

Yapay Zekâ Tabanlı Sistemler

Teori, Uygulama ve Gelecek Perspektifleri-2

Editör
EYYÜP GÜLBANDILAR



BİDGE Yayınları

Yapay Zekâ Tabanlı Sistemler: Teori, Uygulama ve Gelecek Perspektifleri-2

Editör: EYYÜP GÜLBANDILAR

ISBN: 978-625-8567-79-3

1. Baskı

Sayfa Düzeni: Gözde YÜCEL

Yayınlama Tarihi: 2025-12-25

BİDGE Yayınları

Bu eserin bütün hakları saklıdır. Kaynak gösterilerek tanıtım için yapılacak kısa alıntılar dışında yayıncının ve editörün yazılı izni olmaksızın hiçbir yolla çoğaltılamaz.

Sertifika No: 71374

Yayın hakları © BİDGE Yayınları

www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com

Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.

Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya /
Ankara



ÖNSÖZ

Yapay zekâ (YZ) ve derin öğrenme paradigmaları, günümüzde teknik birer araç olmanın ötesine geçerek; sağlık bilimlerinden telekomünikasyona, enerji yönetiminden stratejik güvenlik sistemlerine kadar uzanan geniş bir spektrumda yapısal bir dönüşümün ana aktörleri haline gelmiştir. Hesaplama kapasitesindeki eksponansiyel artış ve veri madenciliğindeki metodolojik devrim, geleneksel analitik yaklaşımların yetersiz kaldığı lineer olmayan ve yüksek boyutlu problemlerin çözümünde YZ tabanlı modelleri vazgeçilmez kılmıştır.

Elinizdeki bu çalışma, yapay zekânın teorik müktesebatını ve uygulama sahasındaki güncel izdüşümlerini disiplinler arası bir sentez ile sunmayı gaye edinmektedir. Eserde; tıbbi görüntüleme üzerinden gerçekleştirilen onkolojik teşhis modellemelerinden biyometrik tabanlı demografik analizlere, yazılım güvenliği mimarilerinden karmaşık zaman serisi kestirimlerine kadar uzanan çeşitlilik, YZ'nin operasyonel esnekliğini somut verilerle tevsik etmektedir.

Kitabın ayırt edici niteliklerinden biri de sentetik veri üretimi, IoT entegreli kestirimci bakım stratejileri ve 6G ağ ekosistemlerinde konumlandırılan dağıtık yapay zekâ mimarileri gibi teknolojik öngörü gerektiren başlıkları derinlemesine işlemesidir. Kablosuz ağlarda deneyim kalitesi (QoE) gibi kritik parametrelerin incelenmesi, okuyucuya yalnızca mevcut teknolojileri değil, aynı zamanda yarının araştırma gündemini de ihtiva eden vizyoner bir perspektif sunmaktadır.

Akademik bir titizlikle hazırlanan bu eser; lisansüstü araştırmacılardan sektör paydaşlarına, sistem tasarımcısı mühendislerden kuramsal çalışma yürüten bilim insanlarına kadar geniş bir kitleye hitap eden bir başvuru kaynağıdır. Eserin, yapay zekânın çok katmanlı doğasının anlaşılmasına hizmet etmesini, farklı disiplinler arasında yeni korelasyonlar kurulmasına zemin hazırlamasını ve bilim dünyasına özgün katkılar sunmasını temenni ederiz.

Dr.Eyyüp GÜLBANDILAR

Editör

İÇİNDEKİLER

ÇEVİK YAZILIM GELİŞTİRMENİN TEMELLERİ	1
<i>BURCU DURMUŞ</i>	

KABLOSUZ AĞLARDA DENEYİM KALİTESİ (QOE): KAVRAMSAL ÇERÇEVE, ÖLÇÜM METODOLOJİLERİ VE YÖNETİM STRATEJİLERİ	24
<i>SEMİH KAHVECİ, RAMAZAN AKKURT</i>	

YAZILIM GÜVENLİK AÇIKLARININ YAPAY ZEKA TABANLI TESPİTİ	42
<i>GÖZDE MİHRAN KAYA, İRFAN KÖSESÖY</i>	

TIBBİ GÖRÜNTÜLERDE DERİN ÖĞRENME UYGULAMALARI: BEYİN TÜMÖRÜ MRI VERİLERİ ÜZERİNDEN BİR İNCELEME	62
<i>YUSUF ÇELİK</i>	

AKILLI TARIM UYGULAMALARI: SENSÖR VE VERİ ANALİTİĞİ İLE BİTKİ VERİMLİLİĞİNİN ANALİZİ	82
<i>HÜSEYİN TÜRKMEN, İNAYET HAKKI ÇİZMECİ</i>	

CUZN ALAŞIMLARININ ÖZDİRENCİNE SODYUM SİTRAT ETKİSİ: İSTATİSTİKSEL BİR ANALİZ	106
<i>HASAN GÜLER, RASİM ÖZDEMİR</i>	

DOĞAL DİL İŞLEMENİN EVRİMİ: VEKTÖRLERDEN AJANLARA	126
<i>MURAT AYDOĞAN</i>	

TÜRKÇE HABER METİNLERİNDE METİN SINIFLANDIRMA YAKLAŞIMLARININ KRONOLOJİK KARŞILAŞTIRMASI	150
<i>ŞÜKRÜ EVRAN, RECEP ÖZDAĞ</i>	

KÜLTÜREL VERİ MADENCİLİĞİ VE TÜRKÇE POP MÜZİKTE TEMATİK ÖRÜNTÜ ÇIKARIMI: MABEL MATİZ DİSKOGRAFİSİ ÜZERİNE BİR MODEL ÖNERİSİ ...	175
<i>ARZUM KARATAS</i>	

AKADEMİK YAZIM İÇİN REFERANS YÖNETİM YAZILIMLARININ KARŞILAŞTIRMALI ANALİZİ	186
--	-----

FURKAN ENES DAĞLI, AYHAN İSTANBULLU, YİĞİT KALAY, ŞİRİN GÜLTEN ÇELİK

KONUŞMA DUYGU TANIMA İÇİN ÖZELLİK ANALİZİ	213
--	-----

KENAN DONUK

SANAL AGORA: SOSYAL VR TABANLI ÖĞRENME VE MÜZAKERE ALANLARI	236
--	-----

EBRU YILMAZ İNCE, MURAT İNCE

YAPAY ZEKÂDA VERİ KİTLİĞİNE DAYANIKLI ÖĞRENME PARADİGMALARI: ÖRNEKSİZ ÖĞRENME, AZ ÖRNEKLE ÖĞRENME VE SENTEZLENMİŞ VERİ YAKLAŞIMLARI	243
--	-----

AYŞE GÜL EKER

DİJİTAL İKİZ TEKNOLOJİSİ: KAVRAM, MİMARİ VE DİSİPLİNLERARASI UYGULAMALAR	260
---	-----

RÜYA AKINCI

HASTA YORUMLARINI KULLANARAK DUYGU ANALİZİ	277
---	-----

İPEK AKPINAR, ALİ METE ÇIPLAK, EDİZ ENGİN, ENES ÇAKMAK, MAHMUTCAN SA

BÖLÜM 1

ÇEVİK YAZILIM GELİŞTİRMENİN TEMELLERİ

BURCU DURMUŞ¹

Giriş

Yazılım sistemleri, günümüzde yalnızca teknik çözümler sunan araçlar olmaktan çıkarak ekonomik, sosyal ve kurumsal yapıların merkezinde yer alan stratejik bileşenler hâline gelmiştir. Dijital dönüşüm, büyük veri, bulut bilişim ve yapay zekâ gibi teknolojik gelişmeler, yazılım projelerinin hem kapsamını hem de karmaşıklığını önemli ölçüde artırmıştır. Bu bağlamda yazılım geliştirme süreçlerinin, değişen gereksinimlere hızlı uyum sağlayabilen, esnek ve sürdürülebilir yaklaşımlar çerçevesinde ele alınması zorunlu hâle gelmiştir (Sommerville, 2020).

Doğrusal ve plan odaklı gibi geleneksel yazılım geliştirme yaklaşımları, gereksinimlerin proje başlangıcında eksiksiz ve değişmez biçimde tanımlanabileceği varsayımına dayanır. Ancak pratikte yazılım projelerinin büyük bir kısmında bu varsayım geçerliliğini yitirmektedir. Kullanıcı beklentilerinin proje süresince değişmesi, pazar koşullarının hızlı evrilmesi ve teknolojik belirsizlikler katı süreç modellerinin başarısını sınırlamaktadır

¹ Öğr. Gör. Dr., Tekirdağ Namık Kemal Üniversitesi, Rektörlük, Orcid: 0000-0002-0298-0802

(Pressman ve Maxim, 2020). Bu durum, yazılım projelerinde zaman ve maliyet aşırımları ile kalite problemlerinin ortaya çıkmasına neden olmaktadır.

Çevik yazılım geliştirme yaklaşımları, bu sorunlara bir yanıt olarak ortaya çıkmış ve yazılım mühendisliğinde önemli bir paradigma değişimini temsil etmiştir. Çevik yaklaşımlar, değişimi kaçınılmaz bir risk olarak değil, yönetilmesi gereken doğal bir unsur olarak ele alır. Kısa geliştirme döngüleri, sürekli geri bildirim ve müşteriyle yakın iş birliği, çevik yazılım geliştirmenin temel yapı taşlarını oluşturmaktadır (Beck vd., 2001).

Bu kitap bölümünün amacı; çevik yazılım geliştirmenin kuramsal temellerini, uygulamada yaygın olarak kullanılan yöntemleri, tasarım ve test yaklaşımlarını ve güncel yapay zekâ destekli çevik süreçleri bütüncül bir bakış açısıyla ele almaktır.

Çevik Yazılım Geliştirmenin Ortaya Çıkışı ve Agile Manifesto

Geleneksel Yazılım Geliştirme Yaklaşımlarının Sınırları

Yazılım mühendisliğinin erken dönemlerinde geliştirilen geleneksel süreç modelleri, yazılım geliştirmeyi ardışık ve doğrusal bir süreç olarak ele almıştır. Özellikle şelale modeli; gereksinim analizi, sistem tasarımı, kodlama, test ve bakım aşamalarının birbirini takip eden safhalar hâlinde yürütülmesini öngörür (Royce, 1970). Bu yaklaşım, belgelendirme ve süreç kontrolü açısından belirli avantajlar sunmasına rağmen değişime kapalı yapısı nedeniyle eleştirilmiştir.

Yazılım projelerinde gereksinimlerin zaman içerisinde değişmesi, doğrusal modellerin en temel zayıflıklarından biridir. Bir safhada yapılan hataların veya eksiklerin, sonraki aşamalarda düzeltilmesi yüksek maliyetlere yol açmaktadır. Ayrıca kullanıcıların yazılımı ancak projenin son aşamalarında

deneyimleyebilmesi, geri bildirimin geç alınmasına ve beklentilerin karşılanmamasına neden olmaktadır (Boehm, 2002).

Bu sorunlar, literatürde ‘yazılım krizi’ olarak adlandırılan durumu gündeme getirmiştir. Projelerin önemli bir kısmının zamanında tamamlanamaması, bütçe aşımaları ve kullanıcı memnuniyetsizliği, yazılım geliştirme süreçlerinin yeniden düşünülmesini zorunlu kılmıştır (Standish Group, 2020).

Çevik Manifesto’nun Ortaya Çıkışı

2001 yılında Kent Beck ve çalışma arkadaşları tarafından yayımlanan Çevik Manifesto, yazılım geliştirme süreçlerine insan merkezli bir bakış açısı kazandırmıştır. Manifesto, yazılım geliştirmede dört temel değeri ön plana çıkarmaktadır (Beck vd., 2001):

- Bireyler ve etkileşimler, süreçler ve araçlardan önce gelir.
- Çalışan yazılım, kapsamlı dokümantasyondan daha değerlidir.
- Müşteri ile iş birliği, sözleşme pazarlıklarından daha önemlidir.
- Değişime yanıt vermek, bir planı takip etmekten daha değerlidir.

Bu değerler, geleneksel yaklaşımların katı süreç anlayışına karşı esnek, uyarlanabilir ve iletişim odaklı bir geliştirme felsefesini temsil etmektedir. Manifestoda ayrıca, bu değerleri somutlaştıran on iki ilke yer almaktadır:

1. Müşteri memnuniyeti için kullanılabilir yazılımın hızlı geliştirilmesi.
2. Değişen gereksinimlere uyum sağlama.
3. Çalışan bir yazılımın sürekli olarak teslim edilmesi.
4. İşbirlikçi iş ve geliştirme süreçleri.
5. Motiveli bireylerin desteklenmesi.

6. Sözlü iletişim önceliği.
7. Çalışan yazılım ilerlemenin birincil göstergesidir.
8. Sürdürülebilir geliştirme sağlanması.
9. Teknik iyileştirme ve iyi tasarımın sürekli teşvik edilmesi.
10. Basitliğe odaklanma.
11. Öz-organize ekiplerin en iyi çözümleri geliştirmesi.
12. Düzenli entegrasyon ve süreç iyileştirme.

Çevik Manifesto, yazılım geliştirmeyi yalnızca teknik bir faaliyet olarak değil, aynı zamanda sosyal ve kültürel bir süreç olarak ele alması bakımından yazılım mühendisliği literatüründe önemli bir dönüm noktasıdır (Highsmith, 2009).

Çevik Süreçler ve Extreme Programming (XP)

Çevik Süreç Kavramı

Çevik yazılım geliştirme, kısa iterasyonlar hâlinde çalışan yazılım üretmeyi, müşteri geri bildirimlerini sürecin merkezine almayı ve sürekli iyileştirmeyi esas alan bir yaklaşımdır. Çevik süreçlerde yazılım geliştirme, önceden detaylı biçimde tanımlanmış sabit bir planı izlemek yerine, öğrenme ve uyum sağlama döngüsü olarak ele alınır (Serrador ve Pinto, 2015).

Çevik süreçlerin temel özellikleri şu şekilde özetlenebilir:

- Gereksinimlerin evrimsel olarak ele alınması
- Kısa geliştirme döngüleri (iterasyonlar)
- Sürekli müşteri geri bildirimi
- Öz-organize ve disiplinler arası takımlar

Bu özellikler, yazılım projelerinin belirsizlik ve değişkenlik içeren doğasına daha uygun bir geliştirme ortamı sunmaktadır.

Extreme Programming (XP)

Extreme Programming (XP), çevik yazılım geliştirme yöntemleri arasında en bilinen ve en teknik odaklı yaklaşımlardan biridir. Kent Beck tarafından 1990'lı yılların sonunda geliştirilen XP, özellikle yazılım kalitesini artırmayı ve değişime hızlı yanıt verebilmeyi hedeflemektedir (Beck, 2000).

XP'nin temelinde beş ana değer yer almaktadır: iletişim, sadelik, geri bildirim, cesaret ve saygı. Bu değerler, yazılım geliştirme sürecinin hem teknik hem de insani boyutlarını kapsayan bütüncül bir yaklaşım sunmaktadır.

XP'nin öne çıkan teknikleri arasında eşli programlama (pair programming), test güdümlü geliştirme (TDD), sürekli entegrasyon, küçük sürümler ve kolektif kod sahipliği yer almaktadır. Bu teknikler, yazılımın sürekli çalışır durumda kalmasını ve hataların erken aşamada tespit edilmesini sağlamaktadır (Fowler, 2018).

Extreme Programming, Scrum gibi yönetim odaklı çevik çerçevelerle birlikte kullanıldığında hem süreç yönetimi hem de teknik uygulamalar açısından güçlü bir sinerji oluşturmaktadır. Bu nedenle birçok yazılım projesinde Scrum ve XP'nin birlikte uygulandığı hibrit yaklaşımlar tercih edilmektedir (Schwaber ve Sutherland, 2020).

Çevik Proje, Sürüm ve İterasyon Planlama

Çevik Proje Planlama Yaklaşımı

Geleneksel yazılım geliştirme yaklaşımlarında proje planlama, çoğunlukla proje başlangıcında yapılan ayrıntılı zaman ve kapsam tahminlerine dayanmaktadır. Bu planlar, gereksinimlerin sabit olduğu varsayımıyla hazırlanır ve proje süresince mümkün olduğunca değişmeden uygulanması beklenir. Ancak çevik yazılım geliştirme yaklaşımı, bu varsayımın gerçekçi olmadığını kabul eder

ve proje planlamasını dinamik, evrimsel bir süreç olarak ele alır (Highsmith, 2009).

Çevik projelerde planlama, tek seferlik bir faaliyet değil, proje süresince sürekli güncellenen ve geliştirilen bir etkinliktir. Bu bağlamda planlama, ‘tahmin et–uygula–geri bildirim al–iyileştir’ döngüsü içerisinde yürütülür. Amaç, belirsizliği tamamen ortadan kaldırmak değil, belirsizlikle etkin biçimde başa çıkabilmektir (Sommerville, 2020).

Çevik proje planlamasının temelinde, geliştirilecek yazılım sisteminin vizyonu yer alır. Bu vizyon, müşterinin iş hedeflerini ve yazılımdan beklentilerini yansıtır. Gereksinimler bu vizyondan türetilir ve kullanıcı hikâyeleri biçiminde ifade edilir. Böylece planlama süreci, soyut teknik gereksinimlerden ziyade, kullanıcı değerine odaklanır.

Sürüm (Release) Planlaması

Sürüm planlaması, çevik projelerde orta vadeli planlama düzeyini temsil eder. Bir sürüm, yazılım sisteminin belirli bir zaman aralığında (genellikle birkaç ay) tamamlanan ve kullanıcı tarafından çalışır durumda kullanılabilen bir versiyonudur. Sürüm planlamasının temel amacı, hangi kullanıcı hikâyelerinin hangi sürümde yer alacağını belirlemek ve proje için bir yol haritası oluşturmaktır (Cohn, 2005).

Extreme Programming (XP) projelerinde sürüm planlama, müşteri ve geliştirici ekibin birlikte yürüttüğü etkileşimli bir süreçtir. Bu süreçte müşteri, iş değeri açısından öncelikli olan kullanıcı hikâyelerini belirlerken; geliştiriciler, bu hikâyelerin teknik uygulanabilirliği ve tahmini eforu hakkında geri bildirim sağlar. Bu yaklaşım, planlamayı tek taraflı bir karar mekanizması olmaktan çıkararak iş birliğine dayalı hâle getirir.

Sürüm planlaması, projenin gidişatını izlemek ve paydaşlara görünürlük sağlamak açısından kritik bir rol oynar. Ancak çevik yaklaşımlarda sürüm planları katı taahhütler olarak değil, değişime açık rehberler olarak değerlendirilir. Gereksinimlerde meydana gelen değişiklikler doğrultusunda sürüm planları güncellenebilir.

İterasyon Planlaması ve Planlama Pokeri

İterasyon planlaması, çevik proje planlamasının en kısa vadeli ve en ayrıntılı düzeyini oluşturur. Bir iterasyon genellikle bir veya iki haftalık bir zaman dilimini kapsar ve bu süre içerisinde tamamlanması hedeflenen kullanıcı hikâyeleri belirlenir. İterasyon planlama toplantılarında, müşteri tarafından seçilen kullanıcı hikâyeleri geliştirme takımı tarafından ayrıntılandırılır ve görevlere bölünür.

Kullanıcı hikâyelerinin uygulanması için gerekli eforun tahmin edilmesi, çevik planlamanın en önemli bileşenlerinden biridir. XP projelerinde bu amaçla yaygın olarak planlama pokeri tekniği kullanılmaktadır. Planlama pokeri, geliştiricilerin kullanıcı hikâyeleri için bağımsız tahminler yapmasını ve bu tahminler üzerinde ortak bir uzlaşuya varılmasını sağlayan etkileşimli bir yöntemdir (Grenning, 2002).

Planlama pokeri, sayısal doğruluktan ziyade, ekip üyeleri arasında ortak anlayışın oluşmasını hedefler. Farklı tahminlerin nedenleri tartışılarak kullanıcı hikâyesinin belirsiz yönleri açığa çıkarılır. Bu süreç hem daha gerçekçi tahminler yapılmasına hem de teknik risklerin erken aşamada fark edilmesine katkı sağlar.

Çevik Takımlarda İletişim ve Çalışma Ortamı

İletişimin Çevik Süreçlerdeki Rolü

Çevik yazılım geliştirme, teknik uygulamaların ötesinde güçlü bir iletişim kültürüne dayanmaktadır. Çevik yaklaşımlarda bilginin bireyler arasında hızlı ve doğru biçimde aktarılması, sürecin

başarısı açısından kritik öneme sahiptir. İletişimin zayıf olduğu ortamlarda, takım üyeleri izole hâle gelmekte ve iş birliği etkin biçimde sağlanamamaktadır (Moe vd., 2010).

Extreme Programming felsefesinde iletişim, temel değerlerden biri olarak kabul edilmektedir. Bu bağlamda çevik süreçler, yazılı dokümantasyondan ziyade yüz yüze iletişimi ve doğrudan etkileşimi teşvik eder. Amaç, yanlış anlamaları azaltmak ve sorunlara hızlı çözümler üretmektir.

Günlük Stand-up Toplantıları

Günlük stand-up toplantıları, çevik projelerde iletişimi güçlendiren temel uygulamalardan biridir. Bu toplantılar genellikle her gün aynı saatte yapılır ve 15 dakikayı geçmeyecek şekilde planlanır. Stand-up toplantılarında her takım üyesi üç temel soruya yanıt verir:

- Dünden beri ne yaptım?
- Bugün ne yapacağım?
- Karşılaştığım engeller nelerdir?

Bu yapı, takım üyelerinin birbirlerinin çalışmalarından haberdar olmasını sağlar ve olası sorunların erken aşamada görünür hâle gelmesine yardımcı olur. Stand-up toplantılarında teknik detaylara girilmemesi, toplantının odaklı ve verimli olmasını sağlar (Schwaber ve Sutherland, 2020).

Retrospektif Toplantıları ve Sürekli İyileştirme

Retrospektif toplantıları, bir iterasyonun veya sürümün ardından gerçekleştirilen değerlendirme toplantılarıdır. Bu toplantıların temel amacı, süreci iyileştirmeye yönelik çıkarımlar yapmaktır. Retrospektifler, suçlama veya bireysel performans değerlendirmesi için değil, öğrenme ve gelişim için düzenlenir (Kerth, 2001).

Çevik retrospektiflerde, ekip üyeleri güvenli bir ortamda görüşlerini paylaşabilmelidir. Bu bağlamda psikolojik güvenlik, retrospektiflerin etkinliği açısından büyük önem taşır. Ekip, nelerin iyi gittiğini, hangi sorunlarla karşılaşıldığını ve gelecekte nelerin farklı yapılabileceğini birlikte tartışır. Bu yaklaşım, çevik süreçlerin sürdürülebilirliğini destekleyen sürekli iyileştirme kültürünü besler.

Çevik Çalışma Ortamı ve Fiziksel Düzen

Çevik yazılım geliştirme, yalnızca süreçleri değil, çalışma ortamlarını da yeniden tanımlamaktadır. XP projelerinde ekip üyelerinin aynı fiziksel ortamda, birbirlerine yakın şekilde çalışması önerilmektedir. Bu düzen, yüz yüze iletişimi kolaylaştırarak bilgi paylaşımını hızlandırır.

Çevik çalışma ortamlarında beyaz panolar, hikâye kartları ve görsel araçlar önemli bir yer tutar. Kullanıcı hikâyelerinin, görevlerin ve ilerleme durumunun görsel olarak sergilenmesi, sürecin şeffaflığını artırır. Bu sayede ekip üyeleri, projenin genel durumu hakkında sürekli farkındalık sahibi olur.

Eşli programlama (pair programming) uygulaması da çevik çalışma ortamının önemli bir bileşenidir. İki geliştiricinin aynı bilgisayar üzerinde birlikte çalışması hem kod kalitesini artırmakta hem de bilgi paylaşımını teşvik etmektedir. Literatürde, eşli programlamanın hata oranlarını düşürdüğü ve takım içi öğrenmeyi desteklediği belirtilmektedir (Williams ve Kessler, 2003).

Müşteri Gereksinimleri ve Evrimsel Gelişim

Müşteri Gereksinimlerinin Çevik Yaklaşımındaki Konumu

Çevik yazılım geliştirme yaklaşımlarının en ayırt edici özelliklerinden biri, müşteri gereksinimlerini sürecin merkezine yerleştirmesidir. Geleneksel yazılım geliştirme modellerinde gereksinimler, proje başlangıcında ayrıntılı biçimde tanımlanmakta ve çoğunlukla proje süresince sabit kalacağı varsayılmaktadır. Oysa

çevik yaklaşımlar, müşterinin ne istediğini başlangıçta her yönüyle bilmesinin çoğu zaman mümkün olmadığını kabul eder. Bu nedenle gereksinimler, proje boyunca evrimsel olarak ele alınır.

Bu yaklaşımda müşteri, yazılım geliştirme sürecinin pasif bir izleyicisi değil, aktif bir paydaşıdır. Müşteri, gereksinimleri tanımlayan, önceliklendiren ve geliştirilen yazılımı düzenli olarak değerlendiren bir rol üstlenir. Bu yaklaşım, geliştirilen yazılımın iş ihtiyaçlarıyla uyumunu artırırken, yanlış veya gereksiz özelliklerin erken aşamada elenmesini sağlar.

Kullanıcı Hikâyeleri ve Gereksinimlerin İfade Edilmesi

Çevik projelerde gereksinimler, genellikle kullanıcı hikâyeleri (user stories) biçiminde ifade edilir. Kullanıcı hikâyeleri, sistemin kullanıcı perspektifinden hangi işlevleri yerine getirmesi gerektiğini kısa ve anlaşılır cümlelerle tanımlar. Bu yaklaşım, teknik ayrıntılardan ziyade kullanıcı değerine odaklanmayı teşvik eder.

Bir kullanıcı hikâyesi çoğunlukla “Bir kullanıcı olarak ... yapmak istiyorum, çünkü ...” biçiminde ifade edilir. Bu yapı hem geliştiricilerin hem de müşterilerin aynı dili konuşmasını sağlar. Kullanıcı hikâyelerinin kısa tutulması, gereksinimlerin zaman içerisinde değişebileceği gerçeğini yansıtır. Ayrıntılı teknik analizler, hikâyenin implemente edileceği iterasyon öncesinde yapılır.

Çevik yaklaşımda kullanıcı hikâyelerinin test edilebilir olması da kritik bir gerekliliktir. Her hikâye, kabul testleriyle doğrulanabilir olmalıdır. Bu durum, gereksinimlerin belirsizliğini azaltırken, yazılımın gerçekten beklenen davranışı sergileyip sergilemediğinin nesnel biçimde değerlendirilmesine olanak tanır.

Keşif Safhası ve Evrimsel Gereksinim Yönetimi

Tipik bir XP projesi, keşif safhası ile başlar. Bu safhada amaç; müşterinin iş ihtiyaçlarını anlamak, temel kullanıcı hikâyelerini belirlemek ve kullanılacak teknolojilere ilişkin genel bir

çerçeve oluşturmaktır. Ancak bu aşama, geleneksel modellerdeki ayrıntılı gereksinim analiziyle karıştırılmamalıdır. Çevik yaklaşımlarda hedef, tüm gereksinimleri baştan eksiksiz tanımlamak değil, en önemli gereksinimleri belirleyerek geliştirmeye başlamaktır.

Evrimsel gereksinim yönetimi, çevik süreçlerin bel kemiğini oluşturur. Geliştirilen her iterasyon, müşteriye somut bir ürün artımı sunar. Müşteri, bu artımı deneyimleyerek geri bildirim verir ve bu geri bildirimler bir sonraki iterasyonda dikkate alınır. Bu döngü, gereksinimlerin sürekli olarak rafine edilmesini sağlar.

Bu yaklaşım, yazılım geliştirmeyi doğrusal bir süreçten ziyade, öğrenmeye dayalı bir dönüşüm süreci olarak ele alır. Müşteri gereksinimlerinin yazılıma dönüştürülmesi, tek seferlik bir ‘çeviri’ değil, proje boyunca devam eden bir etkileşim sürecidir.

İterasyonlar, Kabul Testleri ve Müşteri Onayı

Çevik projelerde her iterasyon sonunda geliştirilen özellikler, müşteri tarafından değerlendirilir. Bu değerlendirme sürecinde kabul testleri (acceptance tests) önemli bir rol oynar. Kabul testleri, sistemin kullanıcı bakış açısından doğru çalışıp çalışmadığını kontrol eder ve gereksinimlerin karşılandığını doğrular.

Müşterinin kabul testlerini tanımlamaya aktif biçimde katılması, yazılımın beklentilerle uyumlu olmasını sağlar. Kabul testleri, geliştirici ekibin varsayımlarını sınırlandırarak, yazılımın gerçek kullanım senaryolarına göre şekillenmesine katkı sunar. Bu durum, çevik yazılım geliştirmenin müşteri odaklı doğasını pekiştiren önemli bir mekanizmadır (Boehm ve Turner, 2004).

Çevik Yazılımda Teknik Altyapı ve Araçlar

Teknik Altyapının Çevik Süreçlerdeki Önemi

Çevik yazılım geliştirme, insan ve iletişim odaklı bir yaklaşım olmakla birlikte, uygun teknik altyapı olmaksızın

sürdürülebilir değildir. Sürekli değişen kod tabanı, sık entegrasyonlar ve hızlı geri bildirim gereksinimi, otomasyon ve araç desteğini zorunlu kılmaktadır. Bu nedenle çevik süreçlerde kullanılan teknik araçlar, yalnızca yardımcı unsurlar değil, sürecin ayrılmaz parçalarıdır.

Çevik projelerde teknik altyapı, mümkün olduğunca sade fakat genişletilebilir olacak şekilde tasarlanmalıdır. Projenin başlangıcında aşırı karmaşık araç setleri kullanmak yerine, temel ihtiyaçları karşılayan araçlarla başlanması ve ihtiyaçlar doğrultusunda altyapının evrimsel olarak geliştirilmesi önerilmektedir.

Versiyon Kontrol Sistemleri

Versiyon kontrol sistemleri, çevik yazılım geliştirme süreçlerinin temel bileşenlerinden biridir. Birden fazla geliştiricinin aynı kod tabanı üzerinde eş zamanlı çalışabilmesi, ancak etkili bir versiyon kontrol mekanizmasıyla mümkün olmaktadır. Versiyon kontrol sistemleri, kod değişikliklerinin izlenmesini, geri alınmasını ve birleştirilmesini sağlar (Spinellis, 2012).

Çevik projelerde geliştiricilerin sık aralıklarla kodlarını merkezi bir depoya göndermesi teşvik edilir. Bu yaklaşım, entegrasyon sorunlarının erken aşamada ortaya çıkmasına yardımcı olur. Açık kaynaklı sistemlerden kurumsal çözümlere kadar geniş bir araç yelpazesi bulunmakla birlikte, çevik felsefe açısından önemli olan aracın kendisi değil, nasıl kullanıldığıdır.

Derleme ve Yapılandırma Araçları

Çevik yazılım geliştirme süreçlerinde, kodun manuel olarak derlenmesi ve paketlenmesi sürdürülebilir değildir. Bu nedenle otomatik derleme (build automation) araçları yaygın olarak kullanılmaktadır. Derleme araçları, kaynak kodun derlenmesi,

testlerin alıřtırılması ve alıřtırılabilir paketlerin oluřturulmasını otomatikleřtirir.

Otomatik derleme sreci, geliřtiricilerin zamanını tekrarlayan manuel iřlemlerden kurtarıırken, hata olasılığını da azaltır. Ayrıca her derleme srecinin aynı adımları izlemesi, ortamdan kaynaklanan tutarsızlıkların nne geer. Bu durum, zellikle evik projelerde sık yapılan entegrasyonlar aısından byk nem tařır (Humble ve Farley, 2010).

evik Ara Ekosistemi ve Entegrasyon

evik yazılım geliřtirme, tek bir aratan ziyade, birbiriyle entegre alıřan aralardan oluřan bir ekosistemi gerektirir. Proje ynetimi, srm kontrol, derleme, test ve dağıtım aralarının uyumlu biimde alıřması, srecin etkinliğini artırır.

Bu entegrasyon, evik ekiplerin yazılımın durumunu anlık olarak izleyebilmesine olanak tanır. Kodun derlenip derlenmediğı, testlerin geip gemediğı ve sistemin dağıtıma hazır olup olmadığı, řeffaf biimde grlebilir. Bu grnrlk, evik srelerin temel ilkelerinden biri olan geri bildirimin hızlanmasına katkı sağılar (Kim vd., 2016).

evik Tasarım: Test Edilebilirlik, SOLID ve Evrimsel Tasarım

evik Tasarım Anlayışı

Geleneksel yazılım geliřtirme yaklařımlarında tasarım, oğunlukla geliřtirme srecinin bařında tamamlanan, ayrıntılı ve kapsamlı bir faaliyet olarak ele alınmaktadır. Bu yaklařımda ama, yazılımın tm mimari ve tasarımsal kararlarını nceden belirlemek ve geliřtirme srecini bu plana sadık kalarak yrtmektir. Ancak evik yazılım geliřtirme yaklařımları, bu anlayışın değıřken gereksinimlere sahip projelerde srdrlebilir olmadığını savunur.

evik tasarım, yazılım tasarımını statik bir ıktı olarak değıl, geliřtirme sreci boyunca evrilen bir yapı olarak ele alır. Bu

yaklaşım, literatürde evrimsel tasarım (evolutionary design) olarak da adlandırılmaktadır. Evrimsel tasarımda amaç, yazılımı başlangıçta mümkün olduğunca sade tutmak, ancak yeni gereksinimler ortaya çıktıkça tasarımı kontrollü biçimde geliştirmektir (Fowler, 2018).

Çevik tasarım anlayışı, gereksiz karmaşıklıktan kaçınmayı, yalın çözümler üretmeyi ve yazılımın değişime açık olmasını temel hedefler arasında görür. Bu bağlamda ‘en basit çözüm’ ilkesi, çevik tasarımın temel taşlarından biridir.

Test Edilebilirlik Kavramı

Test edilebilirlik, bir yazılım sisteminin test edilme kolaylığını ifade eden önemli bir kalite özelliğidir. Çevik yazılım geliştirme süreçlerinde testler, geliştirme faaliyetinin ayrılmaz bir parçası olarak ele alındığından, test edilebilirlik kavramı kritik bir rol oynamaktadır. Test edilebilirliği düşük olan sistemlerde, testlerin yazılması ve sürdürülmesi zorlaşmakta ve bu durum çevik süreçlerin hızını ve etkinliğini olumsuz etkilemektedir (Binder, 2020).

Test edilebilir bir tasarım, bileşenler arasındaki bağımlılıkların gevşek (loosely coupled) olmasını, sorumlulukların net biçimde ayrılmasını ve dış bağımlılıkların kolayca taklit edilebilmesini gerektirir. Bu bağlamda bağımlılık enjeksiyonu, arayüz kullanımı ve modüler tasarım gibi yaklaşımlar, test edilebilirliği artıran temel tasarım teknikleri arasında yer almaktadır.

Çevik projelerde test edilebilirlik, yalnızca teknik bir hedef değil, aynı zamanda stratejik bir gerekliliktir. Sürekli entegrasyon ve test güdümlü geliştirme gibi uygulamalar, test edilebilir tasarımların varlığını zorunlu kılmaktadır.

SOLID Prensipleri ve Çevik Tasarım

SOLID prensipleri, nesne yönelimli yazılım tasarımında sürdürülebilir ve esnek sistemler geliştirmek amacıyla önerilen beş

temel ilkeden oluşmaktadır. Bu ilkeler; tek sorumluluk, açık-kapalı, Liskov yerine geçme, arayüz ayrımı ve bağımlılıkların tersine çevrilmesi prensiplerini kapsamaktadır (Martin, 2018).

Çevik yazılım geliştirme bağlamında SOLID prensipleri, değişime dirençli değil, değişimi kolaylaştıran bir tasarım anlayışını destekler. Özellikle Tek Sorumluluk Prensibi (SRP), çevik projelerde sık karşılaşılan gereksinim değişikliklerine hızlı uyum sağlanmasına katkı sunar. Benzer şekilde Açık-Kapalı Prensibi (OCP), mevcut kodu değiştirmeden yeni davranışların eklenebilmesini mümkün kılar.

SOLID prensiplerinin çevik projelerde uygulanması, yazılımın bakım maliyetlerini düşürmekte ve refactoring faaliyetlerini daha güvenli hâle getirmektedir. Ancak bu prensiplerin aşırı ve bağlamdan kopuk biçimde uygulanması, gereksiz karmaşıklığa yol açabilmektedir.

Evrimsel Tasarım ve Refactoring

Evrimsel tasarım yaklaşımında yazılım, gereksinimler doğrultusunda adım adım geliştirilmektedir. Bu süreçte refactoring, yani kodun dış davranışını değiştirmeden iç yapısının iyileştirilmesi, merkezi bir rol oynamaktadır. Refactoring faaliyetleri, tasarımın zaman içerisinde bozulmasını engelleyerek yazılımın uzun vadede sürdürülebilir olmasını sağlar.

Çevik projelerde refactoring, genellikle her iterasyonun doğal bir parçası olarak ele alınır. Yeni özelliklerin eklenmesi sırasında ortaya çıkan tasarım sorunları, küçük ve kontrollü refactoring adımlarıyla giderilir. Bu yaklaşım, büyük ve riskli yeniden yazım projelerinin önüne geçer.

Birim Testleri, Test Gdml Geliřtirme (TDD) ve Refactoring

Birim Testlerinin evik Yazılımdaki Yeri

Birim testleri, yazılım sisteminin en kk test edilebilir bileřenlerinin doėruluėunu sınamak amacıyla yazılan otomatik testlerdir. evik yazılım geliştirme srelerinde birim testleri, yalnızca kalite gvence aracı deėil, aynı zamanda tasarım ve geliştirme srecini ynlendiren bir mekanizma olarak ele alınmaktadır (Meszaros, 2007).

evik projelerde birim testlerinin otomatik olması, testlerin sık aralıklarla alıřtırılabilmesini ve hızlı geri bildirim alınmasını saėlar. Bu durum, geliřtiricilerin hataları erken ařamada tespit etmesine ve dzeltmesine olanak tanır. Birim testleri, srekli entegrasyon srelerinin de temel yapı tařlarından biridir.

Test Gdml Geliřtirme (TDD)

Test gdml geliştirme (Test-Driven Development – TDD), yazılım geliştirme srecinde testlerin koddan nce yazılmasını ngren bir yaklařımdır. TDD dngs genellikle kırmızı–yeřil–refactor adımlarıyla ifade edilir. İlk adımda bařarısız bir test yazılır (kırmızı), ardından testi geecek en basit kod yazılır (yeřil) ve son olarak kod refactoring ile iyileřtirilir.

TDD yaklařımı, geliřtiricileri gereksinimleri daha net dřnmeye ve test edilebilir tasarımlar retmeye teřvik eder. Testlerin nce yazılması, geliřtiricinin sistemin nasıl davranması gerektiėini aık biimde tanımlamasını saėlar. Bu durum hem kod kalitesini artırmakta hem de gereksinimlerin yanlıř anlařılmasını azaltmaktadır (Janzen ve Saiedian, 2005).

Literatrde TDD’nin hata oranlarını azalttıėı ve kodun bakımını kolaylařtırdıėına dair ok sayıda alıřma bulunmaktadır. Ancak TDD’nin etkili biimde uygulanabilmesi, geliřtiricilerin test

yazma becerilerinin gelişmiş olmasını ve organizasyonel düzeyde bu yaklaşımın desteklenmesini gerektirmektedir.

Birim Testleri ile Tasarım Arasındaki İlişki

Birim testleri ve tasarım arasında çift yönlü bir ilişki bulunmaktadır. İyi tasarlanmış sistemler daha kolay test edilebilirken, test yazma süreci de tasarım kararlarını doğrudan etkilemektedir. TDD uygulanan projelerde, karmaşık ve sıkı bağlı tasarımlar genellikle test yazımını zorlaştırmaktadır. Bu durum geliştiricileri daha modüler ve sade çözümler üretmeye yönlendirmektedir (Ashbacher vd., 2010).

Bu bağlamda birim testleri, yalnızca hataları yakalayan araçlar değil, aynı zamanda tasarımın kalitesini değerlendiren geri bildirim mekanizmaları olarak da görülebilir. Test yazmanın zor olduğu durumlar, çoğu zaman tasarımın yeniden ele alınması gerektiğine işaret etmektedir.

Refactoring ve Sürekli Kalite İyileştirme

Refactoring, çevik yazılım geliştirme süreçlerinde kaliteyi sürdürülebilir kılan temel uygulamalardan biridir. Birim testleri ve TDD, refactoring faaliyetlerinin güvenli biçimde gerçekleştirilmesini mümkün kılar. Testlerin varlığı, geliştiricilerin kodu iyileştirirken sistem davranışının bozulmadığından emin olmasını sağlar.

Çevik projelerde refactoring, büyük ve seyrek yapılan bir faaliyet değil, küçük ve sürekli adımlarla gerçekleştirilen bir süreçtir. Bu yaklaşım, teknik borcun birikmesini engeller ve yazılımın uzun vadeli bakım maliyetlerini düşürür.

Sürekli Entegrasyon ve Sürekli Teslimat (CI/CD)

CI/CD Kavramının Ortaya Çıkışı

Çevik yazılım geliştirme süreçlerinde kodun sık aralıklarla değiştirilmesi, entegrasyon problemlerini kaçınılmaz hâle getirmiştir. Birden fazla geliştiricinin aynı kod tabanı üzerinde paralel çalışması, manuel entegrasyon yaklaşımlarını sürdürülemez kılmıştır. Bu bağlamda Sürekli Entegrasyon (Continuous Integration, CI) ve Sürekli Teslimat (Continuous Delivery, CD) kavramları, çevik yazılım geliştirme süreçlerinin doğal bir uzantısı olarak ortaya çıkmıştır (Humble ve Farley, 2010).

Sürekli entegrasyon, geliştiricilerin kodlarını ortak bir depoya sık aralıklarla göndermesi ve her gönderim sonrası kodun otomatik olarak derlenip test edilmesini ifade eder. Sürekli teslimat ise, başarılı entegrasyonların ardından yazılımın üretim ortamına hazır hâle getirilmesini sağlayan otomasyon süreçlerini kapsar.

Sürekli Entegrasyonun Temel İlkeleri

Sürekli entegrasyon yaklaşımının temelinde erken hata tespiti ve hızlı geri bildirim yer almaktadır. CI süreçlerinde her kod değişikliği otomatik olarak test edilmekte ve entegrasyon hataları erken aşamada görünür hâle gelmektedir. Bu durum, hata düzeltme maliyetlerini önemli ölçüde azaltmaktadır.

CI süreçlerinde yaygın olarak uygulanan ilkeler şunlardır:

- Sık kod birleştirme
- Otomatik derleme
- Otomatik test yürütme
- Hızlı geri bildirim
- Her zaman çalışır durumda bir ana dal

Bu ilkeler, çevik yazılım geliştirme felsefesiyle doğrudan örtüşmektedir.

Sürekli Teslimat ve DevOps Entegrasyonu

Sürekli teslimat, yazılımın üretim ortamına dağıtılmasının teknik ve operasyonel risklerini azaltmayı amaçlar. Dağıtım sürecinin otomatikleştirilmesi, insan kaynaklı hataların önüne geçerken sürüm yönetimini de kolaylaştırmaktadır (Kim vd., 2021).

CI/CD yaklaşımları, DevOps kültürüyle birlikte ele alındığında yazılım geliştirme ve operasyon süreçleri arasındaki sınırları ortadan kaldırmaktadır. Bu bütünleşme, çevik yazılım geliştirme süreçlerinin organizasyonel düzeyde ölçeklenmesine olanak tanımaktadır.

Çevik Yazılımda Yapay Zekâ Kullanımı

AI ve Çevik Yazılımın Kesişim Noktası

Son yıllarda yapay zekâ (AI) tabanlı araçların yazılım geliştirme süreçlerine entegrasyonu, çevik yazılım ekipleri için önemli fırsatlar sunmaktadır. Çevik yazılım geliştirme, hız, geri bildirim ve sürekli iyileştirme üzerine kuruluyken; yapay zekâ otomasyon, tahminleme ve örüntü keşfi gibi yetenekleriyle bu süreçleri desteklemektedir (Amershi vd., 2019).

Çevik Süreçlerde AI Destekli Uygulamalar

AI tabanlı araçlar, çevik yazılım süreçlerinde farklı aşamalarda kullanılabilir. Kullanıcı hikâyelerinin oluşturulması, backlog refinement, sprint tahminleri, kod üretimi ve test otomasyonu bu alanlar arasında yer almaktadır. Özellikle doğal dil işleme teknikleri, kullanıcı gereksinimlerinin analiz edilmesini ve daha net ifade edilmesini mümkün kılmaktadır (Zhao vd., 2021).

Kod üretimi ve refactoring süreçlerinde AI destekli araçların kullanılması, geliştirici verimliliğini artırmakta ve sprint kapasitesini

yükseltmektedir. Bununla birlikte, AI araçlarının önerilerinin eleştirel bir süzgeçten geçirilmesi gerekmektedir.

Etik, Güvenlik ve Sınırlılıklar

AI kullanımının yaygınlaşması, etik ve güvenlik konularını da gündeme getirmektedir. Yanlış öneriler, lisans sorunları ve gizlilik ihlalleri, AI destekli yazılım geliştirme süreçlerinde dikkat edilmesi gereken riskler arasında yer almaktadır (Floridi vd., 2018).

Bu bağlamda yapay zekâ, çevik ekiplerin yerini alan bir unsur değil, karar destek mekanizması olarak konumlandırılmalıdır. İnsan uzmanlığı ve deneyimi, AI destekli sistemlerin sağlıklı biçimde kullanılabilmesi için vazgeçilmezdir.

Çevik Yazılım ve Kariyer Perspektifi

Çevik yazılım geliştirme, yalnızca teknik süreçleri değil, yazılım mühendislerinin kariyer anlayışını da dönüştürmüştür. Günümüz yazılım mühendislerinden yalnızca teknik bilgiye sahip olmaları değil aynı zamanda iletişim, takım çalışması ve sürekli öğrenme becerilerini de geliştirmeleri beklenmektedir (Fitzgerald ve Stol, 2017).

Çevik ekiplerde çalışan yazılım mühendisleri, farklı rolleri deneyimleme ve disiplinler arası iş birliği yapma fırsatı bulmaktadır. Bu durum, kariyer gelişimini hızlandırmakta ve mühendislerin adaptasyon yeteneklerini artırmaktadır. Ayrıca DevOps ve AI destekli yazılım geliştirme yaklaşımları, yazılım mühendisliği kariyerinde yeni uzmanlık alanlarının ortaya çıkmasına yol açmıştır.

Sonuç

Bu kitap bölümünde çevik yazılım geliştirmenin kuramsal temelleri, Extreme Programming yaklaşımı, planlama ve iletişim süreçleri, çevik tasarım ve test uygulamaları, CI/CD ve DevOps

entegrasyonu ile yapay zekâ destekli çevik süreçler bütüncül bir bakış açısıyla ele alınmıştır.

Çevik yazılım geliştirme, değişken ve belirsiz ortamlarda etkili çözümler sunan güçlü bir yaklaşım olmakla birlikte, başarısı teknik uygulamaların ötesinde insan faktörü ve organizasyonel kültüre bağlıdır. Yapay zekâ ve otomasyon teknolojilerinin çevik süreçlere entegrasyonu, bu yaklaşımın gelecekte daha da evrilmesine katkı sağlayacaktır.

Kaynakça

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., ... & Zimmermann, T. (2019, May). Software engineering for machine learning: A case study. In 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP) (pp. 291-300). IEEE.

Beck, K. (2000). Extreme programming explained: embrace change. addison-wesley professional.

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... & Thomas, D. (2001). Manifesto for agile software development.

Binder, R. (2000). Testing object-oriented systems: models, patterns, and tools. Addison-Wesley Professional.

Boehm, B. (2002). Get ready for agile methods, with care. Computer, 35(1), 64-69.

Boehm, B., & Turner, R. (2004, May). Balancing agility and discipline: Evaluating and integrating agile and plan-driven methods. In Proceedings. 26th International Conference on Software Engineering (pp. 718-719). IEEE.

Cohn, M. (2005). Agile estimating and planning. Pearson Education.

Fitzgerald, B., & Stol, K. J. (2017). Continuous software engineering: A roadmap and agenda. *Journal of Systems and Software*, 123, 176-189.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and machines*, 28(4), 689-707.

Fowler, M. (2018). *Refactoring: improving the design of existing code*. Addison-Wesley Professional.

Ashbacher, C., Freeman, S., Pryce, N., & Pearson, A. W. (2010). Growing Object-Oriented Software, Guided by Tests. *J. Object Technol.*, 9(3), 41-42.

Grenning, J. (2002). Planning poker or how to avoid analysis paralysis while release planning. *Hawthorn Woods: Renaissance Software Consulting*, 3, 22-23.

Highsmith, J. (2009). *Agile project management: creating innovative products*. Pearson education.

Humble, J., & Farley, D. (2010). *Continuous delivery: reliable software releases through build, test, and deployment automation*. Pearson Education.

Janzen, D., & Saiedian, H. (2005). Test-driven development concepts, taxonomy, and future direction. *Computer*, 38(9), 43-50.

Kerth, N. (2001). *Project Retrospectives: A Handbook for Team Reviews* Dorset House Publishing Company.

Kim, G., et al. (2021). *The DevOps handbook* (2nd ed.). IT Revolution.

Martin, R. C. (2018). *Clean Architecture: Das Praxishandbuch für professionelles Softwaredesign. Regeln und Paradigmen für effiziente Softwarestrukturierung*. MITP-Verlags GmbH & Co. KG.

Meszaros, G. (2007). *xUnit test patterns: Refactoring test code*. Pearson Education.

Moe, N. B., Dingsøyr, T., & Dybå, T. (2010). A teamwork model for understanding an agile team: A case study of a Scrum project. *Information and software technology*, 52(5), 480-491.

Pressman, R. S., & Maxim, B. R. (2020). *Software engineering: A Practioner; s Appoar*.

Royce, W. (1970, August). The software lifecycle model (waterfall model). In *Proc. Westcon* (Vol. 314).

Schwaber, K., & Sutherland, J. (2020). *Le Guide Scrum*.

Serrador, P., & Pinto, J. K. (2015). Does Agile work?—A quantitative analysis of agile project success. *International journal of project management*, 33(5), 1040-1051.

Sommerville, I. (2020). *Engineering software products* (Vol. 355). London, UK: Pearson.

Spinellis, D. (2012). *Git*. O'Reilly Media.

Standish Group. (2020). *CHAOS report*.

Williams, L., & Kessler, R. R. (2003). *Pair programming illuminated*. Addison-Wesley Professional.

Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K. J., Ajagbe, M. A., Chioasca, E. V., & Batista-Navarro, R. T. (2021). Natural language processing for requirements engineering: A systematic mapping study. *ACM Computing Surveys (CSUR)*, 54(3), 1-41.

BÖLÜM 2

KABLOSUZ AĞLARDA DENEYİM KALİTESİ (QOE): KAVRAMSAL ÇERÇEVE, ÖLÇÜM METODOLOJİLERİ VE YÖNETİM STRATEJİLERİ

SEMİH KAHVECİ¹
RAMAZAN AKKURT²

Giriş

Son yıllarda telekomünikasyon ve ağ teknolojilerinde yaşanan devrim niteliğindeki gelişmeler, küresel internet trafiğinin yapısını ve kullanıcıların dijital tüketim alışkanlıklarını kökten değiştirmektedir. Günümüzde ağ kaynaklarının kullanımı, statik veri transferinden dinamik ve yüksek bant genişliği gerektiren çoklu ortam (multimedya) uygulamalarına kaymıştır. Bu dönüşümün merkezinde ise, internet trafiğinin aslan payını oluşturan çevrimiçi video izleme aktiviteleri yer almaktadır. Özellikle Netflix, Amazon Prime, YouTube ve Disney+ gibi "Over-the-Top" (OTT) video akış hizmetlerinin küresel ölçekte yaygınlaşması, video akışını modern

¹ Arş. Gör. Dr., Mersin Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği, Orcid: 0000-0002-1495-6295

² Arş. Gör., Mersin Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği, Orcid: 0000-0003-2319-9887

ağ sistemlerinin en popüler ve yönetimi en zor uygulaması haline getirmiştir. Ancak, hizmet sağlayıcıların sayısının artmasıyla birlikte oluşan bu yoğun rekabet ortamında, platformların varlıklarını sürdürebilmeleri sadece içerik sunmaya değil, sunulan içeriğin kalitesine bağlı hale gelmiştir. Teknik altyapının kusursuz çalışması tek başına yeterli olmamakla birlikte kullanıcının hizmeti algılama biçimi belirleyici faktör olarak öne çıkmaktadır. İşte bu noktada, hizmet sağlayıcıların kullanıcı memnuniyetini ölçmek, anlamak ve maksimize etmek için başvurdukları Kullanıcı Deneyim Kalitesi (Quality of Experience - QoE) kavramı, ağ yönetiminin ve dijital servislerin sürdürülebilirliğinin en kritik bileşeni olarak literatürde ve endüstride yerini almaktadır (Baraković ve Skorin-Kapov, 2013).

