

T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

METİN ÖZETLEME YÖNTEMLERİNİN KARŞILAŞTIRMALI
İNCELENMESİ VE BASKIN KÜME TABANLI BİR ÇIKARICI YAKLAŞIMIN
GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Said YUNUS

Bilgisayar Mühendisliği Ana Bilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi FATİH OKUMUŞ

OCAK 2025

T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

METİN ÖZETLEME YÖNTEMLERİNİN KARŞILAŞTIRMALI
İNCELENMESİ VE BASKIN KÜME TABANLI BİR ÇIKARICI YAKLAŞIMIN
GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Said YUNUS
(36213619011)

Bilgisayar Mühendisliği Ana Bilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi FATİH OKUMUŞ

OCAK 2025

TEŐEKKÜR VE ÖNSÖZ

Bu tez alıőmasının her aőamasında yardım, öneri, bilgi, tecrübe ve desteklerini esirgemedен beni her konuda yönlendiren danışman hocam Sayın Dr. Öğr. Üyesi Fatih OKUMUŐ'a,

Yaptığım tüm alıőmalarda bana her zaman yol gösteren, destek veren ve değerli katkılarını sunan Sayın Dr. Öğr. Üyesi Cengiz HARK'a,

Ayrıca hayatım boyunca olduđu gibi bu alıőmalarım süresince de benden hiçbir zaman desteklerini esirgemeyen sevgili AİLEME

teőekkür ederim.



ONUR SÖZÜ

Yüksek lisans tezi olarak sunduđum “Metin Özetleme Yöntemlerinin Karşılaştırmalı İncelenmesi ve Baskın Küme Tabanlı Bir Çıkarıcı Yaklaşımın Geliştirilmesi” 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 oluştuđunu belirtir, bunu onurumla doğrularım.

Said YUNUS



İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ	i
ONUR SÖZÜ.....	ii
İÇİNDEKİLER.....	iii
ÇİZELGELER DİZİNİ.....	v
ŞEKİLLER DİZİNİ.....	vi
SEMBOLLER VE KISALTMALAR	vii
ÖZET	viii
ABSTRACT	ix
1. GİRİŞ	1
1.1 Metin Özetlemeye Genel Bakış	2
1.2 Çalışmanın Amacı Ve Kapsamı	3
1.3 Çalışmanın Önemi Ve Literatüre Katkısı	5
1.4 Tezin Organizasyon	6
2. İLİŞKİLİ LİTERATÜR.....	7
2.1 Extractive Metin Özetleme Çalışmaları.....	7
2.2 Abstractive Metin Özetleme Çalışmaları.....	9
2.3 Hibrit Metin Özetleme Çalışmaları.....	11
2.4 Dominating Set Yaklaşımının Özetleme Alanındaki Kullanımı	12
3. YÖNTEMLER VE UYGULAMA SÜREÇLERİ	13
3.1 Extractive Özetleme Yöntemleri	13
3.1.1 TextRank yöntemi	13
3.1.2 LexRank yöntemi	14
3.1.3 LSA yöntemi.....	15
3.1.4 KL-Sum yöntemi	16
3.2 Abstractive Özetleme Modelleri.....	17
3.2.1 Pegasus modeli	18
3.2.2 BART modeli	20
3.2.3 T5 modeli.....	22
3.2.4 LED modeli	23
3.3 Değerlendirme Metrikleri	25
3.3.1 SacreBLEU metriği	26
3.3.2 Meteor metriği	27
3.3.3 BERTScore metriği	28
3.3.4 ROUGE metriği.....	29
3.4 Kullanılan Veri Setleri.....	30
3.5 Extractive Özetleme: Veri İşleme ve Uygulama Süreci	31
3.6 Abstractive Modeller: Kullanım ve Uygulama Süreçleri	33
4. BASKIN KÜME YÖNTEMİ İLE ÖZETLEME: YENİ BİR YAKLAŞIM....	35
4.1 Graf Teorisi.....	35
4.2 Dominating Set Yöntemine Giriş	36
4.3 Dominating Set Çeşitleri.....	37
4.4 Algoritmik Karmaşıklık ve Zorluk	38
4.5 Dominating Set ile Metin Özetleme: Yöntemin Geliştirilmesi	39
4.5.1 Metin ön işleme ve hazırlık	39
4.5.2 Önerilen yaklaşım (DSETS)	40
4.5.3 Graf oluşumu ve dominating set uygulanması	44
4.5.4 Baskın düğüm seçimi ve özetleme sürecindeki örnek kullanımlar	47
5. DENEYSEL SONUÇLAR VE DEĞERLENDİRME	49
5.1 Extractive Yöntemlerin Deneysel Sonuçları.....	49

5.2 Abstractive Modellerin Deneysel Sonuçları	53
5.3 Yöntemlerin Karşılaştırmalı Analizi	57
5.4 Önerilen Yöntemin Diğer Extractive Yöntemlerle Karşılaştırmalı Değerlendirme	64
6. SONUÇLAR VE ÖNERİLER.....	68
6.1 Çalışmanın Genel Bulguları ve Yorumlanması	68
6.2 Çalışmanın Sınırlılıkları.....	69
6.3 Gelecek Çalışmalar için Öneriler.....	69
KAYNAKLAR	71
ÖZGEÇMİŞ	76



ÇİZELGELER DİZİNİ

Çizelge 3.1 : BBC News Summary veri seti özellikleri.....	31
Çizelge 3.2 : PubMed Summarization veri seti özellikleri.	31
Çizelge 3.3 : Kullanılan modellerin detayları.	34
Çizelge 5.1 : BBC News veri seti üzerinde extractive yöntemlerin sonuçları.	50
Çizelge 5.2 : PubMed veri seti üzerinde extractive yöntemlerin sonuçları.....	52
Çizelge 5.3 : BBC News veri seti üzerinde abstractive modellerin sonuçları.....	54
Çizelge 5.4 : PubMed veri seti üzerinde abstractive modellerin sonuçları.	56
Çizelge 5.5 : BBC News veri setinde 50 kelimelik özetler için ROUGE sonuçları. .	65
Çizelge 5.6 : BBC News veri setinde 100 kelimelik özetler için ROUGE sonuçları.	66
Çizelge 5.7 : BBC News veri setinde 150 kelimelik özetler için ROUGE sonuçları.	67



ŞEKİLLER DİZİNİ

Şekil 3.1 : Extractive yöntemlerin uygulama süreci.	32
Şekil 3.2 : Abstractive modellerin uygulama süreci.	34
Şekil 4.1 : Königsberg köprü problemi.	35
Şekil 4.2 : Basit bir graf.	36
Şekil 4.3 : Dominating Set çeşitleri.....	38
Şekil 4.4 : Önerilen yöntemin temel yapı taşları.....	41
Şekil 4.5 : Önerilen DSETS yönteminin detaylı süreç akışı.	42
Şekil 4.6 : “REPORT” kelimesi içeren cümlelerin grafi.	47
Şekil 4.7 : farklı baskın kelimeler için oluşturulan alt graflar.....	48
Şekil 5.1 : BBC ve PubMed veri setlerindeki SacreBLEU sonuçları.	57
Şekil 5.2 : BBC ve PubMed veri setlerindeki Meteor sonuçları.	58
Şekil 5.3 : BBC ve PubMed veri setlerindeki BertScore sonuçları.....	59
Şekil 5.4 : BBC ve PubMed veri setlerindeki Rouge sonuçları (50-kelimelik).....	61
Şekil 5.5 : BBC ve PubMed veri setlerindeki Rouge sonuçları (100-kelimelik).....	62
Şekil 5.6 : BBC ve PubMed veri setlerindeki Rouge sonuçları (150-kelimelik).....	64
Şekil 5.7 : 50 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması. 65	
Şekil 5.8 : 100 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması.66	
Şekil 5.9 : 150 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması.67	

SEMBOLLER VE KISALTMALAR

ATS	: Automatic Text Summarization
NLP	: Natural Language Processing
LSA	: Latent Semantic Analysis
KL-SUM	: Kullback-Leibler Summarization
GSG	: Graph-based Summarization with Google Similarity
PEGASUS	: Pre-training with Extracted Gap-sentences for Abstractive Summarization
BART	: Bidirectional and Auto-Regressive Transformers
T5	: Text-To-Text Transfer Transformer
LED	: Longformer Encoder-Decoder
CatSum	: Cengiz, Ali, Taner Summarization
KUSH	: Karci, Uçkan, Seyyarer, Hark
ÇA-FGA	: Çok Amaçlı Fidan Gelişim Algoritması
DUC	: Document Understanding Conference
TAC	: Text Analysis Conference
C4	: Colossal Clean Crawled Corpus
GPT	: Generative Pre-trained Transformer
BERT	: Bidirectional Encoder Representations from Transformers
DS	: Dominating Set
DSETS	: Dominating Set Extractive Text Summarization
LCS	: Longest Common Subsequence
METEOR	: Metric for Evaluation of Translation with Explicit Ordering
ROUGE	: Recall-Oriented Understudy for Gisting Evaluation
NLTK	: Natural Language Toolkit
SV	: Similarity Value
TF-IDF	: Term Frequency-Inverse Document Frequency
Word2Vec	: Word to Vector
GPU	: Graphics Processing Unit

ÖZET

Yüksek Lisans Tezi

METİN ÖZETLEME YÖNTEMLERİNİN KARŞILAŞTIRMALI İNCELENMESİ VE BASKIN KÜME TABANLI BİR ÇIKARICI YAKLAŞIMIN GELİŞTİRİLMESİ

SAİD YUNUS

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Ana Bilim Dalı

76 + IX sayfa

2025

Danışman: Dr. Öğr. Üyesi Fatih OKUMUŞ

Bu tez, metin özetleme alanında extractive (çıkarcı) ve abstractive (soyutlayıcı) yöntemlerin incelenmesi ve karşılaştırılması üzerine odaklanmaktadır. Çalışmada, metin özetleme yöntemlerinin performanslarını değerlendirmek amacıyla, farklı yapıya sahip iki İngilizce veri seti kullanılmıştır: haber içerikli BBC News Summary ve bilimsel makale içerikli PubMed Summarization. TextRank, LexRank, LSA ve KL gibi çıkarcı yöntemler ile PEGASUS, BART, T5 ve LED gibi soyutlayıcı modeller üzerinde gerçekleştirilen analizlerde, bu yöntemlerin farklı içerik türlerindeki etkileri çeşitli metrikler kullanılarak karşılaştırılmıştır. Elde edilen sonuçlar, her yöntemin belirli durumlarda farklı avantajlar sunduğunu ortaya koymuştur.

Tezin ikinci aşamasında, Dominating Set (Baskın Küme) tabanlı yeni bir çıkarcı özetleme yöntemi önerilmiş ve bu yöntem detaylı olarak incelenmiştir. Bu yöntem, metinlerdeki gereksiz bilgileri filtrelemeyi ve metnin en anlamlı bölümlerini seçerek daha özlü özetler üretmeyi hedeflemektedir. Önerilen yöntemin performansı ROUGE metrikleriyle değerlendirilmiş ve özellikle haber içerikli veri setlerinde diğer yöntemlere kıyasla üstünlük sağladığı görülmüştür. Yöntemin, metin özetleme sürecinde daha verimli ve anlamlı sonuçlar üretebildiği gözlemlenmiştir.

Çalışma, bilgi yoğun veri setleri üzerinde özetleme yöntemlerinin etkinliğini değerlendirmeyi, farklı yaklaşımlar arasında kapsamlı bir karşılaştırma sunmayı ve Dominating Set tabanlı yeni bir yöntemin metin özetleme alanındaki uygulanabilirliğini araştırmayı hedeflemektedir. Bu bağlamda tez, hem teorik hem de uygulamalı boyutlarıyla literatüre katkı sağlamayı amaçlamaktadır.

Anahtar Kelimeler: Metin Özetleme, Graf, Baskın Küme, Çıkarcı Özetleme, Soyutlayıcı Özetleme, ROUGE Metrikleri.

ABSTRACT

Master Thesis

A COMPARATIVE ANALYSIS OF TEXT SUMMARIZATION METHODS AND THE
DEVELOPMENT OF A DOMINATING SET-BASED EXTRACTIVE APPROACH

Said Yunus

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

76 + IX pages

2025

Supervisor: Asst. Prof. FATİH OKUMUŞ

This thesis focuses on the examination and comparison of extractive and abstractive methods in the field of text summarization. To evaluate the performance of text summarization methods, two English datasets with different structures were employed: the news-oriented BBC News Summary and the scientific article-oriented PubMed Summarization. Analyses were conducted on extractive methods such as TextRank, LexRank, LSA, and KL, as well as abstractive models including PEGASUS, BART, T5, and LED, comparing their effectiveness across various content types using diverse evaluation metrics. The results demonstrated that each method offers distinct advantages under specific circumstances.

In the second phase of the thesis, a novel extractive summarization method based on Dominating Set theory was proposed and thoroughly analyzed. This method aims to filter out redundant information in texts and select the most meaningful segments, producing more concise summaries. The performance of the proposed method was evaluated using ROUGE metrics, and it was observed to outperform other methods, particularly on news-oriented datasets. The method demonstrated its capability to generate more efficient and meaningful results in text summarization tasks.

This study aims to assess the effectiveness of summarization methods on information-rich datasets, provide a comprehensive comparison between different approaches, and explore the applicability of the Dominating Set-based method in the field of text summarization. In this context, the thesis seeks to contribute to the literature both theoretically and practically.

Keywords: Text Summarization, Graph, Dominating Set, Extractive Summarization, Abstractive Summarization, ROUGE Metrics.

1. GİRİŞ

Günümüz bilgi çağında, dijitalleşme ile birlikte bilgiye erişim olanakları büyük oranda artmış ve birçok alanda hızlı veri üretimi sağlanmıştır. Özellikle internetin yaygınlaşmasıyla birlikte bireyler ve kurumlar her gün milyonlarca makale, haber, sosyal medya gönderisi ve blog yazısı gibi çok çeşitli kaynaklardan bilgiye maruz kalmaktadır (Aydin, 2023; Guo ve diğ, 2018; Hark ve diğ, 2017). Bu bilgi bolluğu, bir yandan insanlara geniş bir perspektif sunarken, diğer yandan büyük miktarda veriyle baş etme zorunluluğunu da beraberinde getirmektedir. Ancak, her bilgiye ayrıntılı olarak erişme ya da değerlendirme olanağı sınırlı olduğundan, bu verilerin hızlı ve etkili bir şekilde özü çıkarılmış biçimlerde sunulması önemli bir ihtiyaç haline gelmiştir (Alguliyev ve diğ, 2019).

Bu doğrultuda metin özetleme, kapsamlı bilgi içeren metinlerin daha kısa ve anlamlı bir özet halinde sunulması işlemidir. Bu özetleme işlemi sayesinde, metnin ana fikirleri veya önemli bölümleri hızlıca anlaşılabilir hale gelir. Özellikle karar alma süreçlerinde, bilgiye hızlı erişim sağlamak için metin özetleme son derece kullanışlı bir araçtır. Bu özetleme işlemi, manuel olarak yapılabileceği gibi, gelişen teknoloji ile birlikte otomatik hale de getirilmiştir. Otomatik metin özetleme (Automatic Text Summarization - ATS), belirli algoritmalar ve yapay zeka tabanlı teknikler kullanılarak metinlerin insan müdahalesi olmadan özetlenmesini sağlar (Brown ve diğ, 2020; Hark ve Karcı, 2020).

ATS, özellikle bilgi yoğunluğunun yüksek olduğu alanlarda öne çıkmaktadır. Örneğin, bir araştırmacının binlerce makaleyi incelemesi gerektiğinde, ATS algoritmaları bu makalelerin önemli noktalarını sunarak bilgi edinme sürecini hızlandırır. Benzer şekilde, haber siteleri ya da sosyal medya platformları ATS teknolojilerini kullanarak kullanıcılarına daha özlu içerikler sunabilmektedir. Bu nedenle, ATS sistemleri, hem bireysel kullanıcıların bilgiye erişimlerini kolaylaştırmakta hem de büyük veri (big data) ile baş etme sürecinde kurumlara önemli avantajlar sağlamaktadır (Alguliyev ve diğ, 2015; Liu ve Lapata, 2019a).

Büyük veri çağında ATS'nin önemini artıran bir diğer unsur, bilgiye duyulan güvenin ve doğruluğun sağlanması gerekliliğidir. Manuel özetleme süreci, her ne kadar güvenilir sonuçlar sunsa da, zaman alıcı ve maliyetli bir süreçtir. ATS sistemleri, daha hızlı ve tutarlı

sonular retirken, veri analistleri, arařtırmacılar ve dięer bilgi kullanıcıları iin byk bir zaman tasarrufu saęlar (Rautray ve Balabantaray, 2017; Vaswani ve dię, 2017). Bunun yanı sıra, ATS algoritmaları geliřtirilirken hedeflenen doęruluk ve ierik tutarlılıęı sayesinde, zetlenen bilginin orijinal metinle uyumlu olması saęlanmakta ve kullanıcının gvenini kazanmaktadır.

1.1 Metin zetlemeye Genel Bakıř

Metin zetleme, doęal dil iřleme (Natural Language Processing - NLP) alanında nemli bir arařtırma alanıdır ve byk miktarda bilginin zetlenmesini saęlayarak bilgiye hızlı eriřimi mmkn kılar. zetleme teknikleri, bir metnin ana fikirlerini ve temel bilgilerini kısa ve anlamlı bir biimde sunmayı hedefler (Guo ve dię, 2018; Hark ve Karcı, 2020). Geliřen dijital dnya, bilgiye hızla ulařma ihtiyacını artırmıř ve bu doęrultuda metin zetleme, veri analizi ve bilgi ynetimi alanlarında yaygın olarak kullanılmaya bařlanmıřtır. Metin zetlemenin NLP dnyasında ne ıkmasının nedenlerinden biri, byk veri aęında karar alma srelerini hızlandırma ve bilgi yoęunluęunu ynetme yeteneęidir (Alguliyev ve dię, 2015; Manakul ve Gales, 2021).

Metin zetleme, farklı kategorilere ve trlere ayrılmaktadır. zetleme teknikleri, zetleme srecinde izlenen yaklařıma ve oluřturulan zetin doęasına gre sınıflandırılabilir. İlk olarak, zetleme iřlemi genellikle tek belge zetleme ve oklu belge zetleme olarak ikiye ayrılır. Tek belge zetlemede, yalnızca tek bir kaynak metin zetlenir ve zet metni oluřturulurken bu kaynaktaki bilgilerden yararlanır (Nenkova ve McKeown, 2011). oklu belge zetlemede ise aynı konuyla ilgili birden fazla kaynaktan elde edilen bilgilerin bir araya getirilerek btncl bir zet sunulması amalanır. Bu tr zetleme, zellikle haber analizi ve bilimsel makaleler gibi geniř bilgi kaynaęına sahip alanlarda kullanılır (Alguliyev ve dię, 2019; Radev ve dię, 2002).

Metin zetleme trleri ayrıca extractive (ıkarımsal) ve abstractive (soyutlayıcı) yaklařımlar olarak iki ana kategoriye ayrılır. Extractive zetlemede, kaynak metindeki cmlelerin nem sırası belirlenerek en anlamlı olanlar seilir ve zetleme iřlemi gerekleřtirilir. Bu tr zetleme, genellikle cmlelerin sırasını deęiřtirmeden orijinal metinden doęrudan alıntılar yaparak anlamlı bir zet oluřturmayı hedefler (Hark ve dię, 2017; Mani ve Maybury, 1999). Abstractive zetleme ise daha ileri bir sre izler; kaynak metin analiz edilir ve ana fikirler, daha doęal bir dil yapısında yeniden ifade edilerek sunulur. Bu tr zetleme, dil modelleme

ve derin öğrenme yöntemlerinin kullanımını gerektirir ve genellikle daha anlaşılır ve özgün bir özet sunar (Alguliyev ve diğ, 2015; Rush ve diğ, 2015).

Son dönemde, bu iki yöntemi birleştiren hibrit özetleme yöntemleri geliştirilmiştir. Hibrit özetleme, öncelikle extractive bir yaklaşım kullanarak önemli cümleleri seçer, ardından bu cümleleri abstractive bir şekilde yeniden ifade eder. Bu sayede, hem hızlı bir özetleme sağlanır hem de özeti doğal bir yapıda sunulması amaçlanır (Alguliyev ve diğ, 2015; Liu ve Lapata, 2019a). Hibrit özetleme, çoklu belge özetleme ve bilgi yoğun alanlarda etkin sonuçlar vermektedir.

Metin özetleme teknikleri, ayrıca amaca göre de sınıflandırılabilir. Alana özgü özetleme (domain-specific summarization) ve sorgu tabanlı özetleme (query-based summarization) gibi türler, kullanıcının özetleme gereksinimine göre uyarlanmıştır. Alana özgü özetleme, belirli bir konu veya sektör için optimize edilmiştir; örneğin, tıbbi makalelerin özetlenmesi gibi (Mishra ve Mahapatra, 2019). Sorgu tabanlı özetlemede ise özet, belirli bir kullanıcı sorgusuna göre oluşturulur ve metnin yalnızca ilgili bölümleri öne çıkarılır (Alguliyev ve diğ, 2015; Otterbacher ve diğ, 2002). Bu tür özetleme, özellikle arama motorları ve bilgi erişim sistemlerinde kullanılarak kullanıcıların ihtiyaçlarına uygun bilgiye daha hızlı erişmelerini sağlar.

Metin özetleme alanındaki bu çeşitlilik, farklı kullanıcı ihtiyaçlarına yönelik olarak özetleme işlemlerinin özelleştirilmesini mümkün kılar. Özellikle büyük veri analitiği ve doğal dil işleme tekniklerinin gelişimiyle birlikte, metin özetleme hem hız hem de doğruluk açısından büyük ilerleme kaydetmiştir. Bu süreçte, derin öğrenme ve makine öğrenimi tabanlı modellerin de kullanılması, özetlerin daha doğal ve anlaşılır bir yapıda sunulmasına katkı sağlamaktadır (Hark ve diğ, 2017; Rautray ve Balabantaray, 2017; Vaswani ve diğ, 2017). Metin özetlemenin bu tür sınıflandırmaları, kullanıcıların bilgiye hızlı erişimini kolaylaştırmakta ve bilgi yoğunluklu alanlarda verimliliği artırmaktadır.

1.2 Çalışmanın Amacı Ve Kapsamı

Bu çalışmanın temel amacı, farklı metin özetleme tekniklerini karşılaştırarak metin özetleme alanında verimliliği artırmak ve geliştirilmiş Dominating Set (Baskın Küme) tabanlı yeni bir extractive özetleme yönteminin etkinliğini incelemektir. Metin özetleme, günümüzde büyük veri ve bilgi yoğunluğu gibi sorunlara çözüm sunma potansiyeli taşıyan önemli bir araçtır (Guo ve diğ, 2018; Pavan ve Pelillo, 2007). Bu tez kapsamında, özellikle extractive ve

abstractive özetleme yöntemlerinin performanslarını değerlendirmek ve geliştirilen yöntemlerin farklı alanlarda uygulanabilirliğini test etmek amaçlanmaktadır. Bu bağlamda, genel bir abstractive model ile özetleme yapılmakta ve bu model, farklı bir veri setinde fine-tuning¹ yapılmadan kullanılarak geleneksel extractive özetleme yöntemleriyle karşılaştırılmaktadır (Liu ve Lapata, 2019a; Manakul ve Gales, 2021). Graf (Çizge) tabanlı yöntemle yapılan özetlemede, gereksiz kelimeleri elenmiş anlamlı bir özet oluşturulması hedeflenmekte, bu yeni yöntem popüler extractive yöntemlerle kıyaslanarak sonuçları değerlendirilmektedir. Bu yaklaşımla, çalışmanın, özellikle Dominating set algoritmasının metin özetleme alanına getirdiği yenilik ve performans katkısını vurgulaması beklenmektedir (Mihalcea ve Tarau, 2004; Nenkova ve McKeown, 2011).

Bu tez çalışması, metin özetleme alanında extractive ve abstractive yöntemlerin performanslarını değerlendirmek amacıyla tasarlanmıştır. Çalışmada, iki temel İngilizce veri seti olan BBC News Summary ve PubMed Summarization veri setleri kullanılmıştır. BBC veri seti, haber içerikli ve çok kategorili metinlerden oluşurken; PubMed veri seti, bilimsel makale içerikli ve farklı kategorilere sahip bir yapı sunmaktadır (Radev ve diğ., 2002). Bu veri setleri, metin özetleme yöntemlerinin farklı içerik türlerinde nasıl performans gösterdiğini anlamak amacıyla seçilmiştir. Extractive yöntemler arasında TextRank, LexRank, LSA ve KL gibi yaygın olarak kullanılan teknikler yer alırken, abstractive özetleme modelleri olarak PEGASUS, BART, T5-LARGE ve LED kullanılmıştır. Ayrıca, çalışmada kullanılan modellerin fine-tuning yapılmadan test edilmesiyle, metin özetleme alanında genel modellerin performansları da değerlendirilmiştir (Liu ve Lapata, 2019a). Tez çalışması, yalnızca belirli veri setlerine odaklanarak ve kullanılan özetleme yöntemlerini sınırlandırarak belirli bir kapsam çerçevesinde sınırlı kalmıştır. Bu kapsam doğrultusunda, tezde sadece orta ölçekli GPU kaynaklarıyla çalışabilen ve yüksek doğruluk oranı sunan modeller tercih edilmiştir. Çalışmanın teorik ve uygulamalı boyutu, özetleme yöntemlerinin büyük veri analitiğinde daha verimli ve erişilebilir bir hale gelmesine katkıda bulunmayı hedeflemektedir.

¹ Fine-tuning, önceden eğitilmiş bir modelin belirli bir görev veya veri kümesi için özelleştirilmesi sürecidir. Bu yöntem, modelin genel bilgi tabanını korurken, hedef görevde daha iyi performans göstermesini sağlar.

1.3 Çalışmanın Önemi Ve Literatüre Katkısı

Bu tez çalışması, metin özetleme alanında yapılan önceki araştırmalara çeşitli açılardan katkı sağlamayı hedeflemektedir. Literatürdeki birçok çalışma, metin özetleme yöntemlerini yalnızca extractive veya yalnızca abstractive tekniklerle sınırlandırmakta, bu da bazı durumlarda sınırlı bir bakış açısına neden olmaktadır (Guo ve diğ, 2018; Liu ve Lapata, 2019b). Bu çalışmada ise hem extractive hem de abstractive özetleme yöntemleri farklı veri setleri üzerinde karşılaştırmalı olarak ele alınmıştır. Bu sayede, farklı türdeki veri setleri için hangi yaklaşımın daha etkili olduğu konusunda değerli bulgular elde edilmiştir.

Ayrıca, çalışmanın bir diğer önemli katkısı, farklı türde veri setlerinde (Haber içerikli BBC ve bilimsel makale içerikli PubMed) fine-tuning yapılmamış genel bir abstractive modelin extractive yöntemlerle karşılaştırılmasıdır. Literatürde, abstractive modeller genellikle belirli bir veri seti üzerinde fine-tuning edilerek test edilir, ancak bu yaklaşım, modellerin genel performansı hakkında sınırlı bilgi sunmaktadır (Manakul ve Gales, 2021; J. Zhang ve diğ, 2020). Bu tezde ise, fine-tuning yapılmadan kullanılan modellerin, farklı veri türlerinde nasıl performans gösterdiği değerlendirilmiş ve bu sayede genel modellerin potansiyel performansları hakkında önemli bulgular elde edilmiştir. Bu bulgular, özellikle kaynak sınırlı ortamlarda özelleştirilmiş eğitim olmadan abstractive özetleme yapılması gereken durumlar için pratik bilgiler sunmaktadır.

Bu tezde geliştirilen Dominating Set tabanlı graf yöntemi, metin özetlemede özgün bir çözüm sunmaktadır. Graf tabanlı yöntemler genellikle TextRank ve LexRank gibi yaygın algoritmalarla sınırlı kalırken (Erkan ve Radev, 2004; Mihalcea ve Tarau, 2004), bu tez çalışmasında önerilen Dominating Set tabanlı yöntem, cümle seçimi sürecinde gereksiz bilgileri filtreleyerek daha özlü ve anlamlı özetler sunmayı amaçlamaktadır (Nenkova ve McKeown, 2011; Pavan ve Pelillo, 2007). Bu yeni yöntem, özellikle bilgi yoğun ve çok kategorili veri setleri için özetlemenin doğruluğunu artırmakta, metin içeriğinin daha iyi temsil edilmesini sağlamaktadır. Dominating Set tabanlı yaklaşımın, metin özetleme alanında graf tabanlı algoritmalara yenilikçi bir alternatif sunarak literatüre katkı sağlama potansiyeline sahip olduğu düşünülmektedir.

Bu yönleriyle çalışma, metin özetleme alanında graf tabanlı tekniklere yeni bir yaklaşım sunarak ve farklı özetleme yöntemlerini bağımsız bir şekilde karşılaştırarak literatüre katkı sağlamayı hedeflemektedir. Ayrıca, veri setlerinin İngilizce metinlerden oluşması, bu

yöntemlerin farklı dillerde uygulanabilirliği ve adaptasyon süreci için literatürde yapılacak gelecek çalışmalar için bir temel oluşturmaktadır.

1.4 Tezin Organizasyon

Bu tez, metin özetleme yöntemlerinin karşılaştırılması ve Dominating Set tabanlı yeni bir extractive özetleme yaklaşımının geliştirilmesi üzerine yapılan çalışmaları kapsamaktadır. Yapılan tez çalışması, aşağıda özetlenen altı ana bölümden oluşmaktadır.

Birinci bölüm, metin özetleme alanına genel bir bakış sunar ve çalışmanın amacı, kapsamı ve literatüre katkılarını açıklar. Ayrıca, bu bölüm, tezin yapısını tanıtarak okuyucunun her bölümde hangi konuların ele alındığını anlamasına yardımcı olur.

İkinci bölümde, metin özetleme alanında yapılan çalışmalar incelenir. Extractive, abstractive ve hibrit özetleme yöntemleri detaylandırılır ve Dominating Set yaklaşımının metin özetleme alanında nasıl kullanıldığına dair literatürdeki önemli çalışmalar özetlenir. Bu sayede, çalışmanın temelini oluşturan literatürün geniş bir çerçevede anlaşılması sağlanır.

Üçüncü bölüm, çalışmada kullanılan yöntemleri ve değerlendirme sürecini kapsar. Extractive ve abstractive özetleme yöntemleri, kullanılan veri setleri, metin ön işleme adımları ve değerlendirme metrikleri ayrıntılı bir şekilde ele alınır.

Dördüncü bölümde, yeni geliştirilen Dominating Set tabanlı extractive özetleme yöntemi detaylandırılır. Öncelikle, graf teorisi ve düğümler arası benzerlik hesaplama yöntemleri tanıtılır. Daha sonra, Dominating Set yönteminin özetleme için nasıl geliştirildiği ve uygulandığı adım adım açıklanır.

Beşinci bölümde, çalışmada kullanılan yöntemlerin performans analizleri sunulmakta ve farklı veri setleri üzerinde gerçekleştirilen deneylerin sonuçları değerlendirilmektedir. Ayrıca, önerilen yaklaşım ile diğer yöntemlerin karşılaştırmalı analizine yer verilmektedir.

