



**T.C.**

**ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**YALAN HABER YAYILIMININ DUYGU ANALİZİYLE  
İNCELENMESİ VE METİN SINIFLANDIRMADA YAPAY SİNİR  
AĞLARININ BAŞARIMLARININ KARŞILAŞTIRILMASI**

**YÜKSEK LİSANS TEZİ**

**MAİDE FEYZA ER**

**Tez Danışmanı**

**Dr. Öğr. Üyesi YONCA BAYRAKDAR YILMAZ**

**ÇANAKKALE – 2022**





T.C.

ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**YALAN HABER YAYILIMININ DUYGU ANALİZİYLE İNCELENMESİ VE  
METİN SINIFLANDIRMADA YAPAY SİNİR AĞLARININ BAŞARIMLARININ  
KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

MAİDE FEYZA ER

Tez Danışmanı

Dr. Öğr. Üyesi YONCA BAYRAKDAR YILMAZ

ÇANAKKALE – 2022

## ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmada; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

(İmza)

Maide Feyza ER

(Tarih) 08/12/2022

## TEŐEKKÜR

Bu tezin gerekleŐtirilmesinde, alıŐmam boyunca benden bir an olsun yardımlarını esirgemeyen saygı deęer danıŐman hocam Dr. Öğr. Üyesi Yonca BAYRAKDAR YILMAZ'a, tecrübeleriyle her daim yoluma ıŐık tutan sevgili ablam Rezzan Nisa ER'e ve hayatımın her evresinde bana destek olan deęerli aileme sonsuz teŐekkürlerimi sunarım.

Maide Feyza ER  
anakkale, Aralık 2022



## ÖZET

# YALAN HABER YAYILIMININ DUYGU ANALİZİYLE İNCELENMESİ VE METİN SINIFLANDIRMADA YAPAY SİNİR AĞLARININ BAŞARIMLARININ KARŞILAŞTIRILMASI

Maide Feyza ER

Çanakkale Onsekiz Mart Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Tezi

Danışman: Dr. Öğr. Üyesi Yonca BAYRAKDAR YILMAZ

08/12/2022, 71

Sosyal medyanın haber kaynağı olarak kullanılması günümüzde oldukça yaygındır. Ancak karşılaşılan haberlerin doğruluğunun sorgulanmadan ve kontrol edilmeden doğru kabul edilmesi, yalan haberlerin yayılmasının ana nedenlerindedir. Sosyal medya kullanıcıları için, haberin doğruluğu kanıtlardan ziyade fikir ve duyguların kontrolü altında olduğundan, yalan haberlerin yayılmasında hangi duygunun daha etkili olduğu sorusu ortaya çıkmaktadır. Çalışmamız, kullanıcıların yalan haberlere tepkileri ile haberlerin yaygınlığı arasında bir ilişki olup olmadığını araştırmaktadır. Çalışmamızda, Twitter kullanıcılarının 31 Aralık 2019–30 Temmuz 2022 tarihleri arasında yalan haberlere verdikleri tepkiler toplanmış ve 4 ana başlıkta kategorize edilmiştir. Bu tepkiler üzerine duygu analizi yapılmıştır. Yalan haberin yaygınlığıyla habere verilen tepki arasında bir ilişki olduğu ancak yalan haberin türüne ve insanları hangi bağlamda etkilediğine göre değiştiği gözlemlenmiştir. Veri setleri, LSTM yöntemi ve TensorFlow Hub’da önceden eğitilmiş metin gömmelerini kullanan YSA modeliyle eğitilmiştir. Tüm veri setleri için, LSTM’e kıyasla YSA modelinden daha yüksek doğruluk değeri elde edilmiştir. Önceden eğitilmiş metin gömmelerini kullanan YSA modelinin az veriyle bile oldukça iyi sonuç verdiği görülmüş, doğrulama ve test veri seti için sırasıyla 0.9020 ve 0.8998 doğruluk değerleri elde edilmiştir. Modellerin optimum seviyesine ulaşmaları için kullanılan eğitim tur sayıları da farklılık göstermiştir. Bulgular metin sınıflandırma problemi için, önceden eğitilmiş metin gömmelerinin kullanılmasının performansı arttırdığını ve eğitim süresini kısalttığını göstermektedir.

**Anahtar Kelimeler:** Duygu Analizi, Metin Madenciliği, Yapay Sinir Ağları, LSTM, Metin Sınıflandırma

## ABSTRACT

# INVESTIGATION OF FAKE NEWS SPREAD WITH SENTIMENT ANALYSIS AND PERFORMANCE COMPARISON OF NEURAL NETWORKS FOR TEXT CLASSIFICATION

Maide Feyza ER

Çanakkale Onsekiz Mart University

School of Graduate Studies

Master of Science Thesis in Computer Engineering

Advisor: Assist. Prof. Dr. Yonca BAYRAKDAR YILMAZ

08/12/2022, 71

The use of social media as a news source is quite common nowadays. However, the fact that the news encountered is accepted as true without questioning and checking its accuracy is one of the main reasons for the spread of fake news. For social media users, the question arises as to which emotion is more effective in spreading fake news, as the accuracy of the news is under the control of ideas and emotions rather than evidence. From this point of view, our study investigates whether there is a relationship between users' reactions to the news and the prevalence of the news. In our study, the reactions of Twitter users to fake news between 31 December 2019 and 30 July 2022 were collected and categorized under 4 main headings. Sentiment analysis was performed on these reactions. It has been observed that there is a relationship between the prevalence of fake news and the reaction to the news, but it varies according to the type of fake news and the context in which it affects people. The datasets were trained with the LSTM method and the Neural Nets model using pre-trained text embeddings in TensorFlow Hub. For all datasets, higher accuracy value was obtained from the NN model compared to LSTM. It was seen that the NN model using pre-trained text embeddings gave very good results even with little data, and accuracy values of 0.9020 and 0.8998 were obtained for the validation and test dataset, respectively. The number of epochs used to reach the optimum level of the models also differed. The findings show that for the text classification problem, the use of pre-trained text embeddings improves performance and reduces training time.

**Keywords:** Sentiment Analysis, Text Mining, Neural Networks, LSTM, Text Classification

# İÇİNDEKİLER

Sayfa No

ETİK BEYAN.....	i
TEŞEKKÜR.....	ii
ÖZET .....	iii
ABSTRACT .....	iv
İÇİNDEKİLER .....	v
SİMGELER ve KISALTMALAR.....	vii
TABLolar DİZİNİ.....	vii
ŞEKİLLER DİZİNİ.....	ix

## BİRİNCİ BÖLÜM

### GİRİŞ

1

1.1. Yalan Haber	2
1.1.1. Yalan Haber Siteleri	3
1.1.2. Sosyal Medyada Yalan Haber	3
Sosyal Botlar	5
1.2. Sosyal Medya	6
1.2.1. Sosyal Medyanın Tarihçesi	8
1.3. Twitter	8
1.3.1. Neden Twitter?	10
1.4. Duygu Analizi	10
1.4.1. Duygu Analizinin Önemi	11
1.4.2. Duygu Analizi Yaklaşımları	12
Leksikon Tabanlı (Lexicon-Based) Yaklaşım	13
Kural Tabanlı (Rule-Based) Yaklaşım	16
Makine Öğrenimi (Machine Learning) Yaklaşımı	17

## İKİNCİ BÖLÜM

### ÖNCEKİ ÇALIŞMALAR

19

2.1. Makine Öğrenimi Kullanan Çalışmalar	19
2.2. Hibrit Yaklaşım Kullanan Çalışmalar	24



2.3. Derin Öğrenme Kullanan Çalışmalar	26
<b>ÜÇÜNCÜ BÖLÜM</b>	
<b>MATERYAL YÖNTEM</b>	
3.1. Veri Toplama	31
3.2. Metin Normalleştirme	34
3.3. Duygu Analizi (Sentiment Analysis)	40
3.4. Yapay Sinir Ağlarının Eğitimi	42
3.4.1 TensorFlow Girdi Veri Hattı (TF Input Pipeline)	43
3.4.2 YSA Modelinin Oluşturulması	45
3.4.3 LSTM Yöntemi	46
<b>DÖRDÜNCÜ BÖLÜM</b>	
<b>ARAŞTIRMA BULGULARI</b>	
4.1. Duygu Analizi Sonuçları	49
4.2. Modellerin Değerlendirilmesi	55
<b>BEŞİNCİ BÖLÜM</b>	
<b>SONUÇ ve ÖNERİLER</b>	
KAYNAKÇA .....	65
EKLER .....	I

## SİMGELER VE KISALTMALAR

ABD	Amerika Birleşik Devletleri
API	Uygulama Programlama Arayüzü
CERN	Avrupa Nükleer Araştırma Merkezi
CLI	Komut Satırı Arayüzü
CNN	Çeşitli Evrişimli Sinir Ağlarını
COVID	Koronavirüs Hastalığı
CRF	Şartlı Rastgele Alan
CSV	Virgülle Ayrılmış Değerler
GRU	Kapılı Yinelemeli Üniteler
JSON	Javascript Nesnesi Gösterimi
LR	Lojistik Regresyon
LSTM	Uzun Kısa Süreli Bellek
ML	Makine Öğrenmesi
MNB	Çok Terimli Naive Bayes
MPQA	Çok Perspektifli Soru Cevaplama
MaxEnt	Maksimum Entropi
NB	Naive Bayes
NLP	Doğal Dil İşleme
NNLM	Sinir Ağı Dil Modeli
POS	Konuşmanın Bölümleri
RAM	Rastgele Erişimli Bellek
RF	Rastgele Ağaç
RNN	Yinelenen Sinir Ağları
SSWE	Duyguya Özgü Kelime Gömmeleri
SVM	Destek Vektör Makineleri
TF	Tensorflow
URL	Tekdüzen Kaynak Bulucu
UTC	Eşgüdümlü Evrensel Zaman
VADER	Değerlik Duyarlı Duygu Analizi Sözlüğü
YSA	Yapay Sinir Ağları
kNN	K-En Yakın Komşu

## TABLULAR DİZİNİ

<b>Tablo No</b>	<b>Tablo Adı</b>	<b>Sayfa No</b>
<b>Tablo 1</b>	İsim öbeği için önerilen kurallar ve örnek tamlamalar	16
<b>Tablo 2</b>	Görevi, uygulanan metodu ve kullanılan öznitelikleri gösteren önceki çalışmaların özeti	28
<b>Tablo 3</b>	Çalışmada kullanılan, 1 Aralık 2019–30 Temmuz 2022 tarihleri arasında sosyal medyada dolaşıma sokulan yalan haber başlıkları	32
<b>Tablo 4</b>	Twint uygulamasından örnek CLI komutu örnekleri	33
<b>Tablo 5</b>	İngilizcede kullanılan resmi kısaltmalar	37
<b>Tablo 6</b>	Gayri resmi kısaltmalar	38
<b>Tablo 7</b>	İnternet argo kısaltmaları	38
<b>Tablo 8</b>	Sağlık veri setinden rastgele seçilen üç tweetin metin normalleştirmeden önceki ve sonraki hali	39
<b>Tablo 9</b>	Kategorilerine göre yalan haber başlıkları ve bu yalan haberlere verilen yorumların sayısı	50
<b>Tablo 10</b>	YSA ve LSTM modelleri için eğitim, doğrulama ve test verilerinin doğruluk ve hata sonuçları	60

## ŞEKİLLER DİZİNİ

Şekil No	Şekil Adı	Sayfa No
Şekil 1	Çalışmada kullanılan sistem mimarisi	28
Şekil 2	FactCheck.org haber doğrulama sitesi	29
Şekil 3	Twint uygulamasından örnek bir CLI komutu	32
Şekil 4	Metin normalleştirme için uygulanan iş akışı	33
Şekil 5	Duygu sınıflandırmada kullanılan sözde kod	39
Şekil 6	“CERN Opened a Portal to Another Dimension” veri setinin ilk on yorumu için duygu analizi sonuçları	40
Şekil 7	Dört etkileşimli katman içeren LSTM'deki yinelenen modül	45
Şekil 8	Sağlık veri seti için duygu analizi sonuçları	49
Şekil 9	Toplum ve sosyal meseleler veri seti için duygu analizi sonuçları	50
Şekil 10	Sanat ve eğlence veri seti için duygu analizi sonuçları	52
Şekil 11	Siyaset veri seti için duygu analizi sonuçları	53
Şekil 12	“Sağlık” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	54
Şekil 13	“Sağlık” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	54
Şekil 14	“Toplum ve Sosyal Meseleler” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	55
Şekil 15	“Toplum ve Sosyal Meseleler” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	55
Şekil 16	“Sanat ve Eğlence” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	56
Şekil 17	“Sanat ve Eğlence” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	56

<b>Şekil 18</b>	“Siyaset” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	57
<b>Şekil 19</b>	“Siyaset” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri	57



## BİRİNCİ BÖLÜM

### GİRİŞ

Günümüzde sosyal medya, hızlı ve pratik oluşuyla haberin ilk kaynağı olarak kullanılmaktadır. Haberleri sosyal medyadan alma oranı sadece 2021 yılında Avrupa ülkelerinde ortalama %40'ın üzerindeyken Amerika'da %42, Asya ülkelerinde ise bu oran ortalama %50'nin üzerine çıkmaktadır (Newman, vd., 2021). Ancak sosyal medya kullanıcılarının karşılaştıkları haberlerin geçerliliğini sorgulanmadan veya kontrol etmeden doğru kabul etmeleri bu durumun olumsuz sonuçlarındandır. Tam da bu durum yalan haberlerin yayılmasındaki başlıca sebeplerden biridir. Halihazırda sosyal medya; metin, ses, görüntü veya video gibi unsurlarla yalan haberin desteklenebilmesi ve haberlerin teyit edilmesinin son derece zor olmasıyla yalan haberi hızla yayan bir doğaya sahiptir. Bununla birlikte kullanıcılar da haberleri beğenerek veya yeniden paylaşarak yalan haberin yayılmasında büyük rol oynamaktadır. Ayrıca sosyal medyada dolaşan yalan haberler tanıdığımız ve güvendiğimiz kişi ve gruplar aracılığıyla ulaşılabilir hale geldiğinde kolaylıkla meşru bir nitelik kazanabilmektedir (Tandoc, vd., 2018). Bu yönüyle sosyal medya, yalan haberin önu alınamaz bir hızla yayıldığı bir mecra haline gelmiştir.

Yalan haber, insanları birtakım şüphelerle sevk edip bir dönem gündemde kaldıktan sonra etkisini yitiren bir olgu olarak görülmemelidir. Yalan haberin etkilerinin ne denli ciddi boyutlara vardığını önceki yıllarda yaşanan örneklerine bakarak anlayabiliriz. Nitekim 2016 yılında, ABD başkanlık seçimi ve Brexit referandumu ile 2017 yılında Fransa başkanlık seçimleri boyunca, yalan haberler süreci çok yoğun ve ciddi bir biçimde etkilemiştir (Yerlikaya ve Toker, 2019). Hatta geçtiğimiz dönemde tüm insanlığı ilgilendiren COVID-19 salgını ve COVID-19 aşısının sözde zararları ve yan etkileriyle ilgili çıkan yalan haberler ve komplo teorilerinin insanlar üzerindeki etkilerine bakmak bile yalan haberin ulaşabileceği boyutları görmek açısından yeterlidir. Aşı karşıtlığı ve aşı kararsızlığı ile birlikte aşılama hızındaki düşüş insanların sosyal medyadaki COVID-19 aşısıyla ilgili yanlış bilgiye maruz kalmasına bağlanmaktadır (Islam, vd., 2021).

Sosyal medya ekosistemi için, haberlerin doğruluğu ve geçerliliği kanıtlara dayalı verilerden ziyade kanaat ve duyguların kontrolünde olduğundan hangi duygunun yalan

haberi yaymada daha etkili olduđu sorusu ortaya çıkmaktadır. Bu tez çalışmasında, kullanıcıların habere verdikleri tepkiyle haberin yaygınlığı arasında bir ilişki olup olmadığını araştırılmaktadır. Kullanıcılar yalan haberle karşılaştıklarında hangi duygu türü onları bu yalan haberi yaymaya sevk etmekte? Diğer bir deyişle, yalan habere verilen tepkiyle haberin yaygınlığı arasından bir ilişki var mı? Eğer bir ilişki varsa, bu duygu her haber türü için aynı mıdır? soruları çalışmanın araştırma sorularıdır. Araştırma sorularıyla ve içeriğiyle başlı başına özgün olan bu tez çalışmasında, haberin ilk kaynağı ve gündemi öğrenmek için gidilen ilk yer olarak bilinen Twitter platformundaki yorumlar kullanılmıştır. Twitter kullanıcılarının 31 Aralık 2019–30 Temmuz 2022 tarihleri arasında çeşitli kategorilerde yalan haberlere verdikleri tepkiler üzerine duygu analizi yapılmıştır. Bu kategoriler; “Sağlık”, “Toplum ve Sosyal Meseleler”, “Sanat ve Eğlence” ve “Siyaset” olmak üzere dört farklı türden oluşmaktadır. Çok yaygından az yaygına doğru farklı yaygınlık düzeyinde seçilen yalan haber yorumları karşılaştırmalı olarak incelenmiştir. Duygu kutupluluğunu sınıflandırmak için, sözlük ve kural tabanlı bir duygu analizi aracı kullanılmıştır. Veriler TensorFlow Hub’da önceden eğitilmiş metin gömmeleri kullanarak kurulan model ile LSTM yöntemi kullanılarak eğitilmiş, metin sınıflandırma problemi hangi modelin daha iyi sonuç verdiği karşılaştırılmıştır.

### **1.1. Yalan Haber**

Fake news, dilimizde en yaygın kullanımıyla yalan haber, bilgilerin doğruluğunu ve güvenilirliğini sağlamaya yönelik editoryal normlarından ve süreçlerinden yoksun, haber medya içeriğini biçim olarak taklit eden ve yanlış bilgi içeren fabrikasyon bilgiler olarak tanımlanmaktadır (Lazer, vd., 2018). Yalan haberler, iki ana boyutta ele alındığında: birinci olarak, gerçek bilgi içermez, dolayısıyla yanıltıcı ve ikinci olarak da kasten aldatma niyeti taşımaktadır. Yalan haberler, geleneksel haber formatını taklit etme girişimiyle de diğer dezenformasyon –yanlış bilgi verme– biçimlerinden ayrılmaktadır. Dolayısıyla, yalan haberler insanları aldatmaya yönelik üretilir ve bunu gerçek bir haber gibi görünmeye çalışarak yapar (Tandoc, 2019).

### **1.1.1. Yalan Haber Siteleri**

Yalan haberlerin varlığı yeni olmasa da, internet ve sosyal medya, bunların oluşturulma ve yayılma şeklini değiştirmiştir. Yalan haber siteleri güven vermek, meşru ve gerçek bir haber sitesiymiş gibi görünmek için, haber sitelerinin alan adlarını, site tasarımlarını ve içeriklerini mümkün olduğunca kopyalamaktadır. Bununla birlikte, tıklama tuzağı (clickbait) içeren yalan haber siteleri ise insanların ilgisini çekmek ve bu siteleri ziyaret etmeye ikna etmek için merak uyandırıcı ve dikkat çekici manşetlerle göz alıcı ve ilgi uyandıran görseller kullanmaktadır. Yalan haber sitelerinin bu arayüzleri, kaynağının ne olduğuna veya kimin paylaştığına bakılmaksızın insanların sahte bir habere tıklamasına neden olan temel sebeplerdendir.

Kasıtlı olarak yanıltma niyeti taşıyan yalan haberlerin üretilme motivasyonu ideolojik ve ticari sebeplere dayanmaktadır. Bir takım siyasi hareketleri veya fikirleri desteklemek amacıyla bireyler veya grupların yalan haber üretmesi ideolojik bir motivasyon olmakla beraber kamuoyunu manipüle etme misyonu taşır. Bir diğer yalan haber üretimi motivasyonu ise ticaridir. Yalan haberlerin yayınlandığı internet siteleri tıklanma sayıları ile önemli reklam geliri elde edebilmektedir. Bu durumların en çarpıcı örnekleri yine ABD'nin 2016 yılı başkanlık seçimi sürecinde yaşanmıştır. Başkanlık seçiminden önce Facebook'taki en popüler on yalan haberin dördünden sorumlu olan "endingthefed.com" adlı internet sitesinin kurucusu Trump'ın başkanlığı kazanmasına yardım etmek için siteyi kurduğunu iddia etmiştir (Allcott ve Gentzkow, 2017). Diğer yandan, 2016 seçimlerinden sonra BuzzFeed ve Guardian tarafından yapılan ayrı soruşturmalarda, yalan haber yayınlayan yüzden fazla sitenin Makedonya'nın Veles kasabasında gençler tarafından yönetildiği, gençlerin hem Trump hem de Clinton lehine ürettikleri yalan haberlerle on binlerce dolar kazandığı ortaya çıkmıştır (Subramanian, 2017).

### **1.1.2. Sosyal Medyada Yalan Haber**

Günümüzde yalan haber siteleri mümkün olabildiğince çok dikkat çekmek için sosyal medyadan yararlanmaktadır. Sosyal medya, yalan haber sitelerinin erişilebilirliğini



arttıran, ziyaretçi çekmelerini ve dolayısıyla daha çok tıklanma almalarını kolaylaştıran bir yapıya sahiptir. Sosyal medyada dolaşıma giren bir yalan haberin yayılması ise sosyal ağın aktörleri olan kullanıcılarla mümkündür.

Sosyal medya kullanıcılarının özellikle belli davranışları yalan haberin yayılmasındaki ana etkenlerdendir. Bunlardan en önemlisi, kullanıcıların haber makalesini okumadan başlıklar ve etiketler bazında hareket etme eğilimidir. Birçok kullanıcı, haberi okumadan yalnızca başlığa bakarak haber makaleleri hakkında fikir oluşturmaktadır. Haber siteleri de sundukları haber makaleleri hakkında yapılan yorumların çoğunun makalenin kendisinden değil de manşete tepki veren kişilerden geldiğini bildirmektedir (Gabiolkov, vd., 2016). NPR tarafından yapılan bir sosyal deneyle bu durumun endişe verici bir örneğini görmekteyiz. Şaka amaçlı olarak, NPR Facebook sayfasında "Amerika Neden Artık Okumuyor?" başlıklı bir haber linkini paylaştığı deneyde kullanıcılar bağlantıya tıkladıklarında, NPR'nin web sitesinde makalenin şaka olduğunu açıklayan bir sayfaya yönlendirileceklerdi. Ancak birçok takipçinin yazdığı yorumların makaleyi açıkça okumadan hatta haber linkine bile tıklamadan doğrudan başlığa yanıt olarak verdikleri gözlemlenmiştir ("CITS", 2018a). Okumadan paylaşmak, beğenmek ve yorum yazmak haberleri popülerlik kazanıyor veya trend oluyormuş gibi gösterebilir. Yalan haberlerin yayılması ve etkisi açısından bu oldukça zarar vericidir.

