

# Tarım hakkında atılan tweetlerin duygu analizi değerlendirilmesi

## Evaluating tweets about agriculture by using sentiment analysis

Ebru TEMİZHAN<sup>1\*</sup> , Mehmet MENDEŞ<sup>2</sup> 

<sup>1,2</sup>Çanakkale Onsekiz Mart Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü Biyometri ve Genetik ABD, Çanakkale, TÜRKİYE

<sup>1</sup><https://orcid.org/0000-0001-7068-0492>; <sup>2</sup><https://orcid.org/0000-0003-0109-9914>

### To cite this article:

Temizhan, E. & Mendeş, M. (2023). Tarım hakkında atılan tweetlerin duygu analizi değerlendirilmesi. Harran Tarım ve Gıda Bilimleri Dergisi, 27(3): 352-361.

DOI: 10.29050/harranziraat.1260123

### \*Address for Correspondence:

Ebru TEMİZHAN

e-mail:

ebrutemizhan@gmail.com

### Received Date:

04.03.2023

### Accepted Date:

24.08.2023

© Copyright 2018 by Harran University Faculty of Agriculture. Available on-line at [www.dergipark.gov.tr/harranziraat](http://www.dergipark.gov.tr/harranziraat)



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

### Öz

Bu çalışmada Twitter kullanıcılarının iyi tarım, organik tarım ve sürdürülebilir tarım hakkındaki duygu ve düşünceleri duygu analizi tekniğinden yararlanılarak analiz edilmiştir. Bu amaçla, tarımla ilgili üç farklı hashtag grubunda toplam 15984 tweet metin madenciliği tekniği ile değerlendirilmiştir. Yapılan duygu analizi sonucunda, tweet atan bireylerin bu tarım teknikleri hakkındaki duygu ve düşünceleri arasında anlamlı farklılıkların bulunduğu gözlenmiştir. Twitter kullanıcıları için en popüler ve en güvenilir tarım uygulamasının **İyi Tarım** uygulaması olduğu görülmüştür. Sonuç olarak, bütün tweetlere ait duygu analizi sonuçları değerlendirildiğinde Twitter kullanıcıları genel olarak tarım hakkında pozitif duygu ve düşüncelere sahip olduğu belirlenmiştir.

**Anahtar Kelimeler:** Metin madenciliği, Duygu skorları, Tweet, Tarım

### ABSTRACT

In this study, the emotions and thoughts of Twitter users about good, organic and sustainable agriculture practices were analyzed by using sentiment analysis technique. 15984 tweets which were analyzed by this purpose. Sentiment analysis results showed that there were significant differences between the f emotions and thoughts of the tweeting about these agricultural practices. It was observed that the most popular and reliable agricultural application for Twitter the users was the **Good Agriculture** practice. When all results of the sentiment analysis analysis was evaluated, it was determined that the Twitter users generally had positive emotions and thoughts about the agriculture practices.

**Key Words:** Text mining, Emotion score, Tweet, Agriculture

## Giriş

Son verilere göre araştırma, deneme ve projelerden elde verilerin %80'den fazlasının metin, resim, ses, video vb. biçiminde yapılandırılmamış formatta olduğu bildirilmiştir (Kulkarni ve Shivananda, 2019). Telefonda konuşurken, mesaj atarken, resim çekerken, sosyal medya kullanırken, tweet atarken, hikaye ve gönderi paylaşırken, e-ticaret üzerinden alışveriş yaparken ve diğer çeşitli faaliyetlerde

bulunurken kısacası hayatımızın her anında farkında olmadan sayamayacağımız boyutlarda her an veri üretmekteyiz. Üstelik bu verilerin çoğu da metin biçiminde oluşmaktadır (Kulkarni ve Shivananda, 2019). Metin madenciliği ya da başka bir deyişle metin analitiği, metinsel verilerden kaliteli ve işlevsel bilgiler elde etmek için izlenen metodoloji ve süreçtir. Metin madenciliği yapılandırılmamış metin verilerini yapılandırılmış ya da daha yapılandırılmış formlara ayrıştırmak için NLP, bilgi getirme (information retrieval) ve

makine öğrenmesi tekniklerini kullanarak bu verilerden kalıplar ve içgörüler üretmeyi içerir. Metin madenciliği bilgi keşfi, tanımlama ve tahmine yönelik analiz ihtiyaçları için metinden bilgi modellemesi ve çıkarımında kullanılan makine öğrenmesi, dilbilimsel ve istatistiksel teknikleri içermektedir. Metin madenciliği metin sınıflandırma, metin kümeleme, metin özetleme, duygu analizi, benzerlik analizi ve ilişki modelleme gibi bazı temel teknikler ve işlemlere sahiptir (Sarkar, 2016).

Bu araştırmanın asıl amaçlarından birisi özellikle metin madenciliği algoritmalarının tarımsal çalışmalarda etkin bir şekilde kullanılma imkanı olduğu hususunda bir farkındalık oluşturmak ve bu yöntemlerle ilişkili temel bilgiler vermektir. Bu çalışmada, Twitter kullanıcılarının

çeşitli tarım teknikleri hakkında ne düşündüğünü anlamak ve duygularını analiz etmek için metin madenciliğinin duygu analizi tekniğinden yararlanılacaktır.

