

**T.C.  
SÜLEYMAN DEMİREL ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**TWITTER ÜZERİNDE YAPILAN TÜRKÇE PAYLAŞIMLAR  
İÇİN ETİKET ANALİZ ARACI**

**Yunus Emre KARABULUT**

**Danışman  
Doç. Dr. Ecir Uğur KÜÇÜKSİLLE**

**YÜKSEK LİSANS TEZİ  
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI  
ISPARTA - 2018**



© 2018 [Yunus Emre KARABULUT]

## TEZ ONAYI

**Yunus Emre Karabulut** tarafından hazırlanan “**Twitter Üzerinde Yapılan Türkçe Paylaşımlar İçin Etiket Analizi Aracı**” adlı tez çalışması aşağıdaki jüri üyeleri önünde Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü **Bilgisayar Mühendisliği Anabilim Dalı**’nda **YÜKSEK LİSANS TEZİ** olarak başarı ile savunulmuştur.

**Danışman**

**Doç. Dr. Ecir Uğur KÜÇÜKSİLLE**  
Süleyman Demirel Üniversitesi



**Jüri Üyesi**

**Dr. Öğr. Üyesi Turgay AYDOĞAN**  
Süleyman Demirel Üniversitesi



**Jüri Üyesi**

**Dr. Öğr. Üyesi Onur SEVLİ**  
Mehmet Akif Ersoy Üniversitesi



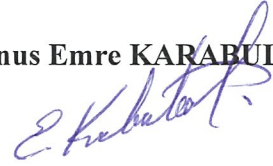
**Enstitü Müdürü**

**Prof. Dr. Yasin TUNCER**

## TAAHHÜTNAME

Bu tezin akademik ve etik kurallara uygun olarak yazıldığını ve kullanılan tüm literatür bilgilerinin referans gösterilerek tezde yer aldığını beyan ederim.

Yunus Emre KARABULUT



# İÇİNDEKİLER

	Sayfa
İÇİNDEKİLER .....	I
ÖZET.....	III
ABSTRACT .....	IV
TEŞEKKÜR.....	V
ŞEKİLLER DİZİNİ.....	VI
SİMGELER VE KISALTMALAR DİZİNİ .....	VIII
1. GİRİŞ .....	1
2. KAYNAK ÖZETLERİ .....	4
3. MATERYALLER VE METOT .....	8
3.1. Kullanılan Materyal ve Teknolojiler .....	8
3.1.1. Python programlama dili.....	8
3.1.2. Tweepy Twitter modülü .....	8
3.1.3. PostreSQL veritabanı yönetim sistemi.....	8
3.1.4. Re (Regex) ile düzenli ifadeler .....	10
3.1.5. Odoo kütüphanesi .....	10
3.1.6. Json kütüphanesi .....	11
3.1.7. URLLib modülü .....	14
3.1.8. NLTK (Natural Language Toolkit).....	15
3.1.8.1. Turkish Stemmer .....	15
3.1.9. Twitter sosyal ağ platformu .....	15
3.1.10. Etiket (hashtag) .....	17
3.1.11. Sosyal medya .....	17
3.1.12. Sosyal medya platformu.....	19
3.1.13. Sosyal medya özellikleri .....	20
3.1.14. Sosyal medya analizi.....	22
3.1.15. Duygu analizi (sentiment analyze).....	22
3.1.16. Cinsiyet analizi.....	24
3.1.17. Bağlantı (url) analizi .....	25
3.1.18. Anahtar kelime analizi .....	25
3.2. Twitter'dan Verilerin Çekilmesi .....	26
3.2.1. Kullanılacak verilerin tespiti .....	27
3.2.2.1. Çekilen verilerin özellikleri ve kaydedilmesi .....	28
3.3. Geliştirilen yazılımın altyapısı .....	29
3.3.1. Yazılımın genel nitelikleri .....	30
3.3.2. Yazılım ile birlikte kullanılan api altyapısı .....	31
3.3.3. Yazılımın veritabanı yapısı .....	32
3.3.4. Twitter'dan çekilen verilerin analizi .....	35
3.3.4.1. Kampanyanın anatomisi.....	35
3.3.4.2. Kampanyanın toplam veri analizi .....	36
3.3.4.3. Tweet atan kullanıcıların cinsiyet analizi.....	37
3.3.4.4. Tweet'lerin duygu durum analizi.....	37
3.3.4.5. Tweetlerde kullanılan bağlantıların analizi.....	38
3.3.4.6. Tweetlerde diğer etiketlerin ve anahtar kelimelerin analizi.....	38
3.4. Analiz Verilerinin İşlenmesi .....	39
3.4.1. Analiz verilerinin saklanması.....	40
4. ARAŞTIRMA BULGULARI .....	31

4.1. Etiket analizi isteđi (iři) oluřturma .....	41
4.2. Kampanya anatomisi analizi .....	42
4.3. Cinsiyet analizi.....	47
4.4. Duygu analizi .....	47
5. TARTIřMA VE SONUÇLAR .....	49
KAYNAKLAR .....	51
EKLER.....	53
EK B. Yazılıma' a Ait Kodların Bir Bölümünün Gösterilmesi .....	54
ÖZGEÇMİř .....	60



## ÖZET

**Yüksek Lisans Tezi**

### **TWITTER ÜZERİNDE YAPILAN TÜRKÇE PAYLAŞIMLAR İÇİN ETİKET ANALİZ ARACI**

**Yunus Emre Karabulut**

**Süleyman Demirel Üniversitesi  
Fen Bilimleri Enstitüsü  
Bilgisayar Mühendisliği Anabilim Dalı**

**Danışman: Doç. Dr. Ecir Uğur Küçüksille**

Sosyal medya kullanımının hızla arttığı günümüzde, bu ortamda paylaşılan bilgilerin analiz edilerek kullanılması da kaçınılmaz hale gelmiştir. Elde edilen sonuçlar hemen hemen her sektörde kullanılmaktadır. Yapılan analizler sayesinde ulaşılmak istenen hedeflere daha kısa sürede ulaşılması nedeni ile bu alana olan ilgi her geçen gün artmaktadır.

En çok kullanılan sosyal medya ortamlarından biri de Twitter'dır. Twitter, kullanıcıların belirli bir karakter sınırlaması altında metin, resim, video ve anket paylaşımı yapabildiği bir sosyal medya platformudur. İnsanların mesajlarını kısa cümleler ve multimedya içerikleri yardımıyla hızlı bir şekilde paylaşabilmeleri nedeni ile dünyanın en çok kullanılan sosyal ağlarından biridir. Bu nedenle burada paylaşılan bilgilerin analiz edilerek sonuçlar elde edilmesi, araştırmacıların üzerinde yoğun bir şekilde çalıştığı ve çalışmaya devam ettiği bir konudur.

Bu çalışmada geliştirilen uygulama yardımı ile Twitter üzerinde kullanıcıların tweet metinlerinde kullandığı etiketler (#hashtag) üzerinden analiz yapılabilmektedir. Bu analiz yardımı ile meydana gelen toplumsal bir hareketin ve kampanyanın detaylarına ulaşılabilir. Oluşan kampanya ile ilgili atılan tweet sayısı, tweet atan gerçek kişi sayısı, kampanyanın başladığı zaman dilimi, kampanyayı ilk başlatan kullanıcı ve tweet, etkileşim sayısı, atılan tweetler üzerinden duygu analizi ve tweet atan kullanıcılar üzerinden cinsiyet analizi yapılabilmektedir. Bir kampanya için atılan tweetler içinde en çok kullanılan diğer etiket ve anahtar kelimeler çıkarılabilmektedir. Ayrıca atılan tweetler içinde kullanılan bağlantılar (url) üzerinde bir analiz yapılarak en çok yönlendirme alan sayfa detayına da ulaşılabilir.

Twitter Etiket Analizi, Etiket Analizi, Sosyal Medya Analizi

**2018, 60 sayfa**

## **ABSTRACT**

**M.Sc. Thesis**

### **HASHTAG ANALYZE TOOL FOR TURKISH SHARES ON TWITTER**

**Yunus Emre KARABULUT**

**Suleyman Demirel University  
Institute of Science  
Department of Computer Engineering**

**Supervisor: Doç. Dr. Ecir Uğur KÜÇÜKSİLLE**

The aim of this thesis is to develop a Twitter tag (hashtag) tracking and analysis tool. For this purpose, a web based software has been developed that draws tweets / sentences of target ethics via Twitter and analyzes them through them. With the developed software, detailed follow-up and analysis processes of a campaign launched on Twitter social network can be done.

After the analyzes made with the developed software, smart conclusions can be made and Twitter social network; detailed analysis of the events and the day can be taken.

As a result of the study, the details of the social movement and the campaign that are taking place on the Twitter social network can be met. Gender analysis can be done through the number of tweets pertaining to the resulting campaign, the number of actual tweets, the time zone at which the campaign starts, the user who initiated the campaign first, the number of interactions, the emotion analysis and tweet over the tweets thrown. The most commonly used tags and keywords in a tweet for a campaign are removed. In addition, in the tweets that are taken, users can reach the details of the page that receives the most referrals by analyzing the links (url).

Thanks to the developed software, details and analyzes that can not be accessed directly via Twitter or another channel can be accessed.

Twitter Hashtag Analyze, Hashtag Analyze, Social Media Analyze

**2018, 60 pages**



## TEŐEKKÜR

Bu arařtırma için beni yönlendiren, karşılařtıđım zorlukları bilgi ve tecrübesi ile ařmamda yardımcı olan deđerli danıřman hocam Doç. Dr. Ecir Uđur KÜÇÜKSİLLE' ye teőekkürlerimi sunarım.

Arařtırmanın yürütülmesinde maddi ve manevi yardımlarını gördüğüm deđerli arkadaşlarım Yahya ERSAN 'a teőekkürlerimi sunarım.

Tezimin her ařamasında bana yardımcı olan arkadaşlarıma ve aileme saygılarımı sunarım.

Yunus Emre KARABULUT  
ISPARTA, 2018



## ŞEKİLLER DİZİNİ

	<b>Sayfa</b>
Şekil 3.1. Json object yapısı .....	12
Şekil 3.2. Json array yapısı.....	12
Şekil 3.3. Json value yapısı .....	13
Şekil 3.4. Json string yapısı.....	13
Şekil 3.5. Json number yapısı.....	14
Şekil 3.6. Twitter api yapısı .....	31
Şekil 3.7. Postgresql veritabanı ve cache yapısı. ....	32
Şekil 3.8. Sistem kullanıcılarının işlemlerinin kaydediliği res_users tablosu.....	33
Şekil 3.9. Raporların oluşturulabilmesi için verilerin tutulduğu tablo yapısı. ....	34
Şekil 3.10.Kullanıcıların analiz kayıt ve türlerinin tutulduğu tablo yapıları.....	34
Şekil 3.11. Etiket analiz isteği (iş) oluşturma ekranı. ....	41
Şekil 3.12. Etiket analiz isteği bilgilerinin girilmesi.....	42
Şekil 3.13. Anatomi analizi birinci bölüm .....	43
Şekil 3.14. Anatomi analizi en çok beğeni alan tweetler. ....	44
Şekil 3.15. Anatomi analizi en çok retweet alan tweetler .....	44
Şekil 3.16. Anatomi analizi en çok kullanılan diğer etiket ve ifadeler .....	45
Şekil 3.17. Anatomi analizi tweetlerde kullanılan bağlantılar .....	45
Şekil 3.18. Anatomi analizi tweetlere gelen yanıtlar ve sayıları.....	46
Şekil 3.19. Anatomi analizi toplam sayılar .....	46
Şekil 3.20. Cinsiyet analizi sonuç ekranı .....	47
Şekil 3.21. Duygu analizi sonuç ekranı.....	48

## SİMGELER VE KISALTMALAR DİZİNİ

NLTK	Natural Language Toolkit
VTYS	Veri Tabanı Yönetim Sistemi
SQL	Structured Query Language
NT	
RE	Regular Expressions
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations
JSON	JavaScript Object Notation
URL	Uniform Resource Locator
HTML	Hypertext Markup Language
CSS	Cascading Style Sheets
MVC	Model View Controller
API	Application Programming Interface

## 1. GİRİŞ

İnternetin dünya üzerinden kullanımının yaygınlaşması ile birlikte insanların internet üzerinden birbirleriyle diyalog kurma, paylaşımında bulunma ihtiyaçlarının neticesinde sosyal medya platformları ortaya çıkmıştır. Bu platformlar her biri farklı kullanıcı ihtiyaç ve isteklerine odaklanarak çeşitli iletişim ve içerik paylaşımı yöntemleri sunmaktadır. Sosyal medya platformlarının sayısının artması ve sıkça kullanılması; insanların aradığı içeriklere, bağlantı kurmak istedikleri kişilere daha kolay ve hızlı şekilde ulaşabilmesine olanak sağlamıştır. Ayrıca bu platformlar üzerinden her bir kullanıcı bir başka kullanıcı ile metin, görsel, video, anket vb. çok çeşitli paylaşımlarda bulunabilmektedir. Zaman ve mekân sınırı olmadan paylaşımların yapılabildiği sosyal ağ platformları birçok kurum ve markanın da bulunduğu alanlar haline gelmiştir. Bu süreç beraberinde, sosyal medyanın kurum ve markalar için özellikle reklam ve doğru hedef kitleye ulaşmak amaçlı olarak kullanılması sonucunu doğurmuştur. Bu nedenle, sosyal medya platformları üzerinden hem veriler hem de kullanıcılar yönünden yapılan analizler her geçen gün önem kazanmaya başlamıştır.

Günümüzde farklı ihtiyaç ve amaçlara hizmet eden onlarca sosyal medya internet sitesi ve uygulaması bulunmaktadır. Dünya üzerinde en çok kullanılan sosyal ağlar araştırıldığında; Facebook, Twitter, Instagram, Swarm, Youtube ve Snapchat öne çıkmaktadır. Bu sosyal ağlardan özellikle Facebook ve Twitter tüm dünya üzerinde oluşturmuş oldukları büyük kullanıcı kitleleriyle öne çıkmaktadır. Facebook ve Twitter, diğer sosyal ağlara göre daha önce geliştirilmiş olmaları sebebiyle hem büyük bir kullanıcı kitlesine hem de bu kullanıcıların oluşturduğu milyonlarca veriye sahiptirler. Bu sosyal ağların kullanıcılardan edindiği verilerin her geçen gün katlanarak artması ile birlikte, elde edilen verileri doğru şekilde analiz etme ve kullanma ihtiyacı doğmuştur.

En büyük sosyal medya platformlarından biri olan Twitter, bir mikro blog sitesidir. İnsanların yakın bir zamana kadar 140 ve günümüzde 280 karaktere kadar metinler yazabildiği bir sosyal ağ platformudur. Twitter'ın tüm dünya üzerinde 330 milyondan fazla kullanıcısı bulunmaktadır. Twitter üzerinde dakikalar içinde milyonlarca gönderi paylaşılmaktadır. Ayrıca dünya üzerinde etiket (#hashtag) kullanımı ilk kez Twitter

üzerinde kullanılmış ve devam eden süreçte tüm sosyal ağlar bu özelliği kendi içinde kullanmaya başlamıştır. Kullanıcılar ilgi alanlarına göre etiketler üzerinden paylaşımlar yapmakta ve bu sayede aynı ya da benzer ilgi alanına sahip topluluklar biraraya gelebilmekte, markalar potansiyel müşterilerine ulaşabilmekte, sivil toplum kuruluşları da bu sayede hedef kitleleri ile etkileşime geçebilmektedir. Gelinen noktada etiketler, tüm dünyada aktif ve sürekli bir şekilde kullanılmaktadır.

Özellikle büyük veri alanındaki gelişmelerin de etkisiyle sosyal medya platformlarındaki veriler üzerinde yapay zekâ ve makine öğrenmesi algoritmaları kullanılarak analizler yapılması yönündeki çalışmalarda artış olmuştur. Bu bağlamda çeşitli sosyal medya analiz ve değerlendirme uygulamaları piyasaya sürülmüştür. Şu an dünya üzerinde en çok kullanılan araçlardan bazıları Keyhole, RiteTag, Hastagify vb. araçlardır. Google, Mashable, Uber, Loreal gibi dünya markaları bu ve benzeri analiz yazılımlarını kullanmaktadır. Bu araçların kullanılması sonucunda elde edilen başarılar da hem firmaların bu alana olan ilgisini artırmış hem de bu alanda yapılan çalışmaların artmasına katkıda bulunmuştur.

Bu çalışmada geliştirilen yazılımda daha önce yayınlamış uygulamalardan farklı olarak birden fazla analiz çeşidi aynı iş altında gerçekleştirilebilmekte ve raporlanabilmektedir. Bu analiz çeşitleri bir etiket kampanyasının sayısal detayları yani anatomi analizi, tweetler üzerinde duygu analizi ve tweet atan kullanıcıların cinsiyet analizi şekliyle gruplandırılmıştır. Twitter analiz araçları arasında en çok kullanılan Keyhole, RiteTag, Hastagify vb. araçların her biri belirli bir analiz türüne odaklanmıştır. Bu araçlardan bazıları, sadece bir konu ile ilgili en çok kullanılan anahtar kelime ve etiketleri listelerken bazıları da girilen bir etiket ile ilgili kaç kişi, kaç tane tweet göndermiş, toplamda ne kadar tweet atılmış vb. gibi verileri sunmaktadır. Yapılan çalışma ile geliştirilen analiz yazılımı ise bir etiket ile ilgili hem sayısal analizleri hem duygu analizini hem de cinsiyet analizini gerçekleştirebilmektedir. Geliştirilen uygulamada; PostgreSQL veri tabanı yönetim sistemi, Python programlama dili ve web altyapısı için Odoo Framework kullanılmıştır.

Birinci bölümde materyal ve metotlardan bahsedilmiştir. Devamında ise, yazılımsal olarak yapılan işlemler ve süreçler açıklanmıştır. Son bölümde ise yapılan analizler ve elde edilen sonuçlar ekran görüntüleriyle birlikte gösterilmiştir.



