

## **BiLSTM Derin Öğrenme Yöntemi ile Uzun Metinlerden Yeni Özet Metinlerin Türetilmesi**

Onur ŞAHİN<sup>1\*</sup>, Rıdvan YAYLA<sup>2</sup>

### **Öz**

Günümüzde, derin öğrenme tekniklerinin yaratıcı metin oluşturma süreçlerine entegrasyonu, araştırmacılar ve yazılım geliştiriciler arasında yaygın olarak kullanılmaktadır. Derin öğrenme, büyük veri setleri üzerinde karmaşık bilgileri öğrenme yeteneği ile bilinir ve bu özellik, dil anlama ve metin üretme konularında önemli avantajlar sunar. Çalışmanın amacı Bidirectional Long Short-Term Memory (BiLSTM) derin öğrenme mimarisi temelinde uzun bir metindeki ifadelerin eğitilerek yeni bir özet metnin türetilmesidir. Bu amaç doğrultusunda Türkçe dilinde yazılmış farklı uzunlukta ve türdeki (Roman, Kişisel Gelişim) iki doküman veriseti olarak kullanılmış, veri setindeki metinler veri temizleme, tokenizasyon ve vektörleştirme gibi bir dizi önışlemeden geçirilmiştir. Çalışma, BiLSTM'nin yanı sıra LSTM, GRU, BiGRU ve CNN gibi diğer derin öğrenme mimarilerini de değerlendirmiş ve BiLSTM modelinin iki farklı kitap türünde ve farklı kelime sayılarında (1.000, 2.000 ve 5.000 kelime) en yüksek METEOR, BLEU ve ROGUE skorlarına sahip olduğunu ortaya koymuştur. Bu bulgular, BiLSTM'nin metin özetleme ve metin üretme için diğer modellere göre daha başarılı sonuçlar ürettiğini göstermektedir. BiLSTM kullanarak belli bir roman veya kişisel gelişim kitabından yaratıcı ve özgün metinler üretme yöntemi araştırmacılar ve yazılım geliştiriciler için ilham verici bir kaynak olup, önerilen yöntemin farklı metin türleri için de uygulanabileceği öngörülmektedir. Bu sayede, metin özetleme ve üretme süreçlerinde BiLSTM mimarisinin başarılı sonuçlar ürettiği gösterilmiştir.

**Anahtar Kelimeler:** BiLSTM, Doğal dil işleme, Metin üretme, Derin öğrenme.

## **New Summary Text Generation from Long Texts With BiLSTM Deep Learning Method**

### **Abstract**

Nowadays, the integration of deep learning techniques into creative text generation processes is widely used among researchers and software developers. Deep learning is known for its ability to learn complex information over large data sets, and this feature offers significant advantages in language understanding and text generation. The aim of the study is to derive a new summary text by training the expressions in a long text on the basis of Bidirectional Long Short-Term Memory (BiLSTM) deep learning architecture. For this purpose, two documents of different lengths and types (Novel, Personal Development) written in Turkish were used as a dataset, and the texts in the dataset were subjected to a series of pre-processes such as data cleaning, tokenization and vectorization. The study evaluated other deep learning architectures such as LSTM, GRU, BiGRU and CNN, as well as BiLSTM, and found that the BiLSTM model had the highest METEOR, BLEU and ROGUE scores in two different book types and different word counts (1,000, 2,000 and 5,000 words). showed. These findings show that BiLSTM produces more successful results than other models for text summarization and text generation. The method of generating creative and original texts from a certain novel or personal development book using BiLSTM is an inspiring resource for researchers and software developers, and it is envisaged that the proposed method can be applied for different text types. In this way, it has been shown that the BiLSTM architecture produces successful results in text summarization and generation processes.

**Keywords:** BiLSTM, Natural language processing, Text generation, Deep learning.

<sup>1</sup>Bilecik Şeyh Edebali Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği ABD. Bilgisayar Mühendisliği Yüksek Lisans Programı, Bilecik, Türkiye, [onur2000sahin@outlook.com](mailto:onur2000sahin@outlook.com)

<sup>2</sup>Bilecik Şeyh Edebali Üniversitesi Mühendislik Bilgisayar Mühendisliği Bölümü, Bilecik, Türkiye, [ridvan.yayla@bilecik.edu.tr](mailto:ridvan.yayla@bilecik.edu.tr)

<sup>1</sup><https://orcid.org/0009-0000-8955-658X> <sup>2</sup><https://orcid.org/0000-0002-1105-9169>

## 1. Giriş

Dil, insanların birbirleriyle iletişim kurmasını, fikirlerini ve duygularını ifade etmesini, bilgi paylaşmasını ve kültürlerini oluşturmasını sağlar. Ayrıca, şiir, müzik, edebiyat ve sanat gibi yaratıcı ifade biçimlerinin de temelini oluşturur. Bu temel oluşması, dilin zengin semantik yapılarının çeşitli biçimlerde kullanılması ile mümkün olmaktadır. İnsan dili, karmaşıklığı ve yaratıcı ifade gücü ile makine üstü bir iletişim aracıdır. Günümüzde sürekli olarak gelişen yapay zekâ teknikleri, dildeki semantik karmaşıklığın detaylı bir biçimde incelenmesine imkân sağlamaktadır.

Derin öğrenme, belli algoritmalarındaki karmaşık görevlerin öğrenilmesini sağlayan yapay sinir ağı temelli bir yapay zekâ alanıdır. Öğrenme modelleri, örüntü tanıma, doğal dil işleme ve makine çevirisi gibi geniş bir uygulama yelpazesine sahiptir. Bu modeller, metin sınıflandırma, bilgi çıkarımı, duygu analizi ve özetleme gibi doğal dil işleme görevlerinde etkin bir şekilde kullanılmaktadır.

Yapay zekâ temelinde oluşturulan modeller temelde yinelemeli sinir ağları (RNN) ile veya buna bağlı eğitilmiş modeller kullanılarak inşa edilir. RNN'ler, önceki girdileri hatırlama ve sonraki girdilerle ilişkilendirme yeteneğine sahiptir (Sherstinsky, 2020).

Bilginin öğrenme modellerinde geçici hafızada tutulması, dilin uzun vadeli bağımlılıklarını anlama açısından önemli bir kriterdir. LSTM, bu kriteri sağlamak için, unutma kapısı, girdi kapısı ve çıkış kapısı olarak üç ana özellikten oluşmaktadır. Bu özellikler, öğrenme modelinde dil karmaşıklığını anlama, yaratıcı ve yeni metinler üretme yeteneğini sağlar.

Gated Recurrent Unit (GRU), LSTM'nin bir türüdür ve RNN'lerdeki problemleri çözmek için tasarlanmıştır. GRU, daha basit bir yapıya sahip olup, özellikle hafıza sorunlarına çözüm sunarak dilin uzun vadeli bağımlılıklarını anlamada etkili olabilir. Bidirectional GRU (BiGRU) ve Bidirectional LSTM (BiLSTM) modelleri, metin içindeki bağlamsal ilişkileri anlamada daha gelişmiş bir yaklaşım sunar. Bu modeller, sadece önceki girdilere değil, aynı zamanda sonraki girdilere de odaklanarak bilginin daha kapsamlı bir şekilde öğrenilmesine olanak tanır. Convolutional Neural Network (CNN), özellikle görüntü işleme alanında popüler olan bir derin öğrenme mimarisidir. Metin işleme için uyarlanmış CNN modelleri, belirli metin özelliklerini vurgulayarak ve öğrenerek dilin yapılarını anlamada kullanılabilir.

Bu çalışmada Türkçe dilinde yazılmış bir roman ve yabancı bir dilde yazılmış ancak Türkçeye çevrilmiş bir kişisel gelişim kitabı olarak iki farklı metin türü ele alınmıştır. Bu, modelin farklı dilsel yapıları ve konuları anlama yeteneğini test etmek açısından önemlidir. Roman metni, karmaşık dil yapısı ve edebi anlatım tarzı ile karakterize edilirken, Kişisel gelişim türünde daha genel bir dil kullanılmaktadır. Ayrıca çalışmanın birinci yazarı tarafından yazılan Türkçe roman metni, önerilen yöntemin yeni fikirler ve yaratıcı metinler çıkarma potansiyeline işaret etmektedir. Bu kapsamda çalışma sadece metin özetleme değil, aynı zamanda yaratıcı ve özgün metinlerin türetilmesi amacını

da içermektedir. Geleneksel özetleme yöntemleri, metindeki önemli bilgileri özetlemek için kullanılan belirli cümleleri seçmeye odaklanırken, BiLSTM, bu çalışma kapsamında metnin dilbilgisi yapısı ve anlamsal içeriğini anlayarak yeni ve özgün cümleler üretmektedir. Bu özellik, edebi yaratıcılık, öyküleme ve otomatik metin üretimi alanlarında yeni olasılıklar sunmaktadır.

Bu çalışmada, BiLSTM mimarisinin yanı sıra LSTM, GRU, BiGRU ve CNN gibi farklı derin öğrenme mimarileri kullanılarak uzun metinlerden oluşan bir verisetinin eğitilmesiyle ayrı modeller oluşturulmuş ve bu modeller yeni üretilecek özet metinler için analiz edilmiştir. Bu analizlerin doğrulukları Meteor, Bleu, Rogue gibi metin üretimi ve özetlemede sıklıkla kullanılan metriklerle gerçekleştirilmiştir.

## 2. Materyal ve Yöntem

Günümüzde makine öğrenmesi ve doğal dil işleme temelinde oluşturulan Yapay zekâ uygulamaları yaygın bir şekilde kullanılmaktadır. Derin öğrenme modellerinin çok katmanlı ve gelişmiş yapısı sayesinde geleneksel makine öğrenmesi yöntemlerine göre daha başarılı sonuçlar üretilebilmektedir. Derin öğrenme mimarileri içerisinde özellikle BiLSTM (Bidirectional Long Short-Term Memory), metin işleme uygulamalarında kullanılan bir tür yinelemeli sinir ağıdır. Bu mimari, analiz edilen verinin etkili bir şekilde öğrenilebilmesini sağlar. BiLSTM, bilgiyi uzun süreli olarak saklayabilme özelliğine sahip olmasıyla ön plana çıkar. Ayrıca, gereksiz bilgileri unutma yeteneği sayesinde, işe yarar özelliklere odaklanır ve bu sayede veri içerisindeki önemsiz detayları filtreler. Ayrıca sadece önceki girdilere değil, aynı zamanda sonraki girdilere de dikkat ederek bilginin daha kapsamlı bir şekilde öğrenilmesine imkân tanır. Bu çalışmada, örnek bir roman ve kişisel gelişim kitabındaki tüm metinler veriseti olarak BiLSTM mimarisine giriş verisi olarak aktarılmış ve bu metinlerden örnek bir özet metin çıkış verisi türetilmiştir.

### 2.1. Literatür Taraması

Literatür çalışmalarında BiLSTM yöntemleri, metin sınıflandırma, duygu analizi ve görüş madenciliği gibi alanlarda daha çok kullanılmaktadır. Ancak metinler arasındaki semantik ilişkilerden yararlanarak özellik çıkarımı için yeni metinler oluşturmak, daha az rastlanan bir çalışma alanıdır. Son yıllarda, derin öğrenme modellerinin, özellikle BiLSTM ve gömme (embedding) katmanlarının dil anlama ve metin üretme yöntemlerinde daha başarılı olduğu gözlemlenmektedir.

