

**T.C.  
ONDOKUZ MAYIS ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ  
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI**



**DÖNÜŞTÜRÜCÜLER VE DERİN ÖĞRENME  
MODELLERİYLE SOSYAL MEDYA DUYGU ANALİZİ**

Yüksek Lisans Tezi

**Hüseyin İLGÜN**

Danışman

**Prof. Dr. Erdal KILIÇ**

SAMSUN  
2022

## TEZ KABUL VE ONAYI

Hüseyin İLGÜN tarafından, Prof. Dr. Erdal KILIÇ danışmanlığında hazırlanan “DÖNÜŞTÜRÜCÜLER VE DERİN ÖĞRENME MODELLERİYLE SOSYAL MEDYA DUYGU ANALİZİ” başlıklı bu çalışma, jürimiz tarafından 13.6.2022 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Yüksek Lisans Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	İmza	Sonuç
<b>Başkan</b>	Doç. Dr. Güzin ULUTAŞ Karadeniz Teknik Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı		<input checked="" type="checkbox"/>
			Kabul
			<input type="checkbox"/>
			Ret
<b>Üye</b> (Danışman)	Prof. Dr. Erdal KILIÇ Ondokuz Mayıs Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı		<input checked="" type="checkbox"/>
			Kabul
			<input type="checkbox"/>
			Ret
<b>Üye</b>	Doç. Dr. Sedat AKLEYLEK Ondokuz Mayıs Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı		<input checked="" type="checkbox"/>
			Kabul
			<input type="checkbox"/>
			Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

ONAY

... / ... / ...

Prof. Dr. Ali BOLAT

Enstitü Müdürü

## BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım yüksek lisans tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi ?

Evet  (Gerekli ise ekler kısmına ekleyiniz)

Hayır

İmza

04/07/2022

Hüseyin İLGÜN

## TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

**Tez Başlığı:** DÖNÜŞTÜRÜCÜLER VE DERİN ÖĞRENME MODELLERİYLE SOSYAL MEDYA DUYGU ANALİZİ

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 12/05/2022 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 18

Tek kaynak oranı : % 4 çıkmıştır.

İmza

04/07/2022

Prof. Dr. Erdal KILIÇ

## ÖZET

### DÖNÜŞTÜRÜCÜLER VE DERİN ÖĞRENME MODELLERİYLE SOSYAL MEDYA DUYGU ANALİZİ

Hüseyin İLGÜN

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Ana Bilim Dalı

Yüksek Lisans, Haziran/2022

Danışman: Prof. Dr. Erdal KILIÇ

Sosyal medya; internet kullanıcılarının herhangi bir konu, işletme, ürün ve durum hakkında olumlu ya da olumsuz fikirlerini belirttiği çevrimiçi bir ağ ortamıdır. Kullanıcıların hızlı erişimi sayesinde ilgili içeriklerin, makalelerin, haberlerin, düşüncelerin, günlük yaşanan olayların ve her türlü görsel ve işitsel materyallerin görüntülenebileceği ve paylaşılabilirdiği bir ortamdır. Sosyal medya duygu analizi, çeşitli sektörler ve akademik çalışmalar için popüler bir alandır. Fikir madenciliği olarak da bilinen bu çalışmalar bir metindeki genel duyguyu sınıflandırmak amacıyla yapılmaktadır.

Bu tez çalışmasında, kullanıcıların sosyal medya platformlarında yaptığı gönderilerin pozitif, negatif ve nötr duygulara göre sınıflandırılması yapılmıştır. Kullanılacak model üç aşamadan oluşmaktadır. İlk olarak veri setleri alınarak ön işleme işlemi ile model için hazır hale getirilmiştir. İkinci aşamada temizlenen veriler dönüştürücü model olarak kullanılacak DistilBERT modeli özellik vektörlerine dönüştürülmüştür. Son aşamada ise makine öğrenmesi ve derin öğrenme modelleriyle sınıflandırma işlemi yapılmıştır. Çalışmada dönüştürücü ve derin öğrenme modellerinin birlikte kullanımının dönüştürücü kullanılarak ya da kullanılmadan oluşturulan makine öğrenmesi modellerine ve literatürde bulunan geleneksel yöntemle oluşturulan modellere göre başarıyı arttırdığı görülmüştür.

**Anahtar Kelimeler:** Derin öğrenme, Makine öğrenmesi, LSTM, CNN, DistilBERT, K en yakın komşu, Lojistik regresyon, Destek vektör makineleri, Duygu analizi, Keras, Google colaboratory, BERT

# ABSTRACT

## SOCIAL MEDIA SENTIMENT ANALYSIS BY USING TRANSFORMERS AND DEEP LEARNING MODELS

Huseyin ILGUN  
Ondokuz Mayıs University  
Institute of Graduate Studies  
Department of Computer Engineering  
Master, June/2022  
Supervisor: Prof. Dr. Erdal KILIÇ

Social media; is an online network environment where internet users express their positive or negative opinions about any subject, business, product, or situation. It is an environment where users can view and share relevant content, articles, news, thoughts, daily events, and all kinds of visual and audio materials thanks to fast access. Social media sentiment analysis is a popular field for various industries and academic studies. These studies are also known as idea mining which is carried out to classify the general feeling in a text.

In this thesis, the posts made by users on social media platforms were classified according to positive, negative and neutral emotions. The model to be used consists of three stages. First, the data sets were taken and made ready for the model by preprocessing. In the second stage, the cleaned data were converted into feature vectors with the help of DistilBERT, which is used as a transformer model. In the last stage, classification was done with machine learning and deep learning models. In the study, it was seen that the use of transformer and deep learning models together increased the success compared to the machine learning models created with or without the transformer and the models created by the traditional method in the literature.

**Keywords:** Deep learning, Machine learning, DistilBERT , LSTM, CNN, KNN, Logistic regression learning, SVM, Sentiment analysis, Keras, Google colaboratory, BERT

## **ÖN SÖZ VE TEŞEKKÜRLER**

Akademik eğitim sürecimde hep yanımda olan, çalışmalarım boyunca yardım ve desteğini benden esirgemeyen yüksek lisans tez danışmanım değerli hocam Sayın Prof. Dr. Erdal Kılıç'a teşekkürlerimi sunarım.

Tez çalışmalarım boyunca desteğini esirgemeyen sevgili annem, babam ve kardeşlerime teşekkür ederim.

Hüseyin İLGÜN

## İÇİNDEKİLER

<b>TEZ KABUL VE ONAYI</b> .....	<b>i</b>
<b>BİLİMSEL ETİĞE UYGUNLUK BEYANI</b> .....	<b>ii</b>
<b>TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI</b> .....	<b>ii</b>
<b>ÖZET</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
<b>ÖN SÖZ VE TEŞEKKÜR</b> .....	<b>v</b>
<b>İÇİNDEKİLER</b> .....	<b>vi</b>
<b>SİMGELER VE KISALTMALAR</b> .....	<b>vii</b>
<b>ŞEKİLLER DİZİNİ</b> .....	<b>ix</b>
<b>TABLolar DİZİNİ</b> .....	<b>xi</b>
<b>1. GİRİŞ</b> .....	<b>1</b>
<b>2. ÖNCEKİ ÇALIŞMALAR</b> .....	<b>3</b>
<b>3. METODOLOJİ VE MATERYALLER</b> .....	<b>7</b>
3.1. Materyal .....	7
3.1.1. Veri Önışleme .....	7
3.1.2. Makine Öğrenmesi .....	8
3.1.2.1. K En Yakın Komşu .....	9
3.1.2.2. Destek Vektör Makineleri .....	10
3.1.2.3. Lojistik Regresyon .....	10
3.1.3. BERT .....	10
3.1.3.1. DistilBERT .....	15
3.1.4. Derin Öğrenme .....	16
3.1.4.1. Yinelenen Sinir Ağları .....	18
3.1.4.2. Uzun Kısa Süreli Bellek .....	19
3.1.4.3. Evrişimli Sinir Ağları .....	22
3.2. Metodoloji .....	24
3.2.1. Kullanılan Teknolojiler .....	25
3.2.2. Veri Setlerinin Hazırlanması .....	26
3.2.3. Dönüşürücü ve Makine Öğrenmesi Modelleri ile Duygu Analizi .....	27
3.2.4. Dönüşürücü ve Derin Öğrenme Modelleri ile Duygu Analizi .....	28
3.2.5. Performans Ölçütleri .....	30
<b>4. BULGULAR VE TARTIŞMA</b> .....	<b>32</b>
4.1. Makine Öğrenmesi ve DistilBERT Modeli Sonuçları .....	31
4.2. Derin Öğrenme ve DistilBERT Modeli Sonuçları .....	34
4.2.1. LSTM ve DistilBERT Model Sonuçları .....	35
4.2.2. CNN ve DistilBERT Model Sonuçları .....	43
<b>5. SONUÇ</b> .....	<b>51</b>

## SİMGELER VE KISALTMALAR

### SİMGELER

$\theta$	: Aktivasyon fonksiyonu
$h_t$	: Saklı durum vektörü
$i_t$	: Giriş kapısı
$o_t$	: Çıkış kapısı
$f_t$	: Unutma kapısı
$E$	: Girdi
$C$	: Çıktı
$h$	: Çıktı fonksiyonu
$t$	: Girdi zamanı
$\sigma$	: Sigmoid fonksiyonu
$W$	: Ağırlık matrisi
$b$	: Sabit
$c$	: Aday durum
$d$	: Uzaklık
$d_{EUC}$	: Öklid uzaklığı
$Q$	: Sorgu Vektörü
$K$	: Anahtar Vektör
$V$	: Değer Vektörü
$d_T$	: Sorgu Vektörü Boyutu

### KISALTMALAR

BERT	: Dönüştürücü ile Çift Yönlü Kodlayıcı Gösterimi (Bidirectional Encoder Representation from Transformers)
LR	: Lojistik Regresyon (Logistic Regression)
IMDB	: İnternet Film Veritabanı (Internet Film Database)
GOP	: Cumhuriyetçi Ulusal Komitesi (Republican National Committee)

ML	: Makine Öğrenmesi (Machine Learning)
MLM	: Maske Dili Modeli (Masked Language Model)
WORD2VEC	: Kelime Temsil Yöntemi
HTML	: Hiper Metin İşaretleme Dili (Hypertext Markup Language)
CLS	: Yapay Belirteş İşlemi (Class Artificial Token)
CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
KNN	: K-En Yakın Komşu (K-Nearest Neighbour)
LSTM	: Uzun Kısa Zamanlı Hafıza (Long Short Term Memory)
RNN	: Yinelenen Sinir Ağı (Recurrent Neural Network)
RNTN	: Özyinelemeli Tensör Sinir Ağı (Recursive Neural Tensor Networks)

## ŞEKİLLER DİZİNİ

Şekil 3.1. Farklı makine öğrenmesi tip ve görevleri .....	8
Şekil 3.2. Dönüştürücü temel mimarisi.....	10
Şekil 3.3. BERT temel mimarisi .....	11
Şekil 3.4. BERT çıktı vektörü gösterimi.....	12
Şekil 3.5. BERT kodlayıcı yapısı.....	13
Şekil 3.6. BERT ile sınıflandırma.....	14
Şekil 3.7. BERT ve DistilBERT karşılaştırılması.....	15
Şekil 3.8. Makine öğrenmesi ve derin öğrenme modeli karşılaştırması .....	16
Şekil 3.9. Derin öğrenme modeli .....	17
Şekil 3.10. RNN modeli yapısı .....	18
Şekil 3.11. LSTM genel mimari.....	20
Şekil 3.12. LSTM iç mimari .....	21
Şekil 3.13. CNN genel mimari.....	22
Şekil 3.14. Aktivasyon fonksiyonları.....	23
Şekil 3.15. Model genel mimarisi .....	24
Şekil 3.16. Kullanılan teknolojiler .....	25
Şekil 3.17. Dönüştürücü ve makine öğrenmesi model mimarisi .....	28
Şekil 3.18. Veri ön işleme aşaması .....	29
Şekil 3.19. DistilBERT ve derin öğrenme model mimarisi .....	30
Şekil 4.1. GOP veriseti LSTM eğitim ve test doğruluğu .....	36
Şekil 4.2. GOP veriseti LSTM eğitim ve test kaybı.....	37
Şekil 4.3. GOP veriseti LSTM karmaşıklık matrisi .....	37
Şekil 4.4. IMDB veriseti LSTM eğitim ve test doğruluğu.....	38
Şekil 4.5. IMDB veriseti LSTM eğitim ve test kaybı .....	39
Şekil 4.6. IMDB veriseti LSTM karmaşıklık matrisi.....	39
Şekil 4.7. Twitter veriseti LSTM eğitim ve test doğruluğu .....	41
Şekil 4.8. Twitter veriseti LSTM eğitim ve test kaybı.....	42
Şekil 4.9. Twitter veriseti LSTM karmaşıklık matrisi .....	42
Şekil 4.10. GOP veriseti CNN eğitim ve test doğruluğu .....	44
Şekil 4.11. GOP veriseti CNN eğitim ve test kaybı.....	45
Şekil 4.12. GOP veriseti CNN karmaşıklık matrisi .....	45
Şekil 4.13. IMDB veriseti CNN eğitim ve test doğruluğu.....	46
Şekil 4.14. IMDB veriseti CNN eğitim ve test kaybı .....	47

Şekil 4.15. IMDB veriseti CNN karmaşıklık matrisi.....	47
Şekil 4.16. Twitter veriseti CNN eğitim ve test doğruluğu.....	49
Şekil 4.17. Twitter veriseti CNN eğitim ve test kaybı.....	50
Şekil 4.18. Twitter veriseti CNN karmaşıklık matrisi.....	50

## TABLolar DİZİNİ

Tablo 3.1. IMDB veri kümesi .....	26
Tablo 3.2. Twitter veri kümesi.....	26
Tablo 3.3. GOP veri kümesi.....	27
Tablo 4.1. IMDB veriseti sonuçları .....	33
Tablo 4.2. Twitter veriseti sonuçları .....	34
Tablo 4.3. GOP veriseti sonuçları.....	34
Tablo 4.4. GOP veriseti LSTM performans metrikleri .....	35
Tablo 4.5. IMDB veriseti LSTM performans metrikleri.....	38
Tablo 4.6. Twitter veriseti LSTM performans metrikleri .....	40
Tablo 4.7. GOP veriseti CNN performans metrikleri .....	43
Tablo 4.8. IMDB veriseti CNN performans metrikleri.....	46
Tablo 4.9. Twitter veriseti CNN performans metrikleri .....	48
Tablo 5.1. IMDB veriseti için modellerin performans karşılaştırması .....	51
Tablo 5.2. GOP veriseti için modellerin performans karşılaştırması.....	51
Tablo 5.3. Twitter veriseti için modellerin performans karşılaştırılması.....	52
Tablo 5.4. Duygu analizi derin öğrenme modelleri performans karşılaştırması.....	52
Tablo 5.5. Duygu analizi dönüştürücü modelleri performans karşılaştırması .....	53

# 1. GİRİŞ

Sosyal ağ, başta Twitter olmak üzere hızlı gelişen ve insanların düşüncelerini görsel, işitsel ve metinsel yollarla paylaştığı çevrimiçi platformdur. Kullanıcılar belirli bir konu hakkında fikirlerini paylaşabilmekte ve başkalarının görüşlerini öğrenebilmektedir. Duygu analizi; bir dokümandaki belirli bir metinde, cümlede ve ya sözcüklerden oluşan bir yapıdaki kutuplaşmayı tanımlamak ve kategorize etmektir. Bir diğer deyişle bir metindeki genel duygunun tanımlanmasıdır. Bu teknik sağlık, e-ticaret, politika gibi pek çok alanda kullanılmaktadır. Örneğin duygu analizi bir şirket için ürünlerinin müşteriler tarafından nasıl yorumlandığının ya da beğenilip beğenilmediğinin izlenmesi için kullanışlıdır. Sosyal medya duygu analizinde ise temel amaç, belirli bir konudaki oluşturulmuş gönderilerin pozitif, negatif ve nötr olarak yorumlanmasıdır. (Nabizath Saleena 2018)

Twitter ve sosyal medya verilerini kullanarak duygu analizi yapmak için pek çok yaklaşım kullanılmaktadır. Makine öğrenmesi algoritmaları ise bunlardan bir tanesidir. Büyük veriler üzerindeki doğal dil işleme problemlerini çözmek için sınıflandırma kuralları ile genel öğrenme algoritmaları birleştirilmiştir. Özellikle Destek Vektör Makineleri ve Naive Bayes algoritmaları kelimeler üzerine duygu sınıflandırması yapmak için kullanılmıştır. (Shiyang Liao 2016) Lakin sınıflandırma yapılırken anahtar kelimeler çıkartılırken bilgi kaybı yaşanabilmektedir. Derin öğrenme modelleri de son yıllarda bilgisayarla görme ve konuşma tanımlama alanlarında mükemmel sonuçlar vermiştir. CNN modeli de dil işleme alanında etkili sonuçlar vermiştir. (Shiyang Liao 2016)

İnternet ve teknolojinin hızlı gelişimi ile birlikte metinler üzerinden duygu analizi yaparak bilgi çıkarma işlemi pek çok araştırmacının dikkatini çekmiştir. (Zhuyi Rao 2020) Görünüm düzeyinde duygu sınıflandırma işlemi duygunun polaritesini (pozitif, negatif ve nötr) bulmayı amaçlar. Derin öğrenme ve sinir ağlarındaki hızlı gelişmeler doğal dil işleme alanında sıkça kullanılmış ve başarı elde edilmiştir. Geleneksel yöntemler ve makine öğrenmesi modelleri sınıflandırma için kullanılmaktadır ama ilgili metnin sınıflandırma modelinde kullanılmasından önce vektör haline getirilmesi gerekmektedir. Bu amaçla word2vec modeliyle özellik çıkarımı yapılarak kelimeler çıkartılmaktadır . (Zhuyi Rao 2020) İlgili derin öğrenme modelleri ile sınıflandırma yapmadan önce ön

eđitim yapılması gerekmektedir. Bu da ilgili metinlerin 6zellik ıkarımının yapılması ile sađlanmaktadır. (Zhuyi Rao 2020)

Literat6rde duygu analizi kapsamında yapılan alıřmalar word2vec, tf-idf gibi 6znitelik ıkaran modeller ile sadece klasik makine 6đrenmesi modelleri ve CNN, LSTM, RNN gibi derin 6đrenme modelleri ile gereklenmiřtir. Bu alıřmalarda bazı sorunlar ortaya ıkabilmektedir bunlar: Klasik y6ntemlerin yetersizliđi, uzun metinlerde bařarı oranının d6řmesi, 6zellik vekt6rlerinin yetersizliđi, kelimelerin c6mle iin analizlerinin iyi yapılamaması bunlardan bazılarıdır.

alıřmamızda diđer alıřmalardan farklı olarak d6n6řt6r6c6 modeli olarak bilinen DistilBERT modeliyle 6znitelikler ıkarılıp sınıflandırma modellerinde kullanılmıřtır. DistilBERT modelinin diđer 6znitelik ıkaran modellere kıyasen avatajları řunlardır :

- ✓ Devasa eđitim setlerine sahip olması
- ✓ Uzun metinler 6zerinde alıřabilmesi
- ✓ Paralel iřlem sayesinde performans sorununu ařması
- ✓ Kelimelerin c6mle ii iliřkilerinin daha ayrıntılı ıkarılması
- ✓ S6zc6k bađlamını daha iyi ıkartabilmesi
- ✓ Aık kaynak kodlu olması

Bu alıřmada sınıflandırma ařamasında Evriřimli Sinir Ađları (CNN) ve Uzun Kısa S6reli Bellek (LSTM) derin 6đrenme modelleri kullanılmıřtır. D6n6řt6r6c6 modelleri ile kullanılan derin 6đrenme modelleri; hızlı olmaları ve dikkat mekanizması ile saklanması gereken verileri ve unutması gereken verileri 6đrenebilmesi ayrıca zaman serileri ile alıřmasından dolayı tercih edilmiřlerdir. Ayrıca d6n6řt6r6c6 kullanılarak oluřturulan makine 6đrenmesi ve derin 6đrenme modellerinin 6nceki alıřmalarda kullanılan klasik makine 6đrenmesi modelleri ile sınıflandırma alıřmalarına g6re ok daha dođru sonular verdiđi ve iyileřtirdiđi g6zlemlenmiřtir.

