

T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

COĞRAFİ REFERANSLI SOSYAL MEDYA  
VERİLERİNDEN DOĞAL DİL İŞLEME VE  
DERİN ÖĞRENME KULLANILARAK  
DUYGU ANALİZİ

**Dilan Gözdem DOLU**

YÜKSEK LİSANS TEZİ

Harita Mühendisliği Anabilim Dalı

Uzaktan Algılama ve CBS Programı

Danışman

Doç. Dr. Alper ŞEN

Temmuz, 2025

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

COĞRAFİ REFERANSLI SOSYAL MEDYA VERİLERİNDEN

DOĞAL DİL İŞLEME VE DERİN ÖĞRENME

KULLANILARAK DUYGU ANALİZİ

Dilan Gözdem DOLU tarafından hazırlanan tez çalışması 04.07.2025 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı, Uzaktan Algılama ve CBS Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Alper ŞEN  
Yıldız Teknik Üniversitesi  
Danışman

**Jüri Üyeleri**

Doç. Dr. Alper ŞEN, Danışman  
Yıldız Teknik Üniversitesi

Prof. Dr. Ali Melih BAŞARANER, Üye  
Yıldız Teknik Üniversitesi

Prof. Dr. Nesibe Necla ULUĞTEKİN, Üye  
İstanbul Teknik Üniversitesi

Danışmanım Doç. Dr. Alper ŞEN sorumluluğunda tarafımda hazırlanan “Coğrafi Referanslı Sosyal Medya Verilerinden Doğal Dil İşleme ve Derin Öğrenme Kullanılarak Duygu Analizi” başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Dilan Gözdem DOLU

İmza



*Aileme*

## TEŐEKKÜR

---

Lisansüstü eğitimim ve çalışmalarım boyunca değerli tavsiyeleri ile her konuda yardım eden, bilgi ve tecrübesini aktaran, bana yol gösteren ve tüm sorularımı cevaplayan danışman hocam Sayın Doç. Dr. Alper ŐEN'e teşekkürlerimi sunarım.

Çalışmam süresince her daim yanımda olan, maddi ve manevi olarak desteklerini hiçbir zaman benden esirgemeyen aileme teşekkür ederim.

Dilan Gözdem DOLU

# İÇİNDEKİLER

---

<b>SİMGE LİSTESİ</b>	<b>viii</b>
<b>KISALTMA LİSTESİ</b>	<b>ix</b>
<b>ŞEKİL LİSTESİ</b>	<b>x</b>
<b>TABLO LİSTESİ</b>	<b>xi</b>
<b>HARİTA LİSTESİ</b>	<b>xii</b>
<b>ÖZET</b>	<b>xiii</b>
<b>ABSTRACT</b>	<b>xv</b>
<b>1 GİRİŞ</b>	<b>1</b>
1.1 Literatür Özeti .....	1
1.2 Tezin Amacı .....	8
1.3 Hipotez .....	9
<b>2 VERİ KAYNAĞI</b>	<b>10</b>
2.1 Veri Kazıma .....	10
2.2 İşaretleme Dili .....	11
2.2.1 Metin Tabanlı İşaretleme Dili .....	12
2.3 Veri Kazımda Kullanılan Python Kütüphaneleri .....	12
2.3.1 BeautifulSoup Kütüphanesi .....	12
2.3.2 Selenium Kütüphanesi .....	13
<b>3 DERİN ÖĞRENME VE SİNİR AĞLARI</b>	<b>15</b>
3.1 Derin Öğrenme .....	15
3.2 Yapay Sinir Ağları .....	15

3.2.1 Uzun Kısa Süreli Bellek (Long Short-Term Memory, LSTM) Derin Ağ Modeli.....	17
3.2.2 Kapı Yinelemeli Birimler (Gated Recurrent Units, GRU) Derin Ağ Modeli.....	18
<b>4 DOĞAL DİL İŞLEME</b>	<b>19</b>
4.1 Doğal Dil İşleme .....	19
4.1.1 Doğal Dil Araç Takımı (Natural Language Toolkit/NLTK) Kütüphanesi.....	20
4.2 Duygu Analizi .....	20
<b>5 MODEL DOĞRULUK METRİKLERİ</b>	<b>23</b>
5.1 Model Doğruluk Değerlendirmesi .....	23
5.2 Çapraz Doğrulama Yöntemi .....	24
<b>6 MEKÂNSAL ANALİZ</b>	<b>26</b>
6.1 Mekânsal Analiz .....	26
6.2 Sıcak Nokta Analizi .....	26
<b>7 UYGULAMA</b>	<b>28</b>
7.1 Veri.....	28
7.1.1 Çalışma Alanı.....	28
7.1.2 Sosyal Medyadan Veri Kazıma.....	30
7.2 İki Etiketli Duygu Analizi.....	33
7.2.1 1. Adım: Duygu Analizi Modeli Oluşturma-Eğitme-Test Etme....	33
7.2.2 2. Adım: Duygu Analizinde Referans ve Tahmin Verileri Üzerinden Sıcak Nokta Analizi.....	35
7.3 Dört Etiketli Duygu Analizi .....	36
7.3.1 1. Adım: Duygu Analizi Modeli Oluşturma-Eğitme-Test Etme....	36
7.3.2 2. Adım: Duygu Analizinde Referans ve Tahmin Verileri Üzerinden Sıcak Nokta Analizi.....	37

<b>8 BULGULAR</b>	<b>38</b>
<b>9 SONUÇ</b>	<b>57</b>
<b>KAYNAKÇA</b>	<b>61</b>
<b>TEZDEN ÜRETİLMİŞ YAYINLAR</b>	<b>66</b>



## SİMGE LİSTESİ

---

$\alpha$	Anlamlılık düzeyi
$\underline{X}$	Değişken ortalama değeri
$h_{t-1}^j$	GRU aday aktivasyon
$U$	GRU ağırlık matrisi
$W$	GRU ağırlık matrisi
$W_r$	GRU ağırlık matrisi
$W_z$	GRU ağırlık matrisi
$\sigma$	GRU aktivasyon fonksiyonu
$x_t$	GRU giriş vektörü
$z_t^j$	GRU güncelleme geçidi vektörü
$h_{t-1}^j$	GRU önceki aktivasyon
$h_t^j$	GRU önceki aktivasyon ile aday aktivasyon arasında t anındaki çıktı vektörü
$r_t$	GRU sıfırlama geçidi
$x_j$	J noktasına ait değişken değeri
$w_{i,j}$	Mekânsal ağırlık değeri
$n$	Örnekleme bulunan nokta sayısı
$\odot$	Vektörler arası elementer çarpım
$G_i^*$	Z-istatistiği değeri

## KISALTMA LİSTESİ

---

API	Application Programming Interface (Uygulama Programlama Arabirimi)
CBS	Coğrafi Bilgi Sistemi
DN	Doğru Negatif
DP	Doğru Pozitif
DVM	Destek Vektör Makineleri (SVM, Support Vector Machine)
EFCM	Extended Fuzzy C-means (Genişletilmiş Bulanık C-ortalamaları)
GPS	Global Positioning System (Küresel Konumlama Sistemi)
GRU	Gated Recurrent Unit (Kapı Yinelemeli Birimler)
GTD	Geo Tweets Downloader (Coğrafi İleti İndiricisi)
HTML	HyperText Markup Language (Hiper Metin İşaretleme Dili)
KNN	K-Nearest Neighbors (K-En Yakın Komşu)
LSTM	Long Short Term Memory (Uzun Kısa Süreli Bellek)
NLP	Natural Language Processing (Doğal Dil İşleme)
NLTK	Natural Language Toolkit (Doğal Dil Araç Takımı)
PHP	Hypertext Preprocessor (Üstünyazı Önışlemcisi)
RNN	Recurrent Neural Network (Yinelenen Sinir Ağları)
SGML	Standard Generalized Markup Language (Standart Genelleştirilmiş İşaretleme Dili)
TF-IDF	Term Frequency-Inverse Document Frequency
TÜİK	Türkiye İstatistik Kurumu
YN	Yanlış Negatif
YP	Yanlış Pozitif
XML	Extensible Markup Language (Genişletilebilir İşaretleme Dili)

## ŞEKİL LİSTESİ

---

Şekil 3.1 Yapay sinir ağları .....	17
Şekil 7.1 Çalışma alanı .....	29
Şekil 7.2 X geliştirici kontrol paneli (URL-10).....	31
Şekil 7.3 Veri kazıma kod akış diyagramı.....	32
Şekil 7.4 İki etiketli duygu analizine ait kod akış diyagramı .....	35
Şekil 7.5 Dört etiketli duygu analizine ait kod akış diyagramı .....	36
Şekil 9.1 ArcGIS dashboards ile oluşturulan duygu analizi rapor ekranı .....	60

## TABLO LİSTESİ

---

<b>Tablo 7.1</b> Çalışma alanının nüfus ve yüz ölçümü bilgileri .....	<b>28</b>
<b>Tablo 7.2</b> Gönderiye ait özellikler ve açıklamaları .....	<b>33</b>
<b>Tablo 8.1</b> İki etiketli derin öğrenme ağında test sonucu modelin performansını değerlendiren hata matrisi.....	<b>40</b>
<b>Tablo 8.2</b> İki etiketli derin öğrenme ağında test sonucu model performansını değerlendiren metrik sonuçları .....	<b>40</b>
<b>Tablo 8.3</b> İki etiketli derin öğrenme ağının örnek ilçe veri setleri üzerinde tahmin performansını değerlendiren metrik sonuçları .....	<b>41</b>
<b>Tablo 8.4</b> İl bazında ileti sayıları.....	<b>42</b>
<b>Tablo 8.5</b> Dört etiketli derin öğrenme ağında k-kat çapraz doğrulama katmanına göre test sonucu modeli performansını değerlendiren hata matrisi ....	<b>46</b>
<b>Tablo 8.6</b> Dört etiketli derin öğrenme ağında test sonucu model performansını değerlendiren genel metrik sonuçları.....	<b>46</b>
<b>Tablo 8.7</b> Dört etiketli derin öğrenme ağının örnek ilçe veri setleri üzerinde tahmin performansını değerlendiren metrik sonuçları .....	<b>47</b>
<b>Tablo 8.8</b> Dört etiketli ve iki etiketli model karşılaştırması.....	<b>47</b>

## HARİTA LİSTESİ

---

<b>Harita 8.1</b>	6 Şubat 2023 depreminden etkilenen illerin nüfusa oranla X uygulamasındaki depremle ilgili ileti sayıları.....	<b>42</b>
<b>Harita 8.2</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen olumsuz ileti sayıları ile ilgili sıcak nokta analizi (referans ve tahmin haritaları aynıdır) ....	<b>44</b>
<b>Harita 8.3</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen empati sınıfı tahminleri ile ilgili sıcak nokta analizi.....	<b>49</b>
<b>Harita 8.4</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen empati sınıfı gerçek iletiler ile ilgili sıcak nokta analizi.....	<b>50</b>
<b>Harita 8.5</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen öfke sınıfı gerçek iletiler ile ilgili sıcak nokta analizi (referans ve tahmin haritaları aynıdır).....	<b>51</b>
<b>Harita 8.6</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen üzüntü sınıfı tahmini iletiler ile ilgili sıcak nokta analizi.....	<b>52</b>
<b>Harita 8.7</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen üzüntü sınıfı gerçek iletiler ile ilgili sıcak nokta analizi.....	<b>52</b>
<b>Harita 8.8</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen korku sınıfı tahmini iletiler ile ilgili sıcak nokta analizi.....	<b>53</b>
<b>Harita 8.9</b>	6 Şubat 2023 depreminden etkilenen illerden edinilen korku sınıfı gerçek iletiler ile ilgili sıcak nokta analizi.....	<b>54</b>

# Coğrafi Referanslı Sosyal Medya Verilerinden Doğal Dil İşleme ve Derin Öğrenme Kullanılarak Duygu Analizi

Dilan Gözdem DOLU

Harita Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Danışman: Doç. Dr. Alper ŞEN

Doğal afet yönetiminde afet öncesi hazırlık, afet anı müdahale ve afet sonrası iyileştirme aşamalarında Coğrafi Bilgi Sistemleri (CBS) kullanılarak uygulanan yönetim biçimleri ile doğal afetlerin yol açtığı zararlar en aza indirgenebilmektedir. Bu çalışmanın amacı, 6 Şubat 2023 tarihli Kahramanmaraş depremleri ile ilgili X (eskiden bilinen adıyla Twitter) sosyal medya verilerinin Python programlama dilinde Selenium ve BeautifulSoup kütüphaneleri kullanılarak kazınması ve doğal dil işleme ve derin öğrenme yöntemleri kullanılarak depremden etkilenen insanların afet sonrası duygu durumlarının incelenmesidir. Bu sayede, deprem anında bölge için genel duygu durumu ve sonrası için yapılacak sosyal ve psikolojik iyileştirme çalışmaları planlamalarının CBS ortamında daha hızlı ve kolay bir şekilde yapılmasına katkı sağlanabilecektir. Bu çalışmada, doğal dil işlemede kapı yinelemeli birimler (Gated Recurrent Units, GRU) derin ağ modeli kullanılarak X platformundan kazınmış ve düzenlenmiş veri kümesi üzerinde iki etiketli duygu modelinde %87, dört etiketli duygu analizi modelinde %79 test doğruluğu ile bir duygu analizi gerçekleştirilmiştir. Ayrıca CBS ortamında sıcak nokta analizi yapılarak yaşanan deprem ile ilgili X iletilerindeki duygu durumuna ait kümelenme örüntüsü istatistiksel olarak incelenmiştir. Böylece, yaşanabilecek depremler ile

ilgili X platformuna ait coğrafi referanslı sosyal medya verilerinin duygu analizinde kullanılabileceği saptanmıştır.

**Anahtar Kelimeler:** Coğrafi referanslı sosyal medya verileri, web kazıma, duygu analizi, derin öğrenme, mekânsal analiz



# **Sentiment Analysis from Georeferenced Social Media Data Using Natural Language Processing and Deep Learning**

Dilan Gözdem DOLU

Department of Geomatic Engineering

Master of Science Thesis

Supervisor: Assoc. Prof. Dr. Alper ŞEN

In natural disaster management, the damages caused by natural disasters can be minimized by using Geographic Information Systems (GIS) in pre-disaster preparation, disaster response and post-disaster recovery stages. The aim of this study is to examine the effects of the Kahramanmaraş earthquakes of February 6, 2023 on the region and people by scraping the social media data of X (formerly known as Twitter) using Selenium and BeautifulSoup libraries in Python programming language, and to examine the post-disaster emotional state of people affected by the earthquake using natural language processing and deep learning methods. Thus, it will be possible to contribute to the planning of the general emotional state of the region at the time of the earthquake and the social and psychological rehabilitation activities to be carried out afterwards in a faster and easier way in the GIS environment. In this study, a sentiment analysis was conducted on the scraped and organized dataset from the X platform using the Gated Recurrent Units (GRU) deep network model in natural language processing. The analysis achieved 87% test accuracy for the binary sentiment classification model and %79 test accuracy for the four-label sentiment analysis model. In addition, by performing hot spot analysis in the GIS environment, the clustering

pattern of the emotional state in X messages related to the earthquake occurred was statistically analyzed. Thus, it was determined that georeferenced social media data of the X platform related to possible earthquakes can be used in sentiment analysis.

**Keywords:** Georeferenced social media data, web scraping, sentiment analysis, deep learning, spatial analysis



## 1.1 Literatür Özeti

Modern teknolojinin hızlı gelişmeleri sayesinde sensör ağları ve uzaktan algılama tekniklerinin kurulması yoluyla doğal afetlerin izlenmesi, önceden uyarılması ve afetlerden sonra yaşanan durumların genel incelenmesi mümkün hale gelmiştir (Feng vd., 2022). Doğal afet yönetiminde de afet öncesi hazırlık, afet anı müdahale ve afet sonrası iyileştirme aşamaları mevcuttur (Neal, 1997). Doğal afetlerin yönetiminde Coğrafi Bilgi Sistemlerinin (CBS) 1960'larda teknoloji ile mantık arasındaki bağlantıyı ortaya çıkarmasıyla başlayan süreç (Clarke, 1997), günümüzde afet yönetimi aşamalarında yaygın bir şekilde diğer sistemlerle entegre edilen CBS'nin kullanımıyla devam etmektedir. Bu bütünleşmiş sistemlerden biri de sosyal medyadır. Kullanıcılar paylaşımlarıyla sosyal medyayı haberleri, etkinlikleri veya ürünleri değerlendirdikleri, kendi düşüncelerini ve duygularını aktardıkları bir platform olarak görürler. Sosyal medya verileri, işletmeciler için ürünleri ve piyasadaki marka izlenimleri hakkında bilgi verebileceği gibi sosyal olaylardan sonra vatandaşların duygu durumlarını da analiz etmek için kullanılabilir. Bu durum, sosyal medya veri işleminin çalışmalarda sık kullanılmasına ve kullanımının artmasına neden olmaktadır (Bhardwaj vd., 2015).

Sosyal medya kaynağının sağladığı çeşitli mekânsal verilerin kullanılması ile ilgi noktalarının belirlenmesi, olayların zamansal ve mekânsal dağılımının değerlendirilmesi, belirli bir zamanda ve mekânda olayların yoğunluğunun belirlenmesi gibi farklı mekânsal problemlerin analizlerine katkı sağlamak mümkündür. Sosyal medya üzerinden anlık veri çekilebilmesi, büyük ve çeşitli veriler sağlanabilmesi nedeniyle potansiyel uygulama alanlarından biri doğal afetlerdir (İlhan & Sağaltıcı, 2020). Sosyal medya içerisindeki bilgiler veri kazıma

ile elde edildikten sonra farklı analizlerde kullanılarak müdahale ve iyileştirme çalışmalarına katkı sağlanabilir. Örneğin, sosyal medyadan elde edilen veriler doğal dil işleme (Natural Language Processing, NLP) tekniklerinden biri olan duygu analizinde kullanılabilir. Bu analiz sonucunda insanların araştırılan konu hakkındaki görüşleri öğrenilebilir. (Adalı, 2012). Duygu analizi yapılırken birçok makine öğrenmesi yöntemi uygulanabilir. Bunlardan biri olan derin evrişimli sinir ağları sayesinde bilgisayarlı görü (computer vision) ve doğal dil işlemede büyük ilerleme sağlanmıştır (Feng vd., 2022).

Sosyal medya veri kazıma çalışmalarından elde edilen metinler üzerinden zamansal ve mekânsal (geometrik ve semantik) veriler elde edilebilir. Elde edilen veriler ile türlerine göre çeşitli analizler gerçekleştirilebilir. Bu tür analizlere örnek olarak mekânsal veri analizinde, verilerin mekânda nasıl etkileşime girdiği yöntemlerle incelenir. Bu işlemlerin yanı sıra verilerin diğer mekânsal vakalarla olan bağlantılarını araştırmaya olanak sağlayan bir yaklaşım sağlar. Bu tür analizler, değişkenlerin mekânsal dağılımını anlamayı, örüntülerin tanımlanmasını, mekânsal kümelenmeyi ve değişkenler arasındaki bağlantıları çözümlenmeyi amaçlayan çeşitli teknikleri içerir (Özgür & Aydın, 2011).

Sosyal medya verileri kullanılarak gerçekleştirilen mekânsal analiz çalışmaları: Alves vd. (2015) bu çalışmalarında, sosyal medya verilerinde duygu analizi yaparken hem zamansal hem de mekânsal boyutları dikkate alarak ayrıntılı bir analiz sunmayı amaçlamışlardır. 2013 FIFA Konfederasyonlar Kupası ile ilgili Portekizce tweetleri kullanarak, Naive Bayes ve Destek Vektör Makineleri (DVM) sınıflayıcıları ile duygu sınıflandırması yapılmıştır. Sonuçlara göre DVM yöntemi daha başarılı bulunmuştur. Analiz sonucunda, futbol müsabakası başlamadan önceki bir zaman dilimine ait duygu sınıflandırılmalı tweetlerden ısı haritaları oluşturulmuştur. Bu sayede negatif, nötr ve pozitif yayılımlar çıkarılmıştır. Çalışma, mekânsal ve zamansal boyutların analize dâhil edilmesiyle, duygu değişimlerinin hem zaman içinde hem de coğrafi bölgeler arasında nasıl farklılık gösterdiğini ortaya koymuştur. Ayrıca, mekânsal ve zamansal duygu analizlerinin, belirli bölgelerde oluşan olumsuz görüşlere hızlı müdahale edilmesini sağlayarak daha ayrıntılı değerlendirmelere imkân tanıdığı sonucuna ulaşılmıştır.

Song ve Xia (2016) bu çalışmada, Curtin Üniversitesi Bentley kampüsü çevresinde coğrafi etiketli tweetler kullanılarak duygusal eğilimlerin mekânsal ve zamansal

özellikleri incelenmiştir. Toplanan tweetler, Starlight Data Engineer yazılımı kullanılarak duygu analizi sürecinden geçirilmiş ve elde edilen veriler FME ve ArcGIS yazılımlarıyla coğrafi veri tabanına aktarılmıştır. Analizde, tweetlerdeki pozitif ve negatif sözcükler üzerinden çeşitli duygu ölçütleri hesaplanmış ve tweetler duygu yönelimlerine göre kategorize edilmiştir. Sonuçlar, tweetlerin yaklaşık %45'inin nötr duygular taşıdığını, pozitif tweet sayısının negatif tweet sayısının yaklaşık iki katı olduğunu göstermiştir. Mekânsal olarak en fazla tweetin eğlence, otopark ve yurt alanlarında paylaşıldığı, kütüphane alanında ise negatif duygu oranının en yüksek olduğu belirlenmiştir. Zaman boyutunda ise, negatif duyguların dönem başında en yüksek seviyede olduğu, sınav dönemi civarında zirveye ulaştığını ve sınav sonrasında ise belirgin şekilde azaldığı görülmüştür. Sonuç olarak çalışmada, bazı tweetlerde duygu sözcükleri bulunmamasıyla birlikte, duygu analizi modelinin bu tür durumları yanlış sınıflandırabileceği ve dolayısıyla duyguların hatalı bir şekilde analiz edilebileceği vurgulanmıştır.

