

## Hikâye Kitapları için Transformatör Tabanlı Bir Özetleme Modeli

Mehtap ÜLKER<sup>1\*</sup>, A. Bedri ÖZER<sup>2</sup>

<sup>1,2</sup> Computer Engineering Department, Firat University, Elazig, Turkey

<sup>\*1</sup> m.ulker@firat.edu.tr, <sup>2</sup> bedriozer@firat.edu.tr

(Geliş/Received: 10/12/2024;

Kabul/Accepted:13/03/2025)

**Öz:** Günümüzde, doküman özetleme için önerilen birçok model olmasına rağmen hikâye kitaplarını ele alan özetleme modelleri sınırlıdır. Bu problemi ele almak için yeni bir transformatör tabanlı özetleme modeli önerilmiştir. Önerilen model, KeyBERT, DistilBERT-NER, TF-IDF ve BART modelinden oluşmaktadır. KeyBERT ve TF-IDF, anahtar kelimeler çıkarılması için kullanılmıştır. DistilBERT-NER ise hikâye kitaplarından varlık isimleri çıkarımı için kullanılmıştır. Varlık isimleriyle bu anahtar kelimeler birleştirilerek İngilizce hikâye kitapları için transformatör tabanlı bir özetleme modeli oluşturulmuştur. Önerilen model, ince ayar yapılarak eğitilmiş T5, BART ve PEGASUS modelleriyle karşılaştırılmış ve sonuçlar insanlar tarafından değerlendirilmiştir. Deneysel sonuçlar, önerilen modelin temel yöntemlere kıyasla Rouge-L - F1 değeriyle hikâye özetlemede daha yüksek değerlere ulaştığı gözlemlenmiştir.

**Anahtar kelimeler:** Metin Özetleme, Soyutlamalı Yöntem, KeyBERT, DistilBERT, TF-IDF.

### Transformer-Based Summarization Model for Story Books

**Abstract:** Recently, although there are many proposed models for document summarization, summarization models for storybooks are limited. To address this problem, a new transformer-based summarization model is proposed. The proposed model consists of KeyBERT, DistilBERT-NER, TF-IDF and BART. KeyBERT and TF-IDF are used for keyword extraction. DistilBERT-NER was used to extract entity names from storybooks. By combining these keywords with entity names, a transformer-based summarization model for English storybooks was created. The proposed model is compared with the fine-tuned T5, BART and PEGASUS models and the results are evaluated by humans. The experimental results show that the proposed model achieves higher values of Rouge-L - F1 in story summarization compared to the baseline methods.

**Key words:** Text summarization, Abstractive method, KeyBERT, DistilBERT, TF-IDF.

### 1. Giriş

Doküman özetleme, günümüzdeki en popüler araştırma konularından biridir. Bu alandaki en temel istekleri kullanıcıların ihtiyaç duydukları bilgilere hızlı bir şekilde ulaşabilmesidir. Ancak, bir dokümandan önemli bilgilerin manuel olarak çıkarılması, zorlu ve zaman alıcı bir süreçtir. Metin özetleme temelde çıkarımlı ve soyutlamalı olmak üzere iki farklı şekilde uygulanmaktadır. Çıkarımlı yöntemlerde, kaynak dokümandaki en önemli cümleler çıkarılarak özetlenmektedir. Soyutlamalı yöntemlerde ise dokümandan yeni kelimelerle yeniden yazılarak özet oluşturulmaktadır. Bu açıdan, soyutlamalı yaklaşımlar insan tarafından oluşturulan özetlere benzemektedir. Dolayısıyla çıkarımlı yaklaşımlara göre uygulanması daha zordur [1, 2].

Derin öğrenme yöntemlerinin gelişmesiyle birlikte, soyutlamalı metin özetleme alanında başarılı sonuçlar elde edilmiştir [3]. Metin özetleme yöntemlerinde, kelimeler arasındaki ilişkileri daha etkili bir şekilde çıkarabildiği için transformatör tabanlı dil modelleri (BERT, ROBERT, vb.) sıklıkla tercih edilmektedir. Ancak, bu dil modelleri metindeki tüm kelimeleri dikkate alarak anlamsal ilişkileri çıkarmada zorluk yaşamaktadır ve dolayısıyla doküman özetleme konusunda performansları da sınırlı kalmaktadır [4, 5]. Doküman özetleme, metinden anlamsal bilgilerin çıkarılması ve bu bilgilerin anlamlı ve tutarlı bir şekilde cümlelere dönüştürülmesi amacıyla taşınmaktadır. Uzun dokümanlarda, cümleler arasındaki anlamsal bütünlüğü koruyarak özet oluşturmak doküman uzunluğu arttıkça daha da zor bir durumdur. Graf tabanlı modeller [3, 6] ile dokümanın hiyerarşik yapısını ele alarak özetlemek mümkün olduğu için uzun dokümanlar da özetlenebilmektedir. Bunun yanı sıra, doküman ait belirleyici önemli kelimelerle uzun dokümanları özetlemek mümkündür. Bu amaçla kullanılan yöntemlerin başında anahtar kelime ve varlık ismi çıkarımı gelmektedir. Anahtar kelimeleri çıkarmak için istatistiksel tabanlı, graf tabanlı (TextRank, LexRank) veya derin öğrenme tabanlı otomatik anahtar kelime çıkarma (AKE) yöntemleri kullanılmaktadır [7]. Anahtar kelimeler, metnin içeriğini hızlı bir şekilde anlamamıza yardımcı olan kelime ya da

\* Sorumlu yazar: [m.ulker@firat.edu.tr](mailto:m.ulker@firat.edu.tr) Yazarların ORCID Numarası: <sup>1</sup> 0000-0001-8680-8518, <sup>2</sup> 0000-0002-8005-7386