Popülerlik temeli üzerine kurulu sosyal medyada, karşılaşılan bilginin popülerliği de onun kabul görmesindeki en büyük etkidir. Bir gönderiye çok sayıda beğeni, paylaşım veya yorum eşlik ettiğinde, başkaları tarafından ilgi görme ve dolayısıyla daha fazla beğenilme, paylaşılma veya yorumlanma olasılığı yüksektir (Tandoc, vd., 2018). Bu sosyal dinamik, sosyal medyadaki popülerlik göstergelerinin doğal sonucu olarak ortaya çıkmıştır. Bu durum, popülerliğin düşüncelerimizi nasıl etkilediğini açıklayan "çoğunluk etkisi" denilen olguyla daha iyi anlaşılabilir. Sosyal medya kullanıcıları, çoğunluk etkisiyle, içeriğin kendisinden ziyade bir şeyin kaç kez paylaşıldığına veya beğenildiğine odaklanmaktadır. Bir haberin popüleritesi, kullanıcıların bilgileri doğrulama sorumluluğunu atlamalarına neden olabilir. Bilginin popülerliğine dair algımız; "binlerce kişi bir haberi paylaştıysa veya beğendiyse mutlaka birileri doğrulamıştır" veya "bu kadar insan yanılıyor olamaz" şeklinde kişinin kendini çoğunluğun etkisine bıraktığı bir

düşünceye evrildiğinde bilginin sorgulanmadan kabul görmesi durumu ortaya çıkmaktadır. Ancak, daha önce de bahsedildiği üzere, yorum ve beğeniler sahipleri tarafından okunmadan da gerçekleşmiş olabilmektedir. Bununla birlikte, yalan haberler kullanıcılara takip ettiği arkadaşlar veya güvenilir hesaplar aracılığıyla da ulaşabilmektedir. Bu şekilde yalan haberi paylaşma, beğenme ve yorum yazma gibi örtülü veya açık desteklemelerle kullanıcıların bu haberlere inanmaları kolaylaşmaktadır.

Sosyal medyada, kaynaktan bağımsız bir habercilik anlayışıyla hareket edilmektedir. Yani, orijinal kaynağın kalitesi, güvenilirliği veya politik eğilimleri ne olursa olsun, çok çeşitli kaynaklardan haber makaleleri toplanmakta ve sunulmaktadır. Haberi kimin yazdığı veya bilgi sağladığı kolayca belirlenemediği için çoğu kullanıcı haber için ayrıntılı bir kaynak kontrolü yapmadan doğru kabul etmektedir. Bu durumu kullanıcıların haber kaynağını kontrol etme alışkanlığını köreltmekle beraber yalan haberin yayılımını hızlandırmaktadır.

### **Sosyal Botlar**

Sosyal medya, insanların yanı sıra yalan haber yayan milyonlarca "sosyal botu" da bünyesinde barındırmaktadır. Programlama becerisine sahip kimseler tarafından oluşturulan ve koddan meydana gelen botlar sosyal medya platformlarının neredeyse hepsinde bulunmaktadır. Çeşitli amaçlar için üretilen botlar, çoğunlukla yalan haberleri yaymaya ve yalan haberlerin sosyal medyadaki görünür popülaritesini artırmak için kullanılmaktadır. 2017 yılındaki bir tahmine göre Twitter'da 23 milyon, Facebook'ta 140 milyon ve Instagram'da ise yaklaşık 27 milyon bot hesap bulundurmaktadır (de Lima Salge ve Berente, 2017).

Sosyal botları Python, Node.js veya Ruby gibi programlama dillerinde programlayarak oluşturmak mümkün olduğu gibi artık hiç kod yazmadan da bot oluşturulabilmektedir (Agarwal, 2017). Botlar, sosyal medyadaki insanların davranışlarını simüle ederek diğer kullanıcılarla etkileşime geçer, bilgi ve mesaj paylaşımı yapabilme yeteneğine sahiptir. Yapay zeka ile desteklendirilen botlar yanıt kalıplarından veya girdi

değerlerinden öğrenerek belirli durumlara yanıt verebilmektedir ("CITS", 2018b). Örneğin, Twitter'da botlar, sıradan insanlar gibi görünmelerini sağlayan bir dizi sosyal etkileşim yeteneğine sahiptir. Başkalarının gönderilerine veya sorularına kullanmaya programlandıkları komut dosyalarına göre yanıt verirler. Etkili veya çok sayıda takipçisi olan Twitter kullanıcılarını ararlar, diğer Twitter kullanıcılarında güven oluşturmak için sorular göndererek onlarla iletişim kurar ve fark edilmeye çalışırlar. Ayrıca, diğer web sitelerinde buldukları konularla ilgili bilgileri arayarak ve tekrarlayarak trend olan konular hakkında paylaşımlar yaparak tartışmalar oluşturabilirler (Ferrara, vd., 2016).

Yalan haberlerin sosyal botlar tarafından yayılması, botların web üzerinde henüz doğrulanmamış bilgileri arama ve alma yetenekleri nedeniyle mümkündür. Sosyal medya sitelerindeki botlar, daha geniş kitlelere ulaşmak için uyguladıkları ana strateji trend konuları ve hashtag'leri kullanmaktır. Yalan haberleri trend konu veya etiketlerle birlikte paylaşarak kullanıcılar tarafından daha görünür olmasına ve dolayısıyla daha da çok yayılmasına sebep olurlar. Sürekli paylaşmaya devam ettikleri bu yalan haberleri gerçek sosyal medya kullanıcılarının gönderilerine yanıt vermek için de kullanırlar ("CITS", 2018b).

Ortalama olarak sosyal medya kullanıcıları, başkalarının paylaştığı gönderileri veya trend olan konuları sorgulamadan veya bilgi paylaşımın kullanıcı profiline dikkatlice bakmadan inanma eğiliminde oldukları için botların bu yöntemleri işe yaramaktadır. Bu nedenle botlar, yüksek hacimli sahte haberler yayımlayarak ve güvenilir görünmesini sağlayarak bundan yararlanır (Ferrara, vd., 2016). Botlar çok fazla yalan haber yaymasına karşın yalan haberlerin yeniden iletimlerinin çoğu gerçek kişiler tarafından sağlanmaktadır (Vosoughi, vd., 2018). Dolayısıyla yalan haberlerin yayılmasında gerçek kişiler önemli bir rol oynamaktadır.

## **1.2. Sosyal Medya**

Sosyal medya terimi, kullanıcıların çevrimiçi platformlara katılmalarına, kullanıcı tarafından oluşturulan içeriğe katkıda bulunmalarına veya çevrimiçi topluluklara

katılmalarına olanak sađlayan geniř bir internet tabanlı ve mobil hizmetler yelpazesini ifade eder (Dewing, 2010). Sosyal medyayla yaygın olarak iliřkilendirilen internet hizmet türleri řunları içerir: Sosyal ađ siteleri, medya paylaşım siteleri, durum güncelleme hizmetleri, bloglar, wikiler, sosyal yer imi siteleri ve sanal dünya içeriđi.

Sosyal ađ siteleri, bireylerin sınırlı bir sistem içinde halka açık veya yarı halka açık bir profil oluřturmasına, bađlantı paylařtıđı diđer kullanıcıların bir listesini oluřturmasına, kendi bađlantı listelerini ve sistem içinde bařkaları tarafından yapılanları görüntülemesine ve gezmesine izin veren web tabanlı hizmetler olarak tanımlanmıřtır. En popüler olanlar arasında Facebook ve LinkedIn bulunmaktadır. Medya paylaşım siteleri, kullanıcıların video veya fotođraf göndermesine izin veren platformlardır. Popüler örnekler arasında YouTube, Pinterest ve Instagram bulunmaktadır. Durum güncelleme hizmetleri ise Twitter gibi mikroblog hizmetleri olarak da bilinmektedir. Kullanıcıların kiřiler veya olaylar hakkında kısa güncellemeler paylařmasına ve bařkaları tarafından oluřturulan güncellemeleri görmesine olanak tanır. Bloglar, web günlüđünün kısaltması olup sayfaların genellikle ters kronolojik sırada görüntülendiđi çevrimiçi bir dergidir. WordPress, Tumblr ve Blogger en yaygın örnekleridir. Wikiler, herhangi bir katılımcının web tarayıcısını kullanarak herhangi bir sayfayı deđiřtirmesine veya yeni bir sayfa oluřturmasına izin verilen toplu bir web sitesidir. İyi bilinen bir örnek olarak wiki teknolojisini kullanan ücretsiz bir çevrimiçi ansiklopedi olan Wikipedia verilebilir. Sosyal yer imi siteleri, kullanıcıların web sitelerinde bađlantılar düzenlemesine ve paylařmasına olanak tanır. Popüler örnekler arasında reddit, StumbleUpon ve Digg bulunmaktadır. Son olarak sanal dünya içeriđi ise, kullanıcıların etkileşimde bulunduđu oyun benzeri sanal ortamlar sunar. Kullanıcıların bařkalarıyla avaturları vasıtasıyla etkileşime girdiđi Second Life sanal dünya içeriđine örnek teşkil etmektedir (Dewing, 2010).

Sosyal medya insanların çevrimiçi etkileşim biçimini etkileyebilecek niteliklerle işbirliđi ve tartışma için yeni yollar açmaktadır. Bu niteliklerden biri kalıcılıktır. Sosyal medya sitelerinde yayımlanan çok sayıda içerik varsayılan olarak orada sürekli olarak kalabilmektedir. Bir diđer niteliđi ise tekrarlanabilir olmasıdır, yani içerik paylařılabilir ve kopyalanabilir. İçeriđin, çevrimiçi arama araçları kullanılarak kolayca bulunabilir olmasıyla aranabilirlik niteliđi de oldukça önemlidir. Son olarak erişilebilirlik özelliđi ile

sosyal medya, internet bağlantısının olduğu her yerde ve her zaman kullanılabilir. Bu nitelikler, çevrimiçi sosyal etkileşimin dinamiklerini şekillendirmektedir (Dewing, 2010).

### **1.2.1. Sosyal Medyanın Tarihçesi**

Kullanıcıların içerik oluşturmaya ve yüklemesine izin veren web siteleri, 1990'ların sonlarında, geniş bant internetin daha popüler hale gelmesiyle ortaya çıkmaya başlamıştır. İlk sosyal ağ sitesi "SixDegrees.com" 1997'de ortaya çıkmıştır. 2002'den itibaren ise çok sayıda sosyal ağ sitesi açılmaya başlamıştır. 2000'lerin sonunda sosyal medya yaygın bir kabul görmüş ve bazı hizmetler çok sayıda kullanıcı kazanmıştır. Örneğin, Kasım 2012'de Facebook, dünya çapında 1 milyar kullanıcıya sahip olduğunu duyururken Temmuz 2012'de Twitter tahmini olarak 517 milyon kullanıcıyı bünyesinde barındırmaktaydı (Dewing, 2010).

Sosyal medya kullanımındaki bu hızlı büyümeye birçok faktör katkıda bulunmuştur. Bunlar, artan geniş bant kullanılabilirliği, yazılım araçlarının iyileştirilmesi ve daha güçlü bilgisayarların ve mobil cihazların geliştirilmesi gibi teknolojik faktörlere bağlanmaktadır. Bununla birlikte, bilgisayar ve yazılım satın alınabilirliğinin artması ve sosyal medya sitelerine karşı artan ticari ilgi gibi ekonomik faktörler ve sosyal medyanın genç yaş grupları tarafından hızlı bir şekilde kabul görmesi gibi sosyal faktörler de bu hızlı büyümede oldukça etkilidir (OECD, 2007).

### **1.3. Twitter**

Twitter bir mikroblog ve sosyal ağ servisi. Bu yönüyle Twitter, insanların istedikleri herhangi bir konu veya kişi hakkında yazdıkları bloglarla benzerlik göstermektedir. Ancak yazılan mesajların belirli bir karakter limitiyle sınırlandırılması bakımından bloglardan ayrılır. Mesaj göndermek, tweet olarak bilinmektedir. Diğer kullanıcılarla etkileşime geçmek tweetler aracılığıyla sağlanmaktadır. Bir tweetin içeriği kişi, kurum, ürün veya olaylardan kişisel görüşlere kadar değişebilmektedir. Tweet'in

bileşenleri ise metin, fotoğraf, video veya bağlantı gibi formatlarda olabilmektedir. Tweetler başta en fazla 140 karakter olabilirken 2017’de bu sınırlama ikiye katlanarak 280 karaktere çıkarılmıştır (Rosen, 2017). Tweetler herkes tarafından görülebileceği gibi sadece kullanıcının onayladığı Twitter takipçileri tarafından görülebilecek şekilde bir gizlilik ayarına sahiptir.

Tweet göndermek için bir kullanıcının platforma kayıtlı olması gerekir. Kullanıcı kayıt sırasında bir kullanıcı adı seçer ve bunu daha sonra tweet göndermek için kullanılır. Bir tweet’teki "mention" olarak bilinen bahsetmeler, gönderinin başka bir kullanıcıdan bahsettiğini ifade eder. Bu referansı bir kullanıcı adına yapmak için, kullanıcılar "@" sembolünü ve ardından başvurdukları belirli kullanıcı adını "@kullanici\_adi" olacak şekilde ilave ederler. Bahsetmeler, tweet gövdesinin herhangi bir yerine yerleştirilebilir. Bir tweet’teki yanıtlar gönderinin başka bir tweete cevap niteliğinde olduğunu belirtmek için kullanılır ve genellikle sohbet oluşturmak için kullanılır. Bahsetmelere benzer şekilde, @ sembolü ve ardından başvurdukları kullanıcı adı kullanılarak oluşturulurlar. Yanıtlar, yanıtı oluşturan kullanıcı adının yanına yerleştirilir. Takipçiler, bir kullanıcının tweetlerini ve aktivitelerini takip eden kullanıcıları ifade etmektedir. Takip etmek, Twitter’daki diğer kullanıcılarla bağlantı kurmanın ana yoludur. Twitter’daki kullanıcılar, takip ettikleri kişilerden güncellemeler alır ve kendi güncellemelerini de onları takip eden kullanıcılara gönderir. Retweet denilen kavram ise yeniden dağıtılan tweetleri ifade etmek için kullanılır. Bir kullanıcı bir tweeti ilginç bulduğunda, retweetleme işlevini kullanarak tweeti yeniden gönderebilir. Retweetleme, bilgiyi yaymak için güçlü bir araç olarak kabul edilir. Paylaşılan tweet değişmeden kalır ve genellikle RT kısaltması ve ardından yazarın kullanıcı adı ile işaretlenir. Retweetleyen kullanıcı ayrıca kısa bir yorum ekleyebilmektedir. Hashtag olarak bilinen etiketler, bir tweetin belirli bir konuyla ilgisini belirtmek için kullanılır. Etiketler "#" karakteri ve ardından konu adı kullanılarak oluşturulur. Hashtag’ler, gönderilen tweetlerdeki bilgileri etiketleme ihtiyacından ortaya çıkmıştır. Etiketler kullanıcılar tarafından eklenebilir ve aynı etiket ile tüm tweetleri almak için kullanılabilir. Çok sayıda tweet’te görülen etiketler trend olan konular olarak nitelendirilmektedir (Giachanou ve Crestani, 2016).

### **1.3.1. Neden Twitter?**

Geçtiğimiz dönemlerde araştırmacılar, çok sayıda fikir, görüş ve duygu içerdikleri için sosyal medya sitelerindeki metinleri analiz etmektedirler. Bu sosyal medya sitelerinden özellikle Twitter'ın kullanılmasının sebebi; kullanıcıların fikirlerini, düşüncelerini ve genel olarak ilgilendikleri herhangi bir konuyla ilgili her türlü bilgiyi paylaşmaları ve çok sayıda kullanıcıyı çekmeyi başarmasından dolayıdır. Twitter'da paylaşılan bilgiler ve güncellemeler sıklıkla ürünler, hizmetler, ünlüler, etkinlikler veya kullanıcının ilgisini çeken herhangi bir şey hakkında kişisel görüşlerini içermektedir. Çeşitli görevler için sosyal medya verilerini analiz eden birçok araştırmacı Twitter'ı kullanmıştır (Giachanou ve Crestani, 2016).

Genel olarak Facebook'tan daha az popüler olmasına rağmen, gazeteciler, siyasetçiler ve politikacılar tarafından yaygın olarak kullanılan Twitter haberlerin ilk yayınlandığı yer olarak bilinmektedir. Bu yönüyle Twitter, haberleri ilk kaynağından öğrenmeyi sağlamakla birlikte insanların habere verdikleri tepkileri görmek, görüşlerini okumak ve atışmaları takip etmek adına başvurdukları ilk yer haline almaktadır. Reuters Enstitüsü Dijital Haber Raporuna (2021) göre sosyal medya platforlarını haber için kullananların çoğu haberi tesadüfen aldıklarını söylemiştir. Buna karşılık Twitter, genellikle haberleri almak için birincil hedef olarak görülmüştür (Newman vd, 2021). Pew Research Center'ın Amerika'da yaptığı ankete göre ise 2020 yılında Twitter kullanıcılarının %59'u siteyi haber kaynağı olarak kullanmıştır (Shearer ve Mitchell, 2021). Bu yönüyle Twitter, yalan haberlerin yayılımını incelediğimiz çalışmamız için veri toplanacak en verimli sosyal medya platformu haline gelmektedir.

### **1.4. Duygu Analizi**

Duygu Analizi (Sentiment Analysis), insanların ürünler, hizmetler, organizasyonlar, kişiler, olaylar, konular ve bunların nitelikleri gibi varlıklara yönelik görüşlerini, değerlendirmelerini, tutumlarını ve duygularını analiz eden çalışma alanıdır (Liu, 2012). Bu alanda, özellikle duygu tespitinde doğal dil işleme teknikleri (Natural Language Processing) kullanılmaktadır. Duygu tespiti, doğal dil işleme ve bilgi alımının (Information

Retrieval) birleştigi bir disiplin olduđu için bilgi çıkarımı, metin madenciliđi ve hesaplamalı dilbilim gibi görevlerin birtakım özellikleri benzerlik göstermektedir (Vinodhini ve Chandrasekaran, 2012).

Duygu Analizi ve Fikir Madenciliđi genellikle birbirinin yerine kullanılmaktadır. Ancak, bazı arařtırmacılar, bu iki kavramın birbirinden biraz farklı iki sorunu çözmeyi amaçladıklarını belirtmektedir. Buna göre, Fikir Madenciliđi bir metin parçasının görüş içerip içermediđini belirlemekle ilgilidir, bu aynı zamanda öznellik analizi olarak da bilinen bir problemdir (Tsytarau ve Palpanas, 2012). Fakat Duygu Analizinin odak noktası, incelenen metnin duygu kutupluluđunun incelenmesidir, yani olumlu, olumsuz veya nötr bir duygu verip vermediđiyle ilgilidir. Daha kapsamlı olarak, Duygu Analizi temel olarak olumlu veya olumsuz duyguları ifade veya ima eden görüşlere odaklanır (Liu, 2012).

#### **1.4.1. Duygu Analizinin Önemi**

Kullanımının artmasıyla birlikte sosyal medya, kullanıcıların kişisel görüşlerini ifade etmeleri için en popüler araç haline gelmiştir. Bireylerin hayatlarındaki günlük olayları paylaştığı, fikir ve duygularını ifade ettiđi bir ortam olmasıyla sosyal medya çeşitli konular hakkında sayısız düşüncenin bulunabileceđi bir mecradır. Paylaşılan bu kadar çok bilginin manuel olarak okunması, analiz edilmesi ve bundan bir anlam çıkarılması mümkün değildir. Zira bu devasa veri yığını insan okumasıyla baş edilemeyecek derecede büyüktür, deđişkendir ve en önemlisi de sürekli artmaktadır. Büyük miktarda veriyle karşı karşıya olmanın yanı sıra veriler birçok farklı kaynaktan üretilmektedir. Bu nedenle bir kişinin farklı kaynaklara bakması ve faydalı bilgiler elde etmesi neredeyse imkansızdır. Dolayısıyla metin kutupluluđunu otomatik olarak tespit edebilen ve çok büyük miktarda veri içinde bile fikirlere dayalı bilgileri etkin bir şekilde arařtırabilen bir yaklaşım olarak duygu analizine ihtiyaç duyulmaktadır. Duygu analizi herhangi bir şey hakkında yapılan görüşleri toplamak ve incelemek için bir sistem oluşturmayı sağlamaktadır (Vinodhini ve Chandrasekaran, 2012).



Kullanıcılardan gelen yorumlar ve görüşler iyi bir fikir kaynağıdır. Bir kavram hakkında kamuoyunu anlamayı gerektiren çeşitli uygulamalar için son derece önemlidir. Örneğin duygu analiziyle, bir ürün veya hizmet hakkındaki görüşleri yakalayabilmek, yeni bir ürünün başarısını yargılamak ve popülerliğini ortaya çıkarmak mümkündür. Bu geri bildirimleri, şirketler hizmetlerinin veya ürünlerinin kalitesini iyileştirmek veya yeni politikalar ortaya koymak için kullanabilmektedir. Bununla birlikte duygu analizi, hükümetlerin farklı sosyal konularda kamunun görüşünü anlaması ve hızlı davranması da mümkün kılmaktadır. Ayrıca, kullanıcıların ilgilendikleri çeşitli konular hakkındaki duygularının gelişimini ve değişimini görebilmeyi de sağlamaktadır.

#### **1.4.2. Duygu Analizi Yaklaşımları**

Herhangi bir doğal dil işleme probleminde olduğu gibi, metinden duyguları analiz etmek, doğal dilin yinelemeli ve belirsiz doğası nedeniyle zordur. Günümüzün dijital dünyasında düzenli olarak yeni kelimeler ve semboller üretilirken dilin sürekli evrim geçirmesi duygu analizi görevine ciddi zorluklar getirir (Kannan, vd., 2016).

Duygu sınıflandırma teknikleri, kabaca leksikon tabanlı yaklaşım, kural tabanlı yaklaşım, makine öğrenimi yaklaşımı ve bunların birkaçını birlikte kullanan hibrit yaklaşım olarak dörde ayrılabilir. Leksikon tabanlı yaklaşım, bilinen ve önceden derlenmiş duygu terimlerinin bir koleksiyonu olan bir duygu sözlüğüne dayanır. Makine öğrenimi yaklaşımı, yaygın kullanılan makine öğrenimi algoritmalarını uygular ve modelin öğrenmesi için dilin özelliklerini kullanır. Hibrit yaklaşım, her iki yaklaşımı birleştirir ve yöntemlerin çoğunda önemli bir rol oynayan duygu sözlüklerinde çok yaygın olarak kullanılmaktadır (Medhat, vd., 2014). Kural tabanlı yaklaşım ise duygu yoğunluğunu ifade etmek ve vurgulamak için önceden tanımlanmış bir dizi dilbilgisi ve sözdizimsel kuralları dikkate almaktadır (Hutto ve Gilbert, 2016).

## Leksikon Tabanlı (Lexicon-Based) Yaklaşım

Leksikon tabanlı yaklaşım, sözcüklerin anlamsal yöneliminden belgenin veya cümlelerin duygu yönelimi hesaplanmaktadır. Sözcüklerin anlamsal yönelimi pozitif, negatif veya nötr olabilir (Gupta ve Agrawal, 2020). Bu yaklaşımda, belgedeki tüm kelimelerin duygu puanlarını toplamak ve belgeyi puanlamak için önceden hazırlanmış bir duygu sözlüğü (sentiment lexicon) kullanılır. Önceden hazırlanmış duygu sözlüğü, kelimeleri ve bu kelimelere karşılık gelen duygu puanlarını içermektedir (Kannan, vd., 2016). Cümledeki kelimelerin duygu puanlarının tek bir puana dönüştürülmesiyle metnin duygu yönelimi hesaplanır.

Aynı kelimelerin olumsuz formu duygu sözlüğüne ayrı bir girdi olarak eklenmekte ve bu kelimelerin olumsuz olmayan haline göre daha yüksek öncelik almaktadır. Olumsuzluk getiren terim veya sözcükleri işlemek için basit kurallar da kullanılabilir. (Kannan, vd., 2016). Örneğin TextBlob, cümlelerin polaritesini hesaplamak için puanları bulunan kelime ve ağırlık sözlüğüdür. Leksikon tabanlı bir duygu analizcisi olmakla birlikte polariteyi hesaplamak için önceden tanımlanmış bazı kurallara da sahiptir. Bu nedenle Leksikon tabanlı duygu analizcilerine "Kural Tabanlı (Rule-Based) Duygu Analizcileri" de denilmektedir (Barai, 2021). Ancak, bu iki kavram birbirinin yerine kullanılsa da farklı metodolojileri ifade etmektedir. Bu nedenle, bir sonraki bölümde Kural Tabanlı Yaklaşım ayrıca incelenecektir.