Zhou ve Xu (2017), New York City - Washington DC bölgesindeki olayların X verileri üzerinden mekânsal ve zamansal dağılımları incelenmiştir. Bu çalışmada, Geo API'den toplanan Küresel Konumlama Sistemi (Global Positioning System, GPS) bilgisini içeren iletiler kullanılmıştır. Geo API'den toplanan iletilerde konum içeriği, zaman dilimleri, yer adları ve GPS ölçümleri gibi en az bir tür konum verisi bulunur. Elde edilen veriler içerisinde yağmur gibi rastgele olayların yanı sıra planlı etkinlikler ile ilgili, örneğin Papa'nın ziyareti gibi, konuların X kullanıcıları üzerindeki etkileşimleri belirlenmiştir. Rastgele orman algoritması ve mekânsal-zamansal analizler kullanılarak incelenen olayların, X platformundaki iletilerde artışa neden olduğu tespit edilmiştir. X üzerinden olaylara olan anlık tepkileri ve ilgili mekânsal-zamansal örüntüleri belirleyerek sosyal medyadaki etkileşimleri anlamaya yönelik bir metodoloji sunmuştur.

Wang vd. (2018) tarafından, doğal afetlerden biri olan sel konusundaki sorunlar hakkında X üzerinden API ile elde edilen veriler kullanılarak NLP ve bilgisayarlı görü teknikleri ile bir yaklaşım belirlenmiş; X tabanlı verilerin sokak adları ölçeğinde hassasiyet sağladığı tespit edilmiştir.

Hourrane vd. (2018) tarafından, IMDB film yorumları ve Twitter veri setleri üzerinde çeşitli makine öğrenmesi yöntemleriyle duygu analizi gerçekleştirilmiştir. Veri temizleme, normalleştirme, kelime birimlerine ayırma gibi ön işlemler

uygulanmış ve TF-IDF (Term Frequency-Inverse Document Frequency) yöntemiyle özellik çıkarımı yapılmıştır. Lojistik Regresyon, Ridge Regresyonu (Ridge Classifier), Perceptron (Algılayıcı), Naive Bayes varyantları, Destek Vektör Makineleri (DVM) ve K-En Yakın Komşu (KNN) gibi sınıflayıcılar kullanılmıştır. DVM ve Ridge Regresyonu modelleri en başarılı sonuçları vermiştir. Elde edilen en iyi dört model birleştirilerek toplu bir sınıflandırıcı oluşturulmuş ve bireysel modellere kıyasla daha yüksek performans sağladığı gözlemlenmiştir.

Gulnerman ve Karaman (2020) tarafından yakınlık analizi kapsamında kent yol ağı poligon kırık noktalarından Voronoi alanları üretilerek bu alanlarda gönderilen iletilerden yer adları elde edilmiş ve bir coğrafi sözlük üretilmiştir.

Gulnerman vd. (2020a) tarafından, coğrafi ileti indiricisi (Geo Tweets Downloader, GTD) adlı sistem ile elde edilen coğrafi etiketli verilerden mekânsal bilgilere ulaşılmıştır. GTD, coğrafi etiketi olmayan iletileri atlayabilen, X API'leri kullanarak gerçek zamanlı olarak genel durum iletilerini sağlayabilen ve bu iletilerin coğrafi etiketleri de dahil olmak üzere çeşitli bilgilerini toplayabilen bir sistemdir. Bu çalışmada, İstanbul'daki kullanıcıların sosyal medya verileri ile davranışlarını anlamaya yönelik bir inceleme yapılmıştır. Veri analizi ile kullanıcı etkinliklerinin temsil düzeylerinin, mekân-zaman önyargılarının ve şehirdeki trendlerin belirlenmesi çalışmaları gerçekleştirilmiştir. Aktif kullanıcıların verilerinin temizlenmesi, farklı temsil düzeylerinin vurgulanması ve mekânsal-zamansal anomali eğilimlerinin tespit edilmesi gibi yöntemler kullanılarak elde edilen bulguların, şehirdeki günlük rutinlerin tarafsız bir şekilde ortaya çıkarılmasına yönelik potansiyel bir çerçeve sunduğu gösterilmiştir. Sonuç olarak sosyal medya verilerinin acil durumlar veya olağan dışı olayların tespitinde ve şehirdeki vatandaş davranışlarını izleme sistemi oluşturulmasında kullanılabilecek önemli verileri sağladığı vurgulanmıştır.

Gulnerman vd. (2020b) tarafından, sosyal medya üzerinden GTD kullanılarak veriler elde edilmiş; coğrafi koordinatlara sahip iletilerin bilgisi mekânsal analiz ve metin analizi ile birleştirilerek olayları anlama sürecine katkı sağlanması amaçlanmıştır. Mekânsal verilerden daha kesin bilgi çıkarımını belirlemek için optimize edilmiş sıcak nokta analizi ile belirli bir zamanda ve mekânda olayların yoğunluğunun belirlenmesi çalışması yapılmıştır. Bu çalışmada, coğrafi referanslı sosyal medya verilerinin acil durum izlemede faydalı olabileceği ele alınmıştır.

Camacho vd. (2021) tarafından kullanıcıların profilinde belirttikleri konum bilgileri ve iletilerin içerdiği yer adları kullanılarak coğrafi kodlama işlemi gerçekleştirilmiştir. Kullanıcıların profil bilgilerinde belirttikleri konumlar ve iletilerin içeriğinde geçen yer adları, bir coğrafi kodlama hizmeti olan Geocodio aracılığıyla koordinatlara dönüştürülerek mekânsal bilgiler elde edilmiştir. Bu çalışmada, sosyal medyadaki doğal gazla ilgili paylaşımların duygusal içeriği incelenerek mekânsal dağılımı ele alınmıştır. Çeşitli duyarlılık analizi yöntemleri ve makine öğrenmesi modelleri kullanılarak paylaşımlar duygusal kategorilere ayrılmış ve coğrafi konumları belirlenmiştir. Amerika Birleşik Devletleri'nin farklı bölgelerinde doğal gazla ilgili paylaşımların duygusal dağılımı gösterilerek, zaman içinde duygusal içeriğin değişimi de ortaya konmuştur. Duygu analizi ve mekânsal analiz yöntemleri entegre edilerek sosyal medyadaki duygusal içeriğin coğrafi dağılımı sunulmuştur.

Feng vd. (2022) tarafından, doğal afetlerle ilgili sosyal medya içeriği analiz edilmiş ve bu verilerden hangi bilgilerin elde edilebileceği ve bunların nasıl kullanılabilirliği incelenmiştir. İlk olarak, sosyal medya üzerindeki doğal afet odaklı yayınlar sekiz farklı adımda sınıflandırılmış ve bu verilerin kullanımındaki zorluklar ve fırsatlar tartışılmıştır. Yapılan derleme çalışmasında, sosyal medyanın doğal afetlerle ilgili bilgi çıkarımında ve kullanımında karşılaşılan çeşitli zorluklara ışık tutmuşlardır. Sosyal medyanın doğal afetlerle ilgili veri toplamadaki ve veri analizindeki potansiyelini aynı zamanda da zorluklarını ortaya koyarak mevcut çözümler ile birlikte yeni araştırma fırsatlarını vurgulamışlardır.

Cardone vd. (2022) tarafından gerçekleştirilen bu çalışmada, kentsel hizmetlere yönelik vatandaş ve turist algılarını sosyal medya iletileri üzerinden analizi gerçekleştirilmiştir. Analiz için Genişletilmiş Bulanık C-ortalamları (Extended Fuzzy C-means, EFCM) kümeleme algoritmasından yararlanılmıştır. Yaklaşık 5000 sosyal medya iletisi, Bologna (İtalya) kentinin altı bölgesinde toplanmış ve değerlendirilmiştir. İletiler, 8 hoş ve 8 hoş olmayan duygu kategorisine göre sınıflandırılmıştır. Bulgular CBS tabanlı tematik haritalar aracılığıyla görselleştirilmiştir. Borgo Panigale-Reno'da sakinlik (%30), Navile'de endişe (%29), Porto-Saragozza'da ilgi (%31) duygusu baskın bulunmuştur. San Donato-San Vitale'de üzüntü (%30), Santo Stefano'da ise iyimserlik (%32) öne çıkmıştır. Savena bölgesinde iletilerde en çok ilgi duygusu (%26) ifade edilmiştir. Santo

Stefano bölgesi genel olarak olumlu algılanan hizmetlerle öne çıkarken, Navile bölgesi olumsuz duyguların yoğunlaştığı bir bölge olarak belirlenmiştir. Çalışma, sosyal medya verilerinden türetilen duygusal analizlerin, şehir hizmetlerinin algılanmasında etkili bir araç olabileceğini göstermektedir.

Qian ve Gkritza (2024) tarafından ABD genelinde 2019–2022 dönemini kapsayan bu çalışmada, sosyal medya platformu X (Twitter) üzerinden elektrikli araçlara (EV) yönelik kamuoyu algısı incelenmiştir. Çalışmada, X Academic Research API kullanılarak elde edilen coğrafi ve zamansal etiketli iletiler analiz edilmiştir. COVID-19 sürecinde EV'lere yönelik ilginin kısa süreli azaldığı, ancak 2020 ortasından itibaren artarak pandemi öncesi seviyeleri aştığı belirlenmiştir. Coğrafi analizler, doğu ve batı kıyılarındaki eyaletlerde çevrim içi tartışmaların yoğun olduğunu; buna karşın bazı kuzeydoğu şehirlerinde yüksek dijital ilgiye rağmen EV sahipliğinin düşük kaldığını göstermektedir. Konu modelleme ile iletilerin %90'ının Tesla, batarya üretimi, şarj altyapısı ve hükümet politikaları gibi 16 ana başlıkta toplandığı görülmüştür. Duygu analizine göre, erken benimseyici eyaletlerde olumlu tutumlar baskınken; geç benimseyicilerde ekonomik kaygılar ve altyapı yetersizlikleri gibi olumsuz yaklaşımlar öne çıkmaktadır. Bu bulgular, EV benimsenmesinin yalnızca teknolojik ve altyapısal koşullarla değil, aynı zamanda toplumsal algı ve bölgesel dijital etkileşimlerle de şekillendiğini ortaya koymaktadır.

Liu vd. (2024) tarafından gerçekleştirilen bu çalışmada, yüksek yoğunluklu inşa edilmiş çevre faktörlerinin insan duygusal algısı üzerindeki etkisi incelenmektedir. Bu doğrultuda, Wuhan kentinin ana kentsel alanındaki 2665 blok üzerinde, sosyal medya verileri (Weibo) kullanılarak coğrafi konumlandırılmış duygusal veriler toplanmıştır. Doğal Dil İşleme (NLP) teknikleriyle analiz edilen bu veriler, olumlu ve olumsuz duyguların mekansal dağılımını ortaya koymuştur. Ayrıca, Çok Ölçekli Coğrafi Ağırlıklı Regresyon (MGWR) modeli kullanılarak, yüksek yoğunluklu inşa edilmiş çevre göstergeleri ile duygusal algı arasındaki ilişkiler incelenmiştir. Bulgular, olumlu duyguların yüksek yoğunluklu inşa edilmiş çevrelerde mekânsal farklılıklar gösterdiğini ortaya koymuştur. Bu sonuçlar, kentsel tasarım ve yönetişimde daha rafine yaklaşımların benimsenmesi gerektiğini ve yukarıdan aşağıya planlama stratejilerinin sınırlamalarını vurgulamaktadır. Yazarlar, bu çalışmanın kentsel planlamada uygulama potansiyeli taşıdığını öne sürmektedir.

Sosyal medya verileri aracılığıyla yapılan bu nicel incelemenin, planlamacıların kent sakinlerinin psikolojik iyilik halini gözeten stratejiler geliştirmelerine katkı sağlayabileceğini belirtmekte; ancak bulguların büyük ölçekli ve ön inceleme niteliğinde olduğunu da ifade etmektedirler.

Aman ve Matisziw (2025) tarafından yapılan bu çalışma, coğrafi etiketli Instagram gönderilerinden çıkarılan duygu verileri ile sokak görüntülerinden elde edilen kentsel tasarım özelliklerini bütünleştirerek, kentsel çevrenin insan duyguları üzerindeki etkisini mekânsal ve zamansal bağlamda analiz etmektedir. Doğal dil işleme (BERT) ve bilgisayarlı görme (PSPNet, Mask R-CNN) tekniklerinin birlikte kullanıldığı analizlerde, 47.000'den fazla sosyal medya gönderisi ve 1.364 sokak görüntüsü değerlendirilmiştir. Bulgular, yeşil alanlar, kaldırımlar ve yoğun yaya etkileşiminin olumlu duygularla, çitlerle çevrili boş alanların ise olumsuz duygularla ilişkili olduğunu göstermektedir. Pandemi döneminde araç ve ticari bölgelerde olumsuz duygular artarken, sonrasında yaya dostu alanlarda pozitif duygular yeniden yükselmiştir. Sıcak-soğuk nokta analizleri, olumlu duyguların parklarda ve sosyal alanlarda yoğunlaştığını, olumsuz duyguların ise araç trafiğiyle ilişkili bölgelerde kümelenildiğini ortaya koymaktadır. Çalışma, geleneksel anket yöntemlerinin yerine ölçeklenebilir ve zaman duyarlı bir çerçeve sunarak, şehir planlamasında fiziksel çevrenin yanı sıra duygusal deneyimlerin de dikkate alınması gerektiğini vurgulamaktadır. Yazarlar, sosyal medya verilerinin sınırlılıklarına dikkat çekerek, daha kapsayıcı ve çeşitli veri kaynaklarıyla desteklenen bütüncül analizlere ihtiyaç duyulduğunu belirtmektedir

Zhang vd. (2025), sosyal medya analitiğini afet yönetimiyle entegre ettikleri çalışmalarında, 2023 yılında Tayfun Doksuri'nin Pekin ve çevresinde yol açtığı aşırı yağışlara odaklanmışlardır. Sina Weibo platformundan elde edilen 5.791 gönderi ve 60.546 yorum, Bert-BiLSTM temelli çoklu duygu sınıflandırma modeli ile analiz edilmiştir. Sekiz afet duygusu kategorisinin belirlendiği çalışmada, afetin öncesi, sırası ve sonrası olmak üzere üç evresine yönelik tematik evrim incelenmiş; K-ortalamalar yöntemi ile bu temalar ayrıştırılmıştır. Ayrıca, afet zinciri analiz yöntemiyle ikincil afet türleri (can kaybı, altyapı hasarı, trafik sorunları vb.) sistematik biçimde sınıflandırılmıştır. Mentougou ve Fangshan gibi bölgelerde tespit edilen mekânsal kümelenmeler, yerel topografya ve altyapı koşullarıyla ilişkilendirilerek bu alanların yüksek risk taşıdığı vurgulanmıştır. Çalışma, teorik

düzyeyde ince taneli duygu sınıflandırması ve çok boyutlu afet çerçevesi sunarken; pratik düzeyde ise yüksek riskli bölgelerde öncelikli müdahale ve kamu iletişim stratejilerinin geliştirilmesine katkı sağlamaktadır.

Yapılan literatür incelemesi, sosyal medya verilerinin mekânsal ve zamansal boyutlarda analiz edilerek kent içi davranışların, afet yönetiminin, duygu değişimlerinin ve toplumsal olaylara verilen tepkilerin anlaşılmasında güçlü bir araç olarak kullanılabildiğini ortaya koymuştur. Farklı çalışmalar; coğrafi etiketleme, duygu analizi, sıcak nokta tespiti, NLP ve makine öğrenmesi gibi yöntemlerle sosyal medya içeriklerinden anlamlı bilgi çıkarımına olanak tanımıştır. Özellikle doğal afetler ve kriz anlarında, sosyal medya verileri anlık ve yerel duyarlılıkların tespiti açısından kritik öneme sahip olup; mekân-zaman önyargıları, kullanıcı temsili ve doğruluk gibi zorluklara rağmen önemli fırsatlar sunmaktadır. Bu kapsamda, sosyal medya platformlarının şehir planlaması, afet müdahale stratejileri ve hizmet değerlendirmeleri gibi alanlarda veri destekli karar mekanizmalarına katkı sağlayabileceği görülmektedir.

## **1.2 Tezin Amacı**

Tez kapsamında, 6 Şubat 2023 Kahramanmaraş depremlerine ait sosyal medya verilerinin kazınmasını yapan ve elde edilen verilere duygu analizi uygulayan bir model geliştirilmiştir. Duygu analizinden elde edilen bilgiler mekânsal analizlerden geçirilerek sonuç haritaları oluşturulmuştur. Amaç, deprem sonrasında sosyal medya platformuna gönderilen coğrafi referans içeren iletilerin derin öğrenme kullanılarak geliştirilen duygu analizinin performansının tespit edilmesidir. Aynı zamanda, sosyal medya verilerinin duygu ve mekânsal analizler ile birlikte kullanımı sayesinde ihtiyaçların ve duygusal tepkilerin yoğunlaştığı bölgelerin belirlenebileceğini göstermektedir. Kaynakların bu ihtiyaçlara ve bölgelere göre verimli kullanılması sayesinde psikososyal destek ekiplerinin güçlerini en etkin şekilde kullanmalarına yardımcı olması da sağlanabilir.

### 1.3 Hipotez

Tez çalışmasının hipotezi şunlardır:

- Sosyal medya verilerinden elde edilen mekânsal veriler, derin öğrenme kullanılarak mekâna dayalı bir duygu analizi modeli geliştirilebilir.
- Sosyal medya verilerine duygu etiketleri ve mekânsal analizler uygulanarak, acil durumlarda psikososyal destek ekiplerinin kaynaklarını en verimli şekilde planlamalarına olanak sağlanabilir.



### 2.1 Veri Kazıma

İnsanların fikirlerini ve bilgi dağarcığını sosyal medya platformları aracılığıyla paylaşımları büyük veri setlerinin keşfedilmesinde önemli rol oynamıştır. Bu veri setleri, makine öğrenmesi ve veri madenciliği gibi yöntemlerle incelenerek ileriye öngörmesi aynı zamanda da gizli örüntüleri çıkarması gibi amaçlar için kullanılmaktadır. Veri kazıma yalnızca veri ayıklama ile sınırlı değildir. Aynı zamanda analiz edilmek üzere veri tabanına kaydedilme görevlerini de gerçekleştirir (Albayrak vd., 2017).

Veri kazıma sonucunda metinlerden geometrik ve semantik veriler ile bir veri seti oluşturabilir. Veri kazıma sonucu elde edilen iletilerden oluşan bu veri kümeleri zaman dilimi, konuları ve kullanıcıların iletilerinin gönderimi sırasında paylaşım açtıkları konum bilgilerini içerisinde barındırabilir (Kaya, 2021). Sosyal medya ortamında mevcut olabilen konum bilgisi, kullanıcılarının iletiyi paylaştıkları yerin coğrafi konum bilgileri, sözel olarak yer alan kullanıcı profilindeki yerleşim yeri, paylaşımlarında bahsedilen konular ve doğrudan paylaşılan enlem-boylam gibi konular aracılığıyla belirlenebilir (Gulnerman & Karaman, 2020).

Mekânsal veri analizi, verilerin mekânda nasıl etkileşime girdiğini yöntemlerle incelerken, aynı zamanda bu verilerin diğer mekânsal vakalarla olan bağlantılarını araştıran bir yaklaşımdır. Bu tür analizler, değişkenlerin mekânsal dağılımını anlamayı, örüntülerin tanımlanmasını, mekânsal kümelenmeyi ve değişkenler arasındaki bağlantıları çözümlenmeyi amaçlayan çeşitli teknikleri içerir (Özgür & Aydın, 2011).

Büyük verilerin hızlı, otomatik ve sürekli bir şekilde elde edilmesini sağlayan internetten veri kazıma işlemi iki adımda yapılabilir. Bu adımlardan ilki internet kaynaklarının elde edilmesi, diğeri ise elde edilen verilerden istenen bilgilerin çıkarılması işlemidir (Zhao, 2017).

X API, X verilerini programlı bir şekilde almak ve analiz etmek, ayrıca X üzerindeki gönderiler için uygulamalar geliştirmek amacıyla kullanılabilir. Yıllar içinde, X API geliştiricilerin kamuya açık gönderileri geliştirmeleri ve araştırmaları için erişimlerini ölçeklendirmelerine olanak tanıyan ek erişim seviyeleri geliştirmiştir. Yakın zamanda X API v2 kullanıma sunulmuştur. X API v2, modern bir temel, yeni ve gelişmiş özellikler ile temel erişime hızlı bir başlangıç sunmaktadır (URL-9).

