

T.C.
MİLLÎ SAVUNMA ÜNİVERSİTESİ
ATASAREN STRATEJİK ARAŞTIRMALAR VE LİSANSÜSTÜ
EĞİTİM ENSTİTÜSÜ
İSTİHBARAT ÇALIŞMALARI ANABİLİM DALI

SOSYAL MEDYA İSTİHBARATI İLE TWITTER
GÜNDEMLERİNİN PANAS-T ÖLÇEĞİNE GÖRE
İNCELENMESİ
YÜKSEK LİSANS TEZİ

TAN HASKOL
1920106106

TEZ DANIŞMANI: DOÇ. DR. BURAK MİL

İSTANBUL
EKİM 2022

T.C.
MİLLÎ SAVUNMA ÜNİVERSİTESİ
ATASAREN STRATEJİK ARAŞTIRMALAR VE LİSANSÜSTÜ
EĞİTİM ENSTİTÜSÜ
İSTİHBARAT ÇALIŞMALARI ANABİLİM DALI

SOSYAL MEDYA İSTİHBARATI İLE TWITTER
GÜNDEMLERİNİN PANAS-T ÖLÇEĞİNE GÖRE
İNCELENMESİ
YÜKSEK LİSANS TEZİ

TAN HASKOL
1920106106

TEZ DANIŞMANI: DOÇ. DR. BURAK MİL

İSTANBUL
EKİM 2022

T.C.
MİLLÎ SAVUNMA ÜNİVERSİTESİ
ATASAREN STRATEJİK ARAŞTIRMALAR VE LİSANSÜSTÜ
EĞİTİM ENSTİTÜSÜ
İSTİHBARAT ÇALIŞMALARI ANABİLİM DALI

SOSYAL MEDYA İSTİHBARATI İLE TWITTER
GÜNDEMLERİNİN PANAS-T ÖLÇEĞİNE GÖRE
İNCELENMESİ
YÜKSEK LİSANS TEZİ

TAN HASKOL
1920106106

Tezin Enstitüye Verildiği Tarih : 27.10.2022

Tezin Savunulduğu Tarih : 28.09.2022

Tez Oy birliği ile başarılı bulunmuştur.

	Unvan Ad Soyad	İmza
Tez Danışmanı	Doç.Dr. Burak MİL	
Jüri Üyeleri	Dr.Öğr.Üyesi Murat SAĞBAŞ	
	Dr.Öğr.Üyesi Caner YAZICIOĞLU	

İSTANBUL
EKİM 2022

ÖZGÜNLÜK RAPORU

Tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve ç) Sonuç kısımlarından oluşan toplam 85 sayfalık kısmına ilişkin, 27 Ekim 2022 tarihinde şahsım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan özgünlük raporuna göre, tezimin benzerlik oranı % 7'dir.

Uygulanan filtrelemeler:

- 1- Kaynakça hariç
- 2- Alıntılar dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Millî Savunma Üniversitesi Atasaren Enstitüsü Lisansüstü Tez Çalışması Özgünlük Raporu Alınması ve Kullanılması Uygulama Usul ve Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Tan HASKOL

27.10.2022

ETİK BEYAN

Millî Savunma Üniversitesi Enstitüleri Dönem Projesi ve Lisansüstü Tez Hazırlama Kılavuzu'nda yer alan kurallarına uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir; aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Bu tezdeki düşünce, görüş, varsayım, sav veya tezler bana aittir; Millî Savunma Bakanlığı, Millî Savunma Üniversitesi ve Atasaren Enstitüsü sorumlu tutulamaz.

Tan HASKOL

27.10.2022

ÖNSÖZ ve TEŞEKKÜR

15.yy'dan günümüze dikey olarak ayrılan bilim dalları, sosyal bilimler açısından niş bir çalışma alanı olan hesaplamalı sosyal bilimler disiplini yeniden birleşme eğilimi ile karşımıza çıkıyor. Bilgisayar bilimleri, davranışsal bilimler, matematik ve istatistiğin bu nadir kompozisyonu araştırmacılar için bilginin ölçülebilirlik ve ulaşılabilirlik sınırlarını esnetiyor.

Mevcut bilimsel trendin ve bunun sunduğu olanakların paralelinde ve istihbarat disiplini merkeze alacak şekilde bir psikometri uygulamasını bilgisayar destekli olacak incelemeye çalıştım. Aynı alandaki işim gereği sektör önceliklerini temel alan ve kendim başta olmak üzere uygulamacılara, geliştirecekleri ürünlerde ya da bunların erken ar-ge aşamalarında somut olarak yol gösterecek bir çalışma oluşturmaya gayret ettim.

Çalışmada büyük emeği geçen DataKapital ekibine teşekkürü bir borç bilirim. Bu alandaki birkaç sayfalık bir makalenin bile farklı disiplinlerden birkaç araştırmacı tarafından yazıldığı düşünüldüğünde, uzman mühendislik desteği olmasaydı bu çalışmanın ortaya çıkması oldukça güçtü.

Disiplini tanımamı ve derinleştirmemi sağlayan Milli Savunma Üniversitesi'nde ders aldığım pek çok hocamın da vizyonuma katkısı bu çalışmayı mümkün kılmıştır. Bunlar dışında veri çalışmalarında basit gibi görünen ancak işin doğası gereği ağır ve bunaltıcı işlerde bize destek olan DK Network üyesi gençlere de teşekkür ederim.

İstanbul; Ekim 2022

Tan HASKOL

İÇİNDEKİLER

ÖZGÜNLÜK RAPORU	Sayfa
ETİK BEYANI	
İTHAF	
TEŞEKKÜR VE ÖNSÖZ	
İÇİNDEKİLER	vi
TABLolar LİSTESİ	viii
ŞEKİLLER LİSTESİ	ix
KISALTMALAR	x
TÜRKÇE ÖZ	xi
İNGİLİZCE ÖZ (ABSTRACT)	xii
1. GİRİŞ	13
1.1. İstihbarat Metodolojisi ve Temel Konseptler.....	13
1.1.2. Rekabetçi Açık Kaynak İstihbaratı	16
1.2. Tezin Amacı, Önemi, Hipotezler ve Yöntem	18
1.2.1. Tezin Sorunsalı	19
1.2.2. Yöntem	19
1.2.3. Hipotez	20
1.2.4. Literatür ve Ölçek Tercihi.....	21
1.2.5. Uygulama	23
1.3 Tezin Sınırlılıkları	24
1.3.1 İlgili Mevzuatlar ve Hukuki Sınırlılıklar	25
1.3.1.1 Veri Kazıma ve Tarama Yön.İlişkin Hukuki Sınırlılıklar	25
1.3.1.2 Verilerin İşlenmesine İlişkin Hususlar	29
1.3.1.3 Verilerin Mülkiyetine İlişkin Esaslar	37
2. PANAS	29
2.1. Tanımlayıcı Kelimeler, Geçerlilik ve Güvenilirlik	31
2.2. Türkçe Uyarlama	32
3. Veri Seti	33
3.1. Büyük Veri Seti İçin Uygulama ve Rassallaştırma	34
3.2. Betimleyici Sıfatların Saydırılması ve Stemming (Kök Bulma)	35
3.2.1. Türkçeleştirme ve Betimleyici Anahtar Kelimelerin Seçimi...	37
3.2.2. Türkçeleştirmede Kapsama Yaklaşımı	39
3.3. Betimleyici Anahtar Kelimelerin Tanımlanması	39
3.3.1 Betimleyici Sıfatların Büyük Veri Kümesinde Analizi	45
4.0 Analiz Konusu Gündemler	48
4.1. Gündemlerin Seçimi ve Tercih Nedenleri	48
4.2. Gündemlerdeki Hakim Duygu Durumlarının Tespiti ve Hesaplama ...	51
4.3. İlk Bulgular Üzerinden Tespit Edilen Sorunlar	60
4.4. Betimleyici Sıfatların İzole Hesaplanması	61
4.4.1. Gündemlerde Skorların İzole Hesaplanması	62
4.5 Yöntemlerin Mukayesesi ve Bulgular	64
4.5.1 Duygu Alt Setleri ile İlgili Bulgular	65
5.0 Anket Verileri	68
6.0 Korelasyon Analizleri	70
6.1 Sentiment (Duygu) Analizi Katmanı ve İyileştirme.....	73
6.2 Sentiment Modülü	74
6.3 Filtrelenmiş İçeriklerin PANAS Skorları ve Korelasyonlar.....	77

7.0 Kullanılan Teknoloji	79
7.1 Verilerin Sayımı ve Analizi İşlemleri	81
7.2 Teknik Akreditasyon	84
8.0 Sonuç	85
KAYNAKÇA	86
EKLER	89
Ek-1 : Büyük Veri Kümesi Randombase.JSON	
Ek-2 : Gündemlere İlişkin Veri Kümesi Events.JSON	
Ek-3 :Türkçe Etkisiz Kelimeler (Stop Words).....	89
Ek-4 : PANAS Betimleyici Anahtar Kelimeler	90
Ek-5 : Gündemlerin İzole Hesaplanmış PANAS Değerleri	91
Ek-6 : Anket Soruları Örneği	95
Ek-7 : Gündemler İçin Anketlerin Oylama Sonuçları.....	96
Ek-8 : TÜR Belgesi.....	98
Ek-9 : Duygu Durumu Frekansları.....	99
ÖZGEÇMİŞ	103

TABLÖLER LİSTESİ

	Sayfa
Tablo 1: PANAS kelimeleri faktör yükleri (İngilizce)	33
Tablo 2: Türkçe Betimleyici Sıfatların Faktör Yükleri	34
Tablo 3: Büyük Veri Kümesi İçindeki Betimleyici Sıfatlar Contains()	44
Tablo 4: Büyük Veri Kümesindeki Pozitif Sıfatlı İçerikler	48
Tablo 5: Büyük Veri Kümesindeki Negatif Sıfatlı İçerikler	49
Tablo 6: Gündemlerdeki Toplam Tweet Sayıları	52
Tablo 7: Suriyeliler Gündemindeki Betimleyici Sıfatlar	53
Tablo 8: Gündemlerdeki İçeriklerin Duygu Boyutuna Göre Tasnifi	57
Tablo 9: Toplama Yöntemine Göre PANAS Skorları	58
Tablo 10: İzole Yöntemle Hesaplanan α Değerleri	64
Tablo 11: Suriyeliler Gündeminde İzole Hesaplanan PANAS Skorları	65
Tablo 12: Tüm Gündemlerin İzole Yöntemle PANAS Skorları	66
Tablo 13: İzole Yönteme Göre En Yüksek Skorlu Duygu Altsetleri	68
Tablo 14: Anketlerin PANAS Skorları	71
Tablo 15: Toplama Yöntemine Göre Pozitif Duygu Durumu Korelasyonları	73
Tablo 16: Toplama Yöntemine Göre Negatif Duygu Durumu Korelasyonları	73
Tablo 17: İzole Yönteme Göre Pozitif Duygu Durumu Korelasyonları	74
Tablo 18: İzole Yönteme Göre Negatif Duygu Durumu Korelasyonları	74
Tablo 19: Sentiment Modülü ile Filtrelenmiş Toplam İçerikler	79
Tablo 20: Filtrelenmiş İçeriklerle Elde Edilen PANAS Skorları	80
Tablo 21: Filtrelenmiş İçeriklerin Pozitif PANAS Değerlerinin Korelasyonları	81
Tablo 22: Filtrelenmiş İçeriklerin Negatif PANAS Değerlerinin Korelasyonları	81

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1:	Cronbach'ın psikolojik test için üç temel elementi 18
Şekil 2:	Rekabetçi Açık Kaynak İstihbaratı..... 20
Şekil 3:	Veri Kazımada Hesaplama Maliyetleri 29
Şekil 4:	Betimleyici Sıfatların Anahtar Kelimelere Dönüştürülmesi 43
Şekil 5:	Gündemlerin P(pozitif) Skorları..... 61
Şekil 6:	Gündemlerin P(negatif) Skorları 61
Şekil 7:	Gündemlerin Sübjektivite Analizi..... 62
Şekil 8:	Sentiment Analiz Yaklaşımlarının Tasnifi 75
Şekil 9:	Sentiment Filtresi..... 78
Şekil 10:	Apache Commons-lang StringUtils.countMatches() metodu..... 84
Şekil 11:	Tweet-Service Sayım Algoritması 85
Şekil 12:	Betimleyici Sıfatların JSON Çıktıları..... 86

KISALTMALAR

API	: Uygulama Programlama Arayüzü
JSON	: Javascript Object Notation – Javascript Nesnesi Gösterimi
GDPR	: Genel Veri Koruma Yönetmeliği
KVKK	: Kişisel Verilerin Korunması Kanunu
3V	: Volume, Velocity, Variety – Hacim, Hız, Çeşitlilik
PANAS	: Pozitif ve Negatif Duygu Durum Ölçeği
PANAS – T	: Twitter için PANAS
PANAS – X	: PANAS’ın genişletilmiş Versiyonu
R.A.K.İ	: Rekabetçi Açık Kaynak İstihbaratı
TÜR	: Teknolojik Ürün Geliştirme Belgesi



ÖZ

Sosyal Medya İstihbaratı ile Twitter Gündemlerinin PANAS-T Ölçeğine Göre İncelenmesi

Tan HASKOL

Millî Savunma Üniversitesi, Atatürk Stratejik Araştırmalar ve Lisansüstü Eğitim
Enstitüsü

İstanbul, Ekim 2022

Açık kaynak istihbaratı ve destekleyici veri analitikleri çalışmalarının başarısı, bilgi hiyerarşinin en alt katmanı olan ham verilerin ne kadar ölçeklendirilebildiği ile doğrudan ilgilidir. Açık sosyal ağlar ise ham verilerin miktarını arttırmak ve ölçeklendirmek için en iyi maliyet etkin çözümleri sunan platformlardır.

Bu çalışmada psikometri disiplinde kullanılan PANAS (Pozitif ve negatif duygu durum ölçeği) temel alınarak, sosyal medya verileri üzerinden bir gündeme ilişkin, kitlelerin duygu durum analitiklerini elde edilmeye çalışılacaktır. PANAS temelde duygu durumu ifade eden anahtar kelimeler üzerinden yapılan testlerle çıktı üreten bir ölçektir. Analize giren temel bileşenler anahtar kelimeler olduğu için bu ölçek sosyal medya istihbaratına kolaylıkla uygulanabilmektedir.

PANAS ölçeği özelinde, bilgisayar destekli uygulamalarla konvansiyonel araştırma yöntemlerinden daha büyük veri kümelerine ulaşarak benzer sonuçlar elde edilip edilemeyeceği ilişkisellik üzerinden test edilecektir.

Anahtar Sözcükler: Sosyal Medya İstihbaratı, Açık İstihbarat, Büyük Veri, Psikometri, PANAS, Doğal Dil İşleme, Sentiment Analizi

Bilim Kodu : 92414

Sayfa Sayısı : 103

Tez Danışmanı : Doç. Dr. Burak MİL

ABSTRACT

Examining Twitter Agendas by Social Media Intelligence According to PANAS-

T Scale

Tan HASKOL

Turkish National Defense University, Atatürk Strategic Studies and Graduate
Institute

İstanbul, October 2022

The success of open source intelligence and supporting data analytics studies are directly related to how scalable raw data, which is the lowest layer of the information hierarchy. Open social networks (OSN) are the platforms that offer the best cost-effective solutions for increasing the amount and scaling of raw data.

In this study, based on the PANAS (Positive and Negative Affect Schedule) used in the discipline of psychometrics, it will be tried to obtain mood analytics of the masses on an agenda through social media data. PANAS is basically a scale that produces output with tests made on keywords expressing mood. This scale can be easily applied to social media intelligence, as the key component in the analysis is keywords.

Specific to the PANAS scale, it will be tested whether similar results can be obtained by reaching larger datasets than conventional research methods with computer applications.

Keywords: Social Media Intelligence, Open Source Intelligence, Big Data, Psychometrics, PANAS, Natural Language Processing, Sentiment Analysis

Science Code : 92414

Pages : 103

Supervisor : Assoc. Prof. Burak MİL, Ph.D.

1. GİRİŞ

Statista'nın 2022 Ocak verilerine göre ¹Türkiye'de yaklaşık 16.7 milyon aktif Twitter kullanıcısı bulunmaktadır ve bu kullanıcılar günde 5 milyonun üzerinde Tweet atmaktadır. 2020 yılında ise bütün dünyada 353 milyon Twitter kullanıcısı olduğu tahminlenmiştir.² Sürekli büyüyen bu muazzam veri kaynağı doğru mühendislik uygulamalarla günümüzde ulaşılabilir ve işlenebilir durumdadır. Bahsedilen veri kaynağından yararlanarak faydalı ve cari istihbarat üretilebilmesi ise, ilgili bağlamda insan davranışının ne kadar sayısallaştırılabildiği ilgilidir.

Bireysel ya da kurumsal klasik araştırmacıların, insan gücüyle iş yapma kapasitesini aşan açık istihbarat faaliyetlerinin tümü bir büyük veri problemidir. Burada tanımladığımız teknik sorun aynı zamanda çalışmanın metodolojisinin ve teknik sınırlılıklarının belirlenmesindeki ana faktör olmuştur ve açık kaynak istihbaratının hangi alanda ilerletilmesi gerektiği ile ilgili fikirleri de içermektedir.

Teknik sınırları ve metodolojiyi belirleyen beş temel teknik konsept/tema vardır. Bunlar farklı araştırmacılar tarafından, farklı anlamlarda ve kontekslerde kullanılmıştır bu yüzden kavramları kısaca özetleyeceğiz ve çalışmada hangi çerçevede ele alındığını belirteceğiz.

Bahsedilen teknik konseptler;

a-Büyük veri

b-Açık istihbarat

c-Sosyal medya istihbaratı

d-Doğal dil işleme

e-Psikometridir.

1.1. İstihbarat Metodolojisi ve Temel Konseptler

¹ <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>

² Erol Başaran BURAL, “Terörizmle mücadelede sosyal medya istihbaratının önemi: pkk terör örgütünün Twitter'daki görünürlüğü üzerine sosyal ağ analizi” (Jandarma ve Sahil Güvenlik Akademisi Güvenlik Bilimleri Enstitüsü), Ankara 2021, 48.

Büyük veri : Büyük veri, klasik yazılım ve işleme yöntemleriyle yönetilemeyecek büyüklükte, sıklıkta ve çeşitlilikte veri problemleri için kullanılan bir tanım ve temadır. Büyük veri temel düzeyde klasik veri problemlerinden üç temel belirleyici sıfatla ayrılmıştır. Bunlar hacim, hız ve çeşitlilik (3V). Hacim klasik veri işleme yazılımlarıyla işlenemeyecek ölçekte büyük veri miktarlarını, hız ise bu miktarların sürekli ve daha önce görülmemiş hızlarda akışını, çeşitlilik ise akışların standart olmayan ve pek çok farklı formattaki veri kaynağından gelmesini ifade eder.

Klasik 3V(hacim, hız, çeşitlilik) tanımına ek olarak (bu alandaki tartışmalar ihmal edilecektir) konseptin daha iyi soyutlanması gerekmektedir. Büyük veri probleminin açık istihbarat bağlamındaki en iyi tanımı, en ayırıcı özelliği bilgi kaynaklarına, depolamaya ve paylaşmaya yönelik teknolojiler gelişmeye devam ederken, bunları işleyecek işlemci teknolojilerinin inovasyona doymuş olmasıdır. Yani mevcut teknolojik kapasite ile büyük veri sorunu/fırsatı kalıcıdır. Bu alanda teknolojik ya da bilimsel bir devrim olmadan mevcut durum değişmez, sorun dağıtık işleme v.b gibi veri işleme tarafındaki inovasyonlarla çözülemez.

Teknik sabit şu şekildedir;

$$\text{Veri (işleme kapasitesi)} \leq \text{Veri (Artış hızı, depolama kapasitesi)}$$

Açık İstihbarat : Kamuya açık ve erişilebilir kaynaklardan istihbarat üretmek amacıyla bilgi toplama faaliyetidir. Açık istihbarat faaliyeti büyük ölçüde spekülative³ ancak güçlü ve tahmine dayalı istihbarat üretilmesini sağlar⁴ ve rekabetçi özellikleri bu bağlamdan gelir. Yöntemler genellikle sayısal, ölçülebilir ve çıktılar bu özelliklere göre metriklenmiş olsa da istihbaratın esas faaliyeti tahmine, t+1 zamandaki duruma yoğunlaşır. Açık istihbaratta gerçek katma değer, rekabetçi özellik tahmin safhasında

³ Spekülative istihbaratla kastedilen Sherman Kent'in Stratejik İstihbarat isimli eserinde kavramlara açıklık getirmek için kullandığı bağlamdır (sf:15 dipnot:2). Aslında temel konseptler ve temalar da aynı amaçla tanımlandı çünkü hem ekol farklılıklarına göre, hem de kullanılan bağlama göre kavramlar farklı anlamlar ihtiva edebilir ve karışıklığa neden olabilir. Spekülative istihbarat kesin bulgular ve bunlara bağlı mekanik analiz dışında kalan her şeydir. Analiz kısmı da kesin bulguların derlenmesinden ibarettir, yorumlanmasından değil. Bulgular dışında insan yorumu devreye girdiği anda sınır geçilmiş olur, bu sınırdan sonrası artık spekülative istihbarattır.

⁴ Uğur Tekin, Açık Kaynaklardan Siber Tehdit İstihbaratı Verisi Elde Edilmesi, (Gazi Üniversitesi Fen Bilimleri Enstitüsü Bilgi Güvenliği Mühendisliği Ana Bilim Dalı) Yüksek Lisans Tezi, Kasım 2021, 12.

ve veri toplama/işleme uygulamasındadır ⁵. Espiyonaj faaliyeti neticesinde elde edilen veri(data) veya bilgi(enformasyon) başlı başına rekabetçi bir katma değer içerir. Diğer taraftan açık istihbarat yoluyla elde edilen bilgilerin konvansiyonel araştırma yöntemlerinin üzerine çıkması ve rekabetçi hale gelmesi için iki unsur belirleyicidir: Toplanacak verilerin sayısının arttırılması ve veri analitiklerinin geleceğe ilişkin tahminlerde, modellerde kullanılması. Verilerin konvansiyonel yöntemleri aşacak şekilde arttırılması ise ancak uygulamalarla ve destekleyici otomasyonlarla mümkündür.

Sosyal Medya İstihbaratı : Açık sosyal ağlardan verilerin toplanarak ve işlenerek istihbarat elde edilmesi faaliyetidir. Genellikle büyük veri temasıyla iç içe bir faaliyettir ancak her sosyal medya istihbaratı uygulaması büyük veri niteliği taşımaz veya büyük verinin belirleyici özelliği olan 3V⁶'yi farklı kompozisyonlarla içinde bulundurur. Bunun yanı sıra istihbarat bağlamda zaman, maliyet ve kazanç açısından rekabetçi bir sosyal medya istihbaratı enstrümanı, hemen her zaman bilgisayar destekli uygulamadır ve bir büyük veri problemidir.

Doğal Dil İşleme : Belli bir dilde yazılmış metin halindeki veya metin haline getirilebilir veriler (örn : sesler) üzerine yapılan, bilgisayar destekli ve çok disiplinli çalışmalardır. İstihbarat bağlamda sosyal medya verilerinin, Türkçe dilindeki morfolojik ve yapısal analizler ve uyarlamalarda dahil olmak üzere ⁷ toplanması, işlenmesi süreçlerindeki uygulama çalışmalarının tamamına referans verir.

Psikometri : Psikometri, insan davranışları ve duygu durumlarına ilişkin istatistik yöntemleri psikoloji disiplinine uygulayarak ölçen ve davranışların yorumlanmasında kullanılacak matematiksel modellerin geliştirilmesini sağlayan bir bilim dalıdır. Psikometri çalışma alanı iki temel üzerine inşa edilmiştir. Bu temellerden ilki psikolojik test ve ölçüm pratikleridir. Diğerleri ise ölçümlere ilişkin istatistik konsept ve

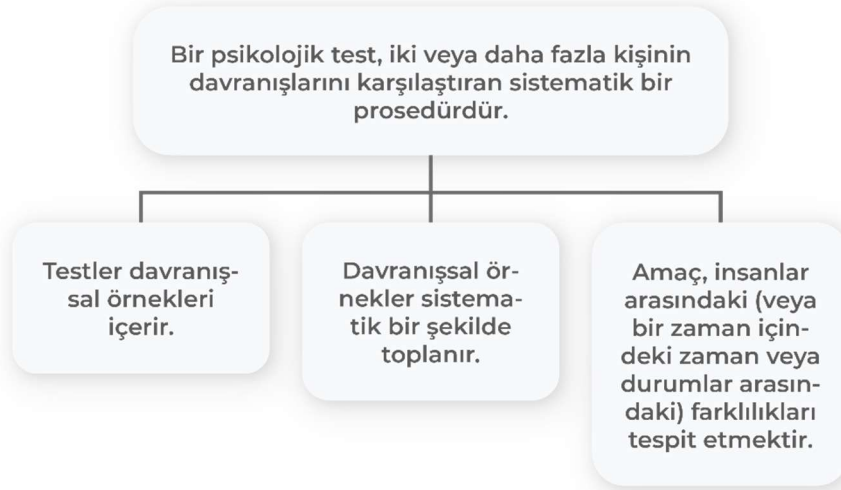
⁵ Uygulama ile kast edilen amaca uygun yazılımlar, uygulayıcı ise bu sürecin tamamına dahil olan ürün sahipleri, ürün geliştiriciler, ar-ge personeli, yazılım geliştiriciler vb.leridir. Uygulayıcı salt yazılım geliştiricilere referans vermez.

⁶ Doug Laney, Meta Group/Gartner “high volume, high velocity and high variety” 2001

⁷ Deniz GÜL ÖZCAN, “Doğal dil işleme tekniklerinin incelenmesi ve seyahat-turizm sesli asistanı için türkçe varlık ismi tanıma aracı geliştirilmesi” (Akdeniz Üniversitesi) Aralık 2020 , Yüksek Lisans tezi, i.

prosedürlerin geliştirilmesidir.⁸ Psikometri temelde insan davranışına ilişkin önceden tanımlanmış faktörlerin niceliksel olarak ölçülmesini sağlar.

Psikometri uygulamalarının/modüllerinin de dahil olduğu tüm psikolojik testler iki veya daha fazla insanın kıyaslandığı sistematik bir prosedürdür.⁹ Bu önermeye göre psikolojik testlerin hali hazırda sistematik olduğu varsayılmaktadır ve kastedilen sistematiklik (mekaniklerin ve niceliklerin önceden belirli olması) başta psikometri olmak üzere psikolojik testler üzerine veri bazlı uygulamalar geliştirilmesini olanaklı kılar.