## 2. ÖNCEKİ ÇALIŞMALAR

Naive Bayes, Destek Vektörleri ve Random Forest algoritmaları ile modeller oluşturulmuş ve bu modellere göre Twitter metinlerinde duygu analizi çalışmaları yapılmıştır. Veri ön işleme yapıldıktan sonra temizlenen veri setlerine geleneksel makine öğrenmesi algoritmaları uygulanmıştır. (Nabizath Saleena 2018) Jocker's sözlüğüne göre tweet konu başlıklarına göre polarite hesaplaması yapılmıştır. (Sujay R. 2018)

Derin öğrenme modelleri, son yıllarda bilgisayarla görme ve konuşma tanımada harika sonuçlar elde etmiştir. CNN, derin öğrenme modellerinden biridir. Görüntü tanımada ve doğal dil işlemede kayda değer sonuçlar elde edilmiştir. Sosyal medya verileri, modelin yapısında bulunan evrişim katmanı sayesinde özellik çıkarımı operasyonuna girer ve metine dair özellikler çıkartılarak eğitim gerçekleştirilir. . (Shiyang Liao 2016)

Literatürde metinleri temsil etmek için en sık kullanılan yöntemler kelimelerin köklerinin, karakter n-gramlarının metinlerdeki geçiş sayıları ve bunların ağırlıklandırılmış halleridir. Kelime köklerinin kullanılabilmesi için bir morfolojik çözümleyiciye ihtiyaç varken, karakter n-gram vektörleri doğrudan kullanılabilir. Eğiticili terim ağırlıklandırma yöntemi ve geleneksel terim ağırlıklandırma yöntemleriyle duygu analizi yapılmaktadır. (Mahmut Çetin 2013)

Türkçe sosyal medya verileri kullanılarak pek çok duygu analizi çalışması da yapılmıştır. Twitter kullanıcılarının ekonomi ile alakalı olarak pozitif ve negatif duygularını analiz etmek için sözlük tabanlı metotlar ile %85 oranında başarı elde edilirken n -gram sözlük tabanlı model ile pozitif, negatif ve nötr duygu analizi yapıp %89 oranında başarı elde edilebilmektedir. Etiketli ve etiketsiz verilerde yarı eğitimli Word2Vec algoritması duygu analizi çalışması yapılmıştır. (Metin Bilgin 2017)

Duygu analizi çalışmalarında kullanılan derin öğrenme tekniklerinden en önemlileri LSTM ve GRU modelleridir. Bu RNN derin öğrenme modelleri yapısında bulunan Attention mekanizması sayesinde metin sunumlarını ve kelime ilişkilerini çıkartan modern yaklaşımlardır. Özellikle hisse senedi üzerine ilgili sosyal medya verileri ve haber verileri BERT ön eğitim ve CNN modelleri ile tahminleme ve sınıflandırma çalışmasında kullanılmıştır. Etiketli olan ve ön işleme tutulan daha sonra da temizlenen veri setleri

BERT dönüştürücü modeli ve ilgili derin öğrenme teknikleriyle borsa ve hisse senetleri üzerine pozitif ya da negatif sınıflandırma yapılmasını sağlamıştır. (Matheus Gomes de Sousa 2019)

Son yıllarda toplum güvenliği kapsamında dünya genelinde pek çok olay gerçekleşmiştir. Güvenliği sağlamak adına her an devasa veri aktarımı gerçekleşmektedir. Bu verileri işlemek geleneksel yollarla mümkün değildir. Bu sebeple word2vec gibi doğal dil işleme yöntemleri ilgili özellik vektörlerini çıkartmak için kullanılmaktadır. Yalnız bu modelin statik ve tek yönlü olması sebebi ile ilgili sözcüklerin temsili yeterli bulunmamaktadır. BERT modeli bu kapsamda sözcüklerin daha zengin ve kapsamlı olarak temsilleri çıkartılıp işlenmesi doğru sonuçlar vermektedir. (Jiaqi Hou 2020)

Soru cevap sistemleri temelde doğal dil işleme ve derin öğrenme modellerini baz alarak geliştirilmektedir. Bu sebeple pek çok organizasyon ve şirket bu sistemleri kullanarak piyasa araştırmaları yapmaktadır. Kelime işleme ve çıkarım yapmak gibi doğal dil işleme görevlerini yapan BERT bu alanda da kullanılmaktadır. Soru cevap sistemleri İngilizce tabanlı sistemler olduğundan Vietnam dilinde de Vietnam Vikipedi özgür kütüphanesini kullanarak ince ayar yapılmış BERT ile çoklu dil soru cevap sistemi geliştirilmiştir. (Nguyen Thi Mai Trang 2020)

Sosyal yazılım, duygu analizi çalışmalarında kullanılan bir türdür. Özellikle metin dosyaları kullanarak duygu sınıflandırması çalışmaları son zamanlarda sıklıkla kullanılmaktadır. Başlarda polarite ve kelime frekansına göre yapılan analiz çalışmaları daha sonra makine öğrenmesi yöntemleri kullanılarak devam etmiştir. Fakat bu modeller başarı ve performans açısından dezavantajları olmasından derin öğrenme modelleri kullanılmaya başlanmıştır. Metinden çıkarım yapabilmek için ön eğitim ve özellik çıkarımı önem arz etmektedir. Bu aşamada BERT ile ön çalışma yapıp sınıflandırma yapılabilmektedir. Sosyal medya alanında duygu analizi kapsamında uluslararası film veri tabanı kullanıcı yorumları kullanılarak sınıflandırma çalışması yapılmış ve başarılı sonuçlar alınmıştır. (Tingyu Zhan 2021)

Dünyada çeşitli veri biçimleri üretilmektedir. Bunlar arasında da metin verileri önemli hal almıştır. Dil işleme alanında ise adlandırılmış varlık tanıma yöntemleri geliştirilmektedir. Adlandırılmış kitlesel duygu analizi için metinler, anlamsal etiketleme

görevleri ve makine çevirisi gibi varlık tanıma teknolojileri çeşitli alanlarda yaygın olarak kullanılmaktadır. Bu nedenle, adlandırılmış varlık tanıma üzerine yapılan araştırmalar, doğal dil işleme alanına büyük önem taşımaktadırlar. Metin verilerinde sınıflandırma çalışmaları için dönüştürücü modelleri kullanılmaktadır. BERT; zengin, semantik özelliklere ve gramer özelliklerine sahip karakter vektörlerini yapı özellikleri ve ardından küresel bilgilerine ayıklar. Çoğu Çince adlandırılmış varlık tanıma yöntemlerine göre karşılaştırıldığında BERT modeli dili temsil etmek için eğitim öncesi dil modelindeki cümlelerin bilgilerini ve sorunlarını çözer. Bunun ardından dönüştürücü modeli kullanarak genel özellikler elde edilir. Derin öğrenme modeli ile sınıflandırma işlemi yapılarak etiketli verilerin kategorize edilmesi işlemi yapılmaktadır. (Yong Gan 2021)

Duygu analizi kapsamında haber başlıklarına göre duyarlılık tespiti ve sınıflandırılması yapılmaktadır. Kullanıcılar çoğu zaman bütün haberi okumadan sadece başlıklarına göre okuyacakları haberleri seçmektedir. Bu noktada ilgili konu başlıklarına, interaktifliklerine, yansıttığı duyguya göre başlıkları sınıflandırmak ve kullanıcıya buna uygun haberleri sunmak amacıyla çalışma yapılmıştır. Çalışmada önceden eğitilmiş BERT modeli ve daha az katmana sahip DistilBERT modelleri kullanılıp dönüştürücü modelleri ve sınıflandırıcıları kullanılmıştır. (Anurag Singh 2021)

Karakter sınıflandırması, psikoloji ve davranış bilimlerinde önemli yer tutmaktadır. Bu alanda sosyal araştırmalar yeterli olmadığından veri işleme ve analizi bilim insanlarını bu alanda da çalışmaya iten motivasyon olmuştur. Kişilik sınıflandırması ilgili duygusal bilgileri analiz etmek ve özetlemek için metindeki kişilik özelliklerini çıkarma mantığına dayanmaktadır. BERT, LSTM, CNN gibi derin öğrenme modelleri kullanılmıştır. Makine öğrenme modelleri zayıf sonuçlar verdiği için onlara kıyasen yukarıda bahsedilen modeller çok daha başarılı sonuçlar vermektedir. Özellikle BERT öznelik çıkarımı sonrası softmax sınıflandırıcı ile sınıflandırma sonuçları tahminleme olarak çok daha doğru sonuçlar vermektedir. (He Jun 2021)

Otomatik konuşma tanıma sistemleri(ASR), günümüzde özellikle film sektöründe altyazı ve çeviri alanlarında kullanılmaktadır. Word2vec ve BERT modelleri ile yazıdan konuşmaya ve konuşmadan yazıya yapılan dönüşümlerde yüksek başarı oranları yakalanmıştır. (Wenkun Li 2021)

Kullanıcıların Twitter gibi sosyal mecralarda oluşturduğu gönderiler, hükümet ve kolluk kuvvetleri tarafından bir olay ve ya felaket tespitinde analiz edilmektedir. Geleneksel makine öğrenmesi algoritmalarının yetersiz kalması sebebiyle BERT, CNN ve LSTM gibi modeller ile gönderilerdeki yaygın duygunun bir felaketi ve kötü bir olayı yansıttığının tespiti kolaylaşmıştır. (Zihan Wang 2020)

Mizah, toplum sağlığı ve insan psikolojisi için gerekli kavramdır. Mizah görsel, yazınsal ve duyuşsal olarak algılanabilmektedir. Metin dosyalarındaki mizahın ve ilgili metnin komik olup olmadığının tespiti çalışma konusu olmuştur. BERT ve Glove modelleri duygu tespitinde kullanılmıştır. (RIDA MIRAJ 2021)

Görüntü yazısı kavramı nesne ve kişilerin görüntülerinin ifade ettiği anlamı çok daha derinleştirmektedir. Görüntü yazısı görüntünün kendisi ve çevresi ile alakalı bilgiler taşımaktadır. Görüntü yazısının bütün bileşenlerinin analiz edilmesi önemli bir araştırma konusu olmuştur. Özellik çıkarımı ve doğal dil işleme safhalarından oluşan bu yöntem, word2vec modelinden çok daha iyi özellik çıkarımı yapan BERT modeli gerçekleştirilmiştir. (Supragya Sonkar, 2021)

İlaç kullanımı son 10 yılda dramatik şekilde artmıştır. Kullanılan ilaçların birbirleriyle olan etkileşimi işleyişi etkileyebilmektedir. Bu etkileşimin önceden tahmin edilebilmesi için çeşitli çalışmalar yapılmıştır. Bu amaçla büyük boyutta tıbbi dokümanın ayrıştırılması ve etki durumuna göre etiketli veriler baz alınarak BERT ile özellik çıkarımı yapılmak suretiyle sınıflandırılması bu etkilerin tahmin tahminlenmesi açısından yüksek başarı oranı sağlamaktadır. (Tanmoy Tapos Datta 2022)

### 3. METODOLOJİ VE MATERYALLER

Bu bölümde, çalışmamızda kullanılmış olan teknolojiler, yapılar, modeller, algoritmalar ve yöntemlerden bahsedilecektir.

#### 3.1. Materyal

Duygu analizi için genellikle geleneksel makine öğrenmesi algoritmaları, derin öğrenme yöntemleri, yapay zeka modelleri ve hepsinden önce veriyi model için hazırlamak adına pek çok kavramdan söz edilmektedir. Çalışmamızda kullanılmış olan yapılar bu bölümde detaylıca anlatılmıştır. Modelimiz; veri ön işleme, dönüştürücü ve sınıflandırma olmak üzere 3 ana katmanda oluşmaktadır.

Bu kısımda çalışmamızda izlediğimiz yol ve yöntemleri içeren veri setleri hazırlama, veri ön işleme adımları, dönüştürücü modelleri ve sınıflandırma için kullanılacak olan makine öğrenme ve derin öğrenme modellerinden bahsedilecektir.

##### 3.1.1. Veri Ön İşleme

Başta sosyal medya platformları olmak üzere çeşitli veri havuzlarından çekilen veriler pozitif, negatif ve nötr olarak sınıflandırılmadan önce ilk olarak ön işleme safhasında gereksiz ve yanlış olan veriler düzeltilmektedir.

Ön işleme, makine öğrenimi çalışmaları için önemli bir adımdır. Çevrimiçi ortamlarda, sosyal medya verileri ve metinler genellikle resmi bir yazı değildir. Özellikle gençler sosyal medya paylaşımlarında emoji kullanımını ve dil kurallarına uyulmadan kısa metin yazımını tercih edebilmektedir. Kısaltılmış kelime biçimleri, özel semboller ve karakterler, kelimeler, dilbilgisi hataları ve ya birleşik kelimelerin yanlış yazımı gibi hataları ve ya eksiklikleri içeren veriler modellere dahil edilmeden önce aşağıdaki ön işleme adımlarından geçmelidir :

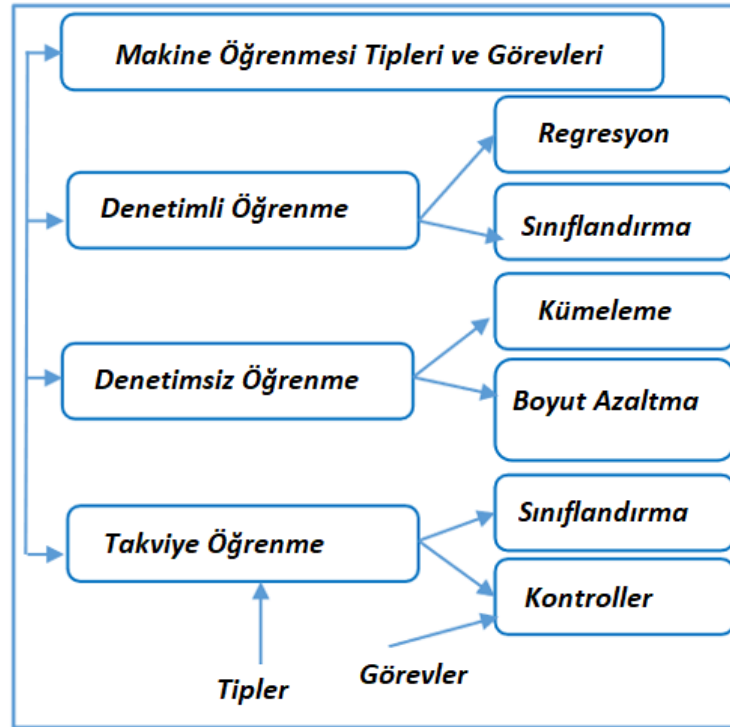
**1. Adım:** 100k gibi karakterler ve yüzde işareti gibi özel karakterler kelime formlarıyla değiştirilmelidir.

**2. Adım:** "=", "?", "#", "+", "-" gibi tüm özel karakterler, "\$", "@", "&" Vb. emojiler kaldırılmalıdır.

**3. Adım:** Kelimelerin kısaltılmış formları orijinal haline dönüştürülmelidir.( Nguyen The Lam 2019)

### 3.1.2. Makine Öğrenmesi

Makine öğrenmesi en yaygın sınıflandırma tekniğidir. Makine öğrenmesi, girdi ve çıktı değişkenleri arasındaki bilinmeyen fonksiyon, ilişki ve yapıları tanımlamaya yaramaktadır. Etiketli, etiketsiz verileri inceleyen yapay zeka dalıdır. Genel yapay zeka modelleri kümeleme, sınıflandırma, boyut küçültme gibi yöntemleri içerir. (Shrikant Tiwari 2022) Şekil 3.1’ de makine öğrenmesi tipleri ve görevleri gösterilmektedir.



Şekil 3.1. Farklı makine öğrenmesi tip ve görevleri

Çalışmamızda dönüştürücü kullanılarak ve kullanılmadan sınıflandırma işlemi yapılmıştır. Sınıflandırma modelleri olarak K En Yakın Komşu, Lojistik Regresyon ve Destek Vektör Makineleri modelleri kullanılmıştır.

### 3.1.2.1. K En Yakın Komşu (KNN)

KNN, yapay zeka hesaplamalarında kullanılan en temel ve esnek algoritmadır. Verisetlerine bağlıdır. Pratikte faydalı olsa da gerçek veri setleri matematiksel varsayımları doğrulamayabilir. Bu da eğitim zamanını hızlı test zamanını da yavaş ve maliyetli yapar. En kötü senaryoda ise bütün noktaların tarandığını düşünürsek çok daha fazla bellek gerekecektir. (Chirag Kariya 2020)

KNN algoritması çok boyutlu örnekleri eğitim setindeki verilerin bir birlerine olan uzaklıklarını hesaplayarak sınıflandırma mantığıyla çalışmaktadır. Sınıflandırma, k tane en yakın noktanın bulunarak buna göre sınıflandırmasından türer. Algoritma çalışma mantığı şu şekildedir : (João Vieira 2019)

1. Algoritmanın ilk adımı olarak noktalar arası uzaklıklar hesaplanmaktadır. En sık kullanılan uzaklık formülleri Manhattan ve Öklid uzaklıklarıdır. Öklid uzaklık hesaplamasına dair temel işlemler 1, 2 ve 3 numaralı formüllerde verilmektedir.

$$d(A, B) = \sum_{i=0}^M (A_i - B_i)^2 \quad (1)$$

$$d_{Euc}(A, B) = \sqrt{\sum_{i=0}^M (A_i - B_i)^2} \quad (2)$$

$$d_{Euc}(A, B) > d_{Euc}(A, C) \rightarrow d(A, B) > d(A, C) \quad (3)$$

Mesafe hesaplamasındaki temel mantık veriler arasındaki en yakın k tane veriyi bulmaktır.

2. Sonraki adımda tam sayı olarak belirlediğimiz k sayısına göre test için verilen veriye ilgili özellik alanlarına göre en yakın k tane veriyi bulmak. Bulunan verilerde sayısal olarak çoğunluğa sahip olan verilerin etiketi, sınıflandırmamızın da sonucunu vermektedir.

### 3.1.2.2. Destek Vektör Makineleri (SVM)

Olasılıksal sınıflandırıcı olarak da bilinen algoritma, eğitim için eğitim verisine ihtiyaç duymaktadır. Doğrusal olmayan haritalama kullanılarak farklı sınıflar arasında en büyük marjinin bulunması hedeflenmektedir. Eğitim zamanı ve performans olarak düşük olsa da doğruluk oranı oldukça yüksektir. Naïve Bayes sınıflandırıcıda olduğu gibi şartlı bağımsızlık aramaz. (Ankit 2018)

### 3.1.2.3. Lojistik Regresyon (LR)

Sınıflandırma için kullanılan regresyon modellerinden biridir. Genellikle kategorik bağımlı değişken ile diğer bağımsız değişkenlerin ilişkisini incelemektedir. LR, sınıflar arası boşluğu maksimize edecek hiper düzlem bulmaya çalışır. (Ankit 2018)

Lojistik regresyon doğrusal regresyona benzemektedir. Ana farkı y değişkeninin tipidir. Doğrusal regresyonda bağımlı değişken olan y miktarsal veri iken lojistik regresyonda kategorik sınıflandırılmış veridir. (Tianxiao Lui 2020)

### 3.1.3. BERT

Dönüştürücü modelleri; kodlayıcı ve çözümleyici olmak üzere iki kısımdan oluşmaktadır. Genellikle çeviri çalışmalarında kullanılan dönüştürücülerin genel yapısı Şekil 3.2' de gösterilmektedir.

