



**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**ŞARKI SÖZLERİNDEN YENİ ŞARKI TÜRETME: DERİN
ÖĞRENME YÖNTEMLERİ İLE SANATSAL ÜRETİM**

Fatma GEZER

**Temmuz-2025
BATMAN**

T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI

YÜKSEK LİSANS TEZİ

ŞARKI SÖZLERİNDEN YENİ ŞARKI TÜRETME: DERİN
ÖĞRENME YÖNTEMLERİ İLE SANATSAL ÜRETİM

Fatma GEZER

Danışman
Doç. Dr. Yılmaz KAYA

Diğer Jüri Yeleri

Doç. Dr. Cüneyt ÖZDEMİR Doç. Dr. Abdulkerim ÖZTEKİN

Temmuz-2025
BATMAN

TEZ KABUL VE ONAYI

Fatma GEZER tarafından hazırlanan “Şarkı Sözlerinden Yeni Şarkı Türetme: Derin Öğrenme Yöntemleri ile Sanatsal Üretim” adlı tez çalışması 09/07/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Ana Bilim/Sanat Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç. Dr. Abdulkerim ÖZTEKİN

.....

Danışman

Doç. Dr. Yılmaz KAYA

.....

Üye

Doç. Dr. Cüneyt ÖZDEMİR

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYANI

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sorumluluğu kabullendiğimi bildiririm.

ETHICAL DECLARATION

I declare that all the information in this thesis has been obtained within the framework of ethical behavior and academic rules, and that the source of any statements and information that do not belong to me in this study prepared in accordance with the thesis writing rules has been fully cited, and I declare that I accept all kinds of legal responsibility in case of any contrary situation.

İmza
Fatma GEZER
09/07/2025

ÖZET

YÜKSEK LİSANS TEZİ

ŞARKI SÖZLERİNDEN YENİ ŞARKI TÜRETME: DERİN ÖĞRENME YÖNTEMLERİ İLE SANATSAL ÜRETİM

Fatma GEZER

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü

Bilgi Teknolojileri Ana Bilim Dalı

Danışman: Doç. Dr. Yılmaz KAYA

2025, 77 Sayfa

Şarkı sözü yazımı yaratıcı bir süreçtir. Geleneksel olarak bu süreç ilham, sezgi ve bireysel yaratıcılığa dayanır. Ancak yapay zekâ teknolojilerinin gelişmesiyle birlikte bu süreç yeni bir döneme girmiştir. Son yıllarda, özellikle yaratıcı metin üretimi alanında, Doğal Dil İşleme (NLP) kapsamında önemli ilerlemeler kaydedilmiştir. Özellikle GAN mimarileri, özgünlük ve rastlantısallık gerektiren içeriklerin üretiminde dikkat çekmektedir. Bununla birlikte, GAN'ların metin üretimine uyarlanması, görüntü üretimine kıyasla daha karmaşık bir süreçtir.

Türkçe gibi sondan eklemeli dillerde hem anlamlı hem de yapısal olarak tutarlı metinler üretmek daha da zordur. TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN ve RelGAN gibi mimariler, Türkçe şarkı sözü üretimi bağlamında sistematik olarak değerlendirilmemiştir. Bu durum, çalışmanın temel problemini oluşturmaktadır. Bu tez çalışmasının amacı, mevcut şarkı sözlerini analiz ederek, özgün ve sanatsal değeri yüksek içerikler üretebilen en uygun yapay zekâ modelini belirlemektir. Bu kapsamda, derin öğrenme modellerinin yanı sıra duygu analizi, ritmik yapı analizi ve yaratıcı metin üretimi teknikleri kullanılmıştır.

Elde edilen sonuçlara göre, düşük BLEU ve Jaccard skorları, modellerin referans metinlerle yüzeysel düzeyde benzerlik kurmakta zorlandığını göstermektedir. Buna karşılık, RelGAN modeli Cosine Similarity metriğinde en yüksek skoru (0.0693) elde ederek, anlamsal olarak en tutarlı çıktıları üretmiştir. MaliGAN modeli ise düşük perplexity değeri (9588.23) ile dil modelleme açısından öne çıkmıştır. Ancak, çeşitlilik metriklerinin (TTR ve Unique Word Ratio) 1.0 değerine ulaşması, anlamlı tekrarların eksikliğini ve doğal akışın zayıf olduğunu göstermektedir. Edebi analizlere göre, yalnızca RelGAN modeli tematik bütünlük ve duygu yoğunluğu açısından tatmin edici şarkı sözleri üretebilmiştir.

Sonuç olarak sayısal başarı ve edebi değerlendirme birlikte ele alındığında, RelGAN en dengeli ve başarılı model olarak öne çıkmıştır. Türkçe gibi eklemeli yapıya sahip dillerde şarkı üretimi için GAN modelleri hâlen sınırlı kalmakta olup, gelecekte Transformer tabanlı yapılarla desteklenmesi önerilmektedir.

Anahtar Kelimeler: GAN, Metin Üretimi, MaliGAN, RelGAN, SeqGAN, TextGAN, RankGAN

ABSTRACT

MASTER THESIS

GENERATING NEW SONGS FROM LYRICS: ARTISTIC PRODUCTION USING DEEP LEARNING METHODS

Fatma GEZER

Batman University Graduate Education Institute

Information Technologies Department of Science

Advisor: Assoc. Prof. Yılmaz KAYA

2025, 77 Pages

Writing song lyrics is a creative process. Traditionally, this process relies on inspiration, intuition, and individual creativity. However, with the advancement of artificial intelligence technologies, this process has entered a new phase. In recent years, significant progress has been made in the field of Natural Language Processing (NLP), particularly in creative text generation. GAN architectures, in particular, have attracted attention for generating content that requires originality and randomness. Nonetheless, applying GANs to text generation remains more complex than their application in image generation.

In agglutinative languages such as Turkish, generating text that is both meaningful and structurally coherent is even more challenging. Architectures such as TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN, and RelGAN have not been systematically evaluated in the context of Turkish song lyric generation. This constitutes the fundamental problem of the study. The aim of this thesis is to identify the most suitable artificial intelligence model capable of generating original and artistically valuable lyrics by analyzing existing song lyrics. In this scope, emotion analysis, rhythmic structure analysis, and creative text generation techniques were employed alongside deep learning models.

According to the results, the low BLEU and Jaccard scores indicate that the models struggled to establish surface-level similarity with reference texts. In contrast, the RelGAN model achieved the highest score in Cosine Similarity (0.0693), producing the most semantically coherent outputs. The MaliGAN model stood out in terms of language modeling with a low perplexity value (9588.23). However, the fact that diversity metrics (TTR and Unique Word Ratio) reached 1.0 suggests a lack of meaningful repetition and a weak natural flow. Based on literary analysis, only the RelGAN model was able to produce song lyrics that were satisfactory in terms of thematic consistency and emotional depth.

In conclusion, considering both numerical performance and literary evaluation, RelGAN emerged as the most balanced and successful model. GAN architectures remain limited in generating poetry in agglutinative languages like Turkish, and future studies are recommended to incorporate Transformer-based models for improved outcomes.

Keywords: GAN, Text Generation, MaliGAN, RelGAN, SeqGAN, TextGAN, RankGAN

ÖN SÖZ

Bu tezin hazırlanma sürecinde, içerik bütünlüğünün sağlanması ve tezde yer alan sayısız ayrıntının birbiriyle uyumlu biçimde şekillenmesinde; bilgi, birikim ve deneyimleriyle bana rehberlik eden; her aşamada yönlendirici katkılarda bulunan; samimiyeti ve desteğiyle çalışmalarımı güçlendiren kıymetli tez danışmanım Sayın Doç. Dr. Yılmaz KAYA'ya en derin saygı ve teşekkürlerimi sunarım.

Ayrıca, eğitimin gücüne yürekten inanan, bu inancı bana da küçük yaşlardan itibaren aşılayan ve her zaman yanımda olan sevgili babam Şafi KARABULUT'a; fedakârlığı, sabrı ve sınırsız desteğiyle yaşam yolumu aydınlatan canım annem Zehra KARABULUT'a; bu zorlu süreçte her türlü desteğiyle daima yanımda olan sevgili eşim Cevat GEZER'e ve hayatıma kattıkları neşe, umut ve ilhamla her anı anlamlı kılan kıymetli çocuklarım Çekdar, Bejan ve Jehat'a en içten teşekkürlerimi sunuyorum.

Bu çalışmanın, alana katkı sunmasını ve konuyla ilgilenecek araştırmacılar için yararlı bir kaynak olmasını temenni ederim.

Fatma GEZER
BATMAN-2025

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖN SÖZ	vi
İÇİNDEKİLER.....	vii
TABLolar LİSTESİ	ix
ŞEKİLLER LİSTESİ	x
KISALTMALAR	iv
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	5
3. MATERYAL VE YÖNTEM.....	29
3.1. Materyal.....	29
3.2. Metot.....	30
3.2.1. Generative Adversarial Networks (GANs).....	30
3.2.2. TextGAN Yöntemi: Metin Üretiminde Derin Temsillerle Adversaryal Öğrenme.....	33
3.2.3. LeakGAN yöntemi: Hiyerarşik yapıda bilgi sızıntısı ile geliştirilmiş adversarial metin üretimi	36
3.2.4. RankGAN yöntemi: Sıralama temelli adversarial metin üretimi.....	38
3.2.5. MaliGAN yöntemi: Maksimum benzerlik tabanlı doğrultulmuş adversarial metin üretimi.....	40
3.2.6. SeqGAN yöntemi: Ayrık sıralı veri üzerinde takviyeli öğrenme tabanlı adversarial metin üretimi	42
3.2.7. RelGAN yöntemi: İlişkisel özelliklerle zenginleştirilmiş adversarial metin üretimi.....	45
3.2.8. Performans metrikleri	47
3.2.8.1. BLEU (Bilingual Evaluation Understudy Score)	48
3.2.8.2. Jaccard similarity	48
3.2.8.3. Perplexity (PP).....	49
3.2.8.4. TTR (Type-Token Ratio).....	49
3.2.8.6. 2-gram coverage	50
3.2.8.7. Kosinüs benzerliği (TF-IDF tabanlı)	50
4. BULGULAR VE TARTIŞMA.....	51
4.1. Genel Bulgular.....	51
4.2. TextGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi	57
4.3. SegGan Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi.....	59
4.4. LeakGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi.....	61
4.5. RankGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi	63
4.6. MaliGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi	65

4.7. RelGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi	67
5. SONUÇLAR VE ÖNERİLER	70
5.1. Genel Değerlendirme	70
5.1.1. Çalışmanın Eksiklikleri	72
5.2 Öneriler	73
5.2.1. Gelecek Çalışmalar İçin Öneriler	73
KAYNAKLAR	75

TABLÖLAR LİSTESİ

Tablo 4.1. Farklı Metin retim GAN modellerin performans Tabloları.....	56
Tablo 4. 2. TextGAN Model Örnek 1	57
Tablo 4. 3. TextGAN Model Örnek 2.....	57
Tablo 4. 4. TextGAN Model Örnek 3 (8 satırlı).....	58
Tablo 4. 5. SegGAN Model Örnek 1	59
Tablo 4. 6. SegGAN Model Örnek 2	60
Tablo 4. 7. LeakGAN Model Örnek 1	61
Tablo 4. 8. LeakGAN Model Örnek 2	62
Tablo 4. 9. RankGAN Model Örnek 1.....	63
Tablo 4. 10. RankGAN Model Örnek 2.....	64
Tablo 4. 11. MaliGAN Model Örnek 1	65
Tablo 4. 12. MaliGAN Model Örnek 2	66
Tablo 4. 13. RelGAN Model Örnek 1	67
Tablo 4. 14. RelGAN Model Örnek 2	68

ŞEKİLLER LİSTESİ

Şekil 3. 1. Verilerin toplandığı ve düzenlendiği Excel dosyasına ait örnek bir ekran görüntüsü	30
Şekil 3. 2. GAN şeması.....	32
Şekil 4. 1. Loss(Kayıp) performans grafikleri	Hata! Yer işareti tanımlanmamış.

KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltma / Sembol	Açıklama
AI	Artificial Intelligence (Yapay Zekâ)
ANN	Artificial Neural Network (Yapay Sinir Ağı)
BLEU	Bilingual Evaluation Understudy
CNN	Convolutional Neural Network (Evrışimli Sinir Ağı)
DFU	Diabetic Foot Ulcer (Diyabetik Ayak Ülseri)
GAN	Generative Adversarial Network (Üretici Çekişmeli Ağ)
LSTM	Long Short-Term Memory
NLP	Natural Language Processing (Doğal Dil İşleme)
ReLU	Rectified Linear Unit
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
RNN	Recurrent Neural Network (Tekrarlayan Sinir Ağı)
SeqGAN	Sequence Generative Adversarial Network
TF-IDF	Term Frequency – Inverse Document Frequency

1. GİRİŞ

Müzik, insanlığın en eski ve en evrensel ifade biçimlerinden biri olarak, tarih boyunca bireylerin duygusal dünyasını şekillendirmiş, toplulukların kültürel kimliğini yansıtmış ve medeniyetlerin iletişim aracına dönüşmüştür. Ritmik yapılar, melodiler ve sözler aracılığıyla insanlar kendilerini ifade etmiş, topluluklarla bağ kurmuş ve ortak duygular etrafında birleşmiştir. Evrensel bir dil olarak kabul edilen müzik, farklı coğrafyalardan gelen bireylerin birbirini anlamasını kolaylaştırmış; sanatın yanı sıra toplumsal ve kültürel bir etkileşim alanı yaratmıştır (Levitin, 2006).

Tarihsel süreçte müzik, medeniyetlerin düşünsel ve dini sistemleriyle iç içe geçmiştir. Çin'de Konfüçyüs'ün (M.Ö. 551–479) öğretilerinde ahlaki eğitimin bir parçası olan müzik, Hindistan'da Vedik ritüellerin temel öğelerinden biri olarak kutsal kabul edilmiştir (Kaufmann, 1967). Zamanla müzik, yalnızca estetik bir uğraş değil, aynı zamanda toplumsal dönüşümlerin, teknolojik gelişmelerin ve insanın çevreyi algılayış biçiminin bir yansıması hâline gelmiştir. Bu dönüşüm, müzik üretiminin temel unsurlarından biri olan şarkı sözü yazımında da kendini göstermektedir. Şarkı sözleri, yalnızca bireysel duyguların ifadesi değil, aynı zamanda toplumsal olaylara karşı duyarlılığın, kültürel mirasın ve sanatsal yaratıcılığın da bir taşıyıcısıdır (Levitin, 2006). Günümüzde, müzik endüstrisinde şarkı sözlerinin önemi yalnızca sanatsal değil; ticari açıdan da belirleyici bir konuma sahiptir (Herremans ve ark., 2017).

Geleneksel olarak bireysel yaratıcılığa dayanan şarkı sözleri yazım süreci, günümüz yapay zekânın sunduğu büyük teknolojik olanaklarla yeniden şekillenmektedir. NLP ve GAN gibi derin öğrenme yöntemleri, ritmik ve anlamsal açıdan tutarlı şarkı sözleri üretme potansiyeli sunmakta; sanatçılar için bir ilham kaynağı ve üretim aracı olarak kullanma fırsatı sunmaktadır (Briot ve ark., 2020; Yu ve ark., 2017). Bu sistemler, ilham eksikliği yaşayan sanatçılara yaratıcı başlangıç noktaları sağlayarak üretkenliği artırmaktadır (Colton ve ark., 2015). Teknolojinin sanatla kesiştiği bu noktada, yapay zekâ yalnızca bir araç değil, aynı zamanda sanatın üretim süreçlerine doğrudan katkı sunan bir bileşen hâline gelmiştir. Beste yapımından şarkı sözü üretimine kadar birçok aşamada aktif rol oynayan yapay zekâ modelleri, armoni ve ritmik analiz gibi teknik görevlerin yanı sıra, duygusal ve tematik tutarlılığın sağlanmasında da kullanılmaktadır

(Herremans ve ark., 2017). GAN gibi yapay zekâ sistemleri, sanatçının üretim sürecini hızlandırmakla kalmayıp, yeni estetik anlayışların gelişmesine de olanak tanımaktadır (Boden, 2016). Bununla birlikte, yapay zekâ tarafından üretilen sanat eserlerinin özgünlüğü, duygusal derinliği ve etik boyutları, sanat kuramcıları arasında tartışılmaya devam etmektedir. İnsan sezgisinin yerini alıp alamayacağı, üretim sürecinde insan ile makine arasındaki sınırın nasıl tanımlanacağı gibi sorular hem felsefi hem de pratik açıdan önem taşımaktadır (Sturm ve ark., 2019). Ancak birçok sanatçı, yapay zekâyı bir tehditten ziyade yaratıcı süreci zenginleştiren bir ortak olarak değerlendirmektedir.

Yapay zekâ ile metin üretimi, özellikle sentetik metin üretimi alanında dikkate değer bir ilerleme göstermektedir. NLP teknikleri, büyük veri setlerinden öğrenerek anlamlı, bağlama uygun ve dilbilgisel olarak tutarlı içerikler oluşturabilmektedir (Jurafsky & Martin, 2021). Bu sayede Transformer, RNN, LSTM ve GAN gibi modeller, insan diline yakın ve estetik değer taşıyan metinlerin üretiminde etkin biçimde kullanılmaktadır (Vaswani ve ark., 2017; Goodfellow ve ark., 2014). NLP teknolojileri; edebiyat, diyalog sistemleri, haber yazımı ve yaratıcı içerik üretimi gibi birçok alanda kullanıma sahiptir (Ghazvininejad ve ark., 2019).

Özellikle GAN modelleri, sanat ve edebiyat alanlarında sundukları olanaklarla dikkat çekmektedir. İki bileşenli yapıları (üretici ve ayırt edici ağlar) sayesinde, bağlamsal tutarlılığı yüksek, özgün ve stilize metinler üretme kapasitesine sahiptir (Elgammal ve ark., 2017; Yu ve ark., 2017). SeqGAN gibi modeller, belirli bir temaya ya da duyguya uygun şarkı sözleri üretmede başarı sağlamaktadır.

GAN tabanlı şarkı sözü üretimi, üretkenlik, zaman tasarrufu, kişiselleştirilmiş içerik oluşturma ve sanatsal çeşitlilik açısından müzik endüstrisine önemli katkılar sunmaktadır. Bu sistemler, farklı müzik türlerinden öğrenerek tarzlara özgü sözler üretebilmekte; kişisel tercihlere, duygu durumlarına veya tematik isteklere uygun içerikler oluşturabilmektedir (Sturm ve ark., 2019). Bu durum hem bireysel sanatçılar hem de büyük ölçekli prodüksiyonlar için yeni fırsatlar sunmaktadır. Bütün bu fırsatlara rağmen Türkçe gibi sondan eklemeli dillerde, şarkı sözü üretimi ayrıca zorluk taşımaktadır. Hem anlamlı hem de biçimsel olarak uyumlu metin üretmek, özel bir problem haline gelmiştir. GAN'ların sunduğu tüm olanakların yanında; duygusal derinlik, bağlam bütünlüğü ve özgünlük gibi bazı alanlarda insan yaratımıyla kıyaslandığında sınırlılıkları da bulunduğu göz ardı edilmemelidir. Ayrıca yapay zekâ üretimi içeriklerin telif hakları, etik sınırları ve sanatsal sahipliği konuları da hâlen netlik kazanmamış olması bir problemdir (Riedl, 2014; Colton ve ark., 2015). Bu problemlerin

yanı sıra TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN ve RelGAN gibi mimarilerin Türkçe şarkı sözü üretiminde performans, yapı ve anlam açısından sistemli biçimde değerlendirilmemiş olması bu çalışmanın hareket noktası ve temel problemini oluşturmaktadır. Ayrıca GAN mimarilerinin Türkçe şarkı sözü üretimindeki etkinliği alan yazında yeterince araştırılmadığı görülmüştür. Bununla beraber bu modellerin üretim çeşitliliği, yapısal bütünlük ve anlamsal yakınlık gibi ölçütler bakımından karşılaştırılmasına dair herhangi bir çalışma olmaması bu çalışmanın gerekliliğini ve ihtiyacının ortaya koyduğu söylenebilir.

GAN tabanlı yapay zekâ sistemleri, müzik ve yaratıcı metin üretimi alanında yenilikçi olanaklar sunarken, insan yaratıcılığı ile bütünleşerek daha anlamlı ve estetik açıdan tatmin edici eserlerin ortaya çıkmasına katkıda bulunma potansiyeline sahiptir. Gelecekte, bu teknolojilerin etik ve sanatsal sınırlarının daha net çizilmesiyle birlikte, yapay zekâ destekli sanatsal üretim süreçlerinin daha da gelişeceği ve alan yazıma katkı sunacağı düşünülmektedir.

Bu tez çalışmasının temel amacı, şarkı sözlerini analiz ederek özgün ve sanatsal değeri yüksek yeni şarkılar türeten bir yapay zekâ modeli geliştirmektir. Bu temel amaç kapsamında çalışmanın alt amaçları şu şekilde sıralanabilir:

1. Derin öğrenme yöntemlerinin yaratıcı şarkı sözü üretiminde nasıl uygulanabilir?
2. Şarkı sözleri üretiminde geleneksel yöntemlerin yanı sıra, yapay zekâ tabanlı yaklaşımlarla yeni sanatsal ürünler oluşturulabilir mi?
3. Müzik endüstrisinde şarkı yazımı süreçlerini hızlandıracak ve çeşitlendirecek yenilikçi bir çözüm sunmak mümkün mü?
4. Şarkı sözleri, yalnızca dilsel bir yapı değil, aynı zamanda duygusal bir bağlama sahiptir. Bu tez, yapay zekâ ile duygusal açıdan tutarlı ve anlamlı metinler üretmeyi hedefleyerek dilsel ve duygusal üretim arasında bir köprü kuracaktır. Buna uygun en iyi üretim yapacak model hangisidir?
5. Müzik endüstrisi, içerik üretimi ve reklamcılık gibi sektörlerde, hızlı ve özgün şarkı sözü üretimi büyük bir ihtiyaçtır. Bu çalışma, bu ihtiyacı karşılayacak teknolojik bir çözüm önerisi sunarak profesyonel yaratıcı süreçlere katkıda bulunabilecek mi?
6. Elde edilen çıktılar, dilsel doğruluk, anlam bütünlüğü ve duygusal tutarlılık açısından değerlendirilecek olup en iyi sonuç hangi modele ait olduğu anlaşılabilir mi?

1.1. Çalışmanın Kapsamı

Bu tez çalışması, mevcut şarkı sözlerinden öğrenilen dilbilimsel ve duygusal örüntüler temelinde, yeni ve anlamlı şarkı sözleri üretebilen yapay zekâ modelini bulmayı amaçlamaktadır. Bu kapsamda, şarkı sözlerinin hem sözdizimsel hem de duygusal yapıları, büyük ölçekli veri setleri üzerinden incelenecek ve NLP ile derin öğrenme (DL) teknikleri kullanılarak dilin anlamsal özellikleri modellenecektir.

Çalışmada kullanılacak veri setleri, tematik açıdan çeşitlilik gösteren ve popüler müzik türlerine ait şarkı sözlerinden oluşacaktır. Bu veri setleri üzerinden, generatif modeller eğitilecek ve dilin yapısal özelliklerini içeren yeni şarkı sözleri üretilenektir. Elde edilen çıktılar, dilsel doğruluk, anlam bütünlüğü ve duygusal tutarlılık açısından değerlendirilecektir.

Bu yönüyle tez çalışması, hem yapay zekâ destekli yaratıcı üretim süreçlerine katkı sağlamayı hem de şarkı sözü yazımında dilin yapısal ve duygusal boyutlarını bütüncül bir yaklaşımla ele almayı hedeflemektedir.

2. KAYNAK ARAŞTIRMASI

Bu bölümde, derin öğrenme yöntemleri kullanılarak gerçekleştirilen metin üretimi ile ilgili çalışmalar özetlenmektedir.

Gill, Lee ve Marwell (2020) tarafından gerçekleştirilen çalışmada, müzikal türlere özgü şarkı sözleri üretebilen bir yapay zeka modelinin geliştirilmesi amaçlanmıştır. Geleneksel dil modellerinin düzyazı üretiminde başarılı olmasına rağmen, şarkı sözleri kendine özgü yapısal ve stilistik özellikler barındırmaktadır. Çalışmada, belirli müzik türlerine özgü şarkı sözleri üretebilmek için kelime seviyesinde bir Uzun Kısa Süreli Bellek (LSTM) ağı kullanılmıştır. Model, her giriş dizisinde 16 kelimelik bir pencere ile sonraki kelimeyi tahmin edecek şekilde tasarlanmış ve şarkı sözlerinin stilistik özelliklerini yakalayabilmesi için belirli müzik türlerinde eğitilmiştir. Modelin başarısını değerlendirmek amacıyla ortalama dize uzunluğu, şarkı kelime çeşitliliği, tür kelime çeşitliliği, bakış açısı (I vs. You) ve kelime tekrar oranı gibi beş temel metrik kullanılmıştır. Çalışmada kullanılan veri seti, Kaggle MetroLyrics ve Genius Lyrics API'den elde edilen yaklaşık 297.876 şarkıdan oluşmakta ve rap, rock, pop, metal, country ve jazz türlerini kapsamaktadır. Deneysel sonuçlar, modelin özellikle rap, pop ve rock türlerinde başarılı şarkı sözleri üretebildiğini, ancak kafiye oluşturma ve ritmik yapı oluşturma konusunda eksiklikleri bulunduğunu göstermektedir. Alternatif yöntemler incelendiğinde, GPT-2 ve BERT gibi transformer tabanlı modellerin müzikal bağlamda yeterli performans gösteremediği belirlenmiş, bunun yerine türlere özgü daha büyük dil modellerinin kullanılmasının faydalı olabileceği öne sürülmüştür. Gelecekte, kafiye ve ritim algılama yeteneklerine sahip modellerin geliştirilmesi, farklı türler için daha geniş veri setleriyle eğitilmesi ve alternatif yöntemlerin (GRU, Markov Modelleri) değerlendirilmesi önerilmektedir. Bu bağlamda, çalışma müzik alanında yapay zeka tabanlı şarkı sözü üretimi üzerine önemli bir katkı sunmakta ve türlere özgü dil modellerinin geliştirilmesine yönelik gelecekteki araştırmalara ışık tutmaktadır. (Gill ve ark., 2020)

Wang, Luo, Wang ve Xing (2016) tarafından yapılan çalışmada, geleneksel Çin şiiri türlerinden biri olan Song iambics (宋词) üretimi için yapay zeka tabanlı bir model geliştirilmiştir. Song iambics, katı ritmik ve yapısal kurallara sahip olup, geleneksel dil

modelleri bu kompleks yapıyı öğrenmekte zorlanmaktadır. Çalışmada, bu tür şiirlerin üretimi için dikkat (attention) mekanizmasına dayalı bir LSTM ağı önerilmiştir. Model, bir çift yönlü (bi-directional) LSTM ağını kullanarak giriş dizisini kodlamakta ve ardından dikkat mekanizmasıyla desteklenen bir LSTM modeli aracılığıyla şiirin geri kalan kısmını üretmektedir. Modelin öğrenme sürecini iyileştirmek için küresel bağlam entegrasyonu, hibrit stil eğitimi ve karakter vektörü başlangıç değerlendirme teknikleri kullanılmıştır. Modelin başarısını değerlendirmek amacıyla BLEU skoru, şiirsellik (poeticness), akıcılık (fluency) ve anlamlılık (meaningfulness) gibi metrikler kullanılmıştır. Veri seti olarak, toplamda 15.689 Song iambics şiiri içeren bir veri kümesi oluşturulmuş ve model 15.001 şiirle eğitilip 688 şiirle test edilmiştir. Deneysel sonuçlar, önerilen modelin geleneksel İstatistiksel Makine Çevirisi (SMT) ve RNN dil modellerinden daha iyi performans gösterdiğini ortaya koymuştur. Özellikle Attention-ALL olarak adlandırılan model, şiirsellik ve anlamlılık açısından yüksek puan alarak insan yazımı şiirlerle rekabet edebilecek bir seviyeye ulaşmıştır. Ancak, insan yazımı şiirlerin hâlâ üstün olduğu ve modelin duygusal ifadeler ile tematik tutarlılığı koruma konusunda geliştirilmesi gerektiği belirtilmiştir. Gelecekte, modelin daha geniş bir veri setiyle eğitilmesi, farklı şiir türleri üzerinde test edilmesi ve özellikle Tang şiiri, Han Fu ve Yuan Qu gibi diğer Çin edebi formlarına uygulanması önerilmektedir. Çalışma, yapay zekanın sanatsal içerik üretme kapasitesini araştırmak açısından önemli bir katkı sunmaktadır. (Wang ve ark., 2020)

Fan, Wang, Zhuang, Wang ve Xiao (t.y.) tarafından yapılan çalışmada, Çince şarkı sözleri üretimi için bağlama duyarlı bir model geliştirilmiştir. Geleneksel dil üretim modelleri kelime bazlı sıralı üretim yaparken, şarkı sözlerinde cümleler arasındaki bağlamın korunması ve anlatım bütünlüğünün sağlanması gerekmektedir. Bu nedenle, çalışmada hiyerarşik dikkat mekanizmasına (hierarchical attention) sahip bir Seq2Seq modeli önerilmektedir. Model, kelime düzeyinde ve cümle düzeyinde bağlam kodlamasını birleştirerek şarkı sözlerinin anlam bütünlüğünü korumakta ve konular arası tutarlılığı artırmaktadır. Modelin eğitimi için büyük ölçekli bir Çince şarkı sözü veri kümesi kullanılmıştır. Çalışmada modelin başarısını değerlendirmek için konu bütünlüğü (topic relevance), akıcılık (fluency) ve anlamsal tutarlılık (semantic coherence) gibi insan değerlendirme kriterlerinin yanı sıra BLEU skoru gibi otomatik değerlendirme metrikleri kullanılmıştır. Modelin eğitildiği veri seti, 100.000 Çince pop şarkı sözünden oluşmakta ve 5.045 kelimelik bir sözlük içermektedir. Deneysel sonuçlara göre, önerilen hiyerarşik Seq2Seq modeli, geleneksel Seq2Seq modeline kıyasla daha iyi performans göstermiştir.

Özellikle, dikkat mekanizması eklenen model, hem insan değerlendirmelerinde hem de BLEU skorlarında en iyi sonuçları elde etmiştir. Çalışma, hiyerarşik kodlama ve dikkat mekanizmalarının, uzun metinlerin bağlam bütünlüğünü koruyarak üretiminde önemli katkılar sunduğunu göstermektedir. Gelecekte, bu yöntemin farklı doğal dil üretimi (NLG) uygulamalarında, özellikle hikâye ve diyalog üretiminde kullanılabileceği öne sürülmektedir. (Fan ve ark., 2020).

