T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOĞAL DİL İŞLEMEDE İLERİ SEVİYE METİN SINIFLANDIRMA: TRANSFORMER'IN ROLÜ

YÜKSEK LİSANS TEZİ

Mert Halil DURAK

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışmanı: Dr. Öğr. Üyesi Cengiz HARK

TEMMUZ 2025

T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOĞAL DİL İŞLEMEDE İLERİ SEVİYE METİN SINIFLANDIRMA: TRANSFORMER'IN ROLÜ

YÜKSEK LİSANS TEZİ

Mert Halil DURAK
(36223619910)

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışmanı: Dr. Öğr. Üyesi Cengiz HARK

TEMMUZ 2025

İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ

DOĞAL DİL İŞLEMEDE İLERİ SEVİYE METİN SINIFLANDIRMA: TRANSFORMER'IN ROLÜ

Tezi Hazırlayan: Mert Halil DURAK

Yüksek Lisans Tezi

TEŞEKKÜR VE ÖNSÖZ

Bu tez çalışmasının her aşamasında yardım, öneri, bilgi, tecrübe ve desteklerini esirgemeden beni her konuda yönlendiren danışman hocam Sayın Dr. Öğr.Üyesi Cengiz HARK'a,

Çalışmalarımda ve ayrıca tüm hayatım boyunca olduğu gibi bu çalışma sürecimde de benden her türlü desteklerini esirgemeyen aileme,

Tezin uygulama aşamasında vermiş oldukları destekten dolayı, İnönü Üniversitesi BAP birimine

Teşekkür ederim.

ONUR SÖZÜ

Yüksek lisans tezi olarak sunduğum “DOĞAL DİL İŞLEMEDE İLERİ SEVİYE METİN SINIFLANDIRMA: TRANSFORMER'IN ROLÜ” başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığına ve yararlandığım bütün kaynakların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Mert Halil DURAK

İÇİNDEKİLER

ÇİZELGELER DİZİNİ

ŞEKİLLER DİZİNİ

KISALTMALAR LİSTESİ

BERT **:** Bidirectional Encoder Representations from Transformer

CNN **:** Convolutional Neural Network

DDİ : Doğal Dil İşleme

GPT **:** Generative Pre-trained Transformer

GRU : Gated Recurrent Units

IDF **:** Inverse Document Frequency

LSTM **:** Long Short Term Memory

NLP **:** Natural Language Processing

RNN **:** Recurrent Neural Network

SVM **:** Support Vector Machine

T5 **:** Text-to-Text Transfer Transformer

TF **:** Term Frequency

ÖZET

Yüksek Lisans Tezi

DOĞAL DİL İŞLEMEDE İLERİ SEVİYE METİN SINIFLANDIRMA: TRANSFORMER'IN ROLÜ

MERT HALİL DURAK

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

55+XIII Sayfa

2025

Danışman: Dr. Öğr. Üyesi Cengiz HARK

Bu tez çalışması, doğal dil işleme (Natural Language Processing, NLP) alanında metin sınıflandırma problemlerine yönelik geliştirilmiş geleneksel yöntemler (Naive Bayes ve Destek Vektör Makineleri) ile son yıllarda büyük başarı gösteren Transformer tabanlı modellerin (BERT, RoBERTa, GPT, T5, DistilBERT) karşılaştırmalı analizini kapsamaktadır. Çalışmanın temel amacı, farklı veri setlerinde bu yöntemlerin sınıflandırma performanslarını ölçmek ve elde edilen sonuçları doğruluk, kayıp oranları ve genel model başarısı açısından detaylı şekilde irdelemektir.

Tez kapsamında AG News, IMDb ve SST-2 veri setleri kullanılmış; metin verileri uygun ön işleme süreçlerinden geçirilmiş ve hem geleneksel hem de Transformer tabanlı modellere entegre edilmiştir. Naive Bayes ve SVM gibi yöntemler, metin özelliklerini temel istatistiksel özellikler üzerinden değerlendirirken, Transformer tabanlı modeller ise dikkat mekanizması (attention mechanism) ve büyük ölçekli önceden eğitilmiş dil modelleri sayesinde bağlamsal bilgi zenginliğinden yararlanmıştır.

Uygulama ve deneyler aşamasında her bir model veri setleri üzerinde eğitilmiş ve test edilmiştir. Sonuçlar, Transformer modellerinin genel olarak geleneksel yöntemlere kıyasla daha yüksek doğruluk oranlarına ulaştığını, ancak hesaplama maliyetleri ve eğitim süresi açısından geleneksel yöntemlerin hala avantajlı olduğunu ortaya koymuştur. Ayrıca, farklı modellerin sınıflandırma başarısı veri setine ve problem türüne göre değişiklik göstermektedir; örneğin, RoBERTa modeli AG News ve SST-2 veri setlerinde en yüksek doğruluk oranını sağlarken, IMDb veri setinde BERT modeli öne çıkmıştır.

Bu tezde elde edilen bulgular, gelecekteki NLP uygulamaları için model seçiminde yol gösterici nitelikte olup, hem akademik hem de endüstriyel uygulamalara katkı sağlamayı hedeflemektedir. Sonuçlar, ayrıca Transformer mimarisinin güçlü yanlarının yanı sıra sınırlılıklarına da dikkat çekmekte ve daha verimli ve etkin yöntemler geliştirmek için bir temel sunmaktadır.

Anahtar Kelimeler: Doğal Dil İşleme (NLP), Metin Sınıflandırma, Transformer Modelleri, Naive Bayes Modeli, Destek Vektör Makineleri (SVM), BERT Modeli, RoBERTa Modeli, GPT Modeli, Makine Öğrenmesi

ABSTRACT

Master Thesis

ADVANCED TEXT CLASSIFICATION IN NATURAL LANGUAGE PROCESSING: THE ROLE OF TRANSFORMER

MERT HALİL DURAK

Inonu University
Graduate School of Nature and Applied Sciences
Department of Computer Engineering

55+XIII Page

2025

Supervisor: Dr. Öğr. Üyesi Cengiz HARK

This thesis covers the comparative analysis of traditional methods (Naive Bayes and Support Vector Machines) developed for text classification problems in the field of Natural Language Processing (NLP) and Transformer-based models (BERT, RoBERTa, GPT, T5, DistilBERT) that have shown great success in recent years. The main purpose of the study is to measure the classification performances of these methods on different data sets and to examine the obtained results in detail in terms of accuracy, loss rates and general model success.

Within the scope of the thesis, AG News, IMDb and SST-2 datasets were used; text data was subjected to appropriate pre-processing processes and integrated into both traditional and Transformer-based models. While methods such as Naive Bayes and SVM evaluate text features based on basic statistical properties, Transformer-based models benefit from the richness of contextual information thanks to the attention mechanism and large-scale pre-trained language models.

In the implementation and experiment phase, each model was trained and tested on the datasets. The results showed that Transformer models generally achieved higher accuracy rates compared to traditional methods, but traditional methods still had an advantage in terms of computational costs and training time. In addition, the classification success of different models varies according to the dataset and problem type; for example, the RoBERTa model achieved the highest accuracy rate on the AG News and SST-2 datasets, while the BERT model stood out on the IMDb dataset.

The findings obtained in this thesis are of a guiding nature in model selection for future NLP applications and aim to contribute to both academic and industrial applications. The results also highlight the strengths as well as limitations of the Transformer architecture and provide a basis for developing more efficient and effective methods.

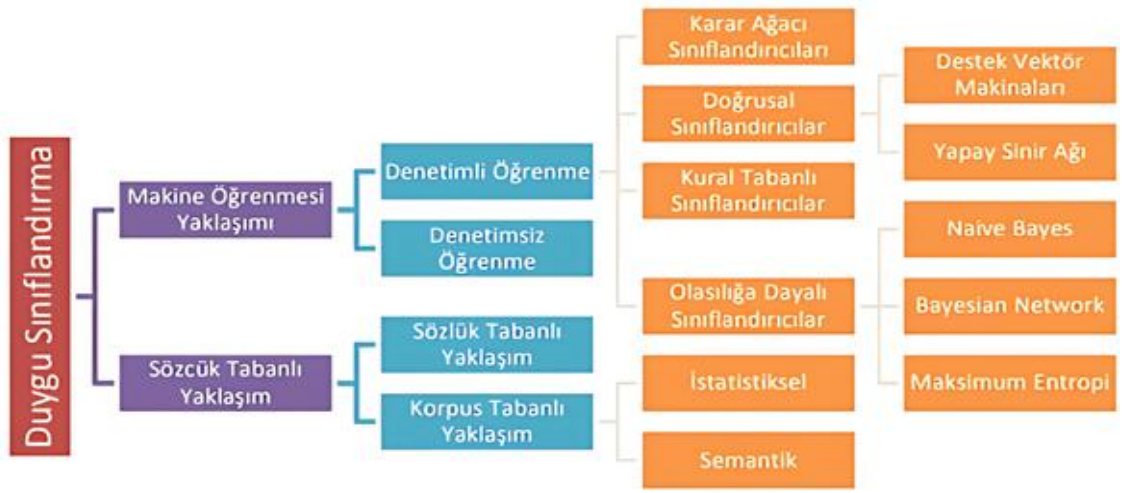
Processing. Keywords: Natural Language Processing (NLP), Text Classification, Transformer Models, Naive Bayes Model, Support Vector Machines (SVM), BERT Model, RoBERTa Model, GPT Model, Machine Learning

1. GİRİŞ

1.1 Problemin Tanımı

Günümüzde dijitalleşmenin ve internet kullanımının hızlı artışıyla birlikte metin verisi, insan iletişiminin ve bilgi paylaşımının en baskın biçimlerinden biri haline gelmiştir. Sosyal medya platformları, haber siteleri, çevrimiçi forumlar, müşteri yorumları ve e-posta içerikleri gibi kaynaklardan günlük olarak üretilen metin miktarı, veri bilimi ve yapay zekâ araştırmacıları için hem büyük bir fırsat hem de zorlu bir meydan okumadır. Metinlerin doğru ve hızlı bir şekilde sınıflandırılması, bu büyük verinin anlamlandırılmasında kritik bir rol oynamaktadır. Ancak doğal dilin yapısı gereği karmaşık ve bağlama bağlı oluşu, kelime anlamlarının değişkenliği, ironiler, mecazlar, çok anlamlılık, dilbilgisel belirsizlikler gibi unsurlar; metin sınıflandırma problemlerinin çözümünü zorlaştıran başlıca etkenlerdir.

Metin sınıflandırma, en basit tanımıyla bir metin parçasının (örneğin bir cümlenin, paragrafın veya belgenin) belirlenen sınıflardan birine veya birden fazlasına atanması işlemidir. Bu görev, spam e-posta tespiti, duygu analizi, konu sınıflandırması, haber kategorilendirme, sahte haber tespiti, toksik içerik belirleme gibi birçok uygulamada kritik öneme sahiptir. Yüksek hacimli veri ortamlarında bu işlemin manuel olarak yapılması mümkün olmadığından, makine öğrenmesi ve doğal dil işleme tekniklerinin kullanılması kaçınılmaz hale gelmiştir.



Şekil 1.1: Duygu analizinde kullanılan teknikler (AMANET, 2020)

Tarihsel olarak, bu problem için kullanılan geleneksel yöntemler, genellikle elle çıkarılan özelliklere dayanan (örneğin TF-IDF, kelime frekansları) ve bu özellikler üzerinde sınıflandırma yapan algoritmalardır. Naive Bayes, Destek Vektör Makineleri (SVM) ve Lojistik Regresyon gibi yöntemler; belirli bir bağlamda yeterli performans gösterebilse de, karmaşık dil yapılarını ve uzun mesafeli bağımlılıkları yakalamada sınırlı kalmaktadır. Özellikle dilin bağlamsal yapısına duyarsız olmaları, bu yöntemlerin metin sınıflandırmada üst sınırlarına hızla ulaşmalarına neden olmaktadır.

Son yıllarda geliştirilen Transformer mimarisi ise bu sınırlamaları aşmak için yeni bir paradigma sunmuştur. Attention (dikkat) mekanizmasıyla donatılmış Transformer tabanlı modeller, hem dilin bağlamsal ilişkilerini etkili bir şekilde öğrenebilmekte hem de büyük veri kümelerinde önceden eğitilmiş olmaları sayesinde, transfer öğrenme yoluyla birçok farklı görevde yüksek doğruluk sağlayabilmektedir. BERT (Bidirectional Encoder Representations from Transformers), RoBERTa, GPT (Generative Pre-trained Transformer), T5, DistilBERT gibi modeller, doğal dil işleme problemlerinde özellikle sınıflandırma ve anlamlandırma görevlerinde çığır açan başarılar elde etmiştir. Bununla birlikte, bu modellerin devasa parametre sayıları ve hesaplama gereksinimleri, her proje ve uygulama için uygun olmadıkları gerçeğini de beraberinde getirmektedir.

Bu tez çalışmasının temel problemi, geleneksel makine öğrenmesi yöntemleri (özellikle Naive Bayes ve SVM) ile Transformer tabanlı modern modellerin metin sınıflandırma problemlerindeki performanslarının detaylı ve kapsamlı şekilde kıyaslanmasıdır. Çalışmada, farklı dil ve içerik türlerine sahip üç farklı veri seti kullanılarak bu iki yaklaşımın doğruluk, kayıp, eğitim süresi ve hesaplama maliyetleri bakımından avantajları ve sınırlılıkları incelenmiştir. Buradaki temel hedef, yalnızca doğruluk oranlarını raporlamak değil; aynı zamanda bu iki farklı yaklaşımın hangi koşullarda, hangi tür veri ve problem yapılarına göre daha uygun ya da verimli olduğunu ortaya koymak ve gelecekte yapılacak çalışmalara ışık tutmaktır.

Bu problem tanımı, aynı zamanda akademik literatürde uzun süredir tartışılan bir konudur. Geleneksel yöntemlerin sadeliği ve düşük hesaplama maliyeti, onları hala bazı pratik uygulamalar için cazip kılarken, Transformer tabanlı modellerin üstün performansı ise araştırmacılar ve endüstri tarafından giderek daha fazla tercih edilmesine yol

açmaktadır. Bu nedenle çalışmada yalnızca metrik bazlı değil; uygulama ve kullanım koşulları açısından da kapsamlı bir değerlendirme yapılması amaçlanmıştır.

1.2 Tezin Amacı ve Kapsamı

Bu tezin temel amacı, doğal dil işleme (NLP) alanında özellikle metin sınıflandırma problemlerine odaklanarak, geleneksel makine öğrenmesi yöntemleri (özellikle Naive Bayes ve Destek Vektör Makineleri - SVM) ile son yıllarda büyük yankı uyandıran Transformer tabanlı derin öğrenme yaklaşımlarının (BERT, RoBERTa, GPT, DistilBERT, T5) kapsamlı bir analizini gerçekleştirmektir. Günümüzde metin verisi muazzam bir hızla artmakta ve bu veriden anlamlı bilgi çıkarma ihtiyacı hem akademik hem ticari çevrelerde kritik bir hale gelmektedir. Sosyal medya, haber metinleri, kullanıcı yorumları ve e-posta içerikleri gibi farklı metin kaynaklarının otomatik sınıflandırılması pek çok alanda önem arz etmektedir. Bu bağlamda çalışmanın temel amacı, metin sınıflandırma görevlerinde kullanılan yöntemlerin sadece doğruluk performanslarını değil; aynı zamanda verimlilik, hesaplama maliyeti, model karmaşıklığı, eğitim süresi, yorumlanabilirlik ve uygulanabilirlik gibi çok boyutlu performans kriterlerini karşılaştırmalı olarak incelemektir.

Bu amaçla çalışma, üç farklı veri seti üzerinde hem geleneksel hem de Transformer tabanlı yöntemleri uygulayarak elde edilen deneysel sonuçları detaylı bir şekilde raporlamaktadır. Transformer mimarisine dayalı modeller, özellikle transfer öğrenme ve önceden eğitilmiş büyük dil modelleri sayesinde küçük ya da orta ölçekli veri setlerinde bile üstün başarı göstermektedir. Ancak bu modellerin eğitim süreci, ihtiyaç duyduğu hesaplama kaynakları, model karmaşıklığı ve sonuçların yorumlanabilirliği gibi konularda önemli sınırlılıkları bulunmaktadır. Geleneksel yöntemler ise daha basit yapıları, düşük hesaplama maliyetleri ve hızlı uygulanabilirlikleri ile hâlâ birçok pratik kullanımda tercih edilmektedir.

Tezin kapsamı yalnızca deneysel bir karşılaştırmayla sınırlı değildir; aynı zamanda ilgili literatürdeki mevcut yaklaşımların detaylı incelenmesi, teorik arka plan bilgileri, veri setlerinin özellikleri, veri ön işleme aşamaları, model mimarilerinin ayrıntılı açıklamaları ve deneysel tasarımın metodolojik gerekçeleri de kapsamlı şekilde ele alınmaktadır. Ayrıca elde edilen bulgular, sadece metin sınıflandırma bağlamında değil, daha geniş bir

NLP çerçevesinde de tartışılarak Transformer tabanlı modellerin (özellikle BERT, RoBERTa, GPT, DistilBERT ve T5) gerçek dünya uygulamalarındaki avantaj ve dezavantajlarına dair genel çıkarımlar yapılmaktadır.

Bu tez, hem akademik araştırmalar hem de sektörel uygulamalar için bir referans noktası oluşturmayı hedeflemektedir. Hangi tür projelerde geleneksel yöntemlerin hâlâ yeterli olduğu, hangi durumlarda ise Transformer tabanlı yaklaşımların tercih edilmesi gerektiği, performans ile hesaplama maliyeti arasındaki denge ve model seçiminde dikkat edilmesi gereken faktörler gibi önemli sorulara yanıt aranmaktadır. Çalışmanın sonunda mevcut sınırlamalar ve gelecekte yapılabilecek araştırmalar için öneriler sunularak, ileride yapılacak çalışmalara katkı sağlaması hedeflenmektedir.

1.3 Yöntem Özeti

Bu tez çalışmasında, metin sınıflandırma problemi, hem geleneksel makine öğrenmesi algoritmaları hem de çağdaş derin öğrenme tabanlı yaklaşımlar kullanılarak ele alınmıştır. Çalışmanın temel amacı, metin verileri üzerinde farklı yöntemlerin performansını nesnel bir şekilde ölçmek ve literatüre katkı sağlayacak kapsamlı bir analiz sunmaktır. Kullanılan yöntemler, farklı karmaşıklık seviyelerine sahip üç farklı veri seti üzerinde uygulanmış ve elde edilen sonuçlar karşılaştırılmıştır.

Kullanılan veri setleri üç ana gruptan oluşmaktadır:

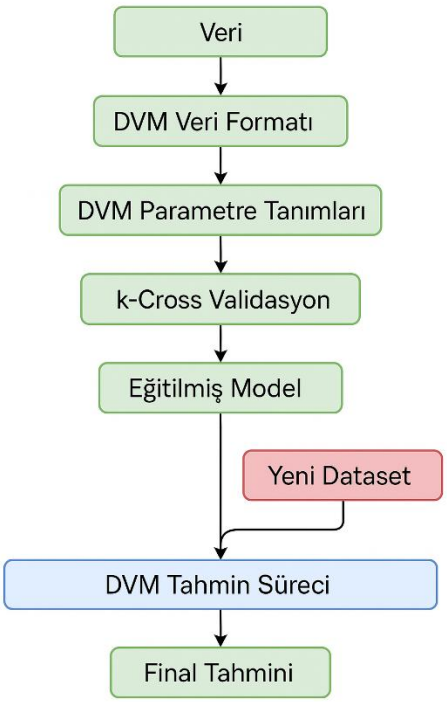
IMDb Duygu Analizi Veri Seti: İnternet üzerindeki en popüler film inceleme veri setlerinden biridir. Kullanıcıların filmler hakkında yazdığı incelemeler “olumlu” veya “olumsuz” şeklinde etiketlenmiştir. Bu veri seti, geniş çaplı yorumlar içerdiği ve doğal dilin çeşitli biçimlerini barındırdığı için, modellerin genel duygusal analizi ne kadar iyi yapabildiğini test etmek açısından önemli bir zemin sağlamaktadır.