Altıncı ve son bölümde ise çalışmanın genel sonuçları tartışılır ve elde edilen bulgular yorumlanır. Ayrıca, çalışmanın sınırlılıkları ele alınarak gelecekte yapılabilecek araştırmalar için öneriler ve iyileştirme alanları sunulur. Bu bölüm, çalışmanın alanına katkılarını özetler ve araştırmanın daha ileriye taşınması için yönlendirici bilgiler sağlar.

2. İLİŞKİLİ LİTERATÜR

Metin özetleme alanında gerçekleştirilen çalışmalar, extractive, abstractive ve hibrit yöntemler gibi farklı yaklaşımlar temelinde şekillenmiştir. Ayrıca, Dominating set gibi yenilikçi teknikler, metin özetleme performansını artırmak için araştırmalara entegre edilmiştir. Bu bölümde, metin özetleme ile ilgili literatür kapsamlı bir şekilde ele alınacak ve ilgili yöntemlerin detayları sunulacaktır.

2.1 Extractive Metin Özetleme Çalışmaları

Metin özetleme çalışmaları, ilk olarak basit istatistiksel tekniklerin kullanıldığı yöntemlerle başlamıştır. Edmundson'un çalışması, cümlelerin sıklık, pozisyon ve kelime ağırlıkları gibi özelliklerini temel alarak önemli cümleleri seçen ilk modellerden birini ortaya koymuştur (Edmundson, 1969). Bu yaklaşım, metin özetlemenin temel taşlarını oluşturarak sonraki çalışmalar için güçlü bir temel oluşturmuştur. Benzer şekilde, Luhn, kelime frekansını kullanarak metindeki önemli kelimeleri belirleyen bir model geliştirerek metin özetleme alanına istatistiksel tabanlı bir bakış açısı kazandırmıştır (Luhn, 1958; Mani ve Maybury, 1999).

Graf tabanlı yaklaşımların bir diğer örneği olan TextRank algoritması, Mihalcea ve Tarau (2004) tarafından geliştirilmiştir ve Google'ın PageRank¹ algoritmasına benzer bir biçimde çalışır. Bu yöntem, cümleleri birer düğüm olarak kabul eder ve düğümler arasındaki kenarlar, cümleler arasındaki benzerliği temsil eder. TextRank, özellikle DUC-2002² gibi çok belgeli özetleme veri setlerinde test edilmiştir ve metindeki en önemli cümleleri seçme başarısı göstermiştir. Mihalcea ve Tarau, bu çalışmada cümle benzerliğini kosinüs benzerliği kullanarak hesaplamış ve cümleler arasındaki bağlantıların yoğunluğunu ölçmüştür. Bu model, bir metni özetlemek için düğümler arasındaki geçiş olasılıklarına göre önem sırası belirleyerek en kritik cümleleri öne çıkarır.

Graf tabanlı özetleme yöntemlerinden PageRank de çeşitli veri setlerinde, örneğin Opinions gibi inceleme metinlerinden oluşan veri setlerinde test edilmiştir. Günes ve Aker (2013) tarihli çalışmasında PageRank algoritması, kullanıcı yorumlarının özeti için kullanılarak

¹ PageRank, Google'ın web sayfası sıralama algoritması, bağlantı analizi ile düğümlerin önemini ölçer.

² DUC-2002 (Document Understanding Conference), 2002 yılında düzenlenen ve tek belgeli özetleme sistemlerinin değerlendirilmesine odaklanan bir doğal dil işleme konferansıdır. Bu konferans kapsamında kullanılan veri seti, özetleme sistemleri için bir benchmark olarak geniş çapta kabul görmüştür.

kullanıcı odaklı içeriklerin daha özlü bir biçimde sunulması amaçlanmıştır. PageRank'in cümle seçme konusunda özellikle inceleme metinlerinde yüksek performans göstermesi, kullanıcı yorumlarının özetlenmesinde önemli avantajlar sunmaktadır (Xiong ve Ji, 2016).

Benzer bir diğer yöntem olan LexRank, Erkan ve Radev (2004) tarafından geliştirilmiştir ve TextRank ile benzer bir mantıkla çalışır; ancak LexRank, düğüm merkeziliği kavramına dayalı olarak daha kapsamlı bir ilişki analizi yapar. Bu yöntemde, cümlelerin merkeziliği hesaplanarak, daha fazla bağlantıya sahip cümleler özet için öncelikli olarak seçilir. LexRank algoritması, özellikle DUC-2004 ve TAC-2008¹ veri setlerinde test edilmiş olup, farklı kaynaklardan gelen haberlerin özeti için başarılı sonuçlar vermiştir. LexRank, ayrıca çoklu belge özetleme görevlerinde önemli cümlelerin belirlenmesinde kullanılmıştır ve metinlerin genel yapısını iyi temsil eden özetler oluşturmuştur.

Graf tabanlı özetleme yöntemlerine ek olarak LSA (Latent Semantic Analysis), Gong ve Liu (2001) tarafından önerilmiştir. Bu yöntem, metindeki semantik ilişkileri analiz eder ve önemli cümleleri seçmek için tekil değer ayrışımı SVD² kullanarak metindeki konuları ortaya çıkarır (Deerwester ve diğ., 1990). PubMed ve Scientific American gibi bilimsel makale veri setleri üzerinde test edilen LSA, özellikle akademik metinlerin özetlenmesinde semantik tutarlılığı koruma yeteneği sayesinde etkili bir performans göstermiştir. Bu yöntem, kelime-cümle ilişkilerini inceleyerek metnin ana temalarını belirlemekte ve semantik yapıya dayalı tutarlı özetler oluşturmakta kullanılmıştır.

Haghighi ve Vanderwende (2009) tarafından geliştirilen KL-SUM yöntemi ise, metin özetleme sürecinde Kullback-Leibler Divergence ölçümünü kullanır ve özet ile orijinal metin arasındaki bilgi farkını minimize etmeyi amaçlar. KL-SUM, özellikle New York Times haber veri setinde denenmiş ve çoklu belge özetleme görevlerinde orijinal metinle yüksek benzerlik oranı sunarak başarılı sonuçlar elde etmiştir. Bu yöntem, özeti bilgi içerik olarak orijinal metne benzer olmasını sağlamayı hedefleyerek farklı uzunluklarda özetler için kullanılabilir bir model sunmaktadır.

Son yıllarda, graf tabanlı özetleme yöntemlerine yönelik farklı yaklaşımlar da geliştirilmiştir. Baydar (2018), genetik algoritmaları kullanarak metin özetlemede üç farklı puanlama yöntemi sunmaktadır. Sabit Puanlı Değerlendirme, Sezgisel Rastgele Puanlama

¹ TAC (Text Analysis Conference), 2008 yılında metin analizi ve özetleme sistemlerinin değerlendirildiği yıllık bir konferanstır.

² SVD (Singular Value Decomposition), matrisleri ayrıştırarak boyut indirgeme ve veri analizi için kullanılan bir lineer cebir yöntemi.

ve Genel Rastgele Puanlama. Bu yöntemler, cümlelerin önem sırasını belirlemede genetik algoritmalarla optimize edilerek özetleme sürecinde daha verimli bir cümle seçimi sağlamaktadır (Spielman ve Teng, 1996).

Bir diğer yenilikçi çalışma, Uçkan (2020) tarafından önerilen bağımsız kümelerle dayalı çizge tabanlı metin özetleme yöntemidir. CatSumm¹ adlı modelde, KUSH² aracı ile normalleştirme işlemi yapıldıktan sonra spektral çizge bölme kullanılarak cümle kümeleri oluşturulur ve merkezilik değeri yüksek olan cümleler seçilerek özet oluşturulur. İkinci model olan Bağımsız Kümeleme Yaklaşımı ise bağımsız kümelerde yer alan düğümler ile ilişkili cümlelerin özetle yer almaması gerektiği varsayımı üzerine kurulmuştur. Bu iki yöntem, metin çizgelerindeki bağımsız kümeleri dikkate alarak yenilikçi bir çıkarımsal yaklaşım sunmaktadır.

Hark (2020) ise entropi ve optimizasyon tabanlı çıkarımsal metin özetleme yöntemleri geliştirmiştir. Bu çalışmada, Karıcı Entropi yöntemi ile metindeki en bilgi içeren cümleler seçilmektedir (Tuğal ve Karıcı, 2019). Shannon Entropi'ye alternatif olarak önerilen bu yöntem, kelime öbeklerinin yerel ve global önemini belirlemeyi kolaylaştırmaktadır. Ayrıca, Çok Amaçlı Fidan Gelişim Algoritması (ÇA-FGA) ile maksimum kapsama ve minimum fazlalığı optimize ederek özetleme kalitesi artırılmaktadır. ÇA-FGA, DUC-2002 ve DUC-2004 veri setlerinde yapılan deneylerle test edilmiştir ve performansı ROUGE metrikleri ile değerlendirilmiştir.

2.2 Abstractive Metin Özetleme Çalışmaları

Abstractive özetleme modelleri, metindeki içeriği doğrudan çıkarmak yerine, metni yeniden ifade ederek özgün bir özet oluşturmayı amaçlar. Bu yöntem, özellikle doğal dil işleme alanındaki derin öğrenme ve dil modelleme gelişmeleriyle önemli bir ilerleme kaydetmiştir. İlk dönem çalışmalarda, kural tabanlı ve şablon tabanlı sistemler kullanılarak belirli bir veri kümesindeki cümle yapıları analiz edilmiş ve belirli kurallar çerçevesinde yeniden ifade edilmiştir (Edmundson, 1969; Reiter ve Dale, 2000). Ancak bu tür sistemler, metnin geniş anlam yapısını veya çeşitli bağlamları yeterince yansıtmada yetersiz kalmıştır.

¹ CatSumm, spektral metin kümeleme ve düğüm merkeziliği ile çoklu doküman özetleyen, üç aşamadan oluşan bir çıkarıcı özetleme modelidir.

² KUSH, Yazım şekilleri farklı, ancak anlamsal olarak benzer kelimeleri işleyerek cümleler arasındaki bağlantıları güçlendiren bir metin işleme aracıdır.

Derin öğrenmenin öne çıkmasıyla birlikte, sıralı-sekans modelleri ve daha sonra Encoder-Decoder¹ modelleri abstractive özetleme için yaygın olarak kullanılmaya başlamıştır. Sutskever ve diğ. (2014) tarafından geliştirilen ilk Encoder-Decoder model, bir dizi kelimeyi başka bir dizi kelimeye çevirme yeteneğiyle doğal dil işleme görevlerinde devrim yaratmış ve özetleme alanında da önemli katkılar sağlamıştır. Bu modelin özetleme çalışmalarına uyarlanmasıyla, metinlerin içerikleri yeniden ifade edilerek özlü ve anlam bütünlüğünü koruyan özetler üretmek mümkün hale gelmiştir (Rush ve diğ, 2015).

Özellikle dikkat mekanizması (attention mechanism) eklentisiyle güçlendirilmiş olan Transformer² tabanlı modeller, abstractive özetleme çalışmalarında çığır açmıştır. Vaswani ve diğ. (2017) sunduğu Transformer modeli, dikkat mekanizmasını kullanarak cümle içindeki her kelimenin diğer kelimelerle ilişkisini incelemiş ve dil modellemede büyük bir esneklik sağlamıştır. Bu modelin uygulamaları arasında, BERT³ (Bidirectional Encoder Representations from Transformers) ve GPT⁴ (Generative Pretrained Transformer) gibi modeller yer almaktadır. Özellikle BERT, metin özetleme için geliştirilmiş ön eğitilmiş modeller arasında öne çıkmış ve Google tarafından geniş ölçekte kullanılmıştır (Devlin ve diğ, 2019). GPT ve sonrasındaki GPT-2, GPT-3 gibi modeller ise, yalnızca özetleme değil, aynı zamanda metin üretiminde ve diyaloga dayalı sistemlerde de yüksek başarı elde etmiştir (Brown ve diğ, 2020).

Abstractive özetleme modellerinin bir diğer önemli ilerlemesi PEGASUS ve T5 (Text-To-Text Transfer Transformer) modelleri olmuştur. Google tarafından geliştirilen PEGASUS modeli, özetleme görevleri için özel olarak tasarlanmış ve büyük ölçekli eğitim veri setleriyle eğitilmiştir. Bu modelde, metindeki bazı cümlelerin maskeleye yoluyla çıkarılması ve ardından eksik cümlelerin tahmin edilmesiyle özetleme performansı artırılmıştır (J. Zhang ve diğ, 2020). T5 ise, dil işleme görevlerini tek bir çerçeveye entegre eden bir model olarak abstractive özetleme çalışmalarına büyük katkı sağlamıştır. Bu modelin “text-to-text” yapısı, metin özetlemenin yanı sıra çeviri ve soru yanıtlama gibi çok sayıda görevi kapsayarak geniş bir uygulama alanı sunmuştur (Raffel ve diğ, 2020).

¹ Encoder-decoder, doğal dil işleme modellerinde kullanılan bir yapıdır. Encoder, girdiyi (örneğin, bir cümleyi) daha soyut bir temsile, yani bir nümerik vektör temsiline dönüştürürken; decoder, bu temsilden hedef çıktıyı (örneğin, bir çeviri veya özet) üretir. Bu yapı, sıklıkla sıralı veri işleme görevlerinde (makine çevirisi, metin özetleme vb.) kullanılır.

² Transformer, doğal dil işleme görevlerinde bağlamı anlamak için kullanılan dikkat mekanizmasına dayalı bir derin öğrenme modeli.

³ BERT, iki yönlü bağlam anlayışıyla metinlerin anlamını öğrenen bir NLP modeli.

⁴ GPT, transformer mimarisine dayanan, doğal dil üretiminde etkili bir derin öğrenme modeli

Daha yakın tarihlerde, BART (Bidirectional and Auto-Regressive Transformer) ve LED (Longformer Encoder-Decoder) gibi modeller, uzun metin özetleme ihtiyacını karşılamada önemli rol oynamıştır. BART modeli, maskeli dil modelleme ve sekans tahminleme görevlerini birleştirerek, metin özetleme gibi dil işleme görevlerinde başarılı sonuçlar elde etmiştir (Lewis ve diğ., 2020). Özellikle haber ve araştırma makaleleri gibi uzun metinlerin özetlenmesinde, BART'in encoder-decoder mimarisi oldukça etkili olmuştur. LED modeli ise, uzun belgelere yönelik olarak Transformer mimarisini optimize eden bir yapı sunarak metin özetlemede geniş metinlerin özetlenmesini mümkün kılmıştır (Beltagy ve diğ., 2020; Cohan ve diğ., 2020).

Abstractive özetlemede önemli olan bir diğer gelişme, transfer öğrenme ve ince ayar (fine-tuning) yöntemleridir. Özellikle GPT-3 gibi büyük modeller, özelleştirilmiş özetleme görevleri için transfer öğrenme yoluyla yeniden eğitilmiş ve daha yüksek başarı oranlarına ulaşmıştır (Brown ve diğ., 2020). Aynı şekilde, BERT ve T5 modelleri de belirli özetleme veri setleriyle ince ayardan geçirilerek çeşitli görevlerde performans artırılmıştır. Bu ince ayar teknikleri, farklı alanlara özgü özetleme uygulamaları için soyutlama yeteneği sağlamakta ve daha genel modellerin belirli görevlere uyum sağlamasına olanak tanımaktadır (Ruder ve diğ., 2019).

2.3 Hibrit Metin Özetleme Çalışmaları

Hibrit özetleme yöntemleri, extractive ve abstractive yaklaşımların en güçlü yönlerini bir araya getirerek özetlemede hem içerik bütünlüğünü hem de dilsel akıcılığı sağlamayı amaçlamaktadır. Bu yöntemlerde genellikle ilk olarak metindeki en önemli cümleler çıkarıcı (extractive) tekniklerle seçilir; ardından, bu cümleler daha doğal ve anlamlı bir dil yapısında yeniden ifade edilmek üzere soyutlayıcı (abstractive) modellere aktarılır (Guo ve diğ., 2018). Cao ve diğ. (2018) sunduğu çalışmada, ilk olarak önemli cümleler çıkarıcı tekniklerle belirlenmiş, ardından bir encoder-decoder modeliyle bu cümleler soyutlayıcı bir tarzda özetlenmiştir. Bu yöntem, özellikle haber özetleme alanında yüksek başarı sağlamış ve ROUGE skorları ile doğrulanmıştır (Narayan ve diğ., 2018). Benzer bir başka hibrit modelde ise Liu ve Lapata (2019a), BERT tabanlı bir encoder kullanarak öncelikle cümleleri çıkarıcı tekniklerle özetlemiş, sonrasında bu cümleleri GPT-2 tabanlı bir decoder ile yeniden ifade ederek özetin akıcılığını ve doğruluğunu artırmıştır (J. Zhang ve diğ., 2020). Hibrit özetleme yöntemleri, bilgi yoğunluğu yüksek belgeler üzerinde yapılan testlerde bilgi kaybını azaltmasıyla dikkat çekmekte ve özellikle bilimsel makale, rapor veya arama motorları gibi

alanlarda etkili sonuçlar sunmaktadır (Manakul ve Gales, 2021). Bu alan, dil modelleme ve içerik çıkarma tekniklerini birleştirerek yenilikçi bir özetleme çözümü sağlamakta ve özetleme arařtırmalarında önemli bir konu olarak öne çıkmaktadır (Gehrmann ve diğ, 2018).

2.4 Dominating Set Yaklaşımının Özetleme Alanındaki Kullanımı

Dominating set (Baskın Küme) yaklaşımı, graf teorisi tabanlı bir teknik olup, metin özetleme alanında önemli bilgi kümelerini belirlemek için kullanılır. Bu yöntem, özellikle çoklu belge özetleme veya bilgi yoğunluğu yüksek metinlerde, metindeki ilişkili bölümleri bir araya getirerek anlamlı kümeler oluşturur. İlk olarak (Pavan ve Pelillo, 2007) tarafından tanıtılan Dominating set teorisi, düğümler arasındaki benzerlikleri ölçerek birbirine bağılı kümeler oluşturmayı amaçlar. Özetleme alanında bu yaklaşımın kullanılması, metindeki en önemli ve birbirine bağılı cümlelerin veya kelime öbeklerinin seçilmesini sağlar.

Metin özetleme için Dominating set, genellikle cümleleri bir graf şeklinde modelleyerek uygulanır; cümleler düğümler olarak ifade edilir ve cümleler arasındaki benzerlikler kenarlarla gösterilir. Kosinüs benzerliği ve Jaccard benzerliği gibi ölçüler kullanılarak, graf üzerinde yoğun bir şekilde bağlantılı düğümler (cümleler) belirlenir ve bu kümelere anlamlı bir özet elde edilir (Coser ve diğ, 2018; Ghosh ve Acharyya, 2011). Örneğin, Moghaddam ve Ester (2011) Dominating set yaklaşımını sosyal medya içeriklerini özetlemek için kullanmış ve yöntemle, çok fazla bilgi içeren verileri düzenleyerek daha özlü bir sunum elde etmiştir.

Bir diğör örnek olarak, Li ve diğ. (2012), Dominating set yaklaşımını haber özetleme için kullanmış ve bu yöntemi çok belgeli özetleme sürecine uygulamıştır. Bu çalışmada baskın kümelerin belirlenmesiyle, birbirini destekleyen veya tamamlayan cümleler seçilmiş ve anlam bütünlüğü sağlayan özlü bir özetleme sağlanmıştır (Li ve diğ, 2012; Van Gerven ve diğ, 2015).

Günümüzde, Dominating set yaklaşımı, derin öğrenme ile entegre edilerek daha da geliştirilmiştir. Zhao ve diğ. (2020), Dominating set teorisini BERT tabanlı bir dil modeli ile birleştirerek, hem derin öğrenme modelinin güçlü yönlerinden yararlanmış hem de Dominating setlerin belirlenmesiyle metindeki önemli bilgilerden oluşan daha nitelikli özetler elde etmiştir (Song ve diğ, 2019).

3. YÖNTEMLER VE UYGULAMA SÜREÇLERİ

3.1 Extractive Özetleme Yöntemleri

Extractive özetleme yöntemleri, metin özetleme sürecinde cümleleri doğrudan seçerek bilgi yoğunluğunu azaltmaya yönelik temel bir yaklaşım sunmaktadır. Bu yöntemler, metindeki ana fikirleri temsil eden cümleleri belirlemek amacıyla geliştirilmiş çeşitli algoritmalara dayanır. Çalışmamızda, TextRank, LexRank, LSA ve KL-SUM gibi yaygın extractive yöntemler Sumy kütüphanesi kullanılarak uygulanmıştır (Belica, t.y.). Sumy, metin özetleme algoritmalarının kolay ve etkili bir şekilde kullanılmasını sağlayan açık kaynaklı bir Python kütüphanesidir. Bu yöntemlerin özetleme performansına katkıları detaylı bir şekilde ele alınacaktır.

3.1.1 TextRank yöntemi

TextRank algoritması, Mihalcea ve Tarau (2004) tarafından Google'ın sayfa sıralama algoritması PageRank'ten esinlenilerek geliştirilmiş, graf tabanlı bir extractive metin özetleme yöntemidir; metin içerisindeki cümleler arasındaki bağlantıları değerlendirerek önemli cümlelerin seçilmesine yönelik bir model sunar ve temel adımları aşağıdaki gibidir. İlk olarak, cümlelerin düğüm olarak temsil edilmesi adımı gerçekleştirilir, bu adımda metindeki her bir cümle graf yapısında bir düğüm olarak kabul edilir ve cümleler arasındaki kenarlar benzerlik ölçütüne göre bağlanır.

Daha sonra, ağırlıklı kenarlarla bağlantı kurma süreci gerçekleştirilir, iki cümle arasında benzerlik varsa bu cümleler arasında bir kenar oluşturulur ve kenar ağırlığı cümlelerin benzerlik derecesini ifade eder. Cümleler arasındaki benzerlik, kelime örtüşmelerine dayalı olarak Denklem (3.1) ile hesaplanır (Mihalcea ve Tarau, 2004). Alternatif olarak, cosine similarity yöntemi kullanılabilir ve bu yöntem, cümlelerin vektör temsillerine dayalı olarak benzerlik hesaplar ve Denklem 3.2 ile tanımlanır (Erkan ve Radev, 2004; Salton ve diğ., 1975).

$$w_{i,j} = \frac{|S_i \cap S_j|}{\log(|S_i|) + \log(|S_j|)} \quad (3.1)$$

Denklem 3.1'de:

- $w_{i,j}$: S_i ve S_j cümleleri arasındaki benzerlik ağırlığını ifade eder.
- $|S_i \cap S_j|$: İki cümlede ortak geçen kelimelerin sayısını ifade eder.

- $|S_i|$ ve $|S_j|$: Sırasıyla S_i ve S_j cümlelerinin kelime sayısını ifade eder.

$$\text{cosine_similarity}(S_i, S_j) = \frac{\vec{S}_i \cdot \vec{S}_j}{\|\vec{S}_i\| \|\vec{S}_j\|} \quad (3.2)$$

Denklem 3.2’de:

- \vec{S}_i ve \vec{S}_j : S_i ve S_j cümlelerinin vektör temsilleridir.
- $\|\vec{S}_i\|$ ve $\|\vec{S}_j\|$: bu vektörlerin normlarını ifade eder.
- $\vec{S}_i \cdot \vec{S}_j$: Vektörlerin skaler çarpımıdır.

Bir sonraki adımda, PageRank algoritmasının uygulanması gerçekleştirilir, bu adımda TextRank algoritması Google’ın PageRank algoritmasından esinlenerek uygulanır (Brin ve Page, 1998). Her bir düğüme (cümleye) başlangıçta rastgele bir değer atanır ve daha sonra bu değerler cümleler arasındaki bağlantıların ağırlıklarına göre iteratif bir şekilde güncellenir. Bir düğüm S_i ’nin TextRank değeri $T(S_i)$, Denklem 3.3 ile hesaplanır.

$$T(S_i) = (1 - d) + d \sum_{S_j \in \text{In}(S_i)} \frac{w_{ji} T(S_j)}{\sum_{S_k \in \text{Out}(S_j)} w_{jk}} \quad (3.3)$$

Denklem 3.3’te:

- $T(S_i)$: S_i -inci cümlelerin (düğümün) TextRank skorunu ifade eder.
- d (Damping Faktörü): Genellikle $d=0.85$ olarak atanır.
- $\text{In}(S_i)$: S_i ’ye gelen bağlantıların (düğümlerin) kümesidir.
- $\text{Out}(S_j)$: S_j ’den çıkan bağlantıların (düğümlerin) kümesidir.
- w_{ji} : S_j ve S_i arasındaki kenarın ağırlığını ifade eder.
- $T(S_j)$: S_j -nin TextRank değerini ifade eder.
- $\sum_{S_k \in \text{Out}(S_j)} w_{jk}$: S_j ’den çıkan tüm kenarların toplam ağırlığını ifade eder.

Son olarak, önemli cümlelerin seçilmesi adımı gerçekleştirilir, iterasyon süreci tamamlandıktan sonra graf üzerindeki düğümler TextRank değerlerine göre sıralanır. En yüksek değerlere sahip düğümler metindeki en önemli cümleleri temsil eder ve bu cümleler, özetleme sürecinde seçilir. Bu yöntem, metindeki anlam bütünlüğünü koruyarak bilgi yoğunluğunu artırmayı hedefler (Mihalcea ve Tarau, 2004).

3.1.2 LexRank yöntemi

LexRank algoritması, Erkan ve Radev (2004) tarafından geliştirilen bir graf tabanlı extractive metin özetleme yöntemidir. Bu yöntem, metindeki cümlelerin önemini belirlemek

için graf teorisini kullanır ve özellikle çoklu belge özetlemesi için tasarlanmıştır. LexRank, metindeki cümleler arasındaki benzerlikleri değerlendirerek, bu benzerliklere dayalı bir merkezilik ölçütü hesaplar ve özetleme sürecinde bu merkezilik değerlerini kullanır.

İlk olarak, cümlelerin düğüm olarak temsil edilmesi adımı gerçekleştirilir. Bu adımda, metindeki her bir cümle bir grafın düğümü olarak ele alınır. Düğümler arasındaki bağlantılar, cümleler arasındaki benzerlikler temel alınarak oluşturulur.

Bir sonraki adımda, eşik değeri ile ağırlıklı bir graf oluşturma süreci gerçekleştirilir. LexRank, düğümler arasındaki bağlantıları belirlemek için bir eşik değeri kullanır. Eğer iki cümle arasındaki benzerlik, belirlenen eşik değerini aşarsa, bu cümleler arasında bir bağlantı oluşturulur. Bu işlem, grafın seyrek bir yapı kazanmasını sağlar ve işlem hızını artırır. Cümleler arasındaki benzerlik cosine formülü ile hesaplanır (bkz. Denklem 3.2).

Bir sonraki aşamada, graf üzerinde düğüm merkeziliği hesaplanır. LexRank, bu hesaplama için eigenvector centrality yöntemini kullanır. Bu yöntem, bir düğümün önemini, kendisine bağlı düğümlerin önemine dayalı olarak belirler. Merkezi düğümler, özetleme sürecinde daha fazla önem taşır (Newman, 2010). Eigenvector centrality, iteratif bir yöntemle hesaplanır ve Denklem 3.4 ile tanımlanır.

$$C(S_i) = \sum_{S_j \in N(S_i)} \frac{w_{ij}}{\sum_{S_k \in N(S_j)} w_{jk}} C(S_j) \quad (3.4)$$

Denklem 3.4'te:

- $C(S_i)$: S_i -inci cümlenin (düğümün) önem skorunu ifade eder.
- $N(S_i)$: S_i 'nin komşu düğümlerinin kümesidir
- w_{ji} : S_j ve S_i arasındaki kenarın ağırlığını ifade eder.
- $C(S_j)$: S_j -nin önem skorunu ifade eder.
- $\sum_{S_k \in N(S_j)} w_{jk}$: S_j 'den çıkan tüm kenarların toplam ağırlığını ifade eder.

Son olarak, önemli cümlelerin seçilmesi süreci gerçekleştirilir. Merkezilik değerleri hesaplandıktan sonra, en yüksek değerlere sahip düğümler (yani cümleler) seçilir. Bu yöntem, özetin metindeki en merkezi ve bilgilendirici cümleleri içermesini sağlar.

3.1.3 LSA yöntemi

LSA (Gizli Semantik Analiz), Gong ve Liu (2001) tarafından özetleme amacıyla önerilen ve metinlerdeki gizli semantik ilişkileri keşfetmeye dayanan bir extractive özetleme

yöntemidir. Bu yöntem, bir metindeki kelime-cümle matrisini düşük boyutlu bir uzaya indirger ve metin içerisindeki önemli cümleleri seçmek için bu gizli semantik yapıyı kullanır. İlk olarak, metin bir kelime-cümle matrisine dönüştürülür. Bu matris, her bir kelimenin her bir cümledeki sıklığını temsil eder. Örneğin, $A_{i,j}$, i 'inci kelimenin j 'inci cümlede kaç kez geçtiğini ifade eder. Matris genellikle Term Frequency-Inverse Document Frequency (TF-IDF)¹ ile ağırlıklandırılır (Deerwester ve diğ., 1990).

Bir sonraki adımda, oluşturulan kelime-cümle matrisi (A), SVD yöntemiyle ayrıştırılır ve Denklem 3.5 ile hesaplanır. Bu süreçte, matrisin yalnızca en büyük birkaç tekil değeri (Σ) korunarak matris düşük boyutlu bir uzaya indirgenir. Bu işlem, metindeki gürültü veya gereksiz bilgilerin ayıklanmasını sağlar ve gizli semantik yapının ortaya çıkarılmasına olanak tanır (Gong ve Liu, 2001).

$$A = U\Sigma V^T \quad (3.5)$$

Denklem 3.5'te:

- U : kelimelerin gizli semantik boyutlarını temsil eden matris,
- Σ : tekil değerleri içeren köşegen matris,
- V^T : cümlelerin gizli semantik boyutlarını temsil eden matris.

Daha sonra, indirgenmiş uzayda her bir cümle bir vektörle temsil edilir. Bu aşamada, cümlelerin önem derecesi bu vektörlerin büyüklüklerine (normlarına) göre belirlenir. Normu en büyük olan cümleler, metindeki en bilgilendirici cümleler olarak seçilir.

Son olarak, özetleme amacıyla norm değeri en yüksek olan belirli sayıda cümle seçilir. Bu cümleler, metnin genel semantik yapısını en iyi temsil eden cümlelerdir.

3.1.4 KL-Sum yöntemi

KL-Sum, Kullback-Leibler (KL) Divergence ölçümüne dayalı bir extractive metin özetleme yöntemidir. Bu yöntem, özeti orijinal metni mümkün olduğunca iyi temsil etmesini sağlamayı hedefler. KL-Sum, özeti ve orijinal metnin olasılık dağılımları arasındaki farkı minimize ederek en bilgilendirici cümleleri seçer. KL Divergence, iki olasılık dağılımı arasındaki farklılığı ölçmek için kullanılan bir istatistiksel metrik olarak öne çıkar (Haghighi ve Vanderwende, 2009; Kullback ve Leibler, 1951).

¹ TF-IDF, kelimelerin metinlerdeki önemini hesaplayan bir metin madenciliği yöntemi.

Öncelikle, metin, kelime frekanslarına dayalı bir olasılık dağılımı ile temsil edilir. Her bir kelimenin metin içerisindeki frekansı, kelimenin önemine ilişkin bir olasılık değeri sağlar. Bu işlem, hem orijinal metin hem de özet için ayrı ayrı gerçekleştirilir.