## Materyal ve Metot

Araştırmanın materyalini 3 Kasım 2020 tarihinde Twitter API ile çekilen tarım uygulamaları ile ilgili 14857 adet İngilizce tweet oluşturmuştur (Çizelge 1). Bu tweetler iyi tarım (n=5263), organik tarım (n=2719) ve sürdürülebilir tarım (n=6875) olarak üç temel tarım uygulaması altında toplanmıştır. İstatistiksel analizlerde Python yazılım programı geliştirme ortamı olan Anaconda Navigator üzerinden Jupyter Notebook (6.0.3) arayüzü kullanılmıştır.

Çizelge 1. Twitter'dan çekilen tarım ile ilgili konu etiketleri ve örnek genişlikleri.

Table 1. Related to agriculture hashtags and sample sizes extracted from Twitter

Tweet konusu <i>Tweet subject</i>	Konu etiketleri <i>Hashtags</i>	Zaman <i>Time</i>	Etiket örnek genişliği <i>Hashtag sample size</i>	Toplam örnek genişliği <i>Total sample size</i>
İyi tarım <i>Good agriculture</i>	#goodagriculture	3.11.2020	3	5263
	Goodagriculture	3.11.2020	2182	
	#goodfarming	3.11.2020	4	
	Goodfarming	3.11.2020	3074	
Organik tarım <i>Organic agriculture</i>	#organicagriculture	3.11.2020	84	3846
	Organicagriculture	3.11.2020	934	
	#organicfarming	3.11.2020	767	
	Organicfarming	3.11.2020	934	
Sürdürülebilir tarım <i>Sustainable agriculture</i>	#sustainableagriculture	3.11.2020	355	6875
	sustainableagriculture	3.11.2020	2593	
	#sustainablefarming	3.11.2020	195	
	Sustainablefarming	3.11.2020	3732	

Bu çalışmada yöntem olarak kelime bulutu ve duygu analizi metin madenciliği teknikleri kullanılmıştır. Analizlerin gerçekleştirilmesinde re, numpy, pandas, tweepy, textblob, matplotlib.pyplot, wordcloud, better\_profanity, textblob.classifiers ve timeit Python kütüphaneleri kullanılmıştır.

### Veri ön işleme

Metin madenciliğinde veri ön işleme, analizlere geçmeden önce veri setinden istenmeyen bazı verilerin temizlenmesi işlemidir (Kaur ve Sharma, 2020). Tweetler için kullanılan veri ön işleme

adımları aşağıda verilmiştir. Veri ön işlemenin amacı, veri kalitesini iyileştirmek ve veriyi analizlere daha uygun hale getirmektir (Anonymous, 2023).

- Tüm tweetlerin küçük harflere dönüşümü
- Sinkaflı kelimelerin sansürlenmesi
- Noktalama işaretlerinin, sembollerin ve rakamların kaldırılması
- Stopwords'lerin kaldırılması
- Özel tweet işaretlerinin (@, # ve emoji) ve URL'lerin (https) kaldırılması

### Kelime bulutu

Kelime bulutu, en sık tekrarlanan kelimelerin resimli bir temsilidir (Kulkarni ve Shivananda, 2019). Kelime frekansı arttıkça resimdeki yazı tipi boyutu da artmaktadır. Kelime bulutu aynı zamanda metnin genel bir özetini oluşturur, okuyucunun metne ilk bakışıdır. Okuyucunun metinde dikkate alacağı anahtar kelimeleri taşımaktadır.

### Duygu analizi

Duygu analizi sosyal ağlar, ürün incelemeleri, haber makaleleri vb. gibi farklı veri kaynaklarından elde edilebilecek bir metinden ölçülebilen bir değer üretme sürecidir (Arumugam ve Shanmugamani, 2018). Bir cümle veya metindeki duygunun kişinin olumlu, olumsuz veya tarafsız görüşünü ve düşüncesinin yansıtılmasını sağlayarak (Arumugam ve Shanmugamani, 2018) bu duyguların sayısal bir ölçekte değerlendirilmesine imkan veren bir analiz yöntemidir. Duygu analizi, Python'un TextBlob kütüphanesi üzerinden gerçekleştirilmiştir. TextBlob, metin verilerini işlemek için kullanılan bir Python kütüphanesidir. Duygu analizi, sınıflandırma, çeviri vb. gibi birçok doğal dil işleme (NLP) yöntemini gerçekleştirmek için basit bir API kullanımı sağlandığından (Anonymous, 2014) duygu analizi için bu kütüphane tercih edilmiştir. Ayrıca hesaplanan doğruluk değerleri yeterli düzeyde açıklanan varyasyon içerdiğinden dolayı Textblob kütüphanesi duygu analizini gerçekleştirmede başarılı bir şekilde kullanılmıştır (Çizelge 2).