Bu bağlamda, X (Twitter) verisi toplama yöntemleri literatürde iki ana grupta ele alınmaktadır. Bunlar, X API'lerinin kullanımı ve web kazıma (scraping) teknikleridir. API tabanlı yöntemler, belgelenmiş ve sürdürülebilir bir veri erişimi sağlar. Ancak kullanıcı başına veri erişimi sınırlamaları içermektedir. Bu kısıtları aşmak için çoklu istemci veya paralel işlem yöntemleri kullanılsa da veri kaybı tamamen önlenememektedir. Web kazıma yöntemleri ise, API ile erişilemeyen verilere ulaşma imkânı sunar. Ancak bu yöntem, teknik ve etik sınırlamalara içermektedir. X'in robots.txt dosyasında hangi sayfaların taranabileceği belirtilmiştir ve her istek arasında en az 1 saniye beklenmesi gerekmektedir. Bu sınırlamalar, verinin yasal ve güvenilir şekilde toplanmasını zorlaştırmaktadır (You vd., 2024).

Veriler elde edilirken X platformundan Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli bölgelerinden 755 adet gönderi hem veri kazıma hem de API üzerinden elde edilebilmiştir.

## **2.2 İşaretleme Dili**

İşaretleme Dili (Markup Language), metinlerin düzenlenmesi, biçimlendirilmesi ve sunulması amacıyla kullanılan bir dil türüdür. Bu dil, metin içindeki öğeleri tanımlamak için belirli işaretler veya etiketler kullanır. Etiketler, metnin nasıl gösterileceğini veya işlenmesi gerektiğini belirler. Yaygın olarak kullanılan

işaretleme dilleri arasında HTML, XML gibi diller bulunmaktadır (Gümüşsu, 2019).

### **2.2.1 Metin Tabanlı İşaretleme Dili**

İnternet sitelerinin metin tabanlı bir işaretleme dili vardır. Bunlardan en yaygın olanı HTML ve XML'dir. Metin tabanlı işaretleme dili sayfanın içeriğini tanımlar ve bu sayede içeriğinde olan evrensel biçimleri ve işaretleme dillerinin etiketleri sayesinde istenen bilgiyi almak kolaylaşır (Gümüşsu, 2019).

HTML, web sayfalarının oluşturulmasında kullanılan bir işaretleme dilidir. Bu dil, bir web tarayıcısında içeriğin nasıl görüneceğini tanımlamak için kullanılır. HTML'in temel amacı, web sayfalarının yapısını tanımlayarak metin, görsel ve diğer içeriklerin tarayıcıda düzenli ve estetik bir biçimde görüntülenmesini sağlamaktır. Ancak, HTML içinde hem verinin hem de görselliğin aynı anda etiketlerle ifade edilmesi sorunlara neden olmuştur. Bu noktada, SGML'nin (Standard Generalized Markup Language) bir alt kümesi ve türevi olan XML (eXtensible Markup Language) geliştirilmiştir (Uzun vd, 2008).

XML, verinin yapısını ve içeriğini tanımlamaya odaklanır. Bu nedenle, verinin görsel sunumundan çok verinin anlamını ve organizasyonunu belirlemek için kullanılır. XML'in özelliklerinden biri, etiketlerin önceden tanımlı olmamasıdır; kullanıcılar ihtiyaçlarına göre kendi etiketlerini oluşturabilir. XML'in hiyerarşik yapısı, verilerin düzenli bir şekilde organize edilmesini sağlar (Uzun vd., 2008).

## **2.3 Veri Kazımada Kullanılan Python Kütüphaneleri**

### **2.3.1 BeautifulSoup Kütüphanesi**

BeautifulSoup kütüphanesi HTML kodundaki bilgileri ayıklamaya ve çıkarmaya yönelik bir modüldür. Sayfa kaynağı HTML veya XML olan belgeleri ayrıştırmak için kullanılır (Kılınç vd., 2022). Bu sayede işlenmekte olan ayrıştırmanın kodlamasını otomatik bir şekilde algılayabilir ve bunu istemci tarafından okunabilir bir kodlamaya dönüştürerek anlaşılabilir olmasını sağlar (Zhao, 2017). Bu işlem gerçekleştirilirken sayfa kaynağında bulunan XPath, sınıf ve işaretleme etiketlerinden yararlanır. İşleme sırası ise ilk olarak sayfa içerisinde bütün bilgileri içeren kaynağı indirip içerisinden özellikle istenilen bilgilerin "find\_all" çağırısıyla elde edilmesi şeklinde ilerlemektedir (URL-6).

BeautifulSoup, özellikle web sayfalarının HTML içeriğini ayrıştırmak için kullanılan bir Python kütüphanesidir. Bu kütüphane, HTML öğelerini bulmak ve düzenlemek için birden fazla ayrıştırıcıya sahiptir. BeautifulSoup, HTML içeriğini anlamlı bir şema gibi dönüştürmek için ayrıştırıcı (parser) kullanır. Parserlar, ham HTML kodunu satır satır okuyarak yapısal olarak düzenler, böylece programlama yoluyla öğeleri kolayca bulabilir veya değiştirebilir. BeautifulSoup birkaç farklı ayrıştırıcı ile çalışabilir. Python'un yerleşik html.parser ayrıştırıcısı, Python 3 sürümleriyle uyumlu olup ek kurulum gerektirmez ve yaygın olarak kullanılır. Diğer yandan, lxml ayrıştırıcısı, hızlı bir işlem gücü sunar ancak ek kurulum gerektirir. Html5lib ise, web sayfalarını bir web tarayıcısı gibi ayrıştırarak daha hassas sonuçlar verir, ancak performansı daha düşüktür (Tanasă & Oprea, 2023).

BeautifulSoup'un temel amacı, HTML verilerini ağaç tabanlı bir yapıya dönüştürmektir. Bu sayede, veriler üzerinde işlem yapmak daha kolay hale gelir. Sayfadan veri çekmek için find ve find\_all (ya da eski sürümlerinde findAll) yöntemleri kullanılır. find() yöntemi, belirtilen öğeyi döndürürken, find\_all() (veya findAll) birden fazla öğeyi döndürür. Bu yöntemlerle, HTML yapısındaki öğelere kolayca ulaşılabilir ve üzerinde işlem yapılabilir. Bu işlevsellik, web verisi çekme (web scraping) işlemlerinde kullanılmaktadır (Oprea & Bâra, 2022).

### **2.3.2 Selenium Kütüphanesi**

Selenium, başlangıçta web sitelerinin otomatik test edilmesi amacıyla geliştirilmiş bir web kazıma aracıdır. Selenium, çeşitli tarayıcıları otomatikleştirerek bir web sayfasını yüklemek, içeriğini almak ve bir kullanıcı gibi tarayıcıda düğmelere veya bağlantılara tıklamak gibi işlemleri gerçekleştirmek için çalışır. Bu nedenle, kullanıcıların gerçekleştirebileceği dinamik işlemlere odaklanan bir web kazıma aracıdır (Tanasă & Oprea, 2023).

Selenium kütüphanesi Google Chrome ve Internet Explorer gibi web tarayıcısı ile bağlantı kurup bunun üzerinden kimlik doğrulama, yeniden yönlendirme, çerezler gibi http istekleri ve sosyal medya platformlarına giriş yapmada istenen bilgilerinin otomatik girilmesini sağlayan Python kütüphanesidir (Zhao, 2017).

Selenium, Java, C#, PHP ve Python gibi programlama dilleri ile kullanılabilir. Ancak Selenium, kendi web tarayıcısıyla birlikte gelmez ve tarayıcılarla etkileşim kurmak için bir tarayıcı sürücüsü (WebDriver) gerektirir. WebDriver'lar, Internet

Explorer, Chrome, Firefox, Edge, Safari gibi tarayıcılar için mevcuttur. Bu WebDriver'lar sayesinde bir tarayıcı penceresi açılır ve Python kodunda belirtilen işlemler simüle edilir. Temelde tarayıcı sürücüleri (WebDriver) üzerinden komutları takip ve kabul eden, bunları bir tarayıcıya gönderen otomasyon çerçevesidir (Tanasă & Oprea, 2023). Bu çalışmada, tarayıcı sürücülerinden Google Chrome seçilmiştir. Bu bağlantılarda önemli olan bir diğer faktör ise kullanılan tarayıcının güncel olmasıdır.



### 3.1 Derin Öğrenme

Derin öğrenme, yapay zekâ ve makine öğrenmesi alanlarında yer edinmiştir. Bu yöntemler, yapay sinir ağlarından türetilmiş olmakla birlikte, daha fazla katman ve nöron içeren karmaşık yapılara sahiptir. Derin öğrenmenin temel amacı, verileri katmanlı bir yapı içerisinde işleyerek, farklı soyutlama seviyelerinde temsil etmektir. Bu yaklaşım, ham veriden başlayarak anlamlı ve yüksek düzeyli özellikler çıkarmayı mümkün kılar.

Birden fazla katmanın bir arada çalıştığı bu sistemlerde, her katman önceki katmandan gelen çıktıları işleyerek yeni temsiller üretir. Böylece, derin öğrenme algoritmaları, genellikle manuel olarak seçilmesi gereken öznelikleri otomatik olarak öğrenebilir ve daha karmaşık veri türlerini analiz edebilir. Günümüzde, metin, görüntü ve ses gibi birçok farklı veri türünün işlenmesinde bu yöntemler kullanılmaktadır. Derin öğrenme yöntemleri, yapay sinir ağlarının temeli üzerine inşa edilmiştir. Yapay sinir ağları, sinir ağı yöntemlerinin temel yapı taşını oluşturur ve nöron adı verilen birimlerden meydana gelir. Her bir nöron, kendisine gelen girdileri belirli ağırlıklarla çarparak birleştirir ve bu doğrusal kombinasyonu doğrusal olmayan bir aktivasyon fonksiyonuna aktarır. Bu mekanizma, verilerin daha karmaşık bir şekilde işlenmesini ve modellenmesini sağlar (Küçük & Arıcı, 2018).

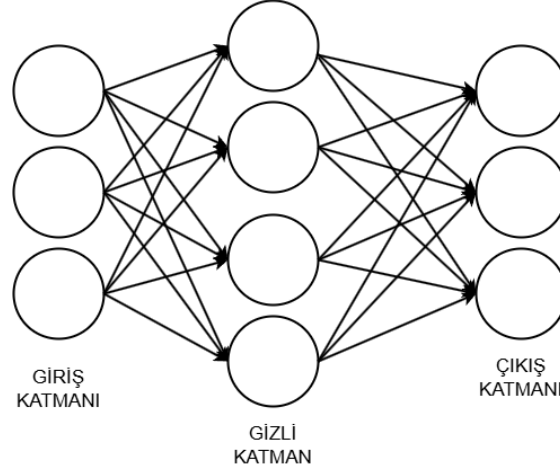
### 3.2 Yapay Sinir Ağları

Yapay sinir ağları başta olmak üzere tüm sinir ağı yöntemleri nöronlara dayanmaktadır. Yapay sinir ağlarında nöronun girdileri farklı ağırlıklarla doğrusal bir şekilde birleştirilir. Bu kombinasyonun sonucunda ise doğrusal olmayan bir

aktivasyon birimine aktarılır (Chen vd., 2020). Ağın yüksek bağlanabilirliği hata etkisinin minimum olmasını sağlar. Yinelenen sinir ağları, veri akışlarını analiz edebilen gizli katmanlara sahip özel bir sinir ağı türüdür. Bu ağlar, çıktılarının önceki hesaplamalara bağlı olduğu problemleri çözmek için oldukça uygun ve etkilidirler. Williams ve Zipser (1989) tarafından tanıtılan bu yöntem, özellikle doğal dil işleme alanında çeşitli problemlerin üstesinden gelmede başarılı olmuştur. Yinelenen sinir ağları (Recurrent Neural Networks, RNN), veri akışlarını dinamik bir şekilde analiz edebilme yetenekleriyle öne çıkarlar. Bu özelliği sayesinde, dilin karmaşıklığını ve sürekli değişen yapısını anlama ve işleme konusunda etkili bir araç haline gelmişlerdir. Bu teknoloji, metin analizi, çeviri, duygu analizi ve daha birçok doğal dil işleme problemine yenilikçi çözümler sunmuştur (Küçük & Arıcı, 2018).

Yapay sinir ağlarına ait nöronlar, özellikle veri akışlarını analiz etme ve doğrusal olmayan ilişkileri öğrenme yetenekleri açısından sinir ağlarında görev almaktadır. Bu bağlamda, yinelenen sinir ağları (RNN) gibi özel ağ yapıları, zaman serisi verilerinin ve dilsel yapılarının işlenmesinde kullanılarak, doğal dil işleme alanında yer edinmiştir (Lagrari & Elkettani, 2021).

Geleneksel sinir ağları, yalnızca mevcut girdilere dayanarak tahmin yapabildikleri için geçmiş bilgileri dikkate alamazlar. Bu durum, özellikle zaman serisi gibi önceki verilerin büyük önem taşıdığı problemler için bir sınırlamadır. Bu sorunu aşmak amacıyla geliştirilen yinelemeli sinir ağları (RNN), her işlem adımında geçmiş bilgileri taşıyan bir gizli durum (hidden state) kullanarak veriler arasındaki zamansal bağıntıları yakalayabilir. Böylece model, yalnızca mevcut girdiye değil, aynı zamanda geçmişte işlenmiş bilgilere de dayanarak daha isabetli tahminlerde bulunabilir. RNN'ler, bu sayede dinamik örüntülere duyarlı hale gelirken, LSTM (Long Short-Term Memory) ve GRU (Gated Recurrent Units) gibi daha gelişmiş yapılarda bu hafıza yeteneği daha da güçlendirilmiş ve uzun vadeli bağımlılıklar daha etkili şekilde öğrenilebilir hâle getirilmiştir. Bu mimaride giriş katmanı veriyi modele alırken, gizli katman zaman içindeki bağıntıları işler ve çıkış katmanı nihai tahmini üretir. Bu yapı sayesinde ağ, önceki adımlardan gelen bilgiyi unutmadan öğrenmeyi sürdürebilir (Şekil 3.1) (Vennerød vd., 2021).



**Şekil 3.1** Yapay sinir ağları

### 3.2.1 Uzun Kısa Süreli Bellek (Long Short-Term Memory, LSTM) Derin Ağ Modeli

Yinelenen sinir ağlarında öğrenme aşamasında karşılaştığı ağırlıkların yok olması sorunu nedeniyle, bu ağların uzun kısa süreli bellek (Long Sort Term Memory, LSTM) (Hochreiter & Schmidhuber, 1997) ve kapı yinelemeli birimler (Gated Recurrent Unit, GRU) (Cho vd., 2014) gibi çözümleri ortaya çıkmıştır (Küçük & Arıcı, 2018). Uzun Kısa Süreli Bellek (LSTM) modelleri, sıralı verilerde uzun vadeli ilişkileri öğrenebilmek için geliştirilen tekrarlayan sinir ağı (RNN) mimarilerinden biridir. Geleneksel RNN'ler, özellikle uzun süreli bağımlılıkları öğrenmede "kaybolan gradyan" (vanishing gradient) problemi nedeniyle sınırlı bir performans sergiler. Bu problem, zaman içinde gradyan değerlerinin üstel olarak azalmasına yol açarak, modelin sıralı girdilerdeki uzun vadeli ilişkileri öğrenmesini zorlaştırır. LSTM mimarisi, bu sorunu çözmek amacıyla, bilgi akışını kontrol eden bir dizi kapı mekanizması ile birlikte bir bellek hücresini içerecek şekilde tasarlanmıştır (William & Altamimi, 2024).

LSTM'nin temel yapı taşlarından biri, uzun süreli bilgiyi saklamak ve gerektiğinde güncellemek için kullanılan hücre durumudur. Bu yapı, modelin önceki bilgileri unutmadan yeni bilgileri öğrenmesini sağlar. LSTM'nin bilgi akışını kontrol eden üç ana kapısı bulunmaktadır. Unutma kapısı (forget gate), hücre durumunda hangi bilginin korunacağına ve hangisinin silineceğine karar verir. Bu işlem, bir sigmoid fonksiyonu ile gerçekleştirilir ve sonuç, tamamen unutma (0) ile tamamen saklama (1) arasında bir değer alır. Giriş kapısı (input gate), yeni bilginin hücre durumuna

ne ölçüde dahil edileceğini belirler. Bu aşamada, giriş bilgileri bir sigmoid fonksiyonu ve hiperbolik tanjant fonksiyonu ile işlenir. Çıkış kapısı (output gate), hücre durumundan gizli duruma ne kadar bilginin aktarılacağını kontrol eder ve gizli durum, o anki girdiye bağlı olarak güncellenir.

LSTM ağları, bu kapılar sayesinde geçmişten gelen uzun vadeli bağımlılıkları öğrenebilir ve sıralı verilerin işlenmesinde üstün performans sergiler. Bu yapı, bilgiyi zaman içinde etkin bir şekilde saklayarak ve gerektiğinde unutarak, geleneksel RNN'lerin sınırlamalarını aşmayı mümkün kılar (Hochreiter & Schmidhuber, 1997).

### 3.2.2 Kapı Yinelemeli Birimler (Gated Recurrent Units, GRU) Derin Ağ Modeli

LSTM birimine benzer şekilde, GRU'nun da birim içinde bilginin akışını modüle eden kapı birimleri vardır; ancak ayrı bir bellek hücresi bulunmaz. Ayrıca LSTM'den farklı olarak sadece iki kapısı vardır. Bunlar, unutulacak geçmiş bilginin miktarına karar vermek için kullanılan sıfırlama kapısı ve atılacak veya eklenecek bilgilere karar veren güncelleme kapısıdır (Sachin vd., 2020). Bu çalışmada model, yinelenen sinir ağındaki gradyanların yok olması probleminde dolayı GRU çözümü ile eğitilmiş ve eğitimde GRU katmanları oluşturulmuştur.

$$h_t^j = (1 - z_t^j)h_{t-1}^j + z_t^j h_{tt}^j \quad (3.1)$$

$$z_t^j = \sigma(W_z x_t + U_z h_{t-1})^j \quad (3.2)$$

$$h_{tt}^j = \tanh(W x_t + U(r_t \odot h_{t-1}))^j \quad (3.3)$$

$$r_t^j = \sigma(W_r x_t + U_r h_{t-1})^j \quad (3.4)$$

Burada;  $x_t$ , giriş vektörünü;  $h_t^j$ , önceki  $h_{t-1}^j$  aktivasyonu ile aday  $h_{tt}^j$  aktivasyonu arasında t anındaki doğrusal enterpolasyonu (çıkış vektörünü);  $z_t^j$ , güncelleme geçidi vektörünü;  $U$ ,  $W_z$ ,  $W_r$ ,  $W$ , ağırlık matrislerini;  $\sigma$ , aktivasyon fonksiyonunu;  $r_t$ , sıfırlama geçidi vektörünü;  $\odot$ , vektörler arası elementer çarpımı ifade eder. Bu eşitlikler kullanılarak, modelin geçmiş bilgileri hatırlaması, anlık bilgiyi işlemesi ve önemli bilgileri vurgulaması sağlanır (Sachin vd., 2020).

# 4

## DOĞAL DİL İŞLEME

---

### 4.1 Doğal Dil İşleme

Diller, makine dili ve insanlar tarafından kullanılan doğal dil olarak ikiye ayrılmaktadır. Doğal dillerin makineler tarafından algılanması ve gerektiğinde diller arası bağlantının kurulup çözümlenmesini doğal dil işleme sağlamaktadır (Şeker, 2015). Doğal dil işleme; yapay zekâ alanında bilgisayarların insan dilini anlama, yorumlama ve bu dili kullanarak metinleri veya konuşmaları analiz etme kabiliyetlerini geliştiren bir alt daldır. Bu teknoloji, metinleri sadece kelime düzeyinde anlamakla kalmayıp, cümlelerin ve paragrafların içerdiği anlamı da çözümlmeyi hedefler (Häberle vd., 2019). Doğal dil işleme günümüzde yaygın olarak diller arası çeviri, mobil telefonlardaki akıllı asistanlar ve çözüm odaklı kullandığımız ChatGPT gibi yapay zekâ asistanlarının temelinde kullanılır (Adalı, 2012). Doğal dil işleme, metinlerin analizlerde kullanılabilir hale getirilmesini sağlayan dil işleme tekniklerini içermektedir (Şeker, 2015).

Hazırlanan veri setine duygu analizi uygulandığında metnin duygusal olarak ne ifade ettiğinin çıkarımı yapılır. Duygu analizi temelde bir metnin pozitif, negatif veya nötr duygularla yazılmış olma durumunu ortaya çıkarır (Albayrak vd., 2017). Pozitif ve negatif duygu çıkarımı yapılabildiği gibi mutlu, üzgün, şaşkın gibi duygu ifadelerinin çıkarımı da yapılabilir (Bhardwaj vd., 2015). Duygu analizinde metin, öncelikle alt kelime birimlerine (tokens) ayrılmaktadır. Metinler içerisindeki kelimelerde edat, bağlaç ve zarfların belirlenmesi ve noktalama işaretlerin temizlenmesi duygu analizinin yapılabilmesi için gerekli işlem adımlarındandır. Böylece metinler üzerinde ayrıştırılmalar yapılır ve yüksek frekansa sahip alt kelime birimlerinin yanlılık (bias) oluşturmalarının önüne geçilerek metinde sadeleştirme sağlanır. Bu sayede metin içinde her bir kelime bir anlamı ifade edecek şekilde

listede yalın şekilde barınmış olur. Bu ön işlemler sonrasında çeşitli makine öğrenmesi sınıflandırma yöntemleri kullanılarak duygu analizi gerçekleştirilmektedir (İlhan & Sağaltıcı, 2020).

