

İş İsteklerindeki Türkçe Metinlerin Doğal Dil İşleme Yöntemleri Kullanılarak  
Analizi

Ceren Özkan

**YÜKSEK LİSANS TEZİ**

Bilgisayar Mühendisliği Anabilim Dalı

Ocak 2024



Analysis Of Turkish Text in Job Requests Using Natural Language Processing Methods

Ceren Özkan

**MASTER OF SCIENCE THESIS**

Department of Computer Engineering

January 2024

İş İsteklerindeki Türkçe Metinlerin Doğal Dil İşleme Yöntemleri Kullanılarak  
Analizi

Ceren Özkan

Eskişehir Osmangazi Üniversitesi

Fen Bilimleri Enstitüsü

Lisansüstü Eğitim ve Öğretim Yönetmeliği Uyarınca

Bilgisayar Mühendisliği Anabilim Dalı

Yapay Zeka Bilim Dalında

YÜKSEK LİSANS TEZİ

olarak hazırlanmıştır.

Danışman: Prof. Dr. Kemal ÖZKAN

Ocak 2024

## ONAY

Bilgisayar Mühendisliđi Anabilim Dalı Yapay Zeka Bilim Dalı Yüksek Lisans öđrencisi Öđrencinin **Ceren ÖZKAN**'ın YÜKSEK LİSANS tezi olarak hazırladıđı “İş İsteklerindeki Türkçe Metinlerin Doğal Dil İşleme Yöntemleri Kullanılarak Analizi” başlıklı bu çalışma, jürimizce lisansüstü yönetmeliđin ilgili maddeleri uyarınca deđerlendirilerek oybirliđi ile kabul edilmiřtir.

Danışman : Prof. Dr. Kemal ÖZKAN

İkinci Danışman :

Yüksek Lisans Tez Savunma Jürisi:

Üye: Prof. Dr. Kemal ÖZKAN

Üye: Prof. Dr. Ahmet YAZICI

Üye: Dr. Öğr. Üyesi Cahit PERKGÖZ

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ..... tarih ve  
..... sayılı kararıyla onaylanmıştır.

Prof. Dr. ....

Enstitü Müdürü

## ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Prof. Dr. Kemal ÖZKAN danışmanlığında hazırlamış olduğum “İş İsteklerindeki Türkçe Metinlerin Doğal Dil İşleme Yöntemleri Kullanılarak Analizi” başlıklı Yüksek Lisans tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallarına uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. 16/01/2024

Ceren ÖZKAN

İmza

## ÖZET

Günümüzde teknolojinin hız kesmeden ilerlemesi, yapay zeka ve doğal dil işleme alanındaki gelişmeleri de artırmaktadır. Yapay zeka, makinelerin öğrenme, problem çözme, dil anlama, algılama ve karar verme gibi karmaşık görevleri gerçekleştirebilme yeteneklerini geliştirmeyi hedeflemektedir. Doğal Dil İşleme, bilgisayarların insan dilini anlama, yorumlama, üretme ve işleme yeteneği üzerine odaklanan bir alanı ifade etmektedir. Doğal dil işleme ve yapay zekanın birlikte kullanılması birçok sektörde inovatif uygulamaların ortaya çıkmasına olanak sağlamıştır. Yapılan tez çalışması kapsamında şirket içerisinde geliştirilip kullanılmakta olan iş istek uygulamasında bulunan iş istek açıklamalarının kategorilerinin tahminlemesi üzerine çalışılmıştır. Girilen açıklama metinlerinden kategori tahminleme yapan bir sistem geliştirilmiştir. Çalışmada ön eğitilmiş doğal dil işleme modellerinden Bert, DistilBERT, Electra ve ConvBERT modelleri kullanılmıştır. Veri seti bu modeller ile eğitildikten sonra kategori tahminlemesi yapılmıştır. Gerçekleştirilen çalışma sonucunda en yüksek doğruluk %89 başarı değeri ile ConvBERT modelinde elde edilmiştir. Çalışmanın ikinci kısmında ConvBERT modelinde elde edilen sonuçlar Lime (Local Interpretable Model-agnostic Explanations) açıklanabilir yapay zeka modeli kullanılarak sistemin tahminlemeyi nasıl yaptığı üzerine çalışılmıştır. Bu sayede modelde ki tahminleme süreci açıklanabilir bir hale getirilerek açıklanabilir yapay zeka ile doğal dil işlemenin birlikte kullanıldığı bir tahminleme sistemi geliştirilmiştir.

**Anahtar Kelimeler:** Doğal Dil İşleme, Metin Sınıflandırma, Açıklanabilir Yapay Zeka

## SUMMARY

Nowadays, the unabated advancement of technology also increases the developments in artificial intelligence and natural language processing. Artificial intelligence aims to improve the ability of machines to perform complex tasks such as learning, problem solving, language understanding, perception and decision making. Natural Language Processing refers to a field that focuses on the ability of computers to understand, interpret, produce and process human language. The combined use of natural language processing and artificial intelligence has enabled the emergence of innovative applications in many sectors. Within the scope of the thesis study, a study was performed to estimate the categories of job request statements in the job request application developed and used within the company. A system that makes category predictions from entered description texts has been developed. In the study, pre-trained natural language processing models Bert, DistilBERT, Electra and ConvBERT were used. Category prediction was made, after the data set was trained with these models. As a result of the study, the highest accuracy was obtained in the ConvBERT model with a success rate of 89%. In the second part of the study, the results obtained in the ConvBERT model were examined and how the system made the prediction using the Lime (Local Interpretable Model-agnostic Explanations) explainable artificial intelligence model. In this way, the prediction process in the model has been made more clear and a the system has been developed in which explainable artificial intelligence and natural language processing are used together.

**Keywords:** Natural Language Processing, Text Classification, Explainable Artificial Intelligence

## TEŞEKKÜR

Yüksek Lisans eğitimim ve tez sürecimde desteklerini, deneyimlerini benden esirgemeyen değerli tez danışmanım Prof. Dr. Kemal ÖZKAN'a ve bu süreçte hep yanımda olan ve beni destekleyen sevgili aileme çok teşekkür ederim. Tez çalışmamda verilerini kullanmama izin veren Özdilek Holding Anonim Şirketi'ne ayrıca teşekkür ederim.

Ceren ÖZKAN

**İÇİNDEKİLER****Sayfa**

<b>ÖZET</b> .....	vi
<b>SUMMARY</b> .....	vii
<b>TEŞEKKÜR</b> .....	viii
<b>İÇİNDEKİLER</b> .....	ix
<b>ŞEKİLLER DİZİNİ</b> .....	x
<b>ÇİZELGELER DİZİNİ</b> .....	xi
<b>SİMGELER VE KISALTMALAR DİZİNİ</b> .....	xii
<b>1. GİRİŞ VE AMAÇ</b> .....	1
<b>2. LİTERATÜR ARAŞTIRMASI</b> .....	5
<b>3. TEORİK BİLGİ</b> .....	14
3.1. Veri Ön İşleme.....	14
3.2. Metin Ön İşleme.....	15
3.3. Çalışmada Kullanılan Modeller.....	17
3.3.1. Bert.....	22
3.3.2. DistilBert.....	24
3.3.3. Electra.....	25
3.3.4. ConvBERT.....	26
3.3.5. Açıklanabilir Yapay Zeka.....	26
3.3.6. Lime (Local Interpretable Model - Agnostic Explanations).....	27
3.4. İncelenen Çalışmalar.....	27
<b>4. DENEYSEL ÇALIŞMA</b> .....	31
<b>5. BULGULAR VE TARTIŞMA</b> .....	36
<b>6. SONUÇ VE ÖNERİLER</b> .....	41
<b>KAYNAKLAR DİZİNİ</b> .....	42
<b>EK AÇIKLAMALAR</b> .....	49
Ek Açıklama-A: Tez Çalışmasından Üretilen Bilimsel Yayın .....	49

## ŞEKİLLER DİZİNİ

### Sekil

### Sayfa

1.1. Yapılan çalışmanın adımları.....	4
2.1. Doğal dil işleminin kullanım alanları.....	8
2.2. Doğal dil işleminin tarihçesi.....	11
3.1. Bir dönüştürücünün yapısı.....	18
3.2. Metin verisinin belirteç yapısına dönüştürülmüşü.....	19
3.3. Metin verisinin belirteç numarası yapısına dönüştürülmüşü.....	19
3.4. BERT modelinin genel gösterimi.....	22
3.5. BERT girdi temsili.....	24
3.6. DistilBERT modelinin genel mimarisi.....	25
3.7. Electra modeli yapısı.....	26
4.1. Google colab ortamı.....	31
4.2. Kullanılan veri seti.....	32
4.3. Veri setinin kategori bazlı dağılımı.....	33
4.4. Metin ön işleme yapılmadan önce veri.....	33
4.5. Metin ön işleme yapıldıktan sonra veri.....	33
4.6. Cümlelerin jetonlara dönüştürülmüş hali.....	35
5.1. Tüm modeller için her iterasyon aşamasındaki eğitim ve doğrulama kayıpları.....	37
5.2. Modelleri başarı yüzdeleri.....	38
5.3. SAP BW kategorisine ait bir cümlenin Lime algoritması ile incelenmesi.....	39
5.4. SAP ERP kategorisine ait bir cümlenin Lime algoritması ile incelenmesi.....	40

## ÇİZELGELER DİZİNİ

<b><u>Cizelge</u></b>	<b><u>Sayfa</u></b>
2.1. ChatGPT ve Bert karşılaştırılması.....	10
3.1. Bert modellerinin karşılaştırılması.....	23
4.1. Kategorilerin sayısal gösterimi.....	34
5.1. İterasyon bazlı modellerin eğitim sonuçları.....	36
5.2. ConvBERT modelinin eğitim sonucu.....	38



## SİMGELER VE KISALTMALAR DİZİNİ

<b>Kısaltmalar</b>	<b>Açıklama</b>
DDİ	Doğal Dil İşleme (Natural Language Processing)
ATN	Artırılmış Geçiş Ağı (Augmented Transition Network)
TF – IDF	Terim Frekansı— Tersine Doküman Frekansı (Term Frequency – Inverse Term Frequency)
CBOW	Sürekli Kelimeler Çantası (Continuous Bag of Words)
RNN	Tekrarlayan Sinir Ağları (Recurrent Neural Network)
NER	Adlandırılmış Varlık Tanıma (Name Entity Recognition)
GPT	Üretken Önceden Eğitilmiş Dönüştürücüler (Generative Pre-trained Transformer)
BERT	Dönüştürücülerden Çift Yönlü Kodlayıcı Temsilleri (Bidirectional Encoder Representations from Transformers)
TPU	Tensör İşleme Birimleri (Tensor Processing Unit)
MLM	Maskelenmiş Dil Modeli (Masked Language Modeling)
NSP	Sonraki Cümle Tahmini (Next Sentence Prediction)
LSTM	Uzun Kısa Süreli Bellek (Long Short – Term Memory)
CNN	Evrişimsel Sinir Ağları (Convolutional Neural Network)
MLP	Çok Katmanlı Algılayıcı (Multi Layer Perceptron)
GPU	Grafik İşlemci Ünitesi (Graphics Processing Unit)
DistilBERT	Transformatörlerden Damıtılmış Çift Yönlü Kodlayıcı Gösterimleri (Distilled Bidirectional Encoder Representations from Transformers)

**SİMGELER VE KISALTMALAR DİZİNİ****Kısaltmalar****Açıklama**

CUDA

Birleşik Hesap Cihazı Mimarisi (Compute Unified Device Architecture)

Electra

Dikkatli Bir Şekilde Token Değişimlerini Sınıflandıran Bir Kodlayıcıyı Etkin Bir Şekilde Öğrenme (Efficiently Learning an Encoder that Classifies Token

Replacements Accurately)

Lime

Yerel Yorumlanabilir Model - Agnostik Açıklamalar (Local Interpretable Model - Agnostic Explanations)

## 1. GİRİŞ VE AMAÇ

Yapay Zeka, bilgisayar sistemlerine insan benzeri düşünme ve öğrenme yetenekleri kazandırmayı amaçlayan bir bilim ve mühendislik dalıdır. Yapay zeka terimi, 1956 yılında düzenlenen Dartmouth Konferansı'nda McCarthy ve diğerleri tarafından ilk kez resmi olarak kullanılmıştır (Lewis, 2014). Bu konferans, yapay zeka alanının resmi başlangıcı olarak kabul edilmekte ve bu alanda çalışmaların ivme kazanmasına öncülük etmektedir. Yapay zeka, bilgisayarların problem çözme yeteneklerini, dil anlama, öğrenme, görsel algı, karar verme ve hatta duygusal tepkiler gibi insan benzeri yeteneklerle donatma amacını taşımaktadır. Yapay zeka sistemleri genellikle algoritmalar, matematiksel modeller ve büyük veri setleri kullanılarak öğrenme süreçlerini gerçekleştirmektedir. Yapay zeka, zaman içinde semantik ağlar, sinir ağları ve derin öğrenme gibi tekniklerle büyük bir evrim geçirmiştir. Yapay zeka disiplini, geniş bir çalışma alanına sahiptir. Bu alanlar arasında, makine öğrenimi, dil işleme, yapay sinir ağları, konuşma sentezi, konuşma anlama, uzman sistemler, bilgi tabanları ve bulanık mantık gibi spesifik alanlar bulunmaktadır (Coşkun ve Gülleroğlu, 2021). Bu tez çalışmasında yapay zekanın bir alt dalı olan doğal dil işleme (DDİ, Natural Language Processing, NLP) yöntemleri kullanılarak iş isteklerindeki Türkçe metinlerin analizi ve yapay zekanın elde etmiş olduğu sonuçlarının açıklanabilirliği üzerine çalışılmıştır.

DDİ, bilgisayar sistemlerinin doğal dildeki metinleri anlaması, yorumlaması, üretmesi ve işleme amacını taşıyan bir yapay zeka alt dalıdır (Şeker, 2016). Bu alandaki temel hedef, bilgisayar sistemlerine dilin karmaşıklığını ve çeşitliliğini anlamak, çözümlenmek ve kullanmak için yetenek kazandırmaktır. Bu sistemler, kelime dağarcığı, dilbilgisi kuralları, semantik anlam çıkarımı ve dilin bağlamını anlama gibi unsurları içeren bir dizi işlemi gerçekleştirebilmektedir. DDİ, metin madenciliği, metin sınıflandırma, çeviri, konuşma tanıma ve duygu analizi gibi birçok uygulama alanında kullanılmaktadır. DDİ günümüzde birçok endüstride kullanılmaktadır. Sosyal medya analizi, müşteri geri bildirimlerini değerlendirme, sağlık sektöründe hastane kayıtlarını analiz etme, finansal raporları çözümlenme gibi birçok uygulama alanına sahiptir. Bu sayede, bilgisayarlar

insanlarla daha etkili bir şekilde iletişim kurabilir ve karmaşık dil yapılarını anlayabilir hale gelmektedir.

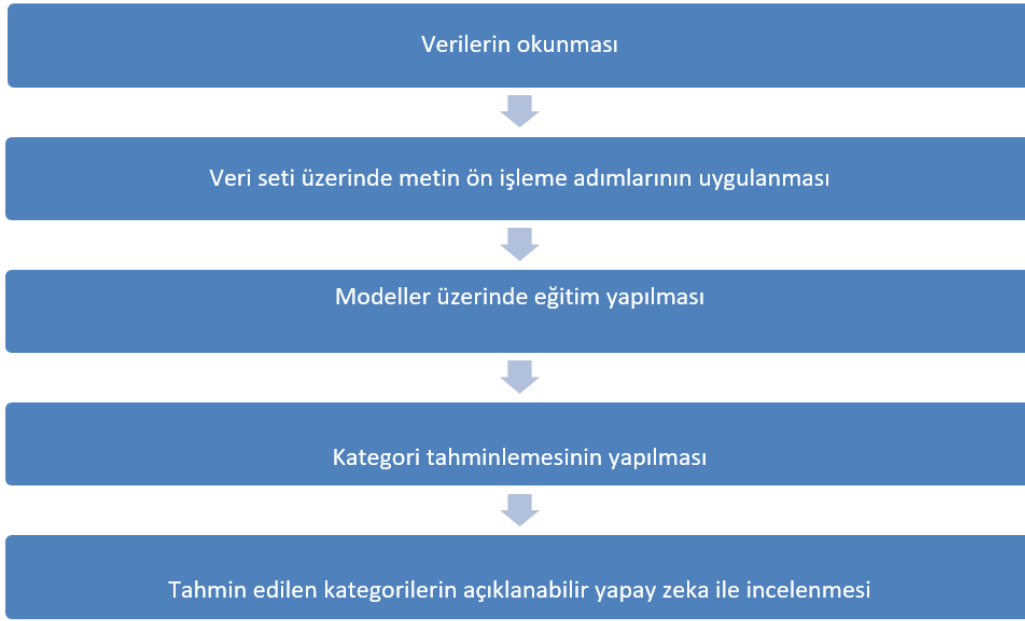
Günümüzde, teknolojideki gelişmelerin hızla artmasıyla birlikte yapay zeka ve DDİ alanlarındaki gelişmeler, insanlık için önemli bir dönüm noktasını temsil etmektedir. Yapay zeka, bilgisayar sistemlerine insan benzeri zekâ ve öğrenme yetenekleri kazandırma amacını taşırken (Altıntop, 2023), DDİ bilgisayarların doğal dildeki insan iletişimini anlaması ve üretmesini sağlayan bir alt dal olarak öne çıkmaktadır.