Bir diğer çalışmada, Ma, Wang, Kan ve Lee (2021) tarafından geliştirilen AI-Lyricist, belirli bir kelime dağarcığı ve MIDI dosyası girdisi kullanarak anlamlı ve müzikal uyumlu şarkı sözleri üretmeyi amaçlayan bir sistemdir. Şarkı sözlerinin yaratıcı ve müzikle uyumlu olmasının yanı sıra, kelime dağarcığı kısıtlamalarına uygun olması, melodik hizalanmayı sağlaması ve belirli bir temaya bağlı kalması gerekmektedir. Bu çalışmada, şarkı sözü üretimi için dört modülden oluşan bir sistem önerilmektedir: (1) Müzik yapı analizörü, MIDI dosyasındaki melodi yapısını ve hece şablonlarını çıkarmaktadır. (2) SeqGAN tabanlı şarkı sözü üreticisi, metin kalitesini ve hece hizalamasını iyileştirmek için çift ayrıştırıcı (twin discriminator) çoklu düşman eğitime (multi-adversarial training) dayanmaktadır. (3) Derin eşlenik müzik-şarkı sözü gömme modeli (deep coupled music-lyrics embedding model), üretilen şarkı sözleri ile müzik arasında anlam ve stil bakımından uyumluluk sağlamak için her iki öğeyi ortak bir uzaya yansıtmaktadır. (4) Polisher modülü, kelime kısıtlamalarını yerine getirmek için maskelenmiş kelime seçimleri ve kelime değiştirme stratejilerini uygulamaktadır. Model, 7.000'den fazla müzik-şarkı sözü çiftinden oluşan bir veri seti üzerinde eğitilmiş ve tematik, duygusal ve türsel etiketlerle zenginleştirilmiştir. Deneysel sonuçlar, hem nesnel hem de öznel değerlendirmelerde AI-Lyricist'in mevcut yöntemlere kıyasla akıcılık, bütünlük, anlamlılık, şiirsel estetik, hece hizalaması ve müzikle uyum açısından üstün performans gösterdiğini ortaya koymuştur. Gelecekte, sistemin daha büyük veri setleriyle eğitilmesi, çok dilli şarkı sözü üretimi için genişletilmesi ve daha sofistike dil üretimi teknikleriyle geliştirilmesi önerilmektedir. (Ma ve ark., 2020)

Manjavacas, Karsdorp ve Kestemont (2019) tarafından yapılan çalışmada ise , hip-hop şarkı sözleri üretimi için koşullu şablonlar (conditional templates) ve hiyerarşik dil modelleri kullanılarak yeni bir yöntem geliştirilmiştir. Geleneksel dil modelleri, şarkı sözlerinin yapısal özelliklerini yakalamakta zorlanırken, bu çalışma özellikle kafiye ve ritim gibi temel özellikleri doğrudan modelleyebilen bir sistem sunmaktadır. Çalışmada, karakter, hece ve kelime seviyelerinde dil modelleme yaklaşımları karşılaştırılmış ve hiyerarşik bir modelin en başarılı sonucu verdiği belirlenmiştir. Önerilen model, koşullu

şablonlar yardımıyla şarkı sözlerinin sözdizimsel ve fonetik özelliklerini koruyarak gerçekçi ve türüne uygun metinler üretebilmektedir. Modelin başarısını değerlendirmek için büyük ölçekli bir kullanıcı değerlendirme testi gerçekleştirilmiş ve sonuçlar, geliştirilen modelin insan tarafından yazılan şarkı sözlerinden ayırt edilmesini zorlaştırdığını ortaya koymuştur. Çalışmada kullanılan veri seti, Original Hip-Hop (Rap) Lyrics Archive (OHHLA)'dan alınan 64.542 şarkıdan oluşmaktadır. Modelin performansı, rölatif kafiye başarısı, kelime çeşitliliği, cümle uzunluğu ve stilistik uyumluluk gibi ölçütler üzerinden değerlendirilmiştir. Deneysel bulgular, hiyerarşik dil modelinin (HLM), geleneksel karakter ve hece tabanlı modellerden daha başarılı olduğunu, ancak koşullu şablon mekanizmasının hiyerarşik model üzerinde tam potansiyelini gerçekleştiremediğini göstermektedir. Gelecek çalışmalar için, modelin daha büyük ölçekli veri kümeleriyle eğitilmesi, hip-hop dışındaki müzik türlerine genişletilmesi ve şarkı sözlerinin anlatım bütünlüğünü artıracak yeni yöntemlerin uygulanması önerilmektedir. (Manjavacas ve ark., 2020)

Ferreira, Limongi ve Fávero (2023) tarafından gerçekleştirilen çalışmada, derin öğrenmeye dayalı dil modellerinin müzik üretimi üzerindeki etkileri incelenmiştir. Çalışmanın temel amacı, Transformer tabanlı modellerin sembolik müzik üretiminde ne kadar başarılı olduğunu insan algısı üzerinden değerlendirmektir. Müzik üretiminde kalite değerlendirmesi genellikle düşük seviyeli matematiksel metriklere dayanırken, bu çalışmada insan merkezli bir yaklaşım benimsenmiş ve 117 katılımcıdan oluşan bir grup üzerinde kör testler uygulanmıştır. Çalışmada kullanılan dört farklı model arasında PerformanceRNN, Music Transformer, MuseNet ve özel olarak geliştirilen bir GRU modeli yer almaktadır. Veri seti olarak, Maestro, JSB Chorales ve MuseData kullanılmış olup, klasik piyano müziğine odaklanılmıştır. Model performanslarını karşılaştırmak amacıyla insan bestesi ile yapay zeka tarafından üretilen besteler yan yana sunulmuş ve katılımcılardan kompozisyonların insana mı yoksa yapay zekaya mı ait olduğunu tahmin etmeleri istenmiştir. Kör test sonuçları, Transformer mimarisine dayalı modellerin, özellikle MuseNet ve Music Transformer'ın, insan bestelerine kıyasla daha yüksek müzikal beğeni kazandığını göstermiştir. Buna karşın, RNN tabanlı modellerin (PerformanceRNN ve Custom GRU) insan bestelerinden daha düşük puan aldığı görülmüştür. Ayrıca, klasik müzik bilgisine sahip katılımcıların ve müzikal duyarlılığı yüksek bireylerin, yapay zeka tarafından üretilen müzikleri insan bestelerinden daha doğru şekilde ayırt edebildiği belirlenmiştir. Gelecek çalışmalar için çok sesli müzik

üretimi, duygusal uyumluluğun artırılması ve daha geniş veri setleriyle eğitilmiş büyük ölçekli modellerin geliştirilmesi önerilmektedir (Ferreira ve ark., 2020).

Chuang, Hsu, Lee ve Chang (2022) tarafından geliştirilen çalışmada ise, RapGAN adlı GAN tabanlı bir model kullanarak rap şarkı sözü üretimi üzerine odaklanmıştır. Rap şarkı sözleri üretimi, serbest biçimli, uzun ve stilistik olarak tutarlı metinler gerektirdiğinden, geleneksel dil üretim teknikleri bu tür için yetersiz kalmaktadır. Çalışmada, GAN tabanlı modellerin klasik LSTM (Long Short-Term Memory) modellerine kıyasla daha orijinal, akıcı ve çeşitli şarkı sözleri üretebileceği öne sürülmektedir. Bu amaçla, PRO (Phrase Roll-Out) ve AREGS (Attention Reward at Every Generation Step) adı verilen iki yeni mekanizma geliştirilmiştir. PRO, GAN eğitim sürecini hızlandırarak yalnızca anlamlı kelime öbeklerini ödüllendirirken, AREGS, dikkat mekanizmasına dayalı bir ödüllendirme sistemi ile modelin orijinallik ve çeşitlilik düzeyini artırmaktadır. Modelin başarısını ölçmek için Sim-EL (mevcut şarkı sözlerine benzerlik), BLEU (dilsel doğruluk), Self-BLEU (çeşitlilik) ve RD (ritim yoğunluğu) gibi metrikler kullanılmıştır. Çalışmada kullanılan veri seti, 160.000 Çince rap şarkı sözünden oluşmakta ve geniş kapsamlı bir eğitim sağlamak adına 3.214 farklı sanatçının eserlerini içermektedir. Deneysel sonuçlar, RapGAN'ın Ghost Writer, SeqGAN ve MaliGAN gibi diğer modellerden daha üstün performans gösterdiğini ortaya koymuştur. Özellikle insan değerlendirme testlerinde, modelin akıcılık, orijinallik ve çeşitlilik açısından önceki yöntemlere kıyasla daha başarılı olduğu belirlenmiştir. Gelecekte, modelin farklı dillerde eğitilmesi, müzikal uyumluluğun artırılması ve daha büyük ölçekli dil modelleriyle karşılaştırmalı analizler yapılması önerilmektedir (Chuang ve ark., 2020).

Yu, Srivastava ve Canales (2021) tarafından gerçekleştirilen çalışmada, şarkı sözlerinden melodi üretimi için koşullu Uzun Kısa Süreli Bellek - Generative Adversarial Network (Conditional LSTM-GAN) tabanlı bir model önerilmektedir. Geleneksel müzik üretim teknikleri, şarkı sözleri ve melodi arasındaki ilişkileri tam olarak modelleyemediğinden, yapay sinir ağlarına dayalı bir yaklaşım benimsenmiştir. Çalışmada, LSTM tabanlı bir jeneratör ve diskriminatör modeli kullanılarak şarkı sözlerine dayalı melodi üretimi gerçekleştirilmiştir. Model, özellikle hece-melodi hizalama (syllable-melody alignment) ilişkisini öğrenerek şarkı sözleriyle müzikal yapının daha iyi eşleşmesini sağlamaktadır. Modelin başarısını değerlendirmek için Maksimum Ortalama Farklılık (MMD), BLEU skoru ve ölçek tutarlılığı (scale consistency) gibi metrikler kullanılmıştır. Çalışmada, 12.197 MIDI şarkısından oluşan büyük ölçekli bir veri seti oluşturulmuş ve söz-melodi eşleşmeleri otomatik olarak

çıkarılmıştır. Deneysel sonuçlar, Conditional LSTM-GAN modelinin rastgele örnekleme ve Maksimum Olasılık Tahminleme (MLE) yöntemlerinden daha başarılı olduğunu göstermektedir. Ayrıca, modelin ölçek tutarlılığı açısından insan bestelerine yakın sonuçlar üretebildiği, ancak müzikal anlatım açısından hâlâ iyileştirme gerektirdiği belirtilmiştir. Gelecekte, modelin çok sesli müzik (polyphonic music) üretimi, eksik sözlerden melodi üretme ve tersine şarkı sözleri tahmini gibi yeni araştırma alanlarına genişletilmesi önerilmektedir (Yu ve ark., 2020).

Saeed, Ilić ve Zangerle (2019) tarafından gerçekleştirilen çalışmada, GAN tabanlı yaratıcı metin üretimi üzerine odaklanılmıştır. Geleneksel dil modelleri, genellikle maksimum olabilirlik tahminine (MLE) dayalı olduğundan, yaratıcı metin üretimi konusunda sınırlamalar içermektedir; çünkü üretilen metinler genellikle tekrarlayıcı, öngörülebilir ve yaratıcı olmayan içerikler barındırmaktadır. Bu çalışmada, şiir, şarkı sözleri ve metafor üretimi için Creative-GAN adlı yeni bir GAN tabanlı model önerilmektedir. Modelin amacı, geleneksel dil modellerinin sınırlamalarını aşarak çeşitlilik, özgünlük ve bağlamsal bütünlüğü artırmaktır. Creative-GAN modeli, AWD-LSTM ve Transformer-XL tabanlı dil modellerini kullanarak eğitilmiş ve ödüllendirme mekanizması ile yaratıcı kelime seçimlerini teşvik eden bir diskriminatör içermektedir. Model, üç farklı veri seti üzerinde test edilmiştir: 740 klasik ve çağdaş İngilizce şiir, 14.950 metafor cümlesi ve 1.500 farklı türde şarkı sözü. Değerlendirme sürecinde, perplexite (perplexity), BLEU skoru ve insan değerlendirmesi gibi ölçütler kullanılmıştır. Sonuçlar, Creative-GAN'ın geleneksel dil modellerinden ve GumbelGAN gibi diğer jeneratif modellerden daha yaratıcı ve özgün metinler üretebildiğini göstermektedir. Özellikle, Transformer-XL tabanlı versiyonun en düşük perplexite skorunu elde ettiği belirlenmiştir. Gelecekte, modelin daha büyük veri setleriyle eğitilmesi, farklı kültürel ve dilsel bağlamlarda test edilmesi ve yaratıcı metinlerin anlam bütünlüğünü artırmaya yönelik yeni mekanizmaların geliştirilmesi önerilmektedir (Saeed ve ark., 2020).

Potash, Romanov ve Rumshisky (2015) tarafından gerçekleştirilen GhostWriter çalışması, rap şarkı sözü üretimi için Uzun Kısa Süreli Bellek (LSTM) tabanlı bir model önermektedir. Çalışmanın temel amacı, belirli bir sanatçının stilini taklit ederek orijinal ancak benzer şarkı sözleri üretmek, yani "ghostwriting" sürecini otomatikleştirmektir. Önceki çalışmalar, şarkı sözü üretimi için şablon tabanlı yaklaşımlar kullanırken, bu model kendi kafiye şemasını, satır uzunluğunu ve kıta yapısını kendiliğinden belirleyebilmektedir. Modelin başarısını değerlendirmek için, rölatif kafiye yoğunluğu

(rhyme density), kelime çeşitliliği ve metrik benzerlik (cosine similarity) gibi metrikler kullanılmıştır. Model, The Original Hip-Hop (Rap) Lyrics Archive (OHHLA)'dan elde edilen 14 sanatçıya ait şarkı sözleriyle eğitilmiş ve deneyler özellikle Fabolous adlı rap sanatçısının stilini taklit etmeye odaklanmıştır. 219 kıta ve her kıtada en az 175 kelime içeren bir eğitim kümesi kullanılmıştır. LSTM modeli, n-gram tabanlı bir Markov modeli ile karşılaştırılmış ve orijinal içerik üretme ve stil taklit etme konusunda daha başarılı olduğu belirlenmiştir. Çalışmada, "<endLine>" ve "<endVerse>" gibi özel belirteçlerin kullanımı, modelin kendi kıta yapısını oluşturmaya sağlamış ve kafiye algılama mekanizmasını iyileştirmiştir. Deneysel bulgular, LSTM tabanlı modelin geleneksel dil modellerine kıyasla daha akıcı, yaratıcı ve stilistik açıdan tutarlı şarkı sözleri üretebildiğini ortaya koymuştur. Gelecek çalışmalar için, modelin farklı sanatçılar üzerinde test edilmesi, fonetik özelliklerin modellenmesi ve ritmik yapının daha iyi analiz edilmesi önerilmektedir (Potash ve ark., 2020).

Chu, Urtasun ve Fidler (2017) tarafından gerçekleştirilen çalışmada, pop müzik üretimi için çok katmanlı bir Hiyerarşik Tekrarlayan Sinir Ağı (Hierarchical Recurrent Neural Network - HRNN) modeli önerilmektedir. Geleneksel müzik üretim modelleri genellikle tek kanallı melodi üretimine odaklanırken, bu çalışma melodi, akorlar ve davulları aynı anda üretebilen çok katmanlı bir sinir ağı tasarımı sunmaktadır. Modelin tasarımı, müziğin yapısal bileşenleri arasındaki hiyerarşik ilişkileri anlamak için müzik teorisine dayalı ön bilgiler içermektedir. Alt katmanlar melodiyi üretirken, üst katmanlar davul ve akorları oluşturarak müziğin tamamlanmasını sağlar. Model, ölçek (scale) türüne dayalı bir koşullama yaparak belirli bir müzik ölçeğinde öğrenmeyi sağlar ve müzikal kuralları göz önünde bulundurarak öğrenme sürecini iyileştirir. Modelin başarısını değerlendirmek amacıyla 100 saatlik pop müzik MIDI verisi üzerinde eğitim yapılmış ve sonuçlar, Google'ın Magenta projesi tarafından geliştirilen modele kıyasla insan katılımcılar tarafından daha yüksek beğeni oranı elde ettiğini göstermiştir. Değerlendirmeler, müziğin akıcılığı, ritmik uyumu ve melodik tutarlılığı gibi kriterlere dayanmaktadır. Çalışmada, önerilen modelin "Neural Dancing and Karaoke" ve "Neural Story Singing" gibi iki farklı uygulama alanında kullanılabileceği gösterilmiştir. İlk uygulama, müziğe eşlik eden dans hareketleri ve şarkı sözleri üretmeye yönelik bir sistem sunarken, ikinci uygulama, görüntülerden otomatik olarak pop şarkıları oluşturmayı amaçlamaktadır. Deneysel sonuçlar, çok katmanlı hiyerarşik bir yapının, tek katmanlı geleneksel modellerden daha başarılı olduğunu ve müzikal bileşenler arasındaki uyumu artırdığını göstermektedir. Gelecek çalışmalar için, modelin enstrüman çeşitliliğinin

artırılması, farklı müzik türlerine uyarlanması ve yapay zeka destekli besteleme sistemlerinin daha büyük veri kümeleriyle eğitilmesi önerilmektedir (Chu ve ark., 2020).

Fukayama, Nakatsuma, Sako, Nishimoto, Ono ve Sagayama (2010) tarafından gerçekleştirilen çalışmada, Japonca şarkı sözlerinden otomatik olarak şarkı besteleme yöntemi geliştirilmiştir. Geleneksel müzik üretim yöntemleri genellikle belirli kurallar çerçevesinde çalışırken, bu çalışmada şarkı sözü prosodisine dayalı bir otomatik beste algoritması önerilmektedir. Önerilen sistem, Japonca şarkı sözlerinin tonlama (pitch accent) özelliklerini analiz ederek, bu bilgilere dayalı olarak melodi üretimi yapmaktadır. Çalışmada geliştirilen Orpheus adlı sistem, kullanıcıların girdiği Japonca şarkı sözlerini analiz ederek, uygun ritmik yapı ve akor dizisi ile uyumlu bir melodi oluşturabilmektedir. Model, dinamik programlama (dynamic programming) yöntemiyle en uygun melodi dizisini belirleyerek prosodiye uygun müzikal yapı oluşturmaktadır. Modelin başarısını değerlendirmek amacıyla bir klasik müzik bestecisi tarafından 59 üretilmiş şarkı üzerinde değerlendirme yapılmış ve bu şarkıların %83,1'inin klasik müzik teorisine uygun olduğu belirlenmiştir. Ayrıca, kullanıcı deneyimi değerlendirmeleri kapsamında, Orpheus sisteminin iki yıl boyunca yaklaşık 87.000 şarkı ürettiği ve 1.378 kullanıcının %70,8'inin üretilen şarkıları çekici bulduğu rapor edilmiştir. Çalışmada, yöntemin yalnızca Japonca gibi "pitch accent" dillerine uygulanabileceği, ancak gelecekte İngilizce gibi "stress accent" dillerine genişletilerek metrik yapının da dikkate alınması gerektiği belirtilmiştir. Bu çalışma, amatör müzisyenlerin şarkı yazım sürecine katılımını artırabilecek yaratıcı bir müzik üretim sisteminin temelini atmaktadır (Fukayama ve ark., 2020).

Lee, Fan ve Ma (2019) tarafından geliştirilen iComposer, Çince pop müziği için otomatik şarkı yazma amacıyla tasarlanmış interaktif bir web tabanlı sistemdir. iComposer, verilen bir metne uygun melodi üretme ve tersine, belirli bir melodiye uygun şarkı sözleri oluşturma yeteneğine sahip iki yönlü bir şarkı yazma modeli sunmaktadır. Modelin temel bileşenleri, melodi, ritim ve şarkı sözlerini ayrı ayrı tahmin eden üç farklı sıralı (sequence-to-sequence) modelden oluşmaktadır. LSTM (Uzun Kısa Süreli Bellek) tabanlı bir sinir ağı yapısı kullanılarak, sistemin geçmiş verilerden öğrenmesi sağlanmış ve bu doğrultuda melodi ve şarkı sözü tahminleri yapılmıştır. Veri seti, 1000 Çince pop müzik parçasından elde edilen MIDI dosyaları ve 300.000 karakter-not eşleşmesi içeren bir sözlük kullanılarak oluşturulmuştur. Deneysel sonuçlarda, iComposer'ın ürettiği şarkılar, insan yapımı ve rastgele oluşturulan müziklerle karşılaştırılmış ve 30 katılımcı tarafından değerlendirilen anket sonuçlarına göre, iComposer tarafından üretilen müziklerin, insan bestelerine yakın kalitede olduğu belirlenmiştir. Katılımcılar, şarkıların

melodik akıcılık, şarkı sözlerinin anlamlılığı ve melodi-şarkı uyumu gibi kriterlere göre değerlendirme yapmış ve iComposer, özellikle şarkı sözleri üretimi konusunda yüksek puan almıştır. Gelecek çalışmalar için modelin, Çince tonal sistemle daha uyumlu hale getirilmesi, nakarat ve kıta yapılarının ayırt edilmesi ve çoklu müzik türlerine genişletilmesi önerilmektedir. Bu sistem, hem profesyonel hem de amatör müzisyenlerin şarkı yazma süreçlerini kolaylaştırarak, yapay zeka destekli yaratıcı müzik üretiminde önemli bir adım atmaktadır (Lee ve ark., 2020).

Güçük (2021) tarafından gerçekleştirilen çalışmada, Üretken Rakip Ağlar GAN kullanılarak Türkçe metin üretimi ele alınmıştır. NLP alanında yaşanan temel sorunlardan biri, etiketli veri eksikliği ve belirli sınıflardaki dengesiz veri dağılımıdır. Bu problem, haber metinlerinin olumlu veya olumsuz olarak sınıflandırılmasında doğrudan bir performans kaybına neden olmaktadır. Çalışmada, bu eksikliği gidermek amacıyla GAN tabanlı bir metin üretim modeli önerilmiştir. Model, LSTM tabanlı bir üretici ağı ve SVM ile TF-IDF tabanlı bir ayırıcı ağı kullanarak çalışmaktadır. Kullanılan veri seti, 3058 haber metninden oluşmakta olup, 2949 olumlu ve 109 olumsuz haber içermektedir. Olumsuz sınıfa ait örneklerin yetersizliği nedeniyle, GAN modeliyle bu sınıf için 50'den 2750'ye kadar yeni veri üretilmiş ve bu veri, sınıflandırma başarısını artırmada kullanılmıştır. Modelin başarısını değerlendirmek amacıyla n-gram, TF-IDF, lojistik regresyon ve destek vektör makinesi gibi farklı makine öğrenmesi yöntemleri kullanılmıştır. Deneysel sonuçlar, GAN tabanlı veri üretiminin sınıflandırma başarısını yaklaşık %47 oranında artırdığını ve toplam başarı oranını %90'ın üzerine çıkardığını göstermektedir. Ayrıca, modelin konuşma diline yakın Türkçe cümleler üretebildiği gözlemlenmiştir. Gelecek çalışmalar için, GAN modellerinin farklı NLP görevlerine genişletilmesi, veri artırma süreçlerinin optimize edilmesi ve farklı metin türleri için test edilmesi önerilmektedir (Güçük, 2020).

Çağlayan (2022) tarafından gerçekleştirilen bu çalışmada, kontrollü çok konulu metin üretimi için derin öğrenmeye dayalı yeni bir yaklaşım önerilmektedir. Çalışmanın temel amacı, GPT tabanlı bir dil modeli kullanarak farklı konulara özel metinler üretebilmek ve metin üretim sürecini daha kontrollü hale getirmektir. Bu doğrultuda, üç farklı yaklaşım önerilmiş ve Türkçe metinler üzerinde kapsamlı deneyler gerçekleştirilmiştir. İlk yaklaşımda, tek katmanlı bir GPT modeli kullanılarak konu bilgisi giriş dizisine eklenmiş ve modelin belirli bir konuya yönelik metinler üretmesi sağlanmıştır. İkinci yaklaşımda, TF-IDF yöntemi ile her konu için çıkarılan anahtar kelimeler giriş dizisine eklenerek modelin konu duyarlılığı artırılmıştır. Son olarak,

üçüncü yaklaşımda konu seçim sınıflandırıcısı (topic selection classifier) içeren yeni bir örnekleme stratejisi önerilmiş ve üretilen kelimelerin belirlenen konuyla uyumlu olup olmadığı doğrulanmıştır. Çalışmada kullanılan veri seti, TC32 adlı 32 farklı konudan oluşan 430.000 Türkçe metin içeren bir veri kümesidir. Bu veri kümesinden, beş farklı konuyu (finans, sağlık, turizm, spor, gıda) içeren TC5 adlı alt veri seti oluşturulmuş ve metin üretim deneylerinde kullanılmıştır. Kontrollü çok konulu metin üretiminin başarısını ölçmek için BLEU skoru, BERTScore ve insan değerlendirme testleri kullanılmıştır. Ayrıca, üretilen metinlerin istenen konuyla ne kadar uyumlu olduğunu belirlemek amacıyla Türkçe için eğitilmiş bir Transformer tabanlı metin sınıflandırma modeli geliştirilmiştir. Deneysel sonuçlar, konu seçim sınıflandırıcısı kullanan örnekleme stratejisinin en iyi performansı gösterdiğini ve modelin belirlenen konulara daha iyi odaklanmasını sağladığını ortaya koymuştur. Ancak, modelin bazen genel ifadeler üreterek belirli bir konunun derinliğini tam olarak yakalayamadığı gözlemlenmiştir. Gelecekte, modelin çok katmanlı GPT mimarileriyle geliştirilmesi, Türkçe dışındaki dillerde test edilmesi ve duygusal bağlamı da içerecek şekilde genişletilmesi önerilmektedir (Çağlayan, 2020).

Ergin (2023) tarafından gerçekleştirilen çalışmada, derin öğrenme algoritmaları kullanılarak karakter tabanlı bir dil modeli geliştirilmesi ve Türkçe dil bilgisi kurallarına uygun anlamlı kelime üretimi amaçlanmıştır. NLP alanında metin üretimi, bilgisayarların insanların iletişim kurma şeklini taklit ederek doğal dilde metin oluşturmasını sağlar. Bu süreçte, kelimeler ve karakterler arasındaki ilişkileri öğrenmek için özyinelemeli yapay sinir ağları (RNN) kullanılmıştır. Çalışmada, LSTM (Uzun Kısa Süreli Bellek), GRU (Geçitli Tekrarlayan Birim), Kodlayıcı-Kod Çözücü ve Dikkat Mekanizması içeren derin öğrenme modelleri karşılaştırılmıştır. Kodlayıcı-Kod Çözücü ve Dikkat modeli, uzun kelime dizileri üretmede diğer modellere göre daha başarılı bulunmuş ve özellikle uzun giriş ve çıkış dizelerinde karakterler arasındaki ilişkileri daha iyi hatırlayarak en iyi sonuçları vermiştir. Modelin başarısını değerlendirmek için farklı epoch değerlerinde (100, 300, 500) sıcaklık örnekleme yöntemi uygulanmış ve en yüksek başarı oranı Kodlayıcı-Kod Çözücü ve Dikkat modeliyle %95,7 olarak elde edilmiştir. Çalışma sonunda, karakter tabanlı geliştirilen modellerin Türkçe dil bilgisi kurallarına uygun kelime ve kelime grupları üretebildiği gözlemlenmiştir. Gelecek çalışmalar için daha büyük veri setleriyle modelin eğitilmesi, kelime tabanlı üretim yöntemleriyle entegrasyon sağlanması ve daha gelişmiş NLP tekniklerinin uygulanması önerilmektedir (Ergin, 2020).

Hussein (2024) tarafından gerçekleştirilen çalışmada, RNN ve LSTM katmanları kullanılarak metin üretimi ele alınmıştır. Bu çalışmanın temel amacı, NLP alanında bağlamsal anlamayı daha iyi yakalayabilen bir dil modeli geliştirmektir. Modelin Nietzsche ve Shakespeare metinleri üzerinde eğitilerek, farklı edebi üsluplarla başa çıkma yeteneğinin değerlendirilmesi amaçlanmıştır. Çalışmada, kelime vektörlerinin kümeleme teknikleri kullanılarak bağlamsal analizinin güçlendirilmesi hedeflenmiştir. Bu yaklaşımın temel hipotezi, bağlamın daha iyi anlaşılmasının metin üretiminin doğruluğunu ve tutarlılığını önemli ölçüde artırabileceğidir. Metodoloji açısından, çalışma ön işlenmiş veri setleri üzerinde kelime bağlam vektörlerini çıkaran bir kümeleme tekniği uygulamaktadır. Bu bağlam vektörleri, geleneksel giriş verileriyle birleştirilerek daha gelişmiş bir metin üretim modeli oluşturulmuştur. Modelin başarısı, Nietzsche ve Shakespeare veri setleri üzerinde eğitilerek ölçülmüş ve LSTM tabanlı RNN modelinin uzun vadeli bağımlılıkları koruma kapasitesi sayesinde başarılı sonuçlar elde edilmiştir. Model, Nietzsche veri setinde %95,21 doğruluk ve 0,25 kayıpla, Shakespeare veri setinde ise %91,25 doğruluk ve 0,3876 kayıpla başarılı bir performans sergilemiştir. Bu sonuçlar, bağlam vektörlerinin dil modeli eğitiminde kullanımının etkinliğini doğrulamakta ve modelin farklı edebi stiller ve dönemler arasında çok yönlü olduğunu göstermektedir. Çalışma ayrıca, yenilikçi bir geri bildirim mekanizması tanıtarak, modelin sürekli değerlendirilmesini ve ayarlanmasını sağlamış, yalnızca yeterli doğruluk ve güvenilirlik seviyesine ulaştığında eğitimin durdurulmasını garanti etmiştir. Gelecek çalışmalar için çok dilli metin üretimi, Transformer tabanlı modellerle karşılaştırmalı analizler ve RNN ile diğer sinir ağı mimarilerinin (örneğin GRU) entegrasyonu önerilmektedir. Çalışma, bağlamsal anlamayı geliştiren dil modeli yaklaşımlarına önemli bir katkı sunarak, daha gelişmiş ve doğru metin üretim teknolojilerine yönelik yeni araştırmalara kapı açmaktadır (Hussein, 2020).

Kırmacı (2015) tarafından gerçekleştirilen bu çalışmada, Türkçe şarkı sözlerinden yola çıkarak müzik üst-verilerini tahmin edebilen bir makine öğrenmesi modeli önerilmektedir. Günümüzde müzik geri getirme sistemleri, genellikle melodik ve aranjman özelliklerine dayanmaktadır; ancak bu çalışmada şarkı sözü metinlerinin, müzik türü, söz yazarı ve yayın tarihi gibi üst-verileri tahmin etmede ne kadar etkili olduğu araştırılmıştır. Çalışmada, Naïve Bayes ve Destek Vektör Makineleri (SVM) gibi makine öğrenmesi algoritmaları kullanılarak şarkı sözü yazarlarını, şarkının türünü ve yayın tarihini tahmin eden bir sistem geliştirilmiştir. Modelin başarısını artırmak için kelime kökleri, karakter n-gramları, sonek n-gramları, satır uzunluğu istatistikleri gibi çeşitli

öznitelik çıkarım teknikleri uygulanmıştır. Kullanılan veri seti, farklı tarzlara sahip geniş bir şarkı sözü koleksiyonundan oluşmaktadır ve 1048 Türkçe şarkı içermektedir. Modelin performansı, doğruluk oranı (accuracy), anma (recall), kesinlik (precision), F1 skoru ve ROC eğrisi gibi metrikler kullanılarak değerlendirilmiştir. Sonuçlar, şarkı sözü metinlerinin müzik üst-verisi tahmini için kullanılabileceğini ve metin madenciliğinin müzik bilgi geri getirimi sistemlerine önemli bir katkı sağlayabileceğini göstermektedir. Gelecekte, modelin daha büyük veri setleriyle eğitilmesi, farklı müzik türlerine uygulanması ve NLP teknikleri ile entegrasyonu önerilmektedir (Kırmacı, 2020).