#### **4.1.1 Doğal Dil Araç Takımı (Natural Language Toolkit/NLTK)**

##### **Kütüphanesi**

NLTK (Natural Language Toolkit), Doğal Dil İşleme alanında kullanılan Python kütüphanelerinden biridir. NLTK, doğal dil metnini işleyebilen bir Python kütüphanesidir. NLTK, romanlar, haberler, ağ sohbet metinleri, film incelemeleri gibi çeşitli türlerde metin materyallerini içeren çok sayıda yerleşik derlem (corpus) sunar. NLTK, WordNet gibi derlem ve sözcük kaynağına erişim sağlayan bir arayüz sunar. Bu kütüphane, sınıflandırma, tokenizasyon, kök bulma (stemming), etiketleme, ayrıştırma (parsing) ve anlamsal çıkarım gibi metin işleme işlemleri için geniş bir araç seti sağlar. Ayrıca, endüstriyel düzeyde doğal dil işleme kütüphanelerine yönelik sarmalayıcılar (wrappers) sunarak, Python'un standart kütüphanesi ve diğer üçüncü taraf kütüphaneleriyle entegrasyon sağlar. Bu sayede, metin işleme sürecinde elde edilen sonuçların ikincil işlemleri gerçekleştirilebilir. NLTK, karmaşık metinlerin işlenmesi için sağlam bir altyapı sunarak, doğal dil işleme (NLP) alanındaki birçok uygulamanın geliştirilmesine olanak tanır (Wang & Hu, 2021).

#### **4.2 Duygu Analizi**

Doğal dil işlemenin öne çıkan kullanım alanlarından biri, duygu analizidir. Duygu analizi, literatürde fikir madenciliği olarak da adlandırılan bir araştırma alanıdır. Bu yöntem, metinlerdeki duygusal içerikleri, algıları ve görüşleri tespit etmeyi amaçlar. Özellikle çevrimiçi platformlarda paylaşılan ifadelerdeki anlamın, duygusal tonun ve görüşlerin sistematik bir şekilde analiz edilmesi için geliştirilmiştir. Duygu analizi, kamuoyu algısının belirli bir konuya yönelik genel eğilimlerini ortaya koyması nedeniyle sosyal medya verilerinin işlenmesinde yaygın olarak kullanılmaktadır. Bir metin üzerinde yapılan analizlerde, metnin kutupsallığının (polarite) belirlenmesi kritik bir rol oynar. Bu süreç, metnin olumlu, olumsuz veya nötr bir duygu taşıyıp taşımadığını ortaya koyar. Duygu analizi, özellikle bireylerin, toplulukların veya markaların algılarının değerlendirilmesi ve

karar destek sistemlerine entegre edilmesi açısından önemli bir yere sahiptir (Hourrane vd, 2018).

Duygu analizi, metin madenciliği kapsamında bir tür sınıflandırma uygulaması olarak değerlendirilir. Metinlerdeki varlıklar, kavramlar, ana fikirler ve kategoriler üzerinde yoğunlaşan duygu analizi sistemleri, doğal dil işleme tekniklerini ve makine öğrenmesi modellerini birleştirerek bu unsurlara duygusal değerler atar.

Bu teknoloji, özellikle büyük ölçekli şirketlerde veri uzmanlarının halkın belirli konulara yönelik algısını analiz etmesine, detaylı pazar araştırmaları yapmasına, marka ile ürün algısını izleyerek müşteri memnuniyetini artırmasına, afetler sonrasında insanların genel durumunu değerlendirmeye ve insanların herhangi bir olay hakkındaki genel düşüncelerini öğrenmek amacıyla kullanılabilir. Bu sayede, şirketler müşteri ihtiyaçlarını daha iyi anlayarak stratejik planlamalarını şekillendirebilirler. Kurum ve kuruluşlar ise, insanlara yardım planlaması yapma ve acil durumları hızlı ve etkili bir şekilde yönetme konusunda destek sağlayabilirler (Talan ve Aktürk, 2021).

Duygu analizi için sözlük tabanlı ve makine öğrenmesi tabanlı olmak üzere iki temel yaklaşım geliştirilmiştir. Sözlük tabanlı yöntemler, kelime ve ifadelerin duygu kutupluluğuna dayanır. Makine öğrenmesi yöntemleri, etiketlenmiş veri örneklerinden öğrenilen modelleri kullanır. Duygu analizi, bazı modellerde geleneksel yöntemlerin kullanılması durumunda zaman alıcı ve karmaşık olabilmektedir. Ancak derin öğrenme yöntemleri, manuel özellik mühendisliği gerektirmeden veri kümelerinden öğrenme yetenekleriyle duygu analizi çalışmalarında önemli avantajlar sunmaktadır.

Duygu analizinde makine öğrenmesi tabanlı modellerde ise temel zorluklardan biri, eğitim verisinin doğru şekilde etiketlenmesidir. Etiketleme süreci, bir metnin hangi duygu durumuna ait olduğunu belirlemek için dikkatli bir değerlendirme gerektirir ve bu süreç genellikle zaman alıcıdır. Özellikle büyük veri kümelerinde bu durum, sürecin daha da karmaşık hale gelmesine neden olabilir. Metinlerin etiketlenmesi için farklı yöntemler bulunmaktadır. En yaygın yöntemlerden biri, metinlerin insanlar tarafından okunarak değerlendirilmesi ve uygun duygu etiketinin manuel olarak atanmasıdır. Bu yöntem yüksek doğruluk sağlayabilir ancak oldukça zaman ve emek gerektirir. Alternatif olarak, yapay zekâ modelleri kullanılarak otomatik

etiketleme yapılabilir. Ancak bu durumda, modellerin eğitimine ve performansına bağlı olarak etiketleme kalitesinde farklılıklar ortaya çıkabilir. Diğer bir yaklaşım ise metinlerdeki belirli kelimeler üzerinden otomatik bir kodlama ile etiketleme yapmaktır. Örneğin, bir metinde "korku" kelimesinin geçmesi, o metni korku duygusuna ait olarak etiketlemek için kullanılabilir. Ancak bu yöntemde, "korku" kelimesi geçmesine rağmen farklı bir anlam taşıyan cümlelerin yanlış şekilde etiketlenme riski bulunmaktadır. Bu nedenle, kelime tabanlı etiketlemede bağlamın dikkatlice incelenmesi ve doğruluk kontrolünün yapılması gereklidir. Sonuç olarak, metin etiketleme süreci, duygu analizi için temel bir adımdır ve kullanılan yöntemlere bağlı olarak çeşitli zorluklar içermektedir. Bu zorlukların üstesinden gelmek için manuel değerlendirme, yapay zekâ modelleri ve otomatik kodlama yöntemleri arasında bir denge kurularak süreç optimize edilebilir.

Duygu analizi alanında geleneksel ve makine öğrenmesi yöntemlerinin yanı sıra, denetimli ve denetimsiz öğrenme yöntemleri de yer almaktadır. Denetimli öğrenme yöntemleri, etiketlenmiş veri setleri kullanılarak model eğitimi sağlar ve bu yöntem, etiketli veri üzerinde doğruluğu yüksek sonuçlar elde etmek için kullanılır. Diğer taraftan, denetimsiz öğrenme yöntemleri etiketsiz veriler üzerinden veri yapısını keşfetmek ve gruplama yapmak için kullanılır. Bu yöntem, özellikle verinin etiketlenmemiş olduğu durumlarda etkili bir çözüm sunar. Dolayısıyla, her iki öğrenme yaklaşımı da duygu analizi gibi farklı veri türleri ve analitik ihtiyaçlar doğrultusunda çözümler sunmaktadır (Lagrari & Elkettani, 2021).

# 5

## MODEL DOĞRULUK METRİKLERİ

---

### 5.1 Model Doğruluk Değerlendirmesi

Sınıflandırma performansı çeşitli metriklerle değerlendirilebilir. Sınıflandırma modellerinin test verileri üzerindeki performansını değerlendirmek için hata matrisi kullanılır. Hata matrisi tahmin sonuçlarını gösterir. Hata matrisi içerisindeki Doğru Pozitif (DP) (True Positive, TP), Doğru Negatif (DN) (True Negative, TN), Yanlış Pozitif (YP) (False Positive, FP) ve Yanlış Negatif (YN) (False Negative, FN), modelin tahminlerinin doğruluğunu belirlemede kullanılır. Bu değerlendirmeler, test edilen modelin, referans (ground truth) veriler üzerindeki başarısını gösterir.  $N$ , test edilen örneklerin toplam sayısıdır. DP, modelin doğru bir şekilde pozitif olarak sınıflandırdığı durumlardır. DN, modelin doğru bir şekilde negatif olarak sınıflandırdığı durumlardır. YP, modelin yanlış bir şekilde pozitif olarak sınıflandırdığı ve YN, modelin yanlış bir şekilde negatif olarak sınıflandırdığı durumlardır (Markoulidakis vd., 2021). Değerlendirmede doğruluk, kesinlik, duyarlılık ve F1 Skor gibi metrikler kullanılmıştır. Doğruluk ölçütü, doğru tahminlerin toplam test sayısına oranıdır ve başarıyı yansıtır.

$$\text{Doğruluk (Accuracy)} = \frac{DP + DN}{N} \quad (5.1)$$

Kesinlik (Precision) ölçütü, pozitif olarak sınıflandırılmış örnekleri değerlendirir. Model ne kadar yüksek bir kesinlikle çalışıyorsa, pozitif olarak sınıflandırılan örneklerin gerçekten pozitif olma olasılığı o kadar yüksektir. Kesinlik, doğru olarak sınıflandırılmış pozitif örneklerin, tahmin edilen toplam pozitif örneklere olan oranını ifade eder.

$$Kesinlik (Precision) = \frac{DP}{DP + YP} \quad (5.2)$$

Duyarlılık (Recall) ölçütü, pozitif olarak sınıflandırılmış örnekleri değerlendirir. Bu ölçü, doğru olarak sınıflandırılmış pozitif örneklerin, referans verisindeki toplam pozitif örneklere olan oranını ifade eder.

$$Duyarlılık (Recall) = \frac{DP}{DP + YN} \quad (5.3)$$

F1-Skor (F1-Score), kesinlik ve duyarlılık ölçütlerinin birlikte hesaplanmasıyla elde edilir. Bu ölçüt, kesinlik ve duyarlılık değerlerinin ağırlıklı harmonik ortalamasının alınmasıyla belirlenir. Yüksek F1-Skor, sınıflandırma performansının kalitesini ifade eder (Gündüz, 2013).

$$F1 \text{ Skor} = \frac{2 * Duyarlılık * Kesinlik}{Duyarlılık + Kesinlik} \quad (5.4)$$

## 5.2 Çapraz Doğrulama Yöntemi

Çapraz doğrulama, makine öğrenmesi ve istatistiksel modellemede modelin doğruluğunu ve genelleme kapasitesini değerlendirmek için yaygın olarak kullanılan bir tekniktir. Bu yöntem, öğrenme setinin eğitim ve test setlerine bölünmesini sağlar. Bu bölme işleminde her örnek, test seti olarak kullanılmak üzere sırasıyla birden fazla kez seçilir. Çapraz doğrulama, modelin farklı veri alt kümeleri üzerinde nasıl performans gösterdiğini değerlendirerek, aşırı uyum (overfitting) sorununu minimize etmeyi amaçlar. Modelin farklı veri setlerine uygulandığında benzer sonuçlar üretmesi, modelin daha sağlam ve genellebilir olduğunu gösterir.

K-katlı çapraz doğrulama, veri kümesinin k eşit alt kümeye bölünmesini içerir. Bu k alt kümelerden her biri, sırasıyla doğrulama seti olarak kullanılırken, geriye kalan alt kümeler eğitim seti olarak kullanılır. Bu işlem k defa tekrarlanır ve her tekrarda farklı bir alt küme test için seçilir, böylece her veri örneği doğrulama setinde bir kez yer almış olur. Bu metodun temel avantajı, tüm veri setinin modelin hem eğitiminde hem de testinde kullanılabilmesidir. Sonuçlar, her iterasyon sonunda

elde edilen performans metriklerinin ortalaması alınarak daha güvenilir bir değerlendirme sağlanır (Berrar, 2019).

Çapraz doğrulamanın amacı, modelin farklı veri alt kümelerinde nasıl çalıştığını gözlemleyerek daha doğru bir performans değerlendirmesi yapmaktır. Bu teknik, modelin çeşitli parametrelerle eğitim ve test edilmesi, modelin genel performansının daha doğru bir şekilde belirlenmesini sağlar. Bu yöntem, veri setindeki farklı çeşitliliği göz önünde bulundurarak sağlam bir model oluşturulmasına katkı sağlar.

Bu doğrulama yöntemi, özellikle küçük veri setleriyle çalışan modellerde, modelin doğruluğunu değerlendirmenin yanı sıra aşırı uyumu (overfitting) engellemeye de yardımcı olur. Çapraz doğrulama ile elde edilen ortalama sonuçlar, modelin genelleme yeteneğini daha doğru bir şekilde yansıtarak, veri setindeki belirli örnekler üzerindeki hatalı tahminlerin etkisini azaltır. Bu yöntem, yalnızca doğruluk değil, aynı zamanda diğer değerlendirme metrikleri (örneğin, F1 skoru, hata oranı) üzerinden de modelin performansını ölçmek için kullanılabilir (Wang vd., 2024).

# 6

## MEKÂNSAL ANALİZ

---

### 6.1 Mekânsal Analiz

Mekânsal analiz, coğrafi verilerin; konumla ilişkilendirilmiş verilerin incelenmesi sürecidir. Bu analiz türü, verilerin yalnızca niceliksel özelliklerine değil, aynı zamanda mekânsal yerleşimlerine de dikkat edilmesini gerektirir. Mekânsal veri analizi, coğrafi nesnelerin veya olayların konumlarına dayalı özelliklerini anlamak amacıyla kullanılan tekniklerden oluşur. Bu analizlerde, harita üzerindeki belirli bir alan içindeki verilerin dağılımı, mekânsal ilişki ve kümelenme gibi faktörler araştırılır. Verilerin yerleşim biçimi, mekânsal değişimleri, örneğin yoğunluk ve kümelenme alanlarını tanımlar (Bailey & Gatrell, 1995).

Mekânsal analizler, verilerin konumsal yönlerini araştırır. Aynı zamanda, mekânsal etkileşimlerin daha derinlemesine anlaşılmasını sağlar. Bir bölgedeki yüksek ya da düşük değerlerin kümelemesi, mekânsal değişimlerin anlaşılması ve bu kümelenmelerin nedenleri üzerine odaklanır. Araştırmanın amacı doğrultusunda, mekânsal analiz için en uygun yöntemler seçilir. Seçilen yöntem veri türüne, örneklem tasarımına ve ölçüm türüne bağlı olarak değişebilir. Mekânsal analiz, coğrafi verilerin doğru yorumlanmasını sağlayarak, farklı alanlarda yerel değişimleri ve ilişkileri anlamada kullanılır (Ergüt & Çilingirtürk, 2022).

### 6.2 Sıcak Nokta Analizi

Sıcak nokta (hot spot) analizi (Getis-Ord  $G_i^*$ ), her bir özneliğin çevresindeki diğer özneliklerle olan ilişkisini değerlendirerek mekânsal kümelenme desenlerini araştırır. Bir noktanın istatistiksel olarak anlamlı bir sıcak nokta olabilmesi için, bu noktanın yüksek bir değere sahip olması gerektiği gibi, aynı zamanda çevresindeki

diğer noktaların da yüksek değerlere sahip olması gerekir. Elde edilen z-istatistiği ve p-değeri, yüksek veya düşük değerlere sahip nesnelerin mekânsal olarak kümelenmesini gösterir (URL-5). Eşitlik 6.1'de verilen  $G_i^*$  yerel istatistiği z-istatistiği değeridir. Örneklemdaki tüm noktalar için  $G_i^*$  hesaplanarak, çalışma alanındaki (belli anlamlılık düzeylerinde) sıcak ve soğuk noktalar belirlenir. Pozitif ve belli bir anlamlılık düzeyinde kritik z değerinden büyük  $G_i^*$  değerlerine sahip noktalar sıcak nokta; negatif ve belli bir anlamlılık düzeyinde kritik z değerinden küçük  $G_i^*$  değerlerine sahip noktalar ise soğuk noktaları oluşturur. Bunun dışında kalan durumlar ise rastlantısal dağılımı gösterir (Çubukçu, 2015).

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j}x_j - \underline{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}} \quad (6.1)$$

$$\underline{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (6.2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\underline{X})^2} \quad (6.3)$$

Burada; n, örneklemdaki nokta sayısını;  $x_j$ , j noktasına ait değişken değerini;  $w_{i,j}$ , i noktası ve j noktası arasındaki yakınlık ilişkisini veren mekânsal ağırlık değerini;  $\underline{X}$ , değişkene ait ortalama değeri belirtir.

### 7.1 Veri

#### 7.1.1 Çalışma Alanı

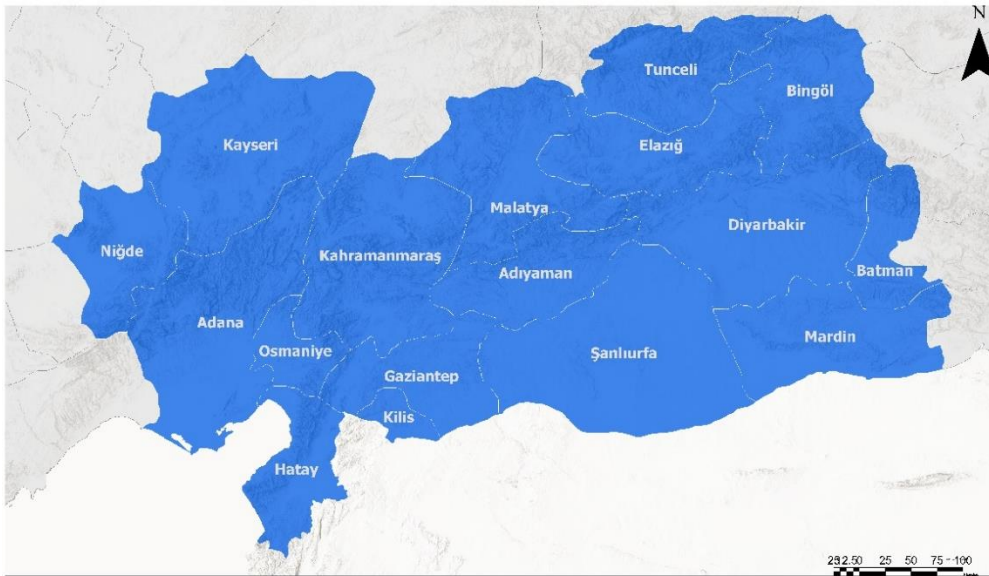
Bu çalışmanın araştırma alanı, 6 Şubat 2023 tarihinde meydana gelen Kahramanmaraş depremleri sonrasında bölgedeki duygu durumunun incelenmesini kapsamaktadır. Çalışma alanı, Kahramanmaraş ve çevresindeki illeri içermekte olup, Şekil 7.1’de gösterildiği gibi Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli illerinden oluşmaktadır. Tablo 7.1 ise çalışma alanlarının 2023 yılına ait nüfus ve yüzölçümü bilgilerini içermektedir (URL-7, URL-8).