Yapay zeka başlangıcından itibaren temelde matematiksel modeller ve algoritmalarla şekillenmiş, ancak son yıllarda derin öğrenme ve büyük veri kullanımının artmasıyla birlikte büyük bir ivme kazanmıştır. Derin öğrenme algoritmaları, büyük veri setlerini analiz ederek karmaşık desenleri tanıma yetenekleriyle öne çıkarken, bu durum DDİ alanında da çığır açıcı sonuçlar doğurmuştur. Bu alanda kullanılan modeller ve yöntemler şu şekilde özetlenebilir:

- ❖ Kural Tabanlı Sistemler, dilin gramatik yapısını tanımlayan kurallar kullanılarak dil analizi yapan sistemlerdir. Ancak, bu yaklaşım genellikle dilin karmaşıklığını tam olarak yansıtmamaktadır.
- ❖ İstatistiksel Modeller, metin verilerindeki desenleri ve ilişkileri çıkarmak amacıyla istatistiksel yöntemleri içermektedir. Bu modeller, büyük miktarda dil verisini analiz ederek dilin yapılarını ve ilişkilerini öğrenir ve bu öğrenilen bilgileri temel alarak çeşitli görevleri gerçekleştirmektedirler. İstatistiksel modellerin temel prensibi, dildeki olasılıkları ve frekansları kullanarak tahminlerde bulunmaktır.
- ❖ Makine Öğrenimi Tabanlı Modeller, bilgisayar sistemlerinin büyük miktarda dil verisi üzerinden öğrenme yeteneğine sahip olduğu bir dizi teknik içermektedir. Bu modeller, metin verilerindeki desenleri ve ilişkileri otomatik olarak çıkarmak amacıyla makine öğrenimi algoritmalarını kullanmaktadır. Bu modeller, DDİ uygulamalarında önemli bir rol oynamakta olup, çeşitli dil işleme görevlerini gerçekleştirmek için geniş bir uygulama alanına sahiptirler.

- ❖ Derin Öğrenme Modelleri, DDİ alanında büyük başarı elde etmiş ve birçok görevde insan benzeri performans göstermiştir. Bu modeller, yapay sinir ağlarının (neural networks) derin ve karmaşık yapılarını kullanarak dil verilerindeki desenleri ve ilişkilerini öğrenmektedir. Derin öğrenme modelleri, büyük veri setleri üzerinde eğitildiklerinde ve uygun şekilde veriler ile uyarlandıklarında, dildeki karmaşıklıkları ve görevleri öğrenme konusunda oldukça başarılı sonuçlar vermektedir.
- ❖ Transfer Öğrenme, bir görevde edinilen bilgilerin, başka bir benzer görevde kullanılması anlamına gelmektedir. Transfer öğrenme modelleri, genellikle büyük bir veri setinde eğitilen bir modelin, daha sonra sınırlı bir veri setinde başka bir görevi öğrenmekte kullanılmasını ifade etmektedir. Bu modeller özellikle genel dil anlama görevlerinde ve az veri senaryolarında başarıyla kullanılmaktadır.

Teknolojinin gelişmesiyle birlikte işletmelerde, ofislerde kullanılan uygulamaların sayısı da büyük ölçüde artmıştır. Yapılan tez çalışmasında dönüştürücü tabanlı modeller kullanılarak işletme içerisinde kullanılan iş istek uygulamasındaki iş açıklamalarının kategorilerinin otomatik olarak tahmin edilmesi üzerine çalışma yapılmıştır. Mevcutta kullanılan uygulamada kişiler iş istek açıklamalarını girip kategori seçerek iş istek oluşturmaktadır. Sistemin arka planında her kategoriye ait ilgili geliştirici bilgileri tanımlıdır. Açıklama girilip kategori seçildikten sonra oluşturulan her iş isteğinin bilgisi o kategorideki tanımlı geliştiricilere gitmektedir. Fakat kullanıcılar kategori seçerken yanlış kategori seçebilmekte ve işlerin yanlış kategoriye atanmasına sebep olmaktadır. Bu durum ile beraber iş istek bilgileri doğru geliştiricilere gitmemektedir. Böylece işlerin çözüm süreleri de uzamaktadır. Tez kapsamında bu probleme çözüm olacak otomatik kategori tahminleme üzerine çalışma yapılmıştır. İş isteklerinin doğru kişilere atanması sayesinde işlerin çözüm süreleri hızlanmış olacaktır. Aşağıdaki şekilde (Şekil 1.1.) yapılan çalışmanın adımları gösterilmiştir. Yapılan çalışma Türkçe DDİ bilimsel literatürüne katkı sağlamıştır.



Şekil 1.1. Yapılan çalışmanın adımları

Yapılan tez çalışması altı bölümden oluşmaktadır. Giriş bölümü olarak adlandırılan birinci bölümde tez çalışması yapılan konu hakkında bilgi verilmiştir. İkinci bölümde DDİ hakkında literatür araştırması yapılarak geçmişten günümüze yaşadığı gelişmelerden bahsedilmiştir. Ayrıca bu bölümde DDİ'nin uygulama alanlarından bahsedilmiştir. Üçüncü bölümde teorik bilgiler hakkında bilgi verilmiştir. Çalışmada kullanılan veri ön işleme, metin ön işleme, kullanılan modeller ve bu modellerle literatürde yapılan çalışmalar incelenerek anlatılmıştır. Dördüncü bölümde yapılan çalışma hakkında bilgi verilmiştir. Beşinci bölümde yapılan çalışma sonucunda elde edilen bilgiler verilmiştir. Altıncı bölümde yani sonuç ve öneriler kısmında çalışmanın sonuçları yorumlanmıştır. Sonuçların geliştirilebilmesi için neler yapılabilirliği hakkında önerilerde bulunulmuştur.

## 2. LİTERATÜR ARAŞTIRMASI

Dil, insanlık tarihinin en önemli yapı taşlarından birini temsil eder. İnsanlar arasında iletişim kurmanın temel aracı olan dil, toplumların, kültürlerin ve bilginin aktarılmasında kritik bir rol oynamaktadır (Kara, 2020). Dil, insanların duygularını, düşüncelerini ve deneyimlerini paylaşmalarını sağlayarak insan topluluklarını bir arada tutan önemli bir bağ olmuştur.

DDİ, bir bilgisayarın insan dilini anlaması, yorumlaması ve işleyebilmesi için kullanılan bir yapay zeka dalıdır (Akarsu, 2022). DDİ, metin ve konuşma verilerini analiz ederek, metinlerin anlamını çıkarma, dilbilgisi yapısı ve anlamını anlama, metindeki duyguyu analiz etme, metin tabanlı cevaplar üretme ve metinleri farklı dillere çevirme gibi çeşitli alt görevleri içermektedir. DDİ, büyük veri analizi, yapay sinir ağları ve dil işleme algoritmaları gibi çeşitli tekniklerin birleşimiyle çalışır ve iş dünyasında, eğitimde, sağlık sektöründe ve daha pek çok alanda önemli uygulamalar sunmaktadır.

DDİ alanı, bilgisayar bilimlerinin doğduğu ve gelişmeye başladığı dönem olan 1950'li yıllarda belirgin bir şekilde ortaya çıkmıştır. DDİ'nin ilk örneklerinden biri Alan Turing'in 1950'li yıllarda yaptığı Turing Testi çalışmasıdır. Yapılan çalışmada bilgisayarlarla düşünme kavramı tanıtılmıştır (Turing, 1950). Turing Testi, bir makinenin insan gibi davranma yeteneğini değerlendirmek amacıyla kullanılan bir testtir. 1954 yılında yapılan bir başka önemli DDİ projesi ise Georgetown Üniversitesi'nde IBM tarafından finanse edilen ve sadece 250 kelime için İngilizce-Rusça çeviri yapmayı amaçlayan "Georgetown-IBM Experiment" çalışmasıdır (Hutchin, 2006). Çalışmada el ile programlanmış kurallara dayalı bir yaklaşım kullanılarak İngilizce metinlerin Rusça'ya çevrilmesi hedeflenmiştir. O dönemdeki bilgisayarlar ve teknoloji yeteri kadar gelişmiş olmadığı için büyük başarılar elde edilememiştir.

1966 yılında geliştirilen ELIZA, Joseph Weizenbaum tarafından MIT'de oluşturulan ilk DDİ bilgisayar programıdır (Üzücek, 2023). ELIZA, temel olarak metin tabanlı bir kullanıcı arayüzü kullanarak yazılı metin yoluyla kullanıcılarla etkileşim kuran bir programdır. Programın çalışma prensibi, kullanıcının sorularını veya ifadelerini anlamak ve

belirli kalıplara veya kurallara dayalı yanıtlar üretmektir. ELIZA, yapay zeka ve DDİ alanlarında bir dönüm noktasıdır. Bu çalışmayla modern sohbet robotlarının temeli atılmıştır.

1970'li yıllarda, istatistiksel yaklaşımlar DDİ için giderek daha fazla önem kazanmıştır. 1971 yılında Terry Winograd tarafından SHRDLU adlı program yazılmıştır. Program blok dünyası adı verilen soyut bir ortamda çalışmaktadır. Kullanıcılar bu blok dünyasında bulunan nesnelere komutlar sayesinde hareket ettirebilmektedir (Winograd, 1971). Program, nesnelere arasındaki ilişkileri anlayabilir ve mantıksal çıkarımlar yaparak kullanıcının talimatlarını yerine getirebilmektedir. SHRDLU programının en dikkat çekici özelliği, komutları anlama ve mantıksal çıkarımlar yapabilmesidir. 1970 yılında NLP araştırmacısı William A. Woods tarafından artırılmış geçiş ağı (Augmented Transition Network, ATN) geliştirilmiştir. Artırılmış geçiş ağı doğal dil girdisini temsil etmektedir ve bilgisayarlara cümlelerin yapılarını anlayabilmesini sağlamaktadır (Woods, 1970). Artırılmış geçiş ağı, dil işleme ve anlam çıkarma süreçlerini modellemek amacıyla kullanılan bir dil işleme çerçevesidir. Artırılmış geçiş ağı'nın temel fikri, dil işleme işlemini bir dizi sıralı adımlarla gerçekleştirmektir. Her adımda bir dil işleme görevi gerçekleştirilir ve sonuçlar bir sonraki adıma iletilir. Bu, dil işleme sürecini modüler hale getirmektedir ve farklı adımların bağımsız olarak geliştirilmesine ve test edilmesine olanak sağlamaktadır. Bu sayede Artırılmış Geçiş Ağı, DDİ alanında teorik bir çerçeve olarak kabul edilir ve daha sonra geliştirilen diğer modellere ilham kaynağı olmuştur.

1980 ve 1990'lı yıllarda DDİ'de istatistiksel modeller ve makine öğrenmesi yöntemleri kullanılmaya başlanmıştır. Bu yıllarda makine öğrenmesi algoritmaları, büyük metin veri setlerini analiz ederek dilin yapısını öğrenmeye başlamıştır. 1980'lerde konuşma sinyallerini metinlere dönüştüren algoritmalar geliştirilmeye başlanmıştır. 1990'lı yıllarda internetin gelişmesiyle birlikte çevrimiçi metinlere erişimin kolaylaşmıştır, böylece dil işleme araştırmacılarının daha geniş ve çeşitli veri setleriyle çalışması sağlanmıştır.

1950 ile 2000 yılları arasında yapılan çalışmalar bilgisayara önceden tanımlanan kurallar sayesinde DDİ süreçlerini gerçekleştirmektedir. 2000'li yıllardan sonra DDİ çalışmalarında yapay sinir ağları kullanılmaya başlanmıştır. 2001 yılında Bengio ve diğerleri, tek gizli katmanlı ileri beslemeli sinir ağından oluşan ilk sinirsel dil modelini

önermişlerdir (Bengio vd, 2003). Bu model ile birlikte dil işleme görevlerinde önemli ilerlemeler kaydedilmiştir. Ayrıca Bengio ve diğerleri kelime gömmelerini de tanıtan ilk kişilerdir.

Ronan Collobert ve Jason Weston çalışmalarında DDİ için derin sinir ağları ve çoklu görev öğrenme yaklaşımını ele almışlardır. Yapılan çalışmada, birçok farklı DDİ görevini aynı anda öğrenmek için çoklu görev öğrenme yaklaşımı önerilmiştir. Bu yaklaşım metin sınıflandırma, dil modellemesi, dil analizi gibi bir dizi görevin bir arada öğrenilmesini içermektedir (Collobert ve Weston, 2008). Yapılan çalışma DDİ'de derin öğrenme ve çoklu görev öğrenmenin kullanılmasının öncüsüdür ve bu alanın gelişmesine büyük katkıda bulunmuştur. Collobert ve Weston'ın çalışması, birçok modern DDİ modelinin temelini oluşturmuş ve çoklu görev öğrenme yaklaşımının DDİ'deki başarısını göstermiştir.

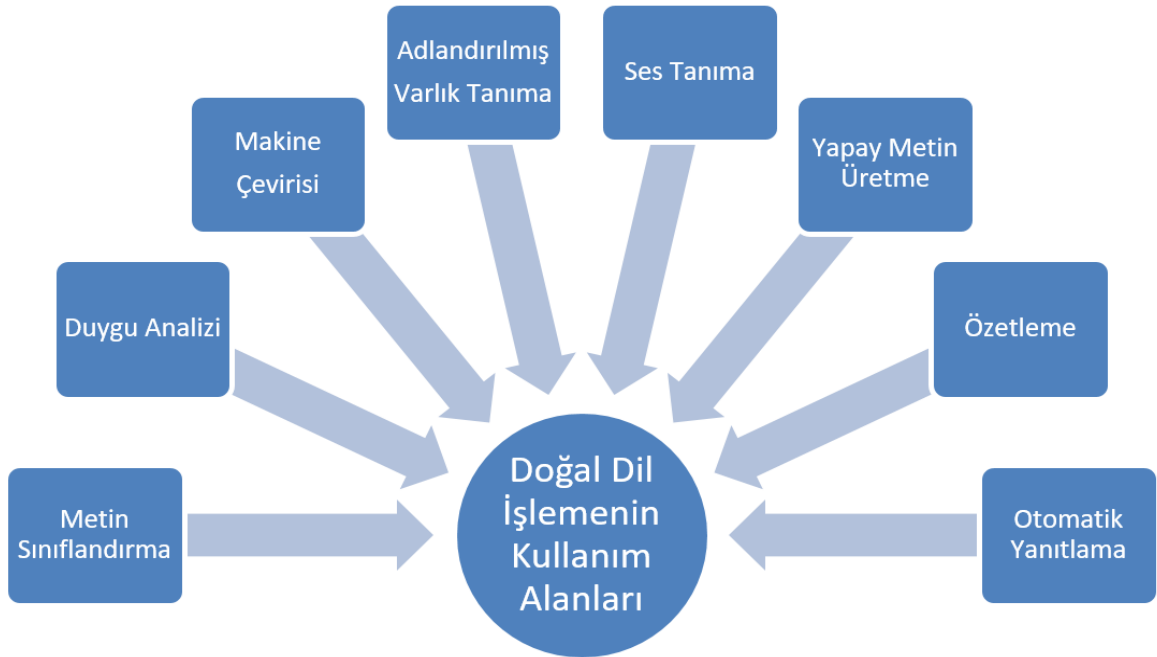
2013 yılında Mikolov ve diğerleri, derin öğrenme ve kelime gömme konusunda önemli bir atılım yaparak Word2Vec modelini tanıtmışlardır (Mikolov vd., 2013). Word2Vec, kelime temsilini öğrenmek için kullanılan bir dizi derin öğrenme modelini içermektedir. Word2Vec, kelime vektörlerini öğrenmek için büyük metin verilerini kullanarak, semantik benzerlikleri ve ilişkileri yakalayabilmektedir. Word2Vec modelinin iki ana modeli bulunmaktadır. Bunlar; bir kelimenin çevresindeki diğer kelimeleri kullanarak o kelimenin tahmin edilmesini amaçlayan Sürekli Kelimeler Çantası (Continuous Bag of Words, CBOW) ve bir kelimenin çevresindeki kelimeleri tahmin etmeyi amaçlayan Skip-Gram modelleridir. Word2Vec, kelimenin semantik anlamını daha iyi yakalamak ve kelime benzerliklerini hesaplamak için kullanışlıdır.

2014 yılında Ilya Sutskever ve diğerlerinin Seq2Seq ismini verdikleri çalışma bir giriş dizisini bir çıkış dizisine dönüştürmek için kullanılan bir derin öğrenme modelidir (Sutskever vd., 2014). Örneğin, bir dildeki cümleyi başka bir dile çevirebilmek için bu model kullanılabilir. Model, tekrarlayan sinir ağları (Recurrent Neural Network, RNN) kullanarak bu görevi gerçekleştirmektedir.

2015 yılında Dikkat Mekanizması (Attention Mechanism) adı verilen bir derin öğrenme bileşeni tanıtılmıştır (Bahdanau vd., 2015). Dikkat mekanizması, metin veya dizi tabanlı giriş ve çıkış verileri arasındaki ilişkiyi modellemek için kullanılan bir derin öğrenme

bileşendir. Dikkat mekanizması, özellikle metin çevirisi ve metin anlama gibi görevlerde, modelin belirli kısımlara "dikkat etmesini" sağlamaktadır. Dikkat mekanizması, özellikle "Sequence-to-Sequence" (Seq2Seq) modelleri ile birleştirildiğinde, metin çevirisi kalitesini önemli ölçüde arttırmaktadır.