Yi, Li ve Sun (2020) tarafından gerçekleştirilen çalışmada, Çince klasik şiir üretimi için RNN Encoder-Decoder mimarisine dayalı yeni bir model geliştirilmiştir. Geleneksel kurallara dayalı şiir üretim teknikleri, belirli kalıplara bağlı kalırken, bu çalışmada derin öğrenme tabanlı bir yöntem kullanılarak, yapay zekanın ritim, tonalite ve yapısal kuralları kendiliğinden öğrenmesi sağlanmıştır. Çalışmanın amacı, verilen bir anahtar kelimeye dayalı olarak, klasik Tang şiirlerine uygun dört mısralık Jueju türünde şiirler üretmek ve bu süreçte hem anlamsal bütünlük hem de şiirsel kurallara uygunluk sağlamaktır. Model, her bir mısranın anlam bütünlüğünü koruması, önceki mısralarla tutarlı olması ve klasik şiir formlarına uygun şekilde tonalite ile ritmik kurallara uyum sağlaması için tasarlanmıştır. Çalışmada, şiir üretimi için Üç aşamalı bir RNN Encoder-Decoder modeli önerilmiştir: (1) Kelime-mısra üretim bloğu (Word-to-Line Block-WPB), anahtar kelimeye dayalı ilk mısrayı üretir. (2) Mısra-mısra üretim bloğu (Sentence-to-Sentence Block-SPB), bir önceki mısraya dayanarak ikinci mısrayı oluşturur. (3) Bağlamsal şiir üretim bloğu (Context-to-Line Block-CPB), şiirin geri kalan mısralarını üretirken önceki iki mısrayı bağlamsal olarak dikkate alır. Ayrıca, dikkat (attention) mekanizması kullanılarak karakterler arasındaki ilişkilerin daha iyi yakalanması sağlanmıştır. Model, Tang Hanedanlığı'ndan günümüze kadar 398.391 şiir içeren bir veri seti ile eğitilmiş ve 10.000 şiir içeren bir test kümesi ile değerlendirilmiştir. Deneysel sonuçlara göre, modelin BLEU skoru ve insan değerlendirme testlerinde önceki sistemlerden daha yüksek puanlar aldığı belirlenmiştir. %95'ten fazla üretilen şiirin tonal kurallara uygun olduğu ve dikkat mekanizması sayesinde karakterler arasındaki anlam ilişkilerinin güçlendiği gösterilmiştir. Özellikle, tersten hedef cümle eğitimi (target inversion training) uygulandığında, modelin şiir üretim kalitesinin önemli ölçüde arttığı gözlemlenmiştir. Gelecekte, modelin diğer klasik şiir türleri (Songci ve Yuefu) üzerine genişletilmesi, şiirsel anlatım zenginliğini artırmak için daha büyük bir bağlamsal anlayış

mekanizması entegre edilmesi ve farklı dillerde klasik şiir üretimi için uyarlanması önerilmektedir (Yi ve ark., 2020).

Wentink (2023) tarafından gerçekleştirilen çalışmada, yüksek kafiye yoğunluğuna sahip rap şarkı sözleri üretmek için bir yapay zeka modeli geliştirilmiştir. Rap müziğinde, kafiye yoğunluğu ve çok heceli kafiyeler, liriklerin akıcılığı ve dinleyici üzerindeki etkisi açısından kritik öneme sahiptir. Bu nedenle, çalışma çok heceli kafiye tespiti eden ve cümleler arasındaki anlamsal bütünlüğü koruyan bir model geliştirmeyi hedeflemektedir. Model, girdi olarak verilen bir cümleye uygun, yüksek kafiye yoğunluğuna sahip ve anlamlı bir yanıt üreten bir sistem tasarlanmaktadır.

Önerilen sistem, Bidirectional Encoder Representations from Transformers (BERT) modelini kullanarak, önceden belirlenen kafiye kelimeleri arasındaki boşlukları doldurmayı amaçlamaktadır. Modelin temel bileşenleri şunlardır: (1) Veri Kümesi: Genius Lyrics'ten toplanan 36 farklı rap sanatçısına ait şarkı sözleri ile oluşturulmuştur. (2) Bi-gram Modeli: Metinleri iki kelimelik gruplara bölerek kafiye ilişkilerini tespit etmektedir. (3) Kelime Gömme (Word Embeddings) Tekniği: FastText kullanılarak kelimelerin anlamsal yakınlıkları hesaplanmaktadır. (4) Kafiye Algılama Modülü: CMU telaffuz sözlüğü kullanılarak kelimeler fonetik olarak analiz edilmekte ve çok heceli kafiye belirlenmektedir. (5) BERT Tabanlı Dil Modeli: Maskelenmiş kelimeleri tamamlayarak, cümlelerin dilbilgisel doğruluğunu artırmaktadır.

Çalışmada modelin başarısını ölçmek için insan katılımcılar tarafından yapılan anketler kullanılmış ve kafiye kalitesi ile anlam bütünlüğü (coherence) açısından değerlendirilmiştir. 40 katılımcı ile yapılan deneylerde, insan yazımı şarkı sözleri ortalama 5.52 kafiye puanı ve 5.0 anlam bütünlüğü puanı alırken, üretilen şarkı sözleri 3.14 kafiye puanı ve 3.26 anlam bütünlüğü puanı almıştır. Sonuçlar, modelin mevcut insan yazımı sözlerle kıyasla daha düşük bir performans sergilediğini, ancak bazı durumlarda katılımcıları etkileyen yaratıcı sonuçlar üretebildiğini göstermektedir.

Çalışmada ayrıca, sistemin geliştirilmesi gereken yönler tartışılmıştır. BERT modelinin rap sözleri ile özel olarak eğitilmemesi, dil modeli çıktılarının zaman zaman anlamsız veya ritim açısından uyumsuz olmasına neden olmuştur. Ayrıca, çok heceli kafiye tespiti için kullanılan CMU telaffuz sözlüğünün bazı rap kelimelerini içermemesi, modelin eksikliklerinden biri olarak görülmüştür. Gelecekte, BERT modelinin rap müziğe özgü bir veri kümesiyle eğitilmesi, daha iyi kafiye tespiti için gelişmiş fonetik modellerin kullanılması ve kafiye içeren kelime önerilerinin daha geniş bir veri tabanından seçilmesi önerilmektedir (Wentink, 2020).

Mao, Shin ve Cottrell (2018) tarafından gerçekleştirilen çalışmada, belirli bir müzik stiline uygun müzik üretimi için derin öğrenme tabanlı bir model olan DeepJ tanıtılmaktadır. Geleneksel müzik üretim sistemleri belirli türlerle sınırlıyken, bu çalışma farklı bestecilerin tarzlarını öğrenebilen ve çoklu müzik stillerinde kompozisyon yapabilen bir model geliştirmeyi amaçlamaktadır. DeepJ, poli-fonik (çok sesli) müzik üretimi için Biaxial LSTM tabanlı bir mimariye dayanmaktadır.

Modelin temel özellikleri arasında müzikal stilin öğrenilmesi, dinamiklerin (nota volümü) modele entegre edilmesi ve stil tutarlılığının sağlanması bulunmaktadır. Çalışmada, modelin öğrenme sürecini iyileştirmek için nota dinamikleriyle (volume) ilgili yeni bir veri gösterimi kullanılmış ve modelin stil öğrenme yeteneği artırılmıştır. Veri seti olarak 23 farklı bestecinin eserlerinden oluşan bir MIDI veri seti kullanılmıştır. Model, her bir besteciye özgü müzikal unsurları öğrenmekte ve bunları bireysel ya da karışık stillerde yeni müzik parçaları üretmek için kullanmaktadır.

Modelin performansı, insan katılımcılar tarafından yapılan değerlendirme anketleri ile ölçülmüştür. Amazon Mechanical Turk üzerinden yapılan kullanıcı testlerinde, DeepJ tarafından üretilen parçaların %70 oranında Biaxial LSTM modelinden daha iyi olduğu belirlenmiştir. Ayrıca, müzikal tarzın korunma yeteneğini ölçmek için yapılan bir deneyde, katılımcıların DeepJ tarafından üretilen müzikleri doğru şekilde sınıflandırabildiği ve modelin stil uyumunu başarıyla yakalayabildiği gösterilmiştir.

Gelecekte, DeepJ modelinin uzun vadeli müzikal yapılar üretme yeteneğinin geliştirilmesi, daha fazla müzik türüyle eğitilmesi ve yaratıcı müzik besteleme süreçlerine entegre edilmesi önerilmektedir. Çalışma, özellikle film yapımcıları, oyun geliştiricileri ve besteciler için özelleştirilebilir müzik üretim sistemleri geliştirilmesine yönelik önemli bir adım olarak değerlendirilmektedir (Mao ve ark., 2020).

Zhang, Mao, Li, Jiang, Chen, Hu, Xi, Fan ve Huang (2020) tarafından geliştirilen Youling, yapay zeka destekli etkileşimli bir şarkı sözü üretim sistemi olarak tasarlanmıştır. Geleneksel şarkı sözü üretim sistemleri, genellikle tamamen otomatik ve tek geçişli metin üretimi yaparken, Youling, müzik yaratıcılarıyla iş birliği içinde çalışan bir yapay zeka asistanı olarak geliştirilmiştir. Sistem, hem tam metin üretim modu hem de etkileşimli üretim modu sunarak, kullanıcıların üretilen sözlerden en uygun olanlarını seçmesine ve düzenlemeler yapmasına olanak tanımaktadır.

Çalışmada, Youling sisteminin başarısını artırmak için GPT-2 tabanlı bir dil modeli kullanılmış ve model 300.000 şarkı sözüyle özel olarak ince ayar yapılmıştır. Modelin eğitimi, Çince internet romanlarından oluşan 30 GB'lık büyük bir metin veri kümesi

üzerinde gerçekleştirilmiştir. Sistemde, şarkı sözlerinin içeriğini ve biçimini kontrol etmek için çok yönlü parametreler bulunmaktadır. İçerik kontrollü özellikler arasında şarkı sözlerinin metin stili, duygusal tonu, teması ve anahtar kelimeler yer alırken, biçim kontrollü özellikler ise akrostiş kelimeler, kafiye düzeni, satır sayısı ve kelime sayısı gibi faktörleri içermektedir.

Youling, kullanıcıların belirtilen tema, duygu ve anahtar kelimelere uygun şarkı sözleri üretmesini sağlamak için ön işleme ve filtreleme teknikleri kullanmaktadır. Sistemin tam metin üretimi ve etkileşimli üretim modları kullanıcıların tercihlerine göre belirlenebilir. Etkileşimli modda, kullanıcılar bir satır yazabilir ve yapay zeka, önceki bağlamı dikkate alarak bir sonraki satır önerilerini üretebilir. Ayrıca, sistemde şarkı sözlerini revize etmek için bir düzeltme modülü bulunmaktadır. Kullanıcılar istenmeyen kelimeleri veya satırları değiştirebilir, sistemin önerdiği alternatifleri değerlendirebilir. Deneysel sonuçlar, Youling'in kullanıcı dostu bir arayüz ve çok yönlü kontrol özellikleriyle geleneksel şarkı sözü üretim sistemlerinden daha fazla esneklik sunduğunu göstermektedir. Yapılan değerlendirmeler, Youling'in ürettiği şarkı sözlerinin, dilbilgisi doğruluğu, anlamsal bütünlük ve stil uyumu açısından yüksek puan aldığını ortaya koymuştur. Gelecekte, sistemin farklı müzik türlerine uyarlanması, çok dilli üretim kapasitesinin genişletilmesi ve sesli melodi entegrasyonunun sağlanması önerilmektedir (Zhang ve ark., 2020).

Reddy, Yu, Harscoët, Canales ve Tang (2020) tarafından gerçekleştirilen çalışmada, şarkı sözü ve melodi üretimini otomatikleştiren bir yapay sinir ağı modeli olan Automatic Neural Lyrics and Melody Composition (AutoNLMC) önerilmektedir. Geleneksel şarkı yazma süreçleri genellikle söz ve melodinin ayrı ayrı üretilmesiyle gerçekleşirken, bu çalışmada her iki süreci birleştiren tam otomatik bir sistem geliştirilmiştir. Model, lyric2vec adlı bir kelime gömme yöntemi kullanarak şarkı sözlerini vektör haline getirmekte ve bu vektörleri bir encoder-decoder mimarisinde kullanarak şarkı sözleriyle uyumlu melodiler üretmektedir. Sistem, hem sıfırdan şarkı oluşturabilmekte hem de var olan şarkı sözlerine uygun melodiler üretebilmektedir.

Çalışmada, modelin eğitimi için 12.023 MIDI dosyasından oluşan geniş ölçekli bir veri kümesi kullanılmıştır. Bu veri kümesi, Lakh MIDI Dataset ve çeşitli Reddit müzik koleksiyonlarından derlenmiştir. Veri, hece, kelime ve cümle seviyelerinde analiz edilerek gömüleme teknikleriyle işlenmiş ve bu sayede modelin şarkı sözlerinin fonetik ve ritmik yapısını öğrenmesi sağlanmıştır. Model, LSTM tabanlı bir Encoder-Decoder

yapısı ile çalışmakta olup, dikkat mekanizması (attention mechanism) kullanarak üretilen melodinin şarkı sözleriyle daha iyi eşleşmesini sağlamaktadır.

Modelin başarısını değerlendirmek için BLEU skoru, özgün ton aralığı (tone span) ve benzersiz nota sayısı (unique tones) gibi metrikler kullanılmıştır. Öznitelik tabanlı (ASW) kelime gömme yönteminin en yüksek başarıyı sağladığı, modelin özellikle uzun vadeli melodik tutarlılığı yakalayabildiği gösterilmiştir. Yapılan subjektif değerlendirmelerde, modelin ürettiği şarkılar insan tarafından yazılan şarkılarla karşılaştırılmış ve AutoNLMC'nin, müzikal uyum açısından yüksek bir performans sergilediği belirlenmiştir. Ancak modelin zaman zaman monoton melodiler üretebildiği ve belirli ritim kalıplarına bağlı kaldığı görülmüştür.

Gelecek çalışmalar için, modelin farklı müzik türleriyle eğitilmesi, armonik yapıların da dikkate alınması ve şarkı sözlerine duygusal içerik katılması önerilmektedir. Bu çalışma, film, oyun ve eğitim teknolojileri gibi alanlarda yapay zeka tabanlı şarkı üretimi için önemli bir adım olarak değerlendirilmektedir (Reddy ve ark., 2020).

Castro ve Attarian (2018) tarafından gerçekleştirilen çalışmada, daha yaratıcı ve çeşitli şarkı sözleri üretmek için iki farklı dil modelinin birleştirildiği bir framework geliştirilmiştir. Geleneksel dil modelleri genellikle şarkı sözlerinin yapısal bütünlüğünü koruyamamakta veya kelime çeşitliliği açısından yetersiz kalmaktadır. Bu çalışmada, şarkı sözlerinin hem yapısal uyumluluğunu hem de kelime çeşitliliğini artıran bir model olan "RichLyrics" önerilmektedir.

Önerilen sistem, iki farklı Transformer Language Model (TLM) kullanmaktadır: LS (Lyric Structure Model) ve LV (Lyric Vocabulary Model). LS modeli, şarkı sözlerinin dizilim yapısını öğrenmek için eğitilirken, LV modeli daha geniş ve zengin bir kelime dağarcığı sağlamak amacıyla eğitilmiştir. LS modeli, farklı müzik türlerinden oluşan geniş bir şarkı sözü veri seti ile eğitilmiş, LV modeli ise Project Gutenberg'den alınan klasik edebiyat eserlerinden türetilmiş bir kelime seti kullanılarak eğitilmiştir. Modelin çıktıları, şarkı sözlerinde hem anlamsal bütünlük hem de estetik çeşitlilik sağlayacak şekilde harmanlanmaktadır.

Çalışmada modelin başarısı BLEU skoru, ortalama kelime uzunluğu, satır tekrar oranı ve kelime tekrar oranı gibi metriklerle ölçülmüştür. RichLyrics modeli, yalnızca şarkı sözlerinden eğitilen dil modellerine kıyasla daha düşük tekrar oranına ve daha geniş kelime çeşitliliğine sahip sonuçlar üretmiştir. Özellikle, PureLyrics modeli (yalnızca şarkı sözleriyle eğitilmiş) %67 oranında satır tekrarına sahipken, RichLyrics modeli bu oranı %48'e düşürerek daha yaratıcı çıktılar elde etmiştir.

Deneyisel sonuçlar, RichLyrics modelinin daha doğal ve çeşitli şarkı sözleri üretme konusunda önceki yöntemlerden üstün olduğunu göstermektedir. Ancak, model hala kafiye yapısı ve fonetik ritim konusunda iyileştirmeye ihtiyaç duymaktadır. Gelecek çalışmalar için, modelin daha büyük edebiyat veri setleriyle eğitilmesi, farklı müzik türlerine özgü varyasyonlar geliştirilmesi ve fonetik kafiye algılama mekanizmalarının entegre edilmesi önerilmektedir. Bu çalışma, yapay zeka destekli yaratıcı müzik üretimi için önemli bir katkı sunmaktadır (Castro ve Attarian, 2020).

Güçük, Durgut ve Fındık (2021) tarafından gerçekleştirilen çalışmada, Türkçe metin üretimi için GAN kullanımı ele alınmıştır. NLP alanında yaşanan temel zorluklardan biri, etiketli veri eksikliği ve sınıflandırma modellerinde belirli sınıfların dengesiz dağılımıdır. Çalışmanın amacı, Türkçe haber metinleri üzerinde GAN kullanarak eksik veri sınıflarını artırmak ve bu yöntemin sınıflandırma başarısına etkisini incelemektir.

Çalışmada, GAN tabanlı bir model geliştirilmiş ve özellikle eksik olan olumsuz haber metinleri üretilmiştir. Kullanılan veri seti, 2017 yılında internetten toplanmış 3058 Türkçe haber metninden oluşmaktadır. Bu haberler, 2949 olumlu ve 109 olumsuz haber içerecek şekilde dengesiz bir yapıya sahiptir. Bu dengesizliği gidermek amacıyla, GAN modeliyle veri artırımı gerçekleştirilmiş ve toplam 2750 ek olumsuz haber metni üretilmiştir. Üretilen metinlerin anlam bütünlüğü ve dilbilgisi doğruluğu açısından değerlendirilmesi için Zemberek NLP aracı kullanılarak metinlerin normalizasyonu sağlanmıştır.

Modelin başarısını değerlendirmek için Destek Vektör Makineleri (SVM), TF-IDF, lojistik regresyon ve n-gram tabanlı teknikler kullanılmıştır. Deneyisel sonuçlara göre, olumsuz haberleri sınıflandırma başarısı %48'den %95'e çıkarılarak yaklaşık %47'lik bir iyileşme sağlanmıştır. Olumlu haberlerin sınıflandırma başarısı ise %99 seviyesinde korunmuştur. Çalışmada ayrıca, GAN ile üretilen metinlerin içerik açısından anlamlı ve Türkçe dil yapısına uygun olduğu gözlemlenmiştir.

Gelecek çalışmalar için, GAN tabanlı modellerin daha büyük veri setlerinde test edilmesi, farklı metin türleriyle uygulanması ve Türkçe dil özelliklerine daha uygun metin üretimi için özel düzenleme mekanizmalarının entegre edilmesi önerilmektedir. Çalışma, GAN'ların doğal dil işleme alanındaki potansiyelini ortaya koyarak, özellikle etiketli veri eksikliği bulunan alanlarda etkin bir çözüm sunabileceğini göstermektedir (Güçük ve ark., 2020).

Joanna Krenz (2024) tarafından gerçekleştirilen çalışmada, Çin’de yapay zeka destekli şiir üretimi ile geleneksel ve modern şiir anlayışı arasındaki ayrım ve birleşme süreçleri ele alınmaktadır. Yapay zeka tarafından üretilen şiirler, hem estetik hem de kültürel kodların makine tarafından yeniden yorumlanmasını içerirken, çalışmada Çin’de geleneksel şiirin modern yapay zeka modelleriyle nasıl üretildiği, bu süreçte estetik, etik ve felsefi unsurların nasıl şekillendiği incelenmektedir.

Çalışma, yapay zeka destekli şiir üretiminin küresel tarihi ve Çin’deki durumu üzerine kapsamlı bir analiz sunmaktadır. Yapay zeka ile şiir üretiminin iki dalga halinde geliştiği belirtilmektedir: birinci dalgada insanlar makinelerden ilham alarak şiir dili ve biçimini genişletmeye çalışırken, ikinci dalgada makineler insan şiirini taklit ederek insan benzeri şiirler üretmeye çalışmaktadır. Özellikle sinir ağları ve NLP tekniklerinin, Çin şiir geleneği ile nasıl bütünleştirildiği araştırılmıştır.

Çalışmada, Çin’de yapay zeka ile üretilen şiirlerin iki ana gruba ayrıldığı belirtilmektedir: klasik şiir ve modern şiir. Klasik şiir üretimi için geliştirilen sistemler, geleneksel biçimlere (örneğin Tang Hanedanlığı’na ait beş karakterli veya yedi karakterli dörtlükler) sıkı sıkıya bağlıdır ve tonalite, ritim ve kafiye kurallarını otomatik olarak takip edebilmektedir. Modern şiir üretimi ise daha serbesttir ancak anlam bütünlüğü, yaratıcı ifade ve duygusal derinlik açısından zorluklar içermektedir.

Çin hükümetinin geleneksel şiir üretimi ve kültürel mirasın korunmasını teşvik eden bir teknoloji politikası izlediği belirtilirken, yapay zeka şiir projelerinin de bu politika doğrultusunda geliştirildiği ifade edilmektedir. Ancak, çalışmada vurgulanan önemli bir nokta, bu şiirlerin mühendisler ve yapay zeka uzmanları tarafından geliştirildiği ve bu nedenle edebi perspektiften çok teknik bir yaklaşımla ele alındığıdır. Çalışma, yapay zeka şiir üretiminin, klasik ve modern şiir arasındaki ayrımı bulanıklaştırarak, geleneksel şiir biçimlerini dijital ortamda yeniden şekillendirdiğini ve kültürel kodları algoritmalara nasıl işlediğini incelemektedir. Çalışma, yapay zeka şiir üretiminde ele alınan anahtar unsurların dilbilgisel kurallar, metin yapıları, algoritmaların eğitildiği veri setleri ve bu süreçte ortaya çıkan kültürel kodlar olduğunu vurgulamaktadır. Gelecek çalışmalar için, yapay zeka şiir modellerinin daha geniş kültürel bağlamlarla bütünleştirilmesi, insan-yapay zeka işbirliğiyle üretilen şiirlerin analiz edilmesi ve şiir üretiminde etik boyutların göz önünde bulundurulması önerilmektedir. Çalışma, yapay zeka şiir üretiminin, geleneksel şiir biçimleri ile modern şiir anlayışlarını nasıl uzlaştırabileceği üzerine önemli katkılar sunmaktadır (Krenz, 2020).

Anonymous Authors (2019) tarafından gerçekleştirilen çalışmada, düşmanca (adversarial) eğitim ile GAN arasındaki ilişki incelenmiştir. Geleneksel GAN modelleri, veri dağılımını öğrenerek gerçekçi görüntüler üretmeye çalışırken, adversarial saldırılar, makine öğrenimi modellerini yanıltmak amacıyla oluşturulan manipüle edilmiş girdileri kullanmaktadır. Çalışmanın temel amacı, adversarial eğitim ile GAN modellerinin birlikte nasıl çalıştırılabileceğini anlamak ve her iki tekniğin birbirini nasıl geliştirebileceğini incelemektir.

Araştırmada, adversarial eğitim yöntemlerinin sınırlamaları analiz edilmiştir. Bu bağlamda, mevcut adversarial eğitim yöntemlerinin modelin genelleme yeteneğini sınırladığı ve modelin test seti üzerindeki performansını düşürdüğü gözlemlenmiştir. Çalışmada, bu sorunu çözmek için, GAN tarafından üretilen "sahte" verilerle veri artırımı (data augmentation) yapılarak, modelin daha dayanıklı hale getirilebileceği önerilmektedir. Ayrıca, GAN eğitimi sırasında diskriminatörün (D) adversarial örneklerle karşı daha dayanıklı hale getirilmesi, üretilen verilerin kalitesini de iyileştirmektedir. Bu doğrultuda, araştırmacılar AdvGAN adlı yeni bir sistem geliştirmiştir. AdvGAN, üç temel bileşenden oluşmaktadır: (1) Üretici (Generator- G), sahte veriler üretir. (2) Diskriminatör (Discriminator- D), gerçek ve sahte verileri ayırt etmek için eğitilir. (3) Adversarial saldırgan (Adversarial Attacker), diskriminatörün zayıf yönlerini tespit etmek için saldırılar oluşturur. Çalışmada, bu üç bileşenin tek bir ağ içinde birlikte eğitildiği ve hem modelin dayanıklılığının artırıldığı hem de üretilen görüntülerin kalitesinin iyileştirildiği gösterilmiştir.

Deneysel sonuçlara göre, AdvGAN'ın mevcut adversarial eğitim tekniklerine kıyasla daha iyi genelleme sağladığı ve daha dayanıklı bir sınıflandırıcı oluşturduğu tespit edilmiştir. CIFAR-10 ve ImageNet veri kümeleri üzerinde yapılan testlerde, AdvGAN ile eğitilmiş modellerin, geleneksel adversarial eğitim tekniklerine göre daha az hassas olduğu ve yüksek doğruluk oranları sağladığı gösterilmiştir. Ayrıca, AdvGAN'ın ürettiği görüntülerin, mevcut GAN modellerine kıyasla daha yüksek kaliteye sahip olduğu belirlenmiştir.

Gelecek çalışmalar için, AdvGAN'ın farklı veri kümelerinde test edilmesi, NLP gibi farklı alanlara uygulanması ve daha büyük ölçekli modellerle karşılaştırılması önerilmektedir. Çalışma, adversarial eğitim ile GAN teknolojilerinin birbirini güçlendirebileceğini ortaya koyarak, hem daha dayanıklı makine öğrenimi modelleri geliştirilmesi hem de yüksek kaliteli üretken modellerin eğitilmesi açısından önemli bir katkı sunmaktadır (Anonymous Authors, 2020).

Xu, Wang, Zhuang, Wang ve Xiao (2019) tarafından gerçekleştirilen çalışmada, Çince şarkı sözleri üretimi için hece yapısına dayalı ve bağlamsal olarak koşullandırılmış bir model önerilmektedir. Geleneksel şarkı sözü üretim modelleri, şarkı sözlerinin melodi ile olan yapısal ve ritmik ilişkisini dikkate almadan çalışmaktadır. Bu çalışmada ise, şarkı sözlerinin melodiyle daha iyi uyum sağlamasını sağlamak için iki kanallı bir Seq2Seq (Sequence-to-Sequence) modeli geliştirilmiştir.

Çalışmanın temel yaklaşımı, şarkı sözlerini yalnızca kelimeler ve anlam açısından değil, aynı zamanda hece yapılarıyla da modellemektir. Bu doğrultuda, iki farklı RNN (Recurrent Neural Network) kodlayıcısı kullanılmıştır: (1) Hece yapısını kodlayan bir encoder, şarkının hece düzenini ve ritmik yapısını öğrenmektedir. (2) Bağlamsal olarak anlam kodlayan bir encoder, şarkı sözlerinin önceki cümleleriyle tutarlılığını sağlamak için çalışmaktadır. Model, dikkat mekanizması (attention mechanism) kullanarak müzik ve dil bileşenleri arasındaki ilişkiyi öğrenmektedir.

Çalışmada, modelin eğitimi için 160.000 şarkıdan oluşan büyük ölçekli bir Çince şarkı sözü veri seti kullanılmıştır. Ayrıca, modelin melodi ile uyumunu artırmak için 50.000 şarkıya ait müzikal notasyonlar crowdsourcing yöntemiyle işlenmiştir. 4,15 milyon cümle-anahtar kelime çiftinden oluşan bir eğitim kümesi kullanılarak, modelin şarkı sözlerinin semantik bütünlüğünü ve ritmik uyumunu öğrenmesi sağlanmıştır.

DeneySEL sonuçlara göre, iki kanallı Seq2Seq modeli, geleneksel Seq2Seq modellerine kıyasla daha yüksek doğruluk oranı ve şarkı-melodi uyumu sağlamıştır. Modelin başarısını değerlendirmek için otomatik ve insan tabanlı değerlendirme testleri uygulanmıştır. BLEU skoru ve melodi uyum ölçütleri kullanılarak yapılan otomatik değerlendirmelerde, modelin melodi yapısına uygunluk oranının %87,6 olduğu belirlenmiştir. İnsan değerlendirmelerinde ise, akıcılık (fluency), anlamlılık (meaningfulness), çeşitlilik (diversity) ve bütünlük (entirety) açısından geleneksel modellere kıyasla çok daha yüksek puanlar aldığı tespit edilmiştir.

Gelecek çalışmalar için, şarkı sözlerinin duygu analizi ile koşullandırılması, farklı müzik türlerine yönelik özelleştirilmiş modellerin geliştirilmesi ve modelin çok dilli şarkı üretimi için genişletilmesi önerilmektedir. Bu çalışma, otomatik şarkı sözü üretiminde dil ve müzik unsurlarını entegre eden yenilikçi bir model sunarak, NLP ve müzik yapay zekası alanlarına önemli bir katkı sağlamaktadır (Xu ve ark., 2020).

He, Zhou ve Jiang (2012) tarafından gerçekleştirilen çalışmada, Çince klasik şiir üretimi için İstatistiksel Makine Çeviri (Statistical Machine Translation- SMT) modelleri kullanılarak yeni bir yöntem geliştirilmiştir. Geleneksel şiir üretim sistemleri genellikle

şablon tabanlı yaklaşımlara dayanırken, bu çalışmada SMT modellerinin şiir üretimine nasıl adapte edilebileceği araştırılmıştır. Amaç, belirli anahtar kelimelere dayalı olarak klasik Tang şiiri formunda dört satırlık şiirler (quatrain) oluşturabilen bir model geliştirmektir.

Çalışmada önerilen sistem, her satırın belirli bir anlam ve yapıya sahip olması gerektiğini göz önünde bulundurarak cümle bazlı SMT modelleri kullanmaktadır. Bu bağlamda, her satır için özel olarak eğitilmiş bir SMT modeli kullanılmış ve böylece şiirin mantıksal tutarlılığı ve akıcılığı artırılmıştır. Ayrıca, daha önce üretilmiş olan satırların bağlamına uygun yeni satırlar üretmek için karşılıklı bilgi (mutual information) tabanlı bir model geliştirilmiştir. Model, dilin tonal yapısını, uyak düzenlerini ve yapısal kısıtlamaları dikkate alarak şiir üretmektedir.

Çalışmanın temel veri seti, Tang, Song, Ming ve Qing hanedanlıklarından derlenmiş 3,5 milyon cümleden oluşan büyük ölçekli bir klasik Çince şiir koleksiyonudur. Bu koleksiyon, toplamda 1,2 milyon cümle çifti içererek SMT modellerinin eğitimi için kullanılmıştır. Dil modellerinin daha geniş bir bağlamda eğitilebilmesi için, 12 milyon cümle içeren ek bir Çince metin koleksiyonu da veri setine eklenmiştir.