Kullanıcı Deneyim Kalitesi (Quality of Experience - QoE), en genel tanımıyla, bir uygulamanın veya hizmetin son kullanıcı tarafından algılanan öznel kabul edilebilirlik düzeyidir. Uluslararası Telekomünikasyon Birliği - Telekomünikasyon Standardizasyon Sektörü (ITU-T) 'nın standartlarında belirtildiği üzere bu kavram, kullanıcının bir hizmetten ne kadar memnun kaldığını, hizmetin kullanıcının beklentilerini ne ölçüde karşıladığını ifade eden, tamamen insan odaklı bir metriktir. QoE'nin tarihsel sahneye çıkışı, 2000'li yılların başına, telekomünikasyon dünyasında yaşanan kritik bir paradigma değişimine dayanır. Başlangıçta mühendislik dünyası sadece ağın teknik performansına odaklanırken, zamanla ağların teknik kapasitesi artmasına rağmen kullanıcı şikayetlerinin azalmadığı fark edilmiştir. Bu durum, teknik mükemmelliğin kullanıcı mutluluğunu garanti etmediği gerçeğini ortaya çıkarmış ve endüstrinin odağını "Ağ Merkezli" (Network-Centric) yaklaşımdan "Kullanıcı Merkezli" (User-Centric) yaklaşıma kaydırmasına neden olmuştur. QoE kavramı, işte bu boşluğu doldurmak ve teknolojiyi insan psikolojisiyle birlikte ele almak ihtiyacından doğmuştur (Taha vd., 2021).

QoE'nin günümüzdeki hayati önemi, servis sağlayıcılar için artık bir tercih değil, rekabetçi piyasada hayatta kalma (survival) meselesi olmasından kaynaklanmaktadır. Modern telekomünikasyon pazarında kullanıcılar, teknik detayları bilmeden sadece hizmetin onlara nasıl hissettirdiğine odaklanmaktadır. Eğer bir kullanıcı kötü bir deneyim yaşarsa, operatörünü veya kullandığı servisi değiştirme eğilimi göstermektedir. Bu nedenle QoE, müşteri sadakatini sağlamanın ve geliri korumanın temel anahtarıdır. Kullanım alanları açısından bakıldığında, özellikle video akış (streaming) hizmetlerinde (YouTube, Netflix gibi) donma veya çözünürlük sorunlarının yönetimi, VoIP (internet üzerinden ses) görüşmelerinde gecikme ve yankının engellenmesi, online oyunlarda ise anlık tepki sürelerinin optimize edilmesi gibi alanlar öne çıkmaktadır. Pratikte QoE yönetimi, ağ kaynaklarının (örneğin bant genişliğinin) rastgele değil, kullanıcının o anki memnuniyetini maksimize edecek şekilde dağıtılmasını hedefler.

Deneyim Kalitesi (QoE) yönetiminin ihmal edildiği bir içerik sağlama senaryosu ele alındığında, son kullanıcı tarafında oluşabilecek hizmet kusurları dramatik sonuçlar doğurmaktadır. Bir kullanıcının video akışını başlatma isteğine rağmen;

- Yanlış içerik gösterimi,
- Oynatma komutuna sistemin tepki vermemesi (başlangıç gecikmesi),
- Akışın başladığı ancak sık sık donmalarla (stalling/rebuffering) kesintiye uğradığı veya video çözünürlüğünde istikrarsız dalgalanmaların (quality switching) yaşandığı durumlar, hizmetin kabul edilebilirliğini ortadan kaldırmaktadır.

Bu tür hizmet kusurları, kullanıcının bilişsel ve duygusal algısında doğrudan bir memnuniyetsizlik yaratmakta ve hizmet sağlayıcıya yönelik güveni sarsmaktadır. Günümüzün rekabetçi

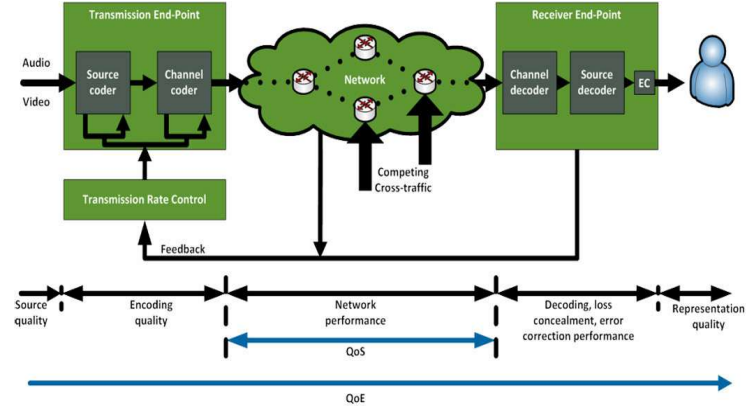
diijital pazarında, video akış hizmetlerinin sürdürülebilirliğı, yalnızca içeriğın kendisine değıl, bu içeriğın son kullanıcıya sunulma üstünlüğüne bağılıdır. Literatürde belirtildiğı üzere, kullanıcının yaşadığı deneyimi anlık olarak izleme ve yönetme yeteneğı, müşteri sadakatini sağlamak ile müşteri kaybı arasındaki ince çizgiyi belirleyen en kritik faktördür (Baraković ve Skorin-Kapov, 2013).

Tarihsel süreçte araştırmacılar, Deneyim Kalitesini (QoE) ölçümlemek için uzun süre yalnızca Hizmet Kalitesi (QoS) parametrelerine güvenmiş ve "yüksek QoS, yüksek QoE ile sonuçlanır" varsayımını kabul etmişlerdir. ITU-T standartları çerçevesinde QoS, üçüncü taraf (Third-Party) veya İçerik Dağıtım Ağları (CDN) dahil olmak üzere tüm altyapının teknik performansını belirleyen özelliklerin toplamı olarak tanımlanmaktadır. Ancak (Dai, 2011) ve (Baraković ve Skorin-Kapov, 2013) tarafından detaylandırıldığı üzere; bant genişliğı, gecikme, paket kaybı, titreşim (jitter) ve önbellek doluluğı gibi tamamen objektif ve makine odaklı bu metriklerin, insan algısındaki sübjektif karşılığını tam olarak yansıtmadığı günümüzde kabul gören bir gerçektir. İki kavram arasındaki bu kritik ayrım, Baraković'in de altını çizdiği üzere ilişkinin doğrusal (non-linear) olmamasından kaynaklanır; teknik kaliteyi (QoS) artırmak her zaman kullanıcı memnuniyetini (QoE) aynı oranda artırmayabilir, zira insan algısı belirli bir eşikten sonraki iyileştirmeleri fark etmezken, küçük bir teknik hata psikolojik olarak büyük bir memnuniyetsizliğe yol açabilir. Özetle QoS, hizmetin teknik teslimat mekanizmasının verimliliğini ölçerken, QoE bu teslimatın kullanıcı zihninde yarattığı nihai tatmini ölçmektedir.

Şekil 1 ile sunulan şema, Hizmet Kalitesi (QoS) ile Deneyim Kalitesi (QoE) arasındaki kapsam ve odak farkını somutlaştırmaktadır. Şekilden de anlaşılacağı üzere QoE, uçtan uca (end-to-end) hizmetin tamamına odaklanan bütüncül bir kavram olarak tanımlanmaktadır. QoS ise bu deneyimin yalnızca teknik dağıtım ve ağ iletim katmanını kapsamaktadır. Bu bağlamda QoS,

QoE ekosisteminin kritik bir bileşeni olmakla birlikte, esasen onun teknik bir alt kümesi niteliğindedir.

Şekil 1: QoE ve QoS farkı (Perkis, 2013)



Literatürde de vurgulandığı üzere, ağ trafiğine dair tüm QoS gereksinimlerinin karşılanması, yüksek bir QoE elde edileceğini garanti etmez (gerekli koulouri ancak yeterli değildir). Zira teknik parametreler ne kadar kusursuz olursa olsun, nihai memnuniyet seviyesinde son kullanıcının sübjektif algısı, psikolojik durumu ve beklentileri belirleyici rol oynamaktadır.

QoE Modelleme ve Yönetim Esasları

Deneyim Kalitesi (QoE) yönetiminde sürdürülebilir bir başarı elde edebilmek, öncelikle insan algısının karmaşık yapısının ve bu algıyı şekillendiren dışsal değişkenlerin derinlemesine anlaşılmasını gerektirmektedir. QoE modelleme süreci, temel olarak ölçülebilir teknik parametreler ile kullanıcının sübjektif memnuniyet seviyesi arasındaki doğrusal olmayan ilişkiyi matematiksel olarak ifade etmeyi amaçlamaktadır. Güvenilir bir modelin nihai hedefi, belirli bir hizmet senaryosu ve koşullar altında kullanıcının olası QoE puanını yüksek doğrulukla tahmin etmektir (Baraković ve Skorin-Kapov, 2013).

QoE Etki Faktörleri (Influence Factors - IFs)

Kullanıcının bir hizmeti deneyimleme sürecinde algıladığı kaliteyi doğrudan veya dolaylı olarak değiştirebilen, kullanıcıya, sisteme, hizmete veya ortam bağlamına ait her türlü karakteristik özellik "Etki Faktörü" olarak tanımlanır (Baraković ve Skorin-Kapov, 2013). Literatürde bu faktörler, yönetilebilirliği kolaylaştırmak adına üç temel kategori altında toplanmıştır:

İnsan Faktörleri (Human IFs): Kullanıcının demografik, fiziksel ve psikolojik özelliklerini kapsar. Bu kategori; cinsiyet, yaş, eğitim düzeyi gibi statik özellikleri içerebileceği gibi; kullanıcının anlık duygusal durumu, motivasyonu, görsel ve işitsel keskinliği gibi dinamik değişkenleri de ifade eder. Örneğin, kullanıcının o anki ruh hali, aynı teknik kalitedeki bir videoyu farklı puanlamasına neden olabilir.

Sistem Faktörleri (System IFs): Bir uygulama veya hizmetin teknik olarak üretilen kalitesini belirleyen deterministik özelliklerdir. İçerik üretiminden (kodlama/encoding), ağ üzerinden iletme (gecikme, paket kaybı, bant genişliği) ve son kullanıcı cihazındaki gösterime (ekran çözünürlüğü, cihaz performansı) kadar olan tüm teknik zinciri kapsar.

Bağlam Faktörleri (Context IFs): Kullanıcının hizmeti aldığı esnadaki fiziksel, zamansal, sosyal ve ekonomik ortamı tanımlar. Kullanıcının konumu (ev/hareketli araç), hizmetin maliyeti, abonelik türü, ortam ışığı veya gizlilik gereksinimleri bu kategoride değerlendirilir (Baraković ve Skorin-Kapov, 2013).

QoE Değerlendirme Metodolojileri

Tasarlanan QoE modellerinin doğrulanması ve kalitenin ölçümü, literatürde Sübjektif (Öznel) ve Objektif (Nesnel) değerlendirme olmak üzere iki ana ekseninde incelenmektedir.

Sübjektif (Öznel) Değerlendirme

Sübjektif değerlendirme, QoE ölçümünün "Altın Standardı" (Ground Truth) olarak kabul edilir. Bu yöntemde, kontrollü bir laboratuvar ortamında veya saha koşullarında gerçek kullanıcılara (test denekleri) orijinal ve işlenmiş video dizileri izletilerek, algıladıkları kaliteyi puanlamaları istenmektedir. ITU-T standartları çerçevesinde en yaygın kullanılan metrik Ortalama Görüş Puanı (Mean Opinion Score - MOS) olup, kullanıcılardan deneyimlerini 1 (Kötü) ile 5 (Mükemmel) arasında derecelendirmeleri beklenmektedir (Dai, 2011). Ancak sübjektif testlerin uygulanabilirliği konusunda ciddi sınırlılıklar mevcuttur:

- Zaman, iş gücü ve finansal kaynaklar açısından yüksek maliyetlidir.
- Sınırlı sayıda denek ve video içeriği ile yapıldığından genelleştirilmesi zordur.
- En önemlisi, gerçek zamanlı (real-time) izleme ve müdahale gerektiren durumlarda kullanılması imkansızdır.

Objektif (Nesnel) Değerlendirme

Objektif değerlendirme yöntemleri, maliyetli ve yavaş olan sübjektif testlere alternatif olarak insan algısını matematiksel algoritmalarla taklit etmek ve teknik parametrelerden tahmini bir QoE puanı üretmek amacıyla geliştirilmiştir. Burada temel amaç, sübjektif testlerden elde edilen verileri referans alarak, nesnel parametreleri (gecikme, çözünürlük vb.) kullanıcı memnuniyetine eşleyen modeller kurmaktır.

Objektif modeller, Orijinal Referans Videosuna duydukları ihtiyaca göre üç sınıfa ayrılır (Seufert vd., 2014):

Tam Referanslı Modeller (Full Reference - FR): Orijinal video ile bozulmuş videoyu kare kare (pixel-by-pixel) karşılaştırır. İnsan algısına en yakın sonuçları veren yöntemdir (Örn: SSIM, VQM). Ancak orijinal videonun kaynakta bulunmasını gerektirdiğinden gerçek zamanlı ağ iletiminde kullanılamamaktadır.

Azaltılmış Referanslı Modeller (Reduced Reference - RR): Orijinal videonun tamamına değil, sadece belirli özniteliklerine (özet bilgilerine) ihtiyaç duymaktadır.

Referanssız Modeller (No Reference - NR): Kaynak videoya ihtiyaç duymadan, sadece alınan sinyal veya ağ verisi üzerinden kalite tahmini yapmaktadır. Hesaplama verimliliği ve pratiklik açısından, özellikle akıllı telefonlar ve gerçek zamanlı izleme sistemleri için en uygun çözümdür.

Makine Öğrenimi Tabanlı Yaklaşımlar Geleneksel NR modelleri, modern video kodlama teknikleri ve karmaşık ağ koşulları karşısında yeterli doğruluk sağlayamamaktadır. (Panahi vd., 2024) tarafından belirtildiği üzere, güncel literatürde Makine Öğrenimi (ML) tabanlı hibrit yaklaşımlar öne çıkmaktadır. Bu yöntemde süreç şu şekilde işler:

- Servis sağlayıcı, temsili video türleri (spor, aksiyon, haber vb.) ile bir başlangıç tahmin modeli eğitir (Offline Training).
- Eğitilen bu model, istemci cihazlara (Client) iletilir.
- İstemci, gerçek zamanlı akış sırasında videonun özniteliklerini çıkarır ve bu model üzerinden anlık kalite tahmini (Qp) yapar.
- Model, yeni video türleri ve kullanıcı davranışlarıyla zaman içinde güncellenerek doğruluğunu artırır.

Bu modellerin başarısı; seçilen algoritmanın (SVM, Random Forest, Deep Learning) yeteneğine, eğitim için kullanılan veri setinin

(Ground-Truth) kalitesine ve videoyu karakterize eden öz niteliklerin doğru seçilmesine bağlıdır (Panahi vd., 2024).

QoE Yönetim Döngüsü

QoE yönetimi, sadece ölçümle sınırlı kalmayıp, izleme ve optimizasyonu da içeren kapalı bir döngü olarak ele alınmalıdır.

İzleme ve Ölçüm (Monitoring)

QoE izleme süreci, ağın ve kullanıcının durumuna dair verilerin toplanmasını içerir. Bu veriler şunlardır:

- Ağ Metrikleri: Bant genişliği, paket kaybı, titreşim.
- Cihaz Yetenekleri: Ekran çözünürlüğü, CPU yükü, pil durumu.
- Uygulama Verileri: Kullanılan codec, video bit hızı, içerik türü.

İzleme işlemi iki noktada gerçekleştirilebilir: Ağ tarafında yapılan izleme (örneğin Derin Paket İnceleme - DPI ile), sorunlara hızlı tepki verilmesini sağlasa da şifreli trafikte yetersiz kalabilir. Son kullanıcı tarafında (Client-side) izleme ise, algılanan kalite hakkında en doğru veriyi sunar ancak bu verinin merkezi sisteme geri raporlanması ek bir trafik yükü ve protokol gerektirir (Baraković & Skorin-Kapov, 2013).

QoE Optimizasyon ve Kontrol

QoE yönetim döngüsünün en kritik ve nihai aşaması olan optimizasyon süreci, (Taha vd., 2021) tarafından önerilen yaklaşımla, statik bir kaynak yönetiminden çıkarılıp dinamik ve uyarlanabilir (adaptive) bir kontrol mekanizmasına dönüştürülmüştür. Geleneksel yöntemler ağ kaynaklarını veri türüne bakmaksızın dağıtırken, güncel optimizasyon stratejileri "QoE Adaptif Yönetim Sistemi" (QoE-AMS) prensibine dayanır. Bu sistemde optimizasyon, sadece ağın o anki durumuna (bant genişliği, gecikme) tepki vermekle

kalmaz; aynı zamanda iletilen multimedya içeriğinin karakteristik özelliklerini de analiz ederek proaktif kararlar alır. Amaç, ağ koşulları bozulmaya başladığında video akışını tamamen kesmek yerine, kullanıcının en az hissedeceği şekilde kalite parametrelerini (çözünürlük, kare hızı) anlık olarak yeniden yapılandırmaktır.

Bu kontrol mekanizmasının merkezinde, videonun mekansal (spatial) ve zamansal (temporal) karmaşıklığının analizi yatar. (Taha vd., 2021), her video karesinin ağ üzerinde yarattığı yükün ve kullanıcı algısındaki öneminin eşit olmadığını vurgulamaktadır. Örneğin, yüksek hareketlilik içeren bir aksiyon sahnesi (yüksek zamansal karmaşıklık) ile durağan bir haber bülteni (düşük zamansal karmaşıklık) farklı optimizasyon stratejileri gerektirir. Geliştirilen algoritmalar, ağda bir daralma (congestion) tespit ettiğinde, videonun en kritik yapı taşları olan I-karelerini (Intra-frames) koruma altına alırken, algısal olarak daha az fark edilen detayların veya ara karelerin (B-kareleri) aktarımından stratejik olarak feragat eder. Böylece, kullanıcı için en büyük memnuniyetsizlik kaynağı olan "donma" (stalling) olayının önüne geçilir.

Sonuç olarak, (Taha vd., 2021) tarafından detaylandırılan bu optimizasyon yaklaşımı, "sınırlı kaynakla maksimum memnuniyet" hedefini matematiksel bir kesinliğe oturtur. Sistem, hem objektif ağ verilerini (QoS) hem de sübjektif video özelliklerini çapraz katmanlı (cross-layer) bir mimaride işleyerek, tahmin edilen MOS (Mean Opinion Score) değerini sürekli olarak belirli bir eşiğin üzerinde tutmayı hedefler. Bu sayede, özellikle kaynakların son derece değişken olduğu kablosuz ağlarda dahi, HD ve UHD gibi yüksek bant genişliği gerektiren servislerin sürdürülebilirliği ve kullanıcı sadakati (loyalty) teknik olarak garanti altına alınmış olur.

Kaynak Araştırmaları

Kablosuz ağ teknolojilerindeki hızlı evrim ve kullanıcıların multimedya içeriklerine olan talebinin artması, Deneyim Kalitesi

(Quality of Experience - QoE) kavramını ağ yönetiminin merkezine yerleştirmiştir. Literatür incelendiğinde, QoE çalışmalarının kavramsal tanımlamalardan başlayıp, adaptif yöntemlere ve son olarak yapay zeka destekli tahmin modellerine doğru evrildiği görülmektedir.

Kavramsal Temeller ve Yönetim Modelleri QoE kavramının teorik altyapısı üzerine yapılan ilk çalışmalarda, kullanıcı algısı ile teknik parametreler arasındaki ayrım netleştirilmiştir. (Dai, 2011), QoE'nin sadece teknik bir sonuç olmadığını, Hizmet Kalitesi'nden (QoS) farklı olarak kullanıcının sübjektif tatminini temsil eden nihai bir hedef olduğunu vurgulamıştır. Benzer şekilde Baraković ve (Skorin-Kapov vd., 2018), kablosuz ağlarda QoE yönetiminin karmaşıklığına dikkat çekerek, deneyimi etkileyen unsurları İnsan, Sistem ve Bağlam (Context) faktörleri olarak sınıflandıran kapsamlı bir yönetim döngüsü önermişlerdir. Bu dönemde (Chen vd., 2014), QoS'tan QoE'ye geçiş sürecindeki metodolojileri derleyerek modelleme çalışmalarına ışık tutmuştur.

Klasik Makine Öğrenimi ve Analitik Yaklaşımlar QoE'yi ölçülebilir kılmak adına yapılan çalışmalarda, başlangıçta analitik formüller ve geleneksel makine öğrenimi yöntemleri kullanılmıştır. (Lloret vd., 2011) ile (Garcia vd., 2009), IPTV sistemlerinde titreşim (jitter), gecikme ve paket kaybı gibi ağ parametrelerine dayalı matematiksel ifadeler geliştirmiştir. (Menkovski vd., 2010) ise bu parametreleri Destek Vektör Makineleri (SVM) ve karar ağaçları gibi algoritmalarla işleyerek, sübjektif kullanıcı puanlarını tahmin etmeye çalışmıştır. (Aroussi ve Mellouk, 2014) tarafından yapılan taramada, Rastgele Orman (RF) ve Naive Bayes gibi algoritmaların QoS-QoE korelasyonunu kurmadaki başarısı ortaya konmuştur.

Adaptif Akış ve İçerik Farkındalığı Ağ koşullarının dinamik yapısı, statik yönetimden adaptif yönetime geçişi zorunlu kılmıştır. (Seufert vd., 2014), ağ koşullarındaki dalgalanmaların video akışında yarattığı sorunları ele alarak, HTTP Uyarlamalı Akış (HAS)

teknolojisinin önemini vurgulamıştır. HAS, videoyu mevcut bant genişliğine göre dinamik olarak uyarlayarak donma (stalling) sürelerini minimize etmekte ve QoE'yi artırmaktadır. Ancak sadece ağa uyum sağlamak yeterli değildir; videonun içeriği de önemlidir. (Taha vd., 2021), videonun mekansal ve zamansal özelliklerini analiz eden içerik farkındalıklı (content-aware) bir sistem önermiştir. Bu sistem, yüksek hareketli aksiyon sahneleri ile durağan sahneleri ayırt ederek kaynak optimizasyonu sağlamaktadır. Ayrıca (Medeiros vd., 2021), heterojen ağlarda hareketlilik (mobility) sorununa odaklanmış ve sadece sinyal gücüne değil, tahmini QoE değerine bakarak karar veren akıllı bir "handover" (baz istasyonu geçiş) mekanizması geliştirmiştir.

Derin Öğrenme ve Modern Yaklaşımlar Son yıllarda, ağ trafiğinin şifreli hale gelmesi ve veri örüntülerinin karmaşıklaşması ile birlikte Derin Öğrenme (Deep Learning) tabanlı yaklaşımlar öne çıkmıştır. (Lopez-Martin vd., 2018), ağ akış paketlerinden (Network Flow Packets) özellik çıkarımı yaparak QoE tahmini gerçekleştiren ilk derin öğrenme modellerinden birini sunmuşlardır. Bu çalışma, otomatik özellik çıkarımı yeteneği ile literatürde bir kırılma noktası olmuştur. Güncel çalışmalarda ise (Zhang ve Li, 2024), deneyimin anlık değil kümülatif bir süreç olduğunu savunarak, video oturumlarındaki "hafıza etkisini" (hysteresis) modelleyen GRU (Gated Recurrent Unit) tabanlı bir yapı önermişlerdir. (Panahi vd., 2024) ise mevcut literatürü sentezleyerek, modern telekomünikasyon ağlarında (5G/6G) QoE yönetiminin tamamen yapay zeka destekli ve otonom çerçeveler (frameworks) üzerinden yürütüleceği vizyonunu ortaya koymuşlardır.

Literatürdeki tarihsel gelişim incelendiğinde, QoE çalışmalarının başlangıçtaki kavramsal tanımlama çabalarından, günümüzdeki yapay zeka destekli otonom yönetim sistemlerine doğru evrildiği açıkça görülmektedir. Bu kapsamda, QoE yönetiminin farklı dönemlerini temsil eden, odaklandıkları

problemler ve geliřtirdikleri metodolojiler ile literatüre yön veren temel çalışmaların karşılařtırılmalı analizi Tablo 1 ile sunulmaktadır.

Tablo 1: QoE literatüründeki temel çalışmalar ve katkıları.

Yazar / Çalışma	Odaklandığı Temel Problem	Yöntem / Teknoloji	Literatüre Katkısı
Dai, 2011	Kavramsal karmařa ve standart eksikliđi.	Kavramsal Analiz	QoS ve QoE arasındaki teorik ayrımı netleřtirdi.
Baraković ve Skorin-Kapov, 2013	Kablosuz ađlarda yönetim modeli eksikliđi.	Taksonomi (Etki Faktörleri)	QoE'yi İnsan, Sistem ve Bađlam faktörleri olarak sınıflandırdı.
Seufert vd., 2014	Ađ dalgalanmalarının video akışını bozması.	HTTP Adaptif Akış (HAS)	Donmayı (stalling) azaltmak için istemci tarafı adaptasyonun önemini kanıtladı.
Lopez-Martin vd., 2018	Şifreli trafik ve karmařık özellik çıkarımı.	Derin Öğrenme (Deep Learning)	Ađ akış paketlerinden QoE tahmini yapan ilk derin öğrenme modellerinden biridir.
Medeiros vd., 2021	Ađ geçiřlerinde (Handover) kalite kaybı.	Algoritma (NS-3 Simülasyonu)	Sinyal gücüne deđil, QoE tahminine dayalı ađ deđiřtirme mekanizması önerdi.
Taha vd., 2021	İçerik türüne bakılmaksızın kaynak israfı.	İçerik Farkındalıđı (Content-Aware)	Videoğun hareketliliđine göre kaynak ayıran adaptif sistem geliřtirdi.
Zhang ve Li, 2024	QoE tahmininde "hafıza etkisinin" yokluđu.	GRU (Gated Recurrent Unit)	Geçmiş deneyimin şimdiki algıya etkisini (Hysteresis) modelledi.
Panahi vd., 2024	Manuel ölçümün imkansızlıđı.	Yapay Zeka / ML Çerçevesi	Modern ML araçlarını ve 5G/6G vizyonunu derledi.

Tablo 1 ile özetlenen çalışmaların kronolojik analizi, QoE literatürünün geçirdiđi üç temel evreyi gözler önüne sermektedir. (Dai, 2011) ve (Baraković ve Skorin-Kapov, 2013) tarafından yapılan öncü çalışmalar, öncelikle kavramsal karmařayı gidererek QoE'yi QoS'tan ayıran teorik temelleri ve etki faktörlerini

sınıflandırmıştır. İlerleyen yıllarda (Seufert vd., 2014) ile (Medeiros vd., 2021) gibi araştırmacılar, bu teorik temelleri "Adaptif Akış" ve "Handover Yönetimi" gibi spesifik ağ sorunlarına uygulayarak pratik mühendislik çözümleri geliştirmişlerdir.

Ancak tablonun son bölümünde yer alan Lopez-Martin vd., (2018) ile başlayan ve (Zhang ve Li, 2024) ile (Panahi vd., 2024) çalışmalarına uzanan süreç, literatürde bir paradigma değişimine işaret etmektedir. (Taha vd., 2021)'nın içerik farkındalığı vurgusuyla desteklenen bu yeni dönemde, artık manuel kural tabanlı sistemlerin yerini; şifreli trafiği analiz edebilen, insan hafızasındaki etkiyi (hysteresis) modelleyen ve Derin Öğrenme (Deep Learning) mimarilerini kullanan otonom yapılar almıştır. Özetle bu tablo, QoE yönetiminin 'reaktif ölçümden', 'proaktif yapay zeka tahminine' doğru evrildiğini kanıtlamaktadır.

Sonuç ve Gelecek Çalışmalar

QoE modelleme, ölçüm ve izleme süreçlerinin literatürdeki evrimi incelendiğinde, telekomünikasyon endüstrisinin sadece teknik parametreleri (QoS) iyileştirmekten, kullanıcı memnuniyetini bütüncül bir şekilde yönetmeye doğru evrildiği görülmektedir. Ancak, modern ağların artan karmaşıklığı ve kullanıcı beklentilerinin dinamik yapısı, mevcut QoE yönetim sistemlerinde yeni araştırma fırsatlarını ve zorlukları beraberinde getirmektedir. Bu bağlamda gelecekteki çalışmaların odaklanması gereken öncelikli alanlardan biri, dinamik ve bağlam-duyarlı metriklerin geliştirilmesidir. Geleneksel QoE modelleri genellikle statik parametrelere dayandığından, zamanla değişen ağ koşulları ve kullanıcı beklentileri karşısında yetersiz kalabilmektedir. Bu nedenle yeni nesil çalışmalar, sadece video kalitesini değil; kullanıcının izleme alışkanlıklarını, günün saatini ve cihaz durumunu içeren dışsal faktörleri de kapsayan, gerçek zamanlı izlemeye uygun ve

kullanıcı katılımının güvenilir bir göstergesi olan metrikler üretmeye odaklanmalıdır.

Metrik geliřtirmelerinin yanı sıra, yapay zeka destekli davranıřsal analiz ve tahmin yöntemlerine duyulan ihtiya da giderek artmaktadır. oęu nesnel kalite modeli, İnsan Grsel Sistemine (HVS) dayalı fiziksel algıya odaklansa da, psikolojik ve davranıřsal boyutlar (beklenti, sabır, hafıza etkisi) hala tam olarak modellenememiřtir. Literatrde (Panahi vd., 2024) ile (Zhang ve Li, 2024) tarafından da vurgulandıęı zere, Derin ęrenme (Deep Learning) algoritmalarının entegrasyonu kaınılmazdır. zellikle videoların bařlama sresine veya donma olaylarına karřı kullanıcının gsterdięi sabır eřięini ve erken terk etme (churn) fenomenini analiz eden regresyon modellerine ihtiya vardır. Bu modellerin, sadece teknik hataları deęil, kullanıcının bu hataya vereceęi duygusal tepkiyi ve hafıza etkisini (hysteresis effect) de ngrebilecek yetkinlikte olması hedeflenmektedir.

Bu geliřmiř modellerin eęitilmesi ve doęrulanması noktasında karřılařılan en byk engellerden biri ise standartlařtırılmıř aık veri seti eksiklięidir. Tketicilerine eriřimin zor olması ve toplanmasının zaman alması nedeniyle, mevcut VQEG ve LIVE gibi veritabanları modern akıř senaryolarını (UHD, VR, řifreli trafik) temsil etmekte yetersiz kalmaktadır. Arařtırma topluluęunun makine ęrenimi modellerini saęlıklı bir řekilde eęitebilmesi iin, gerek dnya kořullarını yansıtan, kapsamlı ve etiketlenmiř aık kaynaklı veritabanlarının oluřturulması kritik bir gerekliliktir.

Ynetim dngsnn tamamlayıcısı olarak, video iletim optimizasyon srelerinin de QoS odaklı yapıdan, ierik farkındalıklı ve ok kullanıcı QoE odaklı yapıya evrilmesi gerekmektedir. Tarihsel olarak bant geniřlięi maksimizasyonu zerine kurulan optimizasyon stratejileri, (Taha vd., 2021) alıřmasının da iřaret ettięi gibi, artık kullanıcı memnuniyetini

maksimize etmeyi hedeflemelidir. Bu dönüşüm, özellikle dar boğaz yaşanan ağlarda her kullanıcının veya her video içeriğinin kalite gereksiniminin aynı olmadığı gerçeğinden hareketle, kaynakların kullanıcıların duyarlılık seviyelerine ve videoların içerik karmaşıklığına göre adaletli bir şekilde dağıtıldığı "QoE Uyumlu Trafik Planlamasını" zorunlu kılmaktadır.

Kaynakça

- Aroussi, S., & Mellouk, A. (2014). Survey on machine learning-based QoE-QoS correlation models. (s. 200-204). IEEE.
- Baraković, S., & Skorin-Kapov, L. (2013). Survey and challenges of QoE management issues in wireless networks. *Journal of Computer Networks and Communications*, 2013(1), 165146.
- Chen, Y., Wu, K., & Zhang, Q. (2014). From QoS to QoE: A tutorial on video quality assessment. *IEEE Communications Surveys & Tutorials*, 17(2), 1126-1165.
- Dai, Q. (2011). A survey of quality of experience. (s. 146-156). Springer.
- Garcia, M., Canovas, A., Edo, M., & Lloret, J. (2009). A QoE management system for ubiquitous IPTV devices. (s. 147-152). IEEE.
- Lloret, J., Garcia, M., Atenas, M., & Canovas, A. (2011). A QoE management system to improve the IPTV network. *International Journal of Communication Systems*, 24(1), 118-138.
- Lopez-Martin, M., Carro, B., Lloret, J., Egea, S., & Sanchez-Esguevillas, A. (2018). Deep learning model for multimedia quality of experience prediction based on network flow packets. *IEEE Communications Magazine*, 56(9), 110-117.
- Medeiros, I., Pacheco, L., Rosário, D., Both, C., Nobre, J., Cerqueira, E., & Granville, L. (2021). Quality of experience and quality of service-aware handover for video transmission in heterogeneous networks. *International Journal of Network Management*, 31(5), e2064.

- Menkovski, V., Exarchakos, G., & Liotta, A. (2010). Machine learning approach for quality of experience aware networks. (s. 461-466). IEEE.
- Panahi, P. H., Jalilvand, A. H., & Diyanat, A. (2024). Enhancing quality of experience in telecommunication networks: A review of frameworks and machine learning algorithms. *arXiv preprint arXiv:2404.16787*.
- Perkis, A. (2013). Quality of experience (QoE) in multimedia applications. *SPIE Newsroom*.
- Seufert, M., Egger, S., Slanina, M., Zinner, T., Hoßfeld, T., & Tran-Gia, P. (2014). A survey on quality of experience of HTTP adaptive streaming. *IEEE Communications Surveys & Tutorials*, 17(1), 469-492.
- Skorin-Kapov, L., Varela, M., Hoßfeld, T., & Chen, K.-T. (2018). A survey of emerging concepts and challenges for QoE management of multimedia services. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 14(2s), 1-29.
- Taha, M., Canovas, A., Lloret, J., & Ali, A. (2021). A QoE adaptive management system for high definition video streaming over wireless networks. *Telecommunication Systems*, 77(1), 63-81.
- Zhang, X., & Li, L. (2024). Assessing the quality of experience in wireless networks for multimedia applications: A comprehensive analysis utilizing deep learning-based techniques. *Heliyon*, 10(9).

BÖLÜM 3

YAZILIM GÜVENLİK AÇIKLARININ YAPAY ZEKA TABANLI TESPİTİ

GÖZDE MİHRAN KAYA¹
İRFAN KÖSESOY²

Giriş

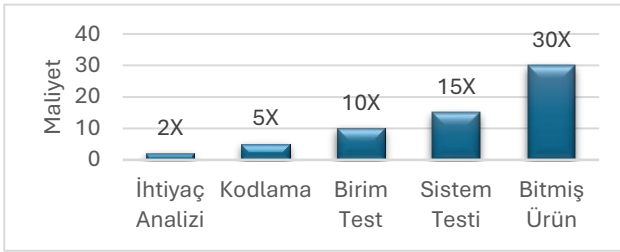
Yazılım ve bilişim sistemlerinin hayatımızdaki rolü gün geçtikçe artmakta, ancak bu yaygınlık beraberinde yeni güvenlik açıklarını da getirmektedir. Yazılım güvenlik açığı, bir saldırgan tarafından istismar edilebilecek yazılım kodundaki veya sistem bileşenlerindeki zafiyetler olarak tanımlanır. Bu zafiyetler saldırganların sistem kontrolünü ele geçirmesine, hassas verileri sızdırmasına veya hizmet dışı bırakma saldırıları gerçekleştirmesine imkân tanıyabilir. Her yıl keşfedilen güvenlik açıklarının sayısı hızla artmaktadır. Örneğin, 2023 yılında NIST ulusal zafiyetler veri tabanına 28.817 açık kaydedilmişken bu sayı 2024 yılında 39.982'ye yükselmiştir. Bu artış trendi, yazılımlardaki güvenlik açıklarının ciddi ve büyüyen bir problem olduğunu göstermektedir.

¹ Öğretim Görevlisi, Kocaeli Sağlık ve Teknoloji Üniversitesi, Bilgisayar Teknolojisi, Orcid: 0000-0003-1522-1159

² Dr. Öğretim Üyesi, Kocaeli Üniversitesi, Yazılım Mühendisliği, Orcid: 0000-0001-5219-5397

Güvenlik açıklarının erken tespit edilmesi, özellikle yazılım geliştirme sürecinin başlarında önlem alınması açısından kritik öneme sahiptir. Çünkü bir güvenlik açığını geliştirme sürecinin ilerleyen safhalarında düzeltmek, erken safhalara kıyasla çok daha maliyetli ve zaman alıcı olabilmektedir. IBM tarafından gerçekleştirilen bir araştırmanın bulgularına göre, yazılım geliştirme sürecinin ilerleyen aşamalarında tespit edilen bir hatanın düzeltme maliyeti katlanarak artmaktadır; başka bir deyişle, bir hata ne kadar erken tespit edilirse giderilme maliyeti de o kadar düşük olmaktadır. Şekil 1, bu bulguyu destekler biçimde, hatanın yazılım geliştirme sürecinin farklı aşamalarındaki maliyet çarpanlarını görsel olarak göstermektedir.

Şekil Hata! Belgede belirtilen stilde metne rastlanmadı. IBM araştırma sonucu maliyet analizi



Geleneksel olarak yazılımlardaki güvenlik açıklarını bulmak, uzmanlar tarafından yapılan manuel kod incelemelerine dayanır. Ancak bu manuel denetimler büyük emek ve uzmanlık gerektirdiği gibi, insan kaynaklı hatalara da açıktır. Bu nedenle, güvenlik açıklarının otomatik araçlarla tespit edilmesi üzerine yoğun araştırmalar yapılmaktadır.

Yazılım güvenlik açığı tespiti için temel yaklaşımlar statik analiz, dinamik analiz ve her ikisini birleştiren hibrit analiz yöntemleri olarak sınıflandırılabilir. Son yıllarda özellikle makine öğrenmesi ve derin öğrenme tabanlı teknikler, bu analiz yöntemlerine entegre edilerek otomatik ve akıllı güvenlik açığı tespiti alanında umut vadeden sonuçlar vermektedir. Bu bölümde, yazılım güvenlik

açıklarının yapay zekâ ile tespitine yönelik yaklaşımlar ele alınacak ve özellikle şu konular üzerinde durulacaktır: (1) Statik ve dinamik analiz tekniklerinin tanımları, kullanım alanları, avantajları ve sınırlılıkları; (2) Doğal dil işleme tabanlı özellik çıkarım tekniklerinin (Word2Vec, FastText, GloVe vb.) kaynak kod analizine uyarlanması; (3) Makine öğrenmesi destekli güvenlik açığı analiz süreçleri ve adımları; (4) Soyut söz dizimi ağacı (AST) tabanlı kod temsiliinin oluşturulması ve sağladığı avantajlar. Tüm bu konular, lisans düzeyindeki okuyucuların da anlayabileceği akademik bir düzende açıklanmaya çalışılacaktır.

GÜVENLİK AÇIĞI TESPİT YAKLAŞIMLARI

Bilgi sistemlerinin güvenliğine yönelik ana tehditlerden biri yazılımlardaki güvenlik açıklarının ortaya çıkmasıdır. Yazılım güvenlik açıkları, saldırganların kendi avantajlarına kullanabilecekleri sistemdeki zafiyetlerdir. Bu zafiyetler, saldırganların hassas bilgileri açığa çıkarmasına veya değiştirmesine, bir sistemi bozmasına veya yok etmesine, bir bilgisayar sisteminin veya programının kontrolünü ele geçirmesine izin veren, bir yazılım parçasındaki belirli kusurlar veya gözetimlerdir. Zafiyetli yazılımların kullanımına sunulması sonrasında bu yazılım açıkları saldırganlar tarafından kullanılabilmekte, sosyal medyada yer edinebilmekte ve bu durum şirketlere ve kurumlara itibar, finansal ve lisans kayıpları gibi zararlar yaşatabilmektedir. Geçmişte güvenlik olaylarına ve kötü

amaçlı kodun çoğalmasına yol açan güvenlik açığı örnekleri bulunmaktadır. Örneğin, birçok farklı bileşene ve işleve sahip, endüstriyel kontrol sistemlerini yeniden programlamak nihai amacına sahip olan Stuxnet, büyük ve karmaşık bir kötü amaçlı yazılımdır. Mantık denetleyicileri (PLC'ler) üzerindeki kodu saldırganın amaçladığı şekilde çalıştırmak için yeniden programlayarak değiştiren ve bu değişiklikleri ekipmanın operatöründen gizleyen işlevlere sahip olan bu güvenlik açığı alanında büyük yankı uyandırmıştır. Bu kötü amaçlı yazılımın yarattığı yankı da göz önünde bulundurulduğunda kod analizinin önemi daha iyi anlaşılmaktadır.

Yazılımdaki hataların bulunması ve giderilmesi için kullanılabilecek teknikler; statik, dinamik ve hibrit yaklaşımlar olarak kategorize edilebilir. Bu tekniklerden, sistem testleri gibi yöntemlerin uygulanması yazılımın çalıştırılmasını gerektirmektedir. Bu yöntemlere dinamik yöntemler denir. Kod gözden geçirme vb. yöntemler ise kod çalıştırılmadan gerçekleştirilen tekniklerdir. Bu tekniklere statik yöntemler adı verilmektedir. Hem statik hem de dinamik yöntemlerin bir arada kullanıldığı hibrit yöntemler de bulunmaktadır. Fuzzy test, taint analiz ve sembolik yürütme gibi yazılım doğrulama alanındaki kavramlar güvenlik açıklarını izlemek için başarıyla uyarlanmıştır. Bununla birlikte, bu yaklaşımların çoğu, belirli koşullar ve güvenlik açığı türleri ile sınırlıdır.

Statik Analiz Teknikleri

Statik kod analizi, yazılımın içerdiği hataların yazılım çalıştırılmadan tespit edilebilmesi ve sahip olduğu çeşitli özelliklerin çıkarılabilmesidir. Statik tekniklerin amacı, çalışma zamanında meydana gelen hataları ve güvenlik açıklarını ortaya çıkarmaktır. Statik teknikler kural tabanlı analiz, kod benzerliği/klonu tespiti ve sembolik yürütme gibi temel olarak kaynak kodun analizine dayalıdır. Güvenlik tespiti gerçekleştirilirken kaynak kod tabanlı

statik analiz önemli bir yaklaşım olmaktadır. Bu yaklaşımda desen tabanlı ve kod benzerliğine dayalı yöntemler kullanılmaktadır. Desen tabanlı yöntemlerde güvenlik açıklarını temsil etmek için uzman kişilerce güvenlik açığı özelliklerinin tanımlanması gerekir. Kod benzerliğine dayalı yöntemlerde ise program kod parçalarına bölünerek, belirteçler, ağaçlar ve grafikler dahil olmak üzere her bir kod parçası soyut biçimde temsil edilir ve soyut temsiller aracılığıyla kod parçaları arasındaki benzerlikler hesaplanır. Kod benzerliğine dayalı yaklaşımın avantajı, hedef programlarda aynı güvenlik açığını tespit etmek adına tek bir güvenlik açığı kodu örneğinin yeterli olmasıdır.

Statik kod analizi, konunun uzmanları tarafından kodun gözden geçirilmesi veya statik kod analiz araçları ile otomatik olarak yapılabilmesi olmak üzere iki şekilde uygulanabilmektedir. İnsanlar tarafından kodun gözden geçirilmesi şeklinde uygulanan yöntemlerde, fonksiyonel hatalar ortaya çıkarılabilir. Statik analiz araçları ile yapılan yöntemlerde ise atama ve denetim gibi programlama hataları tespit edilebilir. Yapılan araştırmalara göre, insanlarda zamanla oluşabilen dikkat dağınıklığı nedeniyle kod gözden geçirme yöntemleri, hataların gözden kaçırılması, nitelikli bir iş gücü gerektirmesi ve hız faktöründen ötürü maliyet açısından olumsuz karşılanmasına sebep olmaktadır (Kalay, 2009).

Dinamik Analiz Teknikleri

Dinamik kod analizinde, uygulama izole bir ortamda yürütülerek uygulamanın davranışları izlenmektedir. Kod yürütüldüğü için kötücül davranışlar gözlemlenebilmektedir. Ağ etkinliği izleme, dosya değişikliğini izleme ve sistem çağrıları gibi yöntemlerle dinamik analizler gerçekleştirilebilir (Utku & Doğru, 2016).

Hibrit Analiz Yöntemleri

Dinamik ve statik analiz teknikleri kendilerine göre avantajlara sahip olsalar da tek başlarına yeterli değildir. Hibrit analiz

yöntemlerinde, otomatize araçların bulamadığı açıkların tespit edilmesi amacıyla statik ve dinamik analizin yanında gözden geçirme yöntemi de dahil edilerek güvenlik testleri gerçekleştirilmektedir.

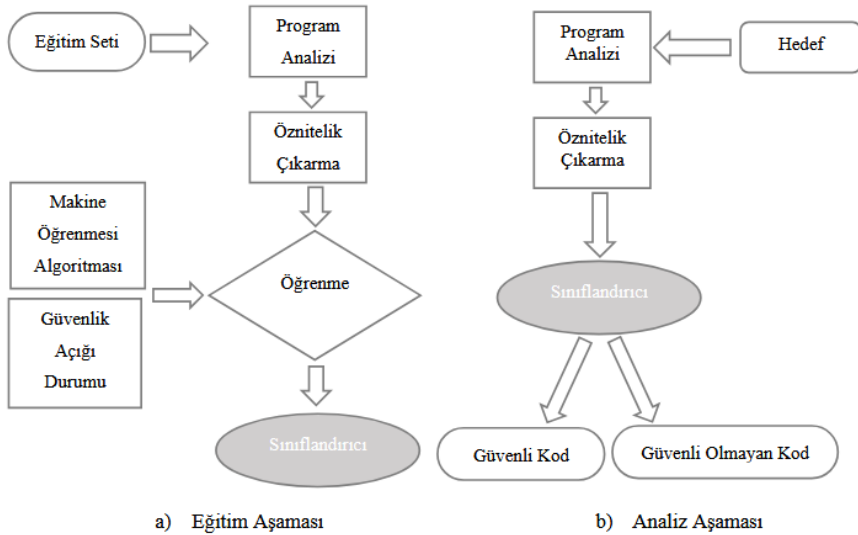
KAYNAK KODDAN ÖZELLİK ÇIKARMA YÖNTEMLERİ

Güvenlik açığı içeren yazılımlar, genellikle birbiriyle karmaşık ilişkiler barındıran kod bloklarından oluşmaktadır. Bu yapısal karmaşıklık, makine öğrenmesi algoritmalarının kaynak koda doğrudan uygulanmasını güçleştirmektedir. Bu nedenle, güvenlik açığı tespitine yönelik makine öğrenmesi tabanlı yaklaşımlarda ilk adım olarak program analizi gerçekleştirilmekte ve ardından koddan anlamlı özellikler çıkarılmaktadır. Özellik çıkarımı tamamlandıktan sonra, elde edilen sayısal temsiller kullanılarak makine öğrenmesi algoritmaları ile güvenli ve güvenli olmayan kodların ayrımı yapılabilmektedir. Bu süreç Şekil 2’de gösterildiği gibi eğitim ve analiz aşamaları olmak üzere iki kısımda incelenebilir.

Makine öğrenmesine dayalı yazılım güvenlik açığı analizi süreci genel olarak iki ana aşamada ele alınmaktadır: eğitim ve analiz. Eğitim aşamasında, güvenli ve güvenlik açığı barındıran kaynak kodlar program analizi ve özellik çıkarımı süreçlerinden geçirilerek sayısal vektörlere dönüştürülür. Bu vektörler, sınıflandırma veya tahmin modellerinin eğitilmesinde kullanılır. Analiz aşamasında ise, daha önce görülmemiş hedef kodlar aynı özellik çıkarım adımlarından geçirilir ve eğitilmiş model yardımıyla

bu kodların güvenlik açığı içerip içermediği belirlenir. Şekil 2 (a)'da yazılım güvenlik açığı analizi sınıflandırıcının aşamaları, Şekil 2 (b)'de hedef kodun güvenlik açığı içerip içermediğinin tespiti sırasında analiz süreci anlatılmaktadır. Her iki süreçte de yer alan özellik çıkarma aşamasında yazılım kodlarına özellik çıkarma metotları uygulanarak kodlar sayısal vektörlere dönüştürülmektedir.

Şekil 2 Makine öğrenmesine dayalı yazılım güvenlik açığı analizi çerçevesi
(Gong et al., 2016)



Makine öğrenmesi tabanlı güvenlik açığı analizi yöntemleri, farklı bakış açılarına göre sınıflandırılabilir (bakınız Şekil 3). Program analizi açısından yöntemler; sözcüksel analiz, sözdizimsel analiz ve anlamsal analiz olmak üzere üç grupta ele alınmaktadır. Sözcüksel analizde kaynak kod, anahtar kelimeler, operatörler ve belirteçler (token) seviyesinde incelenirken; sözdizimsel analizde kodun yapısal bileşenleri ve dil kurallarına uygunluğu değerlendirilir. Çoğu durumda sözdizimsel analiz, sözcüksel analizi de kapsamaktadır. Anlamsal analiz ise programın davranışını, veri ve

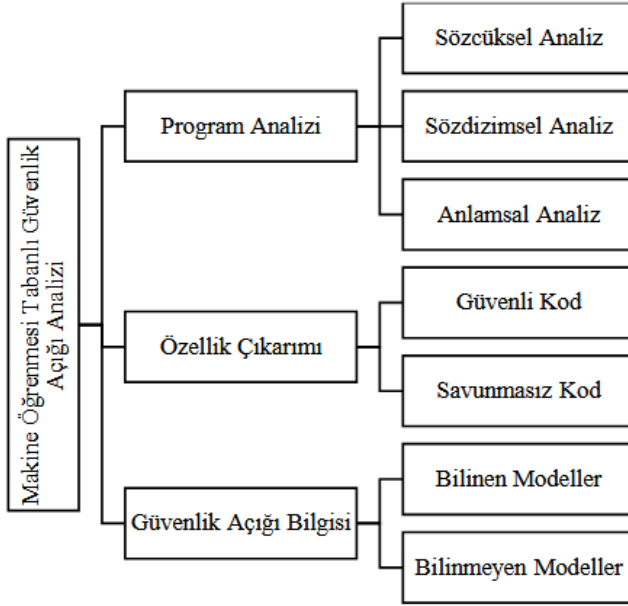
kontrol akışlarını dikkate alarak daha derinlemesine bir inceleme sunmaktadır.

Özellik çıkarımı açısından bakıldığında, yöntemler güvenli ve güvenli olmayan kodlardan ayırt edici niteliklerin çıkarılmasına odaklanmaktadır. Güvenlik açığı bilgisi bakımından ise yaklaşımlar, bilinen güvenlik açığı örüntülerine dayalı modeller ve daha önce görülmemiş güvenlik açıklarını keşfetmeyi amaçlayan geliştirilebilir modeller olarak ikiye ayrılmaktadır.

(Yamaguchi et al., 2012), önceki çalışmalarını temel alarak, kaynak kodun denetimi sırasında güvenlik analistlerine yardımcı olacak statik analize dayalı bir yaklaşım önermişlerdir. Bu yaklaşıma göre, koddan AST çıkarıp, bu ağaçlardaki yapısal örüntüleri belirleyerek, koddaki her işlevi bu örüntülerin bir karışımı olarak tanımlamışlardır. C fonksiyonu ve bu fonksiyonun seri hale getirilmiş AST örneği Şekil 4 ve Şekil 5’te gösterilmiştir. Bu temsil, bilinen bir güvenlik açığını ayrıştırmalarına ve bunu makine öğrenmesi ile kod tabanında tahmin etmelerine olanak tanımaktadır.

