %80’den az olduğunda, verilerin %20’sinden fazlasının hatalı olduğu söylenmektedir. $\bar{P}$ gözlenen uyumluluk ve $\bar{P}_e$ şans uyumluluk oranı olmak üzere kappa istatistiği (k) aşağıdaki gibi hesaplanmaktadır (Bilgen ve Doğan, 2017:66).

$$k = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

Krippendorff Alfa Katsayısı (Krippendorff's Alpha) ise şansın kodlamadaki etkisini minimuma indiren, her türlü veri türüne ve kodlayıcı sayısına uygulanabilir oluşu ile diğer güvenilirlik yöntemlerinin sınırlarını ortadan kaldıran, eksik verilerin söz konusu olduğu kodlamalarda da güvenilirliğin ölçülmesini sağlayan bir yöntemdir. Mükemmel uyum 1 ile ifade edilirken 0, herhangi bir güvenilirlik derecesinin

olmadığını gösterir. 0,9'un üzerinde herhangi bir sonucun her zaman kabul edilebilir olduğuna dair genel bir fikir birliği vardır. Genelde 0.8'in üzeri uygun görülürken, keşif çalışması için 0,7 ve üzeri katsayılar kabul edilebilir (Nili v.d., 2017:4).

Tablo 2.4: Krippendorff Alfa Katsayısı Değer Aralıkları ve Yorumları

Alfa Katsayısı	Uyumun Gücü
<0,67	Zayıf
0,67-0,80	Orta
0,80≤	Güçlü

Kaynak: Bilgen ve Doğan, 2017:67

Alfa katsayısı (α), $D_{gözlenen}$ gözlenen uyumsuzluk (observed disagreement) ve $D_{beklenen}$ beklenen uyumsuzluk (expected disagreement) değerine karşılık gelmekle birlikte aşağıdaki formül yardımı ile hesaplanmaktadır (Krippendorff, 1995:52).

$$\alpha = 1 - \frac{D_{gözlenen}}{D_{beklenen}}$$

ÜÇÜNCÜ BÖLÜM

UYGULAMA

3.1 Araştırmanın Amacı

Sosyal medya hedef kitlelere ulaşmada sağlamış olduğu avantajlar sebebi ile özellikle firmalar için vazgeçilmez bir araç haline gelmiştir. Sosyal medyanın bu denli önem teşkil etmesinin sebebi “Güçteki Temel Değişim” olarak adlandırılan iki yönlü iletişim ile birlikte tüketicilerin, firmaların paylaşmış oldukları kurumsal ve reklam mesajlarının ötesinde, diğer tüketicilerin firma ürünleri ve hizmetleri ile ilgili bilgilere ulaştıkları bir platformlar bütünü oluşudur (Safko, 2012:5). Teşkil ettiği büyük önem neticesinde firmaların sosyal medyayı ne derece başarılı kullandıkları elzem bir konu haline gelmiştir.

Bu çalışmada farklı sektörlerde faaliyet gösteren firmaların sosyal medya istatistiklerinin ve sosyal medyayı kullanım amaçlarının belirlenmesi ve ardından özel olarak kullanıcı-firma etkileşimi yaratma çabalarının etkinliğinin ölçülmesi amaçlanmıştır. Firmalar yalnızca tüketicileri değil, bu sınıfa dâhil olabilecek potansiyel tüketicileri de göz önünde bulundurarak paylaşım yapmakta ve benzer şekilde tüketiciler de ürün ve hizmetlerini satın almadıkları firmalar hakkında da içerik paylaşımı yapmaktadırlar. Araştırmanın temelinde sosyal medyanın var olması sebebi ile tüketici kavramı yerine bu çalışmada kullanıcı kavramı kullanılarak, firmalar ile ilgili Twitter’da paylaşım yapan bireyler ya da kurumlar tasvir edilmiştir.

Araştırmanın amacına ulaşmak için sorulan araştırma soruları aşağıda verilmiştir:

- Firmaların sosyal medya istatistikleri nelerdir, bu istatistikler firmalar arası farklılık ve firmalara göre uyum göstermekte midir?
- Firmalar sosyal medyada ne tür içerikler paylaşmaktadırlar ve bu içerik türlerinden hangileri en yüksek etkileşimi sağlamaktadır?

- Kullanıcı-firma etkileşimi yaratmada firmaların sosyal medya etkinliklerinin ölçümü sonucunda hangi firma ya da sektör daha başarılı olmuştur?

Firmaların Twitter paylaşımları üzerine kurulu ilk iki araştırma sorusu uygulamanın birinci bölümünde, özel olarak belirlenen bir döneme ait firmaların Twitter etkinliklerinin ölçülmesinin amaçlandığı üçüncü araştırma sorusu ise uygulamanın ikinci bölümünde çeşitli analizler gerçekleştirilerek cevaplanmaya çalışılmıştır.

3.2 Araştırmanın Örneklemi

Sosyal medya platformu Twitter, diğer sosyal ağlara kıyasla, kullanıcıların kendi kendilerine organize olmalarını daha fazla teşvik eden bir platformdur. Kullanıcılar içgörü toplamak, tavsiyelerde bulunmak ve genel şikâyetlerini bildirmek amaçlı bu sosyal medya platformunu aktif olarak kullanmaktadırlar (Danneman ve Heimann, 2014:33). Sosyal medya platformları arasında firma ve kullanıcıların katılım düzeylerinin yüksek olması ve kullanıcıların görüşlerinin paylaşılmasını teşvik eden yapısı ile mikroblog siteleri arasında en çok kullanılan sosyal medya platformu olan Twitter, bu uygulamada sosyal medya platformu örnekleme olarak belirlenmiştir. Bu çalışmada yer alan Twitter terimleri ile ilgili gerekli bilgiler Tablo 3.1’de verilmiştir.

Tablo 3.1: Twitter Terim İsimleri ve Açıklamaları

Terim İsmi	Açıklama
Tweet	280 karakterden oluşan Twitter paylaşımlarına verilen isim. Çalışmada tweet paylaşmak (firma/kullanıcı tarafından paylaşılan tweetler vb.) ifadesine sıklıkla başvurulmuştur.
Rt (Retweet)	Türkçesi tekrar tweetlemek olan retweet, başkasına ait olan tweeti kaynak göstererek, kişinin kendi sayfasında paylaşmasıdır. Çalışmanın bundan sonrasında rt olarak ifade edilmiştir.
Beğeni	Tweetlerin altında bulunan kalp işaretine tıklanarak tweet beğeni listesine eklenmiş olur.
Hashtag (#)	“#” sembolü kullanılarak bir konu hakkında paylaşılan içeriklerin kolayca bulunması sağlanabilmektedir.

Mention	Mention, tweet içerisinde kullanıcıların @kullanıcıadı şeklinde birbirlerinden bahsetmesidir.
---------	---

Firmaların Twitter istatistikleri ve paylaşım kategorileri ile ilgili analizleri gerçekleştirmek için örneklem dâhilinde seçilen firmaların 2017 yılı Twitter paylaşımları ele alınmıştır.

Firmaların, kullanıcı-firma etkileşimi yaratma konusundaki Twitter etkinliklerini ölçümlemek amacıyla firmaların Twitter üzerinden kullanıcılarına özel paylaşımlarda bulunduğu bir kampanya dönemini (özel gün) içeren bir aylık zaman dilimi belirlenmiştir. 14 Şubat Sevgililer Günü farklı sektördeki birçok firmanın sosyal medya üzerinden çeşitli kampanyalarla tüketici etkileşimi yaratma çabası gösterdiği bir özel gün olması bakımından bu çalışmanın örneklemini dâhilinde seçilmiştir. Bu özel günü içeren 2018 yılı Şubat ayı boyunca, kullanıcıların firmalar hakkında Twitter platformu üzerinden yapmış oldukları paylaşımlar ile birlikte firmaların belirlenen Şubat ayı Twitter paylaşımları ele alınmıştır.

Sektör ve bu sektörlerdeki rakip firmaların belirlenmesinde ise firmaların Twitter platformu üzerindeki aktif katılımları ve firmaların rakip temelli karşılaştırılabilir oluşu göz önünde bulundurulmuştur. Firmaların sosyal medya katılımları ile ilgili çeşitli karşılaştırma ve raporlamalara olanak tanıyan bir site olan BoomSocial platformuna ait verilerden bu araştırmanın örnekleminin belirlenmesinde faydalanılmıştır. Bu platformda yer alan çeşitli sektör sıralama listeleri gözden geçirilmiş ve alışveriş olarak adlandırılan sektöre ait sıralamada yer alan ilk yirmi firma incelenmiştir. Bu yirmi firma arasından aynı alt sektör içerisinde yer alan ve en fazla takipçi sayısına sahip olan iki rakip firmanın bulunduğu üç alt sektör belirlenmiştir. Bu alt sektörler kozmetik, pazaryeri ve elektronik olarak adlandırılmış ve açıklamaları verilmiştir. Firma isimleri çalışmada açıklanmamış, herhangi bir sıralama gözetmeksizin Tablo 3.2’de gösterildiği üzere numaralandırılmıştır.

Tablo 3.2: Uygulamanın Örneklemini Oluşturan Sektör ve Firma Bilgileri

Sektör Adı	Firma	Twitter Takipçi Sayısı (01.01.2018)	Twitter Başlangıç Tarihi	Firma Kuruluş Tarihi	Kanal	Mağaza Sayısı
Kozmetik	Firma 1	374.654	2010	2009	Tek Kanallı (Mağaza)	500+
Kozmetik	Firma 2	128.784	2010	2005	Çok Kanallı	300+
Pazaryeri	Firma 3	430.313	2009	1998	Tek Kanallı (Online)	-
Pazaryeri	Firma 4	308.045	2012	2013	Tek Kanallı (Online)	-
Elektronik	Firma 5	547.594	2013	2007	Çok Kanallı	63
Elektronik	Firma 6	601.926	2008	2000	Çok Kanallı	210

Firmaların kuruluş tarih bilgileri yurt dışı menşeli firmalar için Türkiye’de faaliyet göstermeye başladıkları yıl ele alınarak yazılmıştır. Twitter başlangıç tarihleri ise firmaların Twitter platformuna katıldıkları tarih bilgisini içermektedir. Kanal sütununda ise firmaların ürünlerini hangi kanal aracılığı ile tüketicilere ulaştırdığının bilgisi verilmiştir. Mağazaya sahip firmaların bilgilerine kurumsal Web sitelerinden ulaşılmıştır.

Çalışmada kozmetik olarak adlandırılan alt sektörde faaliyet gösteren iki rakip firma, güzellik ve kişisel bakım ürünleri perakende zinciridir. Firma 2 yurt dışı, Firma 1 ise yurt içi menşeli bir holding bünyesine aittir. Kişisel bakım market zincirlerinden oluşan bu iki firmadan Firma 1, Türkiye’de mağaza sayısı üstünlüğünü elinde bulundurmaktadır. Firma 2 online alışverişe imkân tanırken, Firma 1 yalnızca mağaza zincirleri aracılığı ile tek kanal üzerinden faaliyet göstermektedir.

Bir diğerk alt sektör olarak belirlenen pazaryerinde faaliyet gösteren iki firma ise yalnızca online alışverişin gerçekleştirilebildiğı e-ticaret firmalarıdır. Firma 4, yurt dışı ve yurt içi iki grubun ortak girişimi sonucunda kurulmuştur. Firma 5 ise yurt içi menşeilili bir holding bünyesinde faaliyet göstermektedir. Çok kategorili ürün yelpazesine sahip olan bu firmalar bünyesinde farklı mağazaları ve markaları bulundurmaktadırlar.

Elektronik olarak adlandırılan alt sektördeki firmalar ise hem online hem de mağaza temelinde çok kanallı bir şekilde faaliyet gösteren e-ticaret firmalarıdır. Çoğunlukla elektronik ve teknoloji ürünlerin satıldığı bu çok kanallı perakende firmalarından Firma 5 yurt dışı, mağaza sayısı üstünlüğünü elinde bulunduran Firma 6 ise yurt içi menşeililidir.

3.3 Araştırmanın Yöntemi

Bu tez çalışmasında yer verilen uygulama iki bölümden oluşmaktadır. Birinci bölüm “Firmaların Twitter Kullanımlarına Ait Değerlendirmeler”, ikinci bölüm ise “Firmaların Twitter Etkinliklerinin Ölçümü” olarak adlandırılmıştır. Twitter veri çekimi, araştırmada kullanılan iki ayrı araçla Twitter API desteğı ile gerçekleştirilmiştir. Öncelikle firmaların sosyal medya kullanımlarına dair çıkarımda bulunmak amacı ile 2017 yılına ait firma tweetleri çekilerek analiz edilmiştir. Daha sonra firmaların kampanya dönemlerinde almış oldukları etkileşimleri değerlendirilmek amacıyla belirlenen özel gün doğrultusunda Şubat ayı boyunca kullanıcılar tarafından firmalar ile ilgili üretilmiş içerikler ve firmaların paylaştıkları içerikler elde edilerek değerlendirmeler gerçekleştirilmiştir.

Araştırmada sosyal medya madenciliğı adımları uygulamanın her iki bölümü için de izlenmiştir. Araştırmada izlenen adımlar Tablo 3.3’de özetlenmiştir.

Tablo 3.3: Uygulamada İzlenen Adımlar

Sosyal Medya Madenciliği Uygulaması	Firmaların Kullanımlarına Değerlendirmeler	Twitter Ait	Firmaların Etkinliklerinin Ölçümü	Twitter
Adım 1: Yetkilendirme	Veri çekimi için kullanılacak olan araca (RapidMiner) gerekli yetkilendirme işlemleri yapılmıştır.		Veri çekimi için kullanılacak olan araca (Tweetera) gerekli yetkilendirme işlemleri yapılmıştır.	
Adım 2: Veri Toplama	2017 yılına ait firma içerikleri elde edilmiştir.	Twitter	2018 Şubat ayına ait kullanıcıların firmalar ile ilgili paylaştıkları Twitter içerikleri ve firmaların Twitter paylaşımları elde edilmiştir.	
Adım 3: Veri Ön işleme	Veri setinden ihtiyaç duyulmayan değişkenler çıkarılarak veri seti indirgenmiş ve nihai veri setini oluşturan değişken açıklamalarına yer verilmiştir.		Veri setinden ihtiyaç duyulmayan değişkenler çıkarılarak veri seti indirgenmiş, belirli kıstaslara bağlı kalınarak veri seti temizlenmiş ve nihai veri setini oluşturan değişken açıklamalarına yer verilmiştir.	
Adım 4: Modelleme ve Analiz	• Tanımlayıcı İstatistikler • Ki-Kare Analizi • Uygunluk Analizi • İçerik Analizi ○ Tweet Analizi ○ Rt Analizi ○ Hashtag Analizi		Hoffman ve Fodor (2010) tarafından yapılan çalışmada yer alan Twitter metrikleri aşağıdaki analizlerle belirlenmiştir: • Tanımlayıcı İstatistikler • Duygu Analizi	
Adım 5: Değerlendirme	Modelleme ve Analiz bölümünde elde edilen bulgular değerlendirilmiştir.		Modelleme ve Analiz bölümünde elde edilen bulgular görselleştirilerek firmaların Twitter etkinlikleri kıyaslanmıştır.	

3.4 Sosyal Medya Madenciliği Uygulaması

3.4.1 Firmaların Twitter Kullanımlarına Ait Değerlendirmeler

Bu bölümde örneklem dâhilinde belirlenen firmaların Twitter istatistiklerini değerlendirmek ve Twitter'ı kullanım amaçlarını belirlemek için 2017 yılı boyunca Twitter platformu üzerinden firma paylaşımları ele alınarak bir veri seti oluşturulmuştur. Verinin elde edilmesi ve analizi sürecinde literatürde değinilen sosyal medya madenciliği süreç adımları izlenmiştir.

3.4.1.1 Yetkilendirme

Sosyal medya madenciliği sürecinin ilk aşaması verinin elde edilmesi aşamasıdır. Bu çalışmanın örneklemine oluşturan firmaların 2017 yılı boyunca Twitter kurumsal hesaplarından yapmış oldukları paylaşımlar, Twitter API desteği ile RapidMiner Studio 8.1 aracı kullanılarak elde edilmiştir.

3.4.1.1.1 Aplikasyon Oluşturma

Twitter platformundan API desteği ile veri çekebilmek için öncelikle bir Twitter hesabına sahip olunması gerekmektedir. Hesap oluşturulduktan sonra aplikasyon oluşturmak için izlenmiş adımlar aşağıda verilmiştir.

- <https://apps.twitter.com> adresinden yeni bir aplikasyon oluşturmak için “Create New App” butonu tıklanır.
- Açılan sayfada aplikasyon detayları girilir; isim (name) bölümüne herhangi bir ad (örneğin; “RapidMiner ile Veri Çekimi”), açıklama (description) bölümüne herhangi bir açıklama (örneğin; “Sosyal Medya Madenciliği ve Bir Uygulama”), Web site bölümüne ise kodu yayınlayacağınızı ilettiğiniz web sitesi adresi (bu bölüme herhangi bir site yazılabilir) girildikten sonra aplikasyon oluşturulur.

- Açılan sayfada “Keys and Access Tokens” tıklanır ve “Your Access Token” başlığı altındaki “Create my access tokens” butonuna tıklanarak API ile veri çekimi için gerekli olan aplikasyon anahtarları elde edilir.

Şekil 3.1: Twitter Veri Çekimi- Twitter Aplikasyonu Oluşturma

RapidMiner ile Veri Çekimi

Test OAuth

Details

Settings

Keys and Access Tokens

Permissions

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)	XXXXXXXXXXXXXXXXXX
Consumer Secret (API Secret)	XXXXXXXXXXXXXXXXXXXXXXXXXXXX
Access Level	Read and write (modify app permissions)
Owner	XXXXXXXXXXXXXXXXXX
Owner ID	XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Application Actions

Regenerate Consumer Key and Secret

Change App Permissions

Your Access Token

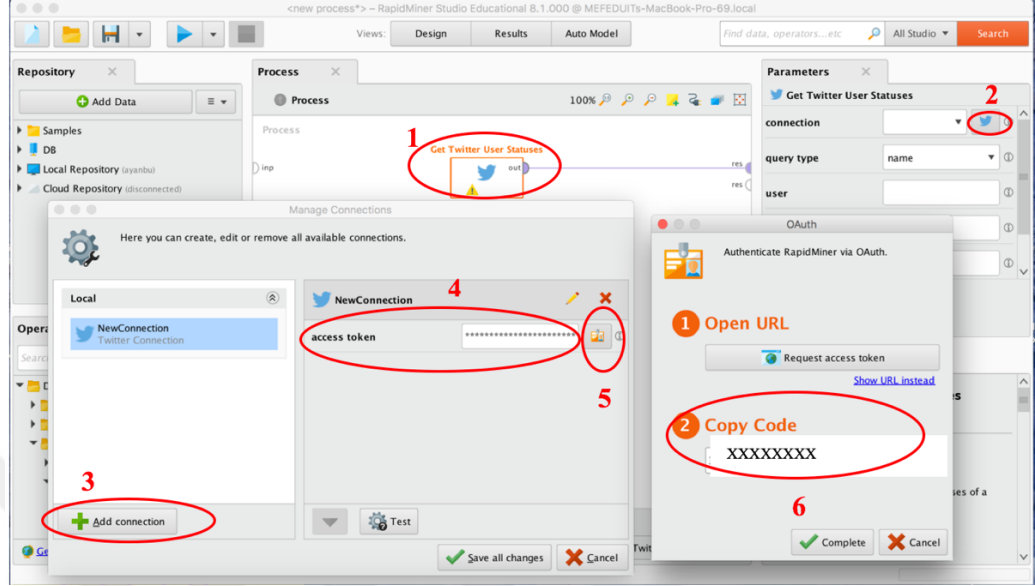
This access token can be used to make API requests on your own account's behalf. Do not share your access token secret with anyone.

Access Token	XXXXXXXXXXXXXXXXXX
Access Token Secret	XXXXXXXXXXXXXXXXXXXXXXXXXXXX
Access Level	Read and write
Owner	XXXXXXXXXXXXXXXXXX
Owner ID	XXXXXXXXXXXXXXXXXXXXXXXXXXXX

3.4.1.1.2 Yetkilendirme İşlemi

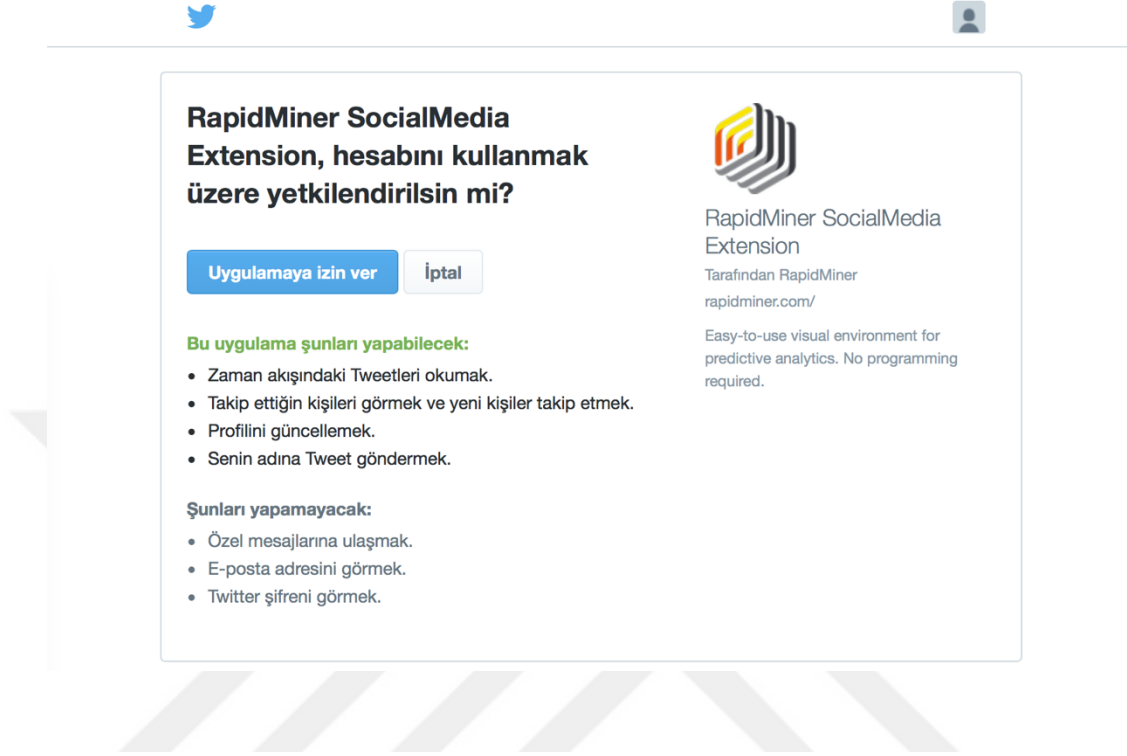
Firmalara ait Twitter verilerinin elde edilmesi için bu çalışmada RapidMiner Studio 8.1 aracında “Get Twitter User Statuses” operatörü kullanılmıştır. Yetkilendirme işlemi için izlenen adımlar Şekil 3.2 üzerinde numaralandırılmış bir şekilde açıklanmıştır.

Şekil 3.2: Twitter Veri Çekimi- RapidMiner Yetkilendirme İşlemi



1. "Get Twitter User Statuses" operatörü ekrana sürüklenmiştir.
2. Parametrelerin yer aldığı bölümde bağlantı sekmesinde (connection) Twitter amblemine tıklanarak yetkilendirme sayfası açılmıştır.
3. Açılan sayfada yeni bir bağlantı açma butonu (add connection) tıklanarak NewConnection adlı bağlantı oluşturulmuştur.
4. Aplikasyon oluşturma adımında elde edilen "access token" değeri ilgili bölüme kopyalanmıştır.
5. OAuth butonu tıklanarak Şekil 3.3 'de ekran görüntüsü verilen yetkilendirme işlemi için uygulamaya izin verilmiştir.
6. İzin ile birlikte elde edilen kod ilgili bölüme kopyalanmış ve süreç tamamlanmıştır.

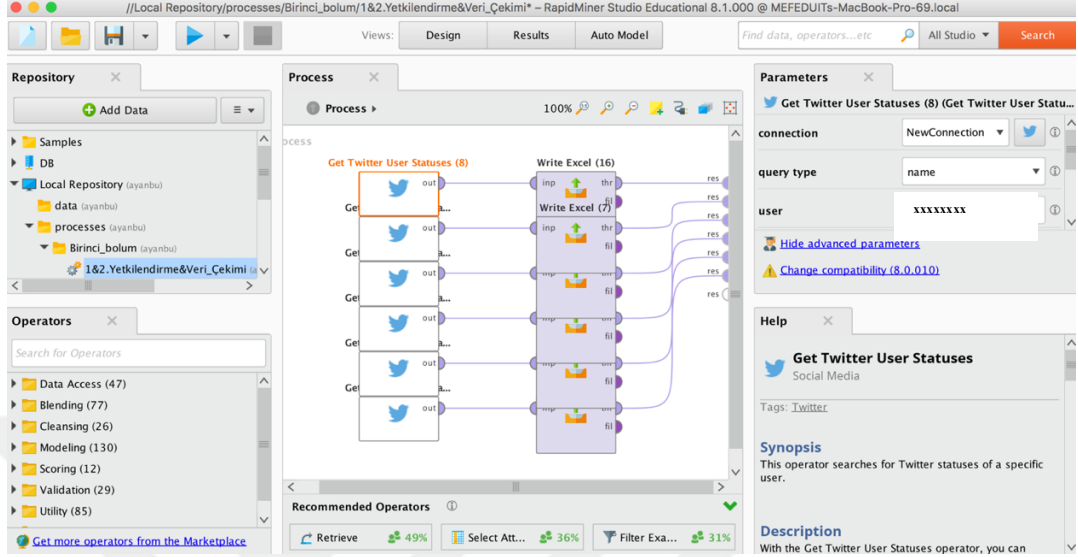
Şekil 3.3: Twitter Veri Çekimi- RapidMiner Yetkilendirme İzin Sayfası



3.4.1.2 Veri Toplama

RapidMiner Studio 8.1 aracında firmalar için ayrı ayrı “Get Twitter User Statuses” operatörü ve çekilen verilerin Excel’e kaydedilmesi için “Write Excel” operatörü süreç ekranına (Process) taşınmış ve “Get Twitter User Statuses” operatöründen “Write Excel” operatörüne ve sonrasında da sonuç (res) bölümüne doğru her firma için bağlantılar kurulmuştur. Öncelikle “Get Twitter User Statuses” operatörü parametrelerinde yer alan sekmeler doldurulmuştur. Sırası ile Şekil 3.4’de görüldüğü üzere bağlantı bölümünde (connection) yetkilendirme işleminde oluşturulan “NewConnection” adlı bağlantı ve sorgu türü (query type) ad (name) olarak seçilmiştir. Kullanıcı adının yazıldığı bölüme (user) firmaların kurumsal Twitter hesaplarındaki kullanıcı adları kopyalanmıştır. Daha sonra dosyaların kaydedilmesi istenen yer “Write Excel” operatörü ile belirlenerek süreç çalıştırılmıştır.

Şekil 3.4: RapidMiner Firmalara Ait Veri Çekim Süreci



Twitter API desteği ile bir kullanıcıya ait en yeni gönderilerden eskilere doğru olmak üzere, ters kronolojik sıra şeklinde, 3200 adet tweet çekilebilmektedir. Bu şekilde firmaların paylaşmış oldukları son 3200 adet tweet elde edilmiştir.

Şekil 3.5: Firma 1'e Ait RapidMiner Twitter Sorgu Sonuç Ekranı Örneği

Row...	Id	Created...	From-User	From-...	To-...	To-...	Lang...	Source	Text	Geo...	Geo...	Retweet
68	960...	Feb 6, 2...	XXXX	16029...	?	-1	tr	<a href...	Procsin by XXX ı Bl...	?	?	0
69	960...	Feb 6, 2...	XXXX	16029...	Esra...	4807...	tr	<a href...	@Esraacakr Merhaba, size ...	?	?	0
70	960...	Feb 6, 2...	XXXX	16029...	ozle...	2683...	tr	<a href...	@ozlemsenyigit85 Merhaba...	?	?	0
71	960...	Feb 6, 2...	XXXX	16029...	?	-1	tr	<a href...	Göz makyajların için harika...	?	?	1
72	960...	Feb 5, 2...	XXXX	16029...	?	-1	tr	<a href...	https://t.co/cVCsRtxero'ye ...	?	?	0
73	960...	Feb 5, 2...	XXXX	16029...	33Mr...	3243...	tr	<a href...	@33Mr Merhaba, Makeup...	?	?	0
74	960...	Feb 5, 2...	XXXX	16029...	?	-1	tr	<a href...	Bırakın, dudaklarınız konu...	?	?	1
75	960...	Feb 5, 2...	XXXX	16029...	?	-1	tr	<a href...	XXX izellik ve Kişisel ...	?	?	0
76	960...	Feb 4, 2...	XXXX	16029...	sym...	4078...	tr	<a href...	@symnr10 Merhabalar, siz...	?	?	0
77	960...	Feb 4, 2...	XXXX	16029...	ayse...	7528...	tr	<a href...	@aysegulkrca Merhabalar, ...	?	?	0
78	960...	Feb 4, 2...	XXXX	16029...	ssse...	3734...	tr	<a href...	@ssseennnaa XXX .	?	?	0
79	960...	Feb 4, 2...	XXXX	16029...	ozle...	2683...	tr	<a href...	@ozlemsenyigit85 Konu ha...	?	?	0
80	959...	Feb 3, 2...	XXXX	16029...	hera...	4173...	tr	<a href...	@heraofzeus Merhabalar, s...	?	?	0
81	959...	Feb 2, 2...	XXXX	16029...	?	-1	tr	<a href...	XXX ığazalarında ve ...	?	?	2
82	959...	Feb 2, 2...	XXXX	16029...	?	-1	tr	<a href...	Ojeyi taşıyarak sürdürünz di...	?	?	4
83	959...	Feb 2, 2...	XXXX	16029...	talha...	8992...	tr	<a href...	@talhaturker9 Size yardım...	?	?	1
84	959...	Feb 1, 2...	XXXX	16029...	?	-1	tr	<a href...	Eski şampuanlara #tbt vakt...	?	?	0
85	959...	Feb 1, 2...	XXXX	16029...	?	-1	tr	<a href...	XXX izellik ve Kişisel ...	?	?	5
86	959...	Feb 1, 2...	XXXX	16029...	?	-1	tr	<a href...	RT @ojemrjumrimelm: Re...	?	?	1
87	958...	Feb 1, 2...	XXXX	16029...	?	-1	tr	<a href...	Yepyeni ürünler ve sürpriz f...	?	?	0

Şekil 3.5’de görüldüğü üzere, elde edilen veri setindeki değişkenler ve açıklamaları aşağıdaki gibidir:

- **Id:** Her tweete özgü verilen numerik bir değerdir.
- **Created-At:** Tweetin atıldığı tarih ve saat bilgisini içermektedir.
- **From-User:** Hangi kullanıcıya ait verilerin çekildiğini gösteren sütundur.
- **From-User-Id:** Verileri çekilen kullanıcının bir id numarası ile temsil edildiği sütundur.
- **To-User:** Tweet metninde bahsedilen kullanıcıya ait id numarasını veren sütundur.
- **To-User-Id:** Tweet metni başında “@” sembolü kullanılarak bir kullanıcıdan bahsedildiyse (mention) bahsedilen kullanıcının isim bilgisini veren sütundur.
- **Language:** Atılan tweet metninin hangi dilde yazıldığının bilgisini içermektedir.
- **Source:** “Android”, “iPhone”, “iPad”, “Web Client” ya da otomatik bir şekilde tweet atmak için kullanılan kaynak (özel hizmet veren sosyal medya danışmanlık şirketleri gibi) bilgilerinin yer aldığı, tweetin hangi araç (kaynak) kullanılarak atıldığının bilgisini içermektedir.
- **Text:** Tweet metnini içermektedir.
- **Geo-Location-Latitude:** Tweetin atıldığı konumun enlem bilgisini içermektedir.
- **Geo-Location-Longitude:** Tweetin atıldığı konumun boylam bilgisini içermektedir.
- **Retweet-Count:** Atılan tweetin almış olduğu retweet sayısını göstermektedir.