**Tablo 7.1** Çalışma alanının nüfus ve yüz ölçümü bilgileri

İl	Nüfus	Yüz ölçümü (km <sup>2</sup> )
Adana	2.270.298	13844
Adıyaman	604.978	7337
Batman	647.205	4477
Bingöl	285.655	8004
Diyarbakır	1.818.133	15101
Elâzığ	604.411	9383
Gaziantep	2.164.134	6803
Hatay	1.544.640	5524

**Tablo 7.1** Çalışma alanının nüfus ve yüz ölçümü bilgileri (devamı)

<b>İl</b>	<b>Nüfus</b>	<b>Yüz ölçümü (km<sup>2</sup>)</b>
<b>Kahramanmaraş</b>	<b>1.116.618</b>	<b>14520</b>
<b>Kayseri</b>	<b>1.445.683</b>	<b>16970</b>
<b>Kilis</b>	<b>155.179</b>	<b>1412</b>
<b>Malatya</b>	<b>742.725</b>	<b>12259</b>
<b>Mardin</b>	<b>888.874</b>	<b>8780</b>
<b>Niğde</b>	<b>377.080</b>	<b>7234</b>
<b>Osmaniye</b>	<b>557.666</b>	<b>3320</b>
<b>Şanlıurfa</b>	<b>2.213.964</b>	<b>19242</b>
<b>Tunceli</b>	<b>89.317</b>	<b>7582</b>



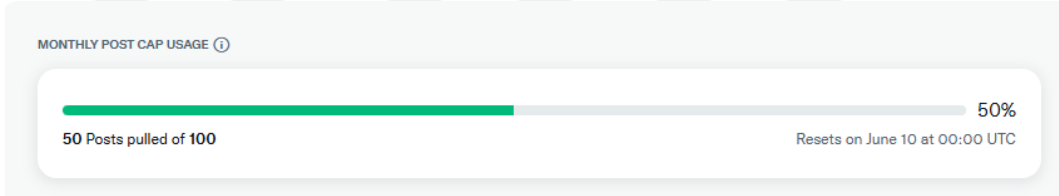
**Şekil 7.1** Çalışma alanı

### 7.1.2 Sosyal Medyadan Veri Kazıma

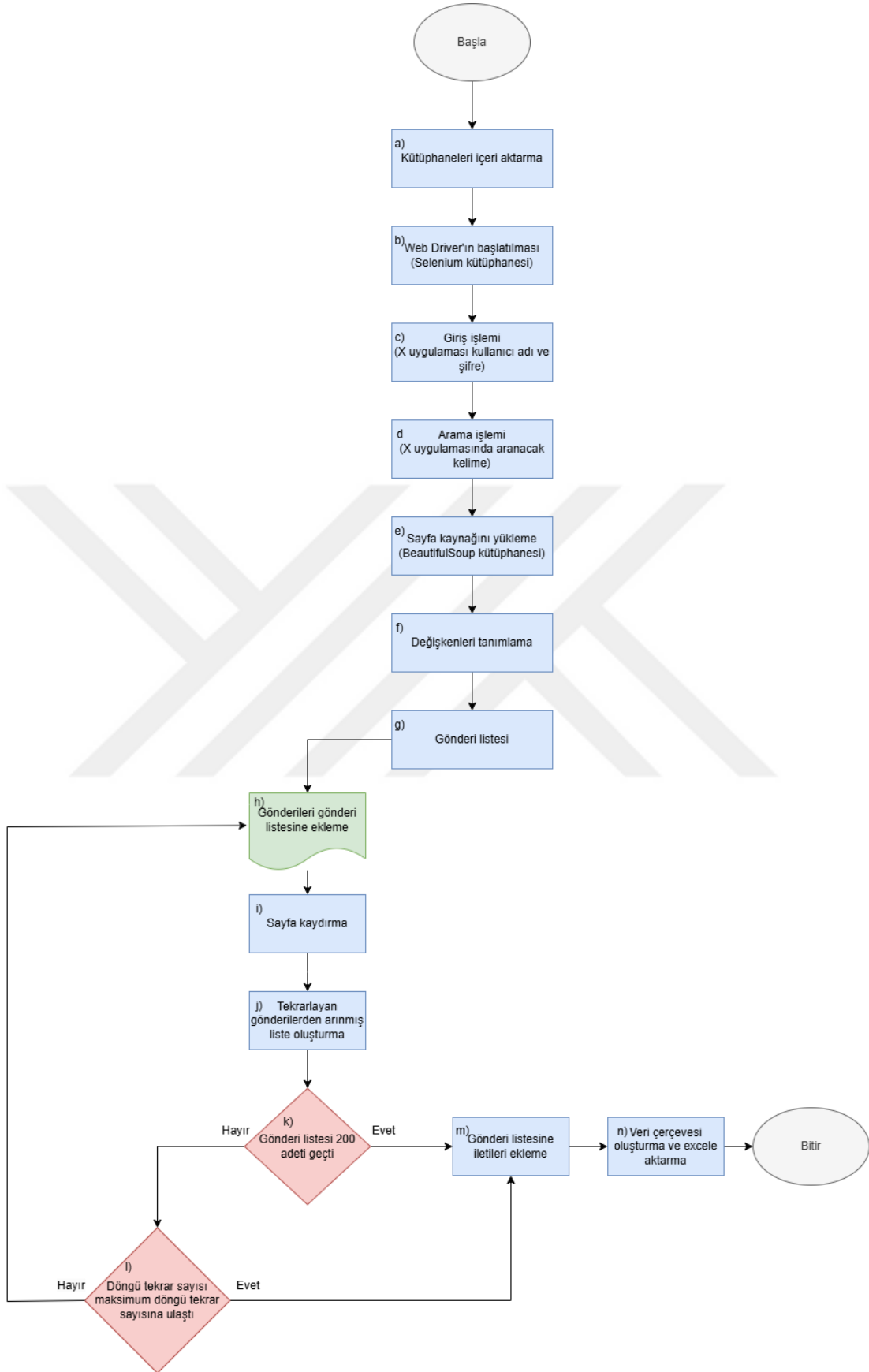
Python programlama dili ile yazılan veri kazıma işleminin akış diyagramı Şekil 7.3'te gösterilmiştir. X platformundan veri kazıma işleminde tarayıcı sunucu üzerinden işlemlerin optimizasyonu için Selenium kütüphanesi, sayfa kaynağını indirmek için BeautifulSoup ve kazıma sonucu elde edilen verileri bir veri çerçevesine aktarmak için Pandas kütüphaneleri eklenir (Şekil 7.3 a). Tarayıcı sürücününün başlatılması işlemi için Google Chrome sunucularına bağlanılarak X uygulamasına ait internet sayfasına bağlanılır (Şekil 7.3 b). Oturum açma işlemi için kullanıcı adı ve şifrenin koda eklenmesi ile internet sayfasında uygun girişlere gerekli bilgilerin girilmesi işlemi gerçekleştirilir (Şekil 7.3 c). 6 Şubat 2023 anahtar kelimesi (örneğin, “6 Şubat 2023” geocode:39.1070000,39.5472000,30km and “6 Şubat 2023”, since:2023-06-01 until:2023-09-30) enlem, boylam, arama çapı ve tarih aralığı gibi bilgilerle birlikte kullanılarak kodun arama sonuçlarını açması işlemi gerçekleştirilir (Şekil 7.3 d). Sayfa kaynağını yüklemek amacıyla BeautifulSoup kütüphanesi kullanılarak XML formatında sayfanın içeriği indirilir. Bu sayede istenilen anahtar kelime ile elde edilen arama sonuçlarına ait sayfa kaynağı sağlanır. Sayfa kaynağı içerisinde XPath'ler sayesinde istenilen metin ve konum bilgilerinin elde edilmesi için ortam hazırlanır (Şekil 7.3 e). X uygulamasından alınacak iletilerin bir liste içine eklenmesi için ve döngülerde kullanılan döngü tekrar başlangıç sayısı ve maksimum döngü tekrar sayısının tanımlanması işlemi gerçekleştirilir (Şekil 7.3 f). İlk sayfadaki iletilerin XPath'leri alınır ve sayfa kaynağında bulunan ilk sayfadaki iletiler listeye eklenerek gönderi listesi oluşturulur (Şekil 7.3 g). İletilerin gönderi listesine ekleme işlemi gerçekleştirilir. İlk gönderiden başlayarak sayfada bulunan tüm iletilerin tek tek listeye eklenmesi için döngü oluşturulur (Şekil 7.3 h). İlk sayfadan alınan iletilerden sonra diğer sayfalardan da iletilerin alınması için otomatik kaydırma işlemi gerçekleştirilir (Şekil 7.3 i). Tekrarlayan gönderilerden arınmış bir liste oluşturmak amacıyla yeni bir gönderi listesi oluşturulur. Yeni gönderi listesine, eski gönderi listesinde birden fazla aynı olan iletilerden yalnızca birinin alınması sağlanır. Böylece tekrarlı iletilerden arındırılır (Şekil 7.3 j). Gönderi listesinin 200 adet iletiyi geçip geçmediği kontrol edilir. 200 adet ileti listeye alındıktan sonra kod çalışmayı durdurur (Şekil 7.3 k). Eğer 200 adetten az ileti varsa döngü 10 kez çalıştırılır. Döngü sayısı maksimum döngü sayısına ulaştığında kod çalışmayı durdurur (Şekil

7.3 l). Döngüden gelen her ileti gönderi listesine eklenir (Şekil 7.3 m). Tekrardan arındırılmış gönderi listesi bir veri çerçevesi içerisine yazdırılır (Şekil 7.3 n). Bu süreçte veri seti, her iki model için de optimize edilerek hem iki hem de dört etiketli sınıflandırma için uygun hale getirilmiştir.

Bu çalışmada, veri kazıma işlemi ile elde edilen verilerin X API üzerinden bir erişim anahtarı (bearer token) kullanılarak da elde edilebildiği kontrol edilmiştir. Erişim anahtarı, bir sistemin veya kullanıcının bir API'ye (örneğin X API) erişmesine izin vermek için kullanılan dijital bir anahtardır. Bu anahtar, bir şifre gibi çalışır; API'ye yapılan her istekte sunulur ve eğer geçerliyse, sistem bu isteği kabul eder. Ücretsiz kullanımda ayda en fazla 100 gönderi (post) alınabilir. Şekil 7.2'de aylık gönderi limitini gösteren X geliştirici panelinden bir parça görülmektedir (URL-10).



Şekil 7.2 X geliştirici kontrol paneli (URL-10)



Şekil 7.3 Veri kazıma kod akış diyagramı

## 7.2 İki Etiketli Duygu Analizi

### 7.2.1 1. Adım: Duygu Analizi Modeli Oluşturma-Eğitme-Test Etme

X platformundan 755 adet iletinin kazıma işlemi ile elde edilen veri kümesi, “6 Şubat 2023” anahtar kelimesini içeren metinlerden ve konumlardan oluşmaktadır. Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli illerinin konumlarında gönderilen iletiler ile veri seti oluşturulmuştur.

İletilerden oluşan veri setinin kalitesinin iyileştirilmesi adına, elde edilen verilerin kontrolü gönderilerin özelliklerine göre yapılmıştır. Tablo 7.2’de bu çalışmada gönderilere ait özellikler verilmiştir. Çalışmada kullanılan iletiler; konusu, beğeni sayısı, gönderi sahibinin benzer gönderilerinden sadece birinin alınması ve gönderi tarih aralıkları gibi özelliklerine göre filtrelenerek seçilmiştir. Bu kontroller sonucunda, iki etiketli modelin sınıfları doğru öğrenebilmesi adına iletiler Tablo 7.2’deki kriterlere göre filtrelenerek 685 adete indirilmiştir.

**Tablo 7.2** Gönderiye ait özellikler ve açıklamaları

<b>Gönderiye ait özellikler</b>	<b>Açıklama</b>
<b>Gönderi konusu</b>	<b>6 Şubat 2023 depremleri hakkında bilgi ve görüş içermesi</b>
<b>Gönderi beğeni sayısı</b>	<b>Gönderilerin genel kullanıcılar arasında desteklenme durumu</b>
<b>Gönderi sahibi</b>	<b>Aynı gönderinin aynı kişiden birden fazla atılması</b>
<b>Gönderi tarihi</b>	<b>Haziran 2023 – Eylül 2023</b>

Seçilen iletiler ikili olarak olumlu veya olumsuz düşünce belirtmesi durumuna göre etiketlenmiştir. Etiketlenen bu iletiler referans veri seti olarak belirlenmiştir. Referans veri seti duygu analizi modelinin eğitiminde kullanılmak üzere hazırlanmıştır.

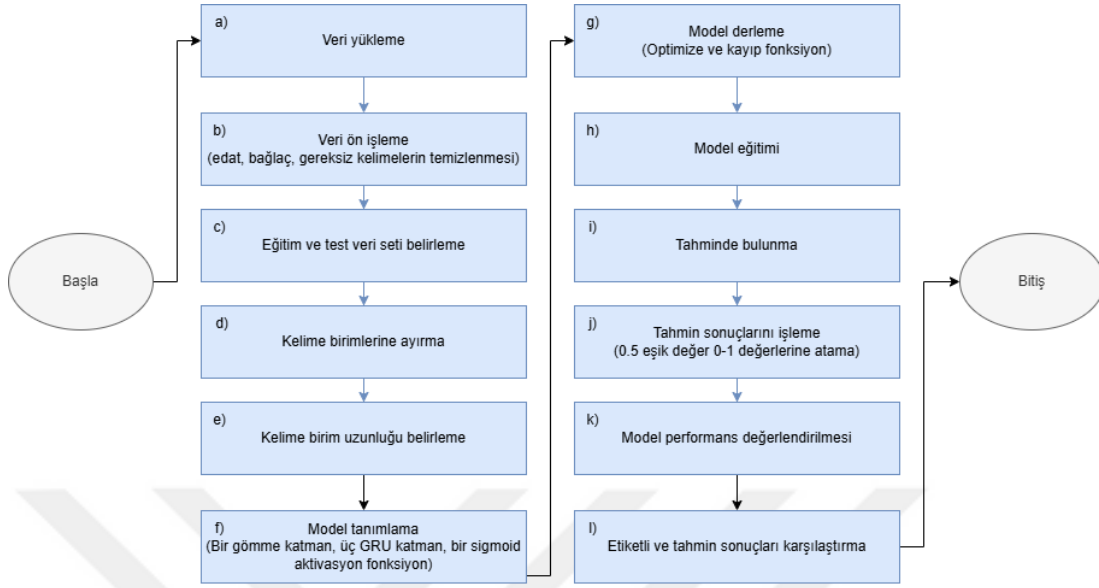
Bu çalışmada oluşturulan duygu analizi modeline ait akış diyagramı ise Şekil 7.4’te gösterilmiştir. Bu aşamada, iletileri ve olumlu-olumsuz anlamda olduğunu ifade

eden etiketleri yükleme işlemi gerçekleştirilir (Şekil 7.4 a). Veri ön işleme amacıyla edat, bağlaç, zarf ve gereksiz kelimeler (stop words) gibi cümlede bulunan fakat tek başına anlamlı olmayan terimlerden verinin arındırılması işlemi gerçekleştirilir (Şekil 7.4 b). K-kat çapraz doğrulama (k-fold cross validation) yöntemi uygulanarak eğitim ve test veri setleri belirlenir (Şekil 7.4 c). Kelime haznesi belirlenir (en sık geçen 10000 kelime belirlenir). Veri kümesindeki metin verisi kullanılarak bir sözlük oluşturulur. Veri setindeki metinler kelime birimlerine ayrılır. Sözlük içerisinde her bir kelimeye benzersiz bir sayı atanır. Bu sayede metin verisi sayısal bir formata dönüştürülür (Şekil 7.4 d). Eşitlik 7.1 ile belirli bir ileti boyutunu belirlemek için kullanılacak maksimum kelime birimi sayısı (maks\_tokens) hesaplanır.

$$\text{maks\_tokens} = \text{ort}(\text{num\_tokens}) + 2 * \text{std}(\text{num\_tokens}) \quad (7.1)$$

Kelime birim sayılarının istatistiksel özellikleri kullanılarak bir maksimum sınırlama getirme amaçlanır. Veri setindeki iletilerin kelime birim sayıları belirlenen eşik değerine göre eşitlenir. Bu işlem, bazı iletilerin sonlarına sıfırlar ekleyerek gerçekleştirilir. Bu sayede tüm iletiler aynı boyuta getirilir. Veri seti, sinir ağı modelinde eğitim ve tahmin etme aşamaları için hazır hale getirilir (Şekil 7.4 e). Model oluşturmak amacıyla modelin ilk katmanına bir gömme katmanı (embedding layer) eklenir. Bu katman, metin verilerini temsili vektörlere dönüştürmek için kullanılır. Üç adet GRU katmanı eklenir. İkili sınıflandırmalı bir sonuç almak için yoğun katman içinde sigmoid aktivasyon fonksiyonu kullanılır (Şekil 7.4 f). Modelin daha iyi öğrenmesi için büyük veri setleri ve karmaşık modellerle çalışırken etkili olan “Adam” optimizasyonu kullanılır. Modelin hazır hale getirilmesi için kayıp fonksiyonu olarak ikili sınıflandırma görevi seçilir. Bu fonksiyon, modelin tahminlerinin referans etiketlerden ne kadar uzak olduğunu ölçer (Şekil 7.4 g). Veri seti, k-kat çapraz doğrulama yöntemi kullanılarak 5 katmana bölünür. Her bir iterasyon için eğitim ve test kümeleri oluşturulur. Model veri üzerinde eğitilir (Şekil 7.4 h). Model, test verileri üzerinde tahminler yapar (Şekil 7.4 i). Tahmin olasılıklarında eğer bir tahmin olasılığı 0,5'ten büyükse olumlu (1) sınıfına atama, küçükse olumsuz (0) sınıfına atama işlemi gerçekleştirilir (Şekil 7.4 j). Bu aşamadan sonra model performans değerlendirilmesi işlemi gerçekleştirilir. Performansın değerlendirilmesi için doğruluk, kesinlik, duyarlılık

ve F-1 Skor sonuçları hesaplama işlemi gerçekleştirilir (Şekil 7.4 k). Etiketli veri ve model tahmin sonuçları il bazında karşılaştırılır (Şekil 7.4 l).



Şekil 7.4 İki etiketli duygu analizine ait kod akış diyagramı

## 7.2.2 2. Adım: Duygu Analizinde Referans ve Tahmin Verileri Üzerinden Sıcak Nokta Analizi

Duygu analizi modeli sonucunda elde edilen Haziran-Eylül 2023 zaman dilimi içerisindeki ileti sayıları, gönderilerin konum bilgileri ve Türkiye İstatistik Kurumu (TÜİK) 2023 illere göre elde edilen nüfus sayıları CBS ortamına (ArcGIS Pro kullanılarak) aktarılmıştır (URL-7). Öncelikle ileti sayılarının nüfusa göre normalize edilmesi işlemi gerçekleştirilmiştir. Böylelikle, 1.000.000 kişi başına düşen ileti sayıları elde edilmiştir. Bu aşamada ileti sayılarının bölgedeki nüfusa oranla seyreklik-yoğunluk durumlarının tutarlı olması amaçlanmıştır. 6 Şubat 2023 deprem bölgeleri için normalize edilmiş olan iletilerin koroplet (renk tonlu) haritası oluşturulmuştur (Harita 8.1). Haritada, 6 Şubat 2023 depreminden etkilenen illere ait nüfus bilgisi, mutlak değerlere dayalı orantılı işaretler ile görselleştirilmiştir. Her bir sembolün büyüklüğü, ilgili ilin toplam nüfus miktarı ile orantılı olacak şekilde oluşturulmuştur. Bu sayede harita üzerinde nüfusun yoğun olduğu illerin üzerindeki işaretleme çevresindeki illerin işaretinden daha büyük olarak gösterilmiştir.

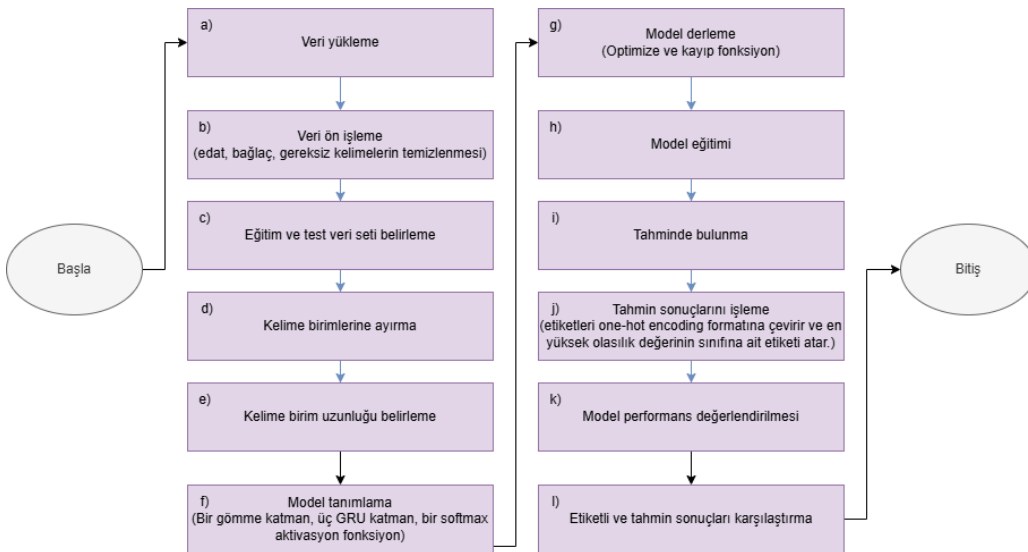
## 7.3 Dört Etiketli Duygu Analizi

### 7.3.1 1. Adım: Duygu Analizi Modeli Oluşturma-Eğitme-Test Etme

Bu çalışmada, X platformundan elde edilen veri kümesi kullanılmıştır. Dört etiketli veri setinde duygu kategorileri arasında daha fazla sayıda ileti bulunmasını sağlamak amacıyla, veri setindeki benzer iletiler gibi filtreleme işlemleri uygulanmamış ve tüm iletiler modele dahil edilmiştir. Bu yaklaşım, sınıflar arası dengeyi koruyarak duygu çeşitliliğini yansıtmayı hedeflemektedir.

Seçilen iletiler dört etiketli olarak korku anlamı taşıyorsa (0), üzüntü anlamını taşıyorsa (1), öfke anlamını taşıyorsa (2), empati anlamını taşıyorsa (3) düşünce belirtmesi durumuna göre etiketlenmiştir. Etiketlenen bu iletiler referans veri seti olarak belirlenmiştir. Referans veri seti duygu analizi modelinin eğitiminde kullanılmak üzere hazırlanmıştır (Şekil 7.5 a).

Bu çalışmada oluşturulan duygu analizi modeline ait akış diyagramı ise Şekil 7.5'te gösterilmiştir. Bu aşamada, iki etiketli duygu analizi modelinde uygulanan adımlar tekranlanmıştır. İki etiketli duygu analizi modelinde izlenen adımlardan farklı olarak, Şekil 7.5 (j)'de sunulan çok etiketli sınıflandırma sürecinde, etiketleme aşamasında bire bir kodlama (one-hot encoding) yöntemi kullanılmıştır. Duygu analizi modeli tarafından üretilen tahmin olasılıkları arasından en yüksek değere sahip olan sınıfa, ilgili örnek atanmıştır. Bu yaklaşım, çok etiketli duygu analizinde karar mekanizmasının olasılık temelli en güçlü sınıfa yönlendirilmesini sağlamaktadır.