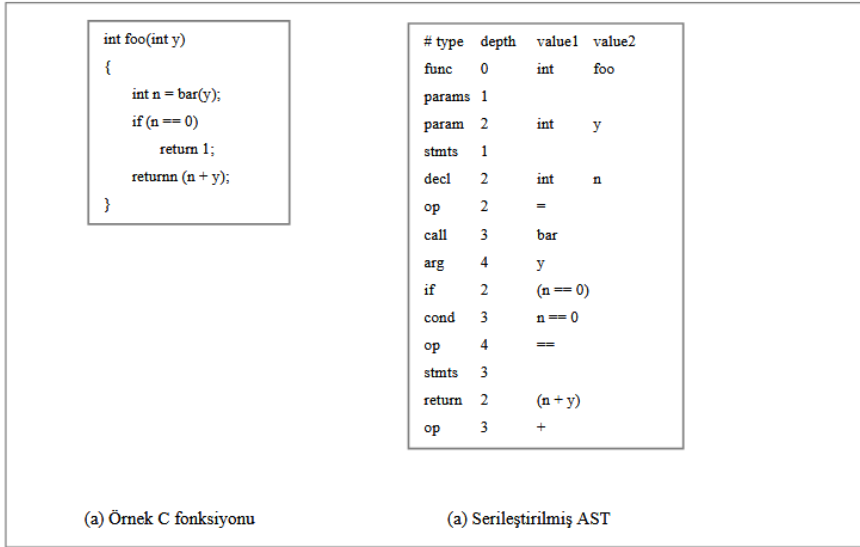
Kaynak koddan özellik çıkarımında doğal dil işleme (NLP) tekniklerinden de yoğun biçimde yararlanılmaktadır. Doğal dil işleme, insanların iletişimde kullandığı yazılı ve sözlü ifadelerin bilgisayarlar tarafından işlenmesini amaçlayan bir alandır. Bu tekniklerin temelinde, metinsel verilerin makinelerin anlayabileceği sayısal temsillere dönüştürülmesi yer almaktadır. Kaynak kod da belirli bir söz dizimine ve anlamsal yapıya sahip olması nedeniyle, doğal dil işleme yöntemleri kullanılarak analiz edilebilmektedir.

Şekil 3 Makine öğrenmesi tabanlı güvenlik açığı analizi yöntemleri sınıflandırması (Gong et al., 2016)

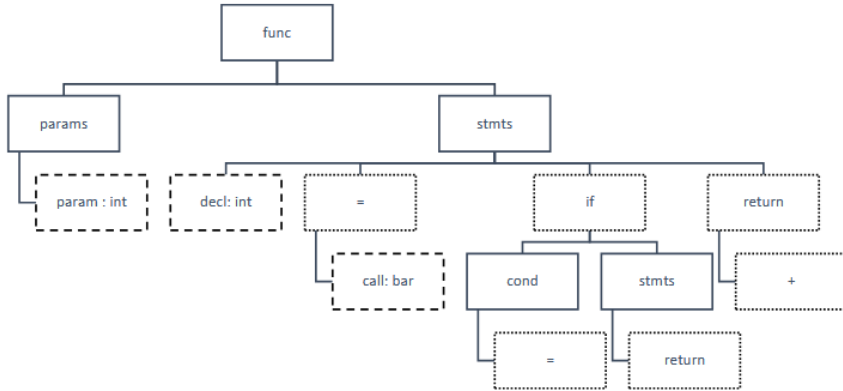


Kaynak koddan özellik çıkarımı sürecinde, doğal dil işleme (Natural Language Processing – NLP) tekniklerinden yaygın biçimde yararlanılmaktadır. Doğal dil işleme, insanların iletişim amacıyla kullandığı yazılı ve sözlü ifadelerin bilgisayar sistemleri tarafından otomatik olarak işlenmesini ve anlamlandırılmasını hedefleyen disiplinlerarası bir alandır. Bu alanın temelinde, metinsel verilerin istatistiksel ve öğrenmeye dayalı yöntemler aracılığıyla makinelerin işleyebileceği sayısal temsillere dönüştürülmesi yer almaktadır. Kaynak kod her ne kadar doğal dilden farklı olarak katı sözdizim kurallarına sahip olsa da, belirli bir söz dizimi, anlamsal yapı ve bağlam ilişkisi içermesi nedeniyle doğal dil işleme yöntemleri kullanılarak analiz edilebilmektedir. Bu benzerlik, NLP tabanlı yaklaşımların yazılım güvenliği alanında etkili bir biçimde kullanılmasına olanak tanımaktadır.

Şekil 4 C fonksiyonu ve seri hale getirilmiş AST örneği (Yamaguchi et al., 2012)



Şekil 5 API düğümleri (kesikli) ve sözdizimi düğümleri (noktalı) içeren soyut sözdizimi ağacı.



Bu çerçevede kelime vektörleri, kaynak kod içerisinde yer alan anahtar kelimelerin, fonksiyon adlarının, değişken isimlerinin ve diğer belirteçlerin sayısal olarak temsil edilmesini sağlamaktadır. Kelime vektörleri, yalnızca sözcüklerin frekans bilgilerini değil, aynı zamanda bulundukları bağlamı ve diğer sözcüklerle olan

ilişkilerini de dikkate alarak anlamsal temsiller üretmektedir. Böylece, benzer işlemlere sahip kod parçaları veya benzer programlama desenlerini içeren yapılar, vektör uzayında birbirine yakın konumlandırılmaktadır. Bu özellik, farklı geliştiriciler tarafından yazılmış olsa bile aynı amaca hizmet eden veya benzer güvenlik riskleri taşıyan kod bölümlerinin tespit edilmesini kolaylaştırmaktadır. Sonuç olarak, makine öğrenmesi modelleri kodlar arasındaki yapısal ve anlamsal benzerlikleri daha etkin bir şekilde öğrenebilmekte ve güvenlik açığı içeren örüntüleri ayırt edebilme kabiliyetini artırmaktadır.

Word2Vec, FastText ve GloVe gibi yöntemler, bu tür vektörel temsillerin oluşturulmasında yaygın olarak kullanılan kelime gömme (embedding) teknikleri arasında yer almaktadır. Bu yöntemler, kaynak kodu oluşturan sözcükleri veya belirteçleri yüksek boyutlu vektörler hâlinde temsil ederek, kodun anlamsal özelliklerini sayısal bir uzaya taşımaktadır. Ancak bu yaklaşımlar sonucunda elde edilen vektör temsilleri çoğunlukla yüksek boyutlu özellik uzayları oluşturmakta ve her bir boyut, modele farklı ölçüde katkı sağlamaktadır. Tüm özelliklerin doğrudan modele dâhil edilmesi, hesaplama maliyetinin artmasına ve modelin gereksiz veya gürültülü bilgiler öğrenmesine yol açabilmektedir. Bu durum, özellikle sınırlı veri kümelerinde aşırı uyum (overfitting) riskini artırmaktadır.

Bu nedenle, hangi özelliklerin modelleme sürecinde kullanılacağına belirlenmesi kritik bir aşama olarak karşımıza çıkmaktadır. Bu aşama özellik seçimi olarak adlandırılmakta olup, modelin performansını en fazla artıran ve ayırt edici gücü yüksek olan özniteliklerin belirlenmesini amaçlamaktadır. Özellik seçimi sayesinde, hem daha sade ve yorumlanabilir modeller elde edilmekte hem de gereksiz özelliklerin elenmesiyle eğitim süresi ve hesaplama maliyeti azaltılabilmektedir.

Özellik seçimi sürecinde genellikle iki temel bileşen yer almaktadır: öznitelik değerlendirici ve arama yöntemi. Öznitelik değerlendirici, veri kümesindeki her bir özelliğin çıktı değişkeni ile olan ilişkisini istatistiksel veya bilgi kuramsal ölçütler aracılığıyla değerlendirerek, ilgili özelliğin modele katkı düzeyini belirlemektedir. Arama yöntemi ise, bu değerlendirmeler doğrultusunda en uygun özellik alt kümelerini bulabilmek amacıyla farklı öznitelik kombinasyonlarını sistematik olarak denemekte veya bu kombinasyonlar arasında gezinmektedir. Uygun bir özellik seçimi stratejisinin uygulanması, makine öğrenmesi modellerinin doğruluğunu ve genelleme yeteneğini artırırken, aynı zamanda daha verimli ve ölçeklenebilir çözümlerin geliştirilmesine katkı sağlamaktadır.

Bu bağlamda, kaynak koddan anlamsal açıdan zengin ve ayırt edici özelliklerin elde edilmesinde kelime gömme (embedding) tabanlı yaklaşımlar önemli bir rol oynamaktadır. Bu yaklaşımlar, kod içerisindeki anahtar kelimeler, fonksiyon adları ve değişken isimleri arasındaki bağlamsal ilişkileri öğrenerek, kodun anlamsal yapısını vektör uzayında temsil edilebilir hâle getirmektedir. Kelime gömme yöntemleri arasında Word2Vec, FastText ve GloVe öne çıkmakta olup, her biri farklı öğrenme stratejileri kullanarak sözcüklerin sayısal temsillerini oluşturmaktadır. Word2Vec ve FastText yöntemleri yerel bağlam bilgisine dayalı olarak kelimelerin birlikte kullanım örüntülerini öğrenirken, GloVe yöntemi daha küresel istatistikleri dikkate alarak kelime temsilleri üretmektedir. Bu farklı yaklaşımlar, kaynak kodun yapısal ve anlamsal özelliklerinin çeşitli açılardan modellenmesine olanak tanımakta ve güvenlik açığı tespitinde kullanılabilecek zengin öznitelik kümelerinin elde edilmesini sağlamaktadır. Aşağıda, bu yöntemlerin her biri ayrıntılı olarak ele alınarak, kaynak koddan özellik çıkarımı sürecindeki rolleri ve sağladıkları avantajlar açıklanmaktadır.

Word2Vec

Word2vec, (Mikolov et al., 2013) tarafından önerilmiş olan yapay sinir ağıları tabanlı bir kelime vektörü temsil yöntemidir. Bu yöntemde ilk olarak kelimeler vektörlere dönüştürülmektedir. Sonraki aşamada ise bu vektörler arasındaki uzaklıklara bakılarak kelimeler arasındaki benzerlikler hesaplanmaktadır. Böylelikle metnin içinde geçen cümleler vektörel olarak temsil edilmektedir.

Bag of Words (BOW, Kelime Torbası), doğal dil işleme alanında kullanılan bir tekniktir. Bu teknik, metin verilerini sayısal bir formatta temsil etmek için kullanılmaktadır. Kelime Torbası tekniği, bir belgeyi kelimelerin bir koleksiyonu olarak ele alır ve kelime sırasını ve bağlamını dikkate almadan, belgedeki her kelimenin frekansını sayar. Bu, belgenin içeriğini basit ve verimli bir şekilde temsil eden bir "kelime torbası" oluşturur. Bir kelime torbası temsilini oluşturmak için, metin öncelikle gereksiz sözcükler (örneğin "ve", "bir", "o" gibi), noktalama işaretleri ve diğer gereksiz unsurlar çıkarılır. Kalan kelimeler, kök veya lemmatizasyon yöntemi kullanılarak temel formlarına indirgenir. Ön işleme işleminden sonra, belgedeki tüm benzersiz kelimelerin bir sözlüğü oluşturulur ve her kelimenin belgedeki frekansı sayılır. Bu sayısal temsil, makine öğrenimi, duyarlılık analizi, metin sınıflandırması gibi çeşitli doğal dil işleme görevlerinde kullanılabilir. Word2vec yönteminde Continious Bag of Word (CBOW) ve Skip-Gram (SG) öğrenme mimarileri kullanılarak kelime koordinatları hesaplanmaktadır. Her iki öğrenme mimarisinde giriş parametresi olarak verilen pencere boyutu (window size) isimli bir parametre bulunmaktadır. CBOW mimarisinde ilgili kelimeye komşu olan kelimelere yani sağındaki ve solundaki pencere boyutu kadar olan kelimelere bakılarak ilgili kelimenin tahmini gerçekleştirilmektedir. SG mimarisinde ise pencere boyutu kadar olan komşu kelimelerin tahmini için mevcut kelimeye bakılmaktadır. Yani kısaca Şekil 6'da gösterildiği gibi CBOW mimarisi bağlama göre mevcut kelimeyi tahmin ederken, SG

mevcut kelime ile çevredeki kelimeleri tahmin etmektedir. Kelimelerin vektörlere dönüştürülmesi sırasında derlem frekansı az olan kelimelerde SG mimarisi daha iyi sonuçlar üretirken, derlem frekansı büyük olan kelimelerde CBOW daha iyi sonuçlar üretmektedir.

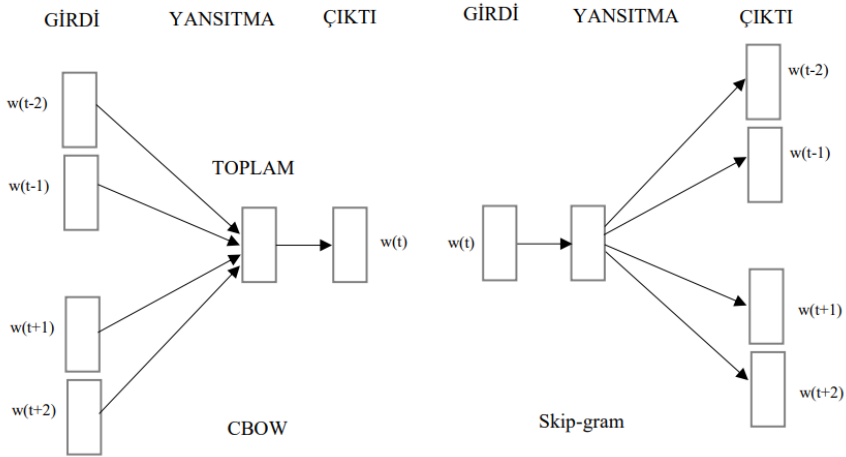
Word2vec modelinde girdi metin derlemi, çıktı ise derlemdeki sözcükleri temsil eden özellik vektörlerinden oluşan bir dizidir. Metin derin sinir ağlarının anlayabileceği sayısal bir forma dönüştürülür. Word2vec yöntemi genlere, kodlara, beğenilere, oynatma listelerine, sosyal medya grafiklerine ve diğer sözlü veya sembolik dizilere uygulanabilmektedir. Bu yöntemde sözcükler arasındaki benzerlikler matematiksel olarak tespit edilmekte ve benzer kelimelerin vektörleri gruplanmaktadır. Yeterli veri, kullanım ve bağlam verildiğinde, geçmişteki görünümlere dayalı olarak, bir kelimenin anlamı ve diğer kelimelerle olan ilişkisinin kurulması veya bir belgenin bir konuya bağlı olarak sınıflandırılmasında doğru tahminler yapabilmektedir. Kelimeler arasındaki ilişkilerde “erkek”, “adam” ile “kız”, “kadın” gibi kelimelerdeki anlamların benzerliğine, “erkek”, “kadın” ile “kral”, “kraliçe” gibi kelimeler arasındaki anlamsal ilişkilerin çıkarımına, kümelerde ise bilimsel araştırma, e-ticaret, müşteri ilişkileri yönetimi gibi çeşitli alanlarda arama, duyarlılık analizi ve önerilerin temelini oluşturabilmektedir.

FastText

Word2Vec’in geliştirilmiş bir türevidir. FastText’te her kelime doğrudan bir vektör olarak ele alınmaz; bunun yerine kelime içindeki karakter n-gram’ları düzeyinde temsil öğrenilir. Bu sayede, daha önce hiç görülmemiş bir kelime bile (örneğin türetilmiş bir değişken adı) alt parçalara bölünerek (n-gram’larına ayrılarak) anlamlı bir vektör elde edilebilir. FastText özellikle dilbilgisel açıdan zengin ve çok ek alan dillerde veya yazım varyasyonları olan ortamlarda (Türkçe gibi, veya programlama dillerinde farklı

adlandırmalar) avantaj sağlar. Kaynak kod analizinde FastText kullanımı, örneğin readFile gibi bir fonksiyon ismini read ve file alt parçalarına ayırıp, bu alt birimlerin vektörlerinden türeterek temsil etme olanağı sağlar. Bu durumda model, readFile görmese bile read kelimesini ve file kelimesini biliyorsa, onların birleşiminden mantıklı bir vektör oluşturabilir. Bu, özellikle büyük bir sözlüğün olduğu (çok sayıda farklı fonksiyon ve değişken ismi içeren) kod tabanlarında genelleme kabiliyetini artırır.

Şekil 6 CBOW ve SG mimarileri (Mikolov et al., 2013)



GloVe

Metindeki kelime çiftlerinin ortak görülme istatistiklerine dayalı bir modeldir. Word2Vec ve türevleri gibi kayan pencere üzerinden örnekler üretmek yerine, tüm veri kümesi üzerinden kelime eşleşmelerinin olasılıklarını hesaplar ve bu olasılıkları faktörize ederek vektörleri öğrenir. GloVe modeli de kaynak koduna uygulanabilir; örneğin, büyük bir kod deposunda hangi tokenların sıklıkla aynı fonksiyonlarda veya dosyalarda birlikte geçtiğine

bakarak vektör temsilleri oluşturulabilir. GloVe, global istatistikleri kullandığı için nadir kelimeler için veriye yeterince sahipse istikrarlı sonuçlar verebilir, ancak genelde Word2Vec/FastText gibi yerel bağlam odaklı yöntemler kadar ince ayrıntı yakalama becerisi olmayabilir.

SONUÇ

Bu bölümde yazılım güvenlik açıklarının yapay zekâ ile tespitine yönelik temel kavramlar ve teknikler ele alındı. Statik ve dinamik analiz yöntemlerinin makine öğrenmesiyle desteklenmesi, kodun vektörel temsillerle incelenmesi ve yapısal özelliklerin öğrenilmesi gibi konuların, günümüzün artan yazılım güvenliği ihtiyacı karşısında ne denli önemli olduğu vurgulandı. Akademik literatürdeki çalışmalar da göstermektedir ki, doğru özellik çıkarımı ve güçlü öğrenme modelleri bir araya geldiğinde, güvenlik açıklarını tespit etmede uzmanlara yardımcı olacak veya bazı durumlarda onların dahi gözünden kaçabilecek zafiyetleri yakalayacak çözümler üretmek mümkün olabilmektedir. Bu alanda ilerleyen araştırmalar, daha az yanlış alarm veren, açıklanabilir sonuçlar üreten ve farklı programlama dillerine/genel kod tabanlarına uyarlanabilir yaklaşımlar geliştirmeye odaklanmaktadır. Böylece yazılım güvenliği alanında proaktif ve akıllı bir savunma hattı oluşturulması hedeflenmektedir.

Kaynakça

- Avgerinos, T., Cha, S. K., Rebert, A., Schwartz, E. J., Woo, M., & Brumley, D. (2014). Automatic Exploit Generation. *Communications of the ACM*, 57(2), 74–84. <https://doi.org/10.1145/2560217.2560219>
- Caruana, G., & Li, M. (2012). A Survey of Emerging Approaches to Spam Filtering. *ACM Computing Surveys*, 44(2), 9. <https://doi.org/10.1145/2089125.2089129>
- Dowd, M., McDonald, J., & Schuh, J. (2006). *The Art of Software Security Assessment - Identifying and Preventing Software Vulnerabilities*.
- Falliere, N., Murchu, L. O., & Chien, E. (2011). W32. Stuxnet Dossier, Symantec Security Response, Version 1.4. *Symantec Security Response*, 4(February), 1–69.
- Garcia-Teodoro, P., Diaz-Verdejo, J., Macia-Fernandez, G., & Vazquez, E. (2009). Anomaly-Based Network Intrusion Detection: Techniques, Systems and Challenges. *Computers and Security*, 28(1), 18–28. <https://doi.org/10.1016/j.cose.2008.08.003>
- Godefroid, P., Levin, M. Y., & Molnar, D. (2012). SAGE: Whitebox Fuzzing for Security Testing. *Communications of the ACM*, 55(3), 40–44. <https://doi.org/10.1145/2093548.2093564>
- Gong, J., Kuang, X. H., & Liu, Q. (2016). Survey on Software Vulnerability Analysis Method Based on Machine Learning. *2016 IEEE 1st International Conference on Data Science in Cyberspace, DSC 2016*, 642–647. <https://doi.org/10.1109/DSC.2016.33>
- Guzella, T. S., & Caminhas, W. M. (2009). A Review of Machine Learning Approaches to Spam Filtering. *Expert Systems with*

Applications, 36(7), 10206–10222.
<https://doi.org/10.1016/j.eswa.2009.02.037>

Jason Brownlee. (2022). *What Is Natural Language Processing?*
<https://machinelearningmastery.com/natural-language-processing/>

Johnson, S. (1977). *Computer Science Technical Report 65, Bell Laboratories.*

Kalay, A. (2009). Statik Kod Analizinin Yazılım Geliştirme Sürecindeki Yeri ve Yazılım Kalitesine Etkisi. 4. *ULUSAL YAZILIM MÜHENDİSLİĞİ SEMPOZYUMU - UYMS'09*, 211–218.

Kumar, C., & Yadav, D. K. (2017). Software Defects Estimation Using Metrics of Early Phases of Software Development Life Cycle. *International Journal of System Assurance Engineering and Management*, 8, 2109–2117. <https://doi.org/10.1007/s13198-014-0326-2>

Li, Z., Zou, D., Xu, S., Ou, X., Jin, H., Wang, S., Deng, Z., & Zhong, Y. (2018). VulDeePecker: A Deep Learning-Based System for Vulnerability Detection. *Network and Distributed Systems Security (NDSS) Symposium 2018*. <https://doi.org/10.14722/ndss.2018.23158>

Lin, G., Zhang, J., Luo, W., Pan, L., Xiang, Y., Member, S., Vel, O. De, & Montague, P. (2018). Cross-Project Transfer Representation Learning for Vulnerable Function Discovery. *IEEE Transactions on Industrial Informatics*, 14(7), 3289–3297.

Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing.*

Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S., & Weaver, N. (2003). Inside the Slammer Worm. *IEEE Security and Privacy*, 1(4), 33–39. <https://doi.org/10.1109/MSECP.2003.1219056>

- Newsome, J., & Song, D. (2005). Dynamic Taint Analysis for Automatic Detection, Analysis, and Signature Generation of Exploits on Commodity Software. *In The Network and Distributed System Security Symposium (NDSS'05)*, 5(3–4).
- NIST. (2025). *National Vulnerability Database*. National Vulnerability Database - Search and Statistics. https://nvd.nist.gov/vuln/search/statistics?form_type=Advanced&results_type=statistics&search_type=all&isCpeNameSearch=false
- Özdemir, E., & Türkoğlu, İ. (2022). Yazılım Güvenlik Açıklarının Evrimsel Sınır Ağları (CNN) ile Sınıflandırılması. *Fırat Üniversitesi Müh. Bil. Dergisi*, 34(2), 517–529. <https://doi.org/10.35234/fumbd.1076870>
- Russell, S., & Norvig, P. (2009). Artificial Intelligence A Modern Approach Third Edition. In *Pearson*.
- Sarıman, G., & Küçüksille, E. U. (2016). Web Servislerinin Yazılım Güvenlik Testleri İçin Önerilen Hibrit Yaklaşım. *SDU International Journal of Technological Science*, 8(2), 1–14.
- Shannon, C., & Moore, D. (2004). The Spread of the Witty Worm. *IEEE Security and Privacy*, 2(4), 46–50. <https://doi.org/10.1109/MSP.2004.59>
- Shin, Y., Meneely, A., Williams, L., & Osborne, J. A. (2011). Evaluating Complexity, Code Churn, and Developer Activity Metrics as Indicators of Software Vulnerabilities. *IEEE Transactions on Software Engineering*, 37(6), 772–787. <https://doi.org/10.1109/TSE.2010.81>
- Sutton, M., Greene, A., & Amini, P. (2007). *Fuzzing - Brute Force Vulnerability Discovery*.

- Utku, A., & Doğru, İ. A. (2016). Mobil Kötücül Yazılımlar ve Güvenlik Çözümleri Üzerine Bir İnceleme. *Gazi University Journal of Science GU J Sci Part C*, 4(2), 49–64.
- Wei, H., Lin, G., Li, L., & Jia, H. (2021). A Context-Aware Neural Embedding for Function-Level Vulnerability Detection. *Algorithms*, 14(335). <https://doi.org/10.3390/a14110335>
- Yamaguchi, F., Lindner, F. F. X., & Rieck, K. (2011). Vulnerability Extrapolation: Assisted Discovery of Vulnerabilities using Machine Learning. *2nd USENIX Workshop on Health Security and Privacy (HealthSec '11)*, 111–120.
- Yamaguchi, F., Lottmann, M., & Rieck, K. (2012). Generalized Vulnerability Extrapolation Using Abstract Syntax Trees. *ACSAC '12: Proceedings of the 28th Annual Computer Security Applications Conference*, 359–368. <https://doi.org/10.1145/2420950.2421003>

BÖLÜM 4

TIBBİ GÖRÜNTÜLERDE DERİN ÖĞRENME UYGULAMALARI: BEYİN TÜMÖRÜ MRI VERİLERİ ÜZERİNDEN BİR İNCELEME

YUSUF ÇELİK¹

Giriş

Tıbbi görüntüleme teknolojilerindeki gelişmeler, hastalıkların erken teşhisi ve doğru sınıflandırılması açısından kritik bir rol oynamaktadır. Manyetik rezonans görüntüleme (MRI), yüksek yumuşak doku kontrastı sunması nedeniyle özellikle beyin tümörlerinin değerlendirilmesinde yaygın olarak kullanılan başlıca görüntüleme yöntemlerinden biridir. Bununla birlikte, beyin tümörlerinin MRI görüntüleri üzerinden sınıflandırılması; tümör tipleri arasındaki morfolojik benzerlikler, sınıf içi heterojenlik ve görüntüleme koşullarındaki farklılıklar nedeniyle deneyimli uzmanlar için dahi zorlayıcı olabilmektedir. Bu durum, otomatik ve güvenilir bilgisayar destekli karar destek sistemlerine olan ihtiyacı artırmaktadır (Litjens vd., 2017).

Makine öğrenmesi ve özellikle derin öğrenme tabanlı yaklaşımlar, son yıllarda tıbbi görüntü analizi alanında önemli bir paradigma

¹ Dr.Öğr.Üyesi, Munzur Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0002-7859-7543

değişimi yaratmıştır. Derin öğrenme modelleri, manuel özellik çıkarımına duyulan ihtiyacı ortadan kaldırarak görüntülerden ayırt edici temsilleri doğrudan öğrenebilmekte ve bu sayede geleneksel yöntemlere kıyasla daha yüksek performans sunabilmektedir (Jordan ve Mitchell, 2015; Litjens vd., 2017). Bu bağlamda, konvolüsyonel sinir ağları (KSA), yerel uzamsal örüntüleri etkili biçimde modelleyebilme yetenekleri sayesinde beyin tümörü MRI görüntülerinin sınıflandırılması probleminde en yaygın kullanılan derin öğrenme mimarileri arasında yer almaktadır.

Literatürde yer alan çok sayıda çalışma, KSA tabanlı derin öğrenme modellerinin MRI görüntülerinde beyin tümörü tespiti ve sınıflandırma görevlerinde yüksek doğruluk ve kararlılık sağladığını göstermektedir. Özellikle KSA tabanlı yaklaşımların, klasik makine öğrenmesi yöntemleriyle karşılaştırıldığında daha güçlü ve genellenebilir temsiller üretebildiği rapor edilmiştir (Pereira vd., 2016; Saeedi vd., 2023). Bu sonuçlar, derin öğrenme temelli yaklaşımların klinik karar destek sistemleri için önemli bir potansiyel sunduğunu ortaya koymaktadır.

Bununla birlikte, derin öğrenme modellerinin başarımı büyük ölçüde eğitilen veri miktarı ve çeşitliliğine bağlıdır. Tıbbi görüntüleme alanında etiketli veri üretiminin maliyetli ve zaman alıcı olması, sıfırdan eğitilen derin ağların genelleme yeteneğini sınırlayabilmektedir. Bu sorunun üstesinden gelmek amacıyla geliştirilen transfer öğrenme yaklaşımı, büyük ölçekli veri setleri üzerinde önceden eğitilmiş modellerin hedef probleme uyarlanmasıyla olanak tanıyarak sınırlı veri koşullarında daha kararlı ve yüksek performanslı çözümler sunmaktadır (Pan ve Yang, 2010; Khan vd., 2019).

Son dönemde, Transformer mimarisinin görüntü alanına uyarlanmasıyla geliştirilen Vision Transformer (ViT) modelleri, self-attention mekanizması sayesinde görüntü içerisindeki küresel bağlamı doğrudan modelleyebilme avantajı sunmaktadır. KSA'ların

yerel alıcı alanlara dayalı yapısının aksine, ViT mimarisi görüntüdeki uzak bölgeler arasındaki ilişkileri etkin biçimde öğrenebilmekte ve bu yönüyle karmaşık tıbbi görüntü sınıflandırma problemleri için güçlü bir alternatif oluşturmaktadır (Dosovitskiy vd., 2020; Shamshad vd., 2023). Bu yaklaşım, özellikle uzun menzilli bağımlılıkların önemli olduğu durumlarda dikkate değer bir temsil gücü sağlamaktadır.

Bu bölümde, dört sınıflı beyin tümörü MRI görüntülerinin sınıflandırılması problemi farklı derin öğrenme paradigmaları karşılaştırmalı olarak incelenmektedir. Sıfırdan tasarlanan temel bir KSA modeli, transfer öğrenme yaklaşımı ile eğitilen ön eğitimli KSA mimarileri ve Vision Transformer tabanlı bir model aynı veri bölünmesi altında değerlendirilmiştir. Amaç, klasik konvolüsyonel yapılardan transformer tabanlı modellere uzanan bu mimari evrimin performans üzerindeki etkisini ortaya koymak ve derin öğrenme yöntemlerinin tıbbi görüntü sınıflandırmadaki güçlü ve sınırlı yönlerini öğretici bir bakış açısıyla sunmaktır.

Derin Öğrenme

Derin öğrenme, çok katmanlı yapay sinir ağları kullanarak veriden anlamlı temsilleri otomatik olarak öğrenebilen bir makine öğrenmesi yaklaşımıdır. Klasik makine öğrenmesi yöntemlerinde özelliklerin insan tarafından tanımlanması gerekirken, derin öğrenme modelleri verinin hiyerarşik yapısını katmanlar boyunca kendisi keşfeder. Bu sayede görüntü, ses, metin, biyomedikal sinyaller ve tıbbi görüntüler gibi karmaşık veri türlerinde geleneksel yöntemlere kıyasla çok daha yüksek performans elde edilebilmektedir.

Derin sinir ağlarının erken dönemlerde karşılaştığı en büyük sorunlardan biri kaybolan gradyan problemidir. Çok katmanlı modellerde hata sinyali geriye doğru yayılırken gradyanın giderek küçülmesi, ağın önceki katmanlarının öğrenmesini zorlaştırmıştır. Hinton ve arkadaşlarının sunduğu katman bazlı ön eğitim yaklaşımı

(Hinton vd., 2006), ReLU ve türevlerinin geliştirilmesi (Glorot vd., 2011; Maas, 2013) ile artık ağların daha derin biçimde eğitilebilmesi mümkün hâle gelmiş; ResNet gibi artık bağlantılar içeren modern yapılar (He vd., 2016) bu sorunu büyük ölçüde ortadan kaldırmıştır.

Günümüzde derin öğrenme, farklı veri türleri ve problem alanları için çeşitli paradigmlar altında uygulanmaktadır. Aşağıda derin öğrenmenin temel bazı paradigmları özetlenmiştir.

Konvolüsyonel Sinir Ağları

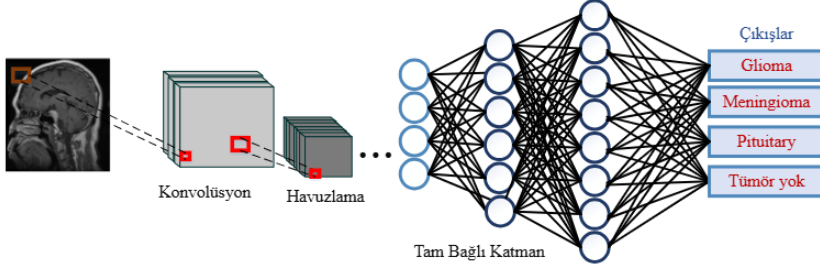
Konvolüsyonel sinir ağları (KSA), özellikle ızgara yapısındaki verilerin, bir boyutlu zaman serileri, iki boyutlu görüntüler ve hacimsel tıbbi veriler gibi işlenmesi için tasarlanmış derin öğrenme mimarileridir. Bu mimariler, giriş verisi üzerinde konvolüsyon ve havuzlama işlemlerini uygulayarak hiyerarşik özellik çıkarımı yapabilmekte ve birçok uygulama alanında başarılı sonuçlar sunmaktadır (Goodfellow vd., 2016). KSA'larda kullanılan bu katmanlı yapı, insan görsel sisteminin çalışma prensiplerinden esinlenilmiş olup, özellikle çok katmanlı yapay sinir ağlarına göre daha verimli bir yerel özellik öğrenme mekanizması sağlar (Kim, 2017; O'Shea ve Nash, 2015).

Geleneksel yapay sinir ağlarında her nöron girişin tamamı ile bağlantılıdır. Bu durum büyük boyutlu görüntülerde hem hesaplama yükünü artırmakta hem de öğrenmeyi zorlaştırmaktadır. Konvolüsyonel ağlar ise yerel alıcı alanlar kullanarak yalnızca girişin küçük bir bölgesini işler; böylece hem hafıza gereksinimini azaltır hem de parametre paylaşımı sayesinde istatistiksel verimliliği artırır (Goodfellow vd., 2016; O'Shea ve Nash, 2015; LeCun vd., 1998).

KSA mimarisinin temel bileşenleri Şekil 1'de özetlenmiştir. Buna göre bir KSA modeli, görüntüden düşük ve orta seviye özellikleri çıkartan konvolüsyon ve havuzlama katmanları ile başlamakta;

ardından bu özneliklerin sınıflandırılması için tam bağı katmanlara aktarılmasıyla sonlanmaktadır.

Şekil 1 Konvolüsyonel sinir ağının genel çalışma modeli



Konvolüsyon Katmanı

Konvolüsyon katmanı, giriş verisini filtre veya çekirdek adı verilen küçük matrislerle tarayarak özellik haritaları üretir. Filtre, girişin belirli bir bölgesi ile çarpılıp toplanarak tek bir çıktı değeri oluşturur ve bu işlem tüm görüntü boyunca kaydırma ile tekrarlanır. Elde edilen özellik haritaları, görüntünün kenar, köşe, doku ve diğer yapısal bileşenlerini temsil eder.

Boyut kaybını önlemek için katmanlar arasında doldurma işlemi uygulanabilir; özellikle sıfır-doldurma çıktı boyutlarının kontrol edilmesini sağlar (O'Shea ve Nash, 2015). Konvolüsyon, yüksek boyutlu görüntülerden anlamlı küçük özellikler çıkarılmasına olanak tanıdığı için KSA'ların başarısının temel bileşenlerinden biridir.

Havuzlama Katmanı

Havuzlama katmanı, konvolüsyon sonucu elde edilen özellik haritalarının belirli bölgelerini özetleyerek boyutunu küçültür. Genellikle maksimum havuzlama veya ortalama havuzlama yöntemleri kullanılır. Bu işlem hem modelin hesaplama yükünü azaltır hem de küçük konumsal değişimlere karşı göreceli bir değişmezlik sağlar (Kim, 2017). Havuzlama, KSA'ların daha soyut ve genel özellikler öğrenmesine katkı sunar.

Tam Bağlı Katman

Konvolüsyon ve havuzlama katmanları tarafından çıkarılan yüksek seviyeli özellikler, sınıflandırma işleminin gerçekleştirilebilmesi için tam bağlı katmana aktarılır. Bu katmanda önceki katmandaki tüm nöronlar bir sonraki katmandaki tüm nöronlarla bağlantılıdır; bu yapı klasik yapay sinir ağlarına benzer (O'Shea ve Nash, 2015).

Tam bağlı katmanın görevi, KSA'nın çıkardığı özellikleri sınıf olasılıklarına dönüştürmektir. Çıkış katmanında genellikle Softmax aktivasyon fonksiyonu kullanılarak her sınıf için 0–1 arası bir olasılık değeri üretilir ve tüm sınıfların toplamı 1 olacak şekilde normalize edilir.

Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları, yapay nöronlara doğrusal olmayan bir yapı kazandırarak derin ağların karmaşık ilişkileri öğrenmesini sağlayan temel bileşenlerdir. Doğrusal olmayan bir dönüşüm uygulanmadığı takdirde, çok katmanlı bir sinir ağı dahi tek katmanlı doğrusal bir modelle aynı ifade gücüne sahip olurdu. Bu nedenle aktivasyon fonksiyonları, ağın öğrenme kapasitesini belirleyen kritik bir rol üstlenmektedir. Ayrıca geri yayılım algoritmasında türevleri kullanıldığından, fonksiyonların türevlenebilir ve hesaplama açısından verimli olması önem taşır (Apicella vd., 2020; Lau ve Lim, 2019).

Transfer Öğrenme ile KSA Modellerinin Kullanımı

Geleneksel makine öğrenmesi yöntemleri, büyük ölçüde yalnızca kendi eğitim veri setinden öğrenen ve yeni bir görev için baştan eğitilmesi gereken istatistiksel modeller kullanır. Oysa derin öğrenme tabanlı KSA mimarileri, önemli miktarda etiketli veri gerektirdiğinden, özellikle tıbbi görüntüleme gibi veri üretiminin zor ve maliyetli olduğu alanlarda bu yaklaşım pratik değildir. Bu noktada transfer öğrenme, bir modelin daha önce benzer bir görev

için öğrenmiş olduğu temsilleri yeni bir probleme aktarmayı mümkün kılan etkili bir strateji olarak öne çıkar.

Bu yaklaşımın temel amacı, bir veya daha fazla kaynak görevden elde edilen bilginin yeni bir hedef görevin performansını artırmak için kullanılmasıdır. Çoklu görev öğrenme yöntemleri, tüm görevleri aynı anda öğrenmeye odaklanırken; transfer öğrenmede amaç, geçmişte öğrenilen bir görevden çıkarılan temsillerin yalnızca hedef görevin doğruluğunu iyileştirmek için kullanılmasıdır (Pan ve Yang, 2010).

Büyük ölçekli veri setleri üzerinde eğitilen modellerin elde ettiği temsil gücü, transfer öğrenmenin başarısının temel nedenlerinden biridir. Örneğin yaklaşık bir milyon görüntü ve 1000 sınıf içeren ImageNet veri seti (Deng vd., 2009), derin öğrenme modellerinin genel görsel özellikleri etkili biçimde öğrenmesini sağlayan en büyük referans veri setlerinden biridir. KSA tabanlı modellerin her bir katmanı, kenarlar, dokular, şekiller ve daha karmaşık nesne bileşenleri gibi farklı seviyelerde özellikler çıkarabilmektedir. Bu özellikler yalnızca ImageNet'e özgü değil, daha küçük veri setlerine sahip farklı görüntüleme görevlerinde de etkili olarak yeniden kullanılabilir (Akçay vd., 2016).

KSA tabanlı modellerde transfer öğrenme genellikle iki yaygın biçimde uygulanır:

Özellik Çıkarımı: Büyük veri setlerinde (örneğin ImageNet) eğitilmiş bir KSA'nın erken ve orta katmanları, genel görüntü özelliklerini öğrenmiş olduğundan değiştirilmez; yalnızca son katmanlar hedef problem için yeniden eğitilir. Bu yaklaşım veri miktarının sınırlı olduğu durumlarda oldukça etkilidir.

İnce Ayar: Modelin tüm katmanları, düşük bir öğrenme oranı ile hedef veri setine uyarlanır. Böylece hem önceden öğrenilmiş temsiller korunur hem de hedef problemin özgün yapısına uygun yeni temsiller geliştirilir.

Tıbbi görüntü sınıflandırma, beyin tümörü tespiti, göğüs röntgeni analizi veya segmentasyon gibi pek çok uygulamada transfer öğrenme, yüksek doğruluk sağlayarak sıfırdan model eğitime ihtiyacını ortadan kaldırır. Bu nedenle, KSA tabanlı modellerde transfer öğrenme hem eğitim süresini kısaltmakta hem de veri yetersizliği sorununu önemli ölçüde azaltmaktadır.

Vision Transformer (ViT)

Konvolüsyonel sinir ağları (KSA), yerel özellikleri yakalamada son derece başarılıdır. Ancak KSA'lar, görüntünün uzak bölgeleri arasındaki uzun menzilli bağımlılıkları modellemekte sınırlı kalır. Bu durum, tüm görüntünün bütüncül yapısının anlaşılmasını gerektiren nesne tanıma ve karmaşık sahne analizi gibi görevlerde önemli bir dezavantaj oluşturur. KSA'nın filtre tabanlı yapısı gereği, özellik çıkarımı yerel pencereler üzerinden gerçekleştiğinden küresel ilişkilerin öğrenilmesi ancak çok derin katmanlar eklenerek mümkün olur ki bu yaklaşım hem hesaplama maliyetini artırır hem de optimizasyon sorunlarına yol açabilir (Vujović, 2021).

Bu sınırlılığı gidermek amacıyla Dosovitskiy ve arkadaşları (2020), doğal dil işlemede büyük başarı göstermiş Transformer mimarisini görüntü alanına uyarlayarak Vision Transformer (ViT) modelini önermiştir (Dosovitskiy vd., 2020). Transformer tabanlı modellerin en önemli avantajı, self-attention mekanizması sayesinde uzak konumlardaki birimlerin ilişkilerini doğrudan modelleyebilmesidir. Doğal dilde bu birimler kelimeler iken, ViT mimarisinde birimler görüntü yamalarıdır. Böylece ViT, bir görüntüyü tıpkı bir metin dizisi gibi ele alarak her yamanın diğer tüm yamalarla olan ilişkisini öğrenir.

ViT, görüntüyü aşağıdaki işlemlerden geçirerek işler:

- Görüntünün Yamalara Bölünmesi: Görüntü, sabit boyutlu küçük yamalara ayrılır. Bu yamalar, model tarafından işlenecek “token” birimlerini oluşturur.

- Lineer Gömme: Her yama bir doğrusal katmandan geçirilerek yüksek boyutlu bir öz nitelik vektörüne dönüştürülür.
- Konumsal Kodlama: Transformer mimarisi dizisel sıralama bilgisine sahip olmadığından, yamaların görüntü üzerindeki konumsal ilişkilerinin korunması için konumsal kodlama uygulanır. Böylece model, yamaların uzamsal düzeni hakkında bilgi sahibi olur.
- Transformer Kodlayıcı: Elde edilen yama vektörleri, çoklu başlıklı kendiliğinden dikkat mekanizması içeren transformer kodlayıcı katmanlarına verilir. Bu mekanizma, tüm yamaların birbirleriyle olan ilişkilerini öğrenerek görüntünün küresel bağlamını yakalar.
- Sınıflandırma Katmanı: Kodlayıcı çıktıları, sınıflandırma görevi için kullanılan bir başlık katmanına aktarılır.

Bu yapı, KSA'lerin aksine görüntüyü yerel parçalara bölerek değil, yamalar arası ilişkileri küresel bağlamda modelleyerek öğrenir. ViT, böylece görüntüdeki genel yapıları, uzak bölgeler arasındaki ilişkileri ve küresel örüntüleri yakalamada üstün bir yetenek sergiler. Bu özellikleri sayesinde ViT, görüntü sınıflandırma görevlerinde oldukça yüksek doğruluk elde etmiş ve modern bilgisayarla görme alanında önemli bir alternatif yaklaşım hâline gelmiştir (Liu vd., 2023).

Derin Öğrenme Modellerinin Tıbbi Görüntü Sınıflandırmadaki Uygulamaları

Bu bölümde, dört sınıflı beyin tümörü MRI görüntülerinin sınıflandırılması probleminde kullanılan farklı derin öğrenme yaklaşımlarının deneysel sonuçları sunulmaktadır. Modellerin adil ve karşılaştırılabilir biçimde değerlendirilebilmesi amacıyla, tüm

deneylerde aynı veri bölünmesi kullanılmıştır. Veri seti, sağlayıcı tarafından eğitim ve test kümeleri ayrı klasörler hâlinde sunulduğundan, bu ayrım korunmuş; eğitim kümesi kendi içinde eğitim (%80) ve doğrulama (%20) alt kümelerine bölünmüştür. Böylece performans farklarının veri bölünmesinden değil, doğrudan model mimarisinden kaynaklanması hedeflenmiştir.

Deneyssel değerlendirme, farklı modelleme paradigmalarının sistematik biçimde karşılaştırılabilmesi amacıyla üç ana aşamada gerçekleştirilmiştir:

- sıfırdan tasarlanan temel bir konvolüsyonel sinir ağı (KSA) modeli,
- ön eğitilmiş KSA modelleri kullanılarak uygulanan transfer öğrenme yaklaşımı,
- görüntüyü yama tabanlı olarak işleyen Vision Transformer (ViT) tabanlı yaklaşım.

Bu yapılandırma, klasik konvolüsyonel mimarilerden başlayarak daha gelişmiş ön eğitilmiş ve transformer tabanlı modellere doğru performans değişiminin net biçimde analiz edilmesine olanak tanımaktadır.

Çalışmada Kullanılan Veri Seti

Bu çalışmada kullanılan “Brain Tumor Classification MRI” veri seti, Kaggle platformu üzerinden topluluk tarafından paylaşılan ve önceden işlenmiş manyetik rezonans görüntülerinden (MRI) oluşmaktadır. Veri setinin tıbbi kaynağı açık şekilde belirtilmemiş olup, hasta bazlı klinik meta-veriler içermemektedir. Bu nedenle elde edilen sonuçlar, herhangi bir tıbbi veya klinik çıkarım iddiası taşımadan, derin öğrenme tabanlı görüntü sınıflandırma yöntemlerinin karşılaştırmalı olarak incelenmesi ve model performanslarının analiz edilmesi amacıyla değerlendirilmiştir (Bhuvaji, 2020).

Veri seti, sağlayıcı tarafından eğitim ve test olmak üzere iki ayrı klasör halinde sunulmuştur ve bu çalışma kapsamında bu ayırım aynen korunmuştur. Her iki alt kümede de dört sınıf yer almaktadır: glioma_tumor, meningioma_tumor, no_tumor ve pituitary_tumor. Eğitim ve test kümelerine ait sınıf bazlı görüntü dağılımları Tablo 1’de görülmektedir.

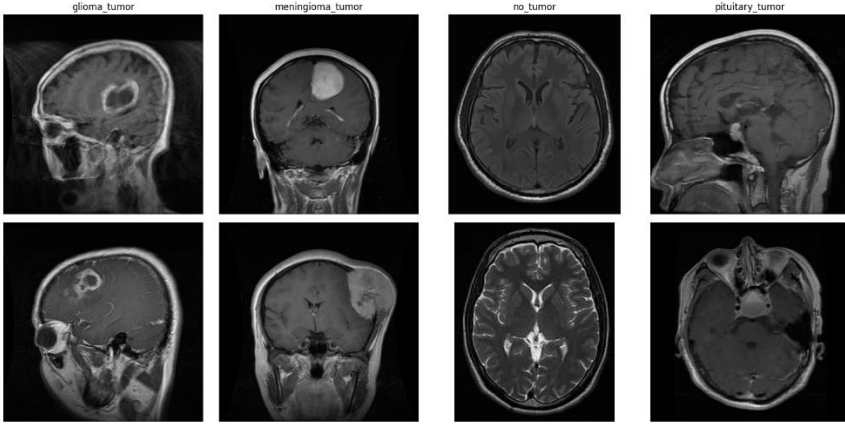
Tablo 1 Brain Tumor Classification MRI veri setinin eğitim ve test kümeleri için sınıf bazlı görüntü dağılımı

Sınıf	Training	Testing	Toplam
glioma_tumor	826	100	926
meningioma_tumor	822	115	937
no_tumor	395	105	500
pituitary_tumor	827	74	901
Toplam	2870	394	3264

Tablo 1 incelendiğinde, veri setinde sınıflar arasında genel olarak dengeli bir dağılımın bulunduğu görülmektedir. Özellikle glioma_tumor, meningioma_tumor ve pituitary_tumor sınıflarının hem eğitim hem de test kümelerinde benzer sayıda örnek ile temsil edildiği; no_tumor sınıfının ise görece daha az sayıda görüntü içermesine rağmen veri setinin bütünlüğünü bozacak düzeyde bir dengesizlik oluşturmadığı anlaşılmaktadır. Bu dağılım, farklı derin öğrenme mimarilerinin karşılaştırmalı performans analizlerinin güvenilir biçimde yapılmasına olanak sağlamaktadır.

Veri setine ait her sınıftan seçilen MRI görüntü örnekleri Şekil 2’de sunulmuştur.

Şekil 2 Brain Tumor Classification MRI veri setindeki dört sınıfa ait temsili MRI görüntü örnekleri.

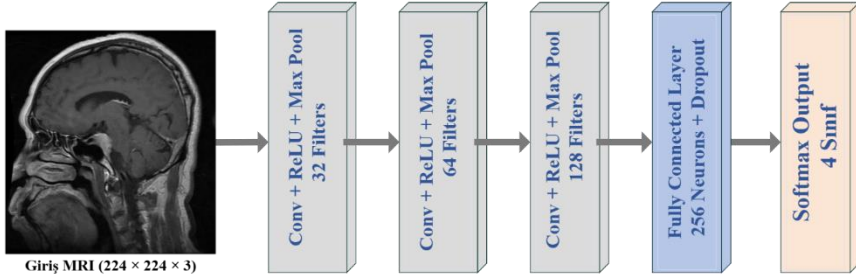


Sıfırdan Tasarlanan KSA Modeli

Bu aşamada, temel bir referans oluşturmak amacıyla sıfırdan tasarlanan sade bir KSA mimarisi değerlendirilmiştir. Önerilen model, art arda yerleştirilmiş dört konvolüsyon–havuzlama bloğundan oluşmakta olup, her blokta filtre sayısı kademeli olarak artırılmıştır (32, 64, 128 ve 256 filtre). Konvolüsyon katmanlarında doğrusal olmayan bir yapı sağlamak amacıyla ReLU aktivasyon fonksiyonu kullanılmış, her konvolüsyon bloğunun ardından uzamsal boyutları azaltmak ve hesaplama yükünü düşürmek için maksimum havuzlama işlemi uygulanmıştır.

Modelin genel mimarisi Şekil 3’te sunulmaktadır. Özellik çıkarım aşamasının ardından elde edilen öznetelik haritaları düzleştirilmiş ve 256 nöron içeren tam bağlantılı bir katman aracılığıyla sınıflandırma başlığına aktarılmıştır. Aşırı öğrenmeyi azaltmak amacıyla bu katmanda %50 oranında dropout uygulanmış, son katmanda ise softmax aktivasyon fonksiyonu kullanılarak dört sınıflı çıktı üretilmiştir.

Şekil 3 MRI görüntüleri üzerinde kullanılan sıfırdan tasarlanan KSA tabanlı sınıflandırma mimarisi.



Bağımsız test kümesi üzerinde gerçekleştirilen değerlendirmede model %66.24 doğruluk elde etmiştir. Sınıf bazlı performans analizi, meningioma_tumor, no_tumor ve pituitary_tumor sınıflarında kabul edilebilir sonuçlar elde edildiğini göstermektedir. Buna karşılık, glioma_tumor sınıfında duyarlılığın %19 seviyesinde kaldığı gözlemlenmiştir. Bu durum, sade konvolüsyonel yapıların karmaşık ve heterojen tümör dokularını ayırt etmede sınırlı kalabildiğini ve daha derin ya da ön eğitilmiş mimarilere duyulan ihtiyacı ortaya koymaktadır.

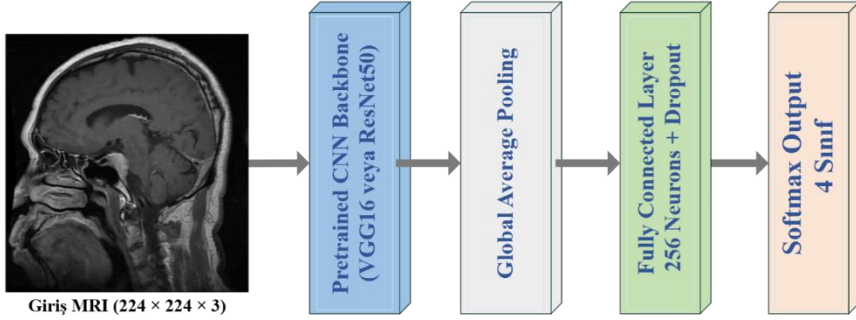
Ön Eğitilmiş KSA Modelleri

İkinci aşamada, ImageNet veri kümesi üzerinde önceden eğitilmiş ResNet50 ve VGG16 mimarileri transfer öğrenme yaklaşımı ile probleme uyarlanmıştır. Bu modeller, düşük seviyeli kenar ve doku özneliklerini önceden öğrenmiş derin konvolüsyonel yapıları sayesinde, sınırlı veri koşullarında daha güçlü temsil öğrenimi sunmaktadır.

Her iki modelde de konvolüsyonel taban katmanlar korunmuş, üst kısma Global Average Pooling, 256 nöronlu tam bağlı katman ve dropout içeren yeni bir sınıflandırıcı başlık eklenmiştir. Eğitim süreci iki aşamalı olarak gerçekleştirilmiş; ilk aşamada taban katmanlar dondurulmuş, ikinci aşamada ise üst konvolüsyonel

katmanlar kısmen açılarak ince ayar uygulanmıştır. Kullanılan transfer öğrenme mimarisinin genel yapısı Şekil 4’te sunulmaktadır.