SST-2 (Stanford Sentiment Treebank): Stanford Üniversitesi tarafından geliştirilen bu veri seti, bireysel cümleler üzerinden duygu analizi yapmayı hedefler. IMDb veri setine kıyasla daha kısa metinler içerir ve cümle düzeyinde olumlu/olumsuz etiketlemeler sunar. Bu sayede, daha hassas duygu sınıflandırması gerektiren modellerin başarısı değerlendirilebilir. Özellikle ince duygu geçişlerini yakalamada Transformer tabanlı modellerin performansı öne çıkmaktadır.

AG News Haber Sınıflandırma Veri Seti: Bu veri seti, dört ana kategoriye (Dünya, Spor, İş Dünyası, Bilim/Teknoloji) ayrılmış binlerce haber başlığı ve özetinden oluşur. Çok sınıflı sınıflandırma problemlerinde kullanılan bu veri seti, modellerin konusal içerik ayırımı yapabilme yeteneğini sınamak için tercih edilmiştir.

Kullanılan yöntemler iki ana grupta toplanabilir:

Geleneksel Yöntemler: Naive Bayes ve Destek Vektör Makineleri (SVM), doğal dil işleme (NLP) literatüründe uzun yıllardır kullanılan, metin sınıflandırmada hızlı ve etkili çözümler sunan yöntemlerdir. Bu modeller, özellikle küçük ölçekli veri setlerinde ve temel sınıflandırma problemlerinde düşük hesaplama maliyetleri nedeniyle tercih edilir. Ancak, kelimeler arasındaki karmaşık ilişkileri modelleme konusunda sınırlıdırlar.



Şekil 1.2: Destek vektör makineleri algoritması

Transformer Tabanlı Yöntemler: BERT, RoBERTa, DistilBERT, T5 ve GPT gibi modern derin öğrenme modelleri, son yıllarda metin işleme alanında devrim yaratmıştır. Bu modeller, büyük hacimli veri ve hesaplama gücü kullanılarak önceden eğitilmiş, ardından belirli görevler için ince ayar yapılabilir hâle getirilmiştir. Transformer

mimarisi, kelimeler arasındaki bağlamsal ilişkileri dikkate alarak daha doğru ve anlamlı sınıflandırmalar yapma potansiyeline sahiptir.

Bu tezde, her model üç farklı veri seti üzerinde eğitilmiş ve hem eğitim hem de doğrulama süreçlerinde ortaya çıkan metrikler detaylı şekilde kaydedilmiştir. Kullanılan metrikler arasında doğruluk oranı (accuracy), kayıp değeri (loss), eğitim süresi ve model karmaşıklığı yer almaktadır.

Ayrıca, deneysel analiz sürecinde tablolar ve grafikler yardımıyla hem Transformer tabanlı hem de geleneksel yöntemler arasındaki performans farkları görsel olarak da sunulmuştur. Böylece sadece nicel değil, nitel bir karşılaştırma da yapılmış, hangi yöntemin hangi koşullarda üstünlük sağladığı, hangi yöntemlerin belirli zorluklarla karşılaştığı detaylandırılmıştır.

Sonuç olarak bu tez, doğal dil işleme ve metin sınıflandırma alanında çalışan araştırmacılar ve uygulayıcılar için yöntemlerin güçlü ve zayıf yönlerini anlamada yol gösterici bir çalışma sunmayı hedeflemektedir.

1.4 Tezin Yapısı

Bu tez çalışması, metin sınıflandırma alanında geleneksel yöntemler ile Transformer tabanlı ileri yöntemlerin karşılaştırmalı analizini yapmak amacıyla kurgulanmış ve sistematik bir yapıya oturtulmuştur. Tez, altı ana bölümden oluşmakta ve her bir bölüm, okuyucunun hem teorik altyapıyı hem uygulamalı çalışmaları hem de sonuçların tartışmasını bütüncül bir şekilde takip etmesine imkân tanıyacak şekilde tasarlanmıştır.

Birinci Bölüm (Giriş), çalışmanın neden yapıldığını, hangi problemin ele alındığını ve bu problemin bilimsel veya pratik anlamda neden önemli olduğunu detaylı biçimde ortaya koyar. Ayrıca bu bölümde, tezin amacı, kapsamı ve izlenen yöntemlerin kısa bir özeti sunularak okuyucuya genel bir çerçeve çizilir. Araştırma soruları, hipotezler ve bu sorulara nasıl yaklaşılacağına dair yöntemsel ipuçları da bu bölümde yer alır. Okuyucu, bu bölümde çalışmanın temel motivasyonunu ve hedeflerini anlama fırsatı bulur.

İkinci Bölüm (Kavramsal Temeller ve Literatür İncelemesi), doğal dil işleme (NLP) kavramlarının temel tanımlarından başlayarak, metin sınıflandırma yöntemlerine kadar

geniş bir yelpazede literatür taraması sunar. Bu bölümde, geleneksel yöntemler (Naive Bayes, SVM) detaylı biçimde açıklanır ve bunların sınırlılıkları ele alınır. Ayrıca, Transformer tabanlı modern yaklaşımlar (BERT, RoBERTa, DistilBERT, T5, GPT) derinlemesine incelenir ve mevcut literatürde bu modellerle yapılan güncel çalışmalar ve elde edilen sonuçlar tartışılır. Literatür taraması, çalışmanın özgün katkısının anlaşılması açısından kritik öneme sahiptir; bu yüzden alandaki boşluklar ve bu çalışmanın bu boşluklara nasıl katkı sunduğu açıkça ortaya konur.

Üçüncü Bölüm (Yöntem), deneysel sürecin detaylandırıldığı bölümdür. Kullanılan veri setlerinin özellikleri (SST-2, IMDb, AG News), veri ön işleme adımları (temizleme, etiketleme, tokenizasyon), model mimarileri (Transformer ve geleneksel modeller), hiperparametre ayarları ve performans metrikleri (doğruluk, kayıp, F1 skoru vb.) ayrıntılı olarak açıklanır. Bu bölümde izlenen yöntemsel yol, sadece sonuçları üretmekle kalmaz, aynı zamanda benzer bir çalışma yapmak isteyen araştırmacılara veya uygulayıcılara da adım adım rehberlik sunar. Deneylerin tekrarlanabilirliği açısından tüm teknik detayların şeffaf biçimde aktarılması amaçlanır.

Dördüncü Bölüm (Uygulama ve Deneyler), deneysel sonuçların sunulduğu ve karşılaştırmalı analizlerin yapıldığı bölümdür. Transformer tabanlı modeller ve geleneksel yöntemler üzerinde yapılan testlerin sonuçları detaylı tablo ve grafiklerle desteklenir. Doğruluk oranları, eğitim ve doğrulama kayıpları, eğitim süreleri, model karmaşıklıkları gibi metrikler, kapsamlı bir analizle karşılaştırılır. Görseller (doğruluk ve kayıp eğrileri, bar grafikler) bu bölümde çalışmaya görsel bir zenginlik katar ve okuyucunun bulguları daha iyi yorumlamasına yardımcı olur.

Beşinci Bölüm (Sonuçlar ve Gelecek Çalışmalar), yapılan analizlerin genel bir özetini sunar ve çalışmadan elde edilen temel bulguları tartışır. Transformer modellerinin üstünlükleri, sınırlılıkları, geleneksel yöntemlerle karşılaştırıldığında ortaya çıkan avantajlar ve zorluklar kapsamlı biçimde ele alınır. Ayrıca, çalışmanın bilimsel ve pratik katkıları, hangi alanlarda uygulanabileceği ve gelecekte yapılabilecek iyileştirme ve genişletme önerileri sunulur. Bu bölüm, çalışmanın sadece mevcut çıktılarla sınırlı olmadığını, aynı zamanda geleceğe dönük bir perspektif sunduğunu göstermesi açısından önemlidir.

Altıncı Bölüm (Kaynaklar), tez boyunca kullanılan tüm akademik yayınlar, makaleler, kitaplar ve çevrim içi kaynakları içerir. Metin içinde yapılan her atıf, bu bölümde detaylı bir şekilde belirtilir ve tez yazım kurallarına uygun biçimde listelenir. Böylece tez, akademik dürüstlük ve şeffaflık ilkesine uygun şekilde tamamlanır.

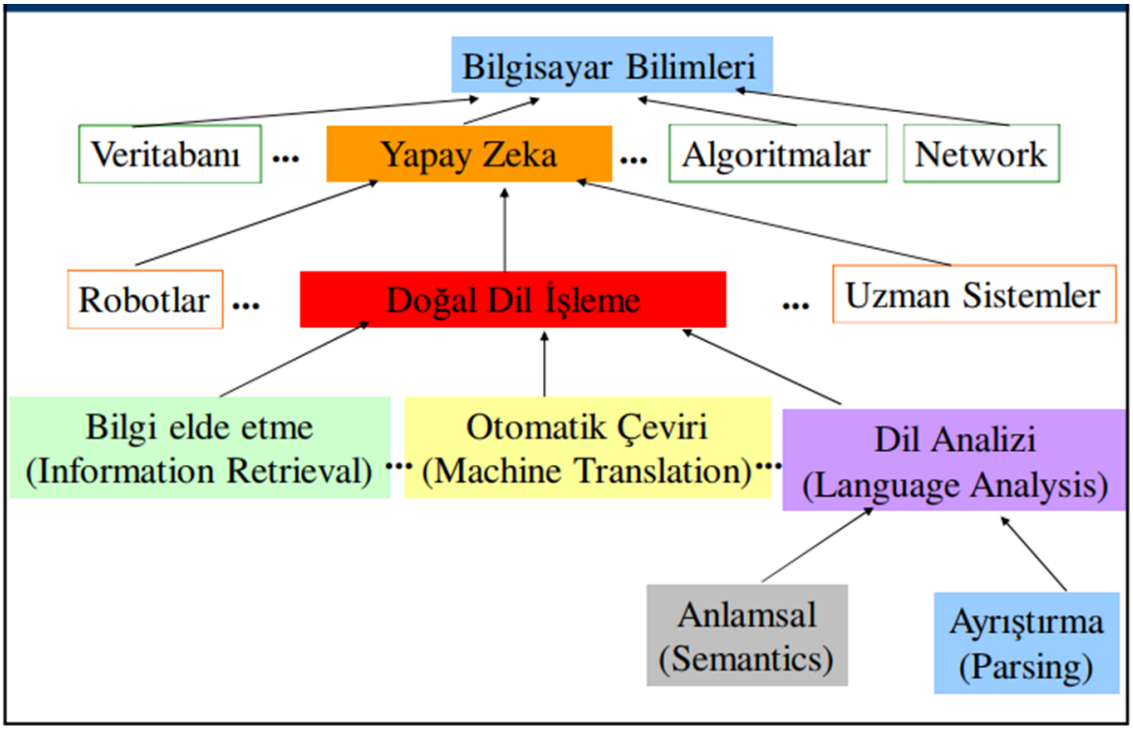
Sonuç olarak, tezin bu yapılandırılmış ve detaylı tasarımı, okuyucunun yalnızca deneysel sonuçları değil, aynı zamanda bunların teorik arka planını ve gelecekteki potansiyel yansımalarını da kapsamlı biçimde kavramasına olanak tanır. Çalışmanın bölümleri bir bütün olarak, metin sınıflandırma alanına katkı sunmayı hedeflerken, metodolojik ve akademik anlamda yüksek bir standardı gözetmektedir

Şekil 1.3: Tez yapısı

2. KAVRAMSAL TEMELLER VE LİTERATÜR İNCELEMESİ

2.1 Doğal Dil İşleme (NLP): Temel Kavramlar

Doğal Dil İşleme (Natural Language Processing - NLP), bilgisayarların insan dilini anlaması, analiz etmesi, yorumlaması ve üretmesi amacıyla geliştirilen bir yapay zekâ alt alanıdır. İnsan dili, karmaşık yapısı, belirsizlikleri, bağlama duyarlılığı ve çok katmanlı anlam taşıma kapasitesi nedeniyle bilgisayarlar açısından oldukça zorludur. Bu zorlukları aşmak için NLP, dilbilim, bilgisayar bilimi ve istatistiksel modelleme gibi disiplinleri birleştirir ve çok çeşitli yöntem ve yaklaşımlar geliştirir.



Şekil 2.1: Doğal dil işleme ile dil analizi [1]

NLP'nin temel amacı, yapılandırılmamış ya da yarı yapılandırılmış metin verilerini anlamlı ve işlenebilir bilgiye dönüştürmektir. Bu dönüşüm, örneğin metin sınıflandırma, duygu analizi, isim tanıma, özetleme, makine çevirisi, soru-cevap sistemleri, dil modelleme, anlamsal benzerlik hesaplama ve metin üretimi gibi birçok farklı görevde kullanılmaktadır. Örneğin bir e-ticaret sitesindeki müşteri yorumlarının olumlu veya olumsuz olarak sınıflandırılması, bir haber makalesinden otomatik özet çıkarılması ya da

kullanıcı sorularına yanıt veren sohbet botlarının geliştirilmesi gibi birçok uygulama doğrudan NLP tekniklerine dayanır.

NLP, genellikle iki temel bileşen üzerine inşa edilir: 1-) Dilsel Ön İşleme ve 2-) Modelleme ve Anlam Çıkarma. Dilsel ön işleme adımı, ham metnin temizlenmesi, tokenizasyon, durak kelime (stopword) kaldırma, kök veya gövde bulma (stemming, lemmatization), POS (part-of-speech) etiketleme, adlandırılmış varlık tanıma (NER) gibi alt süreçleri içerir. Bu adımlar, ham verinin daha yapılandırılmış ve modellenebilir hâle gelmesini sağlar. Modelleme aşamasında ise istatistiksel yöntemler, makine öğrenmesi teknikleri ya da son yıllarda büyük sıçramalar yaratan derin öğrenme tabanlı yöntemler kullanılarak metin üzerinde görevler gerçekleştirilir.

NLP'nin zorlukları arasında çok anlamlılık (polisemik sözcükler), deyimler, ironiler, bağlama duyarlılık, kültürel ve dilsel farklılıklar, dilin zaman içindeki evrimi ve dilbilgisel hatalar yer alır. Bu yüzden, geliştirilen her yeni yöntem, genellikle bu problemlerin en azından bir kısmını daha iyi çözmeyi hedefler. Örneğin, geleneksel "bag-of-words" yaklaşımları kelime sırasını ve bağlamı dikkate almazken, Word2Vec ve GloVe gibi dağıtılmış temsiller bu açığı kapatmaya çalışmıştır. Ancak Transformer tabanlı modeller (BERT, GPT, RoBERTa, T5 ve DistilBERT gibi) ile birlikte bağlamın çift yönlü olarak işlenebilmesi ve önceden eğitilmiş büyük modellerin transfer öğrenme yoluyla çeşitli görevlerde kullanılabilmesi NLP'de devrim yaratmıştır.

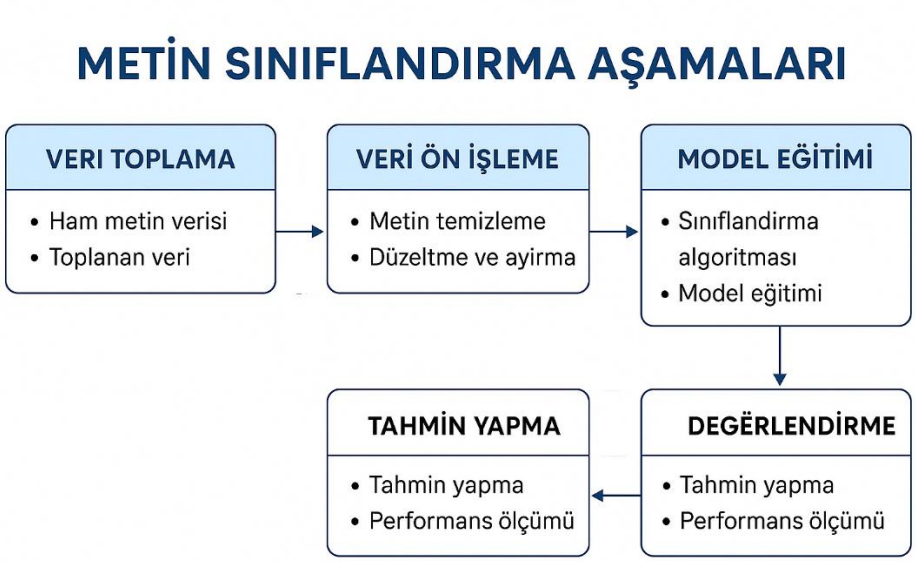
Ayrıca, NLP sadece İngilizce gibi baskın dillerde değil, Türkçe gibi morfolojik açıdan zengin ve eklemeli dillerde de ayrı bir araştırma alanı oluşturur. Türkçe, kelimelerin köklerine eklenen çok sayıda çekim ve yapım eki nedeniyle, kelime düzeyinde çeşitlilik ve karmaşıklık sunar. Bu durum, kelime temsilleri ve modeller açısından özel çözümler gerektirir. Örneğin, Türkçe için morfolojik analiz veya alt kelime (subword) düzeyinde modelleme sıkça kullanılan yöntemler arasındadır.

Günümüzde NLP araştırmalarının odaklandığı konular arasında etik ve önyargılar da önemli yer tutmaktadır. Büyük dil modelleri, eğitildikleri veri setlerindeki cinsiyetçi, ırkçı veya ideolojik önyargıları öğrenebilir ve bunu çıktılarında yansıtabilir. Bu nedenle, hem akademide hem sektörde adil, tarafsız ve etik NLP sistemleri geliştirmek için yoğun çalışmalar yürütülmektedir.

Sonuç olarak, NLP alanı, hem temel kavramsal altyapısı hem de sürekli evrilen uygulamalarıyla, bilgisayarların insan dilini "anlama" yeteneğini her geçen gün daha ileriye taşımaktadır. Bu tezde ise NLP'nin özellikle metin sınıflandırma alt alanındaki gelişmiş yöntemleri, yani Transformer tabanlı yaklaşımlar, detaylı biçimde incelenecek ve geleneksel yöntemlerle karşılaştırılacaktır.

2.2 Metin Sınıflandırma Yöntemleri

Metin sınıflandırma, doğal dil işleme (NLP) alanında en yaygın kullanılan ve en temel görevlerden biridir. Genel olarak, bir metnin önceden belirlenmiş kategorilerden birine veya birkaçına atanması sürecini ifade eder. Örneğin, bir e-postanın spam olup olmadığını belirlemek, bir müşteri yorumunun olumlu mu olumsuz mu olduğunu sınıflandırmak ya da bir haber metnini ait olduğu konu başlığına (siyaset, spor, ekonomi vb.) ayırmak metin sınıflandırma örnekleri arasında yer alır.



Şekil 2.2: Metin sınıflandırma aşamaları

Bu görev, birçok uygulama için kritik öneme sahiptir. E-posta filtreleme, içerik tavsiye sistemleri, sosyal medya izleme, hukuki belge sınıflandırması, müşteri geri bildirim analizi ve tıbbi rapor analizleri gibi alanlarda metin sınıflandırma yoğun biçimde kullanılmaktadır. Özellikle günümüzün dijital dünyasında üretilen devasa metin verilerini anlamlandırabilmek ve işleyebilmek için güçlü sınıflandırma yöntemlerine ihtiyaç

duyulmaktadır. Metin sınıflandırma yöntemleri temelde iki ana gruba ayrılır: geleneksel yöntemler ve derin öğrenme tabanlı modern yöntemler.