kelime gruplarıdır [8]. AKE yöntemlerine ek olarak, varlık isimleri çıkarma (NER) teknikleri de soyutlamalı metin özetlemesinde genellikle tercih edilen yöntemlerdendir [9]. Ancak, metin üretimi genellikle transformatör tabanlı anahtar kelime çıkarma ve NER yöntemlerine odaklanmamıştır. Du ve diğerleri tarafından yapılan çalışmada [10], biyomedikal dokümanları özetlemek için alan bilgisi ve cümle pozisyonu bilgisini kullanan bir model önerilmiştir. Xiao ve diğerleri tarafından yapılan çalışmada [11], cümle füzyonu ve kooperatif pekiştirmeli öğrenmeye dayalı bir özetleme model geliştirmiştir. Cai ve diğerleri tarafından yapılan çalışmada ise [3], SciBERT ve Graph Attention Networks (GAT) kullanan bir özetleme modeli önerilmektedir. Moirangthem ve diğerleri [12], uzun dokümanları özetlemek için zaman temelli hiyerarşik işaretçi üretici ağ tabanlı bir özetleme modeli önermektedir. Liang ve diğerleri, sosyal medya metinlerini özetlemek için seçici pekiştirilmiş sıralıdan sıralı dikkatli ağ modeline dayalı bir özetleme yöntemi önermiştir. BART modelinin gelişmesiyle birlikte, metin üretme son zamanlarda hızla gelişim göstermektedir. BART modeli yapısı gereğince iki yönlü kodlayıcı ve otoregresif kod çözücüyü içermektedir. Otoregresif kod çözücü GPT modellerinde kullanılan bir kod çözücü yöntemidir. Çıktı olarak üretilen özelliği tekrar girdi olarak kullanılmaktadır [14]. Transformatör tabanlı modeller, yerel belirteçler ve global belirteçler arasındaki anlamsal ilişkileri keşfetmede oldukça kullanışlıdır. Ancak dokümanın uzunluğunun artmasıyla anlamsal ilişkilerin yakalanması problemi ortaya çıkmaktadır. Bu problemi çözebilmek için BART modeline dayalı yeni yöntemler önerilmiştir. Wang ve diğerleri tarafından yapılan çalışmada [5], BART ile bütünleşmiş yeni bir konu asistanı modeli önerilmiştir. Yu ve diğerleri tarafından yapılan çalışmada [15], uzman olmayan kişiler özet üreten BART tabanlı bir model geliştirilmiştir. Bajaj ve diğerleri tarafından yapılan çalışmada, BART tabanlı bir modelle sıkıştırılmış dokümanda özet üretme yöntemi önerilmiştir.

Metin özetlemede önerilen yöntem kadar kullanılan giriş metni de oldukça önemlidir. Çoğu model, BART, BERT ve SciBERT gibi modellerin belirli bir belirteç sınırına sahip olması nedeniyle dokümanın tamamını işleyemez [3]. Bu nedenle, özet oluşturulurken Otomatik Anahtar Kelime Çıkartma (AKE) ile çıkarılan anahtar kelimelerden yararlanılmaktadır. Christia ve diğerleri tarafından yapılan çalışmada [17], TF-IDF algoritması ile çıkarılan anahtar kelimelerden özet oluşturmayı önerilmiştir. Deng ve diğerleri [18], TextRank algoritması ile çıkarılan anahtar kelimelere dayalı bir metin özeti oluşturmayı öneren bir model geliştirilmiştir. Li ve diğerleri [19], TextRank algoritması ile çıkarılan anahtar kelimeleri kodlayan bir özetleme modeli önermiştir. TF-IDF, metindeki kelimelerin sıklığına göre anahtar kelimeleri çıkaran güçlü bir algoritmadır. Ancak, kelimelerin anlamsal bilgisini göz ardı etmektedir [20]. Bu nedenle, anahtar kelime çıkartma işlemi için genellikle TextRank algoritmasına geniş bir şekilde odaklanılmaktadır. Anahtar kelimelerin yanı sıra, bu çalışmada girdi metni olarak varlık isimlerine de odaklanılmıştır. Xiao ve diğerleri [11], cümle gruplamada varlıkları kullanan bir metin özetleme modeli önermiştir. Li ve diğerleri [21], graf tabanlı ağlarda varlık isimlerini kullanarak bir özet oluşturmuştur. Koncel-Kedziorski ve diğerleri [6], graf tabanlı ağlardaki varlık isimlerini ve ilişkileri makalenin başlığıyla birleştirerek bir özetleme modeli sunmuşlardır. Xiaoye ve diğerleri [22], BERT kodlayıcı ile çıkarımlı özetleme sürecinde modelin metnin tamamı yerine ana konunun içeriğine odaklanmasını sağlamak için anahtar ifadeleri kullanmayı amaçlayan KeyBERTSUM yöntemini sunmuştur. Yoo ve diğerleri [23], metin özetleme için diyalogdaki ana konuyu yakalamak zorluğuna değinerek anahtar kelimeleri kullanarak diyalog özetleme modeli önermişlerdir. Zhao ve diğerleri [24], hikâye kitapları için eğitim içerikli sorular üretmek adına soru türü tahmini ve olay odaklı özetlemeden oluşan yeni bir model önermiştir. Hikâyedeki paragrafta öne çıkan olayları, eğitimsel sorular üretmek için özetlemek hala çözülmemiş bir problem olarak ifade edilmiştir. Özetleme için BART tabanlı olay odaklı özetleme yöntemi geliştirilmiştir. Eğitim verilerini elde etmek için kural tabanlı bir yöntemle eğitimsel soru-cevap çiftleri oluşturulmuş ve değerlendirme için kullanılmıştır. Ling ve diğerleri [25], metin özetlemede olgusal tutarsızlığın etkilerini azaltmak ve tutarlı özetler oluşturmak amacıyla, PEGASUS modelini temel alan COFACTSUM modelini önermiştir. Soyutlamalı özetleme süreçlerinde metindeki tutarsızlıkların azaltılması ve özetlerin daha tutarlı hale getirilmesi araştırmacıların temel hedefi olarak belirlenmiştir. Upadhyay ve diğerleri [26], film senaryolarını edebi yorumlarla özetlemek için BART tabanlı bir model önermiştir. Ancak, modelin maksimum kelime uzunluğu nedeniyle başarılı özetler üretmediği belirtilmiş ve bu nedenle iyileştirilmesi gerektiği üzerinde durulmuştur.

Bu çalışmada, anahtar kelimeler ve varlık isimleri kullanılarak hikâye özetlerini oluşturmayı amaçlayan yeni bir Transformatör tabanlı model önerilmektedir. Anahtar kelimeler, BERT tabanlı bir model aracılığıyla, varlık isimleri ise DistilBERT-NER modeli ile çıkarılmıştır. Çalışmada, KeyBERT ve NER modelini birleştirerek yeni bir belirteç listesi oluşturulmuştur. Transformatör tabanlı model, İngilizce hikâye kitaplarını özetlemek için eğitilmiştir.