Görsel - 1 Cronbach'ın psikolojik test için üç temel elementi

1.1.2 Rekabetçi Açık Kaynak İstihbaratı

Herhangi bir istihbarat faaliyeti zaman, fayda ve maliyet kısıtlarına tabidir. Açık istihbarat faaliyetleri ise konvansiyonel sayılabilecek araştırma faaliyetleri ile büyük ölçüde kesişmektedir. Bu iki temel sebepten bir açık istihbarat faaliyetinin hem

⁸ Michael Fur, "Psychometrics – An introduction, Sage Publications 4th edition", 46

⁹ Cronbach, 1960.

konvansiyonel arařtırmadan farkını hem de rekabet aısından ayırıcı niteliklerini tanımlamak gerekir.

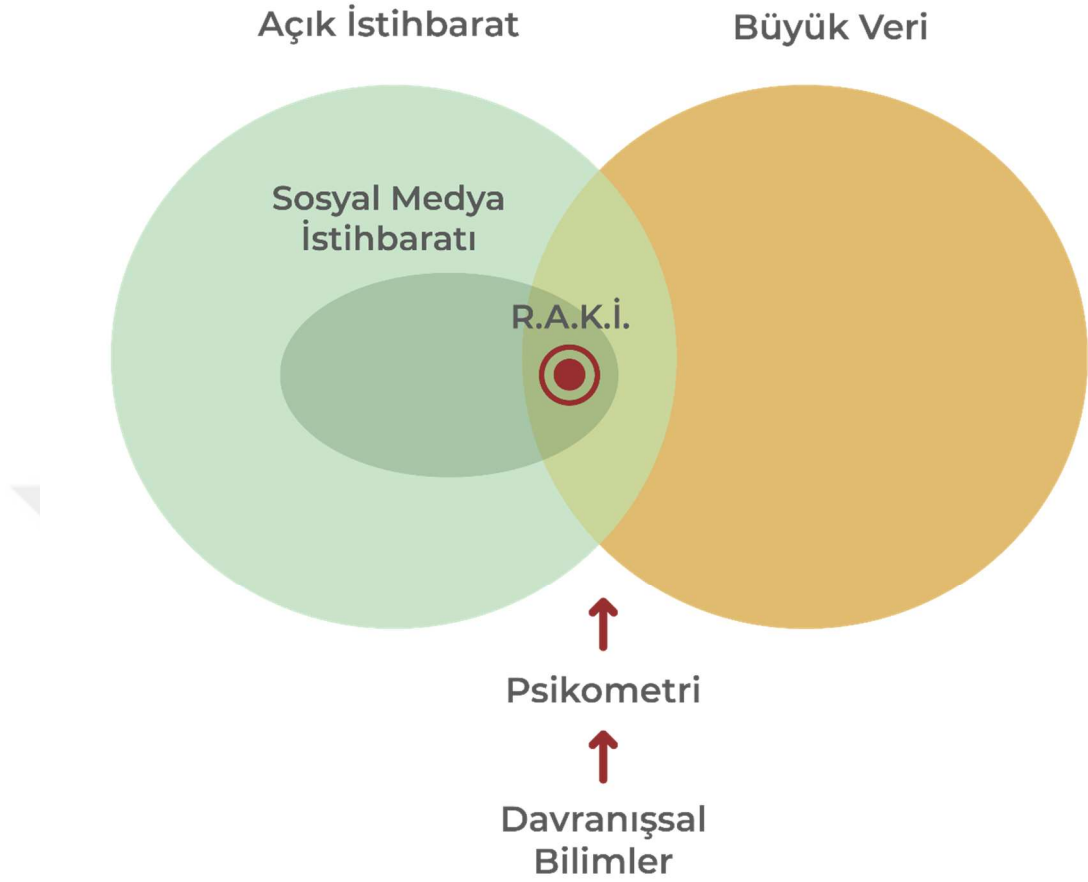
Bu ayırıcı nitelikler;

1-) Uygulama arařtırmacının kapasitesini ařmalıdır,

2-) ıktıya gre uygulamanın hesaplama maliyeti (computational cost) arařtırmacının maliyetinden dřk olmalıdır.

Bahsedilen bu iki temel zellięi optimum řekilde karřılayacak olan bir aık istihbarat faaliyeti bařta belirtilen temaların makul bir kompozisyonuyla saęlanabilir. Bu kompozisyon rekabeti aık kaynak istihbaratı olarak tanımlanacaktır.

Rekabeti niteliklere sahip bir aık kaynak istihbaratı (R.A.K.İ) ancak istihbarat metodolojisi kısmında tanımlanan disiplinlerin ve temaların birbirine entegre edildięi uygulamalar geliřtirerek saęlanabilir.



Görsel – 2 Rekabetçi Açık Kaynak İstihbaratı

Bahsedilen disiplinler arası çalışma alanı farklı ölçeklerde ve yaklaşık olarak son on yılda bilimsel literatüre girmiştir. Alanda bilimsel çalışmalar giderek artsa da bilimsel çalışmaların işleyen, entegre uygulamalara dönme oranı oldukça düşüktür. Buradaki temel problem bilimsel çalışmaların dijital ürünlerin/uygulamaların teknik sınırlılıklarından bağımsız olarak (ya da bunları dikkate almayarak) yapılmasıdır.

1.2 Tezin Amacı, Önemi, Hipotezler ve Yöntem

Giderek yaygınlaşan büyük veri teması, günümüzde herhangi bir konuyla ilgili fikirleri toplamak için doğrudan katılımcıyla iletişime geçme gerekliliğini ortadan kaldırmıştır. Özellikle açık sosyal ağlarda, kullanıcıların belli konularda fikirlerini sürekli olarak ifade etmesi, araştırma konularıyla ilgili kullanılabilir verilere sürekli ve bilgisayar destekli olarak ulaşabilmeyi mümkün kılmaktadır. Ek olarak veri kazıma teknolojileri

ile milyonlarca kullanıcının oluşturduğu verilere çok kısa sürelerde ulaşmak mümkündür. Bu çalışma çerçevesinde yaklaşık on milyon kullanıcının verisi elde edilmiş ve incelenmiştir. Bu ölçekte bir kullanıcı tabanına konvansiyonel araştırma yöntemleriyle ulaşmak mümkün değildir.

1.2.1 Tezin Sorunsalı

Araştırma konusu ile ilgili veriler açık sosyal ağlarda ulaşılabilir olsa da konvansiyonel yöntemlerden farklı olarak dağınık, yapılandırılmamış ve ham bir haldedir. Örneğin bir anket yapılırken katılımcıya araştırma konusu bağlamında doğrudan fikri sorulurken, açık sosyal ağlarda kullanıcı konuyla ilgili fikrini araştırma konusunun bağlamı dışında ifade edebilir ve doğrudan istenen sorulara cevap vermeyebilir, bu sebeple sosyal ağlardan elde edilen kullanıcı verilerinden istenen çıktıların elde edilebilmesi için bu veriler belli şablonlara göre yorumlanmalı ve işlenmelidir.

Tezin temel sorunsalı da burada ortaya çıkmaktadır. Araştırma konusu olan alandaki düzensiz ve yapılandırılmamış veriler, Türkçe dili için hangi psikometri ve doğal dil işleme yaklaşımı ile en isabetli sonuçlar elde edebilir? Bu sorunun cevabı tercih edilen yöntemin sınırları içerisinde verilmeye çalışılacaktır ve inceleme konusu olarak PANAS ölçeği tercih edilmiştir. Çalışmanın konusu olan yöntemin neden tercih edildiği literatür bölümünde izah edilecektir.

1.2.2 Yöntem

Araştırmanın yöntemi niceldir. Toplanan veriler ve bunlardan elde edilen çıktıların ilişkiselliği üzerinden hipotezler test etmektedir. Yöntem alt başlıkta “ilişkisel tarama modelidir”. Bu modelin seçilmesinin nedeni PANAS-T çıktılarının bütün veri kümeleri için ortak bir değerle sembolize edilebilmesi, sayısallaştırılabilmesidir. Literatürde PANAS-T’nin skollama yöntemi hem anket verilerini hem de sosyal medyadan toplanan diğer verileri ortak bir değerlendirme yönteminde eşitleyebilir. Böylece temel hipotezler korelasyon analizleri üzerinden test edilebilir.

Korelasyon analizlerinin yapılabilmesi için üç farklı veri kümesi üzerinde çalışılacaktır. Bu veri kümelerinden ilki analiz edilecek gündemler için kullanıcılardan

toplanan anket verileridir. Bu anket verileri kontrol değişkeni olarak kullanılacaktır. Anketler 18-65 yaş aralığındaki 500 farklı katılımcıdan elde edilmiştir. Demografik özellikler PANAS için anlamlı değildir¹⁰. Gündemler hakkında elde edilen veriler PANAS-T skorlarına dönüştürülerek kullanılacaktır.

Diğer iki veri kümesi ise Twitter'dan toplanan verilerin oluşturduğu veri kümeleridir. Bu verilerden ilki büyük veri kümesidir. 300 milyonun üzerinde içerikten oluşan bu veri kümesi, Twitter'da herhangi bir konu hakkında konuşulduğunda ve normal şartlarda duygu durumlarının nasıl dağıldığını verir. Diğer veri kümesi ise gündemlere ilişkindir. Gündem verilerindeki duygu durumları, normal şartlardaki dağılımla kıyaslanarak elde edilir ve PANAS-T skorlarına çevrilir.

Bu işlemlerden sonra hipotezlerin test edileceği ve korelasyon analizlerinin yapılacağı iki adet işlenmiş veri seti elde edilmektedir. Bu veri setleri PANAS-T şablonuna göre anketlerin duygu durum skorları ve aynı gündemlerin Twitter'daki duygu durum skorlarıdır. Nihayet bu iki skor arasında korelasyon analizi yapılarak hipotezler test edilir ve bulgular değerlendirilir.

1.2.3 Hipotez

Tezin temel hipotezi anket vb. konvansiyonel araştırma yöntemleriyle ulaşılamayacak büyüklükteki veri kümelerine, bilgisayar destekli büyük veri uygulamaları ile ulaşılabilceği ve bu uygulamaların konvansiyonel araştırma yöntemleriyle benzer sonuçlar/çıktılar elde edebileceğidir.

Bu hipotezin doğrulanması ya da yanlışlanması için, anket yoluyla ve bilgisayar destekli uygulama yoluyla iki ayrı veri kümesi elde edilecek ardından elde edilen verilerdeki gündemler PANAS-T ölçeğine göre skorlanacaktır. Bilgisayar destekli oluşturulan veri kümeleri için birden fazla yöntem mukayeseli olarak denenecektir.

Seçilen gündemler için anket verileriyle elde edilen PANAS-T skorlarıyla, bilgisayar destekli uygulama ile sosyal medya verileri üzerinden hesaplanan PANAS-T skorları

¹⁰ John R Crawford , Julie D Henry, "The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample" British Journal of Clinical Psychology. 2004 Sep;43(Pt 3):245-65. doi: 10.1348/0144665031752934.

arasında korelasyon testleri yapılarak sonuçların ne kadar benzeştiği tespit edilmeye çalışılacak ve hipotezin doğruluğu test edilecektir. Bu veri kümeleri arasında 0,60 ve üzeri yüksek seviyede korelasyon aranmaktadır. Bahsedilen değerin elde edilememesi durumunda ilgili yöntem için hipotez reddedilecektir ve veri kümeleri arasındaki PANAS-T skorlarının farklılaşma sebepleri incelenecektir. 0,60 ve üzeri korelasyon değerinin elde edilebildiği yöntemler içinse hipotez doğru kabul edilecektir.

Hipotez iki önerme üzerinden aşağıdaki şekilde kurulmuştur;

H_0 : Anket verileri ile sosyal medya verileri arasındaki PANAS-T değişkenleri arasında yüksek düzeyde ($H_0 < 0,60$ korelasyon) ilişki yoktur.

H_1 Anket verileri ile sosyal medya verileri arasındaki PANAS-T değişkenleri arasında yüksek düzeyde ($0,60 < H_1$ korelasyon) ilişki vardır.

1.2.4 Literatür ve Ölçek Tercihi

Psikometri disiplininde kullanılan yüzlerce farklı ölçek bulunmaktadır. Ancak bu ölçeklerden hangilerinin doğal dil işlemeye ve sosyal medya verilerine uyarlanabileceği belli teknik kısıtlamalarla tabidir.

Bu teknik sınırlamalardan;

- 1- Ölçeğin kişilik yapılarını ya da kişilerin duygu durumları ölçüyor olması.
- 2- Ölçeğin Türkçeye uyarlanabilir olması ve uyarlama için gerekli geçerlilik ve güvenilirlik çalışmalarının yapılmış olması

Zorunlu.

- 3- Ölçeğin, tercihen daha önce herhangi bir dilde sosyal medya istihbaratına uyarlanmış olması

ise çalışmanın uyarlanmasını ve derinleştirilmesi kolaylaştıran ancak opsiyonel bir gerekliliktir.

PANAS ölçeği bu kısıtları baz alarak çalışmada tercih edilmiştir. Yukarıda bahsedilmiş olan şablona uyan ölçeklerin sayısı hali hazırda oldukça azken, sadece PANAS istenen üç yeterliliğinin tümünü karşılamaktadır.

İstenen teknik sınırlara yakın olan bazı ölçeklerin çalışmada neden kullanılmadığı da kısaca açıklanacaktır. Gündemleri, kişilik yapısı yaklaşımlarını genelleyerek ölçmeye çalışan ve sosyal medya istihbaratına daha önce uyarlanmış bilinen en yaygın ölçek beş büyük kişilik faktörüdür. 2018 yılında Cambridge Analytica isimli şirket 87¹¹ milyon Facebook kullanıcısının kişisel verilerini kullanarak bu ölçek üzerinden kullanıcıların gündemlere nasıl tepki verdiğini ölçmüş ve kullanıcılara mikro segmentasyon yapmıştır.

Beş büyük kişilik faktörü gündemlerin ve kişilik yapılarının analizi için etkin bir yöntem olmakla birlikte çok büyük ölçeklerde kullanıcıların hassas verilerine ihtiyaç duymaktadır. Cambridge Analytica istenen ölçekteki hassas verileri yasal olmayan yollardan toplamış ve işlemiştir. Firma bu faaliyet neticesinde 2019 yılında 5 milyar dolar ceza almıştır.¹² Hassas verilerin toplanması ve amacı muğlak bir şekilde işlenmesi Türkiye’de de KVKK’ya (Kişisel verilerin korunması kanunu) uygun olmadığından bu yöntem tercih edilmemiştir. Hassas kullanıcı verileri (yaş, medeni durum vb. gibi özel bilgiler) yeterli büyüklükte elde edilmeden beş büyük faktörü ile yapılan analizler ve regresyonlar istenen sonuçları vermemektedir.

Diğer bir popüler yöntem olan enneagram’ın kitleye ve gündem analizine uyarlanmış çalışmaları¹³ bulunsa da sosyal medya verilerine uyarlanmış bir çalışma bulunmamaktadır. Ek olarak geçerlilik ve güvenilirlik testleri zayıftır ve üzerine ciddi tartışmalar mevcuttur.

¹¹ Annika Richterich, “How Data-Driven Research Fuelled The Cambridge Analytica Controversy”, (Maastricht University, The Open Journal of Sociopolitical Studies) ISSN: 2035-6609, 15 July 2018, 528

¹² Margaret Hu, “Cambridge Analytica’s black box” , Big Data & Society July–December 2020: 1–6, 2

¹³ Naomi Ekas, Ph.D, William C. Moncrief, Ph.D. The enneagram as a psychographic segmentation tool , (Texas Christian University), 2 May 2022

Kullanılmayan ölçeklerle ilgili örnekler çeşitlendirilebilir. Diğer taraftan PANAS istenen teknik kısıtların tümünü karşılamaktadır. PANAS doğrudan kişilik yapısı yerine, kişinin o anki ya da herhangi bir konuyla ilgili duygu durumunu ölçmektedir. Bu da gündemlerdeki hakim duygu durum analizlerinin PANAS şablonuyla yapılmasını kolaylaştırmaktadır. Ölçek 1998 yılında Çağay Dürü tarafından Türkçe'ye uyarlanmıştır, geçerlilik ve güvenilirlik çalışmaları yapılmıştır. Ayrıca literatürde, PANAS'ın doğrudan sosyal medya istihbaratına uyarlanmış bir versiyonu olan PANAS-T (Twitter için PANAS) versiyonu bulunmaktadır. Bahsedilen nitelikler ve literatür zenginliği bu ölçek üzerinden Türkçe ve bilgisayar destekli bir uygulamayı mümkün kılmıştır.

PANAS ve genişletilmiş versiyonu olan PANAS-X daha önce iki kez Twitter'daki veri analizi çalışmalarına uyarlanmıştır. (PANAS-T her iki alt ölçeği de ihtiva etmektedir, uyarlamanın çıktısı PANAS-T'dir) ¹⁴Uyarlamaya ilişkin ilk denemeler 2013 yılında yapılmıştır.¹⁵

Ayrıca ölçek, kişilerin duygu durumlarını kelimeler üzerinden tespit etmeye çalıştığı için anahtar kelimeler üzerinden bilgisayar destekli uygulamayı kolaylaştırmaktadır. Anahtar kelimeler üzerinden veri kazıma ve analiz, doğal dil işleme disiplini ile uyumludur böylece psikometri, büyük veri ve sosyal medya istihbaratı PANAS ölçeğinde bir arada kullanılabilir.

1.2.5 Uygulama

Veri analizi kısmında bahsedilen iddialar test edilecek ve sayısal çıktılarla desteklenmeye çalışılacaktır. Öncelikle büyük ve rastgele bir veri kümesi içerisinde duygu durumu ifade eden içeriklerin, duygu durumlarına göre dağılımı tespit edilecektir. Böylece, örneğin mutluluk ifade eden içeriklerin, rastgele bir dağılımda toplam içeriklerin ne kadarına tekabül ettiği bulunacaktır. Bu rastgele dağılım

¹⁴ Fabricio Benevenuto, Matheus Araújo, "Measuring sentiments in online social networks", WebMedia '13, 2013, 535

¹⁵ Pollyanna Goncalves, Fabricio Benevenuto, Meeyoung Cha "PANAS-t: A Psychometric Scale for Measuring Sentiments on Twitter" Systems, man, and cybernetics, Part C: Applications and reviews, VOL. XX, NO. XX, JUNE 2013

değerlerinin tespitinin ardından seçilen belli gündem başlıkları için aynı işlem uygulanarak gündemlerdeki hakim duygu durumları, rastgele dağılımlara oranlanarak tespit edilecektir. Örneğin mutluluk ifadeleri rastgele veri kümesinde toplam içeriğin %5'ine tekabül ediyorsa ancak seçilen gündem de toplam içeriğin %10'unda mutluluk ifadeleri varsa, o zaman gündemde mutluluk duygu durumunun hakim olduğu söylenebilir. Hesaplama, duygu durumlarının dağılımı, Türkçeleştirme ve stemming (kök bulma) metodolojilerine ilgili başlıklarda detaylı olarak yer verilecektir. Bahsedilen mantıktaki uygulama yöntemi için PANAS-T'nin literatürde kendi ölçüm prosedürü vardır ve çalışmadaki korelasyon hesaplarında bu yöntem kullanılacaktır ölçümün nasıl yapıldığı yine ilgili başlıkta incelenecektir.

1.3 Tezin Sınırlılıkları

Makine öğrenimi ve bunlara ilişkin uygulamaların korelasyonlarla ölçümü üzerinden, iki yönlü epistemolojik tartışmalar bulunmaktadır. Makinenin yorumlama açısından insanın sınırlarını aşıp aşmadığının tespit edilememesi bilgi felsefesi anlamında korelasyon analizini sorgulanabilir yapmaktadır. Örneğin bu çalışmada anket verileri, yani insanların doğrudan yorumladığı veriler kontrol değişkeni olarak kullanılmıştır. Ancak uygulamanın eriştiği veri kümeleri daha büyük olsa da, bu veriler küçük miktarlardaki kontrol değişkeni üzerinden yorumlanmıştır. Bu iki veri kümesinden hangisinin daha iyi yorum sağladığı ya da hangisinin sabit olması gerektiği üzerinden tartışmalar sürmektedir. Bu tartışmalar çalışmada ihmal edilmiştir.

Verilerin mülkiyeti ile ilgili hukuki tartışmalar sürmektedir ancak GDPR da dahil olmak üzere bu konuda uzlaşmış hukuki bir altyapı bulunmamaktadır. Veri mülkiyeti özelinde de tartışmalar çalışmanın dışında bırakılacaktır.

Veri toplama araçları anket ve bilgisayar destekli veri kazıma yöntemleridir. Buna ek olarak verilerin analizinde dağıtık işleme yöntemleri kullanılmıştır. Evren ve örneklem, “Türkçe dilindeki içerikleridir.” Coğrafi bir sınırlama yoktur. Bütün coğrafyalardan üretilmiş Türkçe içeriklerin tamamı analize alınmıştır. Zaman her gündem özelinde farklı seçilse de 2020 yılı ve sonrası içerikler kullanılmıştır. 2020 yılından önceki içerikler analize alınmamıştır.

Araştırma yöntem açısından, verilerin analizi ve bunlardan üretilen varsayımların ilişkisel tarama modeli ile test edilmesiyle sınırlıdır. Metodolojilerin kıyaslanması ve dönüştürülmesine ilişkin varsayımlarda ise farklı akıl yürütme tekniklerini iç içe kullanmaktadır.

Psikometrik yöntemlere ilişkin faktör analizleri yine tezin kapsamı dışındadır. Bu tartışmalar ihmal edilecek ve PANAS için Türkiye’de en yaygın kullanılan, en fazla kabul görmüş faktör analizleri baz alınacaktır.

1.3.1 İlgili Mevzuatlar ve Hukuki Sınırlılıklar

Veri kazıma ve işleme uygulamaları üç farklı alanda ilgili mevzuatlara ve hukuki sınırlılıklara tabidir.

Bunlar :

- a- Veri kazıma ve tarama yöntemlerine ilişkin teknolojilerin, yazılım iskeletlerinin (framework) kullanımına ilişkin hususlar.*
- b- Verilerin işlenmesine ilişkin hususlar.*
- c- Verilerin mülkiyetine ilişkin hususlardır.*

1.3.1.1 Veri Kazıma ve Tarama Yöntemlerine İlişkin Hukuki Sınırlılıklar

Hemen her açık sosyal ağ gibi Twitter verilerinin işlenmesi içinde iki temel yaklaşım bulunmaktadır. Bunlardan ilki hizmet sağlayıcının (mevcut çalışmada Twitter) kendi geliştirdiği API’leri (uygulama programlama arayüzü) ile veri isteklerini belli sınırlar ve ücretlendirme ile karşılamaıdır. Kullanıcılar istek oluşturma, gönderme, alma, hız sınırlarını yönetme ve gerçek zamanlı cevapların, nesne analizlerinin sınırlarını tanımlayarak bu hizmeti alırlar. Bütün ilgili detaylar Twitter’ın geliştiriciler (developers) dokümanında tanımlanmıştır.¹⁶

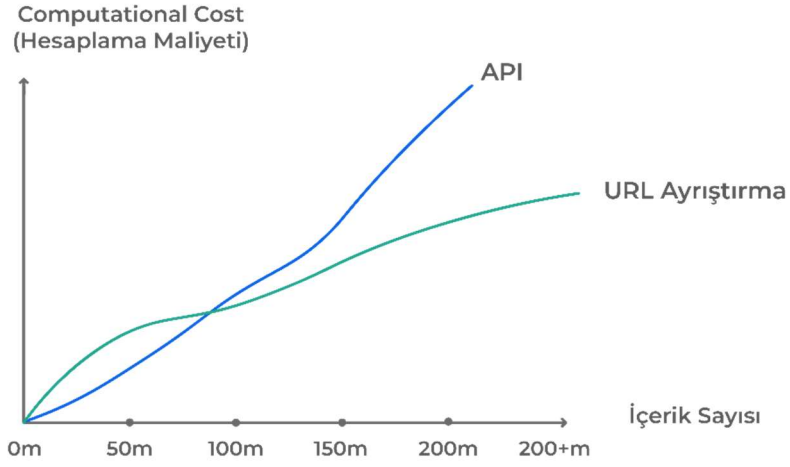
¹⁶ Michael W. Kearney, (School of Journalism, Informatics Institute, University of Missouri) The Journal of Open Source Software, 24 October 2019

Bu yöntemle büyük veri kümelerine istenen veri çeşitliliği ile ulaşmak oldukça güçtür çünkü veri sağlayıcı bütün veri ambarının analiz edilebileceği bir ölçekteki veri miktarını ve çeşitliliğini üçüncü taraflarla paylaşmak istemez. Belli bir dilde bütün verilerin %x miktarına ulaşabilen bir kullanıcı, servis sağlayıcının bütün verilerine ilişkin sağlıklı bir çıkarım yapabilecek altyapıya sahip olabilir.

API'ler küçük veri kümeleri için hazır ve kolaylaştırıcı bir altyapı sağlar bu yüzden Ar-ge'nin ilk aşamalarına uygundur ancak konvansiyonel miktarları aşan büyük veri kümeleri ancak farklı veri kazıma teknolojileri ile etkin olarak elde edilip işlenebilmektedir. Büyük veri kümelerinde API'ler maliyet etkin bir şekilde ölçeklendirilebilir değildir. Hizmet sağlayıcının istediği ücret ödense bile en yüksek ücretli, en büyük paketler bile bu çalışma ölçeğinde ulaşmak istenen veri ölçeğini karşılamamaktadır.

İkinci yöntem ise farklı dns ve ağ teknolojileri kullanmak ve url ayrıştırıcı yazılım iskeletleri ile veri kazıma altyapısını oluşturmaktır. Çalışma bu ikinci yöntemi kullanarak yapılmıştır. Rastgele veri seti için toplamda 308.139.050 Tweet çekilmiş ve JSON formatında kaydedilmiştir.¹⁷ Teknolojik altyapıya ilgili başlıkta detaylı olarak değinilecektir.

¹⁷Ek – 1 Randombase.JSON



Görsel – 3 Veri Kazımada Hesaplama Maliyetleri

Veri kazıma yöntemi de dahil olmak üzere hemen bütün uygulamalar Twitter’ın API ya da diğer sorgu sistemlerinden yararlanırlar ancak url ayrıştırma ile referans verilen sistemlerde yazılım iskeleti Twitter API’lerine odaklanmamıştır ve/veya API’leri farklı protokollerle manipüle ederek sorgu sayısını ve elde edilebilir veri miktarını genişletmektedir.

1.3.1.2 Verilerin İşlenmesine İlişkin Hususlar

Türkiye’de kullanıcı verilerinin nasıl işlenebileceği 7 Nisan 2016 tarihinde yayınlanmış olan 6698 sayılı Kişisel Verilerin Korunması Kanununda tanımlanmıştır. Bu çalışmanın yürürlükteki kanuna uyumluluğunu gerektiren iki konu vardır bunlar:

a-işlenecek verilerin niteliği

b-işlenen verilerin amaca uygunluğudur.