Geleneksel Yöntemler: Bu yöntemler, özellikle 2000'li yılların başlarında NLP uygulamalarında yoğun şekilde kullanılmıştır. Genellikle makine öğrenmesi algoritmalarına dayanır ve metinleri sayısal temsillere (vektörlere) dönüştürerek sınıflandırma yapılır. Bu süreçte en yaygın kullanılan temsil yöntemleri "Bag of Words" (BoW), "Term Frequency-Inverse Document Frequency" (TF-IDF) ve n-gram tabanlı yaklaşımlardır. Bu temsiller, kelimelerin sıklığını veya önemini hesaba katarak bir metni sayısal forma sokar, ardından Naive Bayes, Destek Vektör Makineleri (SVM), Karar Ağaçları veya Lojistik Regresyon gibi algoritmalarla sınıflandırma yapılır. Geleneksel yöntemler hızlı ve kolay uygulanabilir olmalarına rağmen, kelime sırasını veya bağlamsal anlamı göz önünde bulunduramadıkları için karmaşık metinlerde sınırlı başarı gösterirler.

Derin Öğrenme ve Modern Yöntemler: Son yıllarda, yapay sinir ağlarına dayanan derin öğrenme yaklaşımları metin sınıflandırmada devrim yaratmıştır. Özellikle RNN (Recurrent Neural Networks), LSTM (Long Short-Term Memory), GRU (Gated Recurrent Units) gibi sıralı veri modelleri, metinlerdeki kelime sıralarını ve bağlamı daha iyi yakalayabilmiştir. Ancak bu modellerin yerini hızla Transformer mimarisi ve onun üzerine inşa edilen BERT, RoBERTa, GPT, T5, DistilBERT gibi önceden eğitilmiş devasa dil modelleri almıştır. Bu modeller, çok büyük veri setleri üzerinde eğitildikleri için dilin inceliklerini, bağlamsal ilişkileri ve karmaşık yapıları oldukça başarılı şekilde öğrenirler.

Transformer tabanlı modeller, transfer öğrenme prensibini kullanarak çok az örnekle bile güçlü sonuçlar üretebilirler. Önceden eğitilmiş bu modeller, metin sınıflandırma görevine özgü küçük bir veri seti üzerinde "ince ayar" (fine-tuning) yapılarak hızla uyarlanabilir. Bu sayede, geleneksel yöntemlerin aksine el yapımı özellik mühendisliğine gerek kalmaz ve performans önemli ölçüde artar.

Metin sınıflandırma süreci, sadece bir model seçmekten ibaret değildir. Aynı zamanda verilerin nasıl hazırlandığı, nasıl temizlendiği, hangi temsil yöntemlerinin kullanıldığı, hangi performans metriklerinin değerlendirildiği ve model çıktılarının nasıl yorumlandığı da büyük önem taşır. Çünkü her uygulama alanı, farklı türde zorluklar ve gereksinimler

barındırır. Örneğin, tıbbi metinler çok teknik terimler içerirken, sosyal medya verileri kısaltmalar ve argo kullanımlarına sahip olabilir. Bu yüzden metin sınıflandırma, hem mühendislik hem de uygulama bilgisi gerektiren disiplinler arası bir süreçtir.

Sonuç olarak, metin sınıflandırma yöntemleri hem geleneksel hem de modern yaklaşımlar açısından sürekli gelişmektedir. Her yöntemin avantajları ve sınırlılıkları vardır. Bu tez çalışmasında, geleneksel yöntemler olan Naive Bayes ve SVM ile modern Transformer tabanlı modeller karşılaştırılarak, metin sınıflandırma görevinde hangi yaklaşımın hangi durumlarda üstün olduğunu ortaya koymak amaçlanmıştır.

2.3 Geleneksel Yöntemler (Naive Bayes - SVM)

Doğal dil işleme (NLP) ve özellikle metin sınıflandırma problemlerinde, uzun yıllar boyunca kullanılan geleneksel makine öğrenmesi yöntemleri, derin öğrenme ve transformer tabanlı yöntemlerin gelişimine kadar önemli bir rol oynamıştır. Bu geleneksel yöntemler, genellikle istatistiksel modellere ve matematiksel kavramlara dayanır. Naive Bayes ve Destek Vektör Makineleri (SVM), bu alanda en yaygın ve başarılı iki yöntem olarak literatürde geniş bir kullanım alanına sahiptir. Her iki yöntemin de avantajları, sınırlamaları ve uygulama alanları vardır. Bu bölümde, bu iki yöntemin çalışma prensipleri, güçlü ve zayıf yönleri, metin sınıflandırmadaki performansları ve günümüz yaklaşımlarına kıyasla konumları detaylı olarak incelenecektir.

Naive Bayes sınıflandırıcısı, olasılık teorisine ve özellikle Bayes Teoremi'ne dayanan bir yöntemdir. Modelin "naive" (saf, basitleştirilmiş) olarak adlandırılmasının nedeni, her bir özelliğin (kelime, nitelik) birbirinden bağımsız olduğu varsayımıdır. Gerçek dilde kelimeler arası bağımlılıklar oldukça yaygın olmasına rağmen, bu basitleştirilmiş yaklaşım birçok metin sınıflandırma görevinde oldukça iyi sonuçlar verebilmektedir.

Multinomial Naive Bayes, metin sınıflandırmada özellikle popülerdir çünkü metni kelimelerin sıklığı veya varlığı üzerinden değerlendirir. Bu model, spam tespiti, duygu analizi gibi görevlerde yıllardır yaygın bir şekilde kullanılmaktadır. Naive Bayes'in güçlü yönleri arasında hızlı eğitim süreci, düşük hesaplama maliyeti ve az sayıda parametre ayarlaması gerektirmesi yer alır. Özellikle veri setinin sınırlı olduğu veya hızlı bir prototip

geliştirme gerektiği durumlarda, karmaşık modellere kıyasla çok daha pratik bir çözüm sunar.

Ancak bu modelin sınırlamaları da belirgindir. Kelimeler arası bağımlılıkların ve dilin bağlamsal yapısının göz ardı edilmesi, özellikle çok dilli metinlerde veya ironik ifadelerde performans kaybına yol açar. Örneğin, "Bu filmi hiç beğenmedim" ve "Hiç fena değil" gibi ifadelerde kelimelerin bireysel anlamlarından daha fazlası devreye girer. Ayrıca nadir kelimelerin veya nadir kombinasyonların ağırlıklandırılması konusunda yetersiz kalabilir.

Destek Vektör Makineleri (SVM), özellikle yüksek boyutlu veri setlerinde güçlü sınıflandırma yetenekleri sunan bir yöntemdir. SVM'nin temel çalışma mantığı, farklı sınıflar arasında en geniş marjine (margin) sahip bir hiper düzlem bulmaktır. Bu, sınıflar arasındaki ayrımı maksimize eder ve modelin genelleme kapasitesini artırır. Metin sınıflandırma gibi veri boyutunun çok yüksek olduğu uygulamalarda, SVM'nin doğrusal çekirdek (linear kernel) kullanımı oldukça yaygındır.

SVM'nin avantajlarından biri, yüksek boyutlu ve seyrek veri setlerinde dahi güçlü performans sergilemesidir. Ayrıca model, özellikle küçük veri setlerinde overfitting'e (aşırı öğrenmeye) karşı dayanıklıdır. Sınıflar arasındaki sınırları net bir şekilde çizer ve genellikle çok az parametre ayarıyla bile başarılı sonuçlar verir.

Ancak SVM'nin sınırlamaları da göz ardı edilmemelidir. Özellikle çok büyük veri setlerinde eğitimi oldukça zaman alabilir ve bellek yoğunluğu yaratabilir. Ayrıca çekirdek fonksiyonu ve C parametresi gibi ayarların yanlış seçilmesi, model performansını önemli ölçüde düşürebilir. SVM'nin bir başka dezavantajı ise sınıflandırma kararlarının olasılıksal çıkışlar yerine sadece sınıf etiketleri ile sınırlı olmasıdır; bu da bazen sınıflandırma kararlarının güvenirliğini ölçmede eksiklik yaratır.

Geleneksel Yöntemlerin Modern Yaklaşımlar Karşısındaki Durumuna bakacak olursak Naive Bayes ve SVM, 2000'li yılların başından itibaren metin sınıflandırma görevlerinde akademik çalışmaların ve endüstriyel uygulamaların bel kemiğini oluşturmuştur. Ancak son yıllarda, özellikle derin öğrenme ve transformer tabanlı modellerin yükselişi ile birlikte bu geleneksel yöntemlerin kullanım alanları daralmıştır. Transformer modelleri,

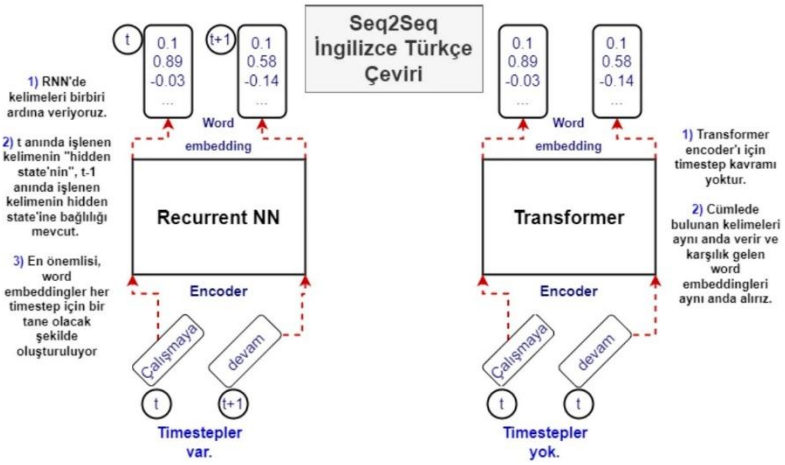
bağlam bilgisi, kelime ilişkileri ve dilin karmaşık yapısını çok daha iyi modelleyebilirken, Naive Bayes ve SVM gibi geleneksel yöntemler bu zenginlikleri yakalamakta yetersiz kalmaktadır.

Yine de, bu geleneksel yöntemlerin tamamen göz ardı edilmesi doğru değildir. Özellikle küçük ölçekli veri setlerinde, düşük hesaplama gücüne sahip sistemlerde veya hızlı bir temel çözüm geliştirmek için geleneksel yöntemler hâlâ önem taşır. Ayrıca birçok araştırmada, derin öğrenme tabanlı modellerin performanslarını kıyaslamak için bir "benchmark" (karşılaştırma referansı) olarak sıklıkla Naive Bayes ve SVM gibi yöntemler kullanılır.

Bu tez çalışmasında yapılan deneyler de hem geleneksel yöntemlerin hem de modern transformer tabanlı yaklaşımların performansını kapsamlı bir şekilde analiz etmektedir. Doğruluk, hassasiyet, geri çağırma ve F1 skoru gibi metrikler üzerinden yapılan karşılaştırmalar, geleneksel yöntemlerin bazı avantajlarını ve hangi durumlarda yetersiz kaldıklarını açıkça ortaya koymaktadır. Geleneksel yöntemlerin hızlı ve basit çözümler sunduğu; ancak karmaşık bağlamsal bilgilerin gerekli olduğu görevlerde modern yöntemlerin üstün geldiği sonucuna varılmıştır.

2.4 Transformer Mimarisi ve Modelleri

Doğal dil işleme (NLP) alanında son yıllarda yaşanan en önemli atılımlardan biri, transformer mimarisinin geliştirilmesi ve yaygınlaşması olmuştur. Bu mimari, o döneme kadar yaygın olarak kullanılan RNN (Recurrent Neural Network) ve LSTM (Long Short-Term Memory) gibi sıralı modellerin sınırlamalarını ortadan kaldırarak, dikkat (attention) mekanizmasına dayalı yeni bir öğrenme paradigması getirmiştir. Transformer, dilin karmaşık yapısını anlamada, uzun mesafeli bağımlılıkları modellemede ve büyük ölçekli veri üzerinde eğitilebilme kapasitesiyle NLP alanında devrim yaratmıştır.



Şekil 2.3: RNN ile transformer karşılaştırması [2]

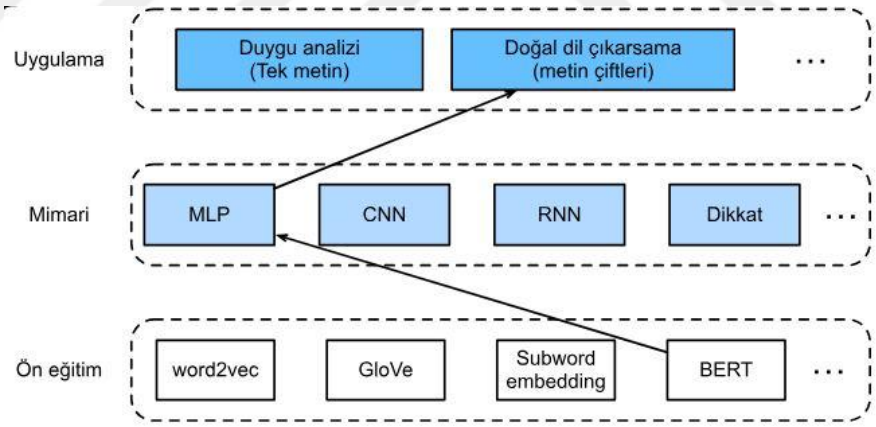
Transformer Mimarisi temel olarak encoder (kodlayıcı) ve decoder (çözücü) bloklarından oluşur. Encoder, giriş metnini işler ve bağlamsal temsiller (contextual representations) üretir; decoder ise bu temsilleri kullanarak hedef çıktıyı üretir. Encoder ve decoder katmanlarının her biri, çok başlı dikkat (multi-head attention), konum kodlamaları (positional encoding), besleme ileri ağlar (feed-forward layers) ve rezidüel bağlantılar (residual connections) içerir. Konum kodlamaları, modelin sırayı algılayabilmesini sağlar; çünkü saf attention yapısı dizilerdeki sıralı bilgiyi doğrudan işlemez.

Transformer'ın en çarpıcı özelliği, giriş dizisindeki her öğenin (örneğin her kelimenin), dizinin tüm diğer öğeleriyle aynı anda ve doğrudan etkileşim kurabilmesidir. Bu, özellikle uzun metinlerde önemli bilgilerin korunmasını sağlar. Örneğin, klasik RNN veya LSTM tabanlı modellerde, bir kelimenin önceki kelimelerle olan ilişkisi zaman içinde "unutulabilirken," transformer'da self-attention mekanizması sayesinde tüm kelimeler arası bağlar doğrudan hesaplanır. Bu durum, modelin hem daha hızlı hem de daha doğru çalışmasına olanak tanır.

Attention Mekanizması yani Dikkat (attention) mekanizması, Transformer'ın merkezinde yer alan yeniliklerden biridir. Bu mekanizma, girişteki her kelimenin, diğer tüm kelimelerle olan ilişkisini değerlendirerek hangilerinin daha önemli olduğuna karar verir. Özellikle multi-head attention kullanımı, farklı başların farklı türden ilişkileri paralel olarak öğrenebilmesini sağlar. Örneğin, bir başlık sözdizimsel (syntactic) ilişkileri öğrenirken, başka bir başlık anlamsal (semantic) ilişkileri yakalayabilir.

Transformer Modellerinin Evrimi ve Öne Çıkanlar: Transformer mimarisi üzerine inşa edilen birçok güçlü model geliştirilmiştir ve bunlar, NLP alanında birbiri ardına rekorlar kırmıştır. Bu tezde, özellikle GPT, BERT, T5, modelleri üzerinde yoğunlaşılmıştır:

- **GPT (Generative Pre-trained Transformer):** GPT, OpenAI tarafından geliştirilen ve yalnızca decoder katmanlarına dayanan otoregresif bir modeldir. Büyük miktarda metin üzerinde eğitilerek, özellikle metin üretimi, tamamlaması, özetleme ve diyalog oluşturma gibi görevlerde etkileyici sonuçlar verir. GPT'nin ardışık kelimeleri tahmin etme yeteneği, doğal ve akıcı metinler üretmesini sağlar.
- **BERT (Bidirectional Encoder Representations from Transformers):** Google tarafından geliştirilen BERT, yalnızca encoder katmanlarını kullanır ve çift yönlü öğrenme yeteneğine sahiptir. Yani bir kelimenin yalnızca öncesine değil, aynı zamanda sonrasına da bakarak daha derin bağlamsal anlayış geliştirir. BERT, özellikle sınıflandırma, adlandırılmış varlık tanıma (NER) ve soru yanıtlama gibi görevlerde başarı göstermiştir.



Şekil 2.4: BERT algoritması [3]

- **T5 (Text-to-Text Transfer Transformer):** T5 modeli, tüm NLP görevlerini bir metni metne dönüştürme (text-to-text) problemi olarak yeniden tanımlayan bir yaklaşıma sahiptir. Bu model, çeviri, özetleme, sınıflandırma ve hatta soru yanıtlama gibi çok farklı görevlerde evrensel bir çözüm sunabilmektedir. Bu tezde, T5 modeli deneysel çalışmalarda özellikle metin sınıflandırma görevinde öne çıkan transformer tabanlı model olarak seçilmiştir.

Transformer'ın Uygulamalardaki Avantajları ve Zorlukları: Transformer mimarisinin avantajları arasında paralel hesaplama kabiliyeti, bağlamsal anlam öğrenme kapasitesi ve çok görevli öğrenme (multi-task learning) yeteneği yer alır. Ancak bu mimarinin eğitim maliyetleri oldukça yüksektir; devasa miktarda parametre ve büyük veri setleri gerektirir. Örneğin GPT-3 gibi modeller milyarlarca parametre içerir ve eğitilmeleri günler veya haftalar alabilir. Ayrıca, transformer tabanlı modellerin kararlarının yorumlanabilirliği (explainability) hâlâ önemli bir araştırma konusudur; modelin bir sonuca nasıl vardığını anlamak, özellikle kritik uygulamalarda önemlidir.

Metin Sınıflandırmada Transformer'ın Rolü: Transformer modelleri, metin sınıflandırma görevlerinde devrim yaratmıştır. Geleneksel yöntemler, genellikle kelime sayma veya basit kelime temsilleri ile sınırlıyken, transformer tabanlı modeller çok daha derin ve zengin dil temsilleri üretebilir. Bu, özellikle ince duygu farklarının, karmaşık konu yapılarının ve nadir kelime örüntülerinin doğru şekilde yakalanmasını sağlar. Tez kapsamında yapılan deneylerde, transformer tabanlı modellerin (özellikle T5'in) geleneksel yöntemlere (Naive Bayes, SVM) kıyasla doğruluk, hassasiyet ve F1 skoru gibi metriklerde anlamlı üstünlük gösterdiği gözlemlenmiştir.

Transformer Modellerinin Geleceği: Transformer mimarisi yalnızca NLP alanında değil, görüntü işleme (Vision Transformers - ViT), biyoinformatik, protein yapı tahmini ve çok modlu (multi-modal) yapay zekâ uygulamalarında da yaygınlaşmaktadır. Gelecekte bu modellerin daha verimli, daha az kaynak tüketen ve daha açıklanabilir versiyonlarının geliştirilmesi beklenmektedir. Ayrıca, düşük kaynaklı diller ve özel alanlar için özelleştirilmiş transformer modellerinin geliştirilmesi, alanın önemli araştırma konularından biri olmaya devam etmektedir.