2018 yılından itibaren "Pretrained Language Models" yani Önceden Eğitilmiş Dil Modelleri, DDİ alanında devrim yaratan bir gelişme olarak ortaya çıkmıştır. Bu modeller, büyük metin veri kümeleri üzerinde eğitilen derin öğrenme tabanlı dil modelleridir ve genellikle geniş bir kelime dağarcığına ve karmaşık dil yapısını anlayabilme yeteneğine sahiptirler. Öğrenilen bilgi, kelime vektörleri ile temsil edilmektedir (Peters vd., 2018). Bu sayede, modelin dil yapısını anlamasına olanak tanınmıştır. BERT, XLNet, T5 ve daha birçok popüler dil modeli, bu tür modellerin örneklerindedir. Bu modeller, metin sınıflandırma, dil çevirisi, metin oluşturma, duygu analizi, metin özetleme, soru-cevap sistemleri ve daha birçok DDİ görevini çözmek için kullanılmaktadırlar. Bu modellerin kullanım alanları Şekil 2.2.'de gösterilmektedir. Önceden Eğitilmiş Dil Modelleri, DDİ alanında büyük bir çığır açmış ve birçok uygulama ve endüstri için önemli bir araç haline gelmiştir (Howard ve Ruder, 2018).



Şekil 2.1. Doğal dil işlemenin kullanım alanları

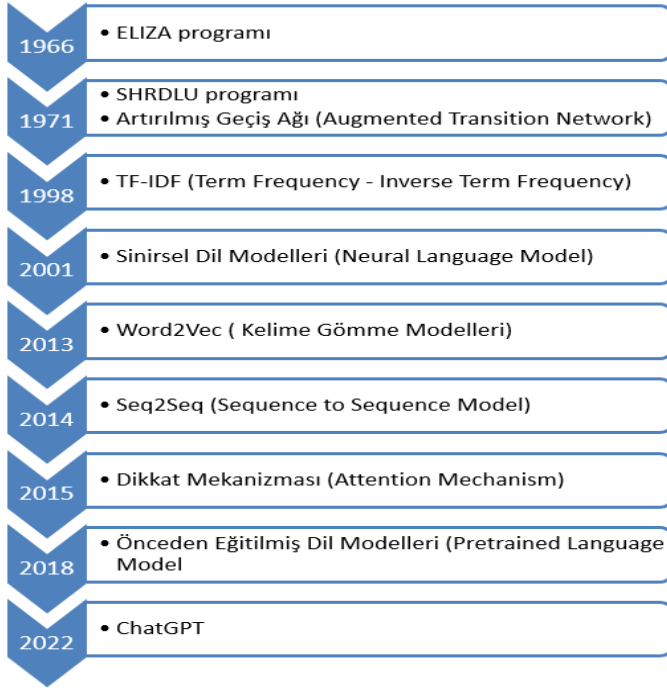
2022 yılında yayınlanan ChatGPT, OpenAI tarafından geliştirilen bir dil modelidir (Çiğdem, 2023). Bu model, büyük miktarda metin verisi üzerinde eğitilmiş ve DDİ alanında yapılan çalışmalara önemli bir yetenek kazandırmıştır (Keskin, 2023). ChatGPT'nin çalışma mimarisinde üç temel aşama bulunmaktadır; eğitim, ince ayar ve çıkarsamadır. Eğitim aşamasında ChatGPT, büyük metin verileri üzerinde eğitilmektedir. Eğitim aşamasından sonra model kullanılacak görevlere göre ince ayar yapılmaktadır. Son adımda yani çıkarsama aşamasında girdiye dayanarak bir sonraki kelime veya cümle tahmin edilmeye çalışılmaktadır. ChatGPT, kullanıcılarla doğal bir dilde sohbet edebilme ve metin tabanlı görevlere cevap verebilme yeteneğiyle öne çıkmaktadır.

Bert ve ChatGPT'nin her ikisi de dönüştürücü tabanlı mimari kullanmakta olmasına rağmen temelde farklı eğitim metodlarına sahiptirler. Bert modeli genellikle duygu analizi, metin sınıflandırma ve dil anlama gibi görevlerde, ChatGPT ise sohbet robotları ve sanal asistanlar oluşturmak için kullanılmaktadır (Afzal, 2023). Bert modelinin metin sınıflandırma gibi doğal dil işleme görevlerini gerçekleştirmek için tasarlanması yapılan tez çalışmasında ChatGPT yerine Bert ve türevlerinin kullanılmasını sağlamıştır. Aşağıdaki tabloda (Çizelge 2.1.) ChatGPT ve Bert'in amaç, kullandıkları mimari, performans ve kullanım alanları açısından karşılaştırmaları gösterilmiştir (Amplework, 2023).

Çizelge 2.1. ChatGPT ve BERT karşılaştırması (Amplework, 2023)

Özellik	ChatGPT	Bert
Amaç	İnsan benzeri metin yanıtları oluşturmak için etkileşimli bir yapay zeka modeli olarak tasarlanmıştır.	Duygu analizi, ve metin sınıflandırma dahil olmak üzere çok çeşitli doğal dil işleme görevlerini gerçekleştirmek üzere tasarlanmıştır.
Mimari	ChatGPT, yalnızca kod çözücü mimarisine sahip bir transformatör mimarisini temel almaktadır.	Bert, çift yönlü kodlamaya sahip bir transformatör mimarisine dayanmaktadır.
Performans	Konuşmaya dayalı yapay zeka uygulamalarında insan benzeri metin yanıtları oluşturmada güçlü bir performans sergilemektedir.	Varlık tanıma, metin sınıflandırma, duygu analizi de dahil olmak üzere çok çeşitli DDİ görevlerinde güçlü performans sergilemektedir.
Kullanım Alanları	ChatGPT, sohbet robotları ve sesli asistanlar gibi konuşmaya dayalı yapay zeka uygulamalarında yaygın olarak kullanılmaktadır.	Bert, duygu analizi, metin sınıflandırma da dahil olmak üzere çeşitli DDİ uygulamalarında yaygın olarak kullanılmaktadır.

Görsel 2.1.'de (Şekil 2.1.) DDİ alanında yapılan çalışmaların tarihsel gelişimi gösterilmektedir. Metin sınıflandırma, DDİ alanında önemli bir görevdir. Bu görev, bir metni belirli bir kategoriye veya sınıfa atama süreci olarak tanımlanmaktadır (Haltaş, 2015). Metin sınıflandırma, birçok uygulama için temel bir bileşendir ve bir metnin içeriğini anlamak, etiketlemek veya sınıflandırmak için kullanılır. Metin sınıflandırması çoğunlukla sosyal medya analizinde, ürün incelemesi yorumlarında ve müşteri geri bildirimleri vb. uygulamalarda kullanılmaktadır. Haber verilerini otomatik olarak kategorize etmek veya elektronik postaların spam veya spam olmayan (normal) olarak sınıflara ayırmak metin sınıflandırmasının önemli görevlerinden biridir.



Şekil 2.2. Doğal dil işleminin tarihçesi

Duygu analizi, metin verilerini inceleyerek içerdikleri duygusal içerikleri belirlemede kullanılan DDİ görevidir. Bu, genellikle bir metnin pozitif, negatif veya nötr bir duygu içerip içermediğini anlamak için yapılmaktadır. Duygu analizi tespiti şirketlerin müşteri yorumları ve geri bildirimleri üzerinde duygu analizi yaparak ürünleri veya hizmetleri hakkındaki müşteri memnuniyetini değerlendirmektedirler. Sosyal medya platformlarındaki kullanıcı yorumları ve paylaşımları üzerinde duygu analizi yapılarak, bir markanın veya ürünün popülerliği, itibarı ve genel algısı hakkında bilgi elde edilmektedir. İş yerlerinde duygu analizi sayesinde çalışanların içsel memnuniyeti tespit edilerek iş performansı üzerinde katkılar sağlanmaktadır. Finans haberleri üzerinde duygu analizi yapılarak, piyasalardaki duygusal değişimleri ve yatırımcıların hissiyatlarını anlamak mümkündür (Özcan, 2021). Duygu analizi, bu ve benzeri alanlarda bilinçli kararlar almak, müşteri memnuniyetini artırmak, kamusal algıyı anlamak gibi birçok uygulamada kullanılan önemli bir araç haline gelmiştir.

Adlandırılmış varlık tanımanın (Name Entity Recognition, NER), temel amacı metin içindeki belirli varlıkları (kişiler, yerler, organizasyonlar, tarihler, sayılar, para birimleri vb.) tanımlamak ve bu varlıklara etiketler atamaktır. Adlandırılmış varlık tanıma, bir metindeki

belirli bilgileri çıkarmak amacıyla metin madenciliği ve otomatik metin anlama, sosyal medya analizi, belge sınıflandırma vb. uygulamalarda önemli bir rol oynamaktadır (Nasiboğlu ve Gencer, 2023). Örneğin, bir metinde "Steve Jobs, Apple şirketinin kurucusudur" ifadesi bulunuyorsa, adlandırılmış varlık tanıma, "Steve Jobs" ismini bir kişi olarak, "Apple" şirketini bir organizasyon olarak tanımlamaktadır. Örneğin bir haber yazısında geçen kişilerin, tarihlerin veya yer adlarının tespit edilmesini sağlamaktadır. Metin içindeki önemli bilgileri otomatik olarak tanımlayarak, büyük miktarda veriye sahip metinlerden anlam çıkarma süreçlerini iyileştirmektedir.

DDİ içinde dil çevirisi, bir dilde yazılmış metni başka bir dilde yazılmış metne çevirme sürecidir (Güner, 2023). Temel amacı, bir dilde ifade edilen bilgiyi anlamak ve aynı anlamı koruyarak başka bir dilde ifade etmektir. Dil çevirisi, yazılı metinleri, belgeleri, web sayfalarını ve diğer çeşitli metin tabanlı içerikleri bir dilden diğerine dönüştürmektedir. Dil çevirisi, insan çevirisi ve makine çevirisi olarak ikiye ayrılmaktadır. İnsan çevirisi, uzman dil bilgisi ve kültürel anlayışa sahip insan çevirmenler tarafından gerçekleştirilen geleneksel çeviri sürecidir. Makine çevirisi, bilgisayar programları veya yapay zeka modelleri tarafından gerçekleştirilen otomatik çeviri sürecidir. Makine çevirisi, büyük veri setlerini kullanarak dil modellerini eğiten ve ardından yeni gelen metinleri çeviren sistemleri içermektedir. Google Translate gibi çeviri hizmetleri bu tür bir makine çevirisi olarak kullanılmaktadır. Dil çevirisi, küreselleşme, işbirliği, bilgi paylaşımı ve kültürlerarası iletişim gibi birçok alanda önemli bir rol oynamaktadır. Makine çevirisi alanında, derin öğrenme ve büyük dil modelleri gibi gelişmiş teknolojilerin kullanılmasıyla birlikte, çeviri sistemlerinin performansı önemli ölçüde artmıştır.

Yapay metin üretme, bilgisayarların belirli bir dilde anlamlı ve mantıklı bir biçimde metin oluşturmasını ifade etmektedir (Gui vd., 2021). Bu, yapay zeka ve DDİ tekniklerinin bir kombinasyonu kullanılarak gerçekleştirilmektedir. Yapay metin üretme, bir dizi farklı görevi içerebilir ve çeşitli uygulama alanlarında kullanılmaktadır. Gelişmiş yapay zeka modelleri, özellikle büyük dil modelleri, dilin yapısını ve bağlamını anlamak üzere eğitildiklerinden, özgün metin üretebilmektedir. Üretken Önceden Eğitilmiş Dönüştürücüler (Generative Pre-trained Transformer, GPT) gibi modeller, büyük bir metin veri setinden öğrenilmiş bilgileri kullanarak anlamlı ve akıcı metinler üretebilmektedir. Chatbotlar, kullanıcılarla doğal dilde etkileşime girmek ve anlamlı yanıtlar vermek için kullanılmaktadır

(İlaslan, 2023). Gelişmiş chatbotlar, karmaşık sorulara veya taleplere anlamlı cevaplar üretebilmektedir. Bazı yapay zeka uygulamaları, öykü yazma, şiir üretme veya şarkı sözleri oluşturma gibi yaratıcı yazılım görevlerinde kullanılmaktadır.

Otomatik yanıtlama, bilgisayar sistemlerinin kullanıcıların sorularına veya girdilerine anında ve otomatik olarak yanıt verme sürecini ifade etmektedir. Genellikle sohbet botları, müşteri hizmeti uygulamaları veya diğer etkileşimli sistemlerde kullanılmaktadır. Otomatik yanıtlama, yapay zeka ve dil işleme tekniklerini içermektedir. Otomatik yanıtlama, müşteri destek hattından sorulara cevap verme, eğlence amaçlı sohbet botları, çevrimiçi öğrenme platformlarındaki yardım sistemleri gibi birçok uygulamada kullanılmaktadır. Bu teknoloji, kullanıcılara hızlı ve etkileşimli bir deneyim sunma, bilgiye erişimi kolaylaştırma ve bazı basit görevleri otomatikleştirme konularında yardımcı olmaktadır.

### 3. TEORİK BİLGİ

Metin verilerinin doğal dil işleme alt görevlerinde kullanılabilmesi için bazı ön işleme adımlarından geçmesi gerekmektedir (Silay, 2020). Bu adımlar, metin verilerini daha anlamlı ve işlenebilir bir formata dönüştürmeyi amaçlamaktadır. Bu kısımda veri ön işleme, metin ön işleme ve yapılan tez çalışmasında kullanılan metin sınıflandırma modelleri anlatılmıştır.

#### 3.1. Veri Ön İşleme

DDİ'de kullanılacak olan verilerin belirli bir formata sahip olması gerekmektedir. DDİ uygulamalarında veri ön işleme, metin verilerinin karmaşıklığını azaltarak, verilerin analiz için daha uygun hale getirilmesini sağlayan önemli bir adımdır (Ersöz ve Çınar, 2021). Bu ön işleme adımları, metin verilerindeki gürültüyü azaltma, anlam karmaşıklıklarını ele alma ve dilin özünü daha iyi anlamak amacıyla gerçekleştirilmektedir. Yapılan işlemler sonucunda verinin kalitesi artarak model performansı iyileştirilmektedir.

Günümüzde kullanılmakta olan bir dizi veri ön işleme yöntemleri bulunmaktadır. Bu yöntemler sayesinde veri setleri modellerde kullanıma hazır hale getirilmektedir. Bu yöntemler veri temizleme (data cleaning), veri birleştirme (data integration), veri dönüştürme (data transformation) ve veri indirgeme (data reduction) olarak sıralanabilir (Oğuzlar, 2003).

**Veri temizleme:** Bu adımda veri setlerindeki hatalı, eksik, çelişkili veya anlamsız veriler tespit edilerek bunların veri setinden kaldırılması işlemi gerçekleştirilmektedir. Bu aşamada veri setindeki eksik verilerin doldurulması, aykırı değerlerin veri setinden hariç tutulması, tutarsız verilerin düzeltilmesi, veri setindeki gürültünün giderilmesi ve veri tiplerinin doğrulanması gibi işlemler yapılmaktadır.

**Veri birleştirme:** Veri birleştirme (data integration), farklı kaynaklardan gelen verileri birleştirerek tek bir veri tabanında veya veri setinde toplamayı amaçlayan bir veri ön işleme adıdır. Bu işlem, veri madenciliği, analitik çalışmalar ve iş zekası uygulamaları gibi veri odaklı işlemler için verileri daha kullanışlı ve bütünlük bir hale getirmek için kullanılmaktadır.

**Veri dönüştürme:** Veri dönüştürme (data transformation), veri setini daha uygun ve anlaşılır bir hale getirmek amacıyla verinin formatını veya yapısını değiştirme işlemidir. Veri dönüştürme aşamasında veri seti üzerinde veri düzeltme, veri birleştirme, veri genelleştirme ve veri normalleştirme gibi işlemler yapılmaktadır (Oğuzlar, 2003). Veri dönüştürme aşamasında, veriyi daha etkili bir şekilde işlemek için çeşitli dönüşümler uygulanmaktadır. Bu dönüşümlere veri setindeki sayısal değerlerin normalleştirilmesi, kategorik değişkenlerin sayısal formata dönüştürülmesi, tarih ve zamanın dönüşümü örnek verilebilir.

**Veri indirgeme:** Veri indirgeme (data reduction), büyük ve karmaşık veri setlerini daha küçük, yönetilebilir ve anlamlı bir formata dönüştürme sürecidir. Bu işlem, veri analizi ve işlem maliyetlerini azaltmak, daha hızlı analiz yapmak ve gereksiz karmaşıklığı ortadan kaldırmak amacıyla uygulanır. Veri indirgemedeki hedef verinin boyutunu küçülterek bellek ve işlemci gücünü optimize etmek, gereksiz veya az önemli bilgileri veri setinden çıkarmaktır. Veri indirgeme özellik seçimi, özellik çıkartma, veri kümeleme ve boyut azaltma gibi yöntemler kullanılarak yapılmaktadır.