Young ve ark. tarafından yapılan çalışmada, derin öğrenme mimarilerinin (LSTM, GRU, CNN) doğal dil işleme alanındaki gelişmeleri incelenerek, çeşitli doğal dil işleme görevlerinde kullanılan modellerin başarı oranları analiz edilmiş ve karşılaştırılmıştır. Ayrıca derin öğrenme modellerinin

doğal dil işleme alanındaki avantajlarına ve gelecekteki çalışmaların hangi alanlara döneceğine ilişkin bilgiler detaylandırılmıştır (Young ve ark, 2018). Sutskever ve ark. çalışmalarında, tekrarlayan sinir ağları (RNNs) üzerindeki zorlukların Free-Hessian optimizasyonu ile aşılabildiğini göstermiş, çalışmada karakter düzeyinde dil modelleme görevlerinde uygulanan yeni bir RNN çeşidi önerilmiştir (Sutskever ve ark., 2011). Çarpımsal bağlantılar içeren bu RNN türünde, standart mimarilere göre daha etkili sonuçlar elde edilebilmektedir. Ayrıca çalışmada, dilin karakter düzeyindeki modellemesinin zorluğuna rağmen, bu yöntemin güçlü bir öğrenme sistemi olduğu vurgulanmıştır. Bu öğrenme sisteminin, kelimeleri kolayca öğrendiği ve eğitim setinde görünmeyen uygun kelimelerin olasılığını hesaplayabildiği belirtilmektedir.

Yu ve ark. Doğal dil işlemenin temel konularından biri olan metinden yeni metin oluşturma için genel ve özel teknikleri açıklamıştır (Yu ve ark., 2022).

Zhao ve ark. çalışmasında, büyük dil modellerinin (LLM) metin oluşturma performanslarını incelemişlerdir. LLM'lerin yüksek kaliteli ve tutarlı metin üretme yeteneklerini değerlendirmişler ve bu modellerin gelecekteki sınırlamalarına ve güvenlik risklerine değinmişlerdir. Büyük dil modellerinin etkililiğini artırmak için daha fazla ince ayar, sağlam eğitim ve kontrollü üretim yöntemleri önerilmiştir (Zhao ve ark., 2023).

Vo ve ark. çalışmasında, meta-öğrenme ve BiLSTM kullanarak açıklanabilir bir özet çıkarma modelini ele almaktadır. Araştırma, sınırlı veri ile performansı artırmayı amaçlayan meta-öğrenme algoritması kullanarak derin öğrenme modelini eğitiyor ve modelin karar verme sürecine dair içgörü sağlamak için SHAP, doğrusal regresyon ve karar ağaçları gibi çeşitli açıklama tekniklerini uygulamıştır. Sonuçlar, modelin performansını değerlendirerek açıklanabilirlik ve etkinlik açısından önemli bulgular sunmaktadır (Vo ve ark., 2024).

Iqbal ve Qureshi tarafından gerçekleştirilen çalışma derin öğrenme alanındaki gelişmelerin özellikle Doğal Dil İşleme (NLP) bağlamında metin üretme modellerine etkisini ele almaktadır. 2015 yılından itibaren yayınlanan bir dizi çalışmayı inceleyerek derin öğrenme modellerinin geçmişi, bugünü ve geleceğini detaylı bir şekilde analiz etmeyi amaçlamaktadır. Çalışma, kelime yerleştirmeleri, RNN'ler, CNN'ler, VAE'ler ve GAN'lar gibi popüler derin öğrenme modellerini inceleyerek bu modellerin metin üretme sorunlarına nasıl yaklaştığını ele almaktadır (Iqbal ve Qureshi, 2020). Li ve ark. tarafından gerçekleştirilen çalışma, önceden eğitilmiş dil modellerin Pre-trained language models (PLM) kullanımına odaklanan bir anket niteliğindedir. Bu çalışma, derin öğrenme yöntemlerinin metin oluşturma alanındaki gelişmeleri ve özellikle PLM'lerin metin üretimine olan katkılarını incelemektedir. Ayrıca, PLM'leri metin oluştururken karşılaşılan temel zorlukları ve bu zorlukların çözüm önerilerini tartışmaktadır. Çalışma PLM'leri metin üretimine uygulamanın üç temel yönünü ele almaktadır. Bunlar, girdi temsili öğrenimi, etkili bir PLM modelinin tasarımı ve referans metni verilen PLM'lerin etkili bir şekilde optimize edilmesiyle ilgili

konuları içermektedir. Her bir yöntem, mevcut zorluklarla birlikte detaylı bir şekilde incelenmiştir (Li ve ark., 2022).

Otter ve ark. tarafından gerçekleştirilen çalışma, doğal dil işleme (NLP) alanındaki derin öğrenme modellerinin kullanımına odaklanarak, alandaki son gelişmeleri sistematik bir şekilde ele almaktadır. Çalışma, derin öğrenme mimarileri ve yöntemleri konusunda bir genel bakış sunarak, çok sayıda güncel araştırmayı incelemekte ve bu araştırmaların alanına katkılarını özetlemektedir. Araştırma, hesaplamalı dilbilim, temel dilsel işleme konuları ve çeşitli uygulama alanlarına odaklanarak geniş bir yelpazede yapılmıştır. Derin öğrenme modellerinin, özellikle transformatör birimleri yığınlarının, NLP görevlerinde üstün performans sergilediği vurgulanmıştır. Çalışma, ön eğitim ve transfer öğrenimi gibi tekniklerin önemine değinerek, derin öğrenme modellerinin mevcut durumunu tartışmaktadır (Otter ve ark. 2021). Babüroğlu ve ark. tarafından gerçekleştirilen çalışma, Türkçe için derin öğrenme tabanlı doğal dil işleme modeli geliştirmeyi amaçlamaktadır. çalışmada, doğal dilin önemi ve karmaşıklığı vurgulanarak, bilgisayar ortamında doğal dil işlemenin zorlukları ele alınmaktadır. Derin öğrenme yöntemlerinin, yapay zekâ teknolojilerinde dil bilimi alanında kullanımına vurgu yapılmış ve özyinelemeli sinir ağları (RNN) özellikle sıralı verilerin işlenmesinde kullanışlı bulunmuştur. Çalışmada önerilen model, iki yönlü uzun-kısa vadeli bellek (BiLSTM) kullanılarak Türkçe sözcük etiketleme amacıyla geliştirilmiştir. Bu model, doğal dil araştırmacılarına kendi analizlerini gerçekleştirme ve kullanabilme imkânı sunan bir platform olarak tasarlanmıştır. BLSTM kullanılarak geliştirilen platform, uzman görüşleri ve geri bildirimlerle sürekli olarak iyileştirilmiş ve sözcük etiketleyici hata oranı azaltılmıştır (Babüroğlu ve ark., 2019). Semeniuta ve ark. tarafından gerçekleştirilen çalışma metin üretimi için Değişken Otomatik Kodlayıcı (VAE) üzerindeki mimari seçimlerin etkisini araştırmaktadır. Geleneksel VAE modellerinden farklı olarak, önerilen hibrit mimari, tekrarlayan bir dil modeli ile tamamen ileri beslemeli evrişimli ve ters evrişimli bileşenleri birleştirmektedir. Bu yeni mimari, hızlı çalışma süresi, yakınsama yeteneği ve uzun dizileri daha etkili bir şekilde işleme özellikleriyle dikkat çekmektedir. Çalışma, öne sürülen mimarinin VAE modellerinin metinsel veriler üzerindeki eğitim zorluklarını aşmada etkili olduğunu göstermektedir. İleri beslemeli kısmın eğitimi, VAE'yi daha kolay hale getirirken, kod çözücünün standart dil modeline dönmesi, örnekleme engelleyen bir sorunu önlemektedir. Ayrıca, gizli vektöre güvenmeyi teşvik etmek için ek bir maliyet terimi önerilmiş ve VAE'nin uzun dizilerde iyi çalıştığı gözlemlenmiştir (Semeniuta ve ark., 2017).

Onan çalışmasında, dikkat mekanizması tabanlı derin öğrenme mimarilerini kullanarak Türkçe duygu analizi için karşılaştırılmalı bir çalışma gerçekleştirmiştir. Çalışmada küresel dikkat mekanizması tabanlı LSTM ve GRU, kendine dikkat mekanizması tabanlı LSTM ve GRU gibi farklı dikkat mekanizmalarını inceleyerek duygu analizi için daha doğru sınıflandırma oranları elde etmiştir (Onan, 2022). Onan diğer bir çalışmasında, Türkçe duygu analizi için farklı derin öğrenme

mimarilerini deęerlendirmiştir. Yapılan deneylerde, çalışan bellek bağlantıları tabanlı uzun kısa süreli bellek (LSTM) mimarisinin dięer mimarilere göre daha yüksek doęruluk oranlarına ulaştığı gözlemlenmiş, GloVe kelime gömme yönteminin kullanımıyla dięer yöntemlere göre daha başarılı sonuçlar elde edilmiştir (Onan, 2022).

Chakraborty ve ark. tarafından yapılan çalışmada, karakter tabanlı kelime üretme modelleri için Uzun Kısa Süreli Bellek hücreleri (LSTM'ler) ve sıralı veri işleme için tek boyutlu Evrişimli LSTM ağlarının performansı incelenmiştir. Çalışmada, LSTM hücre sayısının model başarısına etkisini deęerlendirilmiş; artan hücre sayısının başlangıçta iyileşme sağladığı ancak belirli bir noktadan sonra aşırı uyum sorunlarına yol açtığını belirtilmiştir. Bu kapsamda LSTM hücre sayısı daha az olduğu takdirde daha başarılı sonuçlar elde edileceği ifade edilmiştir (Chakraborty ve ark., 2020). Welleck ve ark. metin oluşturma alanında sıralı olmayan düzenlerde çalışan yeni bir model geliştirmiştir. Geleneksel metin oluşturma yöntemlerinin aksine, önerilen yöntem rastgele bir konumda kelime üretmek ve ardından kelimenin soluna ve sağına kelimeler türeterek ikili ağaç oluşturma yöntemi ile yapılmıştır. Yapılan deneysel çalışmalar, bu yöntemin geleneksel soldan sağına metin oluşturma yöntemi ile rekabet edebilecek bir performans sergilemiştir. Ayrıca, modelin anlamı koruyacak şekilde çeviri yapmayı öğrendiği ve dil modelleme görevlerinde etkili olduğu gözlemlenmiştir (Welleck ve ark., 2019). Zhu ve ark. çalışmasında farklı metin oluşturma platformlarını mukayese eden, metinlerin çeşitliliğini, kalitesini ve tutarlılığını deęerlendiren Taxygen isimli bir metin oluşturma platformu tasarlamıştır (Zhu ve ark., 2018).