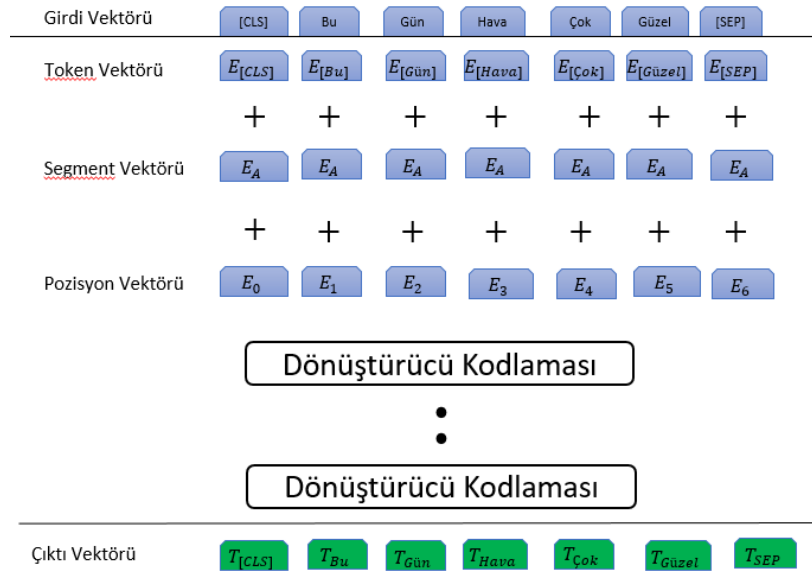


Şekil 3.2. Dönüştürücü Genel Mimarisi

BERT modelinin tam adı Dönüştürücü ile Çift Yönlü Kodlayıcı Gösterimi'dir. BERT modelinde dönüştürücü modelinin kodlayıcı birimi kullanılmaktadır. Bu da

öznitelik vektörü olarak kullanılacak bölüm olmaktadır. Modelin temel mimarisi Şekil 3.3' de gösterilmiştir. BERT modeli Google tarafından yayınlanmış dönüştürücü modeli olup, 2018 yılında alt yapı olarak nitelendirilen yeniden keşfedilmiş iyileştirilmiş Doğal Dil İşleme (NLP) modeli sunmaktadır (Whenyu Chen, 2020). BERT modeli iki temel kısımdan oluşmaktadır: Ön işleme ve ince ayar (fine tuning). Ön işleme bölümünde, büyük boyutta etiketlenmemiş veriler üzerinde dil modellemesi için denetimsiz bir yaklaşım kullanılır. Bu aşamadan sonra ön eğitim bölümünün parametreleri ince ayar bölümünde kullanılarak ve sınıflandırma amacıyla katmanlar eklenmektedir.

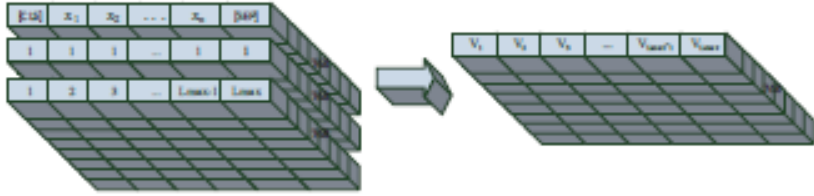
Bir kelimenin sol ve sağ bağlamını birleştirmek için BERT; çift yönlü dönüştürücü kodlayıcı kullanır. Model aynı zamanda "Maskelenmiş Dil Modeli" ve "Sonraki Cümle Tahmini" işlemlerini amaçlar. Bu sayede bir kelimenin cümle içinde ifade ettiği anlamı çok daha detaylı çıkarabilmektedir. (Qing Cai 2019)



Şekil 3.3. BERT Temel Mimarisi

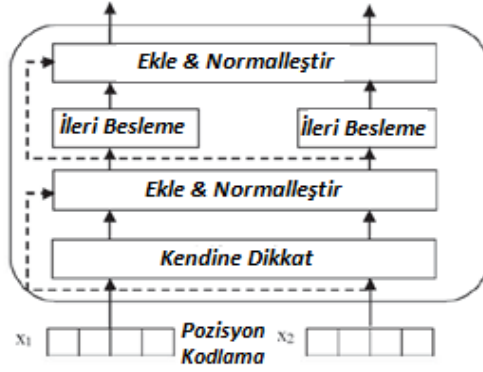
Şekil 3.3.'de temel BERT mimarisi gösterilmektedir. Modele aktarılan Girdi vektörü; token, pozisyon ve segment vektörlerinin çıkarılması sonrası kodlayıcı yapısından geçerek Çıktı vektörü haline getirilmektedir. Bu vektörlerin yapısı şekil 3.4' de gösterilmiştir. Çıktı vektörünü oluşturan yapılardan bahsedecek olursak, token vektörleri metin içerisindeki her bir karakterin sabit boyuta vektörize edilmiş halidir.

Pozisyon vektörü, cümle içerisindeki pozisyon bilgilerini modele bilgi olarak vermektedir. Bu özellik normal dönüştürücü modellerinde bulunmamaktadır. Segment vektörü ise bir cümleden daha geniş bir bağlam oluşturmak ve metin bilgilerini bölmek için kullanılır. Bu üç vektörün toplamı BERT kodlayıcı yapısı için girdi olarak kabul edilmektedir. Dönüştürücü kodlaması olarak belirlenen dikkat tabanlı katmanda dönüştürücü modelinin ana işlemleri yapılmaktadır. Kodlama katmanında token, segment ve pozisyon vektörlerinin cümle içerisinde ağırlıklı toplamları hesaplanmaktadır. Bunun sonucunda sınıflandırma için kullanılacak öznitelik Çıktı vektörleri oluşturulmaktadır.



Şekil 3.4. BERT Çıktı Vektörü Gösterimi (Matheus Gomes de Sousa 2019)

BERT modelinde bulunan dönüştürücü kodlayıcı katmanı, dönüştürücü modelinden dönüştürücü bloklarının kullanımı ile oluşturulur. Temel BERT modelinde bu bloklardan 12 katman halinde kullanılır, her biri 12 çok başlı dikkat bloğundan oluşur. Ayrıca BERT modeli hem önceki hem de gelecekteki bağlamlardan haberdar olmak için Maskeli Dil Modelini (MLM) kullanmaktadır.(Akbar Karimi 2020) BERT modeli dönüştürücü kodlayıcı yapısının işleyişi Şekil 3.5' te gösterilmektedir.

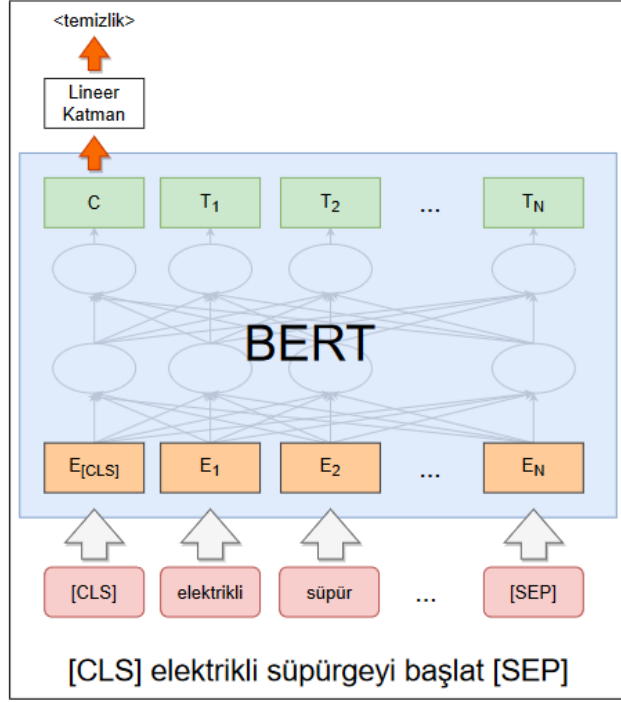


Şekil 3.5. BERT Kodlayıcı Yapısı

Bu kodlayıcı yapısında, kodlayıcı katmanına girdi olarak gelen vektörlerden  $Q$ ,  $K$ ,  $V$  ve  $d_k$  vektörleri oluşturulmaktadır. Katmanlarda bulunan çok başlı dikkat bloklarında denklem 4’ de görüldüğü üzere ilgili vektörlerin softmax aktivasyon fonksiyonunda kullanılmasıyla kelimelerin birbirleriyle ilişkileri ve ağırlıkları çıkartılmaktadır.

$$Attention(Q, V, A) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

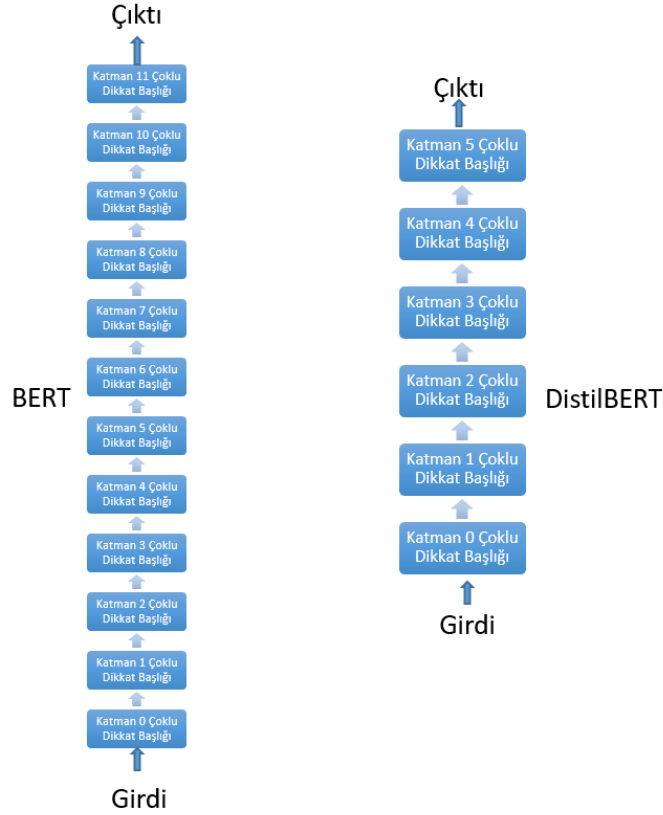
BERT ile oluşturulan öznitelik modellerinde word2vec, glove gibi öznitelik çıkarma işleminde karşılaşılabilecek bilgi eksikliğinin önüne geçilmiş olur. Burada oluşturulan ön eğitilmiş kelime vektörleri BERT modelinin kendi sınıflandırma modelinde de kullanılmaktadır. (Yunxiang Zhang 2020 ) Şekil 3.6 ’da BERT modeli kullanılarak yapılmış öznitelik çıkarma ve sınıflandırma işlemi gösterilmektedir.



Şekil 3.6. BERT ile Sınıflandırma

### 3.1.3.1 DistilBERT

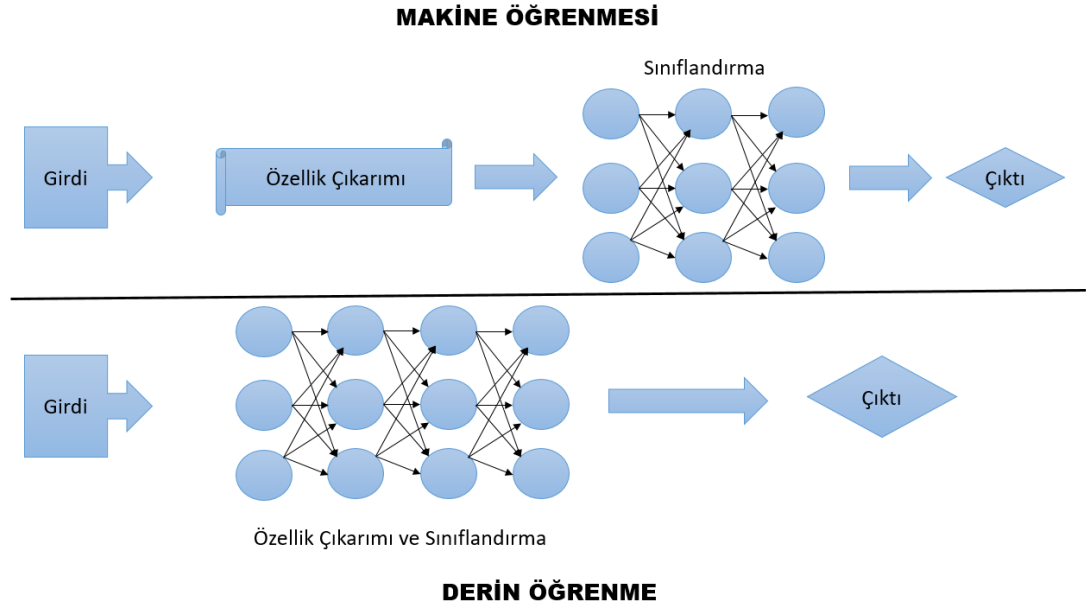
DistilBERT modeli BERT modeliyle aynı yapıya sahiptir. İki model arasındaki farklılıklara değinecek olursak; DistilBERT küçük, hızlı, ekonomik ve çok daha hafif yapılar kullanmaktadır. Aynı zamanda %40 daha az parametre kullanıp %60 daha hızlı çalışmaktadır (Asma UI Hussna, 2021). Ayrıca katman sayısı olarak farklılıklar mevcuttur. BERT modeli on iki katmandan oluşmaktadır. DistilBERT ise çoklu dikkat başlığına (multi head attention) sahip altı katmandan oluşmaktadır. Mevcut BERT modelinin 0, 2, 4, 7, 9 ve 11 katmanlarının parametreleri alınırken DistilBERT' in her katmanının parametrelerinin başlangıç değerleri verilerek ince ayar (fine tuning) işlemi gerçekleştirilmektedir (Jing Bai, 2020). İki model arasındaki yapısal farklılık Şekil 3.7' de gösterilmektedir.



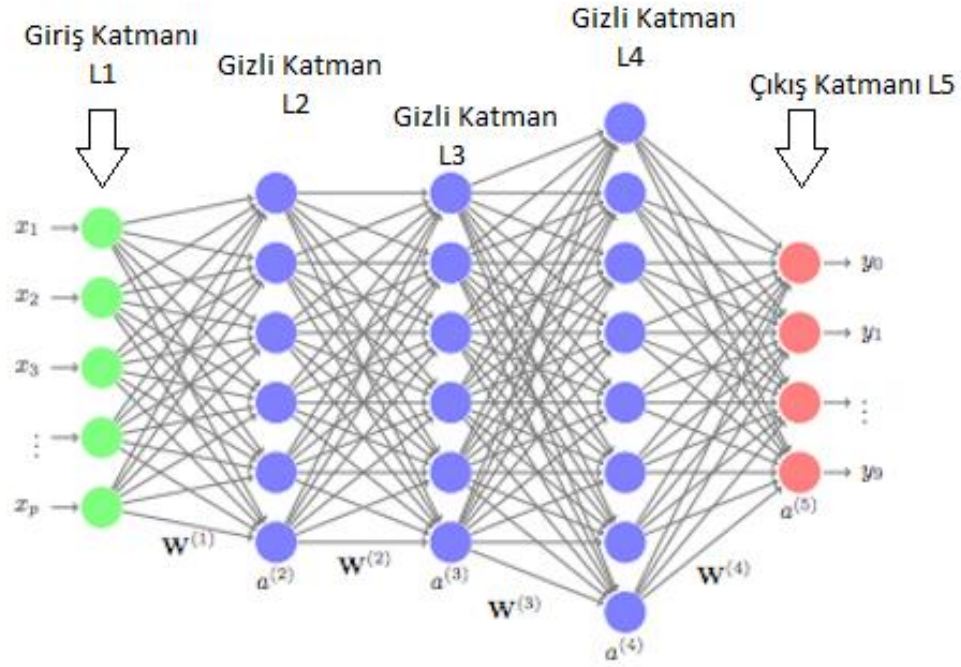
Şekil 3.7. BERT ve DistilBERT Karşılaştırılması

### 3.1.4. Derin Öğrenme

Derin öğrenme makine öğrenmesinin alt sınıfı gibidir ve de geleneksel yöntemlere göre büyük veri işlemesine daha uygundur. Veri hacmi arttıkça performans da artar. Geleneksel makine öğreniminden farklı olarak derin öğrenme, uygulama özelliklerinin yapay olarak belirlenmesine dayanmaz. Bunun yerine, verilerden doğrudan daha yüksek seviyeli özellikler elde etmeye ve özelliklerin çoklu dönüşümleri yoluyla derin seviyeli bir makine öğrenimi modeli elde etmeye çalışır. Derin öğrenme ve makine öğrenmesi modelleri arasındaki fark Şekil 3.8’ de gösterilmektedir. (Yu-nan Dong 2019)



Şekil 3.8. Makine Öğrenmesi ve Derin Öğrenme Modeli Karşılaştırması



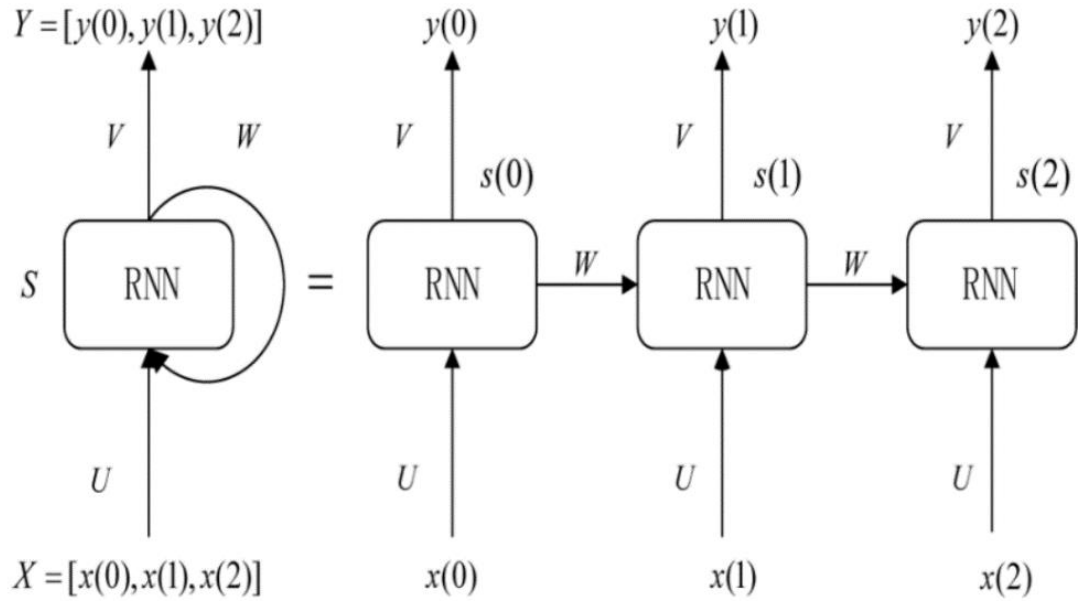
Şekil 3.9. Derin Öğrenme Modeli (UC Business Analytics R Programming Guide)

Son zamanlarda derin öğrenme modelleri, yapay zeka araştırmaları alanında en sık kullanılan modellerdir. Aynı zamanda yapay zeka ağının bir dalıdır. Derin öğrenmenin amacı, örnek verilerin iç kurallarını öğrenmektir ve öğrenme sürecinde elde edilen bilgiler metin, görüntü ve ses verilerini yorumlamada büyük yardımcıdır. Derin öğrenmenin nihai amacı, makine veya bilgisayarı analitik olarak insanlar kadar yetenekli hale getirerek metin, görüntü ve ses gibi verileri tanımalarını sağlamaktır. Şekil 3.9' da görüleceği üzere Derin öğrenme sinir ağı genellikle bir girdi katmanı, bir veya daha fazla gizli katman ve bir çıktı katmanından oluşur. (Jinqiao Liu, 2020)

Bu bölümde RNN modeli ve çalışmamızda kullanılmış olan LSTM ve CNN derin öğrenme modelleri hakkında bilgiler verilecektir.

### 3.1.4.1 Yinelenen Sinir Ağları

Geleneksel yapay sinir ağı yapısı, yalnızca mevcut girdiden hedef vektöre eşlenebilen bitişik katmanlar arasındaki tam bağlantı ile karakterize edilirken Yinelenen Sinir Ağı (RNN) ile önceki girdinin tüm geçmişinden hedef vektörleri eşleştirilebilir. Bu nedenle, geleneksel sinir ağları ile karşılaştırıldığında, RNN; sürekli dizi verilerinin dinamiklerini modellemede daha etkilidir. Genel olarak, RNN yönlendirilmiş döngülerdeki birimler arasında bağlantılar kurar ve dahili durumu aracılığıyla önceki girdileri hatırlar. Spesifik olarak, t-1 zaman adımındaki RNN çıkışı, t zaman adımındaki RNN çıktısını etkileyebilir. Bu, RNN'nin bir geçerli zaman korelasyonel dizisi ve önceki bir dizi oluşturmasını sağlar. Şekil 3.10' da RNN model yapısı gösterilmektedir.