Şekil 7.5 Dört etiketli duygu analizine ait kod akış diyagramı

### 7.3.2 2. Adım: Duygu Analizinde Referans ve Tahmin Verileri Üzerinden Sıcak Nokta Analizi

Duygu analizi modeli sonucunda elde edilen Haziran-Eylül 2023 zaman dilimi içerisindeki ileti sayıları, gönderilerin konum bilgileri ve Türkiye İstatistik Kurumu (TÜİK) 2023 illere göre elde edilen nüfus sayıları CBS ortamına (ArcGIS Pro kullanılarak) aktarılmıştır (URL-7). Dört etiketli duygu analizi modeliyle elde edilen tahmin verileri ve gerçek etiketli ileti sayıları, Türkiye'nin 17 ilinin verileriyle birlikte ArcGIS ortamına aktarılmıştır. Bu iller için korku, üzüntü, öfke ve empati sınıflarına ait ileti sayılarının aritmetik ortalaması alınarak bir eşik değeri belirlenmiştir. Batman, Bingöl, Kilis, Mardin, Niğde ve Tunceli illeri düşük ileti sayıları nedeniyle sıcak nokta analizine dahil edilmemiştir. Sıcak nokta analizi; Adana, Gaziantep, Hatay ve Malatya gibi illerde duygulara ait kümelenmeler göstermiştir. Ancak, öfke sınıfında kümelenme gözlemlenmemiştir. Bu sonuçlar, modelin empati, üzüntü ve öfke sınıflarında yüksek doğruluk gösterdiğini, ancak korku sınıfında rastlantısal dağılımlar sergilediğini ortaya koymuştur.

## 8 BULGULAR

---

Veri kümesi, X platformundan 755 adet iletinin kazıma işlemi ile elde edilmiştir. Veri kazıma aşamasında “6 Şubat 2023” anahtar kelimesini içeren ve Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli illerinden gönderilen iletilerin elde edilmesi sağlanmıştır. BeautifulSoup ve Selenium kütüphanesi ile kazıma işlemi yapıldıktan sonra Pandas kütüphanesi ile bir veri çerçevesi oluşturulmuştur. Veri kazıma sonucu elde edilen veri setinin iyileştirilmesi adına, elde edilen veriler kontrolden geçirilmiştir. Veriler birinci uygulamada ikili olmak üzere olumlu ve olumsuz düşünce belirtmesi durumuna göre etiketlenmiştir. İkinci uygulamada ise dört etiketli olmak üzere korku, üzüntü, öfke ve empati düşünce belirtisine göre etiketlenmiştir. Etiketlenen veri setleri duygu analizi modellerinin eğitim aşamasında kullanılmıştır. Referans veriler eğitim sırasında eğitim verisi ve test verisi olmak üzere ikiye ayrılmaktadır. Bu çalışmada, duygu analizinde modellerinde kullanılan sınıflandırmada veriye k-kat çapraz doğrulama uygulanmış ve performansın değerlendirme sonuçları belirlenmiştir. Modeller, verilen veri setini k-kat çapraz doğrulama yönteminden geçirerek eğitimi bitirmiştir. Toplamda 755 örnekten oluşan veri seti için, veri setinin boyutu ve dengeli dağılımı göz önünde bulundurularak  $k = 5$  alınmıştır. Burada amaç modellerin eğitimini doğru bir şekilde tamamlamış olduğunu test etmektir. İki etiketli duygu analizi modeli eğitimini %87 test doğruluğu ile tamamlamıştır. Dört etiketli duygu analizi modeli ise eğitimini %79 test doğruluğu ile tamamlamıştır.

Eğitilmiş duygu analizi modellerinin tahminler yapması amacıyla kapı yinelemeli birimler derin öğrenme ağları ile veriler sınıflandırılmıştır. Veriler sınıflandırılırken

ilk olarak sinir ağına girdi olarak bir ileti eklenmiştir ve eklenen bu ileti, kelimelerin vektörlerinden oluşan kelime birimleriyle (tokenlerle) temsil edilmektedir. Gömme (embedding) katmanı, ileti içindeki kelimelere karşılık gelen vektörleri içermektedir. Bu vektörler, bir sonraki GRU katmanına çıkış olarak aktarılmaktadır. Bu çalışmada, veri sayısının az olmasından dolayı 3 adet GRU katmanı oluşturulmuştur. Kelime vektörleri, üç GRU katmanında işlenmiştir ve en son çıktı katmanındaki tek bir nörona iletilmiştir. Bu nöron, iki etiketli duygu analizi modelinde, değerleri tek bir sayıya dönüştürmüş ve ardından sigmoid aktivasyon fonksiyonundan geçirerek 0 ile 1 arasında bir değer üretmiştir. Bu değer, modelin analiz edilen ileti (X platformu gönderisi) hakkındaki tahminini temsil etmektedir. Sonuç 1'e yakınsa, model iletiyi olumlu olarak; 0'a yakınsa ise olumsuz olarak sınıflandırmıştır (Chen vd., 2020).

Dörtlü etiketleme modelinde ise çıktı katmanında, her sınıfa ait olasılık değerlerini döndüren softmax aktivasyon fonksiyonu kullanılmıştır. Bu fonksiyon, her bir etiket için olasılık değeri hesaplayarak sonuç üretmiştir. Model, bir iletiye ait dört etiketin olasılık değerlerini hesapladıktan sonra en yüksek olasılık değerine sahip etiketi metnin duygu analizi sonucu olarak belirlemiştir.

Modellerin eğitimleri sonucunda test verilerinin doğrulukları Bölüm 5'te belirtilen metriklere göre belirlenmiştir. Tablo 8.1'de iki etiketli duygu analizi modelinin hata matrisi sonucu verilmiştir. Tablo 8.2'de, iki etiketli duygu analizi modelinin k-kat çapraz doğrulama yöntemiyle derin öğrenme ağı kullanılarak eğitilmesinin ardından, modele verilen iletiler üzerinden elde edilen tahmin performansına ait metrik sonuçlar sunulmuştur.

İkili etiketleme modelinin sonuçlarına göre, referans veride olumsuz (0) olarak etiketli iken modelin de olumsuz (0) olarak tahmin ettiği ileti sayısı 475 olarak bulunmuştur. Referans veride olumsuz (0) olarak etiketli iken modelin olumlu (1) olarak tahmin ettiği bir sonuç bulunmamaktadır. Referans veride olumlu (1) olarak etiketli iken modelin olumsuz (0) olarak tahmin ettiği ileti sayısı 90 olarak bulunmuştur. Referans veride olumlu (1) olarak etiketli iken modelin de olumlu (1) olarak tahmin ettiği ileti sayısı 120 olarak bulunmuştur. Tablo 2'de hata matrisi olumsuz (0) sınıfını doğru bir şekilde tahmin etme yeteneğinin, olumlu (1) sınıfını tahmin etme yeteneğinden daha yüksek olduğu görülmektedir. Bu durum, kesinlik değerinin yüksek (1.00); duyarlılık değerinin ise düşük (0.57) olmasına yol

açmıştır. Bir başka deyişle, referans veride (gerçekte) olumsuz olan iletilerin tümü doğru sınıflandırılmasına rağmen, gerçekte olumlu olan iletilerin %57'si doğru olarak sınıflandırılmıştır. Olumsuz (0) olarak etiketli ileti sınıfının daha net ifadeler ve kısa cümleler içermesi (korkuyoruz, uyuyamıyoruz vb.) modelin bu sınıfı daha iyi öğrenmesine sebep olmuştur. Ancak, olumlu (1) olarak etiketli ileti sınıfının daha karışık ifadeler (örneğin, “üzgünüz ve yanınızdayız” gibi hem olumsuz hem de olumlu duygu durumlarını bir arada bulunduran) ve daha uzun cümleler içermesi, modelin bu sınıfı daha zor öğrenmesine neden olmuştur.

**Tablo 8.1** İki etiketli derin öğrenme ağında test sonucu modelin performansını değerlendiren hata matrisi

		Tahmin		
		Olumlu (Pozitif)	Olumsuz (Negatif)	Toplam
Referans (Gerçek)	Olumlu (Pozitif)	120 (DP)	90 (YN)	210
	Olumsuz (Negatif)	0 (YP)	475 (DN)	475
	Toplam	120	565	685

**Tablo 8.2** İki etiketli derin öğrenme ağında test sonucu model performansını değerlendiren metrik sonuçları

Metrik	Genel sonuç
Doğruluk	0.87
F1 Skor	0.73
Duyarlılık	0.57
Kesinlik	1.00

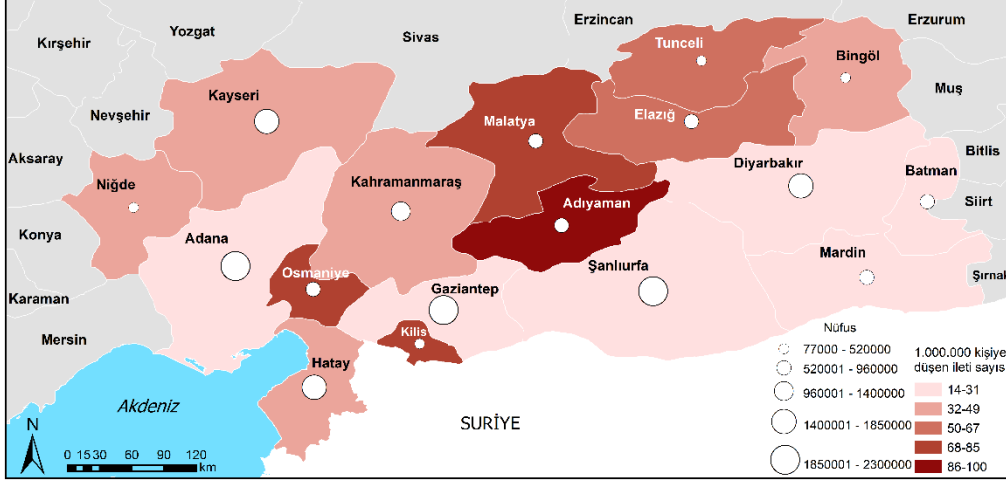
Eğitimini tamamlayan modellere test veri seti olarak Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli konumlarından elde edilen etiketli X iletileri gönderilmiştir. Tablo 8.3'te ise duygu analizi modelinin illere göre ayrılmış iletiler üzerinde tahmin sonuçlarına göre modeli değerlendiren metrik sonuçlar verilmiştir.

İki etiketli modelin derin öğrenme ağının Adıyaman, Hatay ve Kahramanmaraş illeri üzerindeki tahmin doğruluk sonuçları Tablo 8.3'te gösterilmiştir. Adıyaman ili için ortalama test tahmin doğruluğu %82, Hatay için %70 ve Kahramanmaraş ili için ise %90 olduğu tespit edilmiştir. İl bazında ortalama test doğruluklarının birbirine yakın olduğu ve %70-%90 arasında değiştiği görülmüştür. Test doğruluğunun düşük olduğu illerde nötr olarak da değerlendirilebilecek yorumların (olumlu ya da olumsuz ifadelerin daha belirsiz) olması nedeniyle eğitim modelinin daha düşük performans ile çalıştığı gözlemlenmiştir.

**Tablo 8.3** İki etiketli derin öğrenme ağının örnek ilçe veri setleri üzerinde tahmin performansını değerlendiren metrik sonuçları

<b>Metrik</b>	<b>Adıyaman</b>	<b>Hatay</b>	<b>Kahramanmaraş</b>
<b>Doğruluk</b>	<b>0.82</b>	<b>0.70</b>	<b>0.90</b>
<b>F1 Skor</b>	<b>0.74</b>	<b>0.58</b>	<b>0.87</b>
<b>Duyarlılık</b>	<b>0.82</b>	<b>0.70</b>	<b>0.90</b>
<b>Kesinlik</b>	<b>0.67</b>	<b>0.49</b>	<b>0.83</b>

İki etiketli duygu analizi modelinin tahmin ettiği iletilerin ve etiketli iletilerin sayıları, gönderilerin konum bilgileri ve nüfus verileri ArcGIS ortamına aktarılmıştır. Referans olumlu-olumsuz ve tahmin olumlu-olumsuz ileti sayıları ait olduğu ilin nüfusuna göre normalize edilmiştir. 6 Şubat 2023 deprem bölgeleri için nüfusa göre normalize edilmiş olan iletilerin koroplet (renk tonlu) haritası, Harita 8.1'de verilmiştir. Harita üzerine 6 Şubat 2023 depreminden etkilenen illerin nüfus bilgisi orantılı işaret haritası olarak sunulmuştur. Şehirlerin nüfus sayılarına oranla 6 Şubat 2023 anahtar kelimeli iletilerin en az paylaşımının Şanlıurfa, Diyarbakır ve Mardin illerinden gönderildiği ortaya çıkmıştır. Nüfusun daha az olduğu Kilis'te ise nüfusa oranla daha fazla iletilerin paylaşıldığı görülmektedir.



**Harita 8.1** 6 Şubat 2023 depreminden etkilenen illerin nüfusa oranla X uygulamasındaki depremle ilgili ileti sayıları

İllere ait sıcak nokta analizinde kullanılacak olan olumlu-olumsuz ileti sayıları için 17 ilin (Adana, Adıyaman, Batman, Bingöl, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Kilis, Malatya, Mardin, Niğde, Osmaniye, Şanlıurfa ve Tunceli) ileti sayılarının aritmetik ortalaması alınmıştır. Aritmetik ortalama sonucuna göre bir eşik değeri belirlenmiştir. Batman, Bingöl, Kilis, Mardin, Niğde ve Tunceli illerine ait ileti sayılarının ortalama ileti sayısından düşük olmasının verebileceği yanlışlık nedeni ile 6 il sıcak nokta analizinden muaf tutulmuştur.

**Tablo 8.4** İl bazında ileti sayıları

İl	İleti Sayısı
Tunceli	4
Kilis	11
Mardin	12
Niğde	12
Bingöl	14
Batman	15
Elâzığ	36
Osmaniye	46
Diyarbakır	49
Hatay	54

**Tablo 8.4** İl bazında ileti sayıları (devamı)

<b>İl</b>	<b>İleti Sayısı</b>
<b>Gaziantep</b>	<b>58</b>
<b>Şanlıurfa</b>	<b>58</b>
<b>Adana</b>	<b>59</b>
<b>Kahramanmaraş</b>	<b>59</b>
<b>Kayseri</b>	<b>65</b>
<b>Adıyaman</b>	<b>66</b>
<b>Malatya</b>	<b>67</b>

Elde edilen nüfusa ve ileti sayılarına göre normalize edilmiş referans ve tahmin olumlu-olumsuz ileti sayıları üzerinden Adana, Adıyaman, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Malatya, Osmaniye ve Şanlıurfa illeri için sabit mesafe bant (fixed distance band) ilişkisi ile sıcak nokta analizi yapılmıştır. Referans ve tahmin olumlu-olumsuz iletilerin sıcak nokta analiz sonucuna göre her iki ileti grubunun da aynı anlamlılık düzeylerinde ( $\alpha=0.10$ ;  $\alpha=0.05$ ;  $\alpha=0.01$ ) aynı kümelenme sonucunu verdiği gözlemlenmiştir. İllere göre duygu analizi modelinin olumsuz tahmin ileti sayılarının sıcak nokta analizi sonucu Harita 8.2’te verilmiştir. Harita 8.2’te Malatya için 0.05 anlamlılık düzeyinde, Kahramanmaraş için 0.10 anlamlılık düzeyinde diğer illere kıyasla olumsuz iletilerin kümelenmiş (sıcak nokta) olduğu öne çıkmaktadır.



**Harita 8.2** 6 Şubat 2023 depreminden etkilen illerden edinilen olumsuz ileti sayıları ile ilgili sıcak nokta analizi (referans ve tahmin haritaları aynıdır)

Referans ve tahmin ileti grupları üzerindeki sıcak nokta analizlerinin sonucunda aynı anlamlılık düzeyinde sonuç vermesi referans ve duygu analizi modelinin verdiği tahmin sonucu ileti sayılarının arasında anlamlı bir fark olmadığını ortaya çıkarmaktadır. Bu çalışmada, sırasıyla Malatya ve Kahramanmaraş illerinde nüfusa oranla olumsuz ileti sayılarının istatistiki olarak yüksek değer gösterdiği anlaşılmaktadır. Diğer iller nüfusa oranla olumsuz ileti sayısına göre istatistiksel olarak rastlantısal dağılım göstermektedir. Bu bulgular, duygu analizi modelinin tahminlerinin, referans verileriyle anlamlı bir uyum gösterdiğini ortaya koymaktadır. Ayrıca, Malatya ve Kahramanmaraş illerinde olumsuz ileti sayılarının nüfusa oranla istatistiksel olarak yüksek değerler sergilemesi, bu bölgelerdeki sosyal ve duygusal durumların daha derinlemesine analiz edilmesi gerektiğini düşündürmektedir.

Diğer illerde ise olumsuz ileti sayılarının rastlantısal bir dağılım göstermesi, olumsuz duyguların mekânsal yoğunluğunun daha geniş bir alanda yayılabileceğine işaret etmektedir. Bu bulgular, afet sonrası duygusal tepkilerin mekânsal faktörlerle nasıl bir ilişki gösterdiğini anlamaya yönelik yapılacak gelecekteki çalışmalara bir yol haritası sunmaktadır.

İki etiketli modelde elde edilen sonuçlar, afet sonrası duygusal tepkilerin mekânsal faktörlerle ilişkisini belirlemek açısından bulgular sunmuştur. Bu bulgular, duygusal çeşitliliği daha kapsamlı bir şekilde incelemek için etiket sayısının artırılmasının gerekliliğini ortaya koymuştur. Bu bağlamda, olumlu ve olumsuz duyguların ötesinde, temel duyguların (korku, öfke, üzüntü, empati)

sınıflandırıldığı dört etiketli duygu modeli geliştirilmiş ve bu model aracılığıyla bölgedeki duygusal yayılımın daha ayrıntılı bir şekilde incelenmesi hedeflenmiştir. Bu doğrultuda, dört etiketli duygu analizi modelinin geliştirilmesiyle, duygusal yayılımın daha derinlemesine bir şekilde incelenmesi amaçlanmıştır. Dört etiketli duygu analizi modelinde, k-kat çapraz doğrulama yöntemi kullanılarak elde edilen test verisinin hata matrisi ve doğruluk oranları sunulmuştur.

Bölüm 5'te belirtilen metriklere göre model performansı belirlenmiştir. Tablo 8.5'te dört etiketli duygu analizi modelinin hata matrisi sonucu verilmiştir. Tablo 8.6'da dört etiketli duygu analizi modelinin derin öğrenme ağıyla eğitilmesi sonucunda, tüm iletileri içeren test verisi üzerindeki tahmin performansına ait metrik sonuçlar sunulmuştur.

Dörtlü etiketleme modelinde sonuçlarına göre, referans veride korku (0) olarak etiketli iken modelin de korku (0) olarak tahmin ettiği ileti sayısı 220 olarak bulunmuştur. Referans veride korku (0) olarak etiketli iken modelin üzüntü (1) olarak tahmin ettiği ileti sayısı 1 olarak bulunmuştur. Referans veride korku (0) olarak etiketli iken modelin öfke (2) olarak tahmin ettiği ileti sayısı 2 olarak bulunmuştur. Referans veride korku (0) olarak etiketli iken modelin empati (3) olarak tahmin ettiği ileti sayısı 3 olarak bulunmuştur.

Referans veride üzüntü (1) olarak etiketli iken modelin korku (0) olarak tahmin ettiği ileti sayısı 8 olarak bulunmuştur. Referans veride üzüntü (1) olarak etiketli iken modelin de üzüntü (1) olarak tahmin ettiği ileti sayısı 144 olarak bulunmuştur. Referans veride üzüntü (1) olarak etiketli iken modelin öfke (2) olarak tahmin ettiği bir sonuç bulunmamaktadır. Referans veride üzüntü (1) olarak etiketli iken modelin empati (3) olarak tahmin ettiği ileti sayısı 9 olarak bulunmuştur.

Referans veride öfke (2) olarak etiketli iken modelin korku (0) olarak tahmin ettiği ileti sayısı 96 olarak bulunmuştur. Referans veride öfke (2) olarak etiketli iken modelin üzüntü (1) olarak tahmin ettiği ileti sayısı 13 olarak bulunmuştur. Referans veride öfke (2) olarak etiketli iken modelin de öfke (2) olarak tahmin ettiği bir sonuç bulunmamaktadır. Referans veride öfke (2) olarak etiketli iken modelin empati (3) olarak tahmin ettiği ileti sayısı 7 olarak bulunmuştur.

Referans veride empati (3) olarak etiketli iken modelin korku (0) olarak tahmin ettiği ileti sayısı 11 olarak bulunmuştur. Referans veride empati (3) olarak etiketli

iken modelin üzüntü (1) olarak tahmin ettiği ileti sayısı 7 olarak bulunmuştur. Referans veride empati (3) olarak etiketli iken modelin öfke (2) olarak tahmin ettiği bir sonuç bulunmamaktadır. Referans veride empati (3) olarak etiketli iken modelin de empati (3) olarak tahmin ettiği ileti sayısı 234 olarak bulunmuştur.

Modelin test aşamasındaki sonuç matrisine göre; korku, üzüntü, öfke ve empati sınıflarından en iyi korku sınıfını öğrenirken en az öfke sınıfını öğrenebilmiştir.