## 2. KAYNAK ÖZETLERİ

Hao çalışmasında (2011), Twitter'ın şu anda insanların geniş bir konu yelpazesıyla ilgili yorumlarını paylaştıkları, günde yaklaşık 190 milyon tweet atılan bir platform olduğundan bahsetmiştir. Twitter'ın çok sayıda tweet, ürün ve hizmetler hakkındaki fikirleri içerdiğini belirtmiştir. Bununla birlikte, Twitter'ın nispeten yeni bir fenomen olmasından dolayı, bu gönderilerin, müşteri duyarlılığını değerlendirmek için bir kaynak olarak yetersiz kaldığına vurgu yapmıştır. Yüksek hacimli twitter verilerini keşfetmek için, üç yeni zaman bazlı görsel duyarlılık analizi tekniği sunmuştur: Bunların ilki müşteri görüşlerini ayıklayan, haritalayan ve ölçen konu bazlı duyarlılık analizi; ikincisi İlginç tweetleri yoğunluklarına, olumsuzluklarına ve etki özelliklerine göre tanımlayan akış analizi; ve üçüncüsü büyük hacimli veriyi tek bir görünümde görselleştiren piksel hücresi tabanlı duyarlılık takvimleri ve yüksek yoğunluklu coğrafi haritalar. Bu teknikleri, dağıtımlarını ve kalıplarını göstermek ve etkili fikirleri belirlemek için çeşitli heyecan verilerine (ör. Sinema, eğlence parkları ve oteller) uygulamıştır.

Agarwal vd. çalışmalarında (2011), Twitter verilerinde duygu analizini incelemiştir. Twitter üzerinde olaylar ve etiketler üzerinde analizler yapılarak, insanların ya da belirli bir grubun, olay karşısında duygusal tutumlarına yönelik çıkarımlar elde edilmesi amaçlanmıştır. Tweet metinleri üzerinde yapay zeka ve veri madenciliği ile duygusal durumların anlaşılması üzerine bir çalışma yapmışlardır.

Borgmann, Woelm vd. Çalışmalarında (2016), bir sosyal medya platformu olan Twitter üzerinde bir etiket üzerinde analiz gerçekleştirerek katılımcıların tweet stratejilerini ve tweet içeriğini analiz etmeyi amaçlamıştır. Bu çalışma için örnek uygulama olarak 2013'te gerçekleştirilen Avrupa Birliği Üroloji Kongresi (European Association of Urology Congress) ile ilgili olarak belirlenen ve kullanıcıların tweet atarak katılımında bulunduğu #eau2013 etiketi baz alınmıştır. Çalışma sonucunda kullanıcıların kongrenin içeriklerinden ve konu başlıklarından en çok hangileri ile ilgili konuştuğu ve içerik oluşturduğu analizi yapılmıştır.

Terpstra ve Vries çalışmalarında (2012), Twitter'ın gerçek zamanlı ve otomatik analiz imkânlarını araştırmışlardır. Krizler sırasında gönderilen –etikete ait- mesajları incelemiştirler. Analizi bir bilgi çıkarımı uygulayarak gerçekleştirmişlerdir. Belçika'da Pukkelpop 2011 festivalinde çıkan fırtına olayı sırasında ve sonrasında kısa bir süre önce gönderilen yaklaşık 97.000 tweet olduğunu tespit etmişlerdir. Fırtına festivali vurduğunda tweet etkinliği artmış ve üssel olarak dakikada 576 çarpana ulaşmıştır. Extraction Tool ile tweetleri analiz etmişlerdir. Ortaya çıkan önemli konular arasında; 'erken uyarı tweetleri', 'söylentiler', 'selforganizasyon' ve 'Afet Sorunlarının Giderilmesi' vardır. Sonuçlar, bilginin otomatik filtrelenmesinin operasyonel tepki ve kriz iletişimi için değerli bilgiler sağladığını ortaya koymuştur.

Wang, Can çalışmalarında (2012), 2012 ABD seçimlerinde, mikro bloglama hizmeti olan Twitter'da açıklanan başkan adaylarına yönelik kamuoyundaki duyguların gerçek zamanlı analizine yönelik bir sistem geliştirmişlerdir. Twitter'ın, insanların siyasi partiler ve adaylar hakkında görüşlerini dile getirdiği merkezi bir site haline geldiğinden bahsetmişlerdir. Geleneksel içerik analizinin tamamlanmasının günler veya haftalar aldığını, önerilen sistemde ise, seçimle ilgili tüm tweetlerde duygu analizi yapılarak sonuçların anında ve sürekli olarak izlenebildiğini belirtmişlerdir.

Li, Lei vd. çalışmalarında (2012), Twitter üzerinden bir olayın yani kampanyanın tespit edilmesi, bu olay veya kampanyanın mekansal ve zamansal olarak gelişiminin çıkarılması hedeflerini gerçekleştirmek için Twitter özelinde bir "Olay Algılama ve Analiz" sistem geliştirmeyi hedeflemiştir. Bu çalışmalarında geliştirdikleri sistem ile ilgili olarak sistemin genel mimarisini, tweetleri arayan ve tasnif eden yapıyı ve tweet metinlerinden kullanıcıların konum bilgilerini elde eden uygulamalarını sunmuşlardır.

Bruns çalışmasında (2013), Twitter tabanlı iletişimin analizi için yaygın olarak uygulanabilir, standartlaştırılmış metriklerden oluşan bir katalogu özetleyerek Twitter üzerindeki iletişimin analiz edilmemesi sorununa hitap etmiş, özellikle hashtaglere odaklanmıştır. Ayrıca, çeşitli ölçütlerin daha kapsamlı karşılaştırmalarının nasıl elde edebileceğine ilişkin bir gösterge sunarak, bu metriklerin potansiyel kullanımlarını da işaret etmiştir.



Bügel çalışmasında (2013), sosyal medyanın, afetler sırasında veya sonrasında doğal felaketleri tespit edebilecek ve kullanıcıların kriz yönetimini destekleyebilecekleri, olay temelli erken uyarı sistemleri için gittikçe artan bir kaynak haline geldiğini belirtmiştir. Bu çalışma aşamasında, acil durumlar için çok dilli Twitter yayınlarını analiz etme problemlerini incelediğini belirtmiştir. On deprem olayının araştırıldığını ve olayla alâkalı olmayan ilgisiz mesajları filtreleyerek depremleri tespit etmek için kullanılacak dört dile özgü sınıflandırıcıyı tanımladıklarını belirtmiştir. Nihai hedefin, bu çalışmayı daha fazla Akdeniz ülkesi diline yaymak, ilgili bilgileri tweet'ten sınıflara ayırmak ve anahtar kelimeleri tercüme etmek olduğu belirtilmiştir.

Goonetilleke, Sellis vd. çalışmalarında (2014), Twitter verilerini çekmek, saklamak, sorgular ile analizler elde etmek için bir yazılım geliştirme üzerine çalışmışlardır. Mevcut analiz araçlarını da inceleyerek daha yapılandırılmış ve metodolojiye sahip bir araç geliştirmeyi hedeflemişlerdir.

Burnap vd. çalışmalarında (2015), sosyal medyayı; başkalarıyla iletişim kurmak ve kişisel görüş ve eylemlerini belgelemek için giderek artan sayıda kişinin kullandığını söylemişlerdir. Sosyal medyada; çevrimiçi sosyal formasyonlar hakkında bir fikir edinmek için sosyal bilimcilerin çevrimiçi araştırma yöntemlerini kullanma fırsatı sağlayan önemli bir veri akışı oluşturduğunu belirtmişlerdir. Bu çalışmada elde edilen bilgiler, İngiliz polis teşkilatının tanımladığı sosyal gerginlikteki artışları "sosyal medyada bireyler veya gruplar arasındaki normal ilişkinin ciddi şekilde bozulduğunu gösteren eğilimleri" öngörme olasılığını ortaya koymuştur. Birçok Premier Lig futbol maçı sırasında ırksal kötüye kullanım suçlamasıyla ilgili, Twitter'dan toplanan insan özelinde bir veri örneğini kullanarak gerginlik artışlarını saptamak için çeşitli, farklı hesaplama yöntemleri denenmiştir. Konuşma analizi, söz dizimsel ve sözlüğe dayalı veri madenciliği kurallarıyla birleştirilmiş; Duygu analizi ve makine öğrenme yöntemleri olası bir yaklaşım olarak test edilmiştir.

Hoerber, Meseery vd. çalışmalarında (2016), Twitter'dan elde edilen verilerle görsel bir analiz ortamı oluşturmayı hedeflemişlerdir. Bu çalışmada özellikle kullanıcıların belli konularda duygu değişimlerini takip etmeyi ve daha özelde bunu spor ile ilgili tweet metinlerinde yapmayı amaçlamışlardır. Çalışma sonucunda ortaya bir görsel analiz yazılımı çıkarmışlardır. Geliştirilen yazılım ile popüler bir spor müsabakası ile ilgili Twitter üzerinde oluşturulan içerikler analiz edilmekte ve görsel olarak sunulabilmektedir.



### **3. MATERYALLER VE METOT**

#### **3.1. Kullanılan Materyal ve Teknolojiler**

Bu bölümde Twitter Profesyonel İzleme ve Takip Aracı'nın geliştirme sürecinde kullanılan materyal ve teknolojilere yer verilmiştir.

##### **3.1.1. Python Programlama Dili**

Python programlama dili Guido Van Rossum adlı Hollandalı bir yazılım geliştirici tarafından 1990'lı yılların başında geliştirilmeye başlanmıştır.

Python, ilk sürümünden (1990 yıllarından) bu yana dikkatleri üzerine çeken genel amaçlı, yorumlanabilen, nesne tabanlı, temiz kod yazılabilen, hızlı öğrenilebilen, hızlı yazılım geliştirilebilen ve birçok hazır aracı (kütüphaneyi) içinde barındıran bir programlama dilidir (Mustafa Başer, 2001).

Python diğer programlama dillerinin birçoğunda olduğu gibi hem masaüstü hem de web tabanlı olarak nesneye dayalı şekilde yazılımlar geliştirilmesine olanak sağlar.

##### **3.1.2. Tweepy Twitter modülü**

Tweepy Modülü, Python'da twitter için geliştirmeler yapmayı sağlayan bir kütüphanedir. Twitter API'ı kullanılarak yazılmıştır. Tweepy modülü, Python yazılım dili ile birlikte kullanılarak Twitter'ın geliştiriciler için sunmuş olduğu API hizmetine erişim sağlamaktadır. (Twitter Documentation, 2014) Bu çalışmada, Twitter üzerinden veri çekme işlemleri yapılırken kullanılmıştır.

##### **3.1.3. PostgreSQL veri tabanı yönetim sistemi**

PostgreSQL veri tabanı, veri tabanları için ilişkisel modeli kullanan ve standart Sql sorgu dilini destekleyen ve Sql ile sorguların yapılabildiği bir veri tabanı yönetim sistemidir.

PostgreSQL aynı zamanda birçok veri tabanı yönetim sistemine göre iyi performans veren, güvenli ve gelişmiş özellikleri olan bir veri tabanı yönetim sistemidir. Hemen hemen tüm Unix ya da Unix türevi (Linux, FreeBSD gibi) işletim sistemlerinde çalışabilmektedir. Ayrıca NT çekirdekli tüm Windows sistemlerde de çalıştırılabilir. PostgreSQL ücretsiz ve açık kodlu bir VTYS'dir.

PostgreSQL'in geliştirilmesi ve kullanıcılara açık hale gelmesi; 1977'de Berkeley Üniversitesi'nde Ingres adındaki ilişkisel veri tabanının geliştirilmeye başlanmasına dayanmaktadır. 1986-1994 yılları arasında bu veri tabanı yönetim sistemi Postgres adını almıştır. Daha sonra Illustra tarafından satın alınmış ve Informix olarak geliştirilmeye devam edilmiştir. 1994 yılında Sql sorgu dili özelliklerinin eklenmesiyle Postgres95 adını almıştır. 1996'da kod geliştirilmesi için elektronik posta listesi açılmış ve birçok gönüllü Postgres95'i geliştirme çalışmalarına başlamıştır. 1996 yılında Sql veri tabanı dili standartlarını desteklemeye başlamasıyla birlikte PostgreSQL adını almıştır (Wikipedia, 2016).

PostgreSQL'in güvenilirliği bugüne kadar yaygın kullanımı ve güçlü referansları sayesinde kanıtlanmıştır. Çıkarılan her yeni sürümü defalarca kontrollerden geçirilmiş ve her bir deneme sürümü en az bir aylık testlere tabi tutulmuştur. Geniş kullanıcı grubu ve kaynak koduna dünyanın her yerinden erişilebilir olması sebebiyle olası hatalar çok çabuk kapatılabilmektedir (Wikipedia, 2016).

PostgreSQL'in başarısı ve popülaritesi her yeni sürümle birlikte artmaktadır. Son çalışmalar, PostgreSQL' in belirli koşullar sağlandığında diğer ticari veri tabanları ile aynı performansı verdiğini göstermektedir. Bu da daha fazla tercih edilmesini sağlamaktadır (Postresql, 2016).

Bu çalışmada performans, hız ve güvenlik önemli görüldüğünden bu kriterlere en uygun veri tabanı yönetim sistemi olarak PostgreSQL tercih edilmiştir.

### 3.1.4. Re (Regex) ile düzenli ifadeler

Regex yani düzenli ifadeler, günümüzde kullanılmakta olan modern programlama dillerinin bir parçasıdır. Tüm programlama dillerinde bu ifadeler aynı söz dizimine (syntax) sahiptir. Düzenli ifadeler (regular expressions) genellikle harflerden oluşan karakterler dizisinin belirlenen kurallar sınırında kısa yoldan ve esnek bir şekilde belirlenmesini sağlayan bir yapıdır. Bu tez çalışmasında düzenli ifadeler yoğun veriler içinden ihtiyaç duyulan bilgilerin çekilmesi, kullanıcı tarafından verilen girdinin denetimi ve çekilen verilerin kullanım amacına uygun şekilde getirilmesi için kullanılmıştır.

Düzenli ifadeler (regular expressions) hem özel hem de standart karakterler içerebilir. En çok kullanılan karakterler “A”, “a”, “0” ‘dır. Sıradan karakterler de kullanılabilir. (Python, 2018)

### 3.1.5. Odoo Framework

Odoo, Python ile geliştirilmiş bir yazılım çatısıdır. Dünyada iki milyonu aşkın kurumsal kullanıcısı bulunan Odoo, esnek yapısıyla son yıllarda farklı yazılımlar geliştirmede Python çatıları arasında önemli bir yer tutmaktadır. Odoo, Türkçe dâhil olmak üzere yirmi üç farklı dil desteğine sahiptir. Yazılım altyapısında Python, Xml, Xslt bulunmaktadır. Veritabanı ise PostgreSQL ile sağlanmaktadır.

Bu tez çalışması sonucu geliştirilen analiz yazılımında Odoo çatısı arayüz tasarımı ve verilerin düzenli gösterimi amacıyla kullanılmıştır. Odoo açık kaynak kodlu olması, esnek ve düzenlenebilir yapısı sebebiyle tercih edilmiştir. Odoo’nun içinde barındırdığı web sayfası düzenleyici birçok çatıya göre (Django vb.) daha iyi performans göstermesi ve daha iyi optimize edilmiş olması sebebiyle, geliştirilen yazılımın özellikle yönetici paneli arayüzünde tercih sebebi olmuştur.

### 3.1.6. Json Kütüphanesi

Json (JavaScript Object Notation) insanlar tarafından okunabilen bir veri deęişim biçimidir. Xml'in Javascript ile verimli olarak kullanılamamasından dolayı Json tercih edilmektedir. Json'un temel amacı, veri alışverişı yapılırken daha küçük boyutlarda veri alıp gönderilmesini sağlamaktır. Json sayesinde, Xml vb. uzun kodlar yerini anlaşılır ve sade kod satırlarına bırakmıştır (Json, 2006).

Json, tamamen programlama dillerinden bağımsızdır, ancak C ve türevi dillere (C, C++, C#, Java, JavaScript, Perl, Python, Php ve daha pek çoęu), yazılış bakımından çok benzeyen bir veri tanımlama formatıdır. Bu özellikler, Json'u veri deęiş tokuşu için ideal hale getirmektedir (Json, 2006).

JSON temelde iki yapı üzerine kurulmuştur:

Bu yapılar isim ve deęer çifti yapılarıdır. Bazı programlama dillerinde bu, object, record, struct, dictionary, hash table, keyed list veya associative array olarak da tanımlanmaktadır.

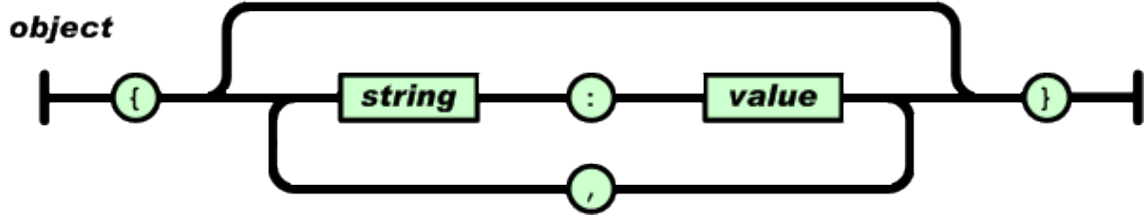
Bu durum programlama kavramları içinde sıralı deęer listesi olarak adlandırılmıştır. Çoęu programlama dilinde bu, "array, vector, list veya sequence" olarak tanımlanır (Json, 2006).

Bu yapılar, tüm dünya üzerinden geliştirciler tarafından kullanılan ve kabul görmüş evrensel veri yapılarıdır. Bütün modern programlama dilleri, bu yapıları, bir şekilde içlerinde barındırmaktadırlar. Programla dilleri arasında veri deęişimi için kullanılan bir formatın, bu yapıları kullanarak oluşturulması da oldukça verimli olmuştur.