### 3.2. Metin Ön İşleme

Metin ön işleme, doğal dil işleme veya metin madenciliği uygulamalarında metin verilerini daha anlamlı, işlenebilir ve analiz edilebilir hale getirmek için yapılan bir dizi işlemi ifade etmektedir. Metin ön işleme aşamasında karakterler sayısal hale dönüştürülerek makine diline çevrilmektedir (Ballı, 2021). Metin ön işlemede kullanılan yöntemler sözcüklere ayırma, küçük harfe çevirme, noktalama işaretlerinin silinmesi, sayıların silinmesi, etkisiz kelimelerin silinmesi ve kök bulma olarak sıralanabilir.

**Sözcüklere Ayırma:** Metin verilerini daha küçük ve işlenebilir parçalara bölme işlemidir. Metin belgeleri veya cümleler, kelimeler olarak ayrıştırılır. Her kelime, bir jeton (token) olarak adlandırılır. Sözcüklere ayırma işlemi, metin verilerini sayısal bir gösterimle işlemek ve metin analizi yapmak için gereklidir.

**Küçük Harfe Çevirme:** Veriyi normalize edebilmek için yapılan ilk çalışma metin içerisinde büyük-küçük harf düzensizliğini ortadan kaldırmaktır (Yılmaz ve Yumuşak, 2021). Metin verilerinde bulunan harf karakterlerini büyük harften küçük harfe dönüştürme işlemidir.

**Noktalama İşaretlerinin Silinmesi:** Dijital ortamda oluşturulmuş metinlerde bulunan noktalama işaretlerinin yanlış kullanılması veya anlamı bozmuş olma ihtimali bulunmaktadır (Adalı, 2012). Bu problemi ortadan kaldırmak için veri seti üzerinden noktalama işaretleri (örneğin, virgül, nokta, ünlem işareti, soru işareti, tırnak işaretleri, kesme işareti vb.) kaldırılmaktadır.

**Sayıların Silinmesi:** Metin verilerinden sayısal değerlerin çıkarılması işlemidir. Bu adım, metin verilerini temizlemek ve metin analizi sırasında sayıların gereksiz gürültüye neden olmasını önlemek amacıyla yapılmaktadır.

**Etkisiz Kelimelerin Silinmesi:** Metin verilerindeki sık kullanılan ve anlamsız kelimelerin, yani etkisiz kelime (stop words) olarak adlandırılan kelimelerin çıkarılması işlemidir. Etkisiz kelimeler, dilde sıkça kullanılan fakat anlam olarak cümleye katkı sağlamayan kelimelerdir. Türkçe için örnek etkisiz kelimeler "ve," "veya," "bir," "ben," "sen," "o," "bu," "şu," "ne," "ama," "şimdi," "ise," "şöyle," vb. kelimeleridir.

**Kök Bulma:** Kelimelerin köklerinin veya temel biçimlerinin bulunması işlemidir. Kelimenin farklı biçimleri (çoğul, zamirler vb.) aynı köke indirgenir. Kök bulma (stemming), metin verilerini daha basit bir formda temsil etmeyi ve aynı anlama gelen farklı varyasyonları birleştirmeyi hedefler. Özellikle metin madenciliği, metin sınıflandırma, duygu analizi ve bilgi çıkarma gibi metin analizi görevlerinde yaygın olarak kullanılmaktadır.

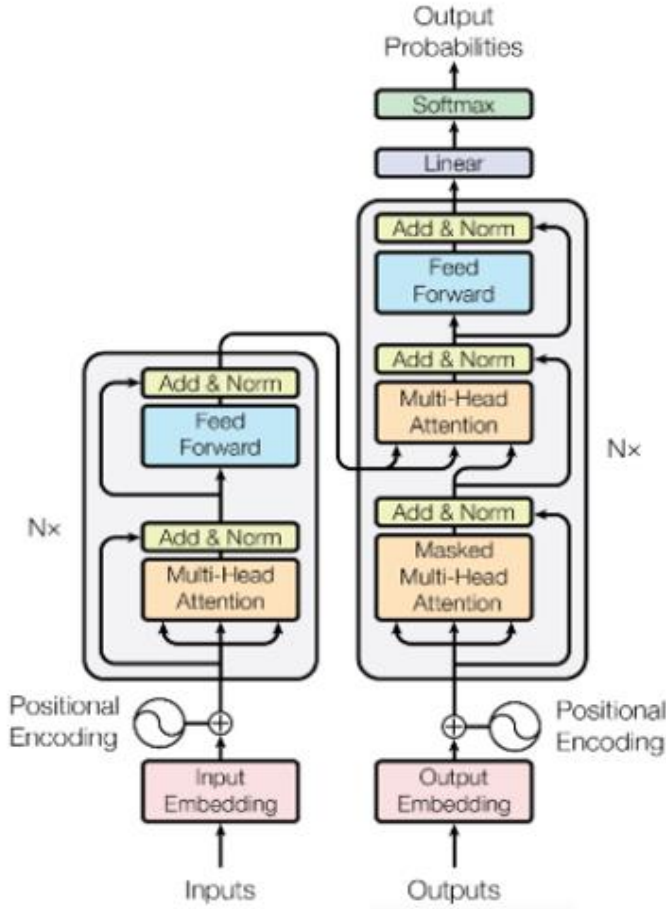
### 3.3. Çalışmada Kullanılan Modeller

Çalışmada iş isteklerinin hangi sınıfa ait olduklarını tahmin edebilmek için dönüştürücü tabanlı Bert, DistilBert, Electra ve ConvBert modelleri kullanılmıştır. Eğitim ve sınıflandırma işlemlerinden sonra modellerin sınıflandırma sonuçları açıklanabilir bir yapay zeka modeli olan Lime algoritması ile incelenmiştir.

#### **Dönüştürücü Mimarisi**

Dönüştürücüler (Transformers), 2017 yılında Google tarafından geliştirilen bir yapay zeka modelidir (Vaswani vd., 2017). Geliştirilen model sayesinde özellikle dil modelleri için büyük bir ilerleme kaydedilmiştir. Geleneksel doğal dil işleme modellerinde, büyük ve karmaşık veri setlerinin eğitim işlemi uzun sürmekte ve buna rağmen başarılı sonuçlar elde edilememektedir. Dönüştürücüler sayesinde karmaşık ve uzun metinler için de başarılı sonuçlar elde eden modeller geliştirilmiştir. Dönüştürücüler, metin verilerini daha iyi anlamak ve işlemek için kullanılmakta, birçok endüstri ve uygulama alanında büyük bir etki yaratmaktadır.

Dönüştürücü mimarisi Şekil 3.1.'de gösterilmektedir. Dönüştürücü olarak adlandırılan mimarinin solunda kodlayıcı (encoder) ve sağında ise kod çözücü (decoder) bloğu bulunmaktadır. Bir dönüştürücü içerisinde altı adet birbirinin aynısı olan kodlayıcı ve kod çözücü bloğu mevcuttur. Her kodlayıcı bloğu bir adet çok başlı dikkat (multi head attention) ve bir adet ileri beslemeli ağ (feed forward network) içermektedir. Her kod çözücü bloğu içerisinde iki adet çok başlı dikkat ve bir adet ileri beslemeli ağ bileşeni bulunmaktadır. Yapının içerisindeki her alt bileşenin etrafında bir artık bağlantı ve normalizasyon katmanı vardır (Sanh vd., 2019).



Şekil 3.1. Bir Dönüştürücünün Yapısı (Vaswani vd., 2017)

Dönüştürücü mimarisi incelendiğinde ilk olarak metin verilerinin model tarafından işlenebilir hale getirilmesi için kullanılan girdi kodlama (input embedding) kısmı bulunmaktadır. Bu aşamada girdi olarak verilen metinler sequence adı verilen sıralı dizi haline getirilmektedir. Elde edilen sıralı dizi belirteç oluşturucular (tokenizer) sayesinde belirteç (token) ve belirteçlerin numarikleştirilmiş hali olan belirteç numaralara (idlere) dönüştürülmektedir. Belirteç oluşturucular, metin işleme görevlerinde metin verilerinin işlenmesini ve model tarafından anlaşılmasını sağlamaktadır. Belirteç oluşturucular genellikle önceden eğitilmiş modellerle birlikte kullanılır ve metin verilerini modelin gereksinimlerine uygun bir şekilde işlerler. Belirteç oluşturucular sayesinde veri setindeki her bir kelime için belirteç ve belirteç numarası bilgileri elde edilmektedir. Şekil 3.2'de metin verisinin belirteç yapısına dönüştürülmüş hali, Şekil 3.3'de belirteç numarasına dönüştürülmüş hali gösterilmiştir.

Bugün hava güneşli olacak .

Şekil 3.2. Metin verisinin belirteç yapısına dönüştürülmüşü

3312 3543 6565 8778 6555

Şekil 3.3. Metin verisinin belirteç numarası yapısına dönüştürülmüşü

Belirteç ve belirteç numarası yapısına dönüştürülen kelimeler için bir sonraki aşamada pozisyon kodlaması yapılmaktadır. Bu aşamada her belirtecin anlamını temsil eden vektörlere pozisyon bilgilerinin bulunduğu sabit vektör eklenir. Bu sayede, her belirtecin pozisyonu metindeki diğer belirteçlerden ayırt edilmiş olmaktadır. Konumu tek sayı olan belirteçler için pozisyon kodlaması değerinin hesaplamasında kosinüs fonksiyonu, konumu çift sayı olan belirteçler için pozisyon kodlaması değerinin hesaplamasında ise sinüs fonksiyonu kullanılmaktadır (Vaswani vd., 2017).

$$PE_{(pos, 2i)} = \sin\left(\frac{pos}{100000\left(\frac{2i}{d_{model}}\right)}\right) \quad (3.1)$$

$$PE_{(pos, 2i+1)} = \cos\left(\frac{pos}{100000\left(\frac{2i}{d_{model}}\right)}\right) \quad (3.2)$$

Formülde kullanılan  $d_{model}$  boyutu ifade etmekte olup oluşturulan belirteç vektörleri ile aynı boyuttadır.  $i$  olarak belirtilen değer pozisyon kodlaması yapılacak kelimenin indisidir.  $Pos$  olarak belirtilen değer ise kodlaması yapılacak kelimenin pozisyonudur.

Pozisyon kodlaması hesabı tamamlandıktan sonra elde edilen konum yerleştirme matrisi, belirteç yerleştirme matrisi ile toplanarak kelimenin hem kimliğini ( $ID$ ) hem de pozisyonunu içeren matris elde edilmektedir. Elde edilen matris öz dikkat (self attention) kısmında bulunan dikkat (attention) adı verilen mekanizmada işleme alınarak, cümle

içerisindeki her bir kelimenin diğer kelimelerle olan ilişkisi yani dikkat değeri hesaplanmaktadır. Bu hesaplamalar yapılırken her kelime / belirteç için sorgu (query), anahtar (key) ve değer (value) vektörleri oluşturulmaktadır. Sorgu vektörleri ( $Q$ ), kelimenin diğer kelimelerle olan ilişkilerinin öğrenilmesi için kullanılmaktadır. Anahtar vektörleri ( $K$ ), sorgu vektörleri ile karşılaştırılarak diğer kelimelerle olan ilişkileri hesaplamak için kullanılmaktadır. Değer vektörleri, sorgu ve anahtar vektörleri arasındaki iç çarpım sonucu, dikkat puanı (attention score) oluşturmak için kullanılmaktadır. Aşağıdaki formül de dikkat skorunun hesaplanış formülü gösterilmektedir (Vaswani vd., 2017).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (3.3)$$

Formülde  $d_K$  olarak gösterilen değer anahtar matrisinin boyutudur. Dikkat puanı hesaplanırken sorgu vektörü ( $Q$ ) matrisi ve anahtar vektörü ( $K$ ) matrisinin transpozu çarpılarak, değer anahtar matrisinin boyutuna bölünmektedir. Elde edilen matris aktivasyon (softmax) fonksiyonu kullanılarak normalleştirilmektedir. Bu, her kelimenin diğer kelimelerle olan ilişkisini 0 ile 1 arasında bir olasılık dağılımı olarak ifade etmektedir. Bu normalleştirilmiş ağırlıklar, metindeki diğer kelimelerin değer vektörleri ile çarpılarak ağırlıklı toplam oluşturulmaktadır. Bu işlem, öz dikkat mekanizmasının temelini oluşturmakta ve her kelimenin diğer kelimelerle olan bağlantısını hesaplamak için kullanılmaktadır. Dikkat puanı, metin verilerinin içeriğini işlerken belirli bir kelimenin metindeki bağlamını ve etkileşimini belirlemenin önemli bir yolunu temsil etmektedir. Bağlam vektörü (context vector), cümledeki kelimelerin anlamları, konumları ve birbirleri ile aralarında olan ilişkileri gibi çok önemli bilgileri barındıran, öz dikkat mekanizmasının sonuç matrisidir. Bu vektör, modelin metin verilerini daha iyi anlamasına yardımcı olmaktadır.

Dönüştürücü mimarisinde çok başlı dikkat, bir kelimenin diğer kelimelerle olan ilişkisini hesaplarken öz dikkat mekanizmasının farklı konumlarından gelen bilgilerin kullanılması anlamına gelmektedir. Bu mekanizma, metin verilerinin daha kapsamlı ve karmaşık ilişkilerini ele almak için kullanılmakta ve dönüştürücü modelinin daha güçlü bir temsilasyon oluşturmaya yardımcı olmaktadır. Çok başlı dikkat, metindeki farklı ilişkileri ve bağlantıları yakalayabilen bir çoklu öğrenme yaklaşımıdır. Dönüştürücü modelinde çok

başlı dikkat, metin işleme görevlerinde daha yüksek başarı elde etmek için önemli bir bileşen olarak kullanılmaktadır.

Her dikkat başlığı için ayrı ayrı öğrenilen matrisler (sorgu, anahtar ve değer matrisleri) oluşturulmaktadır. Her başlık, kendi sorgu, anahtar ve değer matrislerini kullanarak kendi dikkat puanlarını hesaplamaktadır. Her başlığın dikkat puanları, birleştirilip toplanmaktadır. Bu, farklı başlıkların bilgilerini bir araya getirmektedir. Toplanan bilgiler, son bir birleştirme işlemine tabi tutulmaktadır. Bu, farklı başlıklardan gelen bilgilerin daha tutarlı bir temsilini oluşturmaktadır. Elde edilen temsil, modelin sonraki katmanlarına veya işlem katmanlarına iletilmektedir. Çok başlı dikkat skor değerinin hesaplanması aşağıdaki denklemde gösterilmiştir.

$$\begin{aligned} MultiHead(Q, K, V) &= Concat(head_1, \dots, head_h)W^O \\ \text{where } head_i &= Attention(QW_i^Q, KW_i^K, VW_i^V) \end{aligned} \quad (3.4)$$

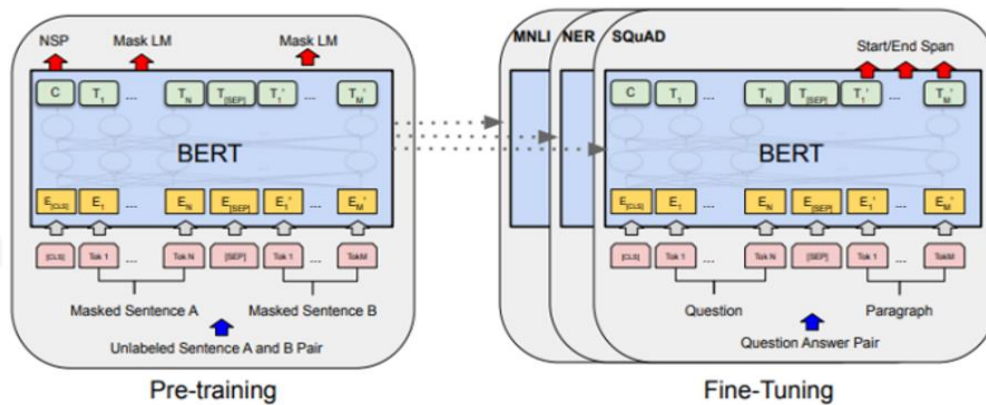
Dönüştürücü mimarisindeki ileri beslemeli ağ (feed-forward network), bir kelimenin veya belirtecin özelliklerini dönüştürmek ve daha karmaşık bir temsili oluşturmak için kullanılan bir bileşendir. Her kodlayıcı ve kod çözücü bloğu arasında tam bağlantılı bir ileri beslemeli ağ bulunmaktadır. İleri beslemeli ağ kendi içerisinde iki doğrusal dönüşüm katmanı ve bu katmanlar arasında bulunan relu aktivasyon kısmı barındırmaktadır (Sanh vd., 2019).