Şekil 4 Ön eğitilmiş VGG16 ve ResNet50 mimarileri kullanılarak oluşturulan transfer öğrenme tabanlı sınıflandırma yapısı



ResNet50 Sonuçları

ResNet50 mimarisi, artık bağlantılar sayesinde derin ağlarda gradyan kaybını azaltan bir yapıya sahiptir. İnce ayar sonrasında model test kümesinde %73.60 doğruluk elde etmiş; glioma_tumor sınıfındaki duyarlılık %25 seviyesine yükselmiştir. Bu sonuç, derin ve artık bağlantılı mimarilerin sade KSA yapısına kıyasla daha ayırt edici özneliteler öğrenebildiğini göstermektedir.

VGG16 Sonuçları

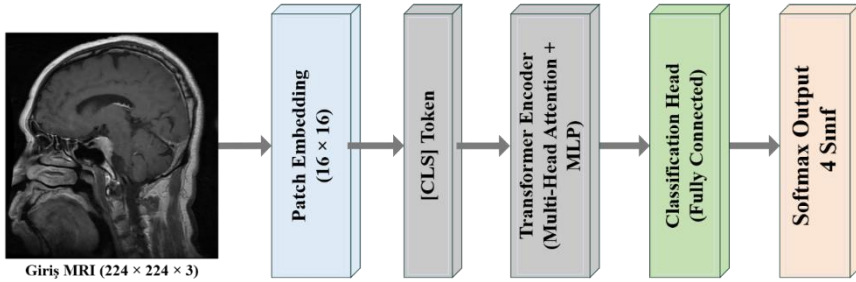
VGG16 modeli, ardışık 3×3 konvolüsyon katmanlarından oluşan daha sade fakat derin bir mimari sunmaktadır. İnce ayar sonrasında model %76.65 test doğruluğu elde etmiş ve özellikle pituitary_tumor ile no_tumor sınıflarında yüksek başarı sağlamıştır. Glioma_tumor sınıfında duyarlılık %28 seviyesine ulaşmış olup, bu değer ResNet50 ve sıfırdan tasarlanan KSA modeline kıyasla daha yüksektir.

Vision Transformer (ViT) Modeli

Üçüncü aşamada, konvolüsyonel yapılardan farklı olarak giriş görüntüsünü 16×16 boyutlu yamalara ayırarak bu yamaları bir token dizisi şeklinde işleyen Vision Transformer (ViT-B/16) mimarisi

değerlendirilmiştir. Model, çok katmanlı kendine dikkat mekanizması sayesinde görüntü içerisindeki küresel bağlamı doğrudan modelleyebilmektedir. Çalışmada ImageNet veri kümesi üzerinde önceden eğitilmiş ViT-B/16 modeli kullanılmış, sınıflandırma başlığı dört sınıf üretecek şekilde yeniden yapılandırılmıştır. Kullanılan ViT mimarisinin genel yapısı Şekil 5'te sunulmaktadır.

Şekil 5 ViT-B/16 mimarisi kullanılarak oluşturulan Vision Transformer tabanlı sınıflandırma yapısı.



Test kümesi üzerinde gerçekleştirilen değerlendirmede ViT-B/16 modeli %77.16 doğruluk elde etmiş, makro F1-skoru 0.75 değeri ile tüm modeller arasında en dengeli performansı sergilemiştir. Özellikle glioma_tumor sınıfında duyarlılığın %31 seviyesine yükselmesi, ViT mimarisinin KSA tabanlı yaklaşımlara kıyasla morfolojik olarak karmaşık sınıflarda daha ayırt edici temsiller öğrenebildiğini göstermektedir. Bu bulgular, küresel bağlam bilgisini doğrudan modelleyebilen Transformer tabanlı yapıların beyin tümörü sınıflandırma problemlerinde önemli bir potansiyele sahip olduğunu ortaya koymaktadır.

Karşılaştırmalı Değerlendirme

Farklı derin öğrenme yaklaşımlarının test kümesi üzerindeki karşılaştırmalı performansları Tablo 2'de sunulmaktadır. Değerlendirme; genel doğruluk, makro F1-skoru, ağırlıklı F1-skoru

ve özellikle klinik açıdan ayırt edilmesi zor olan glioma_tumor sınıfına ait duyarlılık ölçütleri üzerinden gerçekleştirilmiştir.

Tablo 2 Farklı derin öğrenme modellerinin test kümesi üzerindeki karşılaştırmalı performans sonuçları

Model	Doğruluk	Makro F1	Ağırlıklı F1	Glioma sınıfı duyarlılığı
Baseline KSA	0.662	0.63	0.62	0.19
ResNet50 (FT)	0.736	0.71	0.70	0.25
VGG16 (FT)	0.767	0.74	0.73	0.28
ViT-B/16	0.772	0.75	0.74	0.31

Elde edilen sonuçlar, mimari karmaşıklığın ve ön eğitim bilgisinin artmasıyla birlikte sınıflandırma performansının tutarlı biçimde yükseldiğini göstermektedir. Sıfırdan tasarlanan temel CNN modeli referans bir performans sunarken, transfer öğrenme yaklaşımı ile eğitilen ResNet50 ve VGG16 mimarilerinin hem genel doğruluk hem de sınıf bazlı metriklerde belirgin iyileşmeler sağladığı gözlemlenmiştir. Özellikle VGG16 modeli, glioma_tumor sınıfında ResNet50'ye kıyasla daha yüksek duyarlılık elde etmiştir.

Vision Transformer (ViT-B/16) modeli ise tüm metrikler dikkate alındığında en dengeli performansı sergilemiştir. ViT mimarisi, %77.2 doğruluk ve 0.75 macro F1-skoru ile genel sınıflandırma başarısında en yüksek değere ulaşırken, glioma_tumor sınıfında %31 duyarlılık ile CNN tabanlı tüm yaklaşımların üzerinde bir performans sunmuştur. Bu bulgular, küresel bağlam bilgisini doğrudan modelleyebilen Transformer tabanlı yapıların, morfolojik olarak karmaşık sınıfların ayırt edilmesinde önemli avantajlar sağladığını ortaya koymaktadır.

Sonuç ve Değerlendirme

Bu bölümde, dört sınıflı beyin tümörü MRI görüntülerinin sınıflandırılması probleminde farklı derin öğrenme yaklaşımlarının performansı sistematik ve karşılaştırmalı bir biçimde incelenmiştir.

Sıfırdan tasarlanan temel konvolüsyonel sinir ağı, referans bir performans sunarken; transfer öğrenme yaklaşımı ile eğitilen ResNet50 ve VGG16 mimarilerinin hem genel doğruluk hem de sınıf bazlı metriklerde belirgin iyileşmeler sağladığı gözlemlenmiştir. Bu sonuçlar, ön eğitilmiş derin modellerin sınırlı veri koşullarında daha ayırt edici temsiller öğrenebildiğini göstermektedir.

Vision Transformer (ViT-B/16) modeli ise tüm değerlendirme ölçütleri dikkate alındığında en dengeli performansı sergilemiştir. Özellikle glioma_tumor gibi morfolojik açıdan karmaşık ve sınıflandırılması zor sınıflarda elde edilen daha yüksek duyarlılık değerleri, ViT mimarisinin küresel bağlam bilgisini doğrudan modelleyebilme avantajını ortaya koymaktadır. Bu bulgular, transformer tabanlı yaklaşımların yalnızca doğal görüntülerde değil, tıbbi görüntü sınıflandırma problemlerinde de güçlü bir alternatif sunduğunu göstermektedir.

Bununla birlikte, bu çalışmada kullanılan veri setinin topluluk tarafından paylaşılan ve klinik meta-veri içermeyen ön işlenmiş MRI görüntülerinden oluştuğu göz önünde bulundurulmalıdır. Bu nedenle elde edilen sonuçlar, klinik bir karar destek sistemi iddiası taşımamakta; derin öğrenme tabanlı farklı modelleme yaklaşımlarının karşılaştırmalı olarak değerlendirilmesi ve eğitim amaçlı analiz edilmesi kapsamında ele alınmaktadır.

Gelecek çalışmalarda, daha büyük ve iyi tanımlanmış tıbbi veri kullanarak model genelleme yeteneğinin incelenmesi, segmentasyon tabanlı yaklaşımların sınıflandırma modelleriyle birleştirilmesi ve çoklu modalite içeren veri yapılarının değerlendirilmesi önemli araştırma yönleri olarak öne çıkmaktadır. Bu bağlamda, sunulan karşılaştırmalı analizlerin, derin öğrenme yöntemlerinin tıbbi görüntüleme alanındaki uygulamalarına yönelik temel bir referans oluşturmaları hedeflenmektedir.

Kaynakça

Akçay, S., Kundegorski, M. E., Devereux, M., & Breckon, T. P. (2016). Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery. In 2016 IEEE International Conference on Image Processing (ICIP) (pp. 1057–1061). IEEE. <https://doi.org/10.1109/ICIP.2016.7532519>

Apicella, A., Donnarumma, F., Isgrò, F., & Prevete, R. (2020). A survey on modern trainable activation functions. *Neural Networks*, 138, 14–32. <https://doi.org/10.1016/j.neunet.2021.01.026>

Bhuvaji, S. (2020). Brain tumor classification (MRI) [Data set]. Kaggle. <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 248–255). IEEE. <https://doi.org/10.1109/CVPR.2009.5206848>

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Houlsby, N. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv preprint*, arXiv:2010.11929.

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)* (pp. 315–323).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE. <https://doi.org/10.1109/CVPR.2016.90>

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>

Khan, S. U., Islam, N., Jan, Z., Ud Din, I., & Rodrigues, J. J. P. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters*, 125, 1–6. <https://doi.org/10.1016/j.patrec.2019.03.022>

Kim, P. (2017). Convolutional neural network. In *MATLAB deep learning* (pp. 121–147). Apress. https://doi.org/10.1007/978-1-4842-2845-6_6

Lau, M. M., & Lim, K. H. (2019). Review of adaptive activation function in deep neural network. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 686–690). IEEE. <https://doi.org/10.1109/IECBES.2018.8626714>

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2323. <https://doi.org/10.1109/5.726791>

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>

Liu, Y., Zhang, Y., Wang, Y., Hou, F., Yuan, J., & Tian, J. (2023). A survey of visual transformers. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6), 7478–7498. <https://doi.org/10.48550/arXiv.2111.06091>

Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint*, arXiv:1511.08458.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, 35(5), 1240–1251. <https://doi.org/10.1109/TMI.2016.2538465>

Saeedi, S., Rezayi, S., Keshavarz, H., & Niakan Kalhori, S. R. (2023). MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Medical Informatics and Decision Making*, 23, 16. <https://doi.org/10.1186/s12911-023-02114-6>

Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., & Fu, H. (2023). Transformers in medical imaging: A survey. *Medical Image Analysis*, 88, 102802. <https://doi.org/10.1016/j.media.2023.102802>

Vujović, Ž. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6), 599–606.

BÖLÜM 5

AKILLI TARIM UYGULAMALARI: SENSÖR VE VERİ ANALİTİĞİ İLE BİTKİ VERİMLİLİĞİNİN ANALİZİ

HÜSEYİN TÜRKMEN¹
İNAYET HAKKI ÇİZMECİ²

Giriş

Nüfusun hızla artışı dünya genelinde gıda ve tarımsal ürünlere olan talebi arttırmaktadır. Ekilebilir arazilerin günden güne azalması ve doğal kaynakların tükenmesi tarım sektörünü ciddi zorluklarla karşı karşıya bırakmaktadır. Birleşmiş Milletler Gıda ve Tarım Örgütü (FAO) verilerine göre artan nüfus için gıda üretiminin 2050 yılına kadar önemli ölçüde artırılması gerektiği vurgulanmaktadır (Gupta & Nahar, 2023). Fakat ekilebilen arazilerin azalması, su kıtlığı, geleneksel tarım yöntemlerinin verimsizliği gibi nedenlerden dolayı bu hedefe ulaşılabilmesini zorlaştırmaktadır (Büyüközkan & Uztürk, 2024). Geleneksel tarım uygulamaları çiftçilerin deneyimlerine dayandığı için hastalıkların tespiti geç fark

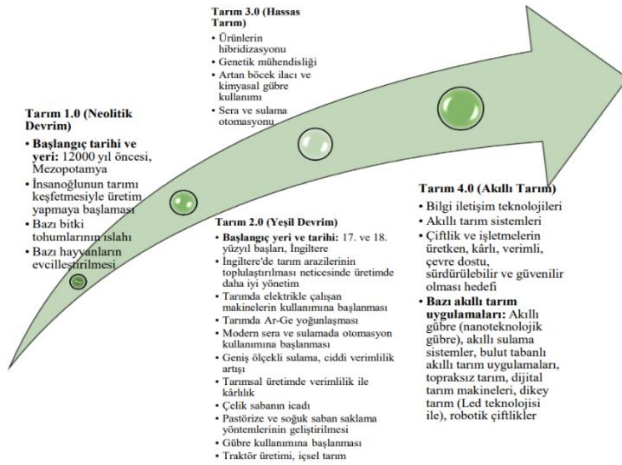
¹ Öğrenci, Afyon Kocatepe Üniversitesi, Bilgisayar ABD, Orcid: 0009-0002-3332-0759

² Dr. Öğr. Üyesi, Afyon Kocatepe Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0001-6202-4807

(Kitap bölümü tez çalışmasından üretilmiştir.)

edilmektedir. Bu durum kaynakların da israfına yol açmaktadır (Rani et al., 2023). Kaynakların verimli kullanılması ve hedeflere ulaşılabilmesine adına “Akıllı Tarım” kavramı ortaya çıkmıştır. Literatürde “Tarım 4.0” veya “Akıllı Tarım” olarak adlandırılan dönemin temel unsurlarını oluşturmakta ve üretimde hem verimliliği hem de sürdürülebilirliği artırma imkânı sağlamaktadır (Çakmakçı & Çakmakçı, 2023). Tarımsal üretimdeki teknolojik dönüşüm ve Tarım 1.0’den Tarım 4.0’a olan ilerleme, Şekil 1 ile özetlenmiştir (Aydınbaş, 2023).

Şekil 1 Tarımsal teknolojilerin evrimi: tarım 1.0’den Tarım 4.0’a geçiş



Kaynak: Aydınbaş 2023

Akıllı tarım, nesnelerin interneti (IoT), büyük veri ve yapay zekâ (AI) teknolojilerinin tarımsal süreçlere entegre edilmesini içermektedir (Ray, 2017). Akıllı tarımda kullanılan bu teknolojiler ile arazilerde gerçek zamanlı veriler toplanarak çiftçilerin doğru karar verebilmesine imkân sağlamaktadır. Özellikle IoT tabanlı sistemler, tarlaların adeta kendilerini ifade edebilmelerini sağlamaktadır. Böylece bitkinin sağlığı ve toprak durumu hakkında anlık bilgi sahibi olunabilmektedir (Ayaz et al., 2019).

Tarım daha çok hastalık tespiti ve verim tahmini gibi alanlarda yoğunlaşmaktadır. Özellikle derin öğrenme teknikleri bitki hastalıklarının tespit edilmesinde yüksek başarı elde etmektedir (Mohanty, Hughes, & Salathé, 2016). Bu tekniklerin karmaşık yapıya sahip olması çiftçilerin bu sistemleri kullanabilmelerini zorlaştırmaktadır. Çiftçilerin daha çok doğal dilde etkileşim kurabilecekleri ve kişileştirilmiş tavsiyeler alabilecekleri sistemlere ihtiyaçları vardır (Kutubuddin & Liyakat, 2024).

Bu çalışmada sensör ve veri analitiği temelli, Büyük Dil Modelleri (LLM) ile desteklenen yenilikçi bir sistem önerilmektedir. Topraktan alınan veriler ve bitki görüntüleri bir mobil uygulama üzerinden toplanmakta ve ChatGPT API' si kullanılarak analiz edilmektedir.

Literatür Bilgileri

Bu bölümde, akıllı tarım teknolojileri, IoT mimarisi, görüntü işleme teknikleri, makine öğrenmesi algoritmaları ve Büyük Dil Modelleri ve karar destek sistemleri üzerine yapılan çalışmalar incelenmiştir.

Akıllı tarım teknolojilerinin evrimi, Tarım 4.0' dan Tarım 5.0' a doğru ilerlemektedir. Javaid ve arkadaşları (2022), Tarım 4.0' ı Büyük Veri, Yapay Zekâ, Robotik, IoT ve Blok Zinciri teknolojilerinin entegrasyonu ile gerçekleşen dördüncü tarım devrimi olarak tanımlamaktadır (Javaid et al., 2022). Mesias-Ruiz ve arkadaşları (2023), bu evrimin makine öğrenmesi algoritmalarıyla desteklenen, otonom karar verme süreçlerinin yer aldığı Tarım 5.0' a doğru ilerlediğini belirtmektedir (Mesías-Ruiz et al., 2023).

IoT tabanlı izleme sistemleri açısından, Yang ve diğerleri (2021) düşük maliyetli Arduino ve Raspberry Pi tabanlı dört katmanlı bir sistem mimarisi önerirken, Jaliyagoda ve arkadaşları (2023) ESP32 ve MQTT protokolü kullanan maliyet etkin bir sera sistemi geliştirmiştir (Jaliyagoda et al., 2023; Yang, Ma, Li, &

Zhang, 2022). Gagliardi ve arkadaşları (2021) ise bağlarda ZigBee tabanlı WSN ve İHA entegrasyonu ile NDVI haritaları oluşturan geniş ölçekli bir mimari sunmuştur (Gagliardi et al., 2021).

Görüntü işleme alanında, Li ve arkadaşları (2021) VGG, ResNet ve GoogLeNet gibi derin öğrenme mimarilerinin bitki hastalık tespitinde geleneksel yöntemlerin yerini aldığını göstermiştir (Li, Zhang, & Wang, 2021). Latif ve arkadaşları (2022), geliştirilmiş VGG19 modeli ile pirinç yapraklarının hastalıklarını %96,08 doğrulukla tespit etmiştir (Latif et al., 2022). Chen ve arkadaşları (2020), YOLOv3 ve LSTM kullanarak AIoT tabanlı zararlı tespit sistemi geliştirmişlerdir (Chen, Huang, Li, Chang, & Huang, 2020).

Makine öğrenmesi uygulamalarında, Gupta ve Nahar (2022) aKNCN ve ELM-mBOA algoritmalarıyla verim tahmini yaparken, Elavarasan ve Vincent (2020) DRQN modeli ile %93,7 doğruluk elde etmiştir (Elavarasan & Durairaj Vincent, 2020). Yang ve arkadaşları (2022), makine öğrenmesi için kullanılan veri tüketimini %5-15 azaltan Kenar Mesafe-Entropisi yöntemini önermiştir (Yang et al., 2022). Elashmawy ve Uysal (2023), toprak sensörlerinden çilek kalitesini tahmin etmiş, Blessy ve diğerleri (2023) ise Transfer Öğrenme ile sulama optimizasyonunda %40-50 su tasarrufu sağlamıştır (Blessy et al., 2023; Elashmawy & Uysal, 2023).

Büyük Dil Modelleri konusunda, Calone vd. (2025) ChatGPT' nin hastalık risk tahminlerini tavsiyelere dönüştürme potansiyelini, Shaikh vd. (2025) LLM' lerin çok modlu veri işleme yeteneğini, Sowmya vd. (2025) ise üretken yapay zekâ platformlarının çiftçilerin teknolojiye erişimini kolaylaştırdığını göstermiştir (Calone et al., 2025; Shaikh, Rasool, Veningston, & Yaseen, 2025; Sowmya et al., 2025).

Literatür incelendiğinde, akıllı tarım uygulamalarının sensör ağları, IoT platformları ve yapay zekâ algoritmalarının entegrasyonu

ile geliştiđi görölmektedir. Ancak Kutubuddin ve Liyakat (2024), AI destekli IoT sistemlerinde karar vermenin önemini vurgulasa da sensör verisi ile görsel veriyi birleştirip bunları LLM aracılığıyla yorumlayan sohbet tabanlı uygulamalar sınırlıdır. Bu çalışma ile sınırlı olan boşluğu doldurmayı hedeflemektedir (Kutubuddin & Liyakat, 2024).

Materyal ve Metot

Bu bölümde, çalışmada kullanılan donanım ve yazılım bileşenleri ayrıntılı biçimde açıklanmış; sistemin çalışma prensibi, veri toplama süreci ve analiz yöntemleri bilimsel çerçevede sunulmuştur.

Donanım Bileşenleri

Bu kısımda sistem tasarımında kullanılan bileşenler tanıtılmaktadır. Çalışmada kullanılan donanım bileşenleri ve temel özellikleri Tablo 1’de devre şeması ise Şekil 2’ de gösterilmiştir.

Sistem, düşük enerji tüketimi ve kablosuz bağlantı özellikleri nedeniyle ESP32 Wi-Fi + Bluetooth Dual Mode geliştirme kartı üzerine inşa edilmiştir. ESP32 hem sensörlerden gelen verilerin işlenmesi hem de kablosuz veri iletimi görevini üstlenmiştir.

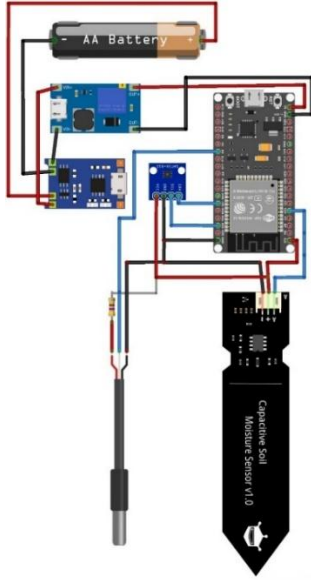
Tablo 1 Çalışmada kullanılan donanım bileşenleri

Bileşen	Model	Kullanım Amacı	Temel Özellikleri
Mikrodenetleyici	ESP32 Dual Mode	Veri toplama ve kablosuz iletişim	Wi-Fi ve Bluetooth destekli, düşük enerji tüketimli
Toprak Sıcaklık Sensörü	DS18B20	Toprak sıcaklığını ölçme	Su geçirmez, dijital çıkış, $\pm 0,5^{\circ}\text{C}$ hassasiyet
Toprak Nem Sensörü	Kapasitif v1.2	Toprak nem oranını tespit etme	Korozyona dayanıklı, analog çıkışlı

Şarj Modülü	TP4056	Pil şarj ve koruma devresi	Aşırı şarj korumalı, 1A çıkış akımı
Güç Kaynağı	18650 Li-ion Pil	Enerji kaynağı	3.7V, 2500 mAh kapasite
Voltaj Regülatörü	MT3608	Düşük voltajlı bir güç kaynağını daha yüksek bir voltaj seviyesine çıkarmak	Giriş voltajı 2V ile 24V arasında, çıkış voltajı ise 28V' a kadar ayarlanabilir
Hava Sıcaklık ve Nem Sensörü	SHT3X	Havanın sıcaklığını ve nemini ölçmek	Giriş voltajı 2.4 V – 5.5 V arasında, -40°C ile +125°C arasında ölçüm
Diğer Elemanlar	Direnç, IC205 sürgülü switch	Devre kararlılığı ve enerji kontrolü	180° sürgülü kontrol mekanizması

Tablo 1’de belirtilen tüm donanım bileşenleri kompakt ve taşınabilir bir sistem oluşturmak üzere bir araya getirilmiştir. ESP32 geliştirme kartı, devre kartı üzerine yerleştirilmiş ve sensör modülleri ile bağlantıları lehimlenerek sağlanmıştır. Koruma amaçlı üç boyutlu yazıcı kullanılarak dış kutusu oluşturulmuş ve içerisine yerleştirilmiştir. Sistemin kutu içerisindeki genel görünümü Şekil 3’te verilmiştir.

Şekil 2 Tasarlanan sistemin devre şeması



Şekil 3 ESP32 tabanlı IoT sisteminin kutu içi yerleşimi



Geliştirilen sistem gerçek ortam koşullarında test edilmek üzere saksı bitkisi üzerine yerleştirilmiştir. Toprağın nemini ve sıcaklığını ölçen sensör doğrudan toprak içerisine yerleştirilirken ortam sıcaklığını ve nemini ölçen sensör ise devre üzerinde konumlandırılmıştır. Böylece toprak içerisindeki koşullar ve bitki çevresindeki hava koşulları izlemektedir. Sistemin saksı içerisindeki durumu Şekil 4’ te gösterilmiştir.

Şekil 4 Sistemin saksı bitkisi üzerinde test edilmesi



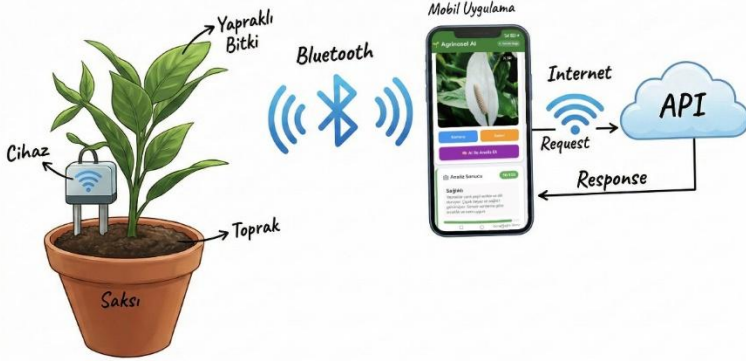
Yazılım Altyapısı

Bu çalışmada geliştirilen akıllı tarım sistemi, yalnızca donanım bileşenlerinden oluşan bir yapı değil, aynı zamanda bu donanımı etkin biçimde yöneten ve veri odaklı kararlar üreten güçlü bir yazılım mimarisi üzerine kurulmuştur. Sistem yazılımı, sensörlerden elde edilen verilerin toplanması, işlenmesi, bulut ortamına aktarılması ve kullanıcıya anlamlı bilgiler olarak sunulması süreçlerini bütüncül bir yaklaşımla yönetmektedir.

Şekil 5’ te sistemin genel çalışmasına ait akış şeması görülmektedir. Bu şemaya göre sensör verileri toplanarak Bluetooth ile mobil uygulamaya aktarılmaktadır. Aktarılan veriler mobil

uygulama vasıtasıyla internet üzerinden API servisine gönderilmektedir. Verilerin analizi sonucunda teşhis ve öneriler mobil uygulamaya gönderilmektedir.

Şekil 5 Geliştirilen akıllı tarım uygulamasının genel akış şeması

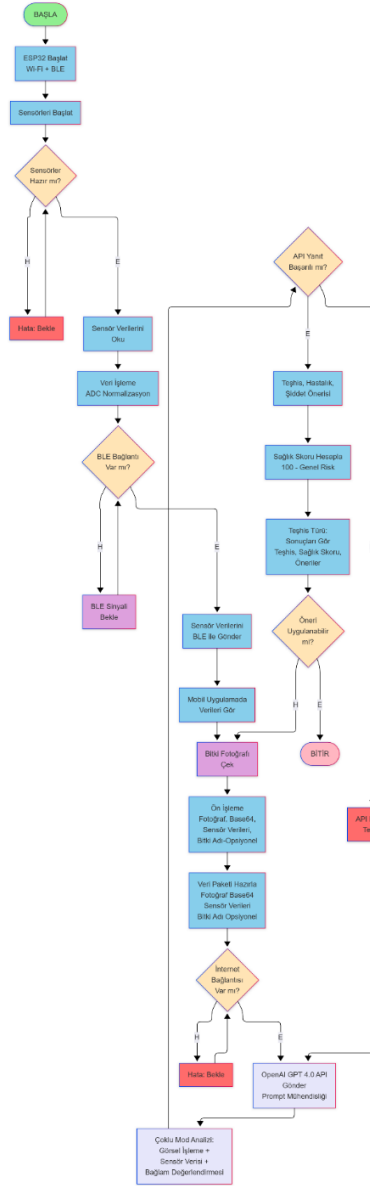


Kaynak: Gemini AI

Şekil 6’ da sisteme ait akış diyagramı görülmektedir. Görüldüğü üzere sistem ESP32 mikrodenetleyicisinin başlatılmasıyla çalışmaya başlamakta ve sensörlerden gelen veriyi toplayarak Bluetooth (BLE) üzerinden mobil uygulamaya aktarmaktadır. Kullanıcının çektiği bitki fotoğrafını ön işlemden geçirilip OpenAI GPT-4 API’ sine gönderilmektedir. Son aşamada çok modlu analiz sonucunda elde edilen teşhis ve önerileri kullanıcıya sunulmaktadır.

Geliştirilen yazılım altyapısı, mobil uygulama geliştirme katmanı ve yapay zekâ ve veri analitiği katmanı (OpenAI GPT-4) olmak üzere iki temel katmandan oluşmaktadır. Bu katmanlı yapı IoT tabanlı akıllı tarım sistemlerinin gerektirdiği modüler, ölçeklenebilir ve güvenilir bir yapıda tasarlanmıştır. Bu sayede sistem hem laboratuvar ortamında hem de gerçek saha koşullarında kesintisiz çalışabilecektir.

Şekil 6 Akıllı tarım sisteminin genel çalışma akış diyagramı



Bu çalışmada geliştirilen kullanıcı arayüzü, tek bir kod tabanı üzerinden hem Android hem de iOS platformlarında çalışabilen bir mobil uygulama olarak React Native platformu üzerinde tasarlanmıştır. React Native, JavaScript diliyle çalışmakta ve “bir kez yaz, her yerde çalıştır” yaklaşımını desteklemektedir. Bu sayede proje geliştirme sürecinde platform bağımsızlığı, kodun yeniden kullanılabilirliği ve bakım-güncelleme maliyetlerinin azaltılması gibi önemli avantajlar sağlanmaktadır (Gülcüoğlu, Ustun, & Seyhan, 2021).

Uygulama ESP32 mikrodenetleyicisinden Bluetooth Low Energy (BLE) protokolü üzerinden gerçek zamanlı sıcaklık, bağıl nemi ve toprak nemi verilerini almaktadır. Ancak sistemin en kritik özelliği bu sayısal verileri bitki görüntüleriyle birleştirmesidir. Veriler yapay zekâ modeline iletilmeden önce mobil cihaz kamerasıyla alınan bitki görüntüleri bir dizi ön işlemden geçirilmektedir. Bu işlemin ana nedeni ise veri iletim hızını artırmak ve işlem maliyetini düşürmek için gerçekleştirilmektedir. Uygulama içerisinde kullanılan expo-image-manipulator kütüphanesi aracılığıyla yeniden boyutlandırma (800 piksel), sıkıştırma (%60 kalite) ve kodlama (Base64 formatı) teknikleri kullanılmıştır. Böylece sensör verileri ve optimize edilmiş görsel veriler paketlenerek analiz katmanına hazır hale getirilmiştir.

Yapay Zekâ Destekli Karar Destek Mekanizması ve Prompt Mühendisliği

Yazılımın analiz ve karar verme süreçleri OpenAI tarafından geliştirilen ve çok modlu özelliklere sahip GPT-4o modeli üzerine inşa edilmiştir. Bu modül klasik görüntü işleme tekniklerinden farklı olarak bitkinin görsel durumu ile çevresel sensör verilerini eş zamanlı olarak değerlendirebilmektedir.

Modelin çıktılarının tutarlılığını ve tarımsal terminolojiye uygunluğunu sağlamak amacıyla Prompt Mühendisliği teknikleri

uygulanmıştır. Sisteme gönderilen istek şu üç bileşenden oluşmaktadır:

1. Sistem İstemi: Modele, “Bitki fizyolojisi ve patolojisi konusunda uzmanlaşmış ileri seviye bir Ziraat Mühendisi” rolü atanmıştır. Ayrıca, saksı bitkileri için “hektar” veya “tarla sulama” gibi uygunsuz önerilerde bulunmaması bunun yerine ev tipi uygulanabilir öneriler sunması kuralı getirilmiştir.
2. Bağlamsal Veri: Kullanıcıdan alınan bitki adı veya optimize edilmiş bitki fotoğrafı sensörlerden alınan anlık sayısal veriler ile modele girdi olarak sunulmaktadır. Bu sayede model, yapraktaki sararmanın sadece bir hastalık belirtisi mi yoksa sensör verisindeki “düşük toprak nemi” kaynaklı bir su stresi mi olduğunu ayırt edebilmektedir.
3. Yapılandırılmış Çıktı: Modelin cevabının mobil uygulama tarafından işlenebilmesi için JSON formatında yanıt vermesi zorunlu kılınmıştır.

Modelden dönen JSON verisi tahmini teşhis, hastalık şiddeti, uygulanabilir öneriler ve risk puanlarını (su stresi, mantar riski vb.) içermektedir. Uygulama, bu veriyi işleyerek “Genel Risk Puanı” üzerinden ters orantılı bir algoritma ile (Sağlık Skoru = 100- Risk Puanı) bitkinin sağlık skorunu hesaplamakta ve kullanıcıya görselleştirilmiş bir rapor sunmaktadır.

Test Protokolü ve Uzman Değerlendirme Metodolojisi

Test için 50 farklı saksı bitkisi seçilmiştir. Bitki örneklerinin dağılımı Tablo 2’de sunulmuştur.

Tablo 2 Test edilen bitki örneklerinin kategori bazlı dağılımı

Kategori	Bitki Sayısı	Yüzde (%)	Açıklama
Sağlıklı Bitki	12	24	Herhangi bir problem belirtisi göstermeyen optimal koşullarda yetişen bitkiler
Su Stresi	10	20	Yetersiz sulama veya aşırı su nedeniyle stres altındaki bitkiler
Fungal/Bakteriyel Hastalık	8	16	Yaprak lekesi, küf, pas gibi patojen kaynaklı hastalıklar
Besin Eksikliği	9	18	Azot, demir, magnezyum vb. element eksikliği belirtileri gösteren bitkiler
Zararlı Atakı	6	12	Yaprak biti, beyaz sinek, trips gibi zararlıların zarar verdiği bitkiler
Çevresel Stres	3	6	Sıcaklık stresi, nem stresi gibi çevresel faktörlere bağlı problemler
Fiziksel Hasar	2	4	Mekanik yaralanma, yaprak yırtılması gibi fiziksel hasarlar
TOPLAM	50	100	

Sistemin farklı görüş açılarındaki tutarlılığını test etmek ve daha geniş bir analiz yapmak amacıyla her bitki örneği için birden fazla görüntü alınmıştır. Bu görüntüler Honor 400 telefonuna ait 200 megapiksele ve f/1.9 diyafram açıklığına sahip kamera ile çekilmiştir. Genel olarak bitki görüntüleri, bitkiden 30 ile 50 cm arası mesafeden çekilmiştir. Her bitkinin farklı açılardan 3 adet görüntüsü alınmıştır. Her görüntü sensör verileriyle birlikte sistem tarafından analiz edilmiş ve teşhis raporu üretilmiştir. Böylece sistemin hem aynı bitkinin farklı açılardan görünümünde tutarlılığı hem de sensör verilerini görsel bilgiyle birleştirme yeteneği test edilmiştir.

Sistemin güvenilirliğini ve teşhis doğruluğunu değerlendirmek için bağımsız iki uzman tarafından çift-kör

değerlendirme ile değerlendirilmiştir. Uzmanların değerlendirme süreçleri verilerin hazırlanması, kör değerlendirme ve sonuçların karşılaştırılması şeklinde gerçekleştirilmiştir.

Veriler hazırlanmasında her bitkiye ait 3 adet görüntü, sensör verileri, sağlık skoru ve GPT-4o analizine ait sonuçlar şeklinde 50 adet bitki için veri seti oluşturulmuştur. Her bir uzman sistemin ürettiği sonuçları ayrı ayrı değerlendirmiştir. Değerlendirirken aşağıdaki performans değerlendirme kriterlerini kullanmışlardır:

- Doğru (1,0 puan): Teşhis kategorisi ve öneri uzman görüşüyle tam örtüşmektedir.
- Kısmen Doğru (0,5 puan): Ana kategori doğru ancak spesifik tanımlama eksik veya genel kalmıştır.
- Yanlış (0,0 puan): Teşhis kategorisi hatalıdır.

Sistemin başarı puanı aşağıdaki denklem ile hesaplanmıştır:

$$\text{Başarı Oranı (\%)} = \frac{[(\text{Doğru}_{\text{sayısı}} \times 1.0) + (\text{Kısmen}_{\text{doğru}} \times 0.5)]}{\text{Toplam}_{\text{örneksayısı}}} \times 100$$

$\text{Doğru}_{\text{sayısı}}$, yapay zekanın verdiği teşhis kategorisi önerisi ile uzman görüşüyle tam örtüştüğü doğru sayısını, $\text{Kısmen}_{\text{doğru}}$, ana kategori doğru ancak spesifik tanımlama eksik veya genel kalmış kısmen doğru sayısını ve $\text{Toplam}_{\text{örneksayısı}}$, toplam analiz edilen bitki sayısını vermektedir.

Uzmanlar arası uyum ve sistem performansını değerlendirmek için Cohen' s Kappa istatistiksel yöntem kullanılmıştır. Kappa değeri, şans faktörü düzeltilmiş uyum oranını gösterir ve -1 ile +1 arasında değer alır (Cohen, 1960).

Bir diğer performans değerlendirmesi olarak aynı bitkiye ait birden fazla görüntünün analiz sonucunun tutarlılığıdır. Yapay zekanın ne kadar tutarlı teşhis koyduğunu ölçmek için tutarlılık skoru geliştirilmiştir. Bu metrik sistemin görüş açısı değişimlerinden

ne kadar etkilendiğini de göstermektedir. Tutarlılık skoru için aşağıdaki denklem kullanılarak hesaplanmıştır.

$$\text{Tutarlılık Skoru (\%)} = \frac{\sum_{i=1}^n P_i}{3 \times n} \times 100$$

P_i , i' inci bitkinin puana katkısı (üç görüntüde aynı teşhis için 3 puan, iki görüntüde aynı teşhis için 2 puan, üç farklı teşhis için 1 puan), n toplam bitki sayısını vermektedir. 1' den n' e kadar olan bitkilere ait skorlar toplanarak maksimum alınabilecek puan 3 olduğundan toplam bitki sayısının 3 ile çarpılmasına bölünmektedir. Böylece tutarlılık skoru elde edilmektedir.

Bulgular

Bu bölümde tasarlanan akıllı tarım alanında sensör izleme ve analiz sisteminin deneysel test sonuçları sunulmuştur. Bulgular donanım bileşenlerinin kararlılığı, mobil uygulamanın performansı ve Büyük Dil Modeli (GPT-4o) destekli yapay zekâ analizinin doğruluğu değerlendirilmiştir.

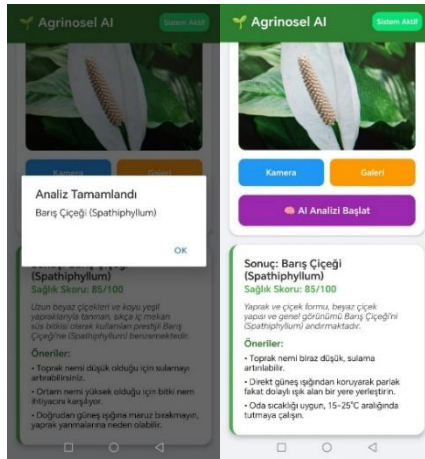
Sistemde kullanılan Kapasitif Toprak Nem Sensörü v1.2 toprak nemi ölçümlerinde sensörden alınan ham analog veriler (ADC), normalizasyon yöntemi ile %0-100 aralığına dönüştürülmüştür. Normalizasyon işlemi aşağıdaki denklem ile elde edilmiştir.

$$Nem = \frac{Ham_Veri - NemRef_{Min}}{NemRef_{Max} - NemRef_{Min}} \times 100$$

Ham_Veri , sensörde okunan ham analog değeri göstermektedir. $NemRef_{Min}$ kurutulmuş topraktaki sensör analog değeri, $NemRef_{Max}$ suya tamamen doymuş topraktaki sensörün analog değerini göstermektedir. Yapılan testlerde toprağa su eklediğimiz andan itibaren nem değişimi ortalama 3 dk sonra kararlı bir sonuca ulaşarak mobil uygulamaya yansıtılmaktadır.

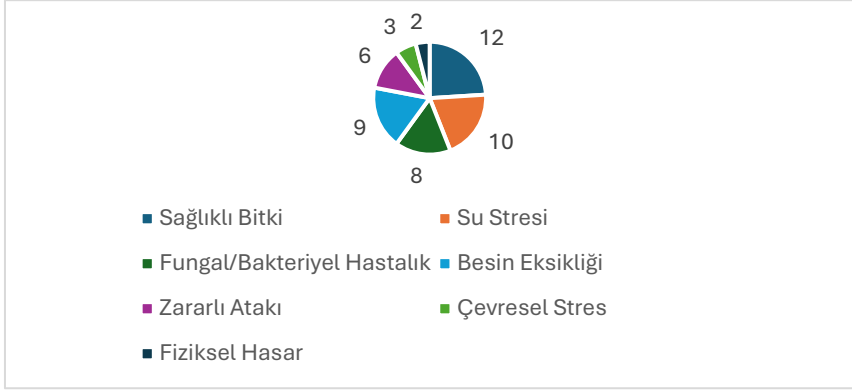
Geliştirilen React Native tabanlı mobil uygulama farklı senaryolar altında test edilmiştir. Uygulamanın en önemli çıktısı, sayısal veriler ile görsel analizi tek ekranda birleştirebilmesidir. Analiz tamamlandıktan sonra yapay zekâ tarafından üretilen detaylı teşhis, hastalık tanımı ve tedavi önerileri, kullanıcının anlayabileceği sade bir arayüz ile sunulmaktadır. Şekil 7’ de sistemin bir bitki için ürettiği teşhis ve çözüm önerilerini içeren sonuç ekranı görülmektedir.

Şekil 7 Yapay zekâ tarafından üretilen analiz sonuç ekranı



Sistem performansını kapsamlı bir şekilde değerlendirmek amacıyla 50 farklı bitki örneği üzerinde test çalışması gerçekleştirilmiştir. Test edilen bitki örneklerinin kategori bazlı dağılım Grafik 1’ de verilmiştir.

Grafik 1 Test edilen kategori bazlı bitki sayısı



Her bitki için 3 farklı açıdan görüntüsü alınmış ve sensör verileriyle birlikte sistem tarafından analiz edilmiştir. Böylece toplam 150 bitki analizi gerçekleştirilmiştir. Sonuçlar Tablo 3’ te sunulmuştur.

Tablo 3 Test edilen bitkilere ait tutarlılık tablosu

Bitki Durumu	Bitki Sayısı (n)	Tam Tutarlı (3/3)	Kısmi Tutarlı (2/3)	Düşük Tutarlı (1/3)	Tutarlılık Skoru (%)
Sağlıklı	12	11	1	0	97,2
Su Stresi	10	9	1	0	96,7
Fungal/Bakteriyel	8	6	2	0	91,7
Besin Eksikliği	9	6	3	0	88,9
Zararlı Atakı	6	5	1	0	94,4
Çevresel Stres	3	3	0	0	100
Fiziksel Hasar	2	2	0	0	100
TOPLAM	50	42	8	0	95,3

Bulgular sistemin 50 bitkinin 42’ sinde (%84) üç görüntüde de aynı teşhisi koyduğunu göstermektedir. 8 bitkide (%16) ise 3 görüntünün 2’ sinde aynı teşhise ulaşılmıştır. Hiçbir bitkide üç farklı teşhis oluşmamıştır. Genel tutarlılık skoru %95,3 olarak hesaplanmıştır. Besin eksikliği kategorisinde tutarlılığın %88,9’ a düşmesi dikkat çekicidir. Bu durum, farklı açılardan çekilen yaprak

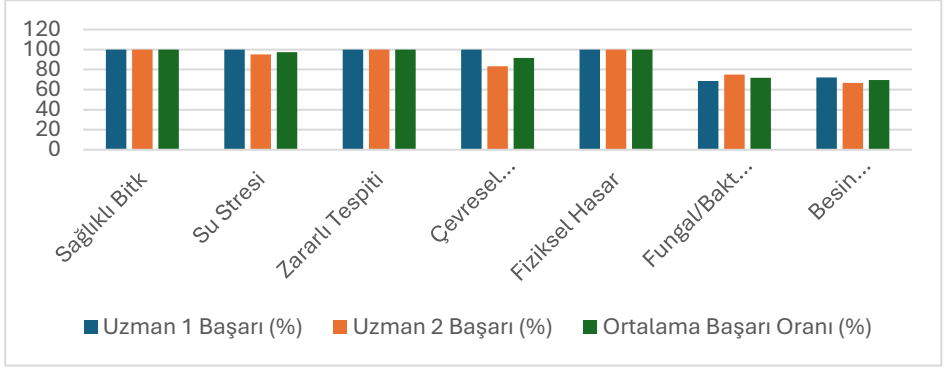
görüntülerinde kloroz (yeşil rengini kaybederek sararması) belirtilerinin görünürlüğünün değişmesiyle açıklanabilir.

İki bağımsız uzmanın 50 bitki üzerindeki değerlendirme uyumu analiz edilmiştir. Uzmanlar birbirlerinin değerlendirmelerini görmeden çalışmış ve her bitki için verilen teşhisi değerlendirmişlerdir. Uzmanlar 50 adet bitkiyi kategorisine göre sınıflandırmışlardır.

İki uzman arasındaki genel uyum oranı %86 olarak bulunmuştur. Cohen' s Kappa değeri 0,86 olup, Landis ve Koch (1977) sınıflandırmasına göre “mükemmel uyum” (0,81-1,00) düzeyindedir. Bu sonuç değerlendirmenin güvenilir ve objektif olduğunu göstermektedir (Landis & Koch, 1977).

Sistemin ürettiği teşhisler, her iki uzmanın değerlendirmeleriyle ayrı ayrı karşılaştırılmıştır. Sistem performansı tam doğru (1,0 puan), kısmen doğru (0,5 puan) ve yanlış (0,0 puan) olmak üzere üç kategoride değerlendirilmiştir. Sistem, Uzman 1 ile %90, Uzman 2 ile %88 başarı oranı göstermiştir. Her iki uzmanla da yüksek düzeyde uyum sağlanması (ortalama %89), sistemin güvenilir teşhis yapabildiğini göstermektedir. Grafik 2' de kategori bazlı sistem performansı (her iki uzman ortalaması) görsel olarak sunulmuştur. Grafik 2, sistemin hangi problem türlerinde daha başarılı olduğunu açıkça göstermektedir.

Grafik 2 Kategori bazlı sistem performansı



Tartışma ve Sonuç

Bu çalışmada nesnelerin interneti (IoT) teknolojisi ile büyük dil modellerinin (LLM) entegrasyonuna dayanan yenilikçi bir akıllı tarım sistemi geliştirilmiştir. ESP32 mikrodenetleyici tabanlı sensör ağı, React Native mobil uygulaması ve GPT-4o yapay zekâ modelinin birleştirildiği bu sistem, bitki sağlığının çok modlu analizini gerçekleştirmektedir.

Sistemin performansı 50 farklı bitki türü üzerinde (toplam 150 analiz) iki bağımsız uzmanın çift-kör değerlendirmesi ile doğrulanmıştır. Uzmanlar arası Cohen' s Kappa katsayısı 0,86 ("mükemmel uyum") olarak hesaplanmış, sistem genel ortalama %89 başarı oranı elde etmiştir. Aynı bitki üzerinde farklı açılardan çekilen üç görüntüdeki tutarlılık skoru %95,3 olarak ölçülmüştür.

Çalışmanın en önemli bulgusu, sensör verilerinin görsel analiz ile birleştirilmesinin teşhis doğruluğunu önemli ölçüde artırdığını göstermesidir. Sistem özellikle abiyotik stres faktörlerini (su stresi, sağlıklı bitki, çevresel stres) yüksek başarı ile tespit edebilmiştir. Geleneksel görüntü tabanlı sistemlerde sararmış yapraklar hem azot eksikliğini hem su stresini hem de bazı

hastalıkları işaret edebilmektedir. Bu sistem düşük toprak nemi verisi ile değerlendirdiğinde durumu doğru sınıflandırabilmiştir.

Literatürdeki CNN tabanlı çalışmalarla karşılaştırıldığında bu sistemin avantajı 50 farklı bitki türünü önceden eğitim gerektirmeden analiz edebilmesidir. Latif ve arkadaşları (2022) VGG19 ile %96 doğruluk elde etmiş ancak sadece pirinç bitkisinde çalışmıştır (Latif et al., 2022). Mohanty ve arkadaşları (2016) %99 doğruluk bildirmiş ancak 54.306 etiketli görüntü kullanmıştır. Bu çalışmada ise kullanıcı sadece fotoğraf çekerek anında teşhis alabilmektedir (Mohanty et al., 2016).

Prompt mühendisliği sistemin başarısında kritik rol oynamıştır. Modele “Ziraat Mühendisi” rolü atanması, JSON formatında yapılandırılmış çıktı talep edilmesi ve sensör verilerinin görsel bilgi ile değerlendirilmesi talimatı, doğru ve kullanılabilir sonuçlar üretilmesini sağlamıştır.

Sistemin sınırlılıkları arasında GPT-4o’ nun genel amaçlı yapısının spesifik patojen tanımlamasında yarattığı zorluklar, internet bağımlılığı ve fungal/bakteriyel hastalıklar ile besin eksikliği kategorilerindeki düşük performans sayılabilir. Buna karşın yaklaşık 700 TL donanım maliyeti, kullanıcı dostu Türkçe arayüz ve geniş bitki yelpazesi desteği sistemin avantajlarından. Ayrıca sistemin ev kullanıcıları ve küçük ölçekli üreticiler için pratik ve erişilebilir bir çözüm sunduğunu da göstermektedir.

Gelecek çalışmalarda cihaz üzerinde çalışan hafif CNN modelleri ile hibrit yaklaşım, genişletilmiş sensör altyapısı (pH, NPK), tarla koşullarında testler ve ekonomik etki analizi önerilmektedir. Sonuç olarak bu çalışma, IoT ve üretken yapay zekâ entegrasyonunun akıllı tarım uygulamalarındaki potansiyelini göstermiş ve çok modlu analiz yaklaşımının etkinliğini bilimsel olarak doğrulamıştır.

Kaynakça/References

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., & Aggoune, E. H. M. (2019). Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. *IEEE Access*, 7, 129551–129583. doi:10.1109/ACCESS.2019.2932609

Aydınbaş, G. (2023). İktisadi Perspektiften Akıllı Tarım (Tarım 4.0) Üzerine Bir İnceleme. *BİLTÜRK Journal of Economics and Related Studies*. doi:10.47103/bilturk.1218500

Blessy, A., Kumar, A., A, P., Md, A. Q., Alharbi, A. I., Almusharraf, A., & Khan, S. B. (2023). Sustainable Irrigation Requirement Prediction Using Internet of Things and Transfer Learning. *Sustainability (Switzerland)*, 15(10). doi:10.3390/su15108260

Büyüközkan, G., & Uztürk, D. (2024). Integrated design framework for smart agriculture: Bridging the gap between digitalization and sustainability. *Journal of Cleaner Production*, 449. doi:10.1016/j.jclepro.2024.141572

Çakmakçı, M. F., & Çakmakçı, R. (2023). Uzaktan Algılama, Yapay Zekâ ve Geleceğin Akıllı Tarım Teknolojisi Trendleri. *European Journal of Science and Technology*, (52), 234–246. doi:10.5281/zenodo.10439935

Calone, R., Raparelli, E., Bajocco, S., Rossi, E., Crecco, L., Morelli, D., ... Bregaglio, S. (2025). Analysing the potential of ChatGPT to support plant disease risk forecasting systems. *Smart Agricultural Technology*, 10. doi:10.1016/j.atech.2025.100824

Chen, C. J., Huang, Y. Y., Li, Y. S., Chang, C. Y., & Huang, Y. M. (2020). An AIoT Based Smart Agricultural System for Pests

Detection. *IEEE Access*, 8, 180750–180761. doi:10.1109/ACCESS.2020.3024891

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46.

Elashmawy, R., & Uysal, I. (2023). Precision Agriculture Using Soil Sensor Driven Machine Learning for Smart Strawberry Production. *Sensors*, 23(4). doi:10.3390/s23042247

Elavarasan, D., & Durairaj Vincent, P. M. (2020). Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. *IEEE Access*, 8, 86886–86901. doi:10.1109/ACCESS.2020.2992480

Gagliardi, G., Lupia, M., Cario, G., Gaccio, F. C., D'angelo, V., Cosma, A. I. M., & Casavola, A. (2021). An internet of things solution for smart agriculture. *Agronomy*, 11(11). doi:10.3390/agronomy11112140

Gülcüoğlu, E., Ustun, A. B., & Seyhan, N. (2021). Comparison of Flutter and React Native Platforms. *Journal of Internet Applications and Management*. doi:10.34231/iuyd.888243

Gupta, A., & Nahar, P. (2023). Classification and yield prediction in smart agriculture system using IoT. *Journal of Ambient Intelligence and Humanized Computing*, 14(8), 10235–10244. doi:10.1007/s12652-021-03685-w

Jaliyagoda, N., Lokuge, S., Gunathilake, P. M. P. C., Amaratunga, K. S. P., Weerakkody, W. A. P., Bandaranayake, P. C. G., & Bandaranayake, A. U. (2023). Internet of things (IoT) for smart agriculture: Assembling and assessment of a low-cost IoT system for polytunnels. *PLoS ONE*, 18(5 May). doi:10.1371/journal.pone.0278440

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Enhancing smart farming through the applications of Agriculture 4.0 technologies. *International Journal of Intelligent Networks*, 3, 150–164. doi:10.1016/j.ijin.2022.09.004

Kutubuddin, K., & Liyakat, S. (2024). *Smart Agriculture based on AI-Driven-IoT (AIIoT): A KSK Approach. Advance Research in Communication Engineering and its Innovations* (Vol. 1).