Leksikon tabanlı yaklaşımın temel avantajı, eğitim verisi gerektirmemesidir (Giachanou ve Crestani, 2016). Ancak bununla birlikte, bu yaklaşımın bazı dezavantajları da bulunmaktadır. Bir metinde, olumlu kelimelerin daha çok bulunması, metnin kesin olarak olumlu olduğu anlamına gelmez. Aynı durum tersi için de geçerlidir. Çoğunlukla farklı alanlara veya konulara ait metinleri puanlamak için aynı sözlüğü kullanmak mümkün değildir. Bunun için, hedef alanın doğasına dayalı olarak yeni bir duygu sözlüğü hazırlanmalıdır (Kannan, vd., 2016). Leksikon tabanlı yaklaşıma yönelik literatürdeki iki farklı yaklaşım vardır. Bunlar Sözlük Tabanlı (dictionary-based) yaklaşım ve Derlem Tabanlı (corpus-based) yaklaşımdır (Liu, 2012).

Sözlük Tabanlı (dictionary-based) yaklaşımdaki süreç, olumlu veya olumsuz duygu yönelimleri bilinen küçük duygu sözcüklerinin manuel olarak toplanmasıyla başlamaktadır. Algoritma daha sonra WordNet'te veya başka bir çevrimiçi sözlükte eş anlamlılarını ve zıt anlamlılarını arayarak bu seti büyütmektedir. Yeni bulunan kelimeler bu listeye eklenir ve bir sonraki yineleme başlar. Yinelemeli süreç, daha fazla yeni kelime bulunamadığında sona erer. İşlem tamamlandıktan sonra listeyi temizlemek için manuel inceleme yapılabilmektedir (Liu, 2012). Bu yöntem temelinde, "olumlu bir kelimenin eş anlamlısı olumlu, olumsuz bir kelimenin eş anlamlısı ise olumsuzdur" ve "olumlu bir kelimenin zıt anlamlısı olumsuz, olumsuz bir kelimenin zıt anlamlısı ise olumludur" fikrine dayanmaktadır.

Sözlük tabanlı yaklaşım kullanmanın avantajı, çok sayıda kelimenin duygu yönelimlerinin kolayca ve hızlı bir şekilde bulunabilmesidir. Ortaya çıkan listede birçok hata bulunabilse de onu temizlemek için manuel bir kontrol yapılabilir. Manuel kontrol bir dezavantaj olarak görülse de sadece bir kerelik bir çaba gerektirir. Sözlük tabanlı yaklaşımın esas dezavantajı, bu şekilde toplanan kelimelerin duygu yönelimlerinin belirli bir alan ve bağlamdan bağımsız olmasıdır. Yani bir kelime, bir bağlamda olumlu anlam katarken farklı bir bağlamda cümleye olumsuz anlam katabilmektedir. Örneğin, "sessiz" kelimesi ele alındığında, makinanın sessiz çalıştığından söz ediliyorsa bu genellikle olumludur. Ancak bir cihaz hoparlörü için sessiz deniyorsa bu olumsuzdur. Bu örnekteki gibi kelimelerin duygu yönelimi, alana veya bağlama bağlı olarak değişmektedir. Derlem tabanlı (corpus-based) yaklaşım bu sorunla başa çıkmaya yardımcı olabilmektedir (Liu, 2012).

Derlem tabanlı yaklaşım, bağlama özgü sözcüklerin duygu yönelimini bulur (Gupta ve Agrawal, 2020). Genel amaçlı duygu sözcüklerinin listesiyle, bir etki alanı derlemi (corpus) kullanarak diğer duygu sözcüklerini ve bunların yönelimlerinin keşfedilmesiyle gerçekleştirilir. Ancak konu, alana özgü bir duygu sözlüğü oluşturmaktan daha karmaşıktır çünkü aynı alanda aynı kelime bir bağlamda olumlu, başka bir bağlamda olumsuz olabilir (Liu, 2012).

Bu yaklaşımın uygulandığı anahtar ve ilk fikirlerden biri (Hazivassiloglou ve McKeown, 1997), bir derlem ve duygu ifade eden başlangıç sıfatlarını kullanarak, derlemdeki diğer duygu ifade eden sıfatların bulunmasını ve böylece bağlama ait duygu sözcüklerinin ortaya çıkmayı önermektedir. Teknik, duygu içeren sıfatların ve bunların derlemdeki yönelimlerini belirlemek için bağlaçlar üzerindeki dilsel kurallardan yararlanmaktadır. Örneğin kurallardan biri, sıfatların genellikle aynı yönelime sahip olduğunu söyleyen “ve” bağlacı ile ilgilidir. “Bu araba güzel ve geniş” cümlesinde “güzel”in olumlu olduğu biliniyorsa, bu derlemde “geniş”in de olumlu olduğu çıkarımı yapılabilir. Bunun nedeni, "ve" bağlacının her iki tarafında da aynı duygunun ifade edilmesidir. “Bu araba güzel ve sürmesi zor” cümlesi pek olası değildir. “Bu araba güzel ama sürmesi zor” olarak değiştirilirse daha kabul edilebilirdir (Liu, 2012). Yine aynı şekilde, öncesinde söylenen olumluysa “ama” bağlacından sonra olumsuz bir ifade geleceğinden veya tersi söz konusu olduğundan, derleme ait duygu sözcükleri bu şekilde çıkartılabilmektedir. Kurallar ayrıca diğer bağlaçlar (or, but, either-or, neither-nor) için de tasarlanmıştır. Daha sonra, sıfatlar arasındaki bağlantılarla bir grafik oluşturur, pozitif ve negatif olmak üzere iki kelime grubu oluşturmak için grafik üzerinde kümeleme yapılır (Medhat, vd., 2014).

Derlem tabanlı yaklaşımı tek başına kullanmak sözlük tabanlı yaklaşım kadar etkili değildir çünkü tüm İngilizce kelimeleri kapsayacak büyük bir derlem hazırlamak zordur. Ancak bu yaklaşımın, bir etki alanı tümcesi kullanarak alana ve bağlama özel fikir sözcüklerini ve bunların yönelimlerini bulmaya yardımcı olabilecek büyük bir avantajı vardır (Medhat, vd., 2014). Son olarak, etki alanına özgü yönelimleri içerse bile bir duygu sözlüğü, bir kelimenin cümle içerisinde ifade ettiğini görüşü veya duyguyu belirlemede yetersiz kalabilmektedir. Örneğin, “Satın almak için iyi bir araba arıyorum” cümlesindeki “iyi” burada herhangi bir araba hakkında olumlu veya olumsuz bir görüş ifade etmez (Liu, 2012).

## Kural Tabanlı (Rule-Based) Yaklaşım

Duygu sözcükleri ve deyimlerinin yanı sıra, duyguları ifade etmek veya ima etmek için kullanılabilecek başka birçok ifade veya dilsel bileşen de vardır. Duygu sözlüğü ile birlikte, his ifadeleri, emoji, belirli deyim listeleri gibi duygu yönelimi değiştiren veya güçlendiren dilsel öğeleri de kullanan kural tabanlı yaklaşım bu listelerdeki terimlere önem atfeden bir dizi kurala sahiptir. Bu listelere ve her bir listeyle ilişkili önemine dayalı olarak bir belgeyi puanlamaktadır (Kannan, vd., 2016).

Örneğin Tablo 1’de, bir sıfatın bir isimle birleştirildiği isim öbeği için kurallar verilmiştir (Klenner, Petrakis, & Fahrni, 2009). Sözcüklerin duygu yönelimi, önceden derlenmiş bir duygu kutupluluğu sözlüğünden gelir. Örneğin, “mükemmel” olumlu sıfatı, olumsuz isim olan “ızdırıp” ile birleştiğinde olumsuz bir isim tamlaması ortaya çıkmaktadır. Aynı şekilde “mükemmel” kelimesi nötr olan “yemek” kelimesiyle birleştiğinde olumlu bir isim tamlaması ortaya çıkmaktadır. Kural tabanlı yaklaşımda, dildeki bu tarz düzenlilikler bir kural dizisi haline getirilerek metnin duygu kutupluluğunu ortaya çıkarmada kullanılmaktadır.

Tablo 1

İsim öbeği için önerilen kurallar ve örnek tamlamalar

Sıfat	İsim	→	İsim Öbeği	Örnek
Negatif	Pozitif	→	Negatif	Yıkılmış umutlar
Negatif	Negatif	→	Negatif	Berbat yalan
Pozitif	Pozitif	→	Pozitif	İyi dost
Pozitif	Negatif	→	Negatif	Mükemmel ızdırıp
Pozitif	Nötr	→	Pozitif	Mükemmel film
Negatif	Nötr	→	Negatif	Berbat film

(Klenner, vd., 2009)

Fiiller de tek başına bir polarite yönelimi taşıyabilir. “Sevmek” fiili olumlu, “nefret etmek” fiili olumsuzdur. “Hoşuna gitmek”, “beğenmek”, aynı zamanda “iğrenmek”, “beğenmemek” gibi fiillerin hepsinin kutupluluğu aşıkardır. Esas soru, fiillerin doğrudan nesnelereyle kombinasyonun, kompozisyonellik açısından nasıl yorumlanması gerektiğidir. “Kötü filmleri sever” cümlesindeki, “kötü filmler” ifadesinin olumsuz olduğu göz önüne alındığında cümlenin olumlu mu yoksa olumsuz mu olduğunu ortaya çıkarmak önemlidir. Aynı durum, “iyi kitaplardan nefret eder” cümlesinde de görülmektedir (Klenner, vd., 2009). Zarflar ise yoğunlaştırıcı görevi yaparak yönelimi etkilemezler ancak gücünü değiştirebilmektedir. Örneğin “çok iyi bir arkadaş”, “iyi bir arkadaştan” daha fazlasını ifade etmektedir. Bu yaklaşımın leksikon tabanlı bir yaklaşıma benzer sınırlamaları vardır. Sözlüğe dayalı ve kurallara dayalı yaklaşımların ana avantajı, yalnızca bir duygu sözlüğü oluşturmak için manuel çabaya ihtiyaç duyulmasıdır (Kannan, vd., 2016).

### **Makine Öğrenimi (Machine Learning) Yaklaşımı**

Makine öğrenmesi yaklaşımını kullanan metin sınıflandırma yöntemleri, kabaca denetimli ve denetimsiz öğrenme yöntemlerine ayrılabilir. Denetimli yöntemler, çok sayıda etiketli eğitim belgesinden yararlanır. Bu etiketli eğitim belgelerini bulmanın zor olduğu durumlarda denetimsiz yöntemler kullanılır. Makine öğrenimi yaklaşımı, duygu analizini sözdizimsel veya dilsel özelliklerden yararlanan düzenli bir metin sınıflandırma problemi olarak çözmek için ünlü makine öğrenmesi algoritmalarına dayanır (Medhat, vd., 2014). Makine öğrenmesi yaklaşımı, önceden tanımlanmış statik kuralları takip edecek şekilde ayarlanmış kurallara dayalı sözlüğün aksine, doğası gereği dinamiktir. Başka bir deyişle, sağlanan veri kümesinden öğrenir ve veriler modelin eğitilmesi için kullanılır. Buna bağlı olarak da modelinin tahmin doğruluğu, ilk etapta eğitim için sağlanan verilerin kalitesine bağlıdır. Genel olarak, derlemin boyutu da önemlidir, boyut ne kadar büyükse model doğru tahminler üretmek için o kadar iyi eğitilir.

Twitter’da duygu analiziyle ilgilenen araştırmacıların önerdiği yöntemlerin çoğu, tweetlerin çeşitli özellikleri üzerinde eğitilmiş makine öğrenimi alanından bir sınıflandırıcı

kullanır. En çok uygulanan sınıflandırıcılardan bazıları Naive Bayes (NB), Maksimum Entropi (MaxEnt), Destek Vektör Makineleri (SVM), Lojistik Regresyon (LR), Random Forest (RF) ve Şartlı Rastgele Alan (CRF) olarak belirtilmektedir (Giachanou ve Crestani, 2016).



## İKİNCİ BÖLÜM

### ÖNCEKİ ÇALIŞMALAR

Bir sınıflandırma problemi olan duygu analizi, başta metnin tamamı üzerinde uygulanarak belge düzeyinde (Turney, 2002) incelenmiş, ardından cümle düzeyinde (Hu ve Liu, 2004) ve daha sonralarında ise ifade düzeyinde (Wilson, vd., 2005) ele alınmıştır. Literatürde duygu analizi dört farklı yöntemle ele alınmıştır. Bunlar; makine öğrenimi tabanlı, leksikon tabanlı, makine öğrenimi ile sözlük tabanlıyı birlikte kullanan hibrit yaklaşım ve yapay sinir ağlarını tabanlı yaklaşımdır.

Makine öğrenimi yaklaşımı, fikir veya duygu içeren tweetleri algılayabilen bir sınıflandırıcı oluşturmak için, çeşitli özellikleri (feature) ve makine öğrenimi yöntemlerini kullanmaktadır. Sözlüğe dayalı yaklaşım ise incelenen metnin polaritesini ortaya çıkarmak için manuel veya otomatik olarak oluşturulmuş bir pozitif ve negatif terimler listesi kullanır. Hibrit yaklaşımı izleyen yöntemler, daha iyi bir performans elde etmek için makine öğrenimi ve sözlük tabanlı yöntemleri birleştirmektedir. Yapay sinir ağı kullanan yaklaşım ise derin öğrenme modelleri kullanmaktadır. Twitter'da duygu analiziyle ilgilenen önceki çalışmaları, görevine, uygulanan metotuna ve kullanılan özneliklerine göre özeti Tablo 2'de verilmiştir.

#### 2.1. Makine Öğrenimi Kullanan Çalışmalar

Twitter'da duygu analizi üzerine ilk çalışma Go ve arkadaşları tarafından gerçekleştirilmiştir (Go, vd., 2009). Go ve arkadaşları sorunu ikili bir sınıflandırma problemi olarak ele alarak tweetleri olumlu veya olumsuz olarak sınıflandırmıştır. Tweetlerin duygu kutupluluğunu manuel olarak etiketlemenin zorluğu nedeniyle Read (2005) tarafından önerilen uzaktan denetim (distant supervision) yaklaşımını kullandılar. Eğitim verilerini, tweetlerin sonuna eklenen his ifadelerine bakarak pozitif veya negatif olarak sınıflandırmıştır. Hem olumlu hem de olumsuz his ifadeleri içeren ve retweetlenen gönderiler veri setinin dışında bırakılmıştır. Nihai eğitim veri seti 1.6 milyon tweetten oluşurken manuel olarak etiketlenmiş 300 tweet de test veri seti olarak kullanılmıştır.



Naive Bayes, MaxEnt ve Destek Vektör Makinesi (SVM) sınıflandırıcılarını kullanarak modeller oluşturulmuştur. Özellik olarak bigrams, unigrams ve POS (part of speech) etiketleri kullanılmıştır. Özellik alanı çok seyrek olduğu için tek başına bigram ve POS etiketlerini kullanmanın polarite sınıflandırması için kullanışlı olmadığı sonucunu çıkarılmıştır. Ayrıca en etkili yöntemin, %82,7'lik bir doğruluk elde etmeyi başaran Naive Bayes sınıflandırıcısında özellik olarak unigram ve bigramları birlikte kullanarak elde edildiğini bildirilmiştir (Go, vd., 2009).

Öznel ve nesnel metinleri kullanarak sınıflama yapan Pak ve Paroubek'in öznel verileri toplama şekli, Go ve arkadaşlarının (2009) çalışmasındaki gibi tweet sonlarındaki his ifadelerine bakmasıyla benzerlik gösterirken nesnel verileri New York Times ve Washington Posts gibi gazete ve dergilerin twitter hesaplarından toplamasıyla önceki çalışmalardan ayrılır. Çalışmalarında, problemi çok sınıflı bir sınıflandırma görevi olarak ele almış olup tweetleri olumlu, olumsuz ve nötr olarak üç sınıfta sınıflandırılmıştır. Özellik olarak unigramlar, bigramlar, n-gramlar ve POS etiketlerini kullanarak SVM, MNB ve CRF'nin performansı karşılaştırılmıştır. Go ve arkadaşları (2009) tarafından sunulan sonuçların aksine POS ve bigramların her ikisinin de sınıflandırıcı performansını arttırmaya yardımcı olduğu bildirilmiştir. Elde edilen sonuçlar, en iyi kombinasyonun n-gram ve POS etiketleri ile birlikte çok terimli Naive Bayes (MNB) olduğu gözlemlenmiştir. Ayrıca daha fazla eğitim verisi ile performansın arttığı belirtilmiştir (Pak ve Paroubek, 2010).

Twitter verilerinde duygu sınıflandırması yapan bir başka önemli çalışma da Barbosa ve Feng'e aittir. Barbosa ve Feng, Twitter duygu analizi problemini iki aşamalı bir sınıflandırıcıyla ele almıştır. İlk adımda mesajın öznel veya nesnel olup olmadığı belirlenirken, ikinci adımda öznel tweet'lerin olumlu veya olumsuz olarak sınıflandırılması gerçekleştirilmiştir. Çalışmalarında, toplanan tweetlere açıklama eklemek için üç farklı duygu algılama aracından gelen bilgileri kullanılmıştır. Algılama araçlarından farklı duygu kutupları atanmış tweetler kaldırılarak 200.000 tweetlik bir eğitim veri seti elde edilmiştir. Sınıflandırıcıları eğitmek için meta-indeks ve sözdizimi özellikleri kullanılmıştır. Meta-indeks özellikleri, MPQA sözlüğünü kullanarak POS etiketlerini ve kelimelerin polaritesini içerirken sözdizimi özellikleri retweetleri, hashtag'leri, URL'leri ve his ifadelerini

içermiştir. Barbosa ve Feng, tweet verilerinde n-gram kullanmanın, Twitter'daki çok sayıda seyrek kelime olması nedeniyle sınıflandırma performansını engelleyebileceğini savunmuştur. Bunun yerine, retweet'ler, hashtag'ler, yanıtlar, noktalama işaretleri ve his ifadeleri gibi mikroblog özellikleriyle beraber kelimelerin önceki polaritesi ve POS gibi özelliklerle birlikte kullanılması önerilmiştir. SVM'leri eğitmek için bu özelliklerin kullanılmasının, yalnızca unigramlardan eğitilen SVM'lere kıyasla duygu sınıflandırma doğruluğunu %2,2 artırdığını keşfedilmiştir. En iyi sonuçlar, nesnellik tespiti için %81,9 ve polarite tespiti için %81,3 doğruluk elde eden SVM sınıflandırıcıyla elde edildiği belirtilmiştir. Çalışmanın ilginç bulgularından biriyse sözdizimi özellikleri öznellik tespiti için daha önemliyken, meta-indeks özellikleri polarite tespiti için daha önemli olduğudur (Barbosa ve Feng, 2010).

K-En Yakın Komşu (kNN) algoritmasına benzer denetimli bir yaklaşım sunan Davidov ve arkadaşlarının çalışmasında önceki yaklaşımların aksine, eğitim verilerini toplamak için tweetlerdeki etiketleri ve his ifadeleri kullanılmıştır. kNN benzeri sınıflandırma stratejisi için kullanılan geleneksel özelliklerin dışında, çalışmalarında ayrıca etiketler, gülen suratlar, noktalama işaretleri ve sık kalıplar kullanarak ikili sınıflandırma için ortalama %86.0'lık bir harmonik F skoru elde edilmiştir (Davidov, vd., 2010).

Özellik seçimi, denetimli makine öğrenmesi yöntemlerinin verimliliği açısından çok önemlidir (Giachanou ve Crestani, 2016). Bu amaçla, bir dizi araştırmacı (Agarwal, vd., 2011; Aisopos vd., 2011; Kouloumpis vd., 2011; Saif vd., 2012; Hamdan vd., 2013), farklı özelliklerin Twitter duygu analizi probleminin üzerindeki etkisini incelemiştir. Agarwal ve arkadaşları, farklı duygu analizi özelliklerini analiz etmek ve farklı yöntemlerin duygu analizi üzerindeki performansını incelemek için bir çalışma yapmıştır. Toplamda üç yöntem karşılaştırılmıştır, bunlar; karşılaştırma için taban görevi görececek olan unigramlar, kısmi ağaç çekirdekleri (partial tree kernels) tabanlı model ve özellik tabanlı model kullanılmıştır. Kısmi ağaç çekirdeği tabanlı modelde, tweetler için yeni bir ağaç temsili tasarlanmıştır. Özellik tabanlı yaklaşımda ise, en önemli özelliklerin, kelimelerin önceki kutuplarını ve onların POS bölümlerini birleştiren özellikler olduğunu ortaya çıkaran özellik analizi yapılmıştır. Hem kısmi ağaç çekirdeğinin hem de özellik tabanlı modelin unigram temel modelinden daha iyi performans gösterdiği belirtilmiştir.

Bu bulguyla beraber, 100 öznitelik kullanan özellik tabanlı modele ve 10.000'den fazla öznitelik kullanan unigram modeline nazaran ağaç çekirdeği tabanlı modelin diğer ikisinden daha iyi performans gösterdiği sonucu ortaya çıkmıştır (Agarwal, vd., 2011).

Farklı özelliklerin verimliliğini araştıran başka bir çalışma da Kouloumpis ve arkadaşları tarafından sunulmuştur. Bu çalışma çoğunlukla his ifadeleri, kısaltmalar ve tümü büyük harflerle yazılması veya karakter tekrarından oluşması gibi abartılmış kelimelerin varlığı gibi mikroblogların anlamsal ve stilistik özelliklerine odaklanılmıştır. Duygu kutupluluğunu tespit etmek için AdaBoost'un kullanıldığı çalışmada, eğitim verilerini meydana getirirken, tweet etiketlemesi için his ifadeleri kullanan Go ve arkadaşlarından (2009) farklı olarak hashtag'ler kullanılmıştır. Hashtag'ler manuel olarak pozitif, negatif ve nötr olarak etiketlenmiştir. Elde edilen sonuçlar, en iyi performansın n-gramları mikrobloglama özellikleri ve kelimelerin önceki kutuplarıyla etiketlendiği sözlük özellikleri ile birlikte kullanmaktan geldiğini göstermiştir. Bununla birlikte bu çalışma, Agarwal ve arkadaşları (2011) tarafından yapılan çalışmanın sonuçlarının aksine, POS'un duygu analizi üzerinde olumsuz bir etkiye sahip olduğunu, performansta bir düşüşe neden olduğunu göstermiştir (Kouloumpis, vd., 2011).

Bir dizi anlamsal özelliği inceleyen Saif ve arkadaşları tarafından yapılan çalışmada eğitim setine ek özellikler olarak semantik özellikleri eklemeye yönelik yeni bir yaklaşım sunulmuştur. Anlamsal özellikler, tweetlerden çıkarılan varlıkları (örn. "London", "Vodafone", "Steve Jobs") temsil eden anlamsal kavramlardan (örn. "şehir", "şirket", "kişi") oluşmaktadır. Tweetlerden çıkarılan her varlık (örn. "iPhone") için semantik konsepti (örn. "Apple ürünü") ayrıca bir özellik olarak eklenmiş ve temsili kavramın negatif veya pozitif duyguyla olan ilişkisi ölçülmüştür. Bu özellikleri kullanmanın ardındaki mantığın, belirli varlıkların ve kavramların olumlu veya olumsuz duygularla daha tutarlı bir korelasyona sahip olma eğiliminde olduğunu belirtmiş, bu korelasyonları bilmenin, anlamsal olarak ilgili veya benzer varlıkların duygu kutupluluğunu belirlemeye ve böylece duygu analizinin doğruluğunu artırmaya yardımcı olabileceğini savunulmuştur. Anlamsal özellikleri bir Naive Bayes sınıflandırıcısına dahil ettikleri mevcut çalışmada anlamsal özellikleri, unigram ve POS dizisi özelliklerinin yanı sıra duygu-konu özellikleriyle karşılaştırılmıştır. Elde edilen sonuçlar, anlamsal özellik modelinin hem

olumsuz hem de olumlu duyguları belirlemek için unigram ve POS taban çizgisinden daha iyi performans gösterdiğini açıkça göstermektedir (Saif, vd., 2012).