Çok başlı dikkat kısmından çıkan vektörler ileri beslemeli ağ kısmında giriş vektörleri olarak kullanılmaktadır. Giriş vektörleri, bir veya daha fazla gizli katmandan (fully connected layers) geçmektedir. Her gizli katman, ağırlık bilgileri ve aktivasyon fonksiyonları içermektedir. Bu katmanlar, kelimenin özelliklerini ve bağlamını daha fazla karmaşıklıkla dönüştürmeye yardımcı olmaktadır. Her gizli katmanın çıktıları, aktivasyon fonksiyonları tarafından işlenmektedir. Bu fonksiyonlar aracılığıyla temsillere lineer olmayan dönüşümler eklenerek temsillerin daha fazla esneklik kazanılması sağlanmaktadır. Son gizli katmandan sonra bir çıkış katmanı bulunmaktadır. Bu katman, kelimenin yeni temsilini oluşturmaktadır. Çıkış katmanından gelen sonuç temsili, kelimenin daha karmaşık bir temsildir. Bu temsil, kelimenin metindeki bağlamı ve özelliklerini daha iyi yansıtmaktadır. İleri beslemeli ağ, metin verilerinin daha yüksek düzeyli temsillerini

oluşturmak ve karmaşıklık eklemek için kullanılmaktadır. Bu sayede, model metindeki ilişkileri ve bağlantıları daha iyi anlayabilmektedir. Dönüştürücü modelinde, öz dikkat katmanları ile birlikte ileri beslemeli ağ, metin işleme görevlerinde kullanılmak üzere metin verilerini daha iyi işlemek için bir arada kullanılmaktadır. Bu, dil işleme görevlerinde büyük bir başarı sağlayan önemli bir bileşendir.

### 3.3.1. Bert

BERT (Bidirectional Encoder Representations from Transformers, Dönüştürücülerden Çift Yönlü Kodlayıcı Temsilleri), doğal dil işleme alanında önemli bir yapay zeka modelidir. BERT, Google tarafından geliştirilen ve 2018'de tanıtılan bir tür dönüştürücü modelidir (Devlin vd., 2019). Bert modelini diğer modellerden ayıran en önemli özellikleri cümleyi hem soldan sağa hem sağdan sola çift yönlü olarak taraması hem de kelimeleri rastgele maskeleyerek tahmin etmeye çalışmasıdır. Bert modeli cümle tahminleme, soru cevaplama, çeviri yapma, sonraki cümleyi tahminleme gibi 11 doğal dil işleme görevi için üstün başarı sergilemektedir (Devlin vd., 2019). Özellikle Google Arama gibi metin tabanlı hizmetlerin daha iyi anlaşılması ve kullanıcıların arama sorgularını daha iyi anlamak için BERT modelini kullanmaktadır (Hoş, 2023). Şekil 3.4. de BERT modelinin genel mimarisi gösterilmektedir.



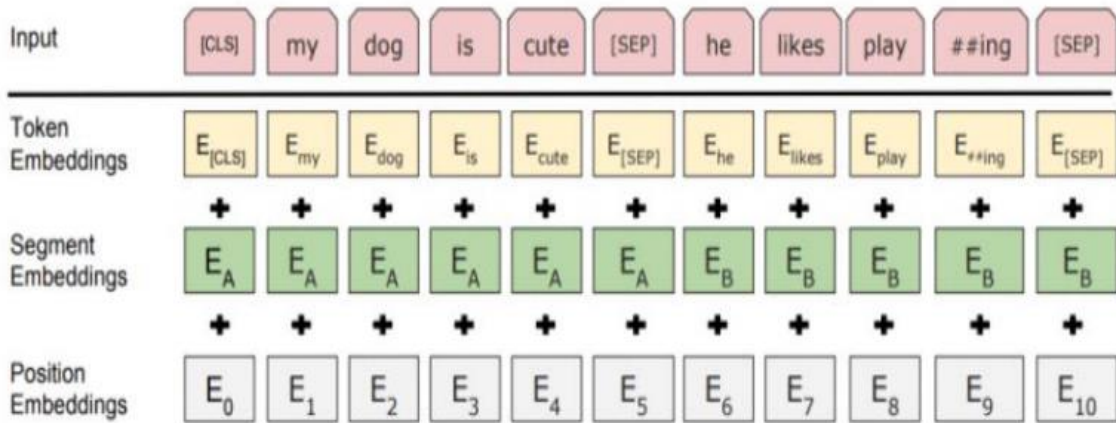
Şekil 3.4. BERT modelinin genel gösterimi (Devlin vd., 2019)

Bert modeli içerisinde 24 adet dönüştürücü bloğu, 1024 adet gizli katman ve 340M adet parametre bulunmaktadır. Model İngilizce kitaplardaki metinlerin veri kütüphanesi olan BooksCorpus (800 milyon kelime) ve İngilizce Wikipedia (2.5 milyar kelime) veri setindeki kelimeler kullanılarak toplam 3.3 milyar kelimelik bir korpus üzerinden 40 epoch ile önceden eğitilerek bert\_large ve bert\_base isimli iki model kullanıcılara sunmuştur. bert\_large modelinin eğitim işlemi aşamasında 16 TPU, bert\_base modelinin eğitim işleminde ise 4 adet TPU kullanılmıştır. Çizelge 3.1’de bert\_base ve bert\_large modellerinin karşılaştırması gösterilmiştir.

Çizelge 3.1. Bert modellerinin karşılaştırılması

	bert_base	bert_large
Transformatör Katmanları	12	24
Gizli Boyut	768	1024
Dikkat Başlıkları	12	16
Parametreler	110 milyon	340 milyon
İşleme	4 TPU	16 TPU
Eğitim Süresi	4 gün	4 gün

Bert dil modeli, girdi olarak verilen cümleleri WordPiece yapısını kullanarak kelime parçacıklarına ayırmaktadır. Şekil 3.5’de kelimelerin girdi temsillerine dönüştürülme adımları gösterilmiştir. Kelimeler, belirteçlere ayrıldıktan sonra her cümlenin başına [CLS], her cümlenin sonuna ise [SEP] belirteci eklenmektedir. Her belirtece hangi cümleye ait olduğunu anlayabilmek için segment bilgisi (segment embeddings) eklenmektedir. Son olarak kelimenin cümledeki konumunu belirleyebilmek için pozisyon bilgisi (position embeddings) eklenmektedir.



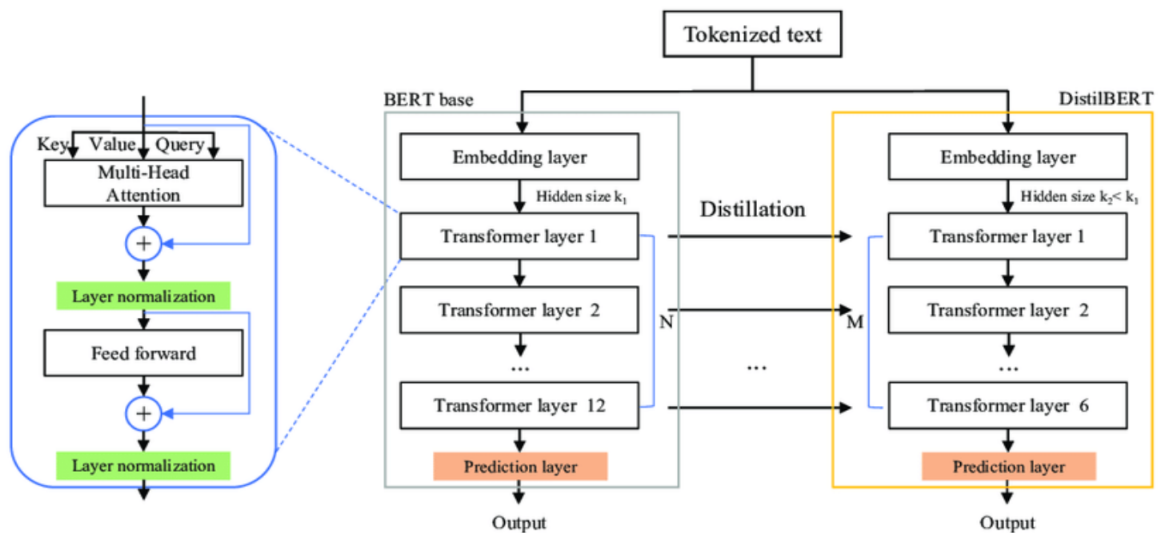
Şekil 3.5. BERT Girdi Temsili (Devlin vd., 2019)

Girdiler hazır hale getirildikten sonra cümleler Maskelenmiş Dil Modeli (Masked Language Modeling, MLM) ve Sonraki Cümle Tahmini (Next Sentence Prediction, NSP) adı verilen iki teknikte eğitilmektedir. Bu kısımda girdi olarak verilen cümledeki kelimeler içerisinde rastgele seçilen %15'lik kısmında maskelenmiş dil modeli tekniği kullanılmaktadır. Seçilen bu %15'lik kısım içerisindeki kelimelerin %80'i [MASK] etiketi ile, %10'u rastgele seçilen başka bir kelime ile kalan %10'lık kısımdaki kelimeler ise hiç değiştirilmeden bırakılmaktadır. Model bu maskelenmiş kelimelerin orijinal hallerini tahmin etmek için eğitilmektedir. Metindeki diğer kelimelerin bilgisi kullanılarak maskelenmiş kelimenin yerine hangi kelimenin en uygun olduğu hesaplanmaktadır. Buna maskelenmiş dil modeli denmektedir. Sonraki cümle tahmini kısmında iki cümle arasındaki ilişki tahmin edilmeye çalışılmaktadır. Bu kısımda girdi olarak iki cümle alınmaktadır. Eğitim aşamasının %50'sinde ikinci cümle değiştirilmeden %50'lik kısımda ise ikinci cümle rastgele değiştirilerek eğitim gerçekleştirilmektedir. Bu sayede eğitim işlemi tamamlandıktan sonra ikinci cümlenin gerçekten birinci cümleden sonra gelip gelmediği tahmin edilerek iki cümle arasındaki ilişki ortaya çıkarılmaktadır.

### 3.3.2. DistilBert

DistilBERT, ilk olarak 2019 yılında tanıtılmıştır (Sanh vd., 2019). Çalışmada araştırmacılar tarafından BERT modelinin daha küçük, daha hızlı, daha ucuz ve daha hafif bir versiyonunu oluşturmak amacıyla geliştirilen DistilBERT'in temel prensipleri ve

performansı ele alınmaktadır. DistilBERT (Distilled Bidirectional Encoder Representations from Transformers), doğal dil işleme alanında sık kullanılan bir öğrenme modelidir. DistilBERT, daha büyük ve karmaşık bir model olan BERT modelinden türetilmiştir. BERT, özgün metinleri anlamak ve işlemek için oldukça başarılı bir model olmasına rağmen çok sayıda parametreye sahip olduğu için eğitim işlemi uzun sürmektedir. DistilBERT, BERT modelinin özünü korurken, daha küçük bir boyutta ve daha hızlı bir şekilde eğitilebilir bir versiyonunu temsil etmektedir (Sanh vd., 2019). Bert modelinde bulunan 12 dönüştürücü katmanı sayısı DistilBERT modelinde altıya düşürülmüştür. Şekil 3.6’da DistilBERT modelinin genel mimarisi gösterilmiştir. DistilBERT, öğrenme sürecinde BERT modelinden gelen bilgiyi özütleyerek (distil ederek) daha az parametre ile aynı dil işleme yeteneklerini sunmaktadır. Bu sayede daha hızlı eğitim, daha düşük bellek ve işlemci gereksinimleri ve daha hızlı tahminler elde edilmektedir.

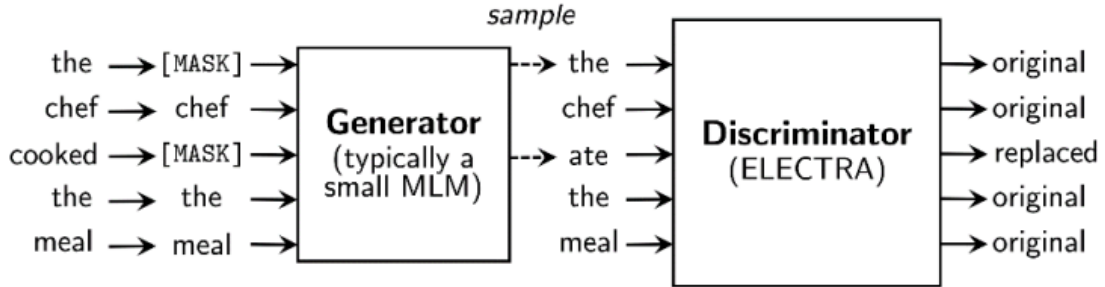


Şekil 3.6. DistilBERT modelinin genel mimarisi (Adel vd., 2022)

### 3.3.3. Electra

ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately) adlı dil modeli, 2019 yılında Google AI tarafından geliştirilmiştir (Manning vd, 2020). Geliştirilen model Bert modeli gibi maskeli dil modeli yerine bir jeneratör (generator) ağı kullanılarak kelimeleri benzer değerlerle değiştirmektedir. Ardından cümle içerisindeki değiştirilen kelimenin farklı bir değer ile değiştirilip değiştirilmediğini tahmin eden model

kullanılmaktadır, Şekil 3.7’de bu model gösterilmiştir. Bu yöntem sayesinde kelimeler arasındaki ilişkiler çıkarılmış olmaktadır.



Şekil 3.7. Electra modeli yapısı (Manning vd., 2020)

Electra dil modeli daha hafif olmasına rağmen diğer dil modellerine benzer veya daha iyi performans göstermektedir. Model daha az bellek ve hesaplama gücü kullanarak daha hızlı eğitim sağlamaktadır. Electra modelinin öğrenme yöntemi diğer dil modellerinin daha etkili ve verimli hale getirilmesine önemli bir katkı sağlamaktadır.

### 3.3.4. ConvBERT

ConvBERT modeli 2020 yılında yayınlanan makalede sunulmuştur. ConvBERT modeli dinamik evrişim ağları ile yerel bağılıklara dayanmaktadır (Jiang vd., 2020). Bert modelindeki dikkat blokları tüm cümleye odaklanarak hesaplamalar yapmaktadır. ConvBERT modelinde yerel bağımlılıkları doğrudan modellemek amacıyla öz-dikkat kafalarının yerini alan yeni bir yayılma tabanlı dinamik evrişim modeli önerilmektedir. Bu model sayesinde hem tüm metinde hem de yerel bağlamda öğrenmede daha verimli olan karma dikkat bloğu önerilmektedir. Deneyler, ConvBERT modelinin daha düşük eğitim maliyeti ve daha az model parametresi ile çeşitli aşağı yönlü görevlerde BERT ve varyantlarından önemli ölçüde daha iyi performansa sahip olduğunu göstermiştir.

### 3.3.5. Açıklanabilir Yapay Zeka

Yapay zeka modellerinin sonuçlarının insanlar tarafından anlaşılabilir bir ifadeye dönüştürülmesi açıklanabilir yapay zeka modelleri tarafından gerçekleştirilmektedir.

Açıklanabilir yapay zeka modelleri iki ana grupta toplanmaktadır. Bunlar tersine mühendislik (reverse engineering) ve şeffaf model tasarımı (transparent model design) olarak sınıflandırılmaktadır (Guidotti vd., 2018). Tersine mühendislik, bir yapay zeka modelinin iç işleyişini anlamak için kullanılan bir tekniktir. Bu yaklaşım, modelin kararlarını etkileyen öznitelikleri belirlemek ve modelin nasıl öğrendiğini anlamak için kullanılır. Şeffaf model tasarımında modelin kendisi baştan itibaren şeffaf ve anlaşılır bir yapıya sahip olacak şekilde tasarlanır. Karmaşık sinir ağları veya derin öğrenme modelleri yerine, daha basit ve anlaşılır modeller tercih edilmektedir. Bu sayede model kullanıcılar tarafından daha kolay anlaşılabilir ve yorumlanabilir hale gelmektedir. Her iki yaklaşım da açıklanabilir yapay zeka için önemli adımları temsil etmektedir. Tersine mühendislik, mevcut karmaşık modellerin anlaşılması için bir araç seti sunarken, şeffaf model tasarımı ise baştan itibaren kullanıcı dostu ve anlaşılır modellerin oluşturulmasını amaçlamaktadır. Bu teknikler, yapay zeka sistemlerini daha güvenilir ve kabul edilebilir hale getirerek, kullanıcıların bu sistemlere duydukları güveni artırmaya yardımcı olmaktadır.

### **3.3.6. Lime (Local Interpretable Model - Agnostic Explanations)**

LIME, karmaşık makine öğrenmesi modellerinin tahminlerini ve kararlarını daha anlaşılır ve yorumlanabilir hale getirmek için kullanılan bir tekniktir (Deliloğlu ve Pehlivanlı, 2021). Lime modelinin modelden bağımsız olması ve her modele uygulanabilmesi modelin kullanılabilirliğini artırmaktadır (Özel, 2020). Lime algoritmasının çalışma prensibinde veri setinden rastgele bir gözlem seçilerek bu gözleme benzer özellikte yeni gözlemler türetilir, ardından değiştirilmiş gözlemler orijinal model kullanılarak tahmin edilmektedir. Lime algoritması, değiştirilmiş gözlemleri kullanarak açıklanabilir bir model oluşturmaktadır.

## **3.4. İncelenen Çalışmalar**

Bu kısımda, dönüştürücü tabanlı DDİ metotlarından BERT, DistilBERT, ConvBERT ve Electra ile ilgili yapılan çalışmalar incelenmektedir. Bu modeller, derin öğrenme ve büyük metin verileri üzerine yoğunlaşan araştırma topluluklarının dikkatini çekmektedir. Yapılan çalışmalar, DDİ alanında büyük bir ilerleme kaydetmiş ve gelecekteki araştırmalara ilham kaynağı olmuştur.