Bir sonraki adımda, metindeki her bir cümle, özeti bir parçası olmaya aday olarak değerlendirilir. Başlangıçta, özet boş bir kümeyle temsil edilir ve algoritma iteratif olarak en bilgilendirici cümleleri bu küme içine ekler.

Daha sonra, KL Divergence hesaplamasıdır. KL Divergence, özeti olasılık dağılımı P ile orijinal metnin olasılık dağılımı Q arasındaki farkı ölçmek için kullanılır. Bu hesaplama, ilerleyen bölümlerde açıklanacak olan Denklem 3.6 ile tanımlanmaktadır. Amaç, $D_{KL}(P||Q)$ değerini minimize etmektir. Böylece, özeti kelime dağılımı orijinal metne daha yakın hale gelir.

$$D_{KL}(P||Q) = \sum_{w \in V} P(w) \log \frac{P(w)}{Q(w)} \quad (3.6)$$

Denklem 3.6'da:

- $D_{KL}(P||Q)$: KL Divergence'ın sonucunu ifade eder.
- $P(w)$: özeti kelime dağılımı.
- $Q(w)$: orijinal metnin kelime dağılımı.
- V : kelime dağarcığını ifade eder.
- $\sum_{w \in V}$: tüm kelimeler (veya öğeler) kümesi V üzerinden toplam alınır.
- $\log \frac{P(w)}{Q(w)}$: iki dağılım arasındaki göreceli farkı ifade eder.

Son olarak, en bilgilendirici cümlelerin seçilir. Her bir iterasyonda, mevcut özet kümesine eklendiğinde $D_{KL}(P||Q)$ 'yi en fazla azaltan cümle seçilir. Bu işlem, belirli bir özet uzunluğuna ulaşıncaya kadar tekrarlanır.

3.2 Abstractive Özetleme Modelleri

Abstractive özetleme modelleri, kaynak metinleri analiz ederek bu metinlerin ana fikirlerini daha doğal ve özgün bir dil yapısında yeniden ifade etmeyi amaçlayan ileri seviye dil modellerine dayanmaktadır. Bu yöntemler, özellikle encoder-decoder mimarisi ve transformer tabanlı yapılar kullanılarak, uzun ve karmaşık metinlerden anlamlı özetler üretebilme yeteneğine sahiptir. Encoder, metni düşük boyutlu bir temsil alanına dönüştürürken; decoder, bu temsili kullanarak metin özeti oluşturur. Pegasus, BART, T5-Large ve LED gibi modeller, bu yapıların en gelişmiş örneklerini sunarak

özetleme süreçlerinde doğruluk ve dilsel akıcılığı artırmayı başarmıştır. Transformer mimarisinin sunduğu dikkat mekanizmaları, modellerin metin bağlamını daha iyi anlamasını sağlayarak, metin özetleme görevlerinde yüksek performans elde edilmesine olanak tanımaktadır (Vaswani ve diğ, 2017).

3.2.1 Pegasus modeli

2019 yılında Google tarafından geliştirilen ve metin özetleme görevlerinde çığır açıcı bir performans sunan Pegasus, transformer tabanlı bir modeldir. Abstractive özetleme için optimize edilen bu model, “gap-sentences generation” (GSG) adı verilen yenilikçi bir ön eğitim yöntemi ile öne çıkmaktadır. Bu yöntemde, kaynak metinlerden önemli cümleler çıkarılarak modelin bu boşlukları yeniden tahmin etmesi sağlanır. Bu süreç, Pegasus’un hem bağlamsal anlamayı hem de dil üretimini güçlendiren bir dil modeline dönüşmesine olanak tanır (Vaswani ve diğ, 2017; J. Zhang ve diğ, 2020).

Pegasus, dilin karmaşıklığını anlamak ve özlü özetler oluşturmak için encoder-decoder mimarisini kullanır. Çalışmamızda, Pegasus’un Hugging Face¹ platformundan indirilen “pegasus-cnn_dailymail” versiyonu kullanılmıştır (Hugging Face, t.y.-e). Model, haber tabanlı CNN/DailyMail² veri seti üzerinde önceden eğitilmiş olup, fine-tuning yapılmadan doğrudan uygulanmıştır. Çalışmamızda, Pegasus’un performansı, özellikle metin özetleme görevlerinde extractive yöntemlerle karşılaştırılmıştır.

Pegasus, temelinde transformer mimarisine dayalı bir encoder-decoder yapısı barındırır. Bu yapı, modelin giriş metinlerini anlamlı bir bağlamsal temsil haline getirmesini ve ardından bu temsil üzerinden özet üretmesini sağlar.

Dikkat mekanizması (attention), Pegasus’un metin bağlamını anlamak için çoklu başlık dikkat mekanizmalarını (multi-head attention) kullanmasını sağlar. Bu mekanizma, modelin farklı bağlamları paralel olarak öğrenebilmesine olanak tanır ve Denklem 3.7 ile tanımlanmıştır.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3.7)$$

Denklem 3.7’de:

¹ Hugging Face, doğal dil işleme (NLP) ve diğer yapay zeka uygulamaları için modeller, araçlar ve platformlar sunan bir açık kaynak topluluğu ve şirkettir.

² CNN/DailyMail, haber özetleme görevleri için kullanılan, CNN ve Daily Mail’den derlenen büyük bir veri kümesi.

- *softmax* : benzerlikleri olasılık dağılımına dönüştürür.
- Q (Query): Sorgu matrisidir ve dikkatin nereye odaklanması gerektiğini belirler.
- K (Key): anahtar matrisidir ve hangi bilginin sorguya uygun olduğunu belirlemek için kullanılır.
- V (Value): değer matrisidir ve dikkatin odaklandığı yerden alınacak bilginin kendisini içerir.
- d_k : boyutları normalize etmek için kullanılan ölçekleme sabiti.
- QK^T : sorgu ve anahtarlar arasındaki benzerlikleri ölçer.

Encoder-decoder bağlantısı, Pegasus'un giriş metinlerini anlamlı bir bağlamsal temsile dönüştürmesini ve bu temsil üzerinden özet üretmesini sağlayan temel yapıdır. Encoder, giriş metnindeki kelimeler arasındaki bağlamı anlamak için self-attention mekanizmasını kullanırken, decoder hem bu bağlamı hem de önceki çıktıları kullanarak hedef metni üretir. Encoder ve decoder arasında çift yönlü dikkat (cross-attention) bulunmaktadır.

Gap-sentences generation (GSG) tekniği, Pegasus'un ön eğitimi sırasında metinlerden önemli cümlelerin çıkarılmasını ve modelin bu boşlukları yeniden tahmin etmesini sağlar. Bu süreç, modelin önemli bilgileri seçme ve dil üretme kapasitesini güçlendiren bir yapıdır (J. Zhang ve diğ, 2020).

Çoklu başlıklar ve paralellik, Pegasus'un uzun metinlerdeki ilişkileri paralel bir şekilde öğrenebilmesini sağlayan çoklu başlık dikkat mekanizmalarına dayanır. Bu özellik, özellikle metin özetleme gibi büyük bağlamsal veri gerektiren görevlerde kritik bir öneme sahiptir.

Pegasus modeli, metin özetleme görevini belirli adımlarla gerçekleştirir. İlk adım, giriş metninin kodlanmasıdır (encoding). Giriş metni, modelin encoder kısmı tarafından kelime vektörlerine dönüştürülür. Bu süreçte self-attention mekanizmaları kullanılarak metnin bağlamı anlaşılır ve bağlamsal bir temsil oluşturulur (Vaswani ve diğ, 2017).

Bir sonraki adım, gap-sentences generation (GSG) yönteminin uygulanmasıdır. Eğitim sırasında, kaynak metindeki önemli cümleler çıkarılır ve model, bu cümleleri yeniden tahmin ederek bağlamsal anlam oluşturmayı öğrenir. Bu yöntem, özetleme sırasında önemli bilgilerin öne çıkarılmasını sağlar (J. Zhang ve diğ, 2020).

Daha sonra, bağlamsal temsillerin özet üretimi için kullanılması gerçekleştirilir (decoding). Encoder tarafından üretilen bağlamsal temsiller, decoder tarafından analiz edilir ve özet metni oluşturulur. Decoder, önceki çıktıları ve encoder'dan gelen bilgiyi kullanarak her kelimeyi sırasıyla tahmin eder.

Son olarak, son özetin üretilmesi süreci tamamlanır. Pegasus, her adımda ürettiği kelimenin bağlamını değerlendirerek, metnin özlü bir özetini oluşturur. Modelin eğitildiği CNN/DailyMail veri seti, bu süreçte modelin doğruluk ve dil akıcılığına katkı sağlamıştır.

3.2.2 BART modeli

2019 yılında Facebook AI tarafından geliştirilen BART, metin üretimi, çeviri ve özetleme gibi doğal dil işleme (NLP) görevleri için optimize edilmiş bir transformer tabanlı modeldir. Encoder-decoder mimarisine dayanan bu model, özellikle “denoising sequence-to-sequence” (gürültü giderici sıralar arası) ön eğitim stratejisi ile öne çıkar. Bu stratejide, giriş metni üzerinde çeşitli gürültüleme işlemleri gerçekleştirilir ve model, bu bozulan metni orijinal haline döndürmek için eğitilir. Bu yöntem, modelin bağlamı anlama ve yeniden ifade etme kapasitesini artırarak, metin özetleme gibi görevlerde üstün performans sunar (Lewis ve diğ., 2020).

Çalışmamızda, Hugging Face platformundaki “bart-large-cnn” versiyonu kullanılmıştır. Bu model, CNN/DailyMail veri seti üzerinde önceden eğitilmiş olup, haber özeti gibi kısa metinlerde etkili bir performans göstermektedir (Hugging Face, t.y.-a). BART, fine-tuning işlemi yapılmadan, farklı veri setlerinde abstractive özetleme performansını değerlendirmek ve extractive yöntemlerle kıyaslama yapmak amacıyla çalışmamızda uygulanmıştır.

BART, doğal dil işleme görevlerinde giriş metnlerinin bağlamsal temsillerini oluşturmak ve dil üretimini sağlamak için transformer tabanlı bir encoder-decoder mimarisi üzerine inşa edilmiştir. Bu yapı, metinlerin bağlamını anlamak ve yeniden üretmek için güçlü bir temel sunar.

Gürültü giderici ön eğitim (denoising pre-training), BART’ın temel ön eğitim stratejisidir. Bu süreçte, giriş metnine çeşitli bozulmalar uygulanır ve model, bozulan metni yeniden yapılandırmak için eğitilir. Kullanılan gürültüleme yöntemleri arasında rastgele kelimelerin maskelenmesi, cümle sıralarının değiştirilmesi ve metnin belirli bölümlerinin silinmesi gibi işlemler bulunmaktadır. Matematiksel olarak, modelin optimizasyonu Denklem 3.8 ile ifade edilir. Bu kayıp fonksiyonu, modelin giriş metni bağlamını anlamasını ve dilsel olarak doğru çıktılar üretmesini sağlar.

$$L_{reconstruction} = - \sum_{i=1}^N \log P(y_i | y_{<i}, X) \quad (3.8)$$

Denklem 3.8’de:

- $L_{reconstruction}$: yeniden yapılandırma kaybı (reconstruction loss),

- N : çıkış dizisinin (örneğin, bir cümlenin) toplam uzunluğunu ifade eder.
- y_i : hedef özetin i . kelimesi,
- $y_{<i}$: hedef özetin i . kelimesine kadar olan kısmı,
- X : encoder tarafından işlenmiş giriş metni.
- $\log P(y_i|y_{<i}, X)$: modelin y_i 'yi tahmin etme olasılığıdır.

Attention mekanizmaları, BART'ın transformer yapısına uygun olarak hem self-attention hem de cross-attention mekanizmalarını kullanmasını sağlar. Encoder kısmında self-attention, giriş metnindeki kelimeler arasındaki bağlamı anlamaya odaklanırken; decoder kısmında cross-attention, encoder tarafından oluşturulan bağlamsal temsilleri kullanarak hedef metni oluşturur. Dikkat mekanizması matematiksel olarak Denklem 3.7'de tanımlanmıştır.

Encoder-decoder bağlantısı, gürültüye maruz kalmış giriş metninin encoder tarafından bağlamsal bir temsile dönüştürülmesi sürecini ifade eder. Bu temsiller, decoder tarafından analiz edilerek temiz bir metne dönüştürülür. Bu süreç, özetleme görevinde giriş metnindeki en önemli bilgilerin seçilmesini sağlar.

Paralel işlem kapasitesi ve çoklu başlık dikkat, modelin paralel işlem yapmasını sağlayan ve her bir başlığın farklı bağlamsal ilişkileri öğrenmesine olanak tanıyan bir özelliktir. Çoklu başlık dikkat mekanizmaları, özellikle uzun ve karmaşık metinlerde bağlam kaybını önlemek için kritik bir rol oynar.

BART modeli, metin özetleme görevini belirli adımlarla gerçekleştirir. İlk olarak, giriş metninin kodlanması süreci gerçekleştirilir (encoding). Modelin encoder kısmı, giriş metnini bağlamsal temsillere dönüştürmek için self-attention mekanizmasını kullanır. Bu temsiller, metindeki kelimeler arasındaki bağlamı anlamaya odaklanır.

Bir sonraki aşama, gürültü giderme sürecidir (denoising). Eğitim sırasında, metne uygulanan gürültüleme işlemleri ile bozulan metni yeniden oluşturmak için model eğitilir. Bu süreç, özetleme görevinde modelin doğru bilgiyi seçmesini sağlar (Lewis ve diğ., 2020).

Daha sonra, bağlamsal temsillerin özet üretimi için kullanılması gerçekleştirilir (decoding). Encoder tarafından üretilen bağlamsal temsiller, decoder tarafından analiz edilir ve özet metni oluşturulur. Decoder, hem encoder'dan gelen bilgiyi hem de önceki çıktıları dikkate alarak kelimeleri sırasıyla tahmin eder.

Son olarak, anlamlı bir özetin üretilmesi süreci tamamlanır. Decoder, her adımda ürettiği kelimenin bağlamını değerlendirerek, metnin özlü bir özetini oluşturur. “bart-large-cnn”

modeli, özellikle CNN/DailyMail veri setinde haber özetleme görevlerinde yüksek doğruluk sunmaktadır.

3.2.3 T5 modeli

T5 (Text-to-Text Transfer Transformer), Google tarafından geliştirilen ve tüm doğal dil işleme (NLP) görevlerini birleştirici bir “text-to-text” paradigması çerçevesinde ele alan, transformer tabanlı bir modeldir. T5’in temel felsefesi, her görevi bir dil modeli olarak çerçevelemektir. Giriş metni (örneğin, “summarize: Bu bir metindir.”) ve hedef çıktı (örneğin, “Bu bir özet.”) tamamen metin formatında ifade edilir (Raffel ve diğ, 2020).

Çalışmamızda, Hugging Face platformundaki T5-Large modeli kullanılmıştır (Hugging Face, t.y.-i). Bu model, çeşitli NLP görevlerinde kapsamlı bir şekilde önceden eğitilmiş olup, özetleme görevlerinde etkili sonuçlar sunmuştur. Çalışmamızda, herhangi bir fine-tuning yapılmadan, T5-Large modelinin performansı farklı veri setlerinde extractive yöntemlerle kıyaslanmıştır. Bu tercih, modelin genel performansını objektif bir şekilde değerlendirmek amacıyla yapılmıştır.

Text-to-text paradigması, T5’in temel yaklaşımıdır. T5, tüm NLP görevlerini “text-to-text” formatında işler. Giriş metni, model tarafından belirli bir görev etiketiyle birlikte işlenir ve çıktı metni doğrudan bu görev için üretilir. Matematiksel olarak, T5’in hedefi giriş metni x verildiğinde hedef metni $\{y_1, y_2, \dots, y_N\}$ üretmektir. Bu süreç Denklem 3.9 ile ifade edilmiştir.

$$P(y|x) = \prod_{i=1}^N P(y_i|y_{<i}, x) \quad (3.9)$$

Denklem 3.9’da:

- $P(y|x)$: giriş metni x verildiğinde hedef çıktı y ’nin olasılığı.
- N : çıkış dizisinin (örneğin, bir cümlenin) toplam uzunluğunu ifade eder.
- y_i : modelin i . kelimeyi tahmin etmesi.
- $y_{<i}$: modelin i . kelimeye kadar olan kısmı (örneğin, $y_{<3} = \text{“Doğal dil”}$)
- x : Giriş metni.
- $\prod_{i=1}^N P(y_i|y_{<i}, x)$: bu denklem, zincir kuralı (chain rule) kullanılarak olasılıkların çarpımı şeklinde ifade edili

Dikkat mekanizmaları, T5’in transformer mimarisinin self-attention ve cross-attention mekanizmalarını kullanmasına olanak tanır. Encoder kısmında self-attention, giriş

metnindeki kelimeler arasındaki bağlamı öğrenirken; decoder kısmında cross-attention, encoder tarafından oluşturulan bağlamsal temsilleri kullanarak hedef metni üretir. Dikkat mekanizması matematiksel olarak Denklem 3.7’de açıklanmıştır.

Transfer learning ve ön eğitim, T5’in dil anlama ve üretme yeteneklerini geliştiren önemli bir süreçtir. Model, önceden “corrupt text” (bozuk metin) onarma görevleri için eğitilmiştir. Bu süreçte, model metindeki belirli kelimeleri maskeler ve bu kelimeleri tahmin ederek bağlamsal anlama kapasitesini artırır (Raffel ve diğ, 2020).

Parametre paylaşımı ve model verimliliği, T5’in encoder ve decoder yapıları arasında parametre paylaşımını kullanarak hesaplama verimliliğini artırmasını sağlar. Bu yaklaşım, modelin büyük veri setlerinde dahi etkili bir şekilde çalışmasını mümkün kılar.

T5-Large modeli, metin özetleme görevini belirli adımlarla gerçekleştirir. İlk olarak, giriş metninin kodlanması süreci gerçekleştirilir (encoding). Modelin encoder kısmı, giriş metnini bağlamsal temsillere dönüştürmek için self-attention mekanizmasını kullanır. Bu temsiller, metindeki kelimeler arasındaki bağlamı anlamaya odaklanır.

Bir sonraki adımda, giriş metni belirli bir görev etiketiyle işlenir. T5, özetleme görevi için giriş metnini “summarize:” etiketiyle başlatır. Bu görev etiketi, modelin hedef çıktığı doğru bir şekilde yönlendirmesini sağlar.

Daha sonra, bağlamsal temsillerin özet üretimi için kullanılması süreci gerçekleştirilir (decoding). Encoder tarafından üretilen bağlamsal temsiller, decoder tarafından analiz edilir ve özet metni oluşturulur. Decoder, hem encoder’dan gelen bilgiyi hem de önceki çıktıları dikkate alarak kelimeleri sırasıyla tahmin eder.

Son olarak, anlamlı ve dilbilgisel olarak doğru bir özetin üretilmesi süreci tamamlanır. Decoder, her adımda ürettiği kelimenin bağlamını değerlendirerek, metin özetleme görevlerinde modelin üstün performansını ortaya koyar.

3.2.4 LED modeli

LED (Longformer Encoder-Decoder), uzun belgeler üzerinde etkili bir şekilde çalışmak için tasarlanmış, transformer tabanlı bir encoder-decoder modelidir. Model, Longformer’in yerel ve küresel dikkat (local-global attention) mekanizmalarını encoder-decoder yapısına entegre ederek, standart transformer modellerinin uzunluk sınırlamalarını aşar. LED, uzun belgelerin özetlenmesi, belge sınıflandırma ve diğer uzun metinlere dayalı doğal dil işleme (NLP) görevlerinde üstün performans sunar (Beltagy ve diğ, 2020).

Çalışmamızda, Hugging Face platformundaki LED modeli kullanılmıştır (Hugging Face, t.y.-c). Model, herhangi bir fine-tuning işlemi yapılmadan, önceden eğitilmiş haliyle özetleme görevlerinde değerlendirilmiştir. Bu yaklaşım, LED'in genel performansını objektif bir şekilde değerlendirmek ve extractive yöntemlerle kıyaslamak amacıyla tercih edilmiştir.

Local-global attention mekanizması, LED'in uzun metinlerde dikkat hesaplamalarını daha verimli hale getirmek için kullandığı bir yapıdır. Local attention, her kelimenin yalnızca belirli bir pencere içindeki komşu kelimelerle olan ilişkisini dikkate alır. Bu pencere, belirli bir bağlam uzunluğu ile sınırlıdır ve kelime yalnızca bu pencere içindeki kelimelere odaklanır. Bu mekanizma, dikkat hesaplamalarının doğrusal zaman karmaşıklığı ($O(n)$) ile yapılmasını sağlar. Global attention ise bazı önemli kelimeleri veya bölümleri (örneğin, başlıklar veya paragrafların ilk cümleleri gibi) "küresel dikkat noktaları" olarak işaretler. Bu sayede model, geniş bağlamları da anlama kapasitesine sahip olur. Local-global attention mekanizması, hem uzun metinlerin işlenmesini mümkün kılar hem de hesaplama maliyetini düşürür. Uzunluk sınırlamalarının aşılması, LED'in diğer transformer modellerine göre öne çıkan özelliklerinden biridir. Standart transformer modelleri genellikle token¹ uzunluğu açısından sınırlıdır. Örneğin, Pegasus ve BART modelleri 1024 token, T5 modeli ise 512 token destekler (Lewis ve diğ, 2020; Raffel ve diğ, 2020; J. Zhang ve diğ, 2020). Buna karşılık, LED 4000 token ve daha uzun diziler üzerinde etkili bir şekilde çalışabilir (Beltagy ve diğ, 2020). Encoder-decoder yapısı, LED'in giriş metinlerini bağlamsal temsillere dönüştürmesini ve bu temsiller üzerinden hedef metinler üretmesini sağlar. Encoder, hem yerel hem de küresel bağlamı anlamak için local-global attention mekanizmasını kullanır. Decoder ise bu bağlamsal temsilleri analiz ederek hedef metni oluşturur.

Hafıza ve verimlilik optimizasyonu, LED'in uzun dizilerde verimli işlem yapabilmesini sağlayan bir diğer önemli özelliğidir. Local attention hesaplamalarının doğrusal zaman karmaşıklığı ($O(n)$) sayesinde model, uzun belgelerde daha az hafıza ve işlem gücü tüketir. Bu optimizasyon, LED'in uzun metinlerde yüksek verimlilikle çalışmasını mümkün kılar.

LED modeli, uzun metinlerin özetlenmesi görevini belirli adımlarla gerçekleştirir. İlk olarak, giriş metninin kodlanması süreci gerçekleştirilir (encoding). Encoder, giriş metinlerini local-global attention mekanizmasını kullanarak bağlamsal temsillere dönüştürür. Yerel dikkat,

¹ token, bir metni daha küçük, işlenebilir birimlere ayıran ve NLP'de metin analizi için kullanılan temel birimdir.

kısa bağlamları işlerken; küresel dikkat, önemli anahtar kelimeleri ve geniş bağlamları dikkate alır.

Bir sonraki aşamada, bağlamsal temsillerin özet üretimi için kullanılması gerçekleştirilir (decoding). Encoder tarafından üretilen bağlamsal temsiller, decoder tarafından analiz edilir. Decoder, hem encoder'dan gelen bilgiyi hem de önceki çıktıları dikkate alarak kelimeleri sırasıyla tahmin eder.

Son olarak, model, uzun metinlerin en önemli bilgilerinden anlamlı ve dilbilgisel olarak doğru özetler üretir. Bu süreç, özellikle uzun belgeler için optimize edilen LED modelinin güçlü yanlarını ortaya koyar.

3.3 Değerlendirme Metrikleri

Metin özetleme görevlerinde, model performansını ölçmek için referans özetlerle karşılaştırma yapan çeşitli değerlendirme metrikleri kullanılmaktadır. Çalışmamızda, özetleme alanında yaygın olarak tercih edilen ROUGE, METEOR, SacreBLEU ve BERTScore metrikleri kullanılmıştır. Bu metrikler, modelin ürettiği özetlerin doğruluğunu hem dilbilgisel hem de anlamsal açıdan değerlendirmek için uygundur. Ayrıca, her bir metrik kapsamında Precision (Kesinlik), Recall (Duyarlılık) ve F1 Score hesaplanarak performans daha kapsamlı bir şekilde analiz edilmiştir.

Değerlendirme sürecinde kullanılan metriklerin sonuçları genellikle Precision, Recall ve F1 Score gibi temel ölçümlerle analiz edilmektedir. Bu ölçümler, modelin referans özetlerle olan benzerliğini hem doğru tahmin oranı hem de kapsayıcılık açısından değerlendirir (Powers, 2011).

Precision (Kesinlik), modelin ürettiği özetlerdeki doğru tahminlerin oranını ölçer ve bu değer Denklem 3.10 ile hesaplanır. Bu ölçüm, modelin ne kadar isabetli tahminler yaptığını anlamak için kritik bir göstergedir.

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (3.10)$$

Recall (Duyarlılık), referans özetten model tarafından doğru tahmin edilen bilgilerin oranını ifade eder ve Denklem 3.11 ile hesaplanır. Bu metrik, modelin referans özetin ne kadarını doğru şekilde kapsayabildiğini ölçer.

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (3.11)$$

F1 Score, Precision ve Recall arasında bir denge sağlayarak genel performansı özetler ve bu değer Denklem 3.12 ile hesaplanır. F1 Score, özellikle Precision ve Recall arasında bir denge kurarak modelin genel başarısını değerlendirmede etkili bir ölçüm sunar.

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (3.12)$$

3.3.1 SacreBLEU metriği

SacreBLEU, makine çevirisi ve metin özetleme gibi NLP görevlerinde, model çıktılarının referans metinlerle olan benzerliğini ölçmek için kullanılan bir metriktir. BLEU metriğinin geliştirilmiş bir versiyonu olan SacreBLEU, hesaplama sürecini standardize ederek farklı platformlar arasındaki tutarsızlıkları azaltır ve sonuçların yeniden üretilebilirliğini artırır (Hugging Face, t.y.-h; Post, 2018).

N-gram¹ benzerliği, SacreBLEU'nun model çıktısı ile referans metin arasındaki n-gram örtüşmesini ölçmesini sağlayan bir yöntemdir. Bu süreç, modelin hem bireysel kelimeler hem de kelime grupları düzeyindeki doğruluğunu değerlendirir.

Kısalık cezası (brevity penalty), modelin ürettiği metnin referansa kıyasla kısa olduğu durumlarda uygulanan bir mekanizmadır. SacreBLEU, bu ceza ile modelin yalnızca kısa ve öz metinler üretmeye odaklanmasının önüne geçer. Kısalık cezası, Denklem 3.13 ile hesaplanır ve genel skorun daha dengeli bir şekilde değerlendirilmesini sağlar.

$$BP = \begin{cases} 1, & c > r \\ e^{(1-r/c)}, & c \leq r \end{cases} \quad (3.13)$$

Denklem 3.13'te:

- c : üretilen metnin uzunluğu,
- r : referans metnin uzunluğu.

BLEU skoru, SacreBLEU tarafından n-gram doğruluğu ile kısalık cezasını birleştirerek hesaplanan genel bir ölçümdür. Bu skor, modelin referans metne ne kadar yakın bir çıktı

¹ N-gram: Bir metindeki ardışık n kelimedenden oluşan birimler ya da gruplar. Örneğin, "Bu bir testtir" cümlesinde 2-gram (bigram) birimleri "Bu bir" ve "bir testtir" şeklindedir.

ürettiğini anlamak için temel bir metrik olarak kullanılır. BLEU skorunun hesaplanması, Denklem 3.14 ile ifade edilir.

$$SacreBLEU = BP \cdot \exp\left(\sum_{n=1}^N w_n \log p_n\right) \quad (3.14)$$

Denklem 3.14'te:

- BP : kısalık cezası.
- p_n : n -gram kesinlik değeri (precision).
- w_n : n -gram ağırlıkları (genellikle eşit dağıtılır, $w_n = \frac{1}{N}$).

3.3.2 Meteor metriği

METEOR (Metric for Evaluation of Translation with Explicit Ordering), makine çevirisi ve metin özetleme gibi doğal dil işleme görevlerinde, model çıktılarının referans metinlerle olan benzerliğini ölçmek için kullanılan bir değerlendirme metriğidir. BLEU metriğinden farklı olarak, kelime eş anlamlılıklarını ve sıralama farklılıklarını dikkate alarak daha ayrıntılı bir analiz sağlar (Banerjee ve Lavie, 2005; Denkowski ve Lavie, 2014; Hugging Face, t.y.-d).

Kelime eşleşmeleri, METEOR'un model çıktısı ile referans metin arasındaki birebir kelime eşleşmelerini değerlendirirken eş anlamlı kelimeleri ve kök benzerliklerini de analiz etmesini sağlar.

Precision ve Recall, METEOR skorunun temel bileşenlerini oluşturur. Eşleşmelerin kesinliğini (precision) ve duyarlılığını (recall), referans metinle model çıktısı arasındaki ilişkinin nicel analizini sağlar.

Ceza fonksiyonu, sıralama farklılıklarını ve bölünmüş eşleşmeleri cezalandırmak için kullanılan bir mekanizmadır. Bu ceza, eşleşmelerin referans metindeki sıralamasına göre ayarlanır ve Denklem 3.15 ile hesaplanır. Ceza fonksiyonu, modelin metni doğru bir sıralama ile üretmesini teşvik eder (Banerjee ve Lavie, 2005).

$$Penalty = 0.5 \cdot \left(\frac{chunks}{matches}\right)^3 \quad (3.15)$$

Denklem 3.15'te:

- $chunks$: bir chunk, referans metindeki ardışık eşleşmeleri temsil eder. Eğer eşleşmeler ardışık değilse, daha fazla chunk oluşur.

- *matches*: eşleşen öğelerin sayısını ölçer ancak bu eşleşmelerin sıralamasını dikkate almaz.
- Çarpan (0.5) ve küp alma: oranın küpü alınarak sıralama uyumsuzluğuna verilen ceza artırılır ve bu, sıralama farklılıklarının ceza fonksiyonunda daha büyük bir etkiye sahip olmasını sağlar. Ceza fonksiyonuna uygulanan 0.5 çarpanı ise bu ceza miktarını dengeler ve diğer değerlendirme bileşenleriyle uyumlu hale getirir (Banerjee ve Lavie, 2005).

METEOR skoru, ceza fonksiyonu ile ağırlıklandırılmış harmonik ortalama olarak hesaplanır. Bu skor, modelin referans metinle ne kadar uyumlu olduğunu ölçmek için kullanılır ve Denklem 3.16 ile ifade edilir (Banerjee ve Lavie, 2005).

$$METEOR = (1 - Penalty) \cdot F_{mean} \quad (3.16)$$

Denklem 3.16'da F_{mean} , Precision ve Recall'un harmonik ortalamasıdır.

3.3.3 BERTScore metriği

BERTScore, metin karşılaştırma görevlerinde, model çıktılarının referans metinlerle olan anlamsal benzerliğini ölçmek için kullanılan bir değerlendirme metriğidir. Transformer tabanlı bir dil modeli olan BERT'in bağlamsal kelime temsillerinden yararlanarak, metinler arasındaki anlam ilişkisini vektör uzayında değerlendirir (Hugging Face, t.y.-b; T. Zhang ve diğ, 2020).

Bağlamsal temsillerin çıkarılması, BERTScore'un ilk adımıdır. Model çıktısı ve referans metin, BERT gibi önceden eğitilmiş bir dil modeli kullanılarak vektör temsillerine dönüştürülür. Bu süreçte her kelime, metindeki bağlamına göre bir vektörle temsil edilir ve metnin bağlamsal anlamı yakalanır.