TextBlob kütüphanesinin çalışma prensibinde polarite ve subjectivite (öznellik) adı altında iki temel ölçüt vardır (Kulkarni ve Shivananda, 2019). Polarite [-1,1] arasında değerler alabilen (Kulkarni ve Shivananda, 2019) ve bir kişinin metni tanımlayan konu hakkında mutlu, mutsuz veya tarafsız olup olmadığını gösteren bir ölçüttür (Arumugam ve Shanmugamani, 2018). Mutlu için 1, mutsuz için -1 ve nötr için 0 gibi ayrı bir değer olarak da ölçülebilir (Arumugam ve Shanmugamani, 2018). Subjectivite ise [0,1] arasında değişen genellikle halk çoğunluğunun

düşüncesini temsil etmektedir, genel olgusal bir fikir değildir. (Kulkarni ve Shivananda, 2019).

Textblob kütüphanesinden hesaplanan polarite değerleri for döngüsünden yararlanılarak pozitif, negatif ve nötr olarak sınıflandırılmıştır (Şekil 1). Daha sonra bu sınıflandırılan duygular üzerinden train (0.70) ve test (0.30) olmak üzere veri seti ikiye bölünerek NaiveBayes classifier algoritması makine öğrenmesi yöntemiyle modelin doğruluk değeri hesaplanarak model performansı değerlendirilmiştir (Çizelge 2).

```
# X tweetler ve y duygu sınıfları olmak üzere
y=[]
X = df["tweet"]
for doc in X:
    score = TextBlob(doc).sentiment.polarity
    if score > 0:
        y.append("Positive")
    elif score < 0:
        y.append("Negative")
    else:
        y.append("Neutral")
```

Şekil 1. Duygu skorları için kullanılan sınıflandırma algoritması

Figure 1. The classification algorithm used for sentiment scores

### Araştırma Bulguları ve Tartışma

Tweetlerdeki duygu analizine ilişkin tanıtıcı istatistikler ve doğruluk değerleri Çizelge 2'de verilmiştir. İyi tarım, organik tarım ve sürdürülebilir tarım tweetlerine ait kelime bulutları Şekil 2'de, tüm tweetlere ait kelime bulutları ise Şekil 3'de verilmiştir. Kelime bulutları incelendiğinde iyi tarım kelime bulutunda iyi niyetler, çiftçi, eşit olarak tasarlamak, yeşil devrim, iyi, yeni, ileri sıçrama, daha iyi toplum gibi pozitif duygu ağırlıklı tweetler göze çarparken kötü niyetli, 40 milyon insan, açlıktan ölmek niyetinde gibi olumsuz duygulara sahip tweetleri de görmekteyiz. Organik tarım kelime bulutunda organik bahçe, bahçe geliştirmek, organik sebze, organik yiyecek, tarım yöntemi, iyi-destek, ihtiyaç, organik bir bakış, organik kullanmak, gelişmiş, sürdürmek, merkezi uygulama, sebzeler, meyveler, silahlı çiftçi, Hintli çiftçi, organiğe başlamak, sürdürülebilir tarım kelime bulutunda

ise iklim biyoçeşitliliği, yaygın tarım, parlamentoda oylanmış, Avrupa parlemantosunu, yeni 7 yıl, açık oturum, iklim felaketi, yeni oylanmış, yaygın 7 yıl, politika başlığı ve açık mektup gibi politika ağırlıklı kelime ve kelime grupları yüksek frekanslarda gözlenmiştir. Bütün tweetlere ait kelime bulutu incelendiğinde ise (Şekil 3) sürdürülebilir tarım, tarımsal politika, amaç, yaygın tarım, iklim biyoçeşitliliği, sürdürülebilir biyoçeşitlilik, iklim felaketi, Avrupa parlemantosunu, 7 yıl, politika başlığı, açık oturum gibi kelime ve kelime grupları ön plana çıkmaktadır. Böylece bu kelime bulutları Twitter kullanıcı yorumlarının bir ön izlemesini ya da ön değerlendirmesini oluşturmuştur.

İyi tarım, organik tarım ve sürdürülebilir tarım tweetlerine ait duygu analizi sonuçları (Çizelge 2 ve Şekil 4) incelendiğinde, pozitif duygular en çok iyi tarım (%92.11) ve organik tarım (%62.95) tweetlerinde görülürken, negatif duygular da en

çok sürdürülebilir tarım (%32.28) tweetlerinde gözlenmiştir. Nötr duyguların ise en çok organik tarım (%25.56) tweetlerinde olduğu gözlenmiştir. Bütün tweetler (Çizelge 2 ve Şekil 5) dikkate alındığında ise üç tarım uygulamasında da pozitif duygu ve düşüncelerin (% 66.83) çoğunluk oluşturduğu söylenebilir.