Çalışmada kullanılan değerlendirme metrikleri, SMT tabanlı çeviri sistemlerinde yaygın olarak kullanılan BLEU skoru ve insan değerlendirmeleri içermektedir. BLEU skoru ile yapılan analizlerde, geliştirilen modelin ürettiği şiirlerin insan tarafından yazılmış şiirlere daha yakın olduğu belirlenmiştir. İnsan değerlendirmelerinde ise, üretken modelin ürettiği şiirlerin akıcılık, uyum ve anlam açısından geleneksel yöntemlerden daha başarılı olduğu tespit edilmiştir.

Çalışmanın sonucunda, SMT tabanlı şiir üretiminin klasik Çin şiir yapısına uygun şekilde gelişebileceği ve makine öğrenmesi ile edebi üretimin daha başarılı bir şekilde birleştirilebileceği gösterilmiştir. Gelecek çalışmalar için, bu yöntemin diğer dillerdeki şiir türlerine uygulanması, kelime seçiminde daha yaratıcı varyasyonların üretilmesi ve anlam bütünlüğünü artırmak için daha gelişmiş dil modellerinin entegrasyonu önerilmektedir (He ve ark., 2020).

Ma, Sharma, Kan, Lee ve Wang (2024) tarafından gerçekleştirilen çalışmada, otomatik şarkı sözü üretiminde bağlam bütünlüğünü artırmak amacıyla KeYric adlı yeni bir sistem geliştirilmiştir. Günümüzde şarkı sözü üretiminde kullanılan dil modelleri, genellikle gramer ve ritim açısından kabul edilebilir sonuçlar verse de, bütünlük (coherence) konusunda eksiklikler taşımaktadır. Özellikle, şarkı sözlerinin kelime, cümle ve tam metin seviyelerinde tutarlılık göstermesi ve müziğin duygusal yapısı ile uyumlu

olması beklenmektedir. Çalışmada, KeYric modeliyle müzikten anahtar kelimeler çıkarılarak, şarkı sözlerinin anlamlı bir çerçevede üretilmesi sağlanmıştır.

KeYric sistemi, denetimsiz öğrenme (unsupervised learning) yöntemiyle müzikten anlamlı anahtar kelimeleri çıkaran bir iskelet çıkarıcı (keyword skeleton extractor) ve grafik tabanlı bir genişletici (graph-based skeleton expander) kullanmaktadır. Model, şarkının stiline ve duygusuna uygun olarak belirlenen bu anahtar kelime iskeletini kullanarak, şarkı sözlerini daha tutarlı hale getirmeyi amaçlamaktadır. KeYric, müziği ve sözel yapıyı üç katmanlı bir tutarlılık mekanizması (word-level, sentence-level ve full-text coherence) ile bütünleştirerek, geleneksel modellere kıyasla bağlam bütünlüğünde %5 oranında, insan değerlendirmelerine göre ise %19 oranında iyileşme sağlamıştır.

Modelin başarısı, nesnel ve öznel değerlendirme yöntemleri kullanılarak test edilmiştir. Nesnel değerlendirmelerde KeYric, Plan2Lyrics ve AI-Lyricist gibi en güncel şarkı sözü üretim modellerine kıyasla daha iyi bağlam bütünlüğü sağlamıştır. Öznel değerlendirmelerde ise, dilbilimciler, şarkı yazarları ve dil eğitmenleri tarafından yapılan değerlendirmelerde, KeYric ile üretilen şarkı sözlerinin dil öğrenimi açısından daha uygun olduğu belirtilmiştir. Özellikle, modelin müziğin türüne uygun öğeleri (örneğin perde ve nota süresi) dikkate alarak şarkı sözü üretmesi, anlam bütünlüğünü güçlendiren en önemli faktörlerden biri olarak gösterilmiştir.

Deneyisel analizler sonucunda, pop ve country müzik türlerinde üretilen şarkı sözlerinin en tutarlı yapıya sahip olduğu, buna karşın klasik müzik ve punk türlerinde bağlam bütünlüğünün daha zayıf olduğu gözlemlenmiştir. Ayrıca, müziğin perde (pitch) bilgisinin kaldırılmasının bağlam bütünlüğünü %13 oranında düşürdüğü, nota süresi (note duration) bilgisinin kaldırılmasının ise %7 oranında olumsuz etkilediği tespit edilmiştir.

Gelecek çalışmalar için, KeYric'in farklı müzik türlerinde test edilmesi, çok dilli şarkı sözü üretimi için genişletilmesi ve insan-makine iş birliği ile daha kişiselleştirilmiş şarkı sözleri üretebilecek şekilde geliştirilmesi önerilmektedir. Çalışma, şarkı sözü üretiminde bağlam bütünlüğünü artırmaya yönelik yenilikçi bir yaklaşım sunarak, hem doğal dil işleme hem de müzik yapay zekâsı alanına önemli katkılar sağlamaktadır (Ma ve ark., 2020).

Nikolov, Malmi, Northcutt ve Parisi (2020) tarafından gerçekleştirilen çalışmada, RAPFORMER adlı bir model kullanılarak koşullu rap şarkı sözü üretimi ele alınmıştır. Geleneksel rap şarkı sözü üretim sistemleri genellikle koşulsuz üretim (unconditional

generation) yöntemlerine dayanmakta ve belirli bir konuya uygun şekilde içerik üretme yeteneğinden yoksun kalmaktadır. Bu çalışmada ise, bir metin girdisine dayalı olarak rap sözleri üretmeye olanak tanıyan bir Transformer tabanlı gürültü giderici (denoising) otomatik kodlayıcı yaklaşımı geliştirilmiştir.

Çalışmada, rap sözlerinin hem içerik hem de biçim açısından korunmasını sağlamak amacıyla yeni bir yaklaşıma yer verilmiştir. Bu yöntem, orijinal rap sözlerinden anahtar içerik kelimelerini çıkararak bir modelin bu kelimelere dayanarak rap sözlerini yeniden oluşturmasını sağlamaktadır. Model, böylece hem anlam bütünlüğünü koruyarak hem de hedeflenen rap tarzını yansıtarak yeni şarkı sözleri üretebilmektedir. Ek olarak, BERT tabanlı bir kafiye geliştirme mekanizması kullanılarak, üretilen şarkı sözlerinde kafiye yoğunluğu %10 oranında artırılmıştır.

Modelin performansı üç farklı metin türü üzerinde test edilmiştir: haber özetleri, film senaryolarının özetleri ve mevcut rap şarkı sözleri. Deney sonuçlarına göre, RAPFORMER, içerik koruma ile stil dönüşümü arasında başarılı bir denge kurarak akıcı ve teknik açıdan yeterli rap sözleri üretmektedir. Yapılan Turing testi benzeri bir deneyde, RAPFORMER tarafından üretilen şarkı sözleri, rap müziği konusunda uzman kişileri %25 oranında yanıltmayı başarmıştır.

Modelin başarısını değerlendirmek için BLEU skoru, içerik örtüşme ölçütleri (content overlap) ve kafiye yoğunluğu (rhyme density) metrikleri kullanılmıştır. Sonuçlar, RAPFORMER'in içerik koruma açısından güçlü bir performans gösterdiğini, ancak özellikle haber metinlerinden rap sözleri üretirken bazı zorluklarla karşılaştığını ortaya koymuştur. Ayrıca, insan değerlendiriciler tarafından yapılan incelemelerde, üretilen rap sözlerinin hem anlam hem de dilbilgisi açısından önceki modellere göre daha üstün olduğu belirtilmiştir.

Gelecek çalışmalar için, farklı müzik türlerine uyarlanabilir rap sözleri üretmek, rap sözlerinin duygu analizi ile koşullandırılmasını sağlamak ve insan-makine iş birliği ile yaratıcı yazım süreçlerini geliştirmek önerilmektedir. Çalışma, rap müziği üretiminde doğal dil işleme ve derin öğrenme tekniklerinin etkin kullanımına yönelik önemli katkılar sağlamaktadır (Nikolov ve ark., 2020).

Zhang, Gan, Fan, Chen, Heno, Shen ve Carin (2017) tarafından gerçekleştirilen çalışmada, metin üretimi için Adversarial Feature Matching (AFM) yöntemine dayalı yeni bir GAN modeli önerilmiştir. GAN tabanlı metin üretimi, genellikle kesikli veri yapılarıyla çalışma zorlukları ve model çökmesi (mode collapse) gibi sorunlarla karşılaşmaktadır. Bu çalışmada, geleneksel GAN yaklaşımlarının zorluklarını aşmak için,

metin üretiminde yüksek boyutlu gizli öznitelik dağılımlarını eşleştiren bir metodoloji geliştirilmiştir.

Çalışmada, TextGAN adlı yeni bir model önerilmiş ve metin üretiminde öznitelik uyumunu artırmaya yönelik bir strateji geliştirilmiştir. Modelde, LSTM tabanlı bir üretici (generator) ve CNN tabanlı bir ayırt edici (discriminator) kullanılmıştır. Geleneksel GAN'ların aksine, TextGAN, gerçek ve sentetik metinler arasındaki öznitelik dağılımlarını karşılaştırarak adversarial eğitimi kolaylaştırmaktadır. Bunu başarmak için, Çoğaltma Çekirdek Hilbert Alanı (Reproducing Kernel Hilbert Space - RKHS) üzerinde kernel tabanlı bir sapma metriği (Maximum Mean Discrepancy - MMD) uygulanmıştır. Modelin başarısı, BookCorpus ve ArXiv veri setleri kullanılarak test edilmiştir. 1 milyon cümleden oluşan eğitim kümesiyle eğitilen model, 50.000 test cümlesi üzerinde değerlendirilmiştir. Performans değerlendirmesinde, BLEU skoru, Kernel Density Estimation (KDE) ve insan değerlendirmeleri kullanılmıştır. Sonuçlara göre, TextGAN, geleneksel GAN ve Sequence GAN (SeqGAN) modellerine kıyasla daha anlamlı ve akıcı metinler üretmiştir. Ayrıca, MMD tabanlı özellik eşleştirme, model çökmesi problemini hafifleterek, daha çeşitli ve doğal görünen cümleler üretmeyi mümkün kılmıştır.

TextGAN ile üretilen metinlerin kalitesi, kelime sıralaması ve gramer açısından değerlendirildiğinde, otomatik metin üretimi modellerinde anlam bütünlüğünü ve akıcılığı önemli ölçüde artırdığı gözlemlenmiştir. Ancak, modelin uzun metinlerde bazen anlamsal hatalar yaptığı ve bağlamsal tutarlılığı tam olarak koruyamadığı belirtilmiştir. Gelecek çalışmalar için, daha büyük ölçekli metin veri setleri ile modelin eğitilmesi, uzun vadeli bağlam bağımlılıklarının daha iyi ele alınması ve metin üretiminde stil transferi gibi ek özelliklerin geliştirilmesi önerilmektedir. Çalışma, GAN tabanlı metin üretiminde adversarial feature matching kullanarak önemli bir ilerleme sağlamakta ve doğal dil işleme alanındaki metin üretimi uygulamalarına katkıda bulunmaktadır (Zhang ve ark., 2020).

3. MATERYAL VE YÖNTEM

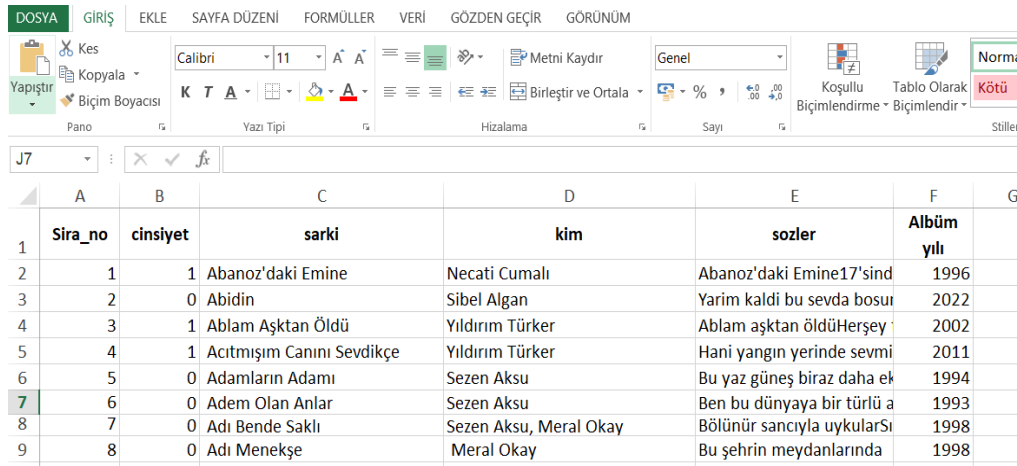
3.1. Materyal

Bu çalışmada, Türkçe dilinde yazılmış, farklı müzik türlerine ve dönemlerine ait şarkı sözlerinden oluşturulan özgün bir veri seti kullanılmıştır. Veri seti, pop, rock, rap, halk müziği ve arabesk gibi çeşitli türlerde; hem geniş kitlelerce tanınan hem de görece daha az bilinen sanatçılara ait şarkı sözlerini içermektedir. Bu türsel ve tematik çeşitlilik, modelin dilsel zenginlikleri, sözcük dağarcığı ve anlatı farklılıklarını öğrenebilmesi açısından tercih edilmiştir. Aynı zamanda, veri setinin oluşturulmasında şarkı sözlerinin yazarlarına ait cinsiyet bilgisi de dikkate alınmıştır. Kadın ve erkek söz yazarlarının üretimlerinin dengeli şekilde temsil edilmesi, modelin cinsiyete özgü anlatım biçimlerini ayırt edebilmesine ve toplumsal cinsiyet temelli dil örüntülerini öğrenebilmesine katkı sağlamaktadır. Bu doğrultuda, veri setinde her bir kayıt, ilgili yazarın cinsiyeti ile eşleştirilerek yapılandırılmış; kadın yazarlar "0", erkek yazarlar ise "1" değeriyle etiketlenmiştir. Bu yaklaşım, modelin yalnızca tek bir perspektiften değil, farklı toplumsal ve duygusal tonları içeren çok boyutlu bir dil üretimini gerçekleştirebilmesini amaçlamaktadır.

Veri setinin oluşturulması sürecinde, şarkı sözleri kamuya açık ve yasal olarak erişilebilir çevrim içi platformlardan elde edilmiştir. Başlıca veri kaynakları arasında Şarkı Sözleri .com, Genius, LyricsTranslate, Musixmatch ve YouTube gibi platformlar yer almakta olup, bu kaynaklardan alınan içeriklerin telif hakkı taşımamasına, doğruluk ve bütünlük açısından güvenilir olmasına dikkat edilmiştir. Veri toplama süreci boyunca gözetilen etik ve bilimsel ilkeler kapsamında; telif hakkı ihlali oluşturmayacak içeriklerin seçilmesi, müzik türleri ve sanatçılar arasında dengeli bir dağılım sağlanması, kadın ve erkek şarkı sözü yazarlarının eşit temsiline özen gösterilmesi ve dilin doğal akışını bozan hatalı ya da eksik içeriklerin ayıklanması temel kriterler olarak belirlenmiştir. Ayrıca, tercih edilen platformların güncel, erişilebilir ve kullanıcıya açık olması, veri güvenilirliğini destekleyen önemli unsurlar arasında değerlendirilmiştir.

Toplamda 3000 şarkı sözü içeren bu veri seti, model eğitimi için yeterli büyüklükte ve çeşitlilikte örnek sunmaktadır. Toplanan şarkı sözleri, Microsoft Excel

yazılımı kullanılarak sistematik biçimde düzenlenmiş ve her bir şarkıya ait bilgiler belirli başlıklar altında yapılandırılmıştır (Şekil 3.1). Excel dosyasında her satırda bir şarkıya ait bilgiler yer almakta olup, şu alanlar tanımlanmıştır: “sıra_no” (her bir örneğe ait sıralama bilgisi), “cinsiyet” (şarkı sözlerini yazan ya da seslendiren kişinin cinsiyeti, kategorik olarak 0 = kadın, 1 = erkek), “sarki” (şarkının adı), “kim” (şarkı sözlerinin yazarı veya seslendiren kişi), “sozler” (şarkının tüm sözleri) ve “albüm yılı” (şarkının yayımlandığı yıl). Bu yapılandırma, modelin eğitimi sırasında içeriklerin türsel, tematik ve dilsel özelliklerinin doğru biçimde analiz edilebilmesine olanak sağlamaktadır.



	A	B	C	D	E	F	G
	Sıra_no	cinsiyet	sarki	kim	sozler	Albüm yılı	
1	1	1	Abanoz'daki Emine	Necati Cumalı	Abanoz'daki Emine17'sind	1996	
2	2	0	Abidin	Sibel Algan	Yarım kaldı bu sevdâ bosu	2022	
3	3	1	Ablam Aşkta Öldü	Yıldırım Türker	Ablam aşktan öldüHerşey	2002	
4	4	1	Acıtmışım Canını Sevdikçe	Yıldırım Türker	Hani yangın yerinde sevmi	2011	
5	5	0	Adamların Adamı	Sezen Aksu	Bu yaz güneş biraz daha ek	1994	
6	6	0	Adem Olan Anlar	Sezen Aksu	Ben bu dünyaya bir türlü a	1993	
7	7	0	Adı Bende Saklı	Sezen Aksu, Meral Okay	Bölünür sancıyla uykularSı	1998	
8	8	0	Adı Menekşe	Meral Okay	Bu şehrin meydanlarında	1998	

Şekil 3. 1. Verilerin toplandığı ve düzenlendiği Excel dosyasına ait örnek bir ekran görüntüsü

3.2. Metot

3.2.1. Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs), 2014 yılında Ian Goodfellow ve arkadaşları tarafından önerilen ve derin öğrenme literatüründe çığır açan bir yapay öğrenme çerçevesidir. GAN’ler, özellikle yüksek boyutlu veri dağılımlarının modellenmesinde, yeni örnekler üretme (örneğin görsel, ses veya metin üretimi) yetenekleri ile dikkat çekmiş; doğal veri dağılımına yakın sentetik örnekler oluşturabilen güçlü generatif modeller olarak tanımlanmıştır (Iqbal ve Qureshi, 2022).

GAN yapısı temelde iki bileşenden oluşur: üretici (Generator, G) ve ayırt edici (Discriminator, D). Bu iki sinir ağı, birbirine karşıt hedefler doğrultusunda eğitilerek rekabetçi bir oyun oynarlar. Generator, verilen bir rastgele gürültü vektöründen (z) sahte fakat gerçekçi örnekler üretmeye çalışırken, Discriminator, kendisine sunulan bir örneğin

gerçek veri dağılımından mı ($z \sim p_{data}(x)$) yoksa Generator tarafından mı üretildiğini $G(z)$ ayırt etmeye çalışır (Shahriar, 2022).

Bu çerçeve, sıfır toplamlı bir oyun teorisi problemi olarak modellenir ve aşağıdaki minimax optimizasyon problemi ile ifade edilir (Zhang ve ark., 2017).

$$\min_G \max_D V(D, G) \left((E_{x \sim p_{data}(x)}[\log D(x)] + E_{z \sim p(z)}[\log(1 - G(z))]) \right) \quad (3.1)$$

Burada:

- $D(x)$, girdinin gerçek veri dağılımına ait olma olasılığını tahmin eden diskriminatif fonksiyondur.
- $G(z)$, rastgele bir gürültü vektörü üzerinden üretilen örnektir.
- $p(z)$, genellikle çok boyutlu bir Gauss veya uniform dağılımdan çekilen gürültü vektörünün dağılımıdır.

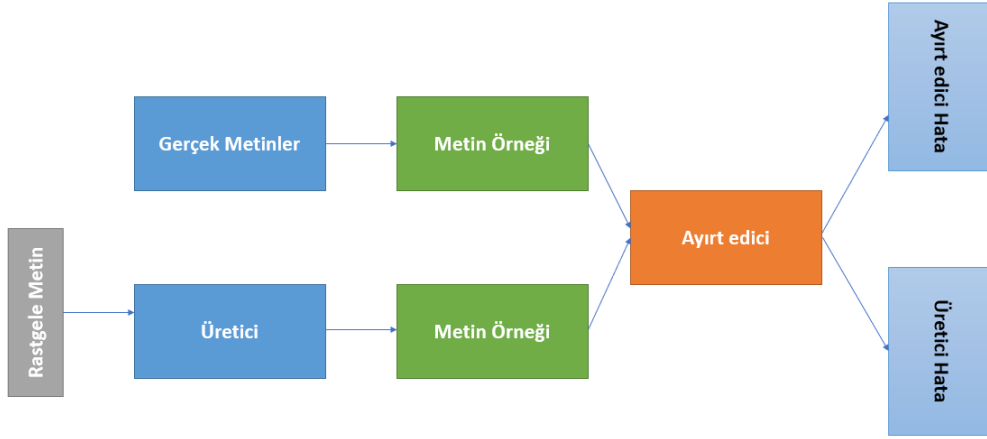
Bu yapı, iki sinir ağının aynı anda ve karşılıklı olarak eğitildiği bir oyun senaryosudur:

- Discriminator gerçek verilerde yüksek, sahte verilerde düşük olasılık vermeye çalışır.
- Generator ise, Discriminator'ı kandıracak şekilde sahte verileri gerçek gibi göstermek amacıyla eğitilir (Shahriar, 2022).

Bu sürecin sonunda ideal olarak:

$$D(x) = \frac{1}{2}, \forall x \in \{x_{gerçek}, G(z)\} \quad (3.2)$$

Genel GAN modellere ait grafiksel gösterim Şekil 3.1'de verilmiştir (Iqbal ve Qureshi, 2022).



Şekil 3. 2. GAN şeması

Her ne kadar GAN yapıları teorik olarak güçlü ve esnek bir çerçeve sunsa da, pratik uygulamalarda ciddi zorluklar barındırmaktadır. Bu zorluklar, hem ağların ortak öğrenme sürecinden kaynaklanan yapısal karmaşıklıklardan hem de generatif modellemenin doğasına özgü problemlerdendir (Xu ve ark., 2018).

İlk olarak, eğitim kararsızlığı (training instability) problemi, GAN'lerin en sık karşılaşılan ve üzerinde en çok çalışılan zorluklarından biridir. Discriminator çok hızlı ve etkili bir şekilde öğrenirse, Generator'dan gelen sahte örnekleri kolayca ayırt edebilir. Bu durumda, Generator'a geri dönen gradyan sinyalleri zayıflar veya anlamını yitirir, dolayısıyla Generator'ın kendini geliştirme kapasitesi engellenir. Bu asimetric öğrenme durumu, sistemin çökmesine veya bir tarafın baskın olduğu dengesiz bir optimizasyon sürecine yol açabilir.

İkinci olarak, mode collapse adı verilen problemde, Generator sınırlı sayıda veri modunu öğrenerek sürekli olarak benzer örnekler üretmeye başlar. Bu durum, üretilecek verilerin çeşitliliğini ciddi ölçüde sınırlar. Örneğin, metin üretimi uygulamalarında model sürekli olarak benzer cümle kalıplarını üretmeye başlar; bu da gerçekçi üretim kabiliyetini düşürür ve modelin yaratıcı kapasitesini sınırlar.

Üçüncü yaygın problem ise konverjans (convergence) sorunlarıdır. GAN'lerdeki minimax oyun doğası, klasik kayıp fonksiyonlarına kıyasla daha karmaşık bir optimizasyon süreci gerektirir. Kayıp fonksiyonları çoğu zaman düzgün bir şekilde azalmadığı gibi, osilasyonlara ve öğrenmenin bir noktada duraksamasına neden olabilir. Bu durum, özellikle hiperparametre ayarlarının hassas olduğu uygulamalarda modelin eğitilememesine neden olur.

Bu zorlukların üstesinden gelmek amacıyla, son yıllarda çok sayıda alternatif GAN mimarisi geliştirilmiştir. Örneğin, Wasserstein GAN (WGAN), klasik GAN'lerdeki kayıp fonksiyon problemini çözmek üzere Wasserstein mesafesini kullanarak daha istikrarlı ve sürekli bir öğrenme sağlar. Least Squares GAN (LSGAN), kareler ortalaması kaybı kullanarak gradyanların daha anlamlı hale gelmesini sağlar. Boundary Equilibrium GAN (BEGAN) ise Generator ve Discriminator arasındaki dengeyi adaptif bir şekilde kurarak öğrenmeyi daha dengeli hâle getirir.

Özellikle sıralı ve ayrık yapıdaki verilerle (örneğin metin) çalışırken, geleneksel GAN yapılarının yetersiz kaldığı görülmüştür. Bu nedenle SeqGAN, RankGAN, TextGAN gibi mimariler geliştirilmiştir. Bu yapılar, metin üretimi gibi ayrık örnekleme gerektiren alanlara özel olarak tasarlanmış olup, örneğin SeqGAN takviyeli öğrenme (reinforcement learning) yaklaşımıyla sıralı verilerde gradyan yayılımı sorununu aşmayı amaçlar. RankGAN ise örneklerin kalitesini sıralamaya dayalı bir şekilde değerlendirmeyi hedefleyerek daha ince ayarlı bir değerlendirme metrik sistemine dayanır.

Sonuç olarak, GAN tabanlı sistemlerin pratikte etkili bir şekilde çalışabilmesi için hem mimari düzeyde hem de öğrenme stratejileri açısından çeşitli düzenlemelere ihtiyaç duyulmaktadır. Bu nedenle çalışmalarda, kullanılacak GAN mimarisi uygulama amacına göre dikkatle seçilmeli, ilgili zorluklara karşı literatürde önerilmiş çözüm stratejileri göz önünde bulundurulmalıdır.

Şimdi bu tez çalışmasında metin üretimi için kullanılan TextGAN, LeakGAN, RankGAN, MaliGAN, SegGan ve RelGAN yöntemlerini açıklayalım.

3.2.2. TextGAN Yöntemi: Metin Üretiminde Derin Temsillerle Adversaryal Öğrenme

TextGAN, NLP alanında metin üretimi görevleri için tasarlanmış, GAN paradigmasını izleyen bir modeldir. Diğer GAN tabanlı metin üretim yöntemlerinden farklı olarak, TextGAN, cümle düzeyinde semantik temsilleri dikkate alan bir yapıyı benimseyerek hem kelime düzeyinde hem de bütünsel anlam düzeyinde üretim kalitesini artırmayı hedefler. Model, LSTM tabanlı bir Generator ve discriminator tarafından yönlendirilen bir encoder-decoder yapısı içerir ve metinlerin hem sentaktik doğruluğunu hem de semantik uyumluluğunu öğrenmeyi amaçlar (Garbacea, 2023).

TextGAN mimarisi, generatif ve diskriminatif bileşenlerden oluşan klasik GAN çerçevesini temel alır; ancak metin üretimine özgü gereksinimleri karşılamak üzere yapısal olarak özelleştirilmiştir. Modelin temelini üç ana bileşen oluşturur: Generator, Discriminator ve Latent Feature Matching mekanizması.

Generator (G_θ), rastgele bir latent vektör $z \sim p_z(z)$ kullanarak sentetik cümleler üretmek üzere yapılandırılmıştır. Bu üretici ağ, genellikle sıralı verilerle başa çıkma yeteneği yüksek olan bir LSTM (Long Short-Term Memory) temelli dil modeli olarak tasarlanır. Amaç, gerçek veri dağılımını taklit eden, gramatik açıdan tutarlı ve semantik olarak anlamlı metin örnekleri üretmektir.

Öte yandan, Discriminator (D_ϕ), kendisine sunulan bir cümlelerin gerçek veri kümesinden mi yoksa Generator tarafından mı üretildiğini ayırt etmeye çalışan bir sınıflandırıcı işlevi görür. Bu yapının implementasyonu genellikle LSTM tabanlı bir sıralı sınıflandırıcı ya da konvolüsyonel (CNN) bir metin sınıflayıcı olarak gerçekleştirilir. Discriminator, yalnızca ikili etiketleme (gerçek/sahte) görevi üstlenmekle kalmaz, aynı zamanda Generator'ın eğitimi için gerekli olan gradyan sinyallerini sağlar.

TextGAN'in mimarisini özgün kılan üçüncü bileşen ise Latent Feature Matching mekanizmasıdır. Bu yaklaşım, yalnızca binary sınıflandırma çıktısı üzerinden öğrenme yerine, Discriminator'ın ara katmanlarında elde edilen yüksek boyutlu semantik temsillerin de dikkate alınmasını sağlar. Böylece Generator, gerçek ve sahte cümleler arasındaki anlam düzeyindeki farkları da minimize etmeye çalışır. Bu yapı, klasik GAN'lerde sık karşılaşılan gradyan yayılımı problemlerini hafifletirken, üretilen metinlerin hem anlamsal bütünlüğünü hem de çeşitliliğini artırır.

TextGAN'in temel optimizasyon hedefi, hem klasik GAN kaybını hem de cümle temsilleri arasındaki semantik mesafeyi minimize etmeye yöneliktir. Kayıp fonksiyonu aşağıdaki gibi klasik GAN kaybı, aşağıdaki gibi tanımlanır.

$$\min_G \max_D V(D, G) = \left(E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p(z)} [\log(1 - G(z))] \right) \quad (3.3)$$

Ancak bu kayıp, ayrık (discrete) doğaya sahip metin verilerinde gradyan yayılımı problemlerine neden olur. Bu sorunu aşmak için TextGAN, feature matching yöntemini de kullanır. TextGAN, yalnızca discriminator'un çıktısını optimize etmek yerine, ara katman (hidden layer) çıktılarının dağılımlarını eşlemeye çalışır. Yani, gerçek ve sahte

cümlelerin discriminator tarafından elde edilen semantik temsilleri birbirine yakınlaştırılır.

$$L_{FM} = \|E_{x \sim p_{data}(x)}[f(x)] - E_{z \sim p_Z(z)}[f(G(x))]\|_2^2 \quad (3.4)$$

Burada: $f(x)$, Discriminator'un ara katmanından elde edilen özellik vektörüdür. $\| \cdot \|_2^2$ L2 normu, yani Öklidyen mesafeyi belirtir. Bu özellik eşleme kaybı, Generator'a daha kararlı ve anlamlı gradyanlar sağlar.

Modelin toplam kaybı aşağıdaki gibi tanımlanır.

$$L_{toplam} = L_{GAN} + \lambda L_{FM} \quad (3.5)$$

Burada λ , iki bileşen arasında dengeyi sağlayan hiperparametredir. Bu yapı, Generator'un yalnızca discriminator'ı kandırmasına değil, aynı zamanda yüksek kaliteli ve çeşitli metinler üretmesine de olanak tanır.

TextGAN modeli, eğitim sürecini iki aşamalı bir yaklaşım çerçevesinde gerçekleştirir: ön eğitim (pretraining) ve adversaryal eğitim (adversarial training). Bu yapı, hem üretici hem de ayırt edici bileşenlerin istikrarlı bir şekilde optimize edilebilmesine olanak tanır.