Eşleşme skorlarının hesaplanması, çıktı ve referans metinler arasındaki kelime benzerliklerini değerlendirmek için kelime çiftlerinin kosinüs benzerliği üzerinden yapılır. Bu işlem, metinlerin yüzeysel ve derin anlamsal ilişkilerini nicel bir şekilde ölçmek için kritik öneme sahiptir (bkz. Denklem 3.2).

Precision, Recall ve F1, BERTScore'un temel bileşenleridir. Bu metrikler, metinler arasındaki eşleşmelerin kesinliğini (precision), duyarlılığını (recall) ve bunların dengeli bir kombinasyonunu (F1 Score) değerlendirir.

Sonuçların normalize edilmesi, kelimeler arasındaki bağlamsal eşleşmelerin normalize edilmiş ortalaması ile gerçekleştirilir. Bu nihai skor, modelin ürettiği metin ile referans metin arasındaki yüzeysel ve anlamsal benzerliği etkili bir şekilde özetler.

3.3.4 ROUGE metriği

ROUGE (Recall-Oriented Understudy for Gisting Evaluation), özetleme modellerinin performansını değerlendirmek için kullanılan en popüler metriklerden biridir. İlk olarak Lin (2004) tarafından önerilmiş ve n-gram tabanlı, sıralı benzerlik ölçümleri sunan metriklerle metinlerin yüzeysel ve yapısal benzerliğini değerlendirmek için geliştirilmiştir. Bu metrikler arasında ROUGE-n, ROUGE-L ve ROUGE-W bulunmaktadır. Ayrıca, çok cümleli özetleme senaryoları için optimize edilmiş bir varyant olan ROUGE-LSum, özellikle haber özetleme ve çoklu belge özetleme gibi alanlarda yaygın olarak tercih edilmektedir.

Bu çalışmada, ROUGE-n, ROUGE-L ve ROUGE-LSum metriklerinin hesaplanmasında Hugging Face tarafından sağlanan değerlendirme kütüphanesi kullanılmıştır (Hugging Face, t.y.-g). ROUGE-W metriği için ise Py-ROUGE¹ aracı tercih edilmiştir (Antognini, t.y.).

ROUGE-n, model çıktısı ile referans metin arasındaki n-gram eşleşmelerini değerlendirir. N-gram, bir metindeki ardışık kelime gruplarını ifade eder. Bu metrik, özetlerin yüzeysel benzerliğini ölçmek için yaygın olarak kullanılır (Lin, 2004). ROUGE-n hesaplamasında kullanılan metrikler arasında Precision (Denklem 3.17), Recall (Denklem 3.18) ve F1 Skoru (Denklem 3.19) yer alır.

$$Precision_n = \frac{Match_n}{Hypothesis_n} \quad (3.17)$$

Denklem 3.17'de $Match_n$, model çıktısı ve referans metin arasındaki ortak n-gram sayısını; $Hypothesis_n$ ise model çıktısındaki toplam n-gram sayısını ifade eder.

$$Recall_n = \frac{Match_n}{Reference_n} \quad (3.18)$$

Denklem 3.18'de $Reference_n$, referans metindeki toplam n-gram sayısını ifade eder.

$$F1_n = 2 \cdot \frac{Precision_n \cdot Recall_n}{Precision_n + Recall_n} \quad (3.19)$$

¹ Py-ROUGE, ROUGE metriklerini hesaplamak için kullanılan bir Python kütüphanesi.

ROUGE-L, model çıktısı ile referans metin arasındaki sıralı kelime eşleşmelerini ölçmek için Longest Common Subsequence (LCS)¹ algoritmasını kullanır. LCS, iki metin arasındaki en uzun sıralı kelime dizisini belirler ve bu sıralı eşleşmelerin uzunluğuna dayalı bir benzerlik değerlendirmesi sunar. Ancak, bu kelimelerin ardışık olması gerekmez, yalnızca sıralı olmaları yeterlidir (Lin, 2004). ROUGE-L hesaplamasında kullanılan metrikler arasında Precision (Denklem 3.20), Recall (Denklem 3.21) ve F1 Skoru (Denklem 3.22) yer alır.

$$LCS\ Precision = \frac{LCS(Hypothesis, Reference)}{Length(Hypothesis)} \quad (3.20)$$

$$LCS\ Recall = \frac{LCS(Hypothesis, Reference)}{Length(Reference)} \quad (3.21)$$

$$LCS\ F1 = 2 \cdot \frac{LCS\ Precision \cdot LCS\ Recall}{LCS\ Precision + LCS\ Recall} \quad (3.22)$$

ROUGE-W, LCS eşleşmelerini ağırlıklandırarak, uzun ardışık kelime eşleşmelerine daha fazla önem verir. Bu metrik, özellikle ardışık eşleşmelerin uzunluğunu ödüllendiren bir yaklaşıma sahiptir (Lin, 2004). Hesaplama formülü Denklem 3.23 ile sunulmuştur.

$$ROUGE-W = \frac{LCS\ Length + \alpha \cdot Continuity}{Total\ Length + \alpha} \quad (3.23)$$

Denklem 3.23'te:

- α : uzun sürekliliklerin önem derecesini belirleyen bir sabittir.
- *Continuity*: sürekli ardışık eşleşme uzunluklarının toplamıdır.

ROUGE-LSum, çok cümleli özetleme senaryoları için optimize edilmiş bir ROUGE-L varyantıdır. Metin düzeyinde birden fazla cümlenin sıralı karşılaştırmasını içerir. Özellikle haber özetleme ve çoklu belge özetleme gibi alanlarda yaygın olarak tercih edilir (Hugging Face, t.y.-g).

3.4 Kullanılan Veri Setleri

Bu çalışmada, metin özetleme yöntemlerini değerlendirmek ve farklı özetleme modellerinin performansını sistematik bir şekilde analiz etmek amacıyla iki farklı veri seti kullanılmıştır.

¹ LCS, iki dizgi arasında sıraları korunmuş en uzun ortak alt diziyi bulmak için kullanılan bir algoritmadır.

BBC News Summary, 2004-2005 yıllarında BBC News web sitesinde yayımlanan 2.225 haber metni ve özetlerinden oluşmaktadır. İş dünyası, eğlence, politika, spor ve teknoloji olmak üzere beş kategoriye ayrılmıştır. Greene ve Cunningham (2006) tarafından metin sınıflandırma ve kümeleme çalışmaları için benchmark olarak tanımlanmış ve araştırmalarda yaygın olarak kullanılmıştır. Bu çalışmada, özetleme yöntemlerinin performansını değerlendirmek için veri setinin tamamı kullanılmıştır (Kaggle, t.y.)¹. Veri setine ilişkin ayrıntılı özellikler Çizelge 3.1’de gösterilmiştir.

Çizelge 3.1 : BBC News Summary veri seti özellikleri.

Tanım	BBC News Summary veri seti
Toplam Haber Sayısı	2225 haber
Toplam Özet Sayısı	2225 özet
Haberlerin Ortalama Kelime Sayısı	384 kelime
Özetlerin Ortalama Kelime Sayısı	165 kelime

PubMed Summarization, PubMed’de yayımlanan bilimsel makalelerden ve özetlerinden oluşmaktadır. Bu veri seti, özellikle uzun metin özetleme çalışmaları için sıkça tercih edilmektedir (Cohan ve diğ, 2018; Hugging Face, t.y.-f). Çalışmada, veri setinden belirli kriterlere göre rastgele seçilmiş 2.000 makale kullanılmıştır. Veri seçimi, 512-1024 kelimelik makalelerden ve 80-300 kelimelik özetlerden oluşan içeriklerden rastgele yapılmıştır. Bu kriterler, uzunluk bakımından dengeli bir veri kümesi oluşturmak ve özetleme modellerinin performansını farklı metin uzunluklarında değerlendirmek amacıyla belirlenmiştir. Veri setine ilişkin ayrıntılı özellikler Çizelge 3.2’de gösterilmiştir.

Çizelge 3.2 : PubMed Summarization veri seti özellikleri.

Tanım	PubMed Summarization veri seti
Toplam Makale Sayısı	2000 Makale
Toplam Özet Sayısı	2000 özet
Makalelerin Ortalama Kelime Sayısı	833 kelime
Özetlerin Ortalama Kelime Sayısı	138 kelime

3.5 Extractive Özetleme: Veri İşleme ve Uygulama Süreci

Extractive özetleme yöntemlerinin uygulanabilmesi için, metinlerin özetlenmeden önce belirli bir hazırlık sürecinden geçmesi gereklidir. Bu çalışmada kullanılan BBC News ve PubMed veri setlerine dayalı olarak, ön işleme süreci üç temel aşamadan oluşmaktadır.

¹ Kaggle, veri bilimi yarışmaları ve veri setleri sağlayan bir çevrimiçi platform.

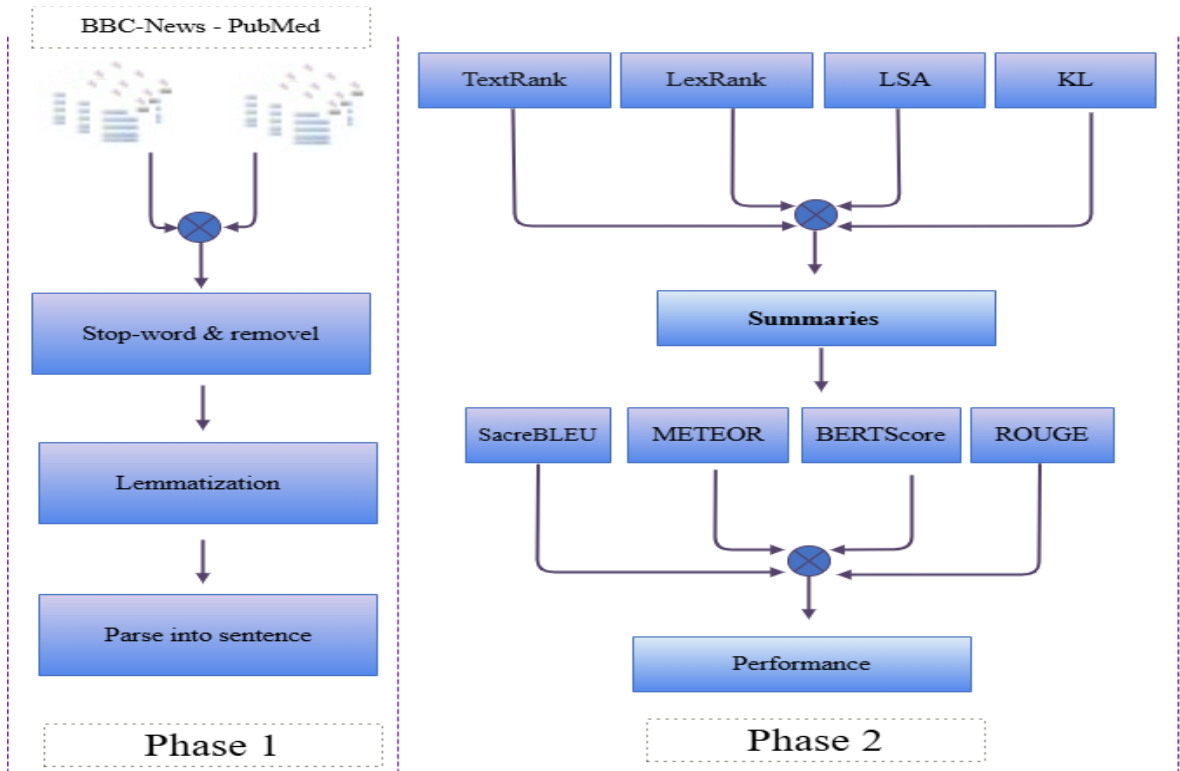
gereksiz kelimelerin çıkarılması (stop-word removal), kelimelerin kök forma dönüştürülmesi (lemmatization) ve cümlelere ayrıştırma (parse into sentence).

Stop-word çıkarımı, metindeki anlamsal olarak düşük katkıya sahip ancak sık tekrarlanan kelimelerin (örneğin, bağlaçlar ve edatlar) çıkarılması işlemidir. Bu adım, daha anlamlı cümlelerin seçilmesini kolaylaştırır ve metnin analizini iyileştirir.

Lemmatization, kelimelerin kök hallerine indirilerek tek bir türev etrafında normalize edilmesi işlemidir. Bu yöntem, kelime köklerinin farklı biçimlerde tekrar etmesini önleyerek metindeki anlamsal tutarlılığı artırır.

Cümlelere ayrıştırma, işlenmiş metinlerin cümle bazında ayrıştırılması işlemidir. Bu aşamada, her bir cümle daha sonraki aşamalarda model tarafından analiz edilmek üzere bağımsız birer birim olarak ele alınır.

Şekil 3.1’de gösterildiği gibi, bu ön işleme adımlarının ardından, extractive yöntemler (TextRank, LexRank, LSA ve KL) kullanılarak metin özetleri oluşturulmuştur. Her yöntem, metin içerisindeki cümleler arasında benzerlik veya önem skorlarına dayalı olarak en anlamlı cümleleri belirlemiştir. Özetleme performansını değerlendirmek için, dört farklı metrik (SacreBLEU, METEOR, BERTScore ve ROUGE) kullanılmıştır. Performans analizleri, üç farklı özet uzunluğu için yapılmıştır. 50 kelime, 100 kelime ve 150 kelime.



Şekil 3.1 : Extractive yöntemlerin uygulama süreci.

3.6 Abstractive Modeller: Kullanım ve Uygulama Süreçleri

Abstractive metin özetleme modellerinde, metinlerin anlamını koruyarak yeniden yazılması hedeflenir. Bu süreçte, extractive yöntemlerin aksine, kapsamlı bir ön işleme aşamasına ihtiyaç duyulmamıştır. Veri setleri, olduğu gibi modellerin girişine verilmiştir.

BBC News veri setinde kullanılan modeller, Pegasus, BART ve T5-Large modellerini içermektedir. Modellerin giriş uzunluğu 512 token olarak ayarlanmış ve bu sınır, BBC News veri setinin ortalama metin uzunluğunu karşılamak amacıyla seçilmiştir. Ancak, 150 kelimelik özetlerde T5-Large modelinin yetersiz kelime üretimi nedeniyle yalnızca Pegasus ve BART modelleri kullanılmıştır.

PubMed veri setinde kullanılan modeller, Pegasus, BART ve LED modellerinden oluşmaktadır. T5-Large, maksimum giriş uzunluğunun 512 token ile sınırlı olması nedeniyle bu veri setinde kullanılmamıştır (Hugging Face, t.y.-i; Raffel ve diğ, 2020). PubMed veri seti metinlerinin büyük bir kısmı bu sınırı aştığı için Pegasus, BART ve LED modelleri tercih edilmiştir. Bu modellerin giriş uzunluğu, PubMed veri setinin metin boyutlarını karşılayacak şekilde 1024 token olarak ayarlanmıştır.

Çizelge 3.3'te kullanılan modellerin detayı aşağıda listelenmektedir:

- PEGASUS: Google tarafından geliştirilen bir encoder-decoder modelidir. Maksimum giriş uzunluğu 1024 token olan model, CNN/Daily Mail veri seti üzerinde eğitilmiştir (J. Zhang ve diğ, 2020).
- BART: Facebook tarafından geliştirilen bir encoder-decoder modelidir. Eksik veya gürültülü metinlerin özetlenmesinde başarılıdır. Model, CNN/Daily Mail veri seti üzerinde eğitilmiştir (Lewis ve diğ, 2020).
- T5-LARGE: Google tarafından geliştirilen Text-to-Text Transfer Transformer modelidir. Maksimum giriş uzunluğu 512 token olan bu model, C4¹ veri seti ile eğitilmiştir (Raffel ve diğ, 2020).
- LED: Allen Institute for AI tarafından geliştirilen LED modeli, uzun metinlerin özetlenmesine odaklanmıştır. Maksimum giriş uzunluğu 4096-16384 token olarak ayarlanmıştır. Arxiv² ve PubMed veri setleri üzerinde eğitilmiştir (Beltagy ve diğ, 2020).

¹ C4 (Colossal Clean Crawled Corpus), GPT ve T5 gibi modellerin eğitimi için kullanılan büyük bir web metin koleksiyonu.

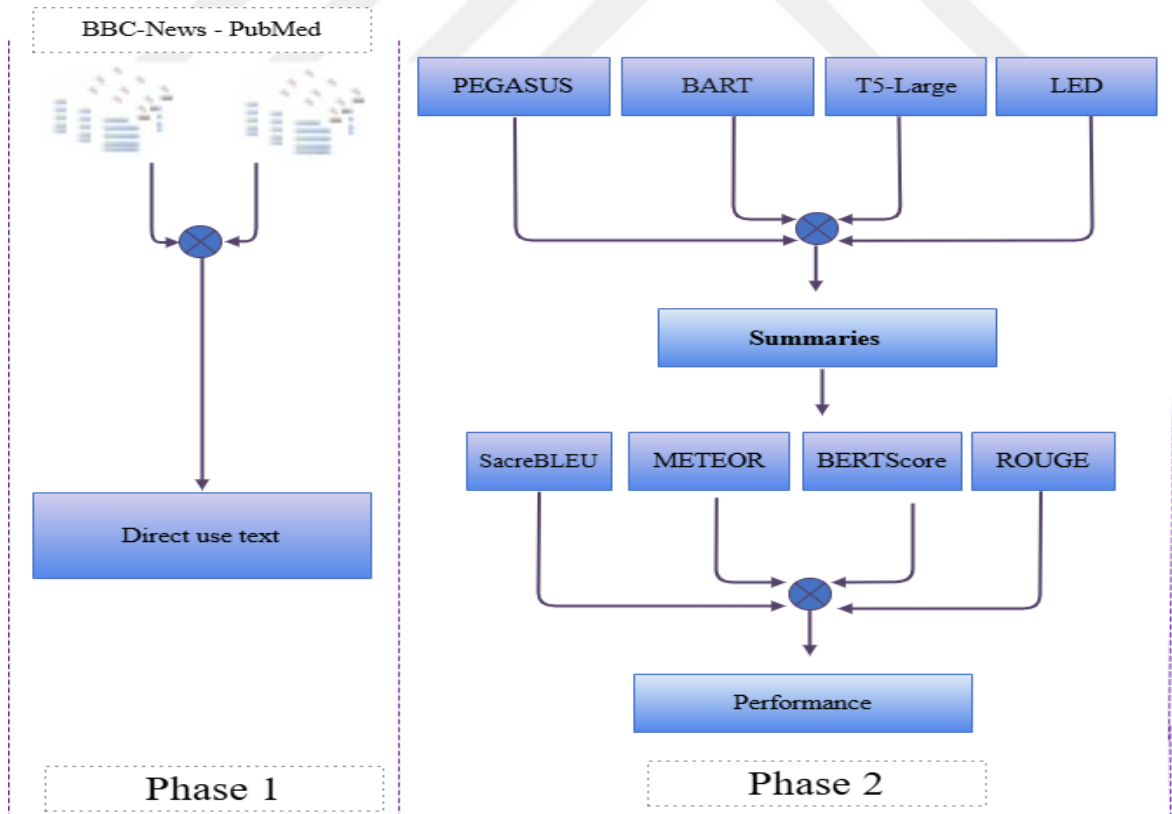
² Arxiv, bilimsel makalelerden oluşan açık erişimli bir veri kümesi (<https://arxiv.org>).

Çizelge 3.3 : Kullanılan modellerin detayları.

Model	Geliştirici	Model Türü	Parametre Sayısı	Maksimum Giriş Uzunluğu	Eğitildiği Veri Seti
Pegasus	Google	Encoder-Decoder	560 milyon	1024 token	CNN/Daily Mail
BART	Facebook	Encoder-Decoder	406 milyon	1024 token	CNN/Daily Mail
T5-Large	Google	Text-to-Text Transfer Transformer (Encoder-Decoder)	738 milyon	512 token	C4 (Colossal Clean Crawled Corpus)
LED	Allen Institute for AI	Encoder-Decoder (Longformer)	Açıklanmamış	4096-16384 token	Arxiv ve PubMed

Uygulama süreci (Şekil 3.2), BBC News ve PubMed veri setlerinin Pegasus, LED, T5-Large ve BART modellerine giriş verisi olarak hazırlanmasıyla başlamıştır. Bu veri setleri, modellenmiş halleriyle farklı boyutlardaki metin özetlerinin üretilmesi amacıyla kullanılmıştır.

Bir sonraki aşamada, yüklenen metinler abstractive modellerle işlenmiş ve anlamlarına dayalı olarak insan yazımına benzer özetler oluşturulmuştur. Bu süreç, özet uzunluğuna göre üç farklı formatta (50 kelime, 100 kelime ve 150 kelime) gerçekleştirilmiştir.



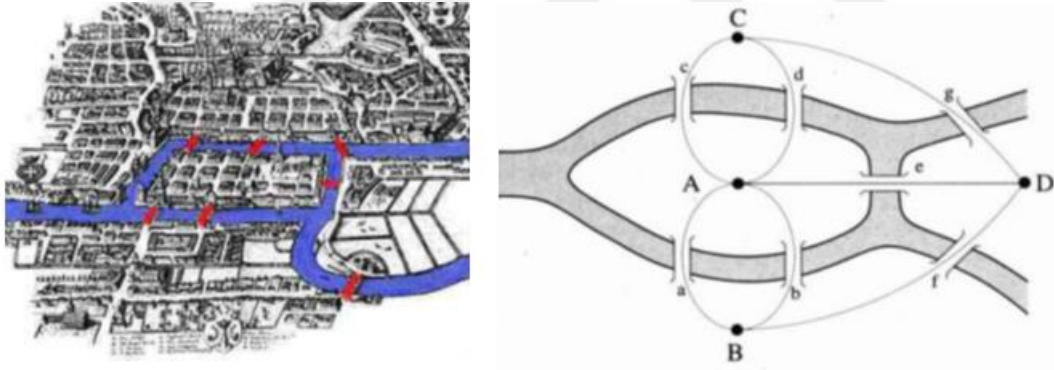
Şekil 3.2 : Abstractive modellerin uygulama süreci.

4. BASKIN KÜME YÖNTEMİ İLE ÖZETLEME: YENİ BİR YAKLAŞIM

4.1 Graf Teorisi

Graf teorisi, düğümler (nodes) ve bu düğümler arasındaki bağlantıları ifade eden kenarlardan (edges) oluşan matematiksel bir yapı olarak, modern bilimde çeşitli problemlerin çözümünde güçlü bir araçtır. İlk olarak 18. yüzyılda Leonhard Euler tarafından tanıtılan graf teorisi, özellikle meşhur Königsberg Köprü Problemi ile dikkat çekmiştir. Bu problem, bir nehrin iki yakasında bulunan dört yerleşim bölgesini birbirine bağlayan yedi köprü üzerinden belirli kurallara göre geçiş yapılabilir mi sorusunu ele almıştır (Şekil 4.1). Euler'ın bu problemi çözümüyle birlikte graf teorisinin temelleri atılmış ve düğüm (yerleşim bölgeleri) ile kenar (köprüler) gibi temel kavramlar tanımlanmıştır (Euler, 1736).

Graf teorisi, ağ analizi, sosyal ağlar, biyoinformatik, lojistik ve iletişim gibi birçok alanda, düğümler arasındaki ilişkileri modellemek ve bu ilişkilerden anlamlı sonuçlar çıkarmak için kullanılmaktadır (West, 2001).



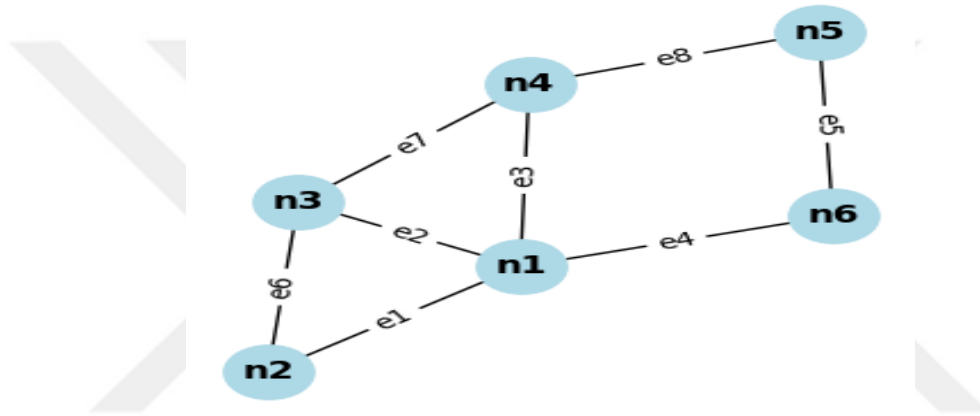
Şekil 4.1 : Königsberg köprü problemi.

Şekil 4.1 Königsberg Köprü Problemi, graf teorisinin tarihi başlangıç noktası olan Königsberg Köprü Problemi'ni görselleştirmektedir. Yerleşim bölgeleri (A, B, C, D) düğümleri, köprüler ise kenarları temsil eder. Euler, bu problemi çözerek graf teorisinin temellerini atmıştır (Euler, 1736; Newman, 2010).

Graf yapıları, yönlü (directed) veya yönsüz (undirected), ağırlıklı (weighted) veya ağırlıksız (unweighted) olabilir. Şekil 4.2'de graf teorisinde kullanılan temel kavramlar görselleştirilmiştir. Burada düğümler (n_1, n_2, \dots, n_6) ve bu düğümleri birbirine bağlayan kenarlar (e_1, e_2, \dots, e_8) gösterilmiştir. Bu basit yapı, bir grafın temel özelliklerini anlamak

için iyi bir model sunmaktadır (Easley ve Kleinberg, 2010). Bir grafın özelliklerini anlamak için aşağıdaki temel terimler açıklanabilir:

- Düğüm (Node): grafın temel birimi; genellikle $V = \{v_1, v_2, \dots, v_n\}$ olarak ifade edilir.
- Kenar (Edge): düğümleri birbirine bağlayan ilişki; $E = \{e_1, e_2, \dots, e_m\}$.
- Komşuluk (Adjacency): bir düğümün doğrudan bağlantılı olduğu diğer düğümler.
- Ağırlık (Weight): kenarların ilişki gücü, maliyet veya mesafe gibi bir değeri ifade eder.
- Derece (Degree): bir düğümün bağlandığı kenarların sayısı.



Şekil 4.2 : Basit bir graf.

4.2 Dominating Set Yöntemine Giriş

Dominating Set (Baskın Küme) algoritması, graf teorisi ve optimizasyon alanlarında önemli bir yer tutar. Bu yöntem, özellikle veri kümelerindeki kümelenme problemlerinin çözümünde etkin bir araç olarak kullanılır. Baskın küme kavramı, sosyal ağ analizi, biyoinformatik, iletişim ağları ve metin özetleme gibi geniş bir uygulama yelpazesine karşımıza çıkar (Cockayne ve Hedetniemi, 1977; Sampathkumar ve Walikar, 1979). Bu bölümde baskın kümenin tanımı, çeşitleri, terminolojisi, matematiksel ifadeleri ve graflar üzerindeki karmaşıklığı detaylı olarak ele alınacaktır.

Bir graf $G = (V, E)$ üzerinde, V düğüm kümesi ve E kenar kümesi olmak üzere, baskın küme $D \subseteq V$, grafın her bir düğümünün ya bu kümenin bir üyesi olduğu ya da bu kümenin bir üyesiyle komşu olduğu bir alt kümedir. Diğer bir deyişle, D kümesindeki her düğüm, D dışında kalan tüm düğümleri kapsar veya onlarla komşudur. Bu tanım, graf teorisi ve

kümelenme problemleri bağlamında kritik bir öneme sahiptir (Cockayne ve Hedetniemi, 1977). Matematiksel olarak baskın küme Denklem 4.1 ile ifade edilir.

$$D = \{v_i \in V \mid \nexists v_j \in V, j \neq i : v_i \text{ tamamen } v_j' \text{yi domine eder ve } v_i \text{ en az bir kriterde } v_j' \text{den daha iyidir}\} \quad (4.1)$$

Denklem 4.1'de:

- D : baskın düğümler kümesini temsil eder.
- V : grafın düğümler kümesi.
- $\nexists v_j$: hiçbir v_j düğümü, v_i 'yi tamamen domine edemez.
- $j \neq i$: düğümün kendisi dışındaki tüm düğümler dikkate alınır.

4.3 Dominating Set Çeşitleri

Dominating Set (Baskın Küme), graf teorisindeki temel baskın küme tanımına dayalı olarak, $D \subseteq V$ kümesi tüm graf düğümlerini kapsar ya da onlarla komşudur. Baskın kümenin genel tanımı Denklem 4.1'de yapılmaktadır (Cockayne ve Hedetniemi, 1977).

Minimum Dominating Set (Minimum Baskın Küme), grafın tüm düğümlerini kapsayan en küçük baskın kümedir. Bu kümenin boyutu, grafın dominasyon sayısı $\gamma(G)$ ¹ olarak adlandırılır. Matematiksel olarak Denklem 4.2 ile sunulmaktadır.

$$D(G) = \min\{|S| : S \subseteq V, \forall v \in V, N[v] \cap S \neq \emptyset\} \quad (4.2)$$

Denklem 4.2'de $N[v]$, v düğümünün kendisi ve komşularını ifade eder (Cockayne ve Hedetniemi, 1977). Minimum baskın küme problemi, graf teorisinde NP-zor olarak sınıflandırılmıştır ve bu nedenle büyük ölçekli graflarda hesaplama açısından zorluklar içermektedir (Guha ve Khuller, 1998).

Connected Dominating Set (Bağlantılı Baskın Küme), baskın kümenin indüklediği alt grafın bağlantılı olduğu bir kümedir. Matematiksel olarak Denklem 4.3 ile sunulmaktadır.

$$CDS : D \subseteq V, \forall v \in V, (N[v] \cap D \neq \emptyset) \text{ ve } G[D] \text{ bağlantılı} \quad (4.3)$$

¹ $\gamma(G)$, grafın tüm düğümlerini kapsayan en küçük baskın kümenin eleman sayısını ifade eder ve grafın yapısal özelliklerini belirleyen önemli bir ölçüttür.

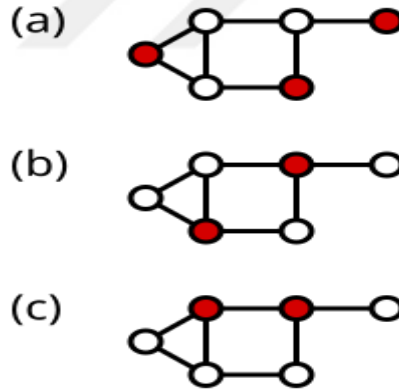
Denklem 4.3'te $G[D]$, D kümesi tarafından oluşturulan alt grafi ifade eder (Guo ve diğ., 2018). Connected Dominating Set problemleri, özellikle kablosuz ağlarda etkin iletişim yollarının belirlenmesi ve enerji verimliliğinin sağlanması gibi uygulama alanlarında kullanılmaktadır (Guha ve Khuller, 1998).

Baskın küme kavramı, Şekil 4.3'te görselleştirilmiştir. Bu şekil, baskın kümenin ve onun çeşitli türlerinin görsel bir açıklamasını sunmaktadır (Wikipedia contributors, t.y.).

Graf (a), Dominating Set'i (Baskın Küme) temsil etmektedir. Bu graf, tüm diğer düğümleri kapsayan veya komşu olan baskın düğümleri kırmızı ile göstermektedir (Sampathkumar ve Walikar, 1979).