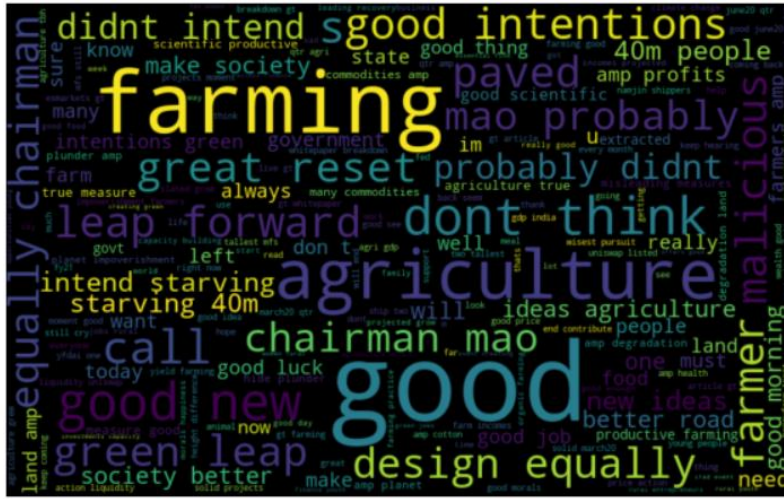
Polarite ve Subjectivite ölçütleri değerlendirildiğinde ise en yüksek polarite ve subjectivite ortalamaları iyi tarımda görülmüştür (Çizelge 2).

Yapılan duygu analizlerinin doğruluk değerleri (accuracy value) ise sırasıyla iyi tarım için % 93.10, organik tarım için % 71.75, sürdürülebilir tarım için % 77.22 ve tüm tarım teknikleri için % 92.35 olarak hesaplanmıştır (Çizelge 2). Ayrıca Python test program zamanları (execution time) sırasıyla iyi tarım için 1.67e-05 s, organik tarım için 1.71e-05 s, sürdürülebilir tarım için 1.77 e-05 s ve tüm tarım teknikleri için 2.03e-05 s sürmüştür.

Çizelge 2. Tweetlerdeki duygu analizine ilişkin tanıttıcı istatistikler ve doğruluk değerleri

Table 2. Descriptive statistics on sentiment analysis in tweets

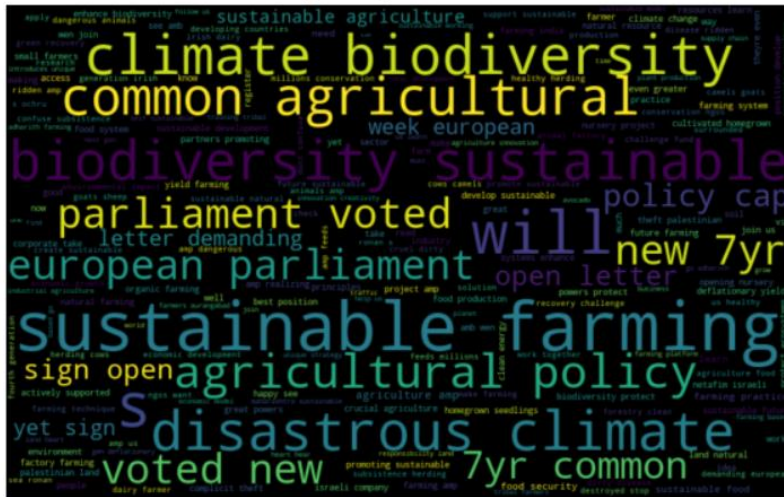
Duygular Sentiments	Pozitif tweet (n, %)	Nötr tweet (n, %)	Negatif tweet (n, %)	Polarite ortalaması ( $\bar{P}$ )	Subjectivite ortalaması ( $\bar{S}$ )	Doğruluk değeri (%)
Tweet konusu Tweet subject	Positive tweet (n, %)	Neutral tweet (n, %)	Negative tweet (n, %)	Polarity average ( $\bar{P}$ )	Subjectivity average ( $\bar{S}$ )	Accuracy value (%)
İyi tarım Good agriculture	4848 (%92.11)	70 (%1.33)	345 (%6.56)	0.34	0.52	93.10
Organik tarım Organic agriculture	2421 (%62.95)	983 (%25.56)	442 (%11.49)	0.17	0.39	71.75
Sürdürülebilir tarım Sustainable agriculture	3413 (%49.64)	1243 (%18.08)	2219 (%32.28)	0.08	0.41	77.22
Tarım Agriculture	10682 (%66.83)	2296 (%14.36)	3006 (%18.81)	0.19	0.44	92.35



Good Agriculture



Organic Agriculture



Sustainable Agriculture

Şekil 2. İyi tarım, organik tarım ve sürdürülebilir tarım tweetlerine ait kelime bulutları  
Figure 2. Word clouds of good agriculture, organic agriculture and sustainable agriculture tweets