3.4.1.3 Veri Ön İşleme

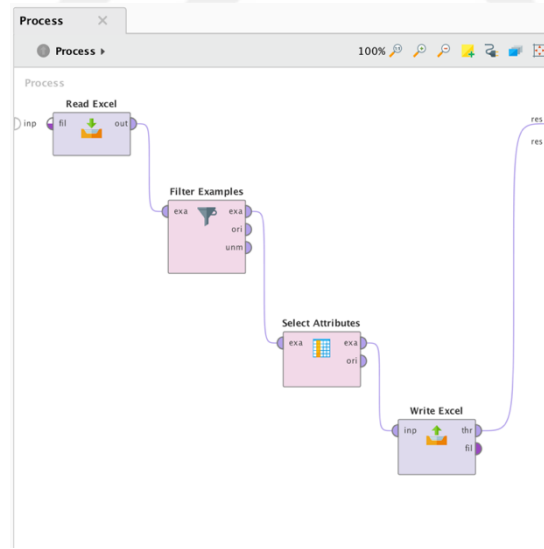
Bu bölümde veri setinde yer alan analize dâhil edilmeyen değişkenler çıkarılarak veri seti indirgenmiş ve nihai veri setinde yer alan değişkenlerin açıklamalarına yer verilmiştir.

3.4.1.3.1 Veri İndirgeme

RapidMiner aracı kullanılarak elde edilen veri setinde ihtiyaç duyulmayan değişkenler ve firmaların 2017 yılı dışındaki yıllara ait veriler temizlenmiştir. Veri indirgeme aşamasında izlenen adımlar aşağıda verilmiştir.

- Veri setindeki 1 Ocak 2017 ile 31 Aralık 2017 tarihleri dışındaki veriler “Filter Examples” operatörü ile veri setinden çıkarılmıştır.
- Veri setindeki “Id”, “From-User-Id”, “To-User”, “To-User-Id”, “Language”, “Source”, “Geo-Location-Latitude”, “Geo-Location-Longitude” sütunları analizde kullanılmayacak değişkenler olduğu için “Select Attributes” operatörü ile veri setinden çıkarılmış ve Write Excel operatörü kullanılarak yeni veri seti olarak kaydedilmiştir.

Şekil 3.6: RapidMiner Veri İndirgeme Süreç Modeli



3.4.1.3.2 Veri Seti

Nihai veri setinde yer alan değişkenlerin isimleri ve açıklamaları Tablo 3.4’de verilmiştir.

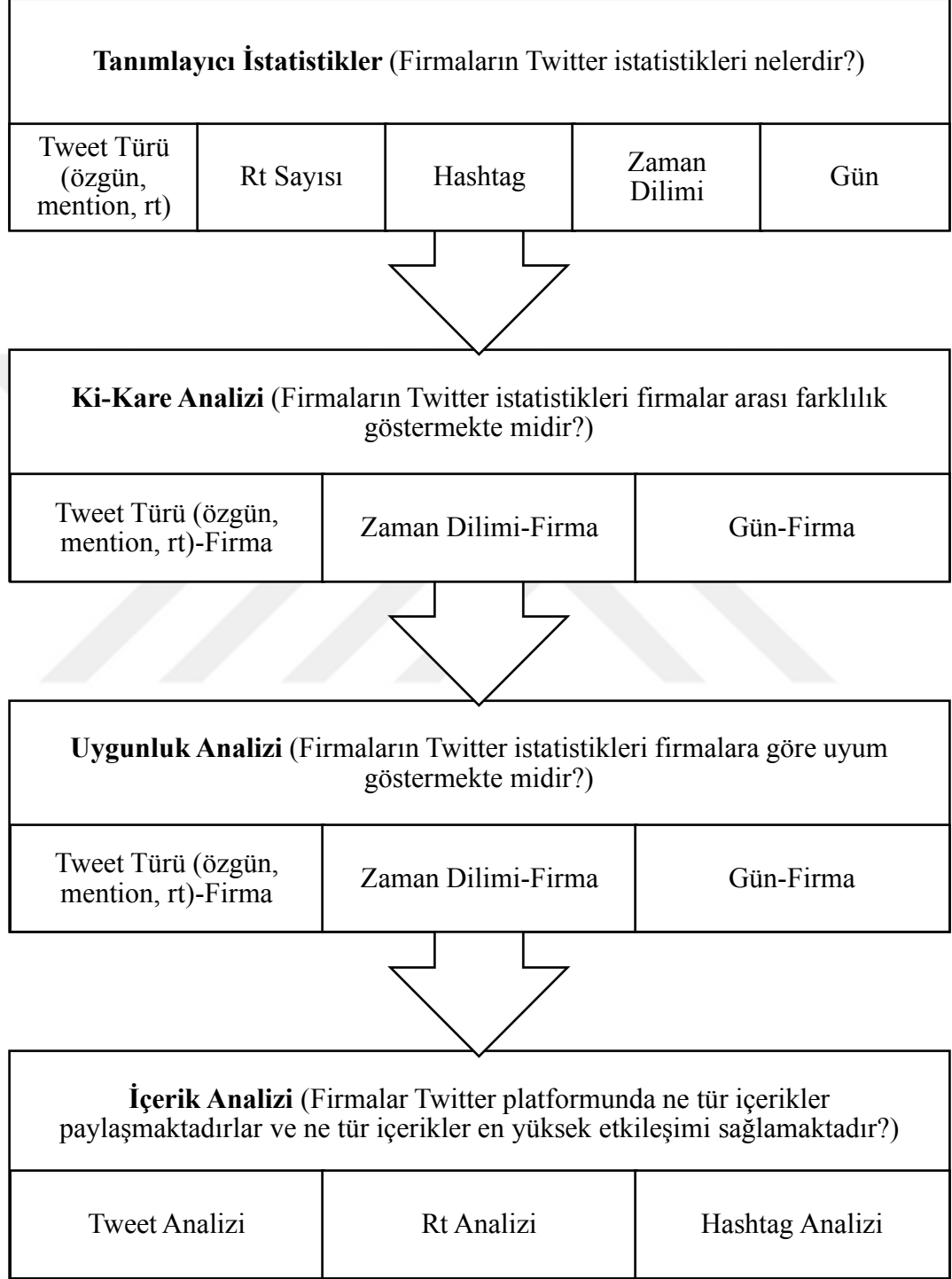
Tablo 3.4: Uygulamanın Birinci Bölümüne Ait Değişken Tablosu

Değişken İsmi	Açıklaması
Tarih	Firma paylaşımlarının hangi tarihte yapıldığının bilgisinin gün/ay/yıl olarak verildiği değişkendir. Veri toplama aşamasında elde edilen ilk veri setinde yer alan “Created-At” sütunundan çekilmiştir.
Saat	Firma paylaşımlarının hangi saatte yapıldığının bilgisinin saat/dakika/saniye olarak verildiği değişkendir. Veri toplama aşamasında elde edilen ilk veri setinde yer alan “Created-At” sütunundan çekilmiştir.
Zaman Dilimi	Firma paylaşımlarının yapıldığı saat bilgisi göz önünde bulundurularak bir zaman dilimi sınıflandırılması yapılmıştır. Belirlenen üç zaman dilimi kapsamında (≥ 05 - < 12 =öğleden önce, ≥ 12 - < 17 = öğleden sonra, ≥ 17 - < 05 =akşam) zaman dilimi değişkeni elde edilmiştir.
Gün	Firma paylaşımlarının atıldığı gün bilgisinin verildiği değişkendir. Tarih değişkeninden elde edilmiştir.
Firma	Paylaşımların hangi firmaya ait olduğunu gösteren değişkendir.
Tweet	Tweet olarak adlandırılan firma paylaşımlarını (Twitter mesajlarını) içeren değişkendir.
Rt Sayısı	Her bir firma paylaşımının almış olduğu retweet miktarını gösteren değişkendir.
Özgün Tweet	Firmaların paylaşmış oldukları özgün içerikler, “rt” ve “mention” içermeyen tweetler olarak belirlenmiştir. Bu değişken eğer tweet özgün ise 1, değil ise 0 olarak hesaplanmıştır.
Rt Tweet	Firmaların paylaşmış oldukları “retweet” içeren tweetlerdir. Bu değişken eğer tweet retweet edilmiş bir paylaşımından oluşuyor ise 1, değil ise 0 olarak hesaplanmıştır.
Mention Tweet	Firmaların paylaşmış oldukları “mention” içeren tweetlerdir. Bu değişken eğer tweet mention içeriyor ise 1, değil ise 0 olarak hesaplanmıştır.
Hashtag	Firma paylaşımlarının içerisinde yer alan “#hashtag” ifadeleri ayrı bir sütunda yer verilerek bu değişken elde edilmiştir.

3.4.1.4 Modelleme ve Analiz

Uygulamanın bu bölümünde izlenen adımların yer aldığı model, Şekil 3.7’de ayrıntılı olarak verilmiştir.

Şekil 3.7: Uygulamanın Birinci Bölümüne Ait Analiz Adımları



Literatür taraması bölümünde değinilen Chae (2015) tarafından önerilen Twitter Analitiği modelinde yer alan Tanımlayıcı İstatistikler Firmaların sosyal medya verilerine uygulanmıştır. Modelde yer verilen tweet türleri (özgün, mention ve rt tweet), rt sayısı, hashtag, zaman dilimi ve gün değişkenlerinin 2017 yılı frekans hesaplamaları gerçekleştirilmiş ve görselleştirilmiştir.

Tanımlayıcı İstatistikler bölümde yer verilen veriler değerlendirildiğinde rakip firma ve sektör temelinde anlamlı farklılıkların söz konusu olup olmadığı Ki-Kare tekniği uygulanarak analiz edilmiştir. Daha sonra, anlamlı farklılık gösteren değişkenlerle firmalar/sektörler arası uyum, Uygunluk Analizi uygulanarak görselleştirilmiştir.

Firmaların hedef kitleleri ile hangi içerikleri paylaşarak Twitter platformunda varlıklarını sürdürdüklerini değerlendirmek amacı ile Chae (2015)'nin önerdiği Twitter Analitiği modelinde yer alan İçerik Analizi çalışmaya uyarlanarak gerçekleştirilmiştir. İçerik Analizi bölümünde firmaların paylaştıkları tweetler metin madenciliği ön işleme aşamalarından geçirilerek kelime frekansları elde edilmiş, daha sonra kelimeler gruplandırılarak kategori başlıkları belirlenmiştir. Her tweetin ait olduğu içerik kategorisi etiketlenmiş ve firma temelinde frekans değerleri verilmiştir. Rt analizi bölümünde ise en çok rt alan tweet kategorilerini saptamak adına Pareto Analizi uygulanmıştır. Hashtag analizi ile ise firmaların hashtag paylaşımları üzerine ayrıntılı değerlendirmelere yer verilmiştir.

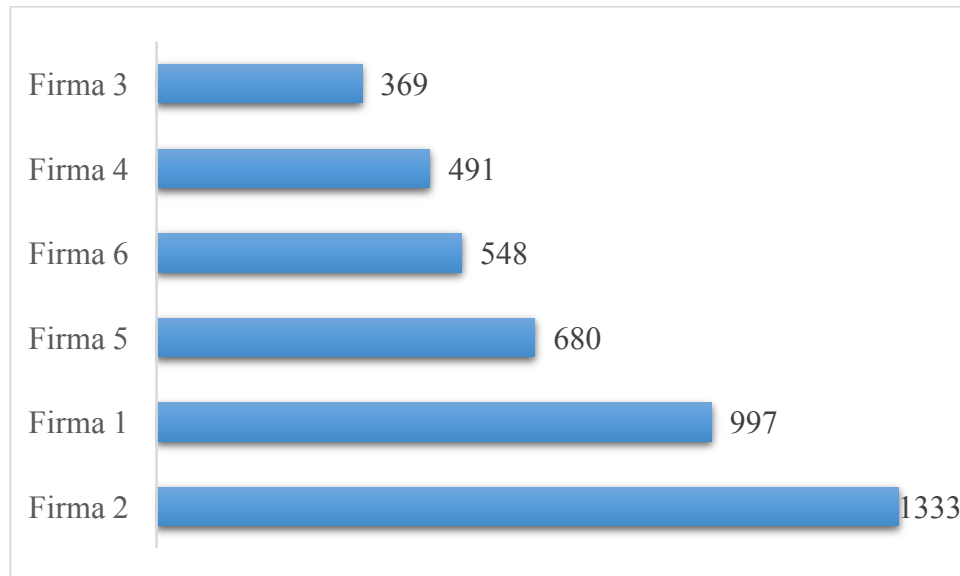
Sosyal medya madenciliği sürecinin son aşaması olan Değerlendirme başlığı altında, çalışmada gerçekleştirilen analizlerin bulgularına yer verilerek genel bir değerlendirme yapılmıştır.

3.4.1.4.1 Tanımlayıcı İstatistikler

Bu bölümde firmaların Twitter paylaşımları ele alınarak yıl temelinde hesaplanan istatistiklere yer verilmiştir. Toplam 4.418 satırdan oluşan veri seti ile ilgili istatistikler bu bölümde ele alınmıştır. Uygulamada yer alan istatistiklerin bir bölümü Chae (2015) tarafından önerilen modelden elde edilmiş, bir bölümü ise bu çalışma için geliştirilmiştir.

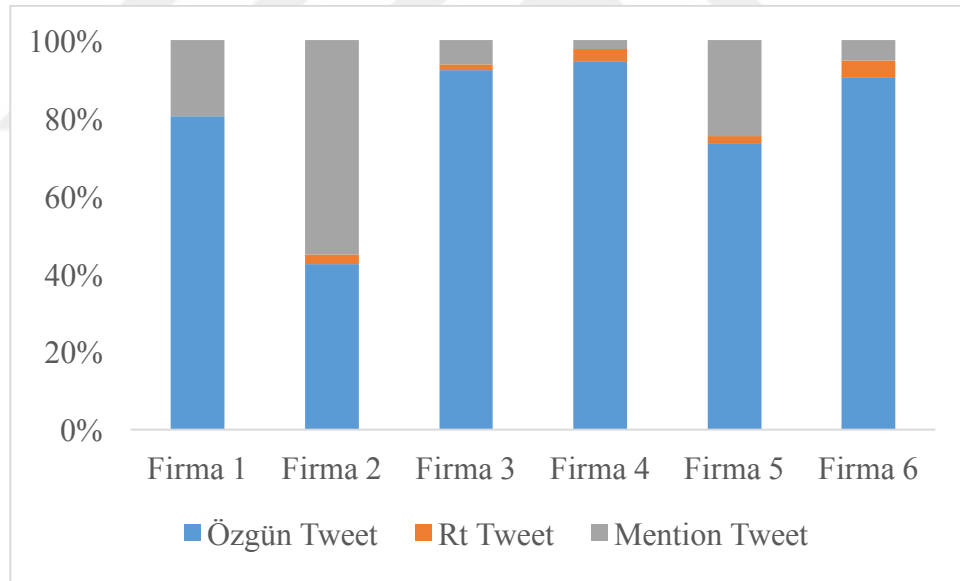
Firmaların 2017 yılı tweet paylaşımlarının frekansları ele alındığında, en çok paylaşım yapan firmaların sırasıyla kozmetik sektöründe faaliyet gösteren Firma 2 ve Firma 1 olduğu görülmektedir. Sıralamada kozmetik sektörü firmalarını elektronik sektöründe faaliyet gösteren Firma 5 ve Firma 6 takip etmiştir. En az paylaşım frekans sayılarına sahip firmalar ise pazaryeri sektörü firmaları olan Firma 3 ve Firma 4'dür. Bu değerlendirmeye göre, Twitter'ı en aktif biçimde kullanan firmaların kozmetik sektöründe ve en az kullanan firmaların ise pazaryeri sektöründe olduğu saptanmıştır.

Şekil 3.8: Firmalar Tarafından Paylaşılan Tweetlerin Frekans Dağılımları



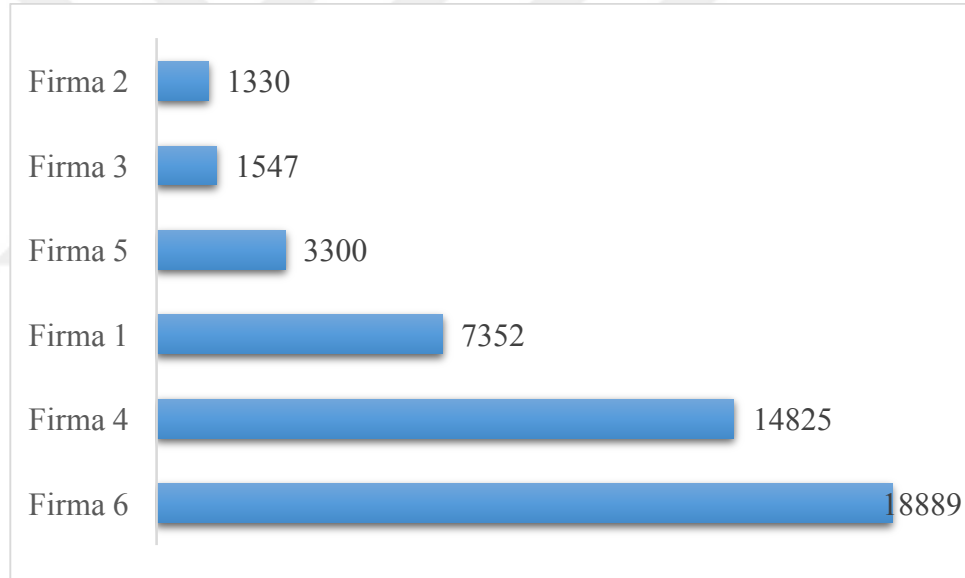
Firmalar tweet paylaşımı yaparken özgün içerikler üretebilmekte, bir kullanıcıdan bahsedebilmekte ya da daha önce başkası tarafından paylaşılan bir içeriği tekrar paylaşabilmektedir. Chae (2015)'nin çalışmasından uyarlanarak tweetlerin türleri özgün, mention ve rt tweet olarak sınıflandırılmış ve frekans değerleri hesaplanmıştır. Şekil 3.9'da görüldüğü üzere kozmetik sektöründe faaliyet gösteren Firma 2 dışındaki tüm firmalar tarafından en yüksek oranla özgün tweet paylaşılmıştır. Rt tweet türü ise hiçbir firma tarafından yüksek oranla paylaşılmamıştır. Mention tweet türü rt tweet türüne göre firmalar tarafından daha çok tercih edilmiştir. Tweet türlerinin firmalara göre değişkenlik gösterdiği gözle saptanmasına karşın, anlamlı farklılık ve uyumlar Ki-Kare ve Uygunluk Analizi uygulanarak detaylı bir şekilde incelenmiştir.

Şekil 3.9: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Frekans Dağılımları



Bir diğerk önemli istatistik olan firmaların Twitter paylaşımlarının almış oldukları rt frekansları incelendiğinde, kozmetik sektörü firması Firma 2’nin 2017 yılı boyunca diğerk firmalara kıyasla daha çok sayıda tweet paylaşmasına rağmen, tezat bir şekilde, toplamda aldığı rt sayısının Şekil 3.10’da verilen sıralamada en düşük sırada yer aldığı gözlemlenmiştir. Elektronik sektöründe faaliyet gösteren Firma 6’nın ise 2017 yılı boyunca paylaşmış oldukları tweetleri en fazla etkileşimi (rt) almıştır. Pazaryeri sektöründe faaliyet gösteren, rt sıralamasında ikinci sırada yer alan Firma 4’ün, rakibi Firma 3’e göre oldukça fazla etkileşim aldığı söylenebilir.

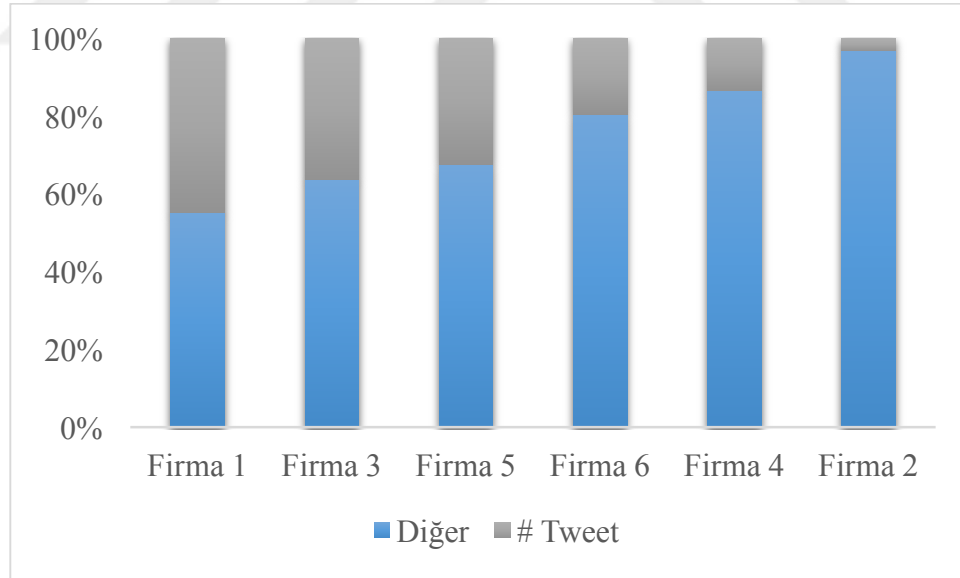
Şekil 3.10: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekans Dağılımları



Rt frekans sıralaması ile tweet frekans sıralamasının benzer olmadığı gözlemlenmiştir. Hangi tweet içeriklerinin daha çok etkileşim aldığı İçerik Analizi bölümünde ayrıntılı olarak ele alınmıştır.

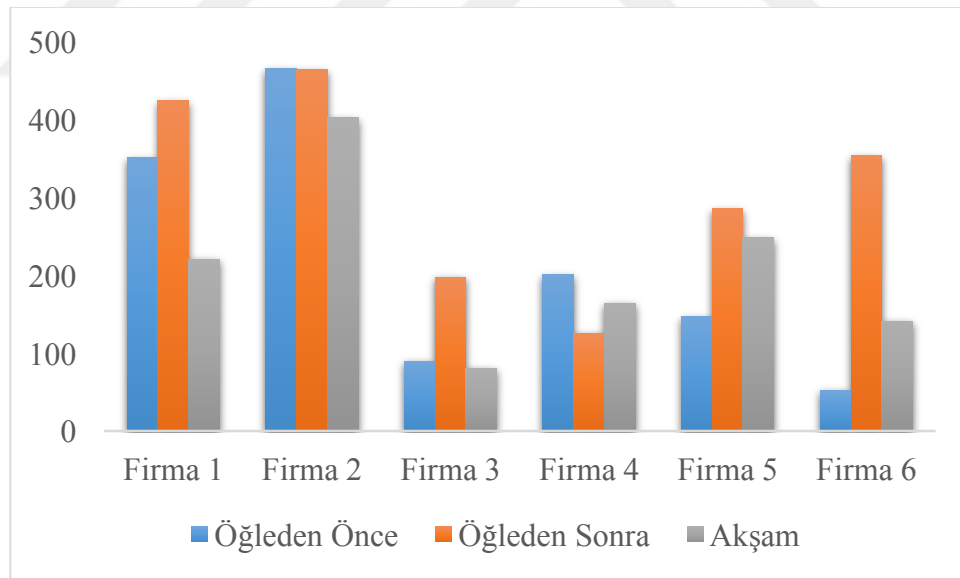
Firmalar Twitter’da paylaşım yaparken çeşitli hashtagler kullanmaktadırlar. Firmaların hashtag kullanımları ele alınarak hashtag kullanım frekansları hesaplanmıştır. Şekil 3.11’de firmaların Twitter platformunda paylaştıkları tweetlerden kaç tanesinin hashtag içerdiği (#Tweet) görselleştirilmiştir. 2017 yılı boyunca firma tweetlerinde hashtag kullanımı ele alındığında kozmetik sektöründe faaliyet gösteren Firma 1’in en yüksek yüzdeyle (%45) tweetlerinde hashtag kullandığı saptanmıştır. Firma 1’in rakibi Firma 2 ise hashtag içeren tweet sıralamasında sonuncu sırada yer almıştır. Pazaryeri sektörü firmaları ele alındığında ise Firma 3’ün (%36) rakibi Firma 4’e (%13) göre daha yüksek bir yüzde ile hashtag içeren tweetlere yer verdiği görülmüştür. Elektronik sektöründe faaliyet gösteren Firma 5 %32’lik, Firma 6 ise %20’lik bir oranla hashtag içeren tweet paylaşımı yapmışlardır. Detaylı inceleme İçerik Analizi bölümünde ele alınmıştır.

Şekil 3.11: Firmalar Tarafından Paylaşılan Hashtag İçeren Tweetlerin Frekans Dağılımları



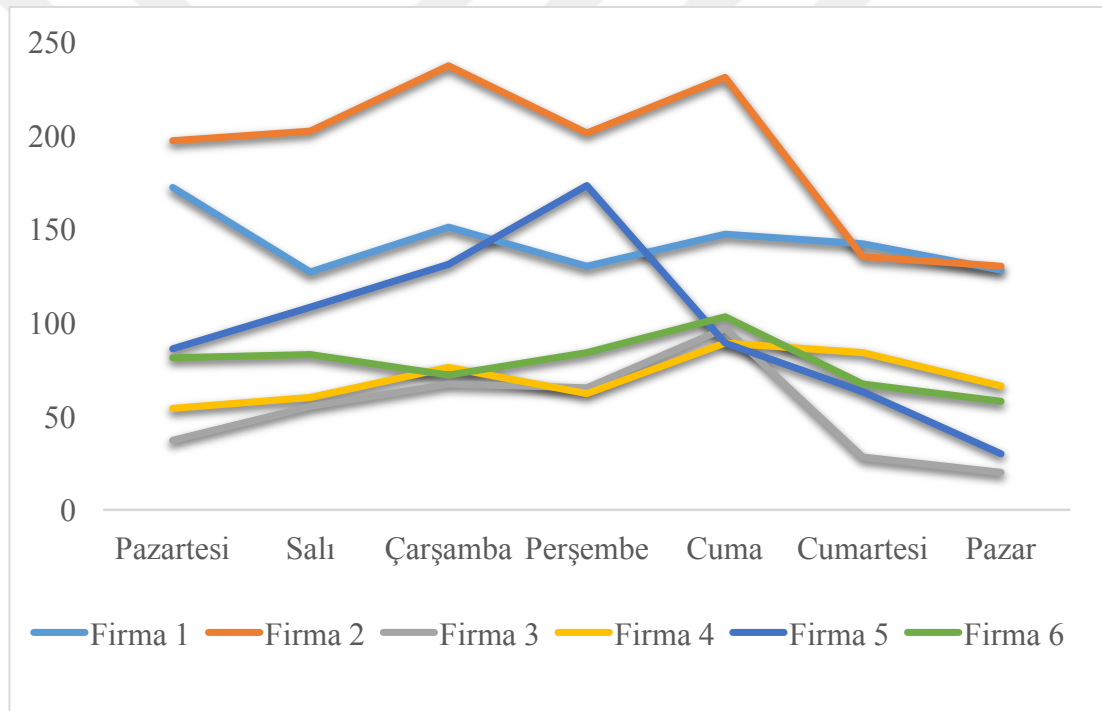
Firmaların tweet paylaşımları gün içerisinde farklı zaman dilimlerinde yapılmaktadır. Firma temelinde yapılan paylaşımların gün içerisindeki dağılımını görmek ve paylaşım açısından günün en yoğun olduğu zaman dilimini belirlemek için, firmaların paylaşımlarının zaman dilimleri Şekil 3.12’deki gibi görselleştirilmiştir. Kozmetik sektöründe faaliyet gösteren Firma 1’in ve Firma 2’nin öğleden önce ve öğleden sonra zaman dilimlerinde daha çok paylaşım yaptığı görülmektedir. Pazaryeri sektöründe faaliyet gösteren firmaların en fazla paylaşım yaptıkları zaman dilimi incelendiğinde Firma 3 öğleden sonra, Firma 4 ise öğleden önce saatlerinde daha çok paylaşım yapmıştır. Elektronik sektöründe faaliyet gösteren firmalar ise daha çok öğleden sonra zaman diliminde Twitter’da içerik paylaşmışlardır. Ayrıntılı analize Ki-Kare ve Uygunluk Analizi bölümünde yer verilmiştir.

Şekil 3.12: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimi Frekans Dağılımları



Firmalar Twitter paylaşımlarını haftanın farklı günlerinde yapmaktadırlar. Yapılan paylaşımların haftanın hangi günü en yoğun olduğunu belirlemek amacı ile veri setinden gün değişkenine ait frekanslar hesaplanmıştır. Firmaların günlere göre tweet paylaşım frekansları Şekil 3.13'deki gibi görselleştirilmiştir. Twitter paylaşımlarının hafta içi günlerinde daha yoğunken hafta sonunda azaldığı görülmektedir. Ayrıntılı analize Ki-Kare ve Uygunluk Analizi bölümünde yer verilmiştir.

Şekil 3.13: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Frekans Dağılımları



3.4.1.4.2 Ki-Kare Analizi

Uygulamanın bu bölümünde, ilk bölümde hesaplanan istatistikler doğrultusunda belirlenen değişkenler arasında anlamlı farkların/ilişkilerin söz konusu olup olmadığı analiz edilmiştir. IBM SPSS Statistics (24.0) programı kullanılarak analizler gerçekleştirilmiştir.

Ki-Kare analizi gözlenen ve beklenen frekanslar arasındaki farkın anlamlı olup olmadığına bakmaktadır. Bu analizin iki ya da daha fazla grup arasında fark olup olmadığının testinde, iki değişken arasında bağ olup olmadığının testinde, gruplar arası homojenlik testinde, uyum iyiliği testinde gibi farklı kullanım alanları bulunmaktadır. (Güngör ve Bulut, 2008:84-85). Ki-Kare analizinin uygulanabilmesi için beklenen frekansların en az %80'inin 5 ve üzeri olması gerekmektedir (McHugh, 2013:144).

Firmaların özgün, mention ve rt tweet paylaşımlarında firmalar arası anlamlı bir farklılığın olup olmadığı incelenmek üzere Ki-Kare testi uygulanmıştır. Tablo 3.5'de görüldüğü üzere Pearson Ki-Kare test istatistiği değeri anlamlı olduğundan ($\chi^2=993,148$; $p=0,00<0,05$), firmaların özgün, mention ve rt tweet paylaşımlarının firmalara göre farklılaştığı anlaşılmaktadır.