Graf (b), Minimum Dominating Set'i (Minimum Baskın Küme) ifade etmektedir. Bu türde, grafin dominasyon sayısını ($\gamma(G)$) sağlayan en küçük düğüm kümesi kırmızı olarak belirtilmiştir (Cockayne ve Hedetniemi, 1977).

Graf (c), Connected Dominating Set'i (Bağlantılı Baskın Küme) temsil etmektedir. Burada, kırmızı düğümler hem baskın olma hem de bağlantılı olma koşulunu sağlamaktadır (Guha ve Khuller, 1998; Sampathkumar ve Walikar, 1979).



Şekil 4.3 : Dominating Set çeşitleri.

4.4 Algoritmik Karmaşıklık ve Zorluk

Minimum baskın küme problemi, NP-zor olarak sınıflandırılan bir optimizasyon problemidir. Bu nedenle, büyük ölçekli graflarda optimal çözüm bulmak hesaplama açısından oldukça zordur. Connected Dominating Set problemleri de NP-zor kategorisine girer ve genellikle yaklaşık algoritmalar veya sezgisel yöntemler kullanılarak çözülür (Guha ve Khuller, 1998). Bu tür problemler, özellikle kablosuz ağlarda enerji verimliliği, iletişim

optimizasyonu ve veri sıkıştırma gibi uygulamalarda yaygın olarak kullanılmaktadır (Wu ve Li, 1999).

4.5 Dominating Set ile Metin Özetleme: Yöntemin Geliştirilmesi

Dominating Set (DS) algoritması, metin özetleme alanında, bir metni graf tabanlı bir modelle temsil ederek anlamlı ve bilginin yoğun olduğu cümlelerin seçimini sağlar. Bu yaklaşımda, metnin her bir cümlesi bir düğüm olarak modellenirken, düğümler arasındaki kenar ağırlıkları, cümleler arası benzerliklerle ilişkilendirilir. Bu benzerlikler genellikle TF-IDF, kosinüs benzerliği veya kelime gömme teknikleri gibi metrikler üzerinden hesaplanır (Pavan ve Pelillo, 2007; Salton ve Buckley, 1988).

Graf üzerindeki baskın kümeler, özetleme için seçilecek cümlelerin belirlenmesinde kullanılır. DS algoritmasının öne çıkan avantajı, grafın içsel özelliklerini optimize ederek yalnızca bilginin yoğun olduğu cümleleri seçmesi ve bu seçimlerin metin genelindeki anlam bütünlüğünü korumasıdır. Özellikle büyük ve uzun metinlerde, tekrar eden ya da ilgisiz bilgilerin dışlanması sağlamak için etkili bir yöntemdir (von Luxburg, 2007).

Metin özetleme sürecinde Dominating Set yöntemi, özetlenmiş cümlelerin seçimini, cümleler arası ilişkilerin optimize edilmesi yoluyla gerçekleştirir. Bu süreç, özetleme performansının ROUGE gibi metrikler üzerinden değerlendirilmesiyle tamamlanır (Lin, 2004). Bu bağlamda, önerilen yaklaşım, DS algoritmasının metin özetleme alanında yenilikçi bir kullanımını sunarak hem doğruluk hem de bilgi yoğunluğu açısından üstün bir performans sağlamayı hedeflemektedir.

4.5.1 Metin ön işleme ve hazırlık

Bu bölüm, metinlerin özetleme işlemi için hazırlanması sürecini kapsamaktadır. Graf tabanlı modelin doğru çalışabilmesi için metinlerin hem cümle hem de kelime seviyesinde işlenmesi kritik bir öneme sahiptir. Bu süreçte bir dizi adım uygulanmıştır.

İlk adım, kaynak metnin cümlelere bölünmesidir. Kaynak metinler dilbilgisel kurallara göre cümlelere ayrılmıştır. Bu işlem, her cümlenin bağımsız bir veri birimi olarak analiz edilmesini sağlar. Graf yapısında düğüm olarak temsil edilecek bu cümleler, nltk kütüphanesinin `sent_tokenize` fonksiyonu ile oluşturulmuştur. Cümlelerin doğru bir şekilde bölünmesi, graf yapısının doğruluğunu ve özetleme sürecinin başarısını doğrudan etkiler.

Bir sonraki adım, stop-word çıkarma işlemidir. Bu adımda, cümleler anlamsal yük taşımayan kelimelerden temizlenmiştir. Stop-word çıkarma işlemi, metindeki kritik olmayan kelimeleri arındırarak anlam taşıyan kelimelerin ön plana çıkmasını sağlar. Bu işlem sırasında, nltk ve spacy kütüphaneleri kullanılmıştır¹.

Daha sonra, etkin olmayan elemanların kaldırılması gerçekleştirilmiştir. Bu süreçte, kelimeler kök formlarına indirgenmiş ve tekrarlayan kelimeler metinden çıkarılmıştır. Lemmatization ve stemming teknikleri, kelimelerin kök halleriyle analiz edilmesini sağlamıştır. Örneğin, “running”, “ran” ve “runs” gibi kelimeler anlam açısından eşdeğer oldukları için “run” olarak normalize edilmiştir. Bu işlemler, nltk kütüphanesinin PorterStemmer (kelimelerin kök hâlini çıkarma aracı) ve spacy kütüphanesinin lemmatization (kelimeleri sözlük biçimine dönüştürme işlemi) araçları kullanılarak gerçekleştirilmiştir.

Son adım, cümle ve kelime havuzlarının oluşturulmasıdır. İşlenen metinler, graf yapısına temel oluşturacak iki ayrı havuzda organize edilmiştir. Cümle havuzu, metindeki her bir cümlenin temizlenmiş ve köklenmiş halini saklarken, kelime havuzu, metindeki her bir kelimenin kök hali ve hangi cümlelerde yer aldığı bilgisiyle yapılandırılmıştır. Bu yapı, kelimelerin cümleler içindeki dağılımını analiz etmeye ve graf modeli güçlendirmeye olanak tanır.

Bu aşamada elde edilen cümle ve kelime havuzları, graf modelinin düğüm ve kenar yapılarını oluşturmak için temel verilerdir. Cümleler arasındaki ilişkiler, kelime vektörlerinin kullanımıyla hesaplanacak olan kosinüs benzerliği metriği üzerinden belirlenir. Kelime vektörleri, Word2Vec² modelinden elde edilmiştir. Bu süreç, kenar ağırlıklarının hesaplanması ve graf yapısının oluşturulmasında kritik bir öneme sahiptir. Detaylı açıklamalar, 4.5.2 ve 4.5.3 bölümlerde sunulacaktır.

4.5.2 Önerilen yaklaşım (DSETS)

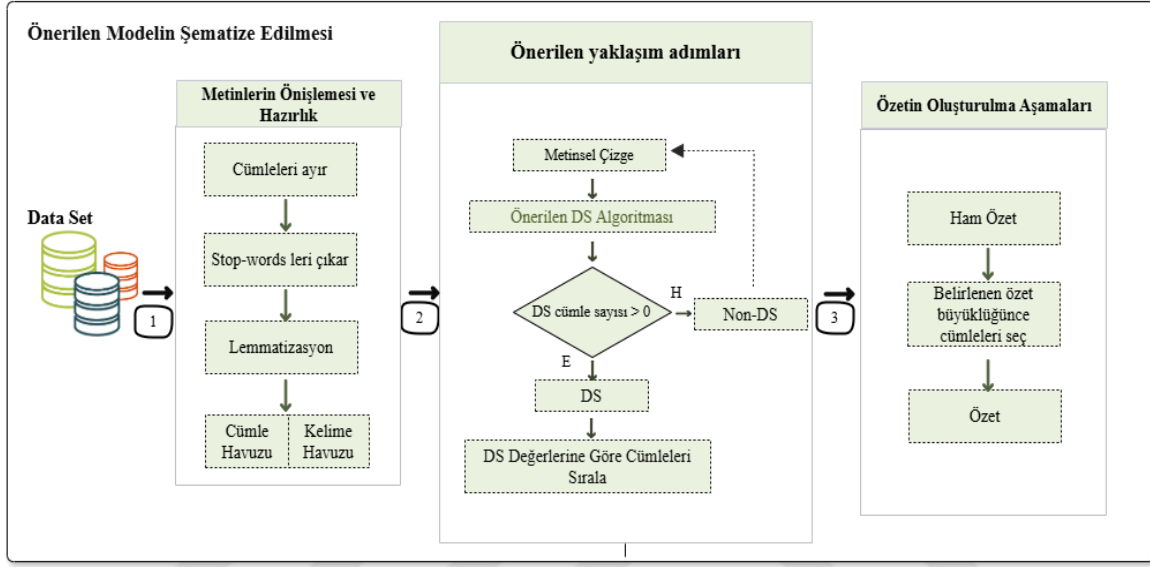
Bu bölümde, önerilen Dominating Set tabanlı Extractive Text Summarization (DSETS) yönteminin temel yapısı ve işleyişi açıklanmaktadır. Yöntemin genel hatları, süreçleri ve kullanılan teknikler iki ana görsel üzerinden ele alınmıştır. Şekil 4.5.2.1, önerilen yaklaşımın genel yapısını şematik bir şekilde sunarken; Şekil 4.5.2.2, bu süreçlerin daha detaylı ve

¹ Nltk ve Spacy, metin işleme ve doğal dil işleme (NLP) için çok çeşitli araçları sağlayan iki Python kütüphanesidir.

² Word2Vec, kelimeler arasındaki anlamsal ilişkileri vektörler olarak temsil eden bir kelime gömme modeli.

teknik bir analizini içermektedir. Aşağıdaki bölümlerde bu görseller açıklanacak ve her bir adım ayrıntılı bir şekilde ele alınacaktır.

Şekil 4.4'te önerilen DSETS yönteminin ana aşamaları, basitleştirilmiş bir akış şeması ile gösterilmektedir. Bu şemada, metinlerin işlenmesinden özet oluşturulmasına kadar olan süreç, anlaşılır bir şekilde sunulmuştur. Metin işleme aşaması, önerilen yöntemin temelini oluşturur ve özletmenin doğruluğunu artırmak için kritik bir ön adımdır.



Şekil 4.4 : Önerilen yöntemin temel yapı taşları.

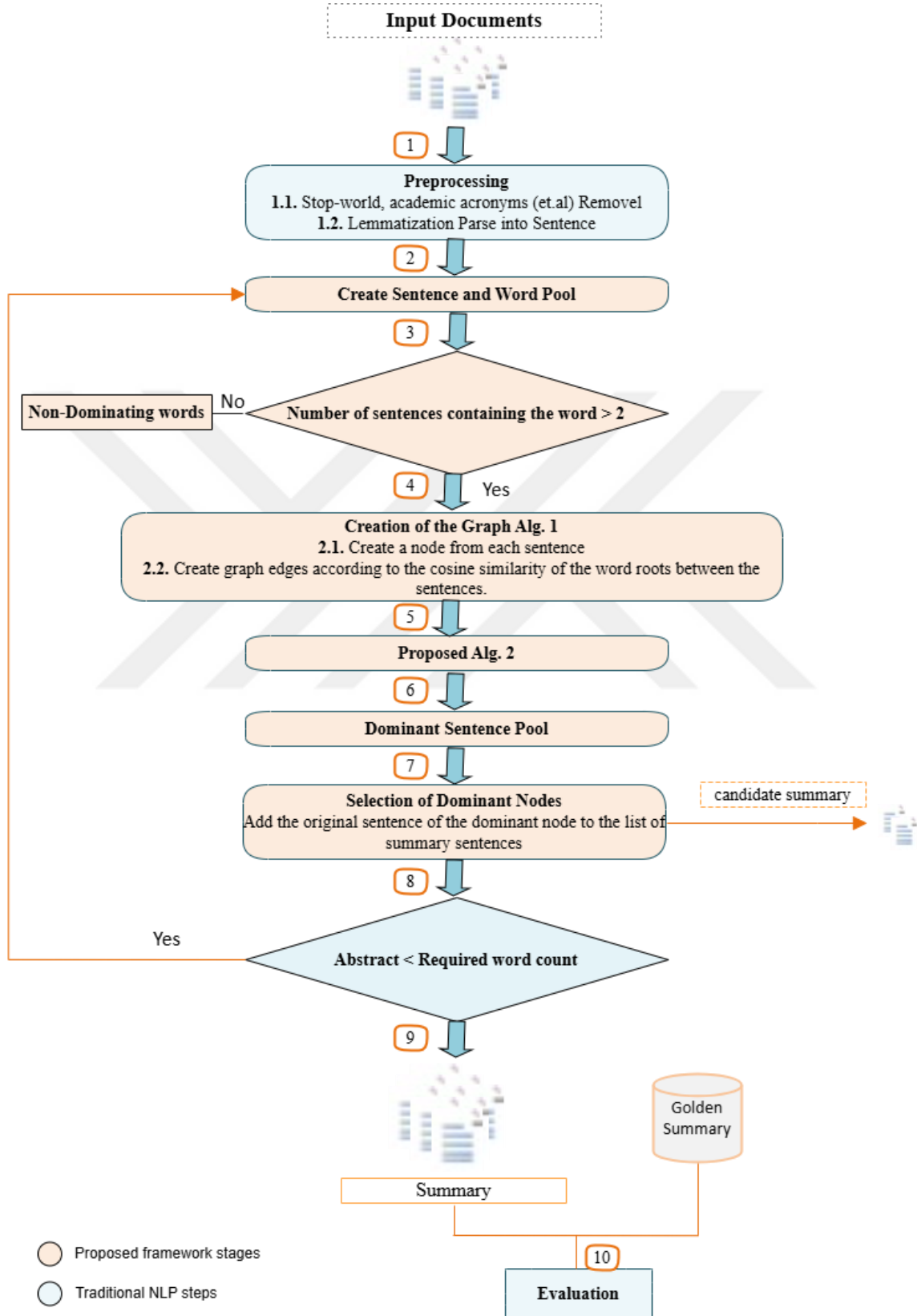
Şekil 4.4'te gösterilen süreç, üç ana adımdan oluşmaktadır. Birinci adım, metinlerin işlenmesi ve hazırlık sürecidir. Bu adım, Bölüm 4.5.1'de detaylandırılmıştır ve metinlerin özetleme için uygun hale getirilmesini sağlar.

İkinci adım, önerilen yaklaşımın adımlarıdır. Bu aşamada, cümleler arasında semantik ilişkileri keşfetmek amacıyla bir graf modeli oluşturulur. Oluşturulan bu graflarda, baskın kelimelerin yer aldığı cümleler analiz edilerek baskın düğümler (baskın cümleler) tespit edilir. Eğer baskın düğümler bulunamazsa, baskın olmayan cümleler de aday özetlere dahil edilebilir.

Üçüncü adım, özeti oluşturulmasıdır. Aday özet cümleleri, belirli bir uzunluk kriterine göre seçilerek bir araya getirilir. Oluşturulan özetler, metnin ana fikrini koruyarak kompakt bir biçimde sunulmaya çalışılır.

Şekil 4.5, önerilen Dominating Set tabanlı Extractive Text Summarization (DSETS) yönteminin teknik detaylarını içeren adımları sunmaktadır. Bu yöntem, ham metinlerin

özetlenmesi sürecinde uygulanan on temel adımdan oluşur. Bu bölümde, her adım ayrıntılı bir şekilde ele alınarak süreçlerin işleyişi açıklanmaktadır.



Şekil 4.5 : Önerilen DSETS yönteminin detaylı süreç akışı.

Input Documents (Girdi Belgeleri) ham metinlerden oluşur ve özetleme işlemine temel oluşturacak giriş verisi olarak alınır. Bu çalışmada kullanılan metinler, farklı konulara ait haberlerden oluşan BBC News veri setinden seçilmiştir. Bu metinler, çeşitli uzunluk ve türde içerikleri barındırarak yöntem için geniş bir test alanı sağlar.

Adım (1) Önişleme. Metinler, doğal dil işleme (NLP) teknikleriyle analiz için hazır hale getirilir. Bu aşama, Bölüm 4.5.1’de detaylıca açıklanmıştır.

Adım (2) Cümle ve kelime havuzlarının oluşturulması. Metinlerdeki cümleler havuza alınır ve her bir cümle kelimelerine ayrılır. Bu aşama, Bölüm 4.5.1’de detaylıca açıklanmıştır.

Adım (3) Kelime frekansı kontrolü. Kelimeler, cümlelerdeki kullanım sıklıklarına göre iki kategoriye ayrılır. Baskın kelimeler (Dominant Words - DS Words), en az üç farklı cümlede geçen kelimelerden oluşur ve baskın kelimeler havuzuna (DS Words Pool) eklenir. Baskın olmayan kelimeler (Non-DS Words), üçten az cümlede yer alan kelimelerden oluşur ve bu kelimelerin geçtiği cümleler baskın olmayan cümleler havuzuna (Non-DS Sentences Pool) eklenir.

Adım (4) Çizge oluşturulması (Algoritma 1). Her baskın kelime için, bu kelimenin geçtiği cümleler arasında bir çizge oluşturulur. Cümleler, çizgenin düğümleri olarak modellenir ve cümleler arasındaki semantik benzerlik, kelime köklerine dayalı kosinüs benzerliği ile hesaplanır. Hesaplama için Word2Vec kullanılarak her kelime 100 boyutlu bir vektöre dönüştürülür. Eğer iki cümle arasındaki kosinüs benzerliği, önceden belirlenmiş bir eşik değerden (Similarity Value - SV)¹ büyük veya eşitse, aralarında bir kenar oluşturulur.

Adım (5) Önerilen Algoritma 2 uygulanması. Çizgelerdeki baskın düğümleri (dominant nodes) tespit etmek için önerilen Dominating Set algoritması çalıştırılır. Bu algoritma, grafin yapısına dayalı olarak baskın cümleleri seçer. Baskın düğümler, çizgede diğer düğümlerle en güçlü ilişkiye sahip olan ve özetleme için en anlamlı cümleleri temsil eder.

Adım (6,7) Baskın düğümlerin seçilmesi. Algoritma tarafından tespit edilen baskın düğümler, özetin aday cümleleri olarak belirlenir. Bu düğümlerin ağırlıkları, özetleme sürecinde dikkate alınır. Eğer belirli bir kelime için baskın düğüm bulunamazsa, bu cümleler baskın olmayan cümleler listesine (Non-DS Sentences Pool) eklenir.

¹ Geliştirilen yöntemde SV değeri 0.05 olarak kullanılmıştır; bu, iki cümle arasında kenar oluşturmak için minimum %5 benzerlik gerekliliğini ifade eder.

Adım (8) Özet oluşturma. Baskın düğümlerden elde edilen cümleler, belirli bir kelime sayısı kriterine göre sıralanır ve birleştirilerek aday özet oluşturulur. Eğer oluşturulan özet, istenen uzunluğun altında kalırsa, baskın olmayan cümleler (Non-DS Sentences) de sürece dahil edilerek özet tamamlanır.

Adım (9) Altın standart özet. Oluşturulan özet, veri setinde bulunan altın standart (golden summary) özet ile karşılaştırılır. Bu karşılaştırma, yöntemin performansını değerlendirmek için referans olarak kullanılır.

Adım (10) Değerlendirme. Yöntemin performansı, ROUGE gibi metrikler kullanılarak ölçülür. Bu metrikler, üretilen özete altın standart özet ile olan benzerliğini değerlendirir ve yöntemin doğruluğunu ve etkinliğini gösterir.

4.5.3 Graf oluşumu ve dominating set uygulanması

Bu bölümde, önerilen yöntemin temel yapı taşlarından olan graf oluşturma, dominating set algoritması ve özet oluşturma modeli detaylandırılmıştır. Çizgeler kullanılarak metinlerin semantik ilişkileri modellenmekte ve özetleme süreci optimize edilmektedir. Aşağıdaki algoritmalar sırasıyla graf oluşturma (Algoritma 1), dominating set belirleme (Algoritma 2) ve özet oluşturma (Algoritma 3) adımlarını açıklamaktadır.

Algoritma 1: graf oluşumuna ilişkin kaba kod

Function create_graph(S)
input: Sentence List
output: Graph

```
1   $G \leftarrow \emptyset$ 
2   $c_s \leftarrow$  similarity treshold // parametric value
3  for each  $s_1$  in  $S$ :
4    for each  $s_2$  in  $S$ :
5      if  $s_1 \neq s_2$ :
6         $c \leftarrow$  Cosinus Similarity between  $s_1$  and  $s_2$ 
7        if  $c \geq c_s$ :
8          add edge between nodes of  $s_1$  and  $s_2$  to  $G$  // with weight as rounded similarity value
9  return  $G$ 
```

Algoritma 1 adımları:

- Adım (1): Boş bir çizge oluşturulur.
- Adım (2): Cümleler arası benzerliği ölçmek için bir eşik değeri (similarity threshold) belirlenir.
- Adımlar (3-8): Her cümle, diğer cümlelerle karşılaştırılarak kosinüs benzerliği (Cosine Similarity) hesaplanır. Eğer bu değer, belirlenen eşik değerden büyük veya eşitse, cümleler arasında bir kenar oluşturulur.
- Adım (9): Çıktı olarak, cümleleri ve ilişkilerini temsil eden bir çizge döndürülür.

Dominating Set Algoritması, oluşturulan graf üzerinde baskın düğümleri (dominant nodes) belirlemek için geliştirilmiştir. Baskın düğümler, özetleme sürecinde kritik öneme sahip olan cümleleri temsil eder.

Algoritma 2: önerilen dominating set algoritması

Function dominating_set(G)

input: Graph
output: Dominating Set as Sentence List

```
1   $D \leftarrow \emptyset$ 
2   $c_s \leftarrow$  similarity treshold // parametric value
3  while  $G \neq \emptyset$ :
4     $best\_node \leftarrow \emptyset$ ,  $best\_weight \leftarrow -\infty$ ,
5    for each  $n$  in  $G$ :
6      if  $n$  has neighbors in  $D$ :
7        continue
8      if weight of  $n$  is greater than  $best\_weight$ :
9        update  $best\_node$ ,  $best\_weight$  with data of  $n$ 
10   add  $best\_node$  to  $D$ 
11   remove  $best\_node$  from  $G$ 
12   remove neighbours of  $best\_node$  from  $G$ 
13  return  $D$ 
```

Algoritma 2 adımları:

- Adımlar (1-2): Boş bir baskın düğüm kümesi (D) oluşturulur ve benzerlik eşiği belirlenir.
- Adımlar (3-12): Graf üzerinde iteratif olarak en yüksek ağırlıklı düğüm ($best_node$) bulunur ve baskın kümesine eklenir. Eklenen düğüm ve onun komşuları graftan çıkarılarak döngü devam eder.
- Adım (13): Baskın düğümleri içeren set döndürülür.

Özetleme Modeli, baskın düğümleri kullanarak özetleme işlemini gerçekleştirir. Belirli kriterlere göre en önemli cümleler seçilerek özet oluşturulur.

Algoritma 3: önerilen özetleme modeli

*input: original text, scope, summary_size**output: summary*

```
1   $T \leftarrow$  original text
2   $s \leftarrow$  scope // size of each part for the segmented original text
3   $MTS \leftarrow$  Multiple Sentence Words of  $T$  // words that appear in more than 3 sentences
4   $LTS \leftarrow$  Limited Sentence Words of  $T$  // words that appear in fewer than 3 sentences
5   $A \leftarrow \emptyset$  // sentences of summary text
6  for each  $w$  in  $MTS$ :
7       $l \leftarrow$  calculate length as ceil of sentences count divided by scope
8      for  $i$  from 0 to  $l - 1$ :
9           $S \leftarrow \emptyset$ 
10         for  $j$  from 0 to  $s - 1$ :
11             if  $(i * s) + j$  is less than sentences count:
12                 add sentence at index  $[(i * s) + j]$  to  $S$ 
13              $G \leftarrow$  create_graph( $S$ )
14              $D \leftarrow$  dominating_set( $G$ )
15             if  $D$  is empty:
16                 add  $w$  to  $LTS$ 
17             else:
18                 add nodes in  $D$  to  $A$ 
19 for each  $w$  in  $LTS$ :
20      $l \leftarrow$  calculate length as ceil of sentences count divided by scope
21     for  $i$  from 0 to  $l - 1$ :
22          $S \leftarrow \emptyset$ 
23         for  $j$  from 0 to  $s - 1$ :
24             if  $(i * s) + j$  is less than sentences count:
25                 add sentence at index  $[(i * s) + j]$  to  $S$ 
26              $G \leftarrow$  create_graph( $S$ )
27              $D \leftarrow$  dominating_set( $G$ )
28             add nodes in  $D$  to  $A$ 
29 return  $A$ 
```

Algoritma 3 adımları:

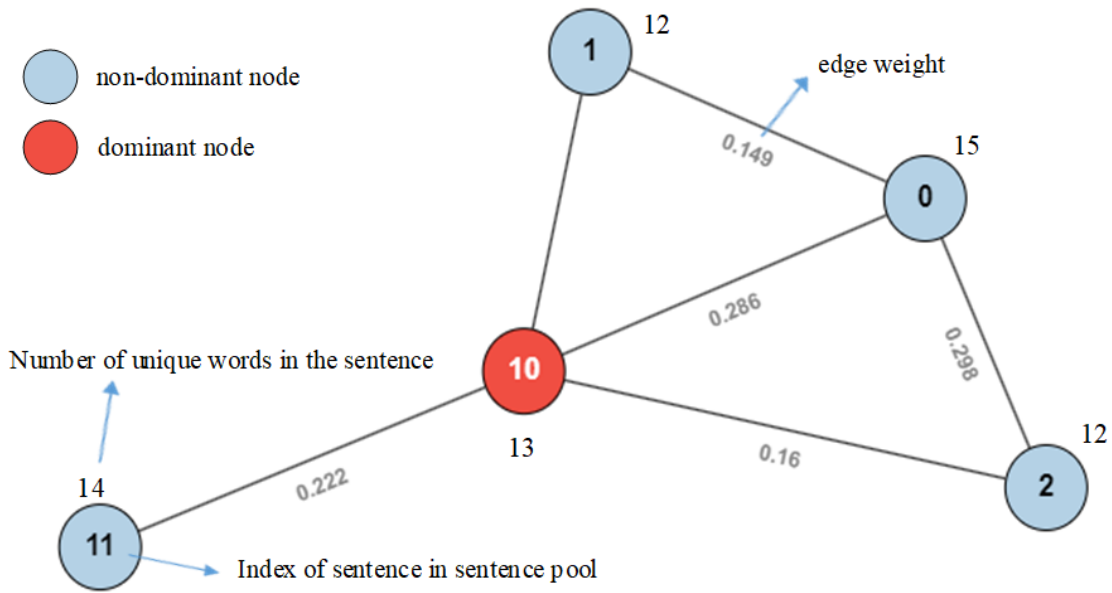
- Adımlar (1-5): Metin, baskın kelimeler (MTS) ve baskın olmayan kelimeler (LTS) olarak iki gruba ayrılır. Baskın kelimeler (MTS), metinde en az üç farklı cümlede geçen kelimelerden oluşur. Bu ayırım, metindeki kelimelerin önem derecesine göre değerlendirilmesini sağlar. Bu adımlar, özetleme işleminin temelini oluşturur ve metni baskın kelimelerle ilişkili cümleler üzerinden işlemeye hazırlar.
- Adımlar (6-18): Her bir baskın kelimenin yer aldığı cümleler, belirli bir scope boyutuna göre bölütlenir. Her bir bölüt için bir graf oluşturularak, bu graf üzerinde baskın kümeler (Dominating Set) hesaplanır. Graf üzerindeki baskın düğümler, özet cümle listesine (A) eklenir ve özeti oluşturulmasında kullanılır. Eğer baskın bir düğüm bulunamazsa, bu kelimelerin yer aldığı cümleler baskın olmayan cümleler listesine (LTS) taşınarak sonraki adımlarda değerlendirilir.
- Adımlar (19-28): Baskın olmayan cümleler için de benzer bir işlem uygulanır. Bu cümleler, belirlenen scope boyutuna göre bölütlenir ve her bir bölüt için bir graf oluşturulur. Oluşturulan grafa baskın kümeler hesaplanarak, önemli cümleler

belirlenir. Bu adımlar, özellikle uzun özetlerin oluşturulmasında kritik bir rol oynar. Bulunan baskın düğümler, özet cümle listesine eklenerek metnin bütünlüğünü koruyan daha kapsamlı bir özet elde edilir.

- Adım (29): Tüm süreç sonunda özet cümle listesini (A) döndürerek nihai özeti oluşturur.

4.5.4 Baskın düğüm seçimi ve özetleme sürecindeki örnek kullanımlar

Bu bölümde, önerilen Dominating Set algoritmasının nasıl çalıştığı ve metin özetleme sürecinde baskın düğümlerin nasıl belirlendiği, graflar üzerinden açıklanmaktadır. Çalışmada kullanılan graflar, belirli kelimeleri içeren cümlelerin ilişkilerini ve bu ilişkilerin güçlerini (kenar ağırlıkları) görselleştirerek baskın düğümlerin seçimini kolaylaştırmaktadır (Şekil 4.6 ve Şekil 4.7).



Şekil 4.6 : "REPORT" kelimesi içeren cümlelerin grafi.

Şekil 4.6'da "REPORT" kelimesini içeren cümleler arasında oluşturulan bir grafi göstermektedir. Bu grafi:

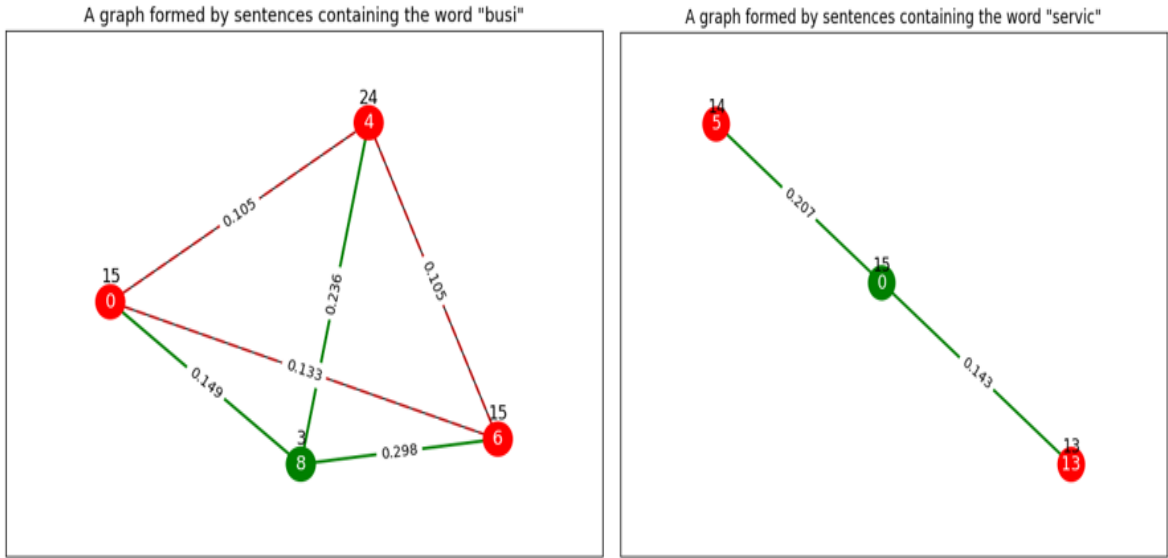
- Düğümler (Nodes): Kelimeyi içeren cümlelerdir. Örneğin, 10. indeksli düğüm, cümle havuzunda 13 benzersiz kelimeye sahip bir cümleyi temsil etmektedir ve kırmızı renkle gösterilmiştir (baskın düğüm).
- Kenarlar (Edges): Cümleler arasındaki semantik benzerlikleri ifade eder. Kenar ağırlıkları, kosinüs benzerliği ile hesaplanmıştır.