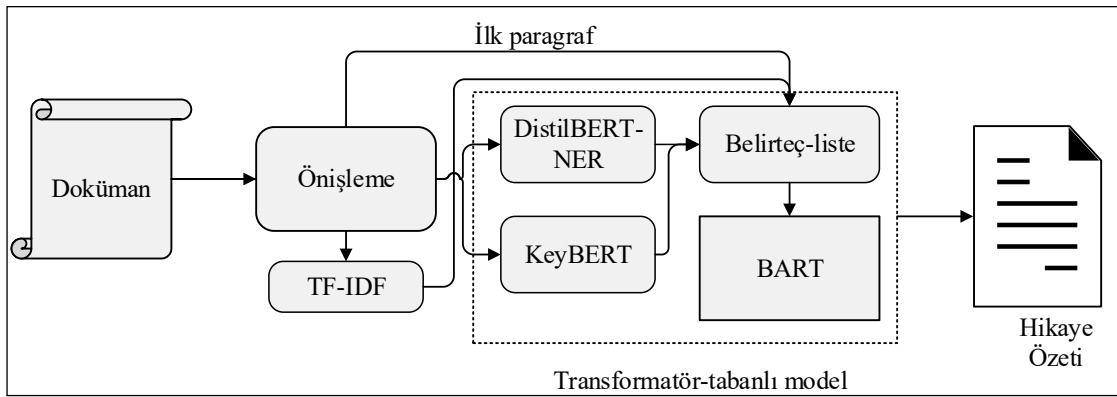
## 2. Önerilen Yöntem

Bu çalışmada, DistilBERT-NER, KeyBERT ve TF-IDF tabanlı bir hikâye kitaplarını özetleme modeli önerilmiştir. Önerilen model, Şekil 1’de gösterilmiştir. İlk olarak, dokümanda bulunan durak kelimeler ve sayılar kaldırılmıştır. Ardından, DistilBERT-NER ile çıkarılan varlık isimleri ve KeyBERT kullanılarak yeni bir belirteç listesi oluşturulmuştur. Ayrıca, bu belirteç listesi dokümanın ilk paragrafına eklenmiştir. Transformatör tabanlı model, bu belirteç listesiyle ince ayar olarak oluşturulmuştur. Önerilen modelin algoritması aşağıda verilmiştir.

### *Doküman özetleme algoritması*

girdi  $\leftarrow$  önişleme(doküman)

1. Varlık ismi  $\leftarrow$  DistilBERT-NER (girdi)
2. Anahtar kelimeler  $\leftarrow$  KeyBERT (girdi, 3-gram) & KeyBERT (girdi, 4-gram) & TF-IDF(girdi)
3. Token-list  $\leftarrow$  Varlık ismi  $\cup$  Anahtar kelimeler  $\cup$  İlk-Paragraf
4. Hikâye Özeti  $\leftarrow$  BART (Token-list)



Şekil 1. Transformatör tabanlı özetleme modeli.

BART, transformatör tabanlı bir dil modelidir. Bu modelin girdisinde anahtar kelime çıkarımı (KeyBERT, TF-IDF) ve varlık ismi çıkarımı (DistilBERT-NER) yöntemleri ile çıkarılan kelimeler kullanılmıştır. BART modelinde, öncelikle girdi bilgisi kopyalanmaktadır. Modelin performansını artırmak için belirteç maskeleyme, belirteç silme, cümle sıralaması değiştirme, doküman döndürme ve metin doldurma gibi fonksiyonlar geliştirilmiştir. Daha az kelime kullanarak bu fonksiyonların işlevselliği de azaltılmıştır. KeyBERT, anahtar kelimeleri çıkarmak için transformatör tabanlı bir yaklaşımdır. Kelimeler arasındaki anlamsal ilişkileri belirlemede oldukça etkili bir yöntemdir. Anahtar kelimeler arasındaki benzerlikler kosinüs benzerliği ile hesaplanmıştır. Bu çalışmada, 3-gram ve 4-gram olacak şekilde anahtar kelimelerin çıkarımı yapılmıştır. Bu kelimeler varlık isimleri ve metnin ilk paragrafıyla birleştirilerek hikâyenin bölümünün özetlenmesi amaçlanmıştır.

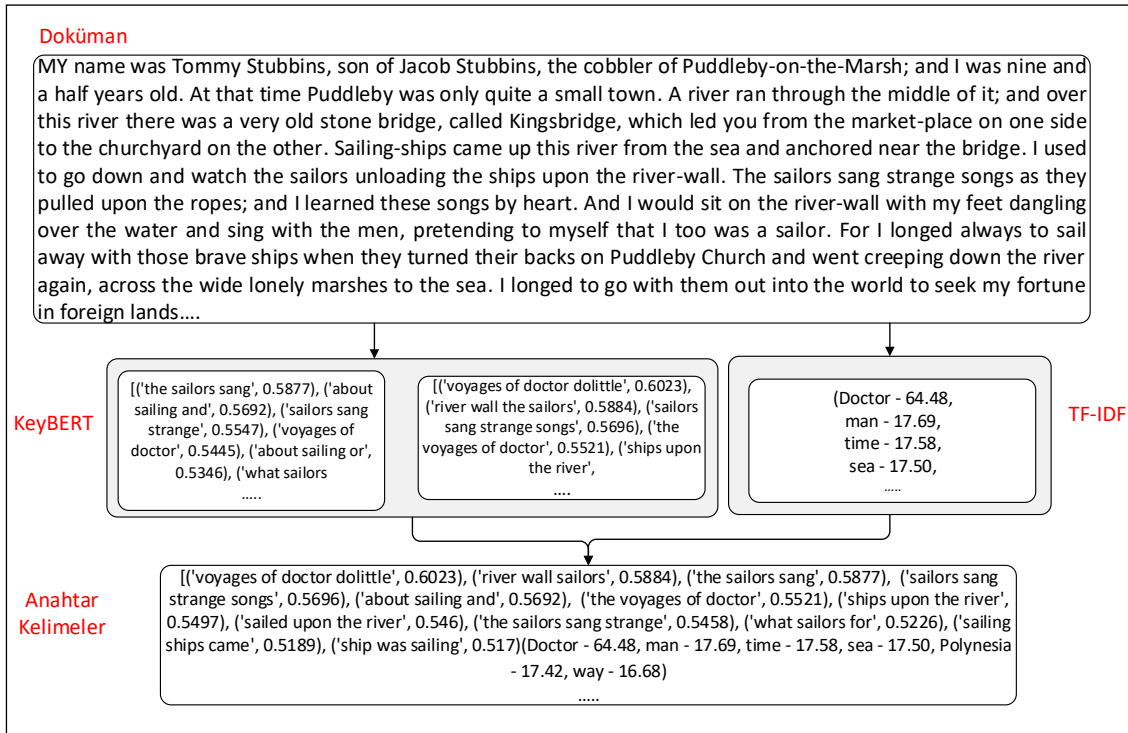
### 2.1. KeyBERT tabanlı anahtar kelime çıkarımı