Tablo 3.5: Firmalara Göre Firma Tweet Türleri Ki-Kare Analizi Sonuçları

		Tweet Türü				
		Özgün Tweet	Mention Tweet	Rt Tweet	Topla m	
Firma	Firma 1	Gözlenen Frekans	803	194	0	997
		Beklenen Frekans	715,6	261,1	20,3	997,0
	Firma 2	Gözlenen Frekans	567	734	32	1333
		Beklenen Frekans	956,8	349,1	27,2	1333,0
	Firma 3	Gözlenen Frekans	341	23	5	369
		Beklenen Frekans	264,8	96,6	7,5	369,0
	Firma 4	Gözlenen Frekans	464	11	16	491
		Beklenen Frekans	352,4	128,6	10,0	491,0
	Firma 5	Gözlenen Frekans	500	167	13	680
		Beklenen Frekans	488,1	178,1	13,9	680,0
	Firma 6	Gözlenen Frekans	496	28	24	548
		Beklenen Frekans	393,3	143,5	11,2	548,0
Toplam		Gözlenen Frekans	3171	1157	90	4418
	Beklenen Frekans	3171,0	1157,0	90,0	4418,0	

$\chi^2=993,148$; $p=0,000$ Contingency coefficient=0,428; $p=0,000$

Firmaların Twitter paylaşım günleri ele alınarak firmalar arası tweet paylaşım günlerine göre anlamlı bir farklılığın olup olmadığı analiz edilmiştir. Tablo 3.6 incelendiğinde, Pearson Ki-Kare test istatistiği anlamlı çıkarak ($\chi^2=181,736$; $p=0,00<0,05$), firmaların tweet paylaşım günlerinin firma temelinde farklılaştığı saptanmıştır.

Tablo 3.6: Firmalara Göre Firma Tweet Paylaşım Günleri Ki-Kare Analizi Sonuçları

		Gün								
		Pazartesi	Salı	Çarşamba	Perşembe	Cuma	Cumartesi	Pazar	Toplam	
Firma	Firma 1	Gözlenen Değer	172	127	151	130	147	142	128	997
		Beklenen Değer	141,5	143,3	165,6	161,4	170,6	117,1	97,5	997,0
	Firma 2	Gözlenen Değer	197	202	237	201	231	135	130	1333
		Beklenen Değer	189,2	191,6	221,5	215,7	228,1	156,6	130,3	1333,0
	Firma 3	Gözlenen Değer	37	55	67	65	97	28	20	369
		Beklenen Değer	52,4	53,0	61,3	59,7	63,1	43,3	36,1	369,0
	Firma 4	Gözlenen Değer	54	60	76	62	89	84	66	491
		Beklenen Değer	69,7	70,6	81,6	79,5	84,0	57,7	48,0	491,0
	Firma 5	Gözlenen Değer	86	108	131	173	89	63	30	680
		Beklenen Değer	96,5	97,7	113,0	110,0	116,4	79,9	66,5	680,0
	Firma 6	Gözlenen Değer	81	83	72	84	103	67	58	548
		Beklenen Değer	77,8	78,8	91,0	88,7	93,8	64,4	53,6	548,0
Toplam	Gözlenen Değer	627	635	734	715	756	519	432	4418	
	Beklenen Değer	627,0	635,0	734,0	715,0	756,0	519,0	432,0	4418,0	

$\chi^2=181,736$; $p=0,000$ Contingency coefficient=0,199; $p=0,000$

Firma tweetlerinin paylaşıldığı zaman dilimleri ilk bölümde görselleştirilmiş ve farklılıklar olduğu gözlemlenmiştir. Bu farklılıkların firmalar arası anlamlı olup olmadığı analiz edilmiş ve Pearson Ki-Kare test istatistiği anlamlı çıkmıştır ($\chi^2=496,443$; $p=0,00<0,05$). Buna bulguya göre firmaların tweet paylaşım zaman dilimlerinin firma temelinde farklılaştığı söylenebilir.

Tablo 3.7: Firmalara Göre Firma Tweet Zaman Dilimleri Ki-Kare Analizi Sonuçları

		Zaman Dilimi			Toplam	
		Öğleden Önce	Öğleden Sonra	Akşam		
Firma	Firma 1	Gözlenen Frekans	351	425	221	997
		Beklenen Frekans	295,2	417,9	283,9	997
	Firma 2	Gözlenen Frekans	466	464	403	1333
		Beklenen Frekans	394,7	558,8	379,6	1333
	Firma 3	Gözlenen Frekans	90	198	81	369
		Beklenen Frekans	109,2	154,7	105,1	369
	Firma 4	Gözlenen Frekans	201	126	164	491
		Beklenen Frekans	145,4	205,8	139,8	491
	Firma 5	Gözlenen Frekans	147	285	248	680
		Beklenen Frekans	201,3	285,1	193,6	680
	Firma 6	Gözlenen Frekans	53	354	141	548
		Beklenen Frekans	162,2	229,7	156	548
Toplam	Gözlenen Frekans	1308	1852	1258	4418	
	Beklenen Frekans	1308	1852	1258	4418	

$\chi^2=304,676$; $p=0,000$ Contingency coefficient=0,254; $p=0,000$

3.4.1.4.3 Uygunluk Analizi

Ki-Kare Analizi bulguları göz önünde bulundurularak bu bölümde, anlamlı farklılıkların bulunduğu değişkenler arasında Uygunluk Analizi uygulanmıştır. Uygunluk Analizi açıklandıktan sonra IBM SPSS Statistics (24.0) programı kullanılarak firmaların 2017 yılı Twitter paylaşımlarından oluşan veri setinde yer alan değişkenlere Uygunluk Analizi uygulanmıştır. SPSS bulgularına yer verilerek tablo ve şekil yorumları yapılmıştır.

Uygunluk Analizi (Correspondence Analysis), bir kontenjans tablosunda yer alan satır ve sütun değişkenleri arasındaki birlikteliğin az boyutlu bir grafik üzerinde gösterimi üzerine kurulu bir tekniktir. Uygunluk Analizinin sonucunda oluşan grafik gösteriminde benzer dağılıma sahip kategoriler birbirlerine yakın olarak konumlanırken, dağılımları birbirlerinden farklı olan kategoriler birbirlerinden uzak konumlanmaktadır. Analizi uygularken izlenen adımlar aşağıdaki gibi sıralanabilir (Çılan, 2013:121-132):

- Kategori profillerinin hesaplanması
- Marjinal oranların (masses) hesaplanması
- Noktalar arasındaki uzaklıkların hesaplanması
- En uygun grafik boyutlarına karar verilmesi

Uygunluk analizinde ilk olarak kontenjans tablosunda yer alan frekanslar oranlara dönüştürülerek kategori profilleri hesaplaması yapılır. Satır ve sütun profilleri hesaplamaları Tablo 3.8’de verilmiştir. Bu tablolarda yer alan nokta (.) ifadesi satır ve sütun toplamını ifade etmek için kullanılmaktadır. Örneğin n.1 ifadesi 1. sütundaki terimlerin toplamını, n2. ifadesi ise 2. satırdaki tüm terimlerin toplamını ifade etmektedir (Kılıç, 2016:5).

Tablo 3.8: Uygunluk Analizi Satır ve Sütun Profillerinin Hesaplanması

Satır Profillerinin Hesaplanması	1	2	...	j	...	c	Satır Profil Toplamı
1	$n_{11}/n_{1.}$	$n_{12}/n_{1.}$	...	$n_{1j}/n_{1.}$	...	$n_{1c}/n_{1.}$	1
2	$n_{21}/n_{2.}$	$n_{22}/n_{2.}$	...	$n_{2j}/n_{2.}$	...	$n_{2c}/n_{2.}$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$n_{i1}/n_{i.}$	$n_{i2}/n_{i.}$	...	$n_{ij}/n_{i.}$	...	$n_{ic}/n_{i.}$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r	$n_{r1}/n_{r.}$	$n_{r2}/n_{r.}$	...	$n_{rj}/n_{r.}$	...	$n_{rc}/n_{r.}$	1
Sütun Profillerinin Hesaplanması	1	2	...	j	...	c	
1	$n_{11}/n_{.1}$	$n_{12}/n_{.2}$	...	$n_{1j}/n_{.j}$	...	$n_{1c}/n_{.c}$	
2	$n_{21}/n_{.1}$	$n_{22}/n_{.2}$	...	$n_{2j}/n_{.j}$	...	$n_{2c}/n_{.c}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	
i	$n_{i1}/n_{.1}$	$n_{i2}/n_{.2}$		$n_{ij}/n_{.i}$		$n_{ic}/n_{.c}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	
r	$n_{r1}/n_{.1}$	$n_{r2}/n_{.2}$	...	$n_{rj}/n_{.j}$	...	$n_{rc}/n_{.c}$	
Sütun Profil Toplamı	1	1	...	1	...	1	

Kaynak: Kılıç, 2016:5

Satır ve sütun profilleri elde edildikten sonra marjinal oranlar (masses) da hesaplanır. Marjinal satır oranları, satırların marjinal frekanslarının (her bir satır toplamının) tüm toplama bölünmesi sonucunda elde edilir. Elde edilen bu marjinal satır oranları, her bir satır profilini ağırlıklandırmak için kullanılır. Benzer şekilde marjinal sütun oranları da her bir sütun toplamının tüm toplama bölünmesi sonucunda elde edilir (Uzgören, 2007:6).

Uygunluk analizinde noktalar arasındaki uzaklıkların hesaplanırken Ki-Kare uzaklığı kullanılmaktadır. Ki-Kare uzaklığı öklit uzaklığın ağırlıklandırılmış halidir (Kılıç, 2016:6). Bir diğer önemli kavram olan inertia (değişkenlik) ise uygunluk analizinde varyans kavramı (Ki-Kare uzaklığı) ile ilişkilidir. Profil noktalarının merkez etrafında nasıl dağıldıklarını gösteren bir mesafe ölçümüdür (Uzgören, 2007:7).

En son aşamada, en uygun boyutun belirlenirken, analizde boyut sayısı kadar özdeğer hesaplanır. Hesaplanan bu özdeğer, boyutların göreceli önemliliklerini ya da toplam inertianın ne kadarlık kısmının açıklanabildiğini ifade eder (Kılıç, 2016:6).

Çalışmada öncelikle, firmaların paylaştıkları tweet türleri (özgün, mention ve rt tweet) ile firmalar arasındaki uyum incelenmiştir. SPSS programında uygulanan Uygunluk Analizi bulguları, literatürde Kılıç (2016) tarafından yapılan tablo gösterimi temel alınarak açıklanmıştır.

Tablo 3.9: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri

					Açıklanan Inertia	
Boyut	Tekil Değer	Inertia	Ki-Kare	α	Açıklanan	Kümülatif
1	0,464	0,215			0,959	0,959
2	0,097	0,009			0,041	1,000
Toplam		0,225	993,148	0,000	1,000	1,000

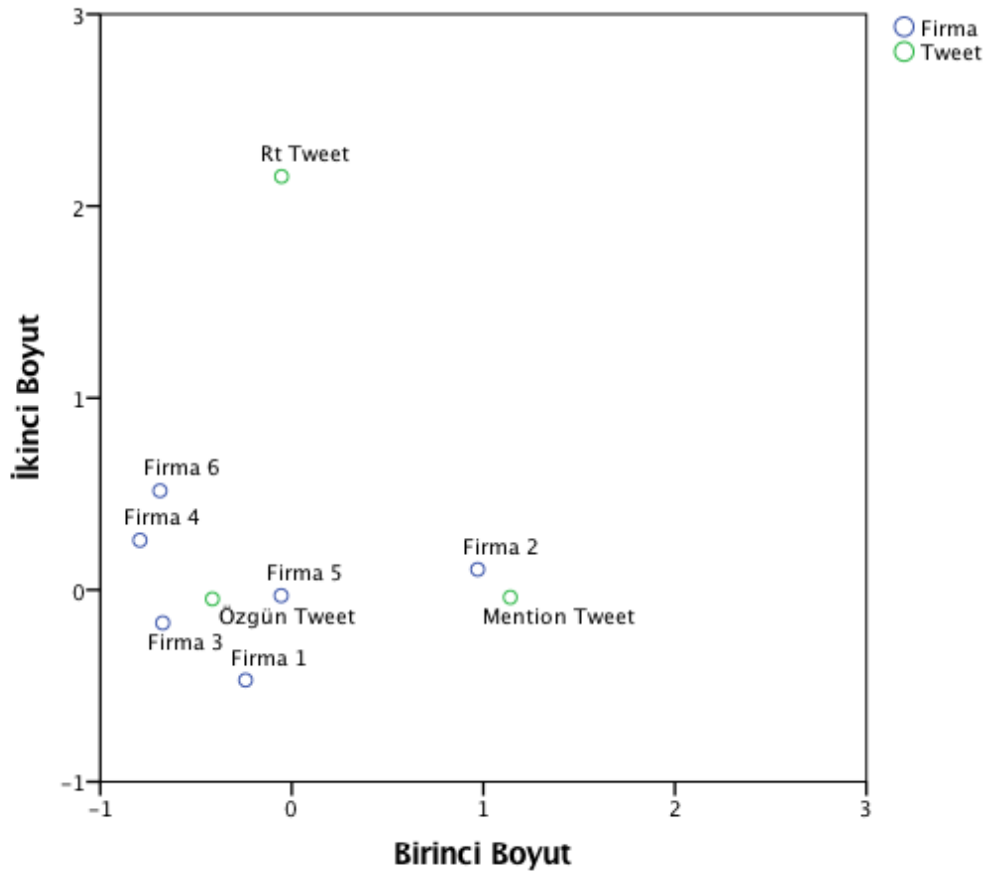
Tablo 3.9’de görüldüğü üzere, inertianın 0’dan farklı olup olmadığının test edildiği Ki-Kare istatistiği değeri ($\alpha=0.00<0.05$) anlamlı çıkmıştır. Buna göre firmalara göre özgün, mention ve rt tweetlerinin dağılımı farklılaşmaktadır. Boyut sayısına göre açıklanan inertia değerleri incelendiğinde, ilk boyutun toplam değişkenliğin %95,9’unu, ikinci boyutun ise %4,1’ini açıkladığı görülmektedir. Böylece iki boyutun toplamı, değişkenliğin tümünü açıklamıştır. Satır ve sütun değişkenlerinin boyutlara olan katkısı, boyutların satır ve sütun değişkenlerine olan katkısı ve inertia değerleri Tablo 3.10’da verilmiştir.

Tablo 3.10: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı

	Değişken	Marjinal Oran	Boyut Puanları		Inertia	Noktaların Boyutlara Katkısı		Boyutların Noktalara Katkısı		Toplam
			1	2		1	2	1	2	
Firma	Firma 1	0,226	-0,241	-0,470	0,011	0,028	0,517	0,559	0,441	1,000
	Firma 2	0,302	0,970	0,107	0,132	0,612	0,036	0,997	0,003	1,000
	Firma 3	0,084	-0,674	-0,172	0,018	0,082	0,026	0,987	0,013	1,000
	Firma 4	0,111	-0,793	0,259	0,033	0,151	0,077	0,978	0,022	1,000
	Firma 5	0,154	-0,056	-0,030	0,000	0,001	0,001	0,943	0,057	1,000
	Firma 6	0,124	-0,688	0,517	0,030	0,126	0,343	0,895	0,105	1,000
	Toplam	1,000			0,225	1,000	1,000			
Tweet	Özgün Tweet	0,718	-0,415	-0,047	0,057	0,266	0,016	0,997	0,003	1,000
	Mention Tweet	0,262	1,141	-0,039	0,158	0,734	0,004	1,000	0,000	1,000
	Rt Tweet	0,020	-0,054	2,154	0,009	0,000	0,980	0,003	0,997	1,000
	Toplam	1,000			0,225	1,000	1,000			

Tablo 3.10’da incelendiğinde Firma 2’nin 0,612 oranla birinci boyutu, Firma 1’in 0,517 oranla ikinci boyutu en çok açıklayan firma değişkenleri olduğu görülmektedir. Tweet kategorileri arasından ise birinci boyuta en fazla katkıyı 0,734 oranla mention tweet kategorisinin, ikinci boyuta olan en yüksek katkıyı 0,980 oranla rt tweet kategorisinin yaptığına ulaşılmıştır. Boyutların noktalara olan katkıları incelendiğinde birinci boyutun 0,997 oranla Firma 2’yi ve özgün tweet kategorisini, ikinci boyutun 0,441 oranla Firma 1 ve 0,997 oranla rt tweet kategorisini açıkladığı görülmektedir. Inertia değerleri ele alındığında, toplam değişkenliği firma kategorilerinden en çok Firma 2 açıklarken, tweet kategorilerinden mention tweet açıklamaktadır. Firma ve tweet kategorilerinin iki boyutlu düzlemde gösterimi Şekil 3.14’de verilmiştir.

Şekil 3.14: Firmalar Tarafından Paylaşılan Tweetlerin Türlerinin ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi



Grafikteki uzaklıklar incelenerek, iki değişkenin kategorileri arasındaki ilişkilerle ilgili çıkarımlarda bulunmak mümkündür. Buna göre tweet kategorileri arasında yer alan mention tweetin bulunduğu konumun uzaklık olarak (Ki-Kare uzaklığı) en yakın olduğu firma, Firma 2’dir. Rt tweet kategorisi ise tüm firmalara uzak konumlanması sebebi ile bu kategorinin firmalar tarafından çok fazla tercih edilmeyen bir tweet türü olduğu söylenebilir. Özgün tweetin ise özellikle Firma 5, Firma 3 ve Firma 1’e oldukça yakın konumlandığı görülmektedir. Tüm bu değerlendirmeler ele alındığında kozmetik sektöründe faaliyet gösteren Firma 2 dışında, diğer tüm firmaların Twitter paylaşımlarının çoğunluğunun özgün tweetlerden oluştuğu söylenebilir.

Bir diğer uyum, firma tweetlerinin paylaşım günleri ile firmalar arasında incelenmiştir. Tablo 3.11 incelendiğinde ilk boyutun toplam değişkenliğin %66,1’ini, ikinci boyutun ise %21,7’sini açıkladığı görülmektedir. İki boyut, toplam değişkenliğin %87,8’ini açıklamaktadır.

Tablo 3.11: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri

					Açıklanan Inertia	
Boyut	Tekil Değer	Inertia	Ki-Kare	α	Açıklanan	Kümülatif
1	0,165	0,027			0,661	0,661
2	0,095	0,009			0,217	0,878
3	0,059	0,004			0,086	0,964
4	0,034	0,001			0,028	0,993
5	0,017	0,000			0,007	1,000
Toplam		0,041	181,736	0,000	1,000	1,000

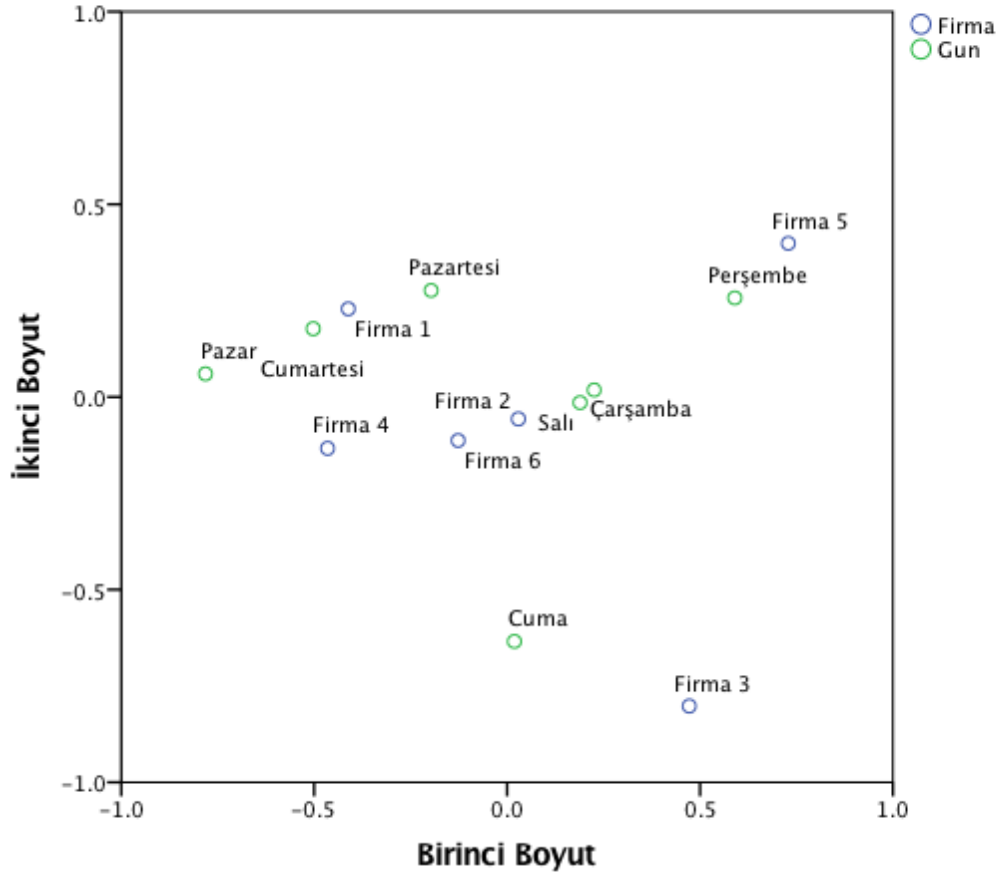
Tablo 3.12 incelendiğinde Firma 5’in 0,496 oranla birinci boyutu, Firma 3’ün 0,568 oranla ikinci boyutu en çok açıklayan firma değişkenleri olduğu görülmektedir. Firma tweetlerinin paylaşım günleri arasından ise birinci boyuta en fazla katkıyı 0,362 oranla pazar gününün, ikinci boyuta olan en yüksek katkıyı 0,728 oranla cuma gününün yaptığına ulaşılmıştır. Boyutların noktalara olan katkıları incelendiğinde birinci

boyutun 0,837 oranla Firma 5'i ve 0,993 oranla pazar gününü, ikinci boyutun 0,618 oranla Firma 3 ve 0,988 oranla cuma gününü açıkladığı görülmektedir. Inertia değerleri ele alındığında, toplam değişkenliği firma kategorilerinden en çok Firma 5 açıklarken, tweet paylaşım günlerinden perşembe açıklamaktadır. Firma ve tweet paylaşım günlerinin iki boyutlu düzlemde gösterimi Şekil 3.15'de verilmiştir.

Tablo 3.12: Firmalar Tarafından Paylaşılan Tweetlerin Paylaşım Günlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı

	Değişken	Marjinal Oran	Boyut Puanları		Inertia	Noktaların Boyutlara Katkısı		Boyutların Noktalara Katkısı		Toplam
			1	2		1	2	1	2	
Firma	Firma 1	0,226	-0,412	0,229	0,008	0,232	0,125	0,822	0,145	0,967
	Firma 2	0,302	0,029	-0,057	0,001	0,002	0,010	0,031	0,068	0,099
	Firma 3	0,084	0,472	-0,802	0,008	0,113	0,568	0,373	0,618	0,991
	Firma 4	0,111	-0,465	-0,133	0,006	0,146	0,021	0,617	0,029	0,646
	Firma 5	0,154	0,729	0,399	0,016	0,496	0,259	0,837	0,144	0,981
	Firma 6	0,124	-0,127	-0,113	0,001	0,012	0,017	0,245	0,110	0,355
	Toplam	1,000			0,041	1,000	1,000			
Paylaşım Günü	Pazartesi	0,142	-0,197	0,277	0,004	0,033	0,115	0,247	0,280	0,527
	Salı	0,144	0,189	-0,015	0,001	0,031	0,000	0,696	0,002	0,698
	Çarşamba	0,166	0,226	0,018	0,002	0,051	0,001	0,607	0,002	0,610
	Perşembe	0,162	0,590	0,257	0,011	0,342	0,113	0,861	0,094	0,955
	Cuma	0,171	0,019	-0,634	0,007	0,000	0,728	0,002	0,988	0,990
	Cumartesi	0,117	-0,502	0,177	0,007	0,180	0,039	0,735	0,052	0,788
	Pazar	0,098	-0,782	0,060	0,010	0,362	0,004	0,993	0,003	0,996
	Toplam	1,000			0,041	1,000	1,000			

Şekil 3.15: Firma Tweetlerinin Paylaşım Günleri ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi



Grafiğe göre Perşembe gününün bulunduğu konumun uzaklık olarak en yakın olduğu firma, elektronik sektöründe faaliyet gösteren Firma 5'dir. Cuma gününün bulunduğu konumun uzaklık olarak en yakın olduğu firma ise, pazaryerinde faaliyet gösteren Firma 3'dür. Kozmetik sektörü firması Firma 1 ve pazaryeri firması Firma 4'ün bulunduğu konum ise Cumartesi gününe en yakındır. Bir diğer kozmetik firması olan Firma 2 ve elektronik firması olan Firma 6 Salı ve Çarşamba günlerine yakın konumda bulunmaktadır. Tüm bu değerlendirmeler ele alındığında aynı sektörde yer alan rakip firmaların birbirlerinden farklı günlerde Twitter'da paylaşım yaptığı söylenebilir.

Son olarak firma tweetlerinin paylaşıldığı zaman dilimleri ile firmalar arasındaki uyum incelenmiştir.

Tablo 3.13: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin Firmalar ile Uyumundaki Boyut Sayısı ve Açıklanan Inertia Değerleri

					Açıklanan Inertia	
Boyut	Tekil Değer	Inertia	Ki-Kare	α	Açıklanan	Kümülatif
1	0,238	0,057			0,822	0,822
2	0,111	0,012			0,178	1,000
Toplam		0,069	304,676	0,000	1,000	1,000

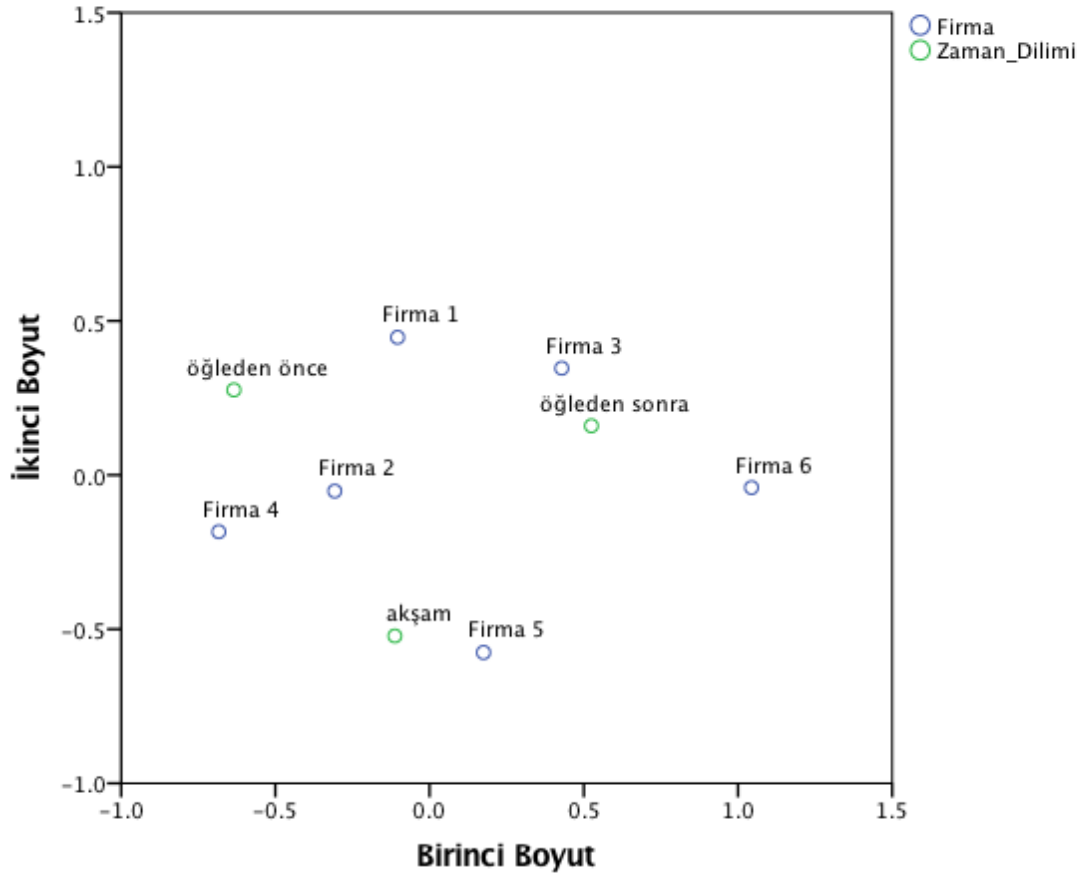
Tablo 3.13’de görüldüğü üzere ilk boyut %82,2’sini, ikinci boyut %17,8’ini olmak üzere iki boyutun tüm değişkenliği açıkladığı görülmüştür. Satır ve sütunların boyutlara olan katkısı ve boyutların satır ve sütunlara olan katkısı Tablo 3.14’de verilmiştir.

Tablo 3.14: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin Firmalar ile Uyumundaki Satırların&Sütunların Boyutlara ve Boyutların Satırlara&Sütunlara Olan Katkısı

	Değişken	Marjinal Oran	Boyut Puanları		Inertia	Noktaların Boyutlara Katkısı		Boyutların Noktalara Katkısı		Toplam
			1	2		1	2	1	2	
Firma	Firma 1	0,226	-0,104	0,447	0,006	0,010	0,406	0,104	0,896	1,000
	Firma 2	0,302	0,308	-0,052	0,007	0,120	0,007	0,987	0,013	1,000
	Firma 3	0,084	-0,428	0,346	0,005	0,064	0,090	0,767	0,233	1,000
	Firma 4	0,111	-0,683	-1,184	0,013	0,218	0,034	0,967	0,033	1,000
	Firma 5	0,154	0,175	-0,576	0,007	0,020	0,460	0,165	0,835	1,000
	Firma 6	0,124	1,044	-0,040	0,032	0,568	0,002	0,999	0,001	1,000
	Toplam	1,000			0,069	1,000	1,000			
Zaman Dilimi	Öğleden Önce	0,296	-0,634	0,276	0,031	0,500	0,204	0,919	0,081	1,000
	Öğleden Sonra	0,419	0,524	0,160	0,029	0,484	0,096	0,959	0,041	1,000
	Akşam	0,285	-0,113	-0,522	0,009	0,015	0,700	0,091	0,909	1,000
	Toplam	1,000			0,069	1,000	1,000			

Firma 6'nın 0,568 oranla birinci boyutu, Firma 5'nin 0,460 oranla ikinci boyutu en çok açıklayan firma değişkenleri olduğu görülmektedir. Firma tweetlerinin zaman dilimleri arasından birinci boyuta en fazla katkısı 0,5 oranla öğleden önce ve ikinci boyuta en fazla katkısı ise 0,7 oranla akşam zaman diliminin yaptığına ulaşılmıştır. Boyutların noktalara olan katkıları incelendiğinde birinci boyutun 0,999 oranla Firma 6'yı ve 0,959 oranla öğleden sonra zaman dilimini, ikinci boyutun 0,896 oranla Firma 1'i ve 0,909 oranla akşam zaman dilimini açıkladığı görülmektedir. Inertia değerleri ele alındığında, toplam değişkenliği firma kategorilerinden en çok Firma 6 açıklarken, tweet zaman dilimlerinden öğleden önce zaman dilimi açıklamaktadır. Firma ve tweet zaman dilimlerinin iki boyutlu düzlemde gösterimi Şekil 3.16'da verilmiştir.

Şekil 3.16: Firmalar Tarafından Paylaşılan Tweetlerin Zaman Dilimlerinin ve Firmaların İki Boyutlu Düzlemde Birlikte Gösterimi



İki boyutlu düzlemde görüldüğü üzere, pazaryeri sektöründe faaliyet gösteren Firma 3 ve elektronik sektöründe faaliyet gösteren Firma 6 öğleden sonra zaman dilimine yakın uzaklıkta konumlanırken, Firma 3'ün rakibi Firma 4 öğleden önce, Firma 6'nın rakibi Firma 5 ise akşam zaman dilimine yakın konumlanmıştır. Kozmetik sektörü firmaları Firma 1 ve Firma 2 ise öğleden önce zaman dilimine yakın uzaklıkta konumlanmıştır. Tüm bu değerlendirmeler ele alındığında kozmetik sektörü dışında aynı sektörde yer alan rakip firmaların birbirlerinden farklı zaman dilimlerinde Twitter'da paylaşım yaptığı söylenebilir.