Bu yapılar ve tanımlamaların Json'da gösterim biçimleri şu şekildedir:

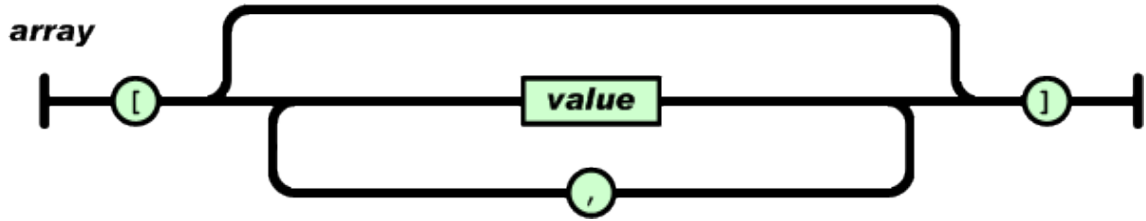
Json yapısında bir nesne (object), isim ve deęer çiftlerinin sırasız birleşiminden oluşmaktadır. Nesne tanımlaması, süslü parantez ile başlar ve süslü parantez ile biter. Her

"isim"den sonra iki nokta üst üste gelir ve isim değer çiftleri virgül işareti ile ayrılarak yazılır (Json, 2006). Şekil 3.1'de Json'da nesne, isim ve değer arasındaki ilişki gösterilmiştir.



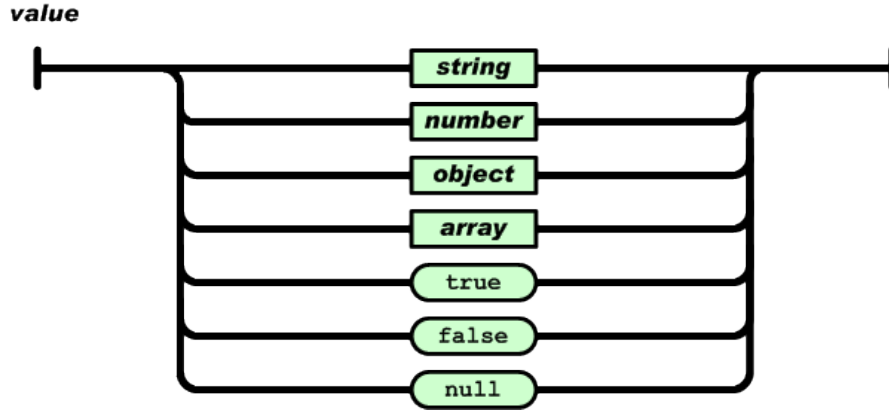
Şekil 3.1. Json object yapısı (Json, 2006)

Programlama dillerinde diziler (arrays), sıralı değerlerden oluşan listelerdir. Bir dizi köşeli parantez ile başlar ve köşeli parantez ile biter. Dizi içinde bulunan değerler, virgül işareti ile ayrılır (Json, 2006). Json array yapısı şekil 3.2'de gösterilmiştir.



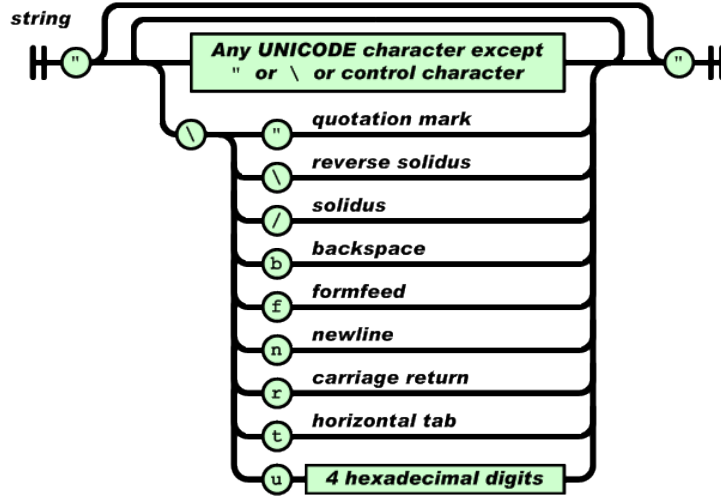
Şekil 3.2. Json array yapısı (Json, 2006)

Bir değer, çift tırnak içinde bir metin (string), sayı, doğru (true), yanlış (false), boş (null), nesne (object) veya dizi (array) olabilir. Bu yapılar aynı zamanda birbirlerinin içinde tekrar edebilirler (Json, 2006). Json'da kullanılan değer yapısı türleri ile birlikte şekilde 3.3'te gösterilmiştir.



Şekil 3.3. Json value yapısı (Json, 2006)

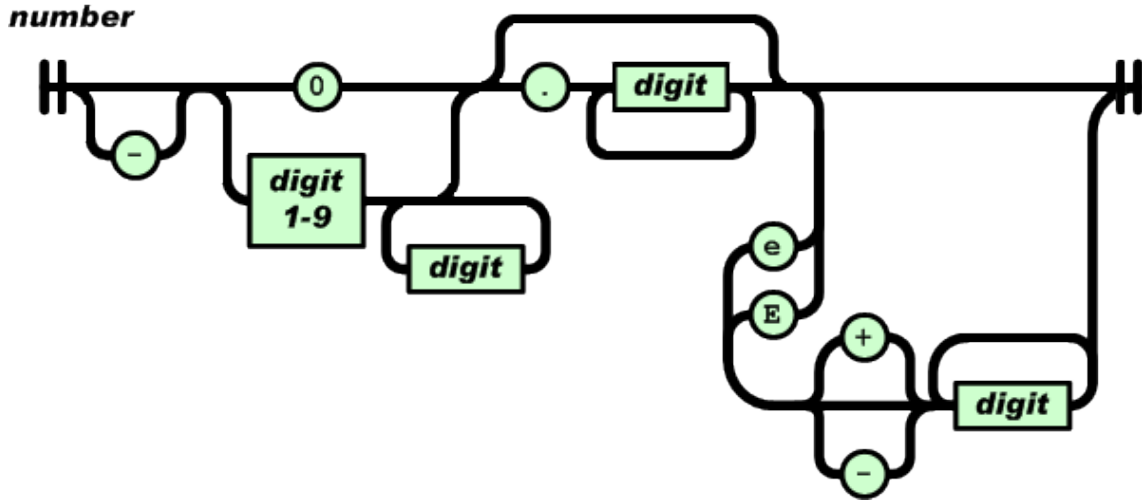
Bir metin (string), çift tırnak içinde, ters bölü (\) işaretlerine ve kod yazımını da izin veren, sıfır veya daha fazla "Unicode" karakterin birleşiminden oluşur. Yazı (string) tipindeki bir karakter tek karakter olarak gösterilebilir. Yazı (string) tipi, C, C++ ve Java dilindeki "string" tipiyle oldukça benzerlik göstermektedir (Json, 2006). Şekil 3.4'te Json string yapısı ve kullanımı görülmektedir.



Şekil 3.4. Json string yapısı (Json, 2006)



Json içinde bir sayı değeri C veya Java dilindeki sayılara çok benzemektedir. Ancak Json'da sekizli ve on altılı (hexadecimal) rakamlar kullanılmamaktadır (Json, 2006). Json'da sayı değerlerinin kullanımı anlatan şema şekil 3.5'te görülmektedir.



Şekil 3.5. Json number yapısı (Json, 2006)

Bu tez çalışması sonucunda geliştirilen araçta Json kullanım amacı Json'da dosya boyutunun XML'e göre küçük olmasıdır. Analizler için çekilen tweetler metin verilerden (string) oluşmaktadır. Xml yapısı gereği etiketlemelerden oluşmaktadır. Xml'de veriler iki etiketin arasında bulunur. Her etiketin başlangıç ve bitişinde etiket tekrarlandığı için büyük boyutta bir veri söz konusu olduğunda dosya boyutu bu tekrarlamalardan dolayı büyümektedir. Aynı büyüklükte bir veri Json'da Xml'e oranla çok daha az yer tutmakta ve dosya boyutu da küçülmektedir.

### 3.1.7. URLLib modülü

URL'leri yani bağlantıları getirmek için kullanılan bir Python modülüdür. Bağlantıların alınması ve listelenmesi için bir arayüz sunduğundan tercih edilmiştir.

Bu tez çalışmasında, URLLib ile farklı domainlerin tespiti için URL'lerin alınması sağlanmıştır.

### **3.1.8. NLTK (Natural Language Toolkit)**

NLTK, insan dili verileriyle çalışacak Python programları oluşturmak için geliştirilmiş bir platformdur. WordNet gibi 50'den fazla şirket ve metinsel kaynak için kullanımı kolay ara yüzler, sınıflandırma, belirleme, kaynak oluşturma, etiketleme, ayrıştırma ve semantik akıl yürütme için bir dizi metin işleme kütüphanesidir. (<http://nltk.org>, 2017)

Bu tez çalışmasında Türkçe tweetler üzerinden duygu analizi yapıldığından, NLTK Kütüphanesi doğal dil işleme amacıyla kullanılmıştır.

#### **3.1.8.1. Turkish Stemmer**

Türkçe sondan eklemeli bir dildir ve çok zengin bir morfolojik yapıya sahiptir. Türkçe'de, bir dizi sonek ekleyerek tek bir kökten çok farklı kelimeler oluşturabilir. Örneğin "doktoruymuşsunuz" kelimesi "Sen onun doktorusun" anlamına gelir. Kelimenin kökü "doktor" dur ve üç farklı son ek; -sU, -ymuş ve -sunuz alır. Turkish Stemmer Kütüphanesi, Türkçe dili için bir kelime kök bulma algoritmasıdır (Github, 2015).

### **3.1.9. Twitter sosyal ağ platformu**

Twitter, kullanıcıların 280 karakter ile sınırlı bir şekilde "tweet" (Türkçe çevirisi ile cıvılda) adı verilen metinler yazabildiği bir sosyal ağ platformudur. Twitter sosyal ağı Jack Dorsey, Noah Glass, Biz Stone ve Evan Williams tarafından Mart 2006'da kurulmuş olup, Temmuz 2006'da tüm dünyada internet kullanıcılarının kullanımına sunulmuştur. 2013'te en çok ziyaret edilen 10 internet sitesinden biri olmuş ve "internetin SMS'i" olarak tanımlanmıştır. Mart 2016'da 310 milyon aktif kullanıcıya ulaşmıştır (Twitter, 2006).

Twitter henüz açılmadan önce 2004 yılında, Twitter'ın da bir alt kuruluşu olarak kurulacağı Odeo isimli bir şirket faaliyete geçirilmiştir. 2005 yılında Odeo, "Podcasting" isimli servisini duyurmuş, fakat beklenen ilgiyi yakalayamamıştır. 2006 Şubat ayında

Jack Dorsey, Noah Glass, Evan Williams ve Biz Stone, insanların durumlarını, duygularını ve yorumlarını mesaj olarak paylaşabileceği bir proje üzerinde çalışmaya başlamışlardır. Bu çalışma sonrasında orijinal fikri Jack Dorsey'e ait olan "Durum" yapısı ortaya çıkmıştır. 2006 yılında son duruma geldiğinde Noah Glass'ın önerisiyle isim olarak Twitter belirlenmiştir (Ayan, 2016).

Twitter sosyal ağı, 2006 yılında Jack Dorsey tarafından geliştirilmesinden bu yana dünya çapındaki önemini ve popülaritesini gün geçtikçe arttırmıştır. Twitter içinde barındırdığı uygulamaların programlama ara yüzünün kısa mesaj gönderim ve alımı konusundaki olanaklarıyla internet dünyasının SMS'i yani kısa mesaj servisi olarak anılmaya başlamıştır. Twitter 25 Nisan 2011 tarihinden itibaren Türkçe olarak kullanılabilir hale gelmiştir. Twitter üzerindeki kullanıcıların etkisinin takipçi sayısından ziyade kullanıcının tweetlerine gelen retweetlere yani etkileşimlere bağlı olduğu ortaya konmuştur (Twitter,2006).

Twitter üzerinde kullanıcılar tarafından atılan tweet metinleri, eğer kullanıcı özel olarak profilini kapatmadıysa herkes tarafından görülebilir. Ancak kullanıcılar attıkları tweetleri sadece kendi takipçilerinin görebileceği şekilde sınırlandırabilir. Kullanıcılar diğer kullanıcıların tweet metinlerini takibe alabilirler. Üye oldukları kişi veya kurumların takipçi sayısını ve bunların kimleri takip ettiklerini görebilirler. Tüm kullanıcılar Twitter'ın internet sayfası aracılığıyla tweet metinlerini gönderebilir ve alabilirler. Kullanıcılar Twitter'ın websitesine uyumlu harici uygulamalar (akıllı telefonlar gibi) veya belirli ülkelerde erişilebilen kısa mesaj servisi tarafından giriş yapabilirler.

Günümüzde Twitter'da her dakika ortalama 98 bin tweet atılmaktadır. Yine yapılan bazı araştırmalara göre Twitter'da profilleri bulunan dev markaların; Google, Twitter, YouTube takipçilerinin %33'ü, Facebook takipçilerinin ise %44'ünün sahte veya spam profillerden oluştuğu tespit edilmiştir (Britannica, 2009).

Bu tez çalışmasında ilk olarak Twitter sosyal ağı üzerinde kullanılmaya başlamış olan etiket analizi özelinde bir analiz aracı geliştirmesi yapılmıştır.

### 3.1.10. Etiket (hashtag)

Bir etiket, hemen hemen bütün mikroblog ve sosyal ağ platformlarında bir sözcük ya da sözün başına diyez işareti (#) eklenerek oluşturulmaktadır. Facebook, Google Plus, Instagram, Twitter, Snapchat ve benzeri sosyal ağ sitelerinde diyez sembolünün kelime veya kelime grubunun önüne eklenerek oluşturulan bağlantılar, anlık bilgileri kategorilere ayırmaya ve büyük kitle ve gruplara ulaştırmaya yardımcı olmaktadır. Etiketlerle yalnızca belirli bir ortama, kategoriye bağlantı verilebilir; etiketler resim, video veya .gif formatına sahip dosyaların içerisinde bağlantı verememektedir. Bir etiket (hashtag), tüm bir cümle ya da cümle içinde birkaç kelimeyi vurgulayabilmektedir (Pbworks, 2011).

Örneğin: "Türkiye'yi sarsan 19 günün saat olarak detaylı şekilde dönüm noktaları twitter üzerinde başlatılan "#yaşarken yazılan tarih" etiketine gönderilerek bu etiket altında herkesin birbirinin paylaştığı mesajları görmesi sağlanmaktadır. Sosyal ağ platformlarının hemen hepsinde, istediğiniz her sözcük ya da sözü bu şekilde diyez işaretiyle yazılarak o etiket (hashtag) altında kişileri bir araya getirilebilir.

Etiket özellikle sosyal ağ Twitter üzerindeki tüm tweet metinlerini kategorize etmek ve Twitter arama sistemi üzerinden daha hızlı bulunmasını kolaylaştırmak için geliştirilmiş bir sistemdir.

Bu çalışma ile geliştirilen analiz aracının temel noktası da Twitter üzerinde kullanılan etiketlere ait verileri çekmek ve bu veriler üzerinde analizler gerçekleştirmektir.

### 3.1.11. Sosyal medya

Sosyal medya, interneti büyük gelişim aşamalarından biri olan Web 2.0'ın kullanıcı hizmetine sunulmasıyla birlikte, kullanıcıların tek yönlü bilgi paylaşımından, çift taraflı ve eş zamanlı bilgi paylaşımına ulaşmalarını sağlayan bir sistem ve geliştirmelerin bütünüdür. Ayrıca sosyal medya; kişilerin internet üzerinde birbirleriyle yaptığı diyaloglar ve paylaşımların tamamından oluşmaktadır. Sosyal ağlar, insanların birbiriyle

içerik ve bilgi paylaşımlarına imkân veren internet siteleri ve uygulamalar sayesinde, herkes aradığı, ilgilendiği içeriklere ulaşabilmektedir. Sosyal medya platformlarında küçük gruplar arasında gerçekleşen diyaloglar ve paylaşımlar, kullanıcı bazlı içerik (İngilizce: UGC-User Generated Content) üretimini giderek arttırmakta, amatör içerikleri dijital dünyada birer değere dönüştürmekte ve önemli bir konuma taşımaktadır (Poynter, 2012).

Sosyal medya platformları zaman ve mekân sınırlaması olmadan (mobil tabanlı), paylaşımın, tartışmanın, iletişim kurmanın temel nokta olduğu bir sosyal ve insani iletişim şeklidir. Sosyal medya platformlarında insanlarla buluşabilir, birlikte hareket edebilir ve iletişimde bulunabilir. Ayrıca insanlara yardım edilebilir, yardım alınabilir, sorularına cevap verilebilir ve sorular sorulabilir. Bu bakımdan sosyal medya resmi olmayan yani informal eğitim yollarından da bir tanesi olarak görülmektedir.

Sosyal ağların teknoloji, telekomünikasyon ve sosyal iletişimin kelimeler, görseller, videolar, ses dosyaları yolu ile sağlandığı bir yapıya sahip olduğu görülmektedir. İnsanlar hikâyelerini ve tecrübelerini bu bağlamda paylaştığı bir çerçeveye de sahiptir.

Sosyal medya, geleneksel medyada olduğu gibi tek bir noktadan çok sayıda kişiye yayın yapmaktan uzaklaşmış, iletişim yönü açısından çoktan çoğa yapısına dayalı "kullanıcı kaynaklı medya" özelliğine sahip bir şekilde büyümesini sürdürmüştür ve sürdürmektedir. Kullanıcı kaynaklı medya yani günümüzde bilinen adıyla sosyal medya geleneksel medyadan farklı olarak, içeriğin bizzat kullanıcıların kendileri tarafından üretilmesine ve paylaşılmasına olanak vermektedir. Sosyal medya aynı zamanda "Kullanıcıların Ürettiği İçerik" ve "Müşterilerin Ürettiği Medya" tanımlarını da ortaya çıkarmış, bu yapısıyla da ticari plandaki özellikle pazarlama alanındaki anlamını güçlendirmiştir (Poynter, 2012).

Ayrıca farklı bir bakış açısı ile Andreas Kaplan ve Michael Haenlein'in 2010'daki tespitlerine göre Sosyal Medya; "İnternetin yeni hali Web 2.0 üzerinde ideolojik ve teknolojik içeriklerin, yapılanmalarını ve niteliklerini kullanıcı merkezli bir şekilde

üretilmesine, yayınlanmasına ve geliştirilmesine izin veren internet tabanlı uygulamaların bütününe” denilmektedir (Poynter, 2012).

Sosyal medya teknolojileri yani sosyal medya platformları, blog siteleri, iş ağları, kurumsal sosyal ağlar, forumlar, fotoğraf paylaşım ortamları, ürün ve hizmet değerlendirmeleri, sosyal imleme, sosyal oyun platformları, video barındırma ve siber kültür gibi birçok farklı şekli alabilir (Aichner vd. , 2015).