Şekil 3. Bütün tweetlere ait kelime bulutu  
Figure 3. Word cloud of all tweets

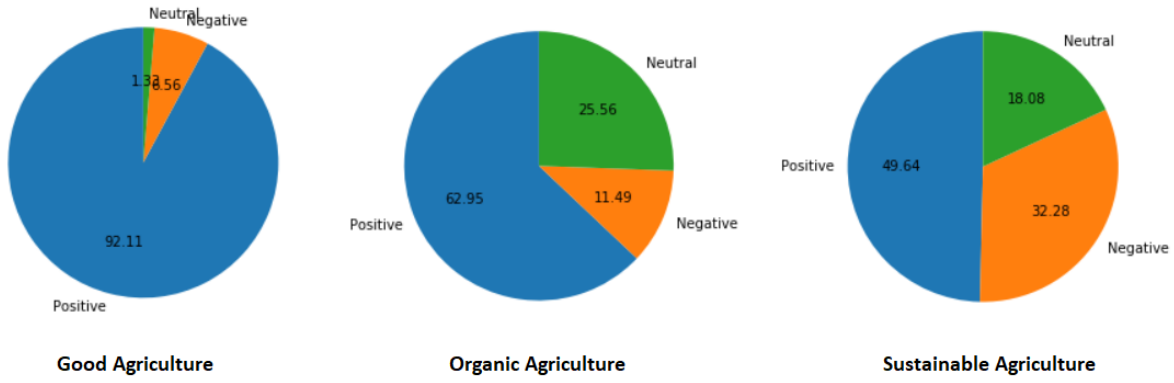
Yapılan literatür taramaları sonucunda bizim çalışmanın amacıyla en çok örtüşen bazı çalışmalar aşağıda verilmiştir.

Abiola ve diğerleri (2023), Nijerya'da koronavirüs pandemisine (COVID-19) verilen duygusal tepkileri analiz etmek için TextBlob ve VADER kütüphaneleri aracılığıyla 1 Milyon tweet üzerinden duygu analizini gerçekleştirmiştir. VADER ile yapılan duygu analizi sonucunda sırasıyla %39.8 pozitif, %28.9 negatif ve %31.3 nötr duygular içerirken, TextBlob ile yapılan duygu analizi sonucunda sırasıyla %36.7 pozitif, %17.3 negatif ve %46.0 nötr duygular içerdiği belirlenmiştir. Diğer bir deyişle halkın COVID-19 hakkındaki görüşleri VADER kütüphanesi çoğunluğun pozitif, TextBlob kütüphanesi ise çoğunluğun nötr duygular içerisinde olduğunu göstermiştir. Ayrıca Latent Dirichlet Allocation ile tweetlerin konu modellemesi yapılmış ve Multi Dimension ölçekleme ile görselleştirilmiştir. Sonuç olarak halkın önemli bir kısmı COVID-19'u kalıcı bir virüs olarak gördüğü, bazılarının ise sonunda bitebileceğine inandığı gözlenmiştir.

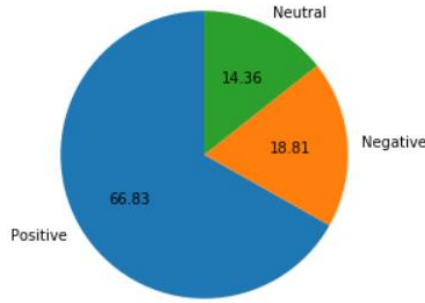
Barzenji (2021), çalışmasında Trump'a ait 41122 adet tweetin polarite değerlerini öğrenmek için duygu analizi yapmıştır. Duygu analizinin ilk aşamasında tweetlerin pozitif, negatif ve nötr olarak sınıflandırılması, ikinci aşamasında tweet

metni makine öğrenimi eğitim prosedürlerine uygun hale getirilmesi için doğal dil işleme tekniklerinden olan Bag of Words metin vektörleştirme yöntemi ile metin sayısal hale dönüştürülmüştür. Daha sonra duygu analizi, üç farklı makine öğrenmesi algoritması kullanılarak gerçekleştirilmiştir. Support Vector Machines ile %89, Random Forest ile %88 ve Gaussian Naive Bayes sınıflandırıcı ile %72 doğruluk değerine ulaşılmıştır. Barzenji (2021), diğer iki sınıflandırıcının yerine Random Forest'ı önermektedir, çünkü Random Forest iyi bir doğruluk değerine sahiptir ve Support Vector Machines 'e göre daha az zamana ihtiyaç duymaktadır.

Diyasa ve diğerleri (2021), TELKOM şirketinin indihome, myindihome, useetv ve wifi.id ürün ve hizmetleri hakkında Twitter kullanıcılarının duygu ve görüşlerini araştırmak için aynı anahtar kelimeleri içeren 3324 adet tweeti analiz etmiştir. Analizlerin yapılmasında çeşitli veri ön işleme teknikleri, duygu analizi, histogram, pasta grafiği ve kelime bulutu kullanılmıştır. Duygu analizi sonuçlarına göre incelenen 3324 tweet'in %34.4'ü pozitif, %16.1'i negatif ve %49.6'sı nötr tweet olduğu tespit edilmiştir. Yapılan duygu analizinin doğruluk değeri ise %77.2 olarak hesaplanmıştır.