Eğitim sürecinin ilk aşaması olan ön eğitim, hem Generator hem de Discriminator için bağımsız olarak uygulanır. Bu aşamada, Generator bir dil modeli ya da autoencoder olarak, gerçek cümleleri taklit etmeyi öğrenir; böylece dil bilgisel yapı, sözdizimsel bütünlük ve temel anlamsal temsiller modelin içine gömülmüş olur. Discriminator da aynı şekilde, gerçek cümleleri doğru sınıflandırma yeteneği kazanacak şekilde eğitilir. Bu ön hazırlık, modelin adversaryal öğrenme sırasında daha anlamlı ve kararlı gradyanlar almasını sağlar.

İkinci aşama olan adversaryal eğitim sürecinde, GAN mimarisine özgü iki oyunculu oyun başlatılır. Generator, latent uzaydan aldığı rastgele vektörleri kullanarak sahte cümleler üretir ve bu cümleleri Discriminator'a sunar. Discriminator ise bu örneklerin gerçek mi yoksa sahte mi olduğunu ayırt etmeye çalışır. Bu süreçte, yalnızca ikili sınıflandırma çıktısı üzerinden öğrenme gerçekleştirilmez; aynı zamanda latent feature matching kaybı da hesaplanarak Generator'un semantik açıdan daha gerçekçi cümleler üretmesi teşvik edilir. Böylece model, sadece “gerçekmiş gibi görünen” değil, anlam olarak da tutarlı metinler üretmeye yönlendirilir.

TextGAN, klasik GAN mimarisine kıyasla daha kararlı ve kaliteli metin üretimi yapabilen özgün bir yapı sunar. Modelin temel avantajlarından biri, yalnızca binary sınıflandırma sinyali ile öğrenmeye çalışmak yerine, cümlelerin gizil temsillerini dikkate alarak daha zengin bir öğrenme mekanizması benimsemesidir. Bu yaklaşım, Generator'un daha anlamlı ve semantik olarak güçlü metinler üretmesini sağlar.

Özellikle kısa ve orta uzunluktaki cümlelerin üretiminde, TextGAN'ın başarı oranı oldukça yüksektir. Model, çeşitlilik (diversity) ve tutarlılık (coherence) arasındaki dengeyi iyi kurarak, hem birbirinden farklı hem de anlamsal olarak sağlam cümleler üretme kapasitesine sahiptir. Bu özellik, özellikle dil modeli olarak kullanılacak sistemlerde önemli bir avantaj sağlar.

Son olarak, TextGAN, metinlerin doğasında bulunan ayrık (discrete) veri yapısı nedeniyle klasik GAN modellerinde karşılaşılan gradyan yayılımı sorunlarını azaltmak üzere özel olarak tasarlanmıştır. Feature matching mekanizması sayesinde, ayrık veri alanında çalışırken bile gradyanların daha etkili bir şekilde geri yayılmasını sağlar. Bu da eğitim sürecinin istikrarını artırır ve öğrenme başarımını yükseltir.

3.2.3. LeakGAN yöntemi: Hiyerarşik yapıda bilgi sızıntısı ile geliştirilmiş adversarial metin üretimi

GAN tabanlı metin üretimi, metinlerin ayrık (discrete) yapısından kaynaklanan gradyan yayılımı problemleri nedeniyle uzun süre zorluklarla karşı karşıya kalmıştır. Bu sorunları çözmeyi amaçlayan LeakGAN, klasik GAN mimarisini, hiyerarşik bir yapay zeka ajanı modeli ve bilgi sızıntısı (information leakage) ilkesiyle geliştirerek, özellikle uzun metinlerin üretiminde önemli başarılar elde etmiştir. LeakGAN, üretilen her kelime sırasında Discriminator'dan daha fazla bilgi edinmeye çalışan bir yapıya sahiptir; bu da, sahte örneklerin kalitesini artırmak için Generator'a daha ayrıntılı rehberlik sağlar (Guo ve ark., 2018).

LeakGAN, klasik GAN yapılarından farklı olarak daha karmaşık ve çok katmanlı bir mimariye sahiptir. Bu yapının temel ayırt edici özelliği, hiyerarşik bir Generator mimarisine sahip olmasıdır. Generator, metin üretim sürecini daha etkili bir şekilde yönlendirebilmek için iki alt modülden oluşur: Manager (Yönetici) ve Worker (Çalışan). Bu iki alt modül, birlikte çalışarak daha uzun, daha anlamlı ve bağlamsal olarak tutarlı metinler üretilmesine olanak tanır.

$$\text{Generator } G = \{\text{Manager}, \text{Worker}\} \quad (3.6)$$

Generator: Her kelime üretim adımında iki aşamalı bir süreç uygular. İlk aşamada Manager, Discriminator tarafından üretim süreci boyunca sağlanan ve "sızdırılmış" olarak adlandırılan yüksek boyutlu semantik özellikleri kullanarak bir hedef vektör (goal vector) üretir. Bu hedef vektör, ilgili zaman adımında Worker modülüne rehberlik edecek olan yönlendirici bir temsildir. Yani Generator artık sadece geçmiş kelimelere değil, Discriminator'un "görüşü" aracılığıyla elde ettiği bilgiye de erişerek daha bilinçli üretim yapabilir.

Bu sürecin matematiksel ifadesiyle, ttt. adımda Manager tarafından oluşturulan hedef vektör şu şekilde tanımlanır.

$$g_t = Manager(f_t, h_{t-1}^M) \quad (3.7)$$

Burada f_t , Discriminator'dan gelen sızdırılmış semantik temsil iken, h_{t-1}^M Manager'ın bir önceki gizil (hidden) durumudur. Bu hedef vektör, bir sonraki kelimenin üretilmesinde Worker modülüne yön gösterici olur.

İkinci aşamada, Worker bu hedef vektörü kullanarak üretilecek bir sonraki kelimeye karşılık gelen olasılık dağılımını oluşturur. Worker, önceki gizil durumu h_{t-1}^W , bir önceki kelime x_{t-1} ve hedef vektör g_t bileşenlerini kullanarak bir karar fonksiyonu üretir. Bu süreç, aşağıdaki şekilde formüle edilir.

$$P(x_t | x_{1:t-1}) = softmax(W.Worker(h_{t-1}^W, x_t, g_t)) \quad (3.8)$$

Bu yapı sayesinde Generator, sadece geçmiş çıktılara değil, geleceğe dair semantik hedeflere de odaklanarak daha stratejik metin üretimi gerçekleştirebilir.

Discriminator : LeakGAN'in Discriminator bileşeni, klasik GAN yaklaşımlarından farklı bir işlevsellik sergiler. Discriminator yalnızca cümlelerin gerçek mi sahte mi olduğunu belirleyen bir sınıflayıcı değildir; aynı zamanda üretim sürecinin her adımında Generator'a yüksek boyutlu semantik geri bildirim sağlayan bir bilgi kaynağıdır. Bu geri bildirim yalnızca cümle tamamlandıktan sonra değil, her üretim adımı sırasında Generator'a "sızdırılır". Bu yaklaşım, Generator'un sadece sonuç odaklı değil, süreç odaklı optimize edilmesini sağlar ve üretimin her aşamasında daha hassas yönlendirme yapılmasına olanak tanır. Böylece, üretim süreci boyunca Discriminator'dan gelen zengin temsiller, Generator'un öğrenmesini derinleştirir ve metinlerin hem anlam bütünlüğünü hem de üretim kalitesini artırır.

3.2.4. RankGAN yöntemi: Sıralama temelli adversarial metin üretimi

Klasik GAN mimarilerinde Discriminator, Generator tarafından üretilen örneklerin gerçek mi sahte mi olduğuna dair ikili (binary) bir karar verir. Ancak bu yaklaşım, özellikle metin üretiminde ayrık (discrete) yapılarla çalışırken yetersiz kalabilmekte ve gradyan yayılımı zayıf olabilmektedir. RankGAN (Ranking-based Generative Adversarial Network), bu sınırlamaları aşmak amacıyla önerilmiş bir modeldir. RankGAN'ın temel yeniliği, ikili sınıflandırma yerine sıralama öğrenimi (learning to rank) yaklaşımını benimseyerek Discriminator'ı gerçek ve sahte örnekleri sıralama temelli değerlendirmeye zorlamasıdır. Bu sayede, Generator çok daha ince ayrıntılar üzerinden yönlendirilir ve metinlerin anlamsal kalitesi belirgin şekilde artar (Juefei ve ark., 2019).

RankGAN'ın mimarisi, diğer GAN türevlerinde olduğu gibi bir Generator (G_θ) ve bir Discriminator yerine geçen Ranker (R_ϕ) bileşenlerinden oluşur. Ancak Ranker, klasik anlamda bir “discriminator” değildir; bunun yerine, verilen bir referans örneğe göre diğer örnekleri sıralamakla (rank etmekle) görevlidir. Modelin yapısal farkı ve başarımı, bu sıralama temelli değerlendirme mekanizmasından kaynaklanır.

Generator (G_θ) : Generator , rastgele bir latent değişkenden ($z \sim p_z(z)$) metin örnekleri üretir. Bu yapı, genellikle LSTM veya GRU tabanlı bir dil modeli biçiminde yapılandırılır. Generator'un temel hedefi, Ranker tarafından yüksek sıralara yerleştirilecek nitelikte metin örnekleri üretmektir.

Ranker (R_ϕ): Ranker, bir referans örneğe göre diğer metin örneklerinin göreceli kalitesini sıralamak üzere eğitilir. Bu bağlamda, Ranker'ın görevi iki farklı örneği (biri gerçek, diğeri sahte) bir referansa göre kıyaslamak ve hangisinin referansa daha yakın olduğunu belirlemektir. Ranker'ın çıktısı, iki metin arasında bir sıralama skorudur ve aşağıdaki biçimde tanımlanabilir.

$$R_\phi(s | u) = \frac{\exp(\gamma \cdot \alpha(s, u))}{\sum_{s'} \exp(\gamma \cdot \alpha(s', u))} \quad (3.9)$$

Burada:

s : karşılaştırılan örnek (örneğin sahte metin),

u : referans örnek (örneğin gerçek metin),

$\alpha(s,u)$: s ve u arasındaki benzerliği ölçen bir fonksiyondur (genellikle cosine similarity), γ : sıcaklık (temperature) parametresi, skorların keskinliğini kontrol eder.

RankGAN, Generator ve Ranker arasında sıralama temelli bir oyun tanımlar. Generator'un hedefi, sahte metin örneklerini gerçek örneklerle benzer olacak şekilde üretmek ve Ranker'ın referans sıralamasında yukarılara yerleşmektir. Eğitim sırasında kullanılan policy gradient temelli kayıp fonksiyonu aşağıdaki gibidir:

$$\nabla_{\theta} E_{s \sim G_{\theta}}[R_{\phi}(s, u)] \approx \sum_t \nabla_{\theta} \log G_{\theta}(s_t | s_{1:t-1}) \cdot R_{\phi}(s, u) \quad (3.10)$$

Bu yapı, Reinforcement Learning ortamında, Ranker tarafından verilen sıralama skorunu bir ödül (reward) olarak kullanır. Böylece Generator, her adımda metin üretimi stratejisini bu sıralama bilgisinden öğrenerek geliştirir.

RankGAN'ın eğitim süreci, klasik GAN mimarilerinden farklı olarak sıralama temelli karşılaştırmalara dayanmaktadır. Modelin eğitimi iki ana aşamada gerçekleştirilir: ön eğitim (pretraining) ve adversaryal eğitim (adversarial training).

Ön eğitim aşamasında, Generator bir dil modeli olarak, Ranker ise sıralama yapan bir değerlendirici olarak ayrı ayrı eğitilir. Bu aşama, her iki bileşenin de yapısal öğrenmeye başlamasını ve üretim-sıralama ilişkisini temellendirmesini sağlar. Özellikle Generator, başlangıçta gramatik olarak doğru ve doğal cümleler üretmeye, Ranker ise cümlelerin göreceli kalitelerini ayırt etmeye odaklanır.

Adversaryal eğitim aşamasında ise Generator, rastgele gürültü vektörlerinden sahte metinler üretir. Bu üretilen cümleler, Ranker tarafından gerçek metinlerle birlikte değerlendirilerek sıralama skorları elde edilir. Ranker, referans bir cümleye göre hangi metnin daha benzer veya daha nitelikli olduğunu belirler. Bu sıralama skoru, Reinforcement Learning (özellikle policy gradient) ilkeleri doğrultusunda Generator'a ödül olarak geri verilir. Böylece Generator, yüksek sıralama skoru elde edebilecek nitelikte metinler üretmeye yönlendirilir ve parametre güncellemeleri bu doğrultuda gerçekleştirilir. Bu yapı, gradyanların doğrudan hesaplanamadığı ayrık veri yapılarında dahi öğrenmenin sürdürülebilmesini sağlar. RankGAN'ın en dikkat çekici yönü, klasik GAN'lerde kullanılan ikili sınıflandırma mekanizması yerine göreceli kalite ölçümüne dayalı sıralama temelli bir öğrenme yaklaşımı sunmasıdır. Bu sayede sahte metinler yalnızca “gerçek mi/sahte mi” şeklinde katı biçimde değerlendirilmez; bunun yerine, gerçek örneklerle göre ne kadar benzer veya anlamlı oldukları açısından sıralanır. Bu

nüanslı değerlendirme yöntemi, benzer yapıya sahip ama kalite açısından farklılık gösteren cümlelerin ayırt edilmesini mümkün kılarak daha hassas bir yönlendirme sağlar. Buna ek olarak, RankGAN özellikle uzun metinlerin üretiminde klasik GAN mimarilerine kıyasla daha başarılı sonuçlar verir. Discriminator yerine sıralama yapan Ranker bileşeninin kullanılması, üretim kalitesine dair daha ayrıntılı geri bildirim sağlar. Bu sayede Generator, hem dil bilgisel hem de anlamsal olarak daha rafine metinler üretme eğilimine girer.

Son olarak, RankGAN'ın Reinforcement Learning ile bütünleşik yapısı, ayrık veri üretimiyle çalışırken ortaya çıkan gradyan yayılımı sorunlarını önemli ölçüde azaltır. Bu özellik, modelin eğitilebilirliğini artırır ve farklı türdeki metin verilerine daha iyi genellenebilmesine imkân tanır. Dolayısıyla RankGAN, hem teknik hem de dilsel açıdan istikrarlı ve etkili bir metin üretim modeli olarak öne çıkmaktadır.

3.2.5. MaliGAN yöntemi: Maksimum benzerlik tabanlı doğrultulmuş adversarial metin üretimi

GAN mimarileri, özellikle görsel verilerde başarılı uygulamalara imza atmış olsa da, metin verilerindeki ayrık (discrete) doğa nedeniyle öğrenme sürecinde çeşitli zorluklarla karşılaşmaktadır. Bu zorlukların başında, gradyan yayılımının bozulması ve yüksek varyanslı ödül sinyalleri gelmektedir. Bu problemleri hedef alan MaliGAN (Maximum-Likelihood Augmented GAN), klasik GAN yapısının içine maksimum benzerlik (maximum likelihood) temelli bir iyileştirme mekanizması entegre ederek daha stabil ve verimli bir eğitim süreci sunmayı amaçlamaktadır. MaliGAN, özellikle yüksek varyansa sahip ödül sinyallerinin normalize edilmesi ve eğitimin daha iyi yönlendirilmesi yönünde yenilikler getirmiştir (Che ve ark., 2017).

MaliGAN, yapısal olarak klasik GAN bileşenlerini korur: bir Generator ($G\theta$) ve bir Discriminator ($D\phi$). Ancak esas fark, öğrenme sürecindeki ödül normalizasyonu stratejisi ve maximum likelihood ilkesi ile yönlendirilmiş gradyan güncellemeleri aracılığıyla ortaya çıkar.

Generator ($G\theta$) : Generator, rastgele gürültü vektörlerinden ya da önceki kelimelerden hareketle cümle üretimi yapar. Standart bir LSTM veya GRU tabanlı dil modeli olarak yapılandırılan bu ağ, klasik GAN'lerde olduğu gibi sahte örnekler üretmeyi hedefler. Ancak eğitim sırasında ödül sinyali olarak doğrudan $D(G(z))$ değeri

kullanılmaz. Bunun yerine, normalize edilmiş bir olasılık dağılımı üzerinden maksimum benzerlik temelli beklentisel bir gradyan hesaplanır.

Discriminator (D_ϕ): Discriminator, klasik GAN'lerde olduğu gibi bir metnin gerçek mi sahte mi olduğuna karar vermeye çalışır. Ancak MaliGAN'de Discriminator'un çıktısı doğrudan Generator'a aktarılmaz; önce normalize edilir ve log-likelihood benzeri bir öğrenme hedefiyle yeniden şekillendirilir.

MaliGAN'in temel önerisi, Generator için aşağıdaki gradyan beklentisini minimize etmektir.

$$\nabla_{\phi} E_{s \sim G_{\phi}} \left[\frac{r_D(x) - b}{\sum_{x'} r_D(x') - b} \nabla_{\phi} \log G_{\phi}(x) \right] \quad (3.11)$$

Burada:

- x : Generator tarafından üretilen metin örneğidir.
- $r_D(x)$: Discriminator'dan türetilmiş ödül skorudur.
- b : varyansı azaltmak için kullanılan bir baseline değeri (örneğin, örneklerin ortalaması).
- $\nabla_{\phi} \log G_{\phi}(x)$: maksimum benzerlik ilkesine dayanan log-olasılık türevidir.

Bu yaklaşım, GAN'lerde sık görülen yüksek varyans ve kararsız gradyan problemlerini yumuşatarak, daha kararlı bir öğrenme süreci sağlar. Ayrıca, bu formülasyon sayesinde Generator, daha etkili ve çeşitli örnekler üretme yönünde teşvik edilir.

MaliGAN'in eğitim süreci iki aşamalıdır. İlk olarak, Generator klasik bir dil modeli gibi eğitilir (pretraining), ardından adversaryal eğitim başlatılır. Adversaryal aşamada, Generator tarafından üretilen cümleler Discriminator'a sunulur ve $D(x)$ çıktılarından türetilen normalize edilmiş skorlar elde edilir. Bu skorlar, örneklerin log-olasılıklarına ağırlık verilerek gradyan hesaplamasında kullanılır. Önemli bir fark olarak, bu yaklaşım sample reweighting mekanizmasıyla her örneğin eğitime katkısını dikkatli şekilde kontrol eder.

Bu reweighting yöntemi, klasik GAN'lerdeki gibi yalnızca "iyi" örneklerin etkili olmasını değil, çeşitli örneklerin dengeli şekilde katkı sağlamasını mümkün kılar. Böylece hem öğrenme sürecinin kararlılığı korunur hem de mode collapse gibi olgulara karşı direnç artar. MaliGAN'in en belirgin avantajı, yüksek varyanslı ve zayıf sinyal

taşıyan klasik GAN ödül fonksiyonlarını normalize ederek daha kararlı ve yönlendirici bir öğrenme süreci sunmasıdır. Discriminator'dan doğrudan gelen skorlar yerine, log-olasılık bazlı normalize edilmiş ödül dağılımı kullanıldığında, modelin istikrarı önemli ölçüde artar.

Ayrıca MaliGAN, maximum likelihood prensibini GAN çerçevesiyle bütünleştirerek, klasik dil modellemenin güçlü yönlerini adversaryal öğrenme avantajlarıyla birleştirir. Bu sayede, metin üretiminde hem çeşitlilik hem de anlamsal doğruluk daha iyi sağlanabilir.

Son olarak, sample reweighting ve baseline dengelemesi gibi stratejiler, ayrık veri ortamlarında eğitimin daha verimli ilerlemesini sağlar. Bu yapı, özellikle sınırlı veriyle çalışan uygulamalarda önemli avantajlar sunar ve klasik GAN'lere kıyasla daha düşük hesaplama maliyetiyle daha etkili sonuçlar elde edilmesini mümkün kılar.

3.2.6. SeqGAN yöntemi: Ayrık sıralı veri üzerinde takviyeli öğrenme tabanlı adversarial metin üretimi

Metin verilerinin ayrık (discrete) doğası, geleneksel GAN mimarileri için ciddi bir zorluk teşkil etmektedir. Görsel verilere başarıyla uygulanan GAN yapıları, metin üretiminde gradyan yayılımı problemleri ve ayrık alanlarda optimizasyon güçlükleri nedeniyle doğrudan uygulanabilir değildir. Bu sorunu çözmeyi hedefleyen SeqGAN (Sequence GAN), metin üretimini bir ardıl karar süreci olarak ele alır ve takviyeli öğrenme (Reinforcement Learning, RL) çerçevesiyle birleştirir. Böylece, metin üretim süreci boyunca gradyanların etkili bir biçimde aktarılması sağlanır. SeqGAN, özellikle sıralı yapıdaki metin üretimi görevlerinde şiir, öykü, altyazı üretimi gibi yüksek kaliteli ve anlamlı çıktılar üretmek üzere geliştirilmiştir (Wu ve ark., 2020).

SeqGAN, temel olarak iki ana bileşenden oluşur: bir Generator (G_θ) ve bir Discriminator (D_ϕ). Ancak klasik GAN'lerin aksine, Generator burada bir ajan (agent) olarak modellenir ve çevresiyle etkileşime girerek bir politika öğrenir. Üretim süreci, her adımda bir kelime (veya token) seçimi olarak tanımlanır.

Generator (G_θ) : Generator, cümleleri kelime kelime (ya da token token) oluşturan bir sıralı modeldir. Genellikle LSTM veya GRU tabanlı bir RNN olarak yapılandırılır. Her adımda önceki kelimeleri dikkate alarak bir sonraki kelimenin olasılık dağılımını üretir. Bu yapı, takviyeli öğrenmedeki bir politika fonksiyonu gibi davranır.

Her t . adımda üretilen kelimeye ilişkin olasılık dağılımı şöyle ifade edilir.

$$P(x_t | x_{1:t-1}) = G_\theta(x_t | x_{1:t-1}) \quad (3.12)$$

Burada $x_{1:t-1}$, o ana kadar üretilmiş kelimeleri temsil eder.

Discriminator (D_ϕ): Discriminator, tamamlanmış bir cümleyi gerçek mi yoksa Generator tarafından üretilmiş mi olduğunu belirlemeye çalışan bir sınıflayıcıdır. Genellikle CNN veya RNN tabanlı bir metin sınıflandırıcısıdır. Ancak en önemli fark, Discriminator'ın yalnızca tam cümleler üzerinde çalışabilmesidir; bu nedenle, tamamlanmamış cümlelerdeki kelimeler için ara değerlendirme yapılamaz.

SeqGAN'ın temel amacı, Generator'ın bir eylem politikasını optimize etmesidir. Takviyeli öğrenme çerçevesinde, her üretilen kelime bir eylem (action), önceki kelimeler bir durum (state), cümlelerin tamamı ise bir episode olarak değerlendirilir. Bir ödül fonksiyonunu kullanır. Ödül fonksiyonu aşağıdaki gibi ifade edilir. Discriminator, tamamlanmış diziler için $D_\phi(x_{1:T})$ değerini verir. Ancak üretim süreci henüz tamamlanmadan ara adımlar için doğrudan bir ödül sinyali elde edilemez. Bu nedenle SeqGAN, ara adımlardaki durumların değerini tahmin etmek için Monte Carlo Rollout yöntemini kullanır.

$$R_t = E_{x_{t+1:T} \sim G_\theta} [D_\phi(x_{1:T})] \quad (3.13)$$

Yani, mevcut $x_{1:t}$ durumundan itibaren farklı tamamlanma örnekleri oluşturulur ve bu örneklerin Discriminator puanlarının ortalaması alınarak o adım için ödül değeri belirlenir.

RL çerçevesinde Generator'ın parametre güncellemesi REINFORCE algoritması kullanılarak gerçekleştirilir:

$$\nabla_\theta j(\theta) \approx \sum_{t=1}^T E_{x_{1:t} \sim G_\theta} [R_t \cdot \nabla_\theta \log G_\theta(x_t | x_{1:t-1})] \quad (3.14)$$

SeqGAN'ın eğitim süreci, üç temel aşamadan oluşan sistematik bir yapı izler: ön eğitim, adversaryal eğitim ve iteratif güncelleme. Bu süreç, hem metinlerin anlamlı ve dil bilgisel olarak doğru biçimde üretilmesini sağlar hem de takviyeli öğrenme ile adversaryal çerçevenin birleşmesini mümkün kılar.

İlk aşama olan pretraining (ön eğitim) sürecinde, Generator geleneksel bir dil modeli gibi ele alınarak gerçek metinler üzerinden maksimum benzerlik (maximum

likelihood estimation - MLE) ilkesine göre eğitilir. Bu eğitim, Generator'ın başlangıçta anlamsız ve yapısal olarak hatalı cümleler üretmesini engeller ve modelin temel dil yapısını öğrenmesine katkı sağlar. Aynı zamanda Discriminator da gerçek ve sahte cümleleri ayırt etmeyi öğrenmek üzere eğitilir ve bu sayede rekabetçi yapının kurulması için gereken altyapı oluşturulur.

İkinci aşama olan adversaryal eğitim sürecinde, Generator artık bir takviyeli öğrenme ajanı olarak değerlendirilir. Generator'ın ürettiği cümleler, tamamlandıktan sonra Discriminator'a sunulur ve sahte olup olmadıklarına dair skorlar alınır. Ancak bu süreçte yalnızca tamamlanmış cümleler değil, üretim sırasında oluşan tamamlanmamış diziler için de ödül tahmini yapılması gereklidir. Bu amaçla, Monte Carlo Rollout tekniği kullanılarak mevcut kısmi dizilerden çoklu tamamlama örnekleri oluşturulur ve bu örneklerin Discriminator skorlarının ortalaması alınarak ara adımlar için dolaylı ödül sinyalleri üretilir. Bu ödüller, REINFORCE algoritması ile Generator'ın parametrelerinin güncellenmesinde kullanılır.

Son aşama olan iteratif güncelleme bölümünde, Generator ve Discriminator sırasıyla güncellenir. Her biri diğerini iyileştirmeye çalışarak rekabetçi bir döngü içinde öğrenmeyi sürdürür. Discriminator daha iyi ayırt ettikçe, Generator daha gerçekçi metinler üretmeyi öğrenir. Bu karşılıklı öğrenme döngüsü, modelin üretim kalitesini adım adım artırır ve metinlerin dil bilgisel tutarlılığını derinleştirir. SeqGAN'ın en temel avantajı, ayrık ve sıralı veri yapılarında gradyan yayılımı problemini çözmek üzere takviyeli öğrenmeyi GAN çerçevesiyle birleştirmesidir. Bu bütünleşik yaklaşım sayesinde, kelime üretimi gibi aşamalı süreçler boyunca adım adım geri bildirim alınabilir ve öğrenme yalnızca son çıktıya bağlı kalmaksızın şekillendirilebilir. Bu, klasik GAN'lerin doğrudan uygulanamadığı metin üretimi problemleri için önemli bir çözüm sunar.

Özellikle Monte Carlo Rollout stratejisi, üretim sürecindeki ara adımlarda ödül sinyallerinin tahmin edilmesini mümkün kılar. Böylece model yalnızca nihai cümle tamamlandığında değil, her kelime üretimi sırasında da dolaylı olarak değerlendirilir. Bu durum, hem öğrenmenin yönlendirilebilirliğini artırır hem de erken aşamalarda meydana gelebilecek hatalı örüntülerin oluşmasını engeller.

Sonuç olarak, SeqGAN yalnızca teknik olarak uygulanabilirliği yüksek bir yöntem olmakla kalmaz; aynı zamanda diyalog üretimi, şiir yazımı, otomatik hikâye oluşturma ve metin tamamlama gibi alanlarda da pratik olarak oldukça başarılı sonuçlar üretir. Dil bilgisel doğruluk, semantik uyum ve üretim çeşitliliği bakımından tatmin edici

çıktılar sağlayarak, metin üretiminde adversaryal yaklaşımların gücünü sıralı veri ortamlarında da ortaya koyar.

3.2.7. RelGAN yöntemi: İlişkisel özelliklerle zenginleştirilmiş adversarial metin üretimi

Klasik GAN tabanlı metin üretiminde karşılaşılan en önemli sorunlardan biri, mode collapse (üretim çeşitliliğinin kaybı) ve gradyan yayılımı sorunlarıdır. Bu tür problemler, Generator'un benzer ya da tekrarlayıcı örnekler üretmesine neden olarak modelin üretim kabiliyetini sınırlar. Bu sorunlara çözüm olarak geliştirilen RelGAN (Relational Generative Adversarial Network), kelime üretiminde ilişkisel dikkat mekanizması (relational attention) kullanarak daha çeşitli, bağlamsal olarak daha tutarlı ve semantik olarak zengin metinler üretmeyi amaçlar. RelGAN, hem mimari hem de öğrenme stratejileri açısından bir dizi iyileştirme sunar: ilişkisel dikkat katmanları, çoklu discriminator yapısı ve takviyeli öğrenmeye dayalı eğitim çerçevesiyle klasik GAN'lere göre daha üstün performans gösterir (Ni et al., 2018).

RelGAN, iki ana bileşenden oluşur: bir Generator (G_θ) ve bir Discriminator (D_ϕ). Ancak bu bileşenler, klasik GAN'lerden farklı olarak ilişkisel yapılar ve çoklu değerlendirme başlıkları içerir. Bu yapı, özellikle uzun metin üretimi ve içerik çeşitliliği açısından dikkat çeker.

Generator (G_θ) : RelGAN'in Generator bileşeni, geleneksel sıralı dil modellerine benzer şekilde çalışır ancak üretim sürecinde ilişkisel dikkat (relational self-attention) mekanizmasıyla güçlendirilmiştir. Her adımda üretilen kelimenin, daha önce üretilmiş kelimelerle olan bağlamsal ilişkileri dikkate alınır ve bu bağlam üzerinden bir sonraki kelime seçilir. RelGAN Generator'ı, kelime üretimini yalnızca kelime gömme (embedding) ve LSTM katmanlarıyla değil, aynı zamanda attention-based relational memory mekanizması ile yürütür. Bu yapı, kelimeler arasındaki ilişkisel bağları açıkça modelleyerek hem daha tutarlı hem de daha çeşitli metinler üretir.

Discriminator (D_ϕ) : Discriminator bileşeni, klasik GAN'lerde olduğu gibi sahte ve gerçek metinleri ayırt etmeye çalışır. Ancak RelGAN'de çoklu başlık (multi-head) değerlendirme prensibi uygulanır. Yani, birden fazla “görüş açısı” ile metinler değerlendirilir. Bu çoklu değerlendirme başlıkları, farklı semantik boyutlarda metinleri analiz ederek daha güvenilir ödül sinyalleri üretir. Discriminator çıktıları daha sonra birleştirilerek Generator'a daha dengeli ve bilgilendirici gradyanlar olarak geri gönderilir.

RelGAN'in Generator bileşeni, kelime üretimini bir politika (π_θ) olarak ele alır. Takviyeli öğrenme bağlamında, her kelime bir eylem (action), cümledeki konum bir durum (state) olarak değerlendirilir. Amaç, cümle tamamlandığında en yüksek ödül sağlayacak eylemler dizisini öğrenmektir. Discriminator tarafından hesaplanan ödül değeri, her tamamlanmış örnek için şöyle ifade edilir.

$$R(x_{1:T}) = \frac{1}{K} \sum_{k=1}^K D_k(x_{1:T}) \quad (3.15)$$

Burada:

K: Discriminator'da kullanılan değerlendirme başlıklarının sayısıdır.

D_k : k. başlığa ait değerlendirme sonucu.

$x_{1:T}$:T: tam bir cümle.