İlgili kanunun 5.maddesine göre kişisel veriler kişinin açık rızası olmadan işlenemez ve büyük miktarlardaki veri kümeleri için kişilerin her birinden açık rıza alınması operasyonel açıdan oldukça güçtür. Ancak kanunun aynı maddesinde verilerin

kişilerden açık rıza alınmadan hangi istisnai durumlarda işlenebileceği de tanımlanmıştır. 5. Maddenin d fıkrasına göre eğer veriler ilgili kişinin kendisi tarafından alenileştirilmiş ise ¹⁸ kişiden açık rıza alınmadan işlenebilir.

Twitter’da verilerin alenileştirilmesi açısından hesap sahipleri için iki farklı güvenlik seçeneği bulunmaktadır. Hesap sahipleri kendi profillerini kapatarak, paylaşımlarını sadece onay verdiği takipçilerinin göreceği şekilde kısıtlayabilir veya profilini herkese açık yaparak paylaşımların herkes tarafından görülmesini sağlayabilir. Kanuna göre kapalı bir profilde kişinin kendisi verilerini alenileştirmemiş aksine saklamıştır diğer taraftan açık profil sahipleri diğer seçeneği uygulamayarak verilerini alenileştirmiştir.

Bahsedilen seçim esas alınarak kazınan verilerin tamamı açık profillerden elde edilmiş ve kapalı profillerdeki veriler, ilgili veri tabanlarına alınmamıştır.

İkinci konu olan verilerin amaca uygun olarak işlenmesidir. Bu da ilgili kanunun 11. Maddesinin c fıkrasında ¹⁹ haklar ve yükümlülükler başlığı altında ilgili kişinin hakları kısmında tanımlanmıştır. Uygulama anlamında kanunun bu maddesi verinin içeriği üzerinden değerlendirilmiştir. Örneğin ekonomi ile ilgili yapılan bir paylaşımdan kişinin politik görüşlerine ilişkin çıktılar üretilemez, veri yine ekonomi çerçevesinde işlenmelidir. İçeriğin niteliğine ilişkin olan kanunun bu maddesi kısmen muğlak olsa da işleyiciye çıkacak sonuçlar, bağlantılar ve veriyi işleme yolları açısından bir fikir vermektedir.

PANAS konteksinde bahsedilen ikinci konuya ilişkin sorun tamamen ortadan kalkmıştır. Çünkü PANAS hali hazırda duygu durumu ifadelerini belli kelimelerle eşleştirmiştir böylece bu kelimeler üzerinden yapılan taramalar ve veri işleme faaliyeti doğrudan amaca uygundur.

Ek olarak verilerin kişilerle bağının koparılması ve homojen bir hale getirilerek kitleleştirilmesi veri işleyicilerde genel bir endüstriyel alışkanlıktır. Böylece hassas bilgiler veri işleme faaliyetinde kullanılırken, elde edilen çıktılarla veri sahiplerinin

¹⁸ Kişisel verilerin korunması kanunu, kanun numarası : 6698 resmi gazete 7/4/2016 sayı : 29677 5. Madde b fıkrası.

¹⁹ Kişisel verilerin korunması kanunu, kanun numarası : 6698 resmi gazete 7/4/2016 sayı : 29677 5. Madde c fıkrası.

geriye dönük bağı kopararak mevzuatla uyum sağlanır. Bu çalışma da aynı sebeplerden kişilerin duygu durumlarına değil seçilen gündemlerin duygu durumuna yoğunlaşmıştır.

1.3.1.3 Verilerin Mülkiyetine İlişkin Esaslar

Günümüzde verilerin mülkiyetine ilişkin tartışmalar, veri kavramı üzerinde tam bir uzlaşma bulunmadığı için muğlaklığını korumaktadır. Dahası veri kavramı ilgili olduğu bağlamın içinde spesifik olarak tanımlanmadıkça ve salt bir karakter seti olarak ele alındığı sürece mülkiyet açısından uzlaşılabilir bir tanımlı bulunmamaktadır.²⁰

Farklı hukuk ekolleri özelinde konu tartışılrsa da GDPR (Genel Veri Koruma Yönetmeliği) de dahil olmak üzere verilerin sahipliğine ilişkin bir düzenleme bulunmamaktadır. Bu yüzden işlenen veriler mülkiyet açısından da kanuna uygun olarak işlenmiştir.

2.PANAS

PANAS (Positive and Negative Affect Schedule), pozitif ve negatif duygu durum ölçeği, 1988 yılında David Watson, Lee Anna Clark, A.Tellegen tarafından geliştirilmiştir bir psikometri ölçeğidir. Yöntem duygunun yapısını kendinden önceki çalışmalardan farklı şekilde çok boyutlu olarak tanımlar. PANAS'a göre duygu yapısı iki baskın ve nispeten bağımsız boyuttan oluşur bu boyutlar negatif ve pozitif duygu durumlarıdır.²¹

Baskın duygu durumunun tespiti için PANAS, duygu durumlarını temsil eden 20 kadar betimleyici sıfat belirlemiştir. Bu kelime setinin yarısı pozitif duygu durumunu, diğer yarısı ise negatif duygu durumlarını temsil eder. Baskın duygu durumu tespit edilmeye çalışılan denek, kelimeleri içeren ve her kelimeyi 1-5 arası puanlayan likert ölçekli bir

²⁰ Professor Dr. Ali M Al-Khouri “Data Ownership: Who Own's My Data”, , International Journal of Management & Information Technology ISSN: 2278-5612 Volume 2, No 1, November, 2012, 1-2

²¹ David Watson and Lee Anna Clark, “Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales” Journal of Personality and Social Psychology 1988. Vol. 54. No. 6, 1063

teste tabi tutulur. Buradan elde edilen çıktı ile deneğin hakim duygu durumu tespit edilir.

PANAS'taki en kritik unsur seçilen her bir terimin, kendi grubundaki duygu durumunu, diğer terimlerden bağımsız ve geçişken olmayacak biçimde ancak yoğun olarak temsil etmesi gerekliliğidir. Seçilen x terim, x faktörünü yoğun ve içinde bulunduğu duygu durumunu yeterli ölçüde temsil ederken diğer terimleri ya da duygu durumunu temsili zayıf (olabildiğince düşük) olmalıdır. Böylece duygu durumlarını ifade eden her bir terim birbirinden olabildiğince bağımsız ve izole olacaktır. Watson, Clark ve Tellegen bu gerekliliği ve gerekliliğe yönelik uygulamayı “saf belirteçler” (pure markers) olarak kavramsallaştırmıştır.

Böylece PANAS terimlerinin geçerlilik, güvenilirlik ve terimler arası bağımsızlık çalışmaları hali hazırda yapılmıştır ve bu çalışmalar neticesinde elimizde hazır bir “duygusal sözlük” bulunmaktadır. Bu da çalışmayı, sözlükteki anahtar kelimeleri kullanarak doğrudan doğal dil işleme disiplinlerine entegre edebilmeyi kolaylaştırmaktadır.

2.1. Tanımlayıcı Kelimeler, Güvenilirlik ve Geçerlilik

PANAS'ın tanımlayıcı kelimelerinin her biri için pozitif duygu durumu ve negatif duygu durumuna ilişkin varimax faktör yükleri tespit edilmiştir.

Pozitif Tanımlayıcı	Pozitif Faktör Yükü	Negatif Faktör Yükü	Pozitif Tanımlayıcı	Pozitif Faktör Yükü	Negatif Faktör Yükü
Enthusiastic	0,75	-0,12	Scared	0,01	0,74
Interested	0,73	-0,07	Afraid	0,01	0,70
Determined	0,70	-0,01	Upset	-0,12	0,67
Excited	0,68	0,00	Distressed	-0,16	0,67
Inspired	0,67	-0,02	Jittery	0,00	0,60
Alert	0,63	-0,10	Nervous	-0,04	0,60
Active	0,61	-0,07	Ashamed	-0,12	0,59
Strong	0,60	-0,15	Guilty	-0,06	0,55
Proud	0,57	-0,10	Irritable	-0,14	0,55
Attentive	0,52	-0,05	Hostile	-0,07	0,52

Tablo - 1 PANAS kelimeleri faktör yükleri (İngilizce)²²

Faktör yükleri incelendiğinde tanımlayıcı kelimelerin tümünün kendi duygu grubu için 0.50 değerinin üzerinde kaldığı ve karşıt duygu grubu için 0,01 değerinin altında olduğu görülür. Buna göre PANAS'ın tanımlayıcı kelimeleri istenen “saf belirteç” durumuna uygun haldedir.

Geliştiriciler doğru tanımlayıcı kelimelere ulaşana kadar birden fazla versiyonu denemiş ve farklı istatistik yöntemlerle test etmiştir. Faktör analizi PANAS'ın tek geçerlilik ve güvenilirlik parametresi değildir ancak önceki versiyonlar ve PANAS'a ilişkin istatistik çalışmalar bu tezin kapsamını aşmaktadır. Netice itibarıyla PANAS

²² David Watson and Lee Anna Clark, “Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales” Journal of Personality and Social Psychology 1988. Vol. 54. No. 6, 1065-1070

günümüzdeki psikometri çalışmalarında en yaygın kullanılan yöntemlerden birisidir ve onlarca dile çevrilerek ve farklı uygulamalara uyarlanarak kullanılmaktadır.

2.2 Türkçe Uyarlama

PANAS 1998 yılında Çağay Dürü tarafından Türkçeleştirilmiş ve faktör analizi yapılmıştır. Yapılan analizde pozitif ifade içeren bir kelime dışında İngilizceden yapılan doğrudan kelime çevirilerinin istenen faktör değerlerini karşıladığı tespit edilmiştir. İç tutarlılık değerleri için Cronbach Alfa katsayıları pozitif duygu durumu için 0.85, negatif duygu durumu için 0.86 olarak tespit edilmiştir.

Pozitif Tanımlayıcı	Pozitif Faktör Yükü	Negatif Faktör Yükü	Pozitif Tanımlayıcı	Pozitif Faktör Yükü	Negatif Faktör Yükü
Hevesli (Enthusiastic)	0,73	-0,26	Ürkmüş (Scared)	-0,15	0,66
İlgili (Interested)	0,57	-0,25	Korkmuş (Afraid)	0,01	0,70
Kararlı (Determined)	0,70	-0,01	Mutsuz (Upset)	-0,34	0,66
Heyecanlı (Excited)	0,13	0,48	Sıkıntılı (Distressed)	-0,23	0,69
İlhamlı (Inspired)	0,67	-0,02	Asabi (Jittery)	0,00	0,60
Uyanık (Alert)	0,63	-0,10	Sinirli (Nervous)	-0,04	0,60
Aktif (Active)	0,61	-0,07	Utanmış (Ashamed)	-0,12	0,59
Güçlü (Strong)	0,62	-0,26	Suçlu (Guilty)	-0,06	0,55
Gururlu (Proud)	0,57	-0,10	Tedirgin (Irritable)	-0,14	0,55

Dikkatli (Attentive)	0,52	-0,05	Düşmanca (Hostile)	0,09	0,61 ²³
-------------------------	------	-------	-----------------------	------	--------------------

Tablo – 2 Türkçe Betimleyici Sıfatların Faktör Yükleri ²⁴

Çağay Dürü'nün Türkçe çevirisinde sıfatlardan bir tanesinin aynı yönlü ve yüksek miktarda faktör yükü aldığı tespit edilmiştir. Bu sıfat İngilizce “alert” sözcüğünden tetikle olarak çevrilen sıfattır ve faktör yükleri pozitif duygu durumu için 0.33, negatif duygu durumu içinde 0.51 olarak bulunmuştur. Bu haliyle duygu durumları için bağımsız belirteç olamayacak sıfatın Türkçe'ye çevirisinde “uyanık” sıfatı tercih edilmiştir ve bu sıfatın faktör yükleri kullanılabilir oranlarda çıkmıştır. Sıfattaki bu değişikliği ve varimaks hesaplamalarına Tülin Gençöz'ün çalışmasında yer verilmiştir ancak bahsedilen sıfat için ilk çeviriyi kimin yaptığı bilgisine ulaşamamıştır.²⁵

3. Veri Seti

Toplamda 6.037.830 kullanıcıdan, 2020-01-01 - 2022-04-28 tarih aralığı için elde edilen 308.139.050 adet Twitter içeriği, ilgili başlıkta bahsedilen veri kazıma teknolojileri ile toplanarak JSON formatında saklanmıştır.²⁶

PANAS bu veri setine iki aşamalı olarak uygulanacaktır. Bunun sebebi doğrudan betimleyici sıfatların Twitter içeriklerinde birden fazla formda bulunabilmesidir. Twitter'daki ifadelerde betimleyici sıfat, anket uygulamasındaki gibi doğrudan ölçülen bağlama yönelik olarak kullanıcıya verilemez ve sıfat Twitter içeriklerinde farklı bağlamlarda bulunabilir.

Büyük veri seti betimleyici sıfatların rastgele ve normal değerlerini tespit etmekte kullanılacaktır. Buna ek olarak 10 adet ayrı gündem belirlenmiş ve her gündem için yaklaşık 10.000 – 500.000 bin arası veri kazınarak, veri tabanına kaydedilmiştir.

²⁴ Tülin Gençöz, “Pozitif ve Negatif Duygu Ölçeği Geçerlilik ve Güvenilirlik Çalışması”, (ODTÜ) Türk Psikolojisi Dergisi 2000, 15(46), 19-26

²⁵ Tülin Gençöz, “Pozitif ve Negatif Duygu Ölçeği Geçerlilik ve Güvenilirlik Çalışması” (ODTÜ) Türk Psikolojisi Dergisi 2000, 15(46), 19-26,

²⁶ Ek-1 randombase.JSON

²⁷Temel amaç rastgele elde edilen veri setiyle, herhangi bir gündeme ilişkin elde edilen veri seti arasındaki tanımlayıcı sıfat farklarını bularak hakim duygu durumlarını PANAS şablonuna göre tespit etmektir.

3.1 Büyük Veri Seti için Uygulama ve Rassallaştırma

308.139.050 adet Twitter paylaşımından oluşan büyük veri seti, amaca yönelik olarak geliştirilen bir uygulama ile elde edilmiştir. Yönteme göre veri kazıyıcı program Twitter'ın arama ara yüzünü ve mantığını kullanarak bir sorgu gönderir. Uygulamaya girilen anahtar kelimeler aynı Twitter'ın arama arayüzü kullanılarak yapılan aramalar gibi sonuç döndürür ve uygulama 1 Ocak 2020 – 28 Nisan 2022 tarih aralığı için yapılan sorguları seçilen anahtar kelimelere dağıtarak içerikleri veri tabanına çeker.

Bu kısımda temel veri setinin rasgele olması en kritik unsurdur ve rasgelelik duygu durumları açısından ele alınmalıdır. Çünkü PANAS şablonu hali hazırda duygu durumlarını tespit etmek için önceden oluşturulmuş ve test edilmiş teknik sınırlara sahiptir. Eğer veri kümesi PANAS'taki duygu durumlarından bir veya birkaç tanesini rasgele olmayacak şekilde içerirse veya beslerse temel veri seti artık PANAS açısından rasgele olmaktan çıkar ve duygu durumu için normal dağılım değerleri tespit edilemez. Temel veri setinin oluşturulmasındaki amaç duygu dağılımına ilişkin normal değerleri elde etmektir.

Çalışmada, veri setini duygu durumu açısından rassal bir hale getirmek için doğal dil işlemede yaygın olarak kullanılan bir kelime filtreleme yöntemi tersten uygulanmıştır. Doğal dil işlemede içerikler duygu durumlarına göre analiz edilmeye çalışılırken, dilde yoğun olarak bulunan ancak cümlelerin anlamını, duygu durumunu etkilemeyen sözcükler tespit edilerek işlenecek veri setinin dışına alınır. Disiplinde bu kelimeler (stop words) etkisiz kelimeler tanımıyla kavramlaştırılmıştır. Örneğin a, ama, beş, bile gibi kelimeler duygu durumu açısından nötrdür bu yüzden etkisiz kelimeler, içerikler bilgisayar tarafından analiz edilirken vektör miktarını ve boyutunu faydasız bir biçimde arttırarak hesaplama maliyetinin verimsizleşmesini sağlar. ²⁸

²⁷ Ek-2 events.JSON

²⁸ Mücahit Altıntaş, “Kullanıcı destek sistemlerinde yardım biletlelerinin otomatik sınıflandırılması” (İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü), yüksek lisans tezi, Eylül 2014, 36

Çalışmadaki yaklaşım ise veri setini rassallaştırmak için etkisiz kelimeleri tersten kullanmıştır. Etkisiz kelimeler içeriğin duygu durumu açısından bir şey ifade etmediğine göre Twitter’ın arama algoritması kullanılarak çekilecek içerikler için tarama etkisiz kelimeler kullanılarak yapılabilir. Böylece elde edilen veri seti PANAS’ın duygu durumlarına ilişkin mekaniği etkilemeyecek, veri tabanına çekilen içerikler duygu durumu açısından rasgele olacaktır. Ardından bu içerikler içerisinde PANAS’ın duygu durumunu içeren kelimeler saydırılarak rasgele bir veri kümesinde duygu durumlarının nasıl dağıldığı bulunabilecektir.

Türkçe dili için kullanılan etkisiz kelimeler Wikipedia’nın paylaştığı Türkçe’de en fazla kullanılan kelimeler verisi baz alınarak tespit edilmiştir ve hemen hemen bütün çalışmalarda benzer etkisiz kelimeler kullanılmaktadır.²⁹ Bu çalışmada ise yine aynı şablon baz alınmış ancak filtre duygu durumu açısından daha sıkı tutulmuştur. Çalışmada kullanılan etkisiz kelimelerin benzer doğal dil işleme çalışmalarında kullanılan kelimelerle %90’ın üzerinde benzeştiği görülecektir.³⁰

Tarama, toplamda seçilen 189 etkisiz kelime ile 1 Ocak 2020 – 28 Nisan 2022 tarihleri arasındaki içeriklerin elde edilmesi için başlatılmıştır. Verinin toplandığı gün sayısı 849 gündür ve veri kümesinin yine rassal olabilmesi için geniş bir tarih aralığı seçilmiştir. Veri seti için tarama yapılırken ekte belirtilen etkisiz kelimelere dağıtım eşit olarak yapılmıştır ve veriler en son tarihten başlanarak geriye doğru çekilmiştir.

3.2 Betimleyici Sıfatların Saydırılması ve Stemming (Kök Bulma)

İngilizce dili için yapılan çalışmalarda, veri tabanına alınan içerikler içinde PANAS’ın betimleyici sıfatları doğrudan saydırılarak içeriğin duygu durumu tespit edilmeye çalışılmıştır. Bu yaklaşım iki aşamalıdır, sıfatlar doğrudan, olduğu gibi (as it is) olarak tespit edilir. “alert” kelimesi aranırken alert içeren (contains) diğer kelimeler analize alınmaz örn : “alerted” analiz dışında tutulur. Diğer bir yaklaşımda ise bunun tam tersi uygulanır, sıfatı içeren kök esas alınır ve kökün kapsadığı bütün içerikler analize dahil edilir ancak uygulama iki aşamalıdır. Birinci aşamada sıfatı içeren bütün içerikler veri

²⁹ Muharrem Önal, “Sözlüksel benzeşim yöntemi ile android tabanlı duygu analizi uygulaması” (Fırat Üniversitesi Fen Bilimleri Enstitüsü) Yüksek lisans tezi, Mart 2017, Ek-1 Stop Words listesi, 69

³⁰ Ek – 3 Çalışmada kullanılan Türkçe Etkisiz Kelimeler (Stop Words)

tabanına toplanırken, ikinci aşamada bu veriler duygu durumunu 1.tekil şahısla ifade edecek şekilde sadeleştirilir ve veri temizlenir.

PANAS-T, PANAS'ın Twitter için özelleştirilmiş bir versiyonudur ve yukarıda bahsedilen metotlar PANAS-T üzerine literatürde bilinen iki çalışma temel alınarak tespit edilmiştir. Aslında PANAS-T üzerine çalışmalar çok nadirdir ve literatürde sadece Güney Amerika'da çalışmalar bulunmaktadır. İngilizce dili için ve doğrudan Twitter'da psikometriye yönelik olarak yapılmış olan tek çalışma veri temizleme aşamasında bahsedilen ikinci yöntemi kullanmaktadır.³¹

Bahsedilen çalışmada 54.981.152 kullanıcıdan 1.755.925.520 adet tweet, veri tabanına çekilmiştir (rassallaştırmaya ilişkin herhangi bir referans yoktur). Bu Tweetler içerisinde PANAS ve PANAS-X'in betimleyici sıfatları kök olarak tespit edilmiştir. Tespit aşamasında mükerrer sayımdan kaçınılmış ve ilk bulunan sıfat, ilgili olduğu duygu durumu kategorisi ile ilişkilendirilmiştir. Örneğin "I'm **alerted** and **happy**" gibi bir içerikte ilk sıfat "**alerted**" olduğu için sadece bu sıfatla ilgili kategoriye koyulmuştur ve "**happy**" sıfatı ihmal edilmiştir. Eğer içerik "I'm **happy** and **alerted**" olsaydı, sadece "**happy**" sıfatını içeren kategoride değerlendirilecek ve "**alerted**" sıfatı ihmal edilecekti.

Aynı çalışmanın ikinci aşamasında ise sıfatın doğrudan kişilerin duygu durumuna ilişkin referans vermediği, betimleyici sıfatın ancak birinci tekil şahısla ifade edilmesi durumunda kişilerin duygu durumuna referans verdiği varsayılmıştır. Bu varsayıma göre veriler "I'm", "I am", "I", "am", "feeling", "me" and "myself" sözcükleri ile filtrelenmiştir. Analize alınacak olan içeriklerin betimleyici sıfatla birlikte bahsedilen filtre kelimelerden birini içermesi gerekmektedir. Benzer bir filtreleme yaklaşımı daha önce İngilizce dilinde Twitter duygu durumu ile borsa hareketleri arasındaki korelasyonun tespiti için kullanılmıştır.³²

³¹ Pollyanna Goncalves, Fabricio Benevenuto, Meeyoung Cha "PANAS-t: A Psychometric Scale for Measuring Sentiments on Twitter" Systems, man, and cybernetics, PART C: Applications and Reviews, VOL. XX, NO. XX, JUNE 2013

³² J. Bollen, A. Pepe, H. Mao, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena" Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM), 2011.

Veri temizleme aşamasından sonra toplam veri seti içerisinde betimleyici sıfatların bulunduğu, duygu durumu ifade eden içerik sayısının toplamı 479.356.536 tweet olarak bulunmuştur. Böylece toplam içeriğin %27'sinin duygu durumu içerdiği tespit edilmiştir.

3.2.1 Türkçeleştirme ve Betimleyici Anahtar Kelimelerin Seçimi

PANAS ya da PANAS-X³³ bilgisayar bilimleri açısından temelde bir sınıflandırma (classification) problemidir. PANAS iki boyutlu ve basit bir sınıflandırmayken (pozitif ve negatif) PANAS-X çok boyutlu (multi-dimensional) bir sınıflandırma problemidir. Ancak yaygın sınıflandırma problemlerinden ve bunları çözmek için kullanılan öğrenme ya da sınıflandırma algoritmalarından farklı olarak PANAS iki katmanlı bir problemidir.

Bahsedilen iki katmanlılık sadece bir problem değil aynı zamanda bir avantajdır ve fark şöyle açıklanabilir. Text içeriklerden müteşekkil bir veri kümesi verilseydi ve bunların yine önceden belirlenmiş sabit bir metoda bağlı sınıflandırmaya tabi tutulması istenseydi Bayesian vb. yaklaşımlar ile sorun tek katmanda çözülebilirdi. Örneğin basit Bayes sınıflandırıcısı, nitelikler sınıfa göre bağımsız olduğunda optimaldir.³⁴ Bilinen en yaygın örnek üzerinden açıklarsak, spam e-postalarda sınıflar içeriğin spam olması veya spam olmamasıdır. Nitelikler ise e-posta üzerinden toplanan verilerdir ve bu veriler kullanıcı tecrübesini içermeyip sadece içerik üzerinden analiz edilseydi, bayesian yaklaşım başarılı olamazdı. Çünkü nitelikler içerikte geçen anahtar kelimelerdir dersek, aynı anahtar kelime her iki sınıfta da bulunabilir ve nitelikler sınıfa göre bağımsız olmaz. Ancak kullanıcı tecrübesi burada ek veri oluşturarak bağımsızlığı sağlar. E-postaları spam ya da değil diye işaretleyen kullanıcı sınıflar için ayırıcı niteliklerin (anahtar kelimelerin) tespit edilmesini sağlar ve sistem çalışır.

³³ David Watson and Lee Anna Clark T"HE PANAS-X Manual for the Positive and Negative Affect Schedule - Expanded Form", (The University of Iowa), 1994

³⁴ Pedro Domingos, Michael Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss" Kluwer Academic Publishers. Manufactured in The Netherlands, 1997, Machine Learning 29, 103–130

PANAS'ta ise klasik sınıflandırma probleminden farklı olarak ayırıcı nitelik önceden verilmiştir.³⁵ Sorunun yine klasik sınıflandırmadan farklı olarak ikinci katmanı ise önceden verilen anahtar kelimelerden yola çıkılarak farklı cümle yapılarının gerçekten o anahtar kelimenin işaret ettiği kategoriye temsil edip etmediğidir. Temelde PANAS'ın anahtar kelimeleri ile analiz yapılırken ankete tabi tutulan kullanıcı, verilen kelime ve duygu durumuna ilişkin nasıl hissettiğini likert bir ölçekte belirtir. Bu soru doğrudan duygu durumuna ilişkindir. Diğer taraftan PANAS-T'de kelimenin farklı versiyonlarda ve uygun kapsayıcı haliyle geçtiği cümleyi doğrudan duygu durumu olarak ele almak durumundayız. Bu yaklaşımda kelimenin ve ifadenin cümle açısından belirleyici olduğu varsayıldı ve bu varsayım sınıflandırmaya ilişkin anahtar kelime yaklaşımını belirledi.

Sınıflandırma problemi içeren ve anahtar kelime seçiminin kritik olduğu bütün bilgisayar destekli veri analizi çalışmalarında (makine öğrenimi dahil), sınıflandırma için optimum anahtar kelime seçimi metodolojisine ilişkin yaklaşım iki ana başlık altında gruplanabilir. Bu başlıklardan ilki “bag of words”³⁶(kelime torbası) diğeri ise “high dimensional data” (çok ve yüksek boyutlu verilerin sınıflandırılması)’dır³⁷. Kelime torbası yaklaşımında içeriği en ayırıcı biçimde ve en fazla sıklıkta temsil eden anahtar kelimeler seçilerek dilbilgisi kuralları, cümle yapısı v.b gibi anlamı etkileyen faktörler ihmal edilir. Cümlelerdeki kelimeler vektörleştirilerek makine öğrenimi modellerine verilir ve anlam açısından etiketlenmiş verilerle desteklenerek bahsi geçen anahtar kelimeler tespit edilmeye çalışılır.