Aisopos ve arkadaşları (2011), sınıflandırma doğruluğunu artırmak için n-gram grafların kullanılmasını önermiştir. Çalışmada, yaklaşık 3 milyon tweet üzerinde değerlendirilen MNB ve C4.5 ağaç sınıflandırıcısı olmak üzere iki sınıflandırma algoritması kullanılmıştır. Eğitim verilerine ek açıklama eklenmesi, his ifadelerinin varlığına veya yokluğuna dayanmaktadır. Kapsamlı deneyler, en iyi modelin, mesafe ağırlıklı kenarlara sahip 4 gramlık bir graf üzerinde eğitilmiş bir C4.5 ağaç sınıflandırıcısı olduğunu göstermiştir. En iyi modelleri ikili sınıflandırma için %66.77 doğruluk elde edilmiştir (Aisopos, vd., 2011).

Asiaee ve arkadaşları tarafından Twitter'da duygu analizi için üç aşamalı kademeli bir sınıflandırıcı çerçevesi sunulmuştur. İlk adımda, ilgilenilen konunun tweetleri belirlenmiş, ikinci adımda duygu içeren tweetler saptanmış ve son adımda ise tweetler duygu polaritesine göre etiketlenmiştir. Bir dizi klasik yöntemin duygu analizi performansı incelenmiş ve ayrıca kNN, NB, ağırlıklı SVM ve Dictionary Learning içeren yeni algoritmalar da önerilmiştir (Asiaee, vd., 2012).

Twitter duygu analizini ele almak için en çok kullanılan sınıflandırıcılardan biri SVM sınıflandırıcıdır (Giachanou ve Crestani, 2016). Bakliwal ve arkadaşları (2012), çalışmalarında 11 özellik üzerinde eğitilmiş bir SVM sınıflandırıcı kullanmıştır. Etkinliklerini ölçmek için tek tek farklı ön işleme tekniklerinin kullanıldığı çalışmada yazım düzeltme, stemming ve stop-words kaldırma gibi ön işlemlerin sınıflandırıcının doğruluğunu artırmayı başardığı belirtilmiştir. Yaklaşımın değerlendirilmesi için Stanford veri kümesi (Go, vd., 2009) ve Mejaj (Bora, 2012) olmak üzere iki farklı veri kümesi kullanılmıştır. En iyi sonuçlar, doğal dil işleme ve Twitter'a özgü özelliklerin birleşimiyle elde edilmiştir (Bakliwal, vd., 2012).

Benzer şekilde Mohammad ve arkadaşları SemEval-2013 tarafından verilen veri setinde bir SVM sınıflandırıcı kullanmıştır. Her tweeti karakter n-gramları, POS,

hashtag'ler, sözlükler, noktalama işaretleri, his ifadeleri, ısrarlı veya vurgulanmış uzatma içeren kelimeler ve olumsuzlama içeren bir özellik vektörü olarak temsil edilmiştir. Bu özellikleri kullanarak eğitilen SVM sınıflandırıcısının, unigramlar üzerinde eğitilen temelden daha iyi performans gösterdiği gözlemlenmiştir. Yöntemleri, mesaj seviyesi analizi için %69.02 ve terim seviyesindeki görev için %88.93 F-skoru elde edilmiştir. Çalışmanın sonucunda, en kullanışlı özelliklerin sözlük özellikleri ve n-gramlar olduğu sonucuna varılmıştır (Mohammad, vd., 2013).

Hamdan ve arkadaşları, Twitter verilerinin duygu analizi üzerinde iyi bir performans elde etmek amacıyla birçok özellik ve kaynağı kullanmayı önerdiler. Çalışmalarının incelenen özellikler arasında DBPedia'dan kavramlar, WordNet'ten fiil gruplarıyla sıfatlar ve SentiWordNet'in senti özellikleri yer almıştır. Ayrıca duygu analizinin doğruluğunu arttırmak için duygular, kısaltmalar ve argo kelimeler sözlüğü kullanılmıştır. Unigramlar üzerinde eğitilmiş SVM ve NB sınıflandırıcıyla karşılaştırıldığında, bu özellikleri dikkate alarak uygulanan yöntemde F-score doğruluğu sırasıyla %2 ve %4 arttırılmıştır (Hamdan, vd., 2013).

Aston ve arkadaşları tweet veri akışlarındaki duygu analizi problemini inceledikleri çalışmada bellek ve işlem süresi sınırlamaları olan farklı denetimli yöntemler üzerine yoğunlaşmıştır. Tweetler, n-gram karakteri kullanılarak temsil edilirken bir gramın en iyi N özelliği seçilerek çok sayıda özellik azaltıldı. En iyi özelliklerin seçimi için altı farklı değerlendirme algoritması (Chi Squared, Filtered Feature, Gain Ratio, Info Gain, One R ve Relief) kullanılan çalışmada Perceptron sınıflandırıcısının üç farklı versiyonunu ve bunların farklı kombinasyonlarını karşılaştırılmıştır. En İyi Öğrenme Oranı ve Voted yöntemlerinin kombinasyonu, en iyi performansı verdiğini, öznellik ve duyarlılık analizi için sırasıyla %85 ve %78 F-skoru elde edilmiştir (Aston, vd., 2014).

## **2.2. Hibrit Yaklaşım Kullanan Çalışmalar**

Twitter verilerinden gelen duyguları ortaya çıkarma girişiminde, hem sözlük hem de derlem tabanlı yöntemlerin avantajlarını NLP ve ML temel tekniklerinin

kombinasyonuyla birleştiren Kumar ve Sebastian karma bir yaklaşım sunmuştur. Fikir kelimelerini, fiil ve zarflarla birlikte sıfatların birleşimi olarak ele aldıkları çalışmada, sıfatların anlamsal yönelimlerini bulmak için derlem temelli yöntem kullanılırken fiillerin ve zarfların anlamsal yönelimlerini bulmak için sözlüğe dayalı yöntem kullanılmıştır. İki unsuru birleştirmedeki motivasyonun; bir sıfatın, söz konusu belgeye duygunun yoğunluğunu tek başına aktaramayacağını ve bu nedenle de sıfatın kuvvetini değiştiren zarf ve fiillerin dikkate alınması gerektiği belirtilmiştir. Zarflar ve fiiller herhangi bir alana bağlı olmadığından, bunların anlamsal yönelimlerini hesaplamak için sözlük yöntemlerini kullanmanın gerekliliği üzerinde durulmuştur. Manuel olarak sık kullanılan zarf, fiil ve his ifadelerine -1 ile +1 arasında değişen ağırlık değerleri atayarak genel tweet duygusu daha sonra duygu yoğunlaştırıcıları da içeren doğrusal bir denklem kullanılarak hesaplanmıştır (Kumar ve Sebastian, 2012).

Liu ve arkadaşlarının çalışmalarında bir görüşün beş unsurdan oluştuğunu, bu unsurların; varlığın adı, varlığın yönelimi, varlığın yönelimine ilişkin duygusu, görüşün sahibi ve görüşün ifade edildiği zaman olduğunu belirtmiştir. Çalışmada, mevcut modellerin sınıflandırma için duygu etiketleri olarak ifadeler veya manuel olarak etiketlenmiş tweetlerin kullanıldığı fark edilmiş ve her iki kaynaktan da bilgi aktaran karma bir sistem önerilmiştir (Liu, vd., 2012).

Belirli bir alanda duygu analizi yaparak, alan bilgisinin duygu sınıflandırmasındaki etkisini belirlemek için yaptıkları çalışmada Neethu ve Rajasree yeni bir özellik vektörü sunmuşlardır. İki adımda özellik çıkarma yapılarak verimli bir özellik vektörünü Naive Bayes, SVM, Maximum Entropy ve Ensemble sınıflandırıcılar gibi farklı sınıflandırıcılar kullanılarak test edilmiştir. Tüm bu sınıflandırıcıların yeni özellik vektörü için neredeyse benzer doğruluğa (NaiveBayes %89,5 iken diğerleri %90 oranında) sahip olduğunu gözlemlenmiştir. Belirli bir alan için spesifik bir özellik vektörü oluşturulunca sınıflandırıcıların duygu analizi kapsamında iyi performansı gösterdiği sonucunu ortaya çıkmıştır (Neethu ve Rajasree, 2013).

### 2.3. Derin Öğrenme Kullanan Çalışmalar

2014 yılında Yapay Sinir Ağlarının yeniden yükselişe geçmesiyle duygu sınıflandırması için Derin Öğrenme kullanan çalışmalar da yapılmıştır. Derin öğrenmeyi kullanan ilk çalışmalardan birini yapan Tang ve arkadaşları (2014) word embeddings oluşturmak için 10 milyon tweet kullanmıştır. Kelimelerin sürekli temsilinde, duygu bilgilerini kodlayan “duyguya özgü kelime gömme” (SSWE) öğrenimiyle ele alınan çalışmada tweetleri uzaktan denetim (distant supervision) kullanarak toplanmıştır. Özellik olarak kullanılan SSWE’yi öğrenmek için üç sinir ağı geliştirdikleri ve SemEval-2013 veri setinde değerlendikleri çalışma pozitif/negatif tweet sınıflandırması görevinde %86 doğruluk elde edilmiştir (Tang, vd., 2014).

Kim, çeşitli evrişimli sinir ağlarını (CNN) incelemiş ve tek bir evrişim katmanıyla bile çoğu mimarinin oldukça iyi performans verdiğini göstermiştir (Kim, 2014). Dong ve arkadaşlarının (2014) ise Uyarlanabilir Özyinelemeli Sinir Ağı (AdaRNN) önerdiği çalışması, 6248 eğitim ve 692 test tweetinden oluşan manuel olarak etiketlenmiş bir veri setinde değerlendirilmiş ve %65,9 F1 skoru almayı başarmıştır (Dong, vd., 2014).

Dünyanın her yerinden çok sayıda kullanıcının düşüncelerine twitter üzerinden kolaylıkla erişebilme imkanı sayesinde, twitter üzerinden duygu analizi, yakın dönemde sosyal ve psikolojik alanda da uygulanan bir çalışma konusu olmuştur. Bu tarz çalışmalardan birini yapan Dzogang ve arkadaşları, İngiltere'den 800 milyon tweet'i analiz etmiş ve 73 psikometrik değişkenin günlük değişimini ölçmüştür. Çalışmada, 24 saatlik gün boyunca zıt zaman noktalarında zirveye çıkan “Kategorik Düşünme” ve “Varoluşçu Düşünme” adlı iki ana faktör bulunmuştur. Bu çalışmadan, dil kullanımını sirkadiyen ritimle ilişkilendiren ek biyolojik bilgiler de sağlanmıştır (Dzogang, vd., 2018).

Bir diğer çalışma ise Sneffjella ve arkadaşları tarafından yapılmıştır. Farklı kültürler arasında belirli dil türlerinin kullanımını ölçmek amacıyla 40 milyon tweet'i analiz ederek Kanada ve ABD arasındaki tweetlerin duygusal varyasyonunu ölçen ve Kanadalıların

Amerikalılardan ortalama olarak daha kibar olduđu klişeyi doğrulayan bir çalışma yapılmıştır (Sneffjella, vd., 2018).

Sosyal ağların popülerleşmesi sonucu yalan haberlerin yayılmasında ve sosyal etkisindeki görülen artış, sosyal ağlar üzerinde yalan haberleri incelemeye olan ilginin de artmasına sebep olmuştur. Yalan haberlerin başarısını desteklemek için yalan haber üreticilerinin kullandıkları hilelerden en kilit olanı okuyucuların duygularını heyecanlandırmaktır. Dolayısıyla duygu analizinin, hem haber öğelerine hem de kullanıcı yorumları gibi ilgili bilgilere uygulandığında, yalan haberleri inceleme kapsamında oldukça yararlı bir yöntem olduđu ortaya çıkmıştır (Alonso, vd., 2021). Bu bağlamda, Vicario ve arkadaşları (2019), sosyal medyadaki olası yanlış bilgilendirme hedeflerine ilişkin erken uyarı için bir çerçeve sunmuştur. Çalışmada yalan haberleri tespit etmek için birkaç klasik makine öğrenimi sınıflandırıcısını, metin özellikleri ve kullanıcı davranışlarına ait duygu temelli özelliklerle beslediler. En iyi performansı, bir Lojistik Regresyon sınıflandırıcısı ile elde edilmiştir (Vicario, vd., 2019). Bir diğerk çalışma da Cui, Wang ve Lee'ye aittir. Sosyal medyadaki kullanıcı yorumlarında bulunan gizli duyguların, sahte haberleri güvenilir içerikten ayırt etmeye yardımcı olup olmadığını araştırılmıştır. Deneysel sonuçlar, sistemin performansına en fazla katkıda bulunan bileşenin duygu analizi olduğunu göstermiştir (Cui, Wang ve Lee, 2019). Anoop, Deepak ve Lajish (2020) ise sağlık alanındaki yalan haberlerin tespiti üzerine bir çalışma yapmıştır. Çalışmada yalan haberi tespit etmek için, sahte ve gerçek sağlık haberlerinde gösterilen farklı türdeki duygusal özellikler incelenmiştir. Klasik ve derin öğrenme sınıflandırıcılarını beslemek için bir leksikondan duygu özellikleri çıkardıkları çalışmalarında duygu bilgisiyle tüm sınıflandırıcıların performansının artırdığını belirtilmiştir. Zhang ve arkadaşları (2021), yalan haberlerle ilgili mevcut çalışmaların çoğunun, viral yayılan sahte haberler için bile haberin kalabalıkta uyandırdığı duygulardan ziyade yayıncılar tarafından iletilen içeriğin duygusal sinyallerine odaklanıldığını belirtmiştir. Bu nedenle çalışmalarında haber yorumlarındaki duyguların ve içeriğin sahte haber tespiti için yararlı olup olmadığını araştırılmıştır. Makalelerinde önerdikleri özellik setinin mevcut yalan haber dedektörleriyle iyi uyumlu olabileceğini ve yalan haberleri tespit etme performansını etkili bir şekilde artırabileceği belirtilmiştir (Zhang, vd., 2021). Iwendi ve arkadaşları (2022) ise COVID-19 ile ilgili sahte haberleri tespit etmek için multimedya metinlerinden 39 özellik kullandıkları



bir model önermiştir. LSTM, GRU ve RNN gibi son teknoloji derin öğrenme modellerinin uygulandığı çalışmada sosyal medya verilerini elde etmek için bir Bilgi Füzyonu (Information Fusion) kullanılmıştır. Yinelemeli Sinir Ağı (RNN) kullandıkları model için %85 oranında kesinlik elde edilirken, sahte haberler için Geçitili Tekrarlayan Birim (GRU) model kullanıldığında en iyi F1 ölçüsünün %83 olduğunu belirtilmiştir (Iwendi, vd., 2022). Sahte haberlerin yayılımının incelendiği bir diğer çalışma da Bodaghi ve Oliveira'ya aittir. Yalan haberlerin yayılması sırasında farklı rollerdeki kullanıcıları inceledikleri çalışmalarında Twitter'da yayılan gerçek ve sahte haberler üzerine analiz gerçekleştirdiler. Bu amaçla, grafları kullanarak toplam 8 milyon düğüm ve 28 milyon bağlantı içeren Twitter'dan toplanan beş sahte haber üzerine inceleme gerçekleştirilmiştir. Çalışmalarıyla sosyal medyada yayılan sahte haberlerin doğasına daha iyi bir anlayış sağlanması amaçlanmıştır (Bodaghi ve Oliveira, 2022).

Günümüzün popüler Duygu Analizi araçları (TextBlob, VADER) leksikon ve kural tabanlı yaklaşımı esas almaktadır. Ancak bu araçları kullanmanın dezavantajı dilin dinamik oluşudur. Dile giren yeni kelimelerin sözlüğe de eklenmesi ve böylece sözlüğün sık sık güncellenmesi gerekmektedir. Çalışılan veri kümesi üzerinde daha iyi performans gösteren yeni dilsel kurallar keşfedildiğinde yeni kuralların da duygu analizcisine eklenmesi gerekmektedir. Tüm güncelleme işlemleri çalışılan dile hakim bir uzman eşliğinde ve manuel olarak yapılmak zorundadır. Bununla birlikte sözlük ve kural tabanlı duygu analizi araçlarının öğrenme kabiliyetleri yoktur. Bu sorunun üstesinden gelmek için çalışmamızda sözlük ve kural tabanlı yöntem ile yapay sinir ağlarını birleştiren hibrit bir sistem önerilmiştir. Duygu analizi aracıyla etiketlenen veriler yapay sinir ağı modelini eğiterek öğrenememe probleminin çözülmesi hedeflenmiştir. Böylece sözlük ve kural tabanlı yöntemin duygu analizi yapma yeteneği doğrudan yapay sinir ağına aktarılmaya çalışılmıştır.

Tablo 2

Görevi, uygulanan metodu ve kullanılan öznitelikleri gösteren önceki çalışmaların özeti

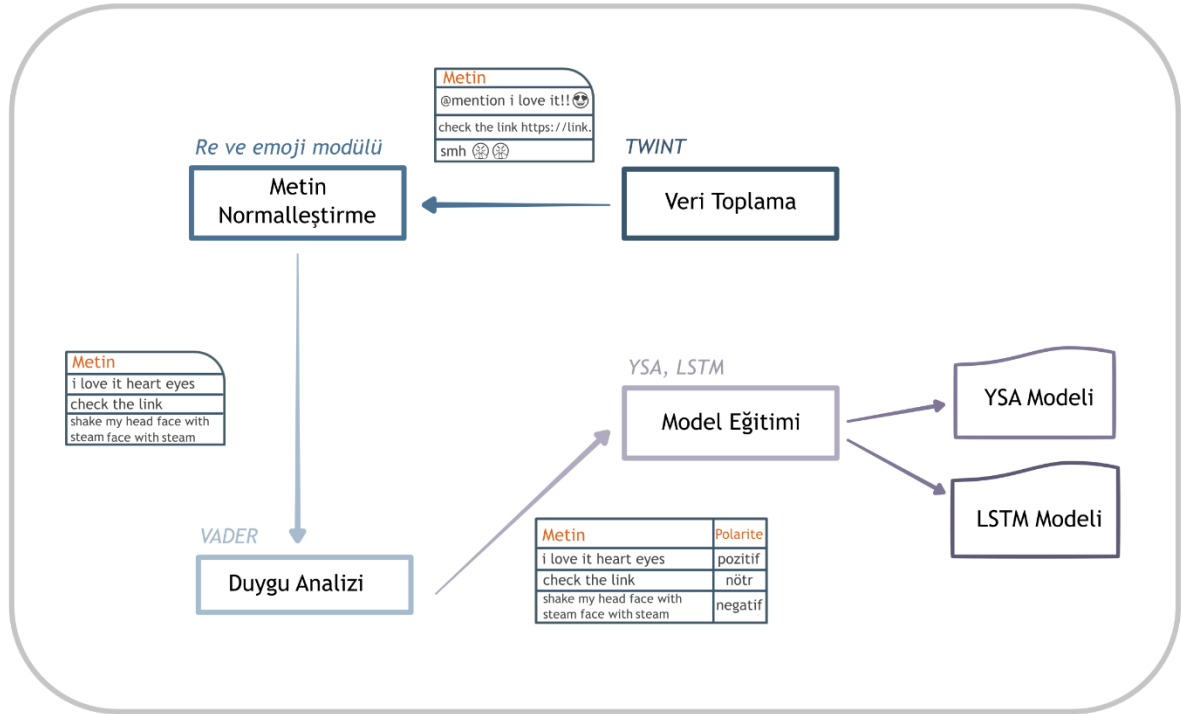
Çalışma	Görev	Metot	Öznitelikler
Go ve ark. (2009)	Twitter'da Duygu Analizi	SVM, MaxEnt, NB	POS, unigram, bigram

Pak ve Paroubek (2010)	Twitter’da Duygu Analizi	CRF, SVM, MNB	POS, unigram, bigram, trigram
Barbosa ve Feng (2010)	Twitter’da Duygu Analizi	SVM	Meta özellikler, Twitter’a özgü özellikler
Davidov ve ark. (2010)	Twitter’da Duygu Analizi	kNN	N-gram, emoji, sözcüksel özellikler (noktalama işaretleri gibi)
Agarwal ve ark. (2010)	Twitter’da Duygu Analizi	SVM	POS, unigram, sözcüksel özellikler (büyük harf, ünlem işareti gibi)
Aisopos ve ark. (2011)	Twitter’da Duygu Analizi	C4.5 Ağacı, MNB	N-gramlar
Kouloumpis ve ark. (2011)	Twitter’da Duygu Analizi	AdaBoost	POS, N-gramlar, Twitter’a özgü özellikler (hashtag gibi)
Asiaee ve ark. (2012)	Twitter’da Duygu Analizi	kNN, NB, Ağırlıklı SVM, Dictionary Learning	POS, N-gramlar
Saif ve ark. (2012)	Twitter’da Duygu Analizi	NB	POS, unigram, semantik özellikler
Kumar ve Sebastian, (2012)	Twitter’da Duygu Analizi	SentiTweet	Semantik özellikler, emojiler, Twitter’a ait özellikler (hashtag gibi)
Bakliwal ve ark. (2012)	Twitter’da Duygu Analizi	NB, SVM	Twitter’a ait özellikler (hashtag ve URL gibi) unigram, bigram, emojiler
Hamdan ve ark. (2013)	Twitter’da Duygu Analizi	NB, SVM	DBPedia’ dan kavramlar, WordNet’ten fiil grupları ve sıfatlar, SentiWordNet’in Senti özellikleri
Mohammad ve ark. (2013)	Twitter’da Duygu Analizi	SVM	POS, n-gramlar, sözcüksel özellikler (büyük harfler, noktalama işaretleri olumsuzlama gibi)
Neethu ve Rajasree, (2013)	Twitter’da Duygu Analizi	SVM, MaxEnt, NB, Ensemble Sınıflandırıcılar	POS, n-gramlar, emojiler, Twitter'a ait özellikler (hashtag, emojiler gibi)
Aston ve ark. (2014)	Twitter’da Duygu Analizi	Perceptron, Voted Perceptron, Ensemble Method	Karakter n-gramları
Tang ve ark. (2014)	Twitter’da Duygu Analizi	Duyguya Özgü Kelime Gömme (SSWE)	POS, n-gramlar, sözcüksel özellikler
Dong ve ark. (2014)	Twitter’da Duygu Analizi	Uyarlanabilir Özyinelemeli Sinir Ağı	Unigram, bigram

## ÜÇÜNCÜ BÖLÜM

### MATERYAL VE YÖNTEM

Çalışmada uygulanan yöntemlerin anlatıldığı bu bölümde, çalışmada gerçekleştirilen adımlara yer verilmiştir. Veri toplama, metin normalleştirme ve duygu analizi olmak üzere üç temel süreçle yürütülen çalışmadaki verilerin işlenişi ve analiz yöntemleri ile yapay sinir ağlarının eğitimi bu bölümde anlatılmıştır. Çalışmada kullanılan sistem mimarisi Şekil 1’de verilmiştir.