### **3.1.12. Sosyal medya platformu**

Sosyal medya platformları; kullanıcıların ortak ilgi alanları, sosyal aktiviteler gibi konularda birbirleri ile bağlantı kurarak sosyalleşmelerine olanak sağlayan mobil ve internet servislerinin üzerinde çalıştığı yazılımlardır. İçeriklerin, kişinin kendisi tarafından yayımlandığı, paylaşıldığı her tür interaktif platformun genel adıdır (Wikipedia, 2016).

Kullanıcıların kişisel deneyimlerini, yaptıkları aktiviteleri ve bir konu hakkındaki yorumlarını internet ortamında paylaşmaları, internet üzerinden kullanıcılara ve potansiyel müşterilere erişmek isteyen markalar için fırsatları beraberinde getirmektedir. Platform sağlayıcıların geliştirdiği bir yazılım türü olan Sosyal Medya Platformları (Social Media Platforms) da bu noktada önemli hale gelmektedir (Wikipedia, 2016).

Sosyal medya platformları üzerinde çeşitli bilgi, eğlence, yaşam, video, hobi, ilgi alanları, mesajlaşma, interaktif, etkinlik, yorum ve içerik paylaşım servisleri geliştirilmiştir ve geliştirilmektedir. Sosyal ağlar üzerinde sınırsız mobil ve internet servisi açılabilirdiği gibi platform sahibine de tüm servislerinin içeriklerini yönetme imkânı verilmektedir. Zengin içerikler ve galeriler, paylaşım ve satışa açılabilir. Sosyal ağlar üzerinde açılan servisler, İnteraktif ve çok yönlü erişim (Wap, Web, Video, IM ve Cep uygulamaları) ile uluslararası sosyal medya platformlarına entegre olabilmektedir (Wikipedia, 2016).

İnternet üzerinde oluşturulabilecek dikey (vertical) ve yatay (horizontal) sosyal ağlar, marka ya da sektör bazlı özelleştirilmiş ve detaylı olarak hedeflenmiş reklam yayınlama

olanağı sağlamaktadır. Günümüzde sosyal medya platformları, markaların ve şirketlerin en çok tercih ettiği reklam kanalı olma özelliğini taşımaktadır.

Mobil internet kullanımının her geçen gün artması ve mobil kullanımın masaüstü internet kullanımını geride bırakması da sosyal medya platformlarının giderek mobile adapte olmasını sağlamıştır. Günümüzde sosyal ağ platformlarının çoğunun bir mobil uygulaması bulunmaktadır. Geliştirilen mobil uygulamalar, mobil oyunlar ve servisler sosyal medya platformlarına entegre olabilmektedir.

### **3.1.13. Sosyal medyanın özellikleri**

**Kolay erişilebilirlik:**

Bir sosyal medya mecrasının içeriğine, internet aracılığıyla her zaman ve her yerden erişim sağlanabilmektedir. Ancak bazı platformlar üyelik olmaksızın içerikleri görüntülenmesine engel olabilmektedir. Mobil internet kullanımının masaüstü internet kullanımını geçmesi ile sosyal ağ platformları, ilk günden mobil cihazlarla uyumlu geliştirilirler ve elektronik posta adresleri ile güncellenebilirler (Wikipedia, 2016).

**Katılım sağlama:**

Sosyal medya, iletişim kurulan kişilerin geri bildirimde bulunabilmesine ve gerekli katkılar sağlayabilmesine imkân vermektedir. Medya ile izleyicisi arasındaki çizgiyi silikleştirerek, patronajı kullananın hizmetine verir (Wikipedia, 2016).

**Konuşma ve topluluklar:**

Geleneksel medya bilindiği üzere tek yönlü bir iletişim içermektedir. Bu yüzden klasik kitle iletişim araçlarında geri bildirimde bulunmakta zordur ve zaman almaktadır. Ancak sosyal medya buna karşılık olarak kullanıcılara, çift yönlü iletişim, rahat ve zamanında geri bildirim sunar. Sosyal medya toplulukların ilgili konu veya kişiler üzerinde hızla oluşmasına ve buna bağlı olarak etkin bir şekilde iletişim kurulmasına izin verir (Wikipedia, 2016).

Bağlanılabilirlik:

Birçok sosyal medya platformu bağlanılabilirlik açısından kolaylık sunar. Bu duruma bağlantısallık denmektedir. Kullanıcılar ve markalar kolay bir şekilde, reklam veya herhangi bir sebepten dolayı diğer sitelere, kaynaklara ve kişilere bağlantı verebilirler (Wikipedia, 2016).

Programlanabilir olma:

Her sosyal ağ platformunun kendine özel ya da belli standartları barındıran bir uygulama geliştirme yapısı (API) bulunmaktadır. Bu sayede sürekli bir takım araçlarla, kullanıcılara ek işlevlerin ve yeni özelliklerin, fonksiyonların sağlanması mümkün olabilmektedir (Wikipedia, 2016).

Ölçeklenebilir:

Bulduğumuz zaman diliminde sosyal ağ platformlarının sayısı binlerle ifade edilmektedir. Bu platformların kullanıcı sayıları milyonları geçmiş olsa da pek çok yeni nesil teknolojiyi verimli ve daha etkin bir şekilde kullanarak büyük kullanıcı kitlelerine hizmet verebilecek kapasitededirler. Bu durum sayesinde sosyal ağ platformları kullanıcı sayısı, etkileşim sayısı, kullanıcılardan elde edilen içerikler vb. veriler üzerinden ölçeklenebilmektedir (Wikipedia, 2016).

Dinamik olma:

Sosyal medyanın en önemli özelliği de her şeyin su gibi akıcı, hızlı, dinamik, sürekli güncelleniyor ve değiştirilebiliyor olmasıdır (Gürsakar, 2009).

Günümüzde en çok kullanılan Facebook, Instagram, Youtube, Twitter, Snapchat gibi sosyal ağlarında dakikalar içinde milyonlarca veri akışı yaşanmakta ve her kullanıcının kendine özgü ana akışı bulunmaktadır.



### **3.1.14. Sosyal medya analizi**

İnternette marka, kurum veya kişilerin birden fazla ve birbirinden farklı odakları olan sosyal ağ platformlarında hesapları bulunmaktadır. Bu durum hem kuruluşlar hem de kişiler için birbirinden farklı platformlarda içerik üretimini, bu içeriklerin yönetimini ve planlamasını gerektirmektedir.

Sosyal ağ hesapları üzerinden içerik paylaşımları yapıldıktan sonra paylaşılan içeriğin başarısının ölçümü işlemleri devreye girmektedir. Sosyal ağlarda paylaşılan bir içeriğin eriştiği kişi sayısı, aldığı etkileşim sayısı, yorum sayısı gibi verileri ile yapılan çalışmalara genel tanımı ile sosyal medya analizi denmektedir.

Sosyal medya analizi, kişi veya kuruluşların sosyal medya üzerinde yaptıkları kampanyaların başarılı olup olmadıklarını ölçmek amacıyla gerçekleştirilmektedir. Sosyal medya analizi için Facebook, Twitter, Instagram, Youtube gibi sosyal ağların kendi içlerinde istatistik ve analiz sayfaları bulunmasına rağmen birçok marka ve kuruluş için bu alanlarda sunulan veriler yetersiz bulunmaktadır.

Bu durumdan ortaya çıkan ihtiyaçla birlikte özel sosyal medya analiz araçları geliştirilmiş ve çeşitli iş modelleri ile kullanıcılara sunulmuştur. Sosyal medya analizi gerçekleştirmekteki amaç kampanyaların daha başarılı olmasını sağlamak, doğru hedef kitleye doğru kampanya ile ulaşmak ve neticede daha fazla dönüşüm sağlamaktır.

Bu tez çalışmasında geliştirilen yazılım ise sosyal medya analizini Twitter üzerinde kullanılan etiketler üzerinden gerçekleştirmekte ve raporlamaktadır.

### **3.1.15. Duygu analizi (sentiment analyze)**

Duygu analizi için genel anlamda yapılan tanımlama, duygu analizinin bir metin işleme (text processing) işlemi olduğudur. Duygu analizi çalışmaları, girilen metnin duygusal olarak anlatmak ve belirtmek istediği durumu belirlemeyi amaçlamaktadır. Duygu analizinin ilk çalışmaları duygusal kutupsallık yani zıtlık (sentimental polarity) olarak

geçmekte olup verilen metni olumlu (positive), olumsuz (negative) ve son olarak nötr (belirsiz) olarak sınıflandırmayı amaçlamaktadır (Seker, 2016).

Zaman geçtikçe ve duygu analizi alanında yapılan çalışmalar arttıkça farklı duygu durumlarını belirten analizlere de izin vermiştir. Duygu durumlarını kodlamak ve geliştirme yapmak için çoğunlukla sınıf etiketleme yapısına benzer olarak veri kümesi üzerinde her metnin tek bir duygu ile etiketlenmesi veya metinlerde birden fazla duygunun etiketlenmesi şeklinde yaklaşımlar görülebilmektedir (Seker, 2016).

Literatürde bulunan çalışmaların büyük çoğunluğu bilgi getirme (information retrieval) alanı altında olup, zaman içinde bu çalışmaların yerini, duygu analizindeki metin etiket bağlantısının dışına çıkarak açık uçlu sonuçlara dönüştüğü fikir madenciliği (opinion mining) çalışmalarına bıraktığı görülmektedir (Seker, 2016).

Sosyal ağlardaki temel problemlerden biri de fikir veya kanaat madenciliği olarak literatüre geçmiş problemdir. Problem kaynaklarda bulunan konumu itibariyle duygu analizi çalışmaları (sentimental analysis) altında geçmektedir. Buna göre bir sosyal ağ verisinin (mesaj, paylaşım, duvar yazısı, haber v.b.) taşımış olduğu fikri anlam bilimsel olarak göstermek ve ifade etmek için yapılan çalışmaya fikir madenciliği denir. Fikir madenciliği için en önemli şartlardan biri, fikir veya kanaat oluşumunun bir topluluk ve bir kullanıcı grubu üzerinde inceleniyor olmasıdır. Özenle belirlenmiş ve özel olarak seçilmiş profesyonellerin fikirleri alınmadığı sürece fikir madenciliğinin ulaşmak istediği sonuç, bir topluluktaki bütün bireylerin fikirlerini anlayabilmek ve ifade edebilmektir. Ancak maalesef bütün bireylere ulaşmanın imkânsızlığı yüzünden genelde bu işlem örneklemelerle yapılmaktadır (Masud, 2006).

Fikir madenciliği (opinion mining) sırasında unutulmaması gereken konulardan biri de fikirlerin kişisel olduğudur. Yani fikir madenciliğinde doğru veya yanlış bir fikir aranmaz, fikir madenciliği mevcut durumun tespitine çalışır (Masud, 2006).

Genelde kaynaklarda ilk uygulamaları ve halen üzerinde en çok çalışılan uygulama duygusal zıtlıktır (sentimental polarity). Bu durumda metinler duygusal olarak olumlu veya olumsuz şeklinde iki zıt gruba ayrılmaya çalışılmaktadır (Masud, 2006).

Geliştirilen araçta Twitter üzerinde gerçekleştirilmiş bir etiket kampanyasına ait tweet metinleri çekilmekte ve bu metinler üzerinde duygu analizi gerçekleştirilmektedir. Gerçekleştirilen duygu analizinde temel olarak üç değer esas alınmıştır. Metinler üzerinden olumlu, olumsuz veya nötr olarak değerlendirme yapılmakta ve bu şekilde rapor sunulmaktadır.

### **3.1.16. Cinsiyet analizi (Gender analyze)**

Cinsiyet analizi toplanan veriyi oluşturan tüm kullanıcıların isimleri üzerinden gerçekleştirilen erkek ve kadın şeklinde belirleme yapılması üzerine geliştirilmiş analiz yöntemidir.

Bu tez çalışmasında belirlenen kampanya etiketi üzerinde başlangıçta Tweetler çekilmekte olduğu belirtilmiştir. Çekilen tweetler ile beraber bu tweetleri yazan kullanıcıların bilgileri de kayıt altına alınmaktadır. Bu kullanıcı bilgileri üzerinden kullanıcı adları ve görünen ad verileri üzerinden cinsiyet analizi yani kadın veya erkek tespiti yapılmaktadır.

Bu analiz sayesinde bir kampanyaya katılım konusunda erkek veya kadınların tepki verme durumları ve katılım durumları incelenebilmektedir.

### **3.1.17. Bağlantı (URL analyze) analizi**

Bağlantı analizi elde edilen bağlantıların ve özellikle kısaltılmış bağlantıların hedef sayfalarının tespiti işlemini gerçekleştiren analiz şeklidir. Bu analiz ile toplanan bağlantılar geliştirilen algoritma sayesinde arka planda tek tek açılarak sonuç olarak hangi sayfaya yönleniyorsa o sayfanın adresinin hedef uzantı ile birlikte kayıt altına alınması çalışmasıdır.

Twitter'ın kendine özgü bağlantı kısaltma servisi kullanması nedeniyle gönderilen tüm tweetlerde kullanılan url'ler benzer bir alan adı üzerinden görüntülenmektedir. Bu durumda geliştirilen araçta url analizi yapılması noktasında problem oluşturmaktadır. Bu yüzden bir bağlantı analizi algoritması geliştirilmiş ve çekilen tweetler içinde kullanılan url'lerin hedef sayfalarının tespiti sağlanmıştır.

Hedef sayfaların tespit edilmesi ile analizi gerçekleştirilen kampanya için gönderilen içeriklerde en fazla hangi alan adı, hangi sayfaya yönlendirme yapıldığı tespit edilebilmektedir.

Bu tez çalışmasında çekilen tweetler içinde çeşitli yönlendirmeler, bağlantılar kullanıldığından; bu bağlantıların analizinin yapılması ihtiyacını karşılamak amacıyla bağlantı analizi gerçekleştirilmiştir.

### **3.1.16. Anahtar kelime analizi**

Anahtar kelime analizi kullanıcıların tweet metinlerinde kullandıkları popüler kelimeleri ve kelime öbeklerini tanımlamayı içeren temel bir işlemdir. Anahtar kelimeleri analiz etmek ve popüler ifadeleri tespit etmek insanların neye daha fazla önem verdiklerini tespit etmekte yarar sağlamaktadır.

Ayrıca bu sayede kampanyayı gerçekleştiren kişi ya da kuruluş için hedef kitlesini daha doğru tespit etme ve tespit ettiği hedef kitleyle daha iyi şekilde etkileşime geçmesi için de çıkarımlar yapma imkânı sunulmaktadır.

Anahtar kelime analizi ile sadece belirli bir anahtar kelimenin kullanım sıklığını tespit etmekle kalmayıp, aynı zamanda kullanıcıların bir fikri veya bir konuyu savunma, belirtme durumlarını tespit etmek mümkün olmaktadır.

Bu tez çalışmasında herhangi bir etiket üzerinde analiz yapılırken çekilen tweet metinlerinde en çok kullanılan ifadeler kayıt altına alınmakta ve sunulmaktadır.

### **3.2. Twitter'dan verilerin çekilmesi**

Bu tez ile geliştirilen araçta, analizinin yapılması istenilen her etiket bir iş olarak kaydedilmektedir. Geliştirilen aracın arayüzünde öncelikle analizinin yapılması istenilen etiket için bir iş kaydı oluşturulur.

Analiz yapılması için kayıt oluşturulurken analizin hangi tarih aralıklarında yapılacağı belirlenir. Twitter'ın kısıtlaması nedeniyle tarih aralığında çalışma sistemi mevcut tarihten ya da girilen tarihten geriye doğru çalışmaktadır.

Her bir iş yani analiz isteği için aynı zamanda hangi analizlerin yapılacağı da seçilebilmektedir. Cinsiyet analizi, etiketin anatomisi ve duygu analizi şeklinde üç farklı analiz gerçekleştirilebilmektedir. İş kaydı oluşturulurken tarih ile birlikte istenilen analiz çeşitleri de işaretlenerek iş kaydı tamamlanır.

Geliştirilen araçta analiz süreci dört aşamada tamamlanmaktadır. İş kaydı oluşturulduktan sonra ilk süreç bekleme sürecidir. Burada alınan istekler tarih önceliğine göre sıralanmaktadır. Bir sonraki aşamada sırası gelen iş analiz işlem sürecine alınmaktadır. Bu aşamada analizi yapılması istenen etiket ile atılmış tweet metinleri belirtilen tarih aralıklarına göre tek tek çekilerek veritabanına kaydı yapılmaktadır. Çekilen her bir tweet metni için veri tabanında o tweet metnine ait özel alanlar oluşturulur. Tweet metnlerinin

kullanıcısı, tweet içinde kullanılan etiketler, bağlantılar, kullanıcı adları kaydedilir. Eğer daha önce yapılan bir analizde aynı kullanıcı oluşturulmuşsa güncelleme yapılır. Tweetlerin çekilmesi aşaması bittikten sonra analiz sürecine geçilir.

Analiz sürecinde iş kaydını oluştururken hangi analizlerin yapılması için tercihte bulunulmuşsa o analizler çekilen tweetler üzerinde uygulanmaya başlar. Gerçekleştirilen her analiz o iş kaydının altında kaydedilir ve istenildiği zaman analiz sonuçlarına ulaşılabilir.

Analiz süreci de tamamlandıktan sonra biten işlemler aşamasında tüm analizler iş kaydı altına kaydedilerek tüm süreç tamamlanmış olur.

### **3.2.1. Kullanılacak verilerin tespiti**

Bu tez çalışmasında twitter üzerinden tweetler çekilirken ve bu tweetlere ait bilgiler kaydedilirken Twitter'ın geliştirici alanında sunmuş olduğu veriler üzerinden seçim yapılmış ve yapılmak istenen analizlere en uygun veriler kaydedilmiştir.

Twitter üzerinden veriler çekilirken Twitter'ın geliştiricilerin kullanımına açmış olduğu "Developer Twitter" alanı üzerinde önce bir geliştirici hesabı açılmıştır. Geliştirici hesabı aktif edildikten sonra bu hesap üzerinden bir uygulama oluşturulmuştur. Oluşturulan uygulama üzerinden Twitter üzerinden verilerin çekilmesini sağlayacak Twitter API'a bağlantı sağlanabilmesi için gerekli anahtarlar ve kodlar elde edilmiştir.