- Kırmızı Düğüm: “REPORT” kelimesi için baskın düğümü temsil eder ve metnin ana fikriyle güçlü bir bağlantıya sahiptir.
- Mavi Düğümler: Baskın olmayan cümleleri temsil eder ve özetleme sürecine dahil edilmez.

Şekil 4.7, baskın düğümlerin belirlenmesinin özetleme sürecine olan etkisini ve ana temaların nasıl temsil edildiğini görselleştirmektedir. Bu görselleştirme, farklı kelimeler için oluşturulan grafları içermektedir.

“Busi” kelimesi için oluşturulan graf, 8. indeksli düğümün baskın cümle olarak belirlendiğini göstermektedir. Ayrıca, 0, 4 ve 6 indeksli düğümler baskın olmayan cümleler olarak tanımlanmıştır.

“Service” kelimesi için oluşturulan graf ise 0. indeksli düğümün baskın cümle olarak seçildiğini göstermektedir. Bunun yanında, 5 ve 13 indeksli düğümler baskın olmayan cümleler olarak sınıflandırılmıştır.



Şekil 4.7 : farklı baskın kelimeler için oluşturulan alt graflar.

5. DENEYSEL SONUÇLAR VE DEĞERLENDİRME

Bu bölümde, çalışmamızda kullanılan extractive ve abstractive özetleme yöntemlerinin performans analizleri sunulmaktadır. Yöntemler, iki farklı içerik türünü temsil eden veri setleri üzerinde değerlendirilmiştir. BBC News Veri Seti, haber içerikli kısa metinlerden oluşurken, PubMed Veri Seti, bilimsel makale içeriğiyle daha uzun ve karmaşık metinler içermektedir. Performans ölçümleri, SacreBLEU¹, METEOR², BERTScore³ ve ROUGE⁴ metrikleri kullanılarak gerçekleştirilmiş ve Precision, Recall ve F1 Skoru ile desteklenmiştir. Ayrıca, önerilen Dominating Set Tabanlı Extractive Text Summarization (DSETS) yöntemi ile diğer extractive yöntemlerin performans karşılaştırmalarına da yer verilmiştir.

5.1 Extractive Yöntemlerin Deneysel Sonuçları

Extractive özetleme yöntemleri, belirlenen metrikler doğrultusunda her iki veri seti üzerinde detaylı bir şekilde değerlendirilmiştir. Bu yöntemler, özellikle kısa (50 kelime) ve orta uzunluktaki (100 kelime) özetlerdeki performanslarıyla dikkat çekmiştir. TextRank, LexRank, KL-SUM ve LSA gibi popüler yöntemlerin sonuçları, hem haber içerikli (BBC News) hem de bilimsel makale tabanlı (PubMed) metinlerde analiz edilmiştir. Performans ölçümleri, yöntemlerin seçilen metrikler (ROUGE, BERTScore, SacreBLEU, METEOR) bağlamında özetlerin doğruluk ve tutarlılığını ne ölçüde sağladığını ortaya koymuştur.

BBC News veri setinde extractive yöntemlerin farklı özet uzunluklarındaki (50, 100 ve 150 kelime) performans sonuçları Çizelge 5.1'de özetlenmiştir. Aşağıda her bir yöntemin dikkat çeken skorları açıklanmıştır.

¹ SacreBLEU skoru, 0 ile 100 arasında bir değer alır. Daha yüksek skor, model çıktısının referans metne daha yakın olduğunu ifade eder.

² METEOR skoru, 0 ile 1 arasında bir değer alır. Daha yüksek skor, model çıktısının referans metne daha yakın olduğunu ifade eder.

³ BERTScore, genellikle 0 ile 1 arasında bir değer alır ve daha yüksek bir skor, daha yüksek anlam benzerliği anlamına gelir.

⁴ ROUGE skoru, 0 ile 1 arasında bir değer alır. Daha yüksek skor, model çıktısının referans metne daha yakın olduğunu ifade eder.

Çizelge 5.1 : BBC News veri seti üzerinde extractive yöntemlerin sonuçları.

Kelime sayısı	Metrik	Extractive yöntemler			
		TextRank	LexRank	KL-Summ	LSA
50-kelime	SacreBleu	4.15	3.564	3.168	3.603
	Meteor	0.271	0.241	0.223	0.247
	Bert Score-Recall	0.614	0.592	0.574	0.597
	Bert Score-Precision	0.809	0.762	0.729	0.77
	Bert Score-F1 Score	0.696	0.664	0.64	0.67
	Rouge-1 Recall	0.3	0.269	0.251	0.273
	Rouge-1 Precision	0.839	0.76	0.705	0.766
	Rouge-1 F1 Score	0.425	0.383	0.356	0.387
	Rouge-2 Recall	0.257	0.211	0.183	0.219
	Rouge-2 Precision	0.725	0.6	0.512	0.617
	Rouge-2 F1 Score	0.365	0.3	0.259	0.311
	Rouge-L Recall	0.242	0.214	0.196	0.22
	Rouge-L Precision	0.676	0.602	0.551	0.613
	Rouge-L F1 Score	0.343	0.304	0.279	0.311
	Rouge-Lsum Recall	0.242	0.214	0.196	0.22
	Rouge-Lsum Precision	0.676	0.602	0.551	0.614
	Rouge-Lsum F1 Score	0.343	0.304	0.279	0.312
100-kelime	SacreBleu	22.294	18.172	16.008	18.421
	Meteor	0.466	0.396	0.365	0.409
	Bert Score-Recall	0.728	0.689	0.662	0.695
	Bert Score-Precision	0.802	0.752	0.726	0.751
	Bert Score-F1 Score	0.761	0.717	0.691	0.72
	Rouge-1 Recall	0.539	0.464	0.428	0.475
	Rouge-1 Precision	0.782	0.699	0.65	0.692
	Rouge-1 F1 Score	0.611	0.537	0.498	0.54
	Rouge-2 Recall	0.468	0.366	0.317	0.379
	Rouge-2 Precision	0.675	0.548	0.475	0.546
	Rouge-2 F1 Score	0.53	0.422	0.367	0.428
	Rouge-L Recall	0.376	0.323	0.299	0.334
	Rouge-L Precision	0.541	0.482	0.45	0.48
	Rouge-L F1 Score	0.425	0.372	0.346	0.377
	Rouge-Lsum Recall	0.377	0.323	0.299	0.334
	Rouge-Lsum Precision	0.541	0.482	0.449	0.48
	Rouge-Lsum F1 Score	0.425	0.373	0.346	0.377
150-kelime	SacreBlue	36.34	28.627	24.655	29.473
	Meteor	0.578	0.477	0.433	0.499
	Bert Score-Recall	0.789	0.735	0.7	0.746
	Bert Score-Precision	0.794	0.746	0.721	0.74
	Bert Score-F1 Score	0.789	0.739	0.709	0.741
	Rouge-1 Recall	0.693	0.578	0.523	0.604
	Rouge-1 Precision	0.74	0.665	0.616	0.644
	Rouge-1 F1 Score	0.694	0.603	0.553	0.603
	Rouge-2 Recall	0.613	0.46	0.39	0.488
	Rouge-2 Precision	0.649	0.526	0.455	0.512
	Rouge-2 F1 Score	0.611	0.478	0.411	0.484
	Rouge-L Recall	0.457	0.379	0.348	0.402
	Rouge-L Precision	0.482	0.432	0.405	0.421
	Rouge-L F1 Score	0.455	0.394	0.366	0.398
	Rouge-Lsum Recall	0.458	0.379	0.348	0.403
	Rouge-Lsum Precision	0.482	0.432	0.405	0.421
	Rouge-Lsum F1 Score	0.455	0.394	0.366	0.398

TextRank, BBC News veri setinde öne çıkan yöntem olmuştur. 50 kelimelik özetlerde, ROUGE-1 Precision(0.839) ve BERTScore-Precision(0.809) ile dikkat çeken sonuçlar elde etmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.782) ve METEOR (0.466)

değerleriyle bağlam uyumunda başarılı bir performans göstermiştir. 150 kelimelik özetlerde ise ROUGE-1 Recall (0.693) ve SacreBLEU (36.34) ile tüm yöntemler arasında en yüksek değerlere ulaşmıştır. Bu sonuçlar, TextRank'ın hem bağlam uyumu hem de metin doğruluğu açısından öne çıktığını göstermektedir.

LexRank, genel olarak dengeli bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.76) ve BERTScore-Precision (0.762) ile tatmin edici sonuçlar elde etmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.699) ve SacreBLEU (18.172) değerleriyle başarılı olmuştur. 150 kelimelik özetlerde ise ROUGE-1 Precision (0.665) ve BERTScore-F1 (0.739) ile özellikle uzun özetlerde bağlam uyumunda dengeli bir performans göstermiştir. Ancak genel olarak TextRank'ın gerisinde kalmıştır.

KL-SUM, özellikle kısa özetlerde düşük performans göstermiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.705) ve BERTScore-Precision (0.729) ile sınırlı başarı göstermiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.650) ve SacreBLEU (16.008) değerleriyle dikkat çekmiştir. 150 kelimelik özetlerde ise ROUGE-1 Recall (0.523) ve BERTScore-Precision (0.721) değerleri öne çıkmıştır. Bu sonuçlar, KL-SUM'un bağlam uyumunu sağlamakta zorlandığını göstermektedir.

LSA, LexRank'a yakın bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.766) ve BERTScore-Precision (0.770) ile başarılı sonuçlar elde etmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.692) ve SacreBLEU (18.421) değerleriyle dikkat çekmiştir. 150 kelimelik özetlerde ise ROUGE-1 Recall (0.604) ve BERTScore-F1 (0.741) değerleri ile bağlam uyumunda tatmin edici sonuçlar sunmuştur. Ancak genel olarak TextRank'ın gerisinde kalmıştır.

PubMed veri setinde extractive yöntemlerin farklı özet uzunluklarındaki (50, 100 ve 150 kelime) performans sonuçları Çizelge 5.2'de özetlenmiştir. Aşağıda her bir yöntemin dikkat çeken skorları açıklanmıştır.

Çizelge 5.2 : PubMed veri seti üzerinde extractive yöntemlerin sonuçları.

Kelime sayısı	Metrik	Extractive yöntemler			
		TextRank	LexRank	KL-Summ	LSA
50-kelime	SacreBleu	4.359	4.197	3.819	4.167
	Meteor	0.171	0.167	0.155	0.166
	Bert Score-Recall	0.546	0.542	0.526	0.538
	Bert Score-Precision	0.656	0.647	0.628	0.642
	Bert Score-F1 Score	0.595	0.588	0.571	0.584
	Rouge-1 Recall	0.2	0.194	0.18	0.192
	Rouge-1 Precision	0.525	0.511	0.478	0.504
	Rouge-1 F1 Score	0.284	0.277	0.257	0.273
	Rouge-2 Recall	0.093	0.088	0.077	0.086
	Rouge-2 Precision	0.253	0.238	0.213	0.234
	Rouge-2 F1 Score	0.134	0.126	0.111	0.124
	Rouge-L Recall	0.145	0.14	0.129	0.139
	Rouge-L Precision	0.383	0.37	0.346	0.367
	Rouge-L F1 Score	0.206	0.199	0.185	0.198
	Rouge-Lsum Recall	0.174	0.169	0.157	0.167
	Rouge-Lsum Precision	0.461	0.447	0.419	0.442
	Rouge-Lsum F1 Score	0.249	0.241	0.224	0.238
100-kelime	SacreBleu	14.476	14.02	12.162	13.507
	Meteor	0.277	0.271	0.248	0.266
	Bert Score-Recall	0.616	0.613	0.591	0.608
	Bert Score-Precision	0.644	0.635	0.616	0.628
	Bert Score-F1 Score	0.629	0.623	0.603	0.617
	Rouge-1 Recall	0.341	0.333	0.306	0.327
	Rouge-1 Precision	0.461	0.45	0.417	0.441
	Rouge-1 F1 Score	0.384	0.375	0.346	0.368
	Rouge-2 Recall	0.157	0.149	0.127	0.142
	Rouge-2 Precision	0.217	0.206	0.177	0.198
	Rouge-2 F1 Score	0.178	0.169	0.144	0.162
	Rouge-L Recall	0.224	0.216	0.197	0.212
	Rouge-L Precision	0.304	0.294	0.27	0.288
	Rouge-L F1 Score	0.252	0.244	0.223	0.239
	Rouge-Lsum Recall	0.288	0.28	0.257	0.275
	Rouge-Lsum Precision	0.392	0.381	0.353	0.373
	Rouge-Lsum F1 Score	0.325	0.316	0.291	0.31
150-kelime	SacreBleu	18.14	17.391	15.01	16.909
	Meteor	0.333	0.326	0.298	0.321
	Bert Score-Recall	0.65	0.648	0.623	0.644
	Bert Score-Precision	0.629	0.62	0.603	0.615
	Bert Score-F1 Score	0.638	0.633	0.612	0.628
	Rouge-1 Recall	0.438	0.427	0.391	0.42
	Rouge-1 Precision	0.402	0.393	0.363	0.386
	Rouge-1 F1 Score	0.409	0.4	0.368	0.393
	Rouge-2 Recall	0.2	0.189	0.159	0.182
	Rouge-2 Precision	0.187	0.177	0.151	0.171
	Rouge-2 F1 Score	0.188	0.178	0.151	0.172
	Rouge-L Recall	0.275	0.265	0.241	0.26
	Rouge-L Precision	0.253	0.244	0.224	0.24
	Rouge-L F1 Score	0.258	0.248	0.227	0.244
	Rouge-Lsum Recall	0.363	0.352	0.323	0.346
	Rouge-Lsum Precision	0.336	0.327	0.302	0.321
	Rouge-Lsum F1 Score	0.341	0.332	0.305	0.325

TextRank, PubMed veri setinde en iyi performans gösteren yöntem olmuştur. 50 kelimelik özetlerde, ROUGE-1 Precision (0.525) ve BERTScore-Precision (0.656) ile en yüksek değerlere ulaşmıştır. 100 kelimelik özetlerde, ROUGE-1 Precision (0.461) ve SacreBLEU

(14.476) ile öne çıkmıştır. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.438) ve Meteor (0.333) değerleriyle bağlam uyumunda üstünlük sağlamıştır. Bu sonuçlar, TextRank'ın genel doğruluk ve bağlam uyumunda lider olduğunu göstermektedir.

LexRank, dengeli bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.511) ve BERTScore-Precision (0.647) değerleriyle dikkat çekmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.450) ve Meteor (0.271) ile makul bir başarı göstermiştir. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.427) ve BERTScore-Recall (0.648) ile özellikle uzun özetlerde iyi sonuçlar vermiştir. Ancak genel olarak TextRank'ın gerisinde kalmıştır.

KL-SUM, tüm özet uzunluklarında en düşük performansı göstermiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.478) ve BERTScore-Precision (0.628) ile sınırlı bir başarı göstermiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.417) ve SacreBLEU (12.162) ile daha düşük bir bağlam uyumu sağlamıştır. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.391) ve Meteor (0.298) ile diğer yöntemlerin gerisinde kalmıştır.

LSA, LexRank'a yakın bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.504) ve SacreBLEU (4.167) değerleriyle dikkat çekmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.441) ve SacreBLEU (13.507) ile bağlam uyumunda dengeli bir sonuç elde etmiştir. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.420) ve BERTScore-F1 (0.628) değerleri ile tatmin edici bir performans göstermiştir. Ancak genel olarak TextRank'ın gerisinde kalmıştır.

5.2 Abstractive Modellerin Deneysel Sonuçları

Bu bölümde, abstractive özetleme modellerinin performans analizleri sunulmaktadır. Analizler, farklı içerik türlerini temsil eden iki veri seti üzerinde gerçekleştirilmiştir. BBC News Veri Seti (haber içerikli) ve PubMed Veri Seti (bilimsel makale içerikli). BBC News veri setinde PEGASUS, BART ve T5-LARGE modelleri kullanılmıştır. PubMed veri setinde ise, metinlerin uzun ve karmaşık yapısını daha iyi ele almak için PEGASUS, BART ve LED modelleri değerlendirilmiştir.

Her iki veri setinde de özetleme modelleri, 50, 100 ve 150 kelimelik özetler oluşturularak analiz edilmiştir. Modellerin performansları, ROUGE, SacreBLEU, METEOR ve BERTScore metrikleri üzerinden değerlendirilmiş, ROUGE ve BERTScore için ayrıca Precision, Recall ve F1 Skoru hesaplanmıştır.

BBC News veri setinde abstractive modellerin farklı özet uzunluklarındaki (50, 100 ve 150 kelime) performans sonuçları Çizelge 5.3'te özetlenmiştir. Aşağıda her bir yöntemin dikkat çeken skorları açıklanmıştır.

Çizelge 5.3 : BBC News veri seti üzerinde abstractive modellerin sonuçları.

Kelime sayısı	Metrik	Abstractive modeller		
		Pegasus	BART	T5-Large
50-kelime	SacreBleu	3.721	3.695	3.692
	Meteor	0.218	0.21	0.224
	Bert Score-Recall	0.579	0.574	0.582
	Bert Score-Precision	0.744	0.728	0.743
	Bert Score-F1 Score	0.65	0.64	0.651
	Rouge-1 Recall	0.248	0.241	0.257
	Rouge-1 Precision	0.72	0.698	0.719
	Rouge-1 F1 Score	0.356	0.346	0.365
	Rouge-2 Recall	0.175	0.163	0.179
	Rouge-2 Precision	0.507	0.473	0.501
	Rouge-2 F1 Score	0.251	0.233	0.254
	Rouge-L Recall	0.185	0.178	0.187
	Rouge-L Precision	0.535	0.514	0.521
	Rouge-L F1 Score	0.265	0.254	0.265
	Rouge-Lsum Recall	0.185	0.178	0.187
	Rouge-Lsum Precision	0.535	0.514	0.52
Rouge-Lsum F1 Score	0.265	0.254	0.265	
100-kelime	SacreBleu	19.747	17.83	11.666
	Meteor	0.354	0.334	0.301
	Bert Score-Recall	0.66	0.65	0.628
	Bert Score-Precision	0.716	0.709	0.694
	Bert Score-F1 Score	0.685	0.677	0.657
	Rouge-1 Recall	0.427	0.403	0.366
	Rouge-1 Precision	0.638	0.636	0.643
	Rouge-1 F1 Score	0.489	0.472	0.446
	Rouge-2 Recall	0.302	0.274	0.243
	Rouge-2 Precision	0.441	0.43	0.418
	Rouge-2 F1 Score	0.342	0.32	0.294
	Rouge-L Recall	0.28	0.265	0.241
	Rouge-L Precision	0.413	0.416	0.42
	Rouge-L F1 Score	0.318	0.31	0.293
	Rouge-Lsum Recall	0.28	0.265	0.242
	Rouge-Lsum Precision	0.413	0.416	0.42
Rouge-Lsum F1 Score	0.318	0.31	0.293	
150-kelime	SacreBleu	29.938	29.302	
	Meteor	0.415	0.408	
	Bert Score-Recall	0.691	0.693	
	Bert Score-Precision	0.691	0.688	
	Bert Score-F1 Score	0.688	0.689	
	Rouge-1 Recall	0.526	0.521	
	Rouge-1 Precision	0.564	0.575	
	Rouge-1 F1 Score	0.519	0.522	
	Rouge-2 Recall	0.373	0.359	
	Rouge-2 Precision	0.387	0.389	
	Rouge-2 F1 Score	0.362	0.356	
	Rouge-L Recall	0.333	0.321	
	Rouge-L Precision	0.349	0.349	
	Rouge-L F1 Score	0.324	0.319	
	Rouge-Lsum Recall	0.333	0.321	
	Rouge-Lsum Precision	0.348	0.349	
Rouge-Lsum F1 Score	0.324	0.319		

PEGASUS, BBC News veri setinde öne çıkan yöntem olmuştur. 50 kelimelik özetlerde, ROUGE-1 Precision (0.720) ve BERTScore-Precision (0.744) ile güçlü bir performans göstermiştir. 100 kelimelik özetlerde, ROUGE-1 Recall (0.427) ve SacreBLEU (19.747) değerleriyle bağlam uyumu açısından en başarılı yöntem olmuştur. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.526) ve Meteor (0.415) değerleriyle öne çıkmıştır. Pegasus, kısa ve orta uzunluklardaki özetlerde bağlam uyumu ve doğruluk açısından liderliğini sürdürmüştür.

BART, Pegasus'un ardından dengeli bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.698) ve Meteor (0.210) ile makul bir başarı elde etmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.636) ve BERTScore-F1 (0.677) değerleriyle dikkat çekmiştir. 150 kelimelik özetlerde ise, ROUGE-1 Precision (0.575) ve BERTScore-Recall (0.693) ile özellikle uzun özetlerde bağlam uyumunda güçlü bir alternatif sunmuştur.

T5-Large, birçok metrikte yüksek değerler sunarak dikkat çekmiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.719) ve BERTScore-Recall (0.582) ile güçlü bir bağlam uyumu göstermiştir. Ayrıca, ROUGE-2 F1 Score (0.254) ile diğer modellere kıyasla kısa özetlerde iyi bir performans sergilemiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.643) ve BERTScore-F1 (0.657) değerleriyle bağlam uyumu ve doğruluk açısından tatmin edici sonuçlar elde etmiştir.

PubMed veri setinde abstractive modellerin farklı özet uzunluklarındaki (50, 100 ve 150 kelime) performans sonuçları Çizelge 5.4'te özetlenmiştir. Aşağıda her bir yöntemin dikkat çeken skorları açıklanmıştır.

Çizelge 5.4 : PubMed veri seti üzerinde abstractive modellerin sonuçları.

Kelime sayısı	Metrik	Abstractive modeller		
		Pegasus	BART	LED
50-kelime	SacreBleu	4.374	4.439	5.648
	Meteor	0.176	0.176	0.185
	Bert Score-Recall	0.557	0.558	0.558
	Bert Score-Precision	0.667	0.658	0.667
	Bert Score-F1 Score	0.606	0.603	0.606
	Rouge-1 Recall	0.217	0.217	0.225
	Rouge-1 Precision	0.532	0.516	0.54
	Rouge-1 F1 Score	0.303	0.299	0.31
	Rouge-2 Recall	0.098	0.095	0.107
	Rouge-2 Precision	0.25	0.234	0.267
	Rouge-2 F1 Score	0.138	0.132	0.15
	Rouge-L Recall	0.155	0.153	0.16
	Rouge-L Precision	0.384	0.366	0.39
	Rouge-L F1 Score	0.217	0.211	0.223
	Rouge-Lsum Recall	0.189	0.186	0.204
Rouge-Lsum Precision	0.464	0.447	0.493	
Rouge-Lsum F1 Score	0.263	0.258	0.283	
100-kelime	SacreBleu	12.039	11.638	14.89
	Meteor	0.266	0.269	0.278
	Bert Score-Recall	0.616	0.621	0.618
	Bert Score-Precision	0.647	0.639	0.648
	Bert Score-F1 Score	0.63	0.629	0.631
	Rouge-1 Recall	0.343	0.345	0.348
	Rouge-1 Precision	0.458	0.447	0.464
	Rouge-1 F1 Score	0.383	0.38	0.387
	Rouge-2 Recall	0.146	0.142	0.161
	Rouge-2 Precision	0.202	0.189	0.221
	Rouge-2 F1 Score	0.165	0.158	0.182
	Rouge-L Recall	0.22	0.22	0.229
	Rouge-L Precision	0.298	0.287	0.309
	Rouge-L F1 Score	0.247	0.243	0.256
	Rouge-Lsum Recall	0.285	0.284	0.316
Rouge-Lsum Precision	0.385	0.371	0.423	
Rouge-Lsum F1 Score	0.32	0.314	0.352	
150-kelime	SacreBleu	15.375	14.224	18.788
	Meteor	0.31	0.312	0.333
	Bert Score-Recall	0.641	0.648	0.65
	Bert Score-Precision	0.631	0.623	0.636
	Bert Score-F1 Score	0.635	0.634	0.642
	Rouge-1 Recall	0.42	0.423	0.435
	Rouge-1 Precision	0.402	0.393	0.412
	Rouge-1 F1 Score	0.401	0.398	0.411
	Rouge-2 Recall	0.174	0.17	0.201
	Rouge-2 Precision	0.171	0.161	0.195
	Rouge-2 F1 Score	0.168	0.161	0.192
	Rouge-L Recall	0.255	0.255	0.276
	Rouge-L Precision	0.246	0.237	0.264
	Rouge-L F1 Score	0.244	0.24	0.262
	Rouge-Lsum Recall	0.343	0.341	0.396
Rouge-Lsum Precision	0.331	0.319	0.376	
Rouge-Lsum F1 Score	0.329	0.322	0.375	

PEGASUS, PubMed veri setinde kısa ve orta uzunluktaki özetlerde güçlü bir performans sergilemiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.532) ve BERTScore-Precision (0.667) ile öne çıkmıştır. 100 kelimelik özetlerde, ROUGE-1 Precision (0.458) ve SacreBLEU (12.039) değerleriyle dikkat çekmiştir. 150 kelimelik özetlerde ise, ROUGE-1

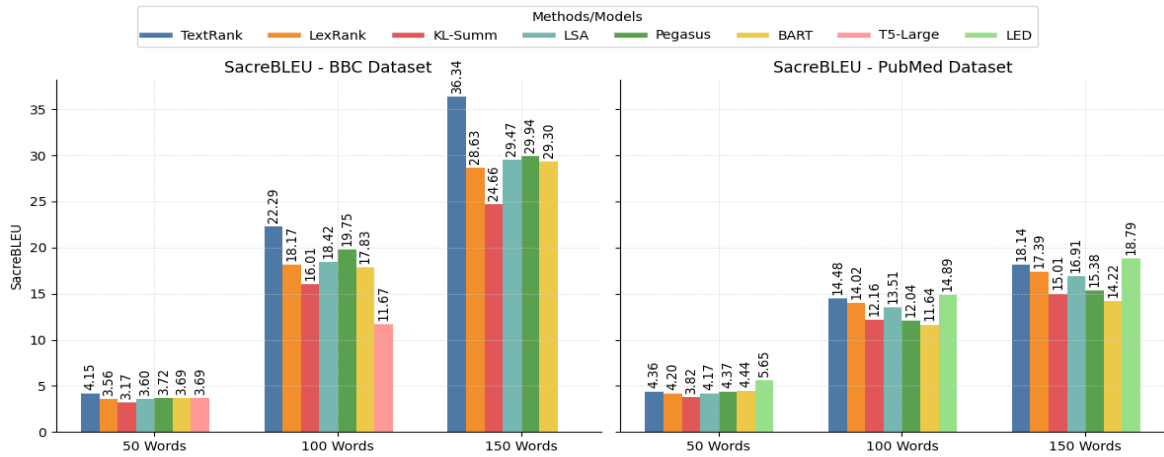
Recall (0.420) ve Meteor (0.310) değerleriyle tatmin edici sonuçlar elde etmiştir. Pegasus, özellikle bağlam uyumunda dengeli bir performans sunmuştur.

BART, kısa özetlerde Pegasus ile benzer sonuçlar elde ederken, uzun özetlerde görece daha düşük bir performans göstermiştir. 50 kelimelik özetlerde, ROUGE-1 Precision (0.516) ve Meteor (0.176) değerleriyle dikkat çekmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.447) ve BERTScore-Recall (0.621) ile dengeli bir performans sergilemiştir. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.423) ve BERTScore-F1 (0.634) ile orta düzeyde bir başarı sağlamıştır.

LED, PubMed veri setinde özellikle uzun özetlerde en başarılı yöntem olmuştur. 50 kelimelik özetlerde, ROUGE-1 Precision (0.540) ve SacreBLEU (5.648) ile dikkat çekmiştir. 100 kelimelik özetlerde, ROUGE-1 Precision (0.464) ve Meteor (0.278) ile diğer yöntemlerin önüne geçmiştir. 150 kelimelik özetlerde ise, ROUGE-1 Recall (0.435) ve SacreBLEU (18.788) değerleriyle üstün bir performans göstermiştir. LED, uzun metinlerin bağlam uyumu ve doğruluğu açısından en etkili yöntemdir.

5.3 Yöntemlerin Karşılaştırmalı Analizi

Şekil 5.1, yöntemlerin ve modellerin BBC ve PubMed veri setleri üzerinde SacreBLEU metriğine göre 50, 100 ve 150 kelimelik özetler için skorlarını göstermektedir.



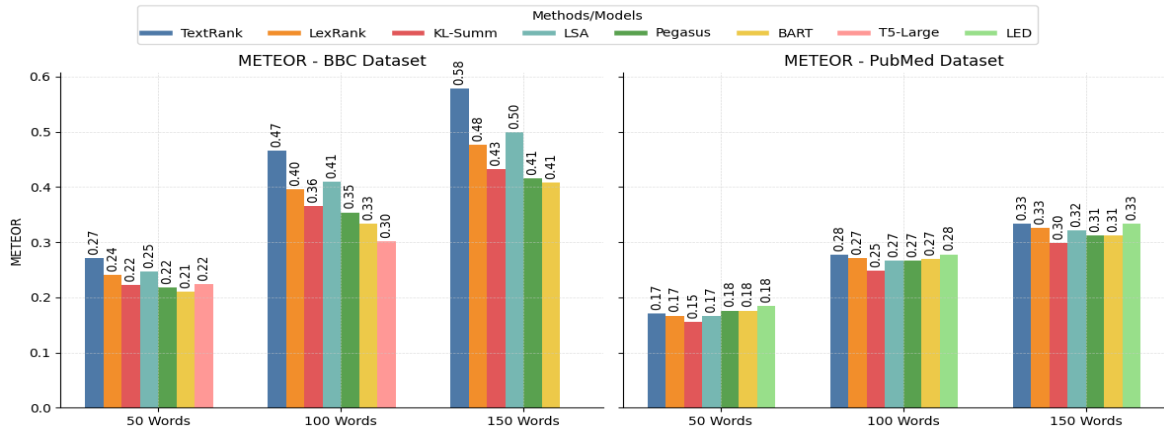
Şekil 5.1 : BBC ve PubMed veri setlerindeki SacreBLEU sonuçları.

BBC News veri setinde yapılan analizlerde, TextRank yöntemi tüm özet uzunluklarında en yüksek SacreBLEU skorlarına ulaşmıştır. Kısa özetlerde TextRank (4.15) lider olurken, Pegasus (3.72) abstractive modeller arasında dikkat çeken bir performans sergilemiştir. Orta uzunlukta özetlerde (100 kelime), TextRank (22.29) açık ara üstünlük sağlamış, Pegasus

(19.75) öne çıkan model olmuştur. Uzun özetlerde (150 kelime), TextRank (36.34) ile üstünlüğünü daha da artırmış, Pegasus (29.94) abstractive modellerde liderliğe devam etmiştir.

PubMed veri setine bakıldığında, kısa özetlerde LED (5.65) en yüksek SacreBLEU skoruna ulaşmış, TextRank (4.36) extractive yöntemler arasında liderliği elde etmiştir. Orta uzunlukta özetlerde LED (14.89), abstractive modellerde en iyi performansı sergilemiş, TextRank (14.48) extractive yöntemlerde öne çıkmıştır. Uzun özetlerde LED (18.79) abstractive modellerde liderliği ele alırken, TextRank (18.14) extractive yöntemlerde en yüksek performansı göstermiştir. Pegasus (18.14) ise abstractive modeller arasında LED'in ardından gelen güçlü bir model olmuştur.