KeyBERT, bir dokümandaki anahtar sözcükleri BERT modeli ve kosinüs benzerliğine dayalı olarak çıkarma yöntemidir [27]. KeyBERT modeli olarak “distilbert-base-nli-mean-tokens” versiyonu kullanılmıştır. KeyBERT’in en önemli avantajlarından biri, anahtar kelime niteliği taşıyan kelime gruplarının çıkarılabilmesidir. Bu çalışmanın kapsamında, 3-gram ve 4-gram olacak şekilde anahtar kelimeler çıkarılmıştır. Ardından, her iki türde ortak olan anahtar ifadeler birleştirilmiştir. TF-IDF (Term Frequency-Inverse Document Frequency) algoritmasında, metindeki kelimelerin sıklığına göre içeriği en iyi şekilde tanımlayan kelimeler çıkarılmaktadır. TF-IDF değerlerini hesaplamak için, terim sıklığı (TF) ve ters doküman sıklığı (IDF) aşağıdaki Denklem 1’deki gibi hesaplanmaktadır.

$$TF - IDF_i = TF_{ij} * \log_e \frac{|D|}{Df_i} \quad (1)$$

Burada,  $TF_{ij}$  i kelimesinin j dokümanda kaç kez geçtiğini ifade etmektedir.  $|D|$  doküman sayısını ifade etmektedir.  $Df_i$  i kelimesini içeren dokümanların sayısını ifade etmektedir. Anahtar kelime seçimi aşağıda maddeler halinde sunulmuştur. Şekil 2’de ise anahtar kelime çıkarımı gösterilmiştir.

- Anahtar kelimeler, 3-gram ve 4-gram olmak üzere sıralama işlemine tabi tutulmaktadır.
- İki terim benzer içeriklerde yer alıyorsa ve aynı skora sahipse, en yüksek skora sahip olan terim seçilmektedir.
- “is, are, were, was” gibi bağlam dışı kelimeler anahtar kelimelerden tamamen çıkarılmaktadır.
- TF-IDF algoritmasıyla elde edilen anahtar kelimeler, skorlarına bakılmaksızın listenin sonuna eklenmektedir.



Şekil 2. KeyBERT-tabanlı anahtar kelime çıkarımı.

## 2.2. DistilBERT-NER

DistilBERT, BERT modelinin mimarisine benzer yapıya sahiptir. Ancak, DistilBERT, BERT’in hızının %97’sini korurken, BERT’te belirteç türü yerleştirmeleri ve havuzlama katmanlarını içermemektedir. NLP’de kelime gömme ve ön eğitimin yöntemlerinin kapsamlı analizi modern doğrusal cebir çerçeveleri, katman normalizasyonu ve doğrusal katman gibi Transformatör mimarisinde kullanılan işlemlerin çoğu için optimize etmiştir. Ayrıca, modelin gizli katmandaki değişikliklerin, katman sayısı gibi diğer parametrelerdeki değişikliklere kıyasla hesaplama verimliliği üzerinde daha az etkisinin olduğu gösterilmiştir. DistilBERT, katman sayısını azaltarak daha verimli bir model ortaya koymaktadır [28].

DistilBERT-NER, dokümandan varlık isimlerini çıkarmak için geliştirilmiş bir modeldir. Bu aşamada, DistilBERT-NER modelinin “distilbert-base-multilingual-cased-ner-hrl” versiyonu kullanılmıştır. Çıktı katmanının aktivasyon fonksiyonu GELU olarak seçilmiş ve nöron sayısı, varlık etiket kategorilerinin sayısına göre eşit olarak belirlenmiştir. Böylece, dokümandan varlık türleri olarak (yerler- LOC, organizasyonlar- ORG ve kişiler- PER) çıkarımı yapılmıştır. DistilBERT-NER ile çıkarılan tüm varlık isimleri, anahtar kelimelerle birlikte kullanılmaktadır.

### 3. Araştırma Bulguları ve Tartışma

Önerilen model, bu çalışma kapsamında hazırlanan StoryBookSum veri seti ile temel yöntemlerle karşılaştırılmıştır.

#### 3.1. Veri seti

Önerilen modeli değerlendirmek için 7500 İngilizce hikâye kitabı bölümünden oluşan StoryBookSum veri setini önerilmiştir. Kitap bölümleri, model tarafından üretilen özetlerin değerlendirilmesi için kullanılmıştır. Veri seti, hazırlanan kitapların tüm bölümlerini içermektedir. Belirteç listesi, bölümün anahtar kelimeleri, varlık isimleri ve bölümün ilk paragrafı çıkarılarak hazırlanmıştır. Tablo 1’de veri seti özelliklerine yer verilmiştir.

**Tablo 1.** StoryBookSum veri kümesi özellikleri.

| Veri seti    | Doküman Sayısı | Kaynak      | Ort. Varlık ismi |
|--------------|----------------|-------------|------------------|
|              |                | Ort. kelime |                  |
| StoryBookSum | 7500           | 15557,4     | 15,14            |

#### 3.2. Değerlendirme Ölçütleri

Rouge metriği (Recall Oriented Understudy of Gisting Evaluation), yazar tarafından oluşturulan özet ile model tarafından oluşturulan özete benzerliğini n-gram tabanlı olarak hesaplayan ölçüm yöntemidir [29]. Yaygın olarak kullanılan üç değerlendirme metriği ROUGE-1, ROUGE-2 ve ROUGE-L’dir. En uzun ortak sıralı söz öbeğinin (LCS- Longest Common Subsequence) örtüşmesine dayalı hesaplama yöntemi Rouge-L özetleme modelinin performansını göstermek için diğerlerinin ölçüm yöntemlerinden daha fazla tercih edilmektedir. ROUGE-N’in hesaplaması, Denklem 2’te verilmiştir.

$$ROUGE - N = \frac{\sum_{S \in \{referans\}} \sum_{gram_n \in S} C_{eşleşme}(gram_n)}{\sum_{S \in \{referans\}} \sum_{gram_n \in S} C(gram_n)} \quad (2)$$