Şekil 3.10. RNN Modeli Yapısı (Qiangqiang Ye, 2019)

Şekil 3.10' da görüleceği üzere,  $X = [x(0), x(1), x(2)]$  dizi vektörleri, ayarlanan zaman adımına göre birer birer RNN modeline geçmektedir. Bu model, tüm dizi vektörlerinin bir kerede modeli beslediği geleneksel ileri beslemeli ağdan açıkça farklıdır. İlgili matematiksel denklemler aşağıdaki gibi tanımlanabilir:

$$S_{(t)} = \sigma(U \cdot x_{(t)} + W \cdot S_{(t-1)} + b) \quad (5)$$

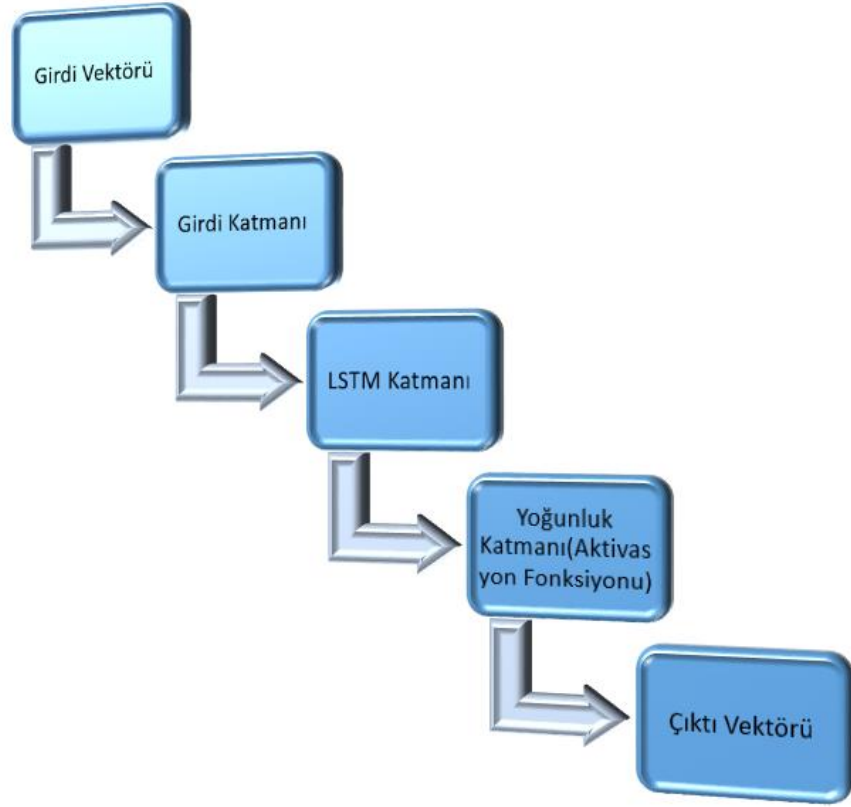
$$y_{(t)} = \sigma(V \cdot s_{(t)} + c) \quad (6)$$

Denklem 5' te  $x(t)$ ;  $t$  zaman adımının girdi değişkenidir,  $W$ ;  $U$  ile  $V$ ' nin bir ağırlık matrisidir,  $b$  ile  $c$  sapma vektörüdür,  $\sigma$  aktivasyon fonksiyonudur, Denklem 6,  $y(t)$   $t$  adım zamanının beklenen çıktısıdır.(Qiangqiang Ye, 2019)

RNN dil modelinin mimarisi genellikle üç katman içerir: giriş katmanı, gizli katman ve çıkış katmanı. Yapı modeli, zamana göre genişletilmiş döngüsel bir sinir ağı dili modeli şeklindedir. RNN dil modelinde verilen bilgi, semantik koşulu altında görünen her bir kelimenin olasılığı, RNN tarafından zaman genişlemesi şeklinde iteratif olarak hesaplanabilir ve kelime dizisindeki bir kelimenin koşullu olasılığı her zaman adımında hesaplanabilir. (Jianqiong Xiao, 2020)

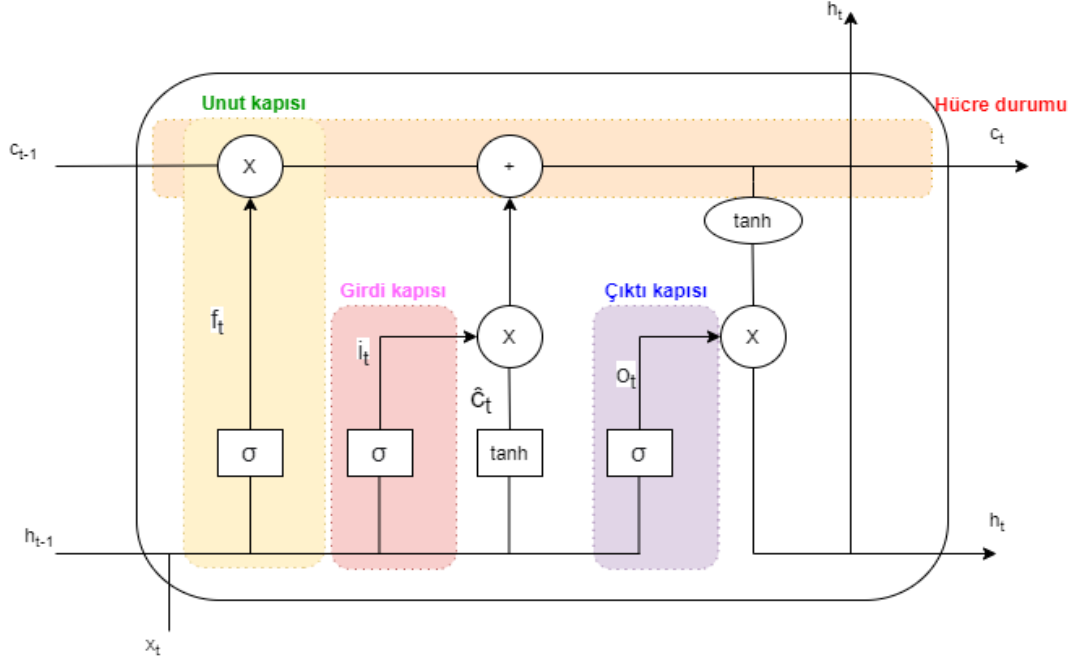
### 3.1.4.2. Uzun Kısa Süreli Bellek

Vanilya RNN' lerinin kaybolan gradyanları sorununa bir çözüm olarak, Uzun Kısa Süreli Bellek modeli olan LSTM 1997'de Sepp Hochreiter ve Jürgen Schmidhuber tarafından önerildi. Ortak bir LSTM birimi, bir hücre, bir giriş kapısı, bir çıkış kapısı ve bir unutma kapısından oluşur. Hücre rastgele zaman aralıklarında değerleri hatırlar ve üç kapı hücrenin içine ve dışına bilgi akışını düzenler. Sezgisel olarak, hücre, giriş dizisindeki öğeler arasındaki bağımlılıkları izlemekten sorumludur. Giriş kapısı, hücreye yeni bir değer ne ölçüde aktığını kontrol eder, unutma kapısı, bir değer hücrede ne kadar kaldığını kontrol eder ve çıkış kapısı, hücredeki değer LSTM ünitesinin aktivasyon çıkışını hesaplamak için ne ölçüde kullanıldığını kontrol eder.(Shayak Chakraborty, 2020)



Şekil 3.11. LSTM Genel Mimari

LSTM'in genel mimari yapısı Şekil 3.11'de verilmiştir. LSTM, tekrar eden bir hücre zinciri şeklini alır. Hücre, ağı geçmiş bilgileri hatırlamasını sağlamak için özel bir şekilde etkileşime giren dört tür etkileşimli sinir ağı içerir. LSTM, giriş kapısı, çıkış kapısı ve unutma kapısı aracılığıyla hücrelerin durumunu korur ve kontrol eder. Şekil 3.12'de LSTM iç mimarisi gösterilmektedir.



Şekil 3.12. LSTM İç Mimari (Ersin Enes Yılmaz, 2021)

Bir LSTM hücresinin içerisindeki hesaplamalar aşağıdaki formüllerde verilmektedir:

$$f_{(t)} = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (7)$$

$$i_{(t)} = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (8)$$

$$\tilde{C}_t = \tan h(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (9)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (10)$$

$$o_{(t)} = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (11)$$

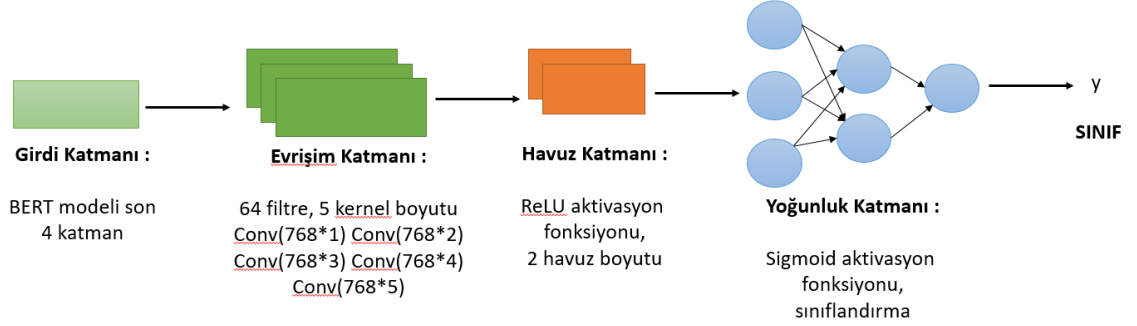
$$h_t = o_t * \tan h(C_t) \quad (12)$$

Denklem 7 unutma kapısı aktivasyon fonksiyonudur. Denklem 8 girdi kapısı aktivasyon fonksiyonunu göstermektedir. Denklem 11 çıkış kapısı aktivasyon kapısını

göstermektedir. Denklem 9 hücre girişini göstermektedir. Denklem 10 hücre durumunu aktivasyon fonksiyonudur. Denklem 12 ise LSTM hücresinden çıkış vektörünü hesaplamaktadır.(Yan Li, 2019)

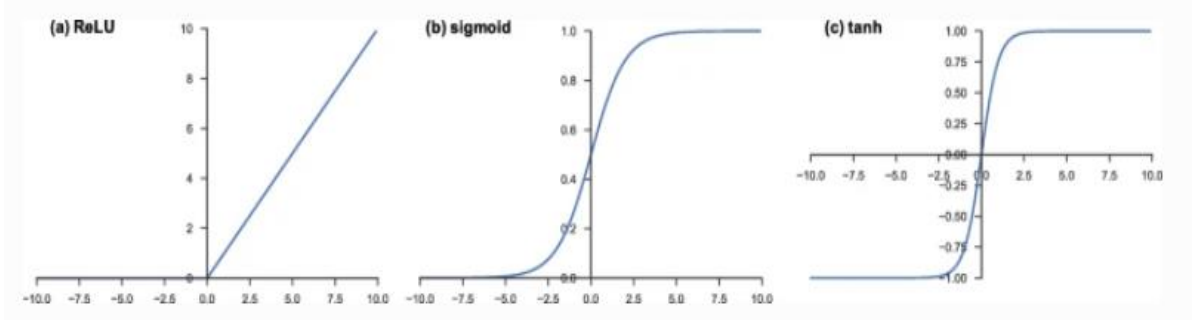
### 3.1.4.3. Evrişimli Sinir Ağları

Evrişimli Sinir Ağı olarak çevrilen CNN modeli, görüntü ve ya metin verilerini işlemek için kullanılan için ızgara(grid) düzenine sahip derin öğrenme modelidir. CNN üç katmandan oluşan matematiksel bir modeldir. Evrişim, havuz ve tam bağlantı katmanı katmanlarından oluşmaktadır. Evrişim ve havuz katmanı öznelik çıkarma işlemi yaparken tam bağlantı katmanı ise sınıflandırma işlemi yapmaktadır (Rikiya Yamashita, 2018). CNN genel mimarisi Şekil 3.13' de gösterilmektedir.



Şekil 3.13. CNN Genel Mimari

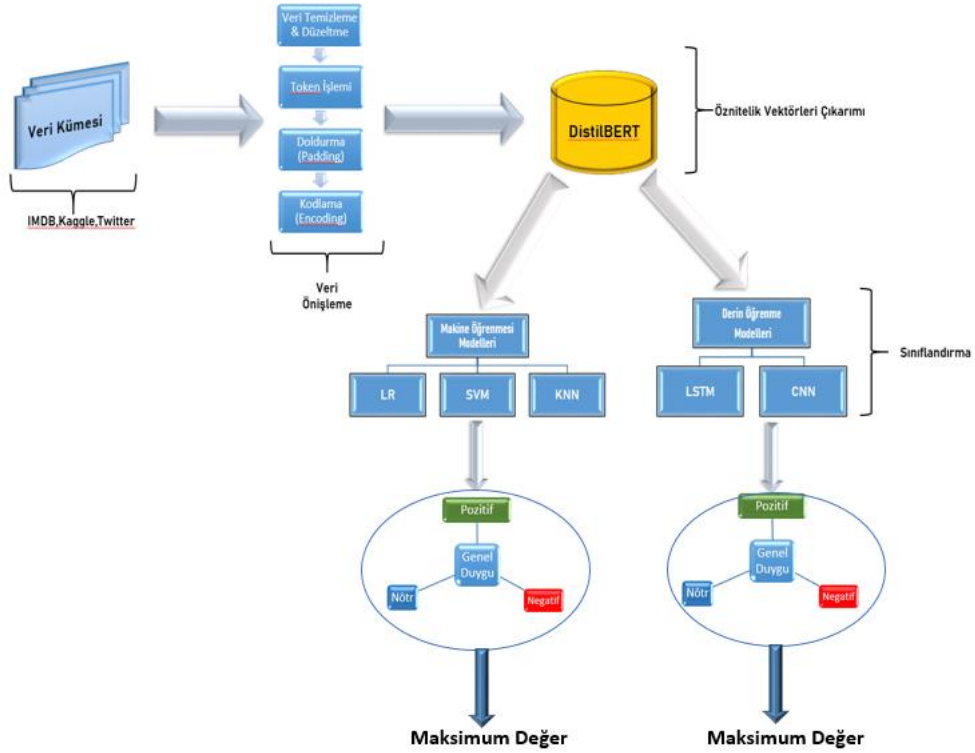
Evrişim katmanında, öznitelik çıkarma için kullanılan, tensör adı verilen bir sayı dizisi olan girdi boyunca kernel (çekirdek) olarak tanımlanan bir sayı dizisinin uygulandığı özel bir doğrusal işlem türüdür. Havuz katmanında özellik haritalarının parametre ve düzlem içi boyutluluğunu azaltan yeni bir örneklem oluşturulmaktadır. Son evrişim ve havuz katmanında tek boyutlu 1D sayı dizisi oluşturulduktan sonra yoğunluk katmanında sınıflandırma işlemi yapılmaktadır. CNN ve LSTM modellerinde yoğunluk ve sınıflandırma çeşitli aktivasyon fonksiyonları kullanılmaktadır. Bu fonksiyonlardan en yaygın kullanılanlar relu, sigmoid ve tanh fonksiyonlarıdır. Bu fonksiyonların grafiksel gösterimleri Şekil 3.14' te gösterilmektedir.



Şekil 3.14. Aktivasyon Fonksiyonları

### 3.2. Metodoloji

Dönüştürücü modeli ile öznitelikleri çıkarılan metinlerin derin öğrenme ve makine öğrenmesi algoritmaları ile duygu analizinin yapılması bu bölümde detaylı şekilde anlatılmıştır. Bu çalışma için kullanılmış teknoloji ve yapılar, veri kümelerinin eğitim ve test aşamasına kadar hazırlanıp ön işlemlerinin yapılması, özniteliklerinin çıkarılıp ilgili modellerde analiz ve tespitlerinin yapılması bu kısımda detaylandırılmıştır. Çalışmamızda kullanılan modelin genel mimarisi Şekil 3.15’ te gösterilmektedir.



Şekil 3.15. Model Genel Mimarisi

### 3.2.1. Kullanılan Teknolojiler

Bu çalışmada derleme ve test aşamalarında Intel i5-6200 2.3 Ghz notebook ve Google Colaboratory çevrimiçi kod yazma ve derleme ortamı kullanılmıştır. Veri setleri için Kaggle platformu üzerinden veriler Drive sistemine aktarılmış ve kullanılmıştır. Visual Studio Code platformu ile birlikte keras, transformers, tensorflow, sklearn, pandas gibi python kütüphanelerinden faydalanılmıştır. Çalışmada kullanılan temel teknolojiler Şekil 3.16' da gösterilmektedir.



Şekil 3.16. Kullanılan Teknolojiler

### 3.2.2. Veri Setlerinin Hazırlanması

Dönüştürücü modeli ile öznitelikleri çıkarıldıktan sonra makine öğrenmesi ve derin öğrenme modelleri ile sınıflandırılması işlemi için 3 farklı veri seti kullanılmıştır. İzleyicilerin ilgili film, dizi, belgesel gibi yapımlar hakkında yorum yapabildiği IMDB platformundan kullanıcı yorumları veri seti olarak alınmıştır. Tablo 3.1.' de görüleceği üzere 2700 pozitif ve 2300 negatif olmak üzere 5000 yorum bulunmaktadır. (Andrew L. Maas 2011)

Tablo 3.1. IMDB alt veri kümesi

	Yorum Sayısı
Pozitif yorumlar	2700
Negatif yorumlar	2300
Toplam yorum sayısı	5000

Günümüzde çoğu kullanıcının belirli bir konuda fikirlerini paylaştığı sosyal medya platformu olan Twitter en önemli veri kaynaklarından bir tanesidir. Çalışmamızda Kaggle üzerinden çektiğimiz, pozitif, negatif ve nötr duyguları temsil eden gönderiler veri seti olarak kullanılmıştır. Tablo 3.2.' de görüleceği üzere 1100 pozitif, 1950 nötr ve 1950 negatif olmak üzere 5000 gönderi bulunmaktadır. (Chaitanya Kaushal 2020)

Tablo 3.2. Twitter alt veri kümesi

	Yorum Sayısı
Pozitif gönderiler	1100
Nötr gönderiler	1950
Negatif gönderiler	1950
Toplam gönderi sayısı	5000

Diğer veri setimiz olan GOP veri seti Amerika başkanlık seçimi öncesi cumhuriyetçi adayların kendi arasındaki seçimi ile alakalı Twitter üzerinden yapılan gönderileri içermektedir. Tablo 3.3.' de görüleceği üzere 2236 pozitif, 3142 nötr ve 8493 negatif olmak üzere toplam 13871 gönderi içermektedir.(Ankit 2018)

Tablo 3.3. GOP veri kümesi

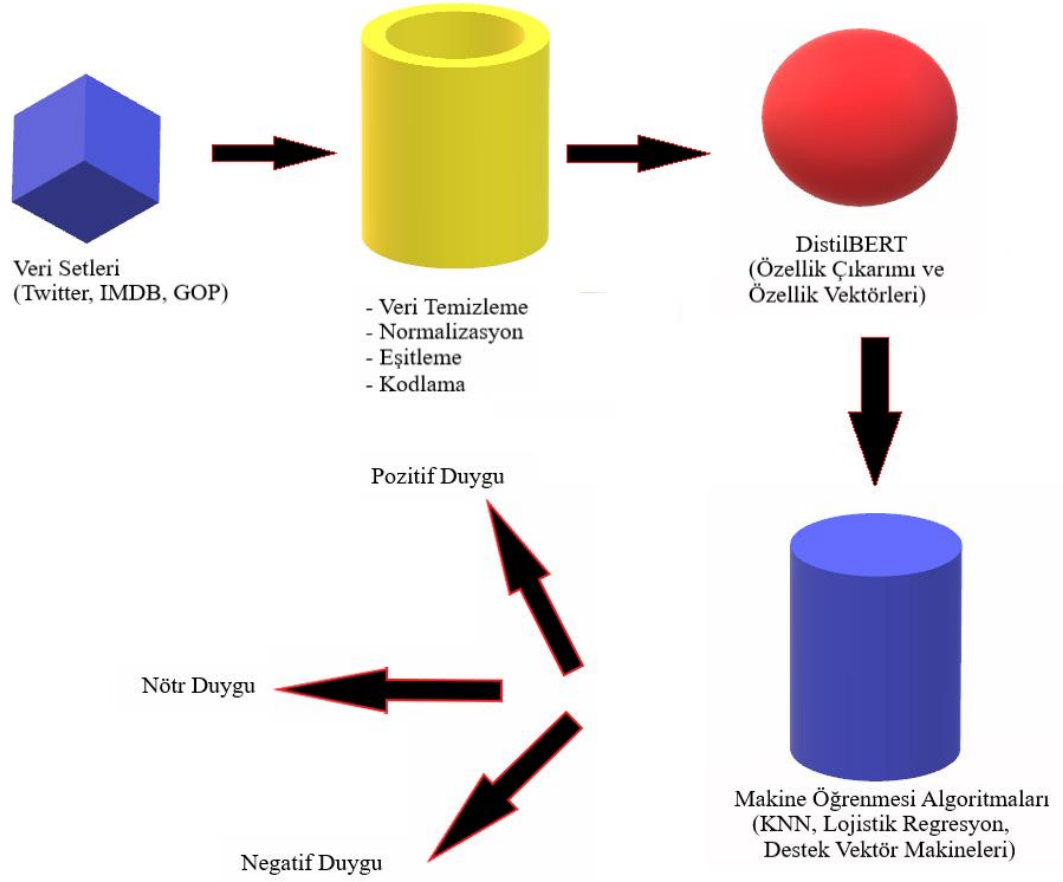
	Yorum Sayısı
Pozitif gönderiler	2236
Nötr gönderiler	3142
Negatif gönderiler	8493
Toplam gödneri sayısı	13871

Çalışmamızda kullanılan veri kümelerinin modellere verilmeden önce bazı ön işleme adımlarından geçmesi gerekmektedir. İlgili adımlar şu şekilde sıralanabilir:

- İngilizce harfler dışındaki karakterler çıkartılır.
- Html etiketleri kaldırılır.
- Noktalama işaretleri silinir.
- Rakamlar kaldırılır.
- Tekli karakterler silinir.
- Birden fazla boşluk kaldırılır.