3.4.1.4.4 İçerik Analizi

Bu bölümde firmaların 2017 yılı boyunca Twitter platformunda yapmış oldukları tweetler ele alınarak öncelikle RapidMiner programında k-means algoritması çalıştırılarak içerikler kümelenmeye çalışılmış, ancak farklı küme sayıları ve uzaklık ölçüleri (kosinüs ve öklit) denenmesine rağmen istenilen dağılım elde edilememiştir. Daha sonra Chae (2015) tarafından yapılan çalışma temel alınarak ayrıntılı kelime analizi gerçekleştirildikten sonra, firmaların paylaşımları belirlenen anahtar kelimeler doğrultusunda kümelenmiştir. Tanımlayıcı istatistiklerde görselleştirilen firma tweetlerinin içermiş oldukları hashtagler, frekans değerlerinin dışında konularına göre incelenerek Hashtag Analizi bölümünde ele alınmıştır. Rt Analizi bölümünde ise Pareto İlkesi kapsamında firma tweet paylaşımlarının almış oldukları rt miktarları analiz edilmiştir.

3.4.1.4.4.1 Tweet Analizi

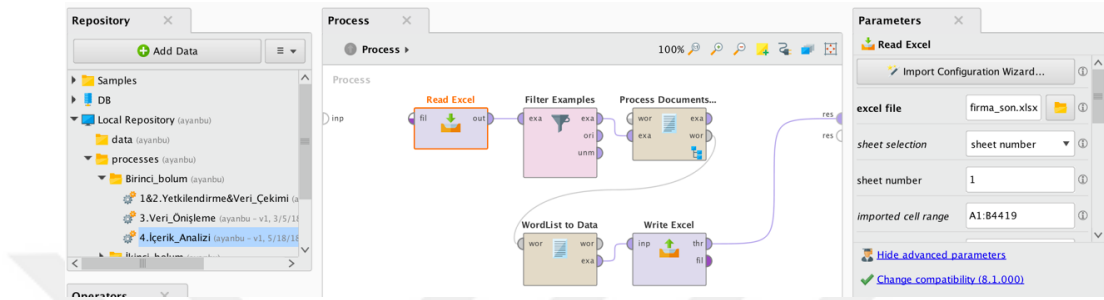
Ayrıntılı kelime analizi gerçekleştirilmeden önce, tweetler bir dizi metin ön işleme aşamalarından geçirilerek temizlenmiştir. Tweetler RapidMiner programında işlenmeden önce Excel’de üç aşamadan geçirilmiştir. Bunlar sırası ile aşağıda verilmiştir:

- **URL Temizleme:** Tweet içerisinde yer alan linkler metin içerisinden temizlenmiştir.
- **Mention (@kullanıcıadı) Temizleme:** Tweet içerisinde yer alan @kullanıcıadı ifadeleri metin içerisinden temizlenmiştir.
- **Küçük Karaktere Dönüştürme:** Tweet içerisindeki tüm kelimeler küçük karaktere dönüştürülmüştür.

Excel dosyasına kayıtlı veriler RapidMiner’a taşınması için “Read Excel” operatörü süreç ekranına sürüklenmiştir. Dosya seçimi ve veri seti ile ilgili ayrıntıları kolayca belirlemek adına parametrelerin yer aldığı ekranda “Import Configuration Wizard”

seçilerek özellikle tweetlerin yer aldığı sütun “text” olarak tanımlanarak gerekli düzenlemeler yapılmıştır.

Şekil 3.17: Tweet Analizi RapidMiner Süreç Modeli



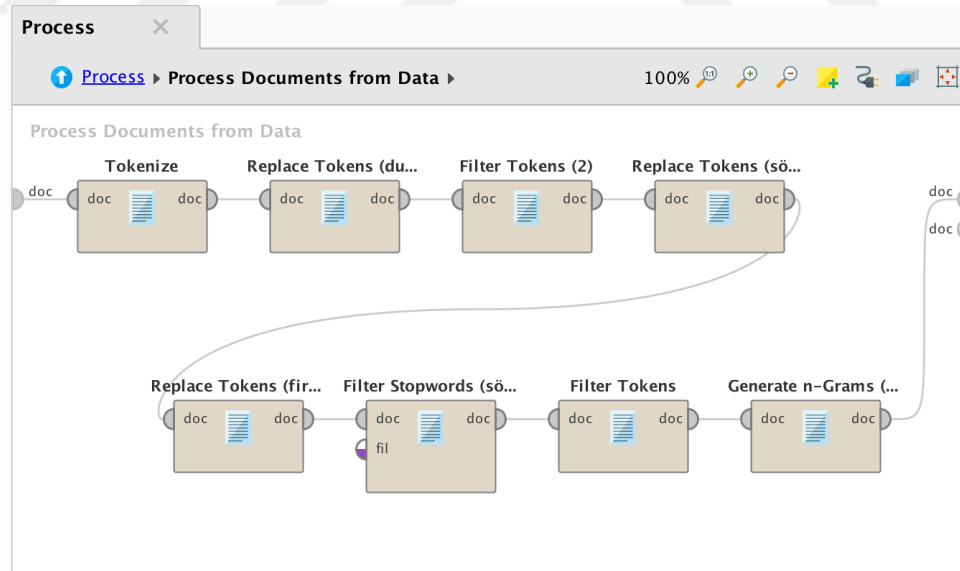
Metni veriye hazırlamak adına RapidMiner programında Process Documents from Data operatörü seçilmiştir. Bu operatör Şekil 3.18’de görüldüğü üzere, içerisinde alt operatörleri de barındırabilen bir çoklu operatördür. Yapılandırılmamış metin verisini temizlemek, yapılandırılmış bir formata dönüştürmek amacı ile bir dizi metin ön işleme işlemleri bu operatörün içerisine taşınmıştır. Kullanılan operatörlerin açıklamaları aşağıdaki gibidir:

- **Parçalara Ayırma (Tokenize):** Bu operatör parçalara (kelimelere) ayırma işlemi için kullanılmıştır. Metin içerisinde yer alan her kelime boşluk ya da noktalama işaretleri tanınarak ayrılmıştır.
- **Durak Sözcüklerini Temizleme (Filter Stopwords-Dictionary):** Tek başına anlam ifade etmeyen çeşitli edatlardan (diye, göre vb.), sayılardan (beş, yirmi vb.), zamirlerden (bizim, size vb.) ve bazı isim ve fiillerden (olduğunda, şey vb.) oluşan sözlük ile metin içerisinde yer alan durak sözcükleri temizlenmiştir.
- **Parçaları Değiştirme (Replace Tokens):** Parçaları değiştirme operatörü ile metin içerisinde yer alan aynı anlama gelen kelimelerin aynı şekilde ele alınması için çeşitli dönüştürmeler yapılmıştır. Örneğin mağaza ve mağazalar kelimelerini birlikte ele almak adına mağazalar kelimesi mağaza kelimesi ile

değiştirilmiştir. EK 1’de bu operatörün içerisinde yer alan sözlükten bir bölüm verilmiştir.

- **Kelime Filtreleme (Filter Tokens-by Length):** Karakter temelinde sınıflandırma yapılarak belirlenen sınırlar dışında bulunan kelimeler analizden çıkarılmıştır. Bu operatör kullanılarak en az 3 karakter bulunma şartı konularak bunun altındaki karakterler elenmiştir.
- **Tamlamalar Yöntemi (Generate N-Grams-Terms):** Bazı kelimeler tek başına değil beraberinde gelen kelime ile tamlama oluşturarak anlam kazandığı için tamlamalar yöntemine başvurulmuştur. İkili ve üçlü tamlamalar modele eklenmiş ve ikili tamlama sonuçlarına göre ikili tamlamalar yönteminin kullanılması uygun bulunmuştur. İkili tamlamalar yöntemine örnek olarak veri setinden “mağaza açmak”, “sosyal medya”, “teşekkür etmek” ve “indirim beklemek” verilebilir.

Şekil 3.18: Tweet Analizi Metin Ön İşleme RapidMiner Süreç Modeli



RapidMiner programında gerçekleştirilen metin ön işleme aşamaları sonucunda kelime listesine ulaşılmıştır. Ön işlemeyen geçirilen veri setine ait kelimeler ve frekansları WordList to Data operatörü kullanılarak Excel’e aktarılmıştır. Kümelere

ayırmak adına liste gözden geçirilmiş ve küme başlıkları belirlenmiştir. Anahtar kelimeler aratılarak küme isimleri veri setinde kodlanmıştır. Kümeler belirlenirken özele inmeden geniş başlıklar seçilerek her firma için ortak olabilecek küme isimleri belirlenmiştir. Her kümeye ait en sık kullanılan beş kelime ve frekansları Tablo 3.15’de verilmiştir.

Tablo 3.15: Kümelere Göre En Sık Kullanılan Kelimeler ve Terim Frekansları

Küme	En Sık Kullanılan Kelime/TF Skoru				
Özel Teklif	indirim/ 865	fırsat/ 366	özel/ 339	ürün/ 298	kaçırmak/ 244
Ürün	yeni/ 194	ürün/ 131	cilt/ 97	renk/ 81	saç/ 78
Yarışma&Etkinlik	kazanmak/ 186	yarışmak/ 115	bilgi/ 96	rica/ 88	etkinlik/ 83
Destek&Geri Bildirim	sevgiler/ 565	merhaba/ 441	rica/ 409	bilgi/ 328	iletmek/ 289
Özel Etkileşim	mağaza/ 184	açmak/ 134	mutlu/ 104	blog/ 99	yarın/ 97
Sosyal	kutlamak/ 65	bayram/ 44	anmak/ 20	dilemek/ 18	anne/ 15

Küme açıklamaları ve örnek firma tweetleri Tablo 3.16’da verilmiştir.

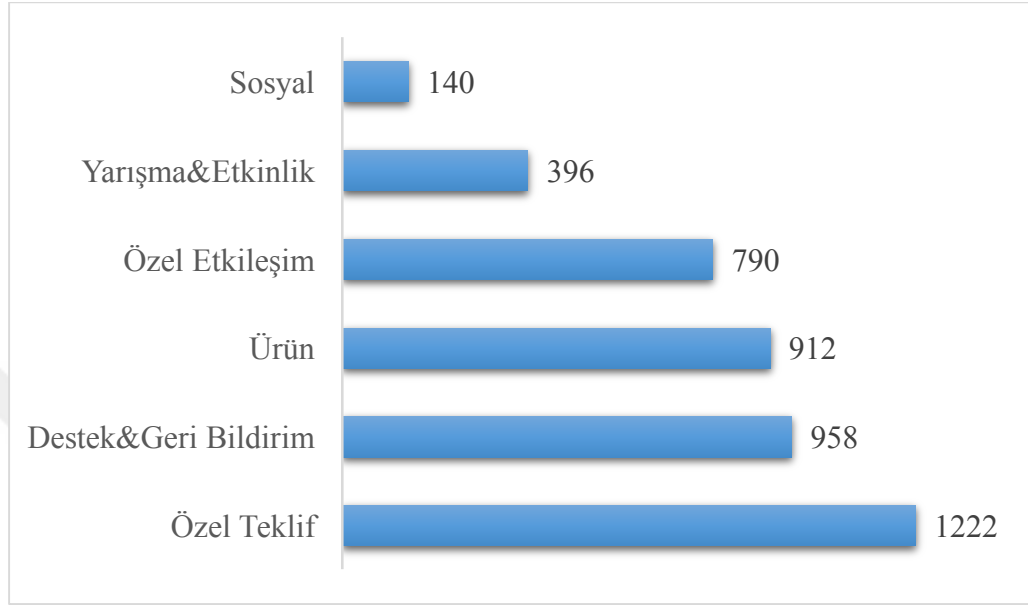
Tablo 3.16: Firmalar Tarafından Paylaşılan Tweetlerin Kümeleri ve Açıklamaları

Küme	Açıklama	Örnek Tweet
Özel Teklif	Firmaların ürün ve hizmetleri ile ilgili paylaştıkları indirim, fırsat, kampanya ve hediye tweetlerini içeren kümedir.	En güzel hafta sonu indirimini 27-29 Ocak tarihleri arasında tüm Firma 1’lerde sizleri bekliyor!
Ürün	İndirim gibi özel teklifler dışında kalan, içerisinde sadece ürünlerle ilgili bilgileri (yeni çıkan ürün, katalog bilgisi, ürünle ilgili çekilmiş video içerikleri vb.) barındıran ürün temelinde yapılan paylaşımlardan oluşan kümedir.	Hindistan cevizinin yenileyen ve güzelleştiren mucizesini barındıran tüm #ürünadı cilt bakım ürünleri mağazalarımızda!

Yarışma&Etkinlik	Takipçilere yönelik çeşitli şekillerle düzenlenen (soru, katılım, başvuru, çekiliş) yarışmalar ile ilgili mesajları, söyleşi ve davet gibi etkinlik tweetlerini ve bu yarışmaların sonuçlarının duyurulduğu tweetleri içeren kümedir.	Hafta sonunun en güzel yarışması Instagram hesabımızda başladı!
Destek&Geri Bildirim	Kullanıcılar tarafından bildirilen sorunlara, sorulara ve bildirimlere yönelik direkt olarak kullanıcılara yanıt verildiği tweetleri içeren kümedir.	@kullanıcıadı Merhabalar, şikâyet kaydınızı oluşturabilmemiz için iletişim bilgilerinizi bizimle D.M yolu ile paylaşmanızı rica ediyoruz.
Özel Etkileşim	Firmaların belirlediği özel içeriklerle (hashtag vb.) takipçilerini katılıma teşvik ettiği tweetleri, kullanıcılara özel olarak paylaşılan teşekkür mesajlarını, ürünler dışında paylaşılan önerileri, etkileşimi arttırmaya yönelik hazırlanan özel videoları ve firma ile ilgili özel bilgilendirme tweetlerini (mağaza açılış, kazanılan ödüller vb.) içeren kümedir.	500. mağaza mutluluğunu hep beraber yaşadık. Sizlerden gelen tüm eğlenceli yorumlar için çok teşekkürler. 💜 Siz de #Firma1veBen hashtagi ile en yaratıcı tweet ve mesajlarınızı atarak bir sonraki videomuzda yer alabilirsiniz. 🤔
Sosyal	Bayram, maç, özel gün gibi kutlama, dilek, baş sağlığı tweetlerini içeren kümedir.	Ortaköy'de yaşanan terör saldırısında hayatını kaybedenlere Allah'tan rahmet, yakınlarına sabır ve tüm yaralılara acil şifalar diliyoruz.

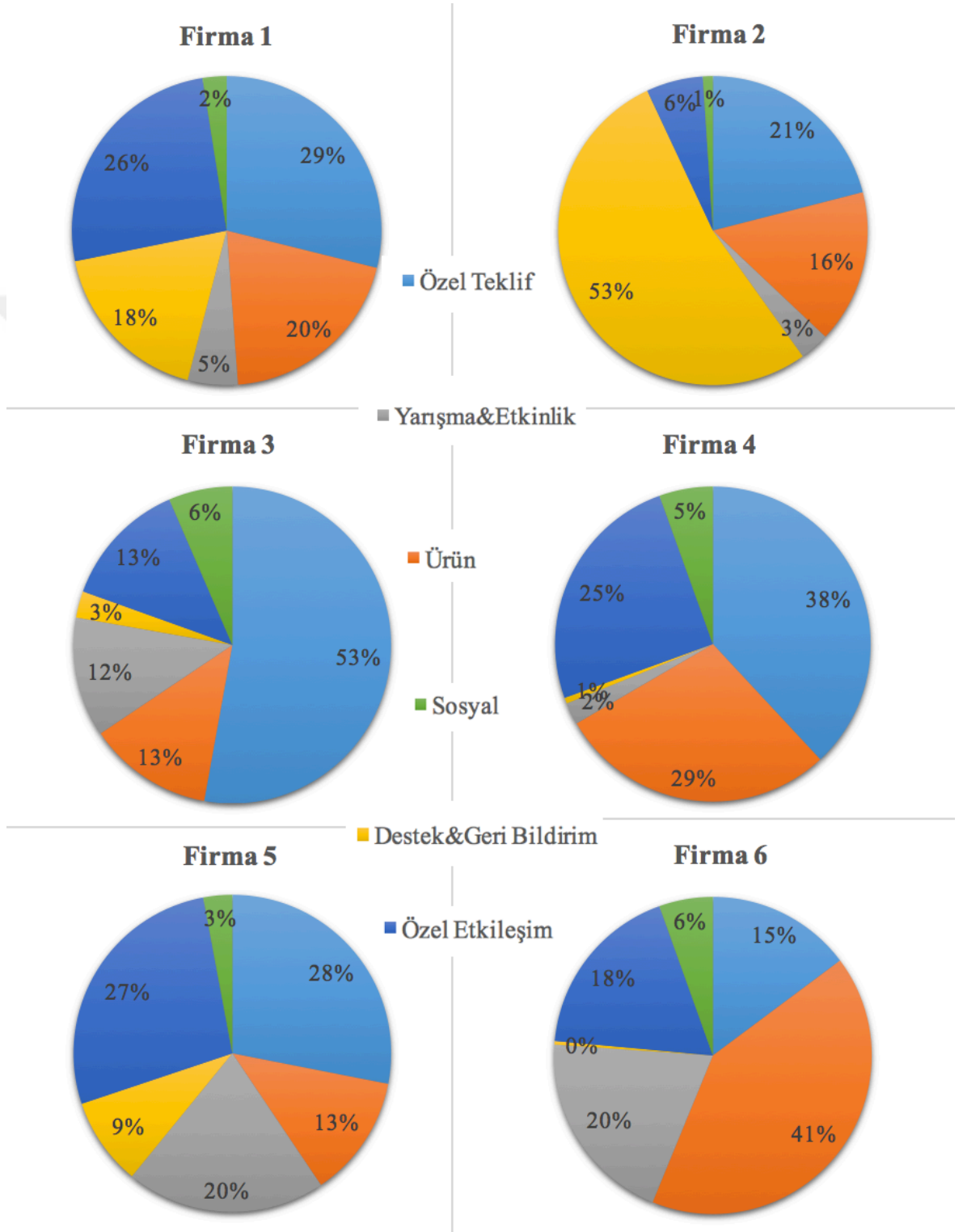
Firma tweetlerinin belirlenen kümelerle göre dağılımları elde edilmiştir. Öncelikle kümelerle göre içeriklerin dağılımları genel anlamda ele alınmış ve yüzdelik frekanslar hesaplanmıştır. Şekil 3.19'da görüldüğü üzere 2017 yılı boyunca altı firmanın Twitter platformu üzerinden paylaştıkları tweetlerin kümelerle göre frekans sıralaması sırasıyla Özel Teklif (%28), Destek&Geri Bildirim (%22), Ürün (%20), Özel Etkileşim (%18), Yarışma&Etkinlik (%9) ve Sosyal (%3) kategorisindedir.

Şekil 3.19: Firmalar Tarafından Paylaşılan Tweetlerin Kümelere Göre Frekans Dağılımları



Bu genel değerlendirme sonrasında, firmaların Twitter paylaşımlarının kategorilerinin yüzdelik dağılımı Şekil 3.20’deki gibi firma temelinde görselleştirilmiştir.

Şekil 3.20: Firmalar Tarafından Paylaşılan Tweetlerin Kümelerinin Firmalara Göre Frekans Dağılım Yüzdeleri



Şekil 3.20 incelendiğinde aynı sektörde (kozmetik) yer almalarına rağmen, Firma 1 ve Firma 2 tweet paylaşımlarının küme yüzdeleri ele alındığında büyük farklılık gözlemlenmiştir. Firma 2 tweetlerinin yarısından fazlasını (%53) Destek&Geri Bildirim kategorisinde paylaşırken, Firma 1 ise bu alanda yalnızca %18 oranında paylaşım yapmıştır. Bu farkın sebebi, Firma 2'nin online alışveriş kanalı ile de hizmet vermesi olduğuna tweet içerikleri ayrıntılı olarak incelendiğinde ulaşılmıştır. Online alışveriş ile ilgili kullanıcıların sordukları sorulara, kullanıcı şikâyet ve sorunlarına, Firma 2 Twitter aracılığı ile yanıt vermiştir. Bu iki firma arasındaki bir diğer büyük farklılık ise Özel Etkileşim kategorisi yüzdelerinde gözlemlenmektedir. Firma 1 paylaşımlarının %26'sını bu kümede yaparken Firma 2 bu kümeye ait yalnızca %6'lık içerik üretmiştir.

Pazaryeri sektöründe faaliyet gösteren Firma 3 ve Firma 4 tweet paylaşımlarının küme yüzdeleri ele alındığında, en çok paylaşım yapılan kategori başlığının (Özel Teklif) aynı olduğu görülmektedir. Özel Teklif kategorisinde Firma 3 %53, Firma 4 ise %38 oranında paylaşım yapmıştır. Ürün ve Özel Etkileşim kategorileri, firma paylaşım yüzdeleri farklılık göstermekle birlikte, iki rakip firma tarafından en çok tercih edilen diğer iki kategoridir. Bu bulgular neticesinde tweet paylaşımlarının kategorileri pazaryeri sektörü firmaları için benzerlik göstermektedir denebilir.

Elektronik sektörü firmaları incelendiğinde, Firma 5'in en çok Özel Teklif (%28), Firma 6'nın ise en çok Ürün (%41) kategorisinde paylaşım yaptığı görülmektedir. Firma 5'in en çok paylaşım yaptığı ikinci kategori (%27) olan Özel Etkileşim, Firma 6 tarafından daha az bir oranla (%18) paylaşılmıştır. Her iki firma da aynı oranlarda Yarışma&Etkinlik kategorisinde paylaşım yapmışlardır. Bu bulgular neticesinde benzerlikler olmakla birlikte tweet paylaşımlarının kategorilerinin elektronik sektörü firmaları için farklılık gösterdiği gözlemlenmiştir.

Pazaryeri ve elektronik sektöründe faaliyet gösteren firmaların Destek&Geri Bildirim kategorisine ait tweet yüzdelerinin oldukça düşük olduğu gözlemlenmiştir. Bu firmaların (Firma 3, Firma 4, Firma 5 ve Firma 6) kurumsal hesapları dışında destek

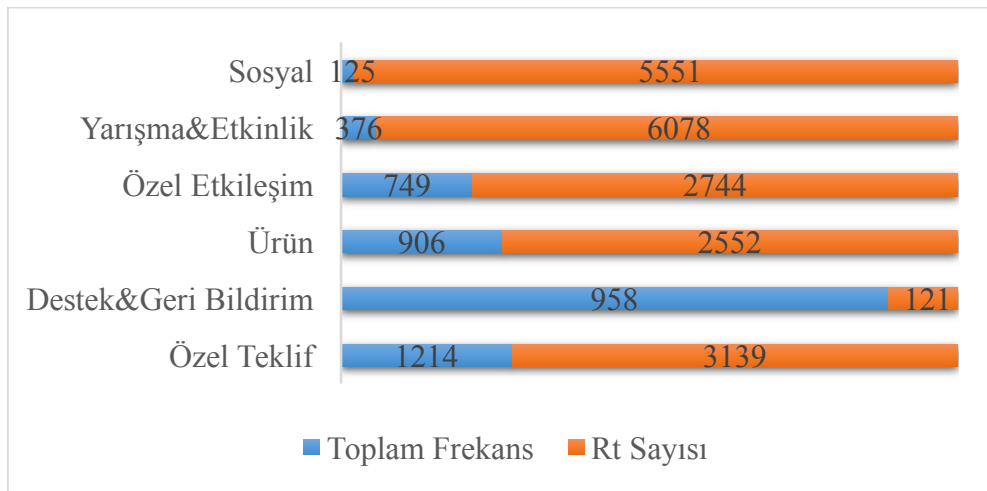
hesapları da mevcuttur. Bu hesaplar incelendiğinde 2017 yılında firmaların Twitter destek hesaplarından kullanıcılara yanıt tweetleri paylaşıldığı görülmüştür. Bu sebeple kullanıcılara yanıt verdikleri mesajlarını kurumsal hesaplarından genellikle paylaşmamayı tercih ettikleri söylenebilir.

3.4.1.4.4.2 Rt Analizi

Bu bölümde firma tweet paylaşımlarının almış oldukları etkileşim sayıları üzerinden gidilerek (rt sayıları) hangi içeriklerin en fazla etkileşim aldıkları analiz edilmiştir. Veri seti incelendiğinde firma tweet paylaşımlarının almış oldukları rt sayıları değerlendirilmiş ve firmaların başka hesapların tweetlerini paylaştıkları (rt ettikleri) tweetlerin istatistikleri yanıtıldığı gözlemlenmiştir. Bu sebeple çalışmanın bu bölümü için rt edilen içeriklerin (Rt Tweet) çıkarılması kararlaştırılmıştır.

Rt tweet türü çıkarıldıktan sonra, tüm firmaların tweet paylaşımları birlikte değerlendirilerek, İçerik Analizi bölümünde belirlenen altı küme kapsamında kümelerin içerdiği tweet frekansları ve o bu tweetlerin almış oldukları rt sayıları Şekil 3.21'deki gibi görselleştirilmiştir.

Şekil 3.21: Firmalar Tarafından Paylaşılan Tweetlerin Kümelere Göre Frekans Dağılımları ve Almış Oldukları Rt Sayıları



Şekil 3.21 incelendiğinde, Yarışma&Etkinlik ve Sosyal kategorileri en düşük tweet frekansına sahip olmasına rağmen en çok rt edilme frekansına sahip olduğu görülmektedir. Sosyal kategoride yer alan paylaşımların (125 tweet) toplamda 5.551 kez, Yarışma&Etkinlik kategorisinde yer alan paylaşımların (376 tweet) ise toplamda 6.078 kez rt edildiği önemli bir bulgudur. Bu sonuca göre firmaların kullanıcılara yönelik paylaşmış olduğu yarışma ve etkinlik mesajlarının ve herkesi ilgilendiren dilek ve kutlama gibi sosyal mesajların, diğer kategori paylaşımlarının almış oldukları etkileşim frekansları ile karşılaştırıldığında, daha fazla kullanıcı tarafından paylaşıldığı söylenebilir.

Kategorilere göre alınan rt sayılarının dışında, firmaların bir yıl boyunca yapmış oldukları paylaşımlar sonucunda elde ettikleri toplam rt sayılarının büyük çoğunluğunun, firmaların tweetlerinin yüzde kaçından oluştuğu incelenmiştir. Bu incelemeyi gerçekleştirmek adına veri seti Pareto İlkesi kapsamında ele alınmıştır.

Pareto İlkesi ilk defa, adını aldığı iktisatçı ve sosyolog Vilfredo Pareto tarafından ülke servetinin yaklaşık yüzde 80'inin popülasyonun yaklaşık yüzde 20'sinde yoğunlaştığını bulması sonucu gözlemlenmiştir. Daha sonra başka araştırmacılar tarafından yapılan incelemeler sonucu 80/20 Kuralı olarak da bilinen bu ilkenin, birçok alanda sirayet ettiği görülmüş ve bir kural haline gelmiştir (Sanders, 1987:37).

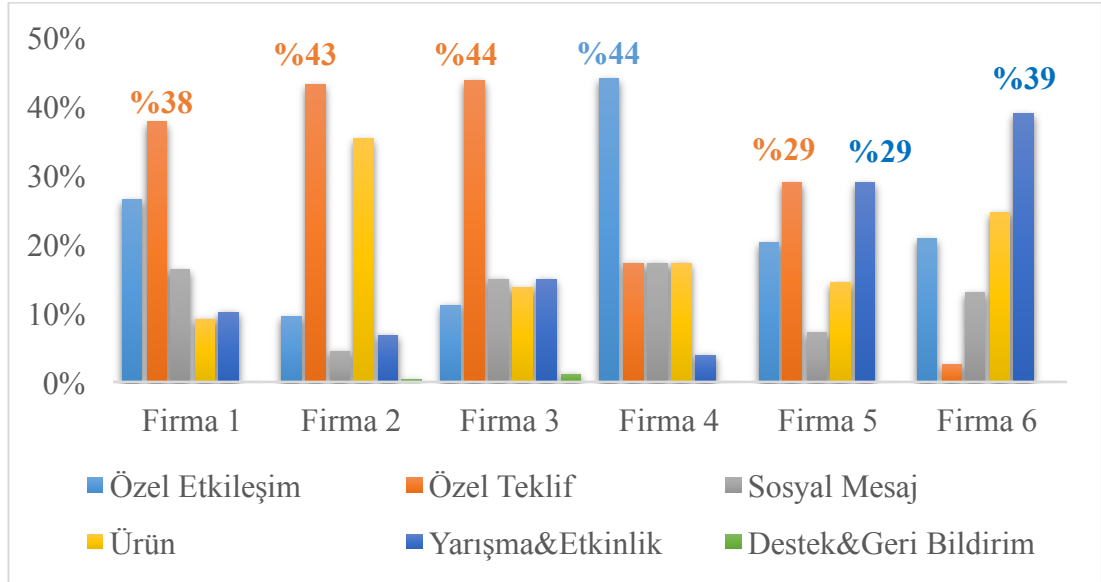
Rt içerikleri temizlenen veri setinde yer alan tweet sayıları, toplamda alınan rt miktarı ve toplam rt miktarının %80'inin ve %90'ının tweetlerin yüzde kaçından oluştuğunun oranları Tablo 3.17'de verilmiştir.

Tablo 3.17: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekanslarının Pareto İlkesi Kapsamında Dağılım Yüzdeleri

	Firma 1	Firma 2	Firma 3	Firma 4	Firma 5	Firma 6
Toplam Tweet Sayısı	997	1301	364	475	667	524
Toplam Rt Sayısı	7352	1180	1013	710	3206	6724
Rt- %80	%10	%17	%22	%16	%10	%15
Rt- %90	%25	%24	%35	%28	%22	%33

Bulgulara göre dağılımın Pareto ilkesine uyum sağladığı görülmektedir. Toplamda elde edilen rt miktarının büyük çoğunluğunun veri setinde yer alan tweetlerin azınlık bölümünden elde edildiğine ulaşılmıştır. Örneğin; Firma 3’ün 2017 yılı boyunca paylaşmış olduğu 364 tweetin almış olduğu toplam rt sayısının (1.013) %80’i yalnızca %22’lik orana denk gelen tweetlerden elde edilmiştir.

Şekil 3.22: Firmalar Tarafından Paylaşılan Tweetlerin Almış Olduğu Rt Frekanslarının Pareto İlkesi Kapsamında Firma Tweet Kümelerine Göre Dağılım Yüzdeleri



Bu bulgular neticesinde 80/20 Kuralı ele alınarak rt miktarının büyük çoğunluğunun elde edildiği tweet kategorileri incelenmiştir. Toplam etkileşimin yüzde 80’inin elde

edildiği tweet kategorileri Şekil 3.22'deki gibi görselleştirilmiştir. Sonuçlara göre kozmetik sektöründe faaliyet gösteren iki firmanın en çok etkileşim aldığı tweetlerin kategorilerinin büyük çoğunluğunun Özel Teklif kategorisinden (Firma 1; %38, Firma 2; %43) oluştuğu söylenebilir. Pazaryeri sektöründe faaliyet gösteren Firma 3 en çok Özel Teklif (%44) ve rakibi Firma 4 en çok Özel Etkileşim (%44) içeren tweetler paylaşarak, elektronik sektöründe yer alan Firma 5 ve Firma 6 ise en çok Yarışma&Etkinlik kategorisinde paylaşımda bulunarak toplam etkileşimlerinin (toplam rt sayılarının) büyük çoğunluğunu elde ettikleri görülmektedir. Firma 5 aynı oranla Özel Teklif (%29) kategorisinde de paylaşımlarda bulunarak yüksek rt sayılarını elde etmiştir.