Landis, J. R., & Koch, G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174.

Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J., & Kazimi, Z. A. (2022). Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model. *Plants*, 11(17). doi:10.3390/plants11172230

Li, L., Zhang, S., & Wang, B. (2021). Plant Disease Detection and Classification by Deep Learning - A Review. *IEEE Access*. Institute of Electrical and Electronics Engineers Inc. doi:10.1109/ACCESS.2021.3069646

Mesías-Ruiz, G. A., Pérez-Ortiz, M., Dorado, J., de Castro, A. I., & Peña, J. M. (2023). Boosting precision crop protection towards agriculture 5.0 via machine learning and emerging technologies: A contextual review. *Frontiers in Plant Science*. Frontiers Media S.A. doi:10.3389/fpls.2023.1143326

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7(September). doi:10.3389/fpls.2016.01419

Rani, S., Mishra, A. K., Kataria, A., Mallik, S., & Qin, H. (2023). Machine learning-based optimal crop selection system

in smart agriculture. *Scientific Reports*, 13(1). doi:10.1038/s41598-023-42356-y

Ray, P. P. (2017). Internet of things for smart agriculture: Technologies, practices and future direction. *Journal of Ambient Intelligence and Smart Environments*, 9(4), 395–420. doi:10.3233/AIS-170440

Shaikh, T. A., Rasool, T., Venington, K., & Yaseen, S. M. (2025, June 1). The role of large language models in agriculture: harvesting the future with LLM intelligence. *Progress in Artificial Intelligence*. Springer Science and Business Media Deutschland GmbH. doi:10.1007/s13748-024-00359-4

Sowmya, B. J., Meeradevi, A. K., Supreeth, S., Pradeep Kumar, D., Ravi Kumar, B. N., Rohith, S., ... Patil, A. U. (2025). Leveraging machine learning for intelligent agriculture. *Discover Internet of Things*, 5(1). doi:10.1007/s43926-025-00132-6

Yang, J., Ma, S., Li, Y., & Zhang, Z. (2022). Efficient Data-Driven Crop Pest Identification Based on Edge Distance-Entropy for Sustainable Agriculture. *Sustainability (Switzerland)*, 14(13). doi:10.3390/su14137825

BÖLÜM 6

CUZN ALAŞIMLARININ ÖZDİRENCİNE SODYUM SİTRAT ETKİSİ: İSTATİSTİKSEL BİR ANALİZ

HASAN GÜLER¹
RASİM ÖZDEMİR²

Giriş

Pirinç olarak bilinen bakır-çinko alaşımları (Cu-Zn), üstün elektriksel iletkenlik ve termal özellikleri sayesinde (Brenner, 1963) elektrik-elektronik ürünlerinde, dekoratif uygulamalarda ve korozyon direnci gerektiren alanlarda yaygın kullanılmaktadır (Juškėnas, Karpavičienė, Pakštis, Selskis, & Kapočius, 2007). Son yıllarda bu alaşımlar, elektronik bileşenler ve lityum iyon bataryaları için potansiyel malzemeler olarak değerlendirilmektedir. Malzemelerin fiziksel özellikleri, yapı ve bileşimleriyle doğrudan ilişkilidir (Porter, 2009). Özellikle elektriksel öz direnç, elektrik devrelerinde kullanılan malzemelerin performansını etkileyen kritik bir parametredir. Cu-Zn alaşımlarının üretiminde elektrokimyasal depolama yöntemi, düşük maliyet, düşük sıcaklık uygulanabilirliği

¹ Dr. Öğr. Üyesi, Kilis 7 Aralık Üniversitesi, Bilgisayar Teknolojileri Bölümü, Orcid: 0000-0001-5312-171X

² Prof. Dr., Kilis 7 Aralık Üniversitesi, Elektrik Bölümü, Orcid: 0000-0003-1439-0444

ve film kalınlığı-homojenliği üzerinde hassas kontrol imkanı sunması nedeniyle tercih edilmektedir (Mohanty, 2011; Pletcher & Walsh, 1993; Walsh & Ponce De León, 2014). Siyanürün toksisitesi sebebiyle, sitrat gibi çevre dostu kompleks oluşturunca ajanlar giderek daha fazla kullanılmaktadır (Liu, Shi, & Reddy, 2020; R. Özdemir & Karahan, 2019). Cu ve Zn'nin indirgenme potansiyelleri arasındaki büyük fark, anormal birlikte depolanma olarak adlandırılan karmaşık elektrokimyasal kinetikler sergiler. Sitrat konsantrasyonu, bakırın indirgenme kinetiğini etkileyerek nihai alaşım özelliklerini belirler (Oulmas vd., 2019). Sülfat anyonları katodik reaksiyonu inhibe ederken, klorür anyonları bakır indirgenmesini hızlandırabilir (Slupska & Ozga, 2014). Elektrodepozisyonla üretilen Cu-Zn alaşımlarında öz direnç değerleri, mekanik yöntemlerle üretilenlere göre %25-45 daha yüksektir ve film bileşimi, faz yapısı ve tane sınırları gibi mikroyapısal faktörlere bağlıdır (Fairbank, 1944). Akım yoğunluğu, sitrat konsantrasyonu, sıcaklık, Cu/Zn oranı ve depolama potansiyeli gibi elektrokimyasal parametreler, alaşımların mikroyapı, kimyasal bileşim ve elektriksel-mekanik özelliklerini önemli ölçüde etkiler (To, Park, Kim, Cho, & Myung, 2022). Elektrodepozisyonla elde edilen CuZn alaşımları genellikle nanokristalin α veya β fazları sergiler (Pereloma & Edmonds, 2012). Morfoloji küresel, dendritik veya karnabakar benzeri yapılar gösterebilir (Chandrasekar & Pushpavanam, 2008). Öz direnç, Matthiessen kuralına göre ısıt titreşim, safsızlık ve kafes kusurlarının toplamıyla açıklanır (Benrazzouq, Ghanbaja, Migot, Milichko, & Pierson, 2025). Üretim parametreleri ile alaşım özellikleri arasındaki karmaşık, doğrusal olmayan ilişkileri sistematik olarak ortaya koymak için ileri istatistiksel yöntemler önemli fırsatlar sunmaktadır (Xiao, Li, Bordas, & Kim, 2023).

Bu çalışmanın amacı, Cu-Zn alaşımlarının morfolojik, yapısal, elektriksel ve mekanik özellikleri ile elektrokimyasal

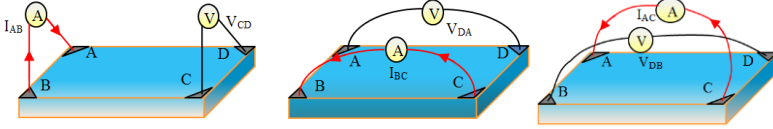
depolama parametreleri (özellikle sitrat konsantrasyonu, bakır çinko madde miktarı oranı, sıcaklık, akım yoğunluğu ve depolanma potansiyeli) arasındaki kompleks ilişkileri sistematik olarak ortaya koymaktır. Bu amaçla: (1) deneysel parametreler üzerinde hassas kontrol sağlanarak elde edilen alaşım özelliklerinin dönüşümlü voltametri (CV), ve mikroyapısal analizler (XRD, SEM, EDX) gibi kapsamlı tekniklerle karakterize edilmesi; (2) veri temizlik, tanımlayıcı analiz ve grup karşılaştırmaları yapılması; (4) SC grupları arasında özdirenç farkları için hipotez testleri (ANOVA/Kruskal–Wallis) ve post-hoc karşılaştırmalar yapılmasıdır.

Deneysel Çalışma

Cu-Zn alaşım ince filmler elektrokimyasal depolama sistemi kullanılarak alüminyum altlıklar üzerine üretildi. Yapısal ve yüzeysel karakterizasyon için çeşitli deneyler gerçekleştirildi: potansiyostat-galvanostat ile dönüşümlü voltametri (CV) deneyleri, taramalı elektron mikroskobu (SEM) görüntüleme, X-ışını difraksiyon spektrometresi (XRD) ölçümleri yapıldı. Film içerisindeki madde miktarlarının dağılımı Enerji Dağılımlı X-Işını Spektroskopisi (EDX) analizi ile incelendi. Elektriksel özdirenç dört nokta kontak metodu kullanılarak ölçüldü. Düşük direnç değerlerinin ölçümünde, alet hataları, bağlantı iletkenleri dirençleri ve temas dirençleri ölçüm sonucunu etkileyerek hatalı değerlere yol açabilir. Bu hataları minimize etmek için dört nokta kontak yöntemi geliştirilmiştir. İnce metal filmler ve yarıiletken malzemelerin elektriksel özelliklerinin ölçümünde sıklıkla kullanılan bu yöntem, örnek üzerinden geçen akım ve belirli bir bölgedeki potansiyel fark ölçümüne dayanır. Van Der Pauw yöntemi, serbest formlu düzlemsel levha şeklindeki malzemelerin özdirenç ölçümlerinde kullanılır. Bu ölçümler için numunenin kenarının dört noktasına omik kontaklar yapılır (Cultrera, Serazio, Fabricius, & Callegaro, 2024; Rasim Özdemir, Karahan, & Karabulut, 2016). Şekil 1'de görüldüğü gibi A,

B, C ve D noktalarından elektrotlar numune yüzeyine temas ettirilir. Ölçülen akım ve gerilim (I, V) değerleri ile elektriksel özellikler hesaplanır.

Şekil 1. Van Der Pauw Metodu ölçme şekilleri



Özdirenç ölçme işleminde sıcaklık 100 K ile 405 K arasında bir kryostat yardımıyla değiştirilerek bilgisayar kontrolünde tam otomasyonlu olarak gerçekleştirildi (Şekil 2). Veriler 10 °C aralıklarla kaydedildi. Bir numunenin ölçme işlemi yaklaşık 6 saat sürmektedir.

Şekil 2. Elektriksel ve manyetik özellikler ölçme düzeneği



CuZn Alaşım İnce Film Deney Üretim Verileri

Elektrokimyasal depolama yöntemiyle üretilen ince filmlerin üretim şartları ve 4 adet ince filme ait banyo içeriği deney verileri Tablo 1'de verildi. Banyo içerisinde sodyum sitrat ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$) 0.3-0.9 Mol.lit⁻¹ arasında değiştirilirken, diğer parametreler sabit tutuldu: bakır sülfat ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) 0.06 Mol.lit⁻¹, çinko sülfat ($\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$) 0.2 Mol.lit⁻¹, pH 5.8, üretim zamanı 60 dakika, üretim akımı 60 mA,

elektrolit sıcaklığı 20 °C. Her bir banyoda depolanan ince filmlerin Dönüşümlü Voltametri Deneyleri (CV) ile depolanmaya başlama potansiyel değerleri tespit edildi. Elektriksel özdirenci ölçülen dört adet Cu-Zn ince film alaşımlarının ölçüm sonuçları grafikte verilmiştir. Grafikte ortam sıcaklığının 100 K ile 405 K arasında değiştirilmesi karşısında ölçülen elektriksel özdirencin sıcaklıkla değişim eğrisi verilmiştir. Cu-Zn alaşım filmlerin XRD analiz verilerinden kristal yapıları belirlendi. XRD verilerinden tane büyüklüğü Debye-Scherrer formülüyle hesaplandı. Ortalama tane boyutunu bulmak için Scherrer formülü kullanılmıştır; formülde L tane boyutunu, 2 θ kırınım pikini, B yarı yükseklik genişliğini ve λ dalga boyunu göstermektedir (Leontyev vd., 2018; Rasim Ozdemir & İsmail Hakkı Karahan, 2023).

Veriseti

Elektriksel özdirencin kapsamlı istatistiksel analizi için deneysel verilerden bir veri seti oluşturuldu. Veri seti elektrokimyasal depolama koşulları, banyo ve film bileşim oranlarını içeren 127 örnekten oluşmakta ve sekiz bağımsız değişken içermektedir: 1) %ECu, 2) %EZn, 3) %FCu, 4) %FZn, 5) tane büyüklüğü (G, nm), 6) depolanma voltajı (V), 7) elektrolit sıcaklığı (St, °C), 8) ölçüm sıcaklığı (T, Kelvin). Bağımlı değişken elektriksel özdirenç (ρ , $\mu\Omega.cm$) değeridir. Analiz güvenilirliğini sağlamak için kapsamlı ön işleme yapılmıştır. Aykırı değer tespiti için çeyrekler arası aralık (IQR) yöntemi uygulanmış (Tukey, 1949) ve veri setinde anlamlı aykırı değer bulunmamıştır.

Dağılım Özellikleri ve Normallik Değerlendirmesi

Elektrokimyasal depolanma prosesi ile üretilen CuZn alaşımlarının fiziksel ve elektriksel özelliklerine ait tanımlayıcı istatistikler sistematik olarak incelenmiştir. Analiz, merkezi eğilim

ölçüleri, dağılım ölçüleri ve dağılım şekil parametrelerini (çarpıklık, basıklık) içermektedir (Tablo 1).

Tablo 1. Veri setinin tanımlayıcı istatistikleri

Parametre	Ortalama	Std Sapma	Min	Max	IQR	Çarpıklık	Basıklık	CV (%)
Sıcaklık (°C)	252.13	90.93	100.00	407.00	151.25	-0.000	-1.185	35.90
FCu (%)	20.89	12.33	6.15	38.31	19.47	0.348	-1.516	54.31
FZn (%)	79.12	12.33	61.69	93.85	19.47	-0.348	-1.516	9.53
SC Kons. (mol/L)	0.598	0.224	0.300	0.900	0.400	0.000	-1.516	23.45
ECu (%)	22.8	0.0	22.8	22.8	0.000	0.000	0.000	0.00
EZn (%)	77.2	0.0	77.2	77.2	0.000	0.000	0.000	0.00
Tane Boyutu (nm)	29.31	1.81	28.00	32.40	0.500	-0.243	-1.516	0.73
Potansiyel (V)	1.533	0.012	1.517	1.549	0.021	-0.176	-1.516	0.57
Resistivity ($\mu\Omega \cdot \text{cm}$)	0.158	0.042	0.091	0.256	0.058	0.088	-0.780	22.77

Veri setinde değişkenler heterojen yapı göstermektedir. Sıcaklığın geniş IQR değeri (151.25°C) farklı termal koşulların araştırıldığını, tane boyutunun dar IQR'si (0.5 nm) ise mikroyapısal homojenliği yansıtmaktadır. Çarpıklık analizinde tüm değişkenlerde $|\text{çarpıklık}| < 0.5$ koşulu sağlanmakta, bu da simetrik dağılımları göstermektedir. Sıcaklık neredeyse mükemmel simetri sergilerken (-0.000), FCu pozitif çarpıklık (+0.348) göstererek düşük konsantrasyonlarda yoğunlaşmaktadır. Varyasyon katsayıları incelendiğinde, FCu'nun yüksek CV değeri (%54.31) kompozisyonel çeşitliliği işaret ederken, tane boyutu (%0.73) ve depolanma potansiyeli (%0.57) düşük CV değerleriyle yüksek tekrarlanabilirlik göstermektedir. ECu% ve EZn% parametrelerinin sıfır standart sapması, elektrolit bileşiminin sabit tutulduğunu ve kompozisyonel değişikliklerin yalnızca proses parametreleri aracılığıyla sağlandığını göstermektedir.

Normalite ve Varyans Homojenliği Analizleri

İstatistiksel analizlerin geçerliliği ve uygun test seçimi, verilerin temel varsayımları sağlayıp sağlamadığına bağlıdır (Field, 2013). Normal dağılım ve varyans homojenliği varsayımları dört

farklı normalite testi (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling, D'Agostino-Pearson) ile kapsamlı olarak değerlendirilmiştir (Razali, Wah, & others, 2011).

Tablo 2 SC konsantrasyon gruplarına göre (Shapiro-Wilk testi) normalite ve varyans homojenliği test sonuçları

Test Kategorisi	Test Türü	SC Konsantrasyonu (mol/L)	İstatistik (W)	p-değeri	Sonuç
Normalite (Grup Bazlı)	Shapiro-Wilk	0.3	0.910	0.011*	Normal değil
		0.5	0.933	0.047*	Normal değil
		0.7	0.930	0.039*	Normal değil
		0.9	0.938	0.071	Normal
Varyans Homojenliği	Levene's Test	Tüm gruplar	-	0.731	Homojen
	Bartlett's Test	Tüm gruplar	-	0.818	Homojen

* $p < 0.05$

Grup bazlı Shapiro-Wilk testi (Tablo 2), yalnızca 0.9 mol/L SC konsantrasyon grubunun normal dağılım sergilediğini göstermiştir ($W=0.938$, $p=0.071$). Diğer konsantrasyon grupları (0.3, 0.5 ve 0.7 mol/L) normallikten anlamlı sapma göstermektedir ($p<0.05$). Bu bulgu, düşük ve orta konsantrasyonlarda elektrokimyasal depolama prosesinin heterojen dinamikleri ile ilişkilendirilebilir (Mohanty, 2011; Walsh & Ponce De León, 2014).

Tablo 3 Tüm değişkenler için kapsamlı normalite testleri sonuçları

Değişken	Shapiro-Wilk		Kolmogorov-Smirnov		Anderson-Darling	D'Agostino-Pearson		Sonuç
	İstatistik	p	İstatistik	p	İstatistik	İstatistik	p	
Resistivity	0.969	0.0059**	0.0566	0.7899	0.7986	9.75	0.0076**	Hayır
Grainsize (nm)	0.619	<0.0001***	0.4203	<0.0001***	23.4279	25.67	<0.0001***	Hayır
Cu%	0.836	<0.0001***	0.2346	<0.0001***	7.6204	196.07	<0.0001***	Hayır
Zn%	0.836	<0.0001***	0.2346	<0.0001***	7.6204	196.07	<0.0001***	Hayır

*** $p < 0.001$, ** $p < 0.01$

Elektriksel özdirenç için testler kısmen tutarsız sonuçlar vermiştir. Shapiro-Wilk ($p=0.0059$) ve D'Agostino-Pearson ($p=0.0076$) testleri normallikten sapma gösterirken, Kolmogorov-

Smirnov testi ($p=0.7899$) normalliği desteklemektedir. Bu tutarsızlık, Kolmogorov-Smirnov testinin düşük istatistiksel gücünden kaynaklanabilir (Steinskog, Tjøstheim, & Kvamstø, 2007). Anderson-Darling istatistiği ($A^2=0.7986$) kritik değerin üzerinde olup normallikten sapmayı doğrulamaktadır. Tane boyutu, tüm normalite testlerinde güçlü ve tutarlı şekilde normallikten sapma göstermektedir (tüm $p<0.0001$). Düşük Shapiro-Wilk istatistiği ($W=0.6196$) ve yüksek Anderson-Darling istatistiği ($A^2=23.4279$), dağılımın ciddi şekilde normal dağılımdan farklılaştığını göstermektedir. Cu% ve Zn% değişkenleri matematiksel olarak birbirini tamamladığından ($\text{Cu\%} + \text{Zn\%} = 100\%$) identik sonuçlar vermiştir. Her iki değişken de tüm testlerde yüksek anlamlılıkta normallikten sapma göstermektedir ($p<0.0001$). Yüksek D'Agostino-Pearson istatistiği ($K^2=196.0661$), kompozisyonel dağılımın normal dağılımla uyumsuz olduğunu göstermektedir (Assaf, El-Seidy, Abou-Krishna, & Eissa, 2015). Varyans homojenliği testleri (Levene $p=0.731$, Bartlett $p=0.818$) tüm SC konsantrasyon grupları arasında varyansların eşit olduğunu göstermiştir (Brown & Forsythe, 1974). Normallik varsayımının ihlal edilmesi nedeniyle, grup karşılaştırmaları için parametrik olmayan Kruskal-Wallis H-testi tercih edilmiştir (Kruskal & Wallis, 1952; Zimmerman, 1998).

Kruskal-Wallis H-Testi ve Etki Boyutu Analizi

SC konsantrasyonunun elektriksel özdirenç üzerindeki etkisini değerlendirmek için Kruskal-Wallis H-testi uygulanmış ve etki boyutu epsilon-squared (ϵ^2) metriği ile ölçülmüştür.

Tablo 4 Kruskal-Wallis h-testi ve etki boyutu sonuçları (resistivity)

Test Parametresi	Değer	p-değeri	Yorum
H-istatistiği	76.37	<0.001***	Son derece güçlü grup etkisi
Serbestlik derecesi	3.00	-	4 grup karşılaştırması
Epsilon-squared (ϵ^2)	0.597	-	Çok büyük etki boyutu
Açıklanan varyans	%59.7	-	Cohen kriterlerine göre

$$***p < 0.001$$

Kruskal-Wallis H-testi (Tablo 4), SC konsantrasyon grupları arasında son derece anlamlı farklar olduğunu ortaya koymuştur ($H=76.37$, $p<0.001$). Epsilon-squared değeri ($\epsilon^2=0.597$), SC konsantrasyonunun özdirenç varyansının %59.7'sini açıkladığını göstermektedir. Cohen (1988) sınıflandırmasına göre bu değer "çok büyük etki" kategorisindedir ($\epsilon^2>0.14$ eşliğinin çok üzerinde), SC konsantrasyonunun kritik bir kontrol parametresi olduğunu doğrulamaktadır (Nakagawa & Cuthill, 2007).

Post-hoc İkili Karşılaştırmalar: Dunn Testi

Hangi konsantrasyon çiftleri arasında anlamlı farklar bulunduğunu belirlemek için Dunn post-hoc testi uygulanmış (Dunn 1964) ve çoklu karşılaştırmalarda Tip I hata kontrolü için Bonferroni düzeltilmesi kullanılmıştır (Bonferroni, 1936).

Tablo 5 Dunn Post-hoc ikili karşılaştırmaları

Karşılaştırma (mol/L)	Medyan Farkı ($\Omega \cdot m$)	U-İstatistik	p-düzeltilmeli	Güven Aralığı (%95)	Anlamlılık
0.3 vs 0.5	-0.020	313.0	0.046*	[-0.0388, -0.0012]	Evet
0.3 vs 0.7	-0.051	106.0	<0.001***	[-0.0698, -0.0322]	Evet
0.3 vs 0.9	-0.086	2.0	<0.001***	[-0.1048, -0.0672]	Evet
0.5 vs 0.7	-0.031	217.0	<0.001***	[-0.0498, -0.0122]	Evet
0.5 vs 0.9	-0.066	32.0	<0.001***	[-0.0848, -0.0472]	Evet
0.7 vs 0.9	-0.035	199.0	<0.001***	[-0.0538, -0.0162]	Evet

$$*p < 0.05, ***p < 0.001$$

Post-hoc Dunn testleri (Tablo 5), tüm SC konsantrasyon çiftleri arasında istatistiksel olarak anlamlı farklar bulunduğunu göstermiştir. En zayıf fark 0.3 vs 0.5 mol/L grupları arasında (medyan fark = -0.020 $\Omega \cdot m$, $p=0.046$), en güçlü fark ise 0.3 vs 0.9 mol/L grupları arasında tespit edilmiştir (medyan fark = -0.086 $\Omega \cdot m$, $p<0.001$). Medyan farklılıklarının negatif işaretli olması, SC konsantrasyonu arttıkça özdirenç değerlerinin sistematik olarak

azaldığını göstermektedir. 0.3 mol/L'den 0.9 mol/L'ye geçişte gözlenen $-0.086 \Omega \cdot \text{m}$ 'lik azalma, başlangıç değerine göre yaklaşık %180'lik bir iyileşmeyi temsil etmektedir.

Parametrik Varyans Analizi (ANOVA) ve Kapsamlı Karşılaştırmalar

Normallik varsayımı ihlal edilse de ANOVA'nın orta dereceli sapmalar karşısında robust yapısı (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017) nedeniyle hem parametrik (ANOVA) hem de parametrik olmayan (Kruskal-Wallis) testler tüm değişkenler için uygulanarak sonuçların tutarlılığı değerlendirilmiştir.

Tablo 6 SC konsantrasyon gruplarına göre tüm değişkenlerin kapsamlı istatistiksel karşılaştırma sonuçları

Değişken	Parametrik Test (ANOVA)			Parametrik Olmayan Test			Yorum
	F	p-değeri	η^2	H	p-değeri	ϵ^2	
Resistivity	62.09	1.64×10^{-24}	0.602***	76.37	1.85×10^{-16}	0.596***	Çok Büyük Etki
Grainsize	-	<0.001	1.000***	126.00	3.94×10^{-27}	1.000***	Mükemmel Ayırım
Cu%	-	<0.001	1.000***	126.00	3.94×10^{-27}	1.000***	Mükemmel Ayırım
Zn%	-	<0.001	1.000***	126.00	3.94×10^{-27}	1.000***	Mükemmel Ayırım

*** $p < 0.001$

Hem parametrik hem de parametrik olmayan testler (Tablo 6), tüm değişkenler için $p < 0.001$ düzeyinde yüksek anlamlı sonuçlar vermiştir. Özdirenç için F-istatistiği (62.09) ve H-istatistiği (76.37) değerleri, her iki yaklaşımın da tutarlı sonuçlar ürettiğini göstermektedir. Parametrik eta-squared ($\eta^2=0.602$) ve parametrik olmayan epsilon-squared ($\epsilon^2=0.596$) değerlerinin birbirine çok yakın olması, ANOVA'nın normallik ihlallerine karşı gösterdiği robust yapıyı doğrulamaktadır. Etki boyutu analizinde en dikkat çekici sonuçlar tane boyutu, Cu% ve Zn% değişkenlerinde elde edilmiştir. Bu üç parametre için hem η^2 hem de ϵ^2 değerleri tam olarak 1.000'e ulaşmıştır. Bu sonuç, söz konusu parametrelerin SC

konsantrasyonuna bağılı ve SC grupları arasında %100 ayırım gücü olduğunu göstermektedir. Başka bir deyişle, bu değişkenlerdeki tüm varyans SC konsantrasyonu değişimiyle açıklanabilmektedir. Bu bulgu, elektrokimyasal depolama prosesinin deterministik karakterini ve SC konsantrasyonunun mikroyapısal özelliklerin kontrolünde kritik rolünü nicel olarak ortaya koymaktadır.

Tukey HSD Post-hoc Analizi

ANOVA sonuçlarını detaylandırmak amacıyla, öz direnç değişkeni için Tukey HSD post-hoc testi uygulanmıştır.

Tablo 7 Tukey HSD Post-hoc karşılaştırmaları (Resistivity)

Grup Karşılaştırması (mol/L)	Ortalama Farkı ($\Omega \cdot m$)	Standart Hata	p-değeri	Güven Aralığı (%95)	Anlamlılık
0.3 vs 0.5	-0.017271	0.0089	0.055	[-0.0348, 0.0002]	Eşik değerde
0.3 vs 0.7	-0.050553	0.0089	<0.001***	[-0.0681, -0.0330]	Yüksek anlamlı
0.3 vs 0.9	-0.085384	0.0089	<0.001***	[-0.1030, -0.0677]	Yüksek anlamlı
0.5 vs 0.7	-0.033282	0.0089	<0.001***	[-0.0508, -0.0158]	Yüksek anlamlı
0.5 vs 0.9	-0.068113	0.0089	<0.001***	[-0.0858, -0.0505]	Yüksek anlamlı
0.7 vs 0.9	-0.034831	0.0089	<0.001***	[-0.0525, -0.0172]	Yüksek anlamlı

*** $p < 0.001$

Tukey HSD analizi (Tablo 7), öz direnç için sistematik ve dozaj-bağımlı bir trend ortaya koymuştur. 0.3 mol/L ile 0.5 mol/L arasındaki karşılaştırma hariç ($p=0.055$), tüm ikili grup karşılaştırmaları $p<0.001$ düzeyinde istatistiksel olarak anlamlıdır. 0.3-0.5 mol/L karşılaştırmasının eşik değerde bulunması, bu konsantrasyon aralığının bir geçiş bölgesi olabileceğini ve 0.5 mol/L'nin kritik bir eşik konsantrasyon seviyesi olarak değerlendirilebileceğini düşündürmektedir (Pletcher & Walsh, 1993). En büyük ortalama fark, en düşük ve en yüksek konsantrasyon grupları arasında gözlenmiştir (0.3 vs 0.9 mol/L: -0.085384 $\Omega \cdot m$). Bu değer, 0.3 mol/L'deki ortalama öz direnç

değerine göre yaklaşık %180'lik bir azalmayı temsil etmektedir. %95 güven aralıkları tüm anlamlı karşılaştırmalarda sıfırı içermemekte olup, sonuçların istatistiksel güvenilirliğini desteklemektedir. Ardışık konsantrasyon grupları arasındaki ortalama farklılıklar incelendiğinde (0.3→0.5: -0.017 $\Omega \cdot m$, 0.5→0.7: -0.033 $\Omega \cdot m$, 0.7→0.9: -0.035 $\Omega \cdot m$), özdirenç azalmasının doğrusal olmadığı, ancak monoton bir trend sergilediği görülmektedir. Bu durum, SC konsantrasyonunun etkisinin özellikle 0.5 mol/L üzerinde daha belirgin hale geldiğini göstermektedir. SC konsantrasyonunun özdirenç varyansının %60'ına yakınına açıklaması ($\epsilon^2=0.596$), bu parametrenin hassas kontrolü ile hedeflenen elektriksel özelliklerin prediktif olarak modellenebileceğini göstermektedir. 0.5 mol/L'nin altındaki konsantrasyonlarda sodyum sitratın kompleks oluşturma kapasitesinin sınırlı olabileceği, bu değer üzerinde ise iyonik ortam ve kompleksleşme dinamiklerinin stabilize olabileceği değerlendirilmektedir (Kruglikov, Kudriavtsev, Vorobiova, & Antonov, 1965). 0.7-0.9 mol/L konsantrasyon aralığı, düşük özdirenç değerleri elde etmek için optimal işletme bölgesi olarak öne çıkmaktadır.

Bu kapsamlı istatistiksel analiz, SC konsantrasyonunun CuZn alaşımlarının elektrokimyasal depolanmasında elektriksel ve mikroyapısal özelliklerin belirlenmesinde dominant rol oynayan kritik bir parametre olduğunu nicel olarak kanıtlamaktadır.

Sitrat Konsantrasyonunun (SC) Özdirenç Üzerindeki Kritik Kontrolü

SC konsantrasyon grupları arasında gerçekleştirilen Kruskal-Wallis H-testi, istatistiksel olarak son derece anlamlı farklar ortaya koymuştur ($H=76.37$, $p<0.001$, $\epsilon^2=0.597$). Epsilon-squared değerinin 0.597 olması, SC konsantrasyonunun özdirenç varyansının yaklaşık %60'ını açıkladığını göstermektedir. Bu etki boyutu, Cohen'in (1988) sınıflandırmasında “çok büyük etki”

kategorisine girmekte ve SC konsantrasyonunun elektrokimyasal depolama prosesinde kritik bir kontrol parametresi olduğunu doğrulamaktadır. Post-hoc Dunn testleri, tüm SC konsantrasyon çiftleri arasında anlamlı farklar bulunduğunu göstermiştir. En güçlü fark, en düşük (0.3 mol/L) ve en yüksek (0.9 mol/L) konsantrasyon grupları arasında gözlenmiştir (medyan fark=-0.086 $\Omega \cdot m$, $U=2.0$, $p<0.001$). SC konsantrasyonunun artmasıyla özdirençte gözlenen monoton azalma trendi, sitrat iyonlarının kompleksleşme mekanizması ile açıklanabilir. Yüksek SC konsantrasyonlarında, bakır ve çinko iyonlarının sitrat kompleksleri oluşturması, elektrodpozisyon kinetiğini değiştirerek daha düzenli kristal yapı oluşumunu teşvik etmektedir. Kritik bir bulgu, 0.3 mol/L ile 0.5 mol/L arasındaki geçişin eşik değerde bulunmasıdır ($p=0.055$). Bu bölge, sitrat kompleksleşme kapasitesinin saturasyona ulaştığı ve elektrokimyasal nükleasyon kinetiğinin modifiye olduğu bir geçiş rejimi olarak değerlendirilebilir. Scharifker-Hills modeli çerçevesinde, bu konsantrasyon rejimi anında nükleasyondan progresif nükleasyona geçişi temsil edebilir (Sun & Zangari, 2023). Parametrik ANOVA testleri ile elde edilen sonuçlar, Kruskal-Wallis testleri ile tutarlılık göstermiştir. Özdirenç için F-istatistiği (62.09) ve H-istatistiği (76.37) değerleri, her iki yaklaşımın da benzer sonuçlar ürettiğini göstermektedir. Parametrik eta-squared ($\eta^2=0.602$) ve parametrik olmayan epsilon-squared ($\epsilon^2=0.596$) değerlerinin birbirine çok yakın olması dikkat çekicidir. Tukey HSD post-hoc analizi, özdirenç için sistematik ve dozaj-bağımlı bir trend ortaya koymuştur. 0.3 mol/L ile 0.5 mol/L arasındaki karşılaştırma hariç ($p=0.055$), tüm ikili grup karşılaştırmaları $p<0.001$ düzeyinde istatistiksel olarak anlamlıdır. En büyük ortalama fark, en düşük ve en yüksek konsantrasyon grupları arasında gözlenmiştir (0.3 vs 0.9

mol/L: $-0.085384 \Omega \cdot m$), bu deęer yaklaşık %180'lik bir azalmayı temsil etmektedir.

Kompozisyonel Parametrelerin Özdirenç ile İlişkisi

XRD ve EDX karakterizasyon sonuçları, SC konsantrasyonunun film kompozisyonunu sistematik olarak etkilediğini göstermiştir. Banyo içerisindeki Cu %22.8 ve Zn %77.2 oranında sabit olmasına rağmen, depolanan CuZn ince film alaşımların içerisindeki Cu oranı %13.01 ile %38.31 arasında, Zn oranı ise %61.69 ile %86.99 arasında değişmektedir. Bu bulgular, SC konsantrasyonunun anormal birlikte depolanma mekanizmasını doğrudan etkilediğini göstermektedir. Bakırın intrinsik elektriksel özellikleri ($\rho_{Cu} \approx 1.68 \mu\Omega \cdot cm$) ve çinkonun daha yüksek intrinsik direnci ($\rho_{Zn} \approx 5.90 \mu\Omega \cdot cm$) nedeniyle, film kompozisyonundaki değişimler özdirenç üzerinde önemli etkiler yaratmaktadır. Nordheim kuralı, ikili alaşımlarda özdirencin kompozisyonla değişimini açıklamaktadır: $\rho_{alaşım} = \rho_{Cu} + C \times Zn^{1-x} \times (Zn^{1-x})$. CuZn sisteminde çinko atomları, bakır matrisinde çözünmüş durumda elektron saçılma merkezleri oluşturarak Fermi yüzeyinde pertürbasyonlara yol açmaktadır (Nordheim, 1931).

Mikroyapısal Özelliklerin Özdirenç ile İlişkisi

XRD analizlerinden Debye-Scherrer formülü ile hesaplanan tane boyutu deęerleri, SC konsantrasyonuna baęlı olarak sistematik deęişim göstermiştir (28.0-32.4 nm aralığında). Tane boyutundaki bu deęişimler, tane sınırı saçılımının elektron taşınımı üzerindeki etkisini ortaya koymaktadır. Klasik Mayadas-Shatzkes modeli, tane sınırı yoğunluğunun artmasıyla elektronların ortalama serbest yolunun kıaldığını ve özdirencin arttığını öngörmektedir: $\rho = \rho_0(1 + (3(1-p)/2)(\lambda/d))$. Burada p tane sınırı specularity parametresi ($0 \leq p \leq 1$), λ elektron ortalama serbest yolu ve d ortalama tane boyutudur. Yüksek SC konsantrasyonlarında nükleasyon merkezlerinin

azalması ve tane büyümesinin desteklenmesi, bu ilişki çerçevesinde daha düşük özdirenç değerlerine katkıda bulunmaktadır.

Sıcaklığın Elektriksel Özdirenç Üzerindeki Etkisi

Özdirenç ölçümleri 100 K ile 405 K sıcaklık aralığında gerçekleştirilmiştir. Sıcaklık artışı ile özdirencin sistematik artışı, fonon-elektron saçılımının yoğunlaşmasıyla açıklanmaktadır. Matthiessen kuralı çerçevesinde, toplam özdirenç termal titreşim, safsızlık saçılımı ve tane sınırı katkılarının toplamı olarak ifade edilir (Rossiter, 1987). Elektrokimyasal depolama yöntemiyle üretilen ince filmlerde, Debye sıcaklığının üzerindeki termal rejimlerde ($T > \theta_D$) fonon popülasyonunun lineer artışı, elektronların ortalama serbest yolunu kısaltarak direnci artırmaktadır. Bu bulgu, CuZn alaşımlarında literatürde rapor edilen $2.8-3.5 \times 10^{-8} \Omega \cdot \text{m}/^\circ\text{C}$ sıcaklık katsayısı aralığı ile uyumludur (Callister, 2019).

Sonuç

Bu çalışma, SC konsantrasyonunun CuZn alaşımlarının elektrokimyasal depolanmasında elektriksel ve mikroyapısal özelliklerin belirlenmesinde dominant rol oynayan kritik bir parametre olduğunu nicel olarak ortaya koymuştur. SC konsantrasyonu, hem film kompozisyonunu (%60 varyans açıklama) hem de tane boyutunu kontrol ederek, nihai özdirenç değerlerini sistematik olarak etkilemektedir. Elektrolit içerisindeki sodyum sitratın 0.7-0.9 mol/L konsantrasyon aralığı, düşük özdirenç değerleri elde etmek için optimal işletim bölgesi olarak belirlenmiştir.

Kaynakça

- Assaf, F. H., El-Seidy, A. M. A., Abou-Krishna, M. M., & Eissa, A. A. (2015). Electrodeposition and Characterization of Zn-Ni-Mn Alloy from Sulfate Bath: Influence of Current Density. *International Journal of Electrochemical Science*, 10(7), 5465-5478. [https://doi.org/10.1016/S1452-3981\(23\)17268-3](https://doi.org/10.1016/S1452-3981(23)17268-3)
- Benrazzouq, S.-E., Ghanbaja, J., Migot, S., Milichko, V. A., & Pierson, J.-F. (2025). Chemically-driven control of electrical resistivity of high-entropy alloys. *Applied Materials Today*, 44, 102726. <https://doi.org/10.1016/j.apmt.2025.102726>
- Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 4(29), 552-557. <https://doi.org/10.7334/psicothema2016.383>
- Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilità*. Seeber. Gelış tarihi gönderen <https://books.google.com.tr/books?id=3CY-HQAACAAJ>
- Brenner, A. (1963). *Electrodeposition of alloys: Principles and practice*. New York: Academic Press.
- Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, 69(346), 364-367. <https://doi.org/10.1080/01621459.1974.10482955>
- Callister, W. D. (2019). *Materials Science and Engineering*. New York: Wiley.
- Chandrasekar, M. S., & Pushpavanam, M. (2008). Pulse and pulse reverse plating—Conceptual, advantages and applications. *Electrochimica Acta*, 53(8), 3313-3322. <https://doi.org/10.1016/j.electacta.2007.11.054>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed., reprint). New York, NY: Psychology Press.
- Cultrera, A., Serazio, D., Fabricius, N., & Callegaro, L. (2024). New IEC standards for the measurement of sheet resistance on

- large-area graphene using the van der Pauw and the in-line four-point probe methods. *Measurement*, 236, 114980. <https://doi.org/10.1016/j.measurement.2024.114980>
- Fairbank, H. A. (1944). The Electrical Resistivity of Copper-Zinc and Copper-Tin Alloys at Low Temperatures. *Physical Review*, 66(9-10), 274-281. <https://doi.org/10.1103/PhysRev.66.274>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock "n" roll* (4th edition). Los Angeles London New Delhi Singapore Washington DC: Sage.
- Juškenas, R., Karpavičienė, V., Pakštas, V., Selskis, A., & Kapočius, V. (2007). Electrochemical and XRD studies of Cu–Zn coatings electrodeposited in solution with d-mannitol. *Journal of Electroanalytical Chemistry*, 602(2), 237-244. <https://doi.org/10.1016/j.jelechem.2007.01.004>
- Kruglikov, S. S., Kudriavtsev, N. T., Vorobiova, G. F., & Antonov, A. Y. A. (1965). On the mechanism of levelling by addition agents in electrodeposition of metals. *Electrochimica Acta*, 10(3), 253-261. [https://doi.org/10.1016/0013-4686\(65\)87023-2](https://doi.org/10.1016/0013-4686(65)87023-2)
- Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583-621. <https://doi.org/10.1080/01621459.1952.10483441>
- Leontyev, I. N., Kuriganova, A. B., Allix, M., Rakhmatullin, A., Timoshenko, P. E., Maslova, O. A., ... Smirnova, N. V. (2018). On the Evaluation of the Average Crystalline Size and Surface Area of Platinum Catalyst Nanoparticles. *Physica Status Solidi (b)*, 255(10), 1800240. <https://doi.org/10.1002/pssb.201800240>
- Liu, A., Shi, Z., & Reddy, R. G. (2020). Mechanism study of Cu-Zn alloys electrodeposition in deep eutectic solvents. *Ionics*,

- 26(6), 3161-3172. <https://doi.org/10.1007/s11581-019-03418-2>
- Mohanty, U. S. (2011). Electrodeposition: A versatile and inexpensive tool for the synthesis of nanoparticles, nanorods, nanowires, and nanoclusters of metals. *Journal of Applied Electrochemistry*, 41(3), 257-270. <https://doi.org/10.1007/s10800-010-0234-3>
- Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews*, 82(4), 591-605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x>
- Nordheim, L. (1931). Zur elektronentheorie der Metalle. II. *Annalen der Physik*, 401(6), 641-678.
- Oulmas, C., Mameri, S., Boughrara, D., Kadri, A., Delhalle, J., Mekhalif, Z., & Benfedda, B. (2019). Comparative study of Cu–Zn coatings electrodeposited from sulphate and chloride baths. *Heliyon*, 5(7), e02058. <https://doi.org/10.1016/j.heliyon.2019.e02058>
- Özdemir, R., & Karahan, İ. H. (2019). Effect of solution Zn concentration on electrodeposition of $\text{Cu}_x\text{Zn}_{1-x}$ alloys: Materials and resistivity characterisation. *Transactions of the IMF*, 97(2), 95-99. <https://doi.org/10.1080/00202967.2019.1570738>
- Özdemir, Rasim, Karahan, İ. H., & Karabulut, O. (2016). A Study on the Electrodeposited Cu-Zn Alloy Thin Films. *Metallurgical and Materials Transactions A*, 47(11), 5609-5617. <https://doi.org/10.1007/s11661-016-3715-0>
- Pereloma, E., & Edmonds, D. V. (2012). *Phase transformations in steels*. Woodhead Publishing Limited. <https://doi.org/10.1533/9780857096111>
- Pletcher, D., & Walsh, F. C. (1993). *Industrial Electrochemistry*. Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-011-2154-5>

- Porter, D. A. (2009). *Phase transformations in metals and alloys* (Third edition; K. E. Easterling & M. Y. Sherif, Ed.). Boca Raton, FL: CRC Press.
- Rasim Ozdemir & İsmail Hakki Karahan. (2023). Effect of Sodium Citrate as Complexing Agent on the Electrodeposited CuZn Alloys: Electrochemical, Morphology, Structure, and Electrical Resistivity Studies. *Protection of Metals and Physical Chemistry of Surfaces*, 59(3), 445-452. <https://doi.org/10.1134/S207020512370051X>
- Razali, N. M., Wah, Y. B., & others. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1), 21-33.
- Rossiter, P. L. (1987). *The Electrical Resistivity of Metals and Alloys* (1. bs). Cambridge University Press. <https://doi.org/10.1017/CBO9780511600289>
- Slupska, M., & Ozga, P. (2014). Electrodeposition of Sn-Zn-Cu alloys from citrate solutions. *Electrochimica Acta*, 141, 149-160. <https://doi.org/10.1016/j.electacta.2014.07.039>
- Steinskog, D. J., Tjøstheim, D. B., & Kvamstø, N. G. (2007). A Cautionary Note on the Use of the Kolmogorov–Smirnov Test for Normality. *Monthly Weather Review*, 135(3), 1151-1157. <https://doi.org/10.1175/MWR3326.1>
- Sun, Y., & Zangari, G. (2023). Commentary and Notes on the Original Derivations of the Scharifker-Hills Model. *Journal of The Electrochemical Society*, 170(3), 032503. <https://doi.org/10.1149/1945-7111/acb618>
- To, D. T., Park, S. H., Kim, M. J., Cho, H.-S., & Myung, N. V. (2022). Effect of solution pH, precursor ratio, agitation and temperature on Ni-Mo and Ni-Mo-O electrodeposits from ammonium citrate baths. *Frontiers in Chemistry*, 10, 1010325. <https://doi.org/10.3389/fchem.2022.1010325>

- Tukey, J. W. (1949). Comparing Individual Means in the Analysis of Variance. *Biometrics*, 5(2), 99. <https://doi.org/10.2307/3001913>
- Walsh, F. C., & Ponce De León, C. (2014). Versatile electrochemical coatings and surface layers from aqueous methanesulfonic acid. *Surface and Coatings Technology*, 259, 676-697. <https://doi.org/10.1016/j.surfcoat.2014.10.010>
- Xiao, S., Li, J., Bordas, S. P. A., & Kim, T.-Y. (2023). Artificial neural networks and their applications in computational materials science: A review and a case study. İçinde *Advances in Applied Mechanics* (C. 57, ss. 1-33). Elsevier. <https://doi.org/10.1016/bs.aams.2023.09.001>
- Zimmerman, D. W. (1998). Invalidation of Parametric and Nonparametric Statistical Tests by Concurrent Violation of Two Assumptions. *The Journal of Experimental Education*, 67(1), 55-68. <https://doi.org/10.1080/00220979809598344>

BÖLÜM 7

DOĞAL DİL İŞLEMENİN EVRİMİ: VEKTÖRLERDEN AJANLARA

MURAT AYDOĞAN¹

Giriş

İnsanlık tarihi, büyük ölçüde dilin ve iletişimin tarihi olarak okunabilir. Medeniyetin şafağından bu yana dil, bilgiyi nesilden nesile aktarmamızı, karmaşık sosyal yapılar kurmamızı, soyut düşünceler geliştirmemizi ve kolektif bilinci oluşturmamızı sağlayan en temel aracımız olmuştur. 20. yüzyılın ortalarında bilgisayar bilimlerinin doğuşuyla birlikte, insanlığın en kadim hayallerinden biri de teknolojik bir hedef haline geldi: Bu en insani yeteneği, yani dili, makinelere kazandırmak. İşte bu cesur ve karmaşık hedefin peşinden giden bilimsel disiplin, Doğal Dil İşleme (Natural Language Processing - NLP) olarak adlandırılır. NLP, bilgisayarların insan dilini (Türkçe, İngilizce, Çince gibi doğal dilleri) anlamasını, işlemesini, yorumlamasını ve üretmesini amaçlayan, yapay zekâ ve bilgisayar biliminin kesişiminde yer alan dinamik bir alandır.

¹ Dr, Fırat Üniversitesi, Yazılım Mühendisliği Orcid: 0000-0002-6876-6454

NLP'nin ilk adımları, büyük ölçüde dilbilimciler tarafından elle hazırlanan karmaşık kurallar ve sözlükler üzerine kurulu kural tabanlı sistemlere dayanıyordu. Örneğin, bir makine çevirisi sistemi, bir dildeki her gramer kuralı ve istisnası için diğer dilde bir karşılık bulmaya çalışırdı. Bu yaklaşım, belirli ve dar alanlarda kısmi başarılar elde etse de, dilin doğasındaki esneklik, belirsizlik (ambiguity) ve sürekli evrim karşısında hızla yetersiz kaldı. Bir sonraki aşamada, istatistiksel yöntemler ön plana çıktı. Bu yaklaşımlar, büyük metin veri kümelerinden (corpus) kelimelerin ve ifadelerin bir arada bulunma olasılıklarını öğrenerek, kural tabanlı sistemlerin kırılma noktasını aştı. Ancak bu modeller de dilin derin anlamsal yapısını ve bağlamını kavramakta zorlanıyor, genellikle yüzeysel desenleri ezberlemekle yetiniyordu.

Bu bölüm, NLP'nin bu erken dönem arayışlarından sıyrılarak günümüzün akıllı sistemlerine uzanan modern evrimini, dört temel paradigma değişimi veya "kırılma noktası" üzerinden anlatmayı hedeflemektedir. Bu yolculuk, sadece daha iyi algoritmalar geliştirme hikayesi değil, aynı zamanda makinelerle "anlamı" nasıl öğretebileceğimize dair anlayışımızın da bir evrimidir (Gururaja ve ark., 2023:16):

1. Vektör Temsilleri: İlk büyük sıçrama, kelimelerin anlamsal ilişkilerini yoğun ve çok boyutlu matematiksel bir uzayda temsil etmeyi başaran Word Embedding yöntemleriyle yaşandı. Bu, makinelerle kelimelerin sadece sembolik kimliklerini değil, aynı zamanda anlamsal yakınlıklarını ve aralarındaki soyut ilişkileri de anlama yeteneği kazandırdı (Mikolov & ark., 2013: 9).
2. Paralel İşleme Devrimi: İkinci kırılma, dilin sıralı doğasının getirdiği hesaplama darboğazlarını aşan Transformer mimarisi ile gerçekleşti. Tekrarlayan yapıları terk edip yalnızca dikkat mekanizmasına dayanan bu mimari, metnin tamamını paralel olarak

işleyerek hem uzun mesafeli anlamsal bağlantıları yakalamada devrim yarattı hem de donanım verimliliğini en üst düzeye çıkardı.

3. Ölçeklenmenin Gücü: Üçüncü ve belki de en çok ses getiren atılım, Transformer mimarisinin devasa veri setleriyle ölçeklendirilmesi sonucu ortaya çıkan Büyük Dil Modelleri (LLM) oldu. Bu modeller, belirli görevleri yerine getiren dar kapsamlı araçlar olmaktan çıkıp, genel amaçlı bir dil anlama ve üretme yeteneği sergileyerek yapay zeka ile olan ilişkimizi temelden değiştirdi.
4. Eyleme Geçen Zeka: Son olarak, bu yolculuğun günümüzdeki en ileri noktasını temsil eden Ajanlar (Agents) konseptini ele alacağız. Ajanlar, LLM'leri birer akıl yürütme motoru olarak kullanarak, harici araçlara ve API'lere erişip dijital ve hatta fiziksel dünyayla etkileşime geçen, hedefe yönelik görevleri otonom bir şekilde tamamlayabilen sistemlerdir.

Bu bölüm, bu dört temel direk etrafında bir yolculuğa çıkarak, NLP'nin vektör temsillerinden akıllı ajanlara uzanan evrimini adım adım gözler önüne serecektir. Her bir teknolojik atılımın ardındaki temel fikirleri, getirdiği yenilikleri ve bir sonraki adımı nasıl hazırladığını inceleyerek, günümüz yapay zekasının neden ve nasıl bu noktaya geldiğine dair kapsamlı bir anlayış sunulması amaçlanmaktadır.

Anlamdan Yoksun Vektörler: Klasik Temsil Yöntemleri

Makinelerin dili anlamlandırabilmesi yolundaki ilk ve en temel engel, kelimelerin doğası gereği sembolik olmasıdır. Bilgisayarlar sayılarla çalışır, kelimelerle değil. Bu nedenle, "kedi", "köpek" gibi kelimeleri bir modelin işleyebilmesi için onları sayısal bir formata dönüştürmek gerekir. Derin öğrenme öncesi dönemde bu

sorun, kelime frekanslarına dayalı yöntemlerle çözülmeye çalışılıyordu.

- Bag-of-Words (BoW - Kelime Torbası): Bu yaklaşımda, bir metin (cümle veya belge), içerdiği kelimelerin bir torbası olarak kabul edilir ve kelimelerin sırası tamamen göz ardı edilir. Her belge, sözlükteki her kelimenin o belgede kaç kez geçtiğini sayan bir vektörle temsil edilir. Bu basit ve hızlı bir yöntem olsa da, kelime sırasını yok saydığı için "Ali Ayşe'yi seviyor" ile "Ayşe Ali'yi seviyor" cümlelerini ayırt edemez (Juluru & ark., 2021: 6) .
- TF-IDF (Term Frequency-Inverse Document Frequency): BoW modelini bir adım ileri taşıyan bu yöntem, bir kelimenin bir belge için önemini hesaplamaya çalışır. Bir kelimenin bir belgedeki frekansı (TF) ne kadar yüksekse ve o kelimenin tüm belgeler koleksiyonundaki (corpus) yaygınlığı (IDF) ne kadar azsa, o kelime o belge için o kadar önemlidir (Zhang, Yoshida & Tang, 2011: 6). Bu, "ve", "ama" gibi sık geçen ama anlamsal yükü az olan kelimelerin etkisini azaltır.
- One-Hot Encoding: Özellikle sinir ağlarının ilk dönemlerinde kullanılan bu yöntemde, sözlükteki her kelimeye benzersiz bir tam sayı (indeks) atanır. Daha sonra her kelime, boyutu sözlükteki toplam kelime sayısına eşit olan bir vektörle temsil edilir. Bu vektörde, kelimenin kendi indeksine karşılık gelen eleman 1, diğer tüm elemanlar ise 0 olur (Patil & ark., 2023: 26).