Ayrıca Twitter'dan veri isteğinde bulunabilmek için gerekli istek ve sorgu kodları da yine twitter geliştirici platformu üzerinden sağlanmıştır.

Örnek istek kodu:

```
$ curl --request GET  
--url 'https://api.twitter.com/1.1/search/tweets.json?q=nasa&result_type=popular'  
--header 'authorization: OAuth oauth_consumer_key="consumer-key-for-app",
```

```
oauth_nonce="generated-nonce", oauth_signature="generated-signature",  
oauth_signature_method="HMAC-SHA1", oauth_timestamp="generated-timestamp",  
oauth_token="access-token-for-authed-user", oauth_version="1.0"
```

Twitter üzerinden çekilen tweetler ait veriler arasında dil, lokasyon, kullanıcı kimliği (user id), kullanıcı adı, görünen ismi, tweet metni, tweetlerin beğeni sayısı, tweetlerin retweet sayısı, tweetlerin hangi platform ve kaynak üzerinden gönderildiği, tweet tarih ve saati verileri çekilebilmiştir.

Twitter API ile sağlanan bağlantı için bu tez konusunda geliştirilen aracın yazılım dili olan Python ile uyumlu Tweepy Kütüphanesi kullanılmıştır. Bu şekilde yukarıda örneklendirildiği şekilde verilerin çekilmesi işlemi sağlanmıştır. Bu verileri düzenli şekilde altyapısı Postgresql olan veritabanına kaydedilmektedir.

Twitter üzerinden çekilen tweetlerden elde edilmek istenen verilerin çekilmesinde Twitter api gönderilen istekle eşleşen tüm tweetleri göndermektedir. Twitter'ın arama api'si ile tweetler doğru bir şekilde çekilmiştir. Çekilen tweetler ile birlikte aynı zamanda analizlerde kullanılacak olan diğer verilerde elde edilmiştir.

### **3.2.2.1. Çekilen verilerin özellikleri ve kaydedilmesi**

Bu tez çalışmasında geliştirilen yazılım ile Twitter'dan çekilen veriler aşağıda gösterilmiştir.

- Dil: Çekilen tweetlerde kullanılan dilin bilgisini içerir.
- Lokasyon: Çekilen tweetlerde eğer tweet atan kullanıcı lokasyon ayarını açık tutmuşsa tweetin atıldığı konum bilgisini içerir.
- Tweet kimliği: Twitter üzerinden atılan her tweetin kendine özel eşsiz bir numarası bulunmaktadır. Bu bilgiyi içerir.
- Kullanıcı kimliği: Twitter üzerinde her kullanıcının eşsiz bir kimlik numarası (user id) bulunmaktadır. Bu bilgiyi içerir.
- Beğeni sayısı: Çekilen her tweetin sahip olduğu beğeni sayısı bilgisini içerir.

- Retweet sayısı: Çekilen her tweetin sahip olduğu retweet sayısı bilgisini içerir.
- Tarih ve saat: Twittter üzerinden çekilen her tweetin gönderildiği tarih ve saat bilgisini içerir.
- Kaynak: Tweetlerin hangi cihaz ya da kaynaklar üzerinden gönderildiği bilgisini içerir. (iOS, Android vb.)

### 3.3. Geliştirilen yazılımın altyapısı

Bu tez çalışması sonucunda geliştirilen yazılımın programlama dili Python'dur. Yazılımda arayüz geliştirmesi için Python ile uyumlu Odoo kütüphanesi kullanılmıştır.

Yazılım ile Twitter'a bağlantı sağlamak ve Twitter'dan istekte bulunarak veri çekmek amacıyla Python programlama dili ile uyumlu olan ve Twitter API bağlantısını sağlamak amacıyla Tweepy kullanılmıştır.

Geliştirilen yazılım ile çekilen tüm verilerin düzenli şekilde kaydedilmesi ve analize hazır hale getirilmesi için gelişmiş ve sorgu sonuçlarını hızlı bir şekilde listeleyebilen Postgresql veritabanı yönetim sistemi kullanılmıştır. Yazılım veritabanı bu yönetim sistemi esaslarına göre düzenlenmiştir.

Çekilen tweetlerde metin veri olduğundan bu metin verilerin düzenlenmesi, doğru bir şekilde listelenmesi için Re – Regex kütüphanesi kullanılmıştır.

Analiz aşamasında aracın özelliklerinden biri olan tweetlerin sahibi kullanıcıların cinsiyet analizinin yapılabilmesi için genderize.io sınıfı yazılmıştır. Yazılımda çekilen tweetler içinde kullanılan bağlantıların (url) analizi de yapılmaktadır. Bu bağlantı analizinin yapılabilmesi için URL Lib kütüphanesinden yararlanılmış ve yazılım altyapısını entegre edilmiştir.

Geliştirilen yazılımda duygu (sentiment) analizi yapılması nedeniyle doğal dil işleme çözümlerine başvurma gereksinimi doğmuştur. Bu sebeple Python programlama dili uyumlu NLTK (Natural Language Toolkit) kütüphanesi ve bu kütüphane ile entegre



edilerek Türkçe kelime köklerinin tespit edilebilmesi için Turkish Stemmer (Github,2015) yapısı kullanılmıştır.

### **3.3.1. Yazılımın genel nitelikleri**

Bu tez sonucu geliştirilen Twitter Profesyonel İzleme ve Takip Aracı yazılımı web tabanlı olarak çalışmaktadır.

Yazılım web üzerinde kararlı bir şekilde çalışabilmesi için Dedicated Server olarak adlandırılan paylaşımsız bir sunucu kullanılmıştır. Yazılım geliştirme sürecinde açık kaynak kodlu bir dil olan Python tercih edildiğinden sunucu tarafında da uyumlu ve kararlı bir yapı sağlanması için işletim sistemi olarak Ubuntu 17.10 sürümü tercih edilmiştir. Kullanılan sunucu 12 Gb Ram belleğe sahiptir.

Yazılımı web tabanlı olarak geliştirilmiş olduğundan ve bir sunucuda tutulduğundan yazılıma erişimde güvenliği sağlamak amacıyla bir giriş ekranı geliştirilmiş ve kullanıcı girişiyle yazılım paneline erişim sağlanabilir hale getirilmiştir.

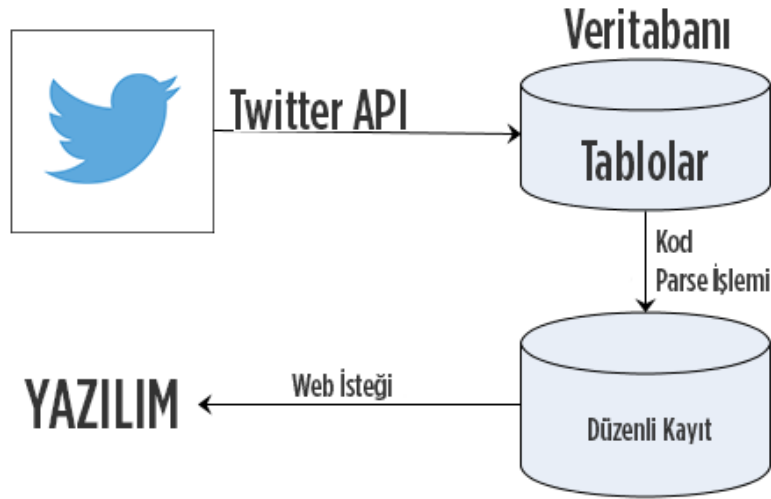
Yazılım MVC yapısı ve standartlarına sadık kalınarak geliştirilmiştir. Yazılımın web tabanlı olması sebebiyle hızlı çalışabilmesi için önbellekleme, html ve css sıkıştırma, veritabanı optimize etme çalışmaları da yapılmıştır. Bu sayede geliştirilen yazılıma ait başlangıç sayfası ve diğer altsayfalarda erişim ve gezinti hızlı bir şekilde yapılabilmektedir. Yazılımın ilerleyen süreçte son kullanıcıya da açılacağı ihtimali düşünülerek gelen trafik karşısında kararlı çalışabilmesi, yavaşlama sorunu yaşamaması adına tarayıcı önbellekleme sistemi de kullanılmıştır.

### 3.3.2. Yazılım ile birlikte kullanılan api altyapısı

Yazılım ile Twitter'dan veri çekebilmek için Twitter'in arama api'na bağlanma gereksinimi duyulmuştur. Api ile bağlantı gerçekleştirilerek Api'dan istekte bulunup tweetler çekilebilmektedir.

Twitter Api kullanımı için öncelikle Twitter'in geliştirici platformu üzerinde bir hesap açılmıştır. Açılan hesap aktif hale geldikten sonra hesap üzerinden bir uygulama oluşturulmuştur. Oluşturulan uygulama içinde Twitter üçüncü taraf yazılımların veri çekmesini kolaylaştırmak için api bağlantı anahtarlarını vermektedir. Bu bağlantı anahtarları alınarak bu tez sonucu geliştirilen yazılımda kullanılan Tweepy kütüphanesi üzerinden Twitter api ile bağlantı sağlanmıştır.

Tweepy'ye verilen api anahtarları sayesinde Twitter arama fonksiyonuna erişim sağlanabilmektedir. Bu erişim sayesinde etiket, kullanıcı aramaları yapılabilmektedir. Şekil 3.6 'da Twitter API ile yazılım arasındaki bağlantı gösterilmiştir.



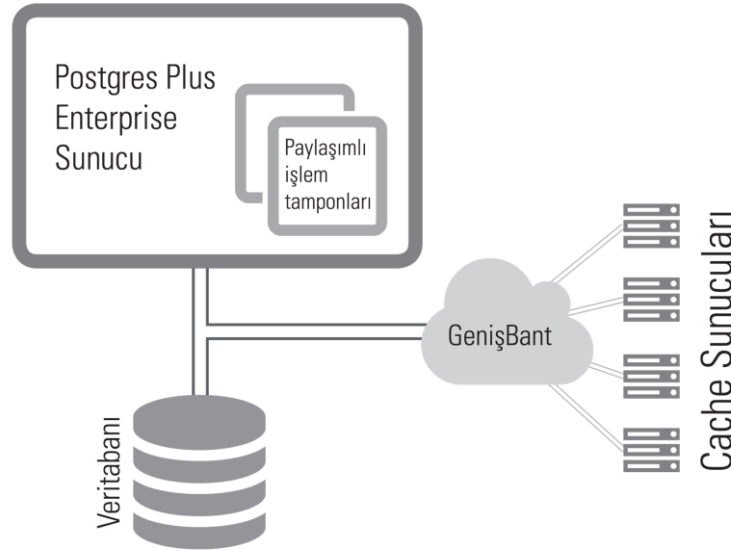
Şekil 3.6. Twiter api yapısı

### 3.3.3. Yazılımın veri tabanı yapısı

Bu tez sonucu geliştirilen yazılımda veritabanı yönetim sistemi olarak Postresql kullanılmıştır. Veri tabanı dili ise sql'dir.

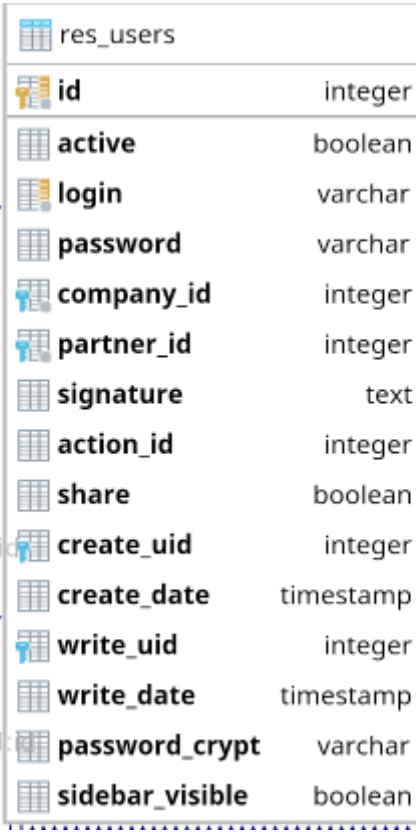
Postresql veritabanı yönetim sistemi yapısı itibariyle MySQL ve diğer açık kaynak kodlu veritabanı yönetim sistemlerine göre daha iyi performans göstermesi sebebiyle tercih edilmiştir. Ayrıca sorgu işlemlerinde gösterdiği performans nedeniyle Postresql kullanılmıştır.

Geliştirilen yazılımda iş kaydı aşamasından sonra işleme alınan her iş kaydı işlem sürecinde tweetler çekilmektedir. Çekilen tweetler direkt olarak veritabanına kaydedilmektedir. Yazılımda hiçbir analiz sonucu işlem tamamlanıp rapor alındıktan sonra silinmemektedir. Şekil 3.7'de PostgreSQL veri tabanı ve sunucu arasında bağlantı gösterilmiştir.



Şekil 3.7. Postresql veritabanı ve ön bellek (cache) yapısı

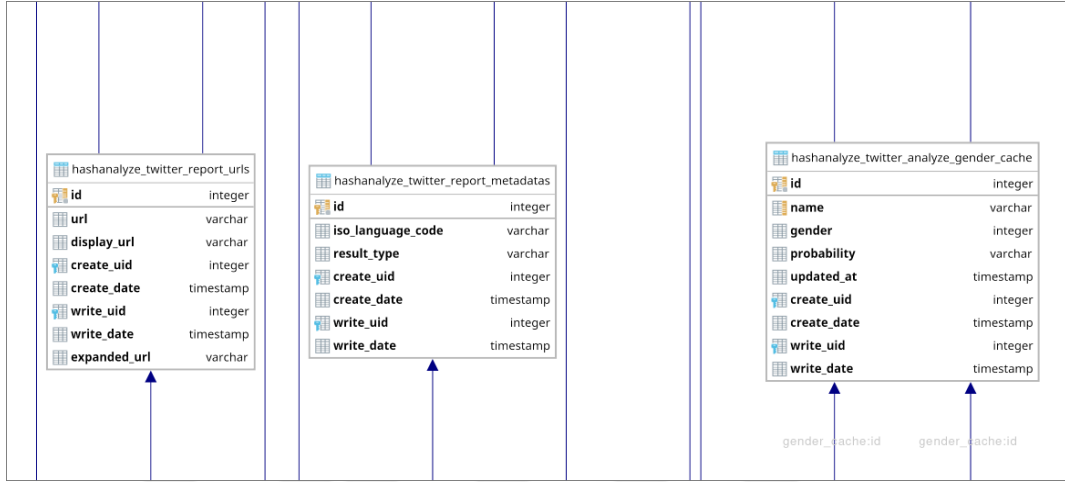
Yazılımın veri tabanında 131 adet tablo kullanılmıştır. Bu tablolar içinde toplamda 163 adet sütun bulunmaktadır. Aşağıda bulunan şekilde 3.8’de çalışma esnasında geliştirilen yazılımda sitem kullanıcılarının yaptıkları tüm işlemlerinin kaydedildiği res\_users tablo yapısı yer almaktadır.



Column Name	Data Type
id	integer
active	boolean
login	varchar
password	varchar
company_id	integer
partner_id	integer
signature	text
action_id	integer
share	boolean
create_uid	integer
create_date	timestamp
write_uid	integer
write_date	timestamp
password_crypt	varchar
sidebar_visible	boolean

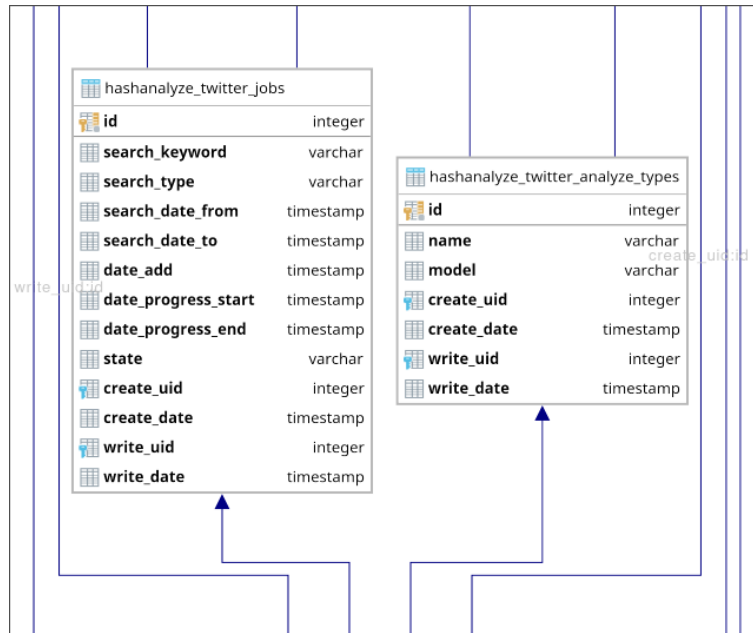
Şekil 3.8. Sistem kullanıcılarının işlemlerinin kaydedildiği res\_users tablosu

Veri tabanı içinde raporların oluşturabilmesi için verilerin tutulduğu yapı örnek olarak şekilde 3.9'da gösterilmiştir. Bağlantıların, tweetlere ait bilgilerin ve kullanıcıların cinsiyet bilgilerinin tutulduğu tablo yapısı örnek olarak sunulmuştur.



Şekil 3.9. Raporların oluşturulabilmesi için verilerin tutulduğu tablo yapısı

Kullanıcıların oluşturduğu analiz kayıtlarının ve bu analizlerin hangi analiz tiplerini içereceğinin bilgilerini tutan tablo yapısı şekilde 3.10'da gösterilmiştir.



Şekil 3.10. Kullanıcıların analiz kayıt ve türlerinin tutulduğu tablo yapıları

### **3.3.4. Twitter'dan çekilen verilerin analizi**

Bu tez konusu sonucu geliştirilen araç ile amaçlandığı üzere etiket (hashtag) analizi yapılmaktadır. Bu analizin gerçekleştirilebilmesi için istenilen etiket kullanılarak gönderilmiş tweetler, Twitter üzerinden arama api 'ı aracılığıyla çekilmektedir.