Bayer ve ark. tarafından gerçekleştirilen çalışmada, sınıflandırıcıların eğitim verilerini yapay olarak artırmak için yeni bir metin üretme yöntemi sunulmuş ve önerilen yöntem hem uzun hem de kısa metin sınıflandırıcılarının performansını deęerlendirmek amacıyla kullanılmıştır. Veri artırma yöntemi, küçük veri analitiği bağlamında %15,53 ve %3,56'ya kadar ilave doęruluk kazançları elde edilmiştir. Bu yaklaşımın düşük veri görevlerinde önemli iyileştirmeler sağladığı da belirtilmiştir. Ayrıca, yöntemin her tür veri kümesine uygun olmadığı durumlar da gözlemlenerek, metin üretme yönteminin çeşitli dönüşümleri başarılı bir şekilde gerçekleştirebildiği belirtilmiştir (Bayer ve ark., 2022). Fang ve ark. kullanıcı gizliliği endişelerinden yola çıkarak, iki kullanıcının şifreli mesaj alışverişini, bir saldırganın bu alışverişini tespit etmeden gerçekleştirmesine olanak tanıyan steganografik bir mesajlaşma sistemi tasarlamıştır. Twitter ve Enron e-posta veri kümeleri üzerinden yapılan araştırmada, şifreleme kapasitesi (kelime başına şifrelenmiş bit sayısı) artırılırken, yüksek kalitede steganografik metin üretilebildiği gösterilmiştir (Fang ve ark., 2017). Fatima ve ark. metin üretme konusunda derin sinir ağı modellerini incelemiştir. Ayrıca çalışmada geçmişten günümüze çeşitli araştırmacıların bu alandaki önemli katkılarına ve metin üretme konusundaki literatür boşluklarına, gelecekte yapılabilecek muhtemel çalışmalara yer verilmiştir. Bulgular, metin üretimi alanındaki çeşitli yönlerdeki eğilimleri ve gelişmeleri ortaya koyarak araştırmacılara kılavuz

olabilecek zengin bir bilgi kaynağı sunmaktadır (Fatima ve ark., 2021). Guo ve ark. tarafından gerçekleştirilen çalışmada, derin öğrenme gelişmeleriyle birlikte metin üretme teknolojisinin nasıl değiştiği ve koşullu metin üretimi üzerine odaklanan yeni araştırma eğilimleri incelenmektedir. Metin oluşturma teknolojisinin antropomorfik ve duygusal öğelerle zenginleştirilmesi, kişiselleştirilmiş metin üretimi gibi konulara odaklanılarak, bu alandaki değişen teknik yönler ele alınmıştır. Ayrıca, mevcut koşullu metin üretimi alanlarını ve genel öğrenme modellerini içeren çeşitli öneriler sunulmuştur. Çalışmada, koşullu metin üretimi alanında devam eden araştırma sorunlarına ve gelecekteki potansiyel araştırma yönlerine de vurgu yapılmaktadır (Guo ve ark., 2020). Li ve Zhang tarafından gerçekleştirilen çalışma, metin oluşturma alanında Long Short-Term Memory (LSTM) tabanlı bir araştırmaya odaklanmaktadır. Çalışmada LSTM ve geliştirilmiş varyantları olan Gated Recurrent Unit (GRU) ve görsel delik bağlantılı LSTM'nin performansı karşılaştırılmıştır. Çalışma LSTM'nin uzun vadeli bağımlılık özelliği ile uzun metinlerde daha iyi performans gösterdiğini belirtmektedir. Gelecekteki dil modellerinin, BERT puanı ve BLEURT gibi otomatik değerlendirme göstergeleri kullanarak daha geniş ölçekli değerlendirmelerle optimize edilebileceği ve iyileştirilebileceği vurgulanmaktadır (Li ve Zhang, 2021).

Pawade ve ark. tarafından gerçekleştirilen çalışma tekrarlayan sinir ağı (RNN) ve uzun kısa vadeli bellek (LSTM) kullanarak metin oluşturma sistemini ele almaktadır. Bu sistem, giriş olarak alınan bir dizi hikâyeden ilham alarak yeni hikayeler oluşturmayı amaçlamaktadır. Farklı konu ve karakterlere sahip hikayelerin yanı sıra aynı hikâyenin farklı ciltleriyle de çalışarak sistemin ürettiği sonuçları dilbilgisi doğruluğu, olayların bağlantısı, ilgi düzeyi ve benzersizlik gibi parametrelerle analiz etmiştir. Sonuç olarak, çalışmada sistem tarafından üretilen hikayelerin insanlar tarafından değerlendirilmesi sonucunda %63 oranında doğruluk elde edildiği belirtilmektedir. Ayrıca, sistemde kullanılan parametrelerin değiştirilerek doğruluğun daha da artırılabilirliği ve sistemin genişletilerek farklı uygulamalarda kullanılabilirliği öne sürülmektedir. Bu çalışma, metin oluşturma alanında yapay zekâ temelli yöntemlerin potansiyelini vurgulamaktadır (Pawade ve ark., 2018). Samant ve ark. tarafından gerçekleştirilen çalışma, doğal dil anlama (NLU) ve doğal dil işleme (NLP) alanındaki derin öğrenme tabanlı dil modellerini sistemli bir şekilde incelemektedir. Çalışma, çok görevli öğrenme (MTL) kullanan dil modellerinin performansını artırmayı hedefleyerek, NLU ve NLP görevlerinde kullanılan çeşitli dil modellerini sistematik bir literatür taraması ile değerlendirmektedir. Sonuçlar, denetimsiz öğrenme yöntemlerinin dil modellerinin performansını artırabileceğini göstermektedir. Ancak, çok görevli NLU için genel amaçlı bir çerçevenin tasarlanmasının zorluğu, dil modelleri için genel bir temsil oluşturma karmaşıklığına işaret etmektedir (Samant ve ark., 2022). Shi ve ark. tarafından gerçekleştirilen çalışma, metin oluşturma alanında çeşitliliği artırmak amacıyla ters takviyeli öğrenmeyi (IRL) kullanmaktadır. Maruz kalma yanlılığı, ödül kıtlığı ve modun çökmesi gibi sorunları çözmek için IRL çerçevesini benimseyen bu yöntem, daha yoğun ödül

sinyalleri üretme yeteneğine sahiptir. Yöntemin iki temel avantajı vurgulanmıştır: Birincisi, ödül fonksiyonu daha yoğun ödül sinyalleri üretebilir, ikincisi ise "entropi düzenlenileştirilmiş" politika eğimiyle eğitilen üretim politikasının daha çeşitli metinler üretmeyi teşvik etmesidir. Yapılan deney sonuçları, önerilen yöntemin önceki modellere göre daha kaliteli metinler üretebildiğini göstermektedir. Çalışma, önerilen yöntemin yanı sıra BLEU puanına dayalı üç yeni değerlendirme ölçüsü de sunarak, metinlerin daha iyi değerlendirilmesine katkıda bulunmaktadır (Shi ve ark., 2018).

Santhanam tarafından gerçekleştirilen çalışma, içerik tabanlı metin üretimi için Long Short-Term Memory (LSTM) ağlarını kullanarak dil modellerini incelemektedir. Bu çalışmada, LSTM birimlerinin doğal dil üretimi modellerindeki etkisi ele alınmakta ve dil modelinin bağlamı anlama yeteneğini artırmak amacıyla bağlam vektörleri kullanılmaktadır. Çalışmada önerilen model, metin oluştururken belirli bir bağlam etrafında çalışabilen dil modellerini eğitmek üzere tasarlanmıştır. Bağlam vektörleri, cümlenin anlamsal anlamını kavrayan bir yapıya sahiptir ve dil modelinin oluşturduğu metinlerde anlamsal tutarlılığı artırmaya yöneliktir. Çalışmada, bağlam vektörlerini çıkarmak için farklı yöntemlerin yanı sıra kelime kümeleme ve kelime önemi gibi faktörlerin etkisi üzerine çeşitli analizler sunulmuştur. Önerilen modelin, bağlam temelli metin üretiminde daha iyi performans gösterdiğini ve sağlanan bağlama anlamsal yakınlık sağladığını göstermektedir. Ayrıca, kelime kümeleme yönteminin bağlam vektörlerini daha etkili bir şekilde çıkardığı ve modelin eğitimini değerlendirmek için kullanılan yöntemin, aşırı öğrenme durumlarını belirlemede etkili olduğu belirtilmiştir (Santhanam, 2020).

Erhandi ve Çallı çalışmasında yapay zekâ ve derin öğrenme teknikleri metin özetleme için başarılı bir şekilde uygulanmamıştır. Yapay sinir ağları temel alınarak geliştirilen derin öğrenme modeli ile, Keras kütüphanesi kullanılarak Türkçe metinler üzerinde başarılı sonuçlar elde edilmiştir. Türkçe ve İngilizce veri setleri üzerinde yapılan gözlemler ve sonuçlar, derin öğrenme modelinin Türkçe için etkili bir şekilde çalıştığını göstermiştir (Erhandi ve Çallı, 2020). Mishra ve ark. tarafından gerçekleştirilen çalışma, biyomedikal alanda metin özetleme konusunda son yıllarda yapılan araştırmaları sistematik olarak incelemektedir. Çalışma, klinisyenler ve klinik araştırmacılar için giderek artan bilgi miktarıyla başa çıkma ihtiyacından yola çıkarak, metin özetleme tekniklerinin biyomedikal alanda nasıl geliştirildiğini ve değerlendirildiğini ele almaktadır. Çalışmanın bulgularına göre, doğal dil işleme ve istatistiksel yöntemlerin bir kombinasyonunu içeren hibrit tekniklerin metin özetleme alanında öne çıktığı belirtilmiştir. Ayrıca, bu alanda yapılan araştırmaların çoğunluğunun içsel değerlendirme gerçekleştirdiği ve gelecekteki araştırmalara yönelik önerilere de yer verildiği görülmektedir (Mishra ve ark., 2014). Ünlü ve Çetin tarafından gerçekleştirilen çalışma derin öğrenme yöntemlerinin anahtar kelime ve anahtar ifade çıkarımı üzerindeki etkilerini incelemeyi amaçlamaktadır. Çalışmada, teknolojik gelişmelerin ardından ortaya çıkan yüksek miktardaki veri ve bu verilerin anlamlandırılma zorluğu ele alınmıştır. Geleneksel istatistiksel yöntemlerin yanı sıra,



derin öğrenme ve yapay sinir ağları gibi tekniklerin anahtar kelime ve anahtar ifade çıkarımı problemlerine uygulanmaya başlandığı belirtilmiştir. Derin öğrenme yöntemlerinin, geleneksel istatistiksel ve Doğal Dil İşleme (NLP) tekniklerine göre daha iyi sonuçlar verdiği vurgulanarak, makine çevirisi, anahtar kelime çıkarımı, özetleme gibi alanlarda derin öğrenme yöntemlerinin kullanımının arttığı ifade edilmiştir. Çalışmada, anahtar kelime ve anahtar ifade çıkarımında kullanılan derin öğrenme yöntemleri kategorik olarak sınıflandırılmış ve literatürdeki bu yöntemlere dair yapılmış çalışmalar incelenmiştir (Ünlü ve Çetin, 2019).