Şekil 4. İyi tarım, organik tarım ve sürdürülebilir tarım tweetlerine ait duygu analizi sonuçları  
Figure 4. Sentiment analysis results of good agriculture, organic agriculture and sustainable agriculture tweets



Şekil 5. Bütün tweetlere ait duygu analizi sonuçları  
Figure 5. Sentiment analysis results of all tweets

Temizhan ve Mendes (2021), Covid-19 pandemisi ile ilgili Twitter üzerinden atılan 40 Bin tweeti metin madenciliği tekniği ile değerlendirerek, halkın bu salgına bakış açısını ve düşüncelerini değerlendirmiştir. #Covid\_19, #Maske, #bilimkurulu ve "Ankara Wuhan" hashtag'lerine ilişkin duygu analizi sonuçları incelendiğinde, duygu skorlarının birbirlerine benzer bir dağılım gösterdiğini ve özellikle negatif, korku, kızgınlık ve beklenti gibi duyguların Covid-19 sürecinde yaşanan baskın duyguları oluşturduğunu saptamıştır. Gündem başlıklarına ilişkin diğer bir analiz değerlendirmesi olan otomatik linear model sonuçları incelendiğinde ise Ankara gündemine ilişkin tahmin denklemi %97.20'lik bir açıklanan varyasyon ile istatistiksel önemliliğe sahip en iyi denklem olarak bulunmuştur. Ankara gündemine ait otomatik linear model denkleminde etki eden yatak, salgın, Wuhan, hastalık, sağlıkçı, maske, düğün, nişan, ihale ve Covid-19 değişkenleri halkın bu salgına bakış açısı ve düşünceleri hakkında özetleyici bir bilgi üretmiştir.

Kaur ve Sharma (2020), kadınların sosyal

sorunlarına ilişkin kullanıcı görüşlerini duygu analizi yapan bir model oluşturmayı amaçlamıştır. Veri kümeleri, Twitter üzerinden #Women ve #Metoo olarak iki kadın hashtag'inde toplanmıştır. Araştırmada hangi hashtag'in daha popüler olduğunu belirlemek ve çeşitli makine öğrenmesi algoritmalarıyla sınıflandırmak için tweetler Python textblob kütüphanesi aracılığıyla polarite ölçütleri belirlenmiştir. Modeli eğitmek ve test etmek için Naive Bayes, Destek Vektör Makinesi, Rastgele Orman ve Lojistik Regresyon makine öğrenmesi algoritmaları kullanılarak elde edilen sonuçlar Precision, Recall, F-score ve Accuracy test parametreleri ile modellerin ve TextBlob kütüphanesinin doğruluğu karşılaştırılmıştır. Elde edilen sonuçlara göre en popüler hashtag #Women bulunurken, en iyi ve en verimli algoritma Destek Vektör Makinesi olarak bulunmuştur.

Bonta ve diğerleri (2019), www.rottentomatoes web sitesinden indirilen film incelemelerini sınıflandırmak için NLTK, Textblob ve VADER duygu analizi kütüphanelerini kullanmıştır. Çalışmanın sonucunda, VADER

(doğruluk değeri = %77) Textblob (doğruluk değeri = %74)'dan daha iyi performans gösterdiği doğrulanmıştır.

Nausheen ve Begum (2018), 2016 ABD başkanlık seçimlerindeki en iyi seçim adaylarından Donald Trump, Hillary Clinton ve Bernie Sanders hakkında toplanan tweetler üzerinden seçim kamuoyu duygularını belirlemek için sözlük tabanlı bir yaklaşım kullanmıştır. Toplanan tweet'ler üzerinden polarite ve öznellik ölçütleri belirlenerek adaylar arasında duygu türü üzerinden bir karşılaştırma yapılmıştır ve kelime bulutları oluşturulmuştur. Böylece her bir aday için halkın görüş ve duyguları belirlenerek, ülkenin potansiyel lideri hakkında ön bir değerlendirme yapılmıştır.

Son yıllarda yapılan duygu analizi araştırmaları incelendiğinde, hangi alanda olursa olsun merak edilen konu hakkında ya da ürün ve hizmet değerlendirmelerinde Twitter kullanıcılarına ait yorumlar üzerinden duygu skorları hesaplanmıştır. Bu duygu skorları Twitter kullanıcılarının duygu ve düşünceleri hakkında bize polarite ölçütlerinin değerlendirilmesini sağlamaktadır. Nitekim Barzenji (2021), Diyasa ve diğerleri (2021), Temizhan ve Mendeş (2021), Kaur ve Sharma (2020) ve Nausheen ve Begum (2018) çalışmalarında çeşitli metin madenciliği yöntemleri kullanarak duygu skorlarını değerlendirmiştir. Ayrıca Temizhan ve Mendeş (2021) hashtag gruplarına ait oluşturdukları linear denklemler üzerinden istatistiksel önemliliğe sahip en iyi denklemi belirleyerek Twitter kullanıcılarının hashtag hakkındaki bakış açısı ve düşünceleri hakkında özetleyici bilgi üretilmesini sağlamıştır. Diyasa ve diğerleri (2021) ve Kaur ve Sharma (2020) duygu skorları üzerinden bir takım sınıflandırma algoritma ya da algoritmaları kullanarak çeşitli test parametreleri hesaplamıştır. Bu test parametreleri polarite sınıflandırma oranları üzerinden modelin doğruluk derecesi ve birçok test değeri hakkında algoritma karşılaştırılmasında kullanılmıştır. Barzenji (2021) ise duygu analizini iki aşamada gerçekleştirerek farklı bir yaklaşım kullanmıştır. Tweetlerin polarite değerlerinin sınıflandırılmasından sonra