Burada,  $gram_n$  n kelimeli sözcükleri,  $C(gram_n)$  hedef özette  $gram_n$  bulunma sıklığı,  $C_{match}(gram_n)$  kelime kümesi  $gram_n$  ‘in hedef ve referans özette bulunma sıklığını ifade etmektedir. Denklem 3, 4 ve 5’te detaylı olarak ROUGE-L hesaplanmasına yer verilmiştir.

$$Anma_{LCS} = \frac{LCS(X, Y)}{m} \quad (3)$$

$$Duyarluluk_{LCS} = \frac{LCS(X, Y)}{n} \quad (4)$$

$$F_{LCS} = \frac{2 * Anma_{LCS} * Duyarluluk_{LCS}}{(Anma_{LCS} + Duyarluluk_{LCS})} \quad (5)$$

Burada,  $X$  model tarafından üretilen özet,  $Y$  orijinal özet ifade etmektedir. LCS ise en uzun ortak alt diziyle çıkarılan kelime sayısını,  $m$  ise orijinal özetteki n gramların sayısı,  $n$  ise hedef özetteki n-gramların sayısını ifade etmektedir.

ROUGE skoru, üretilen özet ile gerçek özetleri n-gramların ile örtüşmesine bağlı olarak karşılaştırmaktadır. Ancak, bu ölçüt tek başına üretilen özetlerin kalitesini kanıtlamak için yeterli değildir. Bu problemi ele almak amacıyla, üretilen özetler insan değerlendirme ölçütüne de tabi tutulmuştur. Fleiss’in kappa katsayısı, iki veya daha fazla gönüllü arasında yapılan karşılaştırmalı anlaşmanın güvenilirliğini ölçmek için kullanılmaktadır. Fleiss’in kappa katsayısı, Denklem 6’da hesaplanmaktadır. Bu denklemle ilgili gerekli hesaplamaları sağlayan eşitlikler Denklem 7, 8, 9 ve 10’da verilmiştir [30].

$$K = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} \quad (6)$$

$$\bar{P} = \frac{1}{N} \sum_i^N P_i \quad (7)$$

$$\bar{P}_e = \sum_{j=1}^k \rho_j^2 \quad (8)$$

$$\rho_j = \frac{1}{N * n} \sum_{i=1}^N n_{ij} \quad (9)$$

$$P_i = \frac{1}{n(n-1)} \sum_{i=1}^N n_{ij} (n_{ij} - 1) \quad (10)$$

Burada  $N$  değerlendirilecek örneklem sayısını,  $n$  toplam değerlendirici sayısını,  $k$  değerlendirmede kullanılacak kategori sayısını ifade etmektedir.  $K$  değeri 0 ile 1 arasında bir değer olarak hesaplanmaktadır. Bu değer 1'e ne kadar yakınsa değerlendiriciler arasındaki tutarlılık o kadar yüksektir.

### 3.3. Temel Yöntemler

Önerilen modelin performansı, temel modellerle karşılaştırılmıştır. Bu modeller, haber makaleleri, finansal metinler ve diyalog metinleri gibi veri setleriyle eğitilerek oluşturulmuş özetleme modelleridir. Bu modeller, hikâye kitaplarıyla karşılaştırılabilir oldukları için tercih edilmiştir. Temel yöntemler aşağıda maddeler halinde açıklanmıştır.

- T5, metin tabanlı dil problemlerini ele almak için geliştirilmiş bir dil modelidir [31]. Bu çalışmada, haber verileriyle eğitilmiş T5 tabanlı dil modeli kullanılmıştır [32].
- **BART<sub>XSUM</sub>**, BBC makalelerinden toplanan 226.711 haber makalesinden oluşan XSum veri setiyle önceden eğitilmiş modeldir [14, 33].
- **BART<sub>Large</sub>**, CNN/DailyMail veri setiyle eğitilmiş bir temel dil modelidir [14, 33]. Bu çalışmada, veri setiyle eğitilmiş BART model kullanılmıştır [34].
- PEGASUS, özellikle soyutlama metni özetleme olmak üzere çeşitli NLP görevlerini gerçekleştirmek için büyük veri kümeleriyle Transformatör tabanlı bir kodlayıcı-kod çözücü modelidir [35]. Bu çalışmada, finansal dokümanlarla eğitilmiş PEGASUS dil modeli kullanılmıştır [36].

### 3.4. Değerlendirme Sonuçları

Önerilen modelin performansı, StoryBookSum veri seti üzerinde temel yöntemlerle karşılaştırılarak değerlendirilmiştir. Değerlendirme sonuçları, Tablo 2'de sunulmuştur. Bu çalışmada, hikâye kitaplarının tamamını okumak yerine, kitaptaki anahtar kelime, varlık isimleri ve ilk paragrafı ile anahtar çıkarımı yapılarak oluşturulan belirleyici kelimelerle özet oluşturulması hedeflenmiştir.

Tablo 2'deki sonuçlara göre, önerilen model, Rouge-L %15,88 F1 değeriyle diğer modellere kıyasla en iyi performansı sergilemektedir. Bu sonuçlardan önerilen modelin kelimeler ve varlık isimleri arasındaki ilişkileri daha etkili bir şekilde çıkarabildiği ve daha başarılı özetler oluşturduğu görülmektedir. Bu çalışma kapsamında, varlık ismi olarak kişi isimleri, yer isimleri gibi bilgi çıkarımları yapılmıştır. Tablo 2'deki sonuçlardan insanlar veya yerler arasındaki ilişkilerin belirlenmesiyle girişten daha az belirteç listesiyle bir kitap özeti oluşturulabilir. Ancak burada önemli olan, varlık isimleri arasındaki anlamsal ilişkilerin belirlenmesidir. Bu ilişki çıkarımı olmadığı için hikâye bölümünün ilk paragrafı belirteç listesine dâhil edilmiştir. T5, BART ve PEGASUS, önerilen modele göre düşük sonuçlar elde etmiştir. Bu fark, modellerin eğitim verileri ve hedef görevlere özgü olmalarından kaynaklanmaktadır.