Geliştirilen yazılımda girilen etiket ile ilgili ve çekilen tweetler üzerinde üç farklı analiz yapılabilmektedir. Etiket ile ilgili girilen etiketin yani kampanyanın nasıl başladığı, ilgili etiketle tweet atan en popüler hesaplar, en çok takipçisi olan hesaplar, kaç farklı kullanıcının tweet attığı, toplam kaç tweet atıldığı, mevcut etiketle ilgili atılan tweetlerin toplamda ne kadar etkileşim aldığı analizleri yapılabilmektedir.

Ayrıca çekilen tweetler üzerinden kullanıcıların bilgileri kaydedildiği için cinsiyet analizi yapılarak katılımında erkek ve kadın oranları belirlenebilmektedir.

Yine çekilen tweetler üzerinden duygu analizi yapılabilmektedir. Tweet metinleri olumlu veya olumsuz (pozitif ve negatif) analizi yapılabilmektedir.

#### **3.3.4.1. Kampanyanın anatomisi**

Kampanya anatomisi Twitter üzerinde bir kampanyanın nasıl yayıldığına anlaşılmasını ve takip edilmesini sağlamak üzere geliştirilmiş bir analiz şeklidir.

İş kaydı olarak girilen bir etiket analizi, isteği kampanya anatomisi analizi aşamasına gelindiğinde, ilgili etiketi içeren tweetler üzerinde ve bu tweetleri gönderen kullanıcılar üzerinde bir analiz gerçekleştirilir. Analiz gerçekleştirilirken ilgili etiketle tweet atan en çok takipçisi olan kullanıcı, toplamda atılan tweet sayısı, tekil olarak kaç kullanıcının ilgili etiketle tweet attığı bilgisi kaydedilmektedir.

Atılan tweetlerin toplam beğeni sayısı, atılan tweetlerin toplam retweet sayısı, bu etiketle ilk hangi kullanıcının tweet attığı bilgisi, en çok beğeni alan ilk on tweet, en çok retweet

alan ilk on tweet, ilgili etiket ile atılan tweetlerde en çok kullanılan diğer etiketler analiz sonuçlarında yer almaktadır.

Ayrıca atılan tweetlerde kullanılan bağlantılar (url) analiz sonuçlarına kaydedilmektedir. Analizi yapılması istenilen etiketle ilgili atılan tweetlere en çok yanıt veren kullanıcılar ve atılan tweetlerde en çok kullanılan kelimeler de analiz sonuçlarında yer almaktadır.

Geliştirilen araçta kampanyanın anatomisi tüm bu verilerin tek sayfada listelendiği analiz sayfasıdır.

#### **3.3.4.2. Kampanyanın toplam veri analizi**

Geliştirilen yazılımda bir kampanya içinde atılan tweetler ile ilgili birçok verinin toplam sayısal değerleri hesaplanarak analiz sonucunda gösterilmektedir.

Kullanıcı analizini yapmak istediği etiketi iş kaydı olarak girdikten sonra yazılım bu iş kaydını işleme alır. İşlem sürecinde öncelikli olarak girilen etiket kullanılarak gönderilen tweetler belirtilen tarih aralığına göre çekilmektedir. Çekilen tweetler veritabanına kaydedilir. Sonra analiz süreci başlatılır.

Analizi yapılan kampanyaya ait veriler toplam olarak gösterilmektedir. Mevcut kampanyaya kaç farklı kullanıcının dâhil olduğunu göstermek amacıyla ilgili etiketle tweet atan tekil kullanıcı sayısı verisi analiz sonucunda yer almaktadır.

Kampanya sürecinde toplam kaç tweet atıldığı analiz sonucunda gösterilmektedir. Atılan tweetlerin toplam kaç beğeni aldığı ve yine atılan tweetlerin toplam kaç retweet aldığı gösterilmektedir.

Kampanya sürecinde ilgili etiketle ilgili kullanıcıların ilgisine dair çıkarım yapabilmek amacıyla atılan tweetlere kaç cevap geldiği ve toplamda atılan tüm tweetlerin aldığı gösterim sayısı toplam veri analizinde yer almaktadır. Tekil olarak atılan tweetlere kaç farklı kişiden cevap geldiği (mention) gösterilmektedir.

Ayrıca analizi istenilen etiketin belirtilen tarih aralıklarında saat kaçta başladığı ve son tweetin saat kaçta atıldığı verileri de analiz sonucunda yer almaktadır.

#### **3.3.4.3. Tweet atan kullanıcıların cinsiyet analizi**

Cinsiyet analizi, çekilen tweetler ve bu tweetlere ait kullanıcı verileri kaydedilerek tüm bu verilerin üzerinde yapılmaktadır.

Cinsiyet analizi ile bir kampanya için tweet atan kullanıcıların kaç tanesinin erkek kaç tanesinin kadın olduğu saptanmaktadır. Ayrıca yüzdelik olarak belirtilmektedir.

Ancak Twitter'da profil ayarlarında cinsiyet detayını belirtmemiş kullanıcılar olabilmektedir.

Geliştirilen yazılımda yapılan cinsiyet analizinde cinsiyetini belirtmemiş kullanıcılar bilinmeyen (unknown) olarak adlandırılmıştır. Cinsiyeti belli olmayan kullanıcı sayısı bu şekilde belirtilmiştir.

Cinsiyet analizi sayesinde bir kampanyanın erkeklerin mi yoksa kadınların mı daha fazla ilgisini çektiği konusunda çıkarımlar yapılabilmektedir.

#### **3.3.4.4. Tweetlerin duygu durum analizi**

Sosyal ağlar üzerinde duygu analizi insanların bir kampanya, konu, marka ya da kurum hakkında olumlu ya da olumsuz konuşma oranlarını tespit etmek amacıyla yapılmaktadır. Bu analiz düzenli periyodik olarak yapılarak marka ya da kurum hakkında insanların görüşlerinde meydana gelen olumlu ya da olumsuz değişimin takibi yapılabilmektedir.



İnsanların düşüncelerinin analizi sosyal ağların temel problemlerinden birisidir. Bu problem sosyal ağlarda varlık gösteren marka ve kurumların da problemi haline gelmiştir. Markalar, kurumlar sosyal ağlarda kendileri hakkında insanların olumlu mu yoksa olumsuz mu düşüncelere sahip olduklarını öğrenmek istemektedirler.

Duygu analizi temelde bir metin işleme işlemidir. Verilen metnin duygusal olarak ifade etmek istediği kategoriyi belirlemeyi amaçlar.

### **3.3.4.5. Tweetlerde kullanılan bağlantıların analizi**

Twitter üzerinde atılan tweetlerde kullanıcılar bağlantı kullanabilmektedir. Yani tweetlerine çeşitli yönlendirmeler ekleyebilmektedirler.

Twitter atılan tweetlere dâhil edilen dışarı yönlenen bağlantıları kendi bağlantı kısaltma servisi ile otomatik olarak kısaltmaktadır. Örneğin: Bir kullanıcı tweete bağlantı eklediğinde bu bağlantı twitter tarafından yirmi üç karakter olarak kısaltılmaktadır. Twitter üzerinde hangi uzunlukta bir bağlantı girilirse girilsin bu bağlantı Twitter bağlantı kısaltma servisi tarafından otomatik olarak yirmiüç karakter olacak şekilde kısaltılmaktadır.

Bağlantılar kısaltılmış olduğu için geliştirilen araçta çekilen tweetler üzerindeki tüm bağlantılar tek alan adı üzerinden görünmektedir. Bu durum da atılan tweetler üzerinde bağlantı analizi yapılmasına engel teşkil etmektedir. Bu sebeple bağlantı analizi yapabilmek amacıyla bir algoritma geliştirilmiş ve tweetlerden gelen tüm bağlantılar arka planda açılarak hangi alan adı üzerinden geldiği tespit edilmiş ve analiz sonuçlarına yansıtılmıştır.

### **3.3.4.6. Tweetlerde diğer etiketlerin ve anahtar kelimelerin analizi**

Bir tweet içinden birden fazla etiket (hashtag) kullanılabilir. Kullanıcılar genellikle gönderdikleri tweet metinlerinden birbiriyle bağlantılı, ilgili etiketleri bir arada kullanmaktadırlar. Bu durum marka ve kurumların kendi etiket kampanyalarını takip ederken aynı zamanda kendi etiketleri dışında kullanıcılar çoğunlukla başka hangi etiketleri kullanmış öğrenmek istemektedirler.

Geliştirilen yazılımda çekilen tweetler üzerinde kullanıcının analizini istediği etiket kullanılarak atılmış tweetlerin içinde geçen diğer etiketlerde kaydedilmektedir. Kaydedilen bu etiketler en fazla kullanım oranına göre listelenmektedir.

Bu sayede kampanya içinde başka hangi etiketler ilgi odağı olmuş ya da organik olarak yayılmış görülmektedir.

Bu tez sonucunda geliştirilen yazılım ile ayrıca tweet metinleri içinde en çok kullanılan kelimelerde kayıt altına alınmaktadır. Bu sayede kampanyanın bir anahtar kelime listesi oluşturulmakta ve analizi yapılabilmektedir.

### **3.4. Analiz verilerinin işlenmesi**

Veriler günümüzde önemli birer metriktir. Bu metrikler de artık alınan kararlarda oldukça etkili bir hale gelmiştir.

Bu sebeple verilerin kullanımları, analiz edilmesi önemli bir durum haline gelmiştir. Verilerin analizi ve işleme işlemi yapılırken kullanılacak verilerin doğru tespit edilmesi, bu verilerin birbirleri ile ilişkilerinin kurulması ve buna uygun olarak bir veritabanı yapısı altında toplanması gerekmektedir. Verilerin bu şekilde düzenli kaydedilmesi sayesinde iyi bir analiz gerçekleştirilebilmektedir.

Geliştirilen yazılımda öncelikli olarak Twitter üzerinden hangi verilerin çekilebileceği araştırması yapılmıştır. Daha sonra bu verilerden hangileri kullanılarak analiz gerçekleştirilebileceğine karar verilmiştir.

Veriler arasında ilişkiler kurulmuş ve geliştirilen yazılımın veritabanı bu ilişkilere göre tasarlanmıştır.

Twitter üzerinden çekilen verilerin analiz edilmesi sürecinde verilerin birbirleri ile ilişkisi üzerinden değerlendirme yapılarak analiz süreci kodlanmıştır.

### **3.4.1. Analiz verilerinin saklanması**

Yazılım içinde çekilen verileri düzenli olarak saklanmaktadır. Kullanıcının analiz talebinde bulunduğu tüm istekler veritabanı üzerine kaydedilmekte ve kullanıcı panelinde gösterilmektedir.

Aynı zamanda girilen isteklerden analiz süreçleri tamamlanan ve sonuçları gösterilen analizlerde veritabanında saklanmaktadır. Hiçbir analiz verisi analiz sonuç raporundan sonra silinmemektedir.

Analiz verilerinin silinmemesindeki amaç daha sonra diğer kullanıcılar tarafından yapılacak benzer analizlerde ya da birebir aynı etiketin analizi istendiğinde daha hızlı cevap verebilmektir.

Ayrıca geliştirilen yazılımla uzun vadede büyük veri üzerinde analiz yapılması mümkündür. Çekilen verilerin, yapılan tüm analiz sonuçlarının kaydedilmesi ile bu tez sonucunda geliştirilen yazılımın kendi içinde bir büyük veri oluşturması mümkündür.

Kaydedilen verilerin sayısal olarak artması ile daha detaylı, sonuçları daha etkili analizler mümkün olacaktır.

## 4. ARAŞTIRMA BULGULARI

Tez çalışmasında geliştirilen yazılım üzerinde örnek analizler yapılmış ve bu analizlerin sonuçları alınmıştır.

Yazılımın amacı doğrultusunda istendiği gibi çalıştığı ve analizleri sunduğu görülmüştür. Bu durum başlıklar halinde verilmiştir.

### 4.1. Etiket analiz (iş) isteğini oluşturma

Geliştirilen yazılımda kullanıcı girişi yapıldıktan sonra analiz isteği girişi yapılması gerekmektedir. Her bir etiketin (kampanyanın) analiz isteği yazılımda iş (jobs) olarak adlandırılmıştır.

Kullanıcı analizlerinin yapılmasını istediği etiketi öncelikle iş olarak oluşturur. Daha sonra kullanıcın oluşturduğu bu iş isteği bekleyen işler sırasına alınır. Şekil 3.11’de “Etiket analiz isteği oluşturma ekranı” gösterilmiştir.

ID	id_str	Kullanıcı	Tweet	Retweet Sayısı	Beğeni Sayısı	Skoru	Duygu Analizi Sonucu	Oluşturulma Tarihi
<a href="#">Bir öge ekle</a>								

Şekil 3.11. Etiket analiz isteği (iş) oluşturma ekranı

Bu ekranda kullanıcı analizini yapmak ve rapor almak istediği etiketin girişini yapmaktadır. Şekil 3.12’de etiket analiz isteği bilgilerinin girildiği ekran gösterilmiştir. Ayrıca girilecek etiket (hashtag) dışında hangi tarihler aralığında analiz yapılmak isteniyorsa o tarih aralıkları da belirtilmektedir.

ID	id_str	Kullanıcı	Tweet	Retweet Sayısı	Beğeni Sayısı	Skoru	Duygu Analizi Sonucu	Oluşturulma Tarihi
<a href="#">Bir öge ekle</a>								

Şekil 3.12. Etiket analiz isteği bilgilerinin girilmesi

Analizinin yapılmasını istenen etiketle ilgili hangi analizlerinin yapılması isteniyorsa türler (types) alanında seçilmelidir.

Bu seçim alanında Kampanyanın Anatomisi (Anatomy), Duygu Analizi (Sentiment) ve Cinsiyet Analizi (Gender) olmak üzere üç çeşit analiz bulunmaktadır. Kullanıcı bunlardan istediği herhangi birini ya da birden fazlasını seçerek analiz işlem kaydını oluşturabilmektedir.

Geliştirilen araçta iş kaydı yani analiz isteğine dair bilgiler girildikten sonra kaydedilerek isteğin iş sırasına alınması sağlanır. İşleme başladığında yazılım seçilen analizleri sırayla gerçekleştirmeye başlamaktadır.

İstenilen tüm analizler tamamlandıktan sonra iş kaydı ekranında “Tamamlananlar” (done) alanında görüntülenebilecektir. Örnek olarak #teknoloji etiketi ile analiz isteği (iş kaydı) oluşturulmuştur. Bu kayıt üzerinden ilgili etiketle tüm analizlerin yapılması için seçimde bulunulmuştur.

İş kaydı oluşturulduktan sonra geliştirilen yazılım isteği iş sırasına almaktadır. Daha sonra bu istek sırası geldiğinde “işleme alındı” şeklinde belirtilmektedir.

#### 4.2. Kampanya anatomisi analizi

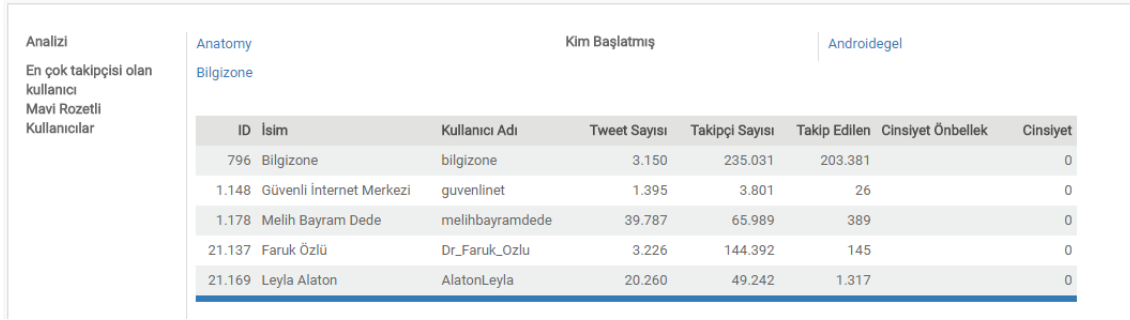
Bu tez çalışmasında geliştirilen yazılımın tasarlandığı gibi çalıştığını göstermek amacıyla “teknoloji” (#teknoloji) etiketi üzerinden analizler gerçekleştirilmiştir.

Teknoloji etiketinin kampanya anatomisi analizi gerçekleştirilmiş ve sonuçlar geliştirilen yazılımda veritabanına kaydedilerek yazılım arayüzünde görüntülenmiştir.

Teknoloji etiketi iş kaydı 09 Nisan 2018 saat 00:00 ile 11 Nisan 2018 saat 16:00 aralığı olarak belirlenmiştir. Analiz tercihleri alanında Anatomi, Duygu ve Cinsiyet analizi seçeneklerinin tamamı seçilmiştir.

İş kaydı işleme alındıktan sonra analizler gerçekleşmiş ve sonuçlar ilgili sayfalarda görüntülenmiştir.

Kampanyanın anatomi analizi sonuçlarında ilk bölüm ilgili etiketle (#teknoloji) belirtilen tarihler arasında ilk tweeti atan hesap tespit edilmiştir. Teknoloji etiketiyle tweet atan en çok takipçi sayısına sahip hesap ve tweet atanlardan doğrulanmış hesap (verified account) yani mavi rozetli olan hesaplar tespit edilmiş ve listelenmiştir. Şekil 3.13’da anatomi analizinin ilk bölümüne yer verilmiştir.



ID	İsim	Kullanıcı Adı	Tweet Sayısı	Takipçi Sayısı	Takip Edilen	Cinsiyet	Önbellek	Cinsiyet
796	Bilgizone	bilgizone	3.150	235.031	203.381			0
1.148	Güvenli İnternet Merkezi	guvenlinet	1.395	3.801	26			0
1.178	Melih Bayram Dede	melihbayramdede	39.787	65.989	389			0
21.137	Faruk Özlü	Dr_Faruk_Ozlu	3.226	144.392	145			0
21.169	Leyla Alaton	AlatonLeyla	20.260	49.242	1.317			0

Şekil 3.13. Anatomi analizi birinci bölüm

Hesaplar kullanıcı adı, görünen adı, tweet sayısı, takipçi sayısı, takip edilen hesap sayısı bilgileriyle gösterilmektedir.

Kampanya anatomisi ikinci bölümünde teknoloji etiketi ile gönderilen ve en çok beğeni alan on tweet listelenmiştir. Şekil 3.14’de en çok beğeni alan tweetlerin listelendiği ekran görülmektedir.