Bu yöntemler, metin sınıflandırma gibi görevlerde belirli bir başarıya ulaşmış olsalar da, temel bir kusuru paylaşıyorlardı: anlamsal bilgi eksikliği. Bu yaklaşımlarla üretilen vektörler, kelimeler arasındaki anlamsal ilişki hakkında hiçbir ipucu taşımaz.

Örneğin, "kedi" ve "köpek" kelimelerinin anlamsal olarak "uçak" kelimesinden daha yakın olduğunu bu temsil ile anlamak imkansızdır. Vektörler son derece seyrek (sparse), yüksek boyutlu ve ortogonaldır (birbirine dik), bu da aralarında bir benzerlik ölçümü yapmayı anlamsız kılar.

Anlamanın Vektörlere Gömülmesi: Word2Vec

İşte bu noktada Word Embedding (Kelime Gömme) kavramı bir devrim yaratmıştır. Word Embedding, kelimeleri seyrek ve anlamsız vektörler yerine, yoğun (dense) ve genellikle daha düşük boyutlu (örneğin 50-300 boyut) sürekli bir vektör uzayında temsil etme fikridir. Bu yaklaşımın temelinde, dilbilimci J.R. Firth'e atfedilen ve daha sonra popülerleşen dağılımsal hipotez (distributional hypothesis) yatar: "Bir kelimeyi, çevresindeki kelimelerle tanıyabilirsin" (you shall know a word by the company it keeps). Bu hipoteze göre, benzer bağlamlarda kullanılan kelimeler, benzer anlamlara sahip olma eğilimindedir. Word Embedding algoritmaları, bu hipotezi matematiksel bir modele dönüştürerek, büyük metin veri kümelerinden kelimelerin anlamsal temsillerini (vektörlerini) öğrenir.

Bu alandaki en popüler ve etkili uygulamalardan biri, 2013 yılında Google'dan Tomas Mikolov ve ekibi tarafından geliştirilen Word2Vec modelidir (Mikolov & ark., 2013: 9). Word2Vec, aslında tek bir algoritma değil, kelime vektörlerini verimli bir şekilde öğrenmek için tasarlanmış iki farklı model mimarisini ve çeşitli optimizasyon tekniklerini içeren bir araç setidir.

Word2Vec Mimarileri: CBOW ve Skip-gram

Word2Vec, dağılımsal hipotezi iki ana model mimarisıyla uygular:

- Continuous Bag-of-Words (CBOW): Bu model, bir kelimeyi çevreleyen bağlam kelimelerinden (context

words) yola çıkarak ortadaki hedef kelimeyi tahmin etmeye çalışır. Örneğin, "hızlı kahverengi ____ çitin üzerinden atladı" cümlesinde, CBOW modeli [hızlı, kahverengi, çitin, üzerinden, atladı] bağlam kelimelerini girdi olarak alır ve ortadaki "tilki" kelimesini tahmin etmeyi öğrenir. Bağlam kelimelerinin vektörlerini birleştirerek (genellikle toplayarak veya ortalamasını alarak) hedef kelimeyi tahmin ettiği için bu modele "sürekli kelime torbası" denir. CBOW, daha hızlı eğitilir ve sık geçen kelimeler için iyi sonuçlar verir.

- Skip-gram: Bu model ise tam tersi bir mantıkla çalışır. Ortadaki bir hedef kelimedenden yola çıkarak onun çevresindeki bağlam kelimelerini tahmin etmeye çalışır. Aynı örnekte, Skip-gram modeli girdi olarak "tilki" kelimesini alır ve çıktı olarak [hızlı, kahverengi, çitin, üzerinden, atladı] kelimelerini tahmin etmeyi öğrenir. Yani, bir kelimedenden yola çıkarak onun olası bağlamlarını tahmin eder. Skip-gram, genellikle daha büyük veri setlerinde daha iyi performans gösterir ve nadir kelimelerin temsillerini öğrenmede CBOW'a göre daha başarılıdır.

Her iki model de aslında basit birer sıg sinir ağıdır (shallow neural network). Genellikle bir girdi katmanı, bir gizli katman (projeksiyon katmanı) ve bir çıktı katmanından oluşurlar. Modelin asıl amacı, bu tahmin görevini (hedef kelimeyi veya bağlam kelimelerini tahmin etme) başarıyla yerine getirmek değildir. Tahmin görevi, sadece bir araçtır. Asıl hedef, bu eğitim süreci tamamlandığında, gizli katmanda oluşan ağırlık matrisini elde etmektir. İşte bu matris, sözlükteki her kelime için öğrenilmiş olan yoğun ve anlamsal vektörleri içerir.

Anlamsal Aritmetik ve Etkisi

Word2Vec'in getirdiği en büyüleyici yeniliklerden biri, bu vektör uzayında anlamlı aritmetik işlemlerin yapılabilmesidir. En meşhur örnek, $\text{Vektör('Kral')} - \text{Vektör('Erkek')} + \text{Vektör('Kadın')}$ işleminin sonucunun, Vektör('Kraliçe') 'ye çok yakın bir vektör vermesidir. Benzer şekilde, $\text{Vektör('Ankara')} - \text{Vektör('Türkiye')} + \text{Vektör('Fransa')}$ işlemi de Vektör('Paris') 'e yakın bir sonuç üretir. Bu durum, modelin sadece kelimelerin benzerliğini değil, aynı zamanda aralarındaki "cinsiyet" veya "başkent-ülke" gibi soyut ilişkileri de öğrendiğini ve bu ilişkilerin vektör uzayında tutarlı yönler olarak kodlandığını göstermiştir. Bu yetenek, NLP uygulamalarının anlamsal çıkarım yapma kapasitesini kökten değiştirmiş ve daha sofistike modellerin geliştirilmesi için sağlam bir temel oluşturmuştur.

Dizisel Veri ve Bellek Sorunları: RNN ve LSTM'in Yükselişi

Word Embedding, kelimelere anlamsal derinlik kazandırarak büyük bir atılım sağlasa da, dilin doğasındaki en temel özelliklerden birini, yani diziselliği (sequentiality) tek başına ele alamıyordu. Cümleler, birbirinden bağımsız kelimelerin bir torbası (bag-of-words) değildir; kelimelerin sırası, cümlenin anlamını tamamen değiştirebilir. "Ali, Ayşe'yi seviyor" ve "Ayşe, Ali'yi seviyor" cümleleri aynı kelimeleri içermesine rağmen tamamen farklı anlamlara gelir. Bu dizisel bağımlılığı, yani bir kelimenin anlamının kendisinden önce gelen kelimelere bağlı olduğunu modelleyebilmek için yeni bir mimariye ihtiyaç vardı.

Bu ihtiyaca cevap veren mimari Tekrarlayan Sinir Ağları (Recurrent Neural Networks - RNN) oldu (Sherstinsky & ark., 2020: 43). Geleneksel ileri beslemeli sinir ağları, girdileri birbirinden bağımsız olarak işler. Oysa RNN'ler, bir döngü yapısı içerir. Bu döngü, ağın bir zaman adımındaki (örneğin, cümlenin bir kelimesi) çıktısının, bir sonraki zaman adımına (bir sonraki kelime) girdi

olarak aktarılmasını sağlar. Bu mekanizma, ağıın gemiřte grdėđđ bilgileri bir tr "bellek" iinde tutarak mevcut girdiyi iřlemesine olanak tanır. Bu bellek, gizli durum (hidden state) olarak adlandırılan bir vektrle temsil edilir. Bir cmlenin kelimelerini sırayla iřleyen bir RNN, her kelimeyi iřlerken sadece o kelimenin vektrn deėil, aynı zamanda bir nceki kelimededen gelen gizli durumu da dikkate alır. Yeni gizli durum, bu iki bilginin bir fonksiyonu olarak hesaplanır. Bu tekrarlayan yapı, RNN'lerin kelimeler arasındaki sıralı baėımlılıkları ėrenmesini saėladı ve makine evirisi, metin retimi, duygu analizi gibi grevlerde nemli bařarılar elde etmesinin yolunu atı.

RNN'lerin Kırılğan Belleėi: Gradyan Kaybolması ve Patlaması

Teoride mkemmel grnse de, pratikte standart RNN'ler ciddi bir sorunla karřılařtı: uzun vadeli baėımlılıkları (long-term dependencies) ėrenememe. Bir cmlenin bařındaki bir kelime, cmlenin sonundaki bir kelimenin anlamını etkileyebilir. rneėin, "Trkiye'nin birok řehrini gezdim, ancak en ok ____'ı sevdim" cmlesinde, bořluėu doldurmak iin cmlenin bařına dnp "Trkiye" kelimesine referans vermek gerekir. RNN'ler, bu tr uzun mesafeli baėlantıları kurmakta zorlanır. Bunun temel nedeni, aėın eėitiminde kullanılan Zaman İinde Geri Yayılım (Backpropagation Through Time - BPTT) algoritmasının doėasından kaynaklanan kaybolan ve patlayan gradyan (vanishing and exploding gradients) problemidir. Eėitim sırasında, hata sinyalleri (gradyanlar) aėın katmanları boyunca geriye doėru yayılır. Uzun dizilerde, bu gradyanlar tekrar tekrar aynı matrisle arpıldıėı iin ya stel olarak klerek sıfıra yaklařır (kaybolma) ya da stel olarak byyerek kontrolden ıkar (patlama). Kaybolan gradyan, aėın gemiřteki bilgilerden ėrenmesini engellerken, patlayan gradyan ise eėitimi tamamen istikrarsızlařtırır.

Akıllı Kapılar: Uzun Kısa Süreli Bellek (LSTM)

Bu bellek sorununu çözmek için 1997 yılında Sepp Hochreiter ve Jürgen Schmidhuber tarafından Uzun Kısa Süreli Bellek (Long Short-Term Memory - LSTM) ağı önerildi (Hochreiter & ark., 1997:43). LSTM, standart RNN'in basit tekrarlayan hücrelerini, daha karmaşık ve akıllı bir yapıyla değiştirir. LSTM hücresinin kalbinde, bilgiyi uzun süre boyunca saklamak için tasarlanmış bir hücre durumu (cell state) bulunur. Bu hücre durumu, ağ boyunca bir "konveyör bandı" gibi uzanır ve bilgiler üzerinde çok az lineer etkileşimle akar. Bu sayede, gradyanların kaybolmadan uzun mesafeler kat etmesi kolaylaşır.

LSTM'in asıl sihri, bu hücre durumuna hangi bilgilerin ekleneceğini, hangilerinin çıkarılacağını ve hangilerinin çıktı olarak kullanılacağını düzenleyen üç adet kapı (gate) mekanizmasıdır. Bu kapılar, sigmoid aktivasyon fonksiyonu sayesinde 0 (hiçbir şey geçirme) ile 1 (her şeyi geçir) arasında değerler üreten küçük sinir ağlarıdır ve şu sorulara cevap verirler:

- Unutma Kapısı (Forget Gate): Geçmiş hücre durumundan hangi bilgileri atmamız? (Örneğin, yeni bir cümle özneleri geldiğinde, eski cümle öznelerini unutmak gerekebilir.)
- Giriş Kapısı (Input Gate): Mevcut girdiden hangi yeni bilgileri hücre durumuna eklemeliyiz?
- Çıktı Kapısı (Output Gate): Hücre durumunun hangi kısmını bu zaman adımının çıktısı olarak kullanmalıyız?

Bu kapı mekanizmaları, LSTM'in hangi bilgiyi ne kadar süreyle "hatırlayacağını" dinamik olarak öğrenmesini sağlar. Bu sayede, hem kısa vadeli hem de uzun vadeli bağımlılıkları etkili bir şekilde modelleyebilir. LSTM ve onun daha basit bir varyasyonu

olan GRU (Gated Recurrent Unit), uzun yıllar boyunca NLP alanındaki en başarılı ve yaygın olarak kullanılan sıralı modeller olmuş ve Transformer mimarisine giden yolda önemli bir basamak teşkil etmiştir (Chung & ark., 2014:9).

Devrimin Merkezi: Transformer Mimarisi ve Dikkat Mekanizması

2017 yılında Google araştırmacıları tarafından yayınlanan "Attention Is All You Need" başlıklı makale (Vaswani & ark., 2017:15), NLP alanında bir deprem etkisi yarattı. Bu makalede tanıtılan Transformer mimarisi, dil modellemesinde sıralı işlemeye dayalı RNN ve LSTM'lerin hakimiyetine son verdi. Transformer, tekrarlayan (recurrent) katmanları tamamen ortadan kaldırıp, bunun yerine metindeki her kelimenin diğer tüm kelimelerle olan ilişkisini doğrudan modelleyen dikkat mekanizmasına (attention mechanism) dayalı bir yapı sunuyordu. Bu, modelin hem uzun mesafeli bağımlılıkları çok daha iyi yakalamasını sağladı hem de RNN'lerin doğası gereği engel olduğu paralelleştirmeye olanak tanıyarak çok daha büyük modellerin ve veri setlerinin eğitilmesinin önünü açtı.

Transformer mimarisi, temel olarak iki ana bileşenden oluşur: bir Encoder (Kodlayıcı) ve bir Decoder (Kod Çözücü). Her ikisi de kendi içlerinde birden çok (orijinal makalede 6 adet) özdeş katmandan oluşan bir yığındır (stack).

- Encoder: Girdi cümlesindeki her kelime için zengin ve bağlamsal bir temsil (embedding) oluşturmaktan sorumludur. Cümlenin tamamını okur ve her kelimenin anlamını, cümlenin geri kalanıyla olan ilişkisine göre ayarlar.
- Decoder: Encoder'dan gelen bu temsilleri ve daha önce ürettiği kelimeleri kullanarak, çıktı cümlesini adım adım (otoregresif olarak) üretir.

Girdi İşleme: Embedding ve Konumsal Kodlama

Modelin bir metni işleyebilmesi için, kelimeler önce Word Embedding katmanından geçirilerek yoğun vektörlere dönüştürülür. Ancak Transformer, tekrarlayan bir yapıya sahip olmadığı için kelimelerin sırasını doğal olarak bilemez. Bu sorunu çözmek için, kelime vektörlerine, kelimenin cümle içindeki konumunu belirten bir Konumsal Kodlama (Positional Encoding) vektörü eklenir. Bu vektörler, sinüs ve kosinüs fonksiyonları kullanılarak deterministik bir şekilde üretilir.

Scaled Dot-Product Attention: Q, K, V Vektörleri

Transformer'ın kalbi, Scaled Dot-Product Attention mekanizmasıdır. Bu mekanizma, bir cümlenin işlenmesi sırasında, modelin hangi kelimelere daha fazla "dikkat etmesi" gerektiğini öğrenmesini sağlar. Her kelime vektörü için üç farklı vektör türetilir:

- Query (Sorgu - Q): Mevcut kelimenin, cümlenin diğer kelimelerinden bilgi arayan temsilidir. "Ben neyim ve ne arıyorum?" sorusunu sorar.
- Key (Anahtar - K): Cümlenin diğer kelimelerinin, sorguyla eşleşmek için sunduğu "etiket" veya "anahtar" temsilidir. "Ben ne tür bir bilgi sunuyorum?" sorusunu temsil eder.
- Value (Değer - V): Cümlenin diğer kelimelerinin asıl içerik temsilidir. "Eğer benimle eşleşirsen, sana bu bilgiyi vereceğim" der.

Dikkat mekanizması şu adımlarla çalışır:

- Skor Hesaplama: Her bir sorgu (Q) vektörü, diğer tüm anahtar (K) vektörleriyle noktasal çarpım (dot product)

işlemine tabi tutulur. Bu, sorgunun her bir anahtarla ne kadar uyumlu olduğunu ölçen bir skor üretir.

- Ölçeklendirme (Scaling): Gradyanların çok büyümesini veya küçülmesini önlemek için bu skorlar, anahtar/sorgu vektörlerinin boyutunun (d_k) kareköküne bölünür.
- Softmax: Ölçeklenmiş skorlar bir Softmax fonksiyonundan geçirilir. Bu, tüm skorları toplamı 1 olan pozitif olasılık değerlerine (dikkat ağırlıkları) dönüştürür. Yüksek bir skor, o kelimeye daha fazla dikkat edileceği anlamına gelir.
- Çıktı: Değer (V) vektörleri, bu dikkat ağırlıklarıyla çarpılarak toplanır. Bu, dikkat mekanizmasının çıktısını oluşturur; bu çıktı, orijinal kelimenin, cümlemin geri kalanından gelen bağlamsal bilgilerle zenginleştirilmiş yeni bir temsildir.

Multi-Head Attention: Çoklu Perspektifler

Tek bir dikkat mekanizması, bir kelimenin diğerleriyle olan ilişkisini tek bir açıdan değerlendirir. Ancak bir kelimenin farklı kelimelerle farklı türde ilişkileri olabilir (örneğin, anlamsal, sözdizimsel vb.). Multi-Head Attention (Çok Başlıklı Dikkat), bu sorunu çözmek için Q, K ve V vektörlerini birden çok (orijinal makalede 8) farklı "temsili alt uzayına" lineer olarak yansıtarak dikkat mekanizmasını paralel olarak birden çok kez çalıştırır. Her bir "baş" (head), farklı bir ilişki türünü öğrenme fırsatı bulur. Son olarak, bu paralel başların çıktıları birleştirilir ve tekrar bir lineer projeksiyondan geçirilerek nihai çıktı oluşturulur. Bu, modelin aynı anda farklı türdeki bağlamsal ilişkileri yakalamasına olanak tanır.

Diğer Katmanlar: Add & Norm ve Feed-Forward

Her bir Transformer katmanı, Multi-Head Attention mekanizmasının yanı sıra iki önemli alt katman daha içerir:

- **Add & Norm (Ekle ve Normalleştir):** Her bir alt katmanın (Multi-Head Attention ve Feed-Forward) çıktısı, alt katmanın girdisiyle toplanır. Bu, residual connection (artık bağlantı) olarak bilinir ve ağı daha derin katmanlara gradyan akışını kolaylaştırarak çok daha derin modellerin eğitilmesini sağlar. Bu toplama işleminin ardından, katman çıktısını stabilize etmek için bir Layer Normalization (Katman Normalizasyonu) adımı uygulanır.
- **Position-wise Feed-Forward Network (FFN):** Her pozisyondaki kelimenin temsiline bağımsız olarak uygulanan, iki lineer dönüşüm ve aralarında bir ReLU aktivasyon fonksiyonu içeren basit bir ileri beslemeli sinir ağıdır.

Decoder Yapısı ve Maskeleme

Decoder yapısı, Encoder'a çok benzer, ancak iki temel farkı vardır:

- **Masked Multi-Head Attention:** Decoder, bir sonraki kelimeyi tahmin ederken gelecek kelimeleri görmemelidir. Bu nedenle, dikkat mekanizmasının Softmax adımından önce, gelecek pozisyonlara karşılık gelen skorlar eksisi sonsuz ($-\infty$) ile maskelenir. Bu, modelin sadece geçmiş kelimelere dikkat etmesini sağlar.
- **Encoder-Decoder Attention:** Decoder'daki ikinci dikkat katmanı, Encoder yığınının çıktısını (K ve V vektörleri) ve bir önceki Masked Multi-Head Attention katmanının çıktısını (Q vektörleri) alır. Bu, Decoder'ın çıktı cümlesini üretirken girdi cümlesinin ilgili kısımlarına odaklanmasını sağlar.

Bu karmaşık ama son derece etkili yapı, Transformer'ı NLP'de bir devrimin merkezi haline getirmiş ve Büyük Dil Modelleri çağına giden yolu açmıştır.

Büyük Dil Modellerinin (LLM) Doğuşu: Ölçeklenmenin Gücü

Büyük Dil Modelleri (LLM) çağı, temel olarak Transformer mimarisinin farklı felsefelerle uygulandığı iki ana akım etrafında şekillendi (Brown & ark.,2020:24):

- GPT (Generative Pre-trained Transformer) Serisi: OpenAI tarafından geliştirilen bu modeller, Transformer'ın yalnızca Decoder kısmını kullanır. Temel görevleri, bir metin dizisi verildiğinde bir sonraki kelimeyi tahmin etmektir ($P(\text{kelime}_n \mid \text{kelime}_1, \dots, \text{kelime}_{\{n-1\}})$). Bu otoregresif yapı, onları metin üretme, özetleme, çeviri gibi üretken görevlerde doğal olarak çok başarılı kılar. GPT modelleri, tek yönlü (unidirectional) bir dikkat mekanizması kullanır; yani bir kelimeyi tahmin ederken sadece ondan önce gelen kelimelere dikkat edebilir (Radford & ark, 2018:12).
- BERT (Bidirectional Encoder Representations from Transformers) (Devlin & ark, 2019:15) ve Türevleri: Google tarafından geliştirilen BERT, Transformer'ın yalnızca Encoder kısmını kullanır. GPT'nin aksine, BERT çift yönlü (bidirectional) bir yapıya sahiptir; yani bir kelimenin temsilini oluştururken hem ondan önce hem de sonra gelen kelimeleri dikkate alır. Bu, modelin derin bir bağlamsal anlayış geliştirmesini sağlar. BERT, iki zekice görev üzerinde ön-eğitilmiştir:
- Maskeli Dil Modeli (Masked Language Model - MLM): Girdi cümlesindeki kelimelerin yaklaşık %15'i rastgele olarak özel bir [MASK] tokeni ile değiştirilir ve modelin

bu maskelenmiş kelimeleri cümlelerin geri kalanına bakarak tahmin etmesi istenir.

- Sonraki Cümle Tahmini (Next Sentence Prediction - NSP): Model, iki cümle alır ve ikinci cümlelerin, ilk cümlelerin metin içinde gerçekten devamı olup olmadığını tahmin eder. Bu, modelin cümleler arası ilişkileri öğrenmesine yardımcı olur.

Bu çift yönlü doğası, BERT'i metin sınıflandırma, soru-cevap, duygu analizi gibi anlama (understanding) odaklı görevlerde son derece güçlü kılmıştır.

Yeni Paradigma: Ön-Eğitim, İnce Ayar ve Komut Verme

LLM'ler, temel modeller (foundation models) olarak bilinen yeni bir konsepti popülerleştirmiştir. Bu modeller, internetin devasa bir kesitinden (trilyonlarca kelime) gelen metin verileri üzerinde, genellikle yüzlerce veya binlerce GPU kullanılarak haftalar veya aylar süren bir ön-eğitim (pre-training) sürecinden geçirilir. Bu süreç, modele dilin gramerini, sözdizimini, anlamsal ilişkilerini ve dünya hakkında geniş bir olgusal bilgiyi öğretir.

Bu devasa ön-eğitilmiş model daha sonra, çok daha az veri ve hesaplama gücü ile belirli görevlere uyarlanabilir. Bu adaptasyonun iki ana yolu vardır:

- İnce Ayar (Fine-Tuning): Model, daha küçük ve belirli bir göreve özel (örneğin, tıbbi rapor analizi, hukuki metin sınıflandırma vb.) etiketli bir veri seti üzerinde kısa bir süre daha eğitilir. Bu süreçte, modelin ağırlıkları yeni göreve daha iyi uyum sağlayacak şekilde hafifçe güncellenir. Bu, ön-eğitime kıyasla çok daha az maliyetlidir.
- Komut Verme (Prompting): Özellikle GPT-3 gibi çok büyük modellerin ortaya çıkmasıyla, modellerin hiçbir

ağırlık güncellemesi yapmadan, sadece doğru bir doğal dil komutu (prompt) verilerek yeni görevleri yerine getirebildiği görüldü. Bu yaklaşıma bağlam-ıçi öğrenme (in-context learning) denir. Komutta hiç örnek verilmiyorsa "zero-shot", birkaç örnek veriliyorsa "few-shot" öğrenme olarak adlandırılır. Bu yöntem, model adaptasyonunu daha da kolaylaştırarak güçlü NLP yeteneklerine erişimi demokratikleştirmiş ve yapay zeka alanında yeni bir inovasyon dalgası başlatmıştır.

Akıllı Yardımcılar ve Otonom Sistemler: Ajanlar (Agents)

Büyük Dil Modelleri (LLM), metin anlama ve üretme konusunda insanüstü bir yetkinliğe ulaşmış olsalar da, temel bir kısıtlılığa sahiptirler: pasif ve kapalı bir sistem olmaları. Bir LLM, yalnızca eğitim verisinde bulunan bilgiyle sınırlıdır (bu durum knowledge cutoff olarak bilinir) ve dış dünya ile anlık, dinamik bir etkileşime geçemez. Bir soruyu yanıtlarken en güncel bilgilere erişemez, bir e-posta gönderemez, bir hesaplama yapamaz veya bir web sitesinden rezervasyon yapamaz. Bu modeller, ne kadar bilgili olurlarsa olsunlar, birer "bilgi kaynağı" veya "metin işlemcisi" olmaktan öteye geçemezler; birer "eylemci" değildirler. NLP evriminin günümüzdeki en heyecan verici ve en hızlı gelişen aşaması, bu durumu değiştirmeyi amaçlayan Ajanlar (Agents) konseptidir (Yao,2022:33).

Ajan Kavramı: Akıl Yürüten ve Araç Kullanan Sistemler

Bir yapay zeka ajanı, belirli bir hedefi gerçekleştirmek için çevresini algılayan, kararlar alan ve eylemlerde bulunan otonom bir varlıktır. LLM tabanlı ajanlar, bu konsepti bir üst seviyeye taşır. Bu ajanlar, LLM'i bir akıl yürütme motoru (reasoning engine) veya "beyin" olarak kullanarak, karmaşık görevleri yerine getirmek için bir dizi harici araç (tool) kullanır. Bu yaklaşım, LLM'in dil anlama, planlama ve problem çözme yeteneklerini, gerçek dünya ile

etkileşime geçebilen somut eylemlerle birleştirir. Bu araçlar çok çeşitli olabilir:

- Bilgi Araçları: Web arama motorları (Google, Bing), Wikipedia API'leri.
- Hesaplama Araçları: Kod yorumlayıcıları (Python, R), hesap makineleri.
- Yazılım Kontrol Araçları: Web tarayıcıları, e-posta istemcileri, takvim uygulamaları.
- Fiziksel Dünya Araçları: Robotik kollar, akıllı ev cihazları, dronlar.

Ajan Döngüsü: ReAct ve Ötesi

Bir ajanın temel çalışma döngüsü, genellikle Düşünce-Eylem-Gözlem (Thought-Action-Observation) prensibine dayanır. Bu alandaki en popüler ve temel çerçevelerden biri olan ReAct (Reason + Act) , bu döngüyü şu şekilde yapılandırır (Yao,2022:33):

- Düşünce (Thought): Ajan, kendisine verilen nihai hedefi (örneğin, "En yakın ve en yüksek puanlı İtalyan restoranından bu akşam için iki kişilik yer ayırt") ve o anki durumu analiz eder. LLM, hedefe ulaşmak için bir plan oluşturur ve atması gereken bir sonraki adımı belirler. Örneğin, "Önce konumumu bulmalı, sonra yakındaki İtalyan restoranlarını aramalı, puanlarını karşılaştırmalı ve en iyisi için rezervasyon yapmalıyım. İlk adım: Konumumu bulmak için bir araç kullanmalıyım." diye düşünür.
- Eylem (Action): Ajan, düşündüğü adımı gerçekleştirmek için uygun bir aracı seçer ve gerekli parametrelerle çağırır. Örneğin, `get_current_location()` gibi bir eylemde bulunur.

- Gözlem (Observation): Araç çalıştırılır ve sonucu (örneğin, "Konum: 41.015, 28.979") ajana bir gözlem olarak geri döndürülür. Bu gözlem, bir web aramasının sonucu, bir API'nin yanıtı veya bir kodun çıktısı olabilir.

Ajan, bu yeni gözlem bilgisini alarak döngünün başına döner. Yeni bir "düşünce" adımıyla durumu tekrar değerlendirir ("Konumumu öğrendim. Şimdi bu konuma yakın İtalyan restoranlarını aramak için Google Haritalar aracını kullanmalıyım.") ve bu döngüyü, nihai hedefe ulaşana kadar tekrar tekrar çalıştırır. Bu basit döngü, ajanlara daha önce mümkün olmayan yetenekler kazandırır:

- Bilgi Sınırlılığını Aşma: Web arama araçlarını kullanarak en güncel bilgilere erişebilirler.
- Fiziksel ve Dijital Eylemler: API'ler aracılığıyla e-posta gönderebilir, takvime etkinlik ekleyebilir, online alışveriş yapabilir veya akıllı ev cihazlarını kontrol edebilirler.
- Karmaşık Problem Çözme: Bir görevi alt adımlara ayırabilir, her adım için farklı araçlar kullanabilir ve sonuçları birleştirerek nihai çözüme ulaşabilirler.

Zorluklar ve Gelecek Yönelimler

Ancak ajan teknolojisi henüz başlangıç aşamasındadır. Güvenilirlik, uzun vadeli planlama yapma, hata yönetimi ve güvenlik gibi konularda aşılması gereken önemli zorluklar bulunmaktadır. Bir ajanın yanlış bir aracı kullanması, bir API'yi hatalı çağırması veya bir döngüye girerek kaynakları tüketmesi olası risklerdir. Ayrıca, bir ajana finansal işlemler yapma veya kişisel verilere erişme yetkisi vermenin getireceği güvenlik ve gizlilik sorunları da dikkatle ele alınmalıdır.

Buna rağmen ajanlar, yapay zekanın pasif bilgi sağlayıcılardan, hedeflerimizi anlayan ve bizim adımıza görevleri

yerine getiren proaktif ortaklara dönüşme potansiyelini temsil etmektedir. Bu, NLP'nin sadece dili anlamakla kalmayıp, dil aracılığıyla dünyayı etkileme hedefine ulaştığı yeni bir evrimsel sıçramadır.

Evrimin Özeti: Vektörlerden Ajanlara

Bu bölüm boyunca, Doğal Dil İşleme'nin (NLP) kelimelerin basit sayısal temsillerinden, karmaşık görevleri otonom olarak yerine getirebilen akıllı ajanlara uzanan evrimsel yolculuğunu inceledik. Bu serüven, birbiri üzerine inşa edilen ve her biri bir öncekinin sınırlarını aşan teknolojik kırılımlarla doludur:

- Word Embedding ile kelimelere anlamsal derinlik kazandırdık ve onları matematiksel bir uzayda temsil edilebilir kıldık.
- RNN ve LSTM ile dilin sıralı doğasını modellemeyi öğrendik, ancak uzun vadeli bellek ve paralelleştirme sorunlarıyla karşılaştık.
- Transformer mimarisi, dikkat mekanizması sayesinde paralel işleme devrimi yaratarak uzun vadeli bağımlılık sorununu çözdü ve ölçeklenmenin kapılarını araladı.
- Bu ölçeklenme, internetin kolektif bilgisini özümseyen Büyük Dil Modelleri'nin (LLM) doğuşuna yol açtı ve "ön-eğitim ve adaptasyon" paradigmasını standart hale getirdi.
- Son olarak, Ajanlar konsepti, bu modelleri pasif bilgi kaynaklarından, akıl yürüten, plan yapan ve araçlar kullanarak dünyayla etkileşime geçen aktif eylemcilere dönüştürmeye başladı.

Geldiğimiz noktada, yapay zeka artık sadece metinleri anlamakla kalmıyor, aynı zamanda bu anlayışı kullanarak planlama

yapabiliyor, araçlar kullanabiliyor ve hedeflere ulaşabiliyor. Bu, insan-bilgisayar etkileşiminin doğasını temelden değiştirme potansiyeline sahip bir paradigma değişimidir.

Gelecek Perspektifi: Ufuktaki Yeni Sıçramalar

Ancak bu yolculuk henüz bitmedi. NLP ve yapay zeka ajanlarının geleceği, mevcut yeteneklerin daha da ötesine uzanan heyecan verici araştırma alanlarına işaret etmektedir. Ufukta beliren bir sonraki büyük sıçramalar şunlardır:

- **Multimodalite ve Embodiment:** Mevcut ajanlar büyük ölçüde metin tabanlı çalışmaktadır. Geleceğin ajanları, metnin yanı sıra görüntüleri, sesleri, videoları ve diğer sensör verilerini de bütünsel olarak anlayıp işleyebilen multimodal sistemler olacaktır. Bir kullanıcının gösterdiği bir resimdeki nesneyi anlayan, o nesne hakkında internette araştırma yapan ve bulgularını sesli olarak özetleyen bir ajan hayal edebiliriz. Bunun bir sonraki adımı ise bedenlenmiş yapay zeka (embodied AI), yani bu multimodal anlayışı fiziksel bir bedende (robot) birleştirmektir. LLM'leri birer beyin olarak kullanan robotlar, dilsel komutları anlayarak fiziksel görevleri yerine getirebileceklerdir. "Mutfaktan bir elma getir" gibi bir komutu anlayan, mutfağın nerede olduğunu bilen, elmayı diğer meyvelerden ayırt edebilen ve onu güvenli bir şekilde getiren bir robot, artık bilim kurgu olmaktan çıkıp ulaşılabilir bir hedef haline gelmektedir.
- **Gelişmiş Akıl Yürütme ve Dünya Modelleri:** Günümüz ajanları, basit ve kısa vadeli görevlerde başarılı olsalar da, karmaşık, çok adımlı ve uzun vadeli planlama gerektiren hedeflerde zorlanabilmektedir. Gelecekteki araştırmalar, ajanların daha sağlam akıl yürütme, soyut düşünme, nedensellik kurma ve karşılaştıkları beklenmedik

durumlara adapte olma yeteneklerini geliştirmeye odaklanacaktır. Bu, LLM'lerin sadece metinsel kalıpları öğrenmekle kalmayıp, dünyanın nasıl işlediğine dair daha tutarlı ve sağlam içsel dünya modelleri (internal world models) oluşturmalarını gerektirecektir. Bu, daha güvenilir ve otonom sistemlerin önünü açacaktır.

- Sosyal ve İşbirlikçi Ajanlar: Geleceğin yapay zeka sistemleri, tekil görevleri yerine getiren izole varlıklar olmaktan çıkıp, hem insanlarla hem de diğer ajanlarla karmaşık sosyal ortamlarda işbirliği yapabilen varlıklar haline gelecektir. Çoklu ajan sistemleri (multi-agent systems), karmaşık bir problemi çözmek için birlikte çalışan, iletişim kuran, müzakere eden ve ortak bir hedefe ulaşan uzmanlaşmış ajanlardan oluşabilir. Bu, bilimsel keşiflerden karmaşık lojistik operasyonlarına kadar birçok alanda devrim yaratma potansiyeline sahiptir.

Sonuç olarak, NLP'nin vektörlerden ajanlara uzanan evrimi, yapay zekanın sadece dünyayı anlamakla kalmayıp, aynı zamanda onunla anlamlı ve hedefe yönelik bir şekilde etkileşime girme arayışının bir yansımasıdır. Önümüzdeki yıllar, bu etkileşimin daha da derinleştiği, daha yetenekli, daha otonom ve hayatımızın her alanına entegre olmuş yapay zeka sistemlerine tanıklık edecektir.

Kaynakça

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).

Gururaja, S., Bertsch, A., Na, C., Widder, D., & Strubell, E. (2023). To Build Our Future, We Must Know Our Past: Contextualizing Paradigm Shifts in Natural Language Processing. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 13310-13325. doi:10.18653/v1/2023.emnlp-main.822

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.

Juluru, K., Shih, H. H., Keshava Murthy, K. N., & Elnajjar, P. (2021). Bag-of-Words Technique in Natural Language Processing: A Primer for Radiologists. *Radiographics*, 41 (5), 1420-1426. doi:10.1148/rg.2021210025

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and

their Compositionality. *Advances in Neural Information Processing Systems*, 26, 3111-3119

Patil, R., Boit, S., Gudivada, V., & Nandigam, J. (2023). A Survey of Text Representation and Embedding Techniques in NLP. *IEEE Access*, 11, 36120-36146. doi:10.1109/ACCESS.2023.3266377

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. doi:10.1016/j.physd.2019.132306

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K. R., & Cao, Y. (2022, October). React: Synergizing reasoning and acting in language models. In *The eleventh international conference on learning representations*.

Zhang, W., Yoshida, T., & Tang, X. (2011). A comparative study of TF*IDF, LSI and multi-words for text classification. *Expert Systems with Applications*, 38 (3), 2758-2765. doi:10.1016/j.eswa.2010.08.066

BÖLÜM 8

Türkçe Haber Metinlerinde Metin Sınıflandırma Yaklaşımlarının Kronolojik Karşılaştırması

**ŞÜKRÜ EVRAN¹
RECEP ÖZDAĞ²**

Giriş

Dijital çağda üretilen metin verilerinin hacmi her yıl katlanarak artmakta; bu durum, verilerin işlenmesi, sınıflandırılması ve anlamlandırılmasına yönelik etkili yöntemlere olan ihtiyacı artırmaktadır. Metin sınıflandırması, dijital belgelerin içeriklerine göre önceden tanımlanmış kategorilere ayrılması süreci olup, bilgi erişimi, kurumsal belge yönetimi, sosyal medya analitiği ve haber akışının takibi gibi birçok alanda merkezi bir rol oynamaktadır (Sebastiani, 2002). Metin temsillerinin niteliği ve kullanılan sınıflandırma algoritmalarının kapasitesi, bu sistemlerin doğruluğunu doğrudan etkilemektedir. Bu nedenle literatürde sunulan çeşitli yöntemler, metnin bağlamsal yapısını daha iyi yansıtan temsiller üretmeyi ve sınıflandırma performansını artırmayı amaçlamıştır.

¹ Van Yüzüncü Yıl Üniversitesi, Yapay Zeka ve Robotik Anabilim Dalı, Orcid: 0009-0009-7767-1295

² Doç. Dr., Van Yüzüncü Yıl Üniversitesi, Bilgisayar Mühendisliği Bölümü, Orcid: 0000-0001-5247-5591

Metin sınıflandırma alanındaki yöntemlerin gelişimi genel olarak üç ana dönemde ele alınmaktadır: istatistiksel yöntemler, klasik makine öğrenmesi yaklaşımları ve derin öğrenme ile Transformer tabanlı modern dil modelleri. Aşağıdaki alt başlıklarda bu gelişimin tarihsel seyri özetlenmektedir.

İstatistiksel Yöntemler ve Naive Bayes

Metin madenciliğinin erken dönemlerinde Naive Bayes (NB), basitliği ve hesaplama verimliliği sayesinde en yaygın kullanılan sınıflandırıcı olmuştur (Sebastiani, 2002). NB, kelimeler arasındaki ilişkileri koşullu bağımsızlık varsayımıyla basitleştirerek her kelimenin sınıf koşullu olasılığını ayrı ayrı değerlendirmektedir. Bu varsayım dilin doğal yapısıyla tam olarak örtüşmese de yüksek boyutlu metin uzayında NB'nin beklenenden iyi performans gösterdiği çeşitli çalışmalar tarafından rapor edilmiştir (Domingos & Pazzani, 1997). Eğitim sürecinin yalnızca terim frekanslarına dayanması, NB'yi hafif ve hızlı bir yöntem hâline getirmiş ve uzun yıllar boyunca temel bir karşılaştırma modeli olarak kullanılmasını sağlamıştır (Rennie & ark., 2003).

Bununla birlikte NB, kelimeler arası bağımlılıkların belirleyici olduğu durumlarda sınırlı başarı göstermektedir (Lewis, 1998). Ekonomi haberlerinde “borsa”, “hisse”, “piyasa” gibi terimlerin birlikte ortaya çıkma ihtimali yüksek olmasına rağmen NB bu ilişkileri doğrudan modelleyemez. (Yang & Liu, 1999) , (Joachims, 1998) tarafından yapılan karşılaştırmalı çalışmalar, veri seti genişledikçe NB'nin daha gelişmiş yöntemlerin gerisinde kaldığını göstermiştir. Buna rağmen NB, küçük veri setleri veya çok yüksek boyutlu temsil uzaylarında hâlen makul bir temel seviye yöntem olarak kullanılmaktadır (McCallum & Nigam, 1998). Bu çalışmada da NB, istatistiksel dönemin temsilcisi olarak değerlendirilmeye dahil edilmiştir.

Klasik Makine Öğrenmesi Yaklaşımları: TF-IDF ve SVM

1990'ların sonu ve 2000'lerin başı, metin sınıflandırma alanında daha gelişmiş özellik çıkarım teknikleri ile güçlü denetimli öğrenme algoritmalarının kullanıldığı bir dönemdir. Bu dönemde TF-IDF (Salton & Buckley, 1988) gibi ağırlıklandırma yöntemleri standart hâle gelmiş; sözcüklerin hem belge içindeki sıklığını hem de tüm korpustaki ayırt ediciliğini dikkate alan temsiller üretmiştir.

Bu temsiller üzerinde etkili biçimde çalışan Destek Vektör Makineleri (SVM), yüksek boyutlu ve seyrek metin verisinde sağladığı başarı ile öne çıkmıştır. SVM, sınıflar arasındaki ayrımı maksimize eden hiperdüzlemi hedefler ve aşırı uyuma karşı dirençli yapısıyla geniş veri setlerinde yüksek doğruluk sağlamıştır (Joachims, 1998). (Yang & Liu, 1999) çalışmalarında, SVM'nin yeterli eğitim verisi ile NB gibi daha basit yöntemlere kıyasla belirgin performans üstünlüğü sağladığını göstermiştir. Reuters-21578 gibi standart veri setlerinde %85'in üzerindeki doğruluklar, bu yöntemi uzun süre alanın referans yaklaşımı hâline getirmiştir.

Bu dönemde k-en yakın komşu, Naive Bayes, karar ağaçları, sinir ağları ve regresyon temelli yaklaşımlar gibi farklı makine öğrenmesi yöntemleri metin sınıflandırmada yaygın biçimde uygulanmıştır (Yang & Liu, 1999). Ancak genel eğilim, TF-IDF + lineer SVM yapısının performans, doğruluk ve hesaplanabilirlik açısından en istikrarlı çözüm olarak benimsenmesi yönündedir. Günümüzde bile belirli veri koşullarında (yüksek boyut, sınırlı örnek sayısı) bu yöntemlerin rekabetçi olduğu bilinmektedir (Kowsari & ark., 2019).

Derin Öğrenme Tabanlı Yaklaşımlar

2010'lara gelindiğinde artan hesaplama gücü ve büyük veri havuzlarına erişimle birlikte derin öğrenme yaklaşımları doğal dil işlemede geniş yer bulmuştur (Minaee & ark., 2021). Bu dönemin en önemli yeniliği, özellik çıkarımının tamamen otomatik hâle

gelmesidir. Word2Vec (Mikolov & ark., 2013) ve GloVe (Pennington, Socher & Manning, 2014) gibi dağıtık kelime temsilleri, kelimeler arasındaki anlamsal ilişkileri vektör uzayında yansıtarak klasik temsillere göre daha güçlü bilgi sağlamıştır.

Bu gömme vektörlerini kullanan CNN (Kim, 2014) ve LSTM tabanlı modeller (Hochreiter & Schmidhuber, 1997) çeşitli sınıflandırma görevlerinde yüksek doğruluk elde ederek derin öğrenmenin metin sınıflandırmadaki potansiyelini ortaya koymuştur. Özellikle dikkat (attention) mekanizmasının RNN tabanlı yapılara eklenmesiyle modeller, karar sürecinde metnin önemli bölümlerine odaklanabilmiş ve uzun metinlerdeki bağlam ilişkilerini daha etkili yakalayabilmiştir (Zhou & ark., 2016).

Transformer Mimarisi ve Önceden Eğitilmiş Dil Modelleri

2017 yılında Transformer mimarisinin (Vaswani & ark., 2017) tanıtılması, doğal dil işlemede önemli bir kırılma noktası oluşturmuştur. Bu mimari, tamamen dikkat mekanizmasına dayalı yapısıyla kelimeler arası ilişkileri paralel olarak işlemekte ve dilin bağlamsal özelliklerini güçlü biçimde temsil etmektedir. Bu avantajları sayesinde Transformer tabanlı modeller kısa sürede pek çok NLP görevinde önceki yöntemlerin yerini almıştır.

Bu gelişmenin ardından önceden eğitilmiş büyük dil modelleri önem kazanmıştır. (Devlin & ark., 2019) tarafından geliştirilen BERT (Bidirectional Encoder Representations from Transformers), iki yönlü bağlam anlayışı ve geniş ölçekli ön eğitim stratejileriyle birçok sınıflandırma görevinde önceki yaklaşımları geride bırakmıştır. Bununla birlikte, bu tür büyük modellerin yüksek hesaplama gereksinimleri, model sıkıştırma ve bilgi damıtma gibi tekniklerin geliştirilmesine yol açmıştır (Hinton, Vinyals & Dean, 2015). TinyBERT (Jiao & ark., 2020), bu doğrultuda büyük bir öğretmen modelin bilgisini daha küçük bir öğrenci modele

aktarmakta ve daha düşük kaynak gereksinimiyle yüksek doğruluk sunmaktadır.

Yöntem

Bu bölümde, çalışmada kullanılan veri seti, uygulanan ön işleme adımları, sınıflandırma yöntemleri ve model performansını değerlendirmek için kullanılan ölçütler ayrıntılı biçimde açıklanmaktadır. Amaç, farklı dönemleri temsil eden metin sınıflandırma yaklaşımlarının hem doğruluk hem de hesaplama verimliliği açısından karşılaştırılmasını sağlamaktır.

1. Veri Seti

Bu çalışmada kullanılan veri seti, 2025 yılına ait toplam 55.000 Türkçe haber metninden oluşmaktadır. Her bir haber metni, verilerin derlendiği ulusal haber ajansı sitelerinde önceden tanımlanmış 13 konu kategorisinden birine aittir. Bu kategoriler; siyaset, spor, ekonomi, sağlık, teknoloji, kültür-sanat, eğitim, dünya, yerel, magazin, emlak, iş dünyası ve trafik başlıklarından oluşmaktadır. Haberler farklı ulusal kaynaklardan elde edildiği için metin uzunluğu, sözcük çeşitliliği ve anlatım biçimi açısından heterojen bir yapı sergilemektedir.

Veri setinde yer alan bir haber metninin ortalama uzunluğu yaklaşık 80 kelime civarındadır. Tüm korpus dikkate alındığında toplam kelime sayısı yaklaşık 2 milyon düzeyindedir. Bu büyüklük, hem klasik makine öğrenmesi yöntemlerinin hem de derin öğrenme tabanlı modellerin eğitimi için yeterli bir örneklem sunmaktadır.

2. Eğitim–Test Ayrımı

Modeller arası karşılaştırmaların adil ve tutarlı biçimde yapılabilmesi amacıyla veri seti eğitim ve test kümelerine ayrılmıştır. Verilerin %80'i (44.000 haber) eğitim, %20'si (11.000 haber) test amacıyla kullanılmıştır. Eğitim süreci içerisinde ayrıca, hiperparametre ayarlamaları ve erken durdurma mekanizmalarının

uygulanabilmesi için eğitim verisinin %10'luk bir bölümü doğrulama kümesi olarak ayrılmıştır. Test kümesi, modellerin nihai performansının değerlendirilmesi amacıyla yalnızca bir kez kullanılmıştır.

3. Ön İşleme Adımları

Türkçe haber metinleri üzerinde sınıflandırma modelleri eğitilmeden önce, metinlerin gürültüden arındırılması ve daha tutarlı bir temsil elde edilmesi amacıyla bir dizi ön işleme adımı uygulanmıştır. Bu adımlar, özellikle Türkçe'nin dilsel özellikleri dikkate alınarak tasarlanmıştır.

a. Tokenizasyon:

Haber metinleri kelime düzeyinde parçalanmış; noktalama işaretleri ve semboller ayrı tokenlar olarak ele alınmıştır. Bu işlem Python programlama dili kullanılarak ve NLTK kütüphanesi aracılığıyla gerçekleştirilmiştir.

b. Küçük harfe dönüştürme:

Türkçe'de büyük/küçük harf farklılıkları anlamsal olarak aynı kelimelerin ayrı özellikler olarak değerlendirilmesine yol açabildiğinden, tüm metinler küçük harfe dönüştürülmüştür. Bu işlem sırasında özellikle Türkçe'ye özgü "İ/i" dönüşümüne dikkat edilmiştir.

c. Stop-word filtreleme:

Anlamsal katkısı sınırlı olan ve metin sınıflandırma performansına genellikle olumlu etki etmeyen sık kullanılan işlevsel kelimeler metinlerden çıkarılmıştır. Bu amaçla Türkçe stop-word listesi kullanılmış ve yaklaşık 100 kelime filtelenmiştir. Bu adım, özellikle TF-IDF tabanlı temsillerde gürültünün azaltılmasına katkı sağlamıştır. Derin öğrenme modellerinde etkisi sınırlı olduğundan bu adım yalnızca klasik yöntemlerde belirleyici olmuştur.

d. Noktalama ve sayı temizleme:

Virgöl, nokta, alıntı işaretleri gibi noktalama karakterleri metinden kaldırılmıştır. Sayısal ifadeler ise özgül değerleri korunmadan genel bir <sayı> tokeni ile temsil edilmiştir. Bu yaklaşım, kelime frekanslarının daha tutarlı biçimde hesaplanmasını amaçlamaktadır.

e. Kök ve gövde çözümleme:

Türkçe'nin eklemeli yapısı nedeniyle aynı kökten türeyen çok sayıda kelime biçimi bulunabilmektedir. Bu çeşitliliği azaltmak ve özellik uzayını daraltmak amacıyla kelimeler kök düzeyine indirgenmiştir. Bu işlemde Zemberek Kütüphanesi kullanılmıştır. Zemberek, kelimelerin çekim ve yapım eklerini ayırarak temel sözlük biçimlerini üretmekte ve Türkçe dilbilgisine özgü kuralları dikkate almaktadır (Akın & Akın, 2007). Kök indirgeme işlemi sonrasında ortaya çıkabilecek anlam kayıpları sınırlı ölçüde manuel olarak kontrol edilmiştir. Bu adım, özellikle Naive Bayes ve SVM modellerinde özellik sayısının azaltılmasına ve öğrenme sürecinin iyileştirilmesine katkı sağlamıştır.

f. Vektörleştirme:

Ön işleme adımlarının ardından elde edilen metinler, sınıflandırma modellerinin girdisi olacak şekilde sayısal vektörlere dönüştürülmüştür. Naive Bayes ve SVM modelleri için, köklenmiş kelimelerden oluşan bag-of-words temsili TF-IDF ağırlıklarıyla kullanılmıştır. Derin öğrenme tabanlı modellerde ise vektörleştirme işlemi modelin kendi tokenizer ve gömme katmanları aracılığıyla gerçekleştirilmiştir. Bu nedenle BERT ve TinyBERT modellerine verilen metinler, ilgili modellerin sözlük ve tokenizasyon yapısına uygun biçimde hazırlanmıştır.

Ön işleme süreci tamamlandıktan sonra her bir haber metni, ya TF-IDF tabanlı bir kelime vektörü ya da derin öğrenme

modellerine uygun token kimlik dizisi şeklinde temsil edilmiştir. Bu temsiller, bir sonraki aşamada açıklanan sınıflandırma modellerinin eğitimi ve değerlendirilmesinde kullanılmıştır.

4. Sınıflandırma Yöntemleri

Bu çalışmada, metin sınıflandırma alanındaki tarihsel gelişimi temsil edecek şekilde dört farklı yaklaşım değerlendirilmiştir. Bu bölümde öncelikle istatistiksel ve klasik makine öğrenmesi yöntemleri ele alınmakta; bir sonraki alt başlıkta ise Transformer tabanlı derin öğrenme modelleri sunulmaktadır.

a. Naive Bayes:

Naive Bayes (NB), metin sınıflandırma alanında yaygın olarak kullanılan en temel olasılıksal yaklaşımlardan biridir. Bu çalışmada, kelime frekanslarına dayalı metin temsilleriyle uyumlu olması nedeniyle Multinomial Naive Bayes varyantı tercih edilmiştir. Model, Bayes teoremi çerçevesinde her bir sınıf için kelime dağılımlarını öğrenmekte ve yeni bir belgenin ait olduğu sınıfı, maksimum posterior olasılığa göre belirlemektedir.