Metin özetleme veya yeni metin yaratımı süreçlerinde performans değerlendirme amacıyla sıkça kullanılan hata metrikleri arasında METEOR BLEU ve ROUGE, öne çıkmaktadır. Literatürde, bu metriklerin birlikte kullanıldığı birçok çalışma bulunmaktadır (Zhang ve ark., 2024; Mao ve ark., 2024; Zhou ve ark., 2023; Wu ve ark., 2023). ROUGE, BLEU ve METEOR metriklerinin aynı anda kullanıldığı projelerde, özetleme ve metin üretimi modellerinin doğruluğu ve etkinliği kapsamlı bir şekilde değerlendirilmiştir. Bu metrikler, farklı açılardan performans analizi yaparak, modellerin referans metinlerle benzerliğini, n-gram düzeyinde doğruluğunu ve anlamsal uygunluğunu ölçmeyi sağlamaktadır. Bu sayede, metin oluşturma ve özetleme modellerinin performansını nesnel bir şekilde değerlendirmek mümkün olmaktadır.

Literatürde verilen örneklerde görüldüğü gibi, makine öğrenimi ve doğal dil işleme alanındaki derin öğrenme modellerinin metin oluşturma üzerindeki etkilerini ve uygulamalarını değerlendirmeyi amaçlamaktadır. Çeşitli çalışmalar, özellikle uzun kısa vadeli bellek (LSTM) ve diğer derin öğrenme mimarilerinin dil anlama, duygu analizi, metin üretimi gibi görevlerde etkileyici başarılar elde ettiğini göstermektedir.

Literatürde incelenen çalışmalardan bazıları, Türkçe duygu analizi için derin öğrenme modellerinin etkinliğini değerlendirirken, diğerleri metin oluşturma, dil modellemesi ve çeşitli NLP görevlerinde derin öğrenme yöntemlerini ele almaktadır. Ayrıca, metin oluşturma alanında yeni modeller ve çerçeveler üzerine yapılan çalışmalardan da bahsedilmiştir.

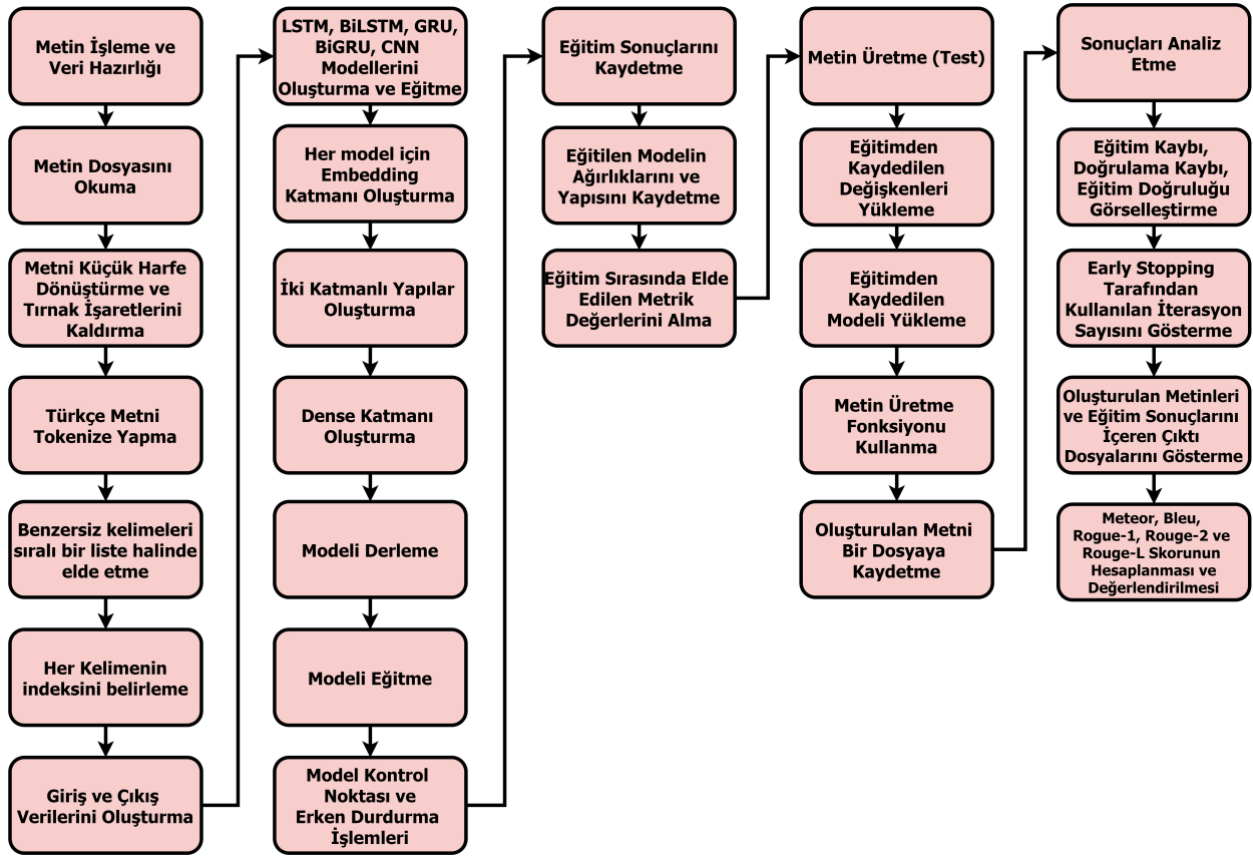
## 2.2. Sistem Tasarımı

Sistem tasarımı bölümü, metin işleme, veri hazırlama, derin öğrenme modellerini oluşturma ve eğitme aşamalarını içermektedir. Bu adımlar, dil modelinin karmaşık dil yapılarını anlamasını sağlamak için belirlenmiş adımlardan oluşmaktadır. Her bir adım, modelin dil öğrenme yeteneklerini artırmayı ve metinden anlamlı sonuçlar üretebilmesini hedeflemek amacıyla tasarlanmıştır.

Geleneksel dil modelleri, sınırlı bir dil yapısı ve metin analizi yöntemlerine sahip oldukları için özellikle uzun metinlerin semantik yapısını çözümleme konusunda zayıf kalmaktadırlar. Bu durum, özellikle metinlerin daha geniş bir anlamsal yapı içinde sınıflandırılmasında birtakım zorluklara

neden olmaktadır. Bu zorluklar, yapay zekâ temelindeki dil modellerinin oluşturulmasına ve derin öğrenme mimarileri temelinde eğitilerek geliştirilmesini sağlamıştır.

Metin özetleme, doğal dil işleme alanında önemli bir konsept olup, çeşitli yöntemlerle gerçekleştirilebilmektedir. Uzun metin içinden özetleme (Extractive Summarization) yönteminde belirli cümleler veya paragraflar, metindeki önem sırasına göre seçilir ve bu seçimler özet olarak kullanılır. Bu yaklaşım, TF-IDF veya benzeri ölçütlerle cümlelerin önemini değerlendirerek metni özetler. Diğer bir önemli yöntem ise metinden özet oluşturma (Abstractive Summarization) işlemidir. Bu yöntemde, metnin bağlamını anlayarak yeni cümleler üretilir ve bu cümleler, metnin özeti olarak kullanılır. Bu çalışmada, metinden özet oluşturma yöntemi benimsenmiştir. Bu sayede, metnin anlamını daha iyi kavramak ve daha özgün özetler oluşturmak hedeflenmiştir. Çalışmanın tüm adımlarını içeren diyagram Şekil 1’de gösterilmektedir.



Şekil 1. Önerilen Yöntemin Çalışma Prensibi

Çalışma metin işleme ve veri hazırlığı ile başlayıp modellerin eğitimi ile devam eder. Kaydedilen eğitim verileri derlendikten sonra metin üretme aşamasına geçilir. Metin üretme aşamasında elde edilen metinler Meteor, Bleu ve Rouge metrikleri ile analiz edilmiş, sonuçlar karşılaştırılmıştır.

### 2.2.1. Veriseti

Çalışmada iki farklı edebi türde kitaplar veriseti olarak kullanılmıştır. Kullanılan verisetinin ilk kitabı, 2020 yılında yayımlanan ve makalenin 1.yazarı tarafından kaleme alınan 250 sayfalık “Roman” türündeki bir kitabın metinlerinden oluşmaktadır. Kitap, 6910 cümle ve 65207 kelimedenden oluşmaktadır. Kullanılan verisetinin ikinci kitabı 2018 yılında yayımlanan 208 sayfalık “Kişisel Gelişim” türündeki bir kitabın metinlerinden oluşmaktadır. Kitap 3897 cümle ve 46077 kelimedenden oluşmaktadır. Metin özetleme, bu iki kitaba ait cümle ve kelimelerden oluşturulan çeşitli derin öğrenme mimarileriyle eğitilen modellerin sonuçlarından elde edilerek geliştirilmiştir.

### 2.2.2. Metin Önışleme

Metin işleme aşaması, dil modelinin derin öğrenme mekanizmalarıyla anlamlandırabileceği bir formata dönüştürme sürecini içermektedir. Bu aşamalardan ilki, metin içeriğindeki tüm karakterlerin küçük harfe dönüştürülmesini içermektedir. Bu adım, modelin büyük harf kullanımına bağlı olarak dil örüntülerini tespit etme konusunda problem yaşamamasını sağlamak amacıyla uygulanmıştır. Büyük ve küçük harflerin modele benzer şekillerde temsil edilmesi, dil öğrenme sürecindeki tutarlılığı arttırmaktadır.

İkinci adımda metin içeriğindeki tırnak işaretleri temizlenmiştir. Bu işlem, modelin tırnak işaretleri arasındaki özel karakterleri anlamdan bağımsız bir şekilde öğrenmesine yardımcı olmaktadır. Tırnak işaretlerinin kaldırılması, dil modelinin doğrudan kelime anlamını öğrenmesine ve dil yapısını daha iyi anlamasına katkı sağlar. Üçüncü olarak kitaplardan elde edilen kelimelerin sayısal veriye dönüştürülmesi (tokenization) işlemini kapsamaktadır. Bu adım, metni anlamlandırmak ve dil örüntülerini çıkarmak için gereklidir. Kelime tokenizasyonu, metni analiz etmek için kelimeleri ayırarak modelin daha iyi anlamasına ve dil yapılarını daha etkili bir şekilde öğrenmesine yardımcı olur. Bu aşamalar, dil modelinin girdi verilerini daha homojen ve anlamlı bir hale getirerek eğitim sürecinin başarılı olmasına katkıda bulunur.

### 2.2.3. Veri Hazırlama

Veri hazırlamanın ilk aşaması, modelin kelimeler arasındaki ilişkileri anlamlandırabilmesi için temel bir ön koşul olan kelime numaralandırma işlemidir. Her bir benzersiz kelimeye bir numara atanmıştır. Her kelimenin benzersiz bir sayısal temsilini kullanmak, modelin kelimeler arasındaki hiyerarşisini anlamasına ve bu kelimeleri sayısal olarak işleyebilmesine olanak tanır. Örneğin,

"merhaba" kelimesine atanan indeks, modelin bu kelimeyle ilişkilendirilmiş diğer kelimelerle olan sıklık ilişkisini ifade eder.