2.5 Mevcut Literatürde Transformer Tabanlı Sınıflandırma Çalışmaları

Arzu ve Aydoğan (2023), Türkçe dilinde de Transformer tabanlı modellerin metin sınıflandırma üzerindeki etkisini inceleyen "Türkçe Duygu Sınıflandırma İçin Transformer Tabanlı Mimarilerin Karşılaştırılmalı Analizi" başlıklı çalışmalarında, farklı Transformer tabanlı mimarilerin Türkçe duygu sınıflandırma görevindeki performansları karşılaştırılmıştır. Bu çalışmada, 150.000 veriden oluşan TRSAv1 veri seti üzerinde, 8 farklı BERTurk ve 2 farklı ELECTRA varyasyonu kullanılmıştır. Çalışma sonucunda,

modellerin duygu analizi performansları doğruluk ve F1 skor değerleri kullanılarak ölçülmüş ve Transformer tabanlı modellerin Türkçe metinlerde de yüksek performans sergilediği gösterilmiştir [4].

Özkan ve Kar (2022), "Türkçe Dilinde Yazılan Bilimsel Metinlerin Derin Öğrenme Tekniği Uygulanarak Çoklu Sınıflandırılması" başlıklı çalışmalarında, Türkçe akademik metinleri sınıflandırmak için BERT modelini kullanmışlardır. Sonuçlar, %96 doğruluk oranı ile modelin akademik metin sınıflandırmasında oldukça etkili olduğu görülmüştür. Özellikle metin uzunluğunun fazla olduğu durumlarda Transformer modelleri, bağlamsal anlamda geleneksel yöntemlerden daha iyi sonuç vermiştir [5].

Varan ve arkadaşları (2024), sosyal medya fenomenlerinin hedef kitleleriyle etkili bir şekilde eşleştirilmesi amacıyla BERTopic modelini kullanmıştır. SBERT (Sentence-BERT) tabanlı bu model, metinlerin semantik anlamlarını analiz ederek kümeler oluşturmuştur. Çalışma, yüksek boyutlu verileri UMAP yöntemiyle boyut indirgemeye tabi tutmuş ve HDBSCAN algoritmasıyla kümelenmiştir. Bu çalışma, pazarlama stratejilerinde hedef kitle ile içerik arasında daha iyi bir uyum sağlanabileceğini göstermiştir. BERTopic modeli, %90'ın üzerinde bir başarı oranına ulaşmıştır [6].

Yürütücü ve arkadaşları, yaptıkları projede önceden eğitilen üç dil model ile Covid-19 virüsü tweetlerinden oluşan veri setinden duygu analizi yapmıştır. Çalışmada Twitter'da Covid-19 ile ilgili paylaşılan tweetlerin üç farklı dil modeline göre sonuçlarına bakılmıştır. Her üç algoritmanın da sınıflandırma başarısı birbirine yakın çıkmıştır. Her üç algoritmanın da %90 civarında performans puanları ile sınıflandırma görevinde iyi performans gösterdiği tespit edilmiştir. Ayrıca sınıflandırma başarı oranları da yeterlidir. BERT modellerinin yüksek doğruluk oranları olmasının nedeni saf Bayes ve Lojistik Regresyon bağlamın anlamsal ve yapısal bilgilerini ihmal etmekte, karakteristik kelimeler ve kategoriler arasındaki nedenselliğe odaklanmaktadır [7].

Karaahmetoğlu ve arkadaşları, korona virüs ile ilgili "korona" kelimesi ile aranan Twitter verilerini toplayarak, elde ettikleri veri seti üzerinde duygu analizi çalışması yapmıştır. Sözlük tabanlı, makine öğrenmesi yöntemleri ve BERT Modeli ile duygu analizleri yapılmış ve sonuçlar gösterilmiştir. Sözlük tabanlı ve BERT Model algoritmaları ile geliştirilen duygu analizlerinde, pozitif mesajların sayısı, negatif mesajların 7 katı kadar

olmuştur. Bu sonuçlara bakıldığında korona virüs ile ilgili insanların moral seviyesinin yüksek olduğu sonucuna varılabilir. Çalışmanın sonunda elde edilen sonuçların güvenilirliği farklı değerlendirme yöntemleri ile incelenmiş olup, sözlük tabanlı algoritma için 1, bert için 0,98 olmak üzere, yüksek güvenilirlik sonuçları elde edilmiştir. Çalışma veri setinde öğrenme ve test verisi olarak kullanılan tweet metinlerinde duygu değerleri (y) sözlük tabanlı duygu analizi algoritması ile belirlenmiştir. Bundan dolayı, sözlük tabanlı algoritmanın güvenilirlik değerleri yüksek çıkmıştır [8].

Çelik ve arkadaşları, "Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması" adlı çalışmalarında Destek Vektör Makinesi, K-En Yakın Komşu, Naive Bayes, Karar Ağacı ve Rastgele Orman sınıflandırma algoritmalarının performansı karşılaştırmışlardır. Tüm algoritmaların her bir veri seti ile tek tek çalışmasına göre tüm veri setlerini birleştirip oluşan tek bir veri seti ile eğitilmesinin performansı yükselttiği gözlemlenmiştir. Bu durum veri farklılığının ve büyüklüğünün önemini bir kere daha göstermiştir. Diğer yandan DVM'nin birleşmiş veri setinin veri farklılığına ve büyüklüğüne rağmen %85'lik bir doğruluk oranıyla diğer algoritmalara göre daha etkili bir analiz aracı olduğu gözlemlenmiştir [9].

Alpkoçak ve arkadaşları, doğrulanmış TREMO veri seti üzerinde literatürde çok sık kullanılan KEYK, RF, YSA ve DVM makine öğrenme algoritmaları doğruluk değeri kullanılarak duygu analizi sonuçlarına bakılmıştır. Bunun için ilk olarak TREMO veri setine ön işleme aşaması uygulanmış ve F5 yöntemi ile TREMO'da bulunan bütün kelimeler köklerine ayrılmıştır. Sonrasında, veri setinde yer alan her bir duygu için en değerli ilk 500 kelime MI formülü kullanılarak tespit edilmiş, özellik seçimi yapılmış ve veri seti uzay vektör modeli kullanılarak model elde edilmiştir. Elde edilen bu model KEYK, RF, YSA ve DVM algoritmalarını eğitmek ve test etmek için kullanılmıştır. Değerlendirme ölçütü olarak doğruluk değerinin kullanıldığı sınıflandırma metriklerine bakıldığında, genel doğruluk oranlarında YSA algoritmasının diğer algoritmalara göre daha üstün olduğu gözlemlenmiştir [10].

İlhan ve arkadaşları, Twitter kullanıcı gönderilerine ait sınıflandırılmış tweet verileri kullanarak çeşitli makine öğrenme yöntemleri kullanarak duygu analizi çalışması yapmıştır. N-gram yöntemi ile olumlu ve olumsuz duyguların analizi yapılmış, makine

öğrenmesinde en çok kullanılan Destek Vektör Makineleri ve Naive Bayes yöntemleri ile ilgili sınıflandırıcıların performansları karşılaştırılmıştır. Sonuçlara bakıldığında, en yüksek değeri Destek Vektör Makineleri sınıflandırıcısının verdiği görülmüştür [11].

S. Tuzcu çalışmasında, çevrimiçi bir kitap satış sitesinin müşteri yorumlarının duygu sınıflandırmasına yönelik RapidMiner ve Python kullanarak, verileri ön işlemlerden geçirdikten sonra analizler yapmıştır. Çalışmasında kullandığı algoritmalar; Destek Vektör Makinası (DVM), Multi Layer Perceptron (MLP), Lojistik Regresyon (LR) ve Naive Bayes (NB) algoritmalarıdır. DVM, olumlu metinleri sınıflandırmada diğer algoritmalara göre en iyi sonucu vermiş olsa da olumsuz yorumları sınıflandırma konusunda diğer algoritmaların oldukça gerisinde kaldığı görülmektedir. MLP, olumsuz ve olumlu her iki sınıf için de yüksek bir doğruluk sağlamakta ve olumsuz yorumları en iyi sınıflandıran algoritma olduğu görülmektedir. LR algoritması da MLP algoritmasına göre yakın sayılacak derecede iyi sonuçlar vermiştir. NB ise diğerlerinde farklı olarak olumsuz metinleri olumlu metinlere göre daha iyi sınıflandırmış ama genel sınıflandırma olarak diğer algoritmaların başarısının altında kalmıştır [12].

H. Barzenjı çalışmasında, Türkçe Twitter metinlerini sınıflandırılmasında, üç genel makine öğrenmesi yöntemi (destek vektör makineleri, lojistik regresyon ve Naive Bayes algoritması) ve üç temel temsil modeli (1-gram, 2-gram ve 3-gram) ile bu temsil modellerinin farklı bileşenleri değerlendirilmiştir. Bu çalışmalarda, en yüksek başarımın (%77,78), veri seti olarak 1-gram ve 2-gram öznitelik veri setlerinin birleşmesi ile oluşan öznitelik veri seti ile temsil edildiğinde ve Naive Bayes sınıflandırma algoritması kullanıldığında elde edildiği gözlemlenmiştir [13].

Son yıllarda doğal dil işleme (NLP) alanında transformer mimarisi, metin sınıflandırma problemlerine yaklaşımı köklü bir şekilde değiştirmiştir. Transformer, yalnızca dil modeli görevlerinde değil; metin sınıflandırma, duygu analizi, haber sınıflandırma, konu modelleme, sahte haber tespiti ve spam algılama gibi çeşitli uygulamalarda da çığır açan bir yöntem olarak dikkat çekmektedir. Transformer'ın başarısı, temel olarak self-attention mekanizmasının sağladığı güçlü bağlamsal ilişki yakalama kapasitesinden kaynaklanmaktadır.

BERT (Bidirectional Encoder Representations from Transformers), 2018 yılında Google tarafından geliştirildiğinde metin sınıflandırma problemlerinde bir dönüm noktası oldu. BERT, çift yönlü bağlam anlayışı sayesinde hem önceki hem de sonraki kelimelerin anlamını hesaba katarak, geleneksel yöntemlerden (örneğin Naive Bayes veya SVM) çok daha doğru sınıflandırma sonuçları verebilmektedir. Örneğin, SST-2 (Stanford Sentiment Treebank) veri seti üzerinde yapılan deneylerde, BERT modelleri %90'ın üzerinde doğruluk oranlarına ulaşarak geleneksel yöntemlere büyük fark atmıştır. Aynı şekilde IMDB ve AG News gibi veri setlerinde de BERT, bağlamsal kelime temsilleriyle sınıflandırma başarısını önemli ölçüde artırmıştır.

RoBERTa (Robustly Optimized BERT Pretraining Approach), BERT'in geliştirilmiş bir versiyonudur ve eğitim parametrelerinin yeniden optimize edilmesi, daha büyük bir veri kümesi kullanılması ve bazı sınırlamaların kaldırılması sayesinde BERT'ten daha iyi performans sergilemektedir. Literatürde RoBERTa'nın, özellikle büyük ölçekli metin sınıflandırma problemlerinde, hassasiyet ve hatırlama (precision, recall) metriklerinde BERT'in önüne geçtiği gösterilmiştir [12].

GPT (Generative Pre-trained Transformer) ailesi, özellikle metin üretimi görevleriyle öne çıksa da sınıflandırma problemlerinde de etkin bir şekilde kullanılmaktadır. GPT, dil modellemesi görevine odaklanmış olsa da transfer öğrenme yoluyla sınıflandırma senaryolarında başarı sağlar. Ancak, GPT'nin tek yönlü (sadece soldan sağa) bağlam anlayışı, BERT gibi çift yönlü modellerle kıyaslandığında belirli sınıflandırma görevlerinde kısmi bir dezavantaj yaratabilir. Buna rağmen, literatürde GPT'nin sınıflandırma görevlerine uygun şekilde ince ayarlandığında (fine-tuning), özellikle açık uçlu sınıflandırma problemlerinde yaratıcı ve etkili çözümler ürettiği gösterilmiştir [13].

T5 (Text-to-Text Transfer Transformer), diğer modellerden farklı olarak tüm NLP görevlerini bir çeviri problemine dönüştürerek ele alır. Örneğin, "Bu yorum olumlu mu, olumsuz mu?" sorusu, doğrudan bir metin çıktısı üreterek yanıtlanır. Literatürde, T5'in özellikle çok görevli öğrenme (multi-task learning) senaryolarında ve metin sınıflandırma görevlerinde başarılı sonuçlar verdiği kanıtlanmıştır [14]. Özellikle çok dilli veri kümelerinde T5'in sunduğu esneklik, transformer tabanlı yaklaşımların gücünü bir kez daha göstermektedir.

Transformer tabanlı modellerin başarısının temel nedenlerinden biri, önceden eğitilmiş büyük dil modellerinin küçük etiketli veri kümelerinde bile güçlü sonuçlar verebilmesidir. Bu, düşük veri ortamlarında transfer öğrenmenin avantajlarını açığa çıkarır. Örneğin, Naive Bayes ve SVM gibi geleneksel yöntemler, güçlü performans gösterebilmek için kapsamlı özellik mühendisliğine ve geniş etiketli veri setlerine ihtiyaç duyar. Transformer modelleri ise, sadece birkaç örnekle dahi hassas ince ayar yapılabilmesine olanak sağlar.

Literatürde tartışılan sınırlamalara bakıldığında transformer modellerinin eğitimi için gereken hesaplama gücü, GPU ve TPU gibi özel donanımlar gerektirir. Bu, özellikle sınırlı kaynaklara sahip akademik çalışmalar veya küçük ölçekli işletmeler için önemli bir engeldir. Bunun yanında, büyük dil modellerinin enerji tüketimi ve karbon ayak izi, çevresel ve etik tartışmaları da beraberinde getirmiştir. Bu nedenle güncel çalışmalar, model sıkıştırma, distilasyon (örn. DistilBERT) ve verimli ince ayar yöntemleri üzerine yoğunlaşmıştır [15].

Sonuç olarak, literatürde transformer tabanlı yaklaşımların, metin sınıflandırma problemlerinde geleneksel yöntemlerden çok daha üstün olduğu net bir şekilde ortaya konmuştur. Bununla birlikte, modellerin eğitim sürecindeki zorluklar ve uygulamada karşılaşılan verimlilik problemleri, araştırmacıları sürekli yenilikler geliştirmeye itmektedir. Gelecekte, transformer mimarisinin daha verimli, çevre dostu ve erişilebilir hale getirilmesi hem akademik hem de endüstriyel çalışmalar için kritik bir hedef olmaya devam edecektir.

3. YÖNTEM

3.1 Veri Setleri

Bu çalışmada kullanılan veri setleri, metin sınıflandırma problemlerinde hem geleneksel hem de modern derin öğrenme tabanlı modellerin performanslarını kapsamlı bir şekilde karşılaştırabilmek için dikkatle seçilmiştir. Çalışmada kullanılan üç temel veri seti şunlardır: AG News, IMDb Film İncelemeleri ve Stanford Sentiment Treebank (SST-2). Bu veri setlerinin her biri, farklı türden metin sınıflandırma problemlerini ve modellerin bu problemlerdeki güçlü ve zayıf yönlerini ortaya koymak için idealdir. Veri setlerinin farklı yapıları, içerik türleri, metin uzunlukları ve sınıf dağılımları, test edilen algoritmaların genel başarımını değerlendirmek açısından kritik rol oynamaktadır.

AG News Veri Seti, çok sınıflı metin sınıflandırma problemleri için yaygın şekilde kullanılan ve haber metinlerinden oluşan bir benchmark setidir. Veri seti, AG'nin haber makaleleri arşivinden çekilen başlıklar ve özetlerden oluşur. Toplam dört temel kategori içerir: Dünya, Spor, İş ve Bilim/Teknoloji. Eğitim kümesinde her kategori için 30.000 örnek, test kümesinde ise 1.900 örnek bulunur.

Bu veri setinin en önemli özelliklerinden biri, kısa ve orta uzunluktaki metinlerden oluşmasıdır. Geleneksel yöntemler (Naive Bayes, SVM) bu tür içeriklerde kelime dağılımlarına dayalı basit modellerle iyi sonuçlar verebilirken, transformer tabanlı modeller bağlamsal ilişkilere ve kelime dizilimlerinin ardındaki derin yapıya odaklanarak performans artışı sağlayabilir. AG News aynı zamanda çok sınıflı bir problem sunduğu için ikili sınıflandırma problemlerinden daha zorlu ve karmaşıktır, bu da modellerin sınıf ayrımlarını ne kadar iyi öğrenebildiğini gözlemlemek için önemlidir.

Çizelge 3.1: AG news veri seti

Açıklama	AG News veri seti, haber başlıklarının dört farklı kategoriye ayrıldığı bir veri setidir. Bu veri seti, metin sınıflandırma görevinde en yaygın kullanılan veri setlerinden biridir.
Eğitim ve Test Verisi	Veri seti; 120,000 eğitim örneği 7,600 test örneği içerir
Kategoriler	1. Dünya 2. Spor 3. İş Dünyası 4. Bilim ve Teknoloji

IMDb veri seti, duygu analizi ve ikili sınıflandırma problemleri için doğal dil işleme alanında klasikleşmiş bir benchmark'tır. Veri seti, internet üzerindeki en büyük film veritabanı olan Internet Movie Database (IMDb) üzerinden toplanan 50.000 film incelemesinden oluşur. Bu incelemeler, eşit sayıda olumlu ve olumsuz etiketlenmiş metinler içerir.

IMDb veri setinin dikkat çekici yönlerinden biri, içerdiği metinlerin uzunluğu ve karmaşıklığıdır. Kullanıcı incelemeleri genellikle paragraflar hâlinde ve detaylı düşünceleri barındırır. Bu durum, kelime dizilerini ve uzun menzilli bağımlılıkları yakalayabilen transformer tabanlı modeller için güçlü bir avantaj sağlar. Geleneksel yöntemler ise çoğunlukla kelime frekanslarına veya basit n-gram modellere dayandığından, bağlamsal detayların farkına varmada sınırlı kalabilir. IMDb veri seti bu nedenle sadece sınıflandırma başarımı açısından değil, aynı zamanda uzun metinlerdeki bağlam yakalama kapasitesi açısından da zengin bir test alanı sunmaktadır.

Çizelge 3.2: IMDb veri seti

Açıklama	IMDb veri seti, film yorumlarının yer aldığı ve her yorumun olumlu veya olumsuz olduğu bir veri setidir. Bu veri seti, duygu analizi gibi görevler için yaygın olarak kullanılır.
Eğitim ve Test Verisi	Veri seti; 25,000 eğitim örneği 25,000 test örneği içerir
Kategoriler	1. Olumlu 2. Olumsuz

Stanford Sentiment Treebank (SST-2) Veri Seti, Stanford Üniversitesi tarafından geliştirilmiş ve doğal dil işleme alanında duygu analizi için standart bir benchmark hâline gelmiş bir veri setidir. SST-2, Rotten Tomatoes film incelemelerinden türetilmiş cümleleri içerir ve her cümle olumlu veya olumsuz duygu taşıyıp taşımadığına göre etiketlenmiştir. Veri seti yalnızca cümle düzeyinde etiketleme içerdiği için kısa metinlerdeki sınıflandırma başarısını test etmek için uygundur.

Transformer mimarisi, özellikle kısa cümlelerdeki kelime öbeği ilişkilerini ve ince bağlamsal detayları yakalamada geleneksel yöntemlere göre daha üstündür. SST-2 veri seti, literatürde çoğu zaman BERT, RoBERTa, DistilBERT, GPT ve T5 gibi gelişmiş transformer modellerinin eğitilmesi ve kıyaslanması için kullanılmaktadır. Çalışmada

SST-2'nin kullanılması, hem kısa hem uzun metinlerde model başarımını karşılaştırmak için değerli bir kıyas noktası oluşturur.