Generator parametreleri, takviyeli öğrenme algoritmaları (örneğin REINFORCE) ile aşağıdaki gibi güncellenir:

$$\nabla_{\theta} J(\theta) = E_{x_{1:t} \sim G_{\theta}} [R(x_{1:T}) \cdot \sum_{t=1}^T \nabla_{\theta} \log G_{\theta}(x_t | x_{1:t-1})] \quad (3.16)$$

RelGAN'in eğitim süreci, üç aşamada gerçekleştirilir. İlk olarak, pretraining (ön eğitim) aşamasında Generator, gerçek metin verileri üzerinden maksimum benzerlik (MLE) ilkesine göre eğitilir. Bu aşama, modelin temel dil yapısını öğrenmesini ve başlangıçta anlamlı cümleler üretmesini sağlar. Discriminator da bu süreçte gerçek ve sahte cümleleri ayırt etme konusunda eğitilir. İkinci olarak, adversaryal eğitim sürecinde Generator, takviyeli öğrenme ajanı olarak değerlendirilir. Ürettiği cümleler Discriminator'a sunulur ve her üretim adımı için gradyanlar, REINFORCE algoritması ile hesaplanır. İlişkisel dikkat katmanları sayesinde, model yalnızca dil bilgisel değil, aynı zamanda anlamsal açıdan da güçlü cümleler üretme eğilimindedir. Son aşamada, multi-head discriminator sayesinde elde edilen ödül sinyalleri, örneklerin farklı semantik boyutlardaki değerlendirmelerini içerdiğinden, Generator'a zenginleştirilmiş ve daha yönlendirici geri bildirimler sunulur. Bu süreç, metin üretiminde hem çeşitliliği artırır hem de tekrarlayan yapıları (mode collapse) engeller.

RelGAN'in en güçlü yönlerinden biri, kelimeler arasındaki bağlamsal ilişkileri dikkat mekanizmasıyla modellemesidir. Bu yaklaşım, sıradan sıralı modellerin kaçırabileceği semantik ilişkileri yakalayarak daha tutarlı cümleler üretmeyi mümkün kılar. Ayrıca, çoklu değerlendirme başlıkları sayesinde Discriminator'dan gelen sinyaller yalnızca "gerçek/sahte" ayrımı değil, anlamsal zenginlik, bağlamsal uyum ve gramatik doğruluk gibi farklı boyutlarda geri bildirim sağlar. Bu durum, Generator'un daha esnek ve güçlü metinler üretmesine olanak tanır. Son olarak, RelGAN, mode collapse sorununu önemli ölçüde azaltarak farklı türde cümlelerin üretimini teşvik eder. Bu yönüyle, metin üretiminde hem çeşitlilik hem de kalite açısından önceki GAN mimarilerine göre daha başarılı sonuçlar ortaya koyar.

3.2.8. Performans metrikleri

GAN tabanlı metin üretiminde, yalnızca görsel ya da sezgisel değerlendirme yöntemleri kullanmak, üretilen metinlerin gramatik doğruluğu, anlamsal bütünlüğü, çeşitliliği ve referans metne benzerliği gibi çok boyutlu kalite kriterlerini nesnel olarak değerlendirmek açısından yetersiz kalmaktadır. Bu nedenle, metin üretim modellerinin başarımını ölçmek amacıyla çeşitli sayısal performans metriklerinin kullanılması zorunludur.

Bu metrikler sayesinde, hem üretilen metnin referans metne ne kadar örtüştüğü (ör. BLEU, Jaccard, 2-gram Coverage gibi n-gram tabanlı benzerlik metrikleri) hem de modelin içerik üretiminde gösterdiği dilsel çeşitlilik (ör. TTR, Unique Word Ratio), anlamsal yakınlık (Cosine Similarity) ve dil modeli başarımı (Perplexity) gibi farklı yönlerden performansı objektif biçimde ölçülebilmektedir. Bu metriklerin çok yönlü olarak kullanımı, yalnızca çıktının doğruluğunu değil, aynı zamanda dilsel zenginlik, ifade gücü ve anlam derinliği gibi daha incelikli dil özelliklerinin de değerlendirilmesini mümkün kılar.

Dolayısıyla, çalışmada hem geleneksel hem de içerik-odaklı metin değerlendirme metriklerine yer verilmiş, model performansı çok boyutlu olarak analiz edilmiştir. Böylece, yalnızca teknik doğruluk değil, aynı zamanda insana yakınlık açısından dilsel kalite de sistematik bir şekilde ölçülmüştür.

3.2.8.1. BLEU (Bilingual Evaluation Understudy Score)

Üretilen cümlelerin referans (gerçek) cümleye ne kadar benzediğini ölçer. Genellikle makine çevirisi ve metin üretiminde kullanılır (Chollampatt ve ark., 2025).

$$BLEU = BP \cdot \exp\left(\sum_{n=1}^N w_n \cdot \log p_n\right) \quad (3.17)$$

Burada;

p_n : n-gram doğruluk oranı (örneğin, 1-gram, 2-gram, ...)

w_n : her n-gram için ağırlık (genellikle eşit dağıtılır: $w_n = \frac{1}{N}$)

BP: Brevity Penalty (kısalık cezası), uzunluk farkını telafi eder.

$$BP = \begin{cases} 1 & \text{eğer } c > r \\ \exp\left(1 - \frac{r}{c}\right) & \text{eğer } c \leq r \end{cases} \quad (3.18)$$

c: üretilen cümle uzunluğu; r: referans cümle uzunluğunu belirtir.

BLEU, referans metinle eşleşen n-gram'ların logaritmik ortalamasını alır ve cümlelerin fazla kısa üretilmesini cezalandırır. 0 ile 1 arasında değer alır (genellikle yüzde olarak 0–100).

3.2.8.2. Jaccard similarity

Üretilen ve referans metinlerin kelime kümeleri arasındaki benzerliği ölçer (Bag ve ark., 2019).

$$\text{Jaccard}(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3.19)$$

A: üretilen cümledeki kelimeler kümesi; B: referans cümledeki kelimeler kümesini belirtir.

1'e yakınsa yüksek benzerlik vardır. Yalnızca kelime düzeyinde çalışır, sözdizimsel sıralamayı dikkate almaz.

3.2.8.3. Perplexity (PP)

Bir dil modelinin bir metni ne kadar iyi öngörebildiğini ölçer. Düşük perplexity değeri, modelin metni daha iyi tahmin ettiğini gösterir. Özellikle RNN, Transformer gibi modellerin çıktılarında sıklıkla kullanılır (Vychegzhanin ve ark., 2023).

$$\text{Perplexity} = \exp \left(-\frac{1}{N} \sum_{i=1}^N \log P(w_i | w_{1:i-1}) \right) \quad (3.20)$$

N: toplam kelime sayısı; $P(w_i | w_{1:i-1})$: her kelimenin bağlamsal olasılığını (dil modeli çıktısı) belirtir.

3.2.8.4. TTR (Type-Token Ratio)

Metindeki sözcük çeşitliliğini ölçer. 0 ile 1 arasında değer alır. TTR değeri yüksekse, metin daha çeşitli kelimeler içermektedir. Kısa metinlerde doğal olarak daha yüksektir, bu nedenle sabit uzunlukta değerlendirme önerilir (Matlach ve ark., 2021).

$$TTR = \frac{\text{Type (benzersiz kelime sayısı)}}{\text{Type (toplam kelime sayısı)}} \quad (3.21)$$

3.2.8.5. Unique word ratio

Üretilen metindeki benzersiz kelimelerin oranını ölçer. TTR ile benzerdir fakat daha geniş bir metin topluluğu (örneğin tüm üretilen metinler) üzerinden hesaplandığında daha genel çeşitlilik ölçümü sağlar (Imperial ve Madabushi, 2022).

$$\text{Unique Word Ratio} = \frac{\text{Type (benzersiz kelime sayısı)}}{\text{Type (toplam kelime sayısı)}} \quad (3.22)$$

3.2.8.6. 2-gram coverage

Üretilen metindeki 2-gram'ların referans metindeki 2-gram'larla ne kadar örtüştüğünü ölçer. Üretilen metindeki ardışık kelime çiftlerinin ne kadarının referansla örtüştüğünü belirler. BLEU'nun özel bir durumudur; sözdizimsel yapıyı yansıtır (Novais ve ark., 2010).

$$2 - \text{gram Coverage} = \frac{|\text{Matched 2-grams}|}{|\text{Total 2-grams in reference}|} \quad (3.23)$$

3.2.8.7. Kosinüs benzerliği (TF-IDF tabanlı)

Üretilen ve referans metinlerin TF-IDF temsilleri arasındaki açıyı ölçerek metin benzerliğini değerlendirir. 1'e yaklaştıkça metinler daha benzerdir. TF-IDF, sözcüklerin önemini (frekans ve ayırım gücü) hesaba kattığı için içerik bazlı benzerlik ölçümünde oldukça etkilidir (Zhu ve ark., 2018).

$$\text{Cosine}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} \quad (3.24)$$

A, B: metinlerin TF-IDF vektörleri; $A \cdot B$: vektörlerin noktasal çarpımı (dot product) ve $\|A\|$: A vektörünün normunu belirtir.

4. BULGULAR VE TARTIŞMA

4.1. Genel Bulgular

Müzik, tarih boyunca insanlığın duygularını ifade etme, toplumsal olaylara tanıklık etme ve kültürel kimlikleri aktarma araçlarından biri olmuştur. Şarkılar, yalnızca melodi ve ritimden ibaret değildir; içerdiği sözlerle duygusal bir derinlik, tematik bir yapı ve anlatsal bir bütünlük sunar. Şarkı sözleri, kimi zaman bireyin iç dünyasını yansıtırken, kimi zaman da kolektif bilincin bir yansıması hâline gelir. Bu bağlamda şarkı sözü yazımı, hem sanatsal hem de kültürel bir üretim sürecidir. Geleneksel anlamda şarkı sözleri, bireysel yaratıcılıkla, ilhamla ve sezgiyle üretilirken; günümüzde yapay zekâ teknolojilerinin gelişimiyle birlikte bu süreç yeni bir döneme girmiştir.

Yapay zekâ, özellikle NLP ve derin öğrenme alanlarındaki ilerlemeler sayesinde yaratıcı yazarlık süreçlerine entegre edilmeye başlanmıştır. Bu dönüşümde GAN önemli bir rol oynamaktadır. GAN'lar, üretici (generator) ve ayırt edici (discriminator) olmak üzere iki sinir ağının karşılıklı etkileşimiyle öğrenen bir yapı sunar. Bu yapı, metin üretiminde hem gerçekliğe yakın hem de anlamlı, bağlamsal olarak tutarlı içerikler üretme potansiyeline sahiptir. GAN mimarileri, başta görsel sanatlar olmak üzere son yıllarda müzik ve edebi üretim alanlarında da yaygınlık kazanmış, özellikle özgünlük, stil tutarlılığı ve içerik çeşitliliği açısından umut vaat eden sonuçlar ortaya koymuştur.

Bu tez çalışmasının amacı, GAN modellerinin şarkı sözü üretimi özelindeki performansını ortaya koymak; hem dilsel tutarlılık hem de duygusal ve sanatsal derinlik bakımından elde edilen sonuçları değerlendirmektir. Bu kapsamda özellikle SeqGAN, RankGAN, RelGAN gibi metin üretimine odaklanan GAN mimarileri incelenmiş; model çıktılarının hem nesnel hem de öznel açılardan analiz edilmesi hedeflenmiştir. Bu analiz yalnızca teknik bir başarı ölçümü değil, aynı zamanda üretilen sözlerin sanatsal yeterliliği ve duygusal etkisi üzerine bir değerlendirme niteliği taşımaktadır.

Bulguların değerlendirilmesinde birden fazla performans ölçütü kullanılmıştır. Her biri, metin üretiminin farklı yönlerini değerlendirmek üzere seçilmiştir:

Modelin performansını değerlendirmek amacıyla bu çalışmada hem dilsel doğruluk hem de çeşitlilik odaklı çeşitli metrikler kullanılmıştır. BLEU (Bilingual Evaluation Understudy) skoru, üretilen şarkı sözlerinin referans sözlerle olan yüzeysel

benzerliğini ölçmekte kritik bir rol üstlenmiştir. BLEU, özellikle n-gram düzeyinde örtüşmeleri dikkate alarak, modelin dilbilgisel olarak doğru ve anlamlı yapılar üretip üretmediğini değerlendirir. Bu metrik sayesinde, modelin öğrenme süreci boyunca gerçek veriye ne derece yakın çıktılar ürettiği gözlemlenmiş, böylece dilsel kalite ölçülmüştür. Özellikle tekrarlayan yapılar, gramatikal doğruluk ve metinsel akış gibi unsurlar, BLEU puanı aracılığıyla yakından takip edilmiştir.

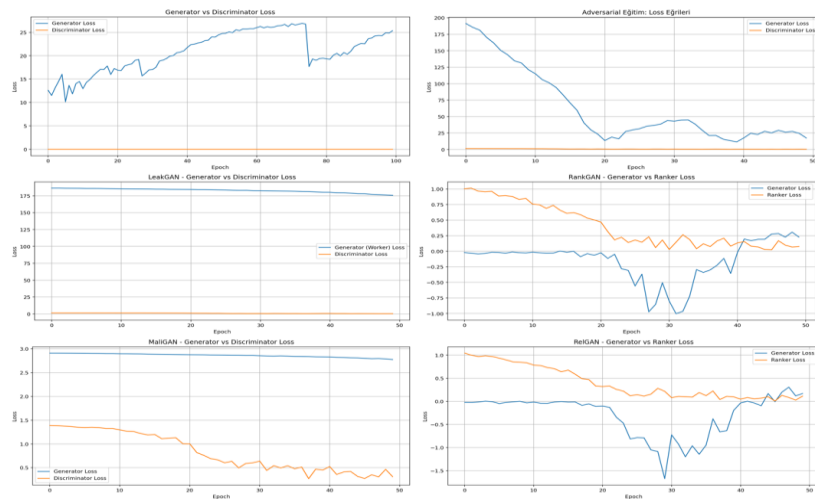
Bunun yanında, Self-BLEU metriği, modelin çeşitlilik kabiliyetini analiz etmek için önemli bir araç olarak kullanılmıştır. Self-BLEU, aynı modelin farklı çıktılarının birbirine benzerliğini ölçerek, üretimdeki tekrar eğilimini ortaya koyar. Düşük Self-BLEU değerleri, modelin monoton yapılar üretmek yerine daha yaratıcı ve çeşitli dizilimler oluşturduğunu gösterir. Bu sayede, modelin yalnızca eğitildiği veri kümesinin kalıplarını taklit etmediği, aynı zamanda yeni varyasyonlar sunabildiği anlaşılır. Bu çeşitlilik, özellikle şarkı sözü gibi sanatsal metinlerde, duygusal zenginlik ve tematik farklılık yaratabilmek açısından oldukça değerlidir. Ayrıca, Perplexity (şaşırtıcılık) metriği, modelin bir sonraki kelimeyi tahmin etme kabiliyeti üzerinden dilsel yetkinliğini ölçmek için kullanılmıştır. Düşük perplexity değeri, modelin dilin istatistiksel yapısını iyi kavradığını, anlamlı ve doğru cümleler kurabildiğini gösterir. Bu bağlamda, şarkı sözlerinin duygusal ve yapısal bütünlüğü, bu metrik ile nesnel olarak değerlendirilmiştir. Cosine Similarity ise, üretilmiş ve gerçek şarkı sözlerinin vektör temsilleri arasındaki anlamsal benzerliği ölçmek için tercih edilmiştir. Bu sayede modelin yalnızca biçimsel olarak değil, anlamsal düzeyde de referans sözlerle ne kadar uyumlu içerikler sunduğu analiz edilmiştir. Son olarak, n-gram kelime frekansı karşılaştırması, modelin sıklıkla tercih ettiği kelimelerin gerçek şarkı sözlerindeki kullanım sıklığıyla örtüşüp örtüşmediğini inceleyerek, dilsel stil ve içerik benzerliğini daha somut biçimde ortaya koymuştur. Tüm bu metrikler bir araya geldiğinde, modelin hem dilsel doğruluk hem de sanatsal çeşitlilik açısından performansı çok boyutlu olarak değerlendirilmiş ve optimize edilmiştir.

Tüm bu metrikler hem model çıktılarının nicel olarak karşılaştırılmasını sağlamış, hem de geleneksel şarkı sözü yazımıyla yapay zekâ üretimi arasındaki farkları nesnel bir zeminde analiz etmeyi mümkün kılmıştır. Aynı zamanda, grafiksel analizlerle gerçek ve sentetik sözlerin yapısal karşılaştırmaları yapılmış; özellikle en sık kullanılan kelimeler ve sözdizimi örüntüleri incelenmiştir.

Bu bölümde, elde edilen bulgular üzerinden GAN tabanlı metin üretiminde hangi mimarilerin daha başarılı olduğu, hangi tematik veya stilistik zorluklarda başarısızlıklar

yaşandığı ve model çıktılarının sanatsal geçerliliği detaylı olarak tartışılacaktır. Bu bağlamda, yapay zekâ ile üretilen şarkı sözlerinin yaratıcı endüstrilerdeki potansiyel kullanımları da göz önüne alınarak, hem akademik hem de sektörel bir perspektifle değerlendirilecektir.

Loss (kayıp) grafiklerinin incelenmesi, derin öğrenme modellerinde özellikle GAN (Generative Adversarial Networks) tabanlı sistemlerde modelin öğrenme sürecini değerlendirmek için kritik bir araçtır. GAN’larda iki temel bileşen olan üreteç (generator) ve ayırtıcı (discriminator veya ranker), birbirine karşılıklı olarak öğrenir. Bu süreçte her bileşenin ne kadar iyi öğrendiğini ve denge sağlayıp sağlayamadıklarını belirlemek için her epoch sonunda hesaplanan kayıp değerleri görselleştirilir. Generator kaybı, üretilen örneklerin ne kadar gerçekçi olduğunu; discriminator veya ranker kaybı ise bu örneklerin gerçek veriden ne derece ayırt edilebildiğini gösterir. Amaç, iki tarafın da birbirine yakın performans sergilemesi, böylece modelin ne aşırı öğrenmiş (overfit) ne de zayıf kalmış (underfit) olmasıdır. Aşağıda TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN ve RelGAN modellerine ait adversarial öğrenim süreçlerinde kaydedilen loss (kayıp) eğrileri karşılaştırmalı olarak verilmiştir. Bu tür GAN tabanlı metin üretim modellerinde hem *generator loss* hem de *discriminator veya ranker loss* eğrileri modelin dengeleme başarısını, öğrenme stabilitesini ve üretim kalitesini değerlendirmek için kullanılır. Her bir modelin performansı, loss eğrisinin seyrine, düşüş eğilimi olup olmamasına, kararlılığa ve öğrenme sırasındaki salınımlara göre analiz edilmiştir.



Şekil 4. 1. Loss(Kayıp) performans grafikleri

Şekil 4.1’te verildiği gibi Adversarial (çekişmeli) öğrenme yaklaşımına dayanan metin üretim modellerinde, eğitim sürecinin başarısını değerlendirmek için kullanılan en

önemli araçlardan biri loss (kayıp) grafikleridir. GAN tabanlı yapılarda, iki ana bileşen olan Generator (üreteç) ve Discriminator ya da Ranker (ayırt edici ya da sıralayıcı), birbirine karşılıklı olarak eğitilir. Bu süreçte her bileşenin ne kadar iyi öğrendiğini gözlemlemek için epoch başına düşen kayıp değerleri izlenir. Generator kaybı, modelin ne kadar gerçekçi örnekler ürettiğini, Discriminator veya Ranker kaybı ise bu örneklerin gerçek veriden ayırt edilme derecesini ifade eder. İdeal durumda bu iki bileşen dengeli bir şekilde ilerlemeli, biri diğerine karşı çok baskın hale gelmemelidir. Aksi halde model ya aşırı güvenli (aşırı öğrenmiş) ya da yeterince üretken olmayan bir yapıya dönüşebilir. Özellikle metin üretimi gibi zorlu görevlerde, bu dengeyi sağlamak oldukça kritiktir. Ayrıca, bu grafiklerin yanında çeşitli metrikler (BLEU, Cosine benzerliği, ROUGE, Distinct, Perplexity vb.) de modelin üretim kalitesini nesnel olarak ölçmekte kullanılır; ancak öğrenme sürecinin dinamiğini anlamada loss eğrileri görsel ve sezgisel olarak temel referans noktasıdır.

İlk grafikte TextGAN modeli gözlemlenmektedir. Burada generator kaybı sürekli dalgalanmakta ve giderek artan bir eğilim göstermektedir. Buna karşın discriminator kaybı sifıra yakın düzeyde sabitlenmiştir. Bu durum discriminator'ın aşırı güçlü hale geldiğini ve generator'ın anlamlı geri bildirim alamadığını gösterir. SeqGAN modelinde ise generator kaybı çok yüksek başlamakta, hızlıca azalarak daha kararlı bir seviyeye inmektedir. Discriminator sabit kaldığı için modelin kontrollü bir şekilde öğrenmeye devam ettiği anlaşılmaktadır. LeakGAN'da generator (worker) kaybı yüksek düzeyde başlamakta ve oldukça yavaş bir düşüş göstermektedir; bu da öğrenmenin etkili olmadığını ve modelin generator-discriminator etkileşiminden yeterince faydalanamadığını gösterir. MaliGAN'da da benzer şekilde generator kaybı oldukça yüksek başlamakta ve azalma oranı oldukça düşüktür; discriminator ise hızla öğrenmekte ve baskın hale gelmektedir. Bu da üretim kalitesini sınırlayan bir öğrenme sürecini yansıtır. Bu modellere kıyasla RankGAN ve RelGAN modelleri daha dengeli grafikler sergilemektedir. RankGAN'da ranker kaybı istikrarlı şekilde düşerken, generator kaybı negatif seviyelere inip toparlanarak dalgalı bir dengeye ulaşmaktadır. Bu yapı, modelin hem gerçeklik hem çeşitlilik açısından üretkenliğini desteklediğini gösterir. RelGAN'da ise generator ve ranker kayıpları birbirini izleyerek zaman içinde uyumlu bir şekilde dengelenmektedir. Bu model, adversarial yapının en sağlıklı işlediği ve öğrenmenin istikrarlı sürdüğü örneklerden biridir.

Tüm bu değerlendirmeler ışığında, adversarial metin üretiminde en istikrarlı ve dengeli performansı sergileyen modelin RelGAN olduğu görülmektedir. Hem generator

hem ranker bileşenleri arasında kurulan dinamik denge, bu modelin hem çeşitli hem anlamlı metinler üretmesini sağlamaktadır. RankGAN da benzer şekilde yüksek performans göstermekte ve yapay örneklerin gerçek veriye yakınlığını başarılı şekilde optimize etmektedir. Buna karşın TextGAN, LeakGAN ve MaliGAN gibi modellerde discriminator'ın çok erken güç kazanması nedeniyle generator etkisiz kalmakta, bu da metinlerin çeşitliliği ve doğallığı açısından sınırlayıcı bir faktör oluşturmaktadır. Dolayısıyla, özellikle sıralayıcı tabanlı (Ranker) yapıların kullanıldığı GAN modelleri, metin üretimi gibi incelikli ve hassas alanlarda daha etkili sonuçlar sunmaktadır.

Aşağıdaki Tablo 4.1'de, farklı metin üretim GAN modellerinin şarkı sözü üretimi üzerindeki dilsel doğruluk, çeşitlilik, anlamsal benzerlik ve yapısal uygunluk gibi çeşitli ölçütlere göre karşılaştırmalı performansı verilmiştir. Her bir satırda bir model yer almakta ve sütunlar sırasıyla kullanılan değerlendirme metriklerine karşılık gelen sayısal değerleri sunmaktadır. Bu tablo, özellikle GAN tabanlı metin üretim modellerinin gerçek şarkı sözlerine olan yakınlığı, içerik çeşitliliği ve anlamsal kalitesi gibi boyutları çok yönlü biçimde analiz etmeye olanak tanır. Model performansını değerlendirmek için kullanılan metrikler, hem dilsel doğruluk hem de çeşitlilik açısından önemli içgörüler sunmaktadır. BLEU ve Jaccard skorları, üretilen dizelerin referans yani gerçek şarkı sözleriyle olan yüzeysel benzerliğini ve ortaklık oranını ölçmektedir. Tablodaki tüm modellerin bu iki metrikte sıfıra yakın değerler alması, model çıktılarının biçimsel olarak referans metinlerle neredeyse hiç örtüşmediğini göstermektedir. Bu durum, modellerin dilin yapısal kalıplarını yeterince öğrenemediğini ve gerçek metinlere benzer çıktılar üretemediğini ortaya koyar.

Bir diğer önemli metrik olan Perplexity, modelin bir sonraki kelimeyi tahmin etme başarısını gösteren bir ölçüttür. Düşük perplexity değerleri, modelin dil yapısını daha iyi öğrendiğini ve gramatikal olarak geçerli, tutarlı cümleler üretebildiğini ifade eder. Bu bağlamda, MaliGAN (9588.23) ve RankGAN (10945.70) modelleri, diğer modellere kıyasla daha iyi dilsel doğruluk sergilemiş ve daha akıcı diziler üretme potansiyeli göstermiştir. Ancak bu başarıya rağmen metinlerin içeriği hâlâ düşük düzeyde kalmaktadır. Ayrıca TTR (Type-Token Ratio) ve Unique Word Ratio değerlerinin tüm modellerde 1.0 çıkması, her kelimenin yalnızca bir kez kullanıldığını ve modelin çıktılarında aşırı çeşitlilik (over-diversity) bulunduğunu gösterir. Bu durum, tutarlılığı azaltarak metinlerin anlam bütünlüğünü zayıflatabilir.

Yapısal örüntülerin incelendiği 2-gram Coverage metriğinde ise tüm modellerin 0.0 değeri alması, üretilen metinlerin gerçek şarkı sözlerinde sıkça rastlanan ikili kelime

yapılarından yoksun olduğunu göstermektedir. Bu, modelin sözdizimsel kalıpları öğrenememesiyle ilişkilidir. Cosine Similarity (TF-IDF) metriği, üretilen ve gerçek metinlerin semantik benzerliğini ölçer; bu metrikte yalnızca RelGAN anlamlı bir fark yaratmıştır (0.0693). Bu sonuç, RelGAN'ın biçimsel olarak yetersiz olsa da anlamsal düzeyde daha tutarlı çıktılar sunduğunu göstermektedir. Son olarak, Word Distribution değeri tüm modellerde yalnızca 0.0002 seviyesinde kalmıştır. Bu düşük oran, üretilen kelimelerin dağılımının gerçek şarkı sözlerindeki dağılımla neredeyse hiç örtüşmediğini ve sözcük kullanımında ciddi bir sapma olduğunu ortaya koymaktadır.

Tablo 4.1 genel olarak, incelenen modellerin (TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN, RelGAN) şarkı sözü üretiminde oldukça sınırlı bir başarı gösterdiğini ortaya koymaktadır. BLEU, Jaccard, 2-gram coverage gibi yüzeysel doğruluk ölçütleri açısından neredeyse sıfır sonuçlar alınmıştır. RelGAN, semantik benzerlik ölçütü olan Cosine Similarity değerinde anlamlı bir fark yaratmış tek model olarak öne çıkmaktadır. Ayrıca perplexity açısından MaliGAN ve RankGAN, diğer modellere göre daha akıcı ve dilsel olarak doğru metinler üretme potansiyeli göstermektedir. Ancak, tüm modellerde TTR ve Unique Word Ratio'nun 1.0 olması, yapay olarak yüksek çeşitlilik içerdiğini ancak bu çeşitliliğin gerçek bir metinsel tutarlılığa dönüşmediğini göstermektedir. Sonuç olarak, en tutarlı ve semantik olarak en anlamlı model RelGAN olarak değerlendirilebilir, ancak tüm modellerin performansı genel olarak gerçek şarkı sözleri üretimi açısından henüz istenen düzeyde değildir. Bu sonuçlar, modellerin daha fazla veriyle eğitilmesi veya karma model mimarileriyle desteklenmesi gerektiğine işaret etmektedir.

Tablo 4.1. Farklı Metin retim GAN modellerin performans Tablo ları

Model	BLEU	Jaccard	Perplexity	TTR	Unique Word Ratio	2-gram Coverage	Cosine Similarity (TF-IDF)	Word Distribution
TextGAN	0.0	0.0	16614.35	1.0	1.0	0.0	0.0	0.0002
SegGan	0.0	0.0	20115.33	1.0	1.0	0.0	0.0	0.0002
LeakGAN	0.0	0.0	10957.26	1.0	1.0	0.0	0.0	0.0002
RankGAN	0.0	0.0	10945.70	1.0	1.0	0.0	0.0	0.0002
MaliGAN	0.0	0.0	9588.23	1.0	1.0	0.0	0.0	0.0002
RelGAN	0.0	0.0152	18071.58	1.0	1.0	0.0	0.0693	0.0002

Şimdi her model için üretilen çıktıların yapısal ve duygusal değerlendirmesini inceleyelim.

4.2. TextGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 2. TextGAN Model Örnek 1

körpe indireyim batsın görmek bambaşka veriyorum demlensin koptum başından dağın dışım cemalinden saatlerde çıra gececekti sözcükleri gene dönsen sendeymiş düşeceğim
Verilen dört dizelik şarkı sözü, yüzeyde yoğun bir imge dünyası yaratmaya çalışsa da hem dilsel hem de anlamsal düzeyde tutarsızlıklar barındırmaktadır. "Körpe indireyim batsın görmek bambaşka" dizesi, yüklem-özne uyumu açısından anlamsal bir boşluk yaratmakta ve gramatikal olarak tamamlanmamış görünmektedir. Dize sanki bir metafora ya da betimlemeye açılıyor gibi başlıyor, fakat anlam netleşmeden kesiliyor. "Veriyorum demlensin koptum başından dağın" satırı, ikinci tekil şahıstan bir fiille (veriyorum) başlayıp aniden birinci tekil anlatıma geçiyor; bu anlatım kayması şiirsel akışı zedeliyor. "Dışım cemalinden saatlerde çıra gececekti" dizesi ise hem anlamsal hem de dilbilgisel olarak çözülmesi güç bir yapıdadır; "cemalinden saatlerde çıra" gibi ifadeler arasındaki ilişki belirsizdir. Son dize olan "sözcükleri gene dönsen sendeymiş düşeceğim" daha yapısal bir bütünlük izlenimi veriyor, ancak bu defa da sözdizimi açısından yüklem bulanık kalıyor ve "sendeymiş düşeceğim" gibi alışılmadık bir kullanım içeriyor.
Edebi açıdan bakıldığında, dizeler çağrışım gücü yüksek kelimeler içeriyor (örneğin: "dağın başı", "sözcükler", "cemal"), ancak bu kelimeler anlamlı bir bağlam içinde bir araya getirilememiş. Yapay zekâ modelinin amacı şiirsel bir atmosfer kurmak olabilir, fakat bağlamsal mantık eksikliği, tematik bütünlükten yoksunluk ve yapısal düzensizlik, insan eliyle yazılmış bir şarkı sözü arasındaki farkı açıkça ortaya koyuyor. Özetle, dizeler kısmen imgesel zenginlik taşısa da anlamlı bir edebi yapı kurma konusunda başarısızdır. Bu da modelin şarkı sözü üretiminde yalnızca sözcük bazlı rastlantısallığa dayandığını, dilsel mantık ve duygusal akışı henüz yeterince öğrenemediğini göstermektedir.