Öğrenme dizileri ve etiketlerin oluşturulması adımıyla ise belirlenen maksimum dizi uzunluğu ve adım aralığı kullanılarak, veriseti içerisinde dolaşarak öğrenme dizileri ve etiketler oluşturulmuştur. Bu adım, modelin dil örüntülerini çıkarmasına ve bir sonraki kelimenin tahminini gerçekleştirmesine olanak tanımaktadır. Her bir öğrenme dizisi, belirlenen maksimum dizi uzunluğunda bir dizi kelime içermektedir. Öğrenme dizilerinin yanı sıra, her bir öğrenme dizisine karşılık gelen etiketler de belirlenir. Bu etiketler, modelin bir sonraki kelimenin tahminini yapmasına yardımcı olur. Öğrenme dizileri ve etiketlerin doğru bir şekilde oluşturulması, modelin dil örüntülerini başarılı bir şekilde öğrenmesine olanak tanır ve dilin yapılarına duyarlılığını artırır.

Bu adımlar, dil modelinin eğitim verilerini etkili bir şekilde oluşturarak, dilin karmaşıklıklarını öğrenmesine ve yeni metinler üretebilmesine olanak sağlar.

#### 2.2.4. Derin Öğrenme Modellerini Oluşturma

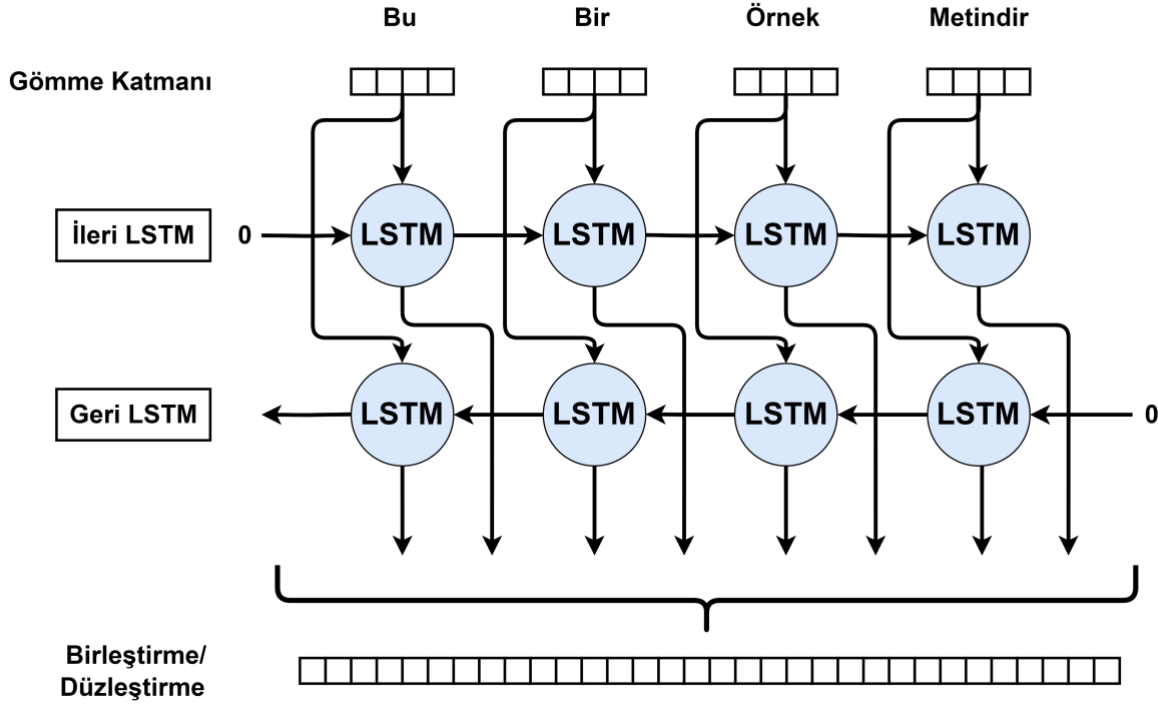
Bu bölümde, metin tabanlı derin öğrenme modelleri kullanılarak dil anlama ve metin üretimi üzerine karşılaştırmalı bir analiz yapılmıştır. Öncelikle, BiLSTM (Bidirectional Long Short-Term Memory) mimarisi temelinde oluşturulacak olan model hazırlanmıştır. Giriş verilerinin kelime gömme (embedding) işlemi için bir Embedding katmanı kullanılmıştır. Embedding katmanı, kelimeler arasındaki semantik ilişkileri ve benzerlikleri daha iyi anlamak adına kelime vektörlerini oluşturur. Bu sayede model, kelimelerin anlamsal bağlamını daha etkili bir şekilde kavrayabilir (Cui ve ark., 2018). Bu katman, benzer anlamlı kelimelerin benzer vektörlere sahip olmasını sağlayarak modelin dil örüntülerini daha etkili bir şekilde öğrenmesine imkân sağlar.

LSTM, (Long Short-Term Memory) derin öğrenme alanında kullanılan bir tür tekrarlayan sinir ağı (RNN) mimarisidir. Standart RNN'lerin aksine, LSTM'ler uzun vadeli bağlantıları daha etkili bir şekilde öğrenebilir ve bilgiyi daha uzun süre saklayabilir. LSTM, girdi, unutma ve çıkış kapıları olarak adlandırılan üç ana kapıdan oluşur. Girdi kapısı, yeni bilgileri kabul etme veya reddetme görevini üstlenir. Unutma kapısı, geçmiş bilgileri hatırlama veya unutma işlevini yerine getirir. Çıkış kapısı ise bilgiyi sonraki birimlere aktarmak için kullanılır. LSTM, bu kapıları kullanarak dilin dil bilgisi yapısını, anlamsal ilişkilerini ve hatta duygusal tonlarını öğrenebilir.

BiLSTM, LSTM'ye dayanan bir yapıdır, ancak bir adım daha ileri giderek metin içindeki bağlamı daha iyi anlamak için çift yönlü işleme olanak tanır. Bir BiLSTM, giriş verilerini hem ileri hem de geri yönde işleyen iki ayrı LSTM katmanını içerir.

İlk olarak, ileri yöndeki LSTM katmanı, metni baştan sona doğru okuyarak dil örüntülerini öğrenir. Bu, tüm zaman adımlarında çıktı üreterek uzun vadeli bağımlılıkları anlama kabiliyetini

sağlar. İkinci olarak, geri yöndeki LSTM katmanı, metni tersinden okuyarak benzer bir şekilde dil örüntülerini öğrenir. Bu iki yönlü işlem, metnin her iki yönünden gelen bağlamı birleştirerek daha kapsamlı bir dil anlama yeteneği sağlar. BiLSTM mimarisinin işleyiş yapısı Sekil 2’de gösterilmektedir.



Şekil 2. BiLSTM Mimarisinin Çalışma Prensibi

BiLSTM'nin bu çift yönlü yaklaşımı, özellikle metin içindeki bağlamın hem önceki hem de sonraki kısımları içermesi durumunda avantajlıdır (Huang ve ark., 2015). Bu, BiLSTM'nin daha uzun vadeli bağımlılıkları ve metnin genel yapısını daha etkili bir şekilde öğrenmesine olanak tanır.

Oluşturulan modelde iki adet BiLSTM katmanı bulunmaktadır. İlk katman, tüm zaman adımlarından çıktıları döndürerek uzun vadeli bağımlılıkları anlama yeteneği ile dil örüntülerini geniş bir bağlamda öğrenmiştir. İkinci katman ise, yalnızca son zaman adımının çıktısını kullanarak daha etkili öğrenme sağlamıştır. Bu yapı, modelin metnin daha önceki kısımlarından aldığı bilgileri koruyarak daha etkili bir öğrenme sağlar.

Modelin çıkış katmanı Tam bağlantılı (Dense) katmandır. Bu katman tüm benzersiz kelimeleri içeren bir çıkış vektörü üretir (Javid ve ark., 2021). Aktivasyon fonksiyonu olarak softmax kullanılarak, olasılık dağılımı elde edilmektedir. Bu sayede, modelin bir sonraki kelimenin olasılığını tahmin etmesi sağlanır ve metin üretim sürecinde çeşitlilik elde edilir (Wang ve ark., 2018).

BiLSTM modeli ile karşılaştırma yapmak üzere, LSTM, GRU, BiGRU ve CNN modelleri de oluşturulmuştur. LSTM, uzun vadeli bağımlılıkları başarıyla ele alabilen bir yapay sinir ağı modelidir.

GRU (Gated Recurrent Unit) ise benzer bir yapıya sahip olmasına rağmen daha az parametre kullanarak hafıza hücrelerini daha etkili bir şekilde güncellemeyi hedefler (Dey ve Salem, 2017). Bu, daha hızlı eğitim süreleri ve daha düşük hesaplama maliyetleri sağlayabilir.

BiGRU modeli, her iki yönde çalışan GRU hücrelerini içerir. Bu, metnin hem ileri hem de geri yönden gelen bağlamını dikkate alarak daha kapsamlı bir anlam çıkarmaya olanak tanır (Li ve ark., 2022). Özellikle metin verilerinde, bu çift yönlü yaklaşım, dil örüntülerini daha etkili bir şekilde kavrama konusunda yardımcı olabilir. Bu model, metin içindeki bağlamsal ilişkileri daha iyi anlayarak performansı artırabilir.

CNN modeli ise metin verilerinde özellik çıkarmak için evrişim katmanları içeren bir yapıdır (Gu ve ark., 2018). Metin içindeki önemli özellikleri vurgular ve dil örüntülerini öğrenir. CNN, özellikle belirli uzunluktaki metinlerde lokal örüntüleri daha etkili bir şekilde tespit edebilir ve bu nedenle belirli metin tabanlı görevlerde başarılı olabilir.

Bu çeşitli modellerin kullanılması, BiLSTM modelinin performansını değerlendirmek ve metin verilerindeki örüntüleri daha iyi anlamak için farklı yaklaşımları keşfetmek açısından önemlidir. Her bir modelin avantajları ve dezavantajları, belirli görevlere uygunluğunu belirlemede yardımcı olabilir.

### 2.2.5. Model Eğitimi

Her bir derin öğrenme mimarisi için model eğitimi, Core Dou i5 işlemci, 32 GB RAM, 8 GB dahili ekran kartı, 1 TB SSD harddisk donanımı ile yaklaşık 16 ila 28 saatlik eğitim süreçlerinden geçirilerek ve farklı iterasyon sayıları (epoch) test edilerek oluşturulmuştur. Her bir epoch, modelin tüm eğitim verilerini bir kez gördüğü bir iterasyonu temsil eder. Batch (grup) büyüklüğü, her iterasyonda kullanılan örnek sayısını belirler (Kandel ve Castelli, 2020). Eğitim sırasında, model giriş verilerini kullanarak tahminlerde bulunur, ardından bu tahminlerle gerçek etiketler arasındaki kaybı hesaplar ve ağırlıkları günceller. Bu süreç, modelin eğitim verilerini öğrenmesini sağlar. Modelin performansını izlemek ve kontrol etmek için geri besleme kullanılır. Model kontrol noktaları, eğitim sırasında belirli bir metrik değerine ulaşıldığında modelin ağırlıklarını kaydeder. Bu yöntem, en iyi performansın elde edildiği durumu korumak için kullanılır. Eğitim belli bir iterasyondan sonra durdurulması, doğrulama kaybındaki belirgin bir iyileşme olmadığı takdirde eğitimi durdurarak aşırı öğrenmeyi önler. Eğitim sürecinin bu adımları, modelin doğru öğrenme, genelleme yeteneği ve istikrarlı performans elde etmesine katkı sağlar. Eğitim süreci, modelin öğrenme başarısını tespit etmek amacıyla 30, 50, 100 ve 200 iterasyon ile test edilmiştir ve model eğitimi 100 iterasyonda en fazla başarıya ulaştığı gözlemlenmiştir.