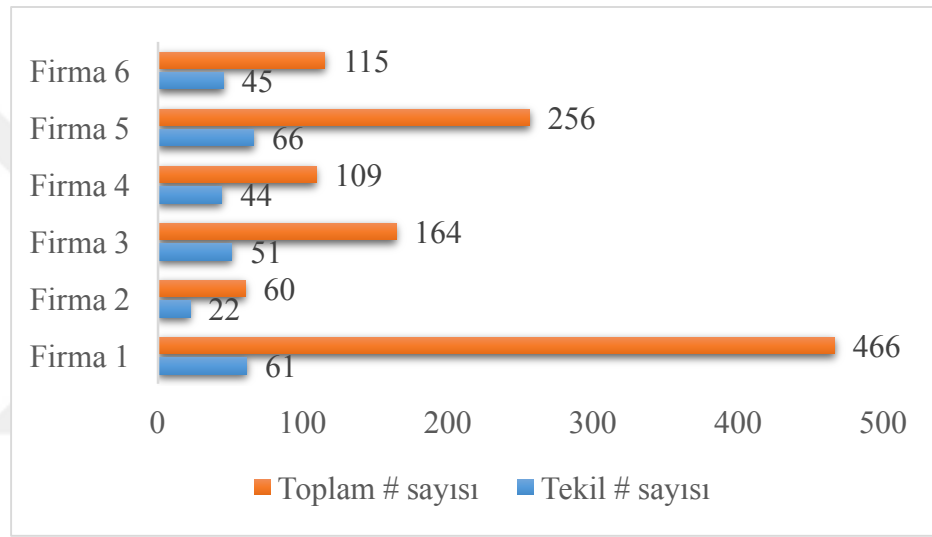
3.4.1.4.4.3 Hashtag Analizi

Firmalar tweetlerinde etkileşimi arttırmak gibi çeşitli amaçlarla hashtag kullanmaktadırlar. Tanımlayıcı istatistikler bölümünde değinildiği üzere, 2017 yılı hashtag içeren paylaşımların en yüksek frekansa sahip olduğu firma Firma 1'dir. Bu bölümde ayrıntılı olarak firmaların hashtag içeren paylaşımlarına ek olarak, toplamda kaç kez hashtag kullandıklarının frekansları ve en fazla tekrar ettikleri hashtagler incelenmiştir.

Bazı firmalar hashtag kullanımlarında aynı hashtag içeriğini hem Türkçe karakter kullanarak hem de kullanmayarak paylaşmışlardır. Çalışmada aynı tweet içerisinde birbirini tekrar eden, Türkçe karakter kullanımından dolayı farklı olarak algılanan iki hashtag tek bir hashtag olarak ele alınıp hesaplamalar gerçekleştirilmiştir. Örneğin bir tweet içerisinde yer alan #guzel ve #güzel hashtagleri tek bir hashtag (#güzel) şeklinde ele alınarak frekans sayısı bir olarak hesaplanırken, farklı tweetlerde yer alan bu iki hashtag yine tek bir hashtag olarak değerlendirilmiş, ancak frekans sayısı iki olarak hesaplanmıştır. Paylaşımlarda kaç farklı hashtag kullanıldığı (tekil # sayısı) ve frekans sayıları göz önünde bulundurularak bir değerlendirme gerçekleştirilmiştir.

Tweetlerde yer verilen hashtagler incelendiğinde kaç farklı hashtag kullanıldığı (tekil # sayısı) her firma için ayrı ayrı hesaplanmıştır. Şekil 3.23’de görüldüğü üzere en fazla farklı hashtag kullanan firma 66 hashtag ile Firma 5 olurken, onu 61 hashtag ile Firma 1 izlemiştir. Hashtaglerin frekans sayıları dikkate alındığında ise Firma 1’in diğer firmalara kıyasla içeriklerinde daha fazla hashtag kullandığı görülmektedir.

Şekil 3.23: Firmalar Tarafından Paylaşılan Tweetlerin İçerdiği Tekil ve Toplam Hashtag Frekans Dağılımları



Hashtaglerin nelerle ilgili olduğu incelendiğinde dört genel kategorinin söz konusu olduğu görülmektedir. Firmaların kendilerine özgü (#mutlulukfirma1de), ürünleri ile ilgili (#xürünüfirma6da), özel günlerle ilgili (#10kasım) ve etkinlik katılımı ile ilgili (#etkinlikadı) paylaşımlardan oluşan hashtag kullanımları olduğu görülmektedir.

Firmaların en çok paylaştıkları hashtaglerin değerlendirilmesi adına en fazla frekansa sahip hashtagler her firma için ayrı olarak ele alınmıştır. Bu değerlendirmeye göre Firma 1’in paylaştığı en yüksek frekansa sahip hashtagin firmanın rengi ile ilgili olduğuna ve bu hashtagin 2017 yılı boyunca paylaşılan tweetlerde toplamda 84 kere geçtiğine ulaşılmıştır. Firma 2’nin en çok paylaştığı hashtagin ise firmanın düzenlemiş olduğu özel etkinlik ile ilgili olduğuna ulaşılmıştır. Bu etkinlik hashtagi bir yıl boyunca toplamda 25 kez tweet içeriklerinde yer almıştır. Firma 3 ise kadınlara

yönelik geliştirdiği programla ilgili olan hashtage toplamda 22 kez yer vermiştir. Firma 4, diğer firmalardan farklı olarak sosyal mesaj barındıran, anneler gününe özel bir hashtag oluşturmuş ve bunu toplamda 19 kez paylaşmıştır. Firma 5 kampanya ile ilgili firmaya özgü bir hashtag geliştirmiş ve bunu bir yıl boyunca toplamda 66 kez paylaşmıştır. Firma 6 ise diğer firmalardan oldukça farklı olarak katıldığı bir fuarın adından oluşan hashtagi toplamda 17 kez paylaşmıştır.

3.4.1.5 Değerlendirme

Literatürde değinildiği üzere sonuçların sunumu olarak adlandırılan Değerlendirme ya da bir diğer adı ile Görselleştirme aşaması, uygulamanın bu bölümü için farklı açıdan ele alınmıştır. Uygulamanın sonucunda elde edilen her analiz bulgusu analizin içerisinde görselleştirilerek açıklandığı için bu bölümde, uygulamada yer verilen analizlerin sonuçları genel bir değerlendirme ile özet şeklinde açıklanmıştır. Ki-Kare Analizi, Uygunluk Analizi ve İçerik Analizi bulguları Tablo 3.18’de genel başlıkları ile özetlenmiştir.

Tablo 3.18: Firmaların Twitter Kullanımlarına Ait Değerlendirmelerin Özet Bulguları

Firma	Tweet Türü	Tweet Paylaşım Günü	Tweet Paylaşım Zaman Dilimi	Tweet Kategorisi	Elde Edilen Rt Sayılarının Büyük Çoğunluğunun Ait Olduğu Tweet Kategorisi
Firma 1	Özgün	Cumartesi	Öğleden Önce	Özel Teklif	Özel Teklif
Firma 2	Mention	Salı	Öğleden Önce	Destek&Geri Bildirim	Özel Teklif
Firma 3	Özgün	Cuma	Öğleden Sonra	Özel Teklif	Özel Teklif
Firma 4	Özgün	Cumartesi	Öğleden Önce	Özel Teklif	Özel Etkileşim
Firma 5	Özgün	Perşembe	Akşam	Özel Teklif	Yarışma&Etkinlik/ Özel Teklif
Firma 6	Özgün	Salı	Öğleden Sonra	Ürün	Yarışma&Etkinlik

Firma 2 haricinde tüm firmaların özgün tweetler paylaştığı bulgular arasında yer almaktadır. Tweet paylaşım günleri ve zaman dilimleri ele alındığında, aynı sektörde yer alan rakip firmaların farklı günlerde paylaşım yaptıklarına ulaşılmıştır. Tweet paylaşım zaman dilimleri ele alındığında ise yalnızca kozmetik sektöründe faaliyet gösteren firmaların aynı zaman diliminde daha fazla paylaşım yaptığı saptanmıştır. Tweet kategorilerine bakıldığında ise firmaların tweetlerin en çok paylaşıldığı kategoriler kapsamında yalnızca pazaryeri sektörü için ortak sonuca ulaşılmıştır. Diğer sektörlerdeki rakip firmalar farklı kategorilerde daha yoğun bir şekilde içerik paylaşmışlardır. Firma paylaşımlarının almış olduğu rt sayıları Pareto İlkesi kapsamında ele alındığında ise etkileşimin büyük çoğunluğunun elde edildiği tweet kategorilerinin kozmetik sektörü için Özel Teklif, elektronik sektörü için ise Yarışma&Etkinlik olduğuna ulaşılmıştır. Pazaryeri sektöründe faaliyet gösteren Firma 3 ve Firma 4'ün ise farklı kategorilerde paylaşımları ile rt sayılarının büyük çoğunluğunu elde ettiği görülmektedir.

Tablo 3.18'de özet olarak sunulan bulguların dışında firmaların tweetlerinin kategorileri toplamda ele alındığında, en çok rt edilen içeriklerin Sosyal ve Yarışma&Etkinlik kategorilerine ait olduğu belirlenmiştir. Firmaların hashtag kullanımlarına bakıldığında ise, Firma 1'in en çok hashtag kullanan firma olduğu saptanmıştır.

3.4.2 Firmaların Twitter Etkinliklerinin Ölçümü

Bu bölümde uygulamanın birinci bölümünden farklı olarak, örneklem dâhilinde belirlenen 2018 Şubat ayı boyunca firmaların ve kullanıcıların firmalar ile ilgili paylaşımları ele alınarak bir Twitter etkinlik (performans) ölçümü yapılması amaçlanmıştır. Birinci bölümde olduğu gibi bu bölümde de verinin elde edilmesi ve analizi sürecinde literatürde değinilen sosyal medya madenciliği süreç adımları izlenmiştir.

3.4.2.1 Yetkilendirme

Firma Twitter paylaşımlarının değerlendirmelerinin yapıldığı ilk bölümde oluşturulan aplikasyon sonucunda elde edilen ilgili değerler (token) bu bölümde de yetkilendirme aşamasında kullanılmıştır. Firmalar ile ilgili 2018 Şubat ayı boyunca kullanıcılar tarafından yapılan Twitter paylaşımlarını elde etmek için C# kodlama dilinde yazılan uygulamaya yetkilendirme işlemi tanınmıştır. 2018 Şubat ayı firma paylaşımları için ise yetkilendirme işlemi ilk bölümde RapidMiner Studio 8.1 aracında gerçekleştirildiği için süreç tekrarlanmamıştır.

3.4.2.2 Veri Toplama

İlk bölümde olduğu gibi bu bölümde de Twitter'ın sağlamış olduğu açık arayüz (API) desteğinden faydalanılarak Twitter'dan veri çekimi gerçekleştirilmiştir. Bu bölümde farklı olarak, bir kullanıcının (firmaların) paylaşmış olduğu tweetler değil, bir konu/kelime (firma) ile ilgili paylaşılmış olan tweetler elde edilmiştir. Twitter API ile son 7 güne ait herkese açık (public) tweetlere ulaşılabilir. Ücretsiz olarak faydalanılan bu arayüz günlük 2.400 tweet sınırlaması dâhilinde tweet getirmektedir. Tweet sınırlaması olmasına rağmen, her anahtar kelime sorgusu için aynı kurallar mevcut olduğundan dolayı, firmalar ile ilgili tweetlerin, karşılaştırılması uygundur.

Tweetlere erişebilmek için .NET için geliştirilmiş bir kütüphane olan “Tweetinvi” adlı kütüphaneden faydalanılmıştır. Bu kütüphaneye referans veren bir Windows form uygulaması, Visual Studio 2015 ortamında C# kodlama dili ile geliştirilmiştir. Bu uygulama “Tweetera” olarak isimlendirilmiş ve ilgili kodlarla birlikte uygulamanın ekran görüntüsü EK 5’de ve EK 6’da verilmiştir. Elde edilmek istenilen firmalarla ilgili kullanıcılar yapılan paylaşımlar, uygulamada anahtar kelimeler firma isimleri olmak üzere aratılarak sorgulanmıştır.

Firmaların 2018 Şubat ayı boyunca yaptığı paylaşımlar ise ilk bölümde açıklandığı üzere RapidMiner Studio 8.1 aracında “Get Twitter User Statuses” operatörü yardımı çekilmiştir.

3.4.2.3 Veri Ön İşleme

Veri setindeki analize dâhil edilmeyen değişkenler çıkarılarak veri seti indirgenmiş ve nihai veri setinde yer alan değişkenlerin açıklamalarına yer verilmiştir.

3.4.2.3.1 Veri İndirgeme

RapidMiner aracı kullanılarak elde edilen veri setinde analize dâhil edilmeyecek olan değişkenler çıkarılmıştır. İlk olarak analize dâhil edilmeyecek olan değişkenler belirlenmiştir. Bu sütunlar birinci bölümde uygulandığı gibi “Select Attributes” operatörü ile veri setinden çıkarılmış ve Write Excel operatörü kullanılarak yeni veri seti olarak kaydedilmiştir.

3.4.2.3.2 Veri Temizleme

Veri seti incelendiğinde, kullanıcılar tarafından paylaşılmış içeriklerin farklı kullanıcılar tarafından tekrar paylaşılması sonucu aynı tweetin veri setinde defalarca tekrarlanmış olduğu görülmüştür. RapidMiner aracı ile “Remove Duplicates” operatörü kullanılarak tekrar eden tweetler, aslı veri setinde muhafaza edilerek, temizlenmiştir. Daha sonra, otomatik hesaplardan (bot) yapılan paylaşımlar, iş ilanları ve ilgisiz tweetlerde hashtag kullanımı ile veri setine dâhil olan tweetler temizlenmiştir. Firmaların kurumsal hesaplarından yapılan tweet paylaşımları da veri setinden çıkarılmıştır.

2018 Şubat ayı boyunca firmalar ile ilgili elde edilen ham veri setinin ve veri temizleme sonucunda oluşturulan nihai veri setinin frekans dağılımı aşağıda verilmiştir.

Tablo 3.19: 2018 Şubat Ayı Boyunca Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Frekans Dağılımları

Sektör Adı	Firma	Ham Veri	Tekrar Eden Tweetler	Kurumsal Firma Tweetleri	İlgisiz Tweetler	Temizlenmiş Veri
		(1)	(2)	(3)	(4)	(5=1-2-3-4)
Kozmetik	Firma 1	10.790	4.272	76	352	6.090
	Firma 2	1.305	122	97	52	1.034
Pazaryeri	Firma 3	1.792	355	34	84	1.319
	Firma 4	2.102	785	36	85	1.196
Elektronik	Firma 5	3.376	1.684	49	31	1.612
	Firma 6	1.073	338	26	38	671
Toplam		20.438	7.556	318	642	11.922

3.4.2.3.3 Veri Seti

İndirgenen ve temizlenen veri setini oluşturan değişkenlerin açıklamaları Tablo 3.20’de verilmiştir.

Tablo 3.20: Uygulamanın İkinci Bölümüne Ait Değişken Tablosu

Değişken İsmi	Açıklaması
Tarih	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının hangi tarihte yapıldığının bilgisinin gün/ay/yıl olarak verildiği değişkendir.
Saat	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının hangi saatte yapıldığının bilgisinin saat/dakika/saniye olarak verildiği değişkendir.
Zaman Dilimi	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının yapıldığı saat bilgisi göz önünde bulundurularak bir sınıflandırılma

	yapılmış ve beş sınıf dâhilinde (≥ 05 - < 09 =sabah, ≥ 09 - < 12 =öğleden önce, ≥ 12 - < 17 = öğleden sonra, ≥ 17 - ≤ 23 = akşam, $00 \geq$ - < 5 =gece) bu değişken hesaplanmıştır.
Gün	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının atıldığı gün bilgisinin verildiği değişkendir. Tarih değişkeninden elde edilmiştir.
Tweet	Tweet olarak adlandırılan kullanıcıların firmalar ile ilgili Twitter mesajlarını içeren değişkendir.
Duygu Durumu	Duygu analizi sonucunda her tweetin duygu durumunun olumlu, nötr ve olumsuz olarak verildiği değişkendir.
Beğeni Sayısı	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının her birinin almış olduğu beğeni (favori) miktarını gösteren değişkendir.
Rt Sayısı	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarının her birinin almış olduğu retweet miktarını gösteren değişkendir.
Kullanıcı Adı	Firmalar ile ilgili Twitter platformunda paylaşım yapan kullanıcıların kullanıcı adlarının verildiği sütundur.
Mention Adı	Kullanıcıların firmalar ile ilgili Twitter paylaşımlarında yer alan mentionların yer verildiği sütundur. Eğer tweet mention içermiyorsa bu sütunda bir ifade verilmemiş, içeriyorsa ise mention edilen kişinin/kişilerin kullanıcı adı verilmiştir.
Kaynak	Kullanıcıların firmalar ile ilgili tweetlerini hangi kaynağı kullanarak paylaştığının bilgisini veren değişkendir. “Android”, “iPhone”, “iPad”, “Web Client” ya da otomatik tweet atmak için kullanılan kaynak (özel hizmet veren sosyal medya danışmanlık şirketleri) bilgileri yer alır.

3.4.2.4 Modelleme ve Analiz

Uygulamanın bu bölümünde firmaların Twitter etkinliklerinin ölçümünün yapılması amaçlanmıştır. Literatürde sosyal medya etkinliğinin ölçümü ile ilgili kabul görmüş çalışma Hoffman ve Fodor (2010) tarafından yapılmıştır. Bu çalışmada her bir sosyal medya platformuna özel belirlenen ölçümlerle (metriklerle) firmaların sosyal medya etkinliğini ölçmek için bir sosyal medya etkinlik ölçüm tablosu geliştirmişlerdir. Twitter platformu için yer alan ölçümler Tablo 3.21’de verilmiştir.

Tablo 3.21: Farklı Amaçlar Doğrultusunda Twitter Etkinliğinin Ölçüm Kriterleri

Sosyal Medya Platformu	Marka Bilinirliği (Brand Awareness)	Marka İlginliği (Brand Engagement)	Ağızdan Ağıza İletişim (Word of Mouth)
Twitter	Marka Hakkındaki Tweet Sayısı Tweetin Değeri (-,+) Takipçi Sayısı	Takipçi Sayısı Yanıt Sayısı (@)	Rt Sayısı

Kaynak: Hoffman ve Fodor, 2010:44

Firma Twitter paylaşımlarının firmanın satışlarını nasıl arttırdığı ile ilgili geleneksel dar bir bakış açısıyla sorulan soruların ya da hedeflerin aksine Hoffman ve Fodor (2010), kullanıcıların kontrolü elinde bulundurduğu sosyal medya ortamında hedeflerin, müşteriye yatırım olarak görülebileceğini ve bu yönde ölçümlemenin gerçekleştirilebileceğini vurgulamıştır. Satışlar, ürün geliştirme ve pazar araştırması gibi amaçlar açık hedefler olmakla birlikte, bunların dışında, sosyal medyanın ayırt edici özelliklerinden faydalanılarak marka bilinirliği, marka bağlılığı ve ağızdan ağıza iletişim hedefleri ölçümlenebilir (Hoffman ve Fodor, 2010:42-43).

Rt sayısının önemi bir başka araştırmacı tarafından da dile getirilmiştir. Sterne (2010)'ye göre firma paylaşımlarının kullanıcılar tarafından retweet edilmesi, kullanıcıların bu tweetlere verdiği değer ile ilgili kesin bir ölçüdür. Bir firmanın hedef kitleyle paylaştığı bir içerik, onu iletmek (yaymak) isteyen başka kişilerce sahiplenilmektedir. Elektronik (sosyal) ağızdan ağıza iletişim, firma erişimini arttırmaktadır. Her retweet aktivitesinde bulunan kullanıcı, aslında farkında olmadan firmanın içeriğinin yayılmasına destek olmaktadır. Retweet ölçümünü takip etmekle birlikte buna firma takipçi sayısı, firma tweetlerini rt eden takipçi sayısı, firma tweetlerini rt eden kişilerin takipçi sayısı vb. kriterler de eklenerek ölçüm zenginleştirilebilir (Sterne, 2010:33-34).

Yukarıda değinilen literatürdeki bu ölçüm kriterleri göz önünde bulundurularak uygulamanın değerlendirme bölümünde ele alınan kriter açıklamaları aşağıda verilmiştir:

- **Firmalar Hakkındaki Tweet Sayısı:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının sayılarına karşılık gelmektedir.
- **Tweet Değeri:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının duygu durumudur.
- **Takipçi Kazanımı:** Firmaların Şubat ayı boyunca takipçi sayısındaki değişimi elde etmek amacı ile BoomSocial adlı Web siteden faydalanılarak veriler elde edilmiştir.
- **Yanıt Sayısı:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının kaçında firmadan bahsettiğinin (@firmaadı) sayısıdır. Başarı sıralamasında yer alacak kriter olduğundan dolayı yanıt tweetlerinden olumsuz duygu barındıran tweetler ayıklandıktan sonra, istatistikler hesaplanarak değerlendirme bölümünde verilmiştir.
- **Rt Sayısı:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının toplamda aldığı rt sayısıdır.
- **Beğeni Sayısı:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının toplamda aldığı beğeni sayısıdır. Beğeni sayısı bu çalışmada ölçüm kriterleri arasında değerlendirilmiştir.

Bu ölçüm kriterler dışında Şubat 2018 paylaşımları ile ilgili Tanımlayıcı istatistikler bölümünde verilen diğer istatistikler aşağıda verilmiştir:

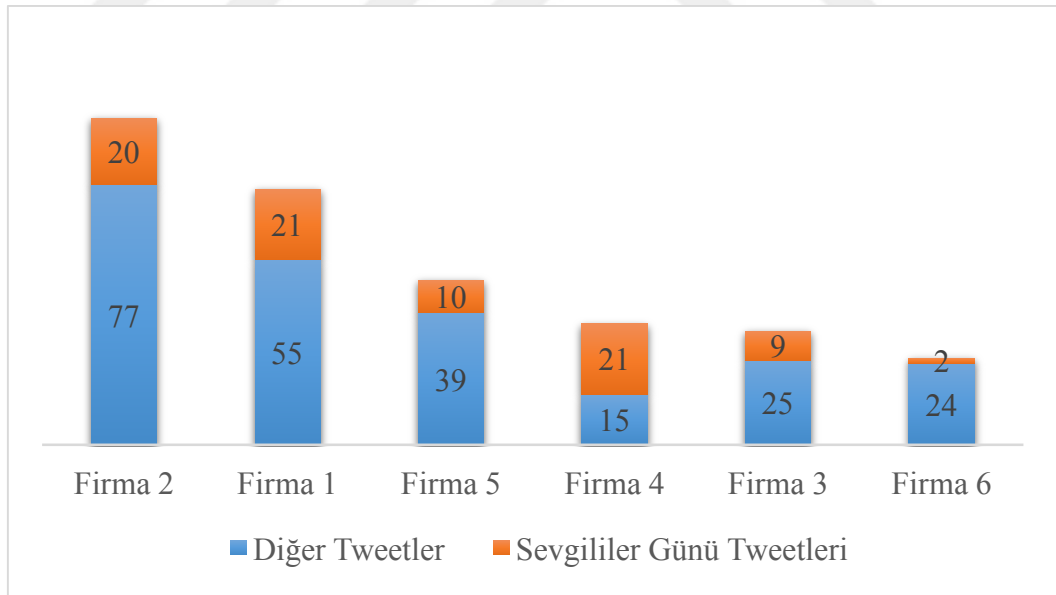
- **Firmaların Tweet Sayısı:** Firmaların Şubat ayı boyunca Twitter paylaşımlarının sayısıdır.
- **Firmaların Rt Sayısı:** Firmaların Şubat ayı boyunca Twitter paylaşımlarının toplamda aldığı rt sayısıdır.

- **Zaman Dilimi:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının hangi zaman diliminde atıldığının bilgisidir.
- **Kaynak:** Kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının hangi kaynaktan atıldığının bilgisidir.

3.4.2.4.1 Tanımlayıcı İstatistikler

Tanımlayıcı istatistikler bölümünde kullanıcıların Şubat ayı boyunca firmalar ile ilgili Twitter paylaşımlarının istatistikleri ile birlikte firmaların paylaşımlarının istatistiklerine de yer verilmiştir. İlk olarak firmaların 2018 Şubat ayı boyunca yapmış oldukları paylaşımlar incelenmiş ve istatistikleri Şekil 3.24’de verilmiştir.

Şekil 3.24: Şubat Ayı Boyunca Firmalar Tarafından Paylaşılan Tweetlerin Frekans Dağılımları

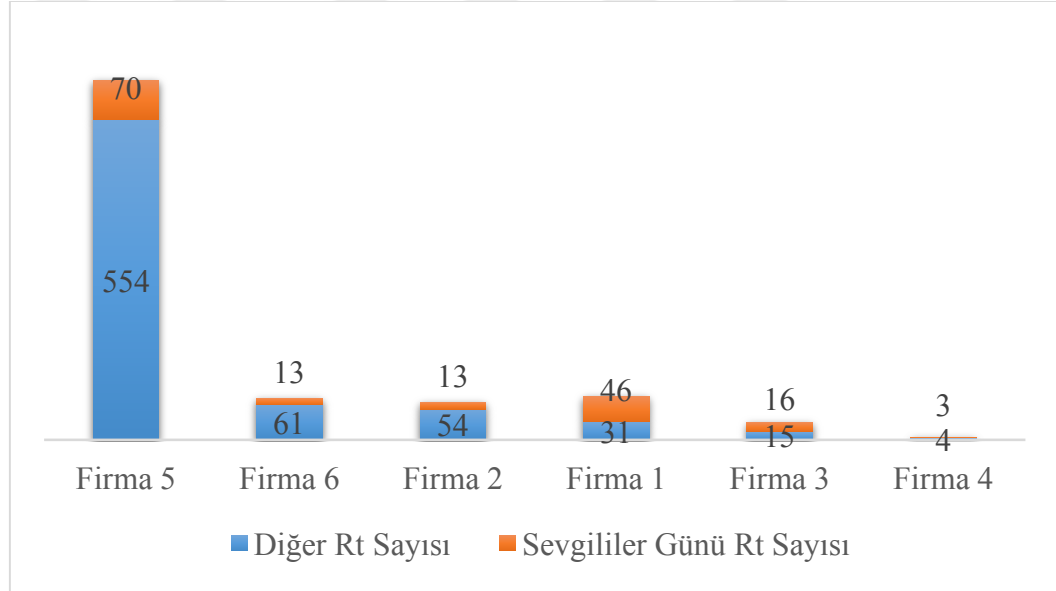


Şekil 3.24 incelendiğinde, Şubat ayı boyunca en fazla paylaşım yapan iki firmanın da kozmetik sektöründe faaliyet gösterdiği görülmektedir. Elektronik sektöründe faaliyet gösteren Firma 5 üçüncü sırada yer alırken, rakibi Firma 6 ise toplamda yalnızca 26 tweet paylaşarak tweet paylaşım frekansı sıralamasında sonuncu olmuştur. Sevgililer

günü ile ilgili paylaşımlara özel olarak bakıldığında ise bu sıralama değişmektedir. Firma 1 ve Firma 4 21 adet, Firma 2 ise 20 adet sevgililer günü ile ilgili tweet paylaşmışlardır.

Firmaların almış oldukları rt sayıları dikkate alındığında ise elde edilen şekil aşağıda verilmiştir. Firmaların paylaşımları sevgililer günü ile ilgili ve diğer olarak ayrılmış ve bu iki kümeye ait tweetlerin elde etmiş oldukları rt sayıları ayrı olarak verilmiştir.

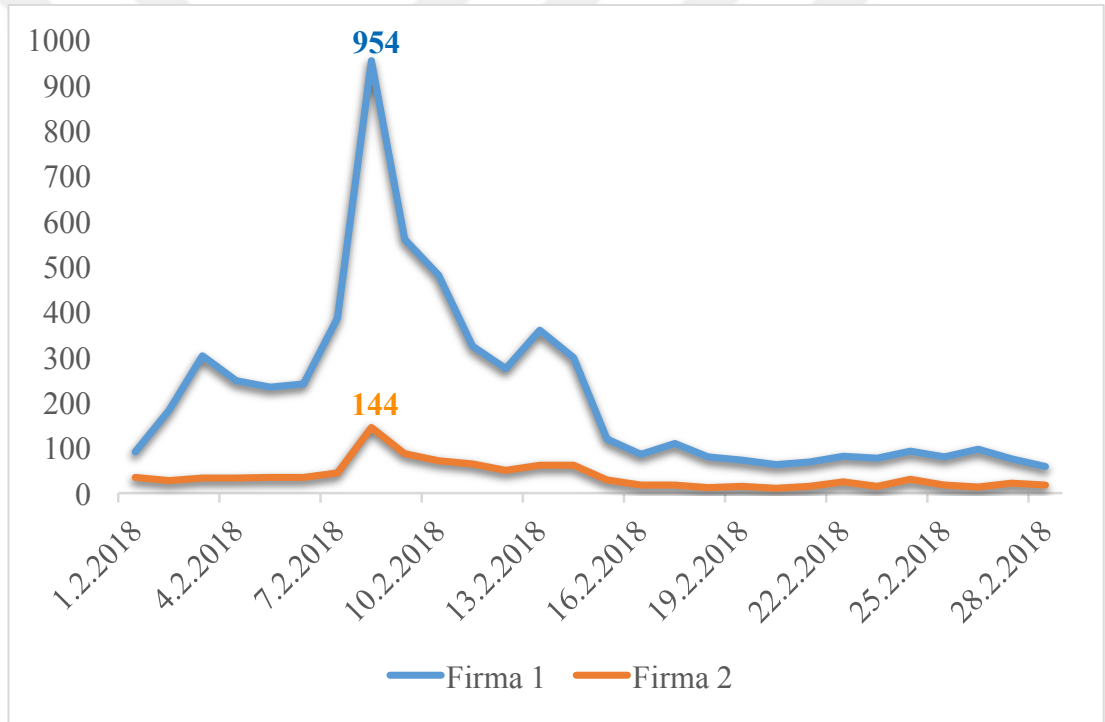
Şekil 3.25: Şubat Ayı Boyunca Firmalar Tarafından Paylaşılan Tweetlerin Almış Oldukları Rt Frekans Dağılımları



Şekil 3.25’de görüldüğü üzere Şubat ayı boyunca paylaşımları en fazla rt edilen firmanın Firma 5 olduğu görülmektedir. İkinci sırada yer alan firmanın ise yine aynı sektörde faaliyet gösteren Firma 6’dır. En az rt alan firmalar pazaryeri sektöründe faaliyet gösteren firmalar olmuştur. Sevgililer günü ile ilgili paylaşımlarının almış olduğu rt sayısı, diğer içerik paylaşımlarının almış olduğu rt sayısından fazla olan firma ise kozmetik sektöründe faaliyet gösteren Firma 1’dir.

Şubat ayı boyunca kullanıcılar tarafından firmalar ile ilgili Twitter’da paylaşılan tweetlerin sayıları günlük olarak hesaplanarak yoğunlukların ay boyunca nasıl bir dağılım gösterdiği sektör temelinde ele alınarak görselleştirilmiştir. Şekil 3.26’da görüldüğü üzere, kozmetik sektöründe faaliyet gösteren her iki firma ilgili kullanıcılar tarafından yapılan paylaşımların sayısında, 7 Şubat tarihinden 8 Şubat tarihine geçişte büyük bir sıçrama gerçekleşmiştir.