Şekil 1. Çalışmada kullanılan sistem mimarisi

Şekil 1, verilerin toplanmasıyla başlayan çalışmada metinlerin normalleştirme sürecinden geçirildiğini, ardından verilerin duygu analizine tabi tutulduğunu ve duygu analizi sonucunda elde edilen etiketlenmiş verilerle iki farklı yapay sinir ağının eğitildiğini göstermektedir. Her bir adım ve uygulanan yöntemler bu bölümün alt başlıklarında açıklanmıştır.

### 3.1. Veri Toplama

Çalışmamızda, yorumlarını toplamak için kullandığımız yalan haberler, bağımsız teyitçilerden olan Snopes, Lead Stories, FactCheck.org ve Checkyourfact gibi haber doğrulama sitelerinden elde edilmiştir.



Şekil 2. FactCheck.org haber doğrulama sitesi

(“FactCheck.org”, 2022)

Haber doğrulama sitelerinden yalan haber başlıkları manuel olarak toplanmıştır. Haber başlıkları “Sağlık”, “Toplum ve Sosyal Meseleler”, “Sanat ve Eğlence” ile “Siyaset” olmak üzere dört farklı kategoriye uygun olacak şekilde seçilmiştir. Çalışmamızda yorumlarını topladığımız yalan haber başlıkları ve yalan haberlerin dahil olduğu kategoriler Tablo 3’de verilmiştir.

Literatürde araştırmacıların çoğunun tweet toplamak için Twitter API’larını kullandıkları bilinmektedir (Giachanou ve Crestani, 2016). Ancak Twitter’ın ücretsiz erişim politikası, fiyatlar ve farklı erişim seçenekleri zaman içinde değiştiğinden, Twitter API’da dakika başına yapılabilecek istek sayısı, istek başına çekilebilecek tweet sayısı ve geçmiş verilere erişimde sınırlamalar bulunmaktadır (Antonakaki, vd., 2021). Bu sınırlamaların üstesinden gelmek için çalışmada tweet toplama aracı olarak Twint kullanılmıştır (Zacharias ve Poldi, 2018).

Tablo 3

Çalışmada kullanılan, 1 Aralık 2019–30 Temmuz 2022 tarihleri arasında sosyal medyada dolaşıma sokulan yalan haber başlıkları

<b>Kategorisi</b>	<b>Yalan Haber Başlıkları</b>
Sağlık	COVID-19 Vaccine Impairs Fertility COVID-19 Vaccine Can Monitor The Human With 5G Chip COVID-19 Vaccine Will Reduce The Planet's Population COVID-19 Vaccine Can Change Human DNA COVID-19 Vaccine Contains Aborted Fetal Cells COVID-19 Vaccine Make Our Body Magnetic Thousands of Americans Died After Receiving Their COVID-19 Vaccine Bill Gates Admits That a COVID-19 Vaccine Could Kill Up To 700K People
Toplum ve sosyal meseleler	Starbucks is Going Cashless in UK, US and Canada Four AI Robots Kill 29 Scientists in Japan Coca-Cola Add Try To Be Less White to Cans New Facebook-Meta Rule Permit Company To Use Your Photos CERN Opened a Portal to Another Dimension George Floyd Appeared on "Judge Judy" Tickets to Trumps 2nd Inauguration Circulate on Social Media Popeyes Hang a Sign Saying They Will Refuse to Serve White People
Sanat ve eğlence	Was Johnny Depp Offered \$301M Disney Deal for Pirates Return The Simpsons Predicted the Future Disney World is removing Cinderella Castle Actor Rowan Atkinson passed away Disney Deny Tim Allen Lightyear Role Due to His Political Beliefs Emma Watson Retiring From Acting The Storyline of "I Am Legend" Takes Place in 2021 Captain America Predicted the Coronavirus Pandemic
Siyaset	Maricopa County Elections Database Deleted Biden Say "White Republican Men Are More Dangerous Than ISIS" Former U.S. President Donald Trump Could Be Reinstated as President by August 2021 Joe Biden is Wanted in Ukraine on Class A Felony Charges The US Navy Has a Flying Aircraft Carrier Jill Biden Gave Speech in Front of a Nazi-Inspired Flag

Twint, Twitter'ın API'sini kullanmadan Twitter profillerinden sınırsız miktarda tweet toplamaya izin veren gelişmiş bir tweet toplama aracıdır. Python programlama dilinde yazılmış olan Twint açık kaynak kodlu bir yazılımdır. Belirli kullanıcılardan gelen tweetleri taramaya, belirli konular, etiketler ve trendlerle ilgili tweetleri toplamaya veya e-

posta ve telefon numaraları gibi tweetlerdeki hassas bilgileri ayıklamayı sağlamak için Twitter'ın arama operatörünü kullanır. Twitter API'sinin aksine hız sınırlaması olmayan Twint, Twitter'a kaydolmayı gerektirmeden anonim olarak da tweetleri çekebilmeği sağlamaktadır (Zacharias ve Poldi, 2018).

Topladığı tweetleri txt uzantılı dosya, CSV, JSON, SQLite ve Elasticsearch olmak üzere beş farklı metin formatında kaydedebilen Twint, komut satırı ara yüzünde (CLI) komutları çalıştırılabilmekle birlikte özel biçimlendirmeyi destekleyen bir modül olarak da kullanılabilir. Twint uygulamasının bazı temel CLI komutları Tablo 4'te verilmiştir.

Tablo 4

Twint uygulamasından örnek CLI komutu örnekleri

<b>Twint CLI Komutu</b>	<b>Kullanımı</b>
-u kullanıcıadi	Kullanıcının zaman çizelgesindeki tüm tweetleri toplar.
-u kullanıcıadi -s "sorgu"	Kullanıcının "sorgu" kelimesini içeren tüm tweetlerini toplar.
-s "sorgu"	Bütün kullanıcılardan "sorgu" kelimesini içeren tweetleri toplar.
-s "sorgu" --replies	Bütün kullanıcılardan "sorgu" kelimesini içeren bir konuya verilen yanıtlarını görüntüler.
-u kullanıcıadi --since "2015-12-20"	Kullanıcının "2015-12-20"den beri mevcut tüm tweetlerini toplar.
-u kullanıcıadi -o file.csv --csv	Kullanıcının tüm tweetlerini toplar ve csv dosyası olarak kaydeder.

(Zacharias ve Poldi, 2018)

Çalışmamızda tweetlerin toplanması Twint'in CLI komutlarını kullanarak gerçekleştirilmiştir. Twint'in CLI komutlarına; arama anahtar kelimesi, tweetlere verilen yanıtlar ve yalan haberin ilk defa görülme tarihi gibi parametrelere ilgili değerler verilerek tweetler toplanmıştır. Örneğin "COVID-19 Vaccine Impairs Fertility" şeklindeki bir yalan haber başlığından "covid vaccine impact fertility", "covid vaccine cause infertility" ve "covid vaccine cause miscarriage" gibi anahtar kelimelerle sorgulama yapılmıştır.

Çeşitli parametreleri içeren Twint uygulamasından örnek bir CLI komutu Şekil 3’te verilmiştir. Anahtar kelimeler “covid vaccine impairs fertility” ile “-s” parametresi arama yapılacağı, “--replies” parametresinin yazılmasıyla yanıtları görüntüleneceği ve “--since” parametresine “2020-12-01” değerinin verilmesiyle belirtilen tarihten itibaren günümüze kadar tüm tweetlerin getirilmesi ve son olarak da “-o” parametresiyle çıktıyı “dataset.txt” isimli dosyaya kaydetmesi komutta belirtilmiştir. Gizlilik gereği kullanıcı isimleri ve bahsetmeler anonimleştirilmiştir.

```
C:\Users\maide\Downloads\twint-master>python cli.py -s "covid vaccine impairs fertility"
--replies --since "2020-12-01" -o dataset.txt
1389551096654401547 2021-05-04 15:02:52 +0300 <[REDACTED]> @[REDACTED] Despite the fact th
at—no doubt—the world would be best served by less procreation from @[REDACTED] and hi
s loyal viewers, he and they should know that COVID-19 itself, not the vaccine, impairs m
ale fertility. https://t.co/I8DC8GMOFE
1363126826860691456 2021-02-20 17:02:15 +0300 <[REDACTED]> @[REDACTED] I read the b
rief statement, watched the short video. The Dr said there's no data showing covid vaccin
es impairs fertility. Then he said FDA is still reviewing the data on fertility. Kind of
a big issue to dismiss so flippantly. I prefer to see the data BEFORE getting the shot.
1354559253944672259 2021-01-28 01:37:47 +0300 <[REDACTED]> @[REDACTED] My friend in t
he UK went for the COVID vaccine.She was asked if she still wanted to have babies cos the
vaccine impairs fertility... How come no one is telling Africans this as we queue up f
or supplies? You see why we need to produce our own vaccines? #vaccine
```

Şekil 3. Twint uygulamasından örnek bir CLI komutu

### 3.2. Metin Normalleştirme

Tweetler toplandıktan sonra veri setleri metin normalleştirme aşamasından geçirilmiştir. Metin temizleme olarak da bilinen metin normalleştirme doğal dil işleme (NLP) için ham metin hazırlama sürecidir. Verilerin analize en uygun olacak şekilde işlenebilmesi açısından metin normalleştirme kritik bir öneme sahiptir.

Twint ile toplanan tweetler sırasıyla tweet ID’sini, tweet’in atılma tarihini, saatini, UTC tanımlayıcısını, kullanıcı adını ve tweet gövdesini içermektedir. Twint’in getirdiği gürültüye ek olarak sosyal medyanın gündelik ve gelişigüzel dili de metin normalleştirme sürecini zorunlu kılmaktadır. Bu kapsamda, uygulanacak analizle ilgisi olmayan nitelikler veri setinden kaldırılmıştır. Metin normalleştirme için izlenen iş akışı Şekil 4’te verilmiştir.



Şekil 4. Metin normalleştirme için uygulanan iş akışı

Twitter'a ait karakterlerin silinmesi ile başlayan metin normalleştirme iş akışı metindeki değerli bilgiyi ortaya çıkarmayı sağlamaktadır. Kişisel verilerin korunması kanunu başta olmak üzere, yapılacak analizle ilgisi olmayan kullanıcı isimleri ve bahsedilmeler silinerek tweetler anonimleştirilmiştir. Konuyu tek başına temsil edecek derecede özet forma indirgenmiş yapıyla etiket verisi önemli bilgiler içerebilmektedir. Dolayısıyla Twitter'da etiketlemek için kelimenin başına yazılan kare (#, hashtag)



karakteri tweetlerden kaldırılmış, etiketler kelimelerden oluşan sade bir forma getirilmiştir. Bilgiyi yaymak için kullanılan retweet duruma verilen tepkileri yakalamak konusunda fayda sağlayacağından retweet metinde tutulurken yeniden tweet'lendiğini ifade eden RT kısaltması silinmiştir. Son olarak metin içinde verilen linklerin kaldırılmasıyla Twitter'a özgü gürültü elimine edilmiştir.

Metne uygulanan NLP analiz araçları leksikon tabanlı olduğu için kelimelerin kullanılan analiz araçlarının sözlüğündeki karşılıklarıyla birebir eşleşmesi gerekmektedir. Örneğin “happy 😊” ile “happy 😄” iki ayrı kelime olarak algılanır. Sözlükte “happy” kelimesinin karşılığı bulunurken “happy 😄” kelime ve emojiden oluşan grubun karşılığı bulunmayacağı için duygu analiz aracı bu grubu etkisiz veya nötr olarak etiketleyecektir. Benzer durum n-gram üretiminde de ortaya çıkmaktadır. Büyük/küçük harf düzeni farklı olan “Great” ve “great” kelimelerine n-gram analizcisi farklı birer sözcükmüş gibi davranmaktadır. Dolayısıyla yazılması gerektiği şekilde yazılmayan her kelime analize zarar verecek ve ifade edilen duygunun ortaya çıkamamasına neden olacaktır. Bu nedenle kelimelere yapışık olarak yazılan his ifadeleri ve emojileri birbirinden ayırarak ve tüm kelimeleri küçük harfe dönüştürerek analiz araçlarının sözcükleri doğru biçimde tanınması sağlanmıştır.

Metindeki duyguyu ortaya çıkarmada en çok katkı sağlayan bileşen emojilerdir. Dolayısıyla emojiler, kelimelerden ayrıldıktan sonra metne kattıkları anlamlarıyla değiştirilmiştir. Örneğin “😲” emojisinin yerini “astonished\_face” veya “👍” emojisinin yerini “thumbs\_up” kelime öbeği almıştır. Böylece, duygu analizcisinin cümlede kutupluluğu etkileyebilecek her türlü bileşeni hesaba katması sağlanmıştır. Bu dönüşüm Python programlama dilinde açık kaynak kodlu olarak geliştirilen Emoji modülü kullanılarak sağlanmıştır (Kim, vd., 2014). Emoji modülü de tıpkı diğer doğal dil işleme araçları gibi leksikon tabanlı bir uygulamadır. Metinde emojiyi tespit ettiğinde, kendi sözlüğünden bu emojiye karşılık gelen sözcük öbeğiyle değiştirerek dönüşüm işlemi gerçekleştirmektedir. Emojilerin yerini alan sözcük öbekleri alt çizgi karakteriyle birbirine bağlıdır. Metin temizliği aşamasında sonradan ortaya çıkan bu durum Emoji modülünün

metne getirdiđi bir gürültü olduđundan kısaltmaların dönüştürülmesi aşamasından sonra temizlenmiştir.

Twitter'daki karakter sınırlaması nedeniyle tweetlerde kısaltmalar sıklıkla kullanılmaktadır. Kısaltmalar, karakter sınırını korumayı ve hızlı yazmayı sağlamasının yanı sıra metnin duygu yönelimi hakkında önemli ipuçları sağlayabilmektedir. Cümlenin sonuna eklenen bir kısaltma konuya yaklaşımın ne yönde olduđunu gösterebilir veya metnin duygu yönelimini tamamıyla tersine çevirebilir. Bununla birlikte metindeki anlam bütünlüğü kısaltmaların ne ifade ettiđinin ortaya çıkartılmasıyla mümkün olmaktadır. Bu nedenle kısaltmaların metne kattıđı anlam ve duyguyu yakalayabilmek için metin üç kademeli bir kısaltma açma sürecine tâbi tutulmuştur. Sırasıyla; İngilizcede kullanılan resmi kısaltmaların açılması, günlük konuşma dilinde kullanılan gayriresmi kısaltmaların açılması ve internet argo kısaltmaların açılması ile gerçekleştirilmiştir. Kısaltmalar manuel olarak oluşturulmuş olup sözlük (dictionary) veri yapısı türünde tutulmuştur. Kısaltma sözlüklerinden örnekler Tablo 5, Tablo 6 ve Tablo 7'de verilmiştir.

Tablo 5

İngilizcede kullanılan resmi kısaltmalar

Kısaltma	Açılımı
's	Is
Aren't	Are not
Can't	Can not
'll	Will
Mustn't	Must not
Won't	Will not

(“*Contractions*”, 2016)

Kısaltmaların açılmasının ardından Emoji modülüyle gelen kelime aralarındaki alt çizgiler temizlenmiştir. Şekil 2'de görüldüğü gibi Twint ile çekilen tweetler; tweet ID'si, tweet'in atılma tarihi, saati ve UTC tanımlayıcısı gibi sayısal veriler içermektedir. Bu veriler metne uygulanacak analizde kullanılmayacağı ve katkı sağlamadığından

kaldırılmıştır. Söz konusu sosyal medya olduğunda noktalama işaretleri ve özel karakterlerin aşırı kullanımı oldukça yaygındır. Metindeki bu tip gürültünün elimine edilmesi için alfa numerik karakterler hariç tüm karakterler silinmiştir. Bu aşamada metindeki noktalama işaretleri de silinmiştir. Son olarak da beyaz boşluk (whitespace) karakterlerinin kaldırılmasıyla metin normalleştirme süreci tamamlanmıştır.

Tablo 6

Gayri resmi kısaltmalar

<b>Kısaltma</b>	<b>Açılımı</b>
Cmon	Come on
Gotta	Got a
Tell'em	Tell them
Supposta	Supposed to
Betcha	Bet you
Ain't	Is not

(*"Informal contractions list"*, 2020)

Tablo 7

İnternet argo kısaltmaları

<b>Kısaltma</b>	<b>Açılımı</b>
d00d	Dude
C&G	Chuckle and grin
L8R	Later
ROFL	Rolling on floor laughing
QQ	Crying
MFW	My face when

(*"Internet Slang"*, 2021)

Sağlık veri setinden rastgele seçilmiş üç tweet için metnin ham hali ile metin normalleştirmeden sonraki halinin örneği Tablo 8'de verilmiştir. Twint ile Twitter'dan

toplandıktan sonraki ilk hali “Ham Metin” sütununda, metin normalleştirme uygulandıktan sonraki hali “Normalleştirilmiş Metin” sütunundadır. Gizlilik gereği kullanıcı isimleri ve bahsetmeler “kullanıcı\_adı” ile değiştirilerek anonimleştirilmiştir.

Tablo 8

Sağlık veri setinden rastgele seçilen üç tweetin metin normalleştirmeden önceki ve sonraki hali

Ham metin	Normalleştirilmiş metin
1385515499413557248 2021-04-23 11:46:51 +0300 <kullanıcı_adı> @kullanıcı_adı As a nurse I wonder why you think its ok to promote a vaccine that is causing death and adverse effects in people. Are you aware of the young woman who are having miscarriages due to the vaccine? Check the data!!! Your promoting this to them 🤔🤔 <a href="https://t.co/mPW1GIMk8M">https://t.co/mPW1GIMk8M</a>	as a nurse i wonder why you think it is okay to promote a vaccine that is causing death and adverse effects in people are you aware of the young woman who are having miscarriages due to the vaccine check the data your promoting this to them face screaming in fear loudly crying face
1385260309993168896 2021-04-22 18:52:49 +0300 <kullanıcı_adı> @kullanıcı_adı @kullanıcı_adı @kullanıcı_adı You started with “no vaccine was ever recommended in pregnancy” b/c they “cause miscarriage”... You've modified your stance as your errors are pointed out... Please do research (using better info sources!) before making sweeping (false) statements about the “rest of the world”.	you started with no vaccine was ever recommended in pregnancy because they cause miscarriage you have modified your stance as your errors are pointed out please do research using better info sources before making sweeping false statements about the rest of the world
1384294248078475265 2021-04-20 02:54:02 +0300 <kullanıcı_adı> @kullanıcı_adı This is 😞 because now whats gonna happen? Lots of ppl r vaccinated now. Elderly essential workers chronically ill folks. and nobody can tell whos who. Schools are starting to open up and so is travel but vaccines are required i think 🤔 how r those ppl gonna pin miscarriage on them	this is woozy face because now whats going to happen lots of people are vaccinated now elderly essential workers chronically ill folks and nobody can tell who is who schools are starting to open up and so is travel but vaccines are required i think loudly crying face how are those people going to pin miscarriage on them

Metnin belirtilen arama desenini içerip içermediğinin kontrol edilmesi ve yerine istenen karakterle değiştirilmesi işlemi Python'ın re modülü ile gerçekleştirilmiştir (“Re — regular”, 2022). Bu modül, bir karakter dizisinin belirli bir düzenli ifadeyle eşleşip eşleşmediğini kontrol etmeyi sağlamaktadır. Düzenli ifade, kendisiyle eşleşen bir dizi karakteri belirtmektedir.

Düzenli ifadeler (Regular expressions) Python'ın içine yerleştirilmiş ve re modülü aracılığıyla kullanıma sunulan küçük, oldukça uzmanlaşmış bir programlama dilidir. Bu dili kullanarak, eşleştirmek istenilen olası kelime dizgeleri için kurallar belirlenebilmektedir. Bu kurallarla karakter dizisinin istenilen kalıpla eşleşip eşleşmediği veya istenilen kalıp için karakter dizisinde bir eşleşme olup olmadığı kontrol edilebilmektedir. Düzenli ifadeler bir dizgiyi değiştirmek veya onu çeşitli şekillerde bölmek için de kullanılabilmektedir (“Re — regular”, 2022).

### **3.3. Duygu Analizi (Sentiment Analysis)**

Verileri en doğru şekilde işleyebilmek için uygulanan metin normalleştirme süreci tamamlandıktan sonra metinlerin duygu kutupluluğunu tespit etmek için duygu analizi gerçekleştirilmiştir. Çalışmamızda metinlerin duygu kutupluluğunu sınıflandırmak için, sözlük ve kural tabanlı olan ve açık kaynak kodlu VADER duygu analizi aracı kullanılmıştır. Tweetler, kullanıcının konu hakkındaki bireysel yorumunun tamamını içerdiği için duygu analizi tweet bazlı yapılmıştır. Dolayısıyla her bir tweet, duygu analizi sonucunda pozitif, negatif veya nötr olarak etiketlenmiştir. Metin normalleştirme sürecinde alfanümerik karakterler hariç tüm karakterler silindiği ve buna noktamala işaretleri de dahil olduğu için birden fazla cümle içeren bir tweetin duygu analizi tek bir cümle şeklinde ele alınmıştır.

VADER, birden fazla bağımsız insan hakem tarafından etiketlenmiş ve bu etiketlerin ortalamasını kelimenin değerlik (valence) puanı olarak kabul eden bir duygu sözlüğü içerir. Cümle düzeyinde metinde algılanan duygu yoğunluğunu hesaplarken makalelerinde açıklanan dilbilgisi ve sözdizimsel kuralları uygulamaktadır. İfade edilen

duyguların hem kutupluluğuna hem de yoğunluğuna duyarlı olan VADER genel olarak tüm alanlarda duygu analizi için uygulanabileceği gibi özellikle sosyal medyada ifade edilen duygulara uyum sağlamaktadır (Hutto ve Gilbert, 2014).

Duygu analiz aracı, bir cümleye uygulandığında “compound” olarak tanımlanan bileşik puan, sözlükteki her kelimenin değerlik puanlarının toplanması, kurallara göre ayarlanması ve ardından -1 (en uç olumsuz) ile +1 (en uç olumlu) arasında olacak şekilde normalleştirilmesiyle hesaplanır. “Pos”, “neu” ve “neg” puanları ise, her kategoriye düşen metin oranlarını göstermektedir ve bu nedenle, bunların toplamı 1 olmaktadır (Hutto ve Gilbert, 2014). Metnin kutupluluk eğilimi ise standart eşik değerlerinin bileşik puana uygulanmasıyla elde edilmiştir. Bu bağlamda duygu sınıflandırmada kullanılan sözde kod Şekil 5’te verilmiştir.

```
IF bileşik puan >= 0.05
    duyguyu pozitif olarak sınıflandır
ELSE IF bileşik puan <= -0.05
    duyguyu negatif olarak sınıflandır
ELSE
    duyguyu nötr olarak sınıflandır
ENDIF
```

Şekil 5. Duygu sınıflandırmada kullanılan sözde kod

Yukarıda açıklandığı şekilde, “compound” ile birlikte “pos”, “neu” ve “neg” değerleri bir cümlenin kutupluluk yönelimini ortaya çıkarmada kullanılan değerlerdir. Kullanılan duygu analizi aracı VADER veri setine uygulandığında her bir yorum için tek tek bu değerleri hesaplamıştır. Şekil 6’da “CERN Opened a Portal to Another Dimension” başlıklı yalan haber yorumlarının on tanesi için duygu analizi sonuçları örnek olarak verilmiştir. Şekil 5 incelendiğinde, tüm puanları içeren değerlerin verildiği “scores” sütunu,

okunabilirliği kolaylaştırmak için ayrı bir sütunda tekrar eden bileşik puan değeri “compound” ve Şekil 5'te açıklanan algoritmanın bileşik puana uygulanmasıyla elde edilen duygu kutupluluğu “sentiment\_type” sütununda verildiği görülmektedir. Buna göre, tüm veri setleri aynı sürece tabi tutulmuş ve yorumların kutupluluk eğilimi elde edilmiştir.

	scores	compound	sentiment_type
0	{'neg': 0.192, 'neu': 0.731, 'pos': 0.077, 'compound': -0.4588}	-0.4588	Negatif
1	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}	0.0000	Nötr
2	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}	0.0000	Nötr
3	{'neg': 0.0, 'neu': 0.954, 'pos': 0.046, 'compound': 0.296}	0.2960	Pozitif
4	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}	0.0000	Nötr
5	{'neg': 0.0, 'neu': 0.855, 'pos': 0.145, 'compound': 0.5267}	0.5267	Pozitif
6	{'neg': 0.046, 'neu': 0.954, 'pos': 0.0, 'compound': -0.296}	-0.2960	Negatif
7	{'neg': 0.0, 'neu': 0.879, 'pos': 0.121, 'compound': 0.5719}	0.5719	Pozitif
8	{'neg': 0.0, 'neu': 0.876, 'pos': 0.124, 'compound': 0.6369}	0.6369	Pozitif
9	{'neg': 0.0, 'neu': 0.89, 'pos': 0.11, 'compound': 0.5719}	0.5719	Pozitif

Şekil 6. “CERN Opened a Portal to Another Dimension” veri setinin ilk on yorumu için duygu analizi sonuçları

### 3.4. Yapay Sinir Ağlarının Eğitimi

Çalışmada kullanılan veri setleri duygu analizine tabi tutulmuş ve veriler pozitif, negatif veya nötr olarak polariteleri belirlenmiştir. Sözlük ve Kural tabanlı duygu analizi aracının öğrenememe probleminin üstesinden gelmek ve metin sınıflandırma problemi için hangi modelin daha etkili olduğunu ortaya çıkarmak için çalışmadan elde edilen duygu analizinden geçirilmiş veri setleri iki farklı modeli eğitmede kullanılmıştır. Modellerin beslendiği veri setleri, normalleştirilmiş metinleri ve bu metinlerin duygu analizi aracıyla belirlenmiş pozitif, negatif ve nötr etiketlerinden oluşmaktadır.