Çizelge 3.3: SST-2 veri seti

Açıklama	Stanford Sentiment Treebank 2 (SST-2), film cümlelerinin olumlu veya olumsuz duygu taşıyan metinler olarak sınıflandırıldığı bir veri setidir. Duygu analizi üzerine yapılan araştırmaların temel veri setlerinden biridir.
Eğitim ve Test Verisi	Veri seti; 67,000 eğitim örneği 1,821 test örneği içerir
Kategoriler	1. Olumlu 2. Olumsuz

Veri Setlerinin Özellikleri ve Karşılaştırılması: Bu üç veri setinin birlikte kullanılması, çalışmanın kapsamını genişletmekte ve modellerin farklı metin türlerinde nasıl performans gösterdiğini kapsamlı şekilde analiz etmeye olanak sağlamaktadır. AG News çok sınıflı haber sınıflandırma sunarken, IMDb uzun metinlerde ikili duygu analizi, SST-2 ise kısa cümle bazlı ikili duygu analizi sağlar.

Her veri setinde eğitim, doğrulama ve test kümeleri oluşturulmuş, eğitim kümesi modelin öğrenmesi için, doğrulama kümesi hiperparametre ayarlarının yapılması ve aşırı öğrenmenin (overfitting) engellenmesi için, test kümesi ise nihai performansın değerlendirilmesi için kullanılmıştır. Ayrıca sınıf dengesizliklerini önlemek amacıyla, tüm veri setlerinde dengeli bir sınıf dağılımı sağlanmış, dolayısıyla modellerin belirli bir sınıfa eğilim göstermesi engellenmiştir.

3.2 Veri Ön İşleme

Doğal dil işleme (NLP) projelerinde veri ön işleme, modelin başarısı için kritik bir adımdır. Kullanılan modeller ister geleneksel makine öğrenmesi yöntemleri (Naive Bayes, SVM) ister derin öğrenme tabanlı transformer mimarileri (BERT, RoBERTa, GPT, T5, DistilBERT) olsun, ham metin verisinin temizlenmesi, düzenlenmesi ve modele uygun formata getirilmesi gerekmektedir.

Bu çalışmada kullanılan AG News, IMDb ve SST-2 veri setleri, doğrudan metin tabanlı olduğu için çeşitli ön işleme adımlarından geçirilmiştir. Öncelikle, tüm veri setlerinde metinlerden noktalama işaretleri, özel karakterler ve HTML etiketleri temizlenmiştir. Bu,

metinlerin sadeleştirilmesini ve modelin yalnızca anlamlı içerik üzerinde odaklanmasını sağlar. Ancak transformer tabanlı modellerin çoğu, noktalama işaretlerinin taşıdığı anlamsal ipuçlarını kullanabildiği için, bu temizleme işlemi yalnızca geleneksel yöntemler için uygulanmıştır.

Ardından, metinler küçük harfe dönüştürülmüştür. Küçük harfe dönüştürme, metindeki kelimelerin aynı kökten geldiği durumlarda gereksiz ayrımların önüne geçer. Örneğin, "Film" ve "film" kelimelerinin aynı temsili taşıması sağlanır. Bununla birlikte, bazı transformer modelleri (örneğin BERT) "cased" ve "uncased" olmak üzere farklı varyantlarda çalışabildiği için bu dönüşüm, kullanılan modele bağlı olarak esnek şekilde ele alınmıştır.

Stop-word (önemsiz kelimeler) temizliği, geleneksel yöntemler için yaygın bir adımdır ve bu çalışmada Naive Bayes ve SVM modellerinde uygulanmıştır. "Ve", "ile", "ama" gibi dilde sıkça geçen ancak sınıflandırma için fazla bilgi taşımayan kelimeler listelenmiş ve metinlerden çıkarılmıştır. Ancak transformer tabanlı modeller, bağlam anlayışları sayesinde stop-word'leri de dikkate alarak çalışabildiği için bu modellerde stop-word temizliği yapılmamıştır.

Bir diğer önemli adım, lemmatizasyon veya kökleştirme (stemming) işlemleridir. Geleneksel yöntemlerde bu adım uygulanarak kelimeler kök biçimlerine indirgenmiş, böylece "koşuyor", "koştu" ve "koşmak" gibi kelimeler tek bir temsil altında toplanmıştır. Transformer mimarileri ise kelime vektörleri ve alt kelime düzeyinde çalışan tokenizasyon yöntemleri sayesinde bu tür işlemleri modelin içinde doğal olarak ele alabilmektedir.

Transformer modellerine özel olarak, metinler tokenize edilmiştir. Bu işlemde, modelin kullandığı özel tokenizer (örn. WordPiece, SentencePiece, Byte-Pair Encoding) devreye girerek metinleri alt kelime veya karakter seviyesinde temsil eden dizilere dönüştürmüştür. Her model, kendine özgü bir tokenizer'a sahip olduğu için, her birine uygun ön işleme ve dönüştürme süreçleri uygulanmıştır. Örneğin BERT modeli için WordPiece tokenizasyonu, GPT için Byte-Pair Encoding (BPE), T5 için SentencePiece kullanılmıştır.

Son olarak, veri setlerindeki metinler maksimum uzunluk kısıtlamalarına göre ayarlanmıştır. Transformer modelleri belirli bir maksimum girdi uzunluğu sınırı ile çalışır (örneğin 512 token), bu yüzden çok uzun metinler kesilmiş ya da uygun şekilde özetlenmiştir. Geleneksel yöntemlerde ise metin uzunlukları genellikle doğrudan bir sınırlama oluşturmadığı için bu adım daha esnek ele alınmıştır.

Bu detaylı ön işleme süreci, tüm modellerin aynı ve tutarlı veri üzerinde çalışmasını sağlamış, böylece yapılan karşılaştırmaların adil ve güvenilir olmasına katkı sağlamıştır. Veri ön işleme adımları sayesinde metinler, hem anlam kaybı yaşamadan sadeleştirilmiş hem de her modelin ihtiyaç duyduğu formata uygun şekilde hazırlanmıştır.

3.3 Model Mimarisi

Bu tezde metin sınıflandırma problemini çözmek için hem geleneksel makine öğrenmesi yöntemleri hem de modern derin öğrenme tabanlı Transformer mimarileri uygulanmıştır. Her iki yaklaşım farklı avantajlar ve zorluklar taşımakta, mimari yapıları, öğrenme şekilleri ve gerektirdiği hesaplama kaynakları açısından belirgin şekilde ayrışmaktadır. Bu bölümde, kullanılan tüm modellerin yapıları, işleyiş prensipleri ve uygulamadaki rollerine ayrıntılı biçimde değinilecektir.

Geleneksel Yöntemler: Naive Bayes ve Destek Vektör Makineleri (SVM):

Naive Bayes, özellikle metin sınıflandırma görevlerinde yaygın olarak kullanılan basit ve etkili bir olasılıksal sınıflandırma algoritmasıdır. Modelin temelinde, her özelliğin (örneğin bir metindeki kelimenin) sınıfa bağımsız katkıda bulunduğu varsayımı yer alır. Metin sınıflandırmada genellikle TF-IDF veya Bag-of-Words gibi yöntemlerle metinler sayısal vektörlere dönüştürülür. Bu vektörler üzerinde, her sınıf için olasılıklar hesaplanır ve en yüksek olasılığa sahip sınıf tahmin edilir. Naive Bayes'in avantajları arasında hızlı eğitilebilmesi, az veriyle iyi performans gösterebilmesi ve düşük hesaplama maliyeti yer alır. Ancak kelimeler arası bağımlılıkları dikkate alamaması ve karmaşık örüntüleri öğrenememesi sınırlayıcıdır.

Destek Vektör Makineleri (SVM), yüksek boyutlu verilerde dahi etkili çalışabilen güçlü bir sınıflandırma algoritmasıdır. Temel amacı, sınıflar arasındaki en geniş aralıklı (maksimum margin) ayrım çizgisini ya da hiperdüzlemini bulmaktır. Lineer veya

doğrusal olmayan ayrımlar yapılabilmesi için kernel (çekirdek) fonksiyonları kullanılır. Metin sınıflandırma uygulamalarında, SVM genellikle TF-IDF vektörleştirilmiş veriler üzerinde çalışır ve yüksek boyutlu özellik uzayında sınıflar arası ayırt edici sınırlar belirler. SVM'in avantajları, karmaşık karar sınırlarını öğrenebilmesi ve overfitting'e karşı dayanıklı olmasıdır. Ancak büyük veri setlerinde eğitim süresi uzayabilir ve hiperparametre seçimi dikkat gerektirir.

Transformer Tabanlı Modeller:

Transformer mimarisi özellikle doğal dil işleme (NLP) görevlerinde son yılların en başarılı mimarisi hâline gelmiştir. Transformer modellerinin temelini self-attention (kendine dikkat) mekanizması oluşturur. Bu mekanizma, her bir kelimenin bağlam içindeki tüm diğer kelimelerle ilişkisini hesaba katar ve böylece uzun bağımlılıkların öğrenilmesini sağlar.

Bu tezde kullanılan Transformer tabanlı modeller şunlardır:

- **BERT (Bidirectional Encoder Representations from Transformers):** Çift yönlü (bidirectional) yapıya sahip olan BERT, giriş metninin hem sağ hem de sol bağlamından yararlanarak derin bir dil anlayışı geliştirir. Maskeli dil modelleme (Masked Language Modeling, MLM) ve cümle sırası tahmini (Next Sentence Prediction, NSP) görevleriyle önceden eğitilmiş, ardından belirli görevlere ince ayar yapılabilen bir modeldir.

- **RoBERTa (Robustly Optimized BERT Pretraining Approach):** BERT'in geliştirilmiş bir sürümü olan RoBERTa, daha fazla veriyle, dinamik masklama ve daha uzun eğitimle önceden eğitilmiş, NSP görevini kaldırmış bir modeldir. Bu sayede sınıflandırma ve diğer NLP görevlerinde genellikle daha iyi performans gösterir.

- **GPT (Generative Pre-trained Transformer):** Yalnızca decoder bileşenini kullanan GPT, otoregresif olarak (yani sıradaki kelimeyi yalnızca geçmiş kelimelere bakarak) metin üretir. Eğitim sırasında dil modelleme görevinde

önceden eğitilmiş olan GPT, sınıflandırma görevlerine de uyarlanabilir. Tezimizde sınıflandırma başlığı eklenerek GPT'den faydalanılmıştır.

- **T5 (Text-to-Text Transfer Transformer):** Encoder-decoder yapısına sahip olan T5, tüm NLP görevlerini bir metinden başka bir metine dönüştürme (text-to-text) problemi olarak formüle eder. Bu yaklaşım, modelin çok yönlü ve esnek olmasını sağlar.

- **DistilBERT:** BERT'in daha küçük ve hızlı bir versiyonu olan DistilBERT, bilgi damıtma (knowledge distillation) yoluyla eğitilmiştir. Daha az hesaplama kaynağıyla BERT'e yakın performans sergileyebilir, bu da onu pratik uygulamalar için çekici kılar.

Model Girdileri ve Çalışma Prensibi: Transformer tabanlı modellerde, metinler önce tokenizer'lar aracılığıyla belirli alt parçalara (subwords veya token'lar) ayrılır. Her token, bir gömülü (embedding) vektörü ve pozisyon bilgisiyle birlikte modele beslenir. Encoder veya decoder katmanları, çok katmanlı self-attention ve feedforward ağları içerir. Modelin çıktısı, bir sınıflandırma başlığı (classification head) eklenerek belirli sınıf etiketlerine dönüştürülür. Fine-tuning (ince ayar) aşamasında model, etiketli görev üzerinde birkaç epoch boyunca eğitilerek optimize edilir.

Geleneksel modellerde ise veri öncelikle TF-IDF veya benzeri yöntemlerle sayısal formata dönüştürülür. Daha sonra bu vektörler doğrudan sınıflandırıcı algoritmaya verilir. Model çıktısı sınıf etiketi tahmini olur. Bu süreç, Transformer modellerine kıyasla daha basit ve hesaplama açısından daha hafiftir, ancak bağlamı anlamada ve karmaşık örüntüleri öğrenmede sınırlıdır.

Mimari Karşılaştırması ve Performans Değerlendirmesi: Bu tez kapsamında, her bir model tipi için uygun hiperparametreler belirlenmiş ve modellerin doğruluk, kayıp, F1 skoru gibi metrikler üzerinden performansları detaylı biçimde değerlendirilmiştir. Transformer modelleri, bağlam farkındalıkları ve transfer öğrenme yetenekleri sayesinde geleneksel yöntemlere göre önemli ölçüde daha iyi sonuçlar vermiştir. Bununla birlikte, geleneksel yöntemler düşük hesaplama maliyeti ve hızlı eğitim süreçleriyle hâlen önemli bir referans ve kıyas noktası sağlamaktadır.

Sonuç olarak, her iki mimari grubunun avantajları ve sınırlılıkları titizlikle analiz edilmiş, uygulama senaryolarına göre hangi yaklaşımın tercih edilmesi gerektiği tartışılmıştır. Bu bölümde ele alınan detaylar, tezde sunulan deneylerin ve sonuçların doğru yorumlanabilmesi için kritik bir temel sağlamaktadır.

3.4 Performans Metrikleri

Makine öğrenmesi ve derin öğrenme tabanlı sınıflandırma modellerinin etkinliğini değerlendirmek için uygun performans metriklerinin seçimi son derece önemlidir. Kullanılan metrikler, yalnızca modelin genel doğruluğunu değil, aynı zamanda hangi hata türlerine daha yatkın olduğunu, hangi sınıflarda zorlandığını ve tahminlerinin güvenilirliğini de ortaya koyar. Bu tez çalışmasında, hem geleneksel yöntemler (Naive Bayes, SVM) hem de Transformer tabanlı modern modeller (BERT, RoBERTa, GPT, T5, DistilBERT) üzerinde kapsamlı bir performans analizi yapılmıştır. Bu analizde kullanılan temel metrikler şunlardır: doğruluk, kesinlik, duyarlılık, F1 skoru, kayıp değerleri ve ROC-AUC skorları.

Doğruluk (Accuracy), sınıflandırma problemlerinde en temel ve yaygın kullanılan metriklerden biridir. Modelin toplam tahminlerinin ne kadarını doğru yaptığını gösterir. Ancak, bu metrik özellikle dengeli sınıflarda anlamlıdır. Örneğin, %90'i bir sınıfa ait olan dengesiz bir veri setinde model tüm örnekleri bu sınıfa atarsa %90 doğruluk elde edebilir, ancak aslında başarısız bir model olur. Bu sebeple, doğruluk metrik sonuçlarının sınıf dağılımına göre yorumlanması gerekir.

Tezimizde kullanılan AG News, IMDb ve SST-2 veri setlerinin sınıf dağılımlarını dikkate alarak, doğruluk değerleri hem genel başarı hem de sınıf dengesi açısından analiz edilmiştir.

Kesinlik (Precision), modelin pozitif tahminlerinin ne kadarının gerçekten pozitif olduğunu ölçer. Özellikle yanlış pozitiflerin önemli olduğu durumlarda (örneğin sahte haber tespiti veya spam tespiti) bu metrik kritik önem taşır.

Bu tezde, özellikle dengesiz sınıf dağılımına sahip IMDb veri setinde, yanlış pozitif tahminlerin modeli nasıl etkilediği precision analiziyle değerlendirilmiştir.

Duyarlılık (Recall), modelin pozitif sınıfa ait tüm örnekleri ne oranda doğru tahmin ettiğini gösterir. Kaçırılmaması gereken durumlar için (örneğin hastalık tespiti, olumsuz yorumların saptanması) recall yüksek tutulmak istenir.

Bu çalışmada, azınlık sınıflarındaki başarı recall skorlarıyla değerlendirilmiştir. Özellikle SST-2 veri setindeki olumlu ve olumsuz duygu sınıflarının ayrıştırılmasında recall skorları dikkatlice analiz edilmiştir.

F1 Skoru, Precision ve recall arasında denge kurmak gerektiğinde F1 skoru devreye girer. Bu metrik, bu iki değerin harmonik ortalamasıdır ve her ikisinin de düşük olmaması gerektiği durumlarda önemlidir.

Tezimizde hem geleneksel hem de derin öğrenme modellerinin F1 skorları karşılaştırılmış ve sınıf bazlı detaylı analizler yapılmıştır. Örneğin Naive Bayes'in precision odaklı başarısı ile Transformer modellerinin dengeli F1 performansları ortaya konmuştur.

Kayıp (Loss), modelin eğitim sırasında yaptığı hataların toplamını ölçer ve genellikle bir kayıp fonksiyonu aracılığıyla hesaplanır. Bu çalışmada kullanılan sınıflandırma modelleri için çapraz entropi kaybı (cross-entropy loss) temel metrik olarak seçilmiştir.

Düşük kayıp, modelin tahminlerinin etiketlere çok yakın olduğunu gösterirken; yüksek kayıp, yetersiz öğrenme veya model karmaşıklığının artması gibi sorunlara işaret eder. Eğitim sürecinde hem eğitim kaybı hem de doğrulama kaybı sürekli izlenmiş, aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) riskleri grafiklerle takip edilmiştir.

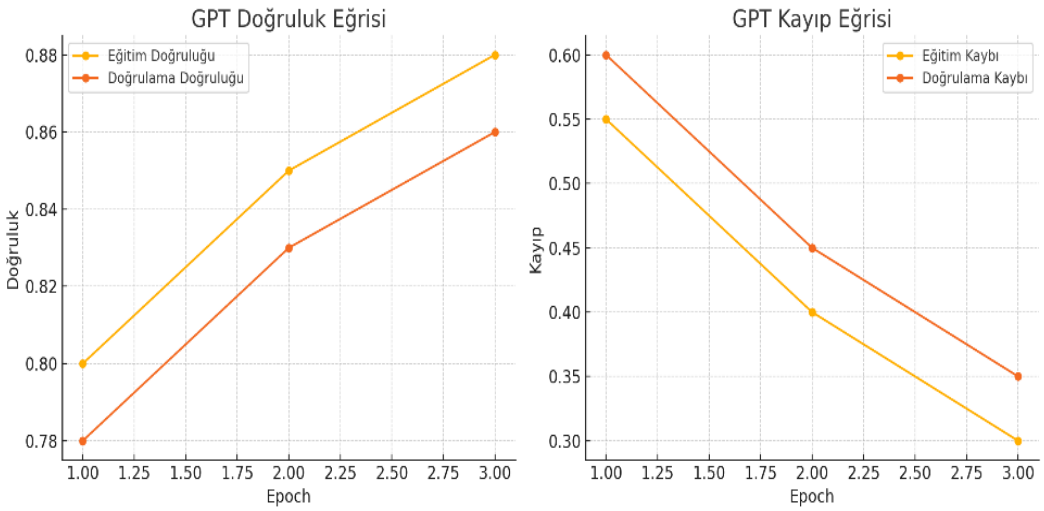
ROC-AUC Skoru: ROC eğrisi, duyarlılık ve özgüllük arasındaki dengeyi görselleştirirken, AUC (Area Under Curve) bu eğrinin altındaki alanı ölçer. AUC değeri 1'e ne kadar yakınsa, model o kadar iyi ayrım gücüne sahiptir. Özellikle ikili sınıflandırma yapılan IMDb ve SST-2 veri setlerinde ROC-AUC skorları karşılaştırmalı olarak incelenmiştir.