Önerilen model, StoryBookSum veri setiyle eğitilmiştir. Bu veri kümesi hikâyenin belirleyici kelimeleriyle eğitilerek oluşturulmuştur. Bu çalışmada, giriş olarak verilen belirteç listesinden hikâyeyi özetleyebilmesi amaçlanmaktadır. Dolayısıyla model, hikâyenin bütünüyle değil hikâyenin bir kısmından özetlenmesi odaklanmaktadır. Haliyle dokümandaki kelimelerin tamamı verilmediği için Rouge-1/2/L değerleri de özetleme

için çok büyük başarı sağlamadığı ancak özetleme için bir temel oluşturduğu ve diğer modellerden daha üstün olduğu da görülmektedir. Ayrıca, hikâyeler bölüm-bölüm ele alınmıştır. Her bir bölüme ait özet bölümü bulunmamaktadır. Model tarafından çıkarılan özetler bölüm bazında karşılaştırıldığı için Rouge-1/2/L değerleri yüksek çıkmamıştır. Bu modelin başarısının artmasında belirteç listesine varlık isimlerinin yanı sıra ilişki çıkarımı ya da anahtar kelime sayısının artırılmasıyla hikâye oluşturmanın başarısını arttıracaktır.

Hikâye kitaplarında bölümlere ait özetlerin olmaması ve her bir bölüme ait özetin bölüm bazında karşılaştırılması Pre-trained modellerin başarısını da etkilediği görülmektedir. Hikâye kitapları, yapısı gereğince bölüm şeklinde düzenlenmiş dokümanlardır. Bölümlerde kişi, konum gibi bilgilerin korunması hikâyedeki bütünlüğün sağlanması için oldukça önemli bir husustur. NER ile bu bilgilerin çıkarımı yapılmış anahtar kelimeler ve ilk paragrafla birleştirilerek modelin oluşturulması gerçekleştirilmiştir.

Tablo 2, önerilen modelin dokümanın tamamı olmadan temel modellerden daha başarılı bir özet üretildiği görülmektedir. Önerilen modelin, diğer yöntemlere karşılaştırıldığında Rouge-L değerinin %15,88 F1 değeriyle özet üretmede daha başarılı olduğu görülmektedir. Kelimeler arasındaki ilişkiyi açığa çıkaran ifadelerin ve belirteç listesinin olmaması, veri seti boyutu, eğitim kümesi özelliği bu değerlerin ortaya çıkmasının ana nedenidir.

**Tablo 2.** Değerlendirme sonuçları.

| Model                      | StoryBookSum Veriseti |                 |                 |
|----------------------------|-----------------------|-----------------|-----------------|
|                            | ROUGE-1<br>(F1)       | ROUGE-2<br>(F1) | ROUGE-L<br>(F1) |
| T5-finetuned               | 12,08                 | 3,96            | 11,56           |
| <b>PEGASUS</b>             | 3,85                  | 0,70            | 3,25            |
| BART-finetuned             | 10,07                 | 5,51            | 9,34            |
| <b>BART<sub>XSUM</sub></b> | 8,92                  | 0,87            | 8,07            |
| <b>Önerilen model</b>      | 19,15                 | 5,36            | 15,88           |

Tablo 2'deki sonuçlardan, T5-finetuned ve BART-finetuned modellerinde yaklaşık 4 ROUGE-L değeri farkıyla başarılı bir şekilde özetler ürettiği görülmektedir. En düşük ROUGE-L puanı ise daha kısa özetler üreten **BART<sub>XSUM</sub>** ve Pegasus modellerinde elde edilmiştir. Önerilen modelde, özet üretmenin, temel yöntemlerle özet üretmekte daha başarılı sonuçlar ürettiği görülmektedir. Tablo 2'de görüldüğü üzere değerlendirme Rouge-N değerleriyle karşılaştırılmıştır. Rouge-L değeri, orijinal özet ile model tarafından üretilen özetin arasındaki en uzun ortak alt diziyi (LCS) temel alarak hesaplanmaktadır. Bu hesaplama, üretilen özet ile orijinal özet arasındaki eşleşen en uzun ortak alt diziyeye dayalı olarak yapılır. Bu hesaplama yönteminde, üretilen hikâye bölümü ile orijinal hikâye bölümündeki özetler arasındaki en uzun ortak alt dizi belirlenir. Denklem 3, 4, ve 5'e göre ROUGE-L değeri hesaplanmaktadır. Bu değer, iki farklı özet arasında örtüşen en uzun kelime ortak alt dizinin kelime sayısına, üretilen özetin uzunluğuna, orijinal özetin uzunluğuna bağlıdır.

Önerilen model, bir kitap bölümünü hızlı bir şekilde özetlemek için, anahtar kelimeler, varlık isimleri ve bölümün ilk paragrafından oluşan bir belirteç listesi ile bir temel oluşturmaktadır. Kelime sayısının sınırlı olması nedeniyle, üretilen özet ile orijinal özet arasındaki en uzun ortak alt dizi (LCS) değeri de sınırlı kalmaktadır. Haliyle LCS değerinin özet uzunluğuna oranı bu değerlerin %15,88 çıkmasına neden olmuştur. Ancak, önerilen modelin hikâye kitabındaki temel niteleyici kelimelerle özet oluşturmak için temel oluşturduğu bu sonuçlardan açık bir şekilde görülmektedir.

Tablo 3'te, önerilen modele göre yapılan insan değerlendirme sonuçları verilmiştir. Değerlendirme kriteri olarak beş ölçüt dikkate alınmıştır. Değerlendirme kriterleri aşağıdaki gibidir:

- Kısıklık: Üretilen özetin gereksiz bilgi içermemesi,
- Bilgilendiricilik: Üretilen özetin orijinal özetteki gibi önemli bilgiler içermesi,
- Tutarlılık: Üretilen özetin içeriği ile orijinal özetin içeriğinin örtüşmesi,
- Okunabilirlik: Üretilen özetin anlaşılabilir olması,
- Dilbilgisi: Üretilen özetin dilbilgisi kurallarına uymasındır.

Yukarıdaki kriterlere göre, beş kişiye modelin ürettiği özetlerin bölümün içeriğiyle ne kadar uyumlu olduğu sorulmuştur. Her bir katılımcıya rastgele 20 doküman seçmesi ve bunları 0 (en kötü) ile 5 (en iyi) arasında bir puanla değerlendirmesi istenmiştir. Değerlendirme sonuçlarına Tablo 3’te verilmiştir.