### 3.2.3. Dönüştürücü ve Makine Öğrenmesi Modelleri ile Duygu Analizi

Bu kısımda dönüştürücü modeli olarak kullanılacak olan DistilBERT modeli ile makine öğrenmesi algoritmalarının sınıflandırma sonuçları ele alınacaktır. Birden fazla ML algoritması başarı durumlarının kıyaslanması için kullanılmıştır. Daha önceden de bahsedildiği üzere duygu analizi bir çok aşamadan oluşmaktadır. “Pandas” kütüphanesi ile sunuma hazır olan veri setleri ilk olarak ön işleme tabi tutulmuştur. Bu aşamada, cümleler token dediğimiz yapılarak dönüştürülmüştür. Veri temizleme ve normalizasyon aşamasında, noktalama işaretleri, html etiketleri silinmiş ve emoji, çoklu boşluk ve rakamlar kaldırılmıştır. Token yapılarına ayrılmış kelimelerden oluşan vektörler padding dediğimiz doldurma yöntemiyle eşitlenmiştir. Son adımda ise, kodlama yapısıyla kelime(token) vektörleri sayısal değerler ile gösterilip özellik çıkarımı aşamasına hazır hale gelmiştir. (Yomna Eid Rizk 2021). Modelin çalışma şekli Şekil 3.17.’ de gösterilmiştir.

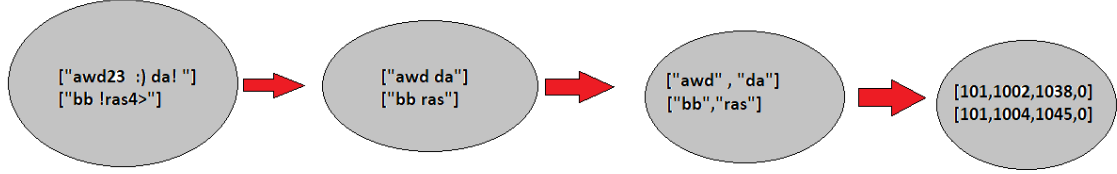


Şekil 3.17. Dönüştürücü ve Makine Öğrenmesi Model Mimarisi

### 3.2.4. Dönüştürücü ve Derin Öğrenme Modelleri ile Duygu Analizi

Derin öğrenme modeli ile veri setlerimizin analizini yapmadan önce makine öğrenmesi modelinde olduğu gibi ön işleme adımları uygulanmaktadır. İlgili metin dosyası programa aktarıldıktan sonra sırasıyla metinlerdeki sayı, özel karakter, emoji, sayfa etiketleri, çoklu boşluklar temizlenmektedir. Token oluşturma safhasında ilgili vektörler kelimelerden oluşan token vektörlere dönüşmektedir. Token vektörler dönüştürücü modele verilip özellik çıkarımı yapmadan önce istenilen formata dönüştürülmektedir. İlgili tokenler kodlama(encoding) yapısı ile sayısal verilere dönüştürülmektedir.

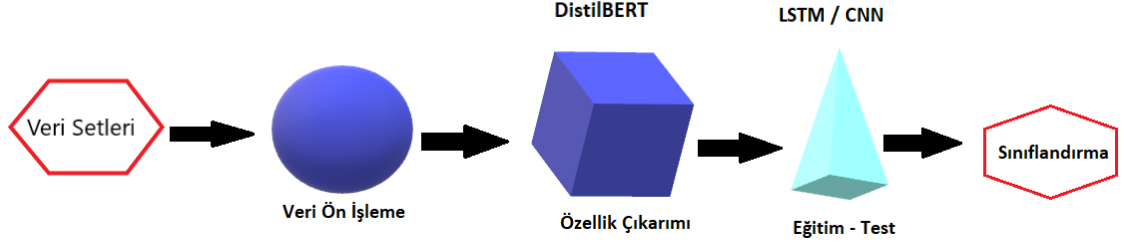
Bu adımda aynı zamanda başlangıç token olduğunu belirten [SEP] ve [CLS] tokenleri başa ve sona yerleştirilmektedir.(Jahanur Biswas, 2021) Bu adımdan sonra da boyut problemini aşmak için dolgu(padding) ile eşit uzunlukta token vektörleri oluşturulmaktadır. İlgili işlem adımları şekil 3.18’ de gösterilmektedir.



Şekil 3.18. Veri Önleme Aşaması

BERT ince ayar süreci oldukça basittir. İnce ayar bölümündeki görevleri yapmak için, eğitim öncesi kısımdan elde edilen gizli [CLS] vektörü, metnin duyarlılığını sınıflandırmak için tamamen bağlantılı bir katmandan geçirilir. Bu safhada, ilk olarak, kelime havuzunu oluşturmak için BERT modeline vermeden önceki önleme adımında tokenizer kullanılarak veri setinin tüm cümleleri tokenize edilir. Bundan sonra tüm tokenler önceden eğitilmiş bir BERT modeline geçirilir. BERT, ön eğitim kısmı için çok büyük bir kelime havuzu gerektirdiğinden ve ön eğitim için yeterli derlemeye sahip olmak ve İngilizce dili için eğitilmiş DistilBERT modelini kullanılır. İnce ayar kısmı yapılırken doğrulama kaybına dayalı olarak en iyi ağırlıklar korunmaktadır. BERT modelinin mimarisi 12 dikkat başlığı, 12 dönüştürücü bloğu içeren n kodlayıcı ve 768 gizli durum vektörü içerir.(Jahanur Biswas, 2021)

Çalışmamızda veri kaynaklarından elde ettiğimiz veri setleri ön işleme adımından geçtikten sonra özellik çıkarımı için DistilBERT modeline verilmektedir. Önceden eğitilmiş kelime korpusu kullanılarak özellik vektörleri çıkarılmaktadır. Daha sonra da sınıflandırma işlemi için kullanılacak olan LSTM ve CNN derin öğrenme modellerine girdi olarak verilmektedir. Modelimizin yapısı Şekil 3.19’ da gösterilmiştir.



Şekil 3.19. DistilBERT ve Derin Öğrenme Model Mimarisi

### 3.2.5. Performans Ölçütleri

Çalışmada istatistiksel analizler farklı metriklerin performansını ölçmek için kullanılmaktadır. Sınıflandırma performansını ölçen metrikler; doğruluk, kesinlik, hassasiyet ve F1-skor’ dur. Sınıflandırma sonucunda tahminlerin doğruluğu karmaşıklık matrisi ile verilmektedir. (Asraf Ali, 2020)

TP : Pozitif etiketli duygunun pozitif olarak tahmin edilmesidir.

TN : Negatif etiketli duygunun negatif olarak tahmin edilmesidir.

FP : Negatif etiketli duygunun pozitif olarak tahmin edilmesidir.

FN : Pozitif etiketli duygunun negatif olarak tahmin edilmesidir.

Bir modelin başarısını ve performansını tespit edebilmek için çeşitli metrikler kullanılmaktadır. Doğruluk metriği doğru tahmin edilen örneklerin bütün örneklem uzayına olan oranını göstermektedir. (Tshephisho Joseph Sefara, 2019)

$$\text{Doğruluk (Accuracy)} = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \quad (13)$$

$$\text{Kesinlik (Precision)} = \frac{t_p}{t_p + f_p} \quad (14)$$

$$\text{Hassasiyet (Recall)} = \frac{t_p}{t_p + f_n} \quad (15)$$

$$\text{F1 – Skor (F1 Score)} = 2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

#### 4. BULGULAR ve TARTIŞMA

“IMDB kullanıcı yorumları”, “Twitter gönderileri” ve “GOP Amerika Başkanlık Öncesi Seçimi” verisetleri çalışmamızda veri kaynakları olarak kullanılmıştır. Bu çalışmada ilk olarak ilgili verisetleri makine öğrenmesi algoritmalarıyla eğitilip test edilmiştir. Dönüştürücü kullanılmadan yapılmış olan özellik çıkarımı operasyonu için “Count Vectorizer” kütüphanesi kullanılmış olup mevcut verilerin %80’ i eğitim için %20’si de test için kullanılmıştır. Sonraki aşamada ise aynı verisetleri dönüştürücü ve makine öğrenmesi modellerinin birlikte kullanımı ile test edilmiştir. Sonuçlar bir sonraki kısımda verilecek ön bilgi olması açısından genel performansa bakacak olursak: IMDB veri seti için dönüştürücü kullanılmadan uygulanan lojistik regresyon algoritmasında 5000 iterasyon yapılmıştır ve %79.16 doğruluk elde edilmiştir. DistilBERT ile öznitelikleri çıkarılmış ve öğrenme algoritmasına verilmiş olan veriler üzerinde ise aynı parametreler ile LR algoritmasında %85.83 doğruluk elde edilmiştir. IMDB veri setini dönüştürücü olmadan destek vektör makineleri algoritması ile c düzenleme parametresini 0.1 olarak verdiğimizde %78.33 doğruluk değeri elde edilmiştir. IMDB veri seti SVM ve DistilBERT modelleri ile kullanıldığında %85.41 doğruluk değeri edilmektedir.

GOP veri setinin “Count Vectorizer” kütüphanesi ile öznitelik çıkarımı yapıldıktan sonra IMDB veri setlerinde olduğu gibi %80’i eğitim ve %20’si test için kullanılmıştır. Dönüştürücü modeli kullanılmadan uygulanan LR modeli ile %81.51 doğruluk elde edilirken DistilBERT dönüştürücü modeli ile öznitelikleri çıkarılan veri setinin aynı model ile %86.17 doğruluk oranı edinilmiştir.(Hüseyin İLGÜN, Erdal KILIÇ, 2021)

Çalışmamızın temeli olan dönüştürücü ve derin öğrenme modelleri ile duygu analizi, DistilBERT ön eğitilmiş ve ince ayarlı(fine-tune) model ve bu modelin çıktısı olarak gelen özellik vektörü LSTM ve CNN derin öğrenme modelleri ile gerçekleştirilmektedir. Dönüştürücü modeli olarak DistilBERT modeli ve bu modelin kendi ayrıştırıcı(tokenizer) yapısı kullanılmıştır. Çıktı gizli katmanları verildikten sonra bu katmanlardaki özniteliklerin ağırlıkları kelimelerin cümle içerisindeki durumları hakkında bilgi vermektedir. Maksimum kelime uzunluğu performans ve bellek problemlerinden dolayı 250 seçilmiştir. Kısa cümlelerde ise bu uzunluk değiştirilmektedir. GOP veriseti için 13871 eğitim 400 test verisi kullanılmıştır. IMDB veriseti içerisinde 5000 eğitim 200

test verisi kullanılmıştır. Twitter gönderilerinin oluşturduğu verisetinde ise 5000 eğitim 400 test verisi kullanılmıştır. Özellik vektörlerini tensör haline getirmeden önce grup boyutu 32 olarak belirlenmiştir.

Performans konusunda bilgi amaçlı olarak doğruluk(acc.) metriği kullanılmıştır. Sonuçların grafik gösterimi için “matplotlib” kütüphanesi, karmaşıklık matrisi ve sınırlandırma raporu (confusion matrix) için “sklearn” kütüphanesi kullanılmıştır. Isı haritası kullanarak da doğru yanlış tahmin sayıları belirtilmiştir.

#### 4.1. Makine Öğrenmesi ve DistilBERT Modeli Sonuçları

İlk olarak IMDB veri seti kullanılmıştır. 5000 veriden oluşan setimiz 2700 pozitif etiketli ve 2300 negatif etiketli olmak üzere kullanıcı yorumlardan oluşmaktadır. %80’ i eğitim %20 test için kullanılmıştır. Lojistik regresyon için 5000 iterasyon yapılmış olup %85 F1-Skor elde edilmiştir. K en yakın komşu modelinde K 3 alınıp %25.3 F1-Skor bulunmuştur aynı model DistilBERT ile kullanıldığında ise %55’ lik bir başarı elde edilmiştir. Modellere dair sonuçlar tablo 4.1.’ de verilmiştir.

Tablo 4.1. IMDB veriseti sonuçları

Teknikler	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR (Teguh 2022)	0.67	0.61	0.82	0.715
SVM (Teguh 2022)	0.79	0.75	0.87	0.81
KNN	0.47	0.171	0.50	0.253
BERT+LR	0.8583	0.86	0.8501	0.85
BERT+SVM	0.8541	0.8401	0.8511	0.8502
BERT+KNN	0.6916	0.39	0.94	0.55

İkincil olarak kullanılan Twitter veriseti, kullanıcıların belirli ve ya belirsiz başlıklar hakkında yorumlarını içeren gönderilerden oluşmaktadır. Verisetimiz 1100 pozitif, 1950 nötr ve 1950 negatif olmak üzere 5000 verimiz bulunmaktadır. Eğitim için %80 test için ise toplam verisetinin %20’ si kullanılmıştır. 5000 iterasyon ile modelimiz eğitilmiş ve test edilmiştir. Sonuçlar ise sınıflandırma raporu ve matrisi ile gösterilmiştir. Lojistik regresyon modeli kullanılarak %68 hassasiyet elde edilmiştir. Aynı model dönüştürücü kullanılmadan %60 hassasiyet vermektedir. Destek vektör makineleri ile yakınsama değeri 1 seçilip %64 hassasiyet elde edilmiştir. Sonuçlar tablo 4.2’ de verilmiştir.

Tablo 4.2. Twitter veriseti sonuçları

Teknikler	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR	0.612	0.73	0.60	0.665
SVM	0.607	0.72	0.62	0.67
KNN	0.473	0.81	0.16	0.485
BERT+LR	0.658	0.72	0.68	0.7
BERT+SVM	0.637	0.72	0.64	0.68
BERT+KNN	0.483	0.63	0.36	0.495

Son olarak GOP veri seti 2236 pozitif, 3142 nötr ve 8493 negatif olmak üzere 13871 kullanıcı gönderilerden oluşmaktadır. Geleneksel makine öğrenmesi modellerine göre çok daha iyi sonuçlar alınmıştır. Modelin başarısı tablo 4.3' te verilmiştir.

Tablo 4.3. GOP veriseti sonuçları

Teknikler	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR (Ankit 2018)	0.8151	0.9901	0.814	0.8935
SVM (Ankit 2018)	0.8344	0.8976	0.8916	0.8946
BERT+LR	0.8617	0.963	0.8812	0.921
BERT+SVM	0.86	0.9409	0.8912	0.91

#### 4.2. Derin Öğrenme ve Dönüştürücü Modeli Sonuçları

Çalışmamızda hem derin öğrenme modelleri hem de dönüştürücü modellerinin kullanıldığı bir model oluşturulmuştur. Bu bölümde dönüştürücü modeli olarak kullanılan DistilBERT ve derin öğrenme modelleri olan CNN ve LSTM modellerinin başarı oranları ve sonuçları gösterilmektedir. Sonuçların epok(epoch) sayısına göre doğruluk(accuracy) ve kayıp(loss) öğrenme eğrileri(learning curve); eğitim(train), test verilerine göre grafiksel olarak gösterilmiştir.

#### 4.2.1. LSTM ve DistilBERT Model Sonuçları

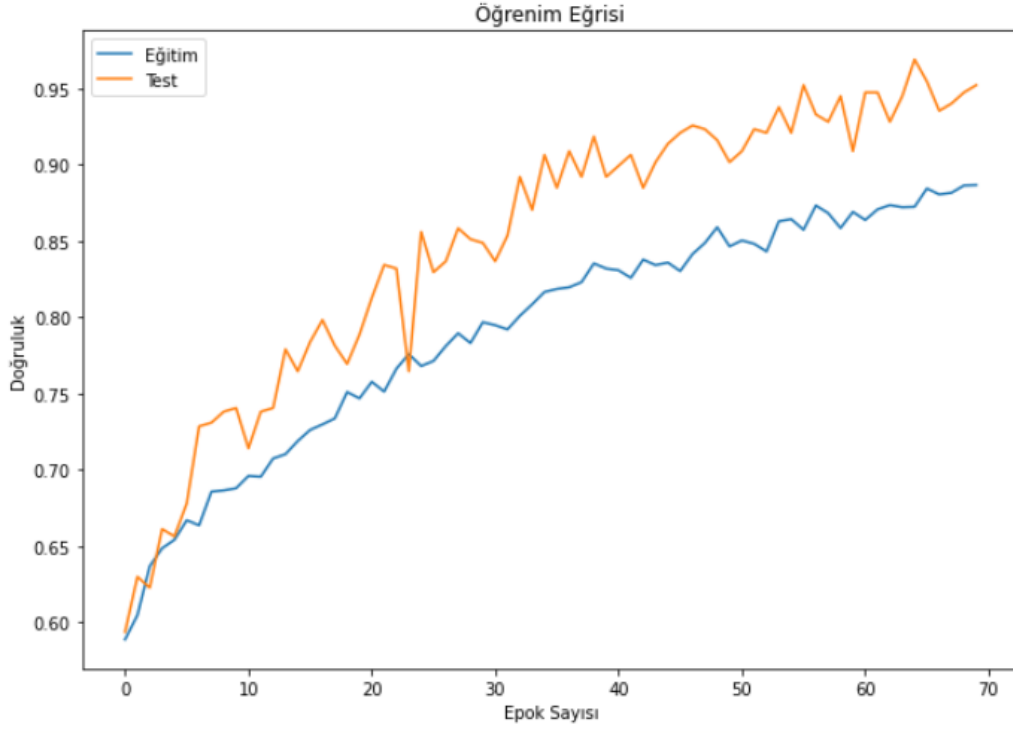
LSTM ve DistilBERT ile oluşturulmuş olan modelde parti(batch) boyutu 32 seçilmiş ve kullanılan eğitim ve test verisetleri karıştırılmıştır. LSTM giriş katmanı ve yoğunluk(dense) gizli katmanı 64 nörondan oluşmaktadır. Ara yoğunluk katmanında “relu” aktivasyon fonksiyonu sınıflandırma katmanında ise “softmax” aktivasyon fonksiyonu kullanılmıştır. Aşırı öğrenmeyi engellemek için Dropout 0.1 olarak belirlenmiştir. Modeli oluştururken “Categorical Crossentropy” fonksiyonu kullanılmış ve yerel minimum noktalarından kaçınmak için optimize edici olarak “Adam” yapısı kullanılmıştır.

İlk veriseti olan GOP veriseti için DistilBERT kelime uzunluğu 200 belirlenmiştir. Epok değeri 70 olarak verilmiştir. Bu veriseti için 13871 eğitim ve 400 test verisi baz alınmıştır. Eğitim ve test süresi 8 saat 40 dakika sürmüştür. Negatif veriler için 0.96 f1-skor değeri, nötr veriler için 0.91 f1-skor değeri ve de pozitif veriler için 0.97 f1-skor değeri bulunmuştur. Tablo 4.4’ te GOP verisetinin LSTM modeli performans metrikleri gösterilmektedir.