Adoma ve diğerlerinin yaptıkları çalışmada International Survey on Emotion Antecedents and Reactions (ISEAR) veri seti üzerinde BERT, RoBERTa, DistilBERT, ve XLNet modelleri kullanılarak duyguları tanıyarak sınıflandırılmaya çalışılmıştır (Adoma vd., 2020). ISEAR veri seti, 37 ülkedeki kültürler arası anket çalışmaları yoluyla oluşturulmuş, kamuya açık bir veri setidir. Yedi farklı duygu etiketine göre sınıflandırılmış 7666 adet cümle içermektedir. Veri setinde bulunan duygu etiketleri sevinç, öfke, üzüntü, utanç, suçluluk, şaşkınlık ve korku olarak sıralanmaktadır. Kullandıkları veri seti dengeli bir veri seti olup, her kategoriye ait 1096 adet veri mevcuttur. Duygu etiketlendirme aşamasında özel karakterler, etiketler ve düzensiz ifadeler, etkisiz kelimeler performansı olumsuz etiketlediği için veri setinden hariç tutulmuştur. Veri setinin %80'i eğitim, %20'si test için ayrılmıştır. Veriler işlemeye hazır hale getirildikten sonra modeller üzerinde eğitilmiştir. Modellerin başarı oranları sırasıyla RoBERTa modelinde 0.7431, XLNet modelinde 0.7299, BERT modelinde 0.7009 ve DistilBERT modelinde 0.6693 olarak elde edilmiştir. Kullanılan veri seti üzerinde en yüksek tanıma doğruluğunun RoBERTa modeline ait olduğu gözlemlenmiştir.

Bir başka çalışma Guven (2023) tarafından yapılmıştır. Çalışmada spam içeren e-postaların tespiti için makine öğrenmesi ve dil öğrenmesi modelleri kullanılarak e-postaların spam olup olmadığı bilgisi tahmin edilmeye çalışılmıştır. Veri seti olarak Kaggle da bulunan Türkçe e-posta spam veri seti kullanılmıştır. Veri seti 502 adet spam olmayan, 517 adet spam olan veriden oluşmaktadır. Verilere model üzerinde eğitilmeden önce bazı ön işlemler uygulanmıştır. Bunlar; metinlerin küçük harfe dönüştürülmesi, metinlerin düzenli ifadeler (Regex) ile temsil edilmesi, sayıların ve noktalama işaretlerinin silinmesi, metinlere kök alma işleminin uygulanması, etkisiz kelimelerin silinmesi işlemleridir. İşlemeye hazır hale getirilen metinler makine öğrenmesi yöntemlerinden Rastgele Orman, Lojistik Regresyon, Naive Bayes, Yapay Sinir Ağları ve dil modelleri yöntemlerinden BERT, ELECTRA, ALBERT, DistilBERT kullanılarak eğitilmiştir. Eğitim işleminin sonunda makine öğrenmesi yöntemleri içerisinde en yüksek doğruluğu %90.15 değeri ile yapay sinir ağları elde ederken, dil modelleri içerisinde en yüksek doğruluk değeri %94.08 doğruluk değeri ile BERT ve ELECTRA modelinde elde edilmiştir. Yapılan çalışma sonucunda dil modellerinin klasik makine öğrenmesi yöntemlerine göre daha başarılı sonuçlar elde ettiği gözlemlenmiştir.

Bu çalışmada Türkçe tweet verilerinden cinsiyet tahmini eden sınıflandırma problemi ele alınmıştır. Çalışmada veri seti olarak farklı kişilere ait 100'er adet gönderi olmak üzere toplamda 5292 adet tweet kullanılmıştır. Verilerin 3368 adedi eğitim, 1924 adedi test olacak şekilde veri seti ikiye bölünmüştür. Metinlerin hepsi küçük harfe çevrildikten sonra modellerde uygulanmıştır. Noktalama işaretleri, emojiler cinsiyet tahminlemede belirleyici özellikler olabileceğinden dolayı veri setinden çıkarılmamıştır. Veri seti işlemeye hazır hale getirildikten sonra problem ikili sınıflandırma olarak ele alınmıştır. Kadınlar için 0, erkekler için 1 etiketi kullanılmıştır. Çalışmada makine öğrenmesi yöntemlerinden Terim Frekansı (Term Frequency, TF) – Ters Belge Frekansı (Inverse Document Frequency, IDF), derin öğrenme yöntemlerinden Uzun Kısa Süreli Bellek (Long Short – Term Memory, LSTM), Evrişimsel Sinir Ağları (Convolutional Neural Network, CNN) ve Türkçe için ön eğitilmiş dil modellerinden Bert, DistilBert, Electra modelleri kullanılmıştır. Yapılan çalışma sonucunda en yüksek başarı değeri Bert modeli ile edilmiştir. Çalışma sonucunda dil modellerinin klasik makine öğrenmesi yöntemlerine ve derin öğrenme yöntemlerine göre doğal dili anlamada başarılı olduğu gözlemlenmiştir (Sel vd., 2021).

Bir başka çalışmada ise kullanıcı yorumları üzerinden ürün kategorisi tahmin edilmeye çalışılmıştır. Çalışmada veri seti olarak bir firmaya ait çevrimiçi satış platformundaki kullanıcı yorumları, ürün kategorileri ve ürün tanımları kullanılmıştır. Veri setinde 77 farklı ürün kategorisine ait 36 bin satır veri bulunmaktadır. Veri setinden gereksiz kelimeler, emojiler ve bilgi içermeyen etiketler hariç tutulmuştur. 5 kelimedenden az kelime içeren yorumların kategori belirlemede belirleyici olmadığı gözlemlendiği için veri setinden 5 kelimedenden az kelime bulunan veri satırları silinmiştir. Veri setinin %80'i eğitim, %20'si ise doğrulama ve test için kullanılmıştır. Veriler hazır hale getirildikten sonra modeller üzerinde eğitilmiştir. Çalışmada dil modellerinden Bert, Electra ve Roberta modelleri kullanılmıştır. Çalışmada Bert modelinin diğer modellere daha yüksek başarı değeri elde ettiği gözlemlenmiştir (Köksal vd., 2022).

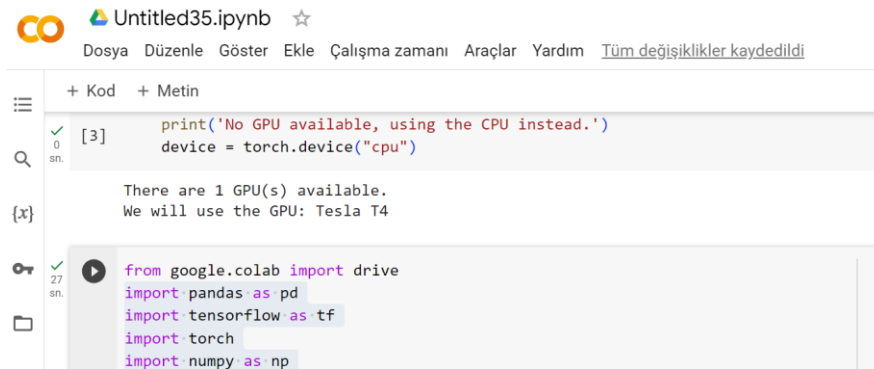
Deliloğlu ve Pehlivanlı (2021) çalışmalarında yapay zeka modelinin açıklanabilirliği üzerinde çalışmışlardır. Yapılan çalışmanın amacı açıklanabilir bir yapay zeka tasarlamak ve ardından farklı veri setleri üzerinde bu tasarlanan modeli çalıştırarak modelin karar vermede kullandığı parametreleri belirleyebilmektir. Çalışma üç aşamadan oluşmaktadır. İlk

aşamada şeffaf model olarak belirtilen açıklanabilirliği yüksek yöntemlerden karar ağaçları, ikinci aşamada hibrit model olarak belirtilen performansı yüksek fakat açıklanabilirliği düşük kara kutu modellerinden rastgele orman metodu ve son aşamada ise açıklanabilirlik için Lime algoritması kullanılmıştır. Çalışmada Taiwan’da bulunan bir bankaya ait kredi kartı verileri kullanılmıştır. Veri setinde 23 bağımsız ve 1 bağımlı değişken bulunmaktadır. Veri setindeki 5000 veri doğrulama için ayrılmıştır. Kalan 25000 veri kendi içerisinde %60’ı eğitim, %40’ı test için ayrılmıştır. Eğitim aşamasından sonra Karar ağacı modelinde %79.89, Rastgele orman modelinde %80.3 ve hibrit modelde (Karar Ağacı ve Rastgele Orman) %80.2 oranında başarı elde edilmiştir. Doğrulama veri seti üzerinden seçilen 100 adet örnekleme Lime algoritması uygulanarak modelin başarısını etkileyen değişkenler araştırılmıştır.

Modelin açıklanabilirliği ile ilgili yapılan bir başka çalışma ise Gabbay ve diğerlerine (2021) aittir. Çalışmanın amacı hastaların tıbbi geçmişine ve laboratuvar test sonuçlarına dayanarak COVID-19 hastalığının düzeyini makine öğrenmesi yöntemleri ile tahmin etmek ve ardından bu tahmini yorumlayabilmektir. Çalışmada veri seti olarak “Mexican Federal Health Secretary through the General Director of Epidemiology.” kullanılmıştır. Bu veri setinde Mayıs 2020’den 28 Ekim 2020’ye kadar toplanan tüm hastaların kayıtları mevcuttur. Yapılan çalışmada COVID-19 tanısı konmayan hastaların verileri hariç tutularak veri setinin %70’i eğitim %30’u test için kullanılmıştır. Çalışmada çok katmanlı algılayıcı (Multi Layer Perceptron, MLP) ve rastgele orman karar ağaçları olmak üzere iki tür makine öğrenmesi modeli kullanılmıştır. Modellerin çıktı katmanını bir nörona sahip ise Model A, çıktı katmanında üç nörona sahipse Model B olarak adlandırılmıştır. Eğitim aşamasından sonra modellerde elde edilen başarı oranları MLP Model A %80, MLP Model B %78, rastgele orman Model A %80 ve rastgele orman Model B %65 olarak gözlemlenmiştir. Sonrasında modellerin açıklanabilirliği için MLP A modelinde LIME algoritması kullanılarak modelin karar verirken hangi değişkene daha yüksek düzeyde önem verdiği gözlemlenmiştir.

#### 4. DENEYSEL ÇALIŞMA

Çalışmanın kodlama kısmı Python yazılım dili kullanılarak yazılmıştır. Çalışma Google Colab ortamında geliştirilmiştir. Google Colab (Colaboratory), Google tarafından 2017 yılında sunulan ücretsiz bulut tabanlı Jupyter Notebook ortamıdır (Google AI Hub, 2022). Şekil 4.1’de Google Colab yazılım geliştirme ortamı gösterilmiştir. Bu platform, Python programlama dilini kullanarak veri analizi, makine öğrenimi, derin öğrenme ve genel programlama gibi birçok görevin gerçekleştirilmesini sağlamaktadır. Colab, Google’ın bulut altyapısını kullanarak kullanıcılara ücretsiz Grafik İşlemci Ünitesi (Graphics Processing Unit, GPU) ve Tensör İşleme Birimi (Tensor Processing Unit, TPU) kaynakları sunmaktadır. Bu sayede Colab özellikle yüksek hesaplama gücü gerektiren makine öğrenimi ve derin öğrenme projeleri için uygun hale gelmektedir. Gerçekleştirilen çalışmada Tesla T4 GPU’su kullanılmıştır. Tesla T4, NVIDIA tarafından üretilen bir grafik işlem birimidir. Bu GPU modeli, geniş bir CUDA çekirdek sayısına ve 8 GB HBM2 belleğe sahiptir, bu da paralel hesaplamalar ve büyük veri setleri ile çalışma yetenekleri açısından etkili bir performans sağlamaktadır.



```

CO Untitled35.ipynb ☆
Dosya Düzenle Göster Ekle Çalışma zamanı Araçlar Yardım Tüm değişiklikler kaydedildi

+ Kod + Metin

[3] print('No GPU available, using the CPU instead.')
     device = torch.device("cpu")

There are 1 GPU(s) available.
We will use the GPU: Tesla T4

from google.colab import drive
import pandas as pd
import tensorflow as tf
import torch
import numpy as np
  
```

Şekil 4.1. Google Colab ortamı

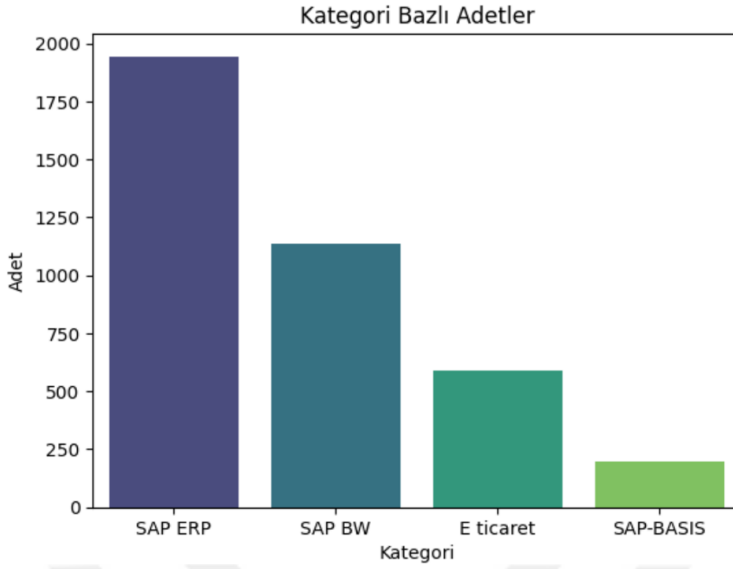
Çalışmada veri seti olarak işletme içerisinde kullanılan iş istek uygulamasındaki iş istek maddeleri kullanılmıştır. Uygulamada kullanıcılar iş istek açıklamalarını girerek ve iş isteğin hangi kategoriye ait olduğu bilgisini seçerek iş istek oluşturmaktadır. İş istek oluşturulduktan sonra hangi kategoriye açıldıysa sistemde o kategorinin sorumlusu olarak eşleştirilen sorumlu kişilere bildirim gitmektedir. Kullanılan veri seti iki sütundan

oluşmaktadır, ilk sütunda iş isteklerin kategori bilgileri yer alırken ikinci sütunda iş isteğine girilen açıklamalar mevcuttur. Şekil 4.2’de veri setinin yapısı gösterilmiştir.

	<b>category</b>	<b>text</b>
0	SAP ERP	firma çeklerinin standart ekrandan girilip oto...
1	SAP ERP	hurdaya çıkış fireye çıkış ürün hareketleri çık...
2	SAP ERP	depolar arası transfer çıktısı çift adım resmi...
3	SAP ERP	fatura girişinde ödeme temel tarihin bulunan s...
4	SAP ERP	satılacak malzemeler için malzeme türü mamül ti...

Şekil 4.2. Kullanılan veri seti

Veri setinde toplamda 3869 adet veri mevcuttur. Veri setinin 3095 adedi eğitim, 386 adedi doğrulama ve 388 adedi test için kullanılmıştır. Her kategoriye açılan iş istek sayısı yeteri kadar olmadığı için veri setinin kategori bazlı dağılımı eşit değildir. En çok veriye sahip kategoriler seçilerek uygulamada kullanılmıştır. Bunlar “SAP ERP”, “SAP BW”, “E ticaret” ve “SAP-BASIS” kategorileridir. “SAP ERP” kategorisine ait 1944, “SAP BW” kategorisine ait 1137, “E ticaret” kategorisine ait 588 ve “SAP-BASIS” kategorisine ait 200 adet veri bulunmaktadır. Veri setinin kategori bazlı dağılımı şekil 4.3’te gösterilmektedir.



Şekil 4.3. Veri setinin kategori bazlı dağılımı

Verilerin modellerde kullanılabilir hale getirilebilmesi için veri seti üzerinde bazı ön işleme adımları gerçekleştirilmiştir. Veri setindeki açıklama sütunundaki veriler öncelikle küçük harfe dönüştürülmüştür, sonrasında duraklama işaretleri, sayılar ve özel karakterler metinlerden çıkarılmıştır. Son olarak veri setinden Türkçe dilinde bulunan etkisiz kelimeler hariç tutulmuştur. Şekil 4.4’de örnek olarak seçilen bir metnin ön işleme yapılmadan önceki hali, Şekil 4.5’de ise ön işleme adımları yapıldıktan sonraki hali gösterilmiştir.

```
deger = df['text'].iloc[100]
print(deger)
```

ZOEQM016 raporuna Z\_Desen karakteristiği eklenecektir Zoesd004 sipariş çıktısına yeni bir sipariş türü eklemesi yapılacaktır

Şekil 4.4. Metin ön işleme yapılmadan önce veri

```
deger = df['text'].iloc[100]
print(deger)
```

zoeqm raporuna zdesen karakteristiği eklenecektir zoesd sipariş çıktısına yeni bir sipariş türü eklemesi yapılacaktır

Şekil 4.5. Metin ön işleme yapıldıktan sonra veri

Verilerin modellerde kullanılabilmesi için kategori sütununda bulunan değerler de sayısal ifadelerle temsil edilmiştir. Çizelge 4.1’de kategorilerin sayısal halleri gösterilmiştir.