**Tablo 3.** İnsan değerlendirme sonuçları.

|              | Kısalık     | Bilgilendiricilik | Tutarlılık | Okunabilirlik | Dilbilgisi  |
|--------------|-------------|-------------------|------------|---------------|-------------|
| Ort/Var.     | 1,60 (0,18) | 2,50(0,21)        | 2,40(0,27) | 4,58(0,21)    | 4,65 (0,12) |
| Fleiss kappa | 0,28        | 0,31              | 0,34       | 0,65          | 0,69        |

Fleiss’in kappa katsayısı 1’e ne kadar yakınsa, gönüllüler arasındaki tutarlılık o kadar yüksektir [27]. Tablo 3’teki sonuçlara göre, önerilen modelin dilbilgisi ve okunabilirlik kriterleri açısından iyi yazılmış cümleler ürettiğini göstermektedir. Diğer kriterler ROUGE L ölçütü ile birlikte değerlendirildiğinde, modelin bölümün özetini oluşturduğu görülmektedir. Ancak bu kriterler için iyileştirme yapılması gerekmektedir. Tablo 4, önerilen model tarafından üretilen özetleri içermektedir. Tablo 4’te orijinal kaynak olarak dokümana ait metinlere yer verilmiştir. Bu metinler hikâye bölümünün bir kısmını içermektedir.

**Tablo 4.** Önerilen modelle üretilen hikâye özetleri.

| Hikaye-<br>No | Orijinal kaynak  | Model tarafından üretilen özet   |
|---------------|--|--|
| #1            | My name was Tommy Stubbins, son of Jacob Stubbins, the cobbler of Puddleby-on-the-Marsh; and I was nine and a half years old. At that time Puddleby was only quite a small town. A river ran through the middle of it; and over this river there was a very old stone bridge, called Kingsbridge, which led you from the marketplace on one side to the churchyard on the other.....   | Tommy Stubbins grew up in Puddleby, a small town on the Marsh. His father Jacob was a cobbler and his mother was a housewife. He recalls sailing with his father and playing with the cats and dogs. He also recalls his friend Matthew Mugg, the cat’s-meat-man, and his father’s mussel-boat. The story ends with a trip to the seaside with Tommy’s father, to buy mussels and lobsters for the family’s dinner.  |
| #2            | The servants at the Hotel de Chalusse, one of the most magnificent mansions in the Rue de Courcelles in Paris, were assembled in the porter’s lodge, a little building comprising a couple of rooms standing on the right hand side of the great gateway. Here, as in all large mansions, the “concierge” or porter, M. Bourigeau, was a person of immense importance, always able and disposed to make any one who was inclined to doubt his authority, feel it in cruel fashion. As could be easily seen, he held all the other servants in his power. He could let them absent themselves without ..... | She insisted on acting alone and persuaded others to follow suit. nearly you She, why am I so exceptional? this quickly from the state of young servants The general was thrilled to discover that his uncle’s friend had been defying foes by undermining. He servants and he smiled miserably. Could He, sir, have been an army there abruptly across the street from the general? Usually, it’s a cunning, ugly life that is revealed. The c’s study calculated startled everyone more, and the gentleman twenty hoped it wasn’t an irascible friend who was worried about matters. He laughed and said, It’s his money man, bachelor, fresh little banks known leave house ..... |
| #3            | There was a heavy silence in the air. The lights were dim, and the shadows deep. A silhouette appeared in the corner of the room, its footsteps echoing softly. The woman, holding the knife in her hand, carefully scanned her surroundings. Her eyes seemed lost in something, as if searching for something unseen. Every movement, every glance, seemed to signal impending danger. Suddenly, a sound broke the silence, a faint whistle, followed by .....  | The waiters’ gloves were neatly folded. After a time of restoration, she murmured to him, and he frowned as he saw his terrified body lying in CY. Her BA brightly lying there made him notice the balcony and intensify the veranda. Emerging hands ordered KA clearly to observe as he tried to take steps and a married figure in the night room lamps heard him signaling some calmness and shuddering blood he lit softly as he traveled. The woman questioned, looking, reached down and sternly nodded down the knife..   |



#### 4. Sonuç

Hikâye kitapları genellikle bölüm şeklinde düzenlenmiş uzun dokümanlardır. Dolayısıyla dokümandaki akışı sağlayarak metni özetlemek oldukça zordur. Bu çalışmada, anahtar kelime çıkarma ve varlık isimlerinin eşleştirilmesi için TF-IDF, KeyBERT ve DistilBERT-NER yöntemlerini kullanarak, hikâye kitabı gibi bölümlerin özetlenmesine yönelik yeni bir Transformer tabanlı model önerilmiştir. Önerilen model, TF-IDF ve KeyBERT'ten çıkarılan anahtar kelimelerin kesişimini, DistilBERT-NER aracılığıyla çıkarılan varlık isimlerini birleştirerek tutarlı özetler üretmeyi amaçlamaktadır. Her bölümün kısa bir bölümle özetlenmesi amaçlanmaktadır. Bu nedenle, belirteç listesi bölümün ilk paragrafı, transformatör tabanlı varlık ismi çıkarma, transformatör tabanlı anahtar kelime çıkarma ve TF-IDF anahtar kelime çıkarma yardımıyla üretilmiştir. Transformatör tabanlı model, belirteç listesi ile ince ayarla eğitilmiştir. Önerilen model, hikâye kitabı bölümlerinden çıkarılan kelimelerle transformatör tabanlı model kullanarak özetleme için temel bir yaklaşım sunmaktadır. Çalışma sonucunda elde edilen veri önerilen modelin performansını ve temel modellere göre üstünlüğünü göstermektedir. Önerilen modelin özetleri daha etkili bir şekilde ürettiğini göstermektedir. İnsan değerlendirme sonuçlarından üretilen özetlerin okunabilirlik ve dil bilgisi açısından uyumlu olduğu gözlemlenmiştir. Bu model, varlık isimleri gibi daha fazla niteleyici kelimelerle anlamsal ilişki çıkarımı modelin başarısının artacağını göstermektedir. Gelecekteki çalışmalarda, özetlere varlık isimlerinin arasındaki anlamsal ilişkilerin eklenmesiyle modelin iyileştirilmesi amaçlanmaktadır.