En çok beğeni alan Tweetler			Retweet Sayısı	Beğeni Sayısı	Skoru	Duygu Analizi Sonucu	Oluşturulma Tarihi
ID_id_str	Kullanıcı	Tweet					
32.870	Güvenli İnternet Merkezi	Youtube Yasadışı Olarak Çocukların Verilerini Topluyor - <a href="https://t.co/awY9JlpY8o">https://t.co/awY9JlpY8o</a> - #internet #haber #teknoloji <a href="https://t.co/LJjm5q77H">https://t.co/LJjm5q77H</a>	63	55	0,00	0	10-04-2018 17:00:41
32.874	Güvenli İnternet Merkezi	Uzmanlar uyardı: Çocuklarınızı bu 10 uygulamadan uzak tutun! - <a href="https://t.co/sdzL2xZhz6">https://t.co/sdzL2xZhz6</a> - #internet #haber #teknoloji <a href="https://t.co/TkKF4sYr94">https://t.co/TkKF4sYr94</a>	68	59	0,01	0	10-04-2018 17:45:02
32.887	Güvenli İnternet Merkezi	Sosyal Medya Devine Boykot Düzenlenecek - <a href="https://t.co/PBfpDHMBVx">https://t.co/PBfpDHMBVx</a> - #internet #haber #teknoloji <a href="https://t.co/gr8v3bXISA">https://t.co/gr8v3bXISA</a>	81	83	0,00	0	09-04-2018 17:45:01
32.889	Güvenli İnternet Merkezi	Türkcell ve MEB siber zorbalıktan çocukları korumak için ortak çalışma başlattı - <a href="https://t.co/gFwwnIZcyq">https://t.co/gFwwnIZcyq</a> - <a href="https://t.co/kTjEuickHy">https://t.co/kTjEuickHy</a>	88	86	0,01	0	10-04-2018 08:08:01
32.925	Güvenli İnternet Merkezi	Çocukların çevrimiçi ve dijital ortamlarda güvenliğinin sağlanması amacıyla düzenlenecek olan Uluslararası Çocuk ve... <a href="https://t.co/BrzVUhnssQ">https://t.co/BrzVUhnssQ</a>	110	93	0,36	1	10-04-2018 11:44:19
32.959	Güvenli İnternet Merkezi	iOS'tan Android'e: En çok radyasyon yayan akıllı telefonlar açıklandı - <a href="https://t.co/0lkZEJza0q">https://t.co/0lkZEJza0q</a> - #internet #haber... <a href="https://t.co/GOPN3yB0IE">https://t.co/GOPN3yB0IE</a>	73	74	0,21	1	09-04-2018 11:44:01
32.989	Güvenli İnternet Merkezi	Google Play Store kullanıcılarının istekleri doğrultusunda getirilen yeni özellik ile artık güncellemelerin sunulduğu... <a href="https://t.co/XNScdxmRy9">https://t.co/XNScdxmRy9</a>	53	57	0,12	1	10-04-2018 16:55:51
32.995	Uzay ve Bilim	Şehir ışıklarının uzaydan görüntüsü.. Uluslararası Uzay İstasyonu'ndan (ISS) güzel bir manzara.. #ISS #uzay... <a href="https://t.co/2z2yxUYcoa">https://t.co/2z2yxUYcoa</a>	66	239	0,24	1	10-04-2018 23:33:46

Şekil 3.14. Anatomi analizi en çok beğeni alan tweetler

Kampanya anatomisi üçüncü bölümde teknoloji etiketi kullanılarak atılan tweetlerden en çok retweet atılan on tweet listelenmiştir. İlgili ekran şekil 3.15'de görülmektedir.

En çok Retweet Edilenler			Retweet Sayısı	Beğeni Sayısı	Skoru	Duygu Analizi Sonucu	Oluşturulma Tarihi
ID_id_str	Kullanıcı	Tweet					
32.870	Güvenli İnternet Merkezi	Youtube Yasadışı Olarak Çocukların Verilerini Topluyor - <a href="https://t.co/awY9JlpY8o">https://t.co/awY9JlpY8o</a> - #internet #haber #teknoloji <a href="https://t.co/LJjm5q77H">https://t.co/LJjm5q77H</a>	63	55	0,00	0	10-04-2018 17:00:41
32.874	Güvenli İnternet Merkezi	Uzmanlar uyardı: Çocuklarınızı bu 10 uygulamadan uzak tutun! - <a href="https://t.co/sdzL2xZhz6">https://t.co/sdzL2xZhz6</a> - #internet #haber #teknoloji <a href="https://t.co/TkKF4sYr94">https://t.co/TkKF4sYr94</a>	68	59	0,01	0	10-04-2018 17:45:02
32.887	Güvenli İnternet Merkezi	Sosyal Medya Devine Boykot Düzenlenecek - <a href="https://t.co/PBfpDHMBVx">https://t.co/PBfpDHMBVx</a> - #internet #haber #teknoloji <a href="https://t.co/gr8v3bXISA">https://t.co/gr8v3bXISA</a>	81	83	0,00	0	09-04-2018 17:45:01
32.889	Güvenli İnternet Merkezi	Türkcell ve MEB siber zorbalıktan çocukları korumak için ortak çalışma başlattı - <a href="https://t.co/gFwwnIZcyq">https://t.co/gFwwnIZcyq</a> - <a href="https://t.co/kTjEuickHy">https://t.co/kTjEuickHy</a>	88	86	0,01	0	10-04-2018 08:08:01
32.925	Güvenli İnternet Merkezi	Çocukların çevrimiçi ve dijital ortamlarda güvenliğinin sağlanması amacıyla düzenlenecek olan Uluslararası Çocuk ve... <a href="https://t.co/BrzVUhnssQ">https://t.co/BrzVUhnssQ</a>	110	93	0,36	1	10-04-2018 11:44:19
32.959	Güvenli İnternet Merkezi	iOS'tan Android'e: En çok radyasyon yayan akıllı telefonlar açıklandı - <a href="https://t.co/0lkZEJza0q">https://t.co/0lkZEJza0q</a> - #internet #haber... <a href="https://t.co/GOPN3yB0IE">https://t.co/GOPN3yB0IE</a>	73	74	0,21	1	09-04-2018 11:44:01
32.989	Güvenli İnternet Merkezi	Google Play Store kullanıcılarının istekleri doğrultusunda getirilen yeni özellik ile artık güncellemelerin sunulduğu... <a href="https://t.co/XNScdxmRy9">https://t.co/XNScdxmRy9</a>	53	57	0,12	1	10-04-2018 16:55:51
32.995	Uzay ve Bilim	Şehir ışıklarının uzaydan görüntüsü.. Uluslararası Uzay İstasyonu'ndan (ISS) güzel bir manzara.. #ISS #uzay... <a href="https://t.co/2z2yxUYcoa">https://t.co/2z2yxUYcoa</a>	66	239	0,24	1	10-04-2018 23:33:46

Şekil 3.15. Anatomi analizi en çok retweet alan tweetler

Kampanya anatomi analiz sonuçları üçüncü kısmında teknoloji etiketi ile atılan tweetlerde en çok kullanılan diğer etiket ve anahtar kelimeler tespit edilerek kullanım sayıları ile birlikte listelenmiştir. Şekil 3.16’te bu ekran gösterilmiştir.

Kullanılan Diğer Etiketler		1-40 / 469 < >	
ID	Anatomi	Etiket	Sayısı
4.455	hashanalyze.twitter.analyze.anatomy,19	internet	551
4.456	hashanalyze.twitter.analyze.anatomy,19	haber	509
4.457	hashanalyze.twitter.analyze.anatomy,19	SonDakika	91
4.458	hashanalyze.twitter.analyze.anatomy,19	mobil	94
4.459	hashanalyze.twitter.analyze.anatomy,19	haberler	90
4.460	hashanalyze.twitter.analyze.anatomy,19	Oyun	15
4.461	hashanalyze.twitter.analyze.anatomy,19	Teknoloji	79
4.462	hashanalyze.twitter.analyze.anatomy,19	Webmaster	15
4.463	hashanalyze.twitter.analyze.anatomy,19	bilim	75
4.464	hashanalyze.twitter.analyze.anatomy,19	uzay	59
4.465	hashanalyze.twitter.analyze.anatomy,19	jetpack	10
4.466	hashanalyze.twitter.analyze.anatomy,19	yerlivemilli	20
4.467	hashanalyze.twitter.analyze.anatomy,19	sanayi	21
4.468	hashanalyze.twitter.analyze.anatomy,19	Eğitim	1
4.469	hashanalyze.twitter.analyze.anatomy,19	indir	1
4.470	hashanalyze.twitter.analyze.anatomy,19	software	4
4.471	hashanalyze.twitter.analyze.anatomy,19	program	1
4.472	hashanalyze.twitter.analyze.anatomy,19	technology	25
4.473	hashanalyze.twitter.analyze.anatomy,19	inovasyon	3

Şekil 3.16. Anatomi analizi en çok kullanılan diğer etiket ve ifadeler

Anatomi analizi dördüncü bölümde teknoloji etiketiyle gönderilen tweetlerde kullanılan bağlantılar (url) tespit edilmiş ve listelenmiştir. Şekil 3.17’te bağlantıların listelendiği ekran görülmektedir.

Kullanılan URL'ler		1-40 / 532 < >	
ID	Anatomi	Url	Sayısı
5.975	hashanalyze.twitter.analyze.anatomy,19	https://t.co/SkPwGzBsGI	3
5.976	hashanalyze.twitter.analyze.anatomy,19	https://t.co/I92E083QIV	1
5.977	hashanalyze.twitter.analyze.anatomy,19	https://t.co/mgGxxlgTg2	4
5.978	hashanalyze.twitter.analyze.anatomy,19	https://t.co/zmOw2pzAaj	3
5.979	hashanalyze.twitter.analyze.anatomy,19	https://t.co/Sxzk13e7SZ	4
5.980	hashanalyze.twitter.analyze.anatomy,19	https://t.co/gEEdzCbFW5	3
5.981	hashanalyze.twitter.analyze.anatomy,19	https://t.co/7DA0xQa90v	2
5.982	hashanalyze.twitter.analyze.anatomy,19	https://t.co/QaSK601qxN	1
5.983	hashanalyze.twitter.analyze.anatomy,19	https://t.co/awYJ9lpY8o	43
5.984	hashanalyze.twitter.analyze.anatomy,19	https://t.co/pGjgHPXYs	36
5.985	hashanalyze.twitter.analyze.anatomy,19	https://t.co/zNoE1ijt9c	1
5.986	hashanalyze.twitter.analyze.anatomy,19	https://t.co/sdzL2xZhz6	46
5.987	hashanalyze.twitter.analyze.anatomy,19	https://t.co/vXJNqUuOEI	16
5.988	hashanalyze.twitter.analyze.anatomy,19	https://t.co/97R9KrmSI1	1
5.989	hashanalyze.twitter.analyze.anatomy,19	https://t.co/PBfpDHMBVx	61
5.990	hashanalyze.twitter.analyze.anatomy,19	https://t.co/gFwwnZcyq	63
5.991	hashanalyze.twitter.analyze.anatomy,19	https://t.co/kTJEuickHy	1
5.992	hashanalyze.twitter.analyze.anatomy,19	https://t.co/iaakcr0faq	1
5.993	hashanalyze.twitter.analyze.anatomy,19	https://t.co/A0kgMh23Jt	1

Şekil 3.17. Anatomi analizi tweetlerde kullanılan bağlantılar



Anatomi analizi beşinci kısımda teknoloji etiketiyle atılan tweetlere yanıt veren kullanıcılar (tweet mention) ve yanıt sayıları listelenmiştir. Şekil 3.18’te tweetlere gelen yanıtlar ve sayılarının yer aldığı ekran görülmektedir.

ID	Anatomi	Kullanıcı	Sayısı
1.678	hasanalyzer.twitter.analyze.anatomy,19	YaşlıHaklarıDerneği	1
1.679	hasanalyzer.twitter.analyze.anatomy,19	Okan Dedeoğlu	1
1.680	hasanalyzer.twitter.analyze.anatomy,19	Mustafa Semerci	9
1.681	hasanalyzer.twitter.analyze.anatomy,19	Can Kardeş	8
1.682	hasanalyzer.twitter.analyze.anatomy,19	Oktay Arslan	13
1.683	hasanalyzer.twitter.analyze.anatomy,19	Uzay ve Bilim	77
1.684	hasanalyzer.twitter.analyze.anatomy,19	Sintek Madencilik	3
1.685	hasanalyzer.twitter.analyze.anatomy,19	Faruk Özlü	15
1.686	hasanalyzer.twitter.analyze.anatomy,19	Dijital Kamu	1
1.687	hasanalyzer.twitter.analyze.anatomy,19	Zeynep Dereli	1
1.688	hasanalyzer.twitter.analyze.anatomy,19	Spigen Türkiye	1
1.689	hasanalyzer.twitter.analyze.anatomy,19	Enterprise Next	1
1.690	hasanalyzer.twitter.analyze.anatomy,19	Konya Isabet Okulu	1
1.691	hasanalyzer.twitter.analyze.anatomy,19	TeknoBalta	2
1.692	hasanalyzer.twitter.analyze.anatomy,19	Melih Bayram Dede	1
1.693	hasanalyzer.twitter.analyze.anatomy,19	Bilgizone	148
1.694	hasanalyzer.twitter.analyze.anatomy,19	Gençlik Kervanı	1
1.695	hasanalyzer.twitter.analyze.anatomy,19	Güncel Ekonomik Konular	1
1.696	hasanalyzer.twitter.analyze.anatomy,19	NakitCoins.Com	1

Şekil 3.18. Anatomi analizi tweetlere gelen yanıtlar ve sayıları

Kampanya anatomi analizi son bölümde teknoloji etiketiyle atılan tweetlere ve kullanıcılara dair toplam veriler ve sayılar listelenmiştir. Şekil 3.19’da ilgili ekran görülmektedir.

Tekil Kullanıcı Sayısı	552	Toplam Tweet Sayısı	1.520
Toplam Beğeni Sayısı	2.137	Toplam Retweet Sayısı	70.805
Toplam yanıt sayısı	1.081	Toplam tekil cevap sayısı	67
Toplam etkileşim ve erişim sayısı	74.023	İş	<a href="#">teknoloji</a>
Bitiş Tarihi	11-04-2018 17:59:02	Tür	<a href="#">Anatomy</a>
Başlangıç Tarihi	11-04-2018 17:58:52	Oluşturulma Tarihi	11-04-2018 17:58:52
Type Id	19		

Şekil 3.19. Anatomi analizi toplam sayılar

Anatomi analizi altı bölümde tamamlanmıştır. Teknoloji (#teknoloji) etiketiyle oluşturulan iş kaydı sonucu gerçekleştirilen analizde bu tez sonucu geliştirilen yazılım başarılı şekilde çalışmış ve sonuçlar alınmıştır.

### 4.3. Cinsiyet analizi

Geliştirilen araçta teknoloji etiketi (#teknoloji) ile yapılan iş kaydı ve tüm analiz isteklerinin kaydedilmesiyle yazılım tweetleri çekmiş ve analizleri gerçekleştirmiştir.

Cinsiyet analizi aşamasında yazılım çekilen tweetlerin kullanıcı adları ve görünen adları üzerinden analiz gerçekleştirmektedir. Gerçekleştirdiği analiz sonucu olarak erkek ve kadın kullanıcı oranlarını sayısal ve yüzdesel olarak göstermektedir. Şekil 3.20’de cinsiyet analizi sonuçlarının görüntülediği ekrana yer verilmiştir.

Analiz	Gender	Toplam Erkek Kullanıcı	209
Toplam Kadın Kullanıcı	70	Cinsiyet Belirtmeyenler	272
Erkek Kullanıcıların Yüzdesi	37,93	Kadın Kullanıcıların Yüzdesi	12,70
Cinsiyet Belirtmemiş Kullanıcıların Yüzdesi	49,36	İş	teknoloji
Bitiş Tarihi	11-04-2018 18:00:36	Tür	Gender
Başlangıç Tarihi	11-04-2018 17:59:02	Oluşturulma Tarihi	11-04-2018 17:59:02
Type Id	24		

Şekil 3.20. Cinsiyet analizi sonuç ekranı

Cinsiyet analizi sonuç ekranında Analiz türü görüntülenmektedir. Sonuçlar toplam erkek kullanıcı ve toplam kadın kullanıcı olarak görülmektedir. Ayrıca sonuçlar erkek kullanıcıların yüzdesi ve kadın kullanıcıların yüzdesi olarak da görülebilmektedir.

Sonuç ekranında sonuçların hangi etiket analizine ve hangi tarihler aralığında atılan tweetlere ait olduğu veriside iş kaydı ekranında alınarak tekrar gösterilmiştir.

### 4.4. Duygu analizi

Bu tez sonucu geliştirilen yazılım ile girilen iş isteğinde bulunan etiketle gönderilen tweet metinleri üzerinde duygu (sentiment) analiz gerçekleştirilebilmektedir. Duygu analizi pozitif, negatif ve nötr olarak gruplandırılmıştır. Tweet metinleri içinde kullanılan Türkçe kelime köklerinin bulunması için NLTK (National Language Toolkit) kütüphanesi kullanılmış ve bu kütüphane içinde Turkish Stemmer sınıfı ile kelime kökleri tespit edilmiştir. Kelime kökleri üzerinde polarite (polarity) skorları bulunmaktadır. Bu skorlar

sıfır ile on aralığındadır. Polarite skoru bir kelimenin olumlu mu yoksa olumsuz anlama mı daha yakın olduğunu göstermektedir.

Tespit edilen kelime kökleri veritabanına polarite puanları ile birlikte kaydedilmektedir. Analizi yapılacak tweet metinleri listesi üzerinde bu polarite puanları üzerinden kelimeler analiz edilmekte ve sonuçlar gösterilmektedir. Şekil 3.21’de duygu analizi sonuç ekranında verilerin yer aldığı ekran görülmektedir.