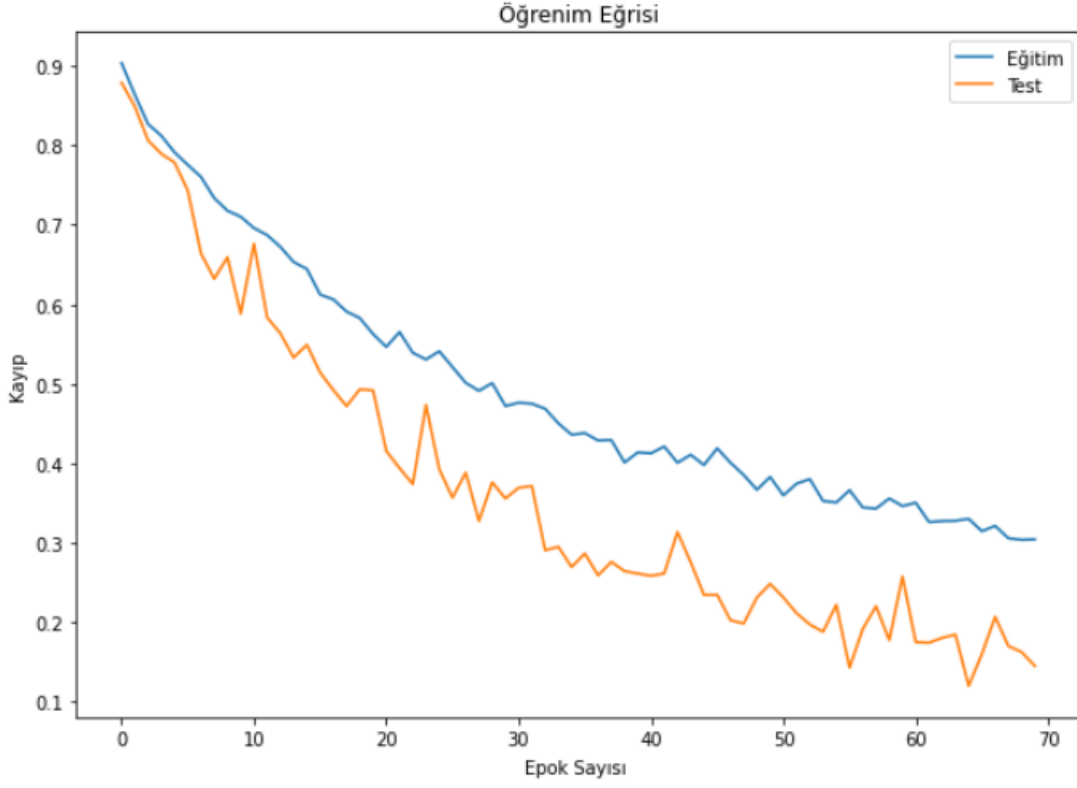
Tablo 4.4. GOP Veriseti LSTM Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	0.94	0.98	0.96	221
1 (Nötr)	0.97	0.86	0.91	102
2 (Pozitif)	0.96	0.97	0.97	77
Doğruluk			0.95	400
Makro Ortalama	0.96	0.94	0.95	400
Ağırlıklık Ortalama	0.95	0.95	0.95	400

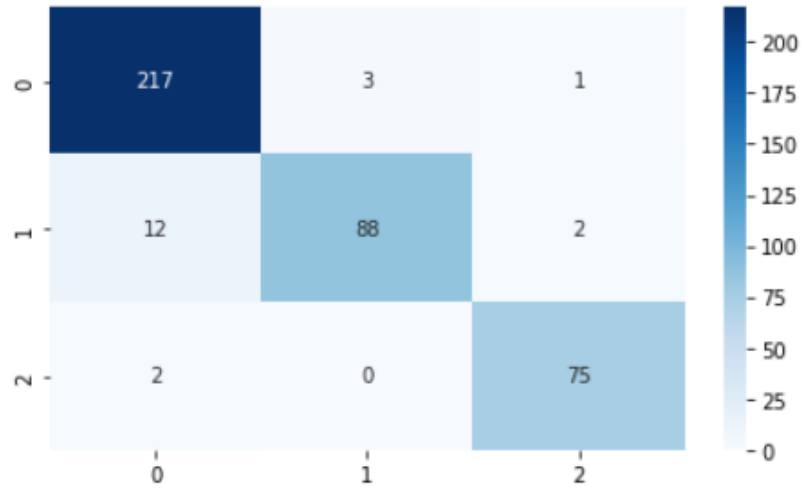
Şekil 4.1’ de GOP verisetinin LSTM modeli doğruluk grafiği gösterilmektedir. Şekil 4.2’ de GOP verisetinin LSTM modeli kayıp grafiği gösterilmektedir. Şekil 4.3’ te ise GOP veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.1. GOP Veriseti LSTM Eğitim ve Test Doğruluğu



Şekil 4.2. GOP Veriseti LSTM Eğitim ve Test Kaybı



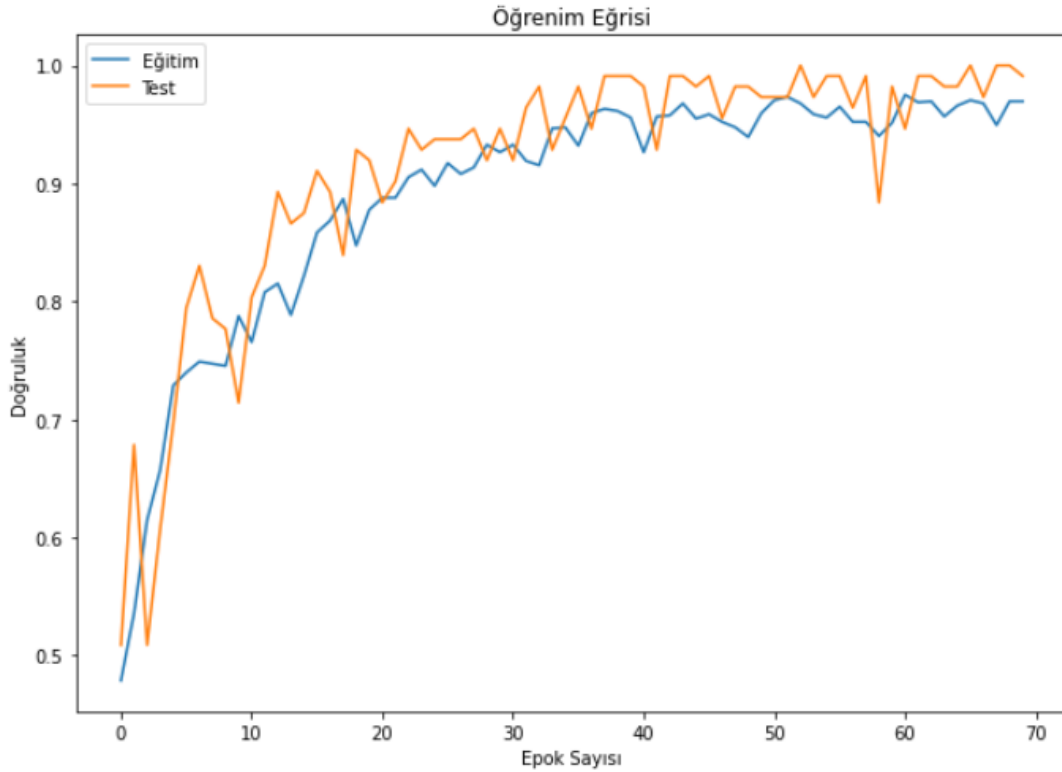
Şekil 4.3. GOP Veriseti LSTM Karmaşıklık Matrisi

IMDB veriseti için DistilBERT kelime uzunluğu 200 verilmiştir. Epok sayısı 70 verilmiş olup toplam eğitim ve test süresi 7 saat 30 dakika sürmüştür. 5000 eğitim ve 400 test verisi kullanılmıştır. Pozitif etiketli veriler için f1-skor 0.99 negatif veriler için ise 0.99 f1-skor değeri elde edilmiştir. Tablo 4.5’ te IMDB verisetinin LSTM modeli performans metrikleri gösterilmektedir.

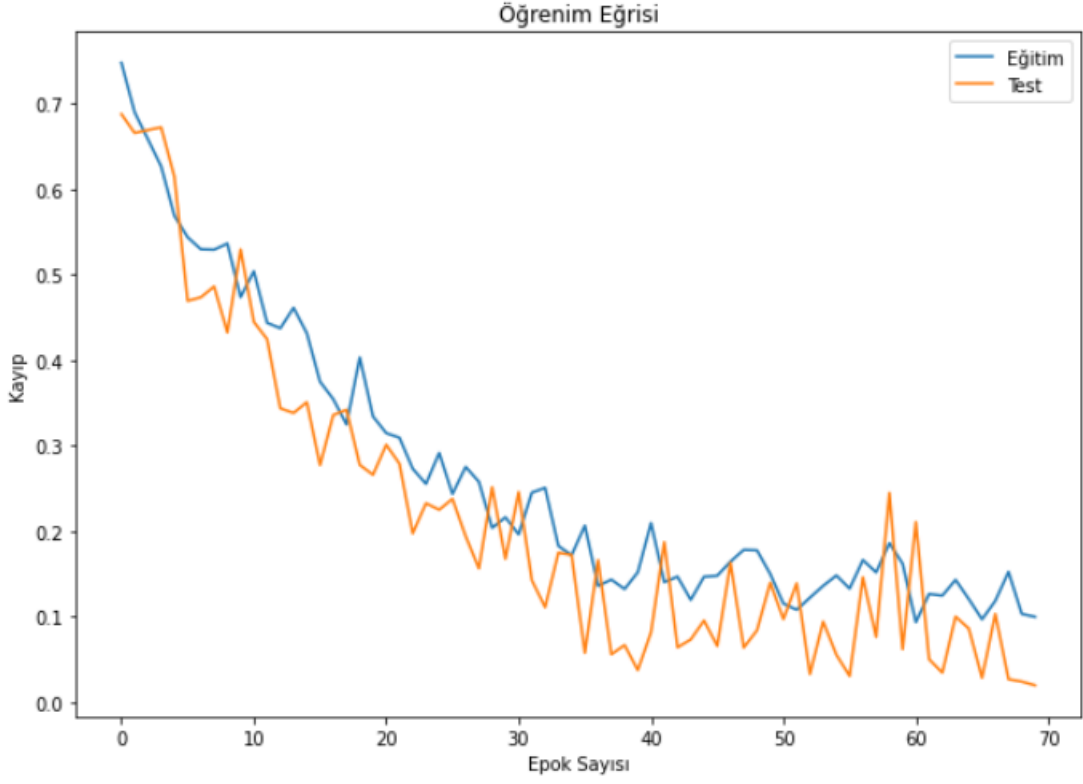
Tablo 4.5 IMDB Veriseti LSTM Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	0.99	0.99	0.99	210
1 (Pozitif)	0.99	0.98	0.99	190
Doğruluk			0.99	400
Makro Ortalama	0.99	0.99	0.99	400
Ağırlıklık Ortalama	0.99	0.99	0.99	400

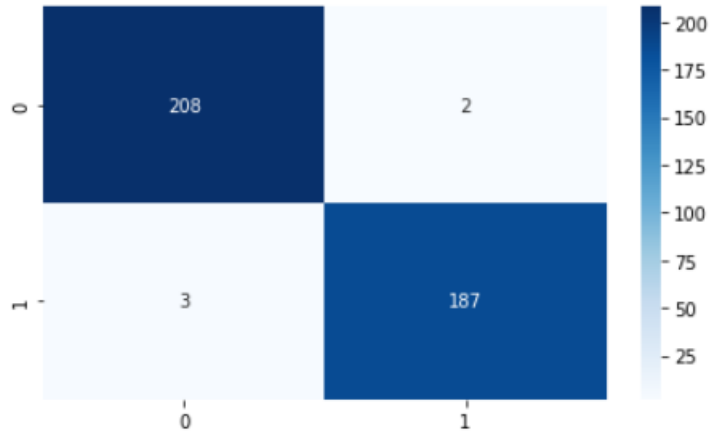
Şekil 4.4’ te IMDB verisetinin LSTM modeli doğruluk grafiği gösterilmektedir. Şekil 4.5’ te IMDB verisetinin LSTM modeli kayıp grafiği gösterilmektedir. Şekil 4.6’ da ise IMDB veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.4. IMDB Veriseti LSTM Eğitim ve Test Doğruluğu



Şekil 4.5. IMDB Veriseti LSTM Eğitim ve Test Kaybı



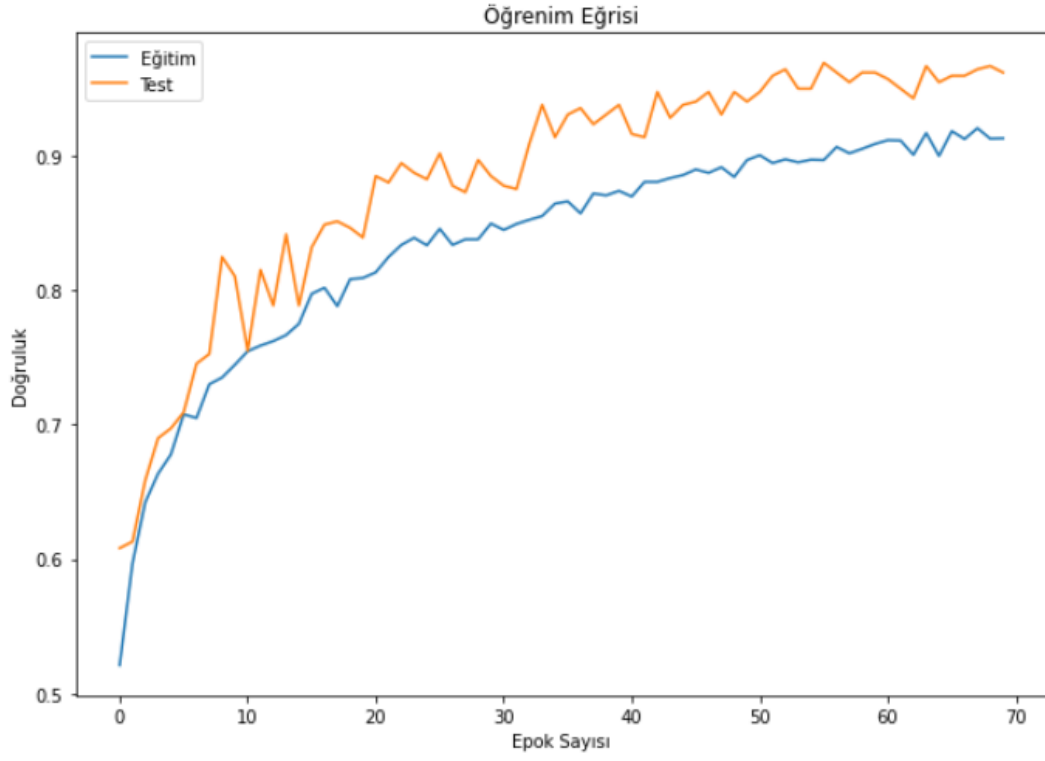
Şekil 4.6. IMDB Veriseti LSTM Karmaşıklık Matrisi

Twitter veriseti için DistilBERT kelime uzunluğu 50 verilmiştir. Epok sayısı 70 verilmiş olup toplam eğitim ve test süresi 7 saat 15 dakika sürmüştür. 5000 eğitim ve 400 test verisi kullanılmıştır. Pozitif etiketli veriler için f1-skor 0.96, nötr veriler için 0.96 ve negatif veriler için ise 0.98 f1-skor değeri elde edilmiştir. Tablo 4.6’ da Twitter verisetinin LSTM modeli performans metrikleri gösterilmektedir.

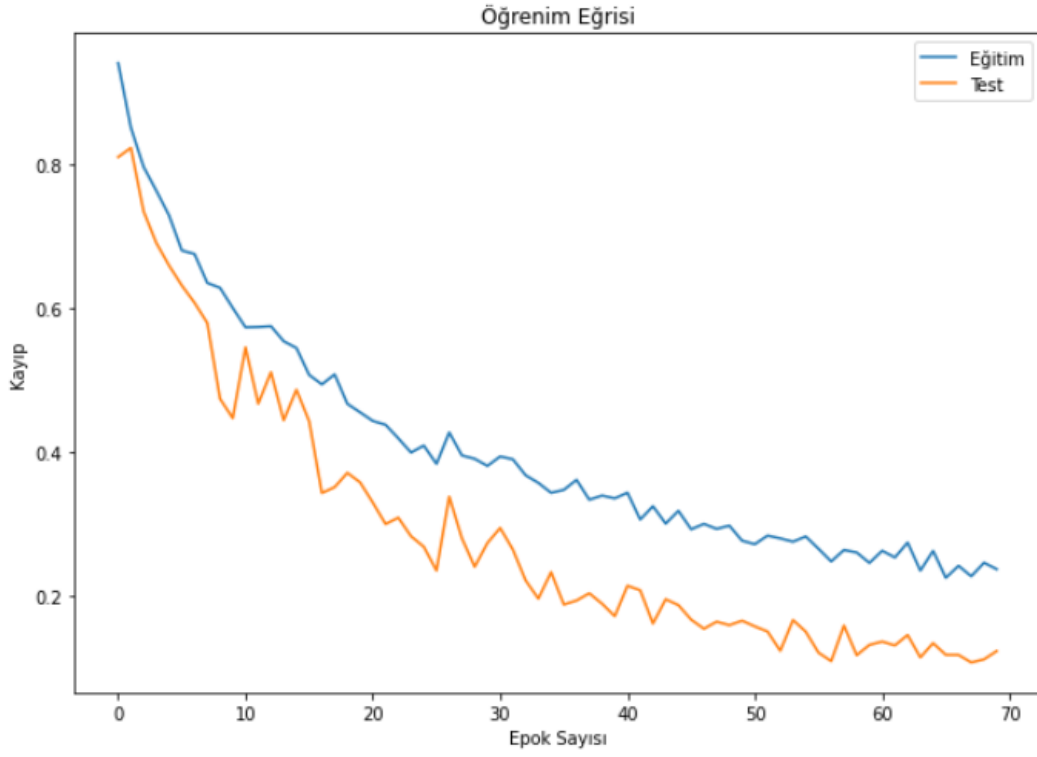
Tablo 4.6. Twitter Veriseti LSTM Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	0.96	0.99	0.98	125
1 (Nötr)	0.94	0.97	0.96	155
2 (Pozitif)	1.00	0.92	0.96	120
Doğruluk			0.96	400
Makro Ortalama	0.97	0.96	0.96	400
Ağırlıklık Ortalama	0.96	0.96	0.96	400

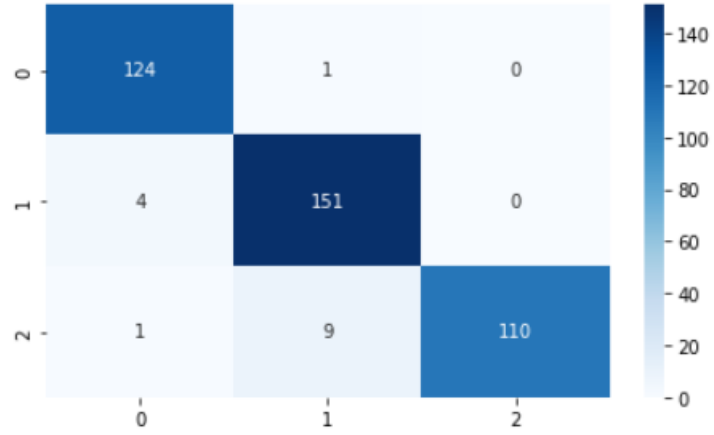
Şekil 4.7’ te Twitter verisetinin LSTM modeli doğruluk grafiği gösterilmektedir. Şekil 4.8’ te Twitter verisetinin LSTM modeli kayıp grafiği gösterilmektedir. Şekil 4.9’ ta ise Twitter veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.7. Twitter Veriseti LSTM Eğitim ve Test Doğruluğu



Şekil 4.8. Twitter Veriseti LSTM Eğitim ve Test Kaybı



Şekil 4.9. Twitter Veriseti LSTM Karmaşıklık Matrisi

#### 4.2.2. CNN ve DistilBERT Model Sonuçları

CNN modeli, görüntünün piksellerini ve renk kanallarını temsil eden 2 boyutlu bir girişinin kabul edildiği görüntü sınıflandırması için geliştirilmiştir. Aynı işlem 1 boyutlu veri dizilerine ve metin verilerine de CNN1D modeli sayesinde yapılabilmektedir. Bu çalışmada CNN1D modeli kullanılarak metin verisetlerimizde sınıflandırma işlemi yapılmaktadır. Genel işleyişi tekrar özetleyecek olursak; evrişim katmanında kullanılacak kelime sayısı ve özellikleri çıkartıldıktan sonra havuz (max pooling) katmanında çıkan veriler konsolide edilmektedir. Çıktı katmanında ise belirli bir aktivasyon kullanılarak sınıflandırma işlemi gerçekleştirilmektedir.