### 2.2.6. Metin Üretimi

Eğitilen model, belirli bir başlangıç kelime dizisi üzerinden yeni metinlerin üretilmesi için kullanılmıştır. Metin üretimi sürecinde, çeşitlilik ve özgünlük elde etmek amacıyla Top-K örnekleme teknikleri uygulanmıştır. Metin üretimi, belirlenen bir başlangıç kelime dizisiyle başlar. Bu dizideki kelimeler, modelin metin üretimini yönlendiren temel unsurları oluşturur. Kullanıcı, isteğe bağlı olarak farklı başlangıç dizileri seçerek modelin üretimini etkileyebilir.

Top-K örnekleme, modele en yüksek olasılığa sahip K kelime seçeneğini belirleme yöntemidir. Bu teknik, metin üretimini belirli bir çeşitlilik seviyesinde tutmak için kullanılır. Modelin en yüksek olasılığa sahip K kelime arasından rastgele bir seçim yapması sağlanır (Fagin ve ark., 2003).

Bu tekniklerin bir araya gelmesi, modelin belirli bir kelime dizisiyle çeşitli, özgün ve anlamlı metinler üretmesini sağlar. Metin üretim süreci, modelin öğrendiği dil yapılarını kullanarak başlangıç dizisi üzerinden tutarlı ve yaratıcı bir dil üretme yeteneğini yansıtır.

### 2.2.7. Model Doğruluğu

Günümüzde metin özetleme, altyazı oluşturma, konuşma tanıma, makine çevirisi vb. alanlarda önemli bir yer tutan METEOR skoru, metin oluşturma süreçlerinde üretilen metinlerin referans metinlerle olan benzerliğini değerlendirmek amacıyla kullanılan bir doğruluk metriğidir. Bu metrik, bir model veya sistem tarafından üretilen metinlerin, bir referans metinle karşılaştırılarak objektif bir benzerlik değerlendirmesi yapılmasına olanak tanır. METEOR skoru, metinler arasındaki kelime düzeni doğruluğunu dikkate alarak çalışmasıyla öne çıkar (Banerjee ve Lavie, 2005). Bu metrik, Precision (Doğruluk) ve Recall (Hatırlama) kavramlarını içeren bir formülle hesaplanır. Precision, sistem tarafından doğru bir şekilde üretilen kelimelerin oranını temsil ederken, Recall referans metinde doğru bulunan kelimelerin oranını ifade eder ve sırasıyla Denklem (1) ve Denklem (2)'de gösterildiği gibi hesaplanır.

$$Precision = \frac{m}{w_t} \quad (1)$$

$$Recall = \frac{m}{w_r} \quad (2)$$

Burada  $m$  sayısı referans metin ve özet metindeki ortak unigram sayısı olarak ifade edilir.  $w_t$  özet metindeki unigram sayısı,  $w_r$  ise referans metindeki unigram sayısını ifade eder. Meteor skorunu

hesaplamak için öncelikle ortalama F değerinin hesaplanması gerekir. *Fort* değerini bulmak için doğruluk ve hatırlama formülleri harmonik ortalama kullanılarak Denklem (3)'deki gibi hesaplanır.

$$Fort = \frac{10PR}{R+9P} \quad (3)$$

Doğruluk ve hatırlama metrikleri harmonik ortalama kullanılarak birleştirilir; Hatırlama metriği, doğruluk metriğinden 9 kat daha fazla olacak şekilde ağırlıklandırılır.

Meteor sonucu hesaplanırken bir diğer parametre parça cezası olarak adlandırılır. Bu metrik bir metin dizisinde özet ve referans kelimelerinin ardışık olarak eşleşmelerine dayanan bir değerlendirme kriteridir. Bu ceza, adaydaki kelimelerin hedef veya referanstaki kelimelerle ne kadar uyumlu olduğunu ölçer. İdeal durumda, tüm kelimelerin tam bir eşleşmesi durumunda sadece bir parça kalır. Ancak, genellikle bu ideal durum gerçekleşmez ve daha fazla parça ortaya çıkar. Sonuç olarak parça cezası, metinler arasındaki ardışık kelime eşleşmelerini değerlendirmek ve bu eşleşmelerin doğruluğunu ölçmek amacıyla kullanılan bir değerlendirme ölçüsüdür. Formülü Denklem (4)'teki gibi hesaplanır.

$$p = 0.5 \left( \frac{c}{u_m} \right)^3 \quad (4)$$

Burada  $p$  sayısı parça cezası,  $c$  sayısı ardışık olarak eşleşen parça sayısı,  $u_m$  ise maksimum parça sayısı (toplam unigram sayısı) olarak ifade edilir. Bu formül, ardışık olarak eşleşen parçaların sayısının maksimum eşleşen parça sayısına oranının küpünün yarısını kullanarak parça cezasını hesaplar. Bu durum, eşleşen parçaların sayısı arttıkça cezanın hızla artmasını ve metinler arasındaki uyumsuzluğun daha belirgin hale gelmesini sağlar. Nihai meteor puanı, Precision (Doğruluk) ve Recall (Hatırlama) verileriyle hesaplanan F puanını parça cezasıyla birleştirir ve 0 ile 1 arasında bir değer oluşturur. Sonuç 1 değerine ne kadar yakınsa o kadar başarılı sayılır. Meteor skoru Denklem (5)'teki formül ile hesaplanır.

$$M = Fort * (1 - p) \quad (5)$$

METEOR skoru, metin oluşturma modellerinin performansını objektif bir biçimde değerlendirmek, farklı modeller arasında karşılaştırma yapmak ve metin üretim kalitesini ölçmek için etkili bir araçtır.

ROUGE (Recall-Oriented Understudy for Gisting Evaluation), metin özetleme sistemlerinin performansını değerlendirmek için yaygın olarak kullanılan başka bir ölçüm aracıdır. ROUGE, özet



metin ile referans metin arasındaki benzerlikleri analiz eder ve farklı türdeki eşleşmeleri ölçen çeşitli metrikler sunar (Lin C.Y., 2004; Grusky M., 2023). Bu metrikler, özetin ne kadar bilgilendirici ve doğru olduğunu belirlemede önemli rol oynar.

ROUGE-N metriği, özet ve referans metinlerde bulunan aynı N kelimelik dizilerin sayısını ölçer. Burada N, 1'den başlayarak herhangi bir pozitif tam sayı olabilir ve genellikle unigram (ROUGE-1), bigram (ROUGE-2) gibi ifadeler kullanılır. ROUGE-N hesaplaması, özet ve referans metinde ortak olan N-gramların sayısının, referans metindeki toplam N-gram sayısına oranlanması ile yapılır. ROUGE-N skoru Denklem (6)'daki formül ile hesaplanır.

$$ROUGE - N = \frac{(\sum_{i=1}^n Count(Unigram(S), Unigram(R)))}{n} \quad (6)$$

Burada: Count(Unigram(S), Unigram(R)): özetlenen metin (S) ve referans metin (R) içinde bulunan aynı N kelimelik dizilerin sayısı.  $n$ : referans metinde bulunan N kelimelik dizilerin sayısı.

ROUGE-L metriği ise, özetlenen metin ile referans metin arasındaki en uzun ortak dizinin (LCS - Longest Common Subsequence) uzunluğunu ölçer. LCS, özetin ve referansın içindeki kelimelerin sırasını göz önünde bulundurarak, iki metin arasındaki en uzun alt diziyi temsil eder. ROUGE-L, iki metin arasındaki benzerliği değerlendirirken sıralı bilgiyi dikkate alır ve Denklem (7)'deki formül ile hesaplanır.

$$ROUGE - L = \frac{2 * LCS(S,R)}{|S| + |R|} \quad (7)$$

Burada: LCS(S, R): Özetlenen metin (S) ve referans metin (R) arasındaki en uzun ortak dizinin uzunluğu.  $|S|$ : Özetlenen metindeki kelime sayısı.  $|R|$ : Referans metindeki kelime sayısı.

ROUGE metrikleri, genellikle F-Measure (F1 skoru) bazlı olarak hesaplanır. F1 skoru, hatırlama (Recall) ve doğruluk (Precision) arasında denge kurar. F1 skoru, özet metnin referans metinle ne kadar örtüştüğünü değerlendirmek için kullanılır. ROUGE skorları 0 ile 1 arasında değer alır; burada 1 tam uyumu, 0 ise hiç uyum olmadığını göstermektedir. Yüksek ROUGE skorları, özet metnin referans metinle daha fazla bilgi içerdiğini ve daha iyi bir performansa sahip olduğunu gösterir.

BLEU (Bilingual Evaluation Understudy): BLEU, özetlenen metinleri bir çeviri olarak değerlendirir. Metrik, referans metin ile özetlenen metin arasındaki kelime eşleşmelerini

değerlendirerek hesaplanır. BLEU, referans metinde bulunan kelimelerin özetlenen metinde kaç kez geçtiğini dikkate alır ve bu bilgiyi bir BLEU skoru oluşturmak için kullanır. BLEU skoru 0 ile 1 arasında değer alır ve yüksek skorlar, özetin referans metne daha yakın olduğunu gösterir (Papineni K. ve ark., 2002).

BLEU, öncelikle, özetlenen metindeki kelimelerin, referans metindeki hangi kelimelerle eşleştiği belirlenir. Daha sonra, bu eşleşmelere dayalı olarak bir kelime eşleşme oranı hesaplanır.

BLEU skoru, Denklem (8)'deki formül ile hesaplanır:

$$BLUE = e^{(\sum_{i=1}^n w_i * \log(pi))} \quad (8)$$

Burada:  $w_i$ :  $i$ -nci  $n$ -gram için bir ağırlık faktörü.  $p_i$ :  $i$ -nci  $n$ -gram için kelime eşleşme oranı. BLEU formülünde, farklı  $n$ -gramlar için farklı ağırlık faktörleri kullanılır. Örneğin, 1-gram (tekli kelimeler) için daha yüksek bir ağırlık faktörü kullanılırken, 4-gram (dört kelimelik diziler) için daha düşük bir ağırlık faktörü kullanılır.

Bu metriklerin kullanılması, özetlenen metinlerin kalitesini objektif bir şekilde değerlendirmek için önemlidir.

### 3. Bulgular ve Tartışma

Bu çalışma, dil modeli eğitimi ve metin üretimi süreçlerini inceleyen bir dizi derin öğrenme modelini içermektedir. Eğitim aşamasında kullanılan Derin öğrenme modelleri, TensorFlow kütüphanesi üzerinde kurulmuş ve Türkçe metin verisi üzerinde eğitilmiştir.

Farklı tür ve lehçelerde yazılan ilgisiz metinlerin bir arada kullanılarak anlamsız metinler üretmesinin önüne geçmek için eğitim verisi olarak "kitap.txt" ve "kitap2.txt" dosyalarındaki kitap metinleri birbirinden bağımsız olarak verisetine aktarılmıştır. Metinler, öncelikle küçük harfe dönüştürülmüş, ardından tırnak işaretleri temizlenmiş ve Türkçe dilinde kelime tokenize işlemine tabi tutulmuştur.