PANAS, psikometri metodolojisi itibarıyla kelime torbası yöntemine birebir uygundur. Ek olarak makine öğrenimi modellerinden farklı olarak hangi kelimenin içeriği en iyi temsil ettiğinin tespit edilmesine gerek yoktur. Çünkü hedeflenen duygu durumlarına ilişkin kelimeler zaten psikometri çalışmaları için tespit edilmiştir. Bu

³⁵ Daha önce belirtilen faktör yükleri ve duygu boyutları.

³⁶ Bag of words, Lexicon/sözlük bazlı yaklaşımdan farklı olarak daha geniş çerçevede bir tasnife işaret etmektedir.

³⁷ Pedro Domingos, Michael Pazzani, “On the Optimality of the Simple Bayesian Classifier under Zero-One Loss” Kluwer Academic Publishers. Manufactured in The Netherlands, 1997, Machine Learning 29, 103–130

kelimelerin geçtiği içeriklerin belirlenen duygu durumlarını temsil ettiği varsayılacaktır.

3.2.2 Türkçeleştirmede Kapsama Yaklaşımı

Türkçe dilindeki zorluk çeviriden ziyade dilin eklemeli yapısı gereği tek bir formdaki anahtar kelimenin duygu durumuna ilişkin temsiline düşük olmasıdır. İngilizce dilinde ise kelimelerin tek formda temsil gücü daha yüksektir.

PANAS'taki herhangi bir betimleyici sıfat ile basit zamanları kapsayacak şekilde bir ifade elde edilmek istenseydi, İngilizce dilinde sıfatın kendisi tüm basit zamanları ve çekimlerini kapsamak için yeterli olacaktı.

Örneğin, “I am afraid”, “I was afraid”, “I will be afraid” ifadelerinin tamamı yalın olarak “afraid” sıfatını içerir ve bu sıfatın bireysel ifade ile beraber tespit edilmesi veri temizleme aşaması için yeterli olacaktır. Diğer taraftan aynı betimleyici sıfatın Türkçe'ye çevrilmiş hali “**korkmuş**” olacaktır. Sıfat, farklı cümle kalıpları ile hem tek başına duygu durumunu temsil eder hem de basit zamanları kapsamaz. Örneğin “korkmuş bir durumdayım” net bir korku ifadesidir. Korkuyorum, korkacağım, korktum gibi korkmak fiilinden türetilen birinci tekil kişi ifadelerinin tamamı da korku ifade eder. Türkçenin bu eklemeli dil yapısı ve zengin morfolojisi gereği betimleyici sıfat tek başına korku ifadelerini kapsamaz.

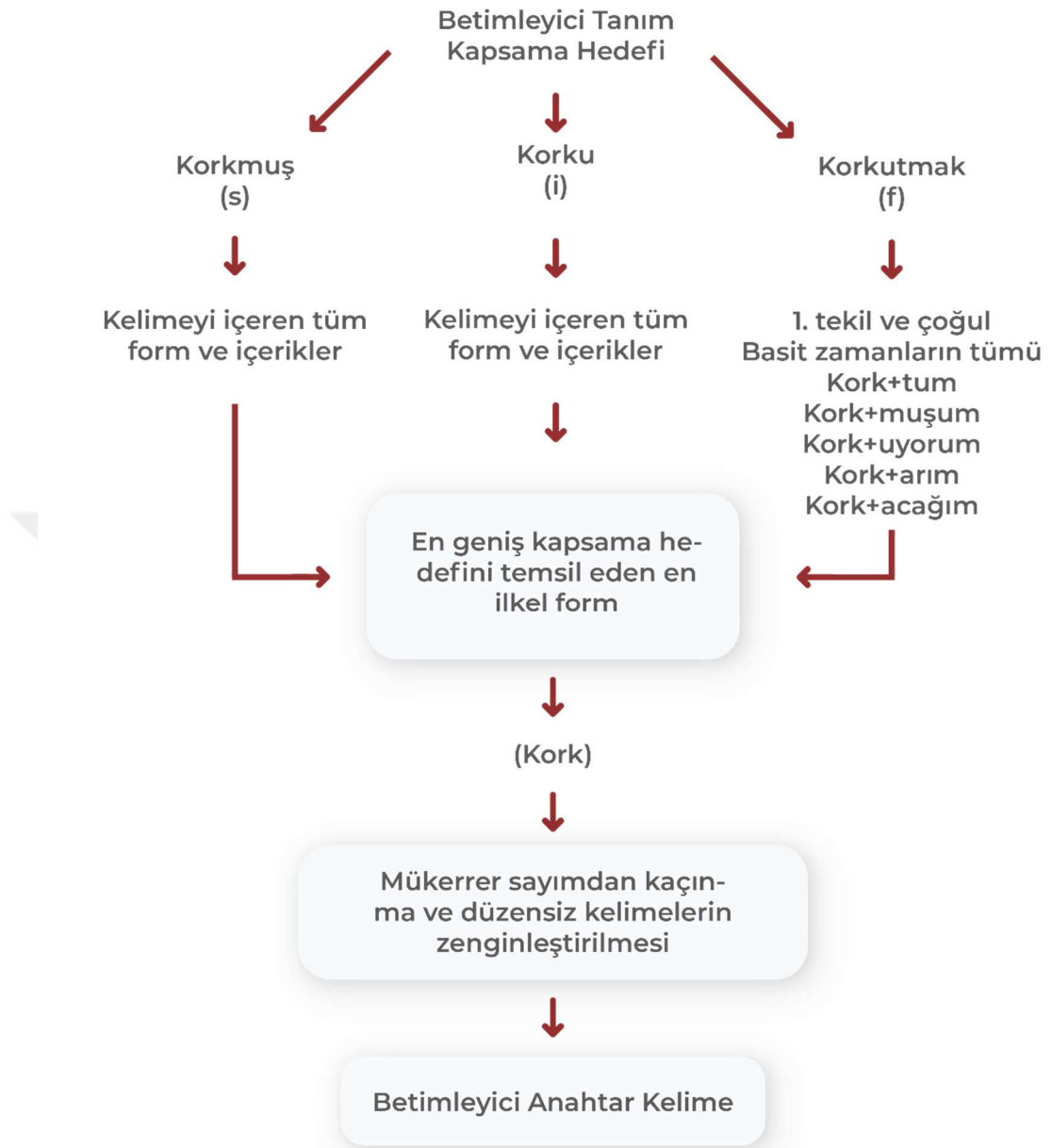
Bu sorunun giderilmesi için yukarıda bahsedilen metodolojilerin tamamı iç içe kullanıldı. Öncelikle PANAS'ın bilgisayar destekli doğal dil işleme uygulaması için “kelime torbası” yaklaşımı seçildi. Ardından Türkçe kapsamanın genişletilmesi içinde sıfatların en yakın isim ve fiil formları kök olarak, veri temizleme ve betimleyici içeriklerin tanımlanması aşamasına eklenmiştir.

3.3. Betimleyici Anahtar Kelimelerin Tanımlanması

Betimleyici içeriklerin tespit edilmesi ve saydırılarak normal değerlerin elde edilmesi için optimize edilmesi gereken iki yönlü bir problem bulunmaktadır. Bu sorun kapsamanın artırılması ile önemsiz içeriklerin analize alınması arasındadır. Kök bulma uygulanırken eğer betimleyici sığatı ile ilgili kapsama geniş tutulursa, gerçekten duygu durumu ifade etmeyen etkisiz içerikler analize dahil olabilir. Örneğin **ürkmüş**

sıfatının kökü ve fiil hali olan **ürk** + **mek** sözcüğündeki mastarsız ilkel halin, **içerme** `contains()` formatında aratılması durumunda **Türk** gibi yoğun kullanılan sözcükler analize girecektir. Eğer bununla başlayan `startswith()` mantıksal operatörü ile kullanılırsa etkisiz içerikleri kısmen dışarıda bırakan makul bir kapsama elde edilebilir. Örnekte görüldüğü gibi her kelimenin yapısı itibariyle, kapsama – etkin olmayan içerik dengesi açısından ayrı ayrı ele alınması ve optimum betimleyici anahtar kelimelerin belli bir metot uygulanarak her kelime için ayrı ayrı seçilmesi gerekmektedir.





Görsel – 4 Betimleyici Sıfatların Anahtar Kelimelere Dönüştürülmesi

Görseldeki metodoloji Türkçe dilinde, PANAS’ın betimleyici sıfatlarının içeriklerde tespit edilmesi için sıfatları, betimleyici anahtar kelimeye dönüştüren süreci temsil etmektedir. Buna göre PANAS’ın betimleyici sıfatları en yakın isim ve fiil halinin duygu ifade eden formlarını kapsayacak şekilde genişletilmiştir.

Bu yöntemin tercih edilmesindeki neden Türkçe ve İngilizce dilinin morfolojik yapısı arasındaki farklardır ve daha önce bahsedilen Türkçe dili için kapsama sorunudur. Bu iddia büyük rassal veri seti içinde test edilerek verilerle desteklenmiştir. İlk sayımlarda PANAS sıfatlarının tümü sadece sıfat olarak ve Ç.Dürü'nün çevirisi baz alınarak saydırıldığında aşağıdaki sonuçlar elde edilmiştir.

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Hevesli	1.452	Ürkmüş	102
İlgili	4.271	Korkmuş	1.310
Kararlı	6.347	Mutsuz	42.696
Heyecanlı	15.491	Sıkıntılı	5.133
İlhamlı	0	Asabi	5.147
Uyanık	3.264	Sinirli	21.229
Aktif	9.403	Utanmış	583
Güçlü	49.575	Suçlu	20.869
Gururlu	7.021	Tedirgin	10.786
Dikkatli	712	Düşmanca	166

Tablo – 3 Büyük Veri Kümesi İçindeki Betimleyici Sıfatlar Contains()

PANAS'ın betimleyici sıfatları sırayla, sıfatı herhangi bir formda içerecek şekilde (contains) büyük veri kümesi içerisinde aratılmıştır. Toplamda 308.139.050 adet içerikte sadece 205.557 adet betimleyici sıfat tespit edilmiştir. Bu rakam toplam içerik sayısının %0.000667'sine tekabül etmiştir ve elde edilen rakam İngilizce dilinde yapılan çalışmaların çok altında kalmıştır. Goncalves, Benevenuto ve Cha'nın çalışmasında PANAS'ın genişletilmiş versiyonu olan PANAS-X'in betimleyici sıfatlarının tespit edildiği içeriğin, toplam içeriğe oranı %27 oranında bulunmuştur ve bu çalışmada hem mükerrer sayım yapılmamış (bir cümlede geçen iki sıfattan sadece ilki alınmıştır) hem de veriler zamir içeriklerine göre filtrelenmiştir. Yukarıdaki tabloda elde edilen rakamlarda ise mükerrer sayım yapılmıştır ve içerikler herhangi ek bir filtre ile filtrelenmemiştir.

PANAS-X'teki sıfat sayısı 60 adettir ve bu sıfatlar, PANAS'taki sıfatların %70'ini içermektedir. Kapsama sorunundan emin olmak için büyük veri kümesi ayrıca PANAS-X'teki betimleyici sıfatların tamamı içinde saydırılmıştır ve bu yöntemle göre yapılan sayımın dillerdeki morfoloji yapısı farklılıkları nedeniyle tahmin edilen kapsama problemini teyit ettiği görülmüştür Bu durum mevcut filtreleme yöntemi üzerine daha fazla çalışmayı gereksiz kılmıştır. Çalışmanın bu kısmına ilişkin diğer veriler paylaşılmayacak sadece gelecekteki araştırmacılara yol göstermesi açısından bilgi notu olarak verilecektir.

Betimleyici sıfatlar yine tam haliyle ve tek bir formda latinize edilerekte saydırılmıştır. Örneğin : “ürkmüş” sıfatı, “urkmus” olarak aratılmıştır. Yazım hataları ya da Türkçe ifadelerin sıklıkla salt latin karakterlerle yazılması gibi ek duygu durumu ifadeleri de kapsanmaya çalışılmıştır. Netice itibariyle bu yöntemle de betimletici sıfatların toplam içeriklerde görülme sıklığı %1.5’i aşmamıştır.

Türkçe kapsama sorununun çözülmesi için görsel – 4’te belirtilmiş olan prosedür uygulamaya alınmıştır. Öncelikle PANAS’ın betimleyici sıfatlarının en yakın isim ve fiil halleri tespit edilmiştir. Örneğin korkmuş sıfatı için korku isim hali ve korkmak fiili seçilmiştir.

İçeriği en iyi şekilde temsil eden hedef betimleyici anahtar kelimelerin:

- Sıfatın kendisi ve birinci tekil ve çoğul şahıs olarak ifadeleri

Örneğin :

Korkmuş bir haldeyim.

Ne çok **korkmuşum**.

- Yalın olarak isim hali ve bunlarla oluşan düzensiz cümle yapılarını

Örneğin :

Korku dolu anlar yaşadım.

- Fiil halinin basit zamanları, bunların birinci tekil ve çoğul hallerini kapsayacak şekilde mastarsız fiil köklerinin

Örneğin :

Gerçekten çok **korktum**

Korkuyoruz zannediyorlar

En optimum şekilde betimleyici sıfatları temsil ettiği gözlemlenmiştir. Bu yöntemde tek kelime olan sıfatların tamamı Java’nın startwith(), bununla başlayan mantıksal operatörü/fonksiyonu ile taranmıştır. Böylece sadece betimleyici anahtar kelime olarak seçilen ifadeleri içeren içerikler sayıma alınmıştır. Seçilen betimleyici anahtar kelimeyi herhangi bir kelime içerse bile eğer kelime verilen ifade ile başlamıyorsa analiz dışında kalmıştır

Örneğin :

Gabon'un **Akork** şehrinde hava durumu. → Bu ifade analize alınmaz.

Korkmaktan ne yapacağımı şaşırdım. → Bu ifade analize alınır.

İkinci aşamada ise aynı sıfat için mükerrer sayımdan kaçınılması gerekmektedir. Bu çalışmadaki yaklaşımda bir içeriğin aynı anda iki duygu durumunu birden ifade edebileceği varsayılmıştır ve aynı cümlede farklı duygu durumları için birden fazla betimleyici sıfat tespit edildiğinde, bu sıfatlar her bir sıfat grubu için mükerrer olarak sayılmıştır. Diğer taraftan aynı duygu durumunu ifade eden aynı betimleyici sıfatın, betimleyici anahtar kelimeler olarak farklı formları mükerrer sayılmamalıdır.

Örneğin:

Korkmuş bir haldeyim.

Bu cümledeki “kork” kökü, korkmuş betimleyici sıfatı için seçilmiş üç betimleyici anahtar kelimedenden üçünü de aynı anda içerir. İlgili sıfat için betimleyici anahtar kelimeler sırasıyla :

- Korkmuş
- Kork
- Korku

Şeklinde. Bu betimleyici anahtar kelimeler ayrı ayrı startwith() fonksiyonu ile saydırıldığında hem **kork** hem de **korkmuş** ve **korku** anahtar kelimeleri ile tespit edilecek ve saydırılacaktır ve bu şekilde aynı sıfat için mükerrer sayım olacaktır.

Bu durumdan kaçınmak için ikinci aşamada betimleyici sıfatı tespit ederken, betimleyici anahtar kelimelerin “*kapsama hedefini temsil eden en düşük formu*” tercih edilmiştir. Görsel – 3 ‘te belirtilen bu aşamada amaç en kapsayıcı kelimeyi tespit edip bir kere saydırarak mükerrer sayımdan kaçınmaktır.

Aynı örnek üzerinden devam edilirse :

Kork fiilinin mastarsız hali betimleyici anahtar kelimelerden hem **korkmuş** sıfatını hem de **korku** ismini kapsar. Bu durumda betimleyici anahtar kelime olarak sadece **kork** seçilirse ve içerikler bu şekilde taratılırsa; hem sıfat, hem isim hem de kapsanması istenen fiilin birinci tekil ve çoğul basit zamanları tespit edilmiş ve mükerrer olarak saydırılmamış olacaktır.

Görsel – 4'teki son aşamada ise mükerrer sayımdan kaçınma ve düzensiz kelimelerin zenginleştirilmesi aşamasıdır. Burada en kapsayıcı ve en ilkel form tespit edildikten sonra kapsama hedefinde olan ancak ilkel formla kapsamaya alınamayan ve/veya çok fazla etkisiz içeriği de analize alabilecek betimleyici anahtar kelimeler düzenlenir.

Örneğin, utanmış betimleyici sıfatı için kapsama hedefi; Utanmış sıfatı, utanç ismi ve utan + mak fiilinin birinci tekil ve çoğul basit zamanlarıdır. Burada utan + mak fiilinin mastarsız hali olan “utan” sözcüğü kapsama hedefindeki her şeyi kapsasa da sözcüğün kısa ve dört harfli olması startswith() fonksiyonuyla bile analizi bozan içerikleri kapsayabilir. Örneğin **utan**agan isimli bir köpek ırkı vardır, **utan**irem ile başlayan ve Twitter algoritması tarafından çok fazla Türkçe algılanan, Azerice içerikler bulunmaktadır vb. Bunlardan kaçınılması için utan gibi özel durumu olan sözcüklerde kapsama için zenginleştirme işlemi yapılır. Buna göre kök kelime alınmaz onun yerine doğrudan ifadeleri içeren anahtar kelimeler tek tek aranır.

Utanmış, utanç, utandım, (*utanmışım, utanmış sıfatı bunu kapsadığı için analize alınmaz*), utanırım, utanıyorum, utanacağım sözcükleri içeriklerde ayrı ayrı aratılır.

Bu kelimelerin sayısı çok fazla olmasa da betimleyici anahtar kelimelerin seçiminde bahsedilen sebeplerden farklı bir yaklaşımla ele alınması gerekmiştir. Her sıfatın karşılığı olan betimleyici anahtar kelimelerin tam listesi eklerde bulunmaktadır.³⁸

3.3.1 Betimleyici Sıfatlarının Büyük Veri Kümesinde Analizi

Betimleyici anahtar kelimelerin tespit edilmesinin ardından. Bu kelimeler iki ayrı veri seti içinde taratılarak, hangi içeriklerde geçtiği bulunacaktır ve betimleyici anahtar kelimeleri içeren ilgili içeriklerin, betimleyici sıfatları temsil ettiği varsayılacaktır. Böylece ilgili içerik betimleyici sıfatın temsil ettiği duygu kategorisinde kabul edilecektir.

Veri setlerinden ilki 308.139.050 tweeti içeren ve duygu durumları için normal dağılım değerlerini tespit edeceğimiz rassal veri kümesidir. Diğer veri seti ise farklı temalarda

³⁸ Ek – 4 PANAS Betimleyici anahtar kelimeler.

Türkiye’deki bazı gündemleri içeren veri setleridir. Gündemlere ilişkin veri setlerine ilgili başlıkta detaylı olarak değinilecektir.

Ek-4’te belirtilen betimleyici anahtar kelimelerle yapılan tarama sonucunda pozitif betimleyici sıfatlar için betimleyici anahtar kelimelerin büyük veri kümesi içerisindeki frekansları aşağıdaki şekilde tespit edilmiştir:

Hevesli	İlgili	Kararlı	Heyecanlı	İlhamlı
Heves : 596.715	İlgili : 27.717	Kararlı : 144.851	Heyecan: 994.703	İlham : 93.708
İlgilen: 738.518				
Toplam : 596.715	Toplam : 766.235	Toplam : 144.851	Toplam : 994.703	Toplam : 93.708
Uyanık	Aktif	Güçlü	Gururlu	Dikkatli
Uyanı : 554.118	Aktif : 340.459	Güçlü : 1.270.759	Gurur : 911.933	Dikkat : 1.312.346
Uyana : 164.832				
Toplam : 718.950	Toplam : 340.459	Toplam:1.270.759	Toplam : 911.933	Toplam :1.312.346

Tablo – 4 Büyük Veri Kümesindeki Pozitif Sıfatlı İçerikler

Büyük veri seti içindeki bütün içerikler içerisinde pozitif duygu durumu ifade eden içeriklerin toplamda 7.150.659 adet olduğu tespit edilmiştir.

Negatif duygu durumu ifade eden ve negatif betimleyici anahtar kelimelerle tespit edilen içeriklerin sayısı şu şekildedir:

Ürkmüş	Korkmuş	Mutsuz	Sıkıntılı	Asabi
Ürkmüş : 1.421	Kork : 3.802.897	Mutsuz : 590.717	Sıkıl : 898.247	Asabi : 25.650
Ürkme : 6.689			Sıkıntı : 1.255.934	
Ürktüm : 1.682				
Ürkmüşüm : 97				
Ürkerim : 333				
Ürküyorum : 2.277				
Ürkeceğim : 4				
Toplam :12.503	Toplam:3.802.897	Toplam:590.717	Toplam : 2.154.181	Toplam : 25.650
Sinirli	Utanmış	Suçlu	Tedirgin	Düşmanca
Sinir : 554.118	Utanmış : 8.862	Suçlu : 546.262	Tedirgin : 104.245	Düşman : 1.631.359
	Utañ : 156.629			
	Utandım : 61.299			
	Utanmışım : 187			
	Utandırım : 12.852			
	Utaniyorum :			
	132.609			
	Utanaçağım :			
	3.153			
Toplam : 554.118	Toplam : 375.591	Toplam : 546.262	Toplam : 104.245	Toplam : 1.631.359

Tablo – 5 Büyük Veri Kümesindeki Negatif Sıfatlı İçerikler

Negatif duygu durumu ifade eden içerikler toplamda 9.797.523 olarak tespit edilmiştir. Bütün büyük veri kümesi içerisinde toplamda 16.948.182 adet içeriğin PANAS duygu durumlarını temsil ettiği görülmektedir. Bu rakam toplam içeriğin %5,50'sine tekabül etmektedir. Elde edilen oran İngilizce dilinde yapılmış olan çalışmaya yakın bir oranı temsil etmektedir.

Elde edilen referans veriler rastgele bir veri kümesi içerisinde negatif ve pozitif duygu durumlarının ne yoğunlukta olduğunu gösteren referans bir değerdir. Bu değerlere dayanarak PANAS-T şablonuna göre Twitter paylaşımlarının %57,8087'sinin negatif duygu durumunu, %42,1913'ünün pozitif duygu durumunu ifade ettiği sonucuna varılmıştır.

4.0 Analiz Konusu Gündemler

Normal değerlerin elde edilmesinin ardından çalışmanın temel amacı olan Twitter gündemlerinin analiz edilebilmesi için aynı betimleyici anahtar kelimelerle, seçilen gündemlerde tarama yapılarak, gündemlerin duygu durumlarına ilişkin dağılım tespit edilmelidir. Bunun ardından da büyük ve rassal veri kümesindeki duygu durumu dağılımı ile gündemlerin duygu durumu kıyaslanarak, gündemlerdeki hakim duygu durumu bulunabilecektir.

Önceki bölümde pozitif ve negatif duygu dağılımları sırasıyla %42,1913 ve %57,8087 olarak tespit edilmiştir. Bu referans değer olmadan gündemlerdeki hakim duygu durumunu hesaplamak mümkün değildir. Örneğin gündemde pozitif ve negatif duygu durumları sırasıyla %50 - %50 olarak tespit edilseydi referans değer olmadan gündemin nötr olduğu söylenebilirdi ancak referans değerlere göre Twitter kullanıcıları bir duygu ifadesinde bulunurken negatif duygu ifade etmeyi daha çok tercih etmişlerdir, bu durum aslında %50 - %50 gibi bir dağılımdaki gündemin pozitif duygu durumuna daha eğilimli olduğu anlamına gelecektir. Yüzde dağılımı sadece örnek olarak verilmiştir, PANAS-T'nin duygu durumunu hesaplamak için kendi metodolojisi bulunmaktadır ve yüzdeler bu hesaplamada doğrudan kullanılmayacaktır.

4.1 Gündemlerin Seçimi ve Tercih Nedenleri

Seçilen gündemler Twitter'da 2021-2022 yılları arasında ve zaman zaman trend topic kategorisine girmiş farklı alanlardaki on kadar gündemden oluşmaktadır. Bu

Gündemlerin tercihinde üç temel kıstas esas alınmıştır. Birincisi gündemlerin olabildiğince polarize olması ve hakkındaki duygu durumlarının apriori olarak tahminlenebilmesidir. Örneğin “orman yangınları” gibi bir gündemdeki hakim duygu durumunun negatif olması beklenir. Bunun sebebi hali hazırda toplanan anket verileri ile ölçülmeye çalışılan PANAS-T ölçeğinin istenen sonuçları verip vermeyeceğini daha net tespit etmek ve anket verileri için olası örneklem tartışmalarını olabildiğince

ortadan kaldırmaktır. Eđer g ndemler daha muęlak, sonularının apriori olarak tahminlenmesi daha g  olsaydı, ortalamaya yakın PANAS-T skorlarına sahip olan g ndemlerin  rneklemi for  in daha detaylı inceleme gerekebilirdi.

 kinci temel kıstas ise seilen g ndemlerde olabildięince fazla ierięin olmasıdır. B ylece alıřmada kullanılacak veri miktarını maksimize ederek daha isabetli sonulara ulařılmaya alıřılmıřtır. Her bir g ndemde seilen zaman periyotları  in 10.000'in altında ierik bulunan g ndemler elenmiřtir. Bahsedilen dięer parametrelerde dikkate alınarak en fazla ierięin bulunduęu g ndemler tercih edilmiřtir.

Son olarak seilen g ndemler olabildięince farklı alanlardan seilerek PANAS'ın farklı konu bařlıkları  in test edilmesi hedeflenmiřtir. evre, siyaset, spor, saęlık vb. bařlıklarla g ndemler eřitlendirilmiřtir.

Gündemlerle ilgili verilen zaman aralıklarında tespit edilen toplam tweet sayısı şu şekildedir;

Gündem	Toplam Tweet Sayısı	Tarih Aralığı	Gündem	Toplam Tweet Sayısı	Tarih Aralığı
Suriyeliler	34.935	Max.	Ukrayna Rusya	106.604	24 Şubat 2022 – 15 Mayıs 2022
Recep Tayyip Erdoğan	495.259	Max.	Corona	34.023	Max
Orman Yangınları	15.892	1 Mayıs 2021 – 1 Ekim 2021	Turkovac	18.402	Max
Enflasyon	165.563	Max.	Dolar tl	124.798	Max
Ramazan	116.232	1 Mart 2022 – 15 Mayıs 2022	Şampiyon Beşiktaş	28.982	1 Mayıs 2021 – 1 Ağustos 2021

Tablo – 6 Gündemlerdeki Toplam Tweet Sayıları

Gündemlere ilişkin veriler toplanırken, gündemlerin kendi bağlamları açısından önemli olduğu tarihlerdeki verilerin çekilmesi tercih edilmiştir. Örneğin Ramazan Ayı 2 Nisan 2022 tarihinde başlamıştır ve 1 Mayıs 2022 tarihinde sonlanmıştır. Ramazan anahtar kelimesiyle tweetler aratılıp veri tabanına aktarılırken bu tarihlere yakın tarihlerde arama yapılmıştır. Diğer gündemlerde de aynı yol izlenmiştir.