Şekil 3.26: Kullanıcılar Tarafından Kozmetik Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)



Firmaların Twitter paylaşımlarına bakıldığında, 7 Şubat tarihinde akşam saatlerinde Firma 1’in ve Firma 2’nin indirimle ilgili bilgilendirme tweeti attığı görülmektedir. Bu tweetler firma bilgisi, tarih bilgisi ve saat bilgisi parantez içinde verilerek aşağıda paylaşılmıştır:

- Sevgililer Günü Aşkına Bu İndirimi Kaçırma! En güzel indirim fırsatları, 14 Şubat’a kadar mağazalarımızda sizleri bekliyor! Detaylar için:... (Firma 1, 07.02.2018, 20:00)

- 🎉Sevgililer Günü İndirimi Başladı! 8-14 Şubat tarihleri arasında mağazalarımız ve <http://firma2.com.tr> 'deki %50'ye varan indirimleri kaçırmayın! Beklemeden alışverişe başlamak için ise 📌... (Firma 2, 07.02.2018, 21:45)

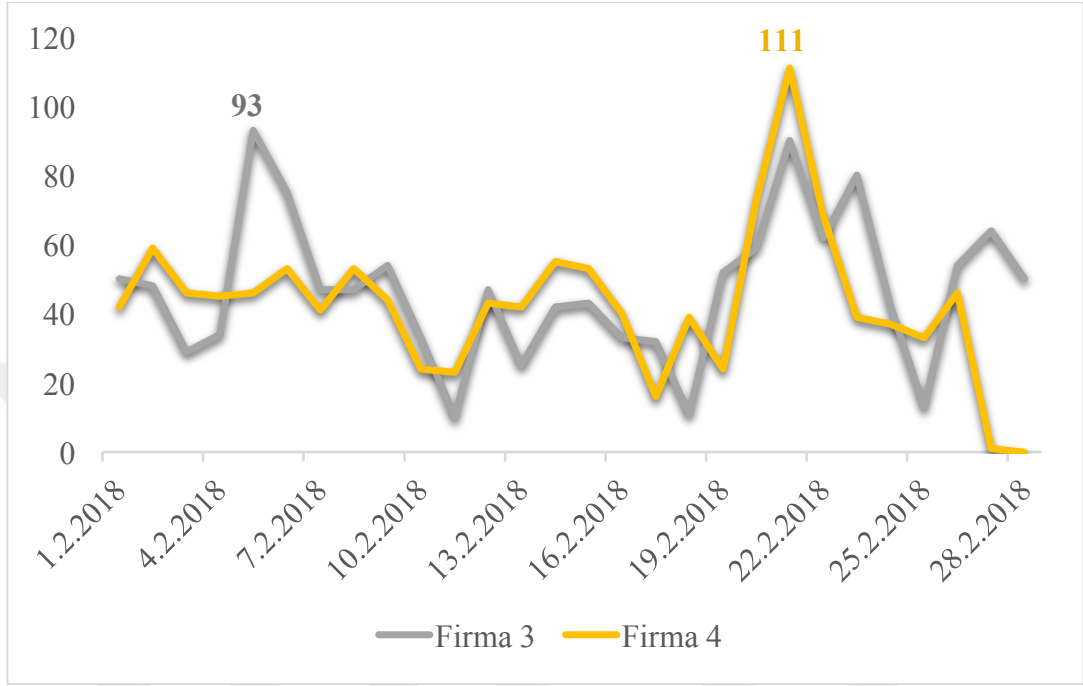
Firmaların paylaşmış oldukları bu tweetler incelendiğinde, her iki firmanın da birbirlerine yakın söylemlerle indirim mesajı attıkları görülmektedir. Bu indirim tweetlerinden sonra kullanıcıların firmalar ile ilgili tweet paylaşımlarının artması, firma indirimlerinin yankı bulduğuna işaret olarak görülebilir. Daha detaylı sonuca ulaşmak adına kullanıcıların bu firmalar ile ilgili atmış oldukları tweetleri içeren veri setinde, sevgililer günü ile ilgili paylaşımlar anahtar kelimelerle aratılarak 7 ve 8 Şubat tarihlerinde atılan tweetlerin yükselişe geçip geçmediği incelenmiştir. Kullanıcıların 8 Şubat tarihinde 14 Şubat ile ilgili en yüksek seviyede paylaşımlarda bulunduğu görülmüştür.

Firma yoğunlukları kıyaslandığında ise Firma 1 ile ilgili paylaşımların yoğunluğunun Firma 2'den Şubat ayı boyunca fazla olduğu görülmektedir. Firmaların Şubat ayı Twitter paylaşımlarından oluşan veri seti incelendiğinde Sevgililer Günü ile ilgili ürün öneri ve indirim tweetlerine Firma 2'nin 7 Şubat, Firma 1'in 2 Şubat tarihinden 14 Şubata kadar paylaşım yaptığına ulaşılmıştır.

14 Şubat kampanya döneminden sonra ise, beklenildiği üzere, kullanıcıların firmaları ile ilgili paylaşımlarının sabit durumda ve minimum düzeyde olduğu görülmektedir. Kampanya döneminde bu iki firmanın da kullanıcı etkileşimini içerik yoğunluğunu arttırarak başarıyla gerçekleştirdiği söylenebilir.

Kozmetik sektörünün aksine pazaryeri sektöründe faaliyet gösteren Firma 3 ve Firma 4 ile ilgili kullanıcılar tarafından yapılan Twitter paylaşımlarının günlük frekansları incelendiğinde, Şekil 3.27'de görüldüğü üzere, Sevgililer Günü döneminde ortak bir içerik artışına rastlanmamıştır.

Şekil 3.27: Kullanıcılar Tarafından Pazaryeri Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)

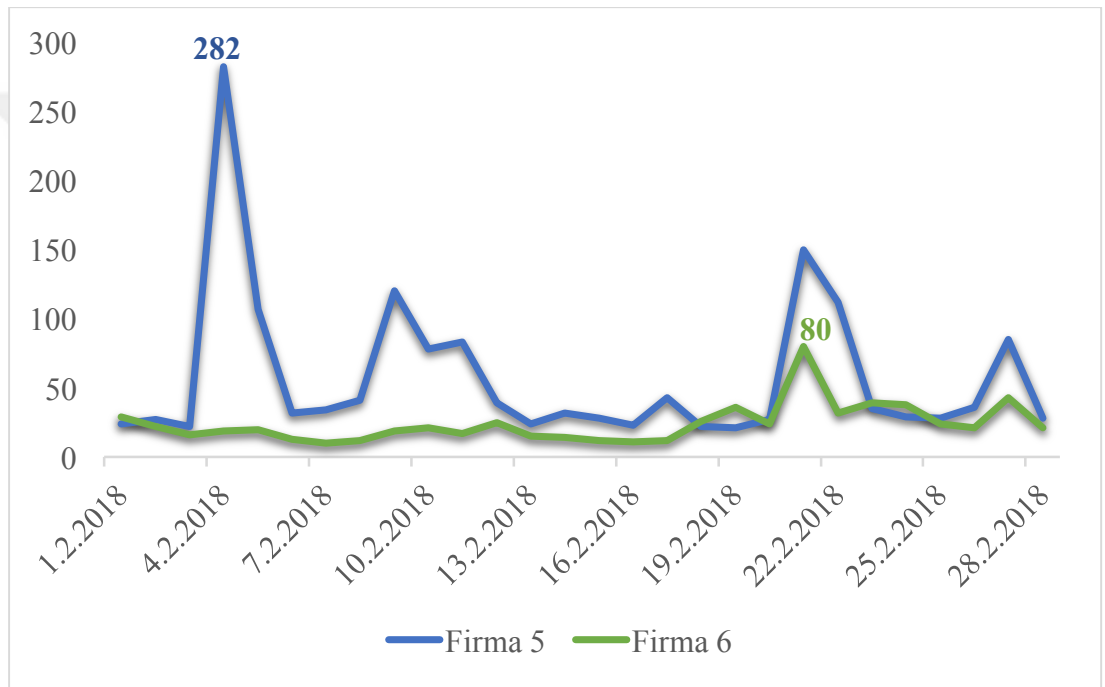


Firma 3 ile ilgili en yoğun paylaşımın yapıldığı 5 Şubat tarihi incelendiğinde, bu tarihte Firma 3'ün Sevgililer Günü kapsamında düzenleyeceği üç farklı etkinliğin (canlı yayın) bilgilendirme tweetinin atıldığı görülmüştür. Kullanıcıların bu firma hakkında paylaştıkları tweetler manuel olarak incelendiğinde, üç etkinlikten yalnızca bir tanesinin düşük bir oranda içerik paylaşımını tetiklediğine ulaşılmıştır.

Firma 4 ise rakibinden farklı olarak 1 Şubat tarihinde ilk Sevgililer Günü indirim tweetini paylaştığı, ancak paylaştığı ilk ve diğer indirim tweetlerinin kullanıcıları içerik üretmeye teşvik etmede zayıf kaldığı görülmüştür. Firma 3 ile ilgili en yoğun paylaşımın yapıldığı 21 Şubat tarihi incelendiğinde ise üretilen içeriklerin firmanın paylaşımlarından bağımsız (ilgisiz), satıcıların/kullanıcıların paylaştığı kupon/indirim ve bilgilendirme mesajlarından oluştuğu görülmüştür.

Elektronik sektöründe faaliyet gösteren firmalarla ilgili yapılan paylaşımların zamana bağlı frekansları Şekil 3.28’de incelendiğinde, Firma 6 ile ilgili paylaşımların, 22 Şubat tarihi dışında, minimum düzeyde olduğu görülmektedir. Firma 5 ile ilgili ise 5 Şubat tarihinde Şubat ayı boyunca en fazla paylaşımın yapıldığı görülmektedir.

Şekil 3.28: Kullanıcılar Tarafından Elektronik Firmaları ile İlgili Paylaşılan Tweetlerin Günlük Frekans Dağılımları (1 - 28 Şubat 2018)



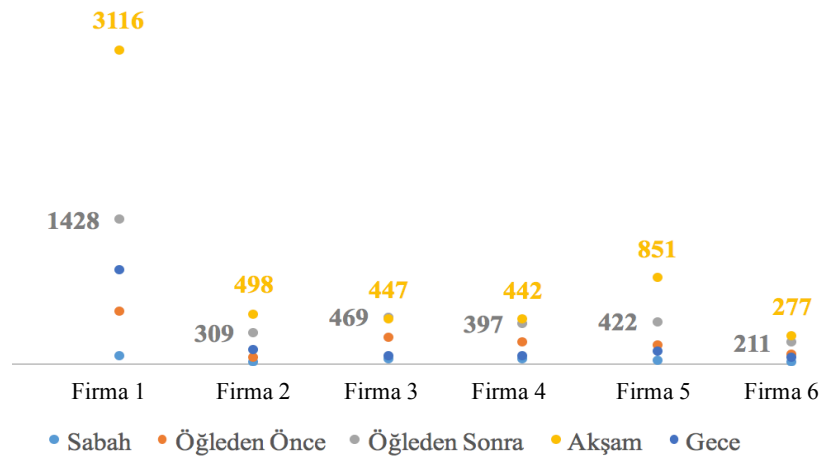
Firmaların Twitter paylaşımlarına bakıldığında, ilk paylaşılmış indirim/etkinlik tweetleri firma bilgisi, tarih bilgisi ve saat bilgisi parantez içinde verilerek aşağıda paylaşılmıştır:

- Sevgililer günü için hala bir hediye fikrin yok mu? O halde hediye bulucumuza gel, soruları cevapla, sevgilin için en uygun hediye anında bul! #AşktaKazananlar Firma 5’de! (Firma 5, 06.02.2018, 15:28)
- Sevgilinize hediye alırken arada kalırsanız aralarında kalmaz. Firma 6’ya gelin, aramızda kalsın. 😎 (Firma 6, 09.02.2018, 12:00)

Firma 5'e ve Firma 6'ya ait Sevgililer Günü tweetleri incelendiğinde iki firmanın da farklı üslup ve yollarla etkileşim yaratma çabalarının olduğu görülmektedir. İçerik yoğunluklarının arttığı günler veri setinde incelendiğinde 14 Şubat ile ilgili firmaların paylaştıkları içeriklerin, kullanıcıları içerik üretmeye teşvik etmediğine ulaşılmıştır. Firma 5 ile ilgili en yoğun paylaşımın yapıldığı 4 Şubat tarihi incelendiğinde, bir müzik grubu hayranlarının organize şekilde firmaya müzik grubunun albümlerinin satılmasını istediklerini belirten tweetler attığı gözlemlenmiştir. Firma 6 ile ilgili en yoğun içeriğin üretildiği 21 Şubat tarihi incelendiğinde, bu tarihte Firma 6'nın düzenlemiş olduğu basın toplantısı ile ilgili kullanıcıların ve haber kaynaklarının üretmiş olduğu içeriklerin yoğunlukta olduğu saptanmıştır. Kullanıcıların Firma 5 ve Firma 6 ile ilgili atmış oldukları tweetler ele alındığında, firmaların yapmış oldukları etkinlik/indirim paylaşımlarının kullanıcılar tarafından üretilen içeriklerde etkili olmadığı görülmüştür.

Bu bölümde son olarak kullanıcıların firmalar ile ilgili paylaşımlarını hangi zaman dilimlerinde ve hangi aracı kullanarak (kaynak) yaptıkları saptanmıştır. Bulgulara göre tüm firmalarla ilgili yapılan paylaşımların büyük çoğunluğu öğleden önce ve akşam saatlerinde yapılmıştır.

Şekil 3.29: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Zaman Dilimlerine Göre Frekans Dağılımları



Kullanıcıların Twitter paylaşımı yaparken kullandıkları araçlar incelendiğinde ise kozmetik ve elektronik sektöründe faaliyet gösteren firmalarla ilgili paylaşımların büyük çoğunluğunun Android cihazlar tarafından yapıldığı görülmüştür. Pazaryeri firmalarında ise daha çok Web ve Diğer araçlar yolu ile paylaşımın yapıldığı saptanmıştır. Ayrıntılı istatistikler Tablo 3.22’de verilmiştir.

Tablo 3.22: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Kaynak Bilgisine Göre Frekans Dağılımları

Kaynak	Firma 1	Firma 2	Firma 3	Firma 4	Firma 5	Firma 6
iPhone	2438	406	292	204	298	182
Android	3112	445	296	150	559	246
Web	230	63	393	272	303	118
Diğer	310	120	338	570	452	125

3.4.2.4.2 Duygu Analizi

Uygulamanın bu bölümünde ele alınan kullanıcıların firmalar hakkında yapmış oldukları Twitter paylaşımları incelendiğinde, oldukça fazla ironi tespit edildiği için sözlük temelli duygu durum belirlemesine başvurulmamıştır. “Kargom gelmedi, sizi tebrik ediyorum, harikasınız” tweeti ele alınacak olduğunda, ağırlıklı olarak olumlu duygu belirten kelimeler kullanılmasına karşın tweet olumsuz bir duygu bildirmektedir. Bu sebeple öncelikle makine öğrenmesi algoritmalarına başvurulmuş, ancak veri setinin büyük bir oranının nötr duygulardan oluşmasından kaynaklı, olumlu ve olumsuz tweetleri saptamada makine öğrenmesi algoritmalarının başarı oranlarının düşük çıkması sebebi ile bu yöntem de uygun bulunmayarak analize eklenmemiştir.

Yukarıda değinilen tüm bu sebeplerden dolayı kullanıcıların firmalar ile ilgili Twitter paylaşımlarının duygu durumlarının belirlenmesi manuel olarak gerçekleştirilmiştir. Tüm veri seti üç araştırmacı tarafından olumlu, nötr ve olumsuz olmak üzere etiketlenmiştir. Araştırmacılar tweetlerin duygu durumlarını belirlerken tweet atılan firma ile ilgili cümlenin genel duygusu üzerinden etiketlemeyi gerçekleştirmiştir.

Puanlayıcılar arası güvenilirliği test etmek adına literatürde değinilen Krippendorff Alfa Katsayısı en kapsamlı ve üç araştırmacı arasındaki güvenilirliği değerlendirebilen bir güvenilirlik ölçütü olduğu için seçilmiştir. Bu katsayının hesaplanması “ReCal OIR” adlı Web tabanlı puanlayıcılar arası güvenilirliği hesaplama sitesi, SPSS’de yazılmış ve Krippendorff tarafından onaylanan makro ve R’da “irr” adlı paket yardımı ile hesaplanabilmektedir. Bu üç hesaplama tekniği Freelon (2013) tarafından yapılan çalışmada karşılaştırılmış ve “ReCal OIR” ve SPSS bulgularının hemen hemen birbirleri ile aynı sonuçları verdiği, ancak R’da çalıştırılan irr paketinin bu iki yöntemin bulgularından farklı olduğuna ulaşılmıştır (Freelon, 2013:14). Bu sebeple uygulamada “ReCal OIR” adlı Web siteden faydalanılarak hesaplamalar gerçekleştirilmiştir. Her firma ile ilgili yapılan paylaşımların bulunduğu altı veri seti için ayrı olarak hesaplanan Alpha katsayıları Tablo 3.23’de verilmiştir.

Tablo 3.23: Araştırmacılar Tarafından Gerçekleştirilen Duygu Durum Kodlamalarının Krippendorff Alfa Katsayısıyla Hesaplanan Güvenilirlik Değerleri

Firma	Alfa Katsayısı Değeri (α)
Firma 1	0,982
Firma 2	0,932
Firma 3	0,988
Firma 4	0,958
Firma 5	0,948
Firma 6	0,964

Üç araştırmacı tarafından kodlanan veri setlerinin güvenilirlik değerleri Tablo 3.23’de verildiği üzere 0,90 üzeri çıkmıştır. Alpha değerleri göz önünde bulundurularak puanlayıcılar arası güvenirlkte yüksek uyumun olduğu söylenebilir.

Puanlayıcılar arası güvenilirlik analizinden sonra nihai duygu durumlarının belirlenmesi için her firmaya ait veri setindeki kodlamalar incelenmiştir. Veri setindeki hiçbir tweet için araştırmacılar tarafından verilen üç ayrı cevaba rastlanmamıştır. Bu sebeple, etiketlenme sonucunda farklılıklar bulunan tweetler için en fazla tekrar eden cevap

seçilmiştir. Örneğin nötr, nötr ve pozitif olarak kodlanan bir tweet, etiketlerin modu alınarak nötr olarak nihai şekilde etiketlenmiştir. EK 2, EK 3 ve EK 4’de firmalarla ilgili paylaşılan tweetlerin duygu durumlarına göre örnek tweetler verilmiştir.

Firmalar ile ilgili kullanıcılar tarafından paylaşılan tweetlerin nihai duygu durumlarının belirlenmesinin ardından duygu durum yüzdelerinin firmalara göre dağılımı Tablo 3.24’de verilmiştir.

Tablo 3.24: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Duygu Durumlarının Firmalara Göre Frekans Dağılım Yüzdeleri

Firma	Tweet Değeri (Olumlu)	Tweet Değeri (Nötr)	Tweet Değeri (Olumsuz)
Firma 1	%31	%54	%14
Firma 2	%30	%57	%13
Firma 3	%4	%63	%32
Firma 4	%6	%87	%7
Firma 5	%1	%90	%9
Firma 6	%4	%62	%35

Tablo 3.24 incelendiğinde tüm firmalar için, paylaşılan tweetlerinin yarısından fazlasının duygu durumlarının nötr olduğu anlaşılmaktadır. Bununla birlikte olumlu tweet paylaşım yüzdeleri incelendiğinde, kozmetik sektöründe faaliyet gösteren firmalar ile ilgili paylaşılan tweet yüzdelerinin diğer sektörlere göre oldukça fazla oranda olduğu görülmektedir. Pazaryeri ve elektronik sektöründe faaliyet gösteren firmalar ile ilgili paylaşılan olumsuz tweetlerin yüzdelerinin olumlu tweet yüzdelerinden fazla olduğu bir başka dikkat çekici bulgudur.

Rakip firmalar karşılaştırmalı olarak ele alındığında kozmetik sektöründeki Firma 1 ve Firma 2 ile ilgili atılan tweetlerin duygu durum yüzdelerinin birbirlerine oldukça benzer olduğu görülmektedir. Pazaryeri sektöründeki Firma 3 ve Firma 4 ile ilgili atılan tweetlerin duygu durum yüzdeleri incelendiğinde ise Firma 3 ile ilgili daha fazla olumsuz tweet paylaşılmıştır. Elektronik sektöründeki Firma 5 ve Firma 6 için de

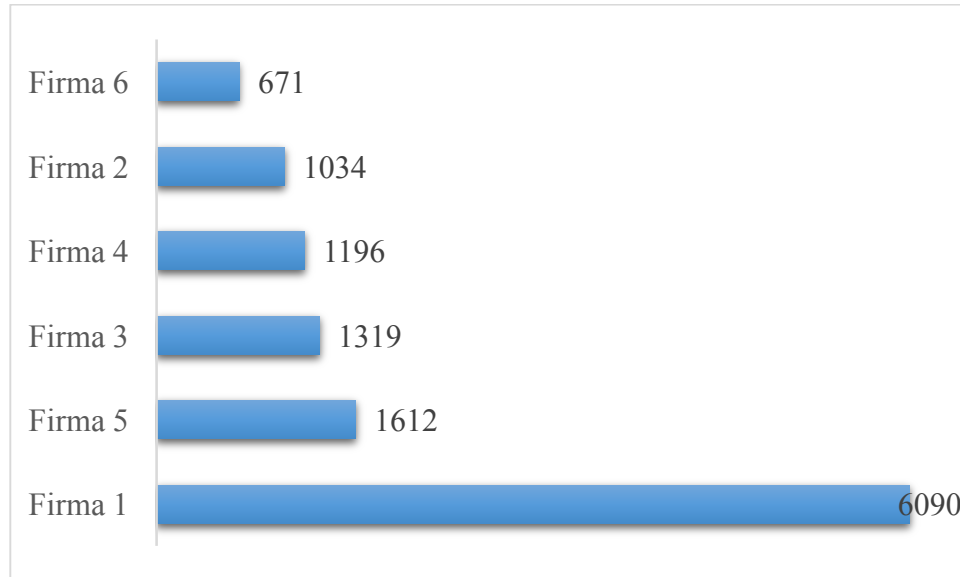
benzer bir durum söz konusudur; Firma 6 ile ilgili daha fazla olumsuz tweetin paylaşıldığı görülmektedir.

3.4.2.5 Değerlendirme

Bu bölümde Twitter etkinliğinin ölçümü ile ilgili değerlendirmeler Modelleme ve Analiz bölümünde açıklanan ölçüm kriterleri ele alınarak gerçekleştirilmiştir.

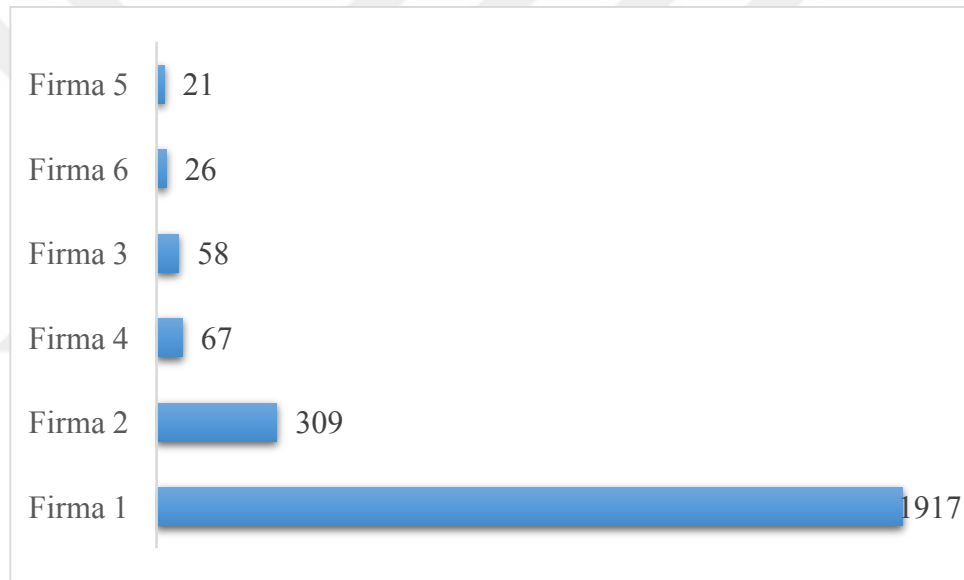
Firmalar ile ilgili paylaşılan tweetlerin frekansları Şekil 3.30'daki gibi görselleştirilmiştir. En fazla konuşulma sayısına sahip firma, Firma 1 olmuştur. Kozmetik sektöründe Firma 1'in rakibi olan Firma 2 ile ilgili olarak, Firma 1'e kıyasla oldukça az içerik üretildiği ve Firma 2'nin sıralamada beşinci olduğu görülmektedir.

Şekil 3.30: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Firmalara Göre Frekans Dağılımları



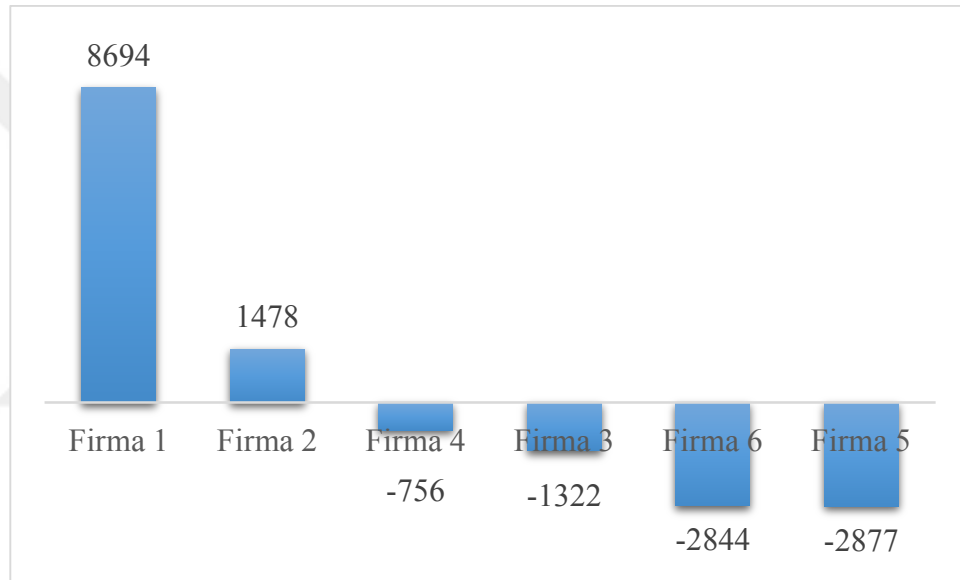
Tweet değeri olarak olumlu tweetler ele alınarak bir sıralama gerçekleştirilmiştir. Buna göre sıralamada birinci olan firma, tweet sayısı ile doğru orantılı olarak Firma 1 olmuştur. Firma 1'i aynı sektörde yer alan Firma 2 takip etmiştir. Elektronik sektöründe yer alan firmalarla ilgili ise minimum seviyede olumlu içerik üretildiği görülmektedir.

Şekil 3.31: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Olumlu Tweetlerin Firmalara Göre Frekans Dağılımları



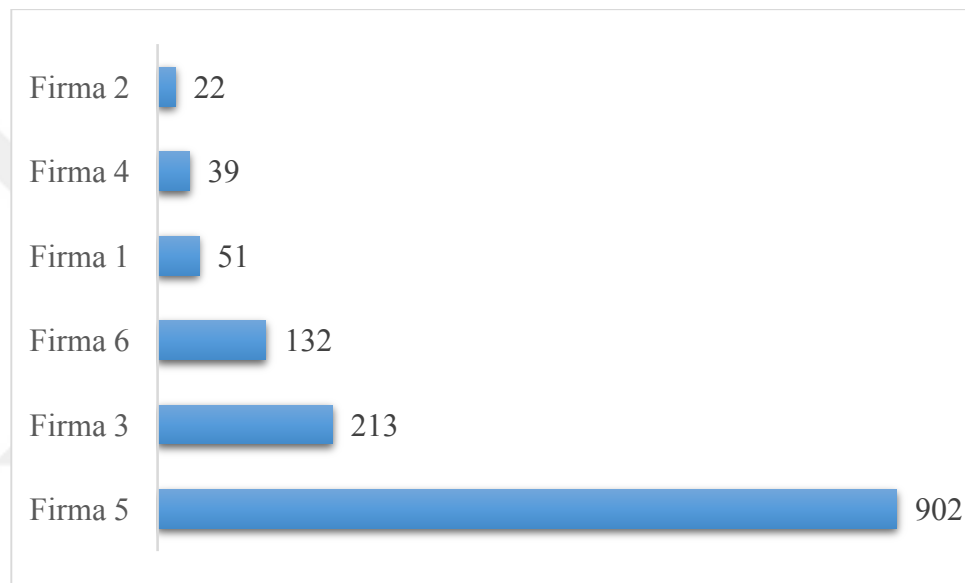
Takipçi kazanımları incelendiğinde ise firmalar arasında takipçi kaybeden firmaların olduğu görülmektedir. Pazaryeri ve elektronik sektörlerindeki firmaların Şubat ayı boyunca takipçi kaybettikleri, yalnızca kozmetik firmalarının takipçi kazandıklarına ulaşılmıştır. En fazla takipçi kazanımı Firma 1 tarafından elde edilirken, en çok takipçi kaybeden iki firma da elektronik sektöründe faaliyet göstermektedir.

Şekil 3.32: Firmaların Şubat Ayı Takipçi Kazanımları



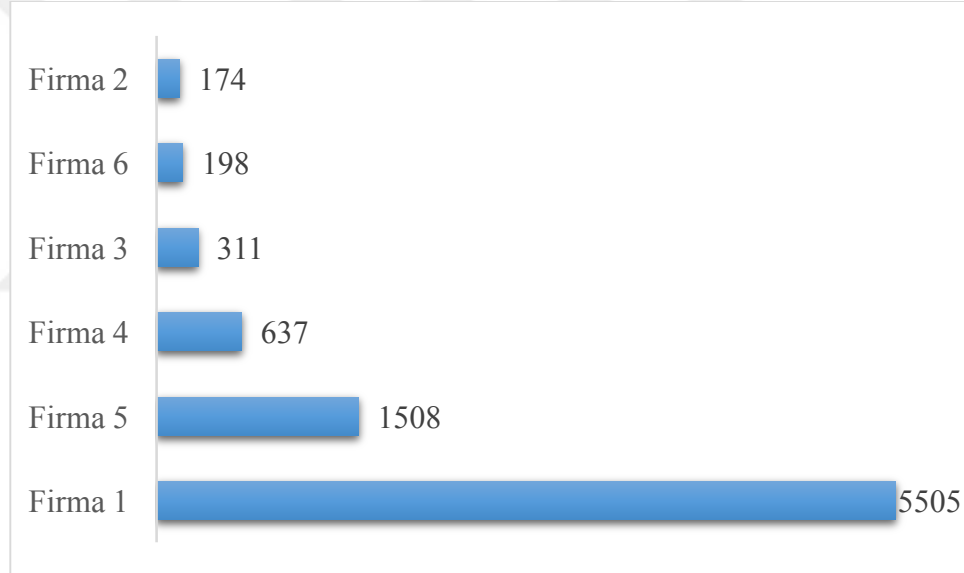
Yanıt tweet sayısı ele alındığında ise en fazla Firma 5'e yönelik paylaşımların yapıldığı görülmektedir. Bu firmayı sırası ile Firma 3, Firma 6, Firma 4, Firma 1 ve Firma 2 izlemiştir.