**Tablo 8.5** Dört etiketli derin öğrenme ağında k-kat çapraz doğrulama katmanına göre test sonucu modeli performansını değerlendiren hata matrisi

		Tahmin				Toplam
		Korku (0)	Üzüntü (1)	Öfke (2)	Empati (3)	
Referans (Gerçek)	Korku (0)	220	1	2	3	226
	Üzüntü (1)	8	144	0	9	161
	Öfke (2)	96	13	0	7	116
	Empati (3)	11	7	0	234	252
	Toplam	335	165	2	253	755

**Tablo 8.6** Dört etiketli derin öğrenme ağında test sonucu model performansını değerlendiren genel metrik sonuçları

Metrik	Genel sonuç
Doğruluk	0.79
F1 Skor	0.73
Duyarlılık	0.79
Kesinlik	0.69

Dört etiketli modelinin derin öğrenme ağının Adıyaman, Hatay ve Kahramanmaraş illeri üzerindeki tahmin doğruluk sonuçları Tablo 8.8’de gösterilmiştir. Adıyaman ili için ortalama test tahmin doğruluğu %70, Hatay için %82 ve Kahramanmaraş ili için ise %79 olduğu tespit edilmiştir. İl bazında ortalama test doğruluklarının birbirine yakın olduğu ve %70-%85 arasında değiştiği görülmüştür. Test doğruluğunun düşük olduğu illerde öfke sınıfının doğru tahmin edilememesinin nedeniyle modelinin daha düşük performans ile çalıştığı gözlemlenmiştir.

**Tablo 8.7** Dört etiketli derin öğrenme ağının örnek ilçe veri setleri üzerinde tahmin performansını değerlendiren metrik sonuçları

<b>Metrik</b>	<b>Adıyaman</b>	<b>Hatay</b>	<b>Kahramanmaraş</b>
<b>Doğruluk</b>	<b>0.70</b>	<b>0.82</b>	<b>0.78</b>
<b>F1 Skor</b>	<b>0.64</b>	<b>0.77</b>	<b>0.72</b>
<b>Duyarlılık</b>	<b>0.70</b>	<b>0.82</b>	<b>0.78</b>
<b>Kesinlik</b>	<b>0.63</b>	<b>0.75</b>	<b>0.67</b>

İki farklı sınıflandırma modeline (iki etiketli ve dört etiketli) ait test performans karşılaştırmaları Tablo 8.8’de sunulmaktadır. Modellerin genel başarı metrikleri (doğruluk, F1 skoru, duyarlılık, kesinlik) ile birlikte Adıyaman, Hatay ve Kahramanmaraş illeri özelinde elde edilen sonuçlar tablo halinde karşılaştırılmıştır.

Ayrıca, veri setindeki sınıf dağılımları ve modellerin sınıflandırmadaki zorlukları da değerlendirilmiştir. İki etiketli modelde verinin büyük kısmının olumsuz etiketli olması model başarısını bazı yönlerden etkilerken, dört etiketli modelde sınıf sayısının artması ve bazı sınıflarda veri dengesizliği öğrenmeyi zorlaştırmıştır. Tabloya göre iki etiketli model genel olarak daha yüksek doğruluk ve kesinlik göstermektedir. Ancak dört etiketli model duyarlılık açısından daha başarılıdır. Özellikle sınıf çeşitliliği arttıkça, modelin bazı sınıflarda daha düşük başarı gösterdiği görülmüştür. İl bazında da performans farklılıkları mevcuttur; örneğin Hatay ili dört etiketli modelde daha yüksek doğruluğa sahiptir.

**Tablo 8.8** Dört etiketli ve iki etiketli model karşılaştırması

<b>Özellik / Model</b>	<b>İki Etiketli Model</b>	<b>Dört Etiketli Model</b>
<b>Çıkış Etiketleri</b>	<b>Olumlu (1) / Olumsuz (0)</b>	<b>Korku (0), Üzüntü (1), Öfke (2), Empati (3)</b>
<b>Çıkış Formatı</b>	<b>Tek nöron (sigmoid aktivasyon)</b>	<b>4 nöron (softmax aktivasyon), one-hot encoding</b>
<b>Kod Akış Diyagramı Örneği</b>	<b>Token → Embedding → 3x GRU → Sigmoid → 0-1</b>	<b>Token → Embedding → 3x GRU → Softmax → One-hot</b>

**Tablo 8.8** Dört etiketli ve iki etiketli model karşılaştırması (devamı)

<b>Özellik / Model</b>	<b>İki Etiketli Model</b>	<b>Dört Etiketli Model</b>
<b>Model Test Doğruluğu</b>	<b>%87</b>	<b>%79</b>
<b>F1 Skoru (Genel)</b>	<b>0.73</b>	<b>0.73</b>
<b>Recall (Duyarlılık)</b>	<b>0.57</b>	<b>0.79</b>
<b>Precision (Kesinlik)</b>	<b>1.00</b>	<b>0.69</b>
<b>İl Bazında Doğruluk (%)</b>	<b>Adıyaman: %82, Hatay: %70, Kahramanmaraş: %90</b>	<b>Adıyaman: %70, Hatay: %82, Kahramanmaraş: %78</b>
<b>İl Bazında F1 Skoru</b>	<b>Adıyaman: %74, Hatay: %58, Kahramanmaraş: %87</b>	<b>Adıyaman: %64, Hatay: %77, Kahramanmaraş: %87</b>
<b>İl Bazında Duyarlılık</b>	<b>Adıyaman: %82, Hatay: %70, Kahramanmaraş: %90</b>	<b>Adıyaman: %70, Hatay: %82, Kahramanmaraş: %79</b>
<b>İl Bazında Kesinlik</b>	<b>Adıyaman: %67, Hatay: %49, Kahramanmaraş: %83</b>	<b>Adıyaman: %63, Hatay: %75, Kahramanmaraş: %67</b>
<b>Veri Setindeki Sınıf Dağılımı</b>	<b>Olumlu: %17.5, Olumsuz: %82.5</b>	<b>Korku: %29.93, Üzüntü: %21.32, Öfke: %15.36, Empati: %33.38</b>
<b>Modelin Sınıflandırma Zorlukları</b>	<b>Duygu belirsizliği ve karmaşık ifadeler</b>	<b>Sınıf sayısının artması nedeniyle öfke gibi az örnekli sınıflarda öğrenme zorluğu ve çoklu duygu yapıları nedeniyle karışıklık</b>

Elde edilen nüfusa ve ileti sayısına göre normalize edilmiş referans ve tahmin korku-üzüntü-öfke-empati ileti sayıları üzerinden Adana, Adıyaman, Diyarbakır, Elâzığ, Gaziantep, Hatay, Kahramanmaraş, Kayseri, Malatya, Osmaniye ve Şanlıurfa illeri için sabit mesafe bant (fixed distance band) ilişkisi ile sıcak nokta analizi yapılmıştır. Referans ve tahmin korku-üzüntü-öfke-empati iletilerin sıcak

nokta analiz ine göre kümelenme sonuçları anlamlılık düzeylerinde ( $\alpha=0.10$ ;  $\alpha=0.05$ ;  $\alpha=0.01$ ) gösterilmiştir.

İllere göre duygu analizi modelinin empati sınıfı tahminleri ileti sayılarının sıcak nokta analizi sonucu Harita 8.3'te verilmiştir. Harita 8.3'te Adana için 0.10 anlamlılık düzeyinde diğer illere kıyasla tahmin edilen empati iletilerin kümelenmiş (sıcak nokta) olduğu öne çıkmaktadır. Malatya için 0.10 anlamlılık düzeyinde diğer illere kıyasla tahmin edilen empati iletilerinin kümelenmiş (soğuk nokta) olduğu öne çıkmaktadır.



**Şekil 8.3** Şubat 2023 depreminden etkilen illerden edinilen empati sınıfı tahminleri ile ilgili sıcak nokta analizi

İllere göre duygu analizi modelinin empati referans (gerçek) ileti sayılarının sıcak nokta analizi sonucu Harita 8.4'te verilmiştir. Harita 8.4'te Adana ve Hatay için 0.10 anlamlılık düzeyinde diğer illere kıyasla tahmin edilen empati iletilerin kümelenmiş (sıcak nokta) olduğu öne çıkmaktadır.



**Harita 8.4** 6 Şubat 2023 depreminden etkilen illerden edinilen empati sınıfı gerçek iletiler ile ilgili sıcak nokta analizi

Empati sınıfına yönelik gerçekleştirilen sıcak nokta analizleri sonucunda, Adana ili için hem referans verilerde hem de model tahminlerinde aynı anlamlılık düzeyinde sonuçlar elde edilmiştir. Bu durum, modelin tahmin ettiği empati içerikli ileti sayıları ile referans etiketler arasında Adana ili özelinde anlamlı bir fark bulunmadığını göstermektedir. Hatay ilinde ise model, referans verilere kıyasla daha düşük sayıda empati sınıfına ait ileti tahmininde bulunmuştur. Haritalar incelendiğinde, Adana ilinin hem modelde hem de referans verilerde, nüfus ve toplam ileti sayısına oranla empati sınıfında istatistiksel olarak anlamlı derecede yüksek değerlere sahip olduğu görülmektedir. Ayrıca, model çıktılarında Malatya ili için empati sınıfına ait iletilerin mekânsal yoğunluğunun düşük olduğu ve bu düşük yoğunluğun da istatistiksel olarak anlamlı düzeyde olduğu belirlenmiştir. Empati sınıfı açısından diğer illerde ise model ve referans veriler birlikte ele alındığında, nüfus ve ileti sayılarına göre istatistiksel olarak anlamlı bir mekânsal örüntü ortaya koymamış; dağılımın rastlantısal olduğu gözlemlenmiştir.

İllere göre duygu analizi modelinin öfke tahmin ileti sayılarının sıcak nokta analizi sonucu Harita 8.5'te verilmiştir. Öfke sınıfında yapılan sıcak nokta analizine göre, 11 ilde modelin tahmin ettiği öfke ileti sayıları, istatistiksel olarak anlamlı bir kümelenme göstermemektedir. Bu durum, öfke içerikli iletilerin bu illerde rastlantısal olarak dağıldığını, belirgin sıcak ya da soğuk noktaların oluşmadığını göstermektedir. Referans (gerçek) etiketlerle yapılan sıcak nokta analizinde de 11

ilin öfke sınıfına ait ileti sayılarında anlamlı bir mekânsal kümelenme gözlemlenmemiştir. Bu durum, öfke temalı iletilerin bu iller özelinde belirli bir mekânsal düzende yoğunlaşmadığını, tersine rastlantısal olarak dağıldığını göstermektedir. Bu durumda, öfke içerikli paylaşımlar bu illerde ne belirgin şekilde artmakta ne de azalmaktadır. Bu sonuç hem modelin tahmin çıktılarıyla hem de referans verilerle tutarlıdır.



**Harita 8.5** 6 Şubat 2023 depreminden etkilen illerden edinilen öfke sınıfı gerçek iletiler ile ilgili sıcak nokta analizi (referans ve tahmin haritaları aynıdır)

Referans ve tahmin ileti grupları üzerindeki sıcak nokta analizlerinin sonucunda analizi gerçekleştirilen tüm illerde öfke sınıfının dağılımı, nüfus ve ileti sayısına oranla istatistiksel olarak rastlantısal bir biçimde gerçekleşmektedir.

İllere göre duygu analizi modelinin üzüntü tahmin ileti sayılarının sıcak nokta analizi sonucu Harita 8.6'da verilmiştir. Harita 8.6'da Gaziantep için 0.05 anlamlılık düzeyinde diğer illere kıyasla tahmin edilen üzüntü iletilerin kümelenmiş (sıcak nokta) olduğu öne çıkmaktadır.



**Harita 8.6** 6 Şubat 2023 depreminden etkilen illerden edinilen üzüntü sınıfı tahmini iletiler ile ilgili sıcak nokta analizi

İllere göre duygu analizi modelinin üzüntü referans (gerçek) ileti sayılarının sıcak nokta analizi sonucu Harita 8.7’de verilmiştir. Harita 8.7’de Gaziantep için 0.10 anlamlılık düzeyinde diğer illere kıyasla tahmin edilen üzüntü iletilerin kümelenmiş (sıcak nokta) olduğu öne çıkmaktadır.



**Harita 8.7** 6 Şubat 2023 depreminden etkilen illerden edinilen üzüntü sınıfı gerçek iletiler ile ilgili sıcak nokta analizi

Referans ve tahmin ileti grupları üzerinde gerçekleştirilen sıcak nokta analizleri sonucunda, üzüntü sınıfı açısından Gaziantep ilinin hem model tahminlerinde hem de referans etiketlerde istatistiksel olarak anlamlı bir sıcak nokta olarak belirlendiği





**Harita 8.9 6** Şubat 2023 depreminden etkilen illerden edinilen korku sınıfı gerçek iletiler ile ilgili sıcak nokta analizi

Korku sınıfı için yapılan sıcak nokta analizinde, model sonuçlarına göre Adana ili 0.10 anlamlılık düzeyinde soğuk nokta olarak kümelmiştir. Bu durum, modelin Adana ilinde korku etiketli iletilerin mekânsal yoğunluğunun düşük olduğunu ve bu düşük yoğunluğun istatistiksel olarak anlamlı olduğunu göstermektedir. Referans verilerde ise tüm iller, anlamlılık düzeylerine ulaşmayan bir dağılım sergilemiştir. Bu farklılık, modelin Adana ili özelinde korku sınıfı iletileri açısından belirli bir örüntü öngördüğünü, ancak gerçek verilerde bu türden bir mekânsal yoğunlaşmanın istatistiksel olarak desteklenmediğini göstermektedir.

Dört etiketli duygu analizi modelinin uygulanması, afet sonrası duygusal tepkilerin mekânsal boyutunu inceleme imkânı sağlamıştır. Bu model, temel duyguların (korku, öfke, üzüntü, empati) sınıflandırılması yoluyla, afet bölgesindeki duygusal yayılımı daha ayrıntılı bir şekilde analiz etme fırsatı sunmuştur. Elde edilen sıcak nokta analizleri, duygusal dağılımların mekânsal olarak nasıl kümelendiğini göstererek, bölgedeki psikososyal destek ihtiyaçlarının belirlenmesine yönelik bir çerçeve sunmaktadır. Ancak bu çalışma, psikososyal destek kapsamına alınacak öncelikli illerin belirlenmesine dair doğrudan bir tespit yapmamaktadır. Bunun yerine, doğal dil işleme ve derin öğrenme yöntemleriyle gerçekleştirilen duygu analizleri ile sıcak nokta analizlerinin, CBS tabanlı psikososyal destek planlaması çalışmalarına nasıl katkı sağlayabileceğine dikkat çekmektedir. Tablo 8.4'te görülen il bazında ileti sayılarının artması, sıcak nokta analizinde kullanılan

doğrulama metriklerinin daha güçlü olmasını sağlayacaktır. Ayrıca, psikososyal destek planlamasında, kitlenin duygu durumu ile birlikte ihtiyaç ve taleplerin mekânsal analizleri de dikkate alınarak planlamaların yapılması gerekmektedir. Bu bağlamda, bu çalışma, psikososyal destek kapsamına alınacak öncelikli illerin belirlenmesine dair doğrudan bir tespit yapmamakla birlikte, duygu analizi ve sıcak nokta analizlerinin birleşiminin, CBS tabanlı psikososyal destek planlaması çalışmalarına önemli katkılar sağlayabileceğini vurgulamaktadır.

Literatürde X platformu üzerinden belli ilgi noktalarından elde edilen veya toplumsal (ünlü birinin bir şehri ziyaret etmesi), siyasal (darbe girişimleri, ülkede yaşanan protestolar vs.) ve doğal olaylar (deprem, sel, yoğun yağış vs.) ile ilgili konular hakkındaki iletilerden yapılan duygu analizi çalışmaları oldukça yaygındır. Bu çalışmanın literatürdeki çalışmalardan farkı, 6 Şubat 2023 depremleri ile ilgili insanların duygu durumlarını ve değişimlerini analiz ederek bu sonuçların dağılımlarını göstermesidir. Bu çalışmada kullanılan veri kazıma ve duygu analizi metodolojisi özgün olarak kurgulanmış ve diğer afet konularına adapte edilebilir esnek bir algoritma ile ortaya çıkmıştır. Başlangıçta, veriler iki sınıf (olumlu ve olumsuz) olarak etiketlenmiş, ancak duygu durumlarının çeşitliliği (örneğin, korkma, üzülme, şaşkınlık, mutluluk gibi) nedeniyle, verilerin daha fazla sınıfa ayrılması gerektiği sonucuna varılmıştır. Bu bağlamda, dört etiketli model geliştirilmiştir. Ancak etiket sayısının artırılmasıyla birlikte, modelin her bir sınıfına ait ileti sayılarının belirgin şekilde azaldığı gözlemlenmiştir. Bu durum, modelin her bir sınıf için yeterli örneklem bulunduramayarak öğrenme oranlarının düştüğünü ve dolayısıyla sınıflandırma başarısının zayıfladığını göstermektedir. Etiket sayısının artması, modelin sınıf sayısını daha fazla ayrıntılı şekilde öğrenme kapasitesini sınırlamakta, ancak aynı zamanda daha karmaşık duygu durumlarının daha doğru bir şekilde tanımlanmasını sağlamaktadır. Bu nedenle, etiket sayısının artırılmasının, modelin doğruluğunda bir dengeyi gerektirdiği; daha ayrıntılı sınıflandırmalar yapılabileceği ancak bunun için daha geniş veri kümelerine ihtiyaç duyulacağı sonucuna varılmaktadır. Sonuç olarak, modelin eğitiminde kullanılacak etiketlerin daha fazla olması modelin tahmin doğruluğunu geliştirebilir. Kullanılan veri kaynağının sosyal medya olması ve kitle kaynaklı toplanan verilerin çelişkili veya anlaşılması güç metinleri içerebilmesi duygu analizinde kullanılan makine öğrenmesi yöntemlerinin performansını olumsuz yönde etkileyen zorluklardan

biridir. Bu alıřmada gzlemlendiđi gibi duygu durumu belirsiz ifadeler sınıflandırıcının performansını dřürebilmektedir. Ayrıca sosyal medya verilerinin CBS ortamında analiz edilmesi ile ilgili bir diđer zorluk olarak kullanıcıların konumu kapalı bir biçimde paylařtıkları iletilerin CBS ortamına dahil edilememesi olarak deđerlendirilebilir.



## 9 SONUÇ

---

Bu çalışmada, X platformu üzerinden Kahramanmaraş depremlerine ait Haziran-Eylül 2023 tarihleri arasındaki 755 adet gönderiye erişilmiştir. Veri kazıma kütüphanelerinden BeautifulSoup ve Selenium kullanılarak otomatik bir şekilde istenen sayıda gönderinin elde edilmesi işlemi gerçekleştirilen bir Python kodu hazırlanmıştır. Hazırlanan derin öğrenme eğitim modeli sayesinde yaşanmış veya yaşanabilecek depremler sonrasında anlık olarak istenilen platform üzerinden veri kazıma ve duygu analizi yapılabilir hale getirilmiştir.

Elde edilen veriler iki ve dört etiketli olmak üzere iki adet duygu analizi modelinde kullanılmıştır. İki etiketli duygu analizi modelinde, veri setindeki iletiler olumlu (1) ve olumsuz (0) etiketleriyle kategorize edilmiştir. Bu veri seti üzerinden doğal dil ön işleme adımları uygulanarak, duygu analizini gerçekleştiren bir model oluşturulmuştur. Model, verilen veri setinin eğitimini ortalama %87 test doğruluğu ile tamamlamıştır. Model eğitimi tamamlayıp veri seti üzerinde tahminlerini gerçekleştirdiğinde ise verinin %82.5'ini olumsuz etiketli, %17.5'ini olumlu etiketli olarak sınıflandırmıştır.

Veri setinde olumlu etikete sahip olan fakat modelin tahmin sonucunda olumsuz etiket sınıfına eklenen gönderiler incelenmiştir. Bu gönderilerin modelin karıştırabileceği duygu durumu belirsiz ifadeler içerdiği fark edilmiştir. İki etiketli modelin sıcak nokta analizi sonuçlarında, ileti sayılarının rastlantısal bir dağılım göstermesi; olumsuz duyguların mekânsal yoğunluğunun daha geniş bir alanda yayılabileceğine işaret etmektedir. Bu bulgular, afet sonrası duygusal tepkilerin mekânsal faktörlerle nasıl bir ilişki gösterdiğini anlamaya yönelik yapılacak gelecekteki çalışmalara bir yol haritası sunmaktadır.

Dört etiketli duygu analizi modelinde ise, veri setindeki iletiler korku (0), üzüntü (1), öfke (2) ve empati (3) etiketleriyle kategorize edilmiştir. Bu veri seti

kullanılarak, verileri doğal dil ön işleme tekniklerinden geçiren ve duygu analizini gerçekleştiren hazır bir model geliştirilmiştir. Model, eğitim sürecini ortalama %79 test doğruluğu ile tamamlamıştır. Model eğitimi tamamlandıktan sonra, veri seti üzerinde yapılan tahminler sonucunda iletilerin %44,37'si (335 ileti) korku, %21,85'i (165 ileti) üzüntü, %0,26'sı (2 ileti) öfke ve %33,51'i (253 ileti) empati etiketleriyle sınıflandırılmıştır. Gerçek duygu dağılımı incelendiğinde ise iletilerin %29,93'ü (226 ileti) korku, %21,32'si (161 ileti) üzüntü, %15,36'sı (116 ileti) öfke ve %33,38'i (252 ileti) empati etiketine sahip olduğu gözlemlenmiştir.