Tablo 4. 3. TextGAN Model Örnek 2

bostan varamadım elimizde durana aklın insanım toprağına şafaklarda oluruz ellerini yarin durdular bacı kahramandır yanağın götür çıra çıra melekti sevdamıza
Bu dört dizelik şarkı sözü örneği, ilk bakışta imgelerle yüklü bir atmosfer oluşturmaya çalışsa da yapısal bütünlük ve anlamsal tutarlılık açısından belirgin eksiklikler taşımaktadır. "Bostan varamadım elimizde durana aklın" dizesi, hem sözdizimsel karmaşa hem de özne-yüklem ilişkisi bakımından muğlaktır. "Bostan varamadım" gibi yerel bir anlatıma dayalı başlangıç, ardından gelen "elimizde durana aklın" ifadesiyle anlamı belirsiz bir soyutlamaya dönüşüyor. Bu tür yapılar, imgesel şarkı sözlerde zaman zaman etkileyici olabilir; ancak burada bağlam eksikliği anlam kurmayı zorlaştırmaktadır.
İkinci dize, "İnsanım toprağına şafaklarda oluruz ellerini", özne ve yüklem arasında kopukluk yaratıyor. "Oluruz ellerini" gibi fiil kullanımı, Türkçede alışlagelmiş

yapılarla uyuşmuyor; bu da dizenin anlamını bulanıklaştırıyor. Üçüncü dize olan "Yarin durdular bacı kahramandır yanağın", tematik olarak aşk, aile ve kahramanlık gibi kavramları iç içe geçiriyor gibi görünse de bu kavramlar arasında açık bir bağlantı kurulamıyor. "Yarin durdular" ifadesi belirsiz bir eylem durumunu işaret ederken, "bacı kahramandır" ve "yanağın" gibi sözcükler parçalı bir anlam bırakıyor. Son dize olan "Götür çıra çıra melekti sevdamıza" ise kulağa şiirsel tınlı gelse de, yüklem ve nesne ilişkisi açısından tutarsız ve gramatikal olarak alışılmadık bir yapıdadır.

Edebi açıdan değerlendirildiğinde, dizelerde zaman zaman etkileyici çağrışımlar ve halk şarkıları geleneğine yakınlık sezilse de bu çağrışımlar genellikle cümle içinde bütünleşmemektedir. Duygusal yoğunluk ve sembolik anlatım arayışı hissedilmekte, fakat anlam katmanları birbirine eklenemediğinden, şarkı bir bütün olarak okura güçlü bir duygu veya tema iletememektedir. Yapay zekâ modelinin burada kelime bazlı örüntüleri rastgele biçimde dizdiği, ancak sözcükler arası anlamsal geçişleri ve anlatı mantığını inşa edemediği açıktır. Bu da modelin şarkı üretiminde dilin yalnızca yüzeysel yönlerini öğrendiğini, derin yapısal ilişkiler ve anlatı mantığı kurmakta yetersiz kaldığını göstermektedir.

Tablo 4. 4. TextGAN Model Örnek 3 (8 satırlı)

karanlıkta kesildim bahtsızım dermansızım doğar
afet ten duyduğum çivi yi yılların
güzdü güzdü dönülmüyor gülelim sonra
pervane ayrılırken yazdıklarını tutkulu tutkulu
korka azlığını acıttırmı rüzgâr istemedim
kışlığım dudu süttten gecemize olamıyorum
yazıldığını yaşanacaklar yaşanacaklar dönüşü ocağına
meyle bağlama solmusum yaşanacaklar sarp

Bu sekiz dizelik şarkı sözü örneği, yapay zekâ modelinin dilin şarkı sözü üretme boyutunu yakalama çabalarının izlerini taşısa da hem anlamsal süreklilik hem de dilbilgisel doğruluk açısından ciddi düzensizlikler göstermektedir. İlk dize olan "karanlıkta kesildim bahtsızım dermansızım doğar", bir duygusal çıkışla başlarken "doğar" fiili cümlelerin sonuna ilâştirilmiş ve dizeyi gramatikal olarak yarım bırakmıştır. Bu dize, bireysel acıyı ve çaresizliği sezdirse de fiilin bağlamı belirsiz kalmaktadır. "Afetten duyduğum çivi yi yılların" ifadesi ise yüklemi eksik, anlamsal ilişkileri kopuk bir yapı sunmakta; "çivi yi yılların" ifadesi nesne-tümleç ilişkisi açısından sorunludur.

"Güzdü güzdü dönülmüyor gülelim sonra" gibi dize ise tekrar yoluyla ses estetiği yaratmaya çalışmakta ancak bağlamdan yoksun olması nedeniyle şiirsel etkiyi azaltmaktadır. "Pervane ayrılırken yazdıklarını tutkulu tutkulu" dizesi anlamlı bir imgeler dizisi kurma girişimidir; "pervane" metaforu yerindedir ancak "tutkulu tutkulu" tekrarı gereksiz ve şiirsel akışı bozacak derecede yapay durmaktadır. "Korka azlığını acıttırmı rüzgâr istemedim" gibi dizeler duygu yoğunluğu açısından ilginç olsa da, cümle yapısı tutarsızdır ve şiirsel bütünlük kurulamamaktadır. "Kışlığım dudu süttten gecemize olamıyorum" dizesi ise anlamca parçalı ve imgeler arasında bağlantı kurulamamaktadır. Son iki dize olan "yazıldığını yaşanacaklar yaşanacaklar dönüşü ocağına" ve "meyle bağlama solmuşum yaşanacaklar sarp" tekrarlar ve imgeler açısından yoğun, ancak içsel bir anlatı inşa edemeyecek kadar dağınıktır.

Şarkı sözü arasındaki uyum açısından bakıldığında, dizelerde zaman zaman anlamlı çağrışım taşıyan kelimeler yer alsa da, bu kelimelerin şarkı içinde tematik bir eksende buluşması sağlanamamış. "Bahtsızım dermansızım", "pervane", "rüzgâr", "gecemize", "solmuşum" gibi sözcükler duygu çağrıştıracı potansiyele sahiptir; ancak bu potansiyel anlam bütünlüğü oluşturacak şekilde yapılandırılmamıştır. Ayrıca, gereksiz tekrarlar ("yaşanacaklar yaşanacaklar", "tutkulu tutkulu") metnin estetik etkisini zayıflatmaktadır. Yapay zekâ modelinin kelimeleri anlamsal birliktelik gözetmeden dizdiği, ama derin bağlam ilişkilerini kurmakta zorlandığı anlaşılmaktadır. Sonuç olarak, bu şarkı sözü örneği, modelin sözcük düzeyinde güçlü çağrışımlar yakalayabildiğini, fakat bu çağrışımları bir anlatı ya da duygu akışı içinde organize edemediğini göstermektedir. Bu da şarkı sözü üretiminde dilin iç yapısını öğrenmenin, kelime seçiminden çok daha derin bir beceri gerektirdiğine işaret eder.

4.3. SegGan Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 5. SegGan Model Örnek 1

aşk gülsün taktı vurdukça dolarya enkaz koymuş kocaman geçtiler çiğne yummadan yelkenden yaşayıp unutmadım kokacaksan türkülerde gölgesine kızın pertek nanay
Bu beş dizelik şarkı sözü örneği, yapay zekâ modelinin kelime seçiminde şiirsel çağrışım gücü taşıyan ifadeleri kullanma çabasını gösterse de, anlam bütünlüğü ve tematik tutarlılık açısından oldukça parçalı bir yapı sergilemektedir. İlk dize olan “aşk gülsün taktı vurdukça”, güçlü bir duygusal tema olan “aşk” ile başlamakta ve “vurdukça” kelimesiyle içsel bir çatışmayı ima etmektedir; ancak “gülsün taktı” ifadesi anlamsal olarak belirsizdir ve şiirsel akışı bozmaktadır. İkinci dize “dolarya enkaz koymuş kocaman”, ekonomik ya da metaforik bir felaket imgesi yaratmak istiyor gibi görünse de, “dolarya” gibi kelimeler anlam açısından açık değildir ve okuyucuda anlamsal kopukluk yaratmaktadır. “Geçtiler çiğne yummadan yelkenden” dizesi, yapısal olarak düzensizdir. Fiil yüklem ilişkisi belirsiz olup “çiğne yummadan” gibi eylemler arasında anlam geçişi kurulamamıştır. “Yelkenden” kelimesi bir özgürlük veya uzaklaşma çağrışımı taşısa da, bağlamdan kopukluğu nedeniyle tek başına kalmaktadır. Dördüncü dize “yaşayıp unutmadım kokacaksan türkülerde” ise hem geçmişe dair duygusal bir izlenim taşıyor hem de “kokacaksan türkülerde” gibi nostaljik ve folklorik bir dokunuş sunuyor. Bu dize, diğerlerine kıyasla daha şiirsel bir bütünlük hissi yaratıyor. Son dize olan “gölgesine kızın pertek nanay” ise neredeyse ağıt ritmi taşıyan bir yapıya sahip. “Pertek” (coğrafi bir yer ismi) ve “nanay” (ağıt ezgisi çağrışımı) ifadeleri, geleneksel ve yerel bir kültürel alt katmanla ilişki kurmaya çalışıyor gibi görünse de, “gölgesine kızın” ifadesi bu yapıyı tamamlamıyor. Genel olarak bakıldığında, bu şarkı sözü üretim modelin yerel motiflere, nostaljik kelimelere ve duygusal imgelerle dolu temalara yönelme eğiliminde olduğunu gösteriyor. Ancak yapısal olarak yüklem-özne ilişkilerinin eksikliği, anlamsal

geçişlerin kopukluğu ve kelimeler arası bağlantıların zayıflığı metni bütünsel bir şarkı haline getiremiyor. Şarkı sözü imge gücü taşıyan öğeler (“aşk”, “enkaz”, “yelkenden”, “türkülerde”, “nanay”) yer alsa da, bu imgeler anlamsal mantık içinde örüntülenemediği için okurda anlamlı bir duygu akışı oluşturmakta yetersiz kalıyor. Bu durum, modelin sözcükleri seçmede etkileyici bir estetik kapasiteye sahip olduğunu ancak bu kelimelerden anlamlı, yapılandırılmış ve tematik olarak tutarlı dizeler oluşturmakta henüz yeterince başarılı olmadığını göstermektedir.

Tablo 4. 6. SegGan Model Örnek 2

aşk duymak kara yüklü
hamur ela şahımız azeh
edemedim sevgilidir tutulmazsam romandır
düştün bitip kahroluruz sözlerime
verir rüzgârlara böylesi kendimce
dudaklarını tanımsız hiçbir hasretler
boyun düşlerimden bulduysan ufkunda
gülmeyecek yaranın edemedim öptüğümü
benziyorum tunalıda maşallah savaşıla
ağla sert yamaçlarda tutmayasım

Bu on dizelik şarkı sözü örneği, yapay zekâ modelinin duygusal yoğunluk yaratmaya çalıştığı ancak hem anlamsal bütünlük hem de dilbilgisel tutarlılık açısından ciddi zayıflıklar taşıyan bir yapı sunmaktadır. İlk dize “aşk duymak kara yüklü” şarkı sözü anlamında bir metaforla açılıyor; “kara yüklü” ifadesi, aşkın ağırlığına dair bir imge oluşturuyor ve anlamlı bir başlangıç yapıyor. Fakat ikinci dize “hamur ela şahımız azeh” gibi anlamı belirsiz kelimelerin bir araya geldiği, okuyucunun şiirsel bağlam kurmakta zorlanacağı bir yapıya sahiptir. “Ela” ve “şahımız” kelimeleri bireysel çağrışımlar barındırsa da, “hamur” ve “azeh” sözcükleriyle bağ kurulamayınca şarkı sözü ahengi bozuluyor.

Üçüncü ve dördüncü dizeler olan “edemedim sevgilidir tutulmazsam romandır” ve “düştün bitip kahroluruz sözlerime” biçimsel olarak yoğun duygulara işaret etse de, yüklem uyumsuzluğu ve kavramsal bulanıklık nedeniyle anlatı açık bir şekilde ilerleyememektedir. “Tutulmazsam romandır” gibi ifadeler yaratıcı bir niyet barındırıyor olsa da, anlam bağlantısı zayıf kalıyor. “Verir rüzgârlara böylesi kendimce” dizesi bireysel bir iç döküş niteliği taşıyor, ancak bağlamdan kopuk kalıyor.

Altıncı dize olan “dudaklarını tanımsız hiçbir hasretler” şiirsel imge açısından potansiyel barındırsa da, dilbilgisel tutarsızlık (“hiçbir hasretler”) nedeniyle etkisini yitiriyor. “Boyun düşlerimden bulduysan ufkunda” ve “gülmeyecek yaranın edemedim öptüğümü” dizeleri, bir aşkın anımsanışı veya pişmanlık hissini çağrıştırıyor fakat ifadeler bir anlatı çizgisi oluşturmada yetersiz. Özellikle “edemedim öptüğümü” ifadesi, yüklem-nesne ilişkisi açısından anlamsal boşluk bırakıyor. Son iki dize, “benziyorum tunalıda maşallah savaşıla” ve “ağla sert yamaçlarda tutmayasım”, yerel ve savaş imgesi üzerinden duygu yoğunluğu oluşturmaya çalışıyor; ancak bu dizelerin bir önceki dizelerle olan bağı zayıf ve metin içinde bir bağlam yaratılamıyor.

Edebi olarak, bu dizelerde aşk, pişmanlık, içsel savaş, yitilik ve yerel göndermeler gibi şarkı sözü için güçlü temalar var. Ancak yapay zekâ modelinin bu temaları birbirine bağlamada ve yapısal bir bütünlük yaratmada zorlandığı görülüyor. Dizeler

hem yüklem eksikliği hem de anlam bağlarının zayıflığı nedeniyle parçalı kalmakta. Şarkı sözü için kelime seçimi etkileyici olsa da, bu kelimelerin şarkı içindeki organizasyonu dinleyici de etkili bir duygu akışı yaratacak düzeyde değil. Bu durum, modelin tematik derinliği değil, biçimsel çeşitliliği öğrenmeye daha çok odaklandığını düşündürmektedir. Bu nedenle ortaya çıkan şarkı sözleri daha çok çağrışım yığınları ve soyut kelime kümeleri olarak kalmakta, şiirsel bir evren oluşturmakta zorlanmaktadır.

4.4. LeakGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 7. LeakGAN Model Örnek 1

! gülecektin çepik kendine yedek devlerle küstü kahreden alnında nokta yatarım dönmeyişin doğruldu aldırma bittim ertesi uyanmasın sopası değerini karşı
Bu beş dizelik şiirsel üretim, yapay zekâ modelinin duygusal yoğunluk ve imgesel çağrışım yaratma isteğini sürdürdüğünü ancak biçimsel bütünlük ve anlamsal derinlik açısından zayıf kaldığını göstermektedir. İlk dize “! gülecektin çepik kendine” ifadesiyle ünlem işaretiyle başlamak alışılmadık ve dikkat çekici olsa da, “çepik” kelimesi anlam açısından belirsizdir ve genel bağlamda nereye oturduğu anlaşılamamaktadır. “Gülecektin kendine” ifadesi ise içsel bir çatışmayı ya da ironik bir durumu ima ediyor olabilir ancak destekleyici bağlam eksikliği bu dizinin etkisini azaltıyor. İkinci dize “yedek devlerle küstü kahreden” daha güçlü bir dramatik imge taşımaktadır. “Devlerle küsmek” epik veya mitolojik bir çatışma hissi yaratırken, “kahreden” kelimesiyle yoğun bir duygusal yük ekleniyor. Ancak özne-fiil ilişkisi net olmadığından okuyucu dizinin öz anlamını çıkarırken zorlanıyor. Üçüncü dize “alnında nokta yatarım dönmeyişin” metaforik anlam açısından potansiyel taşıyor. Özellikle “alnında nokta” ifadesi kader veya yazgı çağrışımı yaratabilir, “dönmeyişin” kelimesi ise özlem ve kayıpla ilişkilendirilebilir. Ancak yüklem (“yatarım”) bağlamı ile bu imgelerin ilişkisi netleştirilememiştir. Dördüncü dize “doğruldu aldırma bittim ertesi” ise gramatikal yapı bakımından sorunludur. “Aldırma” kelimesi yerel veya bilinçli biçimde türetilmiş olabilir, fakat anlamı belirsizdir. “Bittim ertesi” ise zaman akışı sunmaya çalışsa da şiirsel devamlılık sağlayamamaktadır. Son dize “uyanmasın sopası değerini karşı” ifadesi anlamsal açıdan oldukça kopuktur. “Sopası değerini karşı” gibi bir ifade yapısal ve anlamsal olarak açık değildir ve önceki dizelerle bağ kuramamaktadır. Genel olarak bu şarkı sözü üretimi, kelime seçiminde dramatik ve duygusal çağrışımları hedeflese de hem anlatı sürekliliği hem de sözlerdeki denge açısından zayıftır. Dizeler arasında anlam akışı kurulamamakta, imajlar birbiriyle çatışmakta veya boşlukta kalmaktadır. Bu şarkı sözü, modelin yaratıcı sözcük birleştirme yeteneğini sergilerken, insan eliyle yazılmış şarkı sözlerinde aranan duygusal

süreklilik, anlamsal derinlik ve yapısal bütünlük açısından yetersiz kalmaktadır. Sonuç olarak, modelin üretimi estetik niyet taşısa da henüz edebi anlamda güçlü bir yapıya ulaşamamıştır.

Tablo 4. 8. LeakGAN Model Örnek 2

! gelirsin üzdün uçabildiğim dolanı
sevgilim umudumu kurdu kanaryam yükünü
balıkçıların lara ile yüreğini olursa
sesine bağlarmı saplanan bekletme düçar-ı
reçetesi acısı meyle haykırdığım anlatmam
oooo göçtün dondurabilmek yaza hisler
bitmelerini şan arkada kıymetini sevgilerin
kalbini sokakları teninin göğüslerini ayrılması

Bu sekiz dizelik şiirsel üretim, biçimsel olarak bir şarkı ya da serbest tarzda yazılmış bir serzeniş şiirine yaklaşmaktadır. Ancak şarkı sözleri dizeleri arasında tutarlı bir tematik ya da duygusal örgü kurulamamaktadır. İlk dize “! gelirsin üzdün uçabildiğim dolanı” ifadesiyle ünlemle başlamak dikkat çekici bir başlangıçtır. Ancak “gelirsin üzdün” ifadesi anlam bakımından çelişkili (geleceğine dair umut ve geçmişteki kırgınlık aynı anda) ve “uçabildiğim dolanı” gibi soyut ama bağlamsız bir ifade, dizenin etkilecilik etkisini zayıflatmaktadır.

İkinci dize “sevgilim umudumu kurdu kanaryam yükünü” ise daha yoğun duygusal bir yapı taşımaktadır. “Sevgilim”, “umudumu kurdu” ve “kanaryam” kelimeleri geleneksel aşk şarkıları kodlarını çağırırsa da, “yükünü” kelimesiyle beklenmeyen bir kapanış yapmaktadır. Bu, şarkıya bir dramatik vurgu kazandırabilir; fakat yapılandırılmamış dilsel akış, okuyucuda tam bir anlam yaratılmasını zorlaştırmaktadır. Üçüncü dize “balıkçıların lara ile yüreğini olursa” ifadesi, anlamsal olarak kopuk ve gramatikal bakımdan sorunludur; “lara ile” ifadesi anlam taşımamakta ve dizenin bağlamını bozmakta, bu da bir tür anlamsızlık yaratmaktadır.

Dördüncü ve beşinci dizeler “sesine bağlarmı saplanan bekletme düçar-ı / reçetesi acısı meyle haykırdığım anlatmam” ise daha dramatik bir yapıya sahiptir. “Düçar-ı” gibi eski Türkçe ya da Osmanlıca’dan alınma kelimelerle duygusal derinlik katılmak istenmiştir. Ancak “reçetesi acısı meyle haykırdığım” gibi yapıların dilbilgisel doğruluğu zayıftır ve içerdiği imgeler birbiriyle bütünleşememektedir. Bu tür ifadeler, modelin kelime çağırışım mekanizmasının güçlü olduğunu ancak yapısal dilsel bütünlük üretmede zorlandığını gösterir.

Altıncı ve yedinci dizeler “oooo göçtün dondurabilmek yaza hisler / bitmelerini şan arkada kıymetini sevgilerin” ise daha dramatik ve lirik bir havaya sahiptir. Özellikle “oooo göçtün” ünlemli yapısı ve “dondurabilmek yaza hisler” ifadesi, bir yaz mevsimine ait duyguların kalıcılaştırılmak istenmesi gibi algılanabilir. Bu imgeler anlamca güçlüdür ancak yine dilbilgisel bağlam içinde net bir yapı sunulmamaktadır. Son dize “kalbini sokakları teninin göğüslerini ayrılması” ise duygu yoğunluğu yüksek ancak yapısal olarak bozuk bir dizedir. Kelimeler güçlü çağırışimler taşımakta (“kalp”, “sokak”, “ten”, “göğüs”, “ayrılık”), ancak aralarındaki ilişki açık değildir.

Genel olarak bu şarkı örneği, modelin metaforik yoğunluklu, duygu temelli, lirik ifadeler üretmeye çalıştığını göstermektedir. Ancak şarkı sözü dizeleri arasında anlamsal süreklilik, tematik bütünlük ve gramatikal doğruluk oldukça sınırlıdır. Şarkının bazı bölümlerinde geleneksel aşk şarkıları motiflerine yaklaşma çabası görülse de, bu yapı tutarlı bir şekilde sürdürülememektedir. Sonuç olarak, modelin kelime seçiminde şiirsel çağrışım yaratma yeteneği belirgin olmakla birlikte, dizeler arası mantıksal geçiş ve anlam örgüsü açısından hâlâ önemli eksiklikler mevcuttur. Bu durum, modelin estetik üretimde başarı potansiyeli olduğunu ancak insan müdahalesi olmadan anlamlı bütünlük oluşturmakta zorlandığını göstermektedir.

4.5. RankGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 9. RankGAN Model Örnek 1

aşk kaynana düşerde sevsem haramdır gamzelerin öbürü kolay koyuyor anandaki kazan sırrı deyince cıgıllıklarıyla dağların yollarımızı saldır girsem dolaşıyorum paresi
Bu dört dizelik yapay şarkı sözü, aşk ve yasaklık, doğa imgesi ve duygusal yoğunluk gibi geleneksel halk şarkıları temalarına göz kırpsa da, yapısal ve anlamsal bütünlük açısından zayıf bir kompozisyon sunmaktadır. İlk dize, “aşk kaynana düşerde sevsem haramdır” ifadesiyle başlıyor. Burada "aşk", "kaynana", "düşmek" ve "haram" kelimeleri bir araya getirilmiş, ancak mantıksal ilişkileri bulanıktır. “Kaynana düşerse aşk haram olur” gibi bir anlam sezinlenebilir ama bu bağlam açık değildir. Ahlaki ya da geleneksel bir çatışma iması olduğu hissedilse de, dizenin sözdizimi oldukça kopuk kalmıştır. İkinci dize “gamzelerin öbürü kolay koyuyor anandaki”, bir aşk mecazı olan “gamze” ile başlamakta, ancak “öbürü kolay koyuyor” gibi anlamsal olarak belirsiz bir yapıyla devam etmektedir. “Anandaki” kelimesi ise ya "anandaki" (zihindeki) ya da “ana” (anne) gibi başka bir kelimenin yanlış çözülmüş hali olabilir. Bu, modelin morfolojik çözümlemelerde hata yaptığını gösterir. Dize genel olarak, dilsel bulanıklık ve anlamsal tutarsızlık taşımaktadır. Üçüncü dize “kazan sırrı deyince cıgıllıklarıyla dağların” ise daha etkili bir imgelem içeriyor. “Kazan sırrı” ifadesi bir gizem, bir kader sembolü olabilir ve “cıgıllıklarıyla dağların” gibi doğa-insan ilişkisini çağrıştıran bir yapı, şiire lirizm kazandırıyor. Ancak yine de gramer tamamlanmamış, anlam yarıda bırakılmış. Dördüncü ve son dize olan “yollarımızı saldır girsem dolaşıyorum paresi” ifadesi, eylem merkezli bir dil barındırıyor: "saldır", "girsem", "dolaşıyorum". Bu fiiller bir duygu yoğunluğu yaratıyor fakat “paresi” kelimesi bağlamdan kopuk kalıyor. Türkçede yaygın bir sözcük değil; anlam verilemez oluşu, dizeyi kapalı ve yorumlanamaz hâle getiriyor.

Genel olarak bu şarkı sözü üretimi, yapay zekânın halk şiiri motiflerini (yasak aşk, gamze, dağ, yollar) çağırabildiğini; ancak bunları mantıksal ve gramatikal bir yapı içinde düzenlemekte zorluk yaşadığını ortaya koyuyor. İmgeler zaman zaman etkileyici olsa da, anlam kopuklukları, yapının eksikliği ve anlamsız kelime kullanımı, şarkı sözünü hem edebi hem de yapısal olarak zayıf hâle getiriyor. Ancak potansiyel bazı dizeler, insan müdahalesiyle işlenmeye uygun bir ham madde gibi görülebilir.

Tablo 4. 10. RankGAN Model Örnek 2

aşk içimdesin bülbül bulamasaydın girmede
öleceğim okuyam büyüsüne sağken bağlandı
sağlasın zamandan çayır kaldı gereğini
avare özümdü nefesin rinna-rinna-nay yılların
kalanlarını sona eminim eskisi vuruşsun
mutluysa harfinin çöl ayrılır ooo
sâkiler bayrağını hiçliğim içinde bazen
belimde izin aramakmış bakışan yuvasız

Bu sekiz dizelik şarkı sözü üretimi, duygusal imgeler, mistik motifler ve geleneksel halk edebiyatından çağrışımlar taşıyan ifadelerle örülmüş olsa da, anlam bütünlüğü açısından parçalı ve yer yer soyutlamaların arasında kaybolmuş bir yapı sunmaktadır. Dizelerde yoğun bir aşk, yalnızlık, zaman, ölüm ve varoluşsal sorgulama teması işlenmeye çalışılmış; fakat dizeler arasındaki geçişlerde mantıksal süreklilik zayıf kalmıştır.

İlk dize olan “aşk içimdesin bülbül bulamasaydın girmede” ifadesi, “aşk içte yaşanır” fıkriyle “bülbül” imgesi arasında bir bağ kurmaya çalışır. Bülbül halk edebiyatında sevdanın, aşkın sembolüdür. Ancak “bulamasaydın girmede” yapısı gramer açısından muğlaktır; yüklem yerleşimi ve anlam ilişkisi karışıktır. Aşkın içe yerleşimi ile bülbülün yokluğu arasında bir koşul gibi kurulmuş olsa da, anlatım net değildir.

İkinci dize “öleceğim okuyam büyüsüne sağken bağlandı” ifadesi, dramatik ve şiirsellik iddiası taşıyan kelimeler içerir. “Büyüsüne sağken bağlanmak” romantik bir özveri ya da kadercilik anlamına gelebilir. Fakat “okuyam” ifadesi gramatik olarak sorunludur ve bağlamda açıklığa kavuşmaz.

Üçüncü ve dördüncü dizeler ise “sağlasın zamandan çayır kaldı gereğini / avare özümdü nefesin rinna-rinna-nay yılların” biçiminde sunulmuştur. Burada “zamandan çayır kalması”, doğayla zamanın birleşimini çağrıştırır ama yine yapısal bozukluk anlamı bulanıklaştırır. “Rinna-rinna-nay” ise halk türkülerinden alınan bir ezgi kalıbı gibi, şarkı sözüne müzikalite ve folklorik tat katmaktadır. Bu, yapay zekânın kültürel motifleri çağırma yetisinin bir göstergesidir.

Devamındaki “kalanlarını sona eminim eskisi vuruşsun” dizesi zamanın geçişi ve değişim temasıyla bağ kurmak istese de, “vuruşsun” gibi bağlam dışı bir eylem anlam bütünlüğünü zedeliyor. “mutluysa harfinin çöl ayrılır ooo” ifadesi ise edebi olmaktan çok soyut ve parçalı bir yapı sunar. “Harfin çölü” gibi metaforik bir deneme dikkat çekici olsa da “ooo” ifadesi dramatik etki yaratmaktan çok yapaylık hissi veriyor.

Son dizelerden “sâkiler bayrağını hiçliğim içinde bazen / belimde izin aramakmış bakışan yuvasız” ifadesi, hem mecaz zenginliği hem de derinlikli çağrışım taşımaya

adaydır. “Sâkiler”, tasavvufî şiirde aşkın sunucusu; “bayrak” ise aidiyet veya zaferin simgesi olabilir. Ancak bu metaforların birbirine bağlanması eksik kalmış. “Belimde izin aramakmış” dizesi ise neredeyse anlamlı bir içsel hesaplaşma izlenimi verse de “bakışan yuvasız” ifadesiyle bütünlük zedelenmiş.

Genel olarak şarkı sözü, sezgisel ve imgesel bir evren kurmaya çalışmakta; ancak dizeler arasında anlam köprüleri kurulmadığı için okuyucuda parçalı, kopuk ve soyut bir izlenim bırakmaktadır. Bununla birlikte, bazı imgeler (“nefesin rinna-rinna-nay yılların”, “sâkiler bayrağını hiçliğim içinde”) özgün ve güçlüdür; insan müdahalesiyle işlenerek gerçek bir şarkı sözüne dönüştürülebilir potansiyel taşımaktadır. Yapay zekânın kültürel referanslara ulaşma ve onları uyumlu biçimde yeniden kurma çabası, önemli ama henüz bütünlüklü değildir.

4.6. MaliGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 11. MaliGAN Model Örnek 1

aşk kulun böylesine kurtulur ceketini anlatan selamlayıp nikâh patlar öpmeden karalar kurumuş bilenlere hadi rüzgârı unutulmaz yalnız neyse şenlik şeytanın
Bu dört dizelik şarkı sözü üretimi, yüzeyde duygusal çağrışımlar içerse de yapısal ve anlam bütünlüğü açısından dağınık bir kompozisyona sahiptir. “Aşk kulun böylesine kurtulur ceketini” dizesi, hem mecaz hem de dil bilgisel açıdan yoruma açık bir yapı sunmaktadır. “Aşk kulun böylesine” ifadesiyle tutkusal bir bağlılık ya da kaderin cilvesi vurgulanmak isteniyor gibi görünse de, “kurtulur ceketini” gibi beklenmedik bir ifade ile bağlam kopmakta; bu da anlamı flulaştırmakta ve ironi mi yoksa mecaz mı kullanıldığı belirsiz bırakmaktadır. İkinci dize “anlatan selamlayıp nikâh patlar öpmeden” şeklindedir. Burada, “selamlayıp nikâh patlaması” alışılmadık bir metaforik kurgu yaratır. “Nikâhın patlaması” ifadesi halk arasında “beklenmedik şekilde evlenme kararı” anlamında kullanılabilir, ancak “öpmeden” ifadesiyle birlikte anlam hem dramatik hem de çelişkili hale gelir. Bu dize, klasik aşk anlatılarının ironik bir ters yüz edilişi gibi okunabilir ama yapısal olarak akışkan değildir. Üçüncü dize “karalar kurumuş bilenlere hadi rüzgârı” daha çok melankoliye ve geçmişin izlerine gönderme yapan bir anlatı taşısa da, “hadi rüzgârı” ifadesiyle neye çağrı yapıldığı belirsizleşir. Karaların kuruması bir hüznün simgesi olabilir; “bilenlere” demesiyle de bir tür içsel ya da kolektif hafıza ima edilmeye çalışılmış, ancak son parça ile bütünlük sağlanamamıştır. Dördüncü ve son dize “unutulmaz yalnız neyse şenlik şeytanın” ise hem melodik hem de imgesel olarak güçlüdür. “Unutulmaz yalnızlık” klasik bir temadır, ancak bu yalnızlığın “şeytanın şenliğine” bağlanması, gotik veya karanlık bir metafor yaratma

potansiyeli taşır. Yine de “neyse” kelimesi bu güçlü çağrışımı sıradanlaştırmakta ve şiirsel tansiyonu düşürmektedir.