Analiz	Sentiment	Pozitif Polarite Toplamı	395
Negatif Polarite Toplamı	125	Belirsiz / Nötr	1.000
Pozitif Polarite Yüzdesi	25	Negatif Polarite Yüzdesi	8
Nötr Polarite Yüzdesi	65	İş	teknoloji
Bitiş Tarihi	11-04-2018 18:00:45	Tür	Sentiment
Başlangıç Tarihi	11-04-2018 18:00:36	Oluşturulma Tarihi	11-04-2018 18:00:36
Type Id	13		

Şekil 3.21. Duygu analizi sonuç ekranı

Duygu analizi sonuç ekranında analiz türü görülebilmektedir. Analiz sonuçları pozitif polarite toplamı ve negatif polarite toplamı olarak sunulmuştur. Ayrıca pozitif ve negatif polarite oranları da gösterilmiştir.

Duygu analizi sonuç ekranında hangi etiketle gönderilen tweetler üzerinde analizin yapıldığı, hangi tarih aralıklarını kapsadığı ve analizin oluşturulma tarihi bilgileri gösterilmektedir.

## 5. TARTIŞMA VE SONUÇLAR

Bu çalışmada Twitter üzerinde yapılan analiz çalışmalarına farklı ve çok yönlü bir bakış açısı ile yaklaşılmıştır. Çalışma sonucu geliştirilen araç mevcut analiz araçlarından farklı olarak birden fazla farklı analiz türünü aynı panel içinde gerçekleştirebilmektedir. Bir etiketin (hashtag) anatomi analizi, etiketle gönderilen tweet metinlerinin duygu analizi, bu tweet metinlerini yazan kullanıcıların cinsiyet analizi ve kullanıcıların tweet metinleri içinde kullandığı bağlantıların analizi yapılabilmektedir.

Günümüzde internet üzerinde çeşitli Twitter analiz araçları bulunmaktadır. Ancak, bu analiz araçlarının büyük bölümü tek bir analiz türü üzerine odaklanmıştır. Ayrıca mevcut analiz yazılımlarının birçoğunda analizler anlık yapılamamaktadır. Araçların tek bir odağı olduğundan, örneğin; girilen etiket ile ilgili atılan tweet sayısı ve toplam etkileşim sayısı gibi, bu araçlar birçok dijital pazarlama, yazılım geliştirici ve analist profesyonellerin ihtiyacını tam anlamıyla karşılamamaktadır. Ayrıca özellikle Türkiye ve Türkçe özelinde bakıldığında mevcut Twitter analiz araçlarının farklı ülkelerde bulunan firmalar ve geliştiriciler tarafından yayınlanmış olması Türkçe içeriklerin analizinde gerçekçi analizler yapılmasını ve sonuçlar alınmasını imkânsız hale getirmektedir.

Bu tez çalışması sonucu geliştirilen araçta Türkçe dil işleme de kullanılmış ve analizlerin Türkçe içerikler üzerinde doğru sonuçlar vermesi sağlanmıştır. Aracın yönetim panelinde yapılan her bir analiz bir iş kaydı olarak yer aldığından geçmişte yapılan analiz sonuçlarına her zaman ulaşılabilmesi mümkün hale getirilmiştir. Bu sayede eski ve yeni analiz verileri ile kıyaslama yapılarak daha güçlü çıkarımlar yapılabilmesi amaçlanmıştır.

Çalışma esnasında örnek etiket analizi #teknoloji etiketi üzerinden yapılmıştır. Analiz 09.04.2018 ve 10.04.2018 tarihleri aralığında #teknoloji etiketi kullanılarak gönderilen tweetler üzerinde gerçekleştirilmiştir. Toplamda 1.520 tweet üzerinde analiz yapılmış ve bu tweetleri atan kullanıcıların cinsiyet analizi gerçekleştirilmiştir. Atılan tweetlere toplamda 2.137 beğeni, 1.081 yanıt verilmiştir. Ayrıca atılan tweetler 70.805 adet retweet almıştır. Toplamda tweet atan tekil kullanıcı sayısı 552'dir. Tweet atan kullanıcıların %37,93 'ü erkek, %12,70'i kadın olarak tespit edilmiştir. Geriye kalan 49,36 cinsiyet belirtmemiş kullanıcılarıdır. Atılan tweetlerin %25'i pozitif ifadeler %65 'i ise negatif ifadelerdir.

Bu çalışma sonucu geliştirilen Twitter Üzerinde Yapılan Türkçe Paylaşımlar İçin Etiket Analiz Aracı geçmişte geliştirilen benzer araçlara kıyasla içinde daha fazla analiz seçeneği ve fonksiyonu barındırmaktadır.



## KAYNAKLAR

- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011, June). Sentiment analysis of twitter data. In Proceedings of the workshop on languages in social media (pp. 30-38). Association for Computational Linguistics.
- Ayan, B. (2016). Sosyal Ağlar Tarihi. İstanbul: Abaküs Yayınları.
- Başer, M. (2001). Python Programlama Dili. İstanbul: Pusula Yayınları.
- Borgmann, H., Woelm, J. H., Merseburger, A., Nestler, T., Salem, J., Brandt, M. P., ... & Loeb, S. (2016). Qualitative Twitter analysis of participants, tweet strategies, and tweet content at a major urologic conference. *Canadian Urological Association Journal*, 10(1-2), 39.
- Bruns, A., & Stieglitz, S. (2013). Towards more systematic Twitter analysis: metrics for tweeting activities. *International Journal of Social Research Methodology*, 16(2), 91-108.
- Burnap, P., Rana, O. F., Avis, N., Williams, M., Housley, W., Edwards, A., ... & Sloan, L. (2015). Detecting tension in online communities with computational Twitter analysis. *Technological Forecasting and Social Change*, 95, 96-108.
- Hao, M., Rohrdantz, C., Janetzko, H., Dayal, U., Keim, D. A., Haug, L. E., & Hsu, M. C. (2011, October). Visual sentiment analysis on twitter data streams. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on* (pp. 277-278). IEEE.
- Terpstra, T., De Vries, A., Stronkman, R., & Paradies, G. L. (2012). Towards a realtime Twitter analysis during crises for operational crisis management (pp. 1-9). Burnaby: Simon Fraser University.
- Zielinski, A., Bügel, U., Middleton, L., Middleton, S. E., Tokarchuk, L., Watson, K., & Chaves, F. (2012, April). Multilingual analysis of twitter news in support of mass emergency events. In *EGU General Assembly Conference Abstracts* (Vol. 14, p. 8085).
- Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 System Demonstrations* (pp. 115-120). Association for Computational Linguistics.
- Li, R., Lei, K. H., Khadiwala, R., & Chang, K. C. C. (2012, April). Tedas: A twitter-based event detection and analysis system. In *Data engineering (icde), 2012 IEEE 28th international conference on* (pp. 1273-1276). IEEE.
- Goonetilleke, O., Sellis, T., Zhang, X., & Sathe, S. (2014). Twitter analytics: a big data management perspective. *ACM SIGKDD Explorations Newsletter*, 16(1), 11-20.

- Hoeber, O., Hoeber, L., El Meseery, M., Odoh, K., & Gopi, R. (2016). Visual Twitter Analytics (Vista) Temporally changing sentiment and the discovery of emergent themes within sport event tweets. *Online Information Review*, 40(1), 25-41.
- Aichner, T., & Jacob, F. (2015). Measuring the degree of corporate social media use. *International Journal of Market Research*, 57(2), 257-276.
- Seker, S., Duygu Analizi (Sentimental Analysis), YBSAnsiklopedi , v. 3, is. 3, Eylül, 2016 22 , [http://ybsansiklopedi.com/wp-content/uploads/2016/09/duygu\\_analizi.pdf](http://ybsansiklopedi.com/wp-content/uploads/2016/09/duygu_analizi.pdf)
- Poynter, R. (2012: 208). İnternet ve sosyal medya arařtırmaları el kitabı: Pazar arařtırmaları için aralar ve teknikler. (ev: Ü. řensoy). İstanbul: Optimist Yayınları. - Sosyal Medya Nedir
- Masud, M. M., Khan, L., & Al-Shaer, E. (2006, May). Email worm detection using Naïve bayes and support vector machine. In *International Conference on Intelligence and Security Informatics* (pp. 733-734). Springer, Berlin, Heidelberg.
- Twitter, 2018. Eriřim Tarihi; 25.04.2018. <https://developer.twitter.com/en/docs/developer-utilities/twitter-libraries>
- Json, 2018. Eriřim Tarihi; 25.04.2018. <http://json.org>
- Json, 2018. Eriřim Tarihi; 25.04.2018. <https://docs.python.org/2/library/re.html>
- NLTK, 2018. Eriřim Tarihi; 29.10.2015. <http://nltk.org>
- Turkish Stemmer, 2018. Eriřim Tarihi; 29.04.2018. <https://github.com/otuncelli/turkish-stemmer-python>
- Hashtag, 2018. Eriřim Tarihi; 29.04.2018. <http://twitter.pbworks.com>
- Britannica, 2018. Eriřim Tarihi; 26.04.2018. <https://www.britannica.com/topic/Twitter>
- Wikipedia, 2018. Eriřim Tarihi; 26.04.2018. <https://tr.wikipedia.org/wiki/PostgreSQL>
- Postresql, 2014. Eriřim Tarihi; 26.04.2018. <http://postresql.org/docs>
- Github, 2015. Eriřim Tarihi 27.04.2018. <https://github.com/otuncelli/turkish-stemmer-python>

## **EKLER**

### **EK A. Geliştirilen Yazılımın Anatomi Analizi Alanı Kodları**





## EK A. Geliştirilen Yazılımın Anatomi Analizi Alanı Kodları

```
# -*- coding: utf-8 -*-
```

```
from odoo import api, fields, models
```

```
class TwitterAnalyzeAnatomyHashtags(models.Model):
    _name = 'hashanalyze.twitter.analyze.anatomy.hashtags'

    anatomy_id = fields.Many2one('hashanalyze.twitter.analyze.anatomy')
    hashtag_id = fields.Many2one('hashanalyze.twitter.report.hashtags', 'Hashtag')
    count = fields.Integer('Count')
```

```
class TwitterAnalyzeAnatomyMentions(models.Model):
    _name = 'hashanalyze.twitter.analyze.anatomy.mentions'

    anatomy_id = fields.Many2one('hashanalyze.twitter.analyze.anatomy')
    mention_id = fields.Many2one('hashanalyze.twitter.report.users', 'User')
    count = fields.Integer('Count')
```

```
class TwitterAnalyzeAnatomyUrls(models.Model):
    _name = 'hashanalyze.twitter.analyze.anatomy.urls'

    anatomy_id = fields.Many2one('hashanalyze.twitter.analyze.anatomy')
    url_id = fields.Many2one('hashanalyze.twitter.report.urls', 'Url')
    count = fields.Integer('Count')
```

```
class TwitterAnalyzeAnatomy(models.Model):
    _name = 'hashanalyze.twitter.analyze.anatomy'
    _inherits = {'hashanalyze.twitter.analyze': 'analyze_id'}
```

```

analyze_id = fields.Many2one('hashanalyze.twitter.analyze', ondelete="cascade")
who_started = fields.Many2one('hashanalyze.twitter.report.users', 'Who Started')
who_most_follower = fields.Many2one('hashanalyze.twitter.report.users', 'Who have
most follower')
verified_users = fields.Many2many('hashanalyze.twitter.report.users',
'rel_analyze_anatomy_user_verified',
'anatomy_id', 'user_id', 'Verified Users')
most_favorited_tweets = fields.Many2many('hashanalyze.twitter.report.tweets',
'rel_analyze_anatomy_tweet_most_favorite',
'anatomy_id', 'tweet_id', 'Most favorited Tweets')
most_retweeted_tweets = fields.Many2many('hashanalyze.twitter.report.tweets',
'rel_analyze_anatomy_tweet_most_retweet',
'anatomy_id', 'tweet_id', 'Most retweeted Tweets')
used_hashtags = fields.One2many('hashanalyze.twitter.analyze.anatomy.hashtags',
'anatomy_id', 'Used hashtags')
used_urls = fields.One2many('hashanalyze.twitter.analyze.anatomy.urls', 'anatomy_id',
'Used urls')
used_mentions = fields.One2many('hashanalyze.twitter.analyze.anatomy.mentions',
'anatomy_id', 'Used mentions')
unique_user_count = fields.Integer('Unique user count')
total_tweet_count = fields.Integer('Total tweet count')
total_favorite_count = fields.Integer('Total favorite count')
total_retweet_count = fields.Integer('Total retweet count')
total_mention_count = fields.Integer('Total mention count')
total_unique_mention_count = fields.Integer('Total unique mention count')
total_interaction_count = fields.Integer('Total interaction count')

def analyze(self):
    analyze_update_data = {'type_id': self.id, 'total_tweet_count':
len(self.job_id.tweets.ids),
'started_at': fields.Datetime.now()}
    analyze_update_data.update({'who_started': self.job_id.tweets.search([('jobs', '=',
self.job_id.id)], limit=1,

```

```

order="create_date asc").user_id.id))
analyze_update_data.update(
    {'most_favorited_tweets': [(6, 0, self.job_id.tweets.search([('jobs', '=',
self.job_id.id)], limit=10,
order="favorite_count desc").ids]})
analyze_update_data.update(
    {'most_retweeted_tweets': [
        (6, 0, self.job_id.tweets.search([('jobs', '=', self.job_id.id), ('retweeted', '=', False)],
limit=10,
order="retweet_count desc").ids]})

```

```

who_most_follower = {'user_id': 0, 'follower_count': 0}
verified_users = set()
unique_user_count = set()
total_favorite_count = 0
total_retweet_count = 0
total_mention_count = 0
used_hashtags = {}
used_urls = {}
used_mentions = {}
job_search_is_hashtag = self.job_id.search_type == 'hashtag'
job_search_keyword = self.job_id.search_keyword
for tweet_model in self.job_id.tweets:
    total_favorite_count = total_favorite_count + tweet_model.favorite_count
    total_retweet_count = total_retweet_count + tweet_model.retweet_count
    if tweet_model.user_id.followers_count > who_most_follower['follower_count']:
        who_most_follower = {'user_id': tweet_model.user_id.id,
            'follower_count': tweet_model.user_id.followers_count}

    if tweet_model.user_id.verified:
        verified_users.add(tweet_model.user_id.id)
        unique_user_count.add(tweet_model.user_id.id)

    for hashtag in tweet_model.entities_hashtag_ids:

```

```

if job_search_is_hashtag and hashtag.text == job_search_keyword:
    continue
if hashtag.id not in used_hashtags:
    used_hashtags.update({hashtag.id: 0})
used_hashtags.update({hashtag.id: used_hashtags[hashtag.id] + 1})

for url in tweet_model.entities_url_ids:
    if url.id not in used_urls:
        used_urls.update({url.id: 0})
    used_urls.update({url.id: used_urls[url.id] + 1})

total_mention_count = total_mention_count +
len(tweet_model.entities_mention_ids)
for mention in tweet_model.entities_mention_ids:
    if mention.id not in used_mentions:
        used_mentions.update({mention.id: 0})
    used_mentions.update({mention.id: used_mentions[mention.id] + 1})

# used_hashtags
used_hashtag_ids = []
for used_hashtag_id, used_hashtag_count in used_hashtags.items():
    used_hashtag_exist = self.used_hashtags.search(
        [('anatomy_id', '=', self.id), ('hashtag_id', '=', used_hashtag_id)])

    if used_hashtag_exist:
        used_hashtag_exist.write({'count': used_hashtag_count})
    else:
        used_hashtag_exist = self.used_hashtags.create(
            {'anatomy_id': self.id, 'hashtag_id': used_hashtag_id, 'count':
used_hashtag_count})
        used_hashtag_ids.append(used_hashtag_exist.id)
if used_hashtag_ids:
    self.used_hashtags.search([('anatomy_id', '=', self.id), ('id', 'not in',
used_hashtag_ids)]).unlink()

```

```

# used_urls
used_url_ids = []
for used_url_id, used_url_count in used_urls.items():
    used_url_exist = self.used_urls.search(
        [('anatomy_id', '=', self.id), ('url_id', '=', used_url_id)])

    if used_url_exist:
        used_url_exist.write({'count': used_url_count})
    else:
        used_url_exist = self.used_urls.create(
            {'anatomy_id': self.id, 'url_id': used_url_id, 'count': used_url_count})
        used_url_ids.append(used_url_exist.id)
if used_url_ids:
    self.used_urls.search([('anatomy_id', '=', self.id), ('id', 'not in',
used_url_ids)]).unlink()

# used_mentions
used_mention_ids = []
for used_mention_id, used_mention_count in used_mentions.items():
    used_mention_exist = self.used_mentions.search(
        [('anatomy_id', '=', self.id), ('mention_id', '=', used_mention_id)])

    if used_mention_exist:
        used_mention_exist.write({'count': used_mention_count})
    else:
        used_mention_exist = self.used_mentions.create(
            {'anatomy_id': self.id, 'mention_id': used_mention_id, 'count':
used_mention_count})
        used_mention_ids.append(used_mention_exist.id)
if used_mention_ids:
    self.used_mentions.search([('anatomy_id', '=', self.id), ('id', 'not in',
used_mention_ids)]).unlink()

```

```

analyze_update_data.update({
    'type': self.env.ref('twitter.hashanalyze_analyze_anatomy_type').id,
    'type_id': self.id,
    'who_most_follower': who_most_follower['user_id'],
    'verified_users': [(6, 0, list(verified_users))],
    'unique_user_count': len(unique_user_count),
    'total_favorite_count': total_favorite_count,
    'total_retweet_count': total_retweet_count,
    'total_mention_count': total_mention_count,
    'total_unique_mention_count': len(used_mention_ids),
    'total_interaction_count': total_favorite_count + total_retweet_count +
total_mention_count,
    'ended_at': fields.Datetime.now(),
})

self.write(analyze_update_data)

```

## ÖZGEÇMİŞ

Adı Soyadı : Yunus Emre KARABULUT

Doğum Yeri ve Yılı : Sivas, 1991

Medeni Hali : Bekar

Yabancı Dili : İngilizce

E-posta : emrkarabulut@gmail.com

### Eğitim Durumu

Lise : Dikmen Anadolu ve Endüstri Meslek Lisesi

Üniversite: Süleyman Demirel Üniversitesi, Teknik Eğitim Fakültesi,  
Sistemleri Öğretmenliği, 2014

Bilgisayar

### Mesleki Deneyim

Süleyman Demirel Üniversitesi Senirkent Meslek Yüksek Okulu

2015-2015

### Yayımları