Gündemlere ilişkin içeriklerin aratılmasında yine ara katman olarak Twitter’ın kendi arama arayüzü ve bu kısımda işleyen algoritması kullanılmıştır. İçerikler bu arama arayüzüyle Türkçe olarak filtrelenmiştir. Url ayrıştırma yöntemi ile çekilen verilerde

kısmen kayıplar olsa da geçmişe yönelik veri alabilmek sadece bu yöntemle mümkündür. Twitter API'leri güncel tarihten maksimum 15 gün öncesine kadar verilerin çekilmesine izin vermektedir. Url ayrıştırma veri kayıpları analize giren veriyi etkilemez çünkü kaybedilen veriler kısmen değil tamamen kaybedilmektedir. Örneğin ramazan anahtar kelimesi ile belirtilen tarih aralığında 116.232'den bir miktar daha fazla tweet olabilir ancak elde edilen tweetlerin hepsi ya tamdır ya da kaybedilen tweetler hiçbir şekilde analize girmemiştir. Bu da gündemlere ilişkin elde edilen veri setinin temiz ve bağlama uygun olmasını sağlamaktadır.

4.2 Gündemlerdeki Hakim Duygu Durumlarının Tespiti ve Hesaplama

Büyük rassal veri kümesi içerisinde saydırılan aynı betimleyici anahtar kelimeler gündemlerde de saydırılarak, gündemlerin hakim duygu durumları aşağıdaki gibi tespit edilmiştir.

Suriyeliler Gündemi :

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Hevesli	12	Ürkmüş	2
İlgili	10	Korkmuş	403
Kararlı	13	Mutsuz	15
Heyecanlı	3	Sıkıntılı	333
İlhamlı	1	Asabi	2
Uyanık	67	Sinirli	120
Aktif	1	Utanmış	31
Güçlü	44	Suçlu	100
Gururlu	57	Tedirgin	15
Dikkatli	213	Düşmanca	349
Toplam	421	Toplam	1.370

Tablo – 7 Suriyeliler Gündemindeki Betimleyici Sıfatlar

Suriyeliler gündeminde görüldüğü gibi hakim duygu durumu yüksek oranda negatiftir ancak PANAS'ta gündemlerin birbirleriyle mukayese edilebilmesi için gündemlerin hakim duygu durumuna ilişkin skorlar Goncalves, Benevenuto ve Cha'nın PANAS-T (Twitter için PANAS) metninde ortaya attığı hesaplama yöntemi kullanılacaktır. Böylece farklı dildeki çalışmalarla da duygu durumları, kültür ve dil arasındaki farklılıklarda mukayese edilebilir hale gelecektir. Bu çalışmada referans verilen

hesaplama yöntemi her bir duygu boyutu için adım adım aşağıdaki şekilde kullanılmıştır :³⁹

1-) Öncelikle duygu durumu içeren filtrelenmiş içeriklerin toplamı tespit edilir. Daha önce değinilen filtreleme yöntemleri baz alınarak ilgili çalışmada bu rakam $T = 479.356.536$ 'dır filtrelenmiş içeriklerin toplamı "T" ile sembolize edilecektir.

2-) Bunun ardından hesaplanacak duygu durumuna ilişkin toplam içerik sayısı tespit edilecektir. Büyük veri kümesi içerisindeki bu toplam $T_{(ilgili\ duygu\ durumu)}$ şeklinde gösterilmiştir. Örneğin $T_{(Pozitif)}$ büyük veri kümesi içerisindeki toplam pozitif duygu durumu içeren içeriklerin toplamıdır.

3-) Bu iki değer mukayese edilerek (baseline relative occurrence) büyük veri setindeki göreceli frekans elde edilir. Toplam veri setindeki göreceli frekans α işareti ile sembolleştirilmiştir ve şu şekilde hesaplanmıştır : $\alpha(s) = \frac{|Ts|}{|T|}$

4-) Toplam veri seti içerisindeki göreceli frekans bulunduğundan sonra aynı işlem, aynı mantıkla analiz edilen gündem içinde yapılarak (event relative occurrence) gündemdeki duygu durumunun göreceli frekansı bulunur. Bu değer β sembolü ile tanımlanmıştır ve şu formülle hesaplanmaktadır $\beta(s) = \frac{|Ss|}{|S|}$ S ilgili gündemdeki duygu durumu ifade eden içeriklerin tümüdür. Ss ise gündemde hesaplanacak olan ilgili duygu durumunu ifade eden içeriklerin sayısıdır. Örneğin $S_{(pozitif)}$

5-) α ve β değerleri bulunduğundan sonra aşağıdaki formül uygulanarak PANAS-T duygu durumuna ilişkin skor hesaplanır. Skor P değeri ile temsil edilmiştir ve ilgili duygu durumları için fonksiyon olarak değer almaktadır. Örneğin $P_{(pozitif)}$

$$-P(pozitif) = \frac{\alpha(s) - \beta(s)}{\alpha(s)} \text{ Eğer } \alpha(s) > \beta(s) \text{ ise}$$

$$P(pozitif) = \frac{\beta(s) - \alpha(s)}{\beta(s)} \text{ Eğer } \alpha(s) < \beta(s) \text{ ise}$$

PANAS-T orijinal metninde formüldeki negatif ve pozitif değerler tam tersi olarak kullanılmıştır ancak Türkçe'deki hesaplama alışkanlıkları genellikle negatif değeri

³⁹ Pollyanna Goncalves, Fabricio Benevenuto, Meeyoung Cha "PANAS-t: A Psychometric Scale for Measuring Sentiments on Twitter" Systems, man, and cybernetics, PART C: Applications and Reviews, VOL. XX, NO. XX, JUNE 2013

uzak, pozitif değeri yakın kabul ettiği için pozitif ve negatif değerler orijinal formülün tam tersi şeklinde kullanılacaktır.

Bu hesaplama adımları Suriyeliler gündemine uygulanırsa, gündem için aşağıdaki değerler elde edilmektedir:

$$\alpha(\text{pozitif}) = \frac{|7.150.659|}{|308.139.050|} \quad \beta(\text{pozitif}) = \frac{|421|}{|34.935|}$$
$$\alpha(\text{pozitif}) = 0,0232 \quad \beta(\text{pozitif}) = 0,0120$$

Pozitif duygu durumu için alfa ve beta değerlerinin bulunmasının ardından PANAS duygu durumu için ilgili değer hesaplanabilir.

$$-P(\text{pozitif}) = \frac{0,0232 - 0,0120}{0,0232} \quad P(\text{pozitif}) = -0,4827$$

Formül her duygu durumu için -1 ile 1 arasında bir değer üretir ve değerleri farklı gündemler ve duygu boyutları için kıyaslanabilir hale getirir. Yukarıda elde edilen sonuca göre Suriyeliler gündemi pozitif duygu durumu için -0,4827 değerini almıştır buna göre gündem pozitif duygulardan uzaktır. Pozitif duygu durumu sıfır değerini almış olsaydı gündemdeki pozitif duygulara ilişkin durum nötr kabul edilecekti veya gündem pozitif bir değer almış olsaydı, pozitif duygu durumunun baskın olduğu söylenebilecekti.

Bu yöntemdeki en büyük avantajlardan biri duygu durumlarına ilişkin üretilen skorun göreceli olması ve böylece alt sistemlerde bir hata yapılması durumunda hatanın büyük ölçüde tolere edilebilmesidir. Örneğin betimleyici anahtar kelimelerin iyi seçilmediği bir durumda, hem toplam veri kümesi hem de gündemdeki sayım aynı anda değiştiği için göreceli değerlerin değişmesi beklenmez.

İkinci en önemli özellik duygu boyutlarını birbirinden ayrı ele alması ve tamamlayıcı olarak değerlendirmemesidir. Örneğin bir gündemde pozitif duygu durumunun hakim olmaması, negatif duygu durumunun hakim olduğu anlamına gelmez. PANAS'ta negatif ve pozitif duygu durumları birbirinden bağımsız durumları temsil eder ve duygu durumları arasındaki korelasyon katsayısı sırasıyla -0,2 ve -0,3 olarak tespit

edilmiştir⁴⁰. Farklı arařtırmalarda, birden fazla arařtırmacı pozitif ve negatif duygu durumlarının birbirlerinin kutupsal karřıtları olmadığını aıklamıřtır. ⁴¹Her iki duygu boyutu da dřk deęer alabilir veya farklı kompozisyonlarla yksek deęer alabilir.

Aynı gndem iin negatif duygu durumunu da ařaęıdaki řekilde hesaplanmıřtır;

$$\alpha(\text{negatif}) = 0,03179 \quad \beta(\text{negatif}) = 0,03921 \quad \text{ve} \quad P(\text{negatif}) = \frac{0,03921 - 0,03179}{0,03921}$$

$$P(\text{negatif}) = 0,1892$$

Bu deęerlere gre Suriyeliler gndemine iliřkin negatif duygu durumunun gndemde bulunduęu sylenbilir. Dięer taraftan pozitif duygu ifadelerinden kaınma, negatif duygu durumu ifadelerinden daha yksektir.

Btn gndemlere iliřkin saydırılan betimleyici anahtar kelimeler ve bunlarla tespit edilen ierikler ektedir.⁴²

⁴⁰ goston Gyollai, Pter Simor, Ferenc Kteles, Zsolt Demetrovics ; “Psychometric properties of the Hungarian version of the original and the short form of the Positive and Negative Affect Schedule (PANAS)”, Neuropsychopharmacologia Hungarica 2011. XIII. vf. 2. Szm, 74

⁴¹ Diener, E., & Emmons, R. A. . “The independence of positive and negative affect. Journal of Personality and Social Psychology”, (1984), 47(5), 1105–1117

Tüm gündemlerde betimleyici sıfatları içeren içeriklerin toplamı :

Gündem	Pozitif Sıfatlar	Negatif Sıfatlar	Toplam
<i>Suriyeliler</i>	421	1.370	1.791
<i>Recep Tayyip Erdoğan</i>	12.846	7.446	20.292
<i>Orman Yangınları</i>	382	386	768
<i>Enflasyon</i>	3.001	4.248	7.249
<i>Ukrayna Rusya</i>	2.464	2.732	5.196
<i>Corona</i>	600	1.083	1.683
<i>Turkovic</i>	811	160	971
<i>Dolar TL</i>	1.812	2.076	3.888
<i>Şampiyon Beşiktaş</i>	661	438	1.099
<i>Ramazan</i>	2.055	1.455	3.510

Tablo – 8 Gündemlerdeki İçeriklerin Duygu Boyutuna Göre Tasnifi

Ekteki veri setleri baz alınarak tüm gündemlere ilişkin pozitif ve negatif duygu durumu PANAS skorları şu şekilde tespit edilmiştir.

Gündem	P _(pozitif)	P _(negatif)	SJ (Sübjektivite)
<i>Suriyeliler</i>	-0,4827	0,1892	0,3359
<i>Recep Tayyip Erdoğan</i>	0,1042	-0,5281	0,3161
<i>Orman Yangınları</i>	0,0348	-0,2359	0,1353
<i>Enflasyon</i>	-0,2198	-0,1947	0,2072
<i>Ukrayna Rusya</i>	0,0043	-0,1947	0,0995
<i>Corona</i>	-0,2413	0,0003	0,1208
<i>Turkovac</i>	0,4727	-0,7294	0,6010
<i>Dolar TL</i>	-0,3750	-0,4778	0,4264
<i>Şampiyon Beşiktaş</i>	0,0172	-0,5250	0,2711
<i>Ramazan</i>	-0,2413	-0,6067	0,4240

Tablo – 9 Toplama Yöntemine Göre PANAS Skorları

PANAS skorlarına ek olarak doğal dil işlemede sıklıkla kullanılan farklı bir yaklaşım içinde metrik üretilmiştir. Sübjektivite isimli bu metrik, metin işleyen pek çok uygulamada ya da doğal dil işleme kütüphanelerinde kullanılmaktadır. Temelde sübjektivite bir dil bilim terimidir, edebiyat teorisinden uyarlanmıştır ve öznel

yargıların ifadesini kasteder.⁴³Sübjektivite gündeme ilişkin duygusal polarizasyonun tespit edilmesini sağlar. Böylece bir gündemde duygu durumunun yoğunluğuna ilişkin bilgi verir. Bu çalışmadaki yaklaşımda sübjektivite skoru PANAS-T'nin duygu durum değerleri ile üretilmiştir. Yaklaşım bir duygu durumunda kaçınma olmasını da sübjektif ve duygu yoğun kabul eder. Örnekteki “Suriyeliler” gündeminde negatif duyguların ifade edildiği P(neg) skoru yaklaşık 0,18 olarak bulunmuştur. Diğer taraftan pozitif ifadelerden kaçınma eğilimi, negatif ifadelerin kullanılma eğiliminin skor olarak neredeyse iki katıdır -(0,48). Pozitif ifadelerdeki normal dağılımla uyuşmayan bu durum, “ifadelerin eksikliği” üzerinden gündemdeki duygu yoğunluğunu ele verir.

Bu durumda sübjektivite değerinin hesaplanması için PANAS-T'nin duygu durumlarına ilişkin tüm değerleri (pozitif ve negatif) kümülatif olarak ele alınmalıdır. Bu varsayımlara dayanarak PANAS-T için sübjektivite skoru aşağıdaki şekilde geliştirilmiştir:

$$n.SJ = \sum_{x=1}^n |P(X)|$$

SJ, ilgili gündemdeki toplam sübjektivite değeridir ve formüle göre toplamda 0 ile 1 arasında bir değer alır. N toplam duygu durumu sayısıdır ve bu rakam PANAS-T şablonu için 2'dir (pozitif ve negatif) ancak PANAS-X şablonu çok boyutlu bir duygu durum analizidir ve bu analizde toplam duygu durumu sayısı 11'dir. Sübjektivite değerinin farklı şablonlara adapte edilebilmesi için n değeri değişken olarak atanmıştır.

P(x) fonksiyonu sırasıyla duygu durumlarına ilişkin fonksiyondur ve ilgili gündemin ya da veri setinin PANAS duygu durumu değerleridir. Örneğin P(1) = P(pozitif) P(2) = P(negatif) veya tersidir. Netice itibarıyla PANAS duygu durum değerleri 1'den başlar ve her biri betimleyici anahtar kelimelerle elde edilen bir dizi basit fonksiyondur. İfade noksanlıklarının polarizasyon olduğu varsayımı altında P(x) değerlerinin tamamı mutlak değer olarak alınacaktır böylece ilgili duygu durumundaki pozitif ya da negatif sapmalar kümülatif olarak sübjektivite değerini arttıracaktır.

Yine Suriyeliler gündemi üzerinden sübjektivite değeri şu şekilde hesaplanabilir;

⁴³ Rada Mihalcea, “Multilingual sentiment and subjectivity analysis”, Academia (Department of Computer Science University of North Texas) 12 May 2011, 4

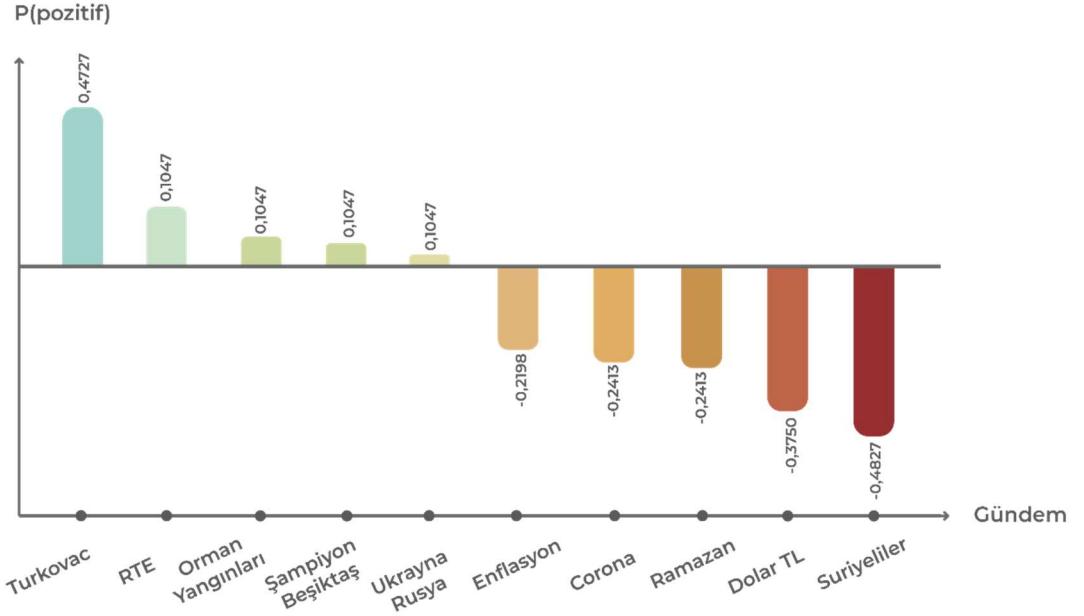
$$\begin{aligned}
2.SJ &= \sum_{x=1}^2 |P(1)| + |P(2)| \\
&= \sum_{x=1}^2 0,4827 + 0,1892 \quad 2SJ = 0,6719 \quad SJ = 0,3359
\end{aligned}$$

Netice itibariyle iki boyutlu duygu durumundan gündemlere ilişkin üç farklı veri üretilmiştir. Buradan elde edilen verilere göre en fazla pozitif duygu durumunu içeren gündem Turkovac gündemidir. Türkiye'nin ürettiği Covid aşısı olan Turkovac'ta negatif duyguların ifadesinden kaçınma da gündemler arasında en yüksek seviyededir. Türkiye için ortak değer kabul edilebilecek bu tip bir başarı hakkındaki duygu durumu, veri setinde beklentilere paralel olarak gerçekleşmiştir.

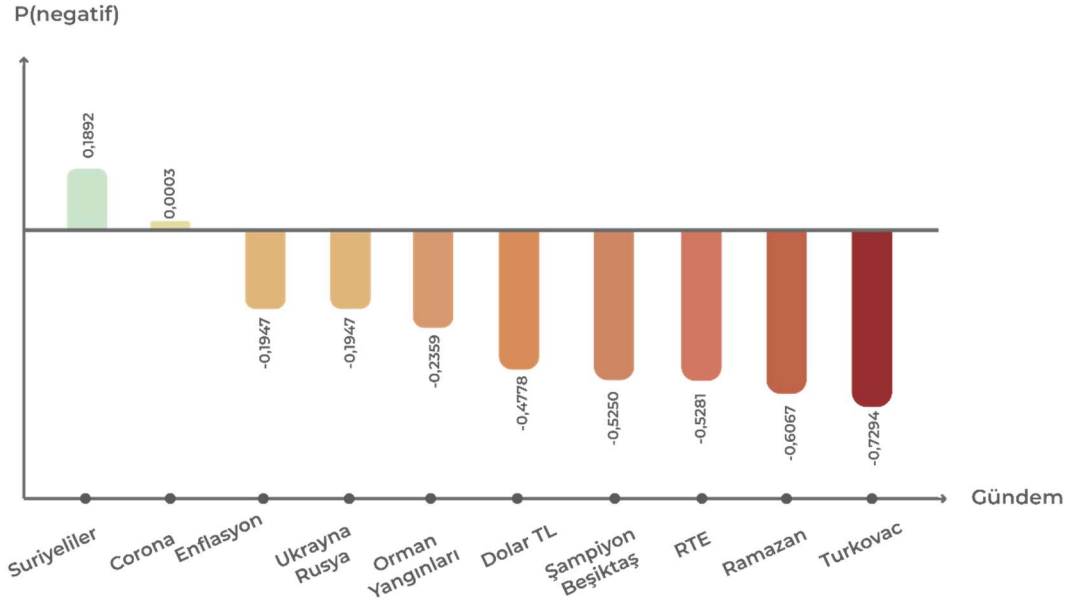
En fazla pozitif duygu durumu ifade etmekten kaçınmanın ve en yüksek negatif duygu durumu ifade etmenin Suriyeliler gündeminde olduğu görülmüştür. Bu da kamuoyundaki genel duruma bakıldığında makul bir sonuçtur.

Torkovac'tan sonra negatif duyguların ifadesinden en fazla kaçınılan gündemler sırasıyla ramazan, Recep Tayyip Erdoğan ve Şampiyon Beşiktaş gündemi olmuştur. Dolar TL ifadesinde duyguların toplam ifadesi düşükken sübjektivite oldukça yüksektir. En düşük sübjektivite yine beklendiği gibi Ukrayna – Rusya savaşı gündemi hakkındadır. Yurtdışı bir başlığın duygusal yoğunluğunun düşük olması da beklentilere paraleldir.

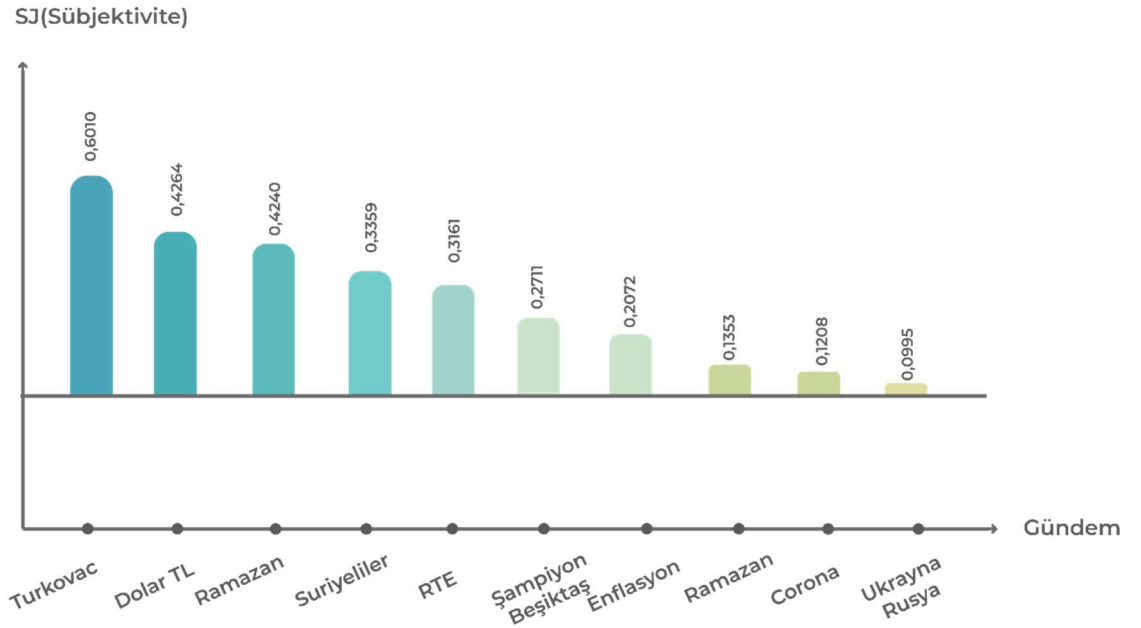
Veri setinin sonuçları incelendiğinde gündemlere ilişkin sezgisel, apriori varsayımlarla sonuçların uyumlu olduğu görülmektedir. Bu da PANAS-T yöntemi ile bilgisayar destekli analizin duygu durumuna ilişkin fikir verebileceğini gösterir niteliktedir. Ancak hipotezin testi için bulgular yeterli değildir.



Görsel – 5 Gündemlerin P(pozitif) Skorları



Görsel – 6 Gündemlerin P(negatif) Skorları



Görsel – 7 Gündemlerin Sübjektivite Analizi ve Skorları

4.3 İlk Bulgular Üzerinden Tespit Edilen Sorunlar

Elde edilen veriler ve PANAS-T skorları gündemlere ilişkin beklentilere paralel sonuçlar verse de duygu durumunun salt iki boyuta indirgenmesi betimleyici sıfatlardan elde edilen çıktı miktarını verimsiz hale getirmektedir. Buradaki en temel sorun sıfatların kendi duygu grubu için bazı sıfatların standart sapmasının çok yüksek olmasıdır. Örneğin büyük rassal veri kümesinde ürkmüş negatif sıfatı 12.503 adet tespit edilmişken korkmuş sıfatı 3.802.897 adet tespit edilmiştir. Bu iki sıfatın adet olarak negatif duygu durumunu aynı oranda temsil ettiği varsayımı ve toplanarak işleme alınması hesaplanan PANAS-T skorlarında daha yüksek frekanslarda görülen sıfatların skora daha fazla etki etmesine sebep olacaktır. Bu da skorun hesaplanmasında sağlıklı bir durum olarak görülmemektedir.

Diğer taraftan PANAS herhangi bir deneğe likert ölçeğinde verilirken de aynı durum söz konusu olduğu için sıfatların bir arada değerlendirilmesi, psikometrik yöntem açısından doğru bir yaklaşım olacaktır. Klasik PANAS Testinde verilen betimleyici sıfatların deneklerin anlam dünyasına göre duygu durumlarını hangi ölçüde temsil ettiğine ilişkin herhangi bir bilgi ya da çalışma yoktur. Örneğin, deneğin bir PANAS testinde 3 ayrı pozitif duygu durumu için yüksek skorlama yaptığını varsayalım bunlar; dikkatli, ilgili ve ilhamlı olsun diğer taraftan deneğin sadece tek bir negatif sıfat için yüksek skor verdiğini varsayalım bu sıfatta utanmış olsun. Deneğin duygu durumuna, utanmış sıfatı diğer üç pozitif sıfatın toplamından daha fazla etki ediyor olabilir ve pozitif PANAS skoru daha yüksek olmasına rağmen deneğin duygu durumu negatif olabilir. PANAS ile ilgili yapılan çalışmalarda, çoğunlukla duygu durumuna yakınlık arz eden psikolojik bozukluk daha güçlü regresyon vermiştir. Örneğin depresyon, negatif duygu durumuna daha yatkın bir psikolojik bozukluktur, PANAS'ın negatif duygu durumu ifade eden sıfatları ile depresyon arasındaki ilişki, PANAS'ın tüm duygu durumuna ilişkin genel ölçüm skorlarından daha güçlü sonuçlar vermiştir⁴⁴. Bu durumda ölçülmesi istenen şeye daha yakın olan sıfatların aslında bağlama göre duygu durumunu daha iyi temsil ettiği sonucuna varılabilir. Bu durumda sıfatları kendi içinde ve izole bir biçimde ölçerek hangi sıfatların gündemde daha baskın PANAS skoruna sahip olduğunu elde etmek ölçülmek istenen gündem açısından daha iyi fikir verecektir.

4.4 Betimleyici Sıfatların İzole Hesaplanması

Bu bölümde sorunun çözümü için PANAS şablonundaki her sıfat ayrı ayrı ele alınacak ve her birinin ilgili duygu durumuna yapacağı katkı eşit hale getirilerek hesaplamaya alınacaktır.