Şekil 3.33: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Yanıt Tweetlerinin Firmalara Göre Frekans Dağılımları



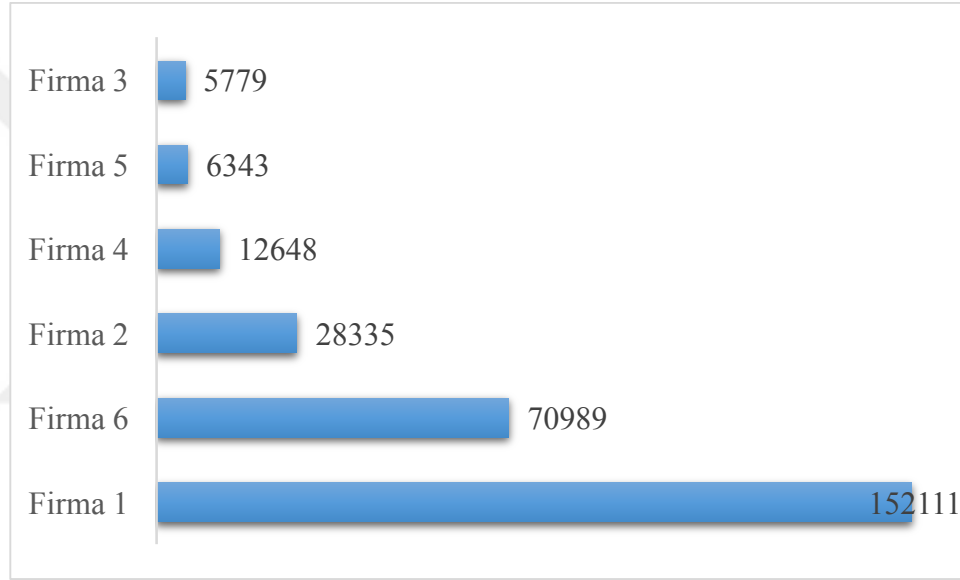
En fazla rt edilen paylaşımların ise kozmetik sektöründeki Firma 1 ile ilgili olduğu Şekil 3.34’de görülmektedir. Aynı sektörde yer almalarına rağmen Firma 2 ile ilgili paylaşılan içerikler en az sayıda rt almıştır. Firma 1’i elektronik sektöründen Firma 5 ile ilgili ilgili paylaşılan tweetler takip etmektedir. Firma 5’in rakibi Firma 6 ile ilgili paylaşımların ise çok daha az bir oranda rt edildiği saptanmıştır. En fazla içeriğin paylaşıldığı ilk iki firma, elde edilen rt miktarı sıralamasında da başlarda yer almıştır.

Şekil 3.34: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Toplamda Aldığı Rt Frekanslarının Firmalara Göre Dağılımları



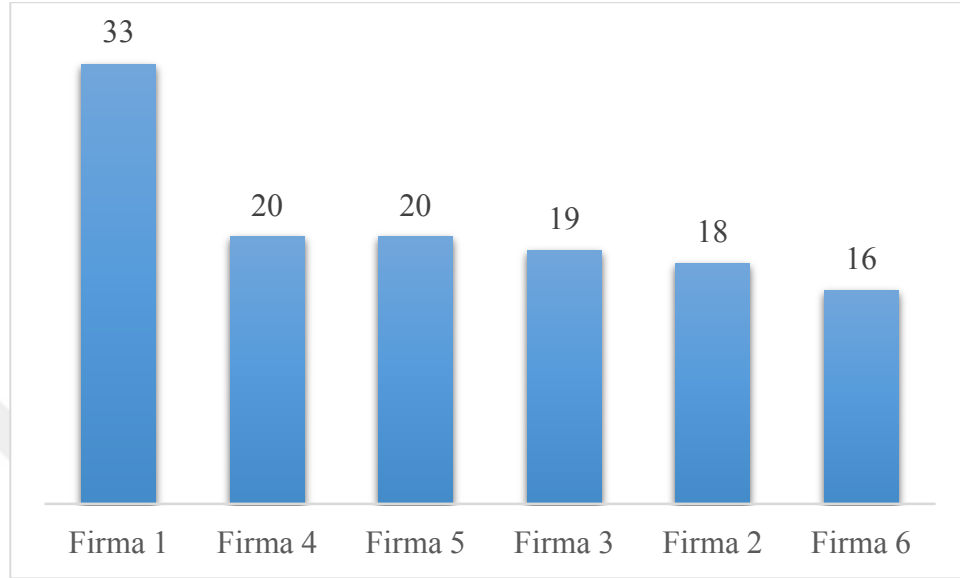
Ölçüm kriterlerine bu çalışma için eklenmesi uygun bulunan beğeni sayısı ele alındığında ise yine Firma 1 ile ilgili paylaşılan tweetlerin, toplamda en çok beğeni aldığı gözlenmiştir. Şubat ayı boyunca hakkında en az içerik paylaşımı yapılmasına karşın Firma 6, beğeni miktarı sıralamasında ikinci sırada yer almıştır.

Şekil 3.35: Kullanıcılar Tarafından Firmalar ile İlgili Paylaşılan Tweetlerin Toplamda Aldığı Beğenilerin Firmalara Göre Frekans Dağılımları



Tüm bu altı kriter ele alınarak gerçekleştirilen Twitter etkinliği ölçümünde sıralama yapabilmek adına, tüm firmaların her kriter için sahip olduğu performans sırası puanlanarak elde ettiği olduğu toplam puan hesaplanmıştır. Puanlama yaparken 1 en düşük puan ve 6 en yüksek puan olmak üzere her kriterdeki performans sırasına göre 1'den 6'ya kadar puan verilmiştir. Firmaların altı kriter için almış olduğu puanlar toplanarak nihai puan elde edilmiştir. Nihai puana göre yapılan Twitter etkinliği sıralaması Şekil 3.36'de verilmiştir.

Şekil 3.36: Firma Twitter Etkinliklerinin Kriterlere Göre Toplam Performans Sıralaması



Şekil 3.36'ya göre Firma 1, neredeyse tüm kriterlerde en yüksek oranı elde ederek, Şubat ayı boyunca en çok etkileşime sahip firma olmuştur. Bu çalışmada ele alınan kriterler temelinde Firma 1'in, rakibi Firma 2'ye kıyasla Twitter'da çok daha fazla konuşulduğu ve daha fazla etkileşim aldığı söylenebilir. Pazaryeri sektöründe faaliyet gösteren Firma 3 ve Firma 4 ele alındığında ise birbirlerine yakın puan aldıkları görülmektedir. Kriter sıralamasında Firma 4'ün rakibi Firma 3'e göre daha üst sırada yer aldığı göz önünde bulundurularak Twitter'ı daha etkin bir şekilde kullandığı söylenebilir. Elektronik sektörü firmaları Firma 5 ve Firma 6 değerlendirildiğinde ise Firma 5'in rakibine kıyasla, Şubat ayı boyunca Twitter'da daha etkin olduğu sonucuna ulaşılmıştır.

SONUÇ

Bu tez çalışmasında, farklı disiplinlerden araştırmacıların dikkatini çekmeye başlaması, güncel ve gelişmekte olan bir alan olması ve özellikle Türkçe literatürde yapılan çalışmaların yetersiz oluşu sebebi ile sosyal medya madenciliği ile ilgili bir uygulama gerçekleştirilmiştir. Sosyal medya platformları arasından Twitter seçilerek, örneklem dâhilinde yer alan sektörlerdeki (kozmetik, pazaryeri ve elektronik) rakip firmalar ile ilgili kullanıcılar tarafından paylaşılan tweetler ve firmaların kendi paylaşımları ele alınarak analizler gerçekleştirilmiştir.

Uygulamanın birinci bölümünde öncelikle, sosyal medya madenciliği adımları izlenerek firmaların 2017 yılına ait Twitter verileri elde edilmiş ve veriler analize hazır hale getirilmiştir. Firmaların Twitter istatistikleri hesaplanarak Tanımlayıcı İstatistikler bölümünde verilmiştir. Firmaların Twitter istatistiklerinin firmalara göre farklılaştığı, Tanımlayıcı İstatistikler bölümünde gözlemlenmiş, anlamlı farklılıkların olup olmadığı Ki-Kare Analizi uygulanarak elde edilmiştir. Anlamlı farklılıklar gözlemlenen değişkenler arasındaki uyum ise Uygunluk Analizi gerçekleştirilerek yorumlanmıştır. Firmaların Twitter'da ne tür içerikler paylaştıklarını ve hangi tür içeriklerin en çok etkileşimi gerçekleştirdiğini analiz etmek için ise İçerik Analizi uygulanmıştır.

Ki-Kare Analizi bulgularına göre, firmaların paylaşmış oldukları tweetlerle ilgili değişkenlerin; tweetlerin türlerinin (özgün, mention ve rt tweet), paylaşım günlerinin (Pazartesi, Salı, Çarşamba, Perşembe, Cuma, Cumartesi ve Pazar) ve zaman dilimlerinin (Öğleden Önce, Öğleden Sonra ve Akşam), firmalara göre farklılık gösterdiği saptanmıştır.

Uygunluk Analizi bölümünde, kozmetik sektöründe faaliyet gösteren Firma 2 dışında tüm firmaların tweet türlerine olan konum uzaklığı değerlendirildiğinde özgün tweet paylaşımına, Firma 2'nin ise mention tweet paylaşımına yakın konumlandığı

görülmüştür. Firma paylaşımlarının günleri ile firmalar arasındaki uyum ele alındığında ise sektör temelli ortak bir paylaşım günü belirlenememiş, her firmanın farklı günlerde Twitter’da paylaşım yaptığına ulaşılmıştır. Tweet paylaşımlarının zaman dilimleri ele alındığında, yalnızca kozmetik sektöründe faaliyet gösteren firmaların aynı zaman diliminde (öğleden önce) daha fazla paylaşım yaptıkları saptanmıştır.

İçerik Analizi bölümünde firmaların tweet paylaşımlarının kategorilerini belirlemek üzere, tweetlere bir dizi metin madenciliği ön işleme metodu uygulanmıştır. Elde edilen kelimeler ve frekansları doğrultusunda belirlenen altı küme dâhilinde firmaların Twitter paylaşımları etiketlenmiştir. Firmaların Twitter paylaşımlarının kategorileri; indirim/fırsat mesajlarından oluşan Özel Teklif, kullanıcılara özel düzenlenen yarışmalar/etkinlikler ile ilgili mesajlardan oluşan Yarışma&Etkinlik, ürünle ilgili bilgilendirmelerden oluşan Ürün, bayram, özel gün vb. gibi kutlama ve dilek mesajlarından oluşan Sosyal, kullanıcılara yönelik bilgilendirme/yanıt mesajlarından oluşan Destek&Geri Bildirim ve firmalara özgü hashtag vb. içerikler paylaşarak kullanıcıların içerik üretmeye teşvik edildiği Özel Etkileşim kategorileri olarak belirlenmiştir. Altı firmanın tweetleri herhangi bir sınıflandırmaya tabi tutulmadan ele alındığında toplamda en fazla Özel Teklif kategorisinde paylaşım yapıldığına ulaşılmıştır. Bu kategoriyi sırasıyla Destek&Geri Bildirim, Ürün, Özel Etkileşim ve Yarışma&Etkinlik izlemiş, en az oranda içeriğin üretildiği küme Sosyal olmuştur.

Her bir firmanın paylaşımlarının kümelere göre yüzdeleri hesaplanmış ve en çok paylaşım yapılan tweet kümesinin pazaryeri sektöründeki rakip firmalara göre aynı, kozmetik ve elektronik sektörü firmalarına göre ise farklı olduğu saptanmıştır. Bulgulara göre kozmetik sektöründe faaliyet gösteren Firma 2’nin Destek&Geri Bildirim, Firma 1’in Özel Teklif kategorisinde daha fazla oranda paylaşım yaptığı belirlenmiştir. Pazaryeri sektöründe faaliyet gösteren Firma 3 ve Firma 4 ise Özel Teklif kategorisinde daha fazla oranda paylaşım yapmışlardır. Elektronik sektörü firmalarından Firma 5’in Özel Teklif kategorisine ait tweetlere daha yüksek bir yüzde

ile yer verdiği, Firma 6'nın ise Ürün kategorisinde daha fazla oranda paylaşım yaptığı bulgular arasında yer almaktadır.

Firmaların paylaşımlarının elde etmiş olduğu rt sayıları Rt Analizi başlığı altında ele alınmıştır. Veri setindeki tüm firmaların paylaştıkları tweetler birlikte değerlendirildiğinde, her ne kadar daha az oranda içerik paylaşımı yapılmış olsa da, en fazla rt edilen içeriklerin Sosyal ve Yarışma&Etkinlik kategorilerine ait olduğu ilgi çekici bir bulgudur. Bu bulgu değerlendirildiğinde, firmaların ürün ve indirim tweetlerinden ziyade, firmalar tarafından kullanıcılara özel olarak paylaşılan yarışma ve sosyal mesaj içeren tweetlerin kullanıcılar tarafından daha çok kabul gördüğü ve başka kullanıcılarla paylaşılabilir görüldüğü söylenebilir.

Firmaların elde etmiş oldukları toplam rt sayılarının Pareto İlkesi ile açıklanabilir oluşu bir başka önemli bulgudur. Firmaların 2017 yılı boyunca elde etmiş oldukları toplam rt sayılarının büyük çoğunluğunun, firmaların paylaştığı tweetlerin azınlık bir bölümünden elde edildiği tespit edilmiştir. Her firma için toplam etkileşimin yüzde 80'inin elde edildiği tweet kategorileri belirlenmiştir. Bulgulara göre kozmetik sektöründe faaliyet gösteren rakip firmaların her ikisinin de en çok etkileşim (rt) aldığı tweetlerin büyük çoğunluğunun Özel Teklif kategorisinden olduğu sonucuna ulaşılmıştır. Pazaryeri sektörü firmalarının almış olduğu rt sayılarının büyük çoğunluğunun ise firmalara göre farklılık gösterdiği, Firma 3'ün Özel Teklif ve Firma 4'ün Özel Etkileşim kategorilerinden oluşan azınlıktaki tweetlerinin firmaların elde etmiş olduğu etkileşim miktarının büyük çoğunluğunu oluşturduğu görülmüştür. Son olarak elektronik sektörü firmalarının Yarışma&Etkinlik kategorilerinde yapmış oldukları paylaşımlarının rt sayılarının büyük çoğunluğunu oluşturduğuna ulaşılmıştır.

Firmalar ile ilgili değerlendirmelerde son olarak firmaların hashtag kullanım frekansları hesaplanarak kozmetik sektöründe faaliyet gösteren Firma 1'in diğer firmaların hashtag paylaşım frekanslarına oranla daha fazla hashtag paylaşımı yaptığı tespit edilmiştir. Bununla birlikte, hashtaglerin hangi içeriklerden oluştuğu incelenerek #mutlulukfirma1de gibi firmalara, #xürünüfirma6da gibi ürünlere,

#10kasım gibi özel günlere ve #etkinlikadı gibi etkinliklere özgü hashtag kullanımlarının olduğu saptanmıştır.

Uygulamanın ikinci bölümünde de ilk bölümde olduğu gibi sosyal medya madenciliği adımları izlenmiş, veri elde edildikten sonra analize hazır hale getirilmiştir. Tanımlayıcı İstatistikler bölümünde, belirlenen kampanya dönemini (Sevgililer Günü) içerisinde bulunduran Şubat ayına ait hem firmaların yapmış oldukları paylaşımlar hem de kullanıcıların firmalarla ilgili yapmış oldukları paylaşımlar ile ilgili istatistiklere yer verilmiştir. Kullanıcılar tarafından paylaşılan tweetlerin günlük frekansları değerlendirilerek, kullanıcılar tarafından firmalar ile ilgili paylaşımların en çok yapıldığı günler ile firmaların paylaşımları arasında bir uyum aranmıştır. Bulgulara göre yalnızca kozmetik sektöründe faaliyet gösteren firmalar belirlenen kampanya dönemi boyunca paylaşmış oldukları tweetler ile kullanıcıları içerik üretmeye teşvik edebilmişlerdir.

Kullanıcıların firmalar ile ilgili Şubat ayı boyunca yapmış oldukları paylaşımlar zaman dilimlerine ve araç bilgisine (kaynak) göre incelenerek sektör temelinde benzerlikler saptanmıştır. Bulgulara göre tüm firmalarla ilgili yapılan paylaşımların en çok öğleden sonra ve akşam saatlerinde yapıldığı görülmüştür. Kullanıcıların hangi kaynağı kullanarak tweet paylaştıkları incelendiğinde ise, kozmetik ve elektronik sektörü firmaları ile ilgili paylaşımların büyük çoğunluğunun Android cihazlar kullanılarak yapılmıştır. Pazaryeri firmalarında ise daha çok Web ve Diğer araçlar yolu ile paylaşımın yapıldığı saptanmıştır. Bu bulgunun sebebinin pazaryeri firmaları ile ilgili üretilen içeriklerin çoğunluğunun pazaryerinde satış gerçekleştiren kişi/kişiler tarafından yapıldığı düşünüldüğünde, bu kişi/kişilerin Web üzerinden faaliyetlerini yürüttüğünü ve aynı zamanda da içerik paylaştıklarını söylemek mümkündür.

Uygulamanın ikinci bölümünün birincil amacı olan firmaların Twitter etkinliklerinin ölçümü, Şubat ayına ait Twitter verileri içerisinde belirlenen kriterler doğrultusunda gerçekleştirilmiştir. Firmalar ile ilgili kullanıcılar tarafından paylaşılan tweetlerle ilgili istatistikler (firmalar hakkındaki tweet sayısı, tweet değeri, takipçi kazanımı, yanıt

sayısı, rt sayısı ve beğeni sayısı) ele alınarak firmaların Twitter etkinliklerinin ölçümü gerçekleştirilmiştir. Tweet değeri kriteri ile ilgili istatistiklerin elde edilebilmesi için duygu analizi yapılmıştır. Üç araştırmacı tarafından veri seti manuel olarak olumlu, nötr ve olumsuz olmak üzere kodlanmış ve araştırmacılar arası güvenilirlik Krippendorff Alfa katsayısıyla test edilerek yüksek uyum saptanmıştır. En çok olumlu içeriğin kozmetik sektörü firmaları ile ilgili üretildiğine ulaşılmıştır. Pazaryeri ve elektronik sektöründe faaliyet gösteren firmalar ile ilgili değerlendirme yapıldığında ise, paylaşılan olumsuz tweetlerin yüzdelerinin olumlu tweet yüzdelerinden fazla olduğu görülmüştür.

Altı kriter gere firmaların etkinlikleri ayrı ayrı sıralanmış, daha sonra her kriterdeki firma sıralamasına göre belirlenen puan skalası doğrultusunda firmaların toplamda elde ettikleri performans sıralaması verilmiştir. Bulgulara göre kozmetik sektöründe faaliyet gösteren Firma 1'in Şubat ayı boyunca diğer firmalara kıyasla çok daha fazla konuşulduğu ve daha fazla etkileşim aldığı ve elektronik sektöründe faaliyet gösteren Firma 6'nın ise etkinlik sıralamasında en düşük sırada yer aldığı tespit edilmiştir. Bu kriter sıralaması yalnızca Twitter etkinlik değerlendirmesi için kullanılmış olup herhangi bir kriter ağırlıklandırması yapılmadan gerçekleştirilmiştir. Gelecek çalışmalarda, kriterlerin ait olduğu başlıklar (Marka Bilinirliği, Marka İlginliği ve Ağızdan Ağıza İletişim) altında firma etkinlikleri kriter ağırlıklandırması da yapılarak gerçekleştirilirse daha güvenilir ve kıyaslanabilir sonuçlar elde edilebilir.

Elde edilen tüm bu bulgular ve gerçekleştirilen kapsamlı literatür taraması ışığında bu tez çalışmanın yeni gelişmekte olan sosyal medya madenciliği alanı ile ilgili literatüre katkı sağlayacağı düşünülmektedir. Bu çalışmada veri kaynağı olarak yalnızca Twitter kullanılmıştır. Daha fazla sosyal medya platformu ele alınarak, değişken sayısı ve veri setindeki sektör ve firma sayısı artırılarak farklı çalışmaların yapılması mümkündür. Gelecek çalışmalarda veri setine firmaların finansal raporlarına ait değişkenler de eklenerek firmaların sosyal medya etkinliği ile finansal başarısı arasındaki ilişki incelenebilir.

KAYNAKÇA

- af Rosenberg, K., “Buzz vs. Sales: Big Social Data Analytics of Style Christina, D., Ida Buhl- Andersen, Line B. Nilsson, Mark P. Rebuild, Raghava R. Mukkamala, ..., Ravi Vatrappu: “Icon Campaigns and Fashion Designer Collaborations on H&M’s Facebook Page.” **Proceedings of the 50th Hawaii International Conference on System Sciences**, 2017, pp.1861-1870.
- Aggarwal, Charu C., “A Survey of Text Clustering Algorithms.” **Mining Text Data**, (eds.) Charu C. Aggarwal, ChengXiang Zhai: Springer, Boston, MA, 2012, pp. 77-128.
- Aghaei, Sareh, Mohammad A. Nematbakhsh, Hadi K. Farsani: “Evolution of The World Wide Web: From Web 1.0 to Web 4.0.” **International Journal of Web&Semantic Technology**, Vol.3, No:1, 2012, pp.1-10
- Akpınar, Haldun: “Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği.” **İÜ İşletme Fakültesi Dergisi**, C.29, No:1, 2000, s.1-22.
- Algozaibi, Abdullelah A., Saleh Albahli, Austin Melton: “World Wide Web: A Survey of Its Development and Possible Future Trends.” **The 16th International Conference on Internet Computing and Big Data**, 2015, pp.79-84
- Alpaydın, Ethem: “Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri.” **Bilişim 2000 Eğitim Semineri**, 2000, s.1-10.
- Andersen, Paul: “What is Web 2.0?: Ideas, Technologies And Implications for Education.” **JISC Technology and Standards Watch**, Vol.1, No:1., 2007, pp.1-64

- Asur, Sitaram, Bernardo A. Huberman: "Predicting the Future with Social Media." **Web Intelligence and Intelligent Agent Technology (WI-IAT)**, Vol. 1, 2010, pp.492-499.
- Atzmueller, Martin: "Mining Social Media: Key Players, Sentiments, and Communities." **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Vol.2, No:5, 2012, pp.411-419.
- Ayata, Değer, Murat Saraçlar, Arzucan Özgür: "Turkish Tweet Sentiment Analysis with Word Embedding and Machine Learning." **25th Signal Processing and Communications Applications Conference (SIU)**, 2017, pp.1-4.
- Baloğlu, Arzu: "Genel Tanımlamalar ve Değişen Devlet Modeli." **Sosyal Medya Madenciliği**, (ed.) Arzu Baloğlu, İstanbul, Beta Yayınları, 2015, s.1-15.
- Barbier, Geoffrey, Huan Liu: "Data Mining in Social Media." **Social Network Data Analytics**, (ed.) Charu C. Aggarwal, Springer, Boston, MA, 2011, pp. 327-352.
- Bello-Orgaz, Gema, Jason J. Jung, David Camacho: "Social Big Data: Recent Achievements and New Challenges." **Information Fusion**, Vol.28, 2016, pp.45-59.
- Benito-Osorio, Diana, Marta Peris-Ortiz, Carlos R. Armengot, Alberto Colino: "Web 5.0: The Future of Emotional Competences in Higher Education." **Global Business Perspectives**, Vol.1, No:3, 2013, pp.274-287.
- Bian, Jiang, Kenji Yoshigoe, Amanda Hicks, Jiawei Yuan, Zhe He, Mengjun Xie, ..., François Modave: "Mining Twitter To Assess The Public Perception of The "Internet of Things"." **Plos One**, Vol.11, No:7, 2016, pp.1-14.

- Bilgen, Özge B., Nuri Doğan: "Puanlayıcılar Arası Güvenirlik Belirleme Tekniklerinin Karşılaştırılması." **Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi**, C.8, No:1, 2017, pp.63-78.
- Boldt, Linda C., Vinothan Vinayagamoorthy, Florian Winder, Melanie Schnittger, Mats Ekran, M., Raghava R. Mukkamala, ..., Ravi Vatrapu: "Forecasting Nike's Sales Using Facebook Data." **2016 IEEE International Conference on Big Data**, 2016, pp. 2447-2456.
- Bollen, Johan, Huina Mao, Xiao-Jun Zeng: "Twitter Mood Predicts The Stock Market." **Journal of Computational Science**, Vol.2, No:1, 2011, pp.1-8.
- Bonzanini, Marco: **Mastering Social Media Mining with Python**, Packt Publishing Ltd., 2016.
- Boyd, Danah M., Nicole B. Ellison: "Social Network Sites: Definition, History, and Scholarship.", **Journal of Computer-Mediated Communication**, Vol.13, No.1, 2007, pp.210-230.
- Bruns, Axel, Mark Bahnisch: "Social Media: Tools for User-Generated Content: Social Drivers Behind Growing Consumer Participation in User-Led Content Generation." **Smart Services CRC**, Vol.1, 2009, pp.1-60.
- Can, Umit, & Bilal Alatas: "Duygu Analizi ve Fikir Madenciliği Algoritmalarının İncelenmesi." **International Journal of Pure and Applied Sciences**, Vol.3, No:1, 2017, pp.75-111.
- Chae, Bongsug K: "Insights from Hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter Data for Supply Chain Practice and Research." **International**

- Journal of Production Economics**, Vol.165, 2015, pp.247-259.
- Chatfield, Akemi T., Uuf Brajawidagda: “Twitter Tsunami Early Warning Network: A Social Network Analysis of Twitter Information Flows.” **23rd Australasian Conference on Information Systems**, 2012, pp.1-10.
- Choudhury, Nupur: “World Wide Web and Its Journey From Web 1.0 to Web 4.0.” **International Journal of Computer Science and Information Technologies**, Vol.5, No:6, 2014, pp.8096-8100.
- Cohen-Almagor, Raphael : “Internet History.” **International Journal of Technoethics**, Vol.2, No:2, 2011, pp.45-64.
- Coşkun, Cengiz, Abdullah Baykal: “Veri Madenciliğinde Sınıflandırma Algoritmalarının Bir Örnek Üzerinde Karşılaştırılması.” **Akademik Bilişim Konferansı**, Malatya, 2011, s.51-58.
- Çilan, Çiğdem Arıcıgil: **Sosyal Bilimlerde Kategorik Verilerle İlişki Analizi-Kontenjans Tabloları Analizi**, 2. bs., İstanbul, Pegem Akademi, 2013.
- Çoban, Önder, Barış Özyer, Gülşah T. Özyer: “Sentiment Analysis for Turkish Twitter Feeds.” **23th Signal Processing and Communications Applications Conference (SIU)**, 2015, pp. 2388-2391.
- Danneman, Nathan, Richard Heimann: **Social Media Mining with R**, Packt Publishing Ltd., 2014.
- Dickey, Irene J., William F. Lewis: “An Overview of Digital Media and Advertising.” **E-Marketing: Concepts, Methodologies, Tools, and Applications**, Information Resources Management Association (USA), 2012, pp.31-61.

- Ding, Chao, Hsing K. Cheng., Yang Duan, Yong Jin: "The Power of The "Like" Button: The Impact of Social Media on Box Office." **Decision Support Systems**, Vol.94, 2017, pp.77-84.
- Dwivedi, Akhilesh, Suresh Kumar, Abhishek Dwivedi, Manjeet Singh: "Current Security Considerations for Issues and Challenges of Trustworthy Semantic Web." **International Journal of Advanced Networking and Applications**, Vol.3, No:1, 2011, pp.978-983.
- Evans, Dave: **Social Media Marketing: An Hour a Day**, Willey Publishing, 2008.
- Fayyad, Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic: "From Data Mining to Knowledge Discovery in Databases." **AI Magazine**, Vol.17, No:3, 1996, pp.37-54.
- Freelon, Deen: "ReCal OIR: Ordinal, Interval, and Ratio Intercoder Reliability as a Web Service." **International Journal of Internet Science**, 2013, Vol.8, No:1, pp.10-16.
- Ghosh, Sayantani, Sudipta Roy, Samir K. Bandyopadhyay: "A Tutorial Review on Text Mining Algorithms." **International Journal of Advanced Research in Computer and Communication Engineering**, Vol.1, No:4, 2012, pp.
- Green, Leila: **The Internet: An Introduction to New Media**, Oxford Berg., 2010.
- Guidry, Jeanine D., Marcus Messner, Yan Jin, Vivian Medina-Messner: "From #mcdonaldsfail to #dominossucks: An Analysis of Instagram Images About The 10 Largest Fast Food Companies." **Corporate Communications: An International Journal**, Vol.20, No:3, 2015, pp.344-359.
- Gundecha, Pritam, Huan Liu: "Mining Social Media: A Brief Introduction." **INFORMS Tutorials in Operations Research**, 2014, pp.1-17.

- Gunelius, Susan: **30 Minute Social Media Marketing: Step By Step Techniques to Spread The Words About Your Business**, Business&Economics, 2011.
- Güngör, Mehmet, Yunus Bulut: “Ki-Kare Testi Üzerine.” **Doğu Anadolu Bölgesi Araştırmaları**, C.7, No:1, 2008, s.84-89.
- Gürel, Ement, Mehmet Yakın: “Ekşi Sözlük: Postmodern Elektronik Kültür.” **Selçuk Üniversitesi İletişim Fakültesi Akademik Dergisi**, C.4, No:4, 2007, s.203-219.
- Gürsakal, Necmi: **Büyük Veri**, 3. bs., Bursa, Dora Yayınları, 2014.
- Han, Jiawei, Jian Pei, Micheline Kamber: **Data Mining: Concepts and Techniques**, 3rd ed., Morgan Kaufmann Publishers, 2011.
- Hand, David J: “Principles of Data Mining.” **Drug Safety**, Vol.30, No:7, 2007, pp.621-622.
- He, Wu, Shenghua Zha, Ling Li: “Social Media Competitive Analysis and Text Mining: A Case Study in The Pizza Industry.” **International Journal of Information Management**, Vol.33, No:3, 2013, pp.464-472.
- Hiremath, B. K., Anand Y. Kenchakkanavar: “An Alteration of the Web 1.0, Web 2.0 and Web 3.0: A Comparative Study.” **Imperial Journal of Interdisciplinary Research**, Vol. 2, No:4, 2016, pp.705-710
- Hoffman, Donna L., Marek Fodor: “Can You Measure the ROI of Your Social Media Marketing?.” **MIT Sloan Management Review**, Vol.52, No:1, pp.41-49.
- Holt, John D., Soon M. Chung: “Multipass Algorithms For Mining Association Rules in Text Databases.” **Knowledge and Information Systems**, Vol.3, No:2, 2001, pp.168-183.
- Holvoet, Katherine: “What is RSS and How Can Libraries Use It To Improve Patron Service?.” **Library Hi Tech News**, Vol.23, No:8, 2006, pp.32-33.