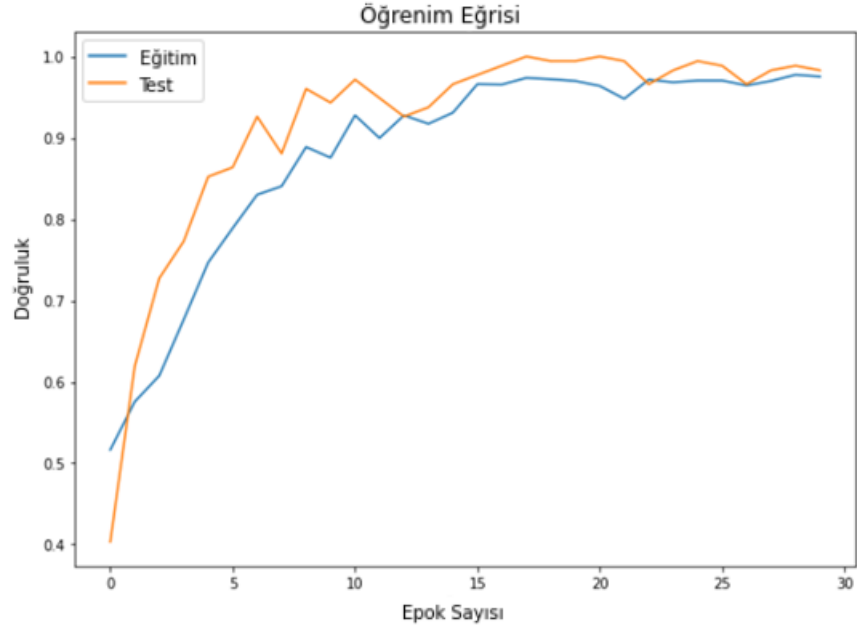
CNN modeline girdi olarak verilecek olan öznitelik vektörü DistilBERT modelinden gelmektedir. Gömülü (embedded) vektör olarak tanımlanan öznitelik vektörü CNN modelinin kelime boyutlarını ve eğitilecek ve test edilecek modelin mimarisi çıkartmaktadır. CNN modeli evrişim katmanı 64 nörondan oluşmaktadır. Çekirdek boyutu 5 verilmiştir. Ara katman kullanılmayan modelde havuz katmanı 2 boyutlu olarak oluşturulmuştur. Yoğunluk katmanı olarak tanımlanan sınıflandırma katmanında sigmoid aktivasyon fonksiyonu kullanılmış olup softmax aktivasyon fonksiyonundan farklı olarak olasılıksal dağılım yerine 0 ve ya 1 değerlerine göre sınıflandırma sonucu hesaplanmıştır.

İlk veriseti olan GOP veriseti için DistilBERT kelime uzuluğu 200 belirlenmiştir. Epok değeri 30 olarak verilmiştir. Bu veriseti için 13871 eğitim ve 400 test verisi baz alınmıştır. Eğitim ve test süresi 5 saat 40 dakika sürmüştür. Negatif veriler için 0.99 f1-skor değeri, nötr veriler için 0.99 f1-skor değeri ve de pozitif veriler için 0.98 f1-skor değeri bulunmuştur. Tablo 4.7' de GOP verisetinin CNN modeli performans metrikleri gösterilmektedir.

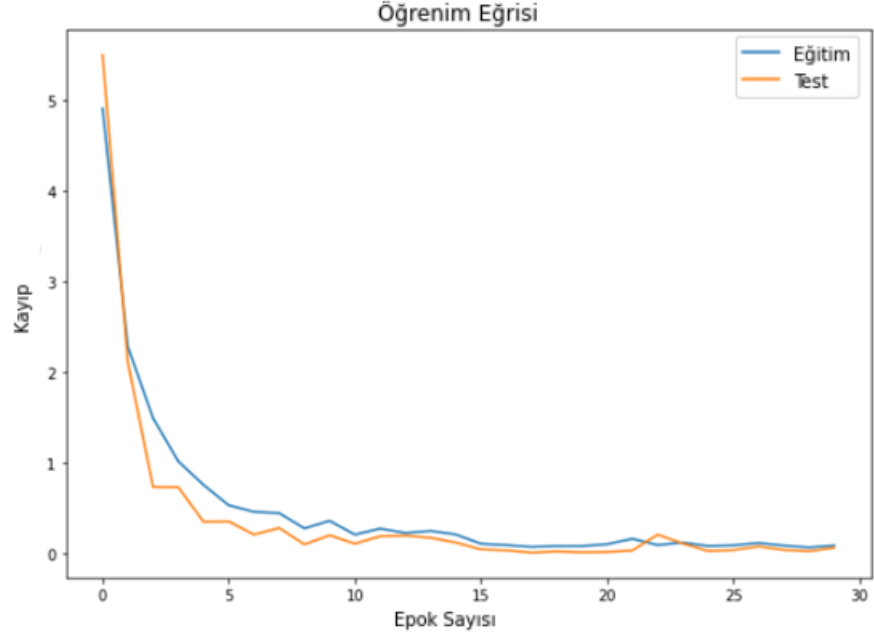
Tablo 4.7. GOP Veriseti CNN Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	0.99	0.99	0.99	231
1 (Nötr)	1.00	0.99	0.99	91
2 (Pozitif)	0.96	1.00	0.98	78
Doğruluk			0.99	400
Makro Ortalama	0.99	0.99	0.99	400
Ağırlıklık Ortalama	0.99	0.99	0.99	400

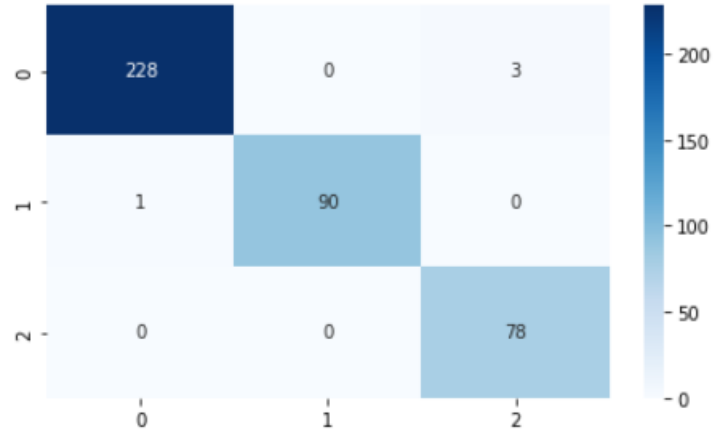
Şekil 4.10' da GOP verisetinin CNN modeli doğruluk grafiği gösterilmektedir. Şekil 4.11' de GOP verisetinin CNN modeli kayıp grafiği gösterilmektedir. Şekil 4.12' te ise GOP veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.10. GOP Veriseti CNN Eğitim ve Test Doğruluğu



Şekil 4.11. GOP Veriseti CNN Eğitim ve Test Kaybı



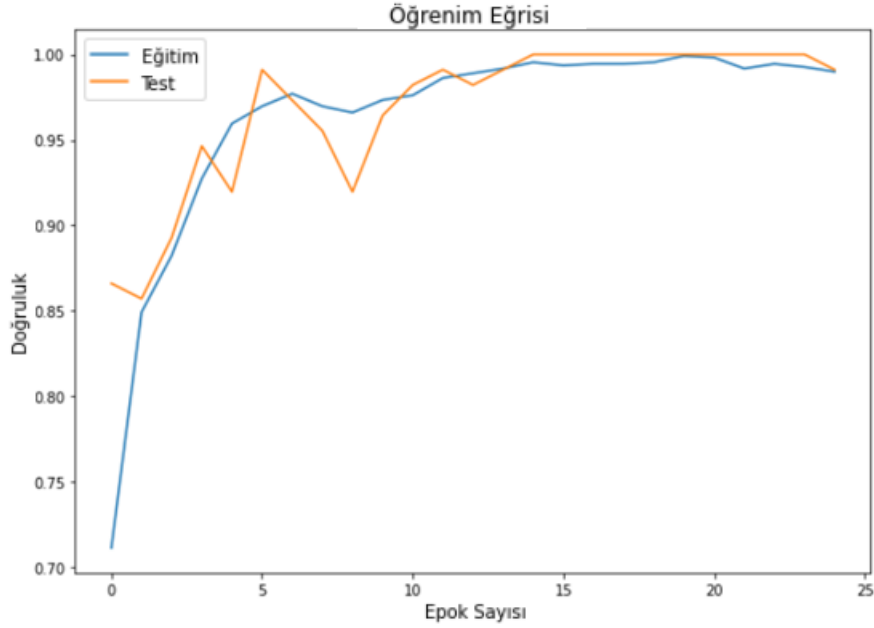
Şekil 4.12. GOP Veriseti CNN Karmaşıklık Matrisi

IMDB veriseti için DistilBERT kelime uzunluğu 200 verilmiştir. Epok sayısı 25 verilmiş olup 5000 eğitim ve 200 test verisi kullanılmıştır. Tablo 4.8’ te IMDB verisetinin CNN modeli performans metrikleri gösterilmektedir.

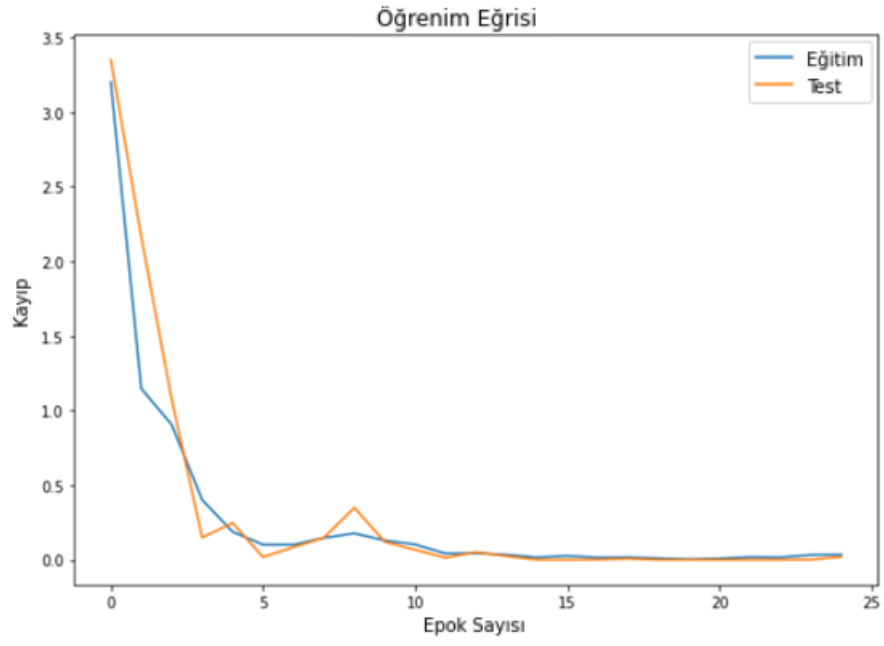
Tablo 4.8. IMDB Veriseti CNN Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	1.00	0.96	0.98	115
1 (Pozitif)	0.94	1.00	0.97	85
Doğruluk			0.97	200
Makro Ortalama	0.97	0.98	0.97	200
Ağırlıklık Ortalama	0.98	0.97	0.98	200

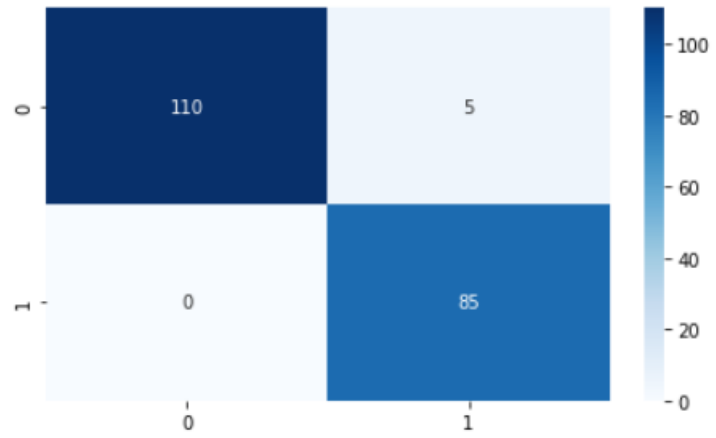
Şekil 4.13’ te IMDB verisetinin CNN modeli doğruluk grafiği gösterilmektedir. Şekil 4.14’ te IMDB verisetinin CNN modeli kayıp grafiği gösterilmektedir. Şekil 4.15’ te ise IMDB veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.13. IMDB Veriseti CNN Eğitim ve Test Doğruluğu



Şekil 4.14. IMDB Veriseti CNN Eğitim ve Test Kaybı



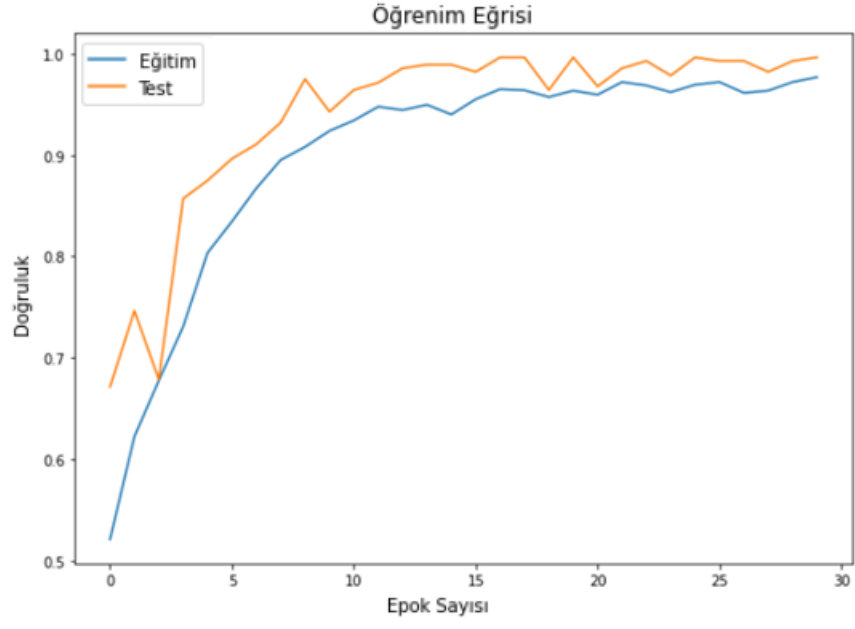
Şekil 4.15. IMDB Veriseti CNN Karmaşıklık Matrisi

Twitter veriseti için DistilBERT kelime uzunluğu 50 verilmiştir. Epok sayısı 30 verilmiş olup 5000 eğitim ve 400 test verisi kullanılmıştır. Tablo 4.9’ da Twitter verisetinin CNN modeli performans metrikleri gösterilmektedir.

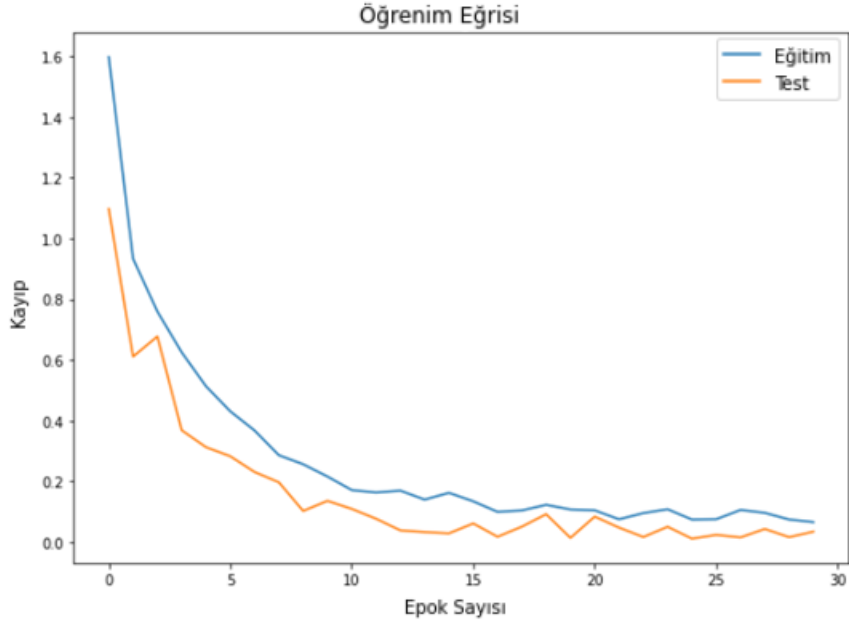
Tablo 4.9 Twitter Veriseti CNN Performans Metrikleri

Tanım	Kesinlik	Hassasiyet	F1-Skor	Destek
0 (Negatif)	0.98	1.00	0.99	120
1 (Nötr)	1.00	0.99	1.00	157
2 (Pozitif)	1.00	0.99	0.98	123
Doğruluk			0.99	400
Makro Ortalama	0.99	0.99	0.98	400
Ağırlıklık Ortalama	0.98	0.99	0.99	400

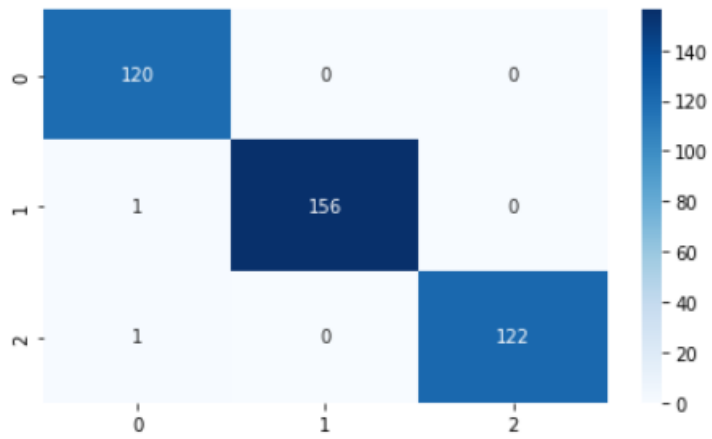
Şekil 4.16’ da Twitter verisetinin CNN modeli doğruluk grafiği gösterilmektedir. Şekil 4.17’ de Twitter verisetinin CNN modeli kayıp grafiği gösterilmektedir. Şekil 4.18’ de ise Twitter veriseti için karmaşıklık matrisi gösterilmektedir.



Şekil 4.16 Twitter Veriseti CNN Eğitim ve Test Doğruluğu



Şekil 4.17 Twitter Veriseti CNN Eğitim ve Test Kaybı



Şekil 4.18. Twitter Veriseti CNN Karmaşıklık Matrisi

## 5. SONUÇ

Bu tez çalışmasında sosyal medya ve kullanıcı gönderilerindeki duygu analizinin tespiti için derin öğrenme modeli ve özellik çıkarımı için de dönüştürücü mimarisini kullanan DistilBERT modeli kullanılmıştır. Keras katman kütüphanesinden LSTM ve CNN derin öğrenme modelleri ile sınıflandırma işlemi yapılmıştır. Çevrimiçi Google Laboratuvarı olarak da bilinen COLAB üzerinden gerçekleştirilmiştir. Sınıflandırma öncesinde denetimli verilerin eğitilmesi ve test edilmesi için dönüştürücü kütüphanesi DistilBERT modeli kullanılmıştır. Bu sayede token haline getirilmiş ve kodlanmış vektörler gizli katman kullanılarak istenen boyutta bilgi ve öznitelik vektörlerini dönüştürülmüştür.

Eğitim aşamasında %90 eğitim ve %10 test verisi olacak şekilde model eğitilmiştir. Bu çalışmada kullanılan verisetlerinin oluşturulan modellere göre performans karşılaştırmaları tablo 5.1, tablo 5.2 ve tablo 5.3' te verilmiştir.