PANAS – T referans metninde geliştirilen hesaplama metodu aslında PANAS'ın genişletilmiş versiyonu olan PANAS-X'e göre geliştirilmiştir. PANAS-X'te duygu durumları daha çeşitli ve birbirinden bağımsız olarak ele alınsa da yine her bir duygu boyutu için birden fazla betimleyici sıfat vardır ve bu sıfatlar yine dağılım ve standart sapma değerlerine bakılmaksızın lineer olarak toplanır ve analize alınır. Bir önceki

⁴⁴ John R. Crawford and Julie D. Henry, "The positive and negative affect schedule(panas): construct validity, measurement properties and normative data in a largenon-clinical sample" (Department of Psychology, King's College, University of Aberdeen, UK) , The British Psychological Society, British Journal of Clinical Psychology (2004), 245–2652004, 43

başlıkta sayılan sebeplerden dolayı bu yöntem analizi bozabilir veya duygu alt setlerinden elde edilecek bilgilerin kaçırılmasına neden olabilir. Her betimleyici sıfat kendi içinde ölçülerek ve bunların PANAS-T skorlarına aynı ağırlıkta katkı yapması sağlanarak sapmalar düşürülebilir ve alt duygu setlerine dair daha geniş bilgilere ulaşılabilir.

Bu iddiaya göre hem büyük veri kümesi içerisinde, hem de gündem özelinde α ve β değerleri sıfatın kendisi için izole olarak hesaplanmıştır ve her bir sıfat kendi duygu grubundaki PANAS-T skorunda eşit olarak ağırlık almıştır.

Örneğin : Büyük veri kümesi içerisinde “**ürkmüş**” sıfatının frekansı 12.503 adettir. İçeriklerin toplamı 308.139.050 adettir. Buna göre;

$$\alpha(\text{korkmuş}) = \frac{|3.802.897|}{|308.139.050|} \text{ ve Suriyeliler gündemi için } \beta(\text{korkmuş}) = \frac{|403|}{|34.935|}$$

$P(\text{korkmuş}) = -0,00656$ olarak elde edilir. Korkmuş sıfatı 10 ayrı negatif duygu durumunu ifade eden sıfatlardan birisidir. Her bir sıfatın kendi içinde hesaplanarak ve aynı oranda PANAS-T duygu skoruna etki edebilmesi için, her bir betimleyici sıfatın değerinin ağırlığı %10 olarak alınacaktır. Böylece her sıfat kendi içinde ölçülerek 10’a bölünecek ardından tamamı toplanarak PANAS-T duygu boyutlarının toplam değeri elde edilecektir. Bu yöntemde ek olarak ilgili gündemde hangi sıfatın hakim olduğu ya da duygu durumlarının hangi alt betimleyici sıfattan/duygu durumundan beslendiği de tespit edilebilecektir.

4.4.1 Gündemlerde Skorların İzole Hesaplanması

Bir önceki başlıkta bahsedilen yöntem uygulanarak büyük veri kümesindeki her duygu durumu için α değerleri şu şekilde hesaplanmıştır;

Pozitif Sıfatlar	$\alpha(\text{poz})$	Negatif Sıfatlar	$\alpha(\text{neg})$
α (Hevesli)	0,00193	α (Ürkmüş)	0,00004
α (İlgili)	0,00249	α (Korkmuş)	0,01234
α (Kararlı)	0,00047	α (Mutsuz)	0,00192
α (Heyecanlı)	0,00322	α (Sıkıntılı)	0,00699

α (İlhamlı)	0,00030	α (Asabi)	0,00008
α (Uyanık)	0,00233	α (Sinirli)	0,00180
α (Aktif)	0,00111	α (Utanmış)	0,00122
α (Güçlü)	0,00412	α (Suçlu)	0,00178
α (Gururlu)	0,00296	α (Tedirgin)	0,00034
α (Dikkatli)	0,00426	α (Düşmanca)	0,00529

Tablo – 10 İzole Yöntemle Hesaplanan α Değerleri

Gündemlerdeki PANAS-T skorunun tek tek hesaplanabilmesi için her bir gündemde de izole β değerlerinin bulunması gereklidir. Tüm gündemlerin izole β değerleri ektedir. Suriyeliler gündemi için izole β değerleri formüle alınarak hesaplanan PANAS skorları şu şekildedir:

Pozitif	$P(x)$	$P(x)$Ağırlık	Negatif	$P(x)$	$P(x)$Ağırlık
Hevesli	-0,8238	-0,0824	Ürkmüş	0,3333	0,0333
İlgili	-0,8835	-0,0884	Korkmuş	-0,0632	-0,0063
Kararlı	-0,0149	-0,0015	Mutsuz	-0,7760	-0,0777
Heyecanlı	-0,9751	-0,0976	Sıkıntılı	0,2665	0,0267
İlhamlı	-0,9000	-0,0900	Asabi	-0,2500	-0,0250
Uyanık	-0,1759	-0,0176	Sinirli	0,4752	0,0475
Aktif	-0,9729	-0,0973	Utanmış	-0,2704	-0,0270
Güçlü	-0,6942	-0,0694	Suçlu	0,3881	0,0388
Gururlu	-0,4493	-0,0449	Tedirgin	0,2093	0,0209
Dikkatli	0,3004	0,0300	Düşmanca	0,4704	0,0470
Toplam		-0,5591	Toplam		0,0782

Tablo – 11 Suriyeliler Gündeminde İzole Hesaplanan PANAS Skorları

İlk hesaplama yönteminden farklı olarak her bir duygu durumu için hesaplanan PANAS değeri, kendi duygu grubunda ona bölünerek ağırlıklandırılmıştır. Ardından

bütün deęerler toplanarak negatif ve pozitif duygu durumuna iliřkin sonu elde edilmiřtir. Bylece her sıfat kendi iinde deęerlendirilerek frekans farklılıkları ortadan kaldırılmıřtır.

Bu yntem ayrıca tikel olarak hakim duygu durumuna iliřkin bilgi de vermektedir. rneęin Suriyeliler gndeminde hakim duygu ifadesi alt setleri dřmanca ve sinirli alt setleridir. En fazla kaınılan duygu ifadeleri heyecanlı, aktif ve ilhamlıdır. Bu ayırıştırma yntemi hem toplam deęer aısından daha isabetli analiz yapılmasını saęlamakta hem de mikro lekte gndemlerin duygu durumlarının tespit edilebilmesini saęlamaktadır.

İzole hesaplama yntemi gndemlerin tamamı iin uygulanmıřtır. Btn gndemlerin duygu alt setleri iin PANAS-T deęerleri ektedir.⁴⁵Bu yntemle elde edilen PANAS-T skorları ise řu řekildedir;

Gndem	P _(pozitif)	P _(negatif)	SJ(Sbjektivite)
<i>Suriyeliler</i>	–0,55910	0,07820	0,31865
<i>R.T Erdoğan</i>	–0,11688	–0,59547	0,35618
<i>Orman Yangınları</i>	–0,24921	–0,34159	0,29540
<i>Enflasyon</i>	–0,36274	–0,39155	0,37715
<i>Ukrayna Rusya</i>	–0,20875	–0,31438	0,26156
<i>Corona</i>	–0,37086	–0,28375	0,32730
<i>Turkovac</i>	–0,38011	–0,69712	0,53861
<i>Dolar TL</i>	–0,46380	–0,51959	0,49170

⁴⁵ Ek-5 Gndemlerin İzole Yntemle Hesaplanmış PANAS Duygu alt setleri deęerleri

<i>Şampiyon Beşiktaş</i>	-0,25146	-0,59329	0,42237
<i>Ramazan</i>	-0,28761	-0,61337	0,45049

Tablo – 12 Tüm Gündemlerin İzole Yöntemle PANAS Skorları

4.5. Yöntemlerin Mukayesesi ve Bulgular

Sıfatların izole olarak hesaplanmasıyla ve önceki verilerle çaprazlanmasıyla önemli bulgular elde edilmiştir. Bunlardan ilki gündemlerdeki duygu ifadelerinden kaçınmaya ilişkin eğilimlerdir. Twitter kullanıcıları seçilen kritik gündemler dışındaki konularda yazarken betimleyici sıfatları daha fazla kullanmaktadır. Gündemlerde yazarken ise duygu ifadelerinden kaçınma tespit edilmiştir ve fark çok belirgindir. Sıfatların toplu olarak değerlendirildiği hesaplama yönteminde, 10 ayrı gündem için pozitif ve negatif duygu durumuna ilişkin PANAS-T skorlarının toplamı sırasıyla -0,9269 ve -3,3028 olarak bulunmuştur. Bu skorlar gündem başına ortalama olarak -0,0926 ve -0,3303 değere tekabül etmektedir ve negatif PANAS-T skorları duygu ifadesinden kaçınma anlamı taşımaktadır. Hali hazırda gündemlerin PANAS-T skorlarının hesaplanmasına büyük ve rassal veri setinden elde edilen değerler girdiği için, elde edilen skor doğrudan rassal dağılımla mukayeseyi vermektedir.

Diğer önemli bir bulgu ise gündemlerde negatif sıfatlardan kaçınmanın, pozitif sıfatlardan kaçınmadan çok daha fazla olmasıdır. Kullanıcılar herhangi bir gündeme ilişkin duygu durumu ifade ederken negatif ifadelerden daha fazla kaçınmıştır.

Bunların dışında yine gündemlerde duygu durumu ifade ederken sıfatların yoğunlaştığı gözlemlenmiştir. Kullanıcılar gündemlerdeki duygu durumunu belli sıfatları yoğun bir şekilde kullanarak ifade etmektedir.

Sıfatların izole hesaplanmasıyla hedeflenen iki amaçtan birine ulaşılmış diğerine ise ulaşamamıştır. Bu amaçlar ölçüm metodolojisini bozmadan dağılımların sapmasını düşürmek ve duygu alt setlerine ilişkin daha geniş verilere ulaşmaktır. Sapmalar bu yöntemde de yüksektir ve duygu durumlarının gündemler için hemen her zaman belli sıfatlarda yoğunlaşarak ifade edildiği görülmüştür. Diğer taraftan izole ölçüm hangi

duygu alt setlerinde yoğunlaşma olduğunun tespit edilebilmesini sağlamış ve gündemlerdeki duygu durumlarına ilişkin detaylı çıktılar oluşturmuştur.

4.5.1 Duygu Alt Setleri ile İlgili Bulgular

Gündemlerdeki duygu ifadelerinden kaçınma ve duygu ifadelerinin tikel yoğunlaşması, gündemdeki hakim duygu alt setinin tespiti edilerek daha detaylı bilgiler elde edilmesini sağlamaktadır. Bu bölümde izole hesaplanmış PANAS-T alt duygu setlerinden en yüksek skora sahip olanlar incelencektir

Sonuçlar şu şekildedir;

(negatif setler kırmızı, pozitif setler yeşil renkle işaretlenmiştir)

Gündem	P _(baskın-1)	P _(baskın-2)
<i>Suriyeliler</i>	<i>Sinirli</i> 0,4752	<i>Düşmanca</i> 0,4704
<i>Recep Tayyip Erdoğan</i>	<i>Kararlı</i> 0,79653	<i>Güçlü</i> 0,58527
<i>Orman Yangınları</i>	<i>Dikkatli</i> 0,63405	<i>Suçlu</i> 0,50372
<i>Enflasyon</i>	<i>Kararlı</i> 0,72504	<i>Dikkatli</i> 0,43711
<i>Ukrayna Rusya</i>	<i>Kararlı</i> 0,68289	<i>Dikkatli</i> 0,56459
<i>Corona</i>	<i>Dikkatli</i> 0,52007	<i>Tedirgin</i> 0,47419
<i>Turkovac</i>	<i>Gururlu</i> 0,83393	<i>Dikkatli</i> 0,78345
<i>Dolar TL</i>	<i>Dikkatli</i> 0,43711	
<i>Şampiyon Beşiktaş</i>	<i>Hevesli</i> 0,64598	<i>Gururlu</i> 0,34010
<i>Ramazan</i>	<i>Heyecanlı</i> 0,4713	<i>İlhamlı</i> 0,25809

Tablo – 13 İzole Yönteme Göre En Yüksek Skorlu Duygu Altsetleri

Hesaplamaya giren 19 duygu alt setinden sadece 3 tanesi negatif duygu grubunda kalmaktadır. En yüksek frekansa sahip baskın duygu alt seti “dikkatli”dir ve baskın duygu alt setlerinin %31,57’si dikkatli duygu alt setinden oluşmaktadır. Dağılımdaki bu anomalinin nedenini anlamak için dikkatli duygu alt setindeki içerikler incelenmiştir. İnceleme neticesinde bazı betimleyici sıfatların, içeriğin bağlamına göre PANAS şablonunda bağlı olduğu duygu grubunun tam tersi bir anlam yüklediği ortaya çıkmıştır.

Buradaki temel sebep PANAS’ın ve PANAS-T’nin uygulama farklılıklarına ilişkindir. PANAS testinde betimleyici sıfat herhangi bir bağlamda verilmez. Deneğin duygu durumuna ilişkin bilgi almak için bağlamdan bağımsız olarak verilir ardından daha önce korelasyon çalışmaları yapılmış bağlamlar için yorumlanır (örneğin depresyon, anksiyete vb.) PANAS-T’de ise betimleyici sıfatların tamamı farklı metotlarla betimleyici anahtar kelimelere dönüştürülür ve betimleyici anahtar kelimelerin tamamı cümledeki bağlamına göre pozitif ya da negatif duygu durumu ifade eder. Hali hazırda PANAS sıfatlarının faktör analizi yapılmış olsa da, PANAS’tan türetilen betimleyici anahtar kelimelerin faktör analizinin yapıldığı bir çalışma bulunmamaktadır. Faktör analizi ile referans verilen betimleyici sıfatların veya anahtar kelimelerin tek bir duygu boyutunu yüksek oranda temsil ederken, karışık duygu boyutunu minimum oranda temsil ettiği bağımsızlık analizleridir. Daha önceki bölümlerde değinildiği gibi. Bu analizler Türkçe’de betimleyici sıfatlar için varimaks analizi şeklinde yapılmıştır.

“Dikkatli” betimleyici sıfatı, betimleyici anahtar kelime olarak hem negatif hem de pozitif duygu ifadelerinde yüksek oranda pay almıştır ve büyük ölçüde dikkatli sıfatındaki yoğunlaşmanın sebebi budur. Örneğin; “*Orman yangınları ihtimaline karşı daha dikkatli olmalıydık.*” ya da “*dikkat edin ormanlarımız yanıyor*” gibi ifadeler negatif olsa da PANAS şablonuna göre pozitif duygu boyutu içerisinde değerlendirilmiştir. Bu da analizde göreceli olarak sapmaya yol açmaktadır.

Sapmaların tespiti ve iyileştirme için uygulanacak çözüm farklı ölçüm metotlarıyla gündemlere ilişkin toplanan anket verileri arasında korelasyon analizleri yapmaktır. Bu analizlere göre anketlerle, Twitter’da toplanan veri setleri arasındaki en güçlü ilişkinin hangi yöntemle elde edildiği tespit edilebilecektir.

5.0 Anket Verileri

Twitter'daki veri setlerinden elde edilen PANAS-T skorlarının, gündemlere ilişkin kamuoyu fikrini ne ölçüde yansıttığını tespit etmek çalışmanın temel amacıdır. Bu tespitin yapılması için ilgili gündemlere ilişkin bir anket yapılmıştır ve anket verilerinden elde edilen PANAS-T ile Twitter verilerinden elde edilen PANAS-T skorları arasındaki korelasyonlar incelenerek hipotez test edilecektir. Anket sorularına ilişkin şablon ekteedir.⁴⁶

Anket, her gündem için katılımcının gündemle ilgili duygu durumunu 1-5 aralığında bir likert ölçeğinde ölçmeye çalışır. Katılımcıdan gündeme ilişkin “*"Suriyeliler" konusu/temasını sizin için hangi oranda pozitif/olumlu duygu durumu çağrıştırmaktadır.*” vb. soruları ve negatif duygu durumuna ilişkin olacak şekilde aynı cümlelerin tersi olarak puanlama yapması istenmektedir. 1-ilgili duygudan uzak, 5-ilgili duyguya yakın olacak şekilde duygu durumu ölçeğe dağıtılmıştır. Her gündem için negatif ve pozitif duygu durumunu içeren iki soru vardır. Daha önceki bölümde değinilmiş olan duygu yapısının birbirinin kutupsal karşıtlığı olmadığına ilişkin bulgulara dayanarak pozitif ve negatif duygu durumları yine ayrı bir şekilde ölçülmüştür. Ardından ölçek üzerinden elde edilen verilerin Twitter PANAS-T skorları ile kıyaslanabilmesi için anket verileri şu şekilde PANAS-T skoruna uyarlanmıştır ;

İlgili Duygudan Çok Uzak - 1 = -1.00

İlgili Duygudan Uzak = -0.50

Nötr = 0

İlgili Duyguya Yakın = +0,50

İlgili Duyguya Çok Yakın = +1,00

Verilen her cevap yukarıdaki şablona göre PANAS-T skoruna çevrilmiştir. PANAS-T şablonu duygu durumlarına ilişkin değeri -1 ve +1 arasında dağıtmaktadır. Buna göre ilgili duygu durumundan en uzak değer -1 en yakın değer +1'dir. Ankette gündeme verilen cevaplar bu skor aralığında likert ölçeğine dağıtılmıştır. Ardından her gündem için elde edilen toplam skorun aritmetik ortalaması alınarak gündemin PANAS-T skoru bulunmuştur.

⁴⁶ Ek – 6 Anket Soruları

Gündem	P _(pozitif)	P _(negatif)	SJ(Sübjektivite)
<i>Suriyeliler</i>	0,0025	0,0170	0,0097
<i>R.T Erdoğan</i>	0,4020	-0,3950	0,3935
<i>Orman</i>	-0,7570	0,7170	0,7370
<i>Yangınları</i>			
<i>Enflasyon</i>	-0,6760	0,6560	0,6660
<i>Ukrayna Rusya</i>	-0,2770	0,2850	0,2810
<i>Corona</i>	-0,6080	0,5640	0,5860
<i>Turkovac</i>	0,4000	-0,4190	0,4095
<i>Dolar TL</i>	-0,6870	0,5810	0,6340
<i>Şampiyon</i>	-0,0238	0,0640	0,0439
<i>Beşiktaş</i>			
<i>Ramazan</i>	0,8490	-0,8500	0,8495

Tablo – 14 Anketlerin PANAS Skorları

Toplamda 500 deneğin katıldığı ankette elde edilen oylama sonuçları ektedir.⁴⁷Bu oylama sonuçlarına göre gündemdeki PANAS-T skorları aşağıdaki şekilde hesaplanmıştır. PANAS’a ilişkin çalışmalarda katılımcıların demografik farklılıkları ile elde edilen ölçüm değerleri arasındaki farklar çok küçük ve ihmal edilebilirdir. Bu yüzden katılımcıların demografik profili önemsizdir.⁴⁸

PANAS’a ilişkin ölçümlerin ve korelasyon analizinin doğru bir şekilde yapılabilmesi için gündemler özellikle duygu durumu kolay tahminlenebilir ve polarize başlıklardan seçilmiştir. Anket verilerinin gündemlere ilişkin genel fikri yansıttığı varsayımı

⁴⁷ Ek – 7 Gündem Anketlerinin Oylama Sonuçları

⁴⁸ The positive and negative affect schedule(panas): construct validity, measurement properties and normative data in a largenon-clinical sample, john r. crawford and julie d. henrydepartment of psychology, king’s college, university of aberdeen, uk , British Journal of Clinical Psychology (2004), 43, 245–2652004 The British Psychological Society

altında, řu ana kadar ölçölen PANAS-T skorlarının korelasyon analizi yapılarak hipotez test edilecektir.

6.0 Korelasyon analizleri

Bu çalışmada PANAS-T skorları iki ayrı yöntemle ölçölmüřtür ve buna ek olarak bazı deęerlerde tespit edilen yüksek sapmalar, betimleyici sıfatlar bazında ve mikro ölçekte ayrıca incelenmiřtir. Bu bölümde ise ölçömlere iliřkin korelasyon analizleri yapılarak ölçüm metotlarının isabeti kıyaslanacak ve sapmalara iliřkin bulgular desteklenmeye çalışılacaktır.

Anket verileri ile her iki ölçüm yöntemi için ve duygu grupları özelinde korelasyon analizleri yapılacaktır. Örneęin anketlerden elde edilen pozitif PANAS-T skorları ile sıfatların saydırılarak ölçöldüęü pozitif PANAS-T skorunun korelasyon analizi yapılacak ardından aynı iřlem negatif duygu grubu içinde yapılacaktır. Böylece iki ayrı ölçüm yöntemi ve iki ayrı duygu grubu için toplamda dört korelasyon analizi incelenecektir.

Sıfatların tümünün toplanarak ölçöldüęü yöntemle, anket verileri arasındaki korelasyon analizleri řu şekildedir;

		Correlations	
		Anket_Poz	Normal_Poz
Anket_Poz	Pearson Correlation	1	,305
	Sig. (2-tailed)		,392
	N	10	10
Normal_Poz	Pearson Correlation	,305	1
	Sig. (2-tailed)	,392	
	N	10	10

Tablo 15 - Toplama Yöntemine Göre Pozitif Duygu Durumu Korelasyonları

		Correlations	
		Anket_Neg	Normal_Neg
Anket_Neg	Pearson Correlation	1	,545
	Sig. (2-tailed)		,103
	N	10	10
Normal_Neg	Pearson Correlation	,545	1
	Sig. (2-tailed)	,103	
	N	10	10

Tablo – 16 Toplama Yöntemine Göre Negatif Duygu Durumu Korelasyonları

Veriler normal dağılmaktadır bu yüzden Pearson korelasyon katsayısı baz alınmıştır. Pozitif duygu grubu için 0,305 değeri ile zayıf ilişki tespit edilmiştir. Negatif duygu grubu içinse 0,545 değeri ile orta düzeyde ilişki tespit edilmiştir. Özellikle pozitif duygu grubu için ilişkinin zayıf olması ve iki duygu grubu arasındaki korelasyon katsayısı farklarının yüksek olması duygu alt setlerinin incelendiği bölümdeki tespitleri teyit eder niteliktedir. Bu bölümde her gündem için en yüksek PANAS-T skoru alan duygu alt setleri incelenmiş ve en yüksek frekansa sahip duygu alt setinin “dikkatli” olduğu görülmüştür. “Dikkatli” betimleyici sıfatından türetilen betimleyici anahtar kelimeler iki yönlü olarak yüksek faktör yükü almaktadır. Yani hem pozitif, hem de negatif olarak cümlelerin bağlamına göre anlamı değişmektedir. Her iki duygu grubu içinde yüksek frekansta kullanılması PANAS-T şablonu üzerinden gündem

analizini bozmaktadır. Her gündem için en baskın iki duygu alt setinin seçilerek incelendiği bölümde tespit edilen 19 baskın duygu alt setinin 5 tanesinin “dikkatli” duygu alt seti olduğu görülmüştür. Bu da baskın duygu alt setlerinin %31,57’sinin “dikkatli” duygu alt setinden oluştuğunu göstermektedir. Alt setteki bu yoğunlaşma “dikkatli” betimleyici sıfatından üretilen, betimleyici anahtar kelimelerin her iki yönde de yoğun olarak kullanıldığını teyit eden diğer önemli bir bulgudur.

Pozitif anlam yüklenmesi beklenen diğer duygu alt setlerinde de benzer sapmalar tespit edilmiştir. Örneğin kararlı duygu alt seti de bağlama göre yüksek oranda negatif anlam almaktadır ancak pozitif duygu grubunda sayılmaktadır. Bu noktada duygu alt setlerine ilişkin daha derin analiz anlamlı değildir çünkü sorun çalışmanın ilerleyen bölümlerinde, farklı bir yaklaşımla tüm duygu alt setleri için çözülecektir.

Correlations

		Anket_Poz	İzole_Poz
Anket_Poz	Pearson Correlation	1	,232
	Sig. (2-tailed)		,518
	N	10	10
İzole_Poz	Pearson Correlation	,232	1
	Sig. (2-tailed)	,518	
	N	10	10

Tablo – 17 İzole Yönteme Göre Pozitif Duygu Durumu Korelasyonları

Correlations

		Anket_Neg	İzole_Neg
Anket_Neg	Pearson Correlation	1	,428
	Sig. (2-tailed)		,217
	N	10	10
İzole_Neg	Pearson Correlation	,428	1
	Sig. (2-tailed)	,217	
	N	10	10

Tablo- 18 İzole Yönteme Göre Negatif Duygu Durumu Korelasyonları

Sıfatların izole olarak ölçüldüğü yönteme göre yapılan korelasyon analizi de benzer sonuçlar vermiştir. Yine pozitif duygu durumu ile ilişki zayıf, negatif duygu durumu ile orta derecede bulunmuştur. Bu verilerde betimleyici anahtar kelimelerin cümlelerin

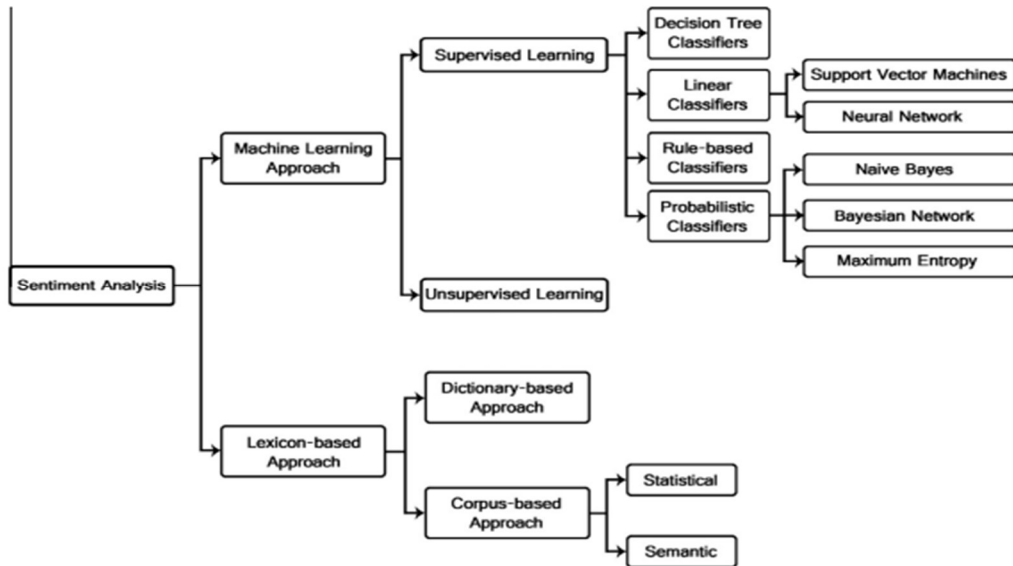
bağlamına göre anlam kaymasına uğradığını ve sorunun birden fazla betimleyici anahtar kelimede olduğunu teyit etmektedir. Çünkü izole ölçüm yönteminde her bir duygu alt seti sadece belli bir oranda PANAS-T skoruna etki edebilmektedir.

Korelasyon verileri, betimleyici anahtar kelimeler yorumlanırken cümlelerin bağlamında analiz edilmesini zorunlu kılmıştır. Bunun sağlanabilmesi için proje içerisinde ikinci bir duygu analizi (sentiment analiz) katmanı geliştirilmiş ve uygulamaya alınmıştır.