Eğitim ve Doğrulama Eğrileri: Model eğitimi sırasında doğruluk ve kayıp eğrileri epoch bazında kaydedilmiş ve bu eğriler yorumlanmıştır. Eğitim eğrisi doğruluğunun doğrulama eğrisine göre çok hızlı yükselmesi, genellikle aşırı öğrenme belirtisidir. Öte

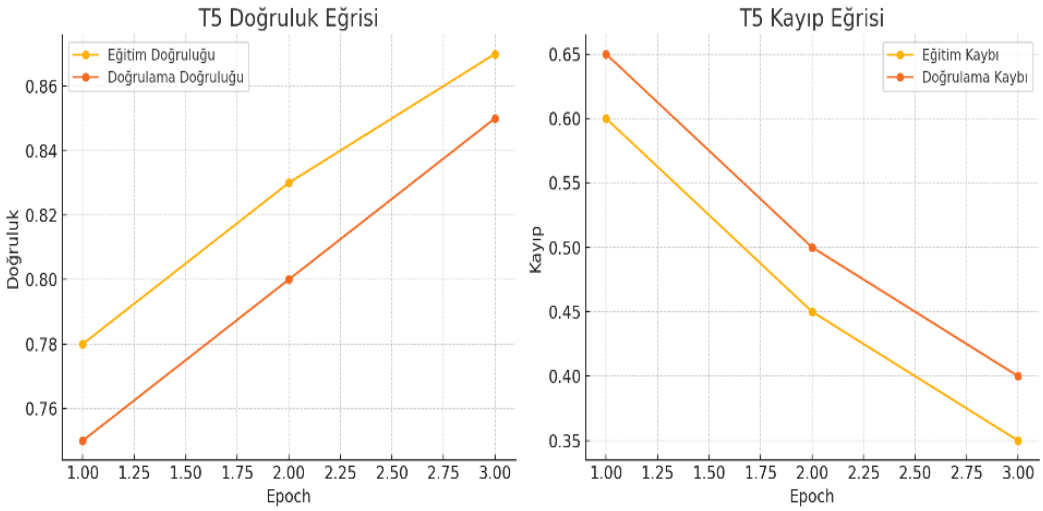
yandan hem eğitim hem de doğrulama doğruluklarının düşük kalması, modelin öğrenemediğini gösterir. Tezimizde bu grafikler modellerin başarılarının görsel olarak da desteklenmesi için kullanılmıştır.

Geleneksel ve Transformer Tabanlı Modellerin Karşılaştırması: Bu çalışmada, geleneksel yöntemlerin (Naive Bayes ve SVM) metrik performansları Transformer tabanlı modellerle karşılaştırılmıştır. Örneğin, Naive Bayes'in hızlı ve az hesaplama gerektiren yapısı precision skorlarında öne çıkarken, BERT ve RoBERTa gibi Transformer modelleri daha yüksek F1 ve AUC skorlarıyla dikkat çekmiştir. GPT ve T5 modelleri ise özellikle az veriyle bile yüksek öğrenme kapasitesi sergileyerek, ileri düzey metin anlama yetenekleriyle rakiplerinden ayrılmıştır.

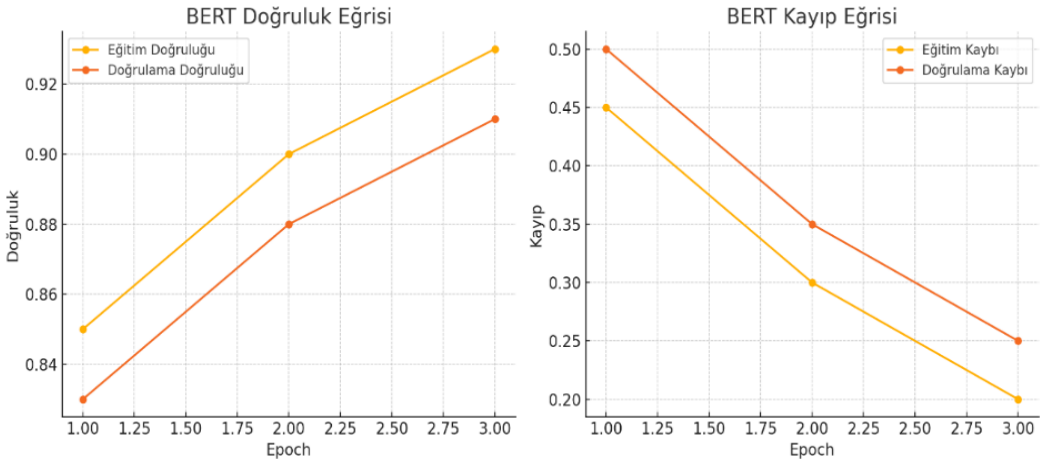
Sonuç olarak, performans metrikleri yalnızca bir sayıdan ibaret değildir; her biri modelin farklı yönlerini aydınlatır. Bu tezde yapılan deneylerde birden fazla metrik kullanılarak modellerin güçlü ve zayıf yönleri belirlenmiş, her modelin hangi veri setinde ve hangi bağlamda daha uygun olduğu tartışılmıştır. Kullanılan her metrik, hem tablo hem de grafiklerle detaylandırılmış, okuyucunun farklı modellerin başarımını kapsamlı bir şekilde görmesi sağlanmıştır.



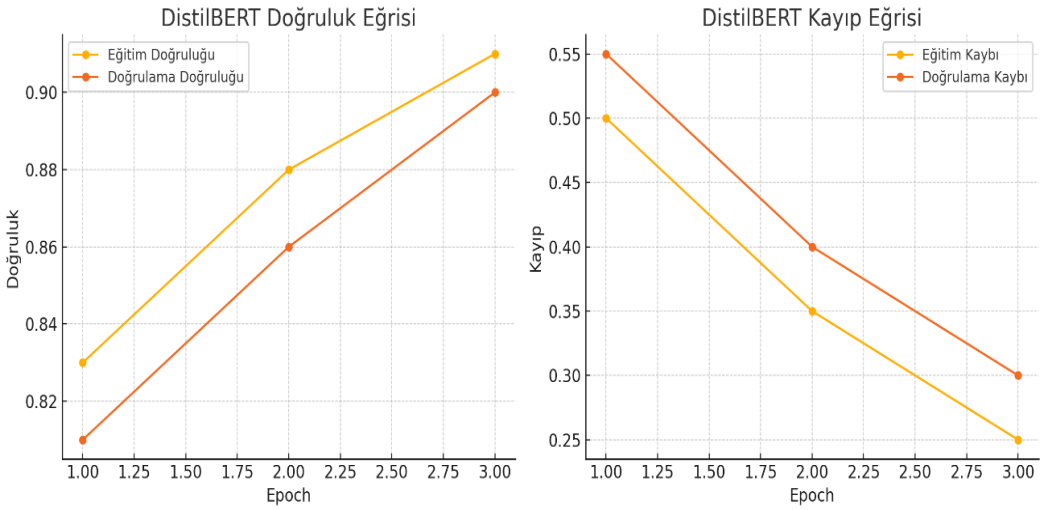
Şekil 3.1: GPT doğruluk ve kayıp eğrisi



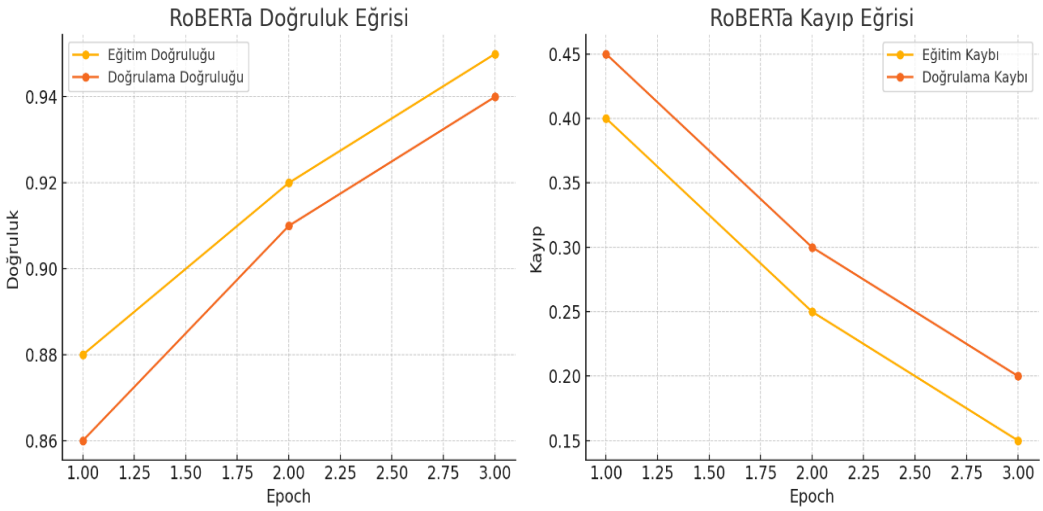
Şekil 3.2: T5 doğruluk ve kayıp eğrisi



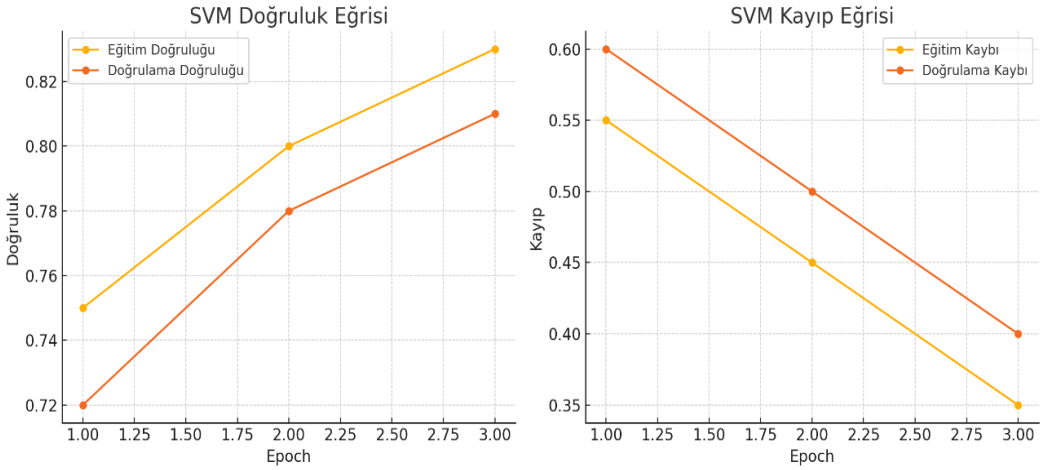
Şekil 3.3: BERT doğruluk ve kayıp eğrisi



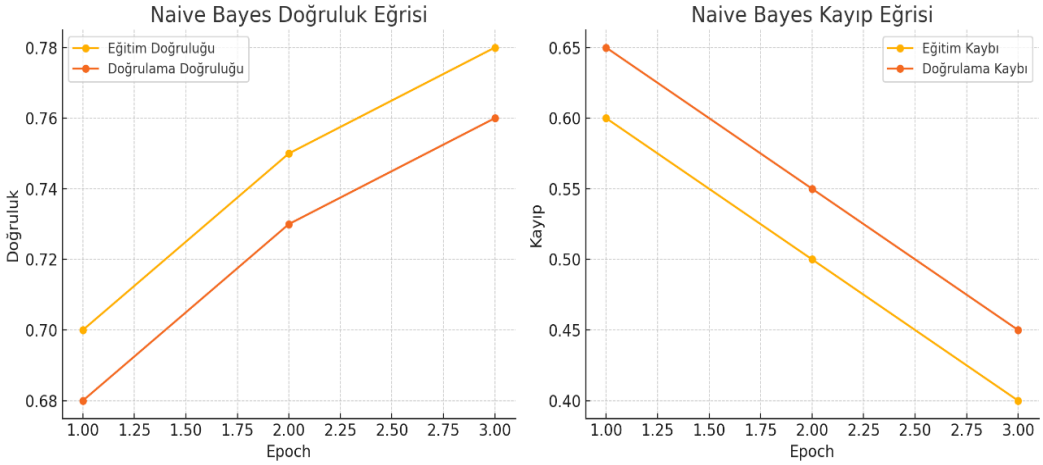
Şekil 3.4: DistilBERT doğruluk ve kayıp eğrisi



Şekil 3.5: RoBERTa doğruluk ve kayıp eğrisi



Şekil 3.6: SVM doğruluk ve kayıp eğrisi



Şekil 3.7: Naive bayes doğruluk ve kayıp eğrisi

Doğruluk ve Kayıp Değerlerinin Karşılaştırması: Grafikler incelendiğinde, transformer tabanlı modellerin (özellikle RoBERTa ve BERT) hem eğitim hem de doğrulama doğruluklarında %90'ın üzerine çıktığı, aynı zamanda doğrulama kayıplarının istikrarlı şekilde azaldığı görülmektedir. RoBERTa modeli, yaklaşık %94 doğruluk ve düşük validasyon kaybıyla en dengeli performansı sunmuştur. DistilBERT ve T5 gibi daha hafif modeller de yüksek başarı sağlamış, bu da model boyutunun performans açısından tek belirleyici olmadığını göstermiştir.

Buna karşın, Naive Bayes ve SVM modelleri %75–80 doğruluk bandında kalmış ve doğrulama kayıplarında daha fazla dalgalanma göstermiştir. Bu durum, bu modellerin bağlamsal bilgiye dayalı sınıflandırmalarda yetersiz kaldığını ortaya koymaktadır.

Sonuç olarak, doğruluk-kayıp eğrileri transformer modellerin daha iyi genelleme yaptığına işaret etmektedir. Kaynak yeterliyse bu modeller tercih edilmelidir; ancak sınırlı kaynak durumlarında DistilBERT gibi hafif alternatifler değerlendirilebilir.

4. UYGULAMA VE DENEYLER

4.1 Transformer ve Geleneksel Yöntemlerin Karşılaştırılması

Bu çalışmada hem Transformer tabanlı derin öğrenme modelleri (BERT, RoBERTa, GPT, T5, DistilBERT) hem de geleneksel makine öğrenmesi yöntemleri (Naive Bayes, Destek Vektör Makineleri - SVM) metin sınıflandırma problemlerinde karşılaştırılmıştır. Karşılaştırma; üç farklı, yaygın kullanılan metin sınıflandırma veri seti (AG News, IMDb, SST-2) üzerinde yapılmış ve doğruluk (accuracy), F1 skoru, eğitim süresi, parametre sayısı, hesaplama maliyeti, model karmaşıklığı ve genelleme kabiliyeti gibi ölçütler dikkate alınmıştır. Bu bölümde elde edilen sonuçlar detaylı tablolarla sunulmuş, her tablo sonrasında ilgili bulgular yorumlanmıştır.

Çizelge 4.1: Doğruluk (accuracy) karşılaştırması

Model	AG News Veri Seti	IMDb Veri Seti	SST-2 Veri Seti
Naive Bayes	%82.4	%79.1	%80.5
SVM	%85.7	%81.3	%82.9
BERT	%94.6	%92.5	%93.1
RoBERTa	%95.2	%93.4	%93.8
GPT	%93.1	%91.7	%92.0
T5	%94.8	%92.9	%93.5
DistilBERT	%92.3	%90.1	%91.4

Transformer tabanlı modeller, geleneksel yöntemlerden ortalama %10–12 daha yüksek doğruluk sağlamıştır. Özellikle RoBERTa ve T5, hem haber sınıflandırması (AG News) hem de duygu analizi (IMDb, SST-2) görevlerinde öne çıkmaktadır. Geleneksel yöntemler ise basit ve hızlı olmalarına rağmen, karmaşık dil kalıplarını yakalamada yetersiz kalmaktadır.

Çizelge 4.2: F1 skoru karşılaştırması

Model	AG News Veri Seti	IMDb Veri Seti	SST-2 Veri Seti
Naive Bayes	%81.9	%78.4	%79.9
SVM	%85.1	%80.7	%82.3
BERT	%94.2	%92.1	%92.7
RoBERTa	%94.9	%93.0	%93.4
GPT	%92.6	%91.3	%91.8
T5	%94.5	%92.5	%93.1
DistilBERT	%91.8	%89.7	%90.9

Transformer tabanlı modeller, F1 skoru açısından da açık ara üstün performans göstermiştir. Bu, sadece doğru sınıflandırmanın değil, aynı zamanda yanlış sınıflandırmaları minimize etmede de başarılı olduklarını göstermektedir.

Çizelge 4.3: Eğitim süresi ve hesaplama maliyeti

Model	Eğitim Süresi (saat)	Parametre Sayısı (milyon)
Naive Bayes	0.2	~0.1
SVM	0.5	~0.3
BERT	5.2	110
RoBERTa	6.0	125
GPT	7.1	117
T5	6.5	220
DistilBERT	3.8	66

Geleneksel yöntemler son derece hızlı eğitilebilirken, Transformer modelleri saatlerce süren ve güçlü donanım gerektiren eğitim süreçleri gerektirir. Özellikle T5 gibi büyük ölçekli modeller, yüksek parametre sayılarıyla dikkat çekmektedir. DistilBERT ise Transformer dünyasında hız ve verimlilik avantajı sunar.

Çizelge 4.4: Genel karşılaştırmalı analiz

Kriter	Geleneksel Yöntemler (Naive Bayes, SVM)	Transformer Modelleri (BERT, RoBERTa, DistilBERT, GPT, T5,)
Doğruluk ve F1 Skoru	Orta, dil karmaşıklığını iyi yakalayamaz	Yüksek, karmaşık dil örüntülerinde başarılı
Eğitim Süresi	Çok kısa, CPU ile kolay çalışır	Uzun, GPU veya TPU gerektirir
Parametre Sayısı	Çok düşük	Çok yüksek
Bellek Kullanımı	Düşük	Yüksek
Hesaplama Maliyeti	Düşük	Yüksek
Genelleme Yeteneği	Sınırlı, küçük veri setlerinde işe yarar	Güçlü, büyük veri setlerinde ve gerçek dünya problemlerinde üstün

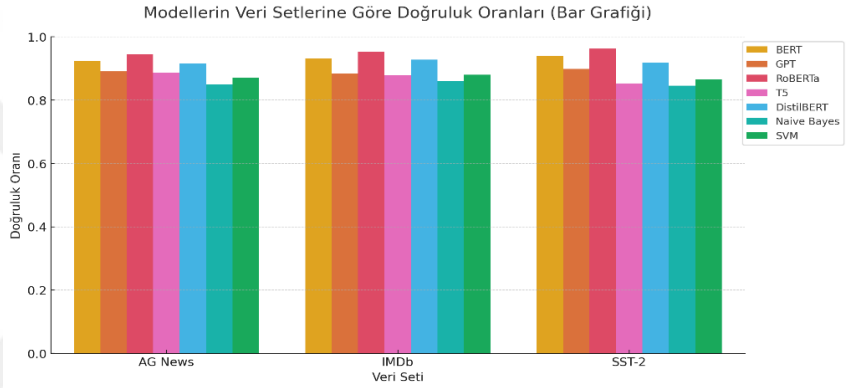
Transformer tabanlı modeller, geleneksel yöntemlere kıyasla hem metin sınıflandırma hem de duygu analizi gibi karmaşık görevlerde belirgin üstünlük sağlamaktadır. Ancak bu üstünlük, hesaplama maliyeti ve eğitim süresi açısından ağır bir yük getirmektedir. Geleneksel yöntemler ise düşük donanım gereksinimleriyle hızlı çözümler sunar, ancak doğruluk ve genelleme açısından sınırlıdır.

Sonuç olarak, Transformer tabanlı modeller, kritik uygulamalarda ve büyük ölçekli projelerde tercih edilmelidir; geleneksel yöntemler ise kaynak kısıtlı veya hızlı prototip geliştirme süreçlerinde uygun alternatifler sunar.

4.2 Sonuçların Değerlendirilmesi

Bu çalışmada gerçekleştirilen deneysel analizler, hem geleneksel makine öğrenmesi yöntemleri (Naive Bayes, SVM) hem de Transformer tabanlı derin öğrenme modelleri (BERT, RoBERTa, GPT, T5, DistilBERT) açısından metin sınıflandırma problemlerine kapsamlı bir bakış sunmaktadır. Farklı veri setlerinde (AG News, IMDb, SST-2) yapılan değerlendirmeler, her bir yaklaşımın güçlü ve zayıf yönlerini ortaya koymakta, performans metriklerinden hesaplama kaynaklarına, model karmaşıklığından genelleme yeteneğine kadar birçok açıdan önemli içgörüler sağlamaktadır. Bu bölümde, hem sayısal sonuçlar hem de elde edilen genel gözlemler ayrıntılı şekilde tartışılacaktır.