Şekil 5.2, yöntemlerin ve modellerin BBC ve PubMed veri setleri üzerinde METEOR metriğine göre 50, 100 ve 150 kelimelik özetler için elde ettiği skorları göstermektedir.

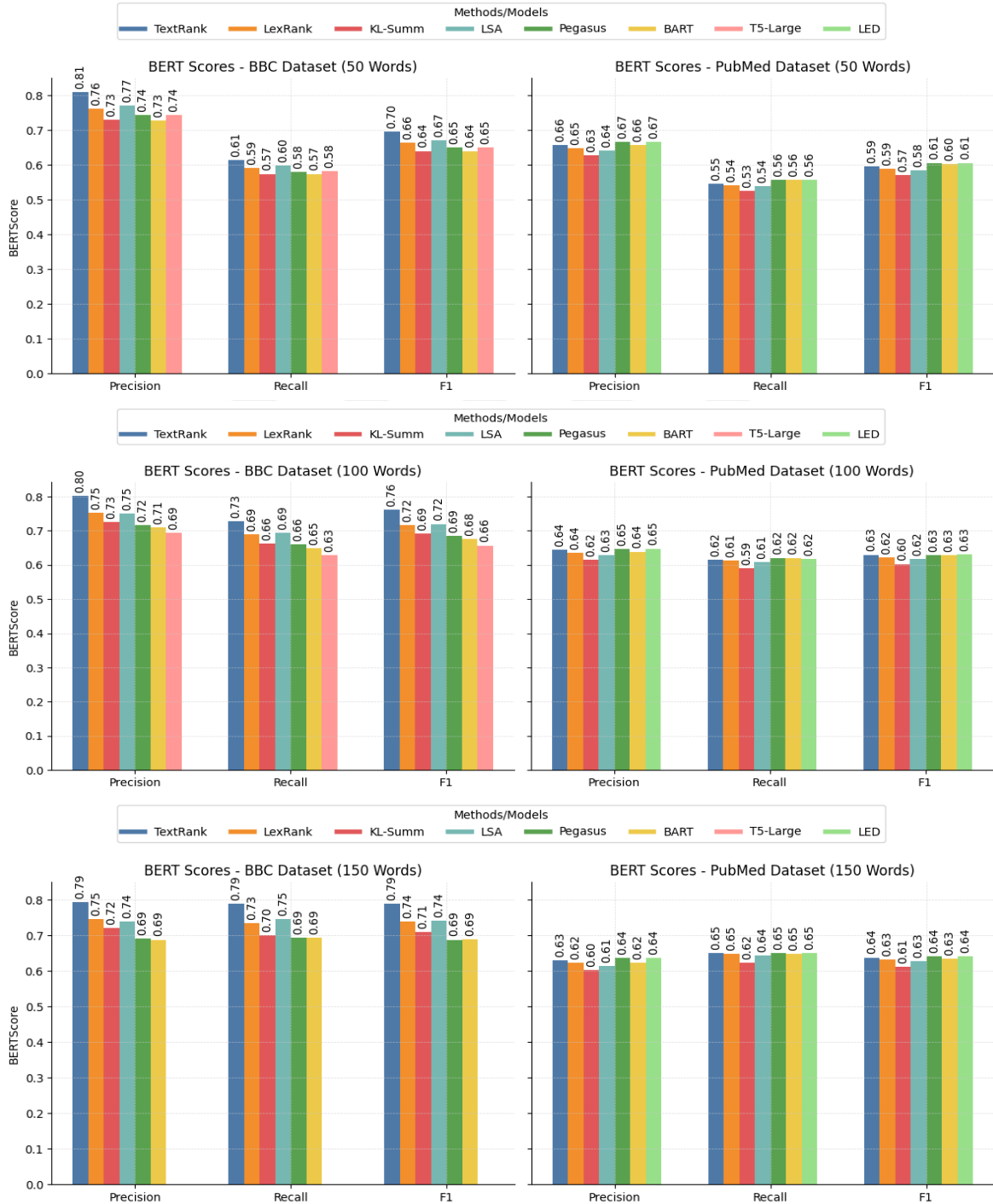


Şekil 5.2 : BBC ve PubMed veri setlerindeki Meteor sonuçları.

BBC News veri setinde yapılan analizlerde, TextRank yöntemi tüm özet uzunluklarında en yüksek METEOR skorlarına ulaşmıştır. Kısa özetlerde TextRank (0.27) lider olurken, LSA (0.25) ikinci sırada yer almış ve extractive yöntemlerde bağlam uyumunu ortaya koymuştur. Abstractive modeller arasında ise T5-Large (0.22), kısa özetlerde en iyi performansı göstermiştir. Orta uzunlukta özetlerde (100 kelime), TextRank (0.47) üstünlüğünü korurken, Pegasus (0.35) öne çıkan model olmuştur. Uzun özetlerde (150 kelime), TextRank (0.58) liderliğini pekiştirirken, LSA (0.50) extractive yöntemlerde ikinci sırayı almıştır. Abstractive modellerde Pegasus ve BART, 0.41 skoruyla uzun özetlerde en yüksek değeri elde etmiştir. PubMed veri setine bakıldığında, Kısa özetlerde Pegasus, BART ve LED, (0.18) METEOR skoruna ulaşarak aynı performansı göstermiş, TextRank ise extractive yöntemler arasında (0.17) skoruyla lider olmuştur. Orta uzunlukta özetlerde LED (0.28), abstractive modellerde

en iyi performansı sergilemiş, TextRank (0.28) extractive yöntemlerde öne çıkmıştır. Uzun özetlerde LED (0.33) ve TextRank (0.33) aynı skoru paylaşarak lider olmuş, Pegasus ve BART (0.31) abstractive modeller arasında dengeli bir performans göstermiştir.

Şekil 5.3, BBC ve PubMed veri setlerinde 50, 100 ve 150 kelimelik özetler için kullanılan yöntem ve modellerin BERTScore metriğine göre Precision, Recall ve F1 değerlerini karşılaştırmaktadır.

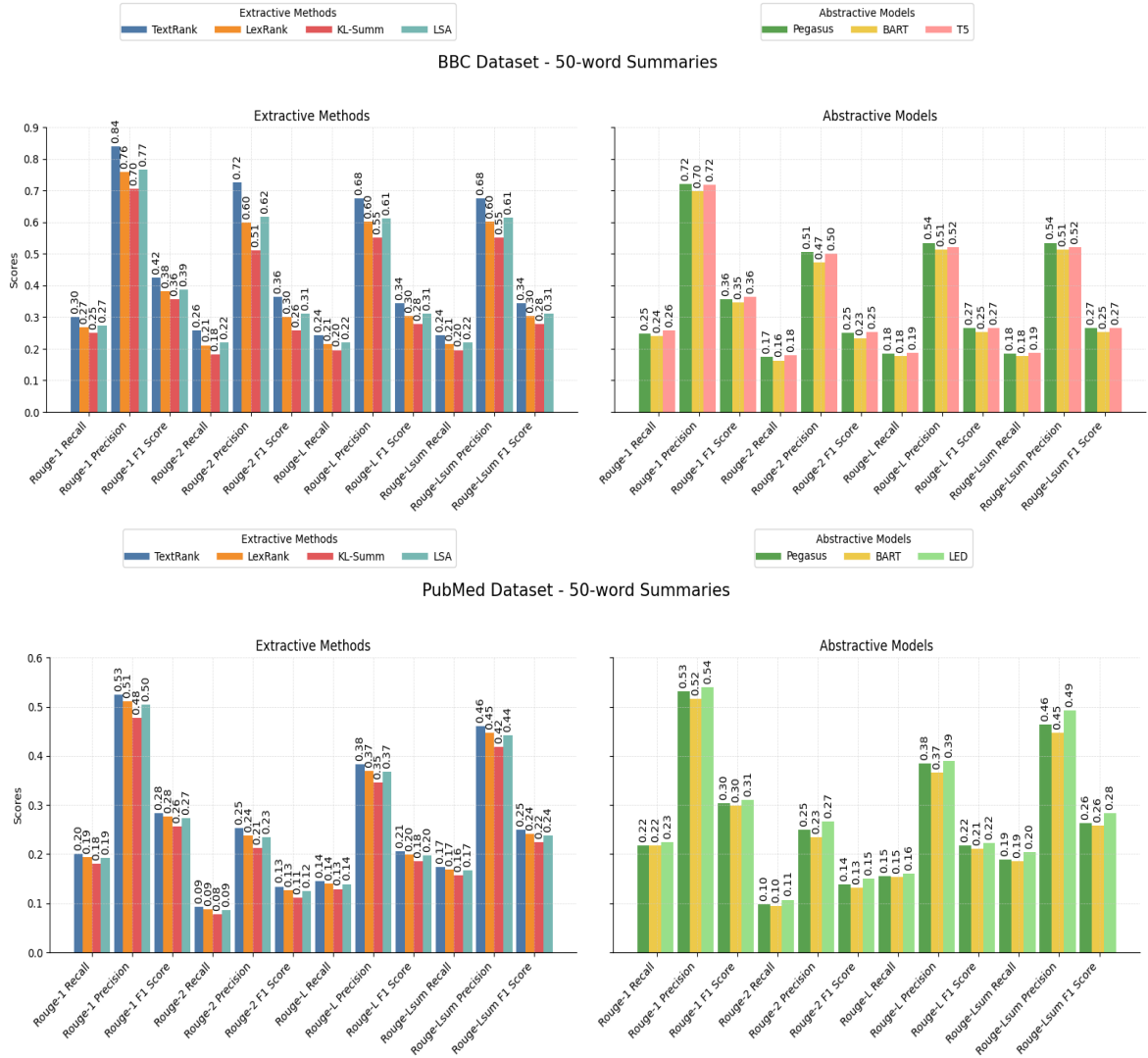


Şekil 5.3 : BBC ve PubMed veri setlerindeki BertScore sonuçları.

BBC-News veri setinde yapılan analizlerde, kısa özetlerde (50 kelime), extractive yöntemlerden TextRank, Precision (0.81), Recall (0.61) ve F1 (0.7) ile tüm metriklerde liderdir. Abstractive modellerde ise Pegasus ve T5, Precision (0.74), Recall (0.58) ve F1 (0.65) değerleriyle liderliği paylaşmıştır. Orta uzunluktaki özetlerde (100 kelime), extractive yöntemlerden yine TextRank, Precision (0.80), Recall (0.73) ve F1 (0.76) ile öne çıkmıştır. Abstractive modellerde ise Pegasus, Precision (0.72), Recall (0.66) ve F1 (0.69) ile liderdir. Uzun özetlerde (150 kelime), extractive yöntemlerde TextRank, Precision (0.79), Recall (0.79) ve F1 (0.79) ile en iyi performansı göstermiştir. Abstractive modellerde ise Pegasus ve BART, Precision (0.69), Recall (0.69) ve F1 (0.69) değerlerinde eşit performans göstermiştir.

PubMed veri setinde yapılan analizlerde, kısa özetlerde (50 kelime), Precision'da Pegasus ve LED (0.67) lider, Recall'da tüm modeller aynı (0.56), F1'de ise Pegasus ve LED (0.61) öndedir. Orta özetlerde (100 kelime), Precision'da yine Pegasus ve LED (0.65) lider, Recall'da tüm modeller aynı (0.62), F1'de ise tüm modeller (0.63) eşit performans göstermiştir. Uzun özetlerde (150 kelime), Precision'da Pegasus ve LED (0.64) lider, Recall'da tüm modeller aynı (0.65), F1'de ise Pegasus ve LED (0.64) ile öndedir.

Şekil 5.4, 50 kelimelik özetler için kullanılan extractive ve abstractive yöntemlerin BBC ve PubMed veri setlerinde elde ettiği ROUGE metriğine dayalı sonuçları göstermektedir. Hem ROUGE-1, ROUGE-2, ROUGE-L hem de ROUGE-Lsum metrikleri açısından Precision, Recall ve F1 değerleri karşılaştırılmıştır.



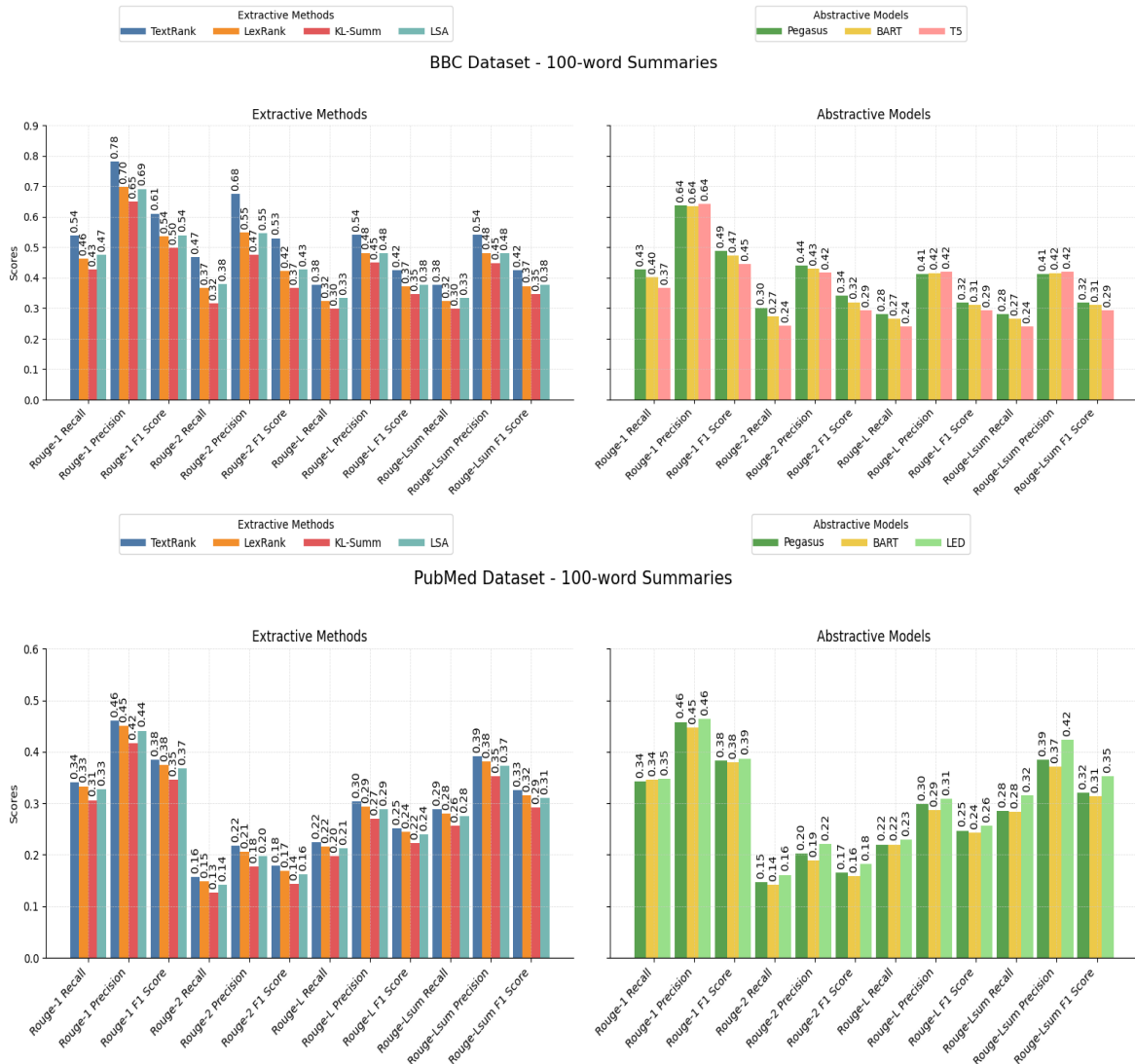
Şekil 5.4 : BBC ve PubMed veri setlerindeki Rouge sonuçları (50-kelime).

BBC NEWS veri setinde, Precision metriklerinde extractive yöntemler arasında TextRank, ROUGE-1 (0.84), ROUGE-2 (0.72) ve ROUGE-L (0.68) ile en iyi performansı göstermiştir. Abstractive modeller arasında ise Pegasus, ROUGE-1 (0.72) ve ROUGE-2 (0.51) ile liderdir. Recall metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.3), ROUGE-2 (0.26) ve ROUGE-L (0.24) ile öne çıkmıştır. Abstractive modellerde ise T5, ROUGE-1 (0.26) ve ROUGE-2 (0.18) ile en iyi sonuçları elde etmiştir. F1 Score açısından extractive yöntemlerde TextRank, ROUGE-1 (0.42) ile liderdir. Abstractive modeller arasında ise hem Pegasus hem de T5, ROUGE-1 (0.36) ile yakın değerlere sahiptir.

PubMed veri setinde, Precision metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.53) ve ROUGE-Lsum (0.46) değerleriyle en iyi sonucu verirken, abstractive modellerde LED, ROUGE-1 (0.54) ve ROUGE-Lsum (0.49) ile liderdir. Recall metriklerinde extractive yöntemlerden yine TextRank, ROUGE-1 (0.2) ve ROUGE-Lsum (0.17) ile öne çıkmıştır.

Abstractive modeller arasında ise LED, ROUGE-1 (0.23) ve ROUGE-Lsum (0.2) değerleriyle bağlam uyumu açısından en yüksek sonuçları elde etmiştir. F1 Score metriklerinde extractive yöntemlerde TextRank, ROUGE-1 (0.28) ile liderken, abstractive modellerde LED, ROUGE-1 (0.31) ve ROUGE-Lsum (0.28) ile öne çıkmıştır.

Şekil 5.5, 100 kelimelik özetler için kullanılan extractive ve abstractive yöntemlerin BBC ve PubMed veri setlerinde elde ettiği ROUGE metriğine dayalı sonuçları göstermektedir. Bu analizde ROUGE-1, ROUGE-2, ROUGE-L ve ROUGE-Lsum metrikleri açısından Precision, Recall ve F1 skorları değerlendirilmiştir.



Şekil 5.5 : BBC ve PubMed veri setlerindeki Rouge sonuçları (100-kelimelik).

BBC veri setinde, Precision metriklerinde extractive yöntemler arasında TextRank, ROUGE-1 (0.78), ROUGE-2 (0.68) ve ROUGE-L (0.54) ile en iyi performansı göstermiştir. Abstractive modellerde, ROUGE-1 değerinde hem Pegasus hem de T5 aynı sonucu (0.64)

verirken, ROUGE-2’de Pegasus (0.44) daha iyi performans göstermiş, ROUGE-L’de ise T5 (0.42) daha iyi sonuç vermiştir. Recall metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.54), ROUGE-2 (0.47) ve ROUGE-L (0.38) ile öne çıkmıştır. Abstractive modellerde ise Pegasus, ROUGE-1 (0.43) ve ROUGE-2 (0.3) değerleriyle en iyi sonuçları almıştır. F1 Score açısından extractive yöntemlerde TextRank, ROUGE-1 (0.61) ile liderdir. Abstractive modellerde ise Pegasus, F1 metriklerinde ROUGE-1 (0.49) ve ROUGE-2 (0.34) ile dikkat çekmiştir.

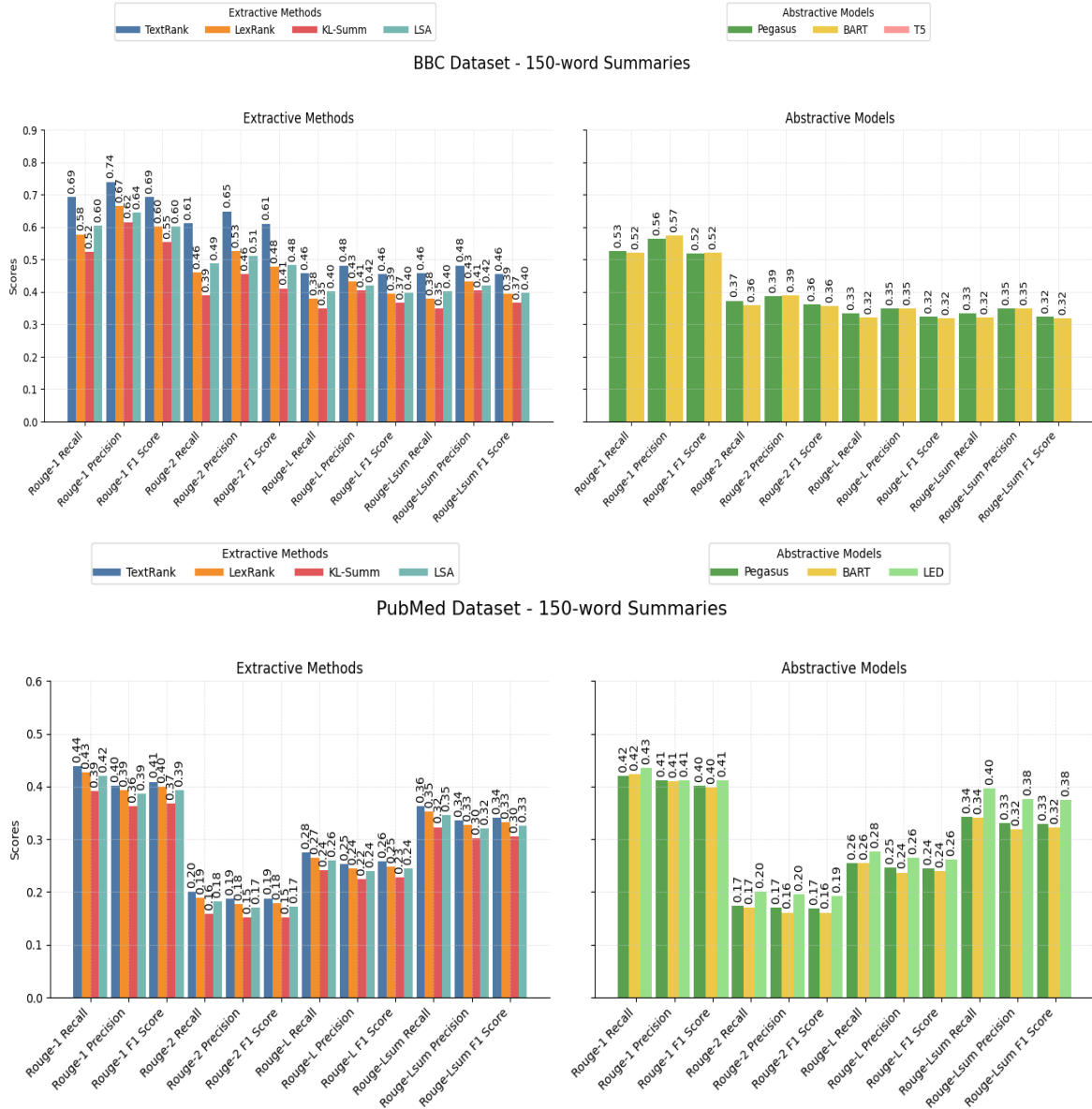
PubMed veri setinde, Precision metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.46) ve ROUGE-Lsum (0.39) ile en iyi sonuçları almıştır. Abstractive modellerde ise LED, ROUGE-1 (0.46) ve ROUGE-Lsum (0.42) ile liderdir. Recall metriklerinde extractive yöntemlerden yine TextRank, ROUGE-1 (0.34) ve ROUGE-Lsum (0.29) ile en yüksek değerlere sahiptir. Abstractive modellerde ise LED, ROUGE-1 (0.35) ve ROUGE-Lsum (0.32) değerleriyle öne çıkmıştır. F1 Score metriklerinde extractive yöntemlerde TextRank ve LexRank, ROUGE-1 (0.38) ile liderken, abstractive modellerde ise LED, ROUGE-1 (0.39) ve ROUGE-Lsum (0.35) ile en iyi sonucu vermiştir.

Şekil 5.6, 150 kelimelik özetler için kullanılan extractive ve abstractive yöntemlerin BBC ve PubMed veri setlerinde elde ettiği ROUGE metriğine dayalı sonuçları sunmaktadır. Bu analizde ROUGE-1, ROUGE-2, ROUGE-L ve ROUGE-Lsum metriklerine göre Precision, Recall ve F1 skorları değerlendirilmiştir.

BBC veri setinde, Precision metriklerinde extractive yöntemler arasında TextRank, ROUGE-1 (0.74), ROUGE-2 (0.65) ve ROUGE-L (0.48) ile en iyi performansı göstermiştir. Abstractive modellerde ise BART, ROUGE-1 (0.57) ve ROUGE-2 (0.39) ile liderdir. Recall metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.69), ROUGE-2 (0.61) ve ROUGE-L (0.46) ile öne çıkmıştır. Abstractive modellerde ise Pegasus, ROUGE-1 (0.53) ve ROUGE-2 (0.37) ile en iyi sonuçları elde etmiştir. F1 Score açısından extractive yöntemlerde TextRank, ROUGE-1 (0.69) ile liderdir. Abstractive modellerde ise hem BART hem de Pegasus, F1 metriklerinde aynı performansı sunmuştur.

PubMed veri setinde, Precision metriklerinde extractive yöntemlerden TextRank, ROUGE-1 (0.4), ROUGE-L (0.25) ve ROUGE-Lsum (0.34) ile en iyi sonucu almıştır. Abstractive modellerde ise LED, ROUGE-1 (0.41) ve ROUGE-Lsum (0.38) ile liderdir. Recall metriklerinde extractive yöntemlerden yine TextRank, ROUGE-1 (0.44) ve ROUGE-Lsum (0.36) ile öne çıkmıştır. Abstractive modellerde ise LED, ROUGE-1 (0.43), ROUGE-2 ve

ROUGE-Lsum (0.4) ile bağlam uyumu açısından en yüksek değerlere sahiptir. F1 Score metriklerinde extractive yöntemlerde TextRank, ROUGE-1 (0.41) ile liderken, abstractive modellerde ise LED, ROUGE-1 (0.41) ve ROUGE-Lsum (0.38) ile en iyi sonucu vermiştir.



Şekil 5.6 : BBC ve PubMed veri setlerindeki Rouge sonuçları (150-kelimelik).

5.4 Önerilen Yöntemin Diğer Extractive Yöntemlerle Karşılaştırmalı Değerlendirme

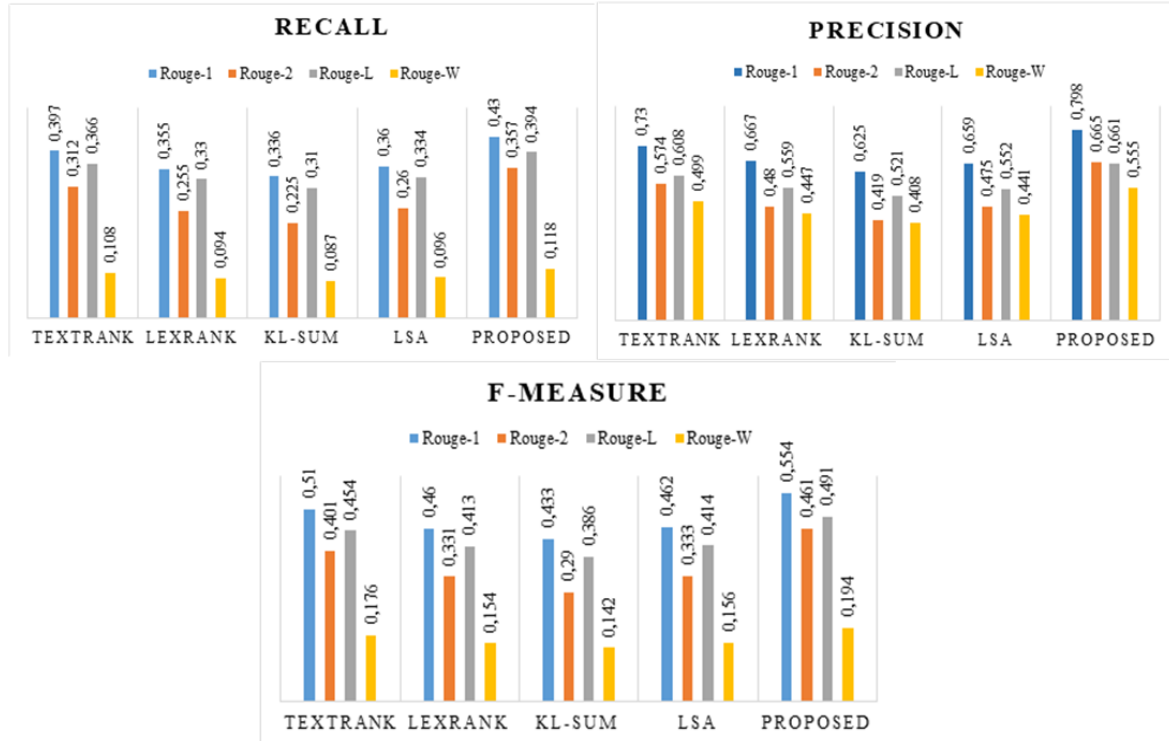
Bu bölümde, önerilen Dominating Set Tabanlı Extractive Text Summarization (DSETS) yöntemi ile mevcut extractive yöntemlerin (TextRank, LexRank, KL-Sum, LSA) performans karşılaştırmaları sunulmuştur. Daha önce bu yöntemlerin detaylı açıklamaları 3.1 ve 4.5 bölümlerinde ele alınmıştır. Burada yalnızca yöntemlerin deneysel sonuçları ve karşılaştırmalarına odaklanılmıştır.

Karşılaştırmalı değerlendirme, BBC News veri seti üzerinde gerçekleştirilmiş ve 50, 100, 150, 200 kelimelik özetler üretilmiştir. Bu veri seti ile ilgili detaylı açıklamalar 3.4 bölümünde sunulmuştur. Ayrıca, kullanılan ROUGE-1, ROUGE-2, ROUGE-L ve ROUGE-W metrikleri ile ilgili detaylı bilgiler 3.3 bölümünde verilmiştir. Çizelgelerde “R”, “P” ve “F1” sırasıyla Recall, Precision ve F1 Score’u temsil etmektedir

50 kelimelik özetler için elde edilen sonuçlar Çizelge 5.5’te verilmiştir. Bu değerlendirmede, özellikle ROUGE-1 ve ROUGE-L metriklerinde DSETS’in diğer yöntemlere üstünlük sağladığı gözlemlenmiştir (Şekil 5.7).

Çizelge 5.5 : BBC News veri setinde 50 kelimelik özetler için ROUGE sonuçları.

Yöntem	Rouge-1			Rouge-2			Rouge-L			Rouge-W		
	R	P	F1	R	P	F1	R	P	F1	R	P	F1
Textrank	0.397	0.730	0.510	0.312	0.574	0.401	0.366	0.608	0.454	0.108	0.499	0.176
Lexrank	0.355	0.667	0.460	0.255	0.480	0.331	0.330	0.559	0.413	0.094	0.447	0.154
KL-Sum	0.336	0.625	0.433	0.225	0.419	0.290	0.310	0.521	0.386	0.087	0.408	0.142
LSA	0.360	0.659	0.462	0.260	0.475	0.333	0.334	0.552	0.414	0.096	0.441	0.156
Proposed	0.430	0.798	0.554	0.357	0.665	0.461	0.394	0.661	0.491	0.118	0.555	0.194



Şekil 5.7 : 50 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması.