#### Kaynaklar

- [1] Alomari A, Idris N, Sabri AQM and Alsmadi I. Deep reinforcement and transfer learning for abstractive text summarization: A review, *Comput Speech Lang* 2022; 71:101276.
- [2] Li P, Lu W and Cheng Q. Generating a related work section for scientific papers: an optimized approach with adopting problem and method information, *Scientometrics* 2022; 127(8), 4397-4417.
- [3] Cai X, Liu S, Yang L, Lu Y, Zhao J, Shen D and Liu T. COVIDSum: A linguistically enriched SciBERT-based summarization model for COVID-19 *J Biomed Inform* 2022; 127:103999.
- [4] Altmami NI and Menai MEB. Automatic summarization of scientific articles: A survey, *J King Saud Univ-Comput Inf Sci* 2020; 34(4), 1011-1028.
- [5] Wang Z, Duan Z, Zhang H, Wang C, Tian L, Chen B, Zhou M. Friendly topic assistant for transformer based abstractive summarization. *Empir Methods Nat Lang Process* 2020; 485-497.
- [6] Koncel-Kedziorski R, Bekal D, Luan Y, Lapata M and Hajishirzi H. Text generation from knowledge graphs with graph transformers, *arXiv preprint arXiv:1904.023422*, 019.
- [7] Huang Z and Xie Z. A patent keywords extraction method using TextRank model with prior public knowledge , *Complex Intell Syst* 2022; 8(1), 1-12.
- [8] Qiu D, Zheng Q. Improving TextRank Algorithm for Automatic Keyword Extraction with Tolerance Rough Set. *Int J Fuzzy Syst* 2022; 24(3).
- [9] Khademi ME, Fakhredanesh M. Persian automatic text summarization based on named entity recognition. *Iran J Sci Technol Trans Electr Eng* 2020; 1-12.
- [10] Du Y, Zhao Y, Yan J and Li Q. UGDAS: Unsupervised graph-network-based denoiser for abstractive summarization in biomedical domain, *Methods* 2022; 203, 160-166.
- [11] Xiao L, He H, Jin Y. FusionSum: Abstractive summarization with sentence fusion and cooperative reinforcement learning. *Knowl-Based Syst* 2022; 243: 108483.
- [12] Moirangthem DS and Lee M. Abstractive summarization of long texts by representing multiple compositionality with temporal hierarchical pointer generator network, *Neural Networks* 2020; 124, 1-11.
- [13] Liang Z, Du J and Li C. Abstractive social media text summarization using selective reinforced Seq2Seq attention model, *Neurocomputing* 2020; 410, 432-440.
- [14] Lewis M, Liu Y, Goyal N, Ghazvininejad M, Mohamed A, Levy O, Stoyanov V, Zettlemoyer L. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension, *arXiv preprint arXiv:1910.13461*, 2019.
- [15] Yu T, Su D, Dai W and Fung P. Dimsum@ laysumm 20: Bart-based approach for scientific document summarization, *arXiv preprint arXiv:2010.09252*, 2020.
- [16] Bajaj A, Dangati P, Krishna K, Ashok Kumar P, Uppaal R, Windsor B, Brenner E, Dotterrer D, Das R, McCallum A. Long Document Summarization in a Low Resource Setting using Pretrained Language Models, *arXiv preprint arXiv:2103.00751*, 2021.
- [17] Christian H, Agus MP, Suhartono D. Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF). *ComTech* 2016; 7(4): 285-294.
- [18] Deng Z, Ma F, Lan R, Huang W and Luo X. A two-stage Chinese text summarization algorithm using keyword information and adversarial learning, *Neurocomputing* 2021; 425:117-126.
- [19] Li C, Xu W, Li S, Gao S. Guiding generation for abstractive text summarization based on key information guide network. *NAACL HLT* 2018; 2: 55-60.

- [20] Salton G, Buckley C. Term-weighting approaches in automatic text retrieval. *Inf Process Manag* 1988; 24(5): 513-523.
- [21] Li J, Tang T, Zhao WX, Wei Z, Yuan NJ and Wen JR. Few-shot knowledge graph-to-text generation with pretrained language models, arXiv preprint arXiv:2106.01623, 2021.
- [22] Xiaoye W and Mizuho I. Extractive Summarization Utilizing Keyphrases by Finetuning BERT-Based Model. In *International Conference on Asian Digital Libraries*, Cham: Springer International Publishing, 2022; 59-72.
- [23] Yoo C, Lee H. Improving Abstractive Dialogue Summarization Using Keyword Extraction. *Appl Sci* 2023; 13(17): 9771.
- [24] Zhao Z, Hou Y, Wang D, Yu M, Liu C and Ma X. Educational question generation of children storybooks via question type distribution learning and event-centric summarization. arXiv preprint arXiv:2203.14187, 2022.
- [25] Ling Z, Xie Y, Dong C and Shen Y. Enhancing Factual Consistency in Text Summarization via Counterfactual Debiasing. In *Proceedings of the 31st International Conference on Computational Linguistics 2025*, pp. 7912-7924.
- [26] Upadhyay A, Bhavsar N, Bhatnagar A, Singh M and Motlicek P. Automatic summarization for creative writing: BART based pipeline method for generating summary of movie scripts. In *Proceedings of The Workshop on Automatic Summarization for Creative Writing*, 2022, pp. 44-50.
- [27] Liu Y, Maier W, Minker W and Ultes S. Empathetic dialogue generation with pre-trained RoBERTa-GPT2 and external knowledge. In *Conversational AI for Natural Human-Centric Interaction: International Workshop on Spoken Dialogue System Technology*, Singapore: Springer Nature Singapore, 2022; 67-81.
- [28] Sanh V, Debut L, Chaumond J and Wolf T. DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter, arXiv preprint arXiv:1910.01108, 2019.
- [29] Lin CY, Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 2024; 74-81.
- [30] Faloutico R, Quatto P. Fleiss' kappa statistic without paradoxes. *Qual Quant* 2015; 49(2): 463-470.
- [31] Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, Zhou Y, Li W, Liu PJ. Exploring the limits of transfer learning with a unified text-to-text transformer. *J Mach Learn Res* 2020; 21(140): 1-67.
- [32] İnternet: T5-base fine-tuned fo News Summarization URL: <https://huggingface.co/mrm8488/t5-base-finetuned-summarize-news>, Son Erişim Tarihi: 08.12.2024.
- [33] Xu J, Desai S and Durrett G. Understanding neural abstractive summarization models via uncertainty. arXiv preprint arXiv:2010.07882, 2020.
- [34] İnternet: T5-base fine-tuned fo News Summarization URL: <https://huggingface.co/philschmid/bart-large-cnn-samsum>, Son Erişim Tarihi: 08.12.2024.
- [35] Zhang J, Zhao Y, Saleh M and Liu P. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *International Conference on Machine Learning*, 2020; 11328-11339.
- [36] İnternet: T5-base fine-tuned fo News Summarization URL: <https://huggingface.co/human-centered-summarization/financial-summarization-pegasus>, Son Erişim Tarihi: 08.12.2024.