Metin sınıflandırmada kullanılmak üzere eğitilen YSA modellerinden ilki Keras ön işleme katmanlarını kullanırken diğeri LSTM (Long Short Term Memory - Uzun Kısa Süreli Bellek) yöntemini kullanmaktadır. Python programlama dilinin sağladığı Numpy ve Pandas kütüphaneleri veri analizi ve veri işleme için kullanılırken verilerin görselleştirilmesi için Matplotlib kütüphanesinden yararlanılmıştır. Model oluşturma ve eğitme için ise TensorFlow ve Keras kütüphaneleri ile eğitilmiş makine öğrenimi modelleri için TensorFlow Hub deposu kullanılmıştır.

### 3.4.1 TensorFlow Girdi Veri Hattı (TF Input Pipeline)

Veri seti; eğitim, doğrulama ve test olmak üzere üç kümeye ayrılmıştır. Eğitim için veri setinin %80'i, doğrulama için %10'u ve test için %10'u alınacak şekilde bölünmüştür. Ardından tf.data API çatısını kullanarak bir girdi veri hattı (input pipeline) oluşturulmuştur.

tf.data API çatısı basit, yeniden kullanılabilir parçalardan karmaşık girdi veri hatları oluşturmaya olanak tanıyan bir API'dir. Bu API aynı zamanda büyük miktarda veriyi işlemeyi, farklı veri formatlarından okumayı ve karmaşık dönüşümler gerçekleştirmeyi mümkün kılar. Böylece, büyük miktarda verinin işlenmesine izin vererek, düşük kaliteli makinelerle hesaplama kolaylığı sağlar ("tf.data: Build TensorFlow", 2022).

Veri setinin tamamı tek seferde rastgele erişimli belleğe (RAM) yüklenip işlenemeyecek kadar büyük olduğu için verileri parçalara bölmek gerekmektedir. Bu nedenle veriler toplu işler (batch) halinde RAM'e yüklenir. Ancak bu verileri Numpy veya Pandas listesi olarak yüklemek yerine bir çeşit özel veri yapısı olan "tf.data.Dataset" kullanılır. Bununla girdi veri hattı oluşturulmaktadır. tf.data.Dataset ile veri seti bir dizi dönüşüm işleminden geçirilebilmektedir. Çalışmamızda her eğitim, doğrulama ve test kümesi bir tf.data.Dataset'e dönüştürülmüş ve bir fonksiyon yardımı ile veriler karıştırma (shuffle), gruplandırma (batch) ve önyükleme (prefetch) gibi dönüşüm işlemlerinden geçirilmiştir.



Verilerin karıştırılması “tf\_dataset.shuffle()” metoduyla gerçekleştirilmiştir. Bu metod, veri kümesinin öğelerini rastgele karıştırmaktadır. Bu veri kümesi bir arabelleği, belirlenen arabellek boyutu kadar öğeyle doldurur, ardından bu arabellekteki öğeleri rastgele örnekler, seçilen öğeleri yeni öğelerle değiştirir. Mükemmel bir karıştırma işlemi için, veri kümesinin tam boyutuna eşit veya daha büyük bir arabellek boyutu seçilmelidir (“tf.data: Build TensorFlow”, 2022). Değerler karıştırılarak veri setinin rastgele olması ve sabit bir sırada olmaması sağlanmıştır. Her iterasyonda arabellek boyutu kadar veri alınmaktadır. Öğeleri arabellek boyutu kadar yüklemekte ve her iterasyonda karıştırmaktadır. Karıştırırken dikkat edilmesi gereken en önemli husus, arabellek boyutunu ifade eden buffer\_size parametresinin yeterince büyük olduğundan emin olmaktır. Bu değer ne kadar büyükse, başlangıçta verileri yüklemek o kadar uzun sürecektir (“Building a data”, 2022).

Gruplandırma (batching) işlemi, veri kümesinin ardışık öğelerini toplu işlerde birleştirmesini kapsamaktadır. Çalışmamızda kullandığımız veri kümesi büyük olduğu için toplulaştırma boyutu da (batch size) büyük seçilmiştir. Veri seti, her biri üç elemanlı iki boyutlu gruplara bölünmüştür. Bu, büyük veri kümelerinin verimli bir şekilde işlenmesine olanak tanır ve çok fazla kaynağa sahip olmayan makinelerde heaplamayı kolaylaştırmaktadır (“Building a data”, 2022).

TensorFlow girdi veri hattının son basamağı olarak önyükleme (prefetch) işlemi yapılmıştır. Önyükleme, bir eğitim adımının önışlemesi ile modelin yürütülmesi işlemlerinin aynı anda yapılmasıdır. Bir diğer deyişle, modelin “s” eğitim adımını yürütürken, girdi veri hattının “s+1” adımı için gerekli verileri okumasıdır. Bu sayede, her zaman bir veri grubu önceden getirilir ve her zaman hazır bir veri olduğundan emin olunur (“Building a data”, 2022). Önceden getirilecek öğelerin sayısı, tek bir eğitim adımı tarafından tüketilen toplu işlerin sayısına eşit veya daha fazla olmalıdır. Bu değer manuel olarak ayarlayabileceği gibi tf.data.AUTOTUNE olarak da ayarlayabilmektedir. Böylece, çalışma zamanında değeri dinamik olarak ayarlamak için “tf.data” çalışma zamanına soracaktır.

TensorFlow girdi veri hattı son derece büyük boyutlu verilerle başa çıkabilmeyi sağladığı gibi modeli eğitmeden önce gerekli olabilecek tüm dönüşüm işlemlerini uygulayabilme olanağı sunar. Veri setlerine bu üç farklı dönüşüm uygulandıktan sonra “tf\_dataset” elde edilmiş ve model beslenirken bu veriler kullanılmıştır.

### 3.4.2 YSA Modelinin Oluşturulması

Girdi verilerimiz cümlelerden oluşmaktadır. Modeli oluşturmadan önce metin türündeki bu verileri modelin anlayacağı bir formata dönüştürmek gerekmektedir. Metinleri sayısal olarak temsil etmenin bir yolu, cümleleri gömme (embedding) vektörlerine çevirmektir. Bunun için, metinlere kelime gömme (word embedding) işlemi uygulanır. Metin gömme, insan tarafından okunabilir metinleri, model tarafından okunabilir sayılara veya vektörlere dönüştürme işlemidir. Böylece bir kelimenin vektör uzayındaki yeri metinden öğrenilebilir. Kelimenin öğrenildiği vektör uzayındaki pozisyona embedding denir. Vektör uzayında daha yakın olan kelimelerin birbirleriyle benzer anlamlara sahip olurlar (Patel, 2021).

Karmaşıklıklarına ve yeteneklerine göre farklılık gösteren kelime gömme teknikleri bulunmaktadır. Çalışmamızda TensorFlow Hub'ın 7 Milyarlık bir korpus olan İngilizce Google News'de eğitilmiş token tabanlı “nnlm-en-dim50” kelime gömmesi kullanılmıştır.

Tensorflow Hub, eğitilmiş makine öğrenimi modellerinin bir deposudur. Temelde bunlar bir kaç parametre ayarına ihtiyaç duyan kullanılmaya hazır modellerdir. TensorFlow Hub'da önceden eğitilmiş birçok metin gömmeleri vardır (“TensorFlow Hub”, 2022). Bu metin gömmeleri, ileri beslemeli Neural Network Language Model (NNLM)'e dayanmaktadır (“nnlm-en-dim50”, 2022).

Metin sınıflandırması için modeli oluştururken eklenecek olan ilk katman, önceden eğitilmiş embedding katmanı olmalıdır. Dolayısıyla önceden eğitilmiş bu “nnlm-en-dim50”

gömmeleri “embedding” değişkenine atanmıştır. Ardından TensorFlow Hub modeli kullanarak bir Keras katmanı oluşturulmuştur.

Doğrusal bir katman yığınının model içinde gruplaması için “Sequential” tekniği kullanılmıştır. Her katmanı “add” yöntemiyle eklenmiş olup ilk olarak “hub\_layer” gömme katmanı eklenmiştir. Bu katman ile cümleler tokenlere bölünür. Anlam olarak birbirine yakın olan tokenler bir araya getirilir. Ardından “Dense” katmanları eklenmiş ve bu katmanlarda relu ve sigmoid aktivasyon fonksiyonları kullanılmıştır.

Aşırı öğrenme (overfit) modelin eğitim verisini oldukça iyi tahmin edebilmesi ancak genelleştirme yetisinin son derece kötü bir olmasıdır. Aşırı öğrenen modeller genelleşememe problemlerine sahip olmakla birlikte yeni verileri iyi tahmin etmede zorluk yaşarlar. Bu nedenle modelin ezberlemesini önlemek için Dropout kamanları eklenmiştir. Dropout temelde iyi performans göstermeyen veya iyi çalışmayan bir kaç düğümü seçip çıkarmaktır. Bu da modelin genelleştirmesine yardımcı olmaktadır.

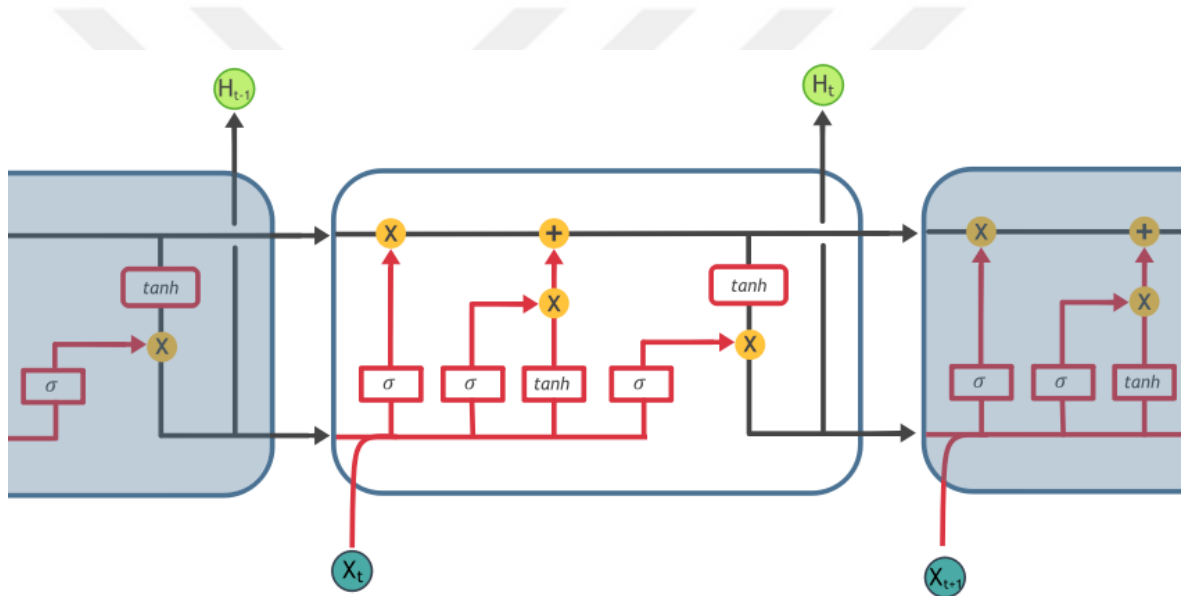
Modeli derlerken, kayıp fonksiyonunu ve bir optimize ediciyi belirtilmiştir. Problemimiz bir sınıflandırma problemi olduğu için kayıp fonksiyonu olarak “Binary Cross Entropy”, optimize edici olarak “Adam” kullanılmıştır. Öğrenme oranı, eğitim tur sayısı ve modelin performansını her eğitim tur sayısında görmek için “accuracy” metriği belirtildikten sonra model eğitilmiştir.

### **3.4.3 LSTM Yöntemi**

Eğitilen bir diğer modelde de LSTM (Long Short Term Memory - Uzun Kısa Süreli Bellek) sinir ağı kullanılmıştır. LSTM uzun süreli bağımlılıkları öğrenebilen ve kısa süreli hafıza problemini çözmeye kullanılan özel bir RNN (Recurrent Neural Network - Yilenen Sinir Ağları) türüdür (Olah, 2015). Tüm RNN’ler, tekrarlayan sinir ağı modülleri zinciri şeklindedir. Standart RNN’lerde, bu yinelenen modül, tek bir tanh katmanı gibi çok basit bir yapıya sahiptir. LSTM’ler zincir benzeri bir yapıya sahiptir,

ancak yinelenen modül bundan farklı bir yapıdadır. Tek bir sinir ağı katmanına sahip olmak yerine, çok özel bir şekilde etkileşime giren dört katmanı vardır.

Bir LSTM'deki yinelenen modül ve dört etkileşimli katman yapısı Şekil 7'de verilmiştir. Diyagram, her hattın, bir düğümün çıktısından diğerlerinin girdilerine kadar bütün bir vektörü taşıdığını göstermektedir. Yeşil daireler, vektör toplama gibi noktasal işlemleri temsil ederken, kutular öğrenilmiş sinir ağı katmanlarıdır. Birleşen hatlar birbirine bağlanmayı belirtirken, hattın çatallaması içeriğinin kopyalandığını ve kopyaların farklı konumlara gittiğini göstermektedir (Olah, 2015).



Şekil 7. Dört etkileşimli katman içeren LSTM'deki yinelenen modül (Olah, 2015)

Diyagramın üstünden geçen yatay çizgi olan hücre durumu (cell state) LSTM'lerin anahtarıdır. Hücre durumu bir tür taşıma bandı gibidir. Sadece bazı küçük doğrusal etkileşimlerle, tüm zincir boyunca dosdoğru çalışır. Bu sayede bilginin değişmeden akması çok kolaydır. LSTM, kapılar (gates) adı verilen yapılar tarafından dikkatlice düzenlenen hücre durumuna bilgi ekleme veya çıkarma yeteneğine sahiptir. Kapılar, isteğe bağlı olarak bilginin geçmesine izin vermenin bir yoludur. Sigmoid sinir ağı katmanından ve noktasal çarpma işleminden oluşurlar (Olah, 2015).

Bir RNN çeşidi olan LSTM ile, kısa süreli belleğin yanında uzun süreli bellek de gelmektedir. Böylece anlamlı, önemli, anahtar sözcük tarzındaki kelimeleri uzun süreli bellekte saklanabilmektedir. LSTM eğitimi esnasında veri setindeki bir çok örnekle karşılaştığı için yeni bir kelime geldiğinde bunu saklayıp saklamayacağını karar verebilir hale gelmektedir.

Çalışmamızda LSTM ile kurulan model için hatırlanacak maksimum sayıdaki kelime 2000 olarak belirlenmiştir. Sözlüğü vektörize etmek için metin özelliklerini tamsayı dizileriyle eşleyen bir Keras ön işleme katmanı kullanılmıştır. Farklı uzunluktaki metinleri standardize etmek ve aynı uzunluğa getirmek için maskeleyme parametresi ve LSTM'in çıktı boyutu metin gömme (text embedding) esnasında belirtilmiştir. Modelin doğruluk ve hata sonuçları dördüncü bölümde verilmiş, bulgular beşinci bölümde tartışılmıştır.

## **DÖRDÜNCÜ BÖLÜM**

### **ARAŞTIRMA BULGULARI**

Çalışmamızda, Tablo 3’de verilen yalan haber başlıklarına ait tepkiler toplanmış ve bu yorumlara duygu analizi uygulanmıştır. Sağlık, Toplum ve sosyal meseleler, Sanat ve eğlence ile Siyaset kategorileri ayrı ayrı incelenmiş olup her yalan habere verilen tepkiler bireysel olarak duygu analizine tabi tutulmuştur. Yalan habere verilen yorum sayısına bağlı olarak veri setinin pozitif, negatif ve nötr yorum olaraları aşağıda verilmiştir. Elde edilen verilerle Yapay Sinir Ağı ve LSTM modelleri eğitilmiş ve metin sınıflandırma problemi için hangi modelin daha iyi sonuç verdiği karşılaştırılarak değerlendirilmiştir.

#### **4.1. Duygu Analizi Sonuçları**

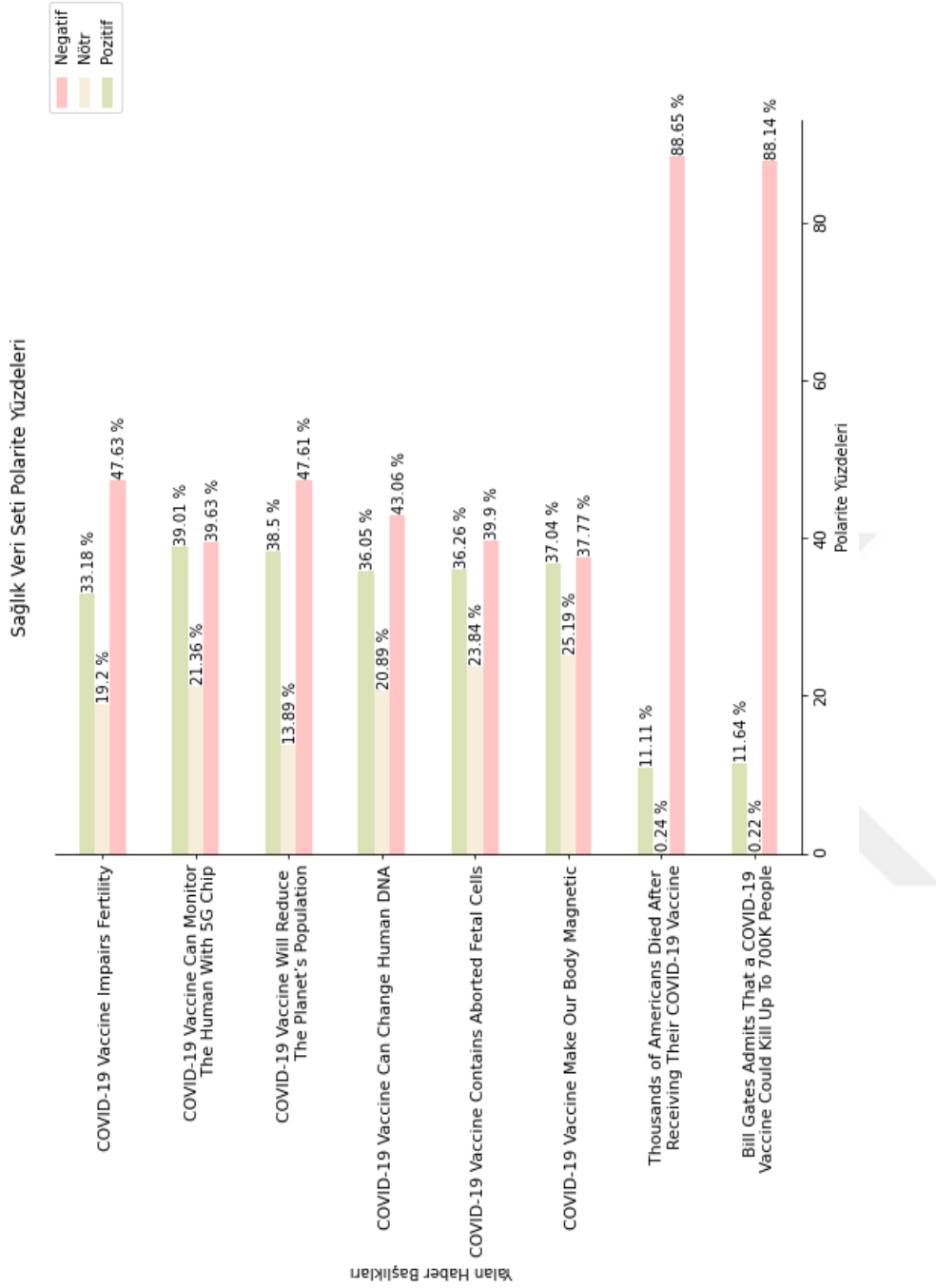
Çalışmada kullanılan yalan haber başlıkları, kategorisi ve her bir veri setine ait toplanan yorum sayısı Tablo 9’da verilmiştir. Yalan haberin yaygınlığı yalan habere yapılan yoruma göre, dolayısıyla veri setindeki veri sayısına göre karar verilmiştir.

Kategorilerine göre veri setlerine uygulanan duygu analizi sonuçları ise Şekil 8, Şekil 9, Şekil 10 ve Şekil 11’deki gibidir. Şekillerde sırasıyla Sağlık, Toplum ve sosyal meseleler, Sanat ve eğlence ile Siyaset kategorilerine ait polarite yüzdeleri yer almaktadır. Dikey ekseninde yalan haber başlıkları bulunurken, grafik aşağıdan yukarıya doğru veri sayısı artan şekilde yalan haber başlıkları yerleştirilmiştir. Dolayısıyla bu gösterim, aşağıdan yukarıya doğru yalan haberin yaygınlığının arttığını ifade etmektedir. Her bir haber başlığının karşısında kendi veri setine ait pozitif, negatif ve nötr polarite yüzdeleri verilmiştir. Şekillerde de görüldüğü gibi, bir yalan haber başlığının karşısındaki bu polarite yüzdeleri toplamı 100 olmaktadır.