kullanacağı makine öğrenmesi yöntemleri için metin vektörleştirme yöntemi kullanmıştır. Böylece tamamen sayısal hale gelen metin üzerinden kolay bir şekilde model eğitim süreci tamamlanmıştır. Hem doğruluk değeri hem de zaman dikkate alınarak Random Forest algoritması önerilmektedir. Nausheen ve Begum (2018) ise duygu skorlarını değerlendirmede polarite ve öznellik ölçütlerini kullanarak hashtagler hakkında değerlendirmelerde bulunmuştur. Ayrıca bu araştırmacılar hashtaglere ait kelime bulutları oluşturularak dikkate alınması gereken kelime ya da kelime gruplarını gözlemlemiştir. Abiola ve diğerleri (2023) ve Bonta ve diğerleri (2019), çalışmalarında Textblob ve VADER duygu analizi kütüphanelerini kullanarak bu iki kütüphanenin performansını karşılaştırmıştır. Sonuç olarak Abiola ve diğerleri (2023) 1 Milyon veride çalıştığı için iki kütüphane arasında belirgin sonuçlar elde ederken, Bonta ve diğerleri (2019) her ne kadar VADER'in performansını (doğruluk değeri = %77) Textblob'dan (doğruluk değeri = %74) üstün olarak kabul etse de, 11861 adet veride çalıştığı için iki kütüphane arasındaki doğruluk değerlerinde dikkate değer belirgin bir fark gözlenmemiştir. Diğer bir deyişle, iki kütüphane arasındaki fark veri sayısı çok çok büyük miktarlara ulaştığında değiştiği söylenebilir. Hangi örnek genişliğine kadar bu kütüphaneler birbirleri için alternatif olabilir? Bunu ileri zamanlarda yapılabilecek kapsamlı çalışmalarda görmemiz mümkün olacaktır.

Biz de bu çalışmada önceki yaptığımız çalışmadan (Temizhan ve Mendeş, 2021) farklı olarak hashtag gruplarına ait duygu skorlarının değerlendirilmesinde, polarite değerlerini dikkate alarak, NaiveBayes classifier algoritması ile hesaplanan %92.35'lik model doğruluk değeri ile Twitter kullanıcılarının duygu ve düşünceleri arasında anlamlı farklılıklar bulunduğunu tespit ettik. Ayrıca bu farklılıklar kelime bulutlarında da gözlemlenmiştir. Sonuç olarak, Twitter kullanıcıları için en popüler ve en güvenilir tarım uygulamasının %93.10'luk doğruluk değeri ile **İyi Tarım** uygulaması olduğu görüldü. Bütün tweetlere ait duygu analizi sonuçları



değerlendirildiğinde Twitter kullanıcıları genel olarak tarım hakkında pozitif duygu ve düşüncelere sahip olduğu belirlenmiştir.

## Sonuçlar

Duygu analizi, tarım teknikleri tweetlerine ait duygular ve düşünceler arasındaki farklılığı göstermiştir. Tweetler iyi tarım ve organik tarım uygulamaları hakkında daha pozitif bir düşünce hakimiyetindedir. Aynı zamanda nötr duyguların ağırlık kazandığı organik tarım uygulaması hakkında belirsiz duyguların arttığı söylenebilir. Sürdürülebilir tarım duyguları incelendiğinde ise pozitif duyguların en aza indiğini ve negatif duyguların en fazla artış gösterdiğini görmekteyiz. Başka bir deyişle sürdürülebilir tarım uygulamasında tweetlerin yarısı pozitif duygulardan oluşuyorken diğer yarısı da negatif ve nötr duygulardan oluşmuştur. Sürdürülebilir tarım tweetlerindeki bu negatif duyguların baskın oluşu, sürdürülebilir tarıma ait kelime bulutu incelendiğinde politik duyguların ağırlıklı frekansta oluşuna atfedilebilir. Ayrıca iyi tarım uygulaması diğer iki tarım uygulamasına göre daha fazla pozitif tweet içerdiğinden ve hesaplanan %93.10'luk doğruluk değeri ile Twitter kullanıcıları üzerinde en popüler ve en güvenilir tarım uygulaması olmuştur.

Sonuç olarak, bütün tweetlere ait duygu analizi sonuçları incelendiğinde hesaplanan %92.35'lik model doğruluk değeri ile Twitter kullanıcılarının tarım hakkında pozitif duygu ve düşüncelere sahip olduğu belirlenmiştir.