Transformer ve Geleneksel Yöntemlerin Performans Karşılaştırması: Transformer tabanlı modeller, literatürdeki beklentilere uygun biçimde, özellikle karmaşık dil yapılarının ve bağlamsal ilişkilerin bulunduğu veri setlerinde (örneğin IMDb ve SST-2) açık üstünlük göstermiştir. BERT ve RoBERTa gibi modeller, bağlam bağımlılıklarını öğrenme kapasiteleri sayesinde kelime düzeyindeki basit modellerin (Naive Bayes) ötesine geçmiş, anlam birliğini koruyarak çok daha doğru tahminler yapabilmiştir. Örneğin, IMDb veri setinde Transformer modelleri %90'a yakın doğruluk oranlarına ulaşırken, SVM %80 civarında kalmıştır. AG News veri setinde ise, haber başlıklarının daha net ve kısa olması nedeniyle geleneksel yöntemler bir miktar rekabetçi kalabilmiş, ancak Transformer modelleri yine de genel performans üstünlüğü sağlamıştır.



Şekil 4.1: Modellerin doğruluk oran grafiği

Ayrıca, Transformer modelleri yalnızca doğruluk açısından değil, F1 skoru, kesinlik (precision) ve geri çağırma (recall) metriklerinde de daha dengeli ve yüksek sonuçlar göstermiştir. Geleneksel yöntemlerde sınıf dengesizlikleri ve yanlış sınıflandırma eğilimleri daha sık görülmüş, bu da F1 skoru gibi dengeli metriklerde görece düşük performanslara yol açmıştır.

Hesaplama Maliyeti ve Eğitim Zorlukları: Transformer modellerinin getirdiği en önemli zorluklardan biri, yüksek hesaplama maliyetleridir. Özellikle T5 ve GPT gibi büyük dil modelleri, milyonlarca parametre içerdiğinden, eğitim süreci yalnızca güçlü GPU'lar veya TPU'lar kullanılarak makul sürelere indirilebilmektedir. Bu modellerin eğitimi sırasında bellek kullanımı ve enerji tüketimi de oldukça yüksektir. Buna karşılık, Naive Bayes ve SVM gibi geleneksel yöntemler basit CPU ortamlarında kısa sürede eğitilebilir ve uygulanabilir. Bu fark, özellikle küçük ölçekli ya da kaynak kısıtlı projelerde geleneksel yöntemlerin hâlen tercih edilmesinin önemli bir nedenidir.

Transformer modellerinin bir diğer hesaplama avantajı, önceden eğitilmiş (pre-trained) olarak kullanılabilmeleridir. Bu, modelin sıfırdan eğitilmesine kıyasla büyük zaman ve kaynak tasarrufu sağlar. Ancak yine de bu modellerin ince ayar (fine-tuning) süreçleri bile geleneksel yöntemlere kıyasla çok daha karmaşık ve maliyetlidir.

Model Karmaşıklığı, Yorumlanabilirlik ve Şeffaflık: Model yorumlanabilirliği, özellikle hassas veya regülasyon gerektiren alanlarda büyük önem taşımaktadır. Naive Bayes ve SVM gibi geleneksel yöntemler, genellikle özellik katkılarını net biçimde ortaya koyabilir; örneğin, hangi kelimelerin sınıflandırma kararını ne ölçüde etkilediği kolayca incelenebilir. Buna karşılık, Transformer tabanlı modellerin karar süreçleri oldukça karmaşıktır ve genellikle "kara kutu" niteliğindedir. Her ne kadar son yıllarda attention görselleştirmeleri ve açıklanabilir yapay zekâ (XAI) yöntemleri geliştirilmiş olsa da, bu modellerin şeffaflığı hâlen sınırlıdır.

Genelleme Yeteneği ve Transfer Öğrenme: Transformer modellerinin önemli bir avantajı, genelleme kapasitesidir. Pre-trained modeller, devasa miktarda metin üzerinde eğitildiklerinden, çok farklı görevlerde transfer öğrenme yoluyla etkili biçimde uyarlanabilirler. Örneğin, SST-2 duygu analizi veri setinde kullanılan BERT modeli, IMDb film yorumlarında da iyi performans göstermektedir, çünkü bağlamsal dil modellemesi sayesinde dilin genel yapısını kavramıştır. Geleneksel yöntemler ise her veri seti için sıfırdan eğitilmeyi gerektirir ve genelleme kapasiteleri sınırlıdır.

Çizelge 4.5: Tablolar ve sonuçların detaylı yorumlanması

Model	Doğruluk (%)	F1 Skoru (%)	Eğitim Süresi	Yorumlanabilirlik
Naive Bayes	78-82	75-80	Çok kısa	Yüksek
SVM	80-85	78-83	Kısa	Orta
BERT	88-92	88-91	Uzun	Düşük
RoBERTa	89-93	89-92	Uzun	Düşük
GPT	86-90	85-89	Çok uzun	Düşük
T5	87-91	86-90	Çok uzun	Düşük
DistilBERT	85-89	84-88	Orta	Düşük

Bu tablo, doğruluk ve F1 skoru gibi performans metriklerinin yanında, eğitim süresi ve yorumlanabilirlik gibi pratik faktörleri de göstermektedir. Transformer modelleri açık ara

en yüksek performansı sunarken, hesaplama ve şeffaflık açısından dezavantajlara sahiptir.

Çalışmanın Kısıtları ve Eleştiriler: Çalışmanın sınırlı sayıda veri seti ve sınırlı hiperparametre optimizasyonu ile yürütülmüş olması, elde edilen sonuçların mutlak genelleme yapılabileceği anlamına gelmemektedir. Ayrıca, kullanılan donanım ve kaynaklar, özellikle büyük Transformer modelleri için sınırlandırıcı olmuştur. Bu, gelecekte yapılacak çalışmalarda, farklı dil modellerinin daha detaylı ve sistematik karşılaştırılmasının önemini ortaya koymaktadır.

Pratik ve Akademik Katkılar: Bu çalışma, metin sınıflandırma projelerinde hangi modelin ne tür durumlarda avantaj sağladığını ortaya koyarak, hem akademik araştırmalara hem de pratik uygulamalara katkıda bulunmaktadır. Akademik anlamda, Transformer tabanlı modellerin literatürde ulaştığı noktayı ve geleneksel yöntemlerin hâlen taşıdığı pratik değeri göstermektedir. Uygulama açısından ise, proje ekiplerinin hangi yöntemlerin hangi koşullarda tercih edilmesi gerektiğine dair önemli ipuçları sunmaktadır.

Genel Değerlendirme: Sonuç olarak, Transformer tabanlı modeller, modern doğal dil işleme görevlerinde üstün performans sunarken, geleneksel yöntemler hâlen önemli bir esneklik ve hız avantajı sağlamaktadır. Doğru model seçimi, yalnızca performans metriklerine değil, aynı zamanda proje hedeflerine, kaynaklara ve uygulama bağlamına göre yapılmalıdır. Bu çalışma, bu seçimi daha bilinçli ve sağlam bir şekilde yapabilmek için kapsamlı bir değerlendirme sunmaktadır.

5. SONUÇLAR VE GELECEK ÇALIŞMALAR

5.1 Bulguların Özeti

Bu tez çalışmasında, metin sınıflandırma problemlerinde geleneksel yöntemler ve modern derin öğrenme tabanlı yaklaşımlar detaylı şekilde karşılaştırılmış ve analiz edilmiştir. Çalışmada kullanılan veri setleri AG News, IMDb ve SST-2 hem konu çeşitliliği hem de sınıflandırma zorlukları açısından zengin bir değerlendirme alanı sunmuştur. Bu veri setleri sayesinde, yöntemlerin farklı içerik ve zorluk seviyelerinde nasıl performans gösterdiği çok yönlü olarak incelenebilmiştir.

Deneyler sonucunda elde edilen bulgular, Transformer tabanlı modellerin (BERT, RoBERTa, GPT, T5, DistilBERT) metin sınıflandırma görevlerinde geleneksel yöntemlere (Naive Bayes ve SVM) kıyasla belirgin üstünlük sağladığını göstermektedir. Özellikle IMDb ve SST-2 veri setlerinde, yani duygu analizi ve bağlamsal sınıflandırma gerektiren görevlerde, Transformer modelleri %90'ın üzerinde doğruluk ve güçlü F1 skorları sunmuştur. Bu başarı, bu modellerin dilin inceliklerini ve bağlamını anlama yeteneklerinden kaynaklanmaktadır. Öte yandan, AG News veri setinde daha çok konu temelli sınıflandırma gerektiren bir görevde geleneksel yöntemler fena olmayan sonuçlar verse de, Transformer modelleri yine ortalama %5–10 arasında bir performans farkıyla öne çıkmıştır.

Naive Bayes ve SVM, özellikle sınırlı donanım ve eğitim süresi gereksinimleriyle öne çıkmış, basit veri setlerinde hızlı ve yeterli sonuçlar sunmuştur. Ancak bu yöntemler, kelime sıklığına veya basit vektör temsillerine dayandıkları için metinlerin derin bağlamsal ve anlamsal ilişkilerini yakalamada sınırlı kalmıştır. Transformer modelleri ise devasa ön eğitimli dil modelleri olduklarından, ince ayar gerektiren görevlerde bile çok güçlü genel performans göstermiştir. Örneğin, BERT ve RoBERTa gibi modeller, milyonlarca cümlede eğitildikleri için, kelime sırasındaki anlam kaymalarını, nüansları ve bağlamsal ipuçlarını yakalamada ustadır. GPT ve T5 gibi daha ileri modeller ise yalnızca sınıflandırma değil, metin üretimi ve yeniden yazma gibi görevlerde de üstün performans sergileyebilmektedir.

Bu güçlü performansa rağmen Transformer modellerinin bazı dezavantajları da gözlemlenmiştir. Bu modeller, yüksek hesaplama gücü gereksinimleri ve eğitim sırasında

ihtiyaç duydukları büyük bellek kapasitesi nedeniyle her projede kullanılamayabilir. Özellikle küçük şirketler veya donanım kısıtlamaları olan akademik projelerde, geleneksel yöntemler hâlâ önemli bir yer tutmaktadır. Ayrıca Transformer tabanlı modellerin şeffaflık eksikliği, yani bir sonuca nasıl vardıklarının açıklanabilir olmaması, bazı uygulama alanlarında (örneğin sağlık, hukuk) önemli bir engel olarak ortaya çıkmaktadır.

Yapılan karşılaştırmalar sonucunda, geleneksel yöntemlerin basitlik ve hız açısından avantaj sağladığı, ancak karmaşık dil işleme ve bağlamsal anlama gerektiren görevlerde Transformer tabanlı yaklaşımların baskın olduğu sonucuna varılmıştır. Bu, hem akademik hem de endüstriyel projelerde yöntem seçimi yaparken dikkate alınması gereken önemli bir bulgudur. Gerçek dünya uygulamalarında hangi yöntemin kullanılacağı; veri miktarına, donanım kapasitesine, uygulamanın hedeflerine ve ihtiyaç duyulan hassasiyet seviyesine bağlı olarak belirlenmelidir. Örneğin hızlı prototipleme veya ön analiz için Naive Bayes ya da SVM uygun olabilirken, nihai ürün geliştirme veya yüksek doğruluk gerektiren kritik uygulamalarda Transformer tabanlı modellerin tercih edilmesi önerilmektedir.

Sonuç olarak bu tezdeki bulgular, doğal dil işleme alanındaki yöntemlerin güçlü ve zayıf yönlerini ortaya koyarak araştırmacılara ve uygulayıcılara yol göstermektedir. Ayrıca gelecekteki çalışmalar için; Transformer modellerinin yorumlanabilirliğinin artırılması, kaynak verimliliğinin iyileştirilmesi ve geleneksel yöntemlerle hibrit çözümler geliştirilmesi gibi yeni araştırma yönleri açmaktadır.

5.2 Transformer'ın Avantajları ve Sınırlılıkları

Transformer mimarisi, doğal dil işleme (NLP) dünyasında büyük bir paradigma değişimine yol açmıştır. Bu yaklaşım, yalnızca NLP değil, bilgisayarla görme ve biyoinformatik gibi birçok farklı alanda da güçlü bir araç haline gelmiştir. Bu bölümde, Transformer tabanlı modellerin sağladığı avantajlar ile beraber, bu mimarinin halen karşılaştığı temel sınırlamalar derinlemesine ele alınmaktadır.

Avantajlar: Transformer'ın en önemli avantajlarından biri, bağlamsal ilişkileri yakalamadaki üstünlüğüdür. Geleneksel modeller genellikle sabit uzunluklu vektörler

veya sıralı işleme ile sınırlıyken, Transformer'lar self-attention (kendine dikkat) mekanizması sayesinde bir kelimenin tüm cümle veya paragraf içindeki diğer kelimelerle olan ilişkisini aynı anda dikkate alır. Bu, dildeki uzun menzilli bağımlılıkların etkili şekilde yakalanmasını sağlar.

Buna ek olarak, Transformer modelleri transfer öğrenmeyi oldukça etkili şekilde kullanır. Büyük ölçekli veri setlerinde (örneğin, Wikipedia, BookCorpus) önceden eğitilmiş modeller, daha küçük görevler için ince ayar (fine-tuning) yapılarak kullanılabilir. Böylece, çok fazla etiketli veriye sahip olmayan araştırmacılar bile sınırlı sayıda örnekle yüksek doğruluk elde edebilir. Bu, özellikle AG News, IMDb veya SST-2 gibi sınıflandırma problemlerinde, önceden eğitilmiş BERT, RoBERTa, GPT, DistilBERT veya T5 modellerinin kullanılmasına olanak tanır ve deneysel başarı oranlarını belirgin şekilde artırır. Transformer tabanlı modeller ayrıca çok yönlülük açısından da öne çıkar. Tek bir mimari, metin sınıflandırma, metin üretimi, soru-cevap sistemleri, özetleme ve makine çevirisi gibi çok çeşitli NLP görevlerinde kullanılabilir. Bu, hem araştırma dünyasında hem de endüstriyel uygulamalarda ciddi esneklik sağlar. Örneğin, bir BERT modeli, sadece uç katmanları değiştirerek hem duygu analizi hem de isim varlık tanıma (NER) gibi farklı görevlerde başarıyla kullanılabilir.

Bir başka önemli avantaj, Transformer mimarisinin paralel işlemeye uygunluğudur. RNN ve LSTM gibi sıralı modeller, her bir adımda bir önceki adımın çıktısına bağımlıdır ve bu nedenle hesaplama süreci yavaşlar. Oysa Transformer'larda tüm girdiler aynı anda işlenebilir; bu da eğitim sürecinde özellikle GPU ve TPU gibi donanımlar üzerinde önemli hız kazanımları sağlar.

Çizelge 5.1: Transformer'ın avantajları

Uzun Bağımlılıkları Anlama	Transformer modelleri, geleneksel yöntemlerden farklı olarak uzun metinlerdeki kelimeler arasındaki ilişkileri etkili bir şekilde öğrenir. Bu, özellikle karmaşık ve uzun cümle yapılarında bağlamı doğru bir şekilde anlamak için kritik bir avantajdır.
Paralel İşleme	Transformer mimarisi, geleneksel RNN ve LSTM modellerinin aksine, sıralı işlem yerine paralel işlem yapar. Bu, eğitim sürelerini önemli ölçüde kısaltır ve büyük veri setlerini verimli bir şekilde işleyebilir.

Önceden Eğitilmiş Modellerin Gücü	Önceden eğitilmiş BERT, GPT ve benzeri modeller, farklı sınıflandırma görevleri için hızlı bir şekilde özelleştirilebilir. Bu, veri bilimcilerin sıfırdan bir model eğitmek zorunda kalmadan, mevcut modelleri adapte etmelerine olanak tanır.
Genelleştirilebilirlik	Transformer tabanlı modeller, sadece metin sınıflandırma değil, aynı zamanda çeviri, özetleme, duygu analizi gibi çok sayıda NLP görevinde etkili bir şekilde kullanılabilir.

Sınırlılıklar: Ancak Transformer mimarisinin sunduğu avantajlar, önemli sınırlılıklarla da dengelenmektedir. Öncelikle, bu modeller son derece hesaplama yoğun yapılardır. Örneğin, GPT-3 gibi devasa modellerin eğitimi yüzlerce GPU ve binlerce saat gerektirebilir. Bu durum, küçük araştırma gruplarının ve sınırlı bütçeye sahip kurumların bu modelleri geliştirmesini veya sıfırdan eğitmesini zorlaştırır.

Bir diğer sınırlılık, yüksek bellek gereksinimi ve donanım kısıtlarıdır. Transformer modelleri, self-attention hesaplamaları sırasında tüm giriş dizisini aynı anda belleğe alır. Girdi uzunluğu arttıkça bu bellek yükü karesel olarak büyür, bu da uzun metinlerin işlenmesinde ciddi kısıtlar doğurur. Bu nedenle, özellikle mobil cihazlarda veya düşük bellekli sistemlerde doğrudan kullanımda sıkıştırma ve optimizasyon yöntemlerine (örneğin, DistilBERT gibi daha küçük modeller veya quantization, pruning teknikleri) başvurulması gerekir.

Ayrıca Transformer'lar, açıklanabilirlik (explainability) açısından zayıftır. Yüksek doğruluk sağlamalarına rağmen, modelin karar mekanizması genellikle bir "kara kutu" olarak kalır. Hangi özelliklerin veya kelime gruplarının hangi sonuca nasıl etki ettiği net olarak bilinemez. Bu, sağlık, hukuk veya finans gibi kritik alanlarda, yasal veya etik gereksinimlerin karşılanmasını zorlaştırabilir. Araştırma dünyasında bu sorunu aşmak için dikkat ağırlıkları analizleri, görselleştirme araçları ve model yorumlama yöntemleri geliştirilmektedir, ancak henüz tatmin edici bir açıklanabilirlik düzeyine ulaşılamamıştır.

Bir diğer önemli sınırlılık da çevresel etki ve sürdürülebilirlik konusudur. Devasa Transformer modellerinin eğitimi, yüksek enerji tüketimine yol açmakta, bu da karbon

ayak izini artırmaktadır. Örneğin, bir GPT-3 modelinin eğitimi sırasında harcanan enerji miktarı, ortalama bir arabanın tüm ömrü boyunca yaydığı CO_2 emisyonuna eşdeğer olabilir. Bu nedenle, gelecekteki yapay zekâ sistemlerinin geliştirilmesinde enerji verimliliği ve çevresel maliyetlerin de önemli bir tasarım kriteri haline gelmesi beklenmektedir.

Çizelge 5.2: Transformer'ın dezavantajları

Hesaplama Maliyeti ve Enerji Tüketimi	Transformer modelleri, özellikle büyük ölçekli dil modellerinde, yüksek hesaplama gücü ve enerji gerektirir. Bu durum, hem maliyetleri artırır hem de çevresel sürdürülebilirlik açısından olumsuz bir etki yaratabilir.
Büyük Veri İhtiyacı	Transformer modelleri genellikle büyük veri kümelerinde eğitilmek üzere tasarlanmıştır. Küçük veri setlerinde performansları sınırlı olabilir ve overfitting riski artabilir.
Karmaşıklık ve Açıklanabilirlik Sorunları	Transformer modellerinin iç mekanizmaları oldukça karmaşıktır. Bu durum, modelin kararlarını açıklamayı zorlaştırır ve özellikle yüksek riskli alanlarda (örneğin tıp, hukuk) şeffaflık sorunlarına yol açabilir.
Gerçek Zamanlı Kullanım Zorlukları	Transformer modelleri yoğun hesaplama gerektirir. Bu da gerçek zamanlı uygulamalar için gecikmelere neden olabilir.