Tablo 9

Kategorilerine göre yalan haber başlıkları ve bu yalan haberlere verilen yorumların sayısı

<b>Kategorisi</b>	<b>Yalan Haber Başlıkları</b>	<b>Veri Sayısı</b>
Sağlık	COVID-19 Vaccine Impairs Fertility	62722
	COVID-19 Vaccine Can Monitor The Human With 5G Chip	56052
	COVID-19 Vaccine Will Reduce The Planet's Population	30792
	COVID-19 Vaccine Can Change Human DNA	21073
	COVID-19 Vaccine Contains Aborted Fetal Cells	13164
	COVID-19 Vaccine Make Our Body Magnetic	11206
	Thousands of Americans Died After Receiving Their COVID-19 Vaccine	5435
	Bill Gates Admits That a COVID-19 Vaccine Could Kill Up To 700K People	3145
Toplum ve sosyal meseleler	Starbucks is Going Cashless in UK, US and Canada	51558
	Four AI Robots Kill 29 Scientists in Japan	50761
	Coca-Cola Add Try To Be Less White to Cans	26942
	New Facebook-Meta Rule Permit Company To Use Your Photos	26276
	CERN Opened a Portal to Another Dimension	12911
	George Floyd Appeared on "Judge Judy"	987
	Tickets to Trumps 2nd Inauguration Circulate on Social Media	546
	Popeyes Hang a Sign Saying They Will Refuse to Serve White People	487
Sanat ve eğlence	Was Johnny Depp Offered \$301M Disney Deal for Pirates Return	37451
	The Simpsons Predicted the Future	34309
	Disney World is removing Cinderella Castle	21464
	Actor Rowan Atkinson passed away	20957
	Disney Deny Tim Allen Lightyear Role Due to His Political Beliefs	18323
	Emma Watson Retiring From Acting	5020
	The Storyline of "I Am Legend" Takes Place in 2021	2003
	Captain America Predicted the Coronavirus Pandemic	542
Siyaset	Maricopa County Elections Database Deleted	38862
	Biden Say "White Republican Men Are More Dangerous Than ISIS"	33372
	Former U.S. President Donald Trump Could Be Reinstated as President by August 2021	12988
	Joe Biden is Wanted in Ukraine on Class A Felony Charges	5068
	The US Navy Has a Flying Aircraft Carrier	2516
	Jill Biden Gave Speech in Front of a Nazi-Inspired Flag	620



Şekil 8. Sağlık veri seti için duygu analizi sonuçları

Şekil 8'deki Sağlık veri seti için duygu analizi sonuçları incelendiğinde, yalan haberin yaygınlığı arttıkça veri setindeki negatif yorumların arttığı, pozitif ve nötr yorum oranlarının ise genellikle aynı oranlarda seyrettiği görülmektedir. Bir diğer ilginç bulgu ise Sağlık veri setinin tüm başlıkları için, negatif yorum oranının, pozitif ve nötr yorum



oranlarına göre daha yüksek oranda bulunmasıdır. Dolayısıyla, sağlık kategorisi için, negatif yorumlar arttıkça haberin yayıldığı, ancak pozitif ve nötr yorumların yaygınlığa bir etkisi olmadığı gözlemlenmiştir.



Şekil 9. Toplum ve sosyal meseleler veri seti için duygu analizi sonuçları

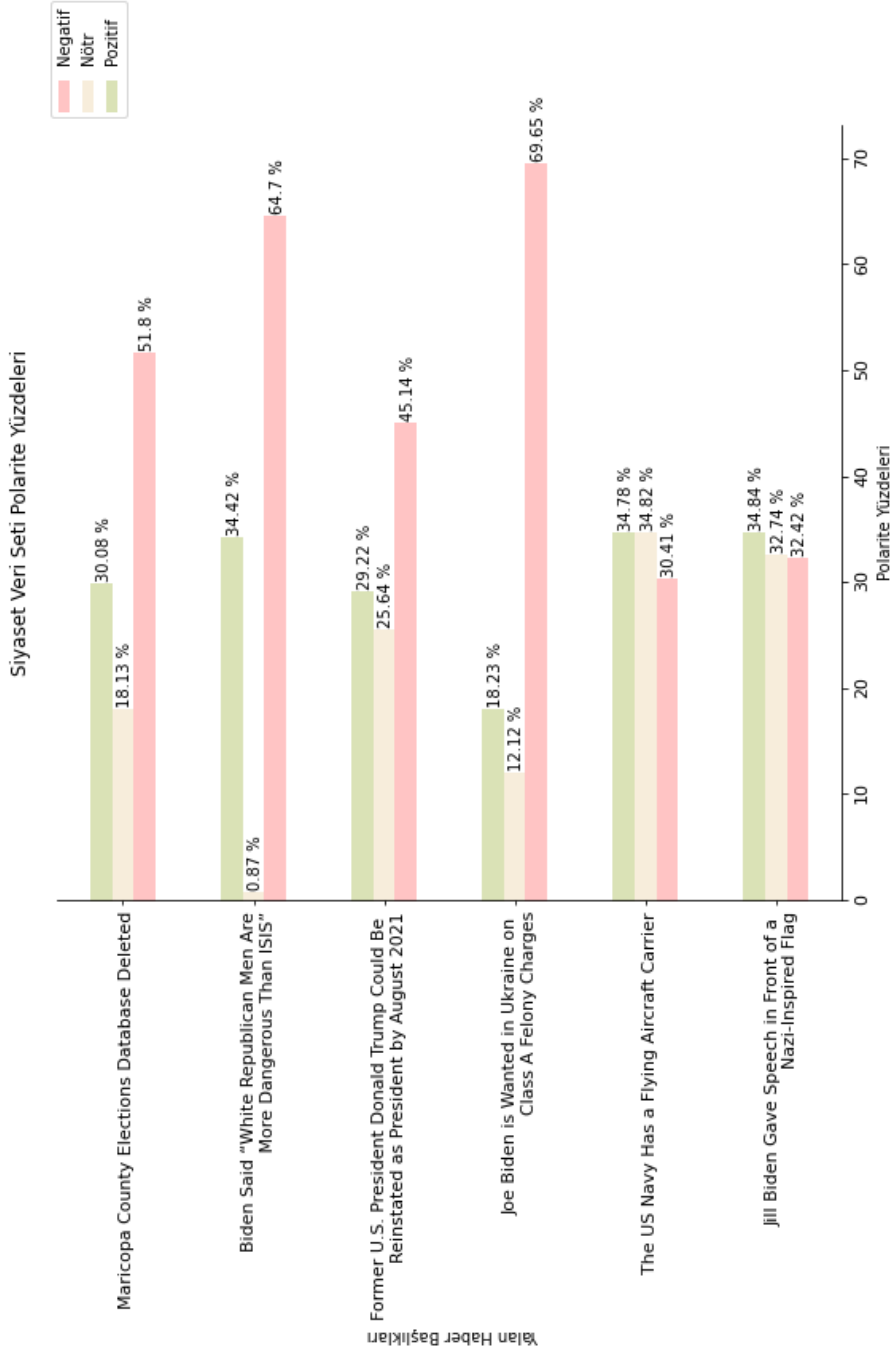
Şekil 9'daki Toplum ve sosyal meseleler veri seti için duygu analizi sonuçları incelendiğinde ise herhangi bir duygu kutupluluğunun belirli bir şekilde öne çıkmadığı gözlemlenmiştir. Yalan haberin içerdiği konuya göre belli duygular ön plana çıkmıştır. Örneğin “Coca-Cola Add Try To Be Less White to Cans”, “George Floyd Appeared on ‘Judge Judy’” veya “Popeyes Hang a Sign Saying They Will Refuse to Serve White People” gibi ırkçılık içeren veri setlerinde negatif duygunun baskın geldiği, ancak “Starbucks is Going Cashless in UK, US and Canada” ve “New Facebook-Meta Rule Permit Company To Use Your Photos” gibi yanlış bilgi içeren veri setlerinde nötr yorumların öne çıktığı görülmektedir. Toplum ve sosyal meseleler üst başlığı, veri setinde görülebileceği üzere, kendi içinde çok çeşitli konuları ihtiva edebilme potansiyeline sahiptir. Bu da her yalan haber için farklı duyguların öne çıkması anlamına gelmektedir. Dolayısıyla bu kategori için, sadece duygu kutupluluğu yalan haberin yaygınlığını açıklamada yetersiz kaldığı görülmektedir.

Şekil 10'daki Sanat ve eğlence veri seti için duygu analizi sonuçları incelendiğinde, nötr yorumlar yalan haberin yaygınlığı ile doğru orantılı bir şekilde artış sergilediği ve veri setindeki başlıkların çoğunda nötr yorumların pozitif ve negatif yorumların oranından daha yüksek olduğu görülmektedir. Haberin yaygınlığı arttıkça negatif yorum oranında az da olsa bir artış trendi saptanmaktadır. Ancak pozitif yorum oranındaki değişimin yaygınlığa herhangi bir etkisi görülmemektedir.

Şekil 11'deki Siyaset veri seti için duygu analizi sonuçları incelendiğinde, tıpkı “Toplum ve Sosyal Meseleler” veri setindeki gibi belirli bir şekilde öne çıkan bir duygu kutupluluğuna rastlanmamıştır. Ancak, haberin yaygınlığı arttıkça negatif yorum oranı azalış göstermiştir. Bununla birlikte pozitif yorum oranının veri seti boyunca yaklaşık aynı oranlarda seyrettiği, nötr yorum oranının ise düzenli bir davranış sergilemediği gözlemlenmiş, dolayısıyla bu iki duygu kutupluluğunun yaygınlığa bir etkisi saptanmamıştır.



Şekil 10. Sanat ve eğlence veri seti için duygu analizi sonuçları

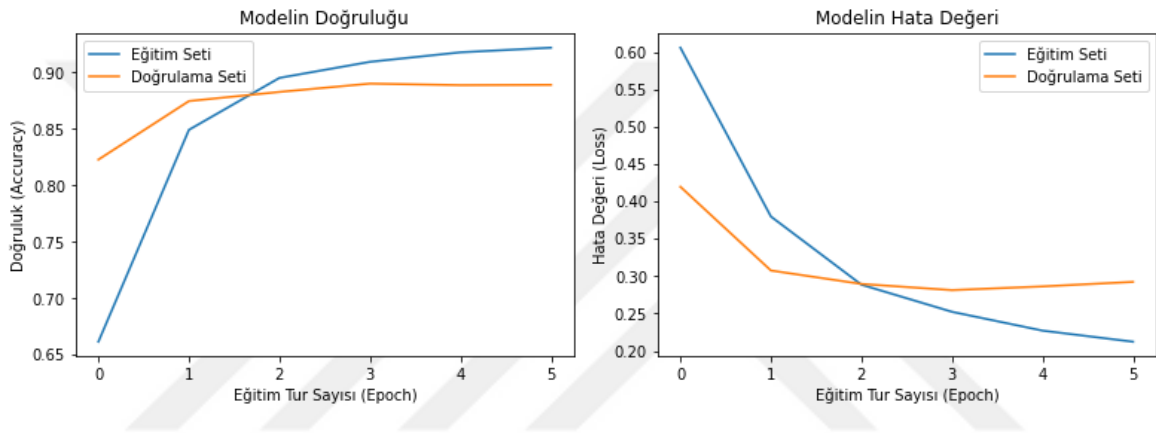


Şekil 11. Siyaset veri seti için duygu analizi sonuçları

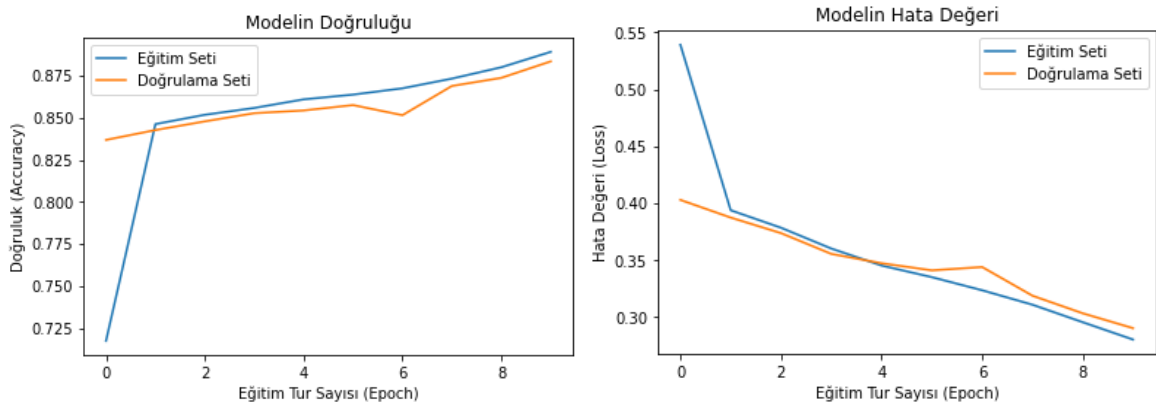
## 4.2. Modellerin Değerlendirilmesi

Bu çalışmadan elde edilen veri setleri kullanılarak Yapay Sinir Ağı ve LSTM modelleri oluşturulmuş ve bu yöntemlerle metin sınıflandırma yapılmıştır. Modeller ikili sınıflandırma yapacak şekilde ayarlanmıştır. İki model de eğitim için fazla veriye ihtiyaç

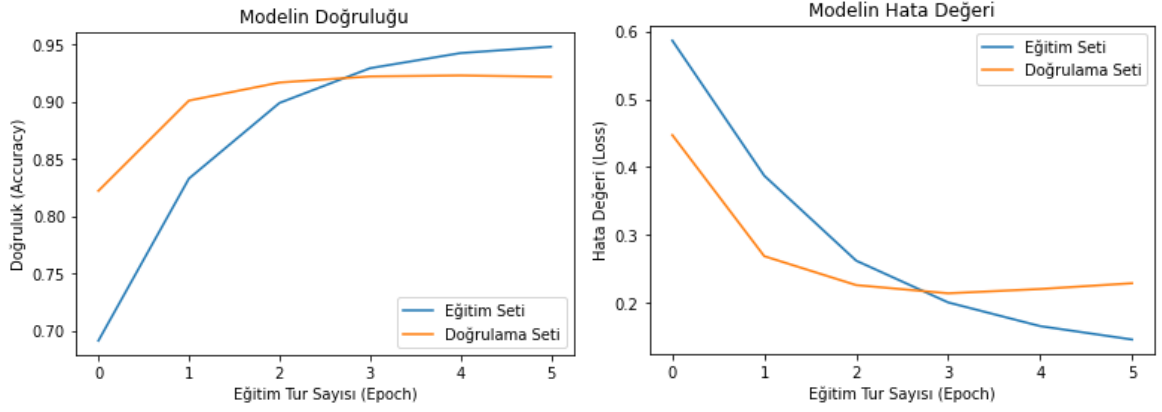
duyduğu için her bir kategorilerine ait veri setleri birleştirilerek modeller eğitilmiştir. Çalışmamızdaki metin sınıflandırma problemi için, tüm sınıflar eşit derecede önemli ve sınıflar dengeli olduğundan modelleri değerlendirmek için yalnızca doğruluk (accuracy) metriği kullanılmıştır. Bununla beraber model performansı hata (loss) fonksiyonu ile değerlendirilmiştir. Yalan haber kategorileri için Yapay Sinir Ağı modelinin eğitim tur sayısına bağlı olarak doğruluk ve hata değerleri Şekil 12, Şekil 14, Şekil 16 ve Şekil 18’de verilmiştir. LSTM modelinin eğitim tur sayısına bağlı olarak doğruluk ve hata değerleri ise Şekil 13, Şekil 15, Şekil 17 ve Şekil 19’da verilmiştir.



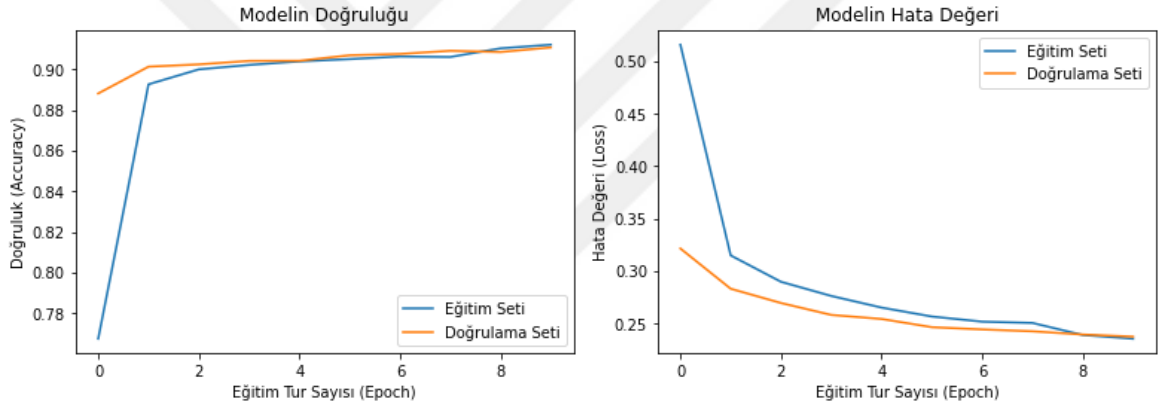
Şekil 12. “Sağlık” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



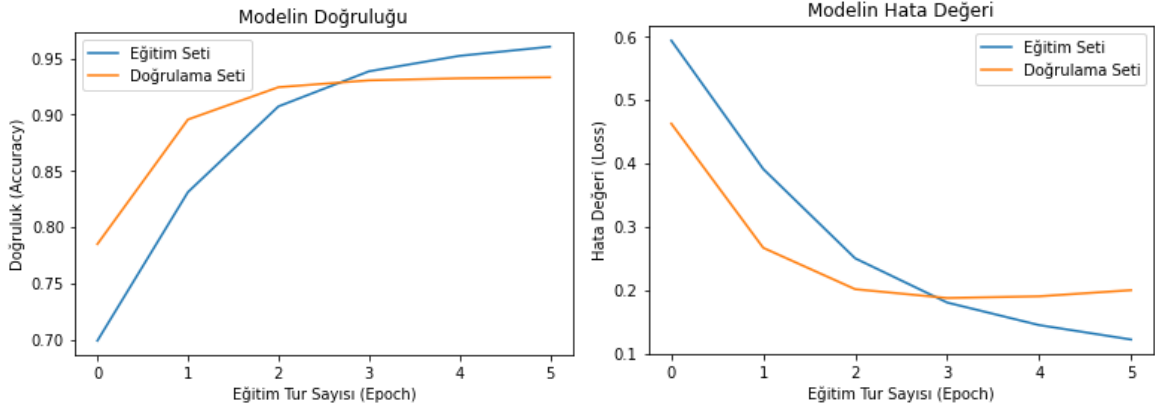
Şekil 13. “Sağlık” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



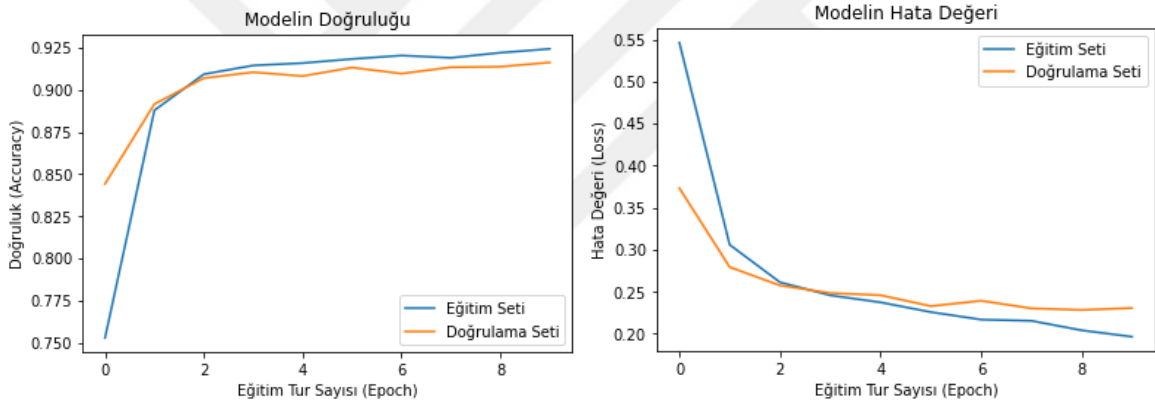
Şekil 14. “Toplum ve Sosyal Meseleler” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



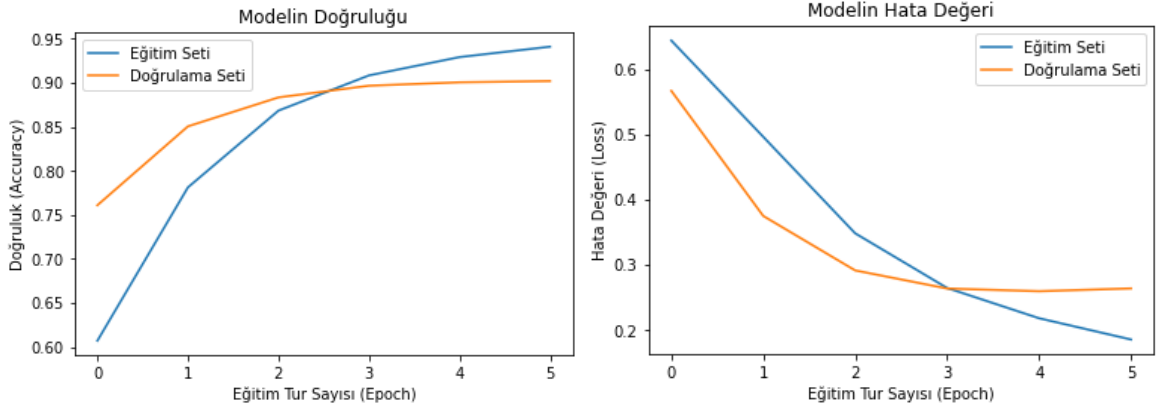
Şekil 15. “Toplum ve Sosyal Meseleler” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



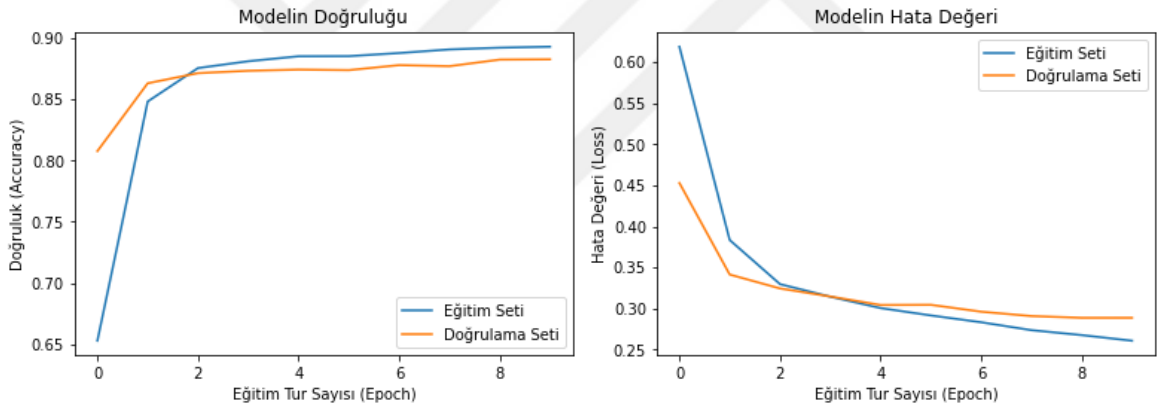
Şekil 16. “Sanat ve Eğlence” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



Şekil 17. “Sanat ve Eğlence” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



Şekil 18. “Siyaset” veri setinin, Yapay Sinir Ağı modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri



Şekil 19. “Siyaset” veri setinin, LSTM modeline göre eğitim seti ve doğrulama seti için Doğruluk (Accuracy) ve Hata (Loss) değeri

YSA ile kurulan modelin başarısını LSTM ile karşılaştırmak için iki modelin de eğitim, doğrulama ve test verilerinin doğruluk ve hata sonuçları Tablo 10'da verilmiştir. Hangi modelin metin sınıflandırmada daha başarılı olduğu doğrulama verilerindeki ve test verilerindeki doğruluk metriğinin yüksek olmasıyla ve hata metriğinin de 0'a olan yakınlığıyla anlaşılmaktadır.