Eğitim aşamasında, her kategori için eğitim kümesinde yer alan kelimelerin frekansları hesaplanmış ve bu frekanslar kullanılarak sınıf koşullu olasılıklar elde edilmiştir. Sınıflandırma sürecinde, bir belgedeki tüm kelimelerin ilgili sınıfta görülme olasılıkları birleştirilerek posterior olasılık hesaplanmıştır. Sayısal kararlılığı sağlamak amacıyla hesaplamalar logaritmik biçimde gerçekleştirilmiş; sıfır frekans problemine karşı Laplace düzeltmesi (Add-1 smoothing) uygulanmıştır.

Naive Bayes modeli, belirgin bir hiperparametre ayarlaması gerektirmediğinden ek optimizasyon yapılmaksızın eğitilmiş ve test edilmiştir. Hesaplama karmaşıklığının düşük olması ve eğitim süresinin kısa olması nedeniyle bu yöntem, çalışmada istatistiksel

dönemi temsil eden bir taban çizgisi (baseline) olarak değerlendirilmiştir.

b. TF-IDF + SVM:

İkinci yaklaşım olarak, klasik makine öğrenmesi yöntemlerini temsil eden TF-IDF tabanlı SVM modeli kullanılmıştır. Bu modelde her bir haber metni, kök indirgeme işlemi uygulanmış kelimelerden oluşan TF-IDF ağırlıklı bag-of-words temsili ile vektörleştirilmiştir. Eğitim korpusunda yer alan benzersiz kök sayısı dikkate alındığında, özellik uzayı yaklaşık 50.000 boyutlu seyrek bir yapı sergilemektedir.

Bu yüksek boyutlu veri üzerinde sınıflandırma için lineer çekirdekli SVM tercih edilmiştir. Çok sınıflı problem, one-vs-rest stratejisi kullanılarak ele alınmış; her sınıf için ayrı bir karar sınırı öğrenilmiştir. Uygulamada, büyük ve seyrek veri yapılarında verimliliği nedeniyle scikit-learn kütüphanesinin LinearSVC implementasyonu kullanılmıştır.

SVM modelinin düzenleme katsayısı C , eğitim verisi üzerinde 5 katlı çapraz doğrulama ile belirlenmiştir. En iyi performans, yaygın olarak kullanılan $C = 1.0$ değeri ile elde edilmiştir. Eğitim sürecinde model, sınıflar arasındaki ayrımı maksimize eden hiper düzlemleri öğrenmiş; test aşamasında her bir haber metni için üretilen karar fonksiyonları karşılaştırılarak en yüksek skoru veren sınıf tahmin edilmiştir.

TF-IDF + SVM yaklaşımı, derin öğrenme yöntemleri öncesi dönemde metin sınıflandırma alanında uzun süre referans yöntem olarak kabul edilmiş; özellikle yüksek boyutlu ve sınırlı örnek sayısına sahip veri setlerinde istikrarlı performansı ile öne çıkmıştır. Bu çalışmada da, klasik makine öğrenmesi paradigmasını temsil eden güçlü bir karşılaştırma modeli olarak değerlendirmeye dâhil edilmiştir. Bu yönleriyle model, derin öğrenme yöntemleri için güçlü bir karşılaştırma zemini sunmaktadır.

c. Transformer Tabanlı Modeller:

Klasik makine öğrenmesi yöntemlerinin kelime frekanslarına dayalı temsillerle sınırlı kalması, bağlamsal ilişkilerin modellenmesinde önemli kısıtlar doğurmaktadır. Bu çalışmada, kelimeler arası bağlamı doğrudan temsil edebilen modern yaklaşımları değerlendirmek amacıyla Transformer tabanlı iki önceden eğitilmiş dil modeli kullanılmıştır: BERT ve TinyBERT. Her iki model de haber metinlerinin bağlamsal temsillerini öğrenmek üzere ince ayar (fine-tuning) süreciyle sınıflandırma görevine uyarlanmıştır.

BERT, çok katmanlı Transformer kodlayıcı mimarisine dayanan ve iki yönlü bağlam bilgisini aynı anda modelleyebilen önceden eğitilmiş bir dil modelidir (Devlin & ark., 2019). Bu çalışmada, Türkçe'yi de kapsayan çok dilli BERT (mBERT) modeli kullanılmıştır. mBERT, farklı dillerdeki Wikipedia metinleri üzerinde eğitilmiş olup yaklaşık 110 milyon parametreye sahiptir.

BERT modelinin haber sınıflandırma görevine uyarlanması için, modelin son katmanına 13 sınıflı bir yoğunluk (dense) katmanı eklenmiş ve softmax aktivasyonu ile sınıf olasılıkları elde edilmiştir. İnce ayar sürecinde, modelin tüm ağırlıkları veri setine özgü olarak güncellenmiştir. Eğitim sırasında mini-batch yaklaşımı kullanılmış; öğrenme oranı $2e-5$ olarak belirlenmiş ve eğitim üç epoch boyunca sürdürülmüştür. Doğrulama kümesinde kaybın iyileşmemesi durumunda erken durdurma uygulanmıştır.

BERT modeli, WordPiece tokenizasyonu kullanmakta olup metinler alt-kelime birimlerine ayrılarak modele sunulmuştur. Girdi dizilerinin maksimum uzunluğu 256 token ile sınırlandırılmış; daha uzun haber metinlerinde yalnızca ilk 256 token dikkate alınmıştır. Sınıflandırma kararı, giriş dizisinin başına eklenen [CLS] tokenına ait bağlamsal temsil üzerinden verilmiştir. Bu yapı sayesinde model,

haber metninin tamamından elde ettiği bilgiyi tek bir vektör temsili üzerinden sınıflandırma sürecine aktarmaktadır.

TinyBERT, BERT modelinin bilgi damıtma (knowledge distillation) yöntemiyle küçültülmüş bir versiyonu olup, daha düşük hesaplama maliyetiyle BERT'e yakın performans sunmayı hedeflemektedir (Jiao & ark., 2020). Bu çalışmada, TinyBERT-6 olarak adlandırılan ve 6 Transformer katmanına sahip sürüm kullanılmıştır. Bu model, BERT'e kıyasla daha az parametre içermesine rağmen benzer mimari prensiplere dayanmaktadır.

TinyBERT modeli de BERT ile benzer şekilde sınıflandırma görevine uyarlanmış; son katmanına 13 sınıflı bir çıktı katmanı eklenerek ince ayar yapılmıştır. Eğitim sürecinde BERT ile aynı öğrenme oranı ve epoch sayısı kullanılmıştır. Modelin daha kompakt yapısı sayesinde eğitim süreci daha kısa sürmüş ve bellek gereksinimi belirgin biçimde azalmıştır.

Bilgi damıtma yaklaşımı sayesinde TinyBERT, öğretmen model olan BERT'in ara katman temsillerini ve çıktı dağılımlarını öğrenerek daha küçük bir ağ yapısında benzer bağlamsal bilgiyi yakalamayı amaçlamaktadır. Bu özellik, TinyBERT'i özellikle hesaplama kaynaklarının sınırlı olduğu senaryolarda veya gerçek zamanlı uygulamalarda tercih edilebilir bir alternatif hâline getirmektedir. Bu çalışmada TinyBERT, doğruluk-verimlilik dengesi açısından BERT ile karşılaştırmalı olarak değerlendirilmiştir.

5. Değerlendirme Metrikleri

Bu çalışmada kullanılan sınıflandırma modellerinin performansını karşılaştırmak amacıyla, hem sınıflandırma başarımını hem de hesaplama verimliliğini yansıtan çeşitli ölçütler kullanılmıştır. Klasik sınıflandırma başarımının değerlendirilmesi için doğruluk, kesinlik, duyarlılık ve F1-skoru hesaplanmıştır. Buna ek olarak, modellerin pratik uygulanabilirliğini değerlendirmek

amacıyla eğitim süresi, çıkarım hızı ve bellek kullanımı gibi hesaplama verimliliği kriterleri de incelenmiştir.

Doğruluk (Accuracy), modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranını ifade ederek genel sınıflandırma başarısına ilişkin özet bir gösterge sunmaktadır. Ancak doğruluk tek başına sınıf bazlı performansı yansıtmakta yetersiz kalabildiğinden, daha ayrıntılı değerlendirme için kesinlik ve duyarlılık ölçütleri de dikkate alınmıştır.

Kesinlik (Precision), modelin belirli bir sınıf için yaptığı pozitif tahminlerin ne kadarının gerçekte doğru olduğunu göstermekte ve yanlış alarm oranı hakkında bilgi vermektedir. Duyarlılık (Recall) ise gerçek pozitif örneklerin ne kadarının model tarafından doğru biçimde yakalanabildiğini ifade etmekte olup, modelin kaçırma eğilimini yansıtmaktadır. F1-skoru, kesinlik ve duyarlılığın harmonik ortalaması olarak, bu iki ölçüt arasında dengeli bir değerlendirme sunmakta ve özellikle dengesiz veri kümelerinde daha güvenilir bir performans göstergesi olarak kullanılmaktadır.

Bu metrikler her bir kategori için ayrı ayrı hesaplanabildiği gibi, genel performansı yansıtmak amacıyla mikro veya makro ortalama değerler üzerinden de raporlanabilmektedir. Bu çalışmada, sonuçlar toplam doğruluk ve makro ortalama F1-skoru üzerinden sunulmuştur.

Sınıflandırma başarımına ek olarak, modellerin hesaplama açısından pratikliğini değerlendirmek amacıyla eğitim süresi, çıkarım hızı ve bellek kullanımı ölçütleri ele alınmıştır. Eğitim süresi, bir modelin eğitiminin tamamlanması için gereken toplam zamanı ifade ederken; çıkarım hızı, eğitilmiş bir modelin yeni haber metinlerini sınıflandırma hızını göstermektedir. Bellek kullanımı ise, özellikle derin öğrenme tabanlı modellerde önemli bir kısıtlayıcı

faktör olduğundan, eğitim ve çıkarım aşamalarında kullanılan bellek miktarı dikkate alınarak değerlendirilmiştir.

Bu ölçütler birlikte ele alındığında, modeller yalnızca doğruluk açısından değil, doğruluk–verimlilik dengesi bakımından da karşılaştırılabilmektedir. Bir sonraki bölümde, dört farklı modelin deneysel sonuçları bu ölçütler çerçevesinde ayrıntılı olarak sunulacak ve karşılaştırmalı biçimde tartışılacaktır.

Bu bölümde tanımlanan değerlendirme ölçütleri doğrultusunda, bir sonraki bölümde Naive Bayes, TF-IDF + SVM, BERT ve TinyBERT modellerinin deneysel sonuçları karşılaştırmalı olarak sunulacak ve her bir yaklaşımın güçlü ve zayıf yönleri tartışılacaktır.

DeneySEL Sonuçlar

Bu bölümde, önceki kısımlarda tanıtılan dört metin sınıflandırma yöntemi olan Naive Bayes, TF-IDF + Linear SVM, TinyBERT ve BERT modellerinin, 55.000 Türkçe haberden oluşan veri seti üzerindeki deneysel sonuçları sunulmaktadır. Veri seti 44.000 eğitim ve 11.000 test örneği olacak şekilde ikiye ayrılmış; modeller makro ortalama doğruluk, kesinlik, duyarlılık ve F1-skoru metrikleri kullanılarak değerlendirilmiştir. Ayrıca, modellerin pratik uygulanabilirliğini incelemek amacıyla eğitim süresi ve çıkarım (inference) maliyeti de analiz edilmiştir.

Dört modelin test kümesi üzerindeki sınıflandırma performansları Tablo 1’de özetlenmiştir. Tablo 1’e göre, en yüksek doğruluk ve F1-skorunun BERT modeline ait olduğu görülmektedir. BERT modeli yaklaşık %79,6 doğruluk ve 0,700 F1-skoru ile çalışmadaki en başarılı yöntem olmuştur. Bu sonuç, Transformer tabanlı modellerin Türkçe haber sınıflandırma görevinde klasik yöntemlere kıyasla daha güçlü bağlamsal temsiller üretebildiğini göstermektedir.

Tablo 1. Dört modelin sınıflandırma performanslarının karşılaştırılması

Model Adı	Doğruluk (Accuracy)	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru
Naive Bayes	0,7257	0,6440	0,4292	0,4723
TF-IDF + Linear SVM	0,7924	0,7314	0,6708	0,6952
TinyBERT	0,7273	0,6252	0,5342	0,5348
BERT	0,7964	0,7280	0,6835	0,6997

TF-IDF + Linear SVM, BERT'e oldukça yakın bir performans sergilemiş; yaklaşık %79,2 doğruluk ve 0,695 F1-skoru elde etmiştir. BERT ile SVM arasındaki fark, doğruluk açısından yaklaşık 0,4 yüzde puan düzeyinde kalmıştır.

Kesinlik ve duyarlılık değerleri birlikte değerlendirildiğinde, SVM modelinin kesinlik açısından BERT'ten bir miktar daha yüksek, buna karşılık BERT'in duyarlılık açısından SVM'i geçtiği görülmektedir. Bu durum, SVM'in daha temkinli (yanlış pozitifleri azaltmaya eğilimli), BERT'in ise daha kapsayıcı (daha fazla gerçek pozitif yakalamaya odaklı) bir sınıflandırma davranışı sergilediğini göstermektedir.

TinyBERT, doğruluk açısından Naive Bayes'e oldukça yakın bir değer elde etmiş; ancak F1-skoru bakımından Naive Bayes'i belirgin biçimde geride bırakmıştır. TinyBERT'in özellikle duyarlılık değerinin daha yüksek olması, daha fazla gerçek pozitif örneği doğru biçimde sınıflandırabildiğini göstermektedir.

Naive Bayes, düşük duyarlılık değeri nedeniyle diğer yöntemlerin gerisinde kalmakla birlikte, %72,6 doğruluk ile anlamlı bir taban çizgisi (baseline) performansı sunmaktadır.

Modellerin yalnızca sınıflandırma başarımı değil, hesaplama maliyetleri açısından da karşılaştırılması gerçek uygulamalar için kritik öneme sahiptir. Bu amaçla, modellerin eğitim süresi, toplam

çıkartım süresi ve örnek başına çıkartım süresi ölçülmüştür. Sonuçlar Tablo 2’de sunulmaktadır.

Tablo 2. Modellerin eğitim ve çıkartım sürelerinin karşılaştırılması

Model Adı	Eğitim Süresi (sn)	Toplam Çıkartım Süresi (sn)	Örnek Başına Çıkartım Süresi (ms)
Naive Bayes	0,1411	0,0140	0,0012
TF-IDF + Linear SVM	4,0036	0,0168	0,0014
TinyBERT (6 katmanlı)	2867,8482	80,6893	6,8888
BERT	5849,4863	162,2674	13,8536

Tablo 2’ye göre Naive Bayes ve SVM modelleri son derece hızlıdır. Naive Bayes modeli yaklaşık 0,14 saniyede eğitilmiş ve 11.000 test örneğini 0,014 saniyede sınıflandırmıştır. SVM modeli ise yaklaşık 4 saniyelik eğitim süresi ile benzer biçimde düşük çıkartım maliyeti sunmaktadır. Her iki yöntem de örnek başına çıkartım süresi açısından milisaniyenin çok altında değerler üretmektedir.

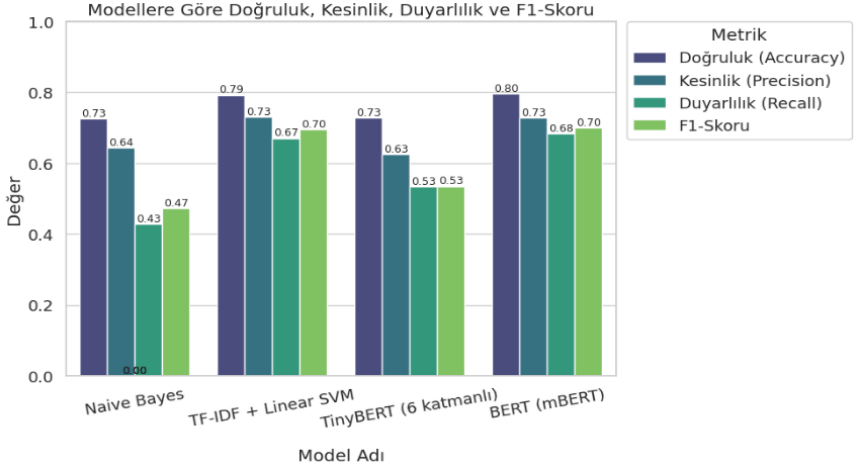
TinyBERT, BERT’e kıyasla yaklaşık iki kat daha hızlı çalışmaktadır. TinyBERT’in eğitim süresi yaklaşık 2.868 saniye ($\approx$ 47,8 dakika) iken, BERT’in eğitim süresi yaklaşık 5.849 saniye ($\approx$ 1,62 saat) olarak ölçülmüştür. Benzer şekilde, TinyBERT test kümesini 80,7 saniyede, BERT ise 162,3 saniyede sınıflandırmıştır.

Örnek başına çıkartım süreleri incelendiğinde, TinyBERT’in yaklaşık 6,9 ms, BERT’in ise 13,9 ms değerlerine sahip olduğu görülmektedir. Bu sonuçlar, TinyBERT’in BERT’e göre belirgin bir hız avantajı sunduğunu doğrulamaktadır.

Klasik yöntemler ile derin öğrenme tabanlı modeller arasındaki hesaplama maliyeti farkı oldukça büyüktür. Örnek başına çıkartım süresi açısından, BERT’in değeri Naive Bayes ve SVM’e kıyasla yaklaşık dört merteye ($\approx$ 10.000 kat) daha yüksektir. Bu

durum, yüksek hacimli veri akışları veya kaynak kısıtlı ortamlar için model seçiminin yalnızca doğruluk değil, hesaplama maliyeti açısından da değerlendirilmesi gerektiğini göstermektedir.

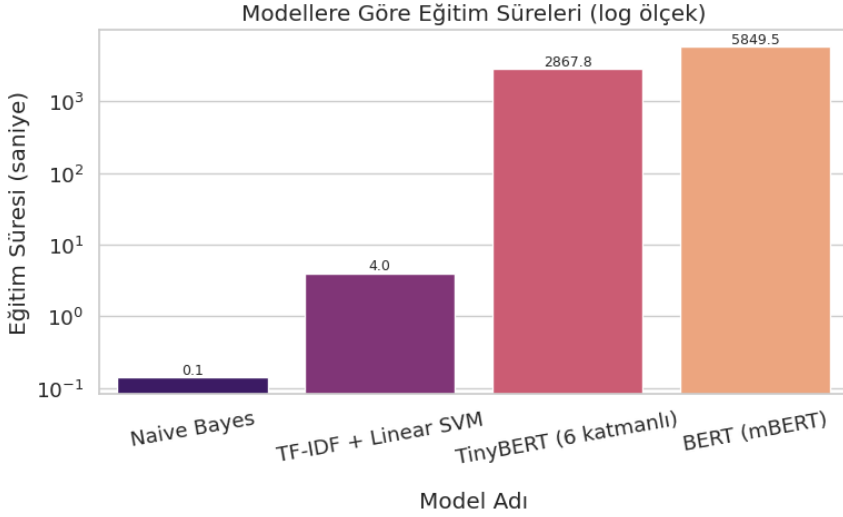
Şekil 1. Modellerin doğruluk, kesinlik, duyarlılık ve F1-skoru karşılaştırması



Şekil 1, Naive Bayes'in temel bir başlangıç noktası sunduğunu; TF-IDF + Linear SVM ve BERT modellerinin genel olarak daha dengeli ve yüksek performans sağladığını; TinyBERT'in ise doğruluk açısından klasik yöntemlere yakın, F1-skoru açısından ise daha güçlü bir profil sunduğunu göstermektedir.

Modellerin eğitim süreleri logaritmik ölçekte karşılaştırmalı olarak Şekil 2'de gösterilmektedir. Logaritmik gösterim, klasik yöntemler ile Transformer tabanlı modeller arasındaki hesaplama maliyeti farkını daha açık biçimde ortaya koymaktadır. Bu fark, derin modellerin sağladığı performans artışının, uygulama senaryosuna bağlı olarak ek zaman ve donanım maliyetleriyle dengelenmesi gerektiğini göstermektedir.

Şekil 2. Modellerin eğitim ve çıkarım sürelerinin karşılaştırılması



Tartışma

Bu çalışmada elde edilen deneysel sonuçlar, metin sınıflandırma yöntemlerinin tarihsel evriminin, hem sınıflandırma başarımı hem de hesaplama maliyeti açısından ne tür kazanımlar ve sınırlılıklar getirdiğini açık biçimde ortaya koymaktadır. Bulgular, modern Transformer tabanlı modellerin sağladığı doğruluk artışının, klasik yöntemlerle karşılaştırıldığında belirli koşullara bağlı olarak değerlendirilmesi gerektiğini göstermektedir.

a. Derin Öğrenme Modellerinin Performans Kazanımı:

BERT modelinin doğruluk ve F1-skoru açısından en yüksek değerlere ulaşması, Transformer tabanlı dil modellerinin metin sınıflandırma görevlerinde sunduğu güçlü bağlamsal temsil yeteneğini desteklemektedir. BERT'in çift yönlü bağlamı aynı anda işleyebilmesi ve geniş ölçekli çok dilli bir korpus üzerinde önceden eğitilmiş olması, özellikle farklı konu kategorilerine sahip haber metinlerinde daha tutarlı ve dengeli genelleme yapabilmesini sağlamaktadır.

Bununla birlikte, BERT'in klasik yöntemlere kıyasla sağladığı performans artışının sınırlı ancak tutarlı olduğu görülmektedir. TF-IDF + Lineer SVM modeli ile karşılaştırıldığında, BERT yaklaşık 0,4 yüzde puan daha yüksek doğruluk ve 0,005 civarında daha yüksek F1-skoru sunmaktadır. Bu fark, akademik karşılaştırmalar açısından anlamlı olmakla birlikte, hesaplama maliyetinin kritik olduğu bazı uygulama senaryolarında klasik yöntemlerin tercih edilmesini dışlamamaktadır.

b. Klasik Makine Öğrenmesinin Süregelen Rekabet Gücü:

TF-IDF + Lineer SVM modeli, bu çalışmada performans–maliyet dengesi açısından öne çıkan yöntemlerden biri olmuştur. BERT'e oldukça yakın doğruluk ve F1-skoru değerleri elde ederken, eğitim ve çıkarım süreleri bakımından çok daha düşük bir hesaplama maliyeti sunmaktadır. Bu durum, SVM tabanlı yaklaşımların metin sınıflandırma alanında uzun süre “state-of-the-art” olarak kabul edilmesiyle uyumlu bir bulgudur.

Özellikle TF-IDF temsillerinin yeterli ayırt ediciliği sağladığı, veri setinin orta ölçekli olduğu ve hesaplama kaynaklarının sınırlı bulunduğu senaryolarda, SVM'nin hâlen güçlü ve pratik bir çözüm sunduğu görülmektedir. Bu çalışma, modern derin öğrenme modellerinin yaygınlaşmasına rağmen, klasik makine öğrenmesi yöntemlerinin tamamen göz ardı edilmemesi gerektiğini nicel bulgularla desteklemektedir.

c. TinyBERT'in Görece Sınırlı Performansı:

TinyBERT modeli, bilgi damıtma yaklaşımı sayesinde BERT'e yakın doğruluk değerlerini daha düşük hesaplama maliyetiyle sunmayı hedeflemektedir. Ancak bu çalışmada TinyBERT'in performansının, beklendiği ölçüde SVM seviyesine yaklaşmadığı gözlenmiştir. TinyBERT, doğruluk açısından Naive Bayes'e yakın, F1-skoru açısından ise Naive Bayes'ten daha iyi bir sonuç üretmiş; ancak TF-IDF + SVM modelinin gerisinde kalmıştır.

Bu durumun olası nedenleri arasında, kullanılan TinyBERT sürümünün Türkçe haber diline tam uyumlu olmaması, ince ayar aşamasında hiperparametrelerin BERT ve TinyBERT için ayrı ayrı optimize edilmemiş olması ve bilgi damıtma sürecinin doğrudan bu veri seti üzerinde yeniden gerçekleştirilmemiş olması sayılabilir. Buna rağmen TinyBERT, BERT'e kıyasla yaklaşık iki kat daha hızlı eğitim ve çıkarım süreleri sunarak, orta düzey performans ve orta düzey maliyet gerektiren uygulamalar için dengeli bir alternatif oluşturmaktadır.

d. Naive Bayes'in Taban Çizgisi Olarak Konumu:

Naive Bayes modeli, bu çalışmada esas olarak tarihsel ve metodolojik bir taban çizgisi (baseline) oluşturmak amacıyla değerlendirilmiştir. Sonuçlar, Naive Bayes'in son derece hızlı eğitilip çalışmasına karşın, düşük duyarlılık ve F1-skoru nedeniyle karmaşık ve çok sınıflı haber sınıflandırma görevlerinde sınırlı kaldığını göstermektedir.

Buna karşın, çok az sayıda örnekle hızlı prototipleme yapılması gereken durumlarda veya donanım kaynaklarının son derece kısıtlı olduğu uç cihaz senaryolarında, Naive Bayes hâlen uygulanabilir bir başlangıç yöntemi olarak değerlendirilebilir.

e. Uygulama Senaryolarına Göre Model Seçimi:

Elde edilen bulgular, model seçiminin doğrudan uygulama senaryosuna bağlı olduğunu ortaya koymaktadır. En yüksek doğruluğun hedeflendiği ve yeterli GPU kaynağının mevcut olduğu araştırma odaklı veya gecikmeye toleranslı çevrimdışı uygulamalarda BERT en uygun tercih olarak öne çıkmaktadır. Üretim ortamlarında hem yüksek doğruluk hem de düşük gecikme gerektiren, büyük hacimli haber akışlarının işlendiği senaryolarda ise TF-IDF + Lineer SVM pratik ve güvenilir bir çözüm sunmaktadır.

Kaynakların daha sınırlı olduđu ancak modern mimarilere yakın bir performansın arandığı durumlarda TinyBERT, hız ve doğruluk arasında dengeli bir seçenek oluşturmaktadır. Son olarak, çok düşük donanım gereksinimleri veya hızlı prototipleme hedeflendiğinde, Naive Bayes düşük maliyeti ve uygulama kolaylığı sayesinde uygun bir taban model olarak tercih edilebilir.

Sonuç ve Gelecek Çalışmalar

Bu çalışmada, 55.000 Türkçe haber metninden oluşan bir veri seti üzerinde, metin sınıflandırma alanındaki tarihsel evrimi temsil eden dört farklı yaklaşım tek ve tutarlı bir deneysel çerçeve altında karşılaştırılmıştır. İncelenen yöntemler; Naive Bayes, TF-IDF + Lineer SVM, BERT ve TinyBERT modellerinden oluşmaktadır. Karşılaştırmalar, yalnızca sınıflandırma başarımı açısından değil, aynı zamanda hesaplama maliyeti ve pratik uygulanabilirlik boyutları da dikkate alınarak gerçekleştirilmiştir.

Elde edilen bulgular, Transformer tabanlı BERT modelinin, doğruluk ve F1-skoru açısından en yüksek performansı sağladığını göstermektedir. Bu sonuç, önceden eğitilmiş derin dil modellerinin, Türkçe haber sınıflandırma görevlerinde de güçlü ve genellenebilir temsiller üretebildiğini ortaya koymaktadır. Bununla birlikte, BERT'in sağladığı performans artışının, klasik yöntemlere kıyasla sınırlı ancak tutarlı olduğu görülmektedir.

TF-IDF + Lineer SVM yaklaşımı, performans–maliyet dengesi açısından çalışmanın en dikkat çekici sonuçlarından birini sunmaktadır. BERT'e oldukça yakın doğruluk ve F1-skoru değerleri elde edilirken, eğitim ve çıkarım sürelerinin birkaç büyüklük mertebesi daha düşük olması, bu yöntemi pek çok gerçek dünya uygulaması için son derece cazip kılmaktadır. Bu bulgu, klasik makine öğrenmesi yöntemlerinin modern derin modellerin varlığına rağmen hâlen güçlü birer alternatif olabildiğini göstermektedir.

TinyBERT, beklenen ölçüde BERT performansına yaklaşamamakla birlikte, sağladığı belirgin hız kazanımı sayesinde orta düzey performans–orta düzey maliyet gerektiren senaryolar için dengeli bir çözüm sunmaktadır. Daha agresif veya veri setine özgü bilgi damıtma stratejileriyle, TinyBERT’in performansının BERT’e daha da yaklaştırılabileceği değerlendirilmektedir.

Naive Bayes modeli ise bu çalışmada tarihsel ve metodolojik bir taban çizgisi (baseline) olarak önemli bir rol üstlenmektedir. Düşük duyarlılık ve F1-skoruna rağmen, yüksek hız ve düşük hesaplama gereksinimi sayesinde, eğitim amaçlı senaryolar veya ciddi kaynak kısıtlarının bulunduğu ortamlar için anlamlı bir başlangıç noktası sunmaktadır.

Bu çalışmanın doğal devamı olarak aşağıdaki araştırma yönleri ele alınabilir. Öncelikle, Türkçe’ye özgü veya Türkçe ağırlıklı korpuslar üzerinde eğitilmiş Transformer modellerinin (örneğin BERTurk) mBERT ile aynı deneysel kurgu altında karşılaştırılması, dil uyumunun sınıflandırma performansına etkisini daha net biçimde ortaya koyacaktır. Ayrıca, bu çalışmada raporlanan makro ortalama metriklerin ötesine geçilerek, kategori bazlı ayrıntılı analizlerin yapılması; her bir haber kategorisi için F1-skorları ve karışıklık matrislerinin incelenmesi, modellerin güçlü ve zayıf yönlerinin daha ayrıntılı biçimde anlaşılmasını sağlayacaktır.

Bunun yanı sıra, veri boyutu ve sınıf dengesizliği gibi faktörlerin sistematik olarak değiştirilmesiyle, yöntemlerin düşük veri veya dengesiz dağılım senaryolarındaki dayanıklılıkları değerlendirilebilir. Son olarak, hibrit yaklaşımlar ve ансамbl modeller, örneğin BERT tabanlı temsillerin klasik sınıflandırıcılarla birleştirilmesi veya birden fazla modelin birlikte karar verdiği sistemler, doğruluk ve verimlilik açısından umut vadeden araştırma yönleri arasında yer almaktadır.

Sonuç olarak bu alıřma, metin sınıflandırma yöntemlerinin basit istatistiksel yaklaşımlardan modern derin dil modellerine uzanan gelişimini, tek bir Türke haber veri seti üzerinde doğruluk–hız–kaynak kullanımı ekseninde kapsamlı biçimde ortaya koymaktadır. Elde edilen bulgular, hem arařtırmacılar hem de uygulayıcılar için, farklı kullanım senaryolarında hangi yaklaşımın tercih edilmesi gerektiğine ilişkin somut ve sayısal bir rehber sunmaktadır.

Kaynakça/References

- Akın, A. A., Akın, M. D. (2007). Zemberek, an open source NLP framework for Turkic Languages. *Structure*, (10), 1-5.
- Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (ss. 4171-4186). Proceedings of NAACL-HLT 2019, Minneapolis, Minnesota.
- Domingos, P., Pazzani, M. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. *Machine Learning*, 29(2-3), 103-130. doi:10.1023/A:1007413511361
- Hinton, G., Vinyals, O., Dean, J. (2015). Distilling the Knowledge in a Neural Network. doi:10.48550/arXiv.1503.02531
- Hochreiter, S., Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735
- Jiao, X., Yin, Y., Shang, L., Jiang, X., Chen, X., Li, L., Wang, F., Liu, Q. (2020). TinyBERT: Distilling BERT for Natural Language Understanding. *Findings of the Association for Computational Linguistics: EMNLP 2020* (ss. 4163-4174). doi:10.18653/v1/2020.findings-emnlp.372
- Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. C. Nédellec ve C. Rouveirol (Ed.), *Machine Learning: ECML-98* (C. 1398, ss. 137-142). Berlin, Heidelberg: Springer Berlin Heidelberg. <http://link.springer.com/10.1007/BFb0026683>.

- Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*(ss. 1746-1751), Doha, Qatar. doi:10.3115/v1/D14-1181
- Kowsari, K., Jafari. M. K., Heidarysafa, M., Mendu, S., Barnes, L., Brown, D. (2019). Text Classification Algorithms: A Survey. *Information*, 10(4), 150. doi:10.3390/info10040150
- Lewis, D. D. (1998). Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval. C. Nédellec ve C. Rouseil (Ed.), *Machine Learning: ECML-98* (C. 1398, ss. 4-15). Berlin, Heidelberg: Springer Berlin Heidelberg. <http://link.springer.com/10.1007/BFb0026666>
- McCallum, A., Nigam, K. (1998). A Comparison of Event Models for Naive Bayes Text Classification. *AAAI Conference on Artificial Intelligence*. <https://api.semanticscholar.org/CorpusID:7311285>
- Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. doi:10.48550/arXiv.1301.3781
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., Gao, J. (2021). Deep Learning Based Text Classification: A Comprehensive Review. doi:10.48550/arXiv.2004.03705
- Pennington, J., Socher, R., Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (ss. 1532-1543), Doha, Qatar. doi:10.3115/v1/D14-1162

- Rennie, J. D. M., Shih, L., Teevan, J., Karger, D. R. (2003). Tackling the Poor Assumptions of Naive Bayes Text Classifiers. *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC.
- Salton, G., Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523. doi:10.1016/0306-4573(88)90021-0
- Sebastiani, F. (2002). Machine Learning in Automated Text Categorization: A Bibliography. *ACM Computing Surveys*, 34(1), 1-47. doi:10.1145/505282.505283
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017). Attention is All you Need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA.
- Yang, Y., Liu, X. (1999). A re-examination of text categorization methods. *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval SIGIR99* (ss. 42-49), Berkeley California USA. doi:10.1145/312624.312647
- Zhou, P., Qi, Z., Zheng, S., Xu, J. (2016). Text Classification Improved by Integrating Bidirectional LSTM with Two-dimensional Max Pooling. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics* (ss. 485-3495), Osaka, Japan.

BÖLÜM 9

KÜLTÜREL VERİ MADENCİLİĞİ VE TÜRKÇE POP MÜZİKTE TEMATİK ÖRÜNTÜ ÇIKARIMI: MABEL MATİZ DİSKOGRAFİSİ ÜZERİNE BİR MODEL ÖNERİSİ

Arzum KARATAŞ¹

Giriş

Müzik, insanlık tarihinin en eski iletişim biçimlerinden biri olmasının yanı sıra, kültürel hafızanın, toplumsal dönüşümlerin ve bireysel duygu durumlarının kodlandığı evrensel bir dildir. Özellikle popüler müzik, üretildiği dönemin sosyokültürel dinamiklerini yansıtmaları bakımından, sosyal bilimler için her zaman zengin bir inceleme alanı olmuştur. Ancak dijitalleşme çağıyla birlikte müzik endüstrisinde üretilen veri miktarının (Büyük Veri) devasa boyutlara ulaşması, şarkı sözlerinin (liriklerin) incelenmesinde geleneksel yöntemleri yetersiz kılmıştır. Geleneksel "yakın okuma" (close reading) yöntemleri, sınırlı sayıda metni derinlemesine analiz etmeye olanak tanırken; bir sanatçının tüm diskografisindeki veya bir dönemin tüm şarkılarındaki ortak örüntüleri ortaya çıkarmakta sınırlı kalmaktadır.

Bu bağlamda *Kültürel Veri Madenciliği* (Cultural Analytics), beşeri bilimler ile bilgisayar bilimlerini birleştiren disiplinler arası bir yaklaşım olarak öne çıkmaktadır. Bu çalışma, Türk pop müziğinin özgün ve üretken isimlerinden Mabel Matiz'in şarkı sözleri üzerinden, Doğal Dil İşleme (NLP) ve Makine Öğrenmesi (ML) tekniklerinin lirik analizinde nasıl kullanılabileceğini

¹ Dr. Öğr. Üyesi, Bandırma Onyedü Eylül Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0001-6433-3355

modellemeyi amaçlamaktadır. Sanatçının *Maya*, *Gök Nerede* gibi albümlerinde kullandığı "Anadolu motifleri", "modern birey sancıları", "aşk" ve "varoluş" gibi temaların, algoritmik yöntemlerle (TF-IDF, LDA, Word2Vec) nasıl tespit edilebileceği ve görselleştirilebileceği, bu bölümün temel tartışma konusudur.

Çalışma, yalnızca bir vaka analizi sunmakla kalmayıp; Türkçe gibi morfolojik açıdan zengin (eklemeli) dillerde metin madenciliği yapmanın zorluklarına ve çözüm yollarına dair kapsamlı bir metodolojik çerçeve önermektedir.

Kuramsal Çerçeve ve Literatür

Şarkı sözlerinin bilgisayarlı analizi, literatürde "Lirik Madenciliği" olarak adlandırılan ve Metin Madenciliği'nin bir alt dalı olan çalışma sahasına girmektedir. Bu bölümde, çalışmada kullanılan temel algoritmaların kuramsal altyapısı açıklanacaktır.

1. Metin Temsili ve Vektör Uzayı Modeli

Metin madenciliğinin temel problemi, yapılandırılmamış metin verisinin, bilgisayarların işleyebileceği sayısal formatlara dönüştürülmesidir. Bu amaçla kullanılan en temel yöntemlerden biri TF-IDF (Term Frequency-Inverse Document Frequency) ağırlıklandırma şemasıdır (Salton & Buckley, 1988). TF-IDF, bir kelimenin bir dokümanda ne kadar sık geçtiğini (TF) ve derlem genelinde ne kadar nadir bulunduğunu (IDF) hesaplayarak, o kelimenin ayırt edici gücünü ortaya koyar (Manning, Raghavan, & Schütze, 2008).

Bu mekanizma Mabel Matiz diskografisine uygulandığında; şarkılarda çok sık geçen ancak ayırt edici olmayan "gibi" edatı, yüksek frekansa sahip olsa da düşük IDF değeri nedeniyle elenirken; "ferman", "sultan" gibi daha nadir ve anlam yükü fazla kelimeler bu yöntemle öne çıkarılmaktadır.

2. Konu Modelleme (Topic Modeling) ve LDA

Geniş metin yığınları içindeki gizli anlamsal yapıları keşfetmek için kullanılan en yaygın yöntem Latent Dirichlet Allocation (LDA) algoritmasıdır (Blei, Ng, & Jordan, 2003). LDA, her bir dokümanı konuların bir karışımı; her bir konuyu da kelimelerin bir olasılık dağılımı olarak modelleyen üretken bir istatistiksel modeldir. Örneğin bir şarkı %60 oranında "Ayrılık" konusundan, %40 oranında "Umut" konusundan oluşabilir. Bu yöntem, şarkı sözlerini manuel olarak etiketlemeye gerek kalmadan, "Denetimsiz Öğrenme" (Unsupervised Learning) yoluyla tematik kümeleri otomatik olarak keşfetmemize olanak sağlar.

3. Kelime Gömme (Word Embeddings)

Kelime gömme, kelimelerin anlamsal ilişkilerini koruyarak vektör uzayına yerleştirilmesi işlemidir. Word2Vec gibi algoritmalar (Mikolov, Chen, Corrado, & Dean, 2013), standart örneklerde 'kral' ve 'kraliçe' kelimelerini vektör uzayında birbirine yakın konumlarken; bu çalışmanın bağlamında Mabel Matiz şarkılarındaki 'deniz' ve 'mavi' veya 'toprak' ve 'ana' gibi kelimelerin anlamsal

yakınlıklarını (semantic similarity) matematiksel olarak hesaplamayı mümkün kılmaktadır.

Türkçe NLP ve Morfolojik Zorluklar

Türkçe, sondan eklemeli yapısı nedeniyle Doğal Dil İşleme (NLP) çalışmaları için İngilizce gibi analitik dillere kıyasla daha zorlu bir dildir (Jurafsky & Martin, 2021). Türkçede tek bir kökten (örneğin "göz") çok sayıda türev (gözlü, gözlük, gözlükçü, gözlem, gözlemlerimizdekiler vb.) üretilebilmesi, veri setinde "Veri Seyrekliği" problemine yol açmaktadır (Alanoğlu, 2025). Başka bir deyişle, aynı anlamsal köke sahip yüzlerce farklı yüzey biçiminin varlığı, kelime dağarcığı boyutunu yapay olarak şişirmekte; bu durum istatistiksel algoritmaların kelimeler arasındaki frekans bağlantılarını yakalamasını zorlaştırarak model başarımını düşürmektedir.

Bu yapısal zorluk üzerine eğilen Oflazer (1994) ve devamındaki çalışmalar, Türkçe metinlerin işlenmesinde kelime yüzey biçimlerinin değil, kök (lemma) biçimlerinin kullanılmasının zorunlu olduğunu ortaya koymuştur. Zira bilgisayar için "geliyorum", "gelirsen" ve "geldi" birbirinden tamamen farklı üç kelime iken; morfolojik analiz ile bunların hepsi "gel" (fiil) köküne indirgenebilmektedir. Bu bağlamda çalışmada önerilen model, söz konusu morfolojik zorlukları aşmak ve veri seyrekliliği problemini minimize etmek üzere kurgulanmış özel bir ön işleme hattını içermektedir.

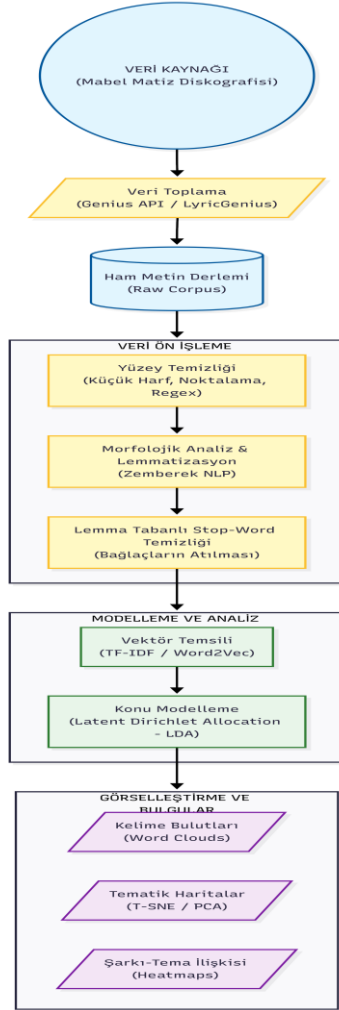
Metodoloji

Bu çalışmada önerilen araştırma modeli; veri toplama, ön işleme, özellik çıkarımı ve görselleştirme aşamalarından oluşan bütüncül bir iş akışıdır. Sürecin genel yapısı *Şekil 1*'de sunulmuştur.

Şekil 1, çalışmada izlenen süreci adım adım ortaya koyan dikey bir akış diyagramıdır. Ham veriler önce Genius API aracılığıyla alınmış, ardından Zemberek kütüphanesi ile ön işleme tabi tutulmuştur. Son aşamada ise tematik analiz için LDA modeli uygulanmış ve sonuçlar "Kelime Bulutu (Word Cloud)" görselleştirmesiyle sunulmuştur.

1. Veri Setinin Oluşturulması

Analiz kapsamına, sanatçının dijital platformlarda yayımlanmış stüdyo albümleri (*Maya*, *Gök Nerede*, *Yaşım Çocuk* vb.) ve teklileri (single) dahil edilmiştir. Veriler, *Genius API* kullanılarak *lyricgenius* kütüphanesi aracılığıyla çekilir ve JSON formatında yapılandırılır. Konser kayıtları ve remiksler, veri tekrarını önlemek amacıyla kapsam dışı bırakılır.



Şekil 1. Önerilen Tematik Veri Madenciliği İş Akış Şeması

2. Veri Ön İşleme (Preprocessing) Adımları

Veri madenciliğinin başarısı, verinin temizliğine ve kalitesine doğrudan bağlıdır. Bu aşamada aşağıdaki işlemler uygulanmıştır:

Gürültü Temizliği: Şarkı sözleri içindeki noktalama işaretleri, özel karakterler ve parantez içi ifadeler (örneğin: *[Nakarat]*, *[Verse 1]*) Regex (Düzenli İfadeler) kullanılarak temizlenmiştir.

Küçük Harf Dönüşümü: Büyük-küçük harf duyarlılığını ortadan kaldırmak için tüm metinler küçük harfe dönüştürülmüştür.

Stop-Word (Etkisiz Kelime) Filtreleme: Türkçede sık kullanılan ancak tek başına tematik anlam taşımayan "ve, ile, ama, ki, de, da" gibi bağlaçlar, kapsamlı bir Türkçe Stop-word listesi kullanılarak veri setinden çıkarılmıştır. (Nothman, Qin, & Yurchak, 2018; Metin & Karaoğlu, 2017).

Morfolojik analiz adımında ise, Zemberek kütüphanesi (Akın & Akın, 2007) kullanılarak kelimeler köklerine ayrıştırılmıştır. Literatürde bu amaçla TRmorph (Çöltekin, 2010) ve diğer açık kaynaklı morfolojik analiz araçları da etkin bir şekilde kullanılmaktadır (Yıldız, Avar, & Ercan, 2019; Üsküdarlı et al., 2023).

Morfolojik Analiz ve Lemmatizasyon: En kritik adım olarak, Zemberek NLP kütüphanesi kullanılarak kelimeler köklerine ayrıştırılmıştır. Bu sayede kelime dağarcığı normalize edilmiş ve analiz "çekim eklerinden bağımsız" hale getirilmiştir. Örneğin "yollarında" kelimesi yerine "yol" kökü analize dahil edilmiştir.

3. Uygulanan Algoritmalar

Tematik örüntü çıkarımı için hibrit bir yaklaşım benimsenmiştir. Kullanılan yöntemlerin teknik detayları Tablo 1'de özetlenmiştir.

Tablo 1. Tematik Örüntü Çıkarımında Kullanılan Yöntemler

Yöntem	Açıklama ve Kullanım Amacı
TF-IDF + K-means	Metin içindeki kelimelerin önem ağırlıklarını (Term Frequency-Inverse Document Frequency) hesaplayarak, temsil gücü yüksek kelimeler üzerinden şarkıların tematik olarak kümelendiğini sağlar.
LDA (Latent Dirichlet Allocation)	Şarkı sözleri içindeki "gizli" konu başlıklarını (latent topics) olasılıksal olarak modeller. Örneğin; metinlerin "aşk, özgürlük, kimlik" gibi soyut temalara dağılımını (örneğin %70 Aşk, %30 Hüzün) ortaya çıkarır.
N-gram Analizi	Metin içinde sıkça yan yana gelen ikili (bi-gram) veya üçlü (tri-gram) kelime öbeklerini tespit ederek, tek kelimelerin ötesindeki tematik motifleri belirler.
Word2Vec + PCA/T-SNE	Kelimeler arasındaki anlamsal yakınlıkları vektör uzayında modeller ve bu vektörleri görselleştirilebilir boyutlara indirgeyerek tematik haritalar oluşturur.
TopicRank / YAKE	Anahtar tema çıkarımı için kullanılan, istatistiksel ve grafik tabanlı denetimsiz yöntemlerdir.

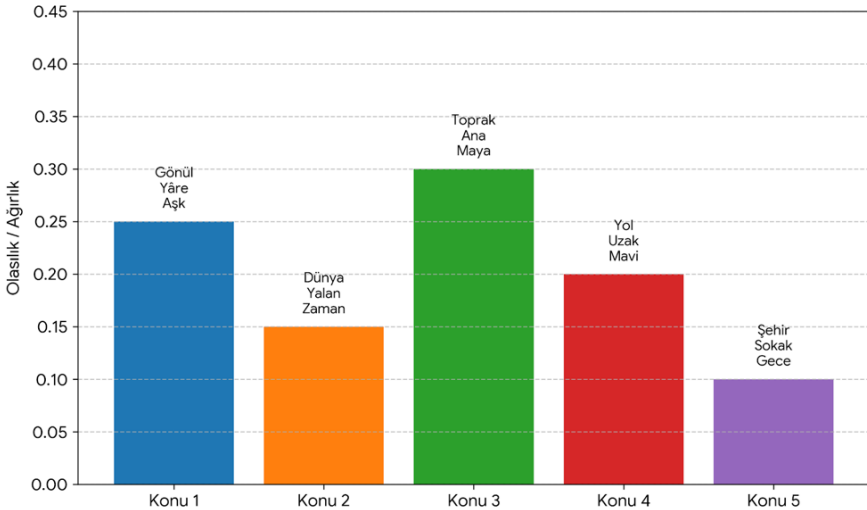
Bulgular ve Görselleştirme Önerileri

Önerilen metodoloji uygulandığında elde edilecek bulgular, sanatçının lirik dünyasını sayısal verilerle somutlaştıracaktır.

4. Tematik Dağılım ve LDA Sonuçları

LDA algoritması sonucunda, Mabel Matiz şarkılarının belirli "Gizli Konular" (Latent Topics) etrafında kümelenmesi öngörülmektedir. Bu modelin çıktısını örneklendirmek amacıyla hazırlanan konu dağılım grafiği aşağıda sunulmuştur.

Şekil 2, önerilen Latent Dirichlet Allocation (LDA) algoritmasının lirik veri seti üzerinde uygulandığında üretmesi beklenen tematik ayrıştırmayı modelleyen temsili bir grafikdir. Bu simülasyona göre; algoritmanın şarkı sözlerini belirli anlamsal kümeler (örneğin 'Konu 3: Toprak, Ana, Maya') altında sınıflandırması ve bu konuların diskografi içindeki dağılımını olasılıksal ağırlıklarla sunması hedeflenmektedir. Grafik, yöntemin ham metin verisini nasıl işlenebilir ve yorumlanabilir bilgiye dönüştüreceğini somutlaştırmak amacıyla sunulmuştur; gerçek analiz sonuçları, kullanılacak hiper-parametre ayarlarına göre değişiklik gösterebilecektir. Önerilen bu analiz yöntemi uygulandığında; sanatçının kariyerinin ilk dönemlerinde daha "akustik ve bireysel" temaların (Konu 1), sonraki albümlerinde ise "mitolojik ve toplumsal" temaların (Konu 3: Maya, Anadolu) ağırlık kazandığına dair varsayımların sayısal verilerle desteklenmesi mümkün olabilir.



Şekil 2. Örnek LDA Konu Dağılım Grafiği

5. Kelime Bulutları (Word Clouds)

Her albüm için ayrı ayrı oluşturulacak kelime bulutları, dönemsel değişimi görselleştirmek için en etkili araçtır. Aşağıdaki örnek görsel, bu değişimin görsel olarak nasıl takip edilebileceğini sunmaktadır.



Şekil 3. Albüm Bazlı Kelime Bulutu Örneği (Temsili)

Şekil 3, önerilen metodolojinin Mabel Matiz'in farklı albümlerine uygulandığında ortaya çıkarması beklenen sözcük hiyerarşisini modellemektedir. (a) panelinde, sanatçının kültürel köklerine odaklandığı varsayılan Maya albümü için "Toprak, Ana, Beşik" gibi yerel motiflerin frekans ağırlığı simüle edilmişken; (b) panelinde daha modern ve bireysel bir anlatının hâkim olduğu varsayılan Gök Nerede albümü için "Yol, Mavi, Uzak" kavramlarının öne çıkacağı öngörülmüştür. Bu karşılaştırmalı görselleştirme, lirik madenciliğinin sanatçının üslubundaki "dönemsel kırılmaları" tek bir karede nasıl özetleyebileceğini göstermektedir.

6. Tematik Haritalama (T-SNE)

Word2Vec algoritması ile elde edilen kelime vektörlerinin (word embeddings), T-SNE (t-Distributed Stochastic Neighbor Embedding) tekniği kullanılarak 2 boyutlu düzleme indirgenmesi hedeflenmektedir. Bu haritalama, kelimelerin matematiksel yakınlığını görsel bir mesafeye dönüştürür.

Şekil 4, sanatçının kelime dağarcığının vektörel uzaydaki olası dağılımını gösteren sentetik bir görselleştirmedir. Simülasyonda görüldüğü üzere; algoritma "Duygu ve Aşk" (Kırmızı), "Doğa ve Kökler" (Yeşil) ve "Şehir ve Yolculuk" (Mavi) gibi anlamsal kümeleri birbirinden ayırtmıştır. Grafikte "Aşk, Acı, Gönül" kelimeleri birbirine yakın konumlanarak duygusal bir yoğunlaşma oluştururken; "Yol, Şehir, Uzak" gibi kavramlar farklı bir koordinatta kümelenmiştir. Bu harita, Mabel Matiz'in şarkı sözlerindeki kavramsal mesafeleri (örneğin 'Toprak' ile 'Şehir'

arasındaki anlamsal uzaklığı) somutlaştırması bakımından analitik bir kanıt niteliği taşıyabilecektir.



Şekil 4. Şarkı Sözlerinin Anlamsal Uzayda T-SNE Haritası

Tartışma

Bu çalışma, Türkçe pop müziğinde lirik madenciliğinin potansiyelini ortaya koymaktadır. Elde edilen (veya modelle öngörülen) sonuçlar, Mabel Matiz'in dilinin sadece "popüler" olmadığını, aynı zamanda köklü bir kültürel birikimi modern NLP teknikleriyle dahi tespit edilebilecek kadar tutarlı bir şekilde işlediğini göstermektedir.