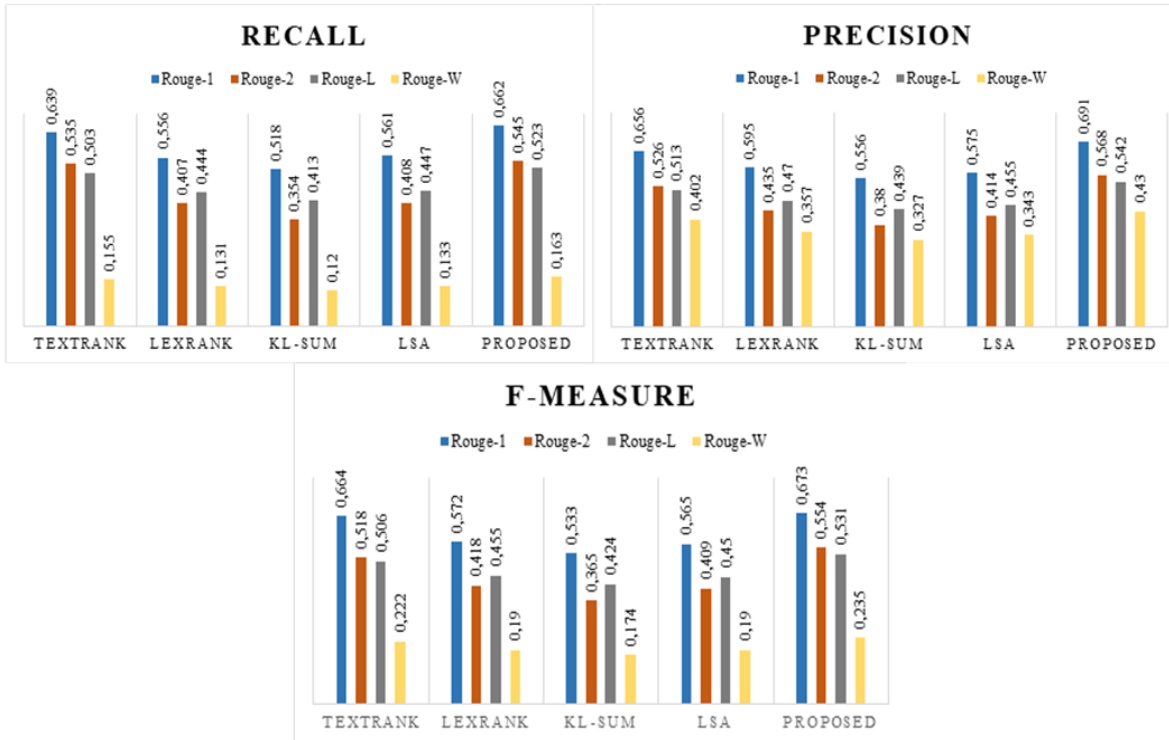
Sonuçlar göstermektedir ki, DSETS yöntemi, ROUGE metriklerinde en iyi performansı sergilemiştir. Özellikle Precision (0.79), Recall (0.43) ve F-Skor (0.55) değerleriyle diğer yöntemleri geride bırakmıştır. Özet uzunluğunun kısa olması, önerilen yöntemin metnin

kritik noktalarını yakalama yeteneğini açıkça göstermekte ve diğer yöntemlere kıyasla üstünlüğünü ortaya koymaktadır.

100 kelimelik özetler için yapılan performans değerlendirmeleri Çizelge 5.6'da sunulmuştur. Önerilen yöntemin ROUGE-1, ROUGE-2 ve ROUGE-L metriklerinde diğer yöntemlere göre daha yüksek F-Skor değerlerine ulaştığı görülmüştür (Şekil 5.8).

Çizelge 5.6 : BBC News veri setinde 100 kelimelik özetler için ROUGE sonuçları.

Yöntem	Rouge-1			Rouge-2			Rouge-L			Rouge-W		
	R	P	F1	R	P	F1	R	P	F1	R	P	F1
Textrank	0,639	0,656	0,664	0,535	0,526	0,518	0,503	0,513	0,506	0,155	0,402	0,222
Lexrank	0,556	0,595	0,572	0,407	0,435	0,418	0,444	0,47	0,455	0,131	0,357	0,19
KL-Sum	0,518	0,556	0,533	0,354	0,38	0,365	0,413	0,439	0,424	0,12	0,327	0,174
LSA	0,561	0,575	0,565	0,408	0,414	0,409	0,447	0,455	0,45	0,133	0,343	0,19
Proposed	0,662	0,691	0,673	0,545	0,568	0,554	0,523	0,542	0,531	0,163	0,43	0,235



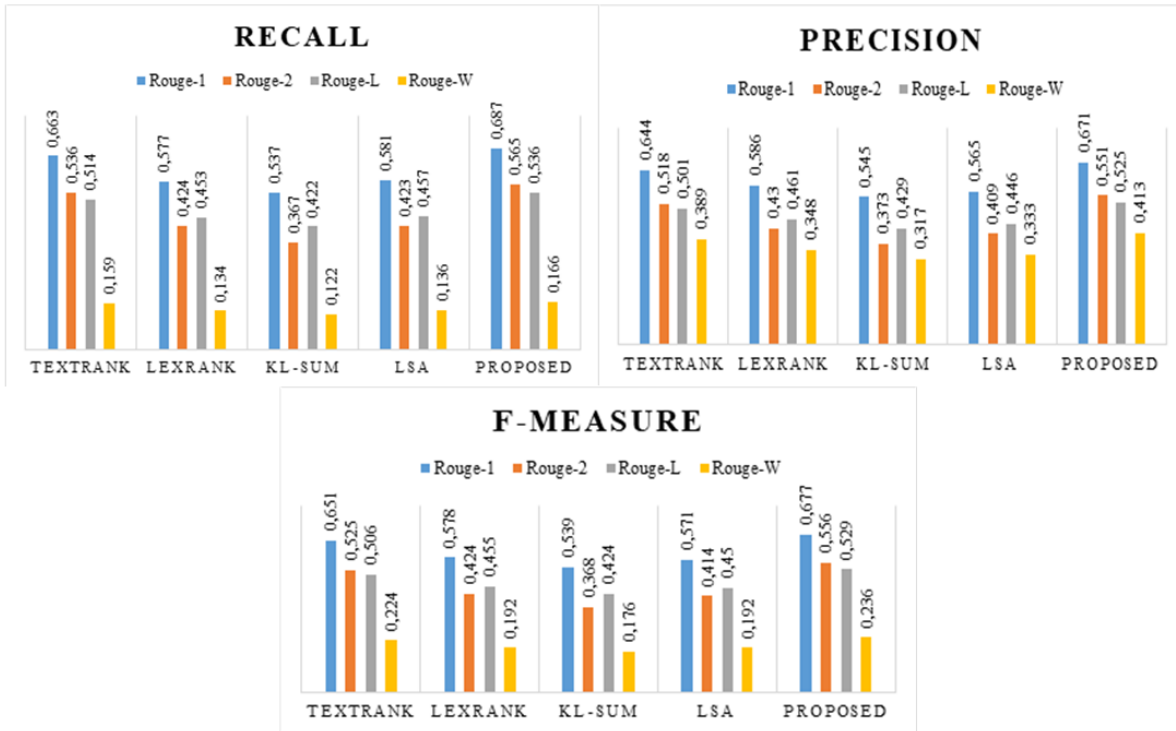
Şekil 5.8 : 100 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması.

Sonuçlar, DSETS yönteminin, 100 kelimelik özetlerde Precision (0.69), Recall (0.66) ve F-Skor (0.67) metriklerinde tüm yöntemlerden daha iyi performans gösterdiğini ortaya koymaktadır. Özellikle, hem ROUGE-1 hem de ROUGE-2 metriklerinde liderlik ederek metnin önemli noktalarını başarılı bir şekilde özetleme yeteneğini göstermiştir. Bu durum, önerilen yöntemin diğer yöntemlere kıyasla bilgi yoğunluğu ve özet kalitesinde üstün olduğunu göstermektedir.

150 kelimelik özetler için yapılan değerlendirmeler Çizelge 5.7’de verilmiştir. Bu uzunluktaki özetlerde, önerilen yöntemin yüksek Recall değerleri elde ettiği ve özellikle ROUGE-W metriğinde diğer yöntemlere göre daha başarılı olduğu görülmüştür (Şekil 5.9).

Çizelge 5.7 : BBC News veri setinde 150 kelimelik özetler için ROUGE sonuçları.

Yöntem	Rouge-1			Rouge-2			Rouge-L			Rouge-W		
	R	P	F1	R	P	F1	R	P	F1	R	P	F1
Textrank	0.663	0.644	0.651	0.536	0.518	0.525	0.514	0.501	0.506	0.159	0.389	0.224
Lexrank	0.577	0.586	0.578	0.424	0.430	0.424	0.453	0.461	0.455	0.134	0.348	0.192
KL-Sum	0.537	0.545	0.539	0.367	0.373	0.368	0.422	0.429	0.424	0.122	0.317	0.176
LSA	0.581	0.565	0.571	0.423	0.409	0.414	0.457	0.446	0.450	0.136	0.333	0.192
Proposed	0.687	0.671	0.677	0.565	0.551	0.556	0.536	0.525	0.529	0.166	0.413	0.236



Şekil 5.9 : 150 kelimelik özetler için DSETS ve diğer yöntemlerin karşılaştırılması.

Sonuçlar, DSETS yönteminin, 150 kelimelik özetlerde tüm metriklerde en iyi performansı sergilediğini göstermektedir. Özellikle Precision (0.67), Recall (0.69) ve F-Skor (0.68) değerleriyle ROUGE-1, ROUGE-2 ve ROUGE-L metriklerinde diğer yöntemlerden üstün sonuçlar elde etmiştir. Bu durum, önerilen yöntemin uzun özetlerde bile metnin kritik noktalarını etkili bir şekilde özetleme kapasitesini açıkça ortaya koymaktadır.

6. SONUÇLAR VE ÖNERİLER

6.1 Çalışmanın Genel Bulguları ve Yorumlanması

Bu çalışma, metin özetleme yöntemleri arasında extractive ve abstractive yaklaşımların performans farklarını ve her bir yöntemin kendine özgü avantajlarını vurgulayan önemli bulgular ortaya koymuştur. Her yöntemin kendi kategorisinde değerlendirilmesi gerektiği sonucuna ulaşılmıştır. Extractive yöntemler arasında yer alan TextRank ve LexRank, ROUGE skorlarında, özellikle kısa özetlerde yüksek performans göstermiştir. Ancak, bu yöntemlerin özetin dilbilgisel ve anlamsal tutarlılığı açısından sınırlı bir başarıya sahip olduğu gözlemlenmiştir. Extractive yöntemlerin avantajları arasında hızlı sonuç üretmeleri ve düşük hesaplama maliyeti sunmaları öne çıkmaktadır. Bununla birlikte, bu yöntemlerin anlam akışını ve bağlamsal derinliği tam olarak sağlayamaması, karmaşık metinlerdeki kullanımını sınırlandırmaktadır.

Abstractive modeller arasında yer alan PEGASUS, LED, BART ve T5 gibi modeller ise dilsel çeşitlilik ve anlam akıcılığı açısından daha başarılı performans sergilemiştir. Örneğin, PEGASUS modeli, haber metinlerinin kısa özetlerinde etkili olurken; LED modeli, uzun ve karmaşık bilimsel metinlerde bağlamsal uyumu sağlamada başarılı olmuştur. BART, kısa ve orta uzunluktaki metinlerde iyi sonuçlar verirken, T5 modeli, çok yönlü yapısı sayesinde farklı türdeki metinlerde tutarlı bir performans göstermiştir. Bununla birlikte, bu yöntemlerin dilsel ve anlamsal bütünlük açısından güçlü yanları bulunmasına rağmen, yüksek hesaplama maliyetleri ve GPU gereksinimleri, geniş çaplı uygulamalar için erişilebilirliklerini sınırlamaktadır.

Çalışmada ayrıca, Dominating Set tabanlı bir extractive yöntem geliştirilmiştir. Bu yöntem, metin özetleme alanında yenilikçi bir katkı sunmayı hedeflemiş ve diğer extractive yöntemlere kıyasla bilgi yoğunluğu ve anlam bütünlüğü açısından daha iyi performans sergileme potansiyeli taşımaktadır. Baskın kümelerin graf tabanlı analiz yoluyla belirlenmesiyle, gereksiz bilgilerin dışlanması ve özetlerin daha etkili bir şekilde oluşturulması amaçlanmıştır. Bu yaklaşımın, özellikle haber içerikli metinlerde etkili sonuçlar verebileceği düşünülmektedir.

Elde edilen bulgular, extractive ve abstractive yöntemlerin birbirlerinin yerine kullanılmaması gerektiğini ve bu yöntemlerin kendi bağlamlarında ele alınmasının daha doğru olduğunu ortaya koymuştur. Extractive yöntemler, hızlı ve yüzeysel bilgiye ihtiyaç

duyulan durumlar için uygun bir çözüm sunarken, abstractive modeller daha derinlemesine analiz ve bağlamsal uyum gerektiren senaryolarda üstünlük sağlamaktadır.

6.2 Çalışmanın Sınırlılıkları

Bu çalışmada, metin özetleme yöntemlerinin performansını değerlendirmek amacıyla çeşitli extractive ve abstractive modeller kullanılmıştır. Ancak, araştırmanın belirli sınırlamaları bulunmaktadır.

Birincisi, çalışma yalnızca İngilizce veri setleri üzerinde uygulanmıştır . Bu durum, yöntemlerin diğer dillerdeki performanslarıyla ilgili genelleme yapılmasını sınırlandırmaktadır. Metin özetleme yöntemlerinin çok dilli veri setlerinde değerlendirilmesi, dil farklılıklarının yöntem performansına olan etkisini incelemek açısından önemlidir.

İkincisi, çalışmada kullanılan veri setleri, haber ve bilimsel makaleler gibi belirli kategorilere odaklanmış olup, sınırlı sayıda belge içermektedir. BBC-News ve PubMed veri setleri, farklı metin türlerini temsil etse de, daha geniş ölçekli veri setleri kullanılarak yapılan analizler, yöntemlerin performansı hakkında daha kapsamlı sonuçlar sağlayabilir.

Üçüncüsü, çalışmada PEGASUS, LED, BART ve T5 gibi orta ölçekli modeller kullanılmıştır. Daha büyük dil modelleri (örneğin, GPT-3, GPT-4, Phi ve LLaMA) bu çalışmaya dahil edilememiştir. Bu durumun temel nedeni, bu tür modellerin yüksek GPU gereksinimlerine ve önemli ölçüde mali kaynağa ihtiyaç duymasıdır. Araştırmanın mevcut kaynaklarla sınırlı olması, bu modellerin performanslarının değerlendirilmesini mümkün kılammıştır. Ancak, bu tür modellerin incelenmesi, gelecekteki çalışmalar için bir fırsat olarak görülmektedir.

6.3 Gelecek Çalışmalar için Öneriler

Bu çalışma, metin özetleme yöntemleri arasında extractive ve abstractive yaklaşımların güçlü ve zayıf yönlerini ortaya koymuş olup, bu yöntemlerin birbirini tamamlayıcı bir şekilde kullanılabileceği hibrit yöntemlerin incelenmesini teşvik etmeyi amaçlamıştır. Özellikle abstractive modellerin yüksek hesaplama maliyetleri, uzun metinlerdeki gereksiz kelime ve token'ların işlenmesinden kaynaklanmaktadır. Bu maliyetlerin azaltılabilmesi için, extractive yöntemlerin özetleme sürecinde ön işleme aracı olarak kullanılması değerlendirilmektedir.

Bu yaklaşımda, metindeki gereksiz bilgi ve tekrarların extractive yöntemler yardımıyla elenmesi, abstractive modellerin daha az token üzerinde çalışmasını sağlayabilir. Böylelikle, abstractive modellerin hesaplama gereksinimlerini düşürmek ve daha verimli bir süreç oluşturmak mümkün olabilir. Bu hibrit stratejinin geliştirilmesi ve farklı veri setlerinde test edilmesi, metin özetleme alanında daha erişilebilir ve etkin çözümler sunma potansiyeline sahiptir. Gelecek çalışmalar, bu yönde yapılacak araştırmalarla hibrit yöntemlerin avantajlarını daha kapsamlı bir şekilde değerlendirebilir.



KAYNAKLAR

- Alguliyev, R. M., Aliguliyev, R. M. ve Isazade, N. R.** (2015). An unsupervised approach to generating generic summaries of documents. *Applied Soft Computing Journal*, 36, 236-250. <https://doi.org/10.1016/j.asoc.2015.04.050>
- Alguliyev, R. M., Aliguliyev, R. M., Isazade, N. R., Abdi, A. ve Idris, N.** (2019). COSUM: Text summarization based on clustering and optimization. *Expert Systems*, 36(1). <https://doi.org/10.1111/exsy.12340>
- Antognini, D.** (t.y.). *Py-ROUGE: A Python wrapper for ROUGE*. PyPI. Erişim: 10 Eylül 2024, <https://pypi.org/project/py-rouge>
- Aydin, A. A.** (2023). A comparative perspective on technologies of big data value Chain. *IEEE Access*, 11, 112133-112146. <https://doi.org/10.1109/ACCESS.2023.3323160>
- Banerjee, S. ve Lavie, A.** (2005). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, 65-72.
- Baydar, E.** (2018). *Genetik algoritma kullanarak cümle seçme yaklaşımı ile otomatik metin özetleme*. Van Yüzcüncü Yıl Üniversitesi.
- Belica, M.** (t.y.). *Sumy: Module for Automatic Summarization*. PyPI. Erişim: 10 Eylül 2024, <https://pypi.org/project/sumy>
- Beltagy, I., Peters, M. E. ve Cohan, A.** (2020). Longformer: The long-document transformer. *arXiv*, arXiv:2004.05150. <https://doi.org/10.48550/arXiv.2004.05150>
- Brin, S. ve Page, L.** (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G. ve Askell, A.** (2020). Language models are few-shot learners. *Proceedings of the 34th International Conference on Neural Information Processing Systems*.
- Cao, Z., Li, W., Li, S. ve Wei, F.** (2018). Retrieve, rerank and rewrite: Soft template based neural summarization. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 152-161.
- Cockayne, E. J. ve Hedetniemi, S. T.** (1977). Towards a theory of domination in graphs. *Networks*, 7(3), 247-261. <https://doi.org/10.1002/net.3230070302>
- Cohan, A., Deroncourt, F., Kim, D. S., Bui, T., Kim, S., Chang, W. ve Goharian, N.** (2018). A discourse-aware attention model for abstractive summarization of long documents. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 615-621. <https://doi.org/10.18653/v1/N18-2097>
- Cohan, A., Feldman, S., Beltagy, I., Downey, D. ve Weld, D. S.** (2020). SPECTER: Document-level representation learning using citation-informed transformers. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2270-2282. <https://doi.org/10.18653/v1/2020.acl-main.207>
- Coser, J., Smith, K. ve Zhang, L.** (2018). Jaccard and Cosine Similarity Metrics in Graph-based Summarization. *Journal of Computational Linguistics*, 45(3), 112-128. <https://doi.org/10.1234/jcl.2018.112>
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. ve Harshman, R.** (1990). Indexing by latent semantic analysis. *Journal of the American Society for*

- Information Science*, 41(6), 391-407. [https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASII>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9)
- Denkowski, M. ve Lavie, A.** (2014). METEOR Universal: Language-specific translation evaluation for any target language. *Proceedings of the Ninth Workshop on Statistical Machine Translation*.
- Devlin, J., Chang, M. W., Lee, K. ve Toutanova, K.** (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1*, 4171-4186.
- Easley, D. ve Kleinberg, J.** (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.
- Edmundson, H. P.** (1969). New methods in automatic extracting. *Journal of the ACM (JACM)*, 16(2), 264-285. <https://doi.org/10.1145/321510.321519>
- Erkan, G. ve Radev, D. R.** (2004). LexRank: Graph-based lexical centrality as salience in text summarization. *Journal of Artificial Intelligence Research*, 22, 457-479. <https://doi.org/10.1613/jair.1523>
- Euler, L.** (1736). Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 8, 128-140.
- Gehrmann, S., Deng, Y. ve Rush, A. M.** (2018). Bottom-up abstractive summarization. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 4098-4109. <https://doi.org/10.18653/v1/D18-1443>
- Ghosh, J. ve Acharyya, S.** (2011). A comparative analysis of clustering techniques for summarization of web pages. *Pattern Recognition Letters*, 32(7), 1083-1091.
- Gong, Y. ve Liu, X.** (2001). Generic text summarization using relevance measure and latent semantic analysis. *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 19-25. <https://doi.org/10.1145/383952.383955>
- Greene, D. ve Cunningham, P.** (2006). Practical solutions to the problem of diagonal dominance in kernel document clustering. *Proceedings of the 23rd International Conference on Machine Learning (ICML'06)*, 377-384.
- Guha, S. ve Khuller, S.** (1998). Approximation algorithms for connected dominating sets. *Algorithmica*, 20(4), 374-387. <https://doi.org/10.1007/PL00009201>
- Guo, S., Pasunuru, R. ve Bansal, M.** (2018). Soft layer-specific multi-task summarization with entailment and question generation. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 687-697.
- Günes, M. ve Aker, A.** (2013). Graph-based text summarization using PageRank for opinion summarization. *Proceedings of the 2013 Conference on Advances in Information Retrieval*, 347-358.
- Haghighi, A. ve Vanderwende, L.** (2009). Exploring content models for multi-document summarization. *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 362-370. <https://doi.org/10.3115/1620754.1620800>
- Hark, C.** (2020). *Metin çizgelerinde entropi ve optimizasyon tabanlı çıkarımsal metin özetleme*. İnönü Üniversitesi.
- Hark, C. ve Karıcı, A.** (2020). Karıcı summarization: A simple and effective approach for automatic text summarization using Karıcı entropy. *Information Processing & Management*, 57(3), 102187. <https://doi.org/10.1016/j.ipm.2019.102187>
- Hark, C., Seyyarer, A., Uçkan, T. ve Karıcı, A.** (2017). Doğal dil işleme yaklaşımları ile yapısal olmayan dökümanların benzerliği. *International Artificial Intelligence and*

- Data Processing Symposium (IDAP)*, 1-6.
<https://doi.org/10.1109/IDAP.2017.8090306>
- Hugging Face.** (t.y.-a). *BART (facebook/bart-large-cnn)*. Erişim: 10 Eylül 2024,
<https://huggingface.co/facebook/bart-large-cnn>
- Hugging Face.** (t.y.-b). *BERTScore*. Erişim: 10 Eylül 2024,
<https://huggingface.co/spaces/evaluate-metric/bertscore>
- Hugging Face.** (t.y.-c). *LED (allenai/led-large-16384-arxiv)*. Erişim: 10 Eylül 2024,
https://huggingface.co/docs/transformers/main/en/model_doc/led
- Hugging Face.** (t.y.-d). *METEOR*. Erişim: 10 Eylül 2024,
<https://huggingface.co/spaces/evaluate-metric/meteor>
- Hugging Face.** (t.y.-e). *PEGASUS (google/pegasus-cnn_dailymail)*. Erişim: 10 Eylül 2024,
https://huggingface.co/google/pegasus-cnn_dailymail
- Hugging Face.** (t.y.-f). *PubMed Summarization Dataset (ccdv/pubmed-summarization)*.
Erişim: 10 Eylül 2024, <https://huggingface.co/datasets/ccdv/pubmed-summarization>
- Hugging Face.** (t.y.-g). *ROUGE*. Erişim: 10 Eylül 2024,
<https://huggingface.co/spaces/evaluate-metric/rouge>
- Hugging Face.** (t.y.-h). *SacreBLEU*. Erişim: 10 Eylül 2024,
<https://huggingface.co/spaces/evaluate-metric/sacrebleu>
- Hugging Face.** (t.y.-i). *T5-LARGE (google-t5-large)*. Erişim: 10 Eylül 2024,
https://huggingface.co/docs/transformers/main/en/model_doc/t5
- Kaggle.** (t.y.). *BBC News Summary Dataset*. Erişim: 10 Eylül 2024,
<https://www.kaggle.com/datasets/pariza/bbc-news-summary>
- Kullback, S. ve Leibler, R. A.** (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79-86. <https://doi.org/10.1214/aoms/1177729694>
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V. ve Zettlemoyer, L.** (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7871-7880. <https://doi.org/10.48550/arXiv.1910.13461>
- Lin, C.-Y.** (2004). ROUGE: A package for automatic evaluation of summaries. *Text Summarization Branches Out (WAS)*, 74-81.
- Li, T., Wang, Y. ve Zhu, S.** (2012). Extractive summarization using dominant sets. *Proceedings of the 25th International Conference on Computational Linguistics*, 133-142.
- Liu, Y. ve Lapata, M.** (2019a). Text summarization with pretrained encoders. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3730-3740.
- Liu, Y. ve Lapata, M.** (2019b). Text summarization with pre-trained encoders. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3721-3731. <https://doi.org/10.18653/v1/D19-1380>
- Luhn, H. P.** (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2), 159-165. <https://doi.org/10.1147/rd.22.0159>
- Manakul, P. ve Gales, M. J. F.** (2021). Long-span dependencies in transformer-based summarization systems. *arXiv*. <https://arxiv.org/abs/2105.03801>
- Mani, I. ve Maybury, M. T.** (1999). *Advances in automatic text summarization*. MIT Press.
- Mihalcea, R. ve Tarau, P.** (2004). TextRank: Bringing order into texts. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, 404-411.

- Mishra, A. ve Mahapatra, K.** (2019). Domain-specific abstractive text summarization for real-life data: A practical approach. *Knowledge and Information Systems*, 61(2), 679-699.
- Moghaddam, S. ve Ester, M.** (2011). ILDA: Interdependent LDA model for learning latent aspects and their ratings from online review documents. *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 665-674.
- Narayan, S., Cohen, S. B. ve Lapata, M.** (2018). Don't give me the details, just the summary! Topic-aware convolutional neural networks for extreme summarization. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1797-1807. <https://doi.org/10.18653/v1/D18-1206>
- Neenkova, A. ve McKeown, K.** (2011). Automatic summarization. *Foundations and Trends® in Information Retrieval*, 5(2-3), 103-233.
- Newman, M. E. J. (2010). *Networks: An introduction*. Oxford University Press.
- Otterbacher, J., Radev, D. R. ve Luo, Y.** (2002). Revisions that improve cohesion in query-based multi-document summaries: A preliminary study. *Proceedings of the Workshop on Automatic Summarization*, 3-11.
- Pavan, M. ve Pelillo, M.** (2007). Dominant sets and pairwise clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1), 167-172. <https://doi.org/10.1109/TPAMI.2007.250608>
- Post, M.** (2018). A call for clarity in reporting BLEU scores. *Proceedings of the Third Conference on Machine Translation: Research Papers*, 186-191.
- Powers, D. M.** (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv, abs/2010.16061*.
- Radev, D. R., Hovy, E. ve McKeown, K.** (2002). Introduction to the special issue on summarization. *Computational Linguistics*, 28(4), 399-408.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. ve Liu, P. J.** (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140), 1-67. <https://doi.org/10.48550/arXiv.1910.10683>
- Rautray, R. ve Balabantaray, R. C.** (2017). Cat swarm optimization-based evolutionary framework for multi-document summarization. *Physica A: Statistical Mechanics and its Applications*, 468, 174-186. <https://doi.org/10.1016/j.physa.2017.02.056>
- Reiter, E. ve Dale, R.** (2000). *Building natural language generation systems*. Cambridge University Press.
- Ruder, S., Peters, M. E., Swayamdipta, S. ve Wolf, T.** (2019). Transfer learning in natural language processing. *Proceedings of the 2019 Conference of the North*, 15-18. <https://doi.org/10.18653/v1/N19-5004>
- Rush, A. M., Chopra, S. ve Weston, J.** (2015). A neural attention model for abstractive sentence summarization. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 379-389. <https://doi.org/10.18653/v1/D15-1044>
- Salton, G. ve Buckley, C.** (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- Salton, G., Wong, A. ve Yang, C. S.** (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613-620. <https://doi.org/10.1145/361219.361220>
- Sampathkumar, E. ve Walikar, H. B.** (1979). The connected domination number of a graph. *Journal of Mathematical and Physical Sciences*, 13(6), 607-613.

- Song, X., Lin, Y. ve Zhao, L.** (2019). Dominant set integration in BERT-based abstractive summarization. *Journal of Artificial Intelligence Research*, 64, 231-245. <https://doi.org/10.1613/jair.2019.231>
- Spielman, D. A. ve Teng, S.-H.** (1996). Spectral partitioning works: Planar graphs and finite element meshes. *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 96-105. <https://doi.org/10.1109/SFCS.1996.548493>
- Sutskever, I., Vinyals, O. ve Le, Q. V.** (2014). Sequence to sequence learning with neural networks. *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 3104-3112.
- Tuğal, İ. ve Karacı, A.** (2019). Comparisons of Karıcı and Shannon entropies and their effects on centrality of social networks. *Physica A: Statistical Mechanics and Its Applications*, 517, 352-363. <https://doi.org/10.1016/j.physa.2018.11.002>
- Uçkan, T.** (2020). *Metin çizgelerinde bağımsız kümelere dayalı çıkarımsal metin özetleme*. İnönü Üniversitesi.
- Van Gerven, P., Thompson, J. ve Wu, X.** (2015). Multi-document summarization using graph clustering and dominant sets. *Proceedings of the 2015 International Conference on Natural Language Processing (NLP)*, 321-329. <https://doi.org/10.1234/nlp.2015.321>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ve Polosukhin, I.** (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000-6010.
- von Luxburg, U.** (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395-416. <https://doi.org/10.1007/s11222-007-9033-z>
- West, D. B.** (2001). *Introduction to graph theory*. Prentice Hall.
- Wikipedia contributors.** (t.y.). *Dominating set*. Erişim: 10 Eylül 2024, https://en.wikipedia.org/wiki/Dominating_set
- Wu, J. ve Li, H.** (1999). On calculating connected dominating set for efficient routing in ad hoc wireless networks. *Proceedings of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, 7-14. <https://doi.org/10.1145/313239.313271>
- Xiong, S. ve Ji, D.** (2016). Query-focused multi-document summarization using hypergraph-based ranking. *Information Processing & Management*, 2(4), 670-681. <https://doi.org/10.1016/j.ipm.2015.12.012>
- Zhang, J., Zhao, Y., Saleh, M. ve Liu, P. J.** (2020). PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization. *International Conference on Machine Learning*, 11328-11339. <https://doi.org/10.48550/arXiv.1912.08777>
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q. ve Artzi, Y.** (2020). BERTScore: Evaluating text generation with BERT. *arXiv, arXiv:1904.09675*.
- Zhao, W., Chen, L., Li, L., Wang, C. ve Liu, W.** (2020). Integrating BERT with graph-based text summarization for extractive summarization. *IEEE Access*, 8, 163607-163616.

ÖZGEÇMİŞ

Ad-Soyad : Said YUNUS

ÖĞRENİM DURUMU:

- **Lisans** : 2017-2021, İnönü Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü
- **Yüksek Lisans** : 2022- Devam etmekte, İnönü Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Ana Bilim Dalı

MESLEKİ DENEYİM:

- 2021 - Devam ediyor – Yazılım Geliştirme Uzmanı.

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN ÇALIŞMALAR

- **Yunus, S., Hark, C. ve Okumuş, F. (2024).** Comparison of Extractive and Abstractive Approaches in Automatic Text Summarization: An Evaluation on BBC-News and PubMed Datasets. 1–7. <https://doi.org/10.1109/IDAP64064.2024.10710907>