6.1 Sentiment(Duygu) Analizi Katmanı ve İyileştirme

Duygu analizi (sentiment analiz), insanların farklı konulardaki duygu durumlarını, görüşlerini, düşüncelerini bilgisayar destekli ve hesaplamalı olarak ele alan disiplindir.

⁴⁹ PANAS'ın doğrusal uyarlaması da temelde bir sentiment analiz çalışmasıdır.



50

Görsel – 8 Sentiment Analiz Yaklaşımlarının Tasnifi

⁴⁹ Walaa Medhat , Ahmed Hassan b , Hoda Korashy ; “Sentiment analysis algorithms and applications: A survey”, Ain Shams Engineering Journal (2014), 5, 1093-1113

⁵⁰ Walaa Medhat , Ahmed Hassan b , Hoda Korashy ; “Sentiment analysis algorithms and applications: A survey”, Ain Shams Engineering Journal (2014), 5, 1093-1113

Görselde görüldüğü gibi sentiment analiz çalışmaları temelde ikiye ayrılır. Uygulamacı birinci aşamada çalışmayı makine öğrenimi ile mi yoksa sözlük tabanlı yaklaşımla mı ele alacağını seçer. Makine öğrenimi yaklaşımında, genellikle uygulamaya verilen etiketlenmiş bir veri seti ile duygu analizine ilişkin vektörler oluşturulur ve farklı yöntemlerle elde edilen bu vektörler yeni gelen veri setlerini yorumlayarak çıktı üretir. Sözlük tabanlı yaklaşımlarda ise kelime özellikleri üzerinden statik bir analiz yapılmaya çalışılır. Bu çalışmanın konusu olan PANAS'ta temelde sözlük tabanlı bir yaklaşımdır. Duygu boyutları açısından farklı çalışmalarda anlamlılığı test edilmiş kelime grupları üzerinden analiz yapar.

Daha önceki bölümlerde korelasyon analizleri ile tespit edilen, PANAS'taki betimleyici anahtar kelimeler ve bağlama göre anlam kayması sorunları iki yöntem kombine edilerek çözülmeye çalışılmıştır. Yine ekte görülen ⁵¹TÜR (Teknolojik Ürün Geliştirme) çalışması çerçevesinde multi-class (çok sınıflı; pozitif, nötr, negatif) analiz yapan bir sentiment modülü geliştirilmiştir ve betimleyici anahtar kelimelerin geçtiği cümlelerin bağlamı tespit edilerek, betimleyici anahtar kelimeler sadece içinde bulundukları duygu grubuyla uyumlu olmaları halinde analize alınmıştır.

6.2 Sentiment Modülü

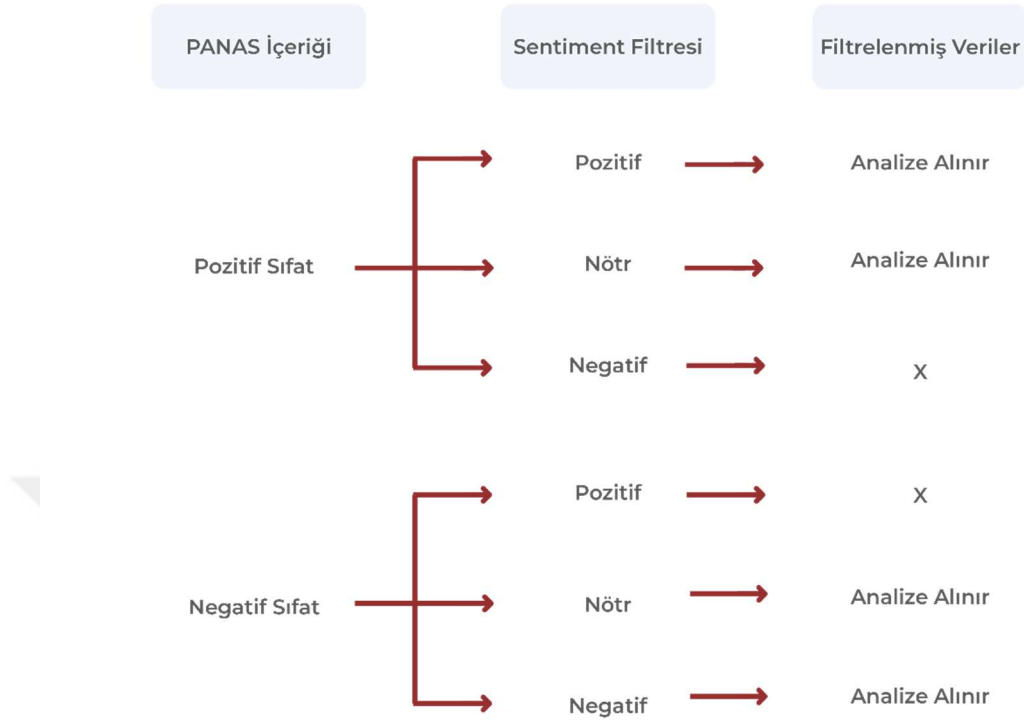
TÜR projesi çerçevesinde oluşturulan sentiment modülünde makine öğrenimi on araştırmacının katılımıyla sağlanmıştır. Yönteme göre Twitter'dan çekilen 25.000 adet veri üç farklı kategoride tek tek etiketlenmiştir. Bu kategoriler pozitif, nötr ve negatif kategorileridir. Araştırmacı kendisine verilen içeriği okur ve en yakın gördüğü duygu durumuna göre içeriği üç kategoriden biri için etiketler. Ardından araştırmacının etiketlediği veriler farklı iki araştırmacı tarafından rasgele seçimlerle kontrol edilir. Kontrolü yapan etiketlenmiş her 1.000 içerikte, 40 tanesini rasgele olarak seçer ve etiketler. Kontrolü yapanlarla, veriyi daha önceden etiketleyen araştırmacı arasındaki etiketleme farklılıkları ve sapmalar tespit edilir. Negatif ve pozitif arasındaki ölçüm sapmaları 5 hata puanı ve nötr ile diğer duygu durumları arasındaki sapmalar 2,5 hata puanı olarak kabul edilir. Toplam 40 örnekte 12,5'in üzerinde hata puanı tespit edilen etiketlenmiş veriler işleme alınmamıştır ve iki kez verileri işleme alınmayan araştırmacı veri etiketleme işleminden çıkarılmıştır.

⁵¹ Ek-8 TÜR belgesi

Etiketlenen 25.000 veri ile üç kategori için yeni içeriklerde duygu analizi %70'in üzerinde bir isabet oranına ulaşmıştır. Makine eğitimi için Google Bert derin öğrenme modeli kullanılmıştır. İsabet oranının artırılması için duygu analizinin dikey (vertical) yöntemlerle iyileştirilmesi gerektiği tespit edilmiştir. Her alanda daha iyi sonuçlar için arama motoru teknolojileri de dahil olmak üzere, doğal dil işleme disiplinleri dikey teknolojileri kullanmaya başlamıştır.⁵²Bu yaklaşım baz alınarak ek 10.000 adet veri betimleyici anahtar kelimeler özelinde veri tabanına çekilmiştir. Örneğin dikkatli betimleyici sıfatını temsil eden, betimleyici anahtar kelimelerle içerikler aratılmış, her bir betimleyici sıfat için 500 kadar içerik etiketlenerek, betimleyici sıfatlar özelinde makine eğitime devam edilmiştir.

En son veri setinin eklenmesiyle birlikte içeriklere ilişkin sentiment analizinin isabet oranı %78,50 olarak tespit edilmiştir. Bu oran mutlak isabet oranı diyebileceğimiz orana oldukça yakındır. Mutlak isabet oranıyla belli içeriklerin farklı araştırmacılar tarafından etiketlenmesindeki isabet oranı ve benzerlikler kast edilmektedir. Örneğin araştırmacılardan birinin etiketlediği verileri diğer araştırmacının test etmesi durumunda isabet oranı yaklaşık %82,5 oranında çıkmaktadır. Bu oran ulaşılacak en yüksek ve mutlak isabet oranıdır çünkü yorumlamadaki görelilik her zaman içeriklerin farklı araştırmacılar tarafından, farklı etiketlenmesine neden olacaktır. Bahsedilen oranlar hata mesafesine göre değil doğrudan ve statik olarak ölçülmüştür. Pozitif etiketlenen bir içeriğin diğer araştırmacı tarafından nötr yorumlanması ile pozitif etiketlenen bir içeriğin diğer araştırmacı tarafından negatif yorumlanması eşit oranda isabetsizlik olarak kabul edilmiştir.

⁵² Bernd Müller · Christoph Poley · Jana Pössel · Alexandra Hagelstein · Thomas Gübitz LIVIVO – “The Vertical Search Engine for Life Sciences,” DOI 10.1007/s13222-016-0245-2 Datenbank Spektrum (2017) 17:29–34



Görsel – 9 Sentiment Filtresi

Sentiment modülü PANAS'taki betimleyici sıfatlar için bir filtre olarak kullanılmıştır. Görseldeki yöntemle göre sentiment filtresi PANAS'ın betimleyici sıfatlarını içeren içeriklerin çekilmesinin ardından, duygu boyutuna göre içeriği değerlendirir. Örneğin “Dikkatli” betimleyici sıfatını içeren içerik, sentiment analize tabi tutulduktan sonra modülün döndürdüğü değere göre analize alınır veya alınmaz. “Dikkatli” sıfatı pozitif duygu boyutunda olduğu için sentiment analiz sonucu pozitif ve nötr ise analize alınır, negatif ise analiz dışında tutulur. Negatif duygu boyutundaki sıfatlar içinse tam tersi uygulanır örneğin, “suçlu” betimleyici sıfatının tespit edildiği cümle sentiment modülünde negatif ya da nötr değer alıyorsa analize alınır, pozitif değer alıyorsa analize alınmaz çünkü “suçlu” betimleyici sıfatı negatif duygu grubundadır.

Filtreleme işlemiyle içeriklerin miktarlarında ve niteliğinde büyük değişimler tespit edilmiştir. Betimleyici anahtar kelimelerle, olumsuz cümlelerde ters anlam yüklenen içeriklerin hepsi temizlenmiştir.

Örneğin;

“*Dikkatli olunmadığı için orman yangınları önlenemedi.*” Cümlesi negatiftir. Olumsuz takıyla pozitif betimleyici sıfat negatif anlam yüklenmiştir. Sentiment modülü bu cümlenin anlamını negatif olarak döndürür ancak sıfat PANAS duygu grubunda pozitif temsil ettiği için analize alınmaz.

“*Düşmanlık etmeyeceğiz kardeşçe yaşayacağız.*” cümlesi pozitifdir. “Düşmanlık” betimleyici sıfatı negatif duygu grubuna aittir. Cümle pozitifdir ve kelimenin bağlı olduğu duygu grubu negatif olduğu için analize alınmaz.

6.3 Filtrelenmiş İçeriklerin PANAS Skorları ve Korelasyonlar

Sentiment modülü ile filtreleme işleminin ardından büyük rassal veri kümesi içerisindeki toplam içerik sayısı aşağıdaki şekilde tespit edilmiştir;

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Hevesli	345.991	Ürkmüş	10.151
İlgili	514.250	Korkmuş	3.173.769
Kararlı	45.012	Mutsuz	479.392
Heyecanlı	727.993	Sıkıntılı	1.681.380
İlhamlı	77.438	Asabi	19.153
Uyanık	400.742	Sinirli	434.626
Aktif	235.562	Utanmış	300.051
Güçlü	980.249	Suçlu	475.783
Gururlu	835.407	Tedirgin	87.943
Dikkatli	311.467	Düşmanca	1.186.141
Toplam	4.474.111	Toplam	7.848.389

Tablo – 19 Sentiment Modülü ile Filtrelenmiş Toplam İçerikler

Elde edilen rakamlar önceki bulgulara paralel niteliktedir. Pozitif sıfatlar ve özellikle “dikkatli” betimleyici sıfatıyla elde edilen içerikler dramatik bir biçimde azalmıştır. Filtrelemeden sonra toplamda 12.322.500 adet içerik tespit edilmiştir ve bu içeriklerin %69,69’u negatif, %31,31’i pozitif duygu grubuna aittir.

Filtreleme yöntemi bütün gündemlere tek tek uygulandıktan sonra elde edilen PANAS skorları şu şekildedir.

Gündem	P _(pozitif)	P _(negatif)	SJ(Sübjektivite)
<i>Suriyeliler</i>	-0,37730	0,68221	0,52976
<i>Recep Tayyip Erdoğan</i>	0,26692	-0,44219	0,35456
<i>Orman Yangınları</i>	-0,32593	0,32214	0,32404
<i>Enflasyon</i>	-0,46715	0,44904	0,45810
<i>Ukrayna Rusya</i>	-0,44330	-0,05980	0,25155
<i>Corona</i>	-0,55912	0,32012	0,43962
<i>Turkovac</i>	0,45366	-0,68027	0,56697
<i>Dolar TL</i>	-0,01083	0,04987	0,03035
<i>Şampiyon Beşiktaş</i>	0,00256	-0,33758	0,17007
<i>Ramazan</i>	0,69270	-0,46681	0,57976

Tablo – 20 Filtrelenmiş İçeriklerle Elde Edilen PANAS Skorları

Sıfatların toplanarak PANAS-T skorlarının hesaplandığı ilk yöntem, izole hesaplama yöntemine göre anket verileriyle daha yüksek korelasyon verdiği için sentiment modülüyle bu yöntem tercih edilmiştir.

Hesaplanan korelasyonlar şu şekilde gerçekleşmiştir.

Correlations

		Anket_Poz	Filtered_Poz
Anket_Poz	Pearson Correlation	1	,833**
	Sig. (2-tailed)		,003
	N	10	10
Filtered_Poz	Pearson Correlation	,833**	1
	Sig. (2-tailed)	,003	
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Tablo – 21 Filtrelenmiş İçeriklerin Pozitif PANAS Değerlerinin Korelasyonları

Correlations

		Anket_Neg	Filtered_Neg
Anket_Neg	Pearson Correlation	1	,715*
	Sig. (2-tailed)		,020
	N	10	10
Filtered_Neg	Pearson Correlation	,715*	1
	Sig. (2-tailed)	,020	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

Tablo – 22 Filtrelenmiş İçeriklerin Negatif PANAS Değerlerinin Korelasyonları

7.0 Kullanılan Teknoloji

Bu bölümde veri kazıma ve analizinde kullanılmış olan teknolojik altyapı kısaca tanıtılacaktır. Veri çekimi modüler bir şekilde üç farklı servis ile yapılmaktadır. Modüler bir mimari ile servislerin bakımı ve ihtiyaca göre geliştirilmesi gerekli olduğu için data ile ilgili işlemlerin farklı servislere bölünmesine karar verilmiştir.

Proxy-Service:

Veri çekim işleminin tek bir i.p adresi üzerinden yapılması mümkün değildir. Çünkü daha önceki bölümlerde değinildiği üzere Twitter API'leri çalışmanın yapılması için

gerekli olan veri miktarının etkin bir biçimde elde edilmesini sağlayamaz. Diğer taraftan tek bir i.p adresi üzerinden yapılan büyük miktarlarda sorguları servis sağlayıcı bloke etmektedir.

Proxy-Service modülü bu sorunun çözülmesi için farklı proxyler sunmaktadır. Servis taban teknoloji olarak python ve fastapi frameworkunu kullanmaktadır ve çalıştırıldıktan sonra belirlenen parametreye göre kullanıma hazır SOCKS proxyleri oluşturmaktadır. Bir microservice olarak Proxy-Service'in görevi sadece proxylerin oluşturulması ve yaşam döngüsü yönetimi ile sınırlıdır. Oluşturulan proxyler port range parametresi üzerinden kullanılabilir. Gelecekteki araştırmacılara yol göstermesi açısından, daha büyük miktarlarda verilerle çalışmak için proxy sorgulama işleminin fastapi entegrasyonlarıyla ağılaştırılması tavsiye edilmektedir.

Scrape-Service:

Scrape-Service ile Twitter arama kriterleri okunup Twitter advanced search query'e uygun sorgular oluşturularak tweetler akış halinde çekilmektedir. Twitter arama motorunu kullanabilmek için servis içerisinde ilk etapta guest token alınmaktadır. Tokenı bütün sorgularda belirli bir akışı çekmek için kullanılmaktadır. Elde edilen token ile arama sorgusuna göre uygun büyük bir data akışını 20 tweet iterasyon ile çekmek mümkün olmaktadır.

Normal şartlarda her bir arama sonucu Twitter tarafından veri akışı şeklinde döndürüldüğü için aramaların paralel bir şekilde yapılması mümkün değildir çünkü 20 tweet sonrasında çekmek için mevcut çekilen 20 tweetin sonunda döndürülen verinin referans değeri olarak kullanılması gerekmektedir. Bu teknik kısıtlamaya rağmen çekim işlemini paralel hale getirmek amacıyla Scrape-Service birden fazla sorguyu işleyebilmektedir ve böylelikle uzun bir arama periyodunu küçük periyodlara bölerek dağıtık bir şekilde işlem yapabilmektedir. Scrape-Service hem tweetlerin yanı sıra, metadataları da işleyip api üzeri döndürebilmektedir (user metadata, tweet metadata). Bir arama oturumunda çekilen tweetler api ile json array olarak döndürülmektedir. Servis python ve fastapi üzerine kurulmuştur ve parametrelerde kullanılacak proxy range belirlenebilmektedir. Çekim işleminde bağlantı ve proxy ile ilgili bir sorun olduğunda servis otomatikman bunu algılayıp proxy rotasyonu yapmaktadır.

Tweet-Service:

Yapılan tweet aramalarını yönetmek ve elde edilen tweetleri arama kriterlerine bağlı veri tabanına kayıt etmek için Java + Spring Framework + Hibernate tabanlı bir mikro servistir. Veri tabanı olarak Postgresql kullanılmaktadır. Her bir tweet arama sorgusu kullanıcı tarafından bu servise iletilmektedir. Paralel bir şekilde tweet çekimini gerçekleştirmek için Tweet-Service her bir isteği periyod olarak günlere bölmektedir ve tweet çekimi için scrape-service'e iletmektedir. Gönderme işlemi eşzamanlı yapılmaktadır ve herhangi bir hata oluşumunda otomatik bir şekilde işleme devam etmek için yeniden başlatma yapmaktadır. Büyük bir tweet akışı çekildiğinde elde edilen mevcut tweetler belirli aralıklarla veri tabanına kaydedilmektedir. Böylelikle herhangi bir sorun oluştuğunda ve herhangi bir periyod için çekim işlemi otomatik olarak yeniden başlatıldığında zaman kayıpları minimize edilmektedir. Çekilen tweetlere, tweet-service'in sunduğu apiler üzerinden JSON formatında her zaman erişebilmek mümkündür.

7.1 Verilerin Sayımı ve Analizi İşlemleri

Veri sayımı için depolanan milyonlarca veri yine akış halinde veri tabanından okunmaktadır. Çünkü bütün tweetlerin anında okunması ana belleğin sınırlı kapasitesinden dolayı mümkün değildir. Sayım işleminde bütün betimleyici anahtar kelimelerin çekilen tweetler içerisinde kaç defa geçtiği aşağıdaki algoritmaya göre kontrol edilmektedir.

Counts how many times the substring appears in the larger string. Note that the code only counts non-overlapping matches.

A null or empty ("") String input returns 0.

```
StringUtils.countMatches(null, *)      = 0
StringUtils.countMatches("", *)       = 0
StringUtils.countMatches("abba", null) = 0
StringUtils.countMatches("abba", "")  = 0
StringUtils.countMatches("abba", "a")  = 2
StringUtils.countMatches("abba", "ab") = 1
StringUtils.countMatches("abba", "xxx") = 0
StringUtils.countMatches("ababa", "aba") = 1
```

Params: str – the CharSequence to check, may be null
sub – the substring to count, may be null

Returns: the number of occurrences, 0 if either CharSequence is null

Since: 3.0 Changed signature from countMatches(String, String) to countMatches(CharSequence, CharSequence)

```
public static int countMatches(final CharSequence str, final CharSequence sub) {
    if (isEmpty(str) || isEmpty(sub)) {
        return 0;
    }
    int count = 0;
    int idx = 0;
    while ((idx = CharSequenceUtils.indexOf(str, sub, idx)) != INDEX_NOT_FOUND) {
        count++;
        idx += sub.length();
    }
    return count;
}
```

Görsel – 10 Apache commons-lang kütüphanesinden StringUtils.countMatches() metodu

PANAS'ın bazı betimleyici anahtar kelimelerin birden fazla kelimeden oluştuğu için işlem anahtar kelimeye göre farklılaşabilmektedir.

Birleşik betimleyici anahtar kelimelerde sayım işlemi tüm tweetler üzerinde yapılmaktadır, tek olan betimleyici anahtar kelimelerin tespiti ise tweetlerin kelimelere bölünüp sayımının yapılmasıyla elde edilmektedir. Böylelikle tweet içerisinde birleşik kelimelerde başarılı bir şekilde tespit edilebilmektedir.

```

public void processTweet(String tweet, PanasKeywordOccurrenceResponse panas, AtomicLong counter) {
    if (!StringUtil.isEmpty(tweet)) {
        String tweetContent = tweet.toLowerCase(Locale.ROOT).trim();
        String[] tweetChunks = tweetContent.split(SPACE);

        for (Map.Entry<String, Long> entry : panas.getKeywordOccurrence().entrySet()) {

            /*
             * process compound panas keywords on whole tweets
             * or process single panas keywords on split tweets
             */

            Long currentOccurrence = entry.getValue();
            String panasKeyword = entry.getKey().toLowerCase(Locale.ROOT).trim();

            if (panasKeyword.contains(SPACE)) {
                int count = StringUtil.countMatches(tweetContent, panasKeyword);
                panas.getKeywordOccurrence().put(panasKeyword, currentOccurrence + count);
            } else {
                for (String chunk : tweetChunks) {
                    if (chunk.startsWith(panasKeyword)) {
                        panas.getKeywordOccurrence().put(panasKeyword, currentOccurrence + 1);
                    }
                }
            }
        }

        panas.setProcessedTweetsCount(counter.incrementAndGet());

        if (ThreadLocalRandom.current().nextInt( origin: 0, bound: 100000) == 50000) {
            log.info("current statistics: {}", panas);
            this.entityManager.clear();
        }
    }
}

```

Görsel – 11 Tweet-Service Sayım Algoritması

Elde edilen sonuçlar JSON olarak döndürülür:

```
{
  "start": "2022-05-29T16:28:13.564401954",
  "end": "2022-05-29T16:46:24.72600959",
  "processedTweetsCount": 34935,
  "keywordOccurrence": {
    "çekindim": 0,
    "morali bozuk": 0,
    "gurur": 57,
    "dikkat": 213,
    "hayret": 21,
    "utanıyorum": 6,
    "cesur": 27,
    "sevin": 69,
    "uyuşuyorum": 0,
    "utanacağım": 0,
    "uyana": 8,
    "üzülürüm": 2,
    "uyurum": 0,
    "kabahatli": 0,
    "konsantre": 0,
    "özgüvenli": 2,
    "ilgilen": 48,
    "kararlı": 13,
    "uyanı": 55,
    "uyuşacağım": 0,
    "üzüldü": 16,
    "üzüleceğim": 1,
    "uyuyorum": 1,
    "uyuşmuşum": 0,
    "sakin": 43,
    "ilgili": 6,
    "üzülmüşüm": 0,
    "uyumuşum": 0,
    "üzüntü": 1,
    "memnun": 48,
  }
}
```

Görsel – 12 Betimleyici Sıfatların JSON Çıktıları

7.2 Teknik Akreditasyon

Teknolojik altyapıyı tanıtan servisler bu çalışmayı yapan kişinin kurucu ortağı olduğu Intelkit firması bünyesindeki (Unvan değişikliğinin ardından Datakapital A.Ş) duygu analizi çalışmalarında kullanılmıştır ve altyapı 2021 yılında Sanayi ve Ticaret bakanlığından Teknolojik Ürün Deneyim Belgesini alarak akredite olmuştur.⁵³

Tübitak projelerinden farklı olarak herhangi bir destek alınmadan öz kaynakla geliştirilen projeler TÜR belgesi ile akredite edilmektedir ve proje yine Tübitak'ın atadığı hakemler tarafından incelenmektedir. Verilerin sayımı ve analizi içinse PANAS şablonuna uygun olarak sıfırdan geliştirme yapılmıştır.

⁵³ Ek - 8

8.0 SONUÇ

Çalışma boyunca üç farklı yöntem için hipotez test edilmiştir.

1-)Bilgisayar destekli uygulama ile elde edilen veriler içinde betimleyici sıfatlar toplu bir şekilde saydırılarak PANAS-T skorları hesaplanmıştır. Bu yöntemde anket verileri ile sosyal medya verilerinin PANAS-T değişkenleri arasında yüksek düzeyde ($0,60 < \text{korelasyon}$) ilişki yoktur. İlgili yöntem için hipotez reddedilmiştir (H_0)

2-)Bilgisayar destekli uygulama ile elde edilen verilerde PANAS-T skorları her bir betimleyici sıfat için ayrı ayrı hesaplanarak ağırlıklandırılmıştır. Bu yöntemde anket verileri ile sosyal medya verilerinin PANAS-T değişkenleri arasında yüksek düzeyde ($0,60 < \text{korelasyon}$) ilişki yoktur. İlgili yöntem için hipotez reddedilmiştir (H_0)

3-) Betimleyici sıfatlar geliştirilen sentiment modülü ile filtrelenmiş ve sadece sentiment değeri PANAS-T sıfat grubundaki bağlamla uyumlu olan sıfatlar analize alınmıştır. Bu yöntemde anket verileri ile sosyal medya verileri arasındaki PANAS-T değişkenleri arasında yüksek düzeyde ($0,60 < \text{korelasyon}$) ilişki vardır. İlgili yöntem için hipotez kabul edilmiştir (H_1)

Elde edilen bulgular şu şekilde özetlenebilir:

- Günümüz teknolojileriyle anket vb. konvansiyonel araştırma yöntemleri ile ulaşılamayacak büyüklükte veri kümelerine ulaşmak ve bu verilerle konvansiyonel yöntemlere paralel sonuçlar elde etmek mümkündür.
- Türkçe dili için PANAS-T ölçeğiyle destekleyici sentiment filtreler kullanılmadan gündemler isabetli bir biçimde analiz edilemez ve konvansiyonel araştırma metotlarıyla benzer sonuçlara ulaşılamaz.
- Yine Türkçe dili için PANAS-T ölçeğindeki betimleyici sıfatlar kendi duygu grubundaki bağlama göre sentiment modülleriyle filtrelenerek bilgisayar destekli uygulamalarla, konvansiyonel araştırma yöntemlerine benzer sonuçlar elde edilebilmektedir

KAYNAKÇA

BURAL, Erol Başaran. Terörizmle Mücadelede Sosyal Medya İstihbaratının Önemi: Pkk Terör Örgütünün Twitter'daki Görünürlüğü Üzerine Sosyal Ağ Analizi (Yüksek Lisans Tezi Jandarma ve sahil güvenlik akademisi güvenlik bilimleri enstitüsü, Ankara 2021)

Kent, Sherman. Stratejik İstihbarat : ASAM Yayınları, 27.05.2005, ISBN :9789756769812, Çeviren : Arıca, NAZLIM ; Özbek, B.Yasemin

Tekin, Uğur. Açık Kaynaklardan Siber Tehdit İstihbaratı Verisi Elde Edilmesi, (Gazi Üniversitesi Fen Bilimleri Enstitüsü, Bilgi Güvenliği Mühendisliği Ana Bilim Dalı, Kasım 2021, Yüksek Lisans Tezi)

Laney, Doug 3D data management: Controlling data volume, variety and velocity (2001 MetaGroup research publication).