Tablo 5.1. IMDB veriseti için modellerin performans karşılaştırılması

Model	Test kaybı	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR (Teguh 2022)		0.67	0.61	0.82	0.715
SVM (Teguh 2022)		0.79	0.75	0.87	0.81
KNN		0.47	0.171	0.50	0.253
KNN + BERT		0.6916	0.39	0.94	0.55
SVM + BERT		0.8541	0.8401	0.8511	0.8502
LR + BERT		0.8583	0.86	0.8501	0.85
BERT + LSTM	0.125	0.92	0.85	1	0.92
BERT + CNN	0.09	0.97	0.98	0.97	0.97

Tablo 5.2. GOP veriseti için modellerin performans karşılaştırılması

Model	Test kaybı	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR (Ankit 2018)		0.8151	0.9901	0.814	0.8935
SVM (Ankit 2018)		0.8344	0.8976	0.8916	0.8946
SVM + BERT		0.86	0.9409	0.8912	0.91
LR + BERT		0.8617	0.963	0.8812	0.921
BERT + LSTM	0.2	0.92	0.95	0.93	0.94
BERT + CNN	0.125	0.99	0.99	0.98	0.99

Tablo 5.3. Twitter veriseti için modellerin performans karşılaştırılması

Model	Test kaybı	Doğruluk	Kesinlik	Hassasiyet	F1-Skor
LR		0.612	0.73	0.60	0.665
SVM		0.607	0.72	0.62	0.67
KNN		0.473	0.81	0.16	0.485
SVM + BERT		0.637	0.72	0.64	0.68
LR + BERT		0.658	0.72	0.68	0.7
KNN + BERT		0.483	0.63	0.36	0.495
BERT + LSTM	0.28	0.93	0.93	0.93	0.93
BERT + CNN	0.08	0.99	0.98	0.99	0.985

İlgili sonuçlara baktığımızda modelimizin başarı oranının geleneksel yöntemlere göre daha başarılı olduğunu görmekteyiz. Dönüştürücü yapısındaki paralel işleme mekanizması performans olarak mevcut analizleri daha performanslı yapılmasına olanak sağlamaktadır. Çalışmamızda parti boyutu(batch-size) 32 olarak epok sayıları da 70, 30 ve 25 olarak verilmiştir.

Sosyal medya duygu analizi kapsamında oluşturduğumuz model, çeşitli performans metrikleri ve analiz sonuçlarına göre dönüştürücü kullanmayan klasik modellere göre daha doğru sonuçlar vermektedir. Uzun metinler üzerinde işlem yapılırken dönüştürücü tekniği sayesinde oluşturulan olan özellik vektörleri ile kelimelerin cümle içi ilişkilerinin analizinin doğru yapılması sağlanmaktadır. Paralel işlem mekanizması sayesinde performans sorunları ortadan kalkmaktadır.

Duygu analizi kapsamında yapılan çalışmalardan derin öğrenme teknikleri kullanılarak alınan sonuçlar tablo 5.4.' te verilmiştir.

Tablo 5.4. Duygu analizi derin öğrenme modelleri performans karşılaştırması

Referans	Veriseti	Model	Doğruluk	F1-Skor
Manjusha, 2021	SemEval-2017 Task 4 Subtask B	LSTM	0.889	0.879
Nhan Cach Dang, 2020	Cornell Film Yorumları	CNN	0.7136	0.7156
		RNN	0.7669	0.7759
Ammar Mohammed, 2019	Twitter Veriseti	CNN	0.7214	0.7131
		LSTM	0.8149	0.817
Elise De Doncker, 2019	SemEval 2016 Twitter Veriseti	CNN	0.8871	0.88
		LSTM	0.8813	0.86
Paramita Ray, 2019	Twitter Elektronik Ürünler Veriseti	CNN	0.713	0.8056
Sani Kamış, 2019	SemEval2014 Task9-SubTask B	CNN W2V	0.49	0.33
		LSTM W2V	0.51	0.39

Duygu analizi kapsamında yapılan çalışmalardan dönüştürücü modelleri kullanılarak alınan sonuçlar tablo 5.5.' te verilmiştir.

Tablo 5.5. Duygu analizi dönüştürücü modelleri performans karşılaştırması

Referans	Veriseti	Model	Doğruluk	F1-Skor
Wenting Li, 2019		BERT + CNN	0.95	0.93
		BERT + BILSTM	0.94	0.92
Matheus Gomes de Sousa, 2019	Web Sayfaları Makale Veriseti	BERT	0.825	0.725
Luca Bacco, 2021	IMDB Veriseti	ExHIT	0.9235	0.9236
		SCC	0.9351	0.9341
Ting Zhang, 2020	GitHub Veriseti	BERT		0.92
		RoBERTa		0.92
		XLNet		0.92
		AIBERT		0.89
Shivaji Alaparthi, 2020	IMDB Veriseti	BERT	0.9231	0.92

Duygu analizi çalışmalarında başta teknik olmak üzere farklı tiplerde sorunlar yaşanabilme ihtimali vardır. Teknik olarak veri çeşitliliği, veri yetersizliği, yazılım kalitesi ve model yetersizliği başta olmak üzere çeşitli sorunlar ortaya çıkabilmektedir.

Bunun yanı sıra güvenlik, gizlilik ve maliyet gibi teknik olmayan sorunlar da ortaya çıkabilmektedir. Yeni bir çözüm olduğu için henüz birkaç senedir dil işleme alanındaki çalışmalarda kullanılan dönüştürücü modelleri fazla işlem gücü gerektirdiğinde grafik işlemci ve ya merkezi işlemcinin güçlü olduğu makinelerde çok daha fazla veri ile daha doğru sonuçlar verecektir. Bizim modellerimizde zaman sorunundan ve işlem kapasitesi sınırlamasından dolayı belirtilen miktarda veri kullanılmış ve ilgili sonuçlar elde edilmiştir. Literatürde dönüştürücü modeli duygu analizi ve çeviri işlemlerinde yapılan çalışmalar ile öne çıkmaktadır. Duygu analizi kapsamında finansal tahminleme, gündem ile alakalı topluluk gönderilerinin analizi, tahminleme çalışmaları gibi önemli konularda kullanılmaktadır.

Bu alıřmada dnřtrc modeli olarak kullanılan DistilBERT modeli makine ğrenmesi modellerinden KNN, SVM ve LR ve de derin ğrenme modellerinden CNN ve LSTM ile birlikte alıřacak bir model oluřturmak suretiyle duygu sınıflandırılmasında kullanılmıřtır. Bu modeller mevcut alıřmada dnřtrc kullanmayan klasik yntemlere gre kıyaslama yapılması iin kullanılmıřtır. Gelecek alıřmalarda aynı modellerin dnřtrc kullanılmadan znelik ıkaran derin ğrenme modellerine gre kıyaslaması yapılabilir.

## KAYNAKLAR

- Ankit, and Nabizath Saleena. «An Ensemble Classification System for Twitter Sentiment Analysis» *International Conference on Computational Intelligence and Data Science*, 2018: 5084--5094.
- Shiyang Liaoa, Junbo Wangb, Ruiyun Yua, Koichi Satob, Zixue Cheng . « CNN for situations understanding based on sentiment analysis of twitter data » *8th International Conference on Advances in Information Technology, IAIT2016(ELSEVIER)* , 2016: 1877-0509.
- Yunxiang Zhang, Zhuyi Rao. « Deep neural networks with pre-train model BERT for aspect-level sentiments classification.» *IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020)*, 2020.
- Sujay Ra, Jagadeesh Pujari b, Vandana Shreenivas Bhat c, Anita Dixit. « Timeline Analysis of Twitter User.» *International Conference on Computational Intelligence and Data Science(Elsevier)*, 2018.
- Mahmut Cetin , M. Fatih Amasyalı . "Egiticili ve Geleneksel Terim Ağırlıklılandırma Yöntemleriyle Duygu Analizi." IEEE, 2013.
- Metin Bilgin, İzzet Fatih Şentürk. "Sentiment Analysis on Twitter data with Semi- Supervised Doc2Vec" *2nc International Conference on Computer Science and Engineering*, 2017.
- Matheus Gomes de Sousa, Kenzo Sakiyama, Lucas de Souza Rodrigues, Pedro Henrique de Moraes , Eraldo Rezende Fernandes. «BERT for Stock Market Sentiment Analysis.» *IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)* (IEEE) ,2019
- Wenting Li, Shangbing Gao,Hong Zhou, Zihe Huang, Kewen Zhang and Wei Li. «The Automatic Text Classification Method Based on BERT and Feature Union.» *IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, 2019.
- Yingyu Wang , Jiaju Qi , Kan Zheng. "A CNN-based Feature Extraction Scheme for Patent Analysis".*IEEE 4th International Conference on Computer and Communications*. 2018.
- Önder Çoban, Gül,sah Tümüklü Özyer. "GTürkçe Twitter Mesajları için LDA ile Duygu Sınıflandırması", 2016.
- Andrew Christian Flores, Rogelyn I. Icoy , Christine F. Peña and Ken D. Gorro . «An Evaluation of SVM and Naive Bayes with SMOTE on Sentiment Analysis Data Set.» *IEEE Transactions on Industrial Electronics*, 2018.
- Sumandeep Kaur, Geeta Sikka, and Lalit Kumar Awasthi. «Sentiment Analysis Approach Based on N-gram and KNN Classifier» *First International Conference on Secure Cyber Computing and Communication*,2018.
- Sani Kamış and Dionysis Goularas «Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data» *International Conference on Deep Learning and Machine Learning in Emerging Applications, I-SMAC 2019*, 2019.

- Wenting Li, Shangbing Gao, Hong Zhou, Zihe Huang, Kewen Zhang Wei Li. «The Automatic Text Classification Method Based on BERT and Feature Union.» *IEEE 25th International Conference on Parallel and Distributed Systems*, 2019.
- Ngoc C. L<sup>ê</sup>, Nguyen The Lam, Son Hong Nguyen, Duc Thanh Nguyen. « On Vietnamese Sentiment Analysis: A Transfer Learning Method.»
- Yunxiang Zhang, Zhuyi Rao. « Deep neural networks with pre-train model BERT for aspect-level sentiments classification.» *IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020)*, 2020.
- Jiaqi Xou , Xin Li, Haipeng Yao. « BERT-Based Chinese Relation Extraction for Public Security» *IEEEACCESS 10.1109/ACCESS.2020.3002863*.
- Shrikant Tiwari, Prasenjit Chanak. «Twitter Sentiment Analysis» *DOI 10.1109/TAI.2022.3142241, IEEE*, 2022.
- Chirag Kariya and , Priti Khodke. «Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data.» *International Conference for Emerging Technology (INCET)*, 2020.
- JOÃO VIEIRA and , RUI P. DUARTE. « kNN-STUFF: kNN STreaming Unit for Fpgas.» *Digital Object Identifier 10.1109/ACCESS.2019.295*.
- Tianxiao Lui. « U.S. Pandemic Prediction Using Regression and Neural Network Models.» *International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI)*, 2020.
- Akbar Karimi, Leonardo Rossi, Andrea Prati and Katharina Full « Adversarial Training for Aspect-Based Sentiment Analysis with BERT.» *arXiv:2001.11316v1 [cs.LG]*, 2020.
- Mayank Ramina, Nihar Darnay, Chirag Ludbe and Ajay Dhruv « Topic level summary generation using BERT induced Abstractive Summarization Model.» *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2*, 2020.
- Zhimin Lin, Dajiang Lei, Yuting Han, Guoyin Wang and Wei Deng « Siamese BERT Model with Adversarial Training for Relation Classification.» *IEEE International Conference on Knowledge Graph (ICKG)*, 2020.
- Dong Yu-nan, Liang Guang-sheng « Research and Discussion on Image Recognition and Classification Algorithm Based on Deep Learning.» *International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, 2019.
- Jinqiao Liu« Survey of the Image Recognition Based on Deep Learning Network for Autonomous Driving Car.» *5th International Conference on Information Science, Computer Technology and Transportation (ISCTT)*, 2020.
- Qiangqiang Ye, Xueqin Yang, Chaobo Chen, Jingcheng Wang «River Water Quality Parameters Prediction Method Based on LSTM-RNN Model.» *The 31th Chinese Control and Decision Conference (2019 CCDC)*.

- Jianqiong Xiao, Zhiyong Zhou «Research Progress of RNN Language Model» *IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA) (2020 IEEE)*, 2020.
- Shayak Chakraborty, Jayanta Banik, Shubham Addhya and Debraj Chatterjee «Study of Dependency on number of LSTM units for Character based Text Generation models.» *International Conference on Computer Science, Engineering and Applications (ICCSEA)*, 2020
- Yan Li, Yifei Lu « LSTM-BA: DDoS Detection Approach Combining LSTM and Bayes.» *Seventh International Conference on Advanced Cloud and Big Data (CBD)*, 2019.
- Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts « Learning Word Vectors for Sentiment Analysis.» *The 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, 2011.
- Yomna Eid Rizk, Walaa Medhat Asa « Sentiment Analysis using Machine Learning and Deep Learning Models on Movies Reviews» *3rd Novel Intelligent and Leading Emerging Sciences Conference*, 2021.
- Acheampong Francisca Adoma, Nunoo-Mensah Henry, Wenyu Chen «Comparative Analyses Of Bert, Roberta, Distilbert, And Xlnet For Text-Based Emotion Recognition» *17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) | /20/\$31.00 ©2020 IEEE | DOI: 10.1109/ICCWAMTIP51612.2020.931*, 2017.
- Asma Ul Hussna, Iffat Immami Trisha, Md. Sanaul Karim, Md. Golam Rabiul Alam« COVID-19 Fake News Prediction On Social Media Data» *IEEE Region 10 Symposium (TENSYP) / 978-1-6654-0026-8/21/\$31.00 ©2021 IEEE | DOI: 10.1109/TENSYP52854.2021.9550957*, 2021.
- Jahanur Biswas, Md. Mahbubur Rahman, Al Amin Biswas, Md. Akib Zabed Khan, Aditya Rajbongsh, Hasnaine Amin Nilo« Sentiment Analysis on User Reaction for Online Food Delivery Services using BERT Model» *7th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2021.
- Jing Bai, Rui Cao, Wen Ma, Hiroyuki Shinnou « Construction of Domain-Specific DistilBERT Model by Using Fine-Tuning» *International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, 2020.
- Razu Ahmed, Asraf Ali, Joy Roy, Shakil Ahmed, N. Ahmed « Breast Cancer Risk Prediction based on Six Machine Learning Algorithms » *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 2020.
- Tshephisho Joseph Sefara« The Effects of Normalisation Methods on Speech Emotion Recognition» *International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, 2019.

Hüseyin İlgin, Erdal Kılıç «Sentiment Analysis using Transformers and Machine Learning Models » *6th International Conference on Computer Science and Engineering (UBMK)*, 2021.

<https://www.kaggle.com/chaitanyakaushal/twitter-sentiment-analysis-using-naive-bayes>

Rahit Kumar Kaliyar « A Multilayer Bidirectional Transformer Encoder for Pre-Trained Word Embedding: A Survey of BERT» *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2020.

Mayank Ramina, Nihar Darnay, Chirag Ludbe, Ajay Dhruv « Topic level summary generation using BERT induced Abstractive Summarization Model» *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)*, 2020.

Nguyen Thi Mai Trang<sup>1</sup> and Maxim Shcherbakov « Vietnamese Question Answering System from Multilingual BERT Models to Monolingual BERT Model» *9th International Conference System Modeling and Advancement in Research Trends (SMART)*, 2020.

Tingyu Zhang, Ruixia Zhan«Revealing the power of BERT for text sentiment classification» *IEEE 4th International Conference on Automation, Electronics and Electrical* , 2021.

Yong Gan, Ruisen Yang, Chenfang Zhang, Dongwei Jia «Chinese Named Entity Recognition based on BERT-Transformer-BiLSTM-CRF Model» *7th International Symposium on System and Software Reliability (ISSSR)*, 2021.

Anurag Singh, Goonjan Jain «Sentiment Analysis of News Headlines Using Simple Transformers» *Asian Conference on Innovation in Technology (ASIANCON)*, 2021.

He Jun, Liu Peng, Jiang Changhui, Liu Pengzheng, Wu Shenke, Zhong Kejia « Personality Classification Based on Bert Model» *IEEE International Conference on Emergency Science and Information Technology (ICESIT)*, 2021.

Wenkun Li, Hui D, Lina Wang, Kazushige Ouchi, Jing Lu« Boost Transformer with BERT and copying mechanism for ASR error correction» *International Joint Conference on Neural Networks (IJCNN)*, 2021.

Zihan Wang, Taozheng Zhu, Shice Mai« Disaster Detector on Twitter Using Bidirectional Encoder Representation from Transformers with Keyword Position Information» *IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)* ,2020.

Rida Miraj, Masaki Aono « Humor Detection Using a Bidirectional Encoder Representations from Transformers (BERT) based Neural Ensemble Model» *8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, 2021.

Rida Miraj, Masaki Aono « Humor Detection Using a Bidirectional Encoder Representations from Transformers (BERT) based Neural Ensemble Model» *8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, 2021.

- Supragya Sonkar, Sriram G. Sanjeevi «Multi-Head Attention on Image Captioning Model with Bert Embedding» *International Conference on Communication, Control and Information Sciences (ICCISc)*, 2021.
- Tanmoy Tapos Datta, Pintu Chandra Shill, Zabir Al Nazi «BERT-D2: Drug-Drug Interaction Extraction using BERT» *International Conference for Advancement in Technology (ICONAT)*, 2022.
- Ersin Enes Yılmaz, Erdal Kılıç «Türkçe İstenmeyen E-postaların Derin Öğrenme ile Tespit Edilmesi» *Samsun Ondokuz Mayıs Üniversitesi*, 2021.
- Yamashita, R., Nishio, M., Do, R.K.G. «Convolutional neural networks: an overview and application in radiology» *Insights Imaging* , 2018.
- Nur Ghaniaviyanto Ramadhan, Teguh Ikhlas Ramadhan « Analysis Sentiment based on IMDB aspects from movie reviews using SVM» *Sinkron : Jurnal dan Penelitian Teknik Informatika* Volume 7, Number 1, 2022.
- Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray « Topic Level Sentiment Analysis of Social Media Data Using Deep Learning» *Applied Soft Computing, Volume 108*, Sempember 2021.
- Nhan Cach Dang, María N. Moreno-García « Sentiment Analysis Based On Deep Learning : A Comperative Study» *Electronics* 2020, 9, 483; doi:10.3390/electronics9030483.
- Ammar Mohammed, Rania Kora1 « Deep learning approaches for Arabic sentiment analysis» *Social Network Analysis and Mining*, 2019.
- Ahmed Sulaiman M. Alharbi, Elise de Doncker «Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information» *ScienceDirect*, 2019.
- Paramita Ray, Amlan Chakrabarti « A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis» *Applied Computing and Informatics Vol. 18 No. 1/2, 2022 pp. 163-178 Emerald Publishing Limited e-ISSN: 2210-8327 p-ISSN: 2634-1964 DOI 10.1016/j.aci.2019.02.002*, 2019.
- Sani Kaniş, Dionysis Goularas « Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data » *International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, 2019.
- Luca Bacco, Andrea Cimino, Felice Dell'Orletta and Mario Merone « Extractive Summarization for Explainable Sentiment Analysis using Transformers » *CEUR Workshop Proceedings*, 2021.
- Ting Zhang, Bowen Xu, Ferdian Thung, Stefanus Agus Haryono, David Lo, Lingxiao Jiang « Sentiment Analysis for Software Engineering: How Far Can Pre-trained Transformer Models Go?» *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2020.
- Shivaji Alaparthi and Manit Mishra « Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey » *arXiv:2007.01127*, 2020.

## ÖZGEÇMİŞ

Hüseyin İlgün, Sakarya Arifiye Anadolu Öğretmen Lisesini bitirdikten sonra Gebze Yüksek Teknoloji Enstitüsü Bilgisayar Mühendisliği bölümünden 02.07.2014 tarihinde mezun oldu. 2014 senesinde Elginkan Holding’ de Bilgisayar Mühendisi olarak işe başladı. 2015 senesinde Vizyoneks A.Ş. şirketinde yazılım mühendisi olarak iş hayatına devam etti. 2016 senesinde VIS Teknoloji firmasında teknoloji danışmanı olarak görev aldı. 2018 senesinde Deniz Harp Okulu’ nda yedek subay olarak görevini tamamladıktan sonra 2019 senesinde Kastamonu Üniversitesi Taşköprü Meslek Yüksekokulu’ nda öğretim görevlisi olarak görev aldı. 2019 yılında Samsun Ondokuz Mayıs Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Ana Bilim Dalı Tezli Bilgisayar Mühendisliği programına başladı. 2021 senesinden bu yana Sermaye Piyasası Kurulu’ nda bilişim uzman yardımcısı olarak görevine devam etmektedir.

### İletişim Bilgileri

ORCID ID : 0000-0002-8414-8445

### Yayımlar:

1. “Sentiment Analysis using Transformers and Machine Learning Models”, 6th International Conference on Computer Science and Engineering (UBMK), 2021

### Sertifikalar:

1. “Yazılım Test ve Kalite Sertifikası ”, Turkcell&YTÜ, Aralık 2016