Günümüzde duygu analizinde kullanılan birçok yaklaşım ve algoritma vardır hala da birçoğu geliştirilmektedir. Bundan sonraki çalışmalarda geliştirilmiş farklı yaklaşım ve algoritmaların farklı veri setlerinde kapsamlı bir şekilde karşılaştırılmasının oldukça yararlı olacağına inanıyoruz.

## Ekler

Bu araştırma, Prof. Dr. Mehmet MENDEŞ danışmanlığında yürütülen Ebru TEMİZHAN'ın

doktora tezinin bir kısmından üretilmiştir.

Prof. Dr. Mehmet MENDEŞ danışmanlığında yürütülen Ebru TEMİZHAN'ın doktora tezi 2018 ile 2022 yılları arasında YÖK 100/2000 Doktora Burs Programı "Gıda Teknolojileri: Organik Tarım" öncelikli tez alanında desteklenmiştir. 2022'den bu yana TÜBİTAK 2211-C Öncelikli Alanlara Yönelik Yurt İçi Doktora Burs Programı kapsamında "Büyük Veri ve Veri Analitiği" tez alanında desteklenmeye devam etmektedir.

**Çıkar Çatışması Beyanı:** Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

**Yazar Katkısı:** Prof. Dr. Mehmet MENDEŞ doktora tez danışmanı olup, araştırmanın planlanması, elde edilen sonuçların yorumlanması ve makalenin son halinin baştan sona incelenmesi aşamalarında katkı sağlarken, Ebru TEMİZHAN ise araştırma için gerekli literatür taramasının yapılması, duygu analizinin yapılması ve makalenin yazılmasına katkı sağlamıştır.

## Kaynaklar

- Abiola O., Alli A.A., Tale O.A., Misra S. and Alli O.A. (2023). Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. *Journal of Electrical Systems and Inf Technol* (2023) 10:5. <https://doi.org/10.1186/s43067-023-00070-9>
- Anonymous (2023). Data Preprocessing in Data Mining (2023). Retrieved from: <https://www.geeksforgeeks.org/data-preprocessing-in-data-mining/>
- Anonymous (2014). Project description. TextBlob: Simplified Text Processing. [https://pypi.org/project/textblob/0.9.0/#:~:text=TextBlob%20is%20a%20Python%20\(2,classification%2C%20translation%2C%20and%20more.](https://pypi.org/project/textblob/0.9.0/#:~:text=TextBlob%20is%20a%20Python%20(2,classification%2C%20translation%2C%20and%20more.)
- Arumugam R., Shanmugamani R. (2018). Hands-On Natural Language Processing with Python. Packt Publishing Ltd. Birmingham B3 2PB, UK. ISBN 978-1-78913-949-5.
- Barzenji H.S.A. (2021). Sentiment Analysis of Twitter Texts Using Machine Learning Algorithms. *Academic Platform Journal of Engineering and Science* 2021; 9-3, 460-471. Doi: 10.21541/apjes.939338
- Bonta V., Kumares N., Janardhan N.(2019). A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis. *Asian Journal of Computer Science and Technology*. ISSN: 2249-0701 Vol.8

- No.S2, 2019, pp. 1-6.
- Diyasa IGSM., Mandenni NMIM., Fachrurrozi MI., Pradika SI., Manab KRN., Sasmita NR. (2021). Twitter Sentiment Analysis as an Evaluation and Service Base On Python Textblob. Workshop on Environmental Science, Society, and Technology (WESTECH 2020). IOP Conf. Series: Materials Science and Engineering 1125 (2021) 012034. DOI: 10.1088/1757-899X/1125/1/012034.
- Kaur C., Sharma A. (2020). Social Issues Sentiment Analysis using Python. 2020 5th International Conference on Computing, Communication and Security (ICCCS), 14-16 October 2020, Patna, India. DOI: 10.1109/ICCCS49678.2020.9277251.
- Kulkarni A., Shivananda A. (2019). Natural Language Processing Recipes, Unlocking Text Data with Machine Learning and Deep Learning using Python. ISBN-13 (pbk): 978-1-4842-4266-7 ISBN-13 (electronic): 978-1-4842-4267-4. <https://doi.org/10.1007/978-1-4842-4267-4>.
- Nausheen F., Begum S.H. (2018). Sentiment Analysis to Predict Election Results Using Python. Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018), 19-20 January 2018, Coimbatore, India, (pp. 1259-1262). DOI: 10.1109/ICISC.2018.8399007.
- Sarkar D. (2016). Text analytics with Python. A Practical Real-World Approach to Gaining Actionable Insights from Your Data. pp. 49.
- Temizhan E., Mendeş M. (2021). COVID-19 Pandemisi ile İlgili Twitter Mesajlarının Metin Madenciliği Tekniği İle Değerlendirilmesi. *Türkiye Klinikleri J Biostat.* 2021;13(2):185-200. DOI: 10.5336/biostatic.2020-79992.