Tüm modeller, belirlenen maksimum kelime sayısı olan 50 kelime uzunluğunda örnek dizileri kullanılarak oluşturulmuştur. BiLSTM mimarisi, iki katmanlı bir BiLSTM yapısını içermekte ve her bir kelimenin gömme (embedding) vektörünü oluşturarak öğrenmeyi gerçekleştirmektedir. Benzer şekilde LSTM, GRU ve BiGRU mimarileri de kendi iki katmanlı yapılarını içermektedir. Bu dört modelde de Overfitting (Aşırı Öğrenme) ile karşılaşmamak için her katmanda %40'lık Dropout kullanılmıştır. CNN mimarisi 3'lük bir çekirdek boyutu ve 'same' dolgusu ile 64 filtre sayısı uygulanarak eğitilmiştir. Eğitim sırasında kullanılan aktivasyon fonksiyonu tüm mimariler için "softmax" olarak belirlenmiş ve optimizasyon için Adam optimizer kullanılmıştır.

Ayrıca, eğitim sürecinde en düşük kayba ve en yüksek doğrulama oranına sahip olan modelin ağırlıkları "model.h5" dosyasına kaydedilmiştir. Eğitim sırasında belirli bir süre boyunca doğrulama kaybında iyileşme olmazsa eğitimi durduran bir erken durdurma (Early Stopping) mekanizması uygulanmıştır. Metin üretimi sırasında Top-K örnekleme yöntemiyle bir sonraki kelime belirlenmektedir. Oluşturulan özet metnin referans metinle ne kadar uyumlu olduğu Meteor, Blue ve Rouge skorlarıyla hesaplanmıştır. Farklı kitap türüne, modellere ve kelime sayılarına göre oluşturulan metin özetlerinin aldıkları Meteor, Bleu, Rouge değerleri Tablo 1’de gösterilmiştir.

**Tablo 1.** Kitap Türü, Model ve Kelime sayısına göre Meteor, Bleu, Rouge-1, Rouge-2 ve Rouge-L Metrik Sonuçları

Kitap Türü	Model	Kelime	Meteor	BLEU	ROGUE-1	ROUGE-2	ROUGE-L
ROMAN	CNN	1.000	0.5497	0.5231	0.6693	0.5402	0.5024
		2.000	0.6798	0.6623	0.7533	0.6632	0.5323
		5.000	0.7558	0.6812	0.8049	0.6823	0.5671
	LSTM	1.000	0.6313	0.5944	0.7123	0.6363	0.5242
		2.000	0.7054	0.6108	0.7843	0.6574	0.5467
		5.000	0.7565	0.6954	0.8299	0.7028	0.5664
	BiLSTM	1.000	0.6609	0.6299	0.7112	0.6424	0.5454
		2.000	0.7158	0.6533	0.7933	0.6509	0.5611
		<b>5.000</b>	<b>0.7702</b>	<b>0.7133</b>	<b>0.8723</b>	<b>0.7201</b>	<b>0.5834</b>
	GRU	1.000	0.6408	0.5902	0.6921	0.6244	0.5145
		2.000	0.7060	0.6290	0.7309	0.6723	0.5402
		5.000	0.7565	0.6892	0.8232	0.6823	0.5498
	BiGRU	1.000	0.6412	0.6030	0.7011	0.6396	0.5033
		2.000	0.7075	0.6186	0.7832	0.6535	0.5335
		5.000	0.7537	0.6877	0.8551	0.7045	0.5721
KİŞİSEL GELİŞİM	CNN	1.000	0.5361	0.5299	0.6923	0.5355	0.4834
		2.000	0.6312	0.6270	0.7751	0.6413	0.5291
		5.000	0.7206	0.6652	0.8140	0.6932	0.5528
	LSTM	1.000	0.6134	0.6043	0.7331	0.6216	0.5123
		2.000	0.6839	0.6333	0.8023	0.6732	0.5392
		5.000	0.7248	0.6824	0.8453	0.7035	0.5622
	BiLSTM	1.000	0.6520	0.6301	0.7412	0.6521	0.5406
		2.000	0.7143	0.6784	0.8237	0.6743	0.5633
		<b>5.000</b>	<b>0.7666</b>	<b>0.7238</b>	<b>0.8619</b>	<b>0.7158</b>	<b>0.5891</b>
	GRU	1.000	0.6403	0.5902	0.7124	0.6223	0.5014
		2.000	0.6823	0.6340	0.7823	0.6674	0.5349
		5.000	0.7320	0.6855	0.8320	0.6810	0.5512
	BiGRU	1.000	0.6397	0.6123	0.7349	0.6403	0.5232
		2.000	0.7082	0.6534	0.8017	0.6734	0.5453
		5.000	0.7469	0.7012	0.8523	0.6923	0.5671

Tablo 1, iki farklı kitap türü için farklı dil modeli türlerinin (CNN, LSTM, BiLSTM, GRU, BiGRU) farklı kelime sayılarında Meteor, Bleu ve Rouge skorlarını içermektedir. Tablo 1’e göre, BiLSTM mimarisi üretilen 1.000, 2.000 ve 5.000 kelime sayılarından oluşan metin özetleme için en yüksek Meteor, Bleu ve Rouge skorlarına sahiptir.

Test sonuçları incelendiğinde, kelime sayısının arttıkça tüm hata metrik skorlarının yükseldiği gözlemlenmiştir. Bu durum, daha büyük veri setlerinin model eğitimini iyileştirebileceğini ve genel performansı artırabileceğini göstermektedir.

LSTM, BiLSTM, GRU ve BiGRU mimarileri genellikle CNN'e göre daha yüksek Meteor, Bleu ve Rogue skorlarına sahiptir. Bu, özellikle dil modelleri ve metin anlama görevleri için uzunluğu daha iyi ele alabilen yineleme (RNN) tabanlı modellerin daha etkili olduğunu göstermektedir. Optimal model seçimi, spesifik bir görev veya veri kümesine bağlı olabilir. Ancak, genel olarak, BiLSTM modeli çeşitli kelime sayıları için daha iyi bir performans göstermektedir.

Roman ve kişisel gelişim kitabı özetlemeleri için kullanılan BiLSTM modeli ile elde edilen 5.000 kelimelik özetlerden oluşan yaklaşık 30 kelimelik örnek metinler Tablo 2'de gösterilmektedir.

**Tablo 2.** Özetlenen Metinden Örnek Metinler

Örneklem	Roman	Kişisel Gelişim
Örnek-1	hayat için çalışacaktı birkaç saatini not defterindeki düşünceleri okuyarak ve okuduklarını zihninde canlandırarak geçirdikten sonra denizle buluşmak için yurttan çıktı yapay zekâ ile o kadar çok büyülenmişti ki zeynep aklının ucundan geçmiyordu	başarıya ulaşmanın sırrı hedef belirlemekten ziyade sürdürülebilir ve etkili bir sistem kurmaktır hedefler bize yön gösterirken sistemler bizi o hedefe ulaştırır bu sistem küçük ve tekrarlanabilir adımlardan oluşmalı
Örnek-2	ancak gece geç saatlere doğru uyuşmuş parmaklarımla son cümleyi yazarken bir tuhaflık hissetti kendi dünyasının dışına çıkmış gibiydi onun hikayesini yazarken kendi hikayesini de yazıyordu sanki tıpkı tüm yazarlar gibi	her ne kadar iyi şeyler hayatımıza renk katsa da unutmamalı ki her iyi şeyin beraberinde getirdiği yüzlerce dezavantaj mevcut bu dezavantajları farkında olmak gerçekçi beklentiler oluşturmaya ve daha bilinçli seçimler yapmamıza yardımcı olur

Bu uygulama, dil modeli eğitimi ve metin üretmede farklı derin öğrenme mimarilerinin bir örneğini sunmaktadır. Elde edilen sonuçlar, dilin karmaşıklığını anlama ve yaratıcı bir şekilde metin üretme potansiyelini değerlendirmek için farklı derin öğrenme mimarilerinin metin özetleme konusundaki karşılaştırmalı bir yaklaşımını göstermektedir.

#### 4. Sonuçlar ve Öneriler

Bu çalışma, dil modelinin eğitiminden metin üretimine uzanan süreci detaylı bir şekilde ele almış ve derin öğrenme tekniklerini kullanarak dilin karmaşıklıklarını anlama ve yaratıcı bir şekilde yeni metinler üretme potansiyelini incelemiştir. Çalışma, Bidirectional Long Short-Term Memory (BiLSTM) mimarisi dahil olmak üzere çeşitli derin öğrenme yöntemlerini değerlendirmiştir.

Araştırma, iki farklı kitap metnini farklı dil modeli mimarilerinin performansını ölçmek için 1000, 2000 ve 5000 kelime sayısı ve Meteor, Bleu, Rogue skorlarını kullanarak kapsamlı bir analiz sunmaktadır. sonuçlara göre, farklı model türleri ve üretilen farklı kelime sayıları arasında belirgin farklar bulunmaktadır.

Roman türü için, BiLSTM modelinin genel olarak daha yüksek bir performans sergilediği görülmüştür. 5.000 kelimelik metinler üzerinde BiLSTM, en yüksek METEOR skoru (0.7702), BLEU skoru (0.7133), ROUGE-1 (0.8723), ROUGE-2 (0.7201) ve ROUGE-L (0.5834) değerlerine ulaşmıştır. Bu değerlere göre BiLSTM modelinin uzun metin üretiminde daha tutarlı ve kaliteli içerik üretebildiğini göstermektedir. LSTM ve GRU modelleri de benzer şekilde yüksek performans sergilemiş, özellikle 5.000 kelimelik metinlerde belirgin iyileşme göstermişlerdir. CNN modeli ise diğer modellere kıyasla daha düşük performans sergilemiş, ancak kelime sayısı arttıkça performansında bir artış gözlenmiştir.

Kişisel gelişim kitapları türü için de benzer bir eğilim gözlenmiştir. BiLSTM modeli yine en yüksek performansı sergileyerek, 5.000 kelimelik metinlerde en yüksek METEOR skoru (0.7666), BLEU skoru (0.7238), ROUGE-1 (0.8619), ROUGE-2 (0.7158) ve ROUGE-L (0.5891) değerlerine ulaşmıştır. LSTM ve GRU modelleri, kişisel gelişim metinlerinde de yüksek performans sergileyerek, uzun metinlerde daha tutarlı ve anlamlı içerikler üretebilmişlerdir. CNN modeli bu türde de diğer modellere göre daha düşük performans sergilemiş, ancak yine de kelime sayısı arttıkça performansında iyileşme gözlenmiştir.

Genel olarak, sonuçlar BiLSTM modelinin hem roman hem de kişisel gelişim kitaplarında en yüksek performansı sergilediğini göstermektedir. Bu modelin üstün performansı, çift yönlü bilgi akışı sayesinde metinlerin bağlamını daha iyi anlamasından kaynaklanmaktadır. BiLSTM, cümle içerisindeki kelimelerin hem önceki hem de sonraki bağlamını dikkate alarak daha tutarlı ve anlamlı metinler üretebilmektedir.