Çizelge 4.1. Kategorilerin sayısal gösterimi

Kategorilerin Metinsel Hali	Kategorilerin Sayısal Hali
SAP ERP	0
SAP BW	1
E ticaret	2
SAP-BASIS	3

Tüm modellerde optimizasyon olarak AdamW optimizasyon algoritması kullanılmıştır. Optimizasyon algoritmaları modellerdeki hata oranını en aza indirmek için kullanılan algoritmalar (Seyyarer vd., 2020). AdamW, "Adam with Weight Decay" ifadesinin kısaltmasıdır. Bu optimizasyon algoritması, özellikle derin öğrenme modellerinde kullanılan bir tür stokastik gradyan iniş (Stochastic Gradient Descent, SGD) optimizasyon algoritmasıdır. Tüm modellerde öğrenme oranı (learning rate) değeri “ $2e-5$ ”, iterasyon sayısı 3, batch size değeri 16 olarak belirlenmiştir. Öğrenme oranı (learning rate), bir makine öğrenimi modelinin eğitimi sırasında kullanılan önemli bir hiperparametredir. Bu değer, her eğitim iterasyonunda model parametrelerinin güncellenme miktarını belirler. Yani, öğrenme oranı, bir tür "adım büyüklüğü" olarak düşünülebilir; ne kadar büyük bir adım atılacağını kontrol eder. Öğrenme oranı, genellikle deneme-yanılma yöntemleri veya otomatik hiperparametre ayarlama teknikleri kullanılarak belirlenir. Doğru öğrenme oranını bulmak, modelin hızlı ve etkili bir şekilde öğrenmesini, aynı zamanda aşırı uyumdan kaçınmasını sağlamak için önemlidir. İterasyon (epochs), bir makine öğrenimi modelinin eğitimi sırasında tüm eğitim veri setinin bir kez geçirilmesini ifade eden bir terimdir (Cansu, 2023). Eğitim süreci boyunca, model genellikle birden çok epoch boyunca eğitilir, bu da ağırlıkların ve parametrelerin veri setindeki desenlere daha iyi uyarlanmasını sağlar. Her epoch, modelin tüm eğitim veri setini bir kez gördüğü ve güncellediği bir dönemdir. Bir epoch tamamlandığında, model, belirli bir öğrenme oranı kullanarak gerçek değerlerle tahminleri karşılaştırarak kaybını (hata) hesaplar ve bu kayba göre modelin parametrelerini günceller. Bu süreç, belirlenen epoch sayısına kadar devam eder. Batch size, bir makine öğrenimi

modelinin eğitimi sırasında kullanılan bir hiperparametredir ve bir seferde işlenen örnek sayısını belirtir. Eğitim veri seti, genellikle büyük miktarda veri içerir ve tüm veriyi bir seferde işlemek hem bilgisayar belleği hem de hesaplama gücü açısından maliyetli olabilir. Bu nedenle, eğitim veri setini küçük parçalara, yani batch'lere ayırmak yaygın bir uygulamadır. Tüm modellerde karakter uzunluğu olarak 48 karakter alınmıştır. Metinler jetonlara dönüştürülerek Şekil 4.6'da gösterilmiştir.

```
Original: firma çeklerinin standart ekrandan girilip otomatik ödeme programına bağlanarak çek bilgilerinin doldurulması çek bordro çıktı formunun oluşturulması
Token IDs: tensor([ 2, 3699, 96033, 1927, 4938, 38608, 111501, 5774, 5003,
 9912, 30434, 2445, 18742, 37868, 2445, 31433, 3984, 40371,
13545, 3, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0])
Attention Mask: tensor([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0])
```

Şekil 4.6. Cümlelerin jetonlara dönüştürülmüş hali

Bert için dönüştürücülerden 'dbmdz/bert-base-turkish-128k-uncased', DistilBERT için 'dbmdz/distilbert-base-turkish-cased' modeli, Electra için 'dbmdz/electra-base-turkish-cased-discriminator' ve ConvBERT için 'dbmdz/convbert-base-turkish-cased' modeli kullanılmıştır. Bert modeli mevcut versiyonunun eğitiminde Türkçe Oscar derleminin filtrelenmiş ve cümle bölümlü versiyonu, güncel Wikipedia verileri, çeşitli OPUS derlemleri Ve Kemal Oflazer tarafından sağlanan özel bir derlem kullanılmıştır (Ozan vd., 2021). DistilBERT modeli Bert modelinin damıtılmış halidir. BERTurk modelinde kullanılan eğitim verisi ile RTX 2080 TI üzerinde 5 gün boyunca eğitilmiştir. Electra modeli, BERTurk modelinde kullanılan eğitim verisi ile eğitilmiştir. ConvBERT modeli, v3-32 TPU üzerinde 512 cümle uzunluğuyla 1M adımda BERTurk modelinde kullanılan eğitim verisi ile eğitilmiştir.

Çalışmanın ikinci kısmında veri seti ConvBERT modeli batch size değeri 16 seçilerek 3 iterasyonda eğitilmiştir. Eğitim işlemi sonucunda modelin kategori tahminlemesini açıklayabilmek için Lime algoritması kullanılmıştır.

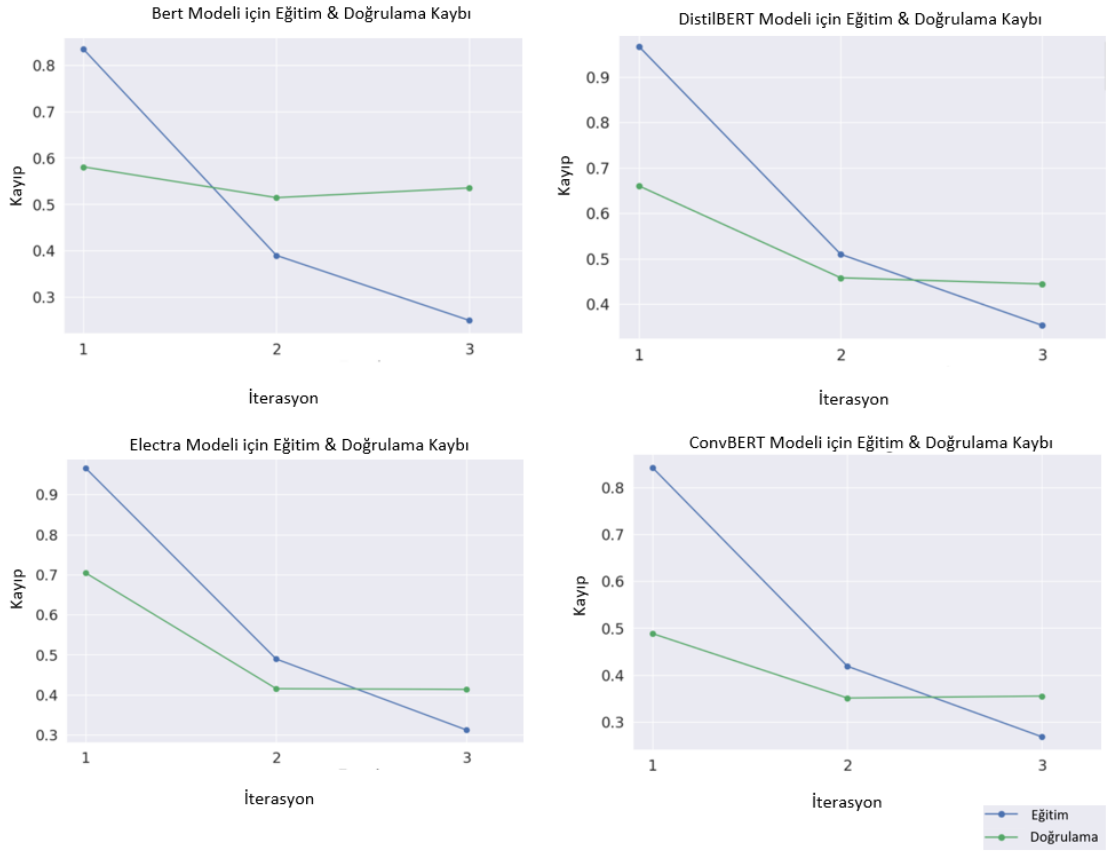
## 5. BULGULAR VE TARTIŞMA

Eğitim işlemleri sonucunda elde edilen eğitim kaybı, doğrulama kaybı ve doğrulama başarısı değerleri model ve iterasyon bazlı olarak Çizelge 5.1’de gösterilmiştir. Her iterasyon aşamasında doğruluk değeri arttığı ve yeteri kadar doğruluk elde edildiği için çalışma 3 iterasyonda tamamlanmıştır.

Çizelge 5.1. İterasyon bazlı modellerin eğitim sonuçları

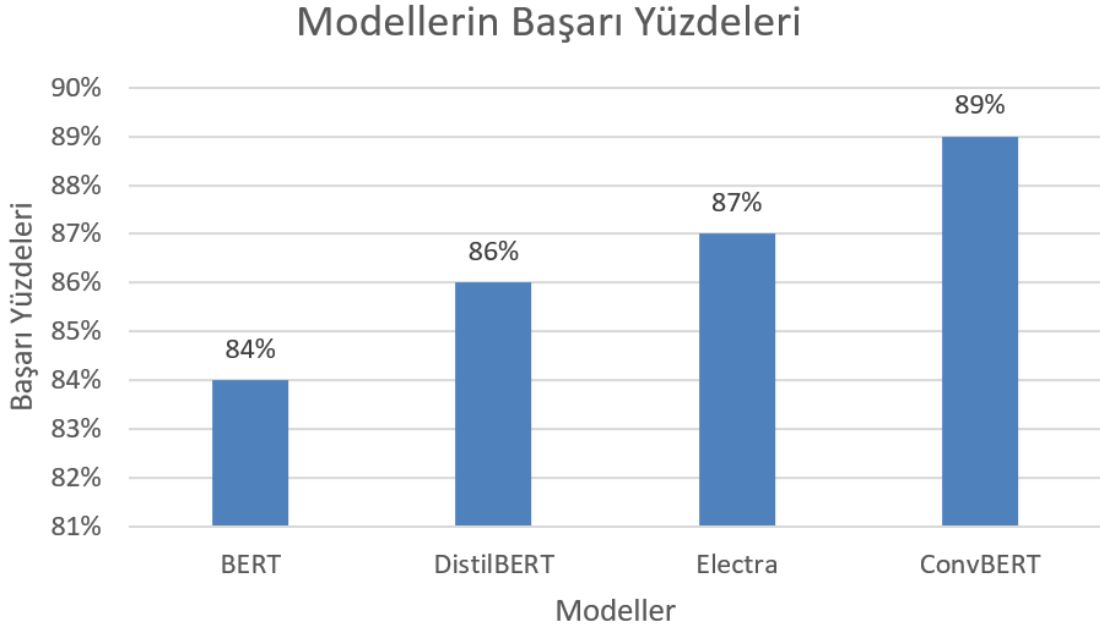
		İterasyon 1	İterasyon 2	İterasyon 3
BERT	Eğitim Kaybı	83%	39%	25%
	Doğrulama Kaybı	58%	51%	54%
	Doğrulama Başarısı	78%	83%	85%
DistilBERT	Eğitim Kaybı	97%	51%	35%
	Doğrulama Kaybı	66%	46%	44%
	Doğrulama Başarısı	78%	85%	86%
Electra	Eğitim Kaybı	96%	49%	31%
	Doğrulama Kaybı	70%	42%	41%
	Doğrulama Başarısı	76%	86%	87%
ConvBERT	Eğitim Kaybı	84%	41%	26%
	Doğrulama Kaybı	49%	35%	35%
	Doğrulama Başarısı	85%	88%	89%

Modellerin 3 iterasyon sonucunda kayıp değerlerinin başlangıca göre düşmesi modelin eğitim sürecinin başarılı bir şekilde gerçekleştirildiğini göstermektedir. Kayıp değeri, modelin tahmin ettiği değer ile gerçek değer arasındaki farkı ölçen bir metriktir. Bu değer ne kadar düşükse, model o kadar iyi çalışıyor anlamına gelmektedir. Aşağıdaki Şekil 5.1’de modellerin her iterasyon aşaması sonunda elde ettikleri eğitim ve doğrulama kayıp değerleri gösterilmiştir.



Şekil 5.1. Tüm modeller için her iterasyon aşamasındaki eğitim ve doğrulama kayıpları

Yapılan çalışmalar sonucunda en yüksek doğruluk değeri ConvBERT modelinde elde edilmiştir. Yapılan çalışmadaki görev için ConvBERT modelinin, BERT, Electra ve DistilBERT modellerinden daha iyi performansa sahip olduğu gözlemlenmiştir. Şekil 5.2’de modellerin başarı yüzdeleri gösterilmiştir.



Şekil 5.2. Modelleri başarı yüzdeleri

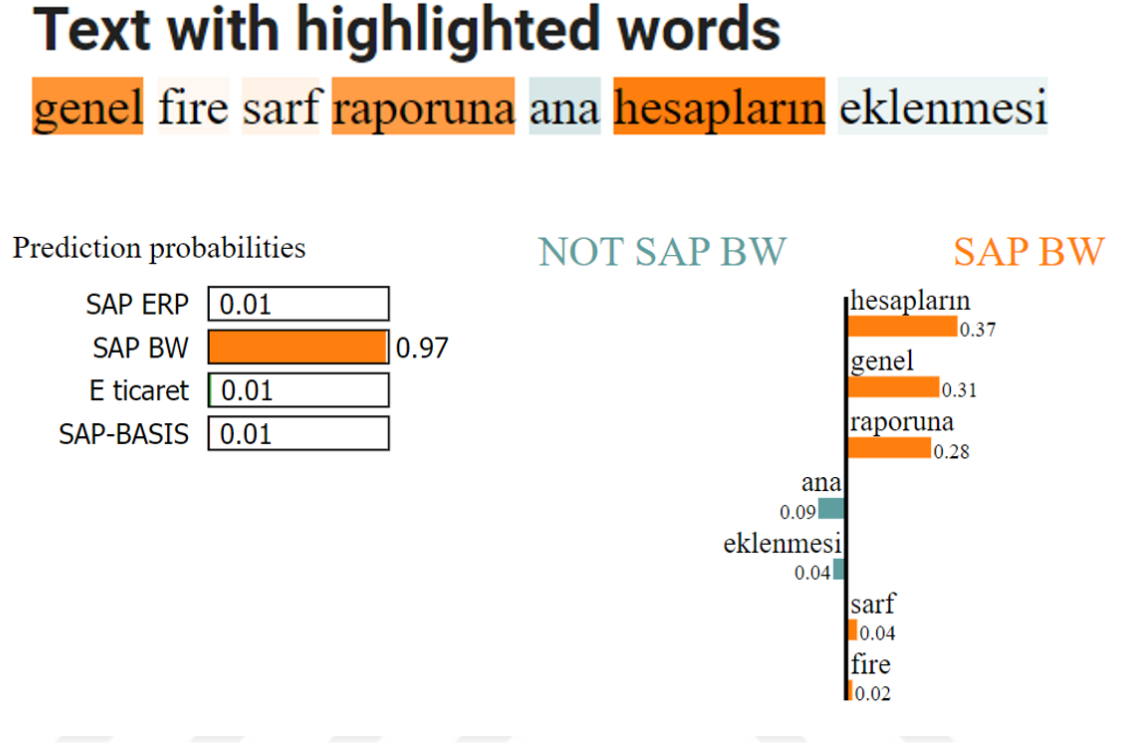
Eğitim işlemi sonucunda %89'luk doğruluk değeri elde edilmiştir. Çizelge 5.2'de ConvBERT modelinin üç iterasyon sonucunda elde ettiği başarı yüzdeleri gösterilmiştir.

Çizelge 5.2. ConvBERT modelinin eğitim sonucu

	Eğitim Kaybı	Doğrulama Kaybı	Doğrulama Başarısı
İterasyon 1	%88	%60	%77
İterasyon 2	%44	%41	%88
İterasyon 3	%28	%40	%89

Çalışmanın ikinci kısmında örnek olarak seçilen cümlelerin tahminlemesine etki eden kelimeler Lime algoritma sayesinde elde edilerek aşağıdaki görsellerde gösterilmiştir. Şekil 5.3'de "SAP BW" kategorisine ait olduğu tahminlenen bir cümlenin hangi kelimeler sayesinde bu kategoriye ait olarak tahminlendiği gösterilmektedir. Seçilen cümle %97 oranında SAP BW kategorisi olarak tahminlenmiştir. Cümlenin bu kategoriye tahminlenmesinde etkisi olan kelimeler aşağıdaki görselde turuncu renklerle gösterilmiştir. Bu cümleyi Lime algoritması ile incelediğimizde "genel, raporuna, hesapların"

kelimelerinin cümleyi SAP BW kategorisine tayin etmeye yüksek oranda etkili olduğu ortaya çıkmıştır.



Şekil 5.3. SAP BW kategorisine ait bir cümlelinin Lime algoritması ile incelenmesi

Bir başka örnekte “SAP ERP” kategorisine ait olduğu tahminlenen bir cümlelinin Lime algoritması ile incelenmesi Şekil 5.4’de gösterilmektedir. Cümle %97 oranında “SAP ERP” kategorisinde tahmin edilmiştir. Cümlelinin bu kategoride tahmin edilmesinde “zorfiw, istenmeyen, kaldırılmalıdır” kelimelerinin etkili olduğu aşağıdaki resimde turuncu renklerle gösterilmiştir. Şekilde “raporunda, yatırım, cariler, detay” gibi kelimelerinde “SAP BW” kategorisi için etkili olduğu gözlemlenmiştir.