Gerçek etiketi öfke olan 116 iletinin 96 ileti (%82,8 oranında) korku etiketi olarak tahmin edilmiştir. Hatalı sınıflandırılan örnekler incelendiğinde, iletiler içerisinde hem korku duygusunun hem de bu korku duygusundan çıkan öfke duygusunun baskın olduğu ifadeler bulunmaktadır (örneğin "Başlatılan ağır hasarlı binaların yıkımları ve kaldırılan enkaz çalışmalarını takip etmiyor musunuz? Şehir toz bulutuna döndü. İnsanların ciğerleri tozlarla doluyor. Yakında hastalıklar artacak. Nerede önlem?" şeklindeki bir ileti). Bu gibi iletiler gerçekte öfke duygusunu ifade etmesine rağmen model tarafından korku olarak sınıflandırılmıştır. Bu durum, iletinin uzun ve karmaşık bir yapıya sahip olması, birden fazla duyguyu (kaygı, korku ve öfke) aynı anda barındırması nedeniyle modelin baskın duyguyu yanlış yorumlamasından kaynaklanmıştır. Veri seti içerisinde bulunan ve çok boyutlu duygusal içerik taşıyan bu gibi iletilerde modelin öğrenme oranını düşürmüştür. Bu çok duygulu iletilerin tek bir duygu seçimi yapması gerektiğinden, baskın duygunun doğru tespit edilememesinde model zorlanmıştır ve hata yapmıştır. Veri seti içerisinde öfke gibi örnek sayısı düşük olan sınıflarda modelin öğrenme kapasitesi sınırlı kalmıştır. Bu durum modelin öfke sınıfında tahmin performansının düşmesine neden olmuştur.

Öfke sınıfında yapılan sıcak nokta analizine göre, 11 ilde modelin tahmin ettiği öfke ileti sayıları, istatistiksel olarak anlamlı bir kümelenme göstermemektedir. Bu durum, öfke içerikli iletilerin bu illerde rastlantısal olarak dağıldığını, belirgin sıcak ya da soğuk noktaların oluşmadığını göstermektedir. Referans (gerçek) etiketlerle yapılan sıcak nokta analizinde de 11 ilin öfke sınıfına ait ileti sayılarında anlamlı bir mekânsal kümelenme gözlemlenmemiştir. Bu durum, öfke temalı iletilerin bu iller özelinde belirli bir mekânsal düzende yoğunlaşmadığını, tersine rastlantısal olarak dağıldığını göstermektedir. Bu durumda, öfke içerikli paylaşımlar bu illerde

ne belirgin şekilde artmakta ne de azalmaktadır. Bu sonuç hem modelin tahmin çıktılarıyla hem de referans verilerle tutarlıdır. Dört etiketli modelin sıcak nokta analizi; Adana, Gaziantep, Hatay ve Malatya gibi illerde duygulara ait kümelenmeler göstermiştir. Ancak, öfke sınıfında kümelenme gözlemlenmemiştir. Elde edilen sonuçlar, veri setindeki sınıf dengesizliğinin ve ifadelerdeki çoklu duygu yapısının model başarısını etkileyen faktörler arasında olduğunu göstermiştir. Gelecek çalışmalarda, veri artırma tekniklerinin kullanılması, sınıf dengesinin gözetilmesi ve çoklu etiketli sınıflandırma yaklaşımlarının denenmesinin, model performansını daha da iyileştirebileceği sonucuna varılmıştır. Bu nedenle, etiket sayısının artırılmasının, modelin doğruluğunda bir dengeyi gerektirdiği; daha ayrıntılı sınıflandırmalar yapılabileceği ancak bunun için daha geniş veri kümelerine ihtiyaç duyulacağı sonucuna varılmaktadır. Sonuç olarak, modelin eğitiminde kullanılacak etiketlerin daha fazla olması modelin tahmin doğruluğunu geliştirebilir.

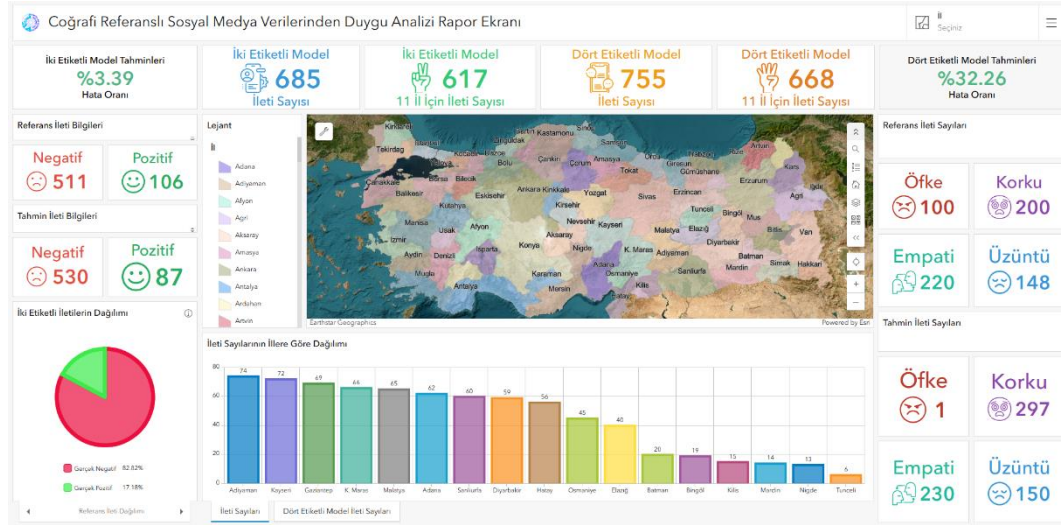
Bu çalışmadan elde edilen sonuçlar ArcGIS Online platformu üzerinde ArcGIS Dashboards aracı kullanılarak bir rapor ekranı üzerinden sunulmuştur (Şekil 9.1). Bu rapor ekranı, iki farklı sınıflandırma modeli (iki etiketli ve dört etiketli duygu analizi) sonucunda elde edilen sosyal medya ileti analizlerinin görsel olarak yorumlanmasını sağlamaktadır.

Rapor ekranı referans veri ile modelin tahmin ettiği etiketlerin karşılaştırmasını, hata oranlarını, illere göre ileti sayılarını ve duygu sınıflarına ait dağılımları aynı ekran üzerinde bütüncül bir biçimde sunmaktadır.

Şekil 9.1'de bulunan rapor ekranındaki bilgilere göre, iki etiketli modelin hata oranı %3.39 ile oldukça düşükken, dört etiketli modelde hata oranı %32.26'ya ulaşmıştır. Bu durum, sınıf sayısının artmasının ve bazı sınıflarda veri dengesizliğinin modelin performansını olumsuz etkilediğini göstermektedir.

Ayrıca, "öfke" gibi düşük örnek sayısına sahip sınıflarda modelin başarısız olduğu dikkat çekmektedir. Örneğin, 100 adet "öfke" etiketli iletiye karşılık, model bu sınıfa yalnızca 1 ileti tahsisi yapmıştır. Bu da dört etiketli modelin, bazı duygusal sınıfları ayırt etmede zorlandığını göstermektedir.

İllere göre ileti dağılımını ise Adıyaman, Kayseri ve Gaziantep gibi illerin analizde öne çıktığını göstermektedir. Bu illerdeki yüksek sosyal medya aktivitesi veya afet sonrası etkilerin daha belirgin olması bu sonucu etkileyen olası faktörlerdir.



Şekil 9.1 ArcGIS dashboards ile oluşturulan duygu analizi rapor ekranı

Bu çalışma, Kahramanmaraş depremleri ile ilgili sosyal medya verilerinin doğal dil işleme ve kapı yinelemeli birimler derin öğrenme modeli kullanılarak duygu analizinin gerçekleştirilebileceği gösterilmiştir. Herhangi bir deprem sonrasında duygu analizi yapabilen bu model sayesinde, afet hakkında duygu analizi güncel olarak gerçekleştirilebilecektir. Ayrıca tahmin verilerinden sıcak nokta analizi ile iletiler arasındaki kümelenme örüntülerinin istatistiki olarak tespit edilebileceği ve bu kümelerin psikososyal planlama aşamasında fikir verebileceği görülmektedir. Gelecek çalışmalarda, sosyal medya üzerinden paylaşılan yeni iletiler de mevcut giriş veri setine eklenerek daha geniş bir sözlük kullanılarak derin öğrenme modelinin eğitimi geliştirilebilir ve sonuçlar iyileştirilebilir. Duygu analizinin sonuçları CBS ortamına aktarılarak duygu durum değişimleri mekân-zaman kapsamında analiz edilebilir.

## KAYNAKÇA

---

- Adalı, E. (2012). Doğal dil işleme. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 5(2).
- Albayrak, M., Topal, K., & Altıntaş, V. (2017). Sosyal medya üzerinde veri analizi: Twitter. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 22(Kayfor 15 Özel Sayısı), 1991-1998.
- Alves, A. L. F., de Souza Baptista, C., Firmino, A. A., de Oliveira, M. G., & Paiva, A. C. (2015). A spatial and temporal sentiment analysis approach applied to Twitter microtexts. *Journal of Information and Data Management*, 6(2), 118.
- Aman, J., & Matisziw, T. C. (2025). Urban sentiment mapping using language and vision models in spatial analysis. *Frontiers in Computer Science*, 7, 1504523.
- Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis* (Vol. 413, No. 8). Essex: Longman Scientific & Technical.
- Berrar, D. (2019). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*. 1, 542-545.
- Bhardwaj, A., Narayan, Y., & Dutta, M. (2015). Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty. *Procedia Computer Science*, 70, 85-91.
- Camacho, K., Portelli, R., Shortridge, A., & Takahashi, B. (2021). Sentiment Mapping: Point Pattern Analysis of Sentiment Classified Twitter Data. *Cartography and Geographic Information Science*, 48(3), 241-257.
- Cardone, B., Di Martino, F., & Sessa, S. (2022). GIS-based fuzzy sentiment analysis framework to classify urban elements according to the orientations of citizens and tourists expressed in social networks. *Evolutionary Intelligence*, 1-10.
- Chen, L. C., Lee, C. M., & Chen, M. Y. (2020). Exploration of Social Media for Sentiment Analysis Using Deep Learning, *Soft Computing*, 24, 8187-8197.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. arXiv preprint arXiv, 1406-1078.
- Clarke, K.C. (1997). *Getting started with geographic information systems* (25-30). New Jersey: Prentice Hall.
- Çubukçu, K.M. (2015). *Planlamada ve Coğrafyada Temel İstatistik ve Mekânsal İstatistik*. Ankara: Nobel Akademik Yayıncılık.

- Ergüt, Ö., & Çilingirtürk, A. M. (2022). Türkiye'nin Yaşam Memnuniyetine Mekansal Etkinin Analizi. *Journal of Life Economics*, 9(2), 65-80.
- Feng, Y., Huang, X., & Sester, M. (2022). Extraction and Analysis of Natural Disaster Related Vgi from Social Media: Review, Opportunities and Challenges. *International Journal of Geographical Information Science*, 36(7), 1275-1316.
- Gulnerman, A. G., & Karaman, H. (2020). Sosyal Medyanın Gönüllü Coğrafi Veri Olarak Kullanımı ve Sosyal Medya Verilerinden Coğrafya Sözlüğü Üretimi. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 20(2), 276-286.
- Gulnerman, A. G., Karaman, H., Pekaslan, D., & Bilgi, S. (2020a). Citizens' Spatial Footprint on Twitter-Anomaly, Trend and Bias Investigation in Istanbul. *ISPRS International Journal of Geo-Information*, 9(4), 222.
- Gulnerman, A. G., Karaman, H., & Basiri, A. (2020b). New age of crisis management with social media. *Open Source Geospatial Science for Urban Studies: The Value of Open Geospatial Data* (131-160). Berlin:Springer.
- Gümüşsu, E. (2019). Metin madenciliği teknikleri ile sosyal medya gönderilerinin analiz edilmesi: ankilozan spondilit hastalığı örneği (Master's thesis, Fen Bilimleri Enstitüsü).
- Gündüz, H. (2013). Borsa İstanbul (BIST) 100 Endeksi Yönünün Ekonomi Haberleri ile Tahmin Edilmesi (Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
- Hourrane, O., Benlahmar, E. H., & Ahmed, Z. (2018). Comparative study of deep learning models for sentiment analysis. *International Journal of Engineering & Technology*, 7 (4), 5726-5731.
- Häberle, M., Werner, M., & Zhu, X. X. (2019). Building Type Classification from Social Media Texts Via Geo-spatial Textmining. 2019 IEEE International Geoscience And Remote Sensing Symposium (10047-10050). Yokohama, Japonya.
- İlhan, N., & Sağaltıcı, D. (2020). Twitter'da Duygu Analizi. *Harran Üniversitesi Mühendislik Dergisi*, 5(2), 146-156.
- Kaya, A. (2021). Bir Araştırma Kaynağı Olarak Arşivlenen Sosyal Medya Verilerinin Kullanımı. *Bilgi ve Belge Araştırmaları*, (16), 49-79.
- Kılınç, M., Aydın, C., & Tarhan, Ç. (2022). Türkiye'de Sosyal ve Dijital Girişimcilik: Veri Kazıma Teknikleriyle Kitle Fonlaması Platformlarının İçerik Analizi, *Acta Infologica*, 6(1), 83-97.
- Koç, M., & Yalçın, S. (2023). Afetlerde Krize Müdahale: Kahramanmaraş Depremi'nde Aile ve Sosyal Hizmetler Bakanlığı'nın Çalışmaları. *Uluslararası Sosyal Hizmet Araştırmaları Dergisi*, 3(2), 93-105.

- Küçük, D., & Arıcı, N. (2018). Doğal Dil İşlemede Derin Öğrenme Uygulamaları Üzerine Bir Literatür Çalışması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2(2), 76-86.
- Lagrari, F. E., & Elkettani, Y., (2021). Traditional and Deep Learning Approaches for Sentiment Analysis: A Survey. *Advances in Science, Technology and Engineering Systems Journal*. 6(5), 01-07.
- Liu, W., Li, D., Meng, Y., & Guo, C. (2024). The Relationship between Emotional Perception and High-Density Built Environment Based on Social Media Data: Evidence from Spatial Analyses in Wuhan. *Land*, 13(3), 294.
- Markoulidakis, I., Kopsiaftis, G., Rallis, I., & Georgoulas, I. (2021). Multi-class Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. In *The 14th Pervasive Technologies Related to Assistive Environments Conference* (412-419). Korfu, Yunanistan.
- Neal, D.M., (1997). Reconsidering the Phases of Disasters. *International Journal of Mass Emergencies and Disasters*, 15 (2), 239–264.
- Oprea, S. V., & Bâra, A. (2022). Why Is More Efficient to Combine BeautifulSoup and Selenium in Scraping for Data Under Energy Crisis. *Ovidius University Annals, Economic Sciences Series*, 22(2), 146-152.
- Özgür E. M. & Aydın, O. (2011). Türkiye’de Evlilik Göçünün Mekânsal Veri Analizi Teknikleriyle Değerlendirilmesi. *Coğrafî Bilimler Dergisi*, 9(1), 29–40.
- Qian, X., & Gkritza, K. (2024). Spatial and temporal variance in public perception of electric vehicles: A comparative analysis of adoption pioneers and laggards using twitter data. *Transport Policy*, 149, 150-162.
- Sachin, S., Tripathi, A., Mahajan, N., Aggarwal, S., & Nagrath, P. (2020). Sentiment analysis using gated recurrent neural networks. *SN Computer Science*, 1, 1-13.
- Song, Z., & Xia, J. C. (2016). Spatial and temporal sentiment analysis of twitter data. *European handbook of crowdsourced geographic information*, 205.
- Şeker, S. E. (2015). Metin Madenciliği (Text Mining). *YBS Ansiklopedi*, 2(3), 30-32.
- Talan, T. & Aktürk, C., (2021). *Bilgisayar Bilimlerinde Teorik ve Uygulamalı Araştırmalar*. Efe Akademik Yayıncılık.
- Tanasă, A. M., Oprea, S. V., & Bâra, A. (2023). Web Scraping and Review Analytics. Extracting Insights from Commercial Data. *Ovidius University Annals, Series Economic Sciences*, 23(2).
- Uzun, E., Kılıçaslan, Y., & Uçar, E. (2008). HTML, XML ve Web Servislerinin İnternet Sunucuları Üzerindeki Etkisinin İncelenmesi. *Trakya Üniversitesi Fen Bilimleri Dergisi*, 8(2), 81-85.
- Vennerød, C. B., Kjærø, A., & Bugge, E. S. (2021). Long short-term memory RNN. *arXiv preprint arXiv:2105.06756*.
- Wang, J., Zhang, Y., Wei, Z., Zhang, J., & Xie, X. (2024, October). Harmonic Current Probabilistic Modeling Based on K-Fold Cross-Validation Kernel

- Density Estimation. In 2024 21st International Conference on Harmonics and Quality of Power (ICHQP) (pp. 421-426). IEEE.
- Wang, M., & Hu, F. (2021). The application of nltk library for python natural language processing in corpus research. *Theory and Practice in Language Studies*, 11(9), 1041-1049.
- Wang, R. Q., Mao, H., Wang, Y., Rae, C., & Shaw, W. (2018). Hyper-Resolution Monitoring of Urban Flooding with Social Media and Crowdsourcing Data. *Computers & Geosciences*, 111, 139-147.
- William, I. O., & Altamimi, E. M. (2024). Hierarchical Long Short-Term Memory (LSTM) Model for News Sentiment Analysis. *2nd International Conference on Scientific and Innovative Studies*, 1(1).
- Williams, R. J., & Zipser, D. (1989). A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation*, 1(2), 270-280.
- You, J., Lee, K., & Kwon, H. Y. (2024). Deepscraper: a complete and efficient tweet scraping method using authenticated multiprocessing. *Data & knowledge engineering*, 149, 102260.
- Zhang, X., Zhang, X., Zhang, Y., Liu, Y., Zhou, R., Raxidin, A., & Li, M. (2025). A Multidimensional Study of the 2023 Beijing Extreme Rainfall: Theme, Location, and Sentiment Based on Social Media Data. *ISPRS International Journal of Geo-Information*, 14(4), 136.
- Zhao, B. (2017). *Web Scraping. Encyclopedia of Big Data*. Cham: Springer.
- Zhou, X., & Xu, C. (2017). Tracing the Spatial-Temporal Evolution of Events Based on Social Media Data. *ISPRS International Journal of Geo-Information*, 6(3), 88.
- URL-1: AFAD (2023), Basın Bülteni. 36. <https://www.afad.gov.tr/kahramanmarasta-meydana-gelen-depremler-hk-36> (Erişim Tarihi: 02 Ocak 2024).
- URL-2: AFAD Türkiye Afet Müdahale Eylem Planı (TAMP) (2014). [https://www.afad.gov.tr/kurumlar/afad.gov.tr/e\\_Kutuphane/Planlar/TAMP.pdf](https://www.afad.gov.tr/kurumlar/afad.gov.tr/e_Kutuphane/Planlar/TAMP.pdf) (Erişim Tarihi: 02 Ocak 2024).
- URL-3: Afet ve Acil Durum Müdahale Hizmetleri Yönetmeliği (2013), Resmî Gazete. <https://www.resmigazete.gov.tr/eskiler/2013/12/20131218-13-1.pdf> (Erişim Tarihi: 02 Ocak 2024).
- URL-4: Aile ve Toplum Hizmetleri Genel Müdürlüğü (2023), Psikososyal Destek Rehberi. <https://www.aile.gov.tr/uploads/athgm/uploads/pages/goc-afet-ve-acil-durumlarda-psikososyal-destek/psi-kososyal-destek-nedi-r.pdf> (Erişim Tarihi: 02 Ocak 2024).
- URL-5: Esri (2024), How Hot Spot Analysis (Getis-Ord Gi\*) Works. <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatialstatistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm> (Erişim Tarihi: 13 Şubat 2024).
- URL-6: Millî Savunma Bakanlığı Harita Genel Müdürlüğü. <https://www.harita.gov.tr/il-ve-ilce-yuzolcumleri> (Erişim Tarihi: 21 Aralık 2024)

URL-7: Python (2021) Real Python. BeautifulSoup: Build a Web Scraper With Python – Real Python (Eriřim Tarihi: 02 Ocak 2024).

URL-8: Türkiye İstatistik Kurumu. <https://nip.tuik.gov.tr/?value=CinsiyeteGoreNufus> (Eriřim Tarihi: 21 Aralık 2024)

URL-9: X Developer Platform. <https://docs.x.com/x-api/getting-started/about-x-api>

URL-10: X Developer Portal. <https://developer.x.com/en/portal/dashboard>



## TEZDEN ÜRETİLMİŞ YAYINLAR

---

### Konferans Bildirileri

1. Dolu, D. G. ve Şen, A., (2023). Kahramanmaraş Depremlerine ait Sosyal Medya Verilerinin Coğrafi Bilgi Sistemlerinde Analizi. 19. Türkiye Harita Bilimsel ve Teknik Kurultayı. 04-06 Ekim. Ankara, Türkiye.
2. Dolu, D. G., & Şen, A., (2024). Coğrafi Referanslı Sosyal Medya Verilerinden Doğal Dil İşleme ve Derin Öğrenme Kullanılarak Duygu Analizi: Kahramanmaraş Depremleri Örneği. YTU Grad Colloquium'24 (ss.48). İstanbul, Türkiye.

### Makaleler

1. Dolu, D. G., & Şen, A. (2024). Coğrafi referanslı sosyal medya verilerinden doğal dil işleme ve derin öğrenme kullanılarak duygu analizi: Kahramanmaraş depremleri örneği. Jeodezi ve Jeoinformasyon Dergisi, 11(1), 51-67. <https://doi.org/10.9733/JGG.2024R0005.T>