Bu bulgular, metin üretiminde model seçiminin kritik olduğunu ve modelin metnin bağlamını ne kadar iyi yakalayabildiğinin, metin kalitesinde büyük bir rol oynadığını göstermektedir. Dil modeli eğitimi ve metin üretimi sürecinde farklı derin öğrenme tekniklerinin etkilerini anlamak isteyen araştırmacılar ve yazılım geliştiriciler için önemli bir kaynak sunmaktadır. BiLSTM'nin yanı sıra LSTM, GRU, BiGRU ve CNN gibi diğer model türlerinin de dikkate alınması, dilin karmaşıklığını anlama ve yaratıcı bir şekilde metin üretme yeteneklerini değerlendirmek için kapsamlı bir yaklaşım sunmaktadır.

Bu çalışmanın sonuçlarına dayanarak, dil modeli eğitimi ve metin üretimi süreçlerinde daha geniş bir perspektif elde etmek için ilerleyen dönemlerde hibrit modeller ortaya koyulabilir. Örneğin, CNN ve BiLSTM'nin birleşimiyle oluşturulan bir model hem lokal hem de uzun vadeli bağımlılıkları ele alabilir. Gelecekteki çalışmalar için, farklı kitap türlerinde ve daha geniş veri setlerinde bu

modellerin performansının test edilmesi önerilmektedir. Ayrıca, modellerin eğitimi sırasında kullanılan hiperparametrelerin optimizasyonu, metin kalitesini daha da artırabilir. Yeni model mimarilerin ve hibrit yaklaşımların, metin üretiminde daha yenilikçi ve etkili çözümleri sunabileceği öngörülmektedir.

Bu öneriler, dil modeli eğitimi ve metin üretimi konularında daha geniş bir araştırma yelpazesi sunarak, derin öğrenme modellerinin dilin zenginliğini anlama ve yaratıcı bir şekilde ifade etme yeteneklerini artırma potansiyelini daha da genişletebilir.

### **Yazarların Katkısı**

Bu çalışmanın veriseti, araştırma, geliştirme ve uygulama süreçleri Yazar1 tarafından, araştırma ve düzenleme süreçleri Yazar2 tarafından yapılmıştır.

### **Çıkar Çatışması Beyanı**

Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

### **Araştırma ve Yayın Etiği Beyanı**

Yapılan çalışmada araştırma ve yayın etiğine uyulmuştur.

### **Kaynaklar**

- Babüroğlu, B., Tekerek, A., & Tekerek, M. (2019). Türkçe İçin Derin Öğrenme Tabanlı Doğal Dil İşleme Modeli Geliştirilmesi. *13th International Computer and Instructional Technology Symposium*
- Banerjee, S. and Lavie, A. (2005). METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. *In Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 65–72, Ann Arbor, Michigan. Association for Computational Linguistics.
- Bayer, M., Kaufhold, M.-A., Buchhold, B., Keller, M., Dallmeyer, J., Reuter, C. (2022). Data augmentation in natural language processing: a novel text generation approach for long and short text classifiers. *International Journal of Machine Learning and Cybernetics*, 14(3), 135-150. <https://doi.org/10.1007/s13042-022-01553-3>
- Chakraborty, S., Banik, J., Addhya, S., & Chatterjee, D. (2020). Study of Dependency on number of LSTM units for Character based Text Generation models. *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*.
- Cui, P., Wang, X., Pei, J., & Zhu, W. (2018). A Survey on Network Embedding. *IEEE Transactions on Knowledge and Data Engineering*, 31(5), 833-852. <https://doi.org/10.1109/TKDE.2018.2849727>
- Dey, R., Salem F. M. (2017). Gate-variants of Gated Recurrent Unit (GRU) neural networks. *Midwest Symposium on Circuits and Systems*, 1597-1600. <https://doi.org/10.1109/MWSCAS.2017.8053243>
- Erhandi, B., Çallı, F. (2020). Derin Özetleme ile Metin Özetleme. *3rd International Conference on Data Science and Applications (ICONDATA'20)*.

- Fagin, R., Kumar, R., Sivakumar, D. (2003). Comparing Top k Lists. *SIAM Journal on Discrete Mathematics* 17(1), 134-160. <https://doi.org/10.1137/S0895480102412856>
- Fang, T., Jaggi, M., & Argyraki, K. (2017). Generating Steganographic Text with LSTMs. <https://doi.org/10.48550/arXiv.1705.10742>
- Fatima, N., Imran, A. S., Kastrati, Z., Daudpota, S. M., & Soomro, A. (2021). A Systematic Literature Review on Text Generation Using Deep Neural Network Models. *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*.
- Grusky, M. (2023, July). Rogue scores. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1914-1934).
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354-377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Guo, B., Wang, H., Ding, Y., Wu, W., Hao, S., Sun, Y., & Yu, Z. (2020). Conditional Text Generation for Harmonious Human-Machine Interaction. *ACM Transactions on Intelligent Systems and Technology* 14, 1-50. <https://doi.org/10.1145/3439816>
- Huang, Z., Xu, W., Yu, K. (2015). Bidirectional LSTM-CRF Models for Sequence Tagging. <https://doi.org/10.48550/arXiv.1508.01991>
- Iqbal, T., Qureshi, S. (2020). The survey: Text generation models in deep learning. *Journal of King Saud University - Computer and Information Sciences* 34/6, 2515-2528. <https://doi.org/10.1016/j.jksuci.2020.04.001>
- Javid, A. M., Das, S., Skoglund, M., & Chatterjee, S. (2021). A ReLU Dense Layer to Improve the Performance of Neural Networks. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2810-2814.
- Kandel, I., Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4), 312-315. <https://doi.org/10.1016/J.ICTE.2020.04.010>
- Li, J., Zhao, W. X., Nie, J.-Y., Wen, J.-R., & Tang, T. (2022). Pre-trained Language Models for Text Generation: A Survey. <https://doi.org/10.48550/arXiv.2201.05273>
- Li, L., Zhang, T. (2021). Research on Text Generation Based on LSTM. *International Core Journal of Engineering*, 7, 2021. [https://doi.org/10.6919/ICJE.202105\\_7\(5\).0067](https://doi.org/10.6919/ICJE.202105_7(5).0067)
- Li, X., Ma, X., Xiao, F., Xiao, C., Wang, F., & Zhang, S. (2022). Time-series production forecasting method based on the integration of Bidirectional Gated Recurrent Unit (Bi-GRU) network and Sparrow Search Algorithm (SSA). *Journal of Petroleum Science and Engineering*, 208, 109309. <https://doi.org/10.1016/J.PETROL.2021.109309>
- Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. In Text summarization branches out (pp. 74-81).
- Mao, Y., Li, X., Li, Z., & Li, W. (2024). Automated Smart Contract Summarization via LLMs. arXiv preprint arXiv:2402.04863.
- Mishra, R., Bian, J., Fiszman, M., Weir, C. R., Jonnalagadda, S., Mostafa, J., & Fiol, G. Del. (2014). Text summarization in the biomedical domain: A systematic review of recent research. *Journal of Biomedical Informatics*. 52, 457-467 <https://doi.org/10.1016/j.jbi.2014.06.009>
- Onan, A. (2022). Türkçe Metin Madenciliği için Çalışan Bellek Bağlantıları Tabanlı Uzun Kısa Süreli Bellek Mimarisi. *European Journal of Science and Technology Special Issue*, 34, 239-246. <https://doi.org/10.31590/ejosat.1080239>
- Onan, A. (2022). Türkçe Metin Madenciliği için Dikkat Mekanizması Tabanlı Derin Öğrenme Mimarilerinin Değerlendirilmesi. *European Journal of Science and Technology Special Issue*, 34, 403-407. <https://doi.org/10.31590/ejosat.1082379>
- Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*. Issue, 2, 604-624.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting of the Association for Computational Linguistics (pp. 311-318).
- Pawade, D., Sakhapara, A., Somaiya, K. J., Jain, M., Jain, N., & Gada, K. (2018). Story Scrambler - Automatic Text Generation Using Word Level RNN-LSTM. *Information Technology and Computer Science*, 6, 44-53. <https://doi.org/10.5815/ijitcs.2018.06.05>

- Samant, R. M., Bachute, M. R., Gite, S., & Kotecha, K. (2022). Framework for Deep Learning-Based Language Models Using Multi-Task Learning in Natural Language Understanding: A Systematic Literature Review and Future Directions.
- Santhanam, S. (2020). Context Based Text - Generation Using LSTM Networks. <https://doi.org/10.48550/arXiv.2005.00048>
- Semeniuta, S., Severyn, A., & Barth, E. (2017). A Hybrid Convolutional Variational Autoencoder for Text Generation. <https://doi.org/10.48550/arXiv.1702.02390>
- Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. <https://doi.org/10.1016/J.PHYSD.2019.132306>
- Shi, Z., Chen, X., Qiu, X., & Huang, X. (2018). Toward Diverse Text Generation with Inverse Reinforcement Learning. <https://doi.org/10.48550/arXiv.1804.11258>
- Sutskever, I., Martens, J., & Hinton, G. (2011). Generating Text with Recurrent Neural Networks.
- Ünlü, Ö., Cetin, A. (2019). A Survey on Keyword and Key Phrase Extraction with Deep Learning. 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 - Proceedings.
- Vo, S. N., Vo, T. T., & Le, B. (2024). Interpretable extractive text summarization with meta-learning and BI-LSTM: A study of meta learning and explainability techniques. *Expert Systems with Applications*, 245, 123045.
- Wang, M., Lu, S., Zhu, D., Lin, J., & Wang, Z. (2018). A High-Speed and Low-Complexity Architecture for Softmax Function in Deep Learning. 2018 *IEEE Asia Pacific Conference on Circuits and Systems, APCCAS 2018*, 223-226. <https://doi.org/10.1109/APCCAS.2018.8605654>
- Welleck, S., Brantley, K., Daumé, H., & Cho, K. (2019). Non-Monotonic Sequential Text Generation. *Proceedings of the 36th International Conference on Machine Learning, PMLR 97:6716-6726*.
- Wu, N., Gong, M., Shou, L., Liang, S., & Jiang, D. (2023, October). Large language models are diverse role-players for summarization evaluation. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 695-707). Cham: Springer Nature Switzerland.
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Computational Intelligence Magazine*. 13, 55-75.
- Yu, W., Zhu, C., Li, Z., Hu, Z., Wang, Q., Ji, H., & Jiang, M. (2022). A Survey of Knowledge-Enhanced Text *ACM Computing Surveys*. 227, 1-38. <https://doi.org/10.1145/3512467>
- Zhang, T., Meng, J., Yang, Y., & Yu, S. (2024). Contrastive learning penalized cross-entropy with diversity contrastive search decoding for diagnostic report generation of reduced token repetition. *Applied Sciences*, 14(7), 2817.
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223.
- Zhou, W., Ye, Z., Yang, Y., Wang, S., Huang, H., Wang, R., & Yang, D. (2023). Transferring pre-trained large language-image model for medical image captioning. In *CLEF2023 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org, Thessaloniki, Greece*.
- Zhu, Y., Lu, S., Zheng, L., Guo, J., Zhang, W., Wang, J., Yu, Y. (2018). Taxygen: A Benchmarking Platform for Text Generation Models. *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* 1097–1100. <https://doi.org/10.1145/3209978.3210080>