Özcan, Deniz Gül : Doğal dil işleme tekniklerinin incelenmesi ve seyahat-turizm sesli asistani için türkçe varlık ismi tanıma aracı geliştirilmesi.(Akdeniz Üniversitesi, Aralık 2020 , Yüksek Lisans tezi.)

Fur, Michael Psychometrics : An introduction, Sage Publications 4th edition.

Benevenuto, Fabricio ; Goncalves, Pollyanna; Cha, Meeyoung : PANAS-t: A Psychometric Scale for Measuring Sentiments on Twitter (Systems, man, and cybernetics, part c: Applications and reviews, vol. xx, no. xx, june 2013)

W. Kearney, Michael : School of Journalism, Informatics Institute, University of Missouri, The Journal of Open Source Software, 24 Ekim 2019

Kişisel Verilerin Korunması Kanunu No : 6698, Resmi Gazete sayı : 29677 7/4/2016

Dr. Al-Khour, Ali : Data Ownership: Who Own's My Data, Professor
(International Journal of Management & Information Technology ISSN: 2278-
5612 Volume 2, No 1, November, 2012)

Watson, David ; Clark, Lee Anna : Journal of Personality and Social Psychology
1988. Vol. 54. No. 6, 1063-1070, Development and Validation of Brief
Measures of Positive and Negative Affect: The PANAS Scales

Gençöz, Tülin : Pozitif ve Negatif Duygu Ölçeği Geçerlilik ve Güvenilirlik
Çalışması, (ODTÜ Türk Psikolojisi Dergisi 2000, 15(46))

Altıntaş, Mücahit : Kullanıcı destek sistemlerinde yardım biletlelerinin otomatik
sınıflandırılması (İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, Yüksek
Lisans tezi, Eylül 2014)

Önal, Muharrem : Sözlüksel benzeşim yöntemi ile android tabanlı duygu analizi
uygulaması, (Fırat Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi,
Mart 2017)

Bollen, J. ; Pepe, A. ;Mao, H. : Modeling public mood and emotion: Twitter
sentiment and socio-economic phenomena, in: Proceedings of the International
AAAI Conference on Weblogs and Social Media (ICWSM), 2011.

Domingos, Pedro ; Pazzani Michael : On the Optimality of the Simple Bayesian
Classifier under Zero-One Loss, Machine Learning, 29, 103–130 (1997) 1997
Kluwer Academic Publishers. Manufactured in The Netherlands.

Mihalcea, Rada Multilingual sentiment and subjectivity analysis, Department of
Computer Science University of North Texas, May 12, 2011, Academia

Watson, David ; Clark, Lee Anna THE PANAS-X Manual for the Positive and
Negative Affect Schedule - Expanded Form, 1994, The University of Iowa

Crawford, John R.; Henry, Julie D. (The positive and negative affect
schedule(panas): construct validity, measurement properties and normative data

in a largenon-clinical sample), British Journal of Clinical Psychology (2004), 43, 245–2652004 The British Psychological Societ. University of aberdeen, UK.

Ágoston Gyollai, Péter Simor, Ferenc Köteles, Zsolt Demetrovics ;
Psychometric properties of the Hungarian version of the original and the short
form of the Positive and Negative Affect Schedule (PANAS)
Neuropsychopharmacologia Hungarica 2011. XIII. évf. 2. Szám

Diener, E., & Emmons, R. A. (1984). The independence of positive and negative
affect. Journal of Personality and Social Psychology, 47(5), 1105–1117

Bernd Müller · Christoph Poley · Jana Pössel · Alexandra Hagelstein · Thomas
Gübitz LIVIVO – the Vertical Search Engine for Life Sciences, DOI
10.1007/s13222-016-0245-2 Datenbank Spektrum (2017) 17:29–34

Fabrizio Benevenuto, Matheus Araújo, Measuring sentiments in online social
networks, WebMedia '13, 2013

How Data-Driven Research Fuelled The Cambridge Analytica Controversy,
Annika Richterich, Maastricht University, The Open Journal of Sociopolitical
Studies, ISSN: 2035-6609, 15 Temmuz 2018

The enneagram as a psychographic segmentation tool, Naomi Ekas, Ph.D.,
William C. Moncrief, Ph.D., Texas Christian University, 2 May1s 2022

EKLER

Ek - 3: Stop Words - Etkisiz Kelimeler

a	acaba	altı	Ama	ancak	artık	az
b	bana	bazen	Bazı	belki	ben	beni
benim	beş	bile	Bir	birçoğu	birçok	birçokları
biri	birisi	birkaç	birkaçı	birşey	birşeyi	biz
bize	bizi	bizim	Böyle	böylece	bu	buna
bunda	bundan	bunu	Bunun	burada	bütün	c
ç	çoğu	çoğuna	çoğunu	çok	çünkü	de
da	daha	de	Demek	diğer	diğeri	diğerleri
diye	dokuz	dolayı	Dört	e	f	fakat
falan	felan	filan	G	gibi	ğ	h
hala	hangi	hangisi	Hem	henüz	hep	hepsi
hepsini	her	ı	İ	için	içinde	iki
ile	ise	işte	J	k	kaç	kadar
kendi	kendine	kendini	Ki	kim	kime	l
m	madem	mı	Mi	mu	mü	n
nasıl	ne	ne kadar	ne zaman	ne zaman	neden	nedir
nerde	nerede	nereden	Nereye	nesi	neyse	niçin
niye	o	on	Ona	ondan	onlar	onlara
onlardan	onların	onu	Onun	orada	oysa	oysaki
ö	öbürü	ön	Önce	ötürü	öyle	p
r	s	sana	Sekiz	sen	senden	seni
senin	siz	sizden	Size	sizi	son	sonra
ş	şayet	şey	şeyden	şimdi	şöyle	şu
şuna	şunlar	şunu	t	tabi	tamam	tüm
tümü	u	ü	üç	üzere	v	var
ve	veya	veyahut	y	ya	ya da	yani
yedi	yerine	yine	yoksa	z	zaten	zira

Ek - 4: PANAS – Betimleyici Anahtar Kelimeler

Betimleyici Sıfat	Betimleyici Anahtar Kelimeler			
<i>İlgili</i>	İlgili	İgilen		
<i>Sıkıntılı</i>	Sıkıntı	Sıkıl		
<i>Mutsuz</i>	Mutsuz			
<i>Heyecanlı</i>	Heyecan			
<i>Güçlü</i>	Güçlü			
<i>Ürkmüş</i>	Ürkmüş	Ürkme	Ürktüm	Ürkmüşüm
	Ürkerim	Ürküyorum	Ürkeceğim	
<i>İlhamlı</i>	İlham			
<i>Sinirli</i>	Sinir			
<i>Kararlı</i>	Kararlı			
<i>Dikkatli</i>	Dikkat			
<i>Tedirgin</i>	Tedirgin			
<i>Aktif</i>	Aktif			
<i>Korkmuş</i>	Kork			
<i>Hevesli</i>	Heves			
<i>Uyanık</i>	Uyanı	Uyana		
<i>Gururlu</i>	Gurur			
<i>Asabi</i>	Asabi			
<i>Utanmış</i>	Utanmış	Utañ	Utandım	Utanmışım
	Utanırım	Utanıyorum	Utanacağım	
<i>Suçlu</i>	Suçlu			
<i>Düşmanca</i>	Düşman			

EK-5 Gündemlerin İzole Hesaplanmış PANAS Değerleri

Ukrayna Rusya Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,71324	P (Ürkmüş)	-1,00000
P (İlgili)	-0,40854	P (Korkmuş)	-0,21854
P (Kararlı)	0,68289	P (Mutsuz)	-0,98046
P(Heyecanlı)	-0,91260	P (Sıkıntılı)	-0,36658
P (İlhamlı)	-0,59351	P (Asabi)	-1,00000
P (Uyanık)	-0,67390	P (Sinirli)	-0,41632
P (Aktif)	0,38370	P (Utanmış)	-0,07733
P (Güçlü)	0,32325	P (Suçlu)	0,43693
P (Gururlu)	-0,74013	P (Tedirgin)	0,42468
P (Dikkatli)	0,56459	P(Düşmanca)	0,05380
Ağr. Toplam	-0,20875	Ağr. Toplam	-0,31438

Recep Tayyip Erdoğan Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,60036	P (Ürkmüş)	-0,94952
P (İlgili)	-0,33344	P (Korkmuş)	-0,55363
P (Kararlı)	0,79653	P (Mutsuz)	-0,78967
P(Heyecanlı)	-0,01237	P (Sıkıntılı)	-0,85181
P (İlhamlı)	-0,19907	P (Asabi)	-0,89904
P (Uyanık)	-0,71316	P (Sinirli)	-0,70498
P (Aktif)	-0,52523	P (Utanmış)	-0,62927
P (Güçlü)	0,58527	P (Suçlu)	-0,08912
P (Gururlu)	0,36235	P (Tedirgin)	-0,48927
P (Dikkatli)	-0,52934	P(Düşmanca)	0,00155
Ağr. Toplam	-0,11688	Ağr. Toplam	-0,59547

Corona Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,74111	P (Ürkmüş)	-0,26520
P (İlgili)	-0,68129	P (Korkmuş)	0,33569
P (Kararlı)	-0,62478	P (Mutsuz)	-0,75507
P(Heyecanlı)	-0,42494	P (Sıkıntılı)	-0,09596
P (İlhamlı)	0,07210	P (Asabi)	-1,00000
P (Uyanık)	-0,35666	P (Sinirli)	0,17241
P (Aktif)	-0,23210	P (Utanmış)	-0,68681
P (Güçlü)	-0,55769	P (Suçlu)	-0,43858
P (Gururlu)	-0,68225	P (Tedirgin)	0,47419
P (Dikkatli)	0,52007	P(Düşmanca)	-0,57773
Ağr. Toplam	-0,37086	Ağr. Toplam	-0,28370

Orman Yangınları Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,93479	P (Ürkmüş)	-1,00000
P (İlgili)	0,34048	P (Korkmuş)	-0,41869
P (Kararlı)	0,25308	P (Mutsuz)	-0,44285
P(Heyecanlı)	-0,86321	P (Sıkıntılı)	-0,34285
P (İlhamlı)	-1,00000	P (Asabi)	-1,00000
P (Uyanık)	-0,37885	P (Sinirli)	0,01360
P (Aktif)	0,26500	P (Utanmış)	-0,43265
P (Güçlü)	-0,19053	P (Suçlu)	0,50372
P (Gururlu)	-0,61735	P (Tedirgin)	-0,25971
P (Dikkatli)	0,63405	P(Düşmanca)	-0,03650
Ağr. Toplam	-0,24921	Ağr. Toplam	-0,34159

Turkovac Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,85922	P (Ürkmüş)	0,26392
P (İlgili)	-0,41075	P (Korkmuş)	-0,66972
P (Kararlı)	-0,65314	P (Mutsuz)	-0,88679
P(Heyecanlı)	-0,79748	P (Sıkıntılı)	-0,95335
P (İlhamlı)	-1,00000	P (Asabi)	-1,00000
P (Uyanık)	-0,74345	P (Sinirli)	-0,75848
P (Aktif)	-0,70626	P (Utanmış)	-0,64366
P (Güçlü)	-0,24818	P (Suçlu)	-0,87788
P (Gururlu)	0,83393	P (Tedirgin)	-1,00000
P (Dikkatli)	0,78345	P(Düşmanca)	-0,44528
Ağr. Toplam	-0,38011	Ağr. Toplam	-0,69712

Enflasyon Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,89047	P (Ürkmüş)	-1,00000
P (İlgili)	-0,27957	P (Korkmuş)	-0,26042
P (Kararlı)	0,72504	P (Mutsuz)	-0,67598
P (Heyecanlı)	-0,85557	P (Sıkıntılı)	0,32124
P (İlhamlı)	-0,83893	P (Asabi)	-0,77350
P (Uyanık)	-0,30009	P (Sinirli)	-0,36580
P (Aktif)	-0,73881	P (Utanmış)	-0,55443
P (Güçlü)	-0,04122	P (Suçlu)	-0,26027
P (Gururlu)	-0,84492	P (Tedirgin)	0,24945
P (Dikkatli)	0,43711	P (Düşmanca)	-0,59581
Ağr. Toplam	-0,36274	Ağr. Toplam	-0,39155

Dolar TL Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,81317	P (Ürkmüş)	-0,59935
P (İlgili)	-0,49155	P (Korkmuş)	-0,48117
P (Kararlı)	-0,13051	P (Mutsuz)	-0,81637
P (Heyecanlı)	-0,80839	P (Sıkıntılı)	-0,37066
P (İlhamlı)	-0,94658	P (Asabi)	-0,89984
P (Uyanık)	-0,01644	P (Sinirli)	-0,39903
P (Aktif)	-0,61018	P (Utanmış)	-0,55994
P (Güçlü)	-0,25705	P (Suçlu)	-0,34726
P (Gururlu)	-0,77531	P (Tedirgin)	-0,17514
P (Dikkatli)	0,21122	P (Düşmanca)	-0,54709
Ağr. Toplam	-0,46380	Ağr. Toplam	-0,51959

Ramazan Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	-0,79940	P (Ürkmüş)	-0,56983
P (İlgili)	-0,59574	P (Korkmuş)	-0,60678
P (Kararlı)	-0,23118	P (Mutsuz)	-0,82972
P (Heyecanlı)	0,47137	P (Sıkıntılı)	-0,55936
P (İlhamlı)	0,25809	P (Asabi)	-0,56983
P (Uyanık)	-0,21350	P (Sinirli)	-0,08230
P (Aktif)	-0,55045	P (Utanmış)	-0,78139
P (Güçlü)	-0,59697	P (Suçlu)	-0,81150
P (Gururlu)	-0,65121	P (Tedirgin)	-0,67104
P (Dikkatli)	0,03291	P (Düşmanca)	-0,65196
Ağr. Toplam	-0,28761	Ağr. Toplam	-0,61337

Şampiyon Beşiktaş Gündemi İzole Hesaplama

Pozitif Sıfatlar	P(poz)	Negatif Sıfatlar	P(neg)
P (Hevesli)	0,64598	P (Ürkmüş)	-1,00000
P (İlgili)	-0,52886	P (Korkmuş)	-0,47433
P (Kararlı)	-0,77976	P (Mutsuz)	-0,80232
P (Heyecanlı)	0,32375	P (Sıkıntılı)	-0,79761
P (İlhamlı)	-0,88499	P (Asabi)	-1,00000
P (Uyanık)	-0,67421	P (Sinirli)	0,03393
P (Aktif)	-0,19179	P (Utanmış)	-0,40608
P (Güçlü)	-0,13740	P (Suçlu)	-0,78677
P (Gururlu)	0,34010	P (Tedirgin)	-0,39110
P (Dikkatli)	-0,62742	P Düşmanca)	-0,30861
Ağr. Toplam	-0,25146	Ağr. Toplam	-0,59329

EK – 6 Anket Soruları Örneği

Örnek PANAS ölçeği

	Hiç yansıtmıyor	Yansıtmıyor	Pek yansıtmıyor	Ne yansıtıyor ne yansıtmıyor	Kısmen yansıtıyor	Yansıtıyor	Tamamen yansıtıyor
1. İlgili	1	2	3	4	5	6	7
2. Sıkıntılı	1	2	3	4	5	6	7
3. Heyecanlı	1	2	3	4	5	6	7
4. Mutsuz	1	2	3	4	5	6	7
5. Güçlü	1	2	3	4	5	6	7
6. Suçlu	1	2	3	4	5	6	7
7. Ürkmüş	1	2	3	4	5	6	7
8. Düşmanca	1	2	3	4	5	6	7
9. Hevesli	1	2	3	4	5	6	7
10. Gururlu	1	2	3	4	5	6	7
11. Asabi	1	2	3	4	5	6	7
12. Uyanık/Tetikte	1	2	3	4	5	6	7
13. Utanmış	1	2	3	4	5	6	7
14. İlahımlı	1	2	3	4	5	6	7
15. Sinirli	1	2	3	4	5	6	7
16. Kararlı	1	2	3	4	5	6	7
17. Dikkatli	1	2	3	4	5	6	7
18. Tedirgin	1	2	3	4	5	6	7
19. Aktif	1	2	3	4	5	6	7
20. Korkmuş	1	2	3	4	5	6	7

54

Anket Soruları Örneği :

"Enflasyon" konusu/temasını sizin için hangi oranda pozitif/olumlu duygu durumu çağrıştırmaktadır.

Pozitif Duygudan Uzak 1 2 3 4 5 Pozitif Duyguya Yakın

"Enflasyon" konusu/temasını sizin için hangi oranda negatif/olumsuz duygu durumu çağrıştırmaktadır.

Negatif Duygudan Uzak 1 2 3 4 5 Negatif Duyguya Yakın

Bütün gündemler için aynı soru tekrarlanmıştır.

⁵⁴ Zafer ÇILGIN, "Olumsuz duygusal yük taşıyan uyarıların eşik altı ve eşik üstü sunumunun sigara kullanıcılarında nikotin stroop testi performansına etkileri" (Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Psikoloji Anabilim Dalı Deneyisel Psikoloji Bilim Dalı), Yüksek lisans tezi, Ankara 2015, 135

EK-7 Gündemler İçin Anketlerin Oylama Sonuçları

Gündem	Duygudan Uzak(1)	(2)	(3)	(4)	Duyguya Yakın(5)	PANAS(ort)
<i>Suriyeliler(Poz)</i>	66	110	133	115	76	0,0025
<i>Suriyeliler(Neg)</i>	73	112	123	109	83	0,0170
<i>RTE (Poz)</i>	50	47	79	99	225	0,4020
<i>RTE(Neg)</i>	219	105	80	44	52	-0,3950
<i>Corona(Poz)</i>	276	91	111	9	13	-0,6080
<i>Corona(Neg)</i>	19	15	119	77	270	0,5640
<i>Turkovac(Poz)</i>	41	28	124	104	203	0,4000
<i>Turkovac(Neg)</i>	216	87	136	22	39	-0,4190
<i>Ramazan(Poz)</i>	8	1	44	28	419	0,8490
<i>Ramazan(Neg)</i>	424	27	36	1	12	-0,8500
<i>Ş.Beşiktaş(Poz)</i>	167	43	214	13	63	-0,0238
<i>Ş.Beşiktaş(Neg)</i>	99	16	224	44	117	0,0640
<i>Dolar TL (Poz)</i>	239	135	88	19	19	-0,6870
<i>Dolar TL(Neg)</i>	23	11	87	120	259	0,5810
<i>Enflasyon(Poz)</i>	297	122	55	12	14	-0,6760
<i>Enflasyon(Neg)</i>	19	20	46	116	299	0,6560
<i>Ukr. Rus. (Poz)</i>	135	107	188	40	30	-0,2770
<i>Ukr. Rus. (Neg)</i>	36	32	185	105	142	0,2850
<i>Orman Yangınları(Poz)</i>	368	60	49	7	16	-0,7570
<i>Orman Yangınları(Neg)</i>	25	7	55	52	361	0,7170

EK-8 TÜR belgesi

 Türkiye Cumhuriyeti Sanayi ve Teknoloji Bakanlığı Milli Teknoloji Genel Müdürlüğü
TEKNOLOJİK ÜRÜN DENEYİM BELGESİ
Belge Numarası : TÜR 2021/1068
Bu belge, 4734 sayılı Kamu İhale Kanunu'nun 10.maddesi f bendine dayanılarak 26.8.2021 tarihinde tanzim edilmiş ve ürünün/hizmetin piyasaya arz tarihinden itibaren 5 yıl süre ile geçerli olmak üzere verilmiştir.
e-imzalıdır Zekeriya ÇOŞTU Genel Müdür

Belge Geçerlilik Tarihi : 1.2.2026 *Firma ve Proje bilgileri belgenin arka yüzündedir.
ZEKERİYA ÇOŞTU 26.08.2021

Kuruluş ve Proje Bilgileri

ÜNVANI	: İNTELKİT VERİ HİZMETLERİ BİLGİSAYAR PROGLAMLAMA VE DANIŞMANLIK HİZMETLERİ ANONİM ŞİRKETİ
VERGİ/TC KİMLİK NUMARASI	: 4581187852
SGK KURUM SİCİL NUMARASI	: 2-7022-0707-1451879-034-41-45/000
ADRESİ	: KÜÇÜKBAKKALKÖY MAH. DEFNE SK. FLORA RESIDANCE NO: 1 İÇ KAPI NO: 365 ATAŞEHİR/ İSTANBUL
BULUNDUĞU İLÇE/İL	: Ataşehir / İstanbul
TELEFON NUMARASI	: 2167097293
FAKS NUMARASI	:
E-POSTA ADRESİ	: info@intelkit.com
TİCARET SİCİL NUMARASI	: 191384-5
WEB ADRESİ	: intelkit.com
PROJENİN GERÇEKLEŞTİĞİ YER	: Şirket Būnyesinde
PROJENİN ADI	: İntelkit Doğal Dil İşleme Projesi
PROJENİN TAMAMLANMA TARİHİ	: 7.1.2021
PROJENİN SEKTÖRÜ (Nace Rev2)	: 63.11.08 Veri işleme, barındırma ve ilgili faaliyetler (veri girişi, verinin işlenmesi, özel raporların oluşturulması, depolanması, vb.)
ÜRÜN/HİZMETİN TİCARİ ADI	: 62.01.01 Bilgisayar programlama faaliyetleri (sistem, veri tabanı, network, web sayfası vb. yazılımları ile müşteriye özel yazılımların kodlanması vb)
ÜRÜN/HİZMETİN PİYASAYA ARZ TARİHİ	: İntelkit Analytics İş İstihbaratı Raporları : 1.2.2021

PANAS Duygu Durumu Frekansları Ek - 9

Recep Tayyip Erdoğan Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Hevesli	382	Ürkmüş	1
İlgili	822	Korkmuş	2.728
Kararlı	1.144	Mutsuz	200
Heyecanlı	1.575	Sıkıntılı	513
İlhamlı	119	Asabi	4
Uyanık	331	Sinirli	263
Aktif	261	Utanmış	224
Güçlü	4.920	Suçlu	803
Gururlu	2.299	Tedirgin	86
Dikkatli	993	Düşmanca	2.624
Toplam	12.846	Toplam	7.446

Orman Yangınları Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Hevesli	2	Ürkmüş	0
İlgili	60	Korkmuş	114
Kararlı	10	Mutsuz	17
Heyecanlı	7	Sıkıntılı	73
İlhamlı	0	Asabi	0
Uyanık	23	Sinirli	29
Aktif	24	Utanmış	11
Güçlü	53	Suçlu	57
Gururlu	18	Tedirgin	4
Dikkatli	185	Düşmanca	81
Toplam	382	Toplam	386

Enflasyon Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	35	Ürkmüş	0
İlgili	297	Korkmuş	1.511
Kararlı	283	Mutsuz	103
Heyecanlı	77	Sıkıntılı	1.705
İlhamlı	8	Asabi	3
Uyanık	270	Sinirli	189
Aktif	48	Utanmış	90
Güçlü	654	Suçlu	218
Gururlu	76	Tedirgin	75
Dikkatli	1.253	Düşmanca	354
Toplam	3.001	Toplam	4.248

Ukrayna Rusya Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	59	Ürkmüş	0
İlgili	157	Korkmuş	1.028
Kararlı	158	Mutsuz	4
Heyecanlı	30	Sıkıntılı	472
İlhamlı	13	Asabi	0
Uyanık	81	Sinirli	112
Aktif	192	Utanmış	120
Güçlü	649	Suçlu	337
Gururlu	82	Tedirgin	63
Dikkatli	1.043	Düşmanca	596
Toplam	2.464	Toplam	2.732

Corona Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	17	Ürkmüş	1
İlgili	27	Korkmuş	632
Kararlı	6	Mutsuz	16
Heyecanlı	63	Sıkıntılı	215
İlhamlı	11	Asabi	0
Uyanık	51	Sinirli	74
Aktif	29	Utanmış	13
Güçlü	62	Suçlu	34

Gururlu	32	Tedirgin	22
Dikkatli	302	Düşmanca	76
Toplam	600	Toplam	1.083

Turkovac Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	5	Ürkmüş	1
İlgili	27	Korkmuş	75
Kararlı	3	Mutsuz	4
Heyecanlı	12	Sıkıntılı	6
İlhamlı	0	Asabi	0
Uyanık	11	Sinirli	8
Aktif	6	Utanmış	8
Güçlü	57	Suçlu	4
Gururlu	328	Tedirgin	0
Dikkatli	362	Düşmanca	54
Toplam	811	Toplam	160

Dolar TL Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	45	Ürkmüş	2
İlgili	158	Korkmuş	799
Kararlı	51	Mutsuz	44
Heyecanlı	77	Sıkıntılı	549
İlhamlı	2	Asabi	1
Uyanık	286	Sinirli	135
Aktif	54	Utanmış	67
Güçlü	382	Suçlu	145
Gururlu	83	Tedirgin	35
Dikkatli	674	Düşmanca	299
Toplam	1.812	Toplam	2.076

Şampiyon Beşiktaş Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	158	Ürkmüş	0
İlgili	34	Korkmuş	188
Kararlı	3	Mutsuz	11
Heyecanlı	138	Sıkıntılı	41
İlhamlı	1	Asabi	0
Uyanık	22	Sinirli	54

Aktif	26	Utanmış	21
Güçlü	103	Suçlu	11
Gururlu	130	Tedirgin	6
Dikkatli	46	Düşmanca	106
Toplam	661	Toplam	438

Ramazan Gündemi

Pozitif Sıfatlar	Toplam	Negatif Sıfatlar	Toplam
Heves	45	Ürkmüş	2
İlgili	117	Korkmuş	564
Kararlı	42	Mutsuz	38
Heyecanlı	708	Sıkıntılı	358
İlhamlı	47	Asabi	4
Uyanık	213	Sinirli	192
Aktif	58	Utanmış	31
Güçlü	193	Suçlu	39
Gururlu	120	Tedirgin	13
Dikkatli	512	Düşmanca	214
Toplam	2.055	Toplam	1.455

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : HASKOL, Tan

Uyruğu : T.C

Eğitim

Derece

Üniversite ve Bölüm

Mezuniyet tarihi

Lisans

Dokuz Eylül Üniversitesi

26.08.2011

İktisadi ve İdari Bilimler Fakültesi

İşletme Bölümü

Yabancı Dil

İngilizce