### Text with highlighted words

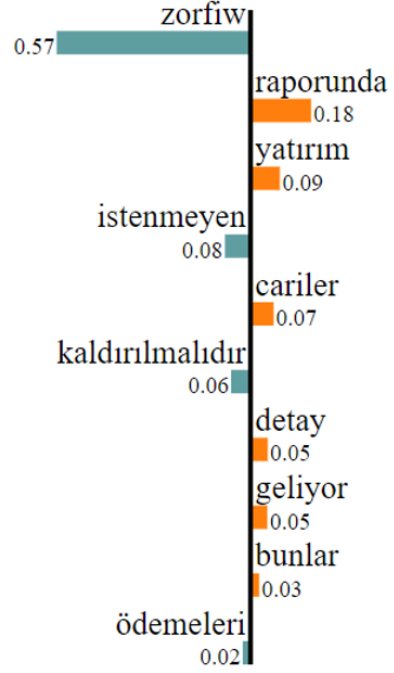
zorfiw yatırım ödemeleri detay raporunda istenmeyen cariler sonuca geliyor bunlar kaldırılmalıdır

#### Prediction probabilities

E ticaret	0.02
SAP BW	0.01
SAP ERP	0.97
SAP-BASIS	0.01

#### NOT SAP BW

#### SAP BW



Şekil 5.4. SAP ERP kategorisine ait bir cümlenin Lime algoritması ile incelenmesi

## 6. SONUÇ VE ÖNERİLER

İşletme içerisinde kullanılan iş istek uygulamasındaki iş isteklerinin yanlış kategoriye açılması işlerin ilgili kişilere gitmemesine ve böylece işlerin çözüm sürelerinin uzamasına sebep olmaktadır. Gerçekleştirilen tez çalışmasında iş istek açıklamalarından kategori tahminlemesi yapılarak işlerin doğru kategorilere atanması sağlanmıştır. Bu sayede açılan iş istekleri doğru bir şekilde sorumlu kullanıcılara gitmekte ve işlerin çözüm süreleri kısalmaktadır.

Yapılan çalışma sonucunda Bert, DistilBERT, ConvBERT ve Electra dil modelleri arasında en yüksek başarı değerinin ConvBERT modelinde elde edildiği sonucuna ulaşılmıştır. Bu durumun sebebi ConvBert modelinin, konvolüsyon tabanlı mimarisi sayesinde metin verilerinden özellikleri daha etkili bir şekilde çıkarma yeteneğine sahip olmasıdır. Yapılan çalışma ile beraber iş isteklerinin yanlış kategorilere atanma problemine çözüm bulunmuştur.

Açıklanabilir yapay zeka modeli olan Lime modelinin sisteme dahil edilmesi modelin açıklanabilirliğini sağlamıştır. Lime modeli ile tüm veriler analiz edildikten sonra elde edilecek kelimeler sayesinde kullanıcılar uzun bir cümle girmek yerine modelin kategori tahminlemesinde en çok etki eden kelimeleri girerek iş istek oluşturabileceklerdir.

Çalışma sonucunda iş isteklerin doğru kategorilere atanması sağlanarak iş isteklerinin ilgili kişilere daha kısa sürede ulaşması sağlanmıştır. Çalışma sayesinde kişilerin yanlış kategori seçmelerinin önüne geçilmiştir. Çalışmada kullanılan modeller, şirket içerisinde farklı birimlerde kullanılan uygulamalarda tahminleme ve sınıflandırma işlemi için örnek teşkil etmektedir. Bu çalışma, sektörel uygulamalar için geniş bir kullanım alanına sahip olabilir, örneğin sosyal medya içerisinde bulunan metin verileri analiz edilebilir, haberlerin otomatik olarak sınıflandırılması, e-postaların hangi kelimeler yüzünden spam olarak sınıflandırmasının analizi gibi çalışmalar yapılabilir. Yapılan çalışma, Türkçe doğal dil işleme ve açıklanabilir yapay zeka alanında yapılan çalışmaların sayısı az olduğu için yapılan çalışma literatüre katkı sağlamaktadır.

## KAYNAKLAR DİZİNİ

- Adalı, E. (2012). *Doğal dil işleme*. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 5(2).
- Adel, H., Dahou, A., Mabrouk, A., Elaziz, M.E.A. (2022). *Improving Crisis Events Detection Using DistilBERT with Hunger Games Search Algorithm*. Mathematics 10(3). DOI:10.3390/math10030447.
- Adoma, A. F., Henry, N. M., Chen, W. (2020). *Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition*. In 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), s.117-121. DOI: 10.1109/ICCWAMTIP51612.2020.9317379.
- Afzal, T. (2023). *Comparing BERT and ChatGPT: Understanding the Differences in NLP Models*. <https://medium.com/@taniaafzal/comparing-bert-and-chatgpt-understanding-the-differences-in-nlp-models-afb78e436105>. Erişim Tarihi: 16.12.2023.
- Akarsu, K. E. (2022). *Doğal Dil İşlemeye Dair Bilmeniz Gereken Her Şey*. <https://www.mespact.com/blog/dogal-dil-islemeye-dair-bilmeniz-gereken-her-sey/>. Erişim Tarihi: 11.12.2023.
- Altıntop, M. (2023). *Yapay Zeka/Akıllı Öğrenme Teknolojileriyle Akademik Metin Yazma: ChatGpt Örneği*. Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 46(2), s.186-211.
- Amplework Software Pvt. Ltd. (2023). *Comparing the AI Giants: ChatGPT vs BERT*. <https://www.linkedin.com/pulse/comparing-ai-giants-chatgpt-vs-bard-amplework-software-pvt-ltd/>. Erişim Tarihi: 16.12.2023.
- Bahdanau, D., Cho, K., Bengio, Y. (2015). *NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE*. Published as a conference paper at ICLR 2015. <https://doi.org/10.48550/arXiv.1409.0473>.

## KAYNAKLAR DİZİNİ

- Ballı, Ç. (2021). *Doğal dil işleme ile Türkçe içerikli paylaşımlardan sosyal medya kullanıcılarının duygu analizi*. [Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü].
- Bengio, Y., Ducharme, R., Vincent, P., Jauvin, C. (2003). *A Neural Probabilistic Language Model*. *Journal of Machine Learning Research* 3, s.1137–1155.
- Cansu O. (2023). *Makine Öğrenmesi Terimleri*. <https://www.linkedin.com/pulse/makine-%C3%B6%C4%9Frenimi-terimleri-orhan-cansu/?originalSubdomain=tr>. Erişim Tarihi: 11.12.2023.
- Clark, K., Luong, M.T., Le, Q.V., Manning, C.D. (2020). *ELECTRA: PRE-TRAINING TEXT ENCODERS AS DISCRIMINATORS RATHER THAN GENERATORS*. Published as a conference paper at ICLR 2020. <https://doi.org/10.48550/arXiv.2003.10555>.
- Collobert, R., Weston, J. (2008). *A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning*. *ICML '08: Proceedings of the 25th international conference on Machine learning*. DOI: 10.1145/1390156.1390177.
- Coşkun, F., Gülleroğlu, H. (2021). *Yapay Zekanın Tarih İçindeki Gelişimi ve Eğitimde Kullanılması*. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 3(54), s.947-966. <https://doi.org/10.30964/auebfd.916220>.
- Çiğdem, Y.N. (2023). *ChatGPT Nedir? Kullanım Alanları Nelerdir?*. <https://www.dia.com.tr/blog/chatgpt-nedir-kullanim-alanlari-nelerdir/#:~:text=ChatGPT%2C%20OpenAI%20taraf%C4%B1ndan%20geli%C5%9Ftirilmi%C5%9F%20bir,olarak%20insana%20yak%C4%B1n%20metinler%20olu%C5%9Fturabilmesidir>. Erişim Tarihi: 11.12.2023.
- Deliloğlu Selim, R. A., Pehlivanlı Çakmak, A. (2021). *Hibrit Açıklanabilir Yapay Zeka Tasarımı ve LIME Uygulaması*. *Avrupa Bilim ve Teknoloji Dergisi*, (27), s.228-236. DOI: 10.31590/ejosat.959030.

## KAYNAKLAR DİZİNİ

- Devlin, J., Chang, M. W., Lee, K., Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.
- Ersöz, F., Çınar, Y. (2021). *Veri Madenciliği ve Makine Öğrenimi Yaklaşımlarının Karşılaştırılması: Tekstil Sektöründe bir Uygulama*. Avrupa Bilim ve Teknoloji Dergisi, (29), s.397-414. DOI: 10.31590/ejosat.1035124.
- Gabbay, F., Bar-Lev, S., Montano, O., Hadad, N.(2021). *A LIME-Based Explainable Machine Learning Model for Predicting the Severity Level of COVID-19 Diagnosed Patients*. Appl. Sci. 2021, 11(21). <https://doi.org/10.3390/app112110417>.
- Google AI Hub. (2022). *Google Colab Nedir ve Nasıl Kullanılır?*. <https://globalaihub.com/google-colab-nedir-ve-nasil-kullanilir/>. Erişim Tarihi: 09.12.2023.
- Gui, T., Wang, X., Zhang, Q., Liu, Q., Zou, Y., Zhou, X., ... & Huang, X. (2021). *Textflint: Unified multilingual robustness evaluation toolkit for natural language processing*. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations, s.347-355. DOI: 10.18653/v1/2021.acl-demo.41.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D. (2018). *A survey of methods for explaining black box models*. ACM Computing Surveys, 51(5), 42. <https://doi.org/10.1145/3236009>.
- Güven, Z. A. (2023). *Türkçe E-postalarda Spam Tespiti için Makine Öğrenme Yöntemlerinin ve Dil Modellerinin Analizi*. Avrupa Bilim ve Teknoloji Dergisi, (47), s.1-6. DOI: 10.31590/ejosat.1234079.

## KAYNAKLAR DİZİNİ

- Haltaş, A., Alkan, A., Karabulut, M., (2015). *METİN SINIFLANDIRMADA SEZGİSEL ARAMA ALGORİTMALARININ PERFORMANS ANALİZİ*. Gazi Üniv. Müh. Mim. Fak. Der, 3(30), s.417-427.
- Hoş, S. (2023). *BERT Algoritması Nedir? Yapay Zekanın Search Ekosistemindeki Hakimiyeti, Örneklerle BERT Algoritması*. Erişim Tarihi: 01.12.2023.
- Howard, J., Ruder, S., (2018). *Universal Language Model Fine-Tuning for Text Classification*. Conference: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. DOI: 10.18653/v1/P18-1031.
- Hutchins, J. (2006). *The first public demonstration of machine translation: the Georgetown IBM system, 7th January 1954*. <https://open.unive.it/hitrade/books/HutchinsFirst.pdf>.
- İlaslan, E. (2023). *Yapay Zeka Sohbet Robotları ve ChatGPT'nin Hemşirelik Eğitiminde Kullanılması*. Akd Hemsirelik D., 2(2), s.73-80. <https://doi.org/10.59398/ahd.1330341>.
- Jiang, Z., Yu, W., Zhou, D., Chen, Y., Feng, J., Yan, S. (2020). *ConvBERT: Improving BERT with Span-based Dynamic Convolution*. 34th Conference on Neural Information Processing Systems (NeurIPS 2020). arXiv:2008.02496.
- Kara, Y. (2020). *Yeni bir yaklaşım olarak dilbilimsel sosyal hizmet*. Toplum ve Sosyal Hizmet, 31(4),s.1703-1718. <https://doi.org/10.33417/tsh.735805>.
- Karakoç Keskin, E. (2023). *Yapay Zekâ Sohbet Robotu Chatgpt Ve Türkiye İnternet Gündeminde Oluşturduğu Temalar*, Yeni Medya Elektronik Dergisi, 2(7), s.114-131.
- Köksal D., Alacan M.M., Olgun E., Sakar C. (2022). *Natural Language Processing-Based Product Category Classification Model for E-Commerce*. 2022 30th Signal Processing and Communications Applications Conference. DOI: 10.1109/SIU55565.2022.9864897.

## KAYNAKLAR DİZİNİ

- Lewis, T. (2014). *A Brief History of Artificial Intelligence*. <https://www.livescience.com/49007-history-of-artificial-intelligence.html>, Erişim tarihi: 26.11.2023.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. , Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality*. NIPS'13: Proceedings of the 26th International Conference on Neural Information Processing Systems, 2, s.3111-3119. <https://doi.org/10.48550/arXiv.1310.4546>.
- Nasiboğlu, R., Gencer, M., (2023). *Adlandırılmış Varlık Tanıma Modelleri ile Türkçe Sosyal Medya Metinlerinde Küfürlü Sözlerin Sansürlenmesi*. Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, s.72-88. DOI: 10.35414/akufemubid.1115786.
- Oğuzlar, A. (2003). *Veri Ön İşleme*. Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 21, s. 67-76.
- Ozan, Ş., Özdil U., Taşar, D.E., Arslan, B., Polat, G. (2021). *BERT Modeli'nin Sınıflandırma Doğruluğunun Sıfır-Atış Öğrenmesi ile Artırılması*. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 14(2), s.99-108.
- Özcan, M. (2021). *Sosyal Medya Platformlarının Hisse Senedi Piyasalarına Etkisi: BIST30 Örneği*. Gümüşhane Üniversitesi Sosyal Bilimler Enstitüsü Elektronik Dergisi, 12(1), s.160-172.
- Özel, U. (2020). *Açıklanabilir Yapay Zeka (Explainable AI)*. <https://www.umutozel.com/explainable-ai>. Erişim Tarihi: 12.12.2023
- Pekçoşkun Güner, S. (2023). *Çevirmen-Bilgisayar Etkileşiminin Kilit Bileşeni: Doğal Dil İşleme*. Uluslararası Filoloji ve Çeviribilim Dergisi, 1(5). DOI: 10.55036/ufced.1306746.

## KAYNAKLAR DİZİNİ

- Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L. (2018). *Deep contextualized word representations*. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, s.2227-2237. DOI: 10.18653/v1/N18-1202.
- Sanh, V., Debut, L., Chaumond, J., Wolf, T. (2019). *DistilBERT, a distilled version of bert: smaller, faster, cheaper and lighter*. Published in arXiv: Computation and Language. arXiv preprint arXiv:1910.01108.
- Sel İ., Hanbay D. (2021). *Ön Eğitimli Dil Modelleri Kullanarak Türkçe Tweetlerden Cinsiyet Tespiti*. Fırat Üniversitesi Müh. Bil. Dergisi, 33(2), s.675-684. <https://doi.org/10.35234/fumbd.929133>.
- Seyyarer, E., Ayata, F., Uçkan, T., Karci,A. (2020). *Derin Öğrenmede Kullanılan Optimizasyon Algoritmalarının Uygulanması ve Kıyaslanması*. Computer Science, 5(2), s.90-98.
- Silay, R. (2020). *NLP (Doğal Dil İşleme) — nedir, ön işlemler ve adımlar*. <https://resulsilay.medium.com/nlp-do%C4%9Fal-dil-i-CC%87%C5%9Fleme-nedir-%C3%B6n-i%C5%9Flemler-ve-ad%C4%B1mlar-7eab97a5056>. Erişim Tarihi: 11.12.2023.
- Sutskever, I., Vinyals, O., Le, Q. (2014). *Sequence to Sequence Learning with Neural Networks*. Advances in Neural Information Processing Systems 27 (NIPS 2014) <https://arxiv.org/abs/1409.3215>.
- Şeker, S. E. (2016). *Duygu Analizi (Sentimental Analysis)*.Yönetim Bilişim Sistemleri Ansiklopedi, 3(3), s.21-36.
- Turing, A. (1950). *Computing Machinery and Intelligence*. Mind, 236, s.433-460. <https://doi.org/10.1093/mind/LIX.236.433>.
- Üzücek, M. (2023). *Joseph Weizenbaum: İlk Chatbot'un Mucidi Nasıl Yapay Zeka Muhalifi Haline Geldi?*. Erişim Tarihi: 23.12.2023.

**KAYNAKLAR DİZİNİ**

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017). *Attention is All You Need*. NIPS 2017, Paper Id: 3058. <https://doi.org/10.48550/arXiv.1706.03762>.

Winograd, T. (1971). *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*. <https://dspace.mit.edu/handle/1721.1/7095>.

Woods, W. A. (1970). *Transition Network Grammars for Natural Language Analysis*. *Communications of the ACM*, 10(13), s.591–606. <https://doi.org/10.1145/355598.362773>.

Yılmaz, H., Yumuşak, S. (2021). *Açık kaynak doğal dil işleme kütüphaneleri*. İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 3(1), s.81-85. <https://doi.org/10.47769/izufbed.879217>.

## EK AÇIKLAMALAR

### Ek Açıklama-A: Tez Çalışmasından Üretilen Bilimsel Yayın

Özkan, C., Özkan, K. (2023, 16-17 Aralık). *İş İsteklerindeki Türkçe Metinlerin Doğal Dil İşleme Yöntemleri ve Açıklanabilir Yapay Zeka Kullanılarak Analizi*. 2. Uluslararası Mühendislik ve Fen Bilimleri Kongresi Bildiri Kitabı, s.701.