- Hotho, Andreas, Andreas Nürnberger, Gerhard Paaß, “A Brief Survey of Text Mining.” **Journal for Computational Linguistics and Language Technology**, Vol.20, No:1, 2005, pp.19-62.
- Hu, Xia, Huan Liu: “Text Analytics in Social Media.” **Mining Text Data**, (eds.) Charu C. Aggarwal, ChengXiang Zhai, Springer, Boston, MA, 2012, pp. 385-414.
- Hu, Yuheng, Lydia Manikonda, L., Subbarao Kambhampati: “What We Instagram: A First Analysis of Instagram Photo Content and User Types.” **Proceedings of the Eighth International AAI Conference on Weblogs and Social Media**, 2014, pp.595-598.
- Hui, Cindy, Yulia Tyshchuk, William A. Wallace, Malik Magdon-Ismail, Mark Goldberg: “Information Cascades in Social Media in Response to A Crisis: A Preliminary Model And A Case Study.” **Proceedings of the 21st International Conference on World Wide Web**, 2012, pp. 653-656.
- Ishikawa, Hiroshi: **Social Big Data Mining**, CRC Press, 2015.
- İşeri, İsmail, Ömer F. Atasoy, Harun Alçiçek: “Sentiment Classification of Social Media Data for Telecommunication Companies in Turkey.” **2017 International Conference on Computer Science and Engineering (UBMK)**, 2017, pp.1015-1019.
- Java, Akshay, Xiaodan Song, Tim Finin, Belle Tseng: “Why We Twitter: Understanding Microblogging Usage and Communities.” **Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis**, 2007, pp.56-65.
- Jham, Bruno C., Gabriela V. Duraes, Howard E. Strassler, Luis G. Sensi: “Joining the Podcast Revolution.” **Journal of Dental Education**, Vol.72, No:3, 2008, pp.278-281.
- Kambil, Ajit: “What is your Web 5.0 strategy?.” **Journal of Business Strategy**, Vol.29, No:6, 2008, pp.56-58

- Kaplan, Andreas M., Michael Haenlein: "Users of The World, Unite! The Challenges and Opportunities of Social Media." **Business Horizons**, Vol. 53, No:1, 2010, pp.59-68.
- Kaplan, Andreas M., Michael Haenlein: "Collaborative Projects (Social Media Application): About Wikipedia, The Free Encyclopedia." **Business Horizons**, Vol.57, No:5, 2014, pp.617-626.
- Khan, M. Laeeq: "Social Media Engagement: What Motivates User Participation and Consumption on Youtube?." **Computers in Human Behavior**, Vol.66, 2017, pp:236-247.
- Kılıç, Abdullah F.: "Uyum Analizi (Correspondence Analysis)." **YBS Ansiklopedi**, C.3, No:1, 2016, s.1-20.
- Ki, Eyun-Jung, Elmie Nekmat: "Situational Crisis Communication and Interactivity: Usage and Effectiveness of Facebook for Crisis Management by Fortune 500 Companies." **Computers in Human Behavior**, Vol.35, 2014, pp.140-147.
- Kietzmann, Jan H., Kristopher Hermkens, Ian P. McCarthy, Bruno S. Silvestre: "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media." **Business Horizons**, Vol. 54, No:3, 2011, pp.241-251.
- Kim, Sora, Xiaochen A. Zhang, Borui W. Zhang: "Self-Mocking Crisis Strategy on Social Media: Focusing on Alibaba Chairman Jack Ma in China." **Public Relations Review**, Vol.42, No:5, 2016, pp.903-912.
- Kim, Won, Jeong, Ok-Ran, Sang-Won Lee: "On Social Web Sites." **Information Systems**, Vol.35, No:2, 2010, pp.215-236.
- Kotsiantis, Sotiris, Dimitris Kanellopoulos: "Association Rules Mining: A Recent Overview." **GESTS International Transactions on Computer**

- Science and Engineering**, Vol.32, No:1, 2006, pp.71-82.
- Kotu, Vijay, Bale Deshpande: **Predictive Analytics and Data Mining: Concepts and Practice with Rapidminer**, Morgan Kaufmann, 2014.
- Krippendorff, Klaus: "On The Reliability of Unitizing Continuous Data." **Sociological Methodology**, Vol.25, 1995, pp.47-76.
- Kuflik, Tsvi, Einat Minkov, Silvio Nocera, Susan Grant-Muller, Ayelet Gal-Tzur, Itay Shoor: "Automating A Framework To Extract and Analyse Transport Related Social Media Content: The Potential and The Challenges." **Transportation Research Part C: Emerging Technologies**, Vol.77, 2017, pp.275-291.
- Kumar, Shamanth, Reza Zafarani, Huan Liu: "Understanding User Migration Patterns in Social Media." **Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence**, 2011, pp.1204-1209
- Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon: "What is Twitter, A Social Network Or A News Media?." **Proceedings of The 19th International Conference on World Wide Web**, 2010, pp.591-600.
- Laningham, Scott: "Tim Berners-Lee," **DeveloperWorks Interviews**, August 22, 2006, (Çevrimiçi) <https://www.ibm.com/developerworks/podcast/dwi/cm-int082206txt.html>, 4 Haziran 2018.
- Lassen, Niels B., Rene Madsen, Ravi Vatrpu: "Predicting iphone Sales From iphone Tweets." **Enterprise Distributed Object Computing Conference (EDOC)**, 2014, pp. 81-90.
- Lee, Kiljae, Won-Yong Oh, Namhyeok Kim: "Social Media for Socially Responsible Firms: Analysis of Fortune 500's Twitter Profiles and Their

- CSR/CSIR Ratings.” **Journal of Business Ethics**, Vol.118, No:4, 2013, pp.791-806.
- Leiner, Barry M., Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch,..., Stephen Wolff : “A Brief History of The Internet.” **ACM SIGCOMM Computer Communication Review**, Vol. 39, No:5, 2009, pp.22-31.
- Lietsala, Katri, Esa Sirkkunen: **Social Media: Introduction to The Tools And Processes of Participatory Economy**, Tampere: University Of Tampere, Hypermedia Laboratory Net Series, 2008.
- Liu, Bing, Lei Zhang: “A Survey Of Opinion Mining and Sentiment Analysis.” **Mining Text Data**, (eds.) Charu C. Aggarwal, ChengXiang Zhai, Springer, Boston, MA, 2012, pp. 415-463.
- Liu, Bing: “Sentiment Analysis and Opinion Mining.” **Synthesis Lectures on Human Language Technologies**, Vol.5, No:1, 2012, pp.1-167.
- Liu, Huan, Fred Morstatter, Jiliang Tang, Reza Zafarani: “The Good, The Bad, And The Ugly: Uncovering Novel Research Opportunities in Social Media Mining.” **International Journal of Data Science and Analytics**, Vol.1, No:3-4, 2016, pp.137-143.
- Lomborg, Stine, Anja Bechmann: “Using APIs for Data Collection on Social Media.” **The Information Society**, Vol.30, No:4, 2014, pp.256-265.
- Mangold, W. Glynn, David J. Faulds: “Social Media: The New Hybrid Element of The Promotion Mix.” **Business Horizons**, Vol.52, No:4, 2009, pp.357-365.
- Manimaran, J., T. Velmurugan: “A Survey of Association Rule Mining in Text Applications.” **2013 IEE International Conference**

- on Computational Intelligence and Computing Research**, 2013, pp. 1-5.
- Mayfield, Antony: “What is Social Media?” iCrossing, 2008, (Çevrimiçi), http://www.icrossing.com/uk/sites/default/files_uk/insight_pdf_files/What%20is%20Social%20Media_iCrossing_ebook.pdf, 1 Mart 2017.
- McCormick, Tyler H., Hedwig Lee, Nina Cesare, Ali Shojaie, Emma S. Spiro: “Using Twitter for Demographic and Social Science Research: Tools for Data Collection and Processing.” **Sociological Methods&Research**, Vol.46, No:3, 2017, pp.390-421.
- McHugh, Mary L.: “Interrater Reliability: The Kappa Statistic.” **Biochemia Medica**, Vol.3, 2012, pp.276-282.
- McHugh, Mary L.: “The Chi-Square Test of Independence.” **Biochemia Medica**, Vol.23, No:2, 2013, pp.143-149.
- Miner, Gary, John Elder IV, Thomas Hill: **Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications**, Academic Press, 2012.
- Murugesan, San: “Web X.0: A Road Map.” **Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications**, (ed.) San Murugesan, Information Science Reference, 2009, pp. 1-11.
- Nath, Keshab, Sourish Dhar, Subhash Basishtha: “Web 1.0 to Web 3.0-Evolution of the Web and Its Various Challenges.” **Proceedings of the 2014 International Conference on Reliability, Optimization and Information Technology**, 2014, pp.86-89.
- Nenkova, Ani, Kathleen Mckeown: “A Survey of Text Classification Algorithms.” **Mining Text Data**, (eds.) Charu C. Aggarwal,

- ChengXiang Zhai, Springer, Boston, MA, 2012, pp. 163-222.
- Neti, Sisira: “Social Media and Its Role in Marketing.” **International Journal of Enterprise Computing and Business Systems**, Vol.1, No:2, 2011, pp.1-15.
- Newman, Russell, Victor Chang, Robert J. Walters, Gary B. Wills: “Web 2.0—The Past and The Future.” **International Journal of Information Management**, Vol.36, No:4, 2016, pp.591-598.
- Nili, Alireza, Mary Tate, Alistair Barros: “A Critical Analysis of Inter-Coder Reliability Methods in Information Systems Research.”, **Australasian Conference on Information Systems**, 2017, pp.1-11.
- O’reilly, Tim: “What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software.” **Communications&Strategies**, Vol.1, No:65, 2007, pp.17-37
- Odlum, Michelle, Sunmoo Yoon: “What Can We Learn About The Ebola Outbreak from Tweets?.” **American Journal of Infection Control**, Vol.43, No:6, 2015, pp.563-571.
- OECD: **Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking**, ABD, 2007.
- Oh, Chong, Yaman Roumani, Joseph K. Nwankpa, Han-Fen Hu: “Beyond Likes and Tweets: Consumer Engagement Behavior and Movie Box Office in Social Media.” **Information&Management**, Vol.54, No:1, 2017, pp.25-37.
- Orsburn, Eva M.: **The Social Media Business Equation: Using Online Connections to Grow Your Bottom Line**, Nelson Education, 2011.

- Özdemir, Selçuk, Halit “Günümüz Teknolojileri Ne Kadar Karalar: “Akıllı”?.” **Türkiye Bilişim Derneği Bilişim Dergisi**, No:146, 2012, s.100-108.
- Patel, Karan: “Incremental Journey for World Wide Web: Introduced with Web 1.0 to Recent Web 5.0–A Survey Paper.” **International Journal of Advanced Research in Computer Science and Software Engineering**, Vol.3, No:10, 2013, pp.410-417.
- Poell, Thomas, & Erik Borra: “Twitter, Youtube, And Flickr As Platforms of Alternative Journalism: The Social Media Account of The 2010 Toronto G20 Protests.” **Journalism**, Vol.13, No:6, 2012, pp.695-713.
- Pournarakis, Demitrios E., Dionysios N. Sotiropoulos, George M. Giaglis: “A Computational Model for Mining Consumer Perceptions in Social Media.” **Decision Support Systems**, Vol.93, 2017, pp.98-110.
- Ravindran, Sharan K., Vikram Garg: **Mastering Social Media Mining with R**, Packt Publishing Ltd., 2015.
- Rossi, Luca, Matteo Magnani: “Conversation Practices and Network Structure in Twitter.” **Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media**, 2012, pp.563-566.
- Runeson, Per, Magnus Alexandersson, Oskar Nyholm: “Detection of Duplicate Defect Reports Using Natural Language Processing.” **Proceedings of the 29th International Conference on Software Engineering**, 2007, pp. 499-510.
- Russell, Matthew A.: **Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More**, O'Reilly Media, Inc., 2013.

- Ryan, Damian: "Understanding Social Media." **Understanding Digital Marketing: Marketing Strategies for Engaging The Digital Generation**, 4th ed., Kogan Page Publishers, 2016, pp.120-152.
- Safko, Lon, David Brake: **The Social Media Bible: Tactics, Tools and Strategies for Business Success**, John Wiley&Son, 2009.
- Safko, Lon: **The Social Media Bible: Tactics, Tools and Strategies for Business Success**, John Wiley&Son, 3rd ed., 2012.
- Sajithra, K., Rajindra Patil: "Social Media–History and Components." **Journal of Business and Management**, Vol.7, No:1, pp.69-74.
- Sanders, Robert: "The Pareto Principle: Its Use and Abuse." **Journal of Services Marketing**, Vol.1, No:2, 1987, pp.37-40.
- Saravanakumar, M., T. SuganthaLakshmi: "Social Media Marketing.", **Life Science Journal**, Vol.9, No:4, pp.4444-4451.
- Seker, Sadi E.: "Sosyal Ağlarda Akan Veri Madenciliği." **YBS Ansiklopedi**, C.1, No:3, 2014, s.21-27 .
- Sert, Fatma, Selim Tüzüntürk, Necmi Gürsakal: "NodeXL ile Sosyal Ağ Analizi: #akademikzam Örneği." **15. Uluslararası Ekonometri, Yöneylem Araştırmaları ve İstatistik Sempozyumu**, Vol.15, 2014, s.464-482.
- Shao, Guosong: "Understanding the Appeal of User-Generated Media: A Uses and Gratification Perspective." **Internet Research**, Vol.19, No:1, 2009, pp.7-25.
- Shearer, Colin: "The CRISP-DM Model: The New Blueprint for Data Mining." **Journal of Data Warehousing**, Vol.5, No:4, 2000, pp.13-22.

- Sterne, Jim: **Social Media Metrics: How to Measure and Optimize Your Marketing Investment**, John Wiley&Sons, 2010.
- Stieglitz, Stefan, Linh Dang-Xuan, Axel Bruns, Christoph Neuberger: “Social Media Analytics.” **Business & Information Systems Engineering**, Vol.6, No:2, 2014, pp.89-96.
- T.C. Gençlik ve Spor Bakanlığı: “Gençlik ve Sosyal Medya Araştırma Raporu: Yönetici Özeti” Ipsos, (ed.) Mehmet Bulut, Ekim 2013, (Çevrimiçi) <https://leventeraslan.files.wordpress.com/2015/09/genclik-sosyalmedyayonetici-ozeti.pdf>, 15 Mayıs 2018.
- Tang, Jiliang, Yi Chang, Huan Liu: “Mining Social Media with Social Theories: A Survey.” **ACM SIGKDD Explorations Newsletter**, Vol.15, No:2, 2014, pp.20-29.
- Tang, Lei, Huan Liu: “Community Detection and Mining in Social Media.” **Synthesis Lectures on Data Mining and Knowledge Discovery**, Vol.2, No:1, 2010, pp.1-137.
- Thelwall, Mike, Kevan Buckley, Georgios Paltoglou: “Sentiment in Twitter events.” **Journal of the Association for Information Science and Technology**, Vol.62, No:2, 2011, pp.406-418.
- Tremayne, Mark: “Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street.” **Social Movement Studies**, Vol.13, No:1, 2014, pp.110-126.
- Tufekci, Zeynep: “Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls.” **Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media**, 2014, pp.505-514
- TÜİK: “Hanehalkı Bilişim Teknolojileri Kullanım Araştırması, 2016”, (Çevrimiçi)

- <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=21779>, 1 Ocak 2018.
- Umesha, Naik, D. Shivalingaiah: “Comparative Study of Web 1.0, Web 2.0 and Web 3.0.” **The 6th International CALIBER**, 2008, pp.499-507.
- Uysal, Alper K., Serkan Gunal: “The Impact of Preprocessing on Text Classification.” **Information Processing&Management**, Vol.50, No:1, 2014, pp.104-112.
- Uzgören, Nevin: “Uyum Analizinin Teorik Esasları Ve Regresyon Analizi İle Benzerliğinin Grafiksel Boyutta Karşılaştırılması.” **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, No:18, 2007, s.1-20.
- van der Meer, Toni G.L.A., Piet Verhoeven: “Public Framing Organizational Crisis Situations: Social Media Versus News Media.” **Public Relations Review**, Vol.39, No:3, 2013, pp.229-231.
- van Mierlo, Trevor: “The 1% Rule in Four Digital Health Social Networks: An Observational Study.” **Journal of Medical Internet Research**, Vol.16, No:2, 2014, pp.1-9.
- Verma, Tanu, Renu, Deepti Gaur: “Tokenization And Filtering Process in Rapidminer.” **International Journal of Applied Information Systems**, Vol.7, No:2, 2014, pp.16-18.
- Wamba, Samuel F., Shahriar Akter, Hyunjin Kang, Mithu Bhattacharya, Mohammed Upal: “The Primer of Social Media Analytics.” **Journal of Organizational and End User Computing**, Vol.28, No:2, 2016, pp.1-12.
- Waters, Richard D., Jia Y. Jamal: “Tweet, Tweet, Tweet: A Content Analysis of Nonprofit Organizations’ Twitter Updates.” **Public Relations Review**, Vol.37, No:3, 2011, pp.321-324.
- We Are Social, Hootsuite: “Digital in 2018 in Western Asia.”, Jan 29, 2018, (Çevrimiçi)

- <https://www.slideshare.net/wearesocial/digital-in-2018-in-western-asia-part-1-northwest-86865983> , 15 May 2018.
- Weber, Sebastian, Jörg Rech: “An Overview and Differentiation of the Evolutionary Steps of the Web X.Y Movement: The Web Before and Beyond 2.0.” **Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications**, (ed.) San Murugesan, Information Science Reference, 2009, pp. 12-39.
- Weinberg, Tamar: **The New Community Rules: Marketing on The Social Web**, O'Reilly Media Inc., 2009.
- Weiss, Sholom M., Nitin Indurkha: **Predictive Data Mining: A Practical Guide**, Morgan Kaufmann, 1988.
- Wirth, Rüdiger, Jochen Hipp: “CRISP-DM: Towards A Standard Process Model for Data Mining.” **Proceedings of the 4th International Conference on The Practical Applications of Knowledge Discovery and Data Mining**, 2000, pp. 29-39.
- Wong, Pak C., Paul Whitney, Jim Thomas: “Visualizing Association Rules for Text Mining.” **Proceedings 1999 IEEE Symposium on Information Visualization**, 1999, pp. 120-123.
- Xiang, Zheng, Qianzhou Du, Yufeng Ma, Weiguo Fan: “A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism.” **Tourism Management**, Vol.58, 2017, pp.51-65.
- Ye, Lan, Ki, Eyun-Jung: “Organizational Crisis Communication On Facebook: A Study of BP’s Deepwater Horizon Oil Spill.” **Corporate Communications: An International Journal**, Vol.22, No:1, 2017, pp.80-92.

Yilmaz, İlker G., Doğuş “Social Media’s Perspective on Industry 4.0: A
Aygün, Zuhul Tanrikulu: Twitter Analysis.” **Social Networking**, Vol.6, 2017,
pp.251-261.

Zafarani, Reza, **Social Media Mining: An Introduction**, Cambridge
Mohammad A. Abbasi, University Press, 2014.

Huan Liu:

Zarrella, Dan: **The Social Media Marketing Book**, O'Reilly Media
Inc., 2009.

Zhang, Wen, Taketoshi “A Comparative Study of TF* IDF, LSI and Multi-
Yoshida, Xijin Tang: Words for Text Classification.” **Expert Systems with
Applications**, Vol.38, No:3, 2011, pp.2758-2765.

Zimbra, David, Tianjun Fu, “Assessing Public Opinions Through Web 2.0: A Case
Xin Li: Study On Wal-Mart.” **Thirtieth International
Conference on Information Systems**, 2009, pp.1-10.

<https://www.boomsocial.com>

<https://apps.twitter.com>

[https://developer.twitter.com/en/docs/tweets/timelines/api-reference/get-statuses-
user_timeline](https://developer.twitter.com/en/docs/tweets/timelines/api-reference/get-statuses-user_timeline)

<https://developer.twitter.com/en/docs/tweets/search/overview>

<https://help.twitter.com/en/rules-and-policies/twitter-limits>

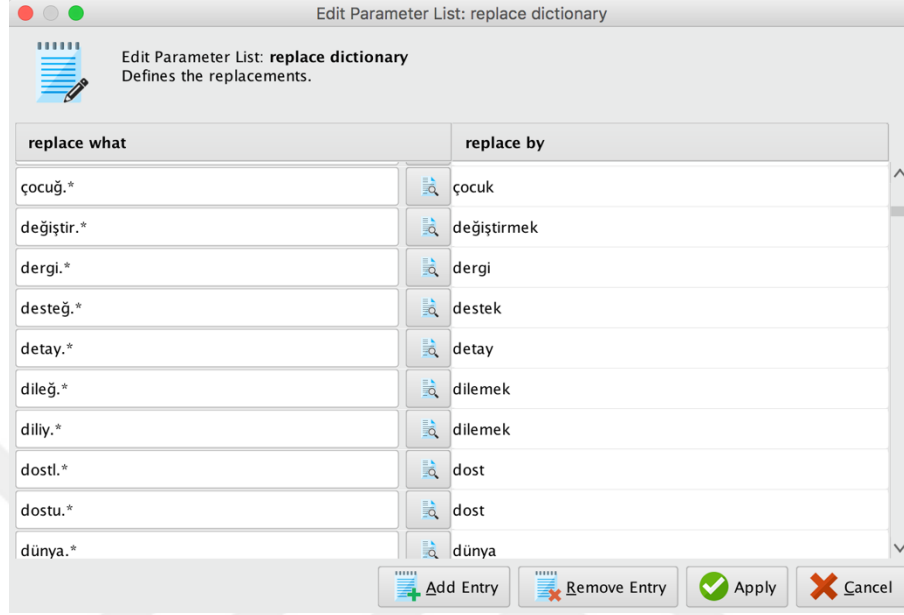
<https://github.com/linvi/tweetinvi>

<https://rapidminer.com>

<http://dfreelon.org/utlis/recalfront/recal3/>

EKLER

EK 1: RapidMiner Parça Değiştirme Operatörü Sözlüğü



EK 2: Kullanıcılar Tarafından Firmalar İle İlgili Paylaşılan Olumlu Tweetlere Örnekler

Tweet	Duygu Durumu
Allaaaah sevgililer günü ve indirimleri geliyor bekle beni Firma1 ❤️❤️❤️❤️	Olumlu
Sevgililer günü için çok heyecanlıyım Firma1 uzun zamandır %50 indirim yapmıyordu.	Olumlu
Biz morun en çok Firma1 tonunu seviyoruz	Olumlu
Firma2 indirimini bekliyorum hemde hemen <3	Olumlu
Firma2 indrimi iyi geldi hee	Olumlu
Konuşamam yalnızca Firma2..😭	Olumlu
Firma3'den sipariş verdim dün akşam ve az önce geldi :)	Olumlu
İptal isteğimi kısa sürede gerçekleştirdiğiniz için teşekkür ederim. @Firma3destek @Firma3	Olumlu
Firma3 müthiş bir alışveriş platformu	Olumlu
@kullanıcıadı Ben de Firma4 den alıcam aşzslsşzşşşcxşşcxşş	Olumlu

@kullanıcıadı Al Firma4 daha güvenilir	Olumlu
İyi akşamlar. Firma4 modakupon herkes güvenerek her şeyi alabilir. Müşteri memnuniyetine yıldızlı beş vermek istiyorum	Olumlu
@ Firma4 Firma4 adamdır. Öpüldünüz :)	Olumlu
@ Firma4 Sorun giderildi teşekkürler	Olumlu
Bugün mükemmel geçti. Evsahipliği için @ Firma4 ekibine ve zaman ayırıp gelen tüm Bursalı teknoloji tutkunlarına sonsuz teşekkürler. Özel video yakında @kullanıcıadı...	Olumlu
@kullanıcıadı Aynen bende heyecanlıyım Firma5 uzun zamandır indirim yapmıyordu	Olumlu
Adana Firma5'in bende ki manevi değeri.. :)	Olumlu
Kulaklığım bozulmuştu teknoşaya götürdüm hemen 2 dk içinde yeni ürün verildi teşekkürler # Firma5 ♥	Olumlu

EK 3: Kullanıcılar Tarafından Firmalar İle İlgili Paylaşılan Nötr Tweetlere Örnekler

Tweet	Duygu Durumu
@kullanıcıadı Firma1 sevgililer günü indiriminden kendime makyaj aldım bu kadar net yalnızız dostum	Nötr
Güzel kadın yoktur; Firma1'de indirim vardır.	Nötr
Erkekler neden #Firma1 in önünde bekliyor neden içeri girmiyorlar anlam verebilmiş değilim :)	Nötr
Firma2 Mağazamızdaki Sweet Cheeks allık paleti 8 yüksek pigmentli yumuşak rengi bir araya getiriyor!	Nötr
Firma2 de parfümler indirimde girmiş	Nötr
Firma2'de sevgililer günü indirimi varmış	Nötr
eFinans Firma3 ile e-Ticaret Zirvesi'nde.\nAyrıntılı bilgi için tıklayınız:...	Nötr
Firma3 de hergün %50 indirim var	Nötr
@ Firma3 tekrar kdv indirimi olacak mı ?	Nötr

@kullanıcıadı Merhaba sizi aradık ulaşamadık konu takibimizde olup tarafınıza bilgilendirme sağlanacaktır. Firma4Destek	Nötr
Yok mu Firma4 kuponu olan bi delikanlı	Nötr
Firma4 e güvenebilir miyim?	Nötr
Firma5 cep telefonuna 10 taksit televizyona 15 taksit kampanyası ...	Nötr
@ Firma5 Sevgili bulucunuz yok mu? Önce ordan başlasak. Hediye alınır yani bi şekilde yine.	Nötr
@ Firma5 # Firma5lefutbol Bülent Korkmaz	Nötr
I'm at @ Firma6 in İzmir	Nötr
Firma6 outlet tekrar açılmış ...	Nötr
Abim Firma6 sakalı bırakmış	Nötr

EK 4: Kullanıcılar Tarafından Firmalar İle İlgili Paylaşılan Olumsuz Tweetlere Örnekler

Tweet	Duygu Durumu
Firma1 çalışanları kendini bilirkişi filan mı zannediyor?	Olumsuz
Firma1 ödemeyi böbrek üzerinden kabul etmeyeceksen mail atıp durma param yok	Olumsuz
Bugün Firma1 beni batırdı	Olumsuz
Kızlar Firma2'yi boykot eden 3 kişi mi kaldık poşetler görüyorum ellerinizde ne çabuk unutuyosunuz her şeyi ya	Olumsuz
Parasıyla rezil olmak isteyen #Firma2 ye buyursun!	Olumsuz
İnternet mağazacılığı gerçekten zayıf. Müşteri hizmetlerinin vermiş olduğu adrese gönderdiğim ürün teslim alınmayıp geri gönderiliyor. Sorun ile ilgili geri dönüş de yapılmadı @Firma2 #Firma2 #şikayet	Olumsuz
@kullanıcıadı @Firma3 @Firma3destek Aldigin urun arizali veya parasini odedigin urun yerine alt model gelmis olabilir..sen musteri olarak potansiyel suçlusun..cagri merkezindeki hukuktan anlamayan elemaanlar zaten bizi bastan mahkum ediyorlar .biz suçluyuz biz	Olumsuz
@Firma3 pişmanlıktır...	Olumsuz

@Firma3 Aman arkadaşlar dikkat! Olmayan ürünü alıp da rezil olmayın sevdiklerinize! Ondan sinra yazarlar bulamadık. Ne diye sattın o zaman?	Olumsuz
#Firma4 youtube'deki reklamınız ne kadar sinir bozucu ? Neredeyse her videodan sonra çıkıyor . Kaç dk dinlemek zorunda kaldım .	Olumsuz
#Firma4 mü? 3 kere düşünürüm bundan sonra	Olumsuz
Firma4 kadar dolandırıcıları barındıran bir site yok. resmen rezalet. kullanmayın sakın.	Olumsuz
@Firma5 müşteri memnuniyetini nasıl sağlıyor acaba? Dün aldığımız ve sadece denediğimiz bir ürünü bugün açıldığı için alamayacaklarını söylüyorlar.	Olumsuz
@Firma5 arıyoruz danışmaya bağlayacaksınız sözde hattan atıp duruyorsunuz bu ne rezilliktir!!!	Olumsuz
@Firma5 İşlemleri inceleyebilmeniz gerekmiyor, işlerinizi düzgun yapmanız gerekiyor. Bilgi formu ve Maillere cevap vermediniz. İnceleyecek olsanız ohooo..	Olumsuz
3 ay evvel aldığım Lenova bozuldu ve kutusu yok diye alamayız diyen yer @Firma6	Olumsuz
@Firma6 müşteri hizmetlerinize ulaşılmıyor ! Siz bana ulaşın.	Olumsuz
Teşhir ürününü kapalı kutu (sıfır ürün) diye teslim eden (etmeye çalışan) @Firma6 @kullanıcıadı	Olumsuz

EK 5: Tweetera Ekranı

The image shows a screenshot of a web form titled "Form1". The form has a light gray background and a blue border. It contains two input fields: "Keyword" and "File Name". Below these fields is a button labeled "Search in Twitter". The form is set against a white background.

EK 6: Tweetera Kodları

```
namespace tweetera
{
    public partial class Form1 : Form
    {
        DataTable dt;
        public Form1()
        {
            InitializeComponent();
        }

        private void button1_Click(object sender, EventArgs e)
        {
            try
            {
                Auth.SetUserCredentials();
                string userName = textBox2.Text;
                followUser(userName);
            }
            catch(Exception ex)
            {
                MessageBox.Show(ex.ToString());
            }
        }

        public void followUser(string userName)
        {
            try
            {
                TweetinviConfig.CurrentThreadSettings.TweetMode
= TweetMode.Extended;
                var searchRequestParameters = new SearchTweetsParameters(userName)
                {
                    Lang = LanguageFilter.Turkish,
                    SearchType = SearchResultType.Recent,
                    MaximumNumberOfResults = 2000000,
                    Since = new DateTime(2018, 1, 1),
                    TweetSearchType = TweetSearchType.All
                };
                var matchingTweets = Search.SearchTweets(searchRequestParameters);

                List<ITweet> tweetList = matchingTweets.Cast<ITweet>().ToList();
            }
            catch { }
        }
    }
}
```

```

string myString = string.Empty;
myString = tweetList[0].ToJson();
createDTandSetColumnNames(myString);

MessageBox.Show("tweetList=" + tweetList.Count.ToString());
#region loop
for (int i=0;i<tweetList.Count;i++)
{
    string jsonString = tweetList[i].ToJson();
    string[] jsonStringArray
= Regex.Split(jsonString.Replace("[", "").Replace("]", ""), "{,}");
    DataRow nr = dt.NewRow();
    int sayac = 0;
    foreach (string jSA in jsonStringArray)
    {
        sayac += 1;
        string[] RowData
= Regex.Split(jSA.Replace("{", "").Replace("}", ""), ",");

        foreach (string rowData in RowData)
        {
            try
            {
                int idx = rowData.IndexOf(":");
                if (idx > 0)
                {
                    string RowColumns = rowData.Substring(0, idx -
1).Replace("\\\"", "");
                    string RowDataString =
rowData.Substring(idx + 1).Replace("\\\"", "");
                    nr[RowColumns] = RowDataString;
                }
            }
            catch (Exception ex)
            {
                continue;
            }
        }
    }
    dt.Rows.Add(nr);
}
#endregion

string directory = String.Empty;

```



```

        string apppath2
= Environment.GetFolderPath(Environment.SpecialFolder.CommonApplicationData
);
        directory = apppath2 + @"REVITeRA";
        string path = Path.Combine(apppath2, directory);
        if (Directory.Exists(path))
        {
        }
        else
        {
            Directory.CreateDirectory(path);
        }
        string fileName = "tweetera.xlsx";
        string dadi = string.Empty;
        fileName = Path.Combine(path, fileName);

        try
        {
            dt.ExportToExcel(fileName);
        }
        catch (Exception ex)
        {
            MessageBox.Show(ex.ToString());
        }

    }
    catch (Exception ex)
    {
    }

}

private void button2_Click(object sender, EventArgs e)
{
    MessageBox.Show("last tweet id");
}
public void createDTandSetColumnNames(string jsonString)
{
    try
    {
        string[] jsonStringArray
= Regex.Split(jsonString.Replace("[", "").Replace("]", ""), "},{");

        List<string> ColumnsName = new List<string>();

```

```

        foreach (string jSA in jsonStringArray)
        {
            string[] jsonStringData
= Regex.Split(jSA.Replace("{", "").Replace("}", ""), ",");
            foreach (string ColumnsNameData in jsonStringData)
            {
                int idx = ColumnsNameData.IndexOf(":");
                if (idx > 0)
                {
                    string ColumnsNameString = ColumnsNameData.Substring(0, idx -
1).Replace("\\"", "");
                    if (ColumnsName != null)
                    {
                        if (!ColumnsName.Contains(ColumnsNameString))
                        {
                            ColumnsName.Add(ColumnsNameString);
                        }
                    }
                }
            }

            }

            dt = new DataTable();
            foreach (string AddColumnName in ColumnsName)
            {
                dt.Columns.Add(AddColumnName);
            }
        }
        catch (Exception ex)
        {
            MessageBox.Show(ex.ToString());
        }
    }
    public DataTable JsonStringToDataTable(string jsonString)
    {
        string[] jsonStringArray
= Regex.Split(jsonString.Replace("[", "").Replace("]", ""), "{,}");

```