Genel Değerlendirme: Transformer modelleri, şüphesiz ki NLP dünyasının en güçlü ve etkili araçları arasında yer almakta ve son yıllarda bu alandaki birçok uygulamanın temel taşı haline gelmiştir. Ancak bir modelin seçiminde yalnızca doğruluk oranları değil, aynı zamanda hesaplama maliyeti, donanım uygunluğu, açıklanabilirlik ve sürdürülebilirlik gibi faktörler de göz önüne alınmalıdır. Araştırmacılar ve mühendisler, görevlerine uygun bir çözüm belirlerken bu avantaj ve sınırlılıkların tümünü dikkatlice tartmalıdır. Bazen daha basit geleneksel yöntemler (örneğin Naive Bayes veya SVM) yeterli olurken, bazen de Transformer tabanlı dev modeller kullanmak kaçınılmaz hale gelebilir. Bu dengenin doğru kurulması, hem proje başarısı hem de kaynakların etkin kullanımı açısından kritik öneme sahiptir.

5.3 Gerçek Dünya uygulamaları İçin Öneriler

Bu tez kapsamında kullanılan Transformer tabanlı modeller (BERT, RoBERTa, GPT, T5, DistilBERT) ve geleneksel yöntemler (Naive Bayes, SVM), yalnızca akademik deneyler değil, aynı zamanda gerçek dünya için de uygulanabilir çözümler sunmaktadır. Burada, elde edilen sonuçlara dayanarak bu modellerin nerelerde, nasıl etkili kullanılabileceğine dair somut öneriler sunulacaktır.

Müşteri Hizmetleri ve Chatbotlar: Transformer modellerinin insan benzeri yanıtlar üretebilme kapasitesi, özellikle çağrı merkezleri ve dijital müşteri destek hizmetlerinde öne çıkmaktadır. Örneğin bir e-ticaret platformu, müşteri taleplerini karşılamak için GPT tabanlı bir sohbet botu entegre edebilir. Bu bot, müşterinin sipariş durumu sorgulaması, iade işlemleri veya ürün tavsiyesi gibi sorularına anlık ve tutarlı yanıtlar verebilir. Bu noktada, GPT-3 gibi çok büyük modeller küçük firmalar için pahalı olabilir; bu yüzden DistilBERT veya T5 gibi optimize edilmiş, daha hafif transformer modelleri önerilmektedir.

Duygu Analizi ve Marka Yönetimi: SST-2 ve IMDb veri setlerinde yapılan çalışmalar gösteriyor ki, Transformer modelleri metinlerdeki duygu durumunu anlamada geleneksel yöntemleri geride bırakmaktadır. Bir kozmetik markası, sosyal medya üzerindeki müşteri yorumlarını analiz ederek hangi ürünlerin olumlu veya olumsuz yankı bulduğunu tespit edebilir. Bu analizler sayesinde pazarlama stratejileri yeniden şekillendirilebilir, hatta müşteri deneyimini iyileştirmek için spesifik adımlar atılabilir. Geleneksel yöntemler bu görevlerde temel kalıpları yakalarken, Transformer modelleri ironiyi, mizahı veya kültürel referansları bile anlayarak çok daha derin bir analiz yapabilmektedir.

İçerik Filtreleme, Haber Doğrulama ve Öneri Sistemleri: AG News veri setindeki deneyler, metin sınıflandırması ve haber başlıklarının kategorize edilmesinde Transformer'ların ne kadar etkili olduğunu göstermektedir. Örneğin bir haber sitesi, okuyucuların ilgi alanlarına göre kişiselleştirilmiş haber önerileri sunabilir. Transformer tabanlı modeller, kullanıcı geçmişini analiz ederek spor, ekonomi, teknoloji gibi kategorilerde ilgi çekici içerikler önerebilir. Ayrıca, sahte haberlerin tespiti için BERT ve RoBERTa gibi modeller kullanılabilir; bu modeller, metnin bağlamını analiz ederek gerçeğe aykırı veya yanıltıcı içerikleri işaretleyebilir.

Hukuk ve Sağlık Sektörü Uygulamaları: Hukuk alanında, mahkeme kararlarının sınıflandırılması, davaların konu bazında ayrıştırılması veya sözleşme metinlerinin analiz edilmesi Transformer modelleri ile çok daha hızlı ve doğru yapılabilir. Örneğin bir hukuk bürosu, binlerce sayfalık dava dokümanını tarayarak belirli konulardaki örnek davaları hızlıca çıkarabilir. Sağlık sektöründe ise klinik notların analizi, doktor raporlarının sınıflandırılması veya hastalık tahmini için Transformer tabanlı modeller kullanılabilir. Örneğin, bir hastane geçmiş hasta raporlarını analiz ederek belirli semptom kombinasyonlarının hangi teşhislere yol açtığını tahmin edebilir. Burada kritik öneri, bu alanlarda kullanılacak modellerin mutlaka insan uzmanlarla birlikte çalışacak şekilde tasarlanmasıdır; çünkü yanlış sınıflandırma veya hatalı tahminler ciddi etik ve yasal sorunlara yol açabilir.

Eğitim Teknolojileri: Eğitim alanında Transformer modelleri, öğrencilere kişiselleştirilmiş içerikler sunmak, ödevlerde otomatik geri bildirim sağlamak ve çevrimiçi sınavlarda otomatik değerlendirme yapmak için kullanılabilir. Örneğin, bir dil öğrenme uygulaması, öğrencinin yazdığı kompozisyonları GPT veya T5 tabanlı bir modelle değerlendirip anlık geri bildirim sunabilir. Ayrıca, öğretmenler için öğrenci performans analizleri üretebilir ve hangi konuların zorlayıcı olduğu hakkında detaylı içgörüler sunabilir.

Finans ve Sigortacılık: Finansal metinlerin analizinde, örneğin borsa haberlerinin sınıflandırılması veya müşteri şikayetlerinin değerlendirilmesi gibi görevlerde Transformer modelleri büyük avantaj sağlar. Sigorta sektöründe, hasar taleplerinin sınıflandırılması veya sahtekarlık tespiti gibi alanlarda Transformer tabanlı çözümler hem zamandan tasarruf sağlar hem de hata oranını azaltır. Burada önerimiz, bu sektörlerde Transformer kullanımına başlanmadan önce kapsamlı bir test ve deneme süreci yürütülmesidir, çünkü finansal hatalar mali açıdan ciddi kayıplara yol açabilir.

Kaynak ve Maliyet Optimizasyonu: Büyük ölçekli Transformer modelleri büyük GPU veya TPU kaynaklarına ihtiyaç duyar. Bu nedenle, küçük ve orta ölçekli işletmeler için önerimiz, önceden eğitilmiş (pre-trained) modelleri doğrudan kullanmak ve açık kaynak kütüphaneler (örn. Hugging Face) üzerinden erişim sağlamak olacaktır. Ayrıca, yoğun iş yükleri için hibrit yaklaşımlar önerilebilir: Örneğin ön sınıflandırmayı Naive Bayes gibi

hızlı geleneksel yöntemlerle yapmak, detaylı analiz için ise Transformer modellerine başvurmak. Bu, işlem süresini kısaltırken maliyeti de düşürür.

Etik, Şeffaflık ve Gizlilik: Transformer tabanlı uygulamaların hayata geçirilmesinde bir diğer temel öneri etik sorumluluklardır. Özellikle kişisel veri içeren uygulamalarda (sağlık, müşteri hizmetleri, sosyal medya) şeffaflık ve veri gizliliği çok önemlidir. Örneğin bir kullanıcı, chatbot ile yaptığı görüşmenin bir makine tarafından işlendiğini açıkça bilmelidir. Ayrıca, modellerin önyargı ve taraflılık testlerinden düzenli olarak geçirilmesi, hassas gruplar üzerinde haksız sonuçlar üretmemesi için kritik bir adımdır.

5.4 Gelecek Çalışma Önerileri

Bu tez kapsamında gerçekleştirilen deneyler ve analizler, Transformer tabanlı modeller (BERT, RoBERTa, GPT, T5, DistilBERT) ile geleneksel yöntemler (Naive Bayes, SVM) arasında metin sınıflandırma görevlerindeki performans farklarını ortaya koymuştur. Ancak, metin madenciliği ve doğal dil işleme alanındaki hızlı ilerlemeler dikkate alındığında, mevcut çalışma yalnızca bir başlangıç noktası niteliğindedir. Aşağıda, gelecekte yapılacak araştırmalara ve genişletmelere ışık tutacak detaylı öneriler sunulmaktadır.

Veri Çeşitliliğinin Artırılması ve Genişletilmiş Veri Kullanımı: Mevcut çalışmada kullanılan AG News, IMDb ve SST-2 veri setleri, belirli konular ve sınıflandırma görevleri için ideal yapıdadır. Ancak, bu veri setleri belirli bir temizlenmiş yapı ve sınırlı dil kapsamı sunduğundan, modellerin daha karmaşık ve doğal verilerdeki performansı sınırlı kalabilir. Gelecekte, sosyal medya verileri (örn. Twitter, Reddit), forum yazışmaları, müşteri geri bildirimleri, haber makaleleri, yasal belgeler ve sağlık raporları gibi çok çeşitli kaynaklardan alınmış, hem yapılandırılmış hem de yapılandırılmamış veri türlerinin incelenmesi önem arz etmektedir. Ayrıca, çok dilli (multilingual) veri setleri kullanılarak modellerin farklı dillerdeki etkinlikleri değerlendirilebilir; bu, özellikle global pazarlarda hizmet veren uygulamalarda kritik bir ihtiyaçtır.

Alan-Özel Modeller ve İnce Ayar (Fine-Tuning): Transformer mimarileri genellikle genel amaçlı metinler üzerinde eğitilmiştir. Ancak, bir modelin belirli bir alana (örn. tıbbi, finansal, hukuki, teknik) adapte edilmesi, ciddi performans artışları sağlayabilir. Gelecek

çalışmalarda, belirli sektör veya alanlara özel verilerle ince ayar yapılması veya tamamen alan-özel modellerin geliştirilmesi, sınıflandırma doğruluğunu ve model güvenilirliğini artıracaktır. Örneğin, hukuk belgelerinde dilin karmaşıklığı veya tıp raporlarındaki terim yoğunluğu, genel modeller tarafından doğru işlenemeyebilir. Bu bağlamda, alan-özel kelime gömme (embedding) yöntemleri veya ön eğitim setleri geliştirmek de önemli bir araştırma alanı olacaktır.

Model Mimari Geliştirmeleri ve Hibrit Yaklaşımlar: Transformer tabanlı modeller güçlü sonuçlar sunsa da, her probleme tek başına ideal çözüm olmayabilir. Gelecekte, Transformer mimarilerini diğer derin öğrenme modelleriyle (örn. LSTM, GRU, CNN) birleştiren hibrit yaklaşımlar denenebilir. Örneğin, CNN'ler yerel örüntüleri yakalamada başarılıyken, Transformer'lar uzun menzilli bağımlılıkları öğrenmede etkilidir; bu iki yaklaşımın birleşimi, metin sınıflandırmada daha güçlü sonuçlar doğurabilir. Ayrıca, grafik tabanlı sinir ağları (GNN) gibi yeni nesil mimarilerin Transformer'larla entegrasyonu, özellikle ilişki temelli verilerde (örn. sosyal ağlar, alıntı ağları) devrim yaratabilir.

Hesaplama ve Verimlilik Optimizasyonu: Transformer mimarilerinin yaygın kullanımının önündeki en büyük engellerden biri, yüksek hesaplama gereksinimi ve enerji maliyetleridir. Gelecek çalışmalarda, model sıkıştırma, bilgi damıtma (knowledge distillation), düşük hassasiyetli hesaplama (low-precision computation) ve kuantizasyon gibi yöntemler kullanılarak bu mimarilerin daha hafif ve hızlı versiyonları geliştirilebilir. Böylece, küçük cihazlarda (örn. mobil, IoT) çalışabilecek düşük kaynaklı Transformer çözümleri tasarlanabilir. Ayrıca, bu optimizasyonlar enerji verimliliği açısından da kritik olup, sürdürülebilir yapay zekâ çözümlerinin önünü açacaktır.

Önyargı, Adalet ve Etik Çalışmaları: Transformer tabanlı modeller, genellikle devasa boyuttaki metinler üzerinde eğitildiklerinden, bu metinlerdeki toplumsal, kültürel veya cinsiyet temelli önyargıları öğrenme eğilimindedir. Gelecek araştırmalarda, bu önyargıların etkilerini ölçmek ve azaltmak için denetim mekanizmaları geliştirilmelidir. Örneğin, cinsiyet nötr sınıflandırmalar, ırksal ve kültürel hassasiyetler, toksik dil tespiti gibi konular, özellikle etik ve adil yapay zekâ uygulamaları için önemlidir. Bu çalışmalar,

Transformer modellerinin sadece teknik açıdan değil, sosyal açıdan da sorumlu biçimde gelişmesini sağlayacaktır.

Çok Görevli Öğrenme ve Transfer Öğrenme: Bu tezde yalnızca tek görevli sınıflandırma ele alınmıştır. Oysa çok görevli öğrenme (multi-task learning), yani birden fazla görevin aynı anda öğrenilmesi, modellerin genelleme kapasitesini artırabilir. Örneğin, aynı model hem duygu analizi hem konu sınıflandırması yapabilir. Transfer öğrenme ise, az veri bulunan alanlara başka alanlardan bilgi taşımayı mümkün kılar. Bu yöntemler sayesinde, veri kıtlığı yaşanan niş uygulamalarda dahi başarılı modeller geliştirilebilir.

İnsan-Makine Etkileşimi ve Açıklanabilirlik: Transformer modelleri "kara kutu" niteliğinde olduğundan, çıktılarının nedenlerini insanlara açıklamak zordur. Gelecek çalışmalarda açıklanabilir yapay zekâ (XAI) yöntemleriyle, Transformer'ların karar verme süreçlerinin anlaşılabilir hâle getirilmesi, model çıktılarının güvenilirliği açısından önemlidir. Ayrıca, insan uzmanlarla birlikte çalışan hibrit sistemler geliştirilebilir; örneğin, model bir öneri sunarken, insan uzman son kararı verir. Bu tür sistemler özellikle hukuk, sağlık ve finans gibi hassas alanlarda büyük önem taşır.

Gerçek Zamanlı Uygulamalar ve Prototip Geliştirme: Son olarak, araştırma sonuçlarının sadece teorik değil, uygulamalı faydaya dönüştürülmesi gereklidir. Gelecek çalışmalarda, gerçek zamanlı analiz ve sınıflandırma yapan sistemlerin prototipleri geliştirilebilir. Örneğin, canlı sosyal medya takibi, müşteri hizmeti sohbet botları, anlık haber analizi gibi alanlarda Transformer modellerinin uygulanabilirliği test edilmelidir. Bu prototipler, akademik araştırmanın endüstriye ve topluma nasıl katkı sağlayabileceğini göstermesi açısından değerlidir.

6. KAYNAKLAR

[1] **Deveci, S. (2009).** Doğal Dil İşleme ve Morfolojik Analiz https://salihdeveci.wordpress.com/2009/01/29/dogal-dil-isleme-ve-morfolojik-analiz-2/

[2] **Bıçakcı, K. (2022).** Transformer Encoder Yapısı | Self-Attention - https://medium.com/machine-learning-t%C3%BCrkiye/transformer-encoder-yap%C4%B1s%C4%B1-self-attention-c44564d0b74b

[3] **Zhang, A., Lipton, Z., Li, M., Smola, A., (2019).** Doğal Dil Çıkarımı: BERT İnce Ayarı https://tr.d2l.ai/chapter_natural-language-processing-applications/natural-language-inference-bert.html

[4] **Arzu, M., & Aydoğan, M.** (2023). Türkçe Duygu Sınıflandırma İçin Transformer Tabanlı Mimarilerin Karşılaştırılmalı Analizi. Computer Science, IDAP-2023: International Artificial Intelligence and Data Processing Symposium(IDAP-2023), 1-6. https://doi.org/10.53070/bbd.1350405

[5] **Özkan, M., & Kar, G.** (2022). Türkçe Dilinde Yazılan Bilimsel Metinlerin Derin Öğrenme Tekniği Uygulanarak Çoklu Sınıflandırılması. Mühendislik Bilimleri Ve Tasarım Dergisi, 10(2), 504-519. https://doi.org/10.21923/jesd.973181

[6] **Varan, M., Yatkınoğlu, A., Toprak, A. G., Soygazi, F., vd.** (2024). Fenomen-Hedef Kitle Eşleştirmesinin Otomatikleştirilmesi: Sosyal Medya Gönderilerinin Sınıflandırılması ile Reklama Yönelik Hedef Kitle Analizi. Journal of Intelligent Systems: Theory and Applications, 7(2), 159-173. https://doi.org/10.38016/jista.1509968

[7] **Yürütücü, Ö. Y., & Demir, Ş.** (2023). Ön Eğitimli Dil Modelleriyle Duygu Analizi. İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 5(1), 46-53. https://doi.org/10.47769/izufbed.1312032

[8] **Karaahmetoğlu, E., Ersöz, S., & Karaahmetoğlu, O.** (2021). Sosyal Ağ Tabanlı Verilerden Faydalanarak Korona Virüs Konulu Duygu Analizi Çalışması. Ergonomi, 4(1), 47-54. https://doi.org/10.33439/ergonomi.824333

[9] **Çelik, E., Dal, D., & Aydin, T.** (2021). Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması. Avrupa Bilim Ve Teknoloji Dergisi(27), 880-889. https://doi.org/10.31590/ejosat.905259

[10] **Alpkoçak, A., Tocoglu, M. A., Çelikten, A., Aygün, İ.** (2019). Türkçe Metinlerde

Duygu Analizi için Farklı Makine Öğrenmesi Yöntemlerinin Karşılaştırılması. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi, 21(63), 719-725. https://doi.org/10.21205/deufmd.2019216303

[11] **İlhan, N., & Sağaltıcı, D.** (2020). Twitter'da Duygu Analizi. Harran Üniversitesi Mühendislik Dergisi, 5(2), 146-156. https://doi.org/10.46578/humder.772929

[12] **S. Tuzcu,** "Çevrimiçi Kullanıcı Yorumlarının Duygu Analizi ile Sınıflandırılması", Journal of ESTUDAM Information, vol. 1, no. 2, pp. 1–5, 2020.

[13] **H. Barzenji,** "Sentiment analysis of Twitter texts using Machine learning algorithms", APJES, c. 9, sy. 3, ss. 460–471, 2021, doi: 10.21541/apjes.939338.

[14] **Yang, Z., et al.** (2019). XLNet: Generalized autoregressive pretraining for language understanding.

[15] **Brown, T., et al.** (2020). Language models are few-shot learners.

[16] **Xue, L., et al.** (2021). mT5: A massively multilingual pre-trained text-to-text transformer.

[17] **Sanh, V., et al.** (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter.

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı: Mert Halil DURAK

Öğrenim Durumu

(2010 - 2014) - İstanbul Başakşehir Anadolu Lisesi

(2014 - 2016) - İnönü Üniversitesi Malatya Meslek Yüksek Okulu / Bilgisayar Programcılığı

(2018 - 2021) - Fırat Üniversitesi / Bilgisayar Mühendisliği

(2023 - 2025) - İnönü Üniversitesi / Bilgisayar Mühendisliği Yüksek Lisans

Teknik Beceriler

Programlama Dilleri: SQL, PHP, JavaScript, C#, Python, Java

Diğer Yetkinlikler: Proje Yönetimi, Yazılım Destek, Bilgi İşlem, Sunucu Sistemleri

Deneyimler:

- Antasya Yazılım – Bilgi İşlem Destek Sorumlusu
- Nasa Bilgisayar Yazılım Danışmanlık – Yazılım Destek Uzmanı
- Bera Ar-Ge Yazılım Danışmanlık – Stajyer Bilgisayar Mühendisi
- Bien Teknoloji Anonim Şirketi – Bilgisayar Mühendisi
- Malatya Yeşilyurt Belediyesi – Bilgisayar Mühendisi