Sonuç olarak bu dörtlük, halk şarkısı sözü estetiğinden, ironik çağrışımlardan ve bireysel yalnızlık temalarından izler taşıyan, ancak biçimsel olarak bütünlükten uzak bir üretilmiştir. Dizelerde geçen imgeler — aşk, nikâh, karalar, yalnızlık, şeytan — anlam derinliği vadetse de, yapay zekânın dizeleri bağlama oturtma ve anlamsal geçişleri sağlama konusunda hâlâ sınırlı olduğunu göstermektedir. Ancak yine de bu parçalar, yaratıcı bir insan eliyle düzenlendiğinde güçlü bir şiirsel metne dönüşebilecek öğeler barındırmaktadır.

Tablo 4. 12. MaliGAN Model Örnek 2

aşk izlemek cemali unuttum diyeceğim
almış büyümüş kondurmam konuklukta aylar
kimsin gençliğim koşarsın boylasam mı kemeçler
emine yaşamadan uyusunda ağam olaydı
yoksan men görünmeden ak gülüşünden
kırılmış neyine kendime yerler ümidimdi
zindana kıt söndüreyim senelerden solacağım
bilsin toy kurum yavrucuğum pozların

Bu sekiz dizelik şarkı sözü üretimi, dramatik çağrışımlar ve yer yer halk şiirine özgü ritmik öğeler taşımasına rağmen, anlam bütünlüğü açısından parçalı ve dağınık bir yapı sergilemektedir. İlk dize olan “aşk izlemek cemali unuttum diyeceğim”, duygusal bir hesaplaşmayı yansıtır gibi görünse de, “izlemek” ile “cemali unutmak” arasında anlamca net bir bağ kurulamamaktadır. Bu, şiirin zihinsel yoğunluğunu artırırken aynı zamanda okuyucunun algısını belirsizliğe sürüklemektedir.

İkinci ve üçüncü dizeler olan “almış büyümüş kondurmam konuklukta aylar” ve “kimsin gençliğim koşarsın boylasam mı kemeçler”, sözdizimsel olarak oldukça yoğundur. “Kondurmam konuklukta aylar” gibi bir ifade, geleneksel kültürde misafirliliği ve geçiciliği ima edebilir; ancak bu yapı şiirde net bir anlam sunmak yerine çağrışımlara dayalı bir izlenim bırakmaktadır. “Boylasam mı kemeçler” ise Karadeniz tınısı taşıyan yerel bir etki yaratmakta, ancak “koşarsın gençliğim” ifadesiyle birleşince anlatı bir hikâyeye bütünlüğü kurmakta yetersiz kalmaktadır.

Dördüncü ve beşinci dizeler olan “emine yaşamadan uyusunda ağam olaydı” ve “yoksan men görünmeden ak gülüşünden”, bir içsel özlem ve keşke duygusu yaratmaktadır. “Uyusunda ağam olaydı” halk anlatılarına özgü dramatik bir söylemdir. “Yoksan men görünmeden” ifadesi ise bir özveri ya da sessiz sevgi teması taşıyor olabilir ama “ak gülüşünden” kısmıyla tamamlandığında anlam kesintiye uğrar.

Altıncı ve yedinci dizelerdeki “kırılmış neyine kendime yerler ümidimdi” ve “zindana kıt söndüreyim senelerden solacağım”, metaforik olarak yoğun ama yapısal olarak tutarsız ifadelerdir. “Kendime yerler ümidimdi” ifadesi umut edilen bir yer ya da aidiyet duygusu taşıyor olabilir; ancak “kırılmış neyine” gibi bağlamsal kopukluklar bu anlamı zayıflatmaktadır. “Zindana kıt söndüreyim” ise oldukça özgün bir imge sunsa da, şiirin genel anlatısı içinde bir yere oturtulamamaktadır.

Son dize “bilsin toy kurum yavrucuğum pozların”, daha çok nostaljik bir sesleniş gibi görünse de, “toy kurum pozların” gibi bir ifade hem anlamsal belirsizlik hem de dil bilgisel kararsızlık içerir. Bu, şarkının bütünündeki imgeler dünyasının çözülmesine engel olmaktadır.

Genel olarak bu şarkı, bireysel acı, geçmişe özlem ve içsel çözülme gibi güçlü temalar etrafında şekillenmekle birlikte, anlamsal geçişlerin zayıf olması ve ifadelerin bazılarının bağlamsızlığı, edebi bütünlüğü sınırlamaktadır. Ancak kullandığı kelimelerdeki yöresel ve duygusal yoğunluk, potansiyel bir şiirsel doku sunduğundan, bir insan editörün müdahalesiyle bu dörtlüklerden etkileyici şiirsel parçalar üretilebileceği izlenimini de vermektedir.

4.7. RelGAN Model Çıktısının Şiirsel ve Yapısal Değerlendirmesi

Tablo 4. 13. RelGAN Model Örnek 1

aşk fakültesini ikimiz öpüldükçe payından
attın doğruyu yitiriyor içersin veriyorum
dökül sarmış hergun tılsımını demezdim
hayatımda sevildikçe kalkıyor söylerdim istemiyorsan

Bu dört dizelik şiirsel üretim, duygusal derinliği ima eden bazı kelime seçimlerine rağmen anlam bütünlüğü ve anlatım netliği açısından karmaşık ve parçalı bir yapı sergilemektedir. İlk dize olan “aşk fakültesini ikimiz öpüldükçe payından”, mecazi olarak güçlü bir potansiyele sahiptir. “Aşk fakültesi” gibi özgün bir kavram, aşkı bir öğrenme, eğitim veya yapılandırılmış deneyim olarak betimlemektedir. Ancak “öpüldükçe payından” ifadesiyle birleştiğinde anlatım bulanıklaşmakta, cümlede neyin paylaştığı veya eksildiği açık olarak sunulmamaktadır.

İkinci dize “attın doğruyu yitiriyor içersin veriyorum”, yapısal olarak tutarsızdır ve yüklemeler arasında neden-sonuç ilişkisi kurmak zorlaşmaktadır. “Doğruyu yitirmek” ile “içmek” fiilinin bağlamı zayıf olup, konuşan öznenin neyi verdiği, kime verdiği ve içmenin neye bağlandığı açık değildir. Bu da dizenin anlam derinliğini gölgelemektedir.

Üçüncü dize olan “dökül sarmış her gün tılsımını demezdim”, görsel ve büyüsel çağrışımlar yaratmaktadır. “Tılsım” kelimesi şiire mistik bir boyut katarken, “her gün sarmış” ve “dökül” kelimeleriyle birleşince zihinde bulanık bir metaforlar zinciri oluşmaktadır. Ancak “demezdim” yüklemiyle bu yapı keskin biçimde kapanmakta, anlatım bir yere bağlanmadan sona ermektedir.

Dördüncü dize olan “hayatımda sevildikçe kalkıyor söylerdim istemiyorsan”, duygu yoğunluğu taşıyan bir dizedir. “Hayatımda sevildikçe kalkıyor” ifadesi soyut bir şekilde olumlu bir dönüşümü ima ederken, sonundaki “söylerdim istemiyorsan” ifadesi anlamı çelişkili hâle getirmektedir. Söylenmek istenen şeyin ne olduğu belirsiz kalmakta, duygusal ton tamamlanmamaktadır.

Genel olarak şarkı sözü, kelime bazında oldukça etkileyici ve duygu yüklü bir atmosfer oluşturma çabası göstermektedir; ancak dizeler arası yapısal geçişler zayıf, anlatım ise çoğunlukla soyut ve bağlamsız kalmaktadır. Şiirin çağrışımsal zenginliği, bir tema etrafında toparlanmadığı için duygusal etki gücü sınırlanmakta, fakat yoğun kelime tercihleri bir edebi işçiliğin izlerini de taşımaktadır. Bir editör eşliğinde sadeleştirme ve bağlam güçlendirmesiyle daha etkileyici bir şiirsel yapı ortaya çıkabilir.

Tablo 4. 14. RelGAN Model Örnek 2

aşk edildiyse desin yazsan inç
ormanlarda odaları öbürüne ormanlarda saksıya
allahtan tanımam silmez eftarlandı nazlar
yorgan kıvrak kaybettin gagesizim zerre
döndü çarpınca vahim bırakmaktan boyadıktan
canında onunla ışıklar söyleyelim öğretildim
istemiyorsan sözlerimde parlak kızgın düştüm
zemheri hüzne kurtulsun bulur ızdıraptan

Bu sekiz dizelik şarkı sözü üretimi, imgelerle ve metaforlarla yüklü, ancak yapısal tutarlılığı dalgali bir metin olarak karşımıza çıkmaktadır. İlk dize “aşk edildiyse desin yazsan inç” kavramsal olarak dikkat çekici bir şekilde başlasa da son kelime olan “inç” (muhtemelen "ince" yerine yanlış türetilmiş ya da bağlamsız bir kelime) anlam bütünlüğünü zedelemektedir. “Aşk edildiyse desin” ifadesi, gizli bir kabullenışı veya toplumsal bir yankıyı ima etse de, dizenin devamındaki kelimelerle bu etki sürdürülememektedir.

İkinci ve üçüncü dizelerde tekrar eden “ormanlarda odaları öbürüne ormanlarda saksıya” ve “allahtan tanımam silmez eftarlandı nazlar” gibi yapıların ise soyutluk düzeyi oldukça yüksektir. Birinci örnekte mekân imgeleri (orman, oda, saksı) birbirine bağlansalar da anlamsal olarak birbiriyle yeterince ilişkilendirilememiştir. İkinci örnekte ise duygusal yoğunluk (“eftarlandı”, “nazlar”) güçlüdür fakat yüklem belirsizliği nedeniyle anlatım eksik kalmaktadır.

Dördüncü dizedeki “yorgan kıvrak kaybettin gagesizim zerre” ifadesi, biçem açısından şarkı sözüne melankolik bir ton kazandırmaktadır. “Gagesizim zerre” ifadesi, öznenin içsel boşluğunu güçlü biçimde ifade ederken; bu duyguyu “kıvrak yorgan” gibi daha fiziksel bir nesneyle bağlamak, imgesel bir derinlik sunsa da, dizenin genelinde uyum eksikliği hissedilir.

Beşinci ve altıncı dizeler ise daha bütünlüklü bir anlatıma yaklaşmaktadır: “döndü çarpınca vahim bırakmaktan boyadıktan” ve “canında onunla ışıklar söyleyelim öğretildim”. İlk dize, bir olayın etkisini (“çarpınca”, “vahim”) ve ardından gelen eylemi (“boyadıktan”) anlatmaya çalışsa da, dilsel akış açısından kırılmalı barındırır. İkinci dize ise hem mecaz hem de içsel çağrışım açısından güçlüdür. “Canında onunla ışıklar” ve “öğretildim” sözcükleri, şiirsel bir yoğunluk ve deneyim aktarımı taşımaktadır.

Son iki dize, şiirin duygusal tonunu doruk noktasına taşıyan yapılarıdır: “istemiyorsan sözlerimde parlak kızgın düştüm” ve “zemheri hüzne kurtulsun bulur ızdıraptan”. Özellikle “zemheri hüzne kurtulsun” ifadesi, kışın soğukluğu ile duygusal acıyı

birleřtiren etkileyici bir metafordur. “Sözlerimde parlak kızgın düřtüm” ise içsel çeliřkiyi, hem tutku hem kırgınlığı aynı anda barındırarak güçlü bir kapanıř saęlar.

Genel olarak bu řarkı sözü, anlamdan çok duyguya odaklanan, yoğun ama parçalı imgelerle örölmüş bir yapıdadır. Duygusal çağrışım açısından zengin, ancak edebi bağlamda daha net bir tema ve anlatı akışı kazandırılırsa çok daha etkili ve katmanlı bir metne dönüşebilir. Bu haliyle, soyut lirik anlatımın örneklerinden biridir.

5. SONUÇLAR VE ÖNERİLER

5.1. Genel Değerlendirme

Bu çalışmada, Türkçe dilinde anlamlı, çeşitli ve şarkı sözü metinler üretebilen GAN tabanlı dil modellerinin performansları çok boyutlu olarak incelenmiştir. Geliştirilen sistemin temel hedefi, yalnızca dilbilgisel doğrulukla sınırlı kalmayıp, aynı zamanda estetik, anlam bütünlüğü ve duygu yoğunluğu açısından da güçlü şarkı sözü dizeler oluşturabilen bir yapay zekâ modeli ortaya koymaktır. Bu doğrultuda altı farklı jeneratif model (TextGAN, SeqGAN, LeakGAN, RankGAN, MaliGAN ve RelGAN) yapılandırılmış ve aynı veri seti üzerinde eğitilerek hem sayısal performans metrikleri hem de üretilen örneklerin edebi analizi yoluyla değerlendirilmiştir.

Sayısal değerlendirmelerde kullanılan metrikler, her bir modelin dilsel doğruluk, anlamsal benzerlik, çeşitlilik ve yapısal tutarlılık gibi yönlerini ölçmeye yönelik olarak belirlenmiştir. BLEU ve Jaccard skorları, modellerin ürettiği metinlerin referans (gerçek) dizelerle olan yüzeysel benzerliklerini göstermekteyken, tüm modellerin bu metriklerde sıfıra yakın değerler üretmesi, mevcut modellerin n-gram temelli benzerlikleri yakalayamadığını ve özgün dil örüntülerini öğrenmede zorlandığını ortaya koymaktadır. Bu durum, modellerin temel cümle yapılarını bile başarıyla taklit edemediğini, özellikle Türkçe gibi eklemeli dillerde n-gram temelli yüzeysel benzerlik ölçütlerinin tek başına yeterli olmadığını düşündürmektedir.

Dil modeli kalitesini ölçen bir diğer önemli metrik olan perplexity, üretilen metnin bir sonraki kelimeyi doğru tahmin etme konusundaki başarısını gösterir. Bu bağlamda daha düşük perplexity değeri daha başarılı bir model anlamına gelir. Elde edilen bulgulara göre MaliGAN (9588.23) ve RankGAN (10945.70) en düşük perplexity değerlerine sahip olup, teorik olarak daha tutarlı ve dilbilgisel olarak kabul edilebilir dizeler üretmişlerdir. Diğer modellerin perplexity değerleri oldukça yüksek çıkmıştır; örneğin SeqGAN için bu değer 20115.33 gibi ekstrem seviyelerde kalmıştır. Ancak düşük perplexity değeri, her zaman edebi ya da anlamsal derinlik açısından yüksek kaliteli içerik anlamına gelmemektedir.

Çeşitlilik ölçütleri olan Type-Token Ratio (TTR) ve Unique Word Ratio metriklerinin tüm modellerde 1.0 olarak çıkması, her kelimenin yalnızca bir kez

kullanıldığı ve aşırı çeşitliliğe gidildiği anlamına gelmektedir. Bu sonuçlar, modellerin dilin doğal tekrar ve akışkanlık yapısından uzaklaştığını ve gerçek metinlerdeki sözcük dağılımına benzemediğini göstermektedir. Gerçek şiirlerde ve şarkı sözlerinde belirli tematik sözcüklerin tekrar kullanımı, duygusal vurgu açısından önemli bir yapı taşıyken; bu modellerin hepsi gereğinden fazla çeşitlilik üretmiş, bu da anlam kurma sürecini zorlaştırmıştır.

Semantik yakınlık açısından en anlamlı sonuçları yalnızca RelGAN modelinin verdiği görülmüştür. TF-IDF tabanlı Cosine Similarity değerlerine göre RelGAN'ın 0.0693'lük değeri, anlam düzeyinde referans metinlere en çok yaklaşan model olduğunu ortaya koymuştur. Diğer tüm modellerin cosine değerleri ise sıfıra eşit olup, anlam dünyası açısından ciddi bir boşluk içinde olduklarını göstermektedir. n-gram coverage değerlerinin tüm modellerde sıfır olması da yine modellerin doğal dil örüntülerini öğrenemediğini desteklemektedir.

Bununla birlikte, bu sayısal bulgular yalnızca model performansının teknik boyutunu yansıtmaktadır. Edebi üretim, teknik doğruluktan öteye geçen bir alandır ve dilin imgesel, duygusal, kültürel yönlerini de içermektedir. Bu nedenle, modeller tarafından üretilen şarkı sözü dizeleri ayrıca insan gözüyle, tematik ve yapısal bütünlük açısından değerlendirilmiştir.

Edebi analiz sonucunda, RelGAN tarafından üretilen bazı şarkı sözlerinin diğer modellere göre daha tutarlı ve anlamlı bir atmosfer sunduğu gözlemlenmiştir. Örneğin:

*“aşk fakültesini ikimiz öpüldükçe payından
attın dođruyu yitiriyor içersin veriyorum
dökül sarmış hergün tılsımını demezdim
hayatımda sevildikçe kalkıyor söyledim istemiyorsan”*

Bu dizeler tematik bir birliktelik taşımakta, sevgi, kayıp, aidiyet ve pişmanlık gibi güçlü duyguları yansıtan kelimelerle örülmüştür. Elbette bazı sözdizimsel kırılmalar gözlenmekte, ancak şiirsel atmosfer ve metafor üretimi açısından diğer modellere göre belirgin bir başarı ortaya konmuştur. Bu örnek, yapay zekâ modelinin imgesel düşünceye ne kadar yaklaşabildiğini göstermesi bakımından önemlidir. Öte yandan SeqGAN ve TextGAN gibi modellerin çıktıları, sıkça dilin anlamlı yapısından sapan ve kelime yığınları halinde örgütsüz dizeler içermiştir. Örneğin:

*“bostan varamadım elimizde durana aklın
insanım toprağına şafaklarda oluruz ellerini*

*yarın durdular bacı kahramandır yanağın
götür çıra çıra melekti sevdamıza”*

Bu dizelerde imgeler hem birbirinden kopuk hem de içsel bir akış sunmaktan uzaktır. Belirli kelimeler arasındaki ilişki rastlantısal bir şekilde kurulmuş gibi görünmektedir. Bu durum, düşük BLEU ve Jaccard skorları ile de uyumludur.

Sonuç olarak, hem teknik metrikler hem de edebi değerlendirme dikkate alındığında, en başarılı model RelGAN olarak öne çıkmıştır. Hem anlamsal benzerlik hem de şiirsel yoğunluk bakımından diğer modellere üstünlük sağlamıştır. MaliGAN ise dil modelleme başarımı açısından düşük perplexity değeri ile dikkat çekmiş, ancak estetik üretim başarımı sınırlı kalmıştır.

Bu tez çalışması, şarkı sözü metin üretiminde GAN tabanlı modellerin sınırlılıklarını ortaya koyarken aynı zamanda yeni nesil mimarilerin bu alandaki potansiyelini de gözler önüne sermektedir. Özellikle Transformer tabanlı modellerin, daha uzun bağlam ilişkilerini kavrayabilme yetenekleri sayesinde şarkı sözü üretimi gibi karmaşık dil görevlerinde daha başarılı olacağı öngörülmektedir. Türkçe gibi yapısal olarak zengin, agglutinatif dillerde bu tür modellerin kullanımı, anlamlı ve sanatsal metin üretimi açısından gelecekte daha umut verici sonuçlar doğuracaktır.

5.1.1. Çalışmanın Eksiklikleri

1. Veri Kümesinin Sınırlılığı: Kullanılan şarkı sözü veri kümesi sınırlı tematik çeşitliliğe sahiptir ve yalnızca belirli bir tür (örneğin halk şarkısı ya da popüler müzik) üzerinden seçilmiştir. Bu durum, modellerin dilin farklı poetik yapılarıyla karşılaşmasını engellemiş olabilir.
2. Yüzeysel Metriklere Aşırı Bağlılık: BLEU, Jaccard ve Perplexity gibi metrikler kullanılmış olsa da, bu metrikler şiirsel kaliteyi tam olarak yansıtamamaktadır. Estetik, duygu aktarımı ve imgeler gibi kavramlar bu tür metriklerle ölçülememektedir.
3. Uzun Bağlamların Kullanılamaması: Modellerin çoğu kısa dizeler üretmeye uygun yapılarla eğitilmiştir. Daha uzun şarkı sözü metinlerinde gerekli olan bağlamsal bütünlük ve anlatı derinliği sağlanamamıştır.
4. Transformer Tabanlı Modellerin Dahil Edilmemesi: Çalışma yalnızca GAN temelli mimarilere odaklanmış, son yıllarda dil üretiminde büyük başarı gösteren

Transformer tabanlı (örneğin GPT, BART) modellerle karşılaştırma yapılmamıştır.

5. Edebi Değerlendirmelerin Sayısal Olmaması: Şarkı örneklerinin edebi analizi yapılmış olsa da bu analizler nicel olarak kodlanmamış, nesnel ölçütlerle değerlendirilmemiştir. İnsan değerlendirmelerinin daha sistematik hale getirilmesi gerekirdi.
6. Aşırı Sözcük Çeşitliliği ve Tutarlılık: TTR ve Unique Word Ratio değerlerinin 1.0 çıkması, aşırı çeşitlilik ve anlamsal tutarlılık sorunu yaratmıştır. Bu durum modellerin denetimsiz şekilde kelime üretmesine neden olmuştur.
7. Duygu ve Anlam Yoğunluğu Analizi Yapılmamıştır: Üretilen dizelerin hangi duygu spektrumunda yer aldığı (örneğin hüznü, neşeli, öfkeli) analiz edilmemiş, anlam derinliği açısından sınıflandırma yapılmamıştır.
8. Geniş Ölçekli İnsan Değerlendirmesi Eksikliği: Modellerin çıktıları uzman edebiyatçılar veya şairler tarafından değerlendirilememiştir. Bu tür uzman değerlendirmeleri, Şarkı sözü niteliğinin doğruluğunu güçlendirebilirdi.
9. Gerçek Şarkı Sözü Yapılarıyla Karşılaştırma Yetersizliği: Üretilen şarkıların yapısal olarak gerçek şarkı türleriyle (pop, arabesk, serbest) biçimsel karşılaştırmaları yapılmamıştır.
10. Metinlerde Tematik Tutarlılığın Ölçülmemesi: Her dizinin benzer tema içinde kalıp kalmadığı değerlendirilmemiş, anlam bütünlüğü ölçülmemiştir.

5.2 Öneriler

5.2.1. Gelecek Çalışmalar İçin Öneriler

1. Transformer Tabanlı Modellerin Entegrasyonu: GPT, BART veya T5 gibi modern dil modelleri ile GAN tabanlı yaklaşımların karşılaştırmalı analizi yapılarak şarkı üretiminde anlam derinliği ve bağlamsal tutarlılık artırılabilir.
2. Anlamsal Değerlendirme İçin Gelişmiş Ölçütlerin Kullanılması: Cosine similarity yerine BERTScore, SBERT benzerliği, semantic coherence ve entailment tabanlı metrikler gibi derin anlam analizine dayalı ölçütlerin uygulanması önerilmektedir.
3. İnsan Temelli Çoklu Değerlendirme Süreci: Edebi değerlendirme için farklı yaş, meslek ve edebi birikimden gelen katılımcılarla çoklu yorumlayıcı analizler

(expert judgment) yapılarak şarkıların estetik değeri daha güvenilir şekilde ölçülebilir.

4. Daha Zengin ve Tematik Veri Setleri ile Eğitim: Farklı şarkı türlerini (halk, arabesk, pop vb.) içeren geniş ve tematik olarak etiketlenmiş veri setleri oluşturularak modelin içerik üretme çeşitliliği artırılabilir.
5. Türkçeye Özgü Dil Özelliklerinin Dahil Edilmesi: Morfolojik analiz ve kök tabanlı dil işleme teknikleriyle Türkçeye özgü yapıların modellenmesi, şarkıların daha doğal ve tutarlı olmasına katkı sağlayacaktır.
6. Anlamsızlık ve Tutarsızlık Filtreleme Modülü Geliştirilmesi: Üretilen dizelerde anlamsız, bağlamsız ya da tekrar eden kalıpları otomatik tespit eden bir filtreleme sistemi ile çıktı kalitesi iyileştirilebilir.
7. Uzun Şarkı Yapıları İçin Bağlam Takibi: 4–8 satırlık şarkı üretimi yerine uzun şarkılar üretmek için context-aware decoder yapıları veya attention temelli yapıların adaptasyonu önerilmektedir.
8. Gerçek Zamanlı Uygulamalar İçin Optimize Edilmiş Modeller: Hafifletilmiş modellerin (TinyGAN, Distil-GAN, MobileBERT gibi) mobil cihazlarda gerçek zamanlı şarkı üretimi yapabilecek şekilde optimize edilmesi hedeflenebilir.
9. Yaratıcılık ve Duygu Analizi Üzerine Odaklanma: Gelecek çalışmalarda üretilen dizelerin yaratıcı ifade gücü ve duygusal etkisi (affective computing) üzerine özel analizler yapılması önerilmektedir.
10. Etkileşimli Şarkı Üretim Sistemlerinin Geliştirilmesi: Kullanıcının başlangıç kelimesi, duygu durumu veya temaya göre yönlendirme yapabileceği interaktif bir şarkı üretim arayüzü geliştirilebilir.

KAYNAKLAR

- Bag, S., Kumar, S. K., & Tiwari, M. K. (2019). An efficient recommendation generation using relevant Jaccard similarity. *Information Sciences*, 483, 53–64. <https://doi.org/10.1016/j.ins.2018.09.065>
- Boden, M. A. (2016). *Artificial intelligence: A very short introduction*. Oxford University Press.
- Briot, J. P., Hadjeres, G., & Pachet, F. (2020). *Deep learning techniques for music generation*. Springer.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Che, T., Li, Y., Zhang, R., Hjelm, R. D., Li, W., Song, Y., & Bengio, Y. (2017). Maximum-likelihood augmented discrete generative adversarial networks. *arXiv preprint arXiv:1702.07983*. <https://arxiv.org/abs/1702.07983>
- Chollampatt, S., Pham, M. Q., Indurthi, S. R., & Turchi, M. (2025, January). Cross-lingual evaluation of multilingual text generation. In *Proceedings of the 31st International Conference on Computational Linguistics* (pp. 7766–7777).
- Colton, S., Wiggins, G. A., & Guckelsberger, C. (2015). Computational creativity: The final frontier? In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 46–52).
- De Novais, E. M., Dias Tadeu, T., & Paraboni, I. (2010). Improved text generation using n-gram statistics. In *Advances in Artificial Intelligence–IBERAMIA 2010* (pp. 316–325). Springer.
- Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). *CAN: Creative adversarial networks, generating “art” by learning about styles and deviating from style norms* [arXiv preprint]. *arXiv*. <https://arxiv.org/abs/1706.07068>.
- Garbacea, G. C. (2023). *Neural language generation for content adaptation: Explainable, efficient low-resource text simplification and evaluation* (Doctoral dissertation).
- Ghazvininejad, M., Shi, X., Priyadarshi, J., & Knight, K. (2019). Hafez: An interactive poetry generation system. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 43–48).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.

- Graves, A. (2013). Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850. <https://arxiv.org/abs/1308.0850>.
- Guo, J., Lu, S., Cai, H., Zhang, W., Yu, Y., & Wang, J. (2018, April). Long text generation via adversarial training with leaked information. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).
- Herremans, D., Chuan, C. H., & Chew, E. (2017). A functional taxonomy of music generation systems. *ACM Computing Surveys*, 50(5), 1–30.
- Huang, C. Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., ... & Eck, D. (2020). Music transformer: Generating music with long-term structure. In Proceedings of the 7th International Conference on Learning Representations (ICLR).
- Imperial, J. M., & Madabushi, H. T. (2022). Uniform complexity for text generation. arXiv preprint arXiv:2204.05185. <https://arxiv.org/abs/2204.05185>.
- Iqbal, T., & Qureshi, S. (2022). The survey: Text generation models in deep learning. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 2515–2528.
- Juefei-Xu, F., Dey, R., Boddeti, V. N., & Savvides, M. (2019). RankGAN: A maximum margin ranking GAN for generating faces. In *Computer Vision – ACCV 2018* (pp. 3–18). Springer.
- Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.). Pearson.
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4401–4410).
- Kaufmann, W. (1967). Musical references in the Chinese classics. *Information Coordinators*.
- Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. Dutton/Penguin.
- Manjavacas, E., Karsdorp, F., & Kestemont, M. (2021). Synthetic literature: Writing science fiction in the style of William Gibson. *Digital Scholarship in the Humanities*, 36(1), 66–86.
- Matlach, V., Krivochen, D. G., & Milička, J. (2021). A method for the comparison of general sequences via type-token ratio. In *Language and Text* (pp. 37–54). John Benjamins.
- Nie, W., Narodytska, N., & Patel, A. (2018, September). RelGAN: Relational generative adversarial networks for text generation. In *International Conference on Learning Representations*.

- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8), 1–7.
- Riedl, M. O. (2014). The Lovelace 2.0 test of artificial creativity and intelligence. *arXiv preprint arXiv:1410.6142*. <https://arxiv.org/abs/1410.6142>.
- Shahriar, S. (2022). GAN computers generate arts? A survey on visual arts, music, and literary text generation using generative adversarial network. *Displays*, 73, 102237.
- Sturm, B. L., Iglesias, M., Ben-Tal, O., & Korshunova, I. (2019). Artificial intelligence and music: Open questions of copyright law and engineering praxis. *Arts*, 8(3), 115.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)* (pp. 5998–6008).
- Vychegzhanin, S., Kotelnikova, A., Sergeev, A., & Kotelnikov, E. (2023, September). Controllable story generation based on perplexity minimization. In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 154–169). Springer Nature.
- Wu, Y., & Wang, J. (2020). Text generation service model based on truth-guided SeqGAN. *IEEE Access*, 8, 11880–11886.
- Xu, J., Ren, X., Lin, J., & Sun, X. (2018). Diversity-promoting GAN: A cross-entropy based generative adversarial network for diversified text generation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 3940–3949).
- Yu, L., Zhang, W., Wang, J., & Yu, Y. (2017). SeqGAN: Sequence generative adversarial nets with policy gradient. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)* (pp. 2852–2858).
- Zhang, Y., Gan, Z., Fan, K., Chen, Z., Henao, R., Shen, D., & Carin, L. (2017, July). Adversarial feature matching for text generation. In *International Conference on Machine Learning* (pp. 4006–4015). PMLR.
- Zhu, X., Su, S., Fu, M., Liu, J., Zhu, L., Yang, W., ... & Guo, Y. (2018). A cosine similarity algorithm method for fast and accurate monitoring of dynamic droplet generation processes. *Scientific Reports*, 8(1), 9967.