Tablo 10

YSA ve LSTM modelleri için eğitim, doğrulama ve test verilerinin doğruluk ve hata sonuçları

<b>Kategorisi</b>	<b>Model</b>	<b>Metrik</b>	<b>Eğitim verileri</b>	<b>Doğrulama verileri</b>	<b>Test verileri</b>
Sağlık	YSA	Doğruluk	0.9393	0.8890	0.8924
		Hata	0.1615	0.2923	0.2925
	LSTM	Doğruluk	0.8981	0.8834	0.8853
		Hata	0.2576	0.2904	0.2931
Toplum ve sosyal meseleler	YSA	Doğruluk	0.9637	0.9218	0.9166
		Hata	0.0957	0.2295	0.2433
	LSTM	Doğruluk	0.9154	0.9107	0.9082
		Hata	0.2191	0.2376	0.2475
Sanat ve eğlence	YSA	Doğruluk	0.9751	0.9332	0.9318
		Hata	0.0771	0.1999	0.1994
	LSTM	Doğruluk	0.9288	0.9164	0.9151
		Hata	0.1814	0.2307	0.2333
Siyaset	YSA	Doğruluk	0.9612	0.9020	0.8998
		Hata	0.1252	0.2639	0.2888
	LSTM	Doğruluk	0.8999	0.8824	0.8855
		Hata	0.2403	0.2887	0.2924

Tablo 10'daki sonuçlar incelendiğinde, tüm kategorilerde, eğitimli TensorFlow Hub kullanan YSA modelinin LSTM modeline göre daha yüksek doğruluk elde ettiği görülmektedir. YSA modeli, En az verisi olan Siyaset kategorisi için bile doğrulama veri setinde 0.9020, test veri setinde 0.8998 doğruluk elde etmiştir. Bu da eğitimli TensorFlow Hub kullanan YSA modelinin az veriyle de oldukça iyi performans gösterdiğini ortaya koymaktadır.

Şekil 12, Şekil 14, Şekil 16 ve Şekil 18'deki Modelin Doğruluğu grafikleri incelendiğinde, eğitimi TensorFlow Hub kullanan YSA modelinin çok hızlı bir şekilde eğitiminin tamamlandığını ve altı eğitim tur sayısında (epoch) optimum performansına ulaştığını söyleyebilmekteyiz.

Yine aynı şekilde Tablo 10'daki sonuçlar LSTM için incelendiğinde, LSTM kullanan modelin doğruluk değerlerinin, diğer modelin altında kaldığı görülmektedir. Öte yandan eğitim, doğrulama ve test verileri için doğruluk değerleri oldukça tutarlıdır. Örneğin Sağlık kategorisi için, eğitim verilerinin doğruluk değeri 0.8981 iken doğrulama verilerinin doğruluk değeri 0.8834 ve test verilerinin doğruluk değeri 0.8853'tür. Ancak YSA modelindeki eğitim verilerinin doğruluk değeri 0.9393 iken doğrulama ve test verilerinde doğruluk değeri sırasıyla 0.8890 ve 0.8924 olarak bulunmuştur. Bu da LSTM'in daha iyi genelleştirdiğini, YSA'nın ise ezberlemeye (overfit) yatkın olduğunu göstermektedir.

Şekil 13, Şekil 15, Şekil 17 ve Şekil 19'deki Modelin Doğruluğu grafikleri incelendiğinde, LSTM kullanan modelin eğitimi daha yavaş tamamlanmıştır. Dokuz eğitim tur sayısında (epoch) optimum performansına ulaşmıştır. Bu da LSTM modelinin eğitiminin YSA modeline göre daha uzun sürdüğünü göstermektedir.

## BEŞİNCİ BÖLÜM

### SONUÇ VE ÖNERİLER

Sosyal medyanın haber kaynağı olarak kullanılması durumu günümüzde oldukça yaygındır. Ancak sosyal medyada karşılaşılan haberlerin doğruluğu ve geçerliliği sorgulanmadan ve kontrol edilmeden doğru kabul edilmesi, yalan haberlerin yayılmasının ana nedenlerinden biridir. Sosyal medya ekosistemi için, haberlerin doğruluğu ve geçerliliği kanıtlara dayalı verilerden ziyade fikir ve duyguların kontrolü altında olduğundan, yalan haberlerin yayılmasında hangi duygunun daha etkili olduğu sorusu ortaya çıkmaktadır. Bu noktadan hareketle çalışmamızda, kullanıcıların yalan haberlere verdikleri tepkiler ile yalan haberlerin yaygınlığı arasında bir ilişki olup olmadığı araştırılmıştır.

Bu çalışmada, yalan haberin yaygınlığıyla kullanıcı yorumlarının duygu kutuplulukları arasında bir ilişki olup olmadığını ortaya çıkartmak amaçlanmıştır. Bunun için yalan haberler çevrimiçi bağımsız teyitçilerden tespit edilip bu yalan haberlere verilen kullanıcı yorumları Twitter'dan toplanmıştır. Yalan haberin yaygınlığı bu yalan habere yapılan yorum sayısına göre karar verilmiştir. Toplanan veriler dört ana kategoride analiz edilmiştir. Bunlar; “Sağlık”, “Toplum ve Sosyal Meseleler”, “Sanat ve Eğlence” ve “Siyaset” kategorileridir. Her kategorilerdeki yalan haber başlıklarına ait yorumlara duygu analizi uygulanmıştır. Çalışmadan elde edilen verilerle sözlük ve kural tabanlı duygu analizi aracının öğrenememe probleminin üstesinden gelmek ve metin sınıflandırma problemi için hangi modelin daha etkili olduğunu ortaya çıkarmak için iki yapay sinir ağı modeli eğitilmiştir. Eğitilen YSA modellerinden ilki Keras ön işleme katmanlarını kullanırken diğeri LSTM yöntemini kullanılmıştır.

Özellikle Sağlık kategorisinde yalan haberin yaygınlı arttıkça veri setindeki negatif yorumların arttığı gözlemlenirken nötr ve pozitif yorumların genellikle birbirine yakın oranlarda seyrettiği, dolayısıyla Sağlık kategorisi için nötr ve pozitif yorumun yaygınlığa bir katkısı olmadığı gözlemlenmiştir. Öte yandan Sanat ve Eğlence kategorisinde, nötr yorumlar yalan haberin yaygınlığıyla doğru orantılı bir şekilde artış sergilemektedir.

Bununla birlikte bu kategorinin veri setlerinde nötr yorumlar pozitif veya negatif yorum oranından daha yüksek oranda görülmektedir. İlk bakışta, sanat ve eğlencenin insandaki pozitif duyguları harekete geçirme konusunda önyak olduğu bilinse de, bu kapsamdaki yalan haberlerin yayılımında pozitif yorum oranındaki değişimin yaygınlığa herhangi bir etkisi görülmemiştir. Toplum ve sosyal meseleler kategorisinde ise yalan haberin yaygınlığıyla herhangi bir duygu kutupluluğu belirli bir şekilde öne çıkmamıştır. Yalan haberin başlığına, yani içerdiği konuya göre belli duygular önplana çıkmıştır. Örneğin toplumun sinir uçlarına dokunan ve ırkçılık içeren “Coca-Cola Add Try To Be Less White to Cans” ve “George Floyd Appeared on ‘Judge Judy’” veri setlerinde negatif duygunun baskın geldiği, ancak “Starbucks is Going Cashless in UK, US and Canada” ve “New Facebook-Meta Rule Permit Company To Use Your Photos” gibi yanlış bilgi içeren veri setlerinde nötr yorumlar öne çıkmıştır. Siyaset kategorisinde ise tıpkı “toplum ve sosyal meseleler” veri setindeki gibi belirli bir şekilde öne çıkan bir duygu kutupluluğuna rastlanmamıştır. Bununla birlikte, veri seti boyunca yaklaşık aynı oranlarda seyreden pozitif polaritenin ve düzenli bir davranış sergilemeyen nötr polaritenin yaygınlığa bir etkisi saptanmamıştır. Siyaset kategorisiyle ilgili bir diğer ilginç tespit ise, yalan haberin yaygınlığı arttıkça negatif yorum oranında azalma görülmesidir.

Bu sonuçlar ışığında, yalan haberin yaygınlığıyla yalan habere verilen tepki arasında bir ilişki olduğu ortaya çıkmıştır. Ancak yalnızca bir duygunun yalan haberi yaymada tek başına yeterli olmadığı, yalan haberin türüne ve insanları hangi bağlamda etkilediğine göre değiştiği gözlemlenmiştir.

Tüm veri setleri için Keras ön işleme katmanlarından eğitilmiş TensorFlow Hub kullanan YSA modelinin LSTM modelinden daha yüksek doğruluk değerine sahip olurken yine YSA modelinden daha düşük hata değeri elde edilmiştir. Eğitilmiş TensorFlow Hub kullanan modelin ise az veriyle bile doğruluk değerinin LSTM'den daha yüksek olduğu görülmektedir. Modellerin optimum seviyesine ulaşmaları için kullanılan eğitim tur sayıları da farklılık göstermiştir. Keras katmanlı YSA 6 eğitim tur sayısında optimuma ulaşırken bu durum LSTM için 9 eğitim tur sayısında gerçekleşmektedir. Bu da, bir metin sınıflandırma problemi söz konusu olduğunda, önceden eğitilmiş metin gömmelerinin kullanılmasının performansı arttırdığını ve eğitim süresini kısalttığını göstermektedir.

Çalışmanın ilave verilerle güncellenmeye açık ve hazır bir yapısı bulunmaktadır. Gelecekteki çalışmalarda, modelin daha iyi genelleyebilmesi ve daha yüksek doğruluk elde edilebilmesi için daha büyük veri setleri kullanılabilir. Yalan haber başlıklarına ait pozitif, negatif ve nötr yorum oranları, haberin veri sayısı gibi parametreler öznitelik olarak kullanılarak makine öğrenmesi yöntemleriyle bir haberin yayılıp yayılmayacağı tahminlenebilir. Bir diğer çalışma ise, yalan haber yorumlarındaki polarite oranlarının rastgele olup olmadığını açığa çıkarmaya yönelik olabilir. Örneğin, negatif yorumlar belli bir orandayken pozitif veya nötr yorumların belli bir değerin üstünde veya altında olması yalan haberi yayıp yaymayacağını, böyle bir gizli ilişki olup olmadığı araştırması yapılabilir. Böyle bir çalışma makine öğrenmesi yapabilecek kadar yalan haber başlıkları ve bu yalan haberlere ait yorumların toplanmasıyla mümkün olabilir.

## KAYNAKÇA

- Agarwal, A., Xie, B., Vovsha, I., Rambow, O. ve Passonneau, R. J. (2011, June). “Sentiment analysis of twitter data”. *In Proceedings of the workshop on language in social media (LSM 2011)*. 30-38.
- Agarwal, A. (2017, 19 Temmuz). How to Write a Twitter Bot in 5 Minutes [Blog yazısı]. Erişim adresi: <https://www.labnol.org/internet/write-twitter-bot/27902/>
- Aisopos, F., Papadakis, G. ve Varvarigou, T. (2011, November). “Sentiment analysis of social media content using n-gram graphs”. *In Proceedings of the 3rd ACM SIGMM international workshop on Social media*. 9-14.
- Allcott, H. ve Gentzkow, M. (2017). “Social Media and Fake News in the 2016 Election”. *Journal of Economic Perspectives*. 31(2). 211-236. doi:10.1257/jep.31.2.211
- Alonso, M. A., Vilares, D., Gómez-Rodríguez, C. ve Vilares, J. (2021). “Sentiment analysis for fake news detection”. *Electronics*. 10(11). 1348.
- Anoop, K., Deepak, P. ve Lajish, V. L. (2020). “Emotion cognizance improves health fake news identification”. *In IDEAS*. 24.
- Asiaee T, A., Tepper, M., Banerjee, A. ve Sapiro, G. (2012, October). “If you are happy and you know it... tweet”. *In Proceedings of the 21st ACM international conference on Information and knowledge management*. 1602-1606.
- Aston, N., Liddle, J. ve Hu, W. (2014). “Twitter sentiment in data streams with perceptron”. *Journal of Computer and Communications*. 2014..
- Bakliwal, A., Arora, P., Madhappan, S., Kapre, N., Singh, M. ve Varma, V. (2012, Temmuz). “Mining sentiments from tweets”. *In Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*. 11-18.
- Barai, M. K. (2021, 20 Ekim). Sentiment Analysis with TextBlob and Vader [Blog yazısı]. Erişim adresi: <https://www.analyticsvidhya.com/blog/2021/10/sentiment-analysis-with-textblob-and-vader/>

- Barbosa, L. ve Feng, J. (2010, Ağustos). “Robust sentiment detection on twitter from biased and noisy data”. *In Coling 2010: Posters*. 36-44.
- Bodaghi, A. ve Oliveira, J. (2022). “The theater of fake news spreading, who plays which role? A study on real graphs of spreading on Twitter”. *Expert Systems with Applications*, 189, 116110.
- Bora, N. (2012). “Summarizing public opinions in tweets”. *International Journal of Computational Linguistics and Applications*. 3(1). 41-55.
- Building a data pipeline. (2022). Erişim adresi: <https://cs230.stanford.edu/blog/datapipeline/#an-overview-of-tfdata>
- CITS. (2018a, 15 Eylül). “Why We Fall for Fake News”. Erişim adresi: <https://www.cits.ucsb.edu/fake-news/why-we-fall>
- CITS. (2018b, 15 Eylül). “How is Fake News Spread? Bots, People like You, Trolls, and Microtargeting”. Erişim adresi: <https://www.cits.ucsb.edu/fake-news/spread>
- Cui, L., Wang, S. ve Lee, D. (2019, Ağustos). “Same: sentiment-aware multi-modal embedding for detecting fake news”. *In Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*. 41-48.
- Davidov, D., Tsur, O. ve Rappoport, A. (2010, Ağustos). “Enhanced sentiment learning using twitter hashtags and smileys”. *In Coling 2010*. 241-249.
- de Lima Salge, C. A. ve Berente, N. (2017). “Is that social bot behaving unethically?”. *Communications of the ACM*. 60(9). 29-31.
- Dewing, M. (2010). “Social media: An introduction”. Rapor. Erişim: 3 Şubat 2010, [https://publications.gc.ca/collections/collection\\_2013/bdp-lop/eb/2010-3-eng.pdf](https://publications.gc.ca/collections/collection_2013/bdp-lop/eb/2010-3-eng.pdf).
- Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M. ve Xu, K. (2014, Haziran). “Adaptive recursive neural network for target-dependent twitter sentiment classification”. *In Proceedings of the 52nd annual meeting of the association for computational linguistics*. 49-54.

- Dzogang, F., Lightman, S. ve Cristianini, N. (2018). "Diurnal variations of psychometric indicators in Twitter content". *PloS one*, 13(6), e0197002.
- Ferrara, E., Varol, O., Davis, C., Menczer, F. ve Flammini, A. (2016). "The rise of social bots". *Communications of the ACM*. 59(7). 96-104.
- Gabrielkov, M., Ramachandran, A., Chaintreau, A. ve Legout, A. (2016, June). "Social clicks: What and who gets read on Twitter?". *In Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science*. 179-192.
- Giachanou, A. ve Crestani, F. (2016). "Like it or not: A survey of twitter sentiment analysis methods". *ACM Computing Surveys (CSUR)*. 49(2). 1-41.
- Go, A., Bhayani, R. ve Huang, L. (2009). "Twitter sentiment classification using distant supervision". *CS224N project report*. Stanford. 1(12), 2009.
- Gupta, N. ve Agrawal, R. (2020). "Application and techniques of opinion mining". S. Bhattacharyya, V. Snášel, D. Gupta, ve A. Khanna (Ed.). içinde *Hybrid Computational Intelligence*. (s. 1-23). Yayınevi: Academic Press.
- Hamdan, H., Béchet, F. ve Bellot, P. (2013, Haziran). "Experiments with DBpedia, WordNet and SentiWordNet as resources for sentiment analysis in micro-blogging". *In Second Joint Conference on Lexical and Computational Semantics (\*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*. 455-459.
- Hu, M. ve Liu, B. (2004, Ağustos). "Mining and summarizing customer reviews". *In Proceedings Of The Tenth ACM SIGKDD International Conference On Knowledge Discovery And Data Mining*. 168-177.
- Hutto, C. ve Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225)*.
- Islam, M. S., Kamal, A. H. M., Kabir, A., Southern, D. L., Khan, S. H., Hasan, S. M., ... ve Seale, H. (2021). "COVID-19 vaccine rumors and conspiracy theories: The need



for cognitive inoculation against misinformation to improve vaccine adherence”. *PloS one*. 16(5). e0251605.

Iwendi, C., Mohan, S., Ibeke, E., Ahmadian, A. ve Ciano, T. (2022). “Covid-19 fake news sentiment analysis”. *Computers and electrical engineering*. 107967.

Kannan S., Karuppusamy S., Nedunchezhian A., Venkateshan P., Wang P., Bojja N. ve Kejariwal A. (2016). “Big Data Analytics for Social Media”. R. Buyya, R. N. Calheiros ve A. V. Dastjerdi (ed). içinde *Big Data Principles and Paradigms* (s. 63-94). Yayınevi: Morgan Kaufmann.

Kim, Y. (2014). “Convolutional Neural Networks for Sentence Classification”. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha. Qatar. Association for Computational Linguistics. 1746-1751.

Klenner, M., Petrakis, S., ve Fahrni, A. (2009, Eylül). “Robust compositional polarity classification”. In *Proceedings of the International Conference RANLP-2009*. 180-184.

Kouloumpis, E., Wilson, T. ve Moore, J. (2011, Temmuz). “Twitter sentiment analysis: The good the bad and the omg!”. In *Proceedings of the International AAAI Conference on Web and Social Media*. 5(1).

Kumar, A. ve Sebastian, T. M. (2012). “Sentiment analysis on twitter”. *International Journal of Computer Science Issues (IJCSI)*. 9(4). 372.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... ve Zittrain, J. L. (2018). “The science of fake news”. *Science*, 359(6380), 1094-1096. doi: 10.1126/science.aao2998.

Liu, B. (2012). “Sentiment analysis and opinion mining”. *Synthesis lectures on human language Technologies*. 5(1). 1-167.

Liu, K. L., Li, W. J. ve Guo, M. (2012, Temmuz). “Emoticon smoothed language models for twitter sentiment analysis”. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 26(1).

- Medhat, W., Hassan, A. ve Korashy, H. (2014). "Sentiment analysis algorithms and applications: A survey". *Ain Shams Engineering Journal*. 5(4). 1093-1113.
- Mohammad, S. M., Kiritchenko, S. ve Zhu, X. (2013, 28 Ağustos). "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets". *In Proceedings of the seventh international workshop on Semantic Evaluation Exercises (SemEval-2013)*. doi: 10.48550/arXiv.1308.6242.
- Neethu, M. S. ve Rajasree, R. (2013, Temmuz). "Sentiment analysis in twitter using machine learning techniques". *In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. 1-5.
- Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T. ve Nielsen, R. K. (2021). *Reuters Institute Digital News Report 2021*. Erişim: 2021, [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital\\_News\\_Report\\_2021\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf).
- nnlm-en-dim50. (2022). TensorFlow Hub. Erişim adresi: <https://tfhub.dev/google/nnlm-en-dim50/2>
- OECD. (2007). *OECD Yıllık rapor*. Erişim adresi: 31 Mart 2007, <https://www.oecd.org/newsroom/38528123.pdf>.
- Olah, C. (2015, 27 Ağustos). Understanding LSTM Networks [Blog yazısı]. Erişim adresi: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Pak, A. ve Paroubek, P. (2010, May). "Twitter as a corpus for sentiment analysis and opinion mining". *In LREc*. 10. 1320-1326.
- Patel, D. (2021, 24 Mart). "Word embedding using keras embedding layer | Deep Learning Tutorial 40 (Tensorflow, Keras & Python)" [Video]. Erişim adresi: <https://www.youtube.com/watch?v=Fuw0wv3X-0o>
- Read, J. (2005, June). "Using emoticons to reduce dependency in machine learning techniques for sentiment classification". *In Proceedings of the ACL student research workshop*. 43-48.

- Rosen, A. (2017, 7 Kasım). Tweeting Made Easier. [Blog yazısı]. Erişim adresi: [https://blog.twitter.com/official/en\\_us/topics/product/2017/tweetingmadeeasier.html](https://blog.twitter.com/official/en_us/topics/product/2017/tweetingmadeeasier.html)
- Saif, H., He, Y. ve Alani, H. (2012, Kasım). “Semantic sentiment analysis of twitter”. In *International semantic web conference*. Springer, Berlin, Heidelberg. 508-524.
- Sarlan, A., Nadam, C., & Basri, S. (2014, Kasım). “Twitter sentiment analysis”. In *Proceedings of the 6th International conference on Information Technology and Multimedia*. IEEE. 212-216.
- Shearer, E. ve Mitchell, A. (2021). News use across social media platforms in 2020. Pew Research Center. Erişim adresi: <https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/>
- Sneffjella, B., Schmidtke, D. ve Kuperman, V. (2018). “National character stereotypes mirror language use: A study of Canadian and American tweets”. *PloS one*. 13(11). e0206188.
- Subramanian, S. (2017, 15 Şubat). Inside the Macedonian Fake-News Complex. Wired. Erişim adresi: <https://www.wired.com/2017/02/veles-macedonia-fake-news/>
- Tandoc, E. C., Lim, Z. W. ve Ling, R. (2018). “Defining “fake news” A typology of scholarly definitions”. *Digital journalism*. 6(2). 137-153.
- Tandoc, E. C. (2019). “The facts of fake news: A research review”. *Sociology Compass*. 13(9). e12724. doi:10.1111/soc4.12724.
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T. ve Qin, B. (2014, June). “Learning sentiment-specific word embedding for twitter sentiment classification”. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. 1555-1565.
- TensorFlow Hub. (2022). Erişim adresi: <https://www.tensorflow.org/hub>
- tf.data: Build TensorFlow input pipelines. (2022, 9 Haziran). Erişim adresi: <https://www.tensorflow.org/guide/data>

- Turney, P. D. (2002). "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews". *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Philadelphia, Pennsylvania. 417-424. doi: 10.48550/arXiv.cs/0212032.
- Tsytsarau, M. ve Palpanas, T. (2012). "Survey on mining subjective data on the web". *Data Mining and Knowledge Discovery*. 24(3). 478-514.
- Vicario, M. D., Quattrociochi, W., Scala, A. ve Zollo, F. (2019). "Polarization and fake news: Early warning of potential misinformation targets". *ACM Transactions on the Web (TWEB)*. 13(2). 1-22.
- Vinodhini, G. ve Chandrasekaran, R. M. (2012). "Sentiment analysis and opinion mining: a survey". *International Journal*. 2(6). 282-292.
- Vosoughi, S., Roy, D., ve Aral, S. (2018). "The spread of true and false news online". *Science*. 359(6380). 1146-1151.
- Wilson, T., Wiebe, J. ve Hoffmann, P. (2005). "Recognizing contextual polarity in phrase-level sentiment analysis". In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*. 347-354.
- Yerlikaya, T. ve Toker, S. (2019). "Seim Manipulasyonları ve Yalan Haberler | 31 Mart 2019 Yerel Seimleri". *SETA Raporu*. Eriřim: 28 Mart 2019, [https://setav.org/assets/uploads/2019/03/R133\\_Manipulasyon.pdf](https://setav.org/assets/uploads/2019/03/R133_Manipulasyon.pdf).
- Zhang, X., Cao, J., Li, X., Sheng, Q., Zhong, L. ve Shu, K. (2021). "Mining dual emotion for fake news detection". In *Proceedings of the Web Conference 2021*. 3465-3476.

## **EKLER**

“Sosyal Medya Kullanıcılarının Duygu Durumlarının Yalan Haber Yayılımına Etkisinin İncelenmesi”, ulusal bildirisi bu tez çalışmasından üretilmiştir.