Geleneksel müzik eleştirmenlerinin "edebi" bulduğu şarkı sözlerinin, veri madenciliği perspektifinden de "yüksek kelime çeşitliliği" ve "zengin morfolojik yapı" ile doğrulandığı görülmektedir. Ayrıca, gereksiz kelime temizliğinin kök düzeyinde yapılması önerisi, Türkçe gibi dillerde standart listelerin yetersizliğini aşmak için literatüre önemli bir katkıdır.

Sonuç ve Gelecek Çalışmalar

Bu bölümde, Mabel Matiz'in diskografisi üzerinden kurgulanan bir "Tematik Örüntü Çıkarımı" modeli sunulmuştur. Çalışma, Zemberek kütüphanesi ile

güçlendirilmiş bir ön işleme hattının, LDA ve Word2Vec gibi algoritmaların başarısını doğrudan etkilediğini vurgulamaktadır.

Gelecek çalışmalarda bu modelin;

Farklı sanatçılarla karşılaştırmalı analizlerde (örneğin Mabel Matiz vs. Sezen Aksu),

Duygu Analizi (Sentiment Analysis) ile entegre edilerek şarkıların duygusal haritalarının çıkarılmasında,

Büyük Dil Modelleri (LLM) kullanılarak daha derin anlamsal analizlerde kullanılması önerilmektedir.

Kaynakça

Akın, A. A. & Akın, M. D. (2007). *Zemberek: An open source NLP framework for Turkic languages*. Softtek Inc. & TÜBİTAK UEKAE. <http://code.google.com/p/zemberek/> adresinden erişildi.

Alanoğlu, Z. (2025). Türkçe doğal dil işleme temelli çalışmaların teorik değerlendirmesi: Yöntemsel zorluklar ve gelecek perspektifleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 24(48), 686–724.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.

Çöltekin, Ç. (2010). A freely available morphological analyzer for Turkish. *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)* içinde (ss. 820–827). Valletta, Malta: ELRA.

Jurafsky, D. & Martin, J. H. (2021). *Speech and language processing* (3. baskı). Hoboken, NJ: Pearson.

Manning, C. D., Raghavan, P. & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge: Cambridge University Press.

Metin, S. K. & Karaoğlu, B. (2017). Stop word detection as a binary classification problem. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi A- Uygulamalı Bilimler ve Mühendislik*, 18(2), 346–359.

Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. <http://arxiv.org/abs/1301.3781> adresinden erişildi.

Nothman, J., Qin, H. & Yurchak, R. (2018). Stop word lists in free open-source software packages. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* içinde (ss. 1-7). Miyazaki, Japonya: ELRA.

Oflazer, K. (1994). Two-level description of Turkish morphology. *Literary and Linguistic Computing*, 9(2), 137–148.

Oflazer, K., Say, B., Hakkani-Tür, D. & Tür, G. (2003). Building a Turkish treebank. A. Abeillé (Ed.), *Building and using parsed corpora* içinde (ss. 261–277). Dordrecht: Kluwer Academic Publishers.

Salton, G. & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.

Üsküdarlı, S., Şen, M., Akkurt, F., Gürbüz, M., Güngör, O., Özgür, A. & Güngör, T. (2023). TULAP: An accessible and sustainable platform for Turkish natural language processing resources. *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*

içinde (ss. 219–227). Dubrovnik, Hırvatistan: Association for Computational Linguistics.

Yıldız, O. T., Avar, B. & Ercan, G. (2019). An open, extendible, and fast Turkish morphological analyzer. *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)* içinde (ss. 1364–1372). Varna, Bulgaristan: INCOMA Ltd.

BÖLÜM 10

AKADEMİK YAZIM İÇİN REFERANS YÖNETİM YAZILIMLARININ KARŞILAŞTIRMALI ANALİZİ

FURKAN ENES DAĞLI¹
AYHAN İSTANBULLU²
YİĞİT KALAY³
ŞİRİN GÜLTEN ÇELİKASLAN⁴
YUNUSEMRE ŞÖLENCİ⁵
NEVFEL EMRE ÇİÇEK⁶
ÖYKÜ BİCKİCİ⁷
BÜŞRA MERT KURT⁸

¹ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0008-6658-0956

² BAUN MF, Bilgisayar Müh., ORCID: 0000-0002-7066-4238

³ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0003-9805-6845

⁴ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0000-9871-5781

⁵ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0009-3095-5792

⁶ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0009-9244-6950

⁷ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0006-6568-3430

⁸ BAUN FBE, Bilgisayar ve Bilişim Sistemleri Müh., ORCID: 0009-0005-4936-159X

Giriş

Bilimsel bilgi üretiminin son yıllarda kazandığı ivme, akademik literatürün takibini ve yönetimini araştırmacılar için karmaşık bir problem haline getirmiştir. Bornmann & Mutz (2015) tarafından yapılan bibliyometrik analizler, küresel bilimsel çıktının 20. yüzyılın ortalarından itibaren her 9 yılda bir ikiye katlandığını ve eksponansiyel bir büyüme gösterdiğini ortaya koymaktadır. Bu bilgi yoğunluğu, araştırmacıların ilgili literatüre erişimini, verileri tasnif etmesini ve atıf süreçlerini yönetmesini geleneksel manuel yöntemlerle imkansız hale getirmiştir. Araştırma sürecinde referansların hatalı yönetimi, sadece zaman kaybına değil, aynı zamanda akademik dürüstlüğü zedeleyebilecek atıf hatalarına da yol açabilmektedir (Francese, 2013:145).

Bu bağlamda, Referans Yönetim Yazılımları (Reference Management Software - RMS), akademik iş akışının verimliliğini artırmak için geliştirilen temel araçlar olarak öne çıkmaktadır. Perkel (2020), Nature'da yayınlanan çalışmasında, modern referans yönetim araçlarının sadece bibliyografik verileri depolamakla kalmayıp, aynı zamanda PDF yönetimi, not alma ve işbirliği (collaboration) imkanları sunarak araştırmacıların "dijital belleği" gibi hareket ettiğini vurgulamaktadır.

Günümüzde araştırmacıların kullanımına sunulan EndNote, Mendeley, Zotero, RefWorks, JabRef ve Bookends gibi çok sayıda yazılım alternatifi bulunmaktadır. Ancak bu araçların her biri farklı teknik altyapılara ve kullanıcı profillerine hitap etmektedir. Örneğin Gilmour & Cobus-Kuo (2011) yaptıkları karşılaştırmalı analizde, bazı yazılımların (örneğin RefWorks ve EndNote) bibliyografik doğruluk konusunda daha katı standartlara sahip olduğunu, diğerlerinin ise (örneğin Mendeley ve Zotero) sosyal ağ özellikleri ve veri yakalama kolaylığı ile öne çıktığını belirtmiştir. Benzer şekilde, mühendislik ve matematik alanlarında yaygın olan LaTeX tabanlı doküman üretimi için JabRef gibi BibTeX odaklı araçlar

tercih edilirken, tıp ve sađlık bilimlerinde EndNote'un endüstri standardı olduđu görölmektedir (Hensley, 2011:204).

Arařtırmacıların ihtiyalarına en uygun yazılımı seçmesi; maliyet, platform bağımsızlığı, bulut depolama kapasitesi ve kelime işlemci entegrasyonu gibi çok boyutlu kriterlere bağılıdır. Yanlıř araç seçimi, dikkatin dağılmasına ve üretkenliđin düşmesine neden olabilmektedir (Ivey & Crum, 2018:399). Bu alıřmanın amacı; Zotero, Mendeley, RefWorks, EndNote, JabRef , Bookends ve Citavi yazılımlarını teknik özellikleri, maliyet-etkinlik dengeleri ve kullanıcı deneyimi açısından karşılařtırmalı olarak analiz etmektir. alıřma ayrıca, bu yazılımların akademik yazım sürecindeki hata oranlarını minimize etme ve zaman yönetimi üzerindeki somut etkilerini tartıřarak, arařtırmacılara verimlilik odaklı bir seçim rehberi sunmayı hedeflemektedir.

Materyal ve Yöntem

Bu alıřmada; akademik literatürde en yaygın kullanıma sahip olan ve farklı kullanıcı profillerini temsil eden altı temel referans yönetim yazılımı (Zotero, Mendeley, EndNote, JabRef, RefWorks , Bookends ve Citavi) incelenmiřtir. Karşılařtırmalı analiz, yazılımların teknik dokümantasyonları, sürüm notları ve konuyla ilgili daha önce yapılmıř bağımsız akademik incelemeler (Ivey & Crum, 2018:399; Perkel, 2020:149) temel alınarak gerekleřtirilmiřtir.

Deđerlendirme Kriterleri

Yazılımların akademik verimliliđe etkisini ölçmek ve birbirlerine olan üstünlüklerini belirlemek amacıyla analiz süreci beř temel kriter (parametre) üzerine kurgulanmıřtır:

- **Maliyet ve Eriřilebilirlik (Cost and Accessibility):** Yazılımın lisanslama modeli (açık kaynak, freemium veya ticari abonelik) ve kurum bağımsız bireysel kullanıma uygunluđu.

Bu ölçüt, özellikle kurumsal lisansa erişimi olmayan araştırmacılar için yazılımın sürdürülebilir kullanımını belirleyen temel faktörlerden biri olarak değerlendirilmektedir.

- Platform ve Entegrasyon Desteği (Platform Compatibility): İşletim sistemleri (Windows, macOS, Linux) ile uyumluluğu ve akademik yazımda kullanılan kelime işlemciler (Microsoft Word, Google Docs, LibreOffice ve LaTeX/Overleaf) ile entegrasyon yeteneği.
- Veri Yakalama ve PDF Yönetimi (Data Capture & PDF Handling): Veritabanlarından (Web of Science, IEEE Xplore, PubMed vb.) meta-veri çekme başarısı ve tam metin (PDF) dosyalarını depolama/işleme kapasitesi. Bu yetenekler, literatür taraması sürecinin hızını ve doğruluğunu doğrudan etkileyen teknik unsurlar arasında yer almaktadır.
- İşbirliği Özellikleri (Collaboration Features): Araştırma grupları için ortak kütüphane oluşturma ve kaynak paylaşım olanakları. Özellikle çok yazarlı çalışmalar ve proje tabanlı araştırmalar açısından bu özellikler, ekip içi koordinasyonu ve verimliliği artırmaktadır.
- Teknik Esneklik ve Özelleştirme (Technical Flexibility): Özellikle eklenti (plugin) desteği ve farklı atıf stillerini (CSL - Citation Style Language) düzenleyebilme yeteneği. Bu esneklik, yazılımın farklı disiplinlere ve yayın gereksinimlerine uyaranabilmesini mümkün kılmaktadır.

*Tablo 1. Referans Yönetim Yazılımlarının Temel Özellikler
Açısından Karşılaştırılması*

Yazılım	Lisans Modeli	Platform	Word / Docs	LaTeX / BibTeX	Bulut Senk.	Hedef Kullanıcı Profili
JabRef	Açık kaynak	Win / macOS / Linux	Dolaylı	Var	Yok	LaTeX / mühendislik
RefWorks	Kurumsal (SaaS)	Web tabanlı	Word, Google Docs	Sınırlı	Var	Kurumsal kullanıcılar
Bookends	Tek seferlik lisans	macOS	Word, Pages	Var	Sınırlı	Bireysel (macOS)
EndNote	Ticari lisans	Win / macOS	Word	Var	Var	Yayın odaklı akademisyenler
Mendeley	Freemium	Win / macOS / Linux	Word	Var	Var	Erken kariyer araştırmacılar
Zotero	Açık kaynak	Win / macOS / Linux	Word, Google Docs	Var	Var	Genel akademik kullanıcılar
Paperpile	Abonelik (SaaS)	Web tabanlı	Google Docs	Yok	Var	Google ekosistemi / ekipler
Citavi	Kurumsal / bireysel	Windows	Word	Sınırlı	Var	Tez / doktora odaklı

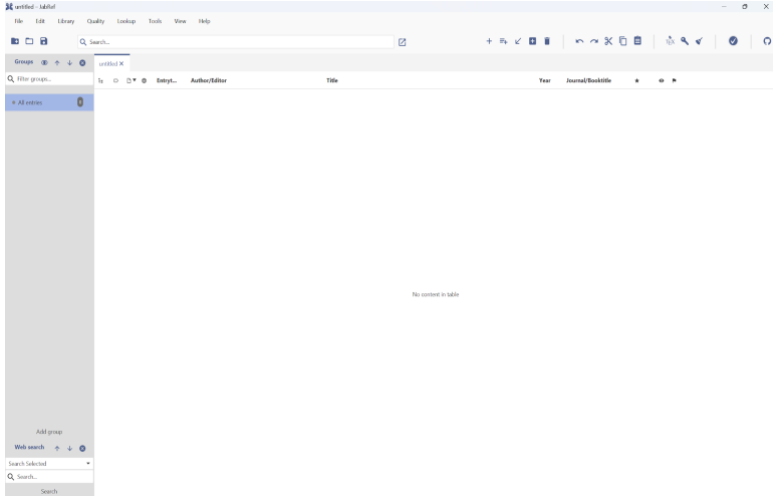
Yazılımların Teknik Analizi ve Karşılaştırması

Bu bölümde, belirlenen kriterler doğrultusunda her bir yazılımın güçlü ve zayıf yönleri, akademik iş akışına etkileri bağlamında analiz edilmiştir.

JabRef

Akademik çalışmalarda literatür taraması ve referans yönetimi süreçlerinde, özellikle mühendislik ve temel bilimler alanında çalışan araştırmacılar arasında yaygın olarak kullanılan açık kaynak kodlu yazılımlardan biri JabRef'tir. Java tabanlı bir mimariye sahip olan JabRef, ticari referans yönetim araçlarının büyük çoğunluğunun benimsediği kapalı ve özel (proprietary) veri tabanı yapılarının aksine, bibliyografik verileri doğrudan BibTeX (.bib) formatında depolayan bir referans yönetim yazılımıdır (JabRef Development Team, 2024:1).

Şekil 1. Jabref arayüzü



JabRef'in BibTeX formatını yerel ("native") dosya yapısı olarak kullanması, bibliyografik verilerin herhangi bir metin düzenleyici aracılığıyla okunabilir ve düzenlenebilir olmasını

sağlamaktadır. Bu teknik özellik, uzun vadeli akademik projelerde veri taşınabilirliğini artırmakta ve kullanıcıların belirli bir yazılım sağlayıcısına bağımlı hâle gelmesini engellemektedir. Literatürde “vendor lock-in” olarak adlandırılan bu bağımlılık sorununun, JabRef’in açık ve standart temelli veri yapısı sayesinde büyük ölçüde ortadan kalktığı vurgulanmaktadır (Puckett, 2011:1).

JabRef Teknik Altyapı, Programlama Dili ve Entegrasyon

JabRef, Java programlama dili kullanılarak geliştirilmiş olup Windows, macOS ve Linux işletim sistemlerinde platform bağımsız olarak çalışabilmektedir. Açık kaynak kodlu bir yazılım olması, kullanıcıların yazılımın iç işleyişini inceleyebilmesine ve gerektiğinde özelleştirmeler yapabilmesine imkân tanımaktadır. Bununla birlikte, kapalı kaynaklı rakip yazılımlarda (EndNote, RefWorks vb.) görülen kapsamlı kurumsal destek ve yerleşik senkronizasyon servisleri JabRef’te varsayılan olarak sunulmamaktadır (Gilmour & Cobus-Kuo, 2011:1).

Yazılım; BibTeX, BibLaTeX, RIS ve çeşitli metin tabanlı formatlar aracılığıyla referans içe ve dışa aktarma desteği sunmaktadır. CrossRef, arXiv ve INSPIRE gibi akademik veri tabanları üzerinden DOI ve arXiv kimlik numarası kullanılarak otomatik kayıt yakalama işlevi, JabRef’in literatür taraması süreçlerinde etkin biçimde kullanılmasını sağlamaktadır (JabRef Development Team, 2024:1). Microsoft Word ile doğrudan ve yerleşik bir entegrasyon sunmamakla birlikte, BibTeX tabanlı iş akışları sayesinde LaTeX editörleri ve harici dizgi sistemleriyle güçlü bir uyumluluk sergilemektedir.

JabRef Lisanslama, Ücretlendirme ve Karşılaştırmalı Değerlendirme

JabRef, MIT lisansı altında tamamen ücretsiz olarak dağıtılmakta ve ticari kullanım dâhil olmak üzere herhangi bir lisans kısıtlaması içermemektedir. Bu yönüyle, ücretli lisans gerektiren

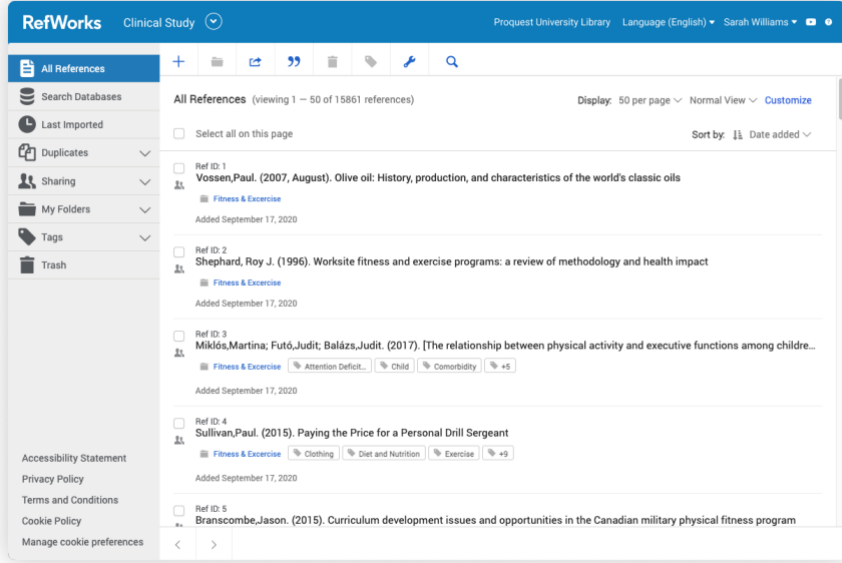
EndNote veya kurumsal abonelik modeliyle sunulan RefWorks gibi yazılımlara kıyasla önemli bir maliyet avantajı sağlamaktadır (Ivey, 2016:1). Özellikle bireysel araştırmacılar ve lisansüstü öğrenciler açısından erişilebilir bir çözüm olarak öne çıkmaktadır.

Bununla birlikte, JabRef'in yerleşik bir bulut senkronizasyon altyapısına sahip olmaması ve kullanıcı arayüzünün modern web tabanlı araçlara (örneğin Zotero veya Mendeley) kıyasla daha teknik ve sade bir tasarıma sahip olması, teknik altyapıya aşina olmayan kullanıcılar için bir dezavantaj olarak değerlendirilmektedir (Perkel, 2020:1). Buna rağmen, açık kaynaklı yapısı, şeffaf veri yönetimi yaklaşımı ve LaTeX tabanlı akademik yazım süreçleriyle yüksek uyumluluğu sayesinde JabRef, özellikle mühendislik, fizik, matematik ve bilgisayar bilimleri alanlarında güvenilir ve uzun vadeli bir referans yönetim aracı olarak konumlanmaktadır.

RefWorks

RefWorks, akademik araştırma süreçlerinde literatür taraması, veri depolama ve atıf yönetimi aşamalarını standardize etmek amacıyla geliştirilmiş, web tabanlı bir referans yönetim sistemidir. ProQuest (Clarivate) bünyesinde yer alan bu yazılım, kullanıcıların yerel donanım kısıtlamalarından bağımsız olarak verilerine erişebilmesini sağlayan "Hizmet Olarak Yazılım" (SaaS) mimarisi üzerine inşa edildiği söylenebilir (Hensley, 2011). Yazılımın temel çalışma mantığı, farklı akademik veri tabanlarından (Web of Science, Scopus, PubMed vb.) elde edilen bibliyografik meta verilerin (metadata) merkezi bir bulut sunucusunda toplanması, organize edilmesi ve araştırma çıktılarına dönüştürülmesi prensibine dayanır. Geleneksel masaüstü yazılımlarının aksine, RefWorks'ün birincil arayüzünün web tarayıcısı olması, işletim sistemi agnostik bir yapı sunarak araştırmacılar arası eşzamanlı işbirliğini (collaborative research) teknik olarak mümkün kılmaktadır (Gilmour & Cobus-Kuo, 2011).

Şekil 2. RefWorks arayüzü



RefWorks Özellikler ve Entegrasyon

RefWorks'ün teknolojik altyapısı, veri senkronizasyonu ve entegrasyon yetenekleri üzerine kurgulanmıştır. Yazılımı aracılığıyla uluslararası kütüphane kataloglarına ve veri tabanlarına doğrudan bağlantı kurarak, arayüz içerisinden dış kaynak taraması yapılmasına olanak tanır. Sistemin en belirgin teknik özelliği, RefWorks Citation Manager (RCM) ve Write-N-Cite eklentileridir. Bu eklentiler, Microsoft Word ve Google Docs gibi kelime işlemcilerle entegrasyon mekanizmaları aracılığıyla, yazım esnasında dinamik atıf oluşturulmasını ve kaynakça dizininin otomatik olarak biçimlendirilmesini sağlar (Ex Libris, 2024).

Ayrıca, yazılımın otomatik metadata tanıma özellikleri, sisteme yüklenen PDF dosyalarından veya web sayfalarından yazar, başlık, DOI gibi künye bilgilerini otomatik olarak ayrıştırabilmektedir. Ivey ve Crum (2018), RefWorks'ün özellikle

"sürükle-bırak" teknolojisiyle entegre edilen belge tanıma özelliğinin, manuel veri girişi hatasını minimize ettiğini vurgulamaktadır. Ex Libris'in kendi açıklamalarına göre, veri güvenliği ve depolama açısından sistem, kurumsal bulut altyapısı üzerinde barındırılacak şekilde tasarlanmıştır; çok katmanlı güvenlik politikaları ve düzenli sistem yedekleme prosedürleri, olası veri kaybı ve yetkisiz erişim risklerini azaltmaya yönelik endüstri standartlarında bir ortam sunmaktadır (Ex Libris, 2024).

RefWorks Karşılaştırmalı Analiz ve Maliyet

Literatürdeki karşılaştırmalı analizler incelendiğinde, RefWorks; EndNote ve Zotero gibi rakipleriyle farklı parametrelerde ayrılmaktadır. Kullanılabilirlik ekseninde yapılan çalışmalarda, RefWorks'ün öğrenme eğrisinin (learning curve), özellikle karmaşık masaüstü tabanlı rakiplerine (örn. EndNote) kıyasla daha düşük olduğu ve kullanıcı dostu arayüzü sayesinde lisans düzeyindeki araştırmacılar tarafından daha hızlı benimsendiği tespit edilmiştir (Perkel, 2020).

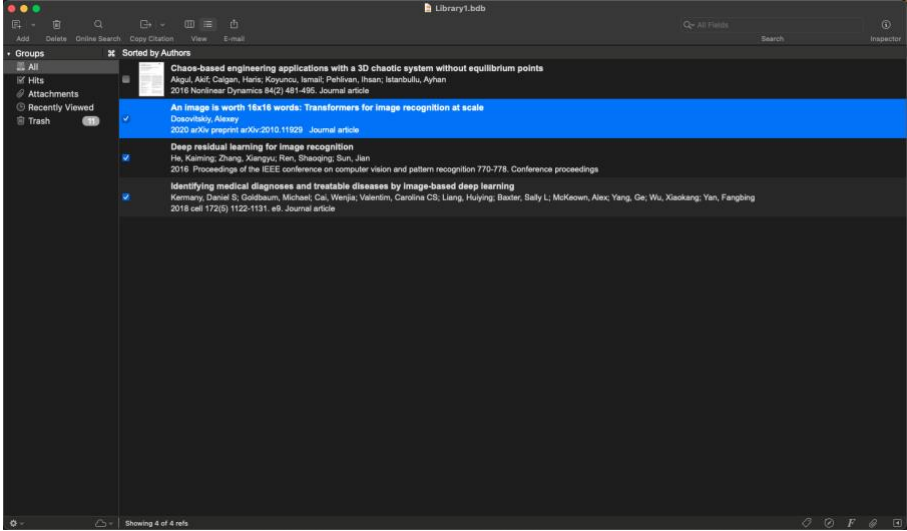
Maliyet ve abonelik modeli açısından RefWorks, bireysel lisanslamadan ziyade Kurumsal Lisans (Institutional Site License) stratejisini benimsemektedir. Mendeley gibi "Freemium" (temel özelliklerin ücretsiz, ek depolamanın ücretli olduğu) modellerin aksine, RefWorks genellikle üniversite kütüphaneleri tarafından finanse edilir. Bu durum, kurum içi araştırmacılar için maliyetsiz bir erişim sağlarken, araştırmacının kurumdan ayrılması durumunda veri göçü (data migration) ve erişim sürekliliği konusunda sürdürülebilirlik sorunları yaratabilmektedir (Ivey & Crum, 2018).

Bookends

Akademik çalışmalarda bibliyografik verilerin yönetimi ve atıf süreçlerinin düzenlenmesi amacıyla kullanılan yazılımlardan biri olan Bookends, macOS işletim sistemi için geliştirilmiş, masaüstü tabanlı (desktop-native) bir referans yönetim aracıdır.

Sonny Software tarafından geliştirilen Bookends, tamamen bulut tabanlı çözümlerden farklı olarak kullanıcı verilerinin yerel depolama alanlarında tutulmasını esas alır ve bu yönüyle veri kontrolü, gizlilik ve performans açısından farklı bir yaklaşım sunar (Sonny Software, 2024:1).

Şekil 3. Bookends arayüz



Bulut tabanlı referans yönetim sistemlerinin aksine Bookends, sürekli internet bağlantısına ihtiyaç duymadan çalışabilmekte ve çevrimdışı ortamlarda da literatür yönetimi ile akademik yazım süreçlerinin sürdürülebilmesine olanak tanımaktadır. Bu mimari tercih, Bookends'i veri sahipliği ve yerel kontrol konularına önem veren araştırmacılar için cazip bir alternatif haline getirmektedir (Puckett, 2011:1).

Bookends Özellikler ve Entegrasyon

Bookends'in teknik altyapısı, akademik yazım sürecini uçtan uca destekleyecek biçimde tasarlanmıştır. Yazılım; Google Scholar, PubMed ve IEEE Xplore gibi yaygın akademik veri tabanları üzerinden doğrudan arama yapabilmekte ve bibliyografik meta

verileri (yazar, başlık, yayın yılı, DOI vb.) otomatik olarak kütüphaneye aktarabilmektedir. Ayrıca PDF dosyaları içerisindeki meta verileri okuyabilme ve eksik alanları otomatik tamamlama yeteneği, manuel veri girişine duyulan ihtiyacı önemli ölçüde azaltmaktadır (Gilmour & Cobus-Kuo, 2011:1).

Bookends, “Cite While You Write” yaklaşımı çerçevesinde Microsoft Word ve Apple Pages ile entegre çalışarak metin içi atıfların ve kaynakça listelerinin otomatik olarak oluşturulmasını sağlar. Bunun yanı sıra BibTeX ve BibLaTeX formatlarını desteklemesi sayesinde LaTeX tabanlı akademik yazım ortamlarıyla da uyumlu bir biçimde çalışabilmektedir. Bu çoklu entegrasyon yapısı, Bookends’in hem klasik kelime işlemciler hem de teknik dokümantasyon araçları kullanan araştırmacılar için esnek bir çözüm sunmasına olanak tanımaktadır (Sonny Software, 2024:1).

Bookends Karşılaştırmalı Analiz ve Maliyet

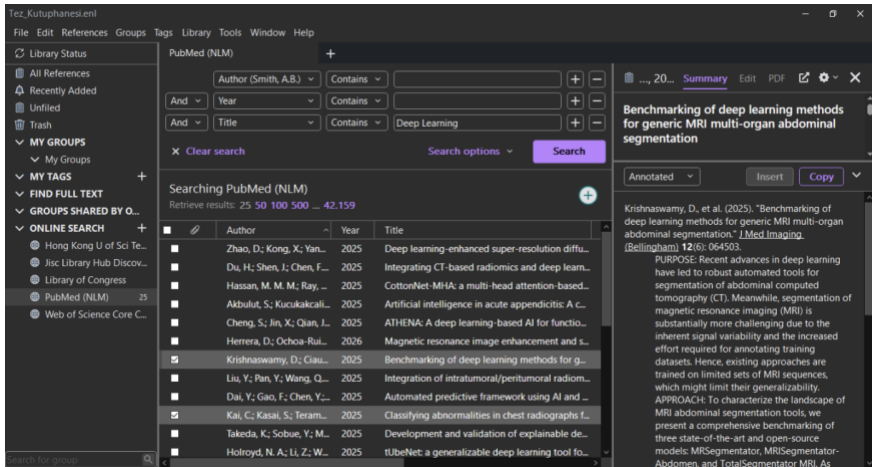
Benzer referans yönetim yazılımlarıyla karşılaştırıldığında Bookends, Zotero gibi açık kaynak kodlu ya da Mendeley gibi freemium lisans modelini benimseyen araçlardan farklı olarak tek seferlik lisanslama modeliyle sunulmaktadır. Bu yaklaşım, uzun vadede abonelik bağımlılığını ortadan kaldırarak bireysel kullanıcılar için maliyetlerin öngörülebilir olmasını sağlamaktadır (Ivey, 2016:1). Teknik açıdan değerlendirildiğinde Bookends’in yalnızca macOS platformuna yönelik geliştirilmiş olması, yazılımın performans ve kararlılık açısından güçlü bir avantaja sahip olmasını sağlarken; platform bağımsızlığı açısından bir sınırlılık oluşturmaktadır. Ayrıca bulut tabanlı, eş zamanlı ekip çalışmasına yönelik özelliklerin sınırlı olması, çok kullanıcı ve kurumsal projelerde RefWorks gibi SaaS tabanlı çözümlerin gerisinde kalmasına neden olmaktadır. Buna karşın, bireysel tez ve makale çalışmalarında, veri kontrolü ve yazım entegrasyonu açısından

Bookends literatürde güçlü ve güvenilir bir alternatif olarak konumlanmaktadır (Puckett, 2011:1; Ivey, 2016:1).

EndNote

Akademik çalışmalarda bibliyografik verilerin yönetimi ve atıf süreçlerinin sistematik biçimde yürütülmesi amacıyla kullanılan köklü yazılımlardan biri olan EndNote, Clarivate Analytics tarafından geliştirilen ve uzun yıllardır akademik literatürde yaygın biçimde kullanılan bir referans yönetim aracıdır. EndNote, hem masaüstü tabanlı (EndNote Desktop) hem de web tabanlı (EndNote Online) bileşenleri bir arada sunarak hibrit bir kullanım modeli sağlamaktadır (Clarivate, 2024:1). Bu mimari yapı, yazılımın çevrimdışı çalışabilmesini mümkün kılarken, aynı zamanda farklı cihazlar arasında senkronizasyon desteği sunmasına olanak tanımaktadır. Hibrit mimarinin sağladığı bu esneklik, kullanıcıların referans kütüphanelerine farklı çalışma ortamlarından erişebilmesini ve akademik üretim sürecini kesintisiz biçimde sürdürebilmesini desteklemektedir.

Şekil 4. EndNote arayüzü



EndNote, özellikle yüksek etki faktörlü (SCI/SCIE) dergiler için yayın hazırlayan araştırmacılar arasında yaygın olarak tercih edilmektedir. Bunun temel nedenleri arasında uzun yıllara dayanan kurumsal geçmişi, kapsamlı atıf stili desteği ve akademik yayıncılarla olan güçlü entegrasyonu yer almaktadır. Yazılım, bireysel araştırmacıların yanı sıra üniversiteler ve araştırma kurumları tarafından sağlanan kurumsal lisanslar aracılığıyla da yaygın biçimde kullanılmaktadır (Brigham, 2014:1).

EndNote Teknik Altyapı, Programlama Dili ve Entegrasyon

EndNote'un masaüstü sürümü, platforma özgü derlenmiş bir yazılım olarak geliştirilmiş olup Windows ve macOS işletim sistemlerini desteklemektedir. Yazılımın kaynak kodu açık değildir; ticari ve kapalı kaynaklı (proprietary) bir mimariye sahiptir. Clarivate tarafından geliştirilen EndNote'un arka planında C/C++ tabanlı çekirdek bileşenler ve platforma özgü kullanıcı arayüzü katmanları bulunduğu, ayrıca web tabanlı bileşenlerinde modern web teknolojilerinin kullanıldığı belirtilmektedir (Clarivate, 2024:1).

EndNote, Web of Science, PubMed ve CrossRef gibi majör akademik veri tabanlarıyla doğrudan entegrasyon sağlayarak bibliyografik meta verilerin (yazar bilgileri, başlık, yayın yılı, dergi adı, DOI vb.) otomatik olarak içe aktarılmasına imkân tanımaktadır. Ayrıca BibTeX formatında içe ve dışa aktarma desteği sunması sayesinde LaTeX tabanlı akademik yazım ortamlarıyla da uyumlu şekilde çalışabilmektedir (Gilmour & Cobus-Kuo, 2011:1).

Microsoft Word ile entegrasyon, "Cite While You Write (CWYW)" eklentisi aracılığıyla sağlanmakta; bu sayede metin içi atıflar ve kaynakça listeleri yazım sırasında otomatik olarak oluşturulabilmektedir. IEEE, APA, Vancouver ve Chicago gibi binlerce atıf stilinin desteklenmesi, EndNote'un farklı disiplinlerde kullanımını mümkün kılmaktadır.

EndNote Lisanslama, Ücretlendirme ve Karşılaştırmalı Değerlendirme

EndNote, Zotero gibi açık kaynak kodlu yazılımlardan veya Mendeley gibi freemium modelini benimseyen araçlardan farklı olarak, ticari lisanslama modeli ile sunulmaktadır. Yazılım, bireysel kullanıcılar için ücretli bir lisans gerektirirken; birçok üniversite ve araştırma kurumu EndNote erişimini kurumsal lisans kapsamında sağlamaktadır (Ivey, 2016:1). Bu durum, EndNote'un bireysel kullanıcılar açısından maliyetli bir çözüm olarak algılanmasına neden olabilmektedir.

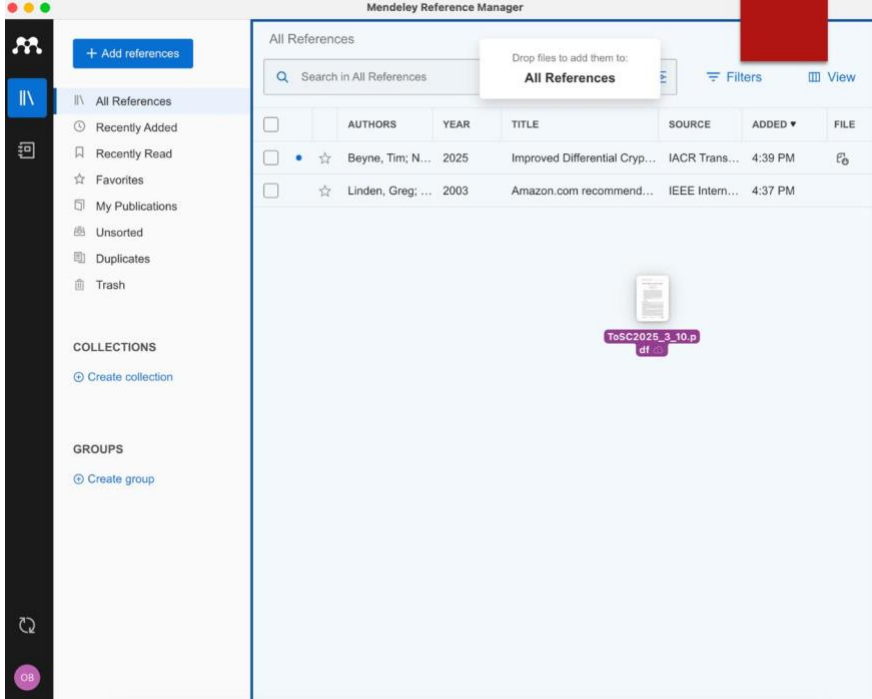
Teknik açıdan değerlendirildiğinde, EndNote'un büyük ölçekli kütüphaneleri yönetebilme kapasitesi ve gelişmiş filtreleme özellikleri önemli bir avantaj sunmaktadır. Bununla birlikte, kapalı kaynaklı yapısı nedeniyle kullanıcıların yazılım üzerinde özelleştirme yapma imkânlarının sınırlı olması ve lisans maliyetleri, açık kaynaklı alternatiflere kıyasla dezavantaj olarak değerlendirilmektedir (Puckett, 2011:1). Ayrıca kullanıcı arayüzünün görece karmaşık olması, yeni kullanıcılar için öğrenme eğrisini artırabilmektedir. Buna rağmen EndNote, güçlü Word entegrasyonu, kurumsal kabul görmüş yapısı ve uzun süreli akademik kullanım geçmişi sayesinde kapsamlı ve uzun soluklu akademik projelerde güvenilir bir referans yönetim aracı olarak konumlanmaktadır.

Mendeley

Akademik çalışmalarda bibliyografik verilerin yönetimi ve araştırma çıktılarının organize edilmesi amacıyla yaygın biçimde kullanılan referans yönetim yazılımlarından biri olan Mendeley, Elsevier tarafından geliştirilen ve hem masaüstü hem de web tabanlı bileşenlere sahip bir referans yönetim aracıdır. İlk olarak 2008 yılında piyasaya sürülen Mendeley, zaman içerisinde yalnızca bir referans yöneticisi olmaktan çıkarak, bulut tabanlı senkronizasyon

ve akademik sosyal ağ özelliklerini bir arada sunan bütünsel bir platform haline gelmiştir (Elsevier, 2024:1).

Şekil 5. Mendeley arayüzü



Mendeley, özellikle disiplinler arası çalışan araştırmacılar ve erken kariyer akademisyenleri arasında yaygın olarak tercih edilmekte; ücretsiz temel sürümü sayesinde bireysel kullanıcılar için erişilebilir bir çözüm sunmaktadır. Bulut tabanlı mimarisi, kullanıcıların farklı cihazlar üzerinden kütüphanelerine erişebilmesine ve verilerin otomatik olarak senkronize edilmesine olanak tanımaktadır (Puckett, 2011:1).

Mendeley Teknik Altyapı, Programlama Dili ve Entegrasyon

Mendeley, hibrit bir mimariye sahiptir. Masaüstü uygulaması, çok platformlu çalışmayı mümkün kılmak amacıyla C++ ve Qt framework tabanlı olarak geliştirilmiştir; web arayüzü ve

arka uç servislerinde ise modern web teknolojileri ve bulut altyapıları kullanılmaktadır. Yazılım kapalı kaynak kodlu (proprietary) olup Elsevier ekosistemiyle doğrudan entegre biçimde çalışmaktadır (Elsevier, 2024:1).

Mendeley, PDF dosyaları üzerinden otomatik meta veri çıkarımı yapabilmekte; başlık, yazar bilgileri, yayın yılı ve DOI gibi bibliyografik verileri kullanıcı müdahalesi olmadan kütüphaneye ekleyebilmektedir. Ayrıca Scopus ve ScienceDirect gibi Elsevier'e ait akademik veri tabanlarıyla güçlü bir entegrasyona sahiptir. Bu entegrasyon, özellikle Elsevier dergilerinde yayın yapan araştırmacılar için veri tutarlılığı açısından avantaj sağlamaktadır (Gilmour & Cobus-Kuo, 2011:1).

Microsoft Word ile entegrasyon, "Mendeley Cite" eklentisi aracılığıyla sağlanmakta; bu sayede metin içi atıflar ve kaynakça listeleri otomatik olarak oluşturulabilmektedir. Ayrıca BibTeX formatında dışa aktarma desteği sunması, Mendeley'in LaTeX tabanlı akademik yazım ortamlarıyla da birlikte kullanılabilmesine olanak tanımaktadır.

Mendeley Lisanslama, Ücretlendirme ve Karşılaştırmalı Değerlendirme

Mendeley, freemium lisans modeli ile sunulmaktadır. Temel sürüm ücretsiz olup sınırlı bulut depolama alanı ve temel referans yönetim özelliklerini içermektedir. Daha geniş depolama alanı ve gelişmiş iş birliği özellikleri ise ücretli abonelik planları kapsamında sunulmaktadır (Elsevier, 2024:1). Bu model, bireysel kullanıcılar için düşük giriş maliyeti sağlarken, uzun vadeli ve yoğun kullanım senaryolarında abonelik bağımlılığı yaratabilmektedir.

Karşılaştırmalı olarak değerlendirildiğinde, Mendeley'in en güçlü yönlerinden biri bulut tabanlı senkronizasyon ve grup çalışmasına yönelik özellikleridir. Buna karşın, Elsevier ekosistemine olan sıkı bağımlılığı ve kapalı kaynak kodlu yapısı,

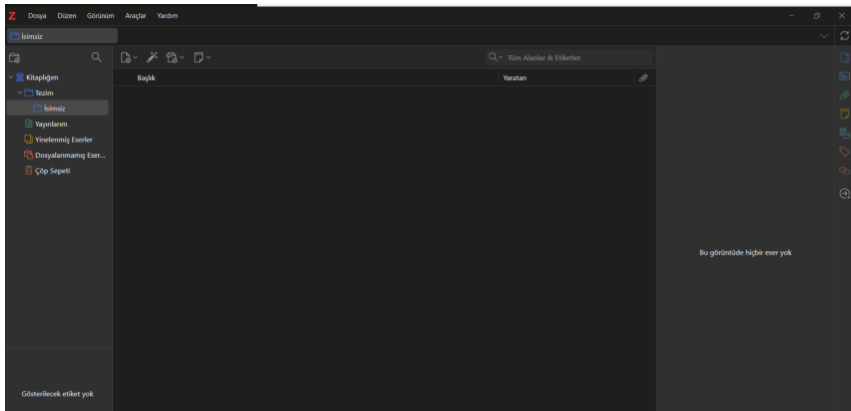
veri taşınabilirliği ve uzun vadeli erişim açısından literatürde eleştiri konusu olmuştur (Puckett, 2011:1; Ivey, 2016:1). Ayrıca önceki sürümlere kıyasla masaüstü uygulamasının özelleştirilebilirliğinin azalması, teknik kullanıcılar için bir dezavantaj olarak değerlendirilmektedir.

Bununla birlikte, kullanım kolaylığı, çok platformlu destek ve ücretsiz temel sürüm sunması sayesinde Mendeley, özellikle bireysel ve erken kariyer araştırmacıları için pratik ve erişilebilir bir referans yönetim aracı olarak akademik literatürde önemli bir konumda yer almaktadır.

Zotero

Akademik çalışmalarda bibliyografik verilerin toplanması, düzenlenmesi ve atıf süreçlerinin yönetilmesi amacıyla kullanılan referans yönetim yazılımlarından biri olan Zotero, açık kaynak kodlu ve ücretsiz bir yazılım olarak geliştirilmiştir. İlk olarak Roy Rosenzweig Center for History and New Media (George Mason University) bünyesinde geliştirilen Zotero, günümüzde küresel ölçekte akademisyenler tarafından yaygın biçimde kullanılan bir referans yönetim aracıdır (Puckett, 2011:1).

Şekil 6. Zotero arayüzü



Zotero'nun temel ayırt edici özelliđi, tamamen açık kaynaklı olması ve herhangi bir zorunlu lisans ücreti gerektirmemesidir. Bu yönüyle Zotero, özellikle bireysel araştırmacılar, lisansüstü öğrenciler ve bütçe kısıtları bulunan akademik kurumlar için erişilebilir bir çözüm sunmaktadır (Ivey, 2016:1). Masaüstü uygulaması ve bulut tabanlı senkronizasyon özellikleri sayesinde kullanıcılar, farklı cihazlar üzerinden kütüphanelerine erişebilmektedir.

Zotero Teknik Altyapı, Programlama Dili ve Entegrasyon

Zotero, çok platformlu çalışmayı destekleyecek şekilde tasarlanmış olup Windows, macOS ve Linux işletim sistemlerinde kullanılabilir. Yazılımın masaüstü sürümü büyük ölçüde JavaScript tabanlı olup HTML ve CSS ile oluşturulmuş kullanıcı arayüzü katmanları içermektedir. Açık kaynak kodlu yapısı sayesinde Zotero, geliştirici topluluđu tarafından sürekli olarak güncellenmekte ve genişletilebilmektedir (Gilmour & Cobus-Kuo, 2011:1).

Zotero'nun tarayıcı tabanlı bağlayıcıları (browser connectors), kullanıcıların akademik makaleleri, kitap kayıtlarını ve web tabanlı kaynakları tek tıkla kütüphaneye eklemesine olanak tanımaktadır. PDF dosyaları üzerinden otomatik meta veri çıkarımı yapılabilmekte; başlık, yazar bilgileri, yayın yılı ve DOI gibi alanlar otomatik olarak doldurulmaktadır. Ayrıca PDF üzerinde vurgulama (highlight) ve not alma özellikleri sunulması, Zotero'nun yalnızca bir referans yöneticisi değil, aynı zamanda bir literatür inceleme aracı olarak da kullanılabilmesini sağlamaktadır (Puckett, 2011:1).

Kelime işlemci entegrasyonu açısından Zotero; Microsoft Word, Google Docs ve LibreOffice ile uyumlu çalışmakta ve yazım sırasında metin içi atıfların ve kaynakça listesinin otomatik olarak oluşturulmasına imkân tanımaktadır. APA, IEEE, Chicago ve MLA gibi yaygın referans stillerinin desteklenmesi, Zotero'nun farklı

disiplinlerde kullanılabilirliğini artırmaktadır (Gilmour & Cobus-Kuo, 2011:1).

Zotero Lisanslama, Ücretlendirme ve Karşılaştırmalı Değerlendirme

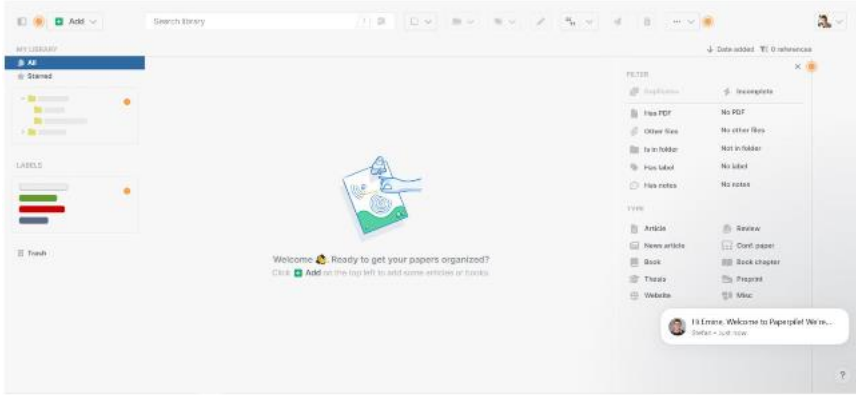
Zotero, açık kaynak lisansı altında dağıtılan ücretsiz bir yazılımdır. Temel kullanım için herhangi bir lisans ücreti talep edilmemekte; yalnızca ek bulut depolama alanı gereksinimi durumunda isteğe bağlı ücretli planlar sunulmaktadır. Bu yönüyle Zotero, tamamen ticari lisanslama modelini benimseyen EndNote ve freemium yapıya sahip Mendeley'den ayrılmaktadır (Ivey, 2016:1).

Karşılaştırmalı değerlendirmelerde Zotero'nun en güçlü yönleri arasında platform bağımsızlığı, açık kaynak kodlu yapısı ve geniş entegrasyon desteği yer almaktadır. Buna karşın, kurumsal ölçekte merkezi yönetim ve gelişmiş ekip iş birliği özellikleri açısından RefWorks gibi SaaS tabanlı çözümlerin gerisinde kaldığı literatürde belirtilmektedir (Puckett, 2011:1). Bununla birlikte, sunduğu özellikler ve maliyetsiz erişim imkânı göz önünde bulundurulduğunda Zotero, özellikle bireysel tez ve makale çalışmalarında akademik literatürde en yaygın benimsenmiş referans yönetim araçlarından biri olarak konumlanmaktadır.

Paperpile

Akademik çalışmalarda bibliyografik verilerin yönetimi, PDF dosyalarının düzenlenmesi ve metin içi atıf süreçlerinin otomatikleştirilmesi amacıyla kullanılan referans yönetim araçlarından biri olan Paperpile, tamamen bulut tabanlı (SaaS) bir mimariye sahip modern bir yazılımdır. Paperpile, özellikle Google Docs ile derin entegrasyonu sayesinde, çevrim içi akademik yazım süreçlerini desteklemek üzere tasarlanmıştır. Bu yönüyle Paperpile, klasik masaüstü tabanlı referans yönetim yazılımlarından farklı bir kullanım paradigması sunmaktadır (Ivey, 2016:1).

Şekil 7. Paperpile arayüzü



Kaynak: Paperpile (n.d.)

Paperpile'in temel yaklaşımı, kullanıcıların akademik kaynaklarını yerel depolama yerine bulut ortamında yönetmelerine olanak tanımak ve çoklu cihazlar arasında kesintisiz senkronizasyon sağlamaktır. Bu mimari yapı, özellikle çevrim içi ve iş birliğine dayalı akademik çalışma ortamlarında verimliliği artırmaktadır. Yüklenen kullanım rehberinde de belirtildiği üzere, Paperpile Google hesabı üzerinden çalışmakta ve Google Drive ile bütünleşik bir yapı sunmaktadır

Paperpile Teknik Altyapı, Programlama Dili ve Entegrasyon

Paperpile, masaüstü odaklı bir yazılım yerine web tabanlı bir platform olarak geliştirilmiştir. Arka planda modern web teknolojileri (JavaScript tabanlı istemci mimarisi ve bulut servisleri) kullanılarak tasarlanan sistem, kullanıcıların tarayıcı üzerinden erişim sağlamasına imkân tanımaktadır. Paperpile'in en ayırt edici teknik özelliği, Google Docs ile yerel bir eklenti (add-on) şeklinde çalışmasıdır.

Yüklenen dokümanda ayrıntılı biçimde gösterildiği üzere, Paperpile Chrome uzantısı aracılığıyla web üzerinden makale toplama, PDF yükleme ve otomatik meta veri çıkarımı işlemlerini

desteklemektedir. Kullanıcı yalnızca PDF dosyasını sisteme yüklemekte; Paperpile, makalenin başlığı, yazarları, yayınlandığı dergi ve yıl gibi bibliyografik bilgileri otomatik olarak tanımlayarak kütüphaneye aktarmaktadır. Bu özellik, literatürde modern referans yönetim yazılımlarının en önemli verimlilik kazanımlarından biri olarak değerlendirilmektedir (Gilmour & Cobus-Kuo, 2011:1).

Paperpile, Google Docs içerisinde metin yazımı sırasında otomatik atıf ekleme ve kaynakça oluşturma işlevlerini doğrudan belge üzerinde sunmaktadır. APA, IEEE, Chicago ve Harvard gibi çok sayıda referans stilinin desteklenmesi, yazılımın farklı disiplinlerde kullanılabilirliğini artırmaktadır. Ayrıca PDF okuma, vurgulama (highlight) ve not alma özelliklerinin bulutla senkronize edilmesi, Paperpile'ı yalnızca bir referans yöneticisi değil, aynı zamanda bir literatür inceleme aracı haline getirmektedir

Paperpile Lisanslama, Ücretlendirme ve Karşılaştırmalı Değerlendirme

Paperpile, abonelik tabanlı ticari lisanslama modeli ile sunulmaktadır. Yazılım, kullanıcılarına sınırlı süreli deneme imkânı sağlamakta; deneme süresi sonrasında aylık veya yıllık abonelik gerektirmektedir. Bu yönüyle Paperpile, Zotero gibi tamamen ücretsiz ve açık kaynaklı çözümlerden ayrılmakta; Mendeley ve EndNote gibi ticari yazılımlarla benzer bir konumda yer almaktadır (Ivey, 2016:1).

Karşılaştırmalı değerlendirmelerde Paperpile'ın en güçlü yönü, Google Docs ile kusursuz entegrasyonu ve bulut tabanlı çalışma modelidir. Özellikle ortak belgeler üzerinde eş zamanlı çalışan araştırma grupları için önemli avantajlar sunmaktadır. Buna karşın, Microsoft Word ve çevrimdışı çalışma senaryoları açısından masaüstü tabanlı çözümlerin (EndNote, Bookends) gerisinde kaldığı literatürde belirtilmektedir (Puckett, 2011:1).

Sonuç olarak Paperpile, Google ekosistemi merkezli akademik yazım yapan araştırmacılar için güçlü ve modern bir çözüm sunarken; veri kontrolünü yerel ortamda tutmak isteyen veya çevrimdışı çalışmaya ihtiyaç duyan kullanıcılar için daha sınırlı bir alternatif olarak değerlendirilmektedir.

Citavi

Akademik araştırma süreçlerinde yalnızca bibliyografik verilerin yönetilmesi değil, aynı zamanda literatürden elde edilen bilgilerin sistematik biçimde analiz edilmesi ve yapılandırılması da büyük önem taşımaktadır. Bu ihtiyaca yönelik olarak geliştirilen Citavi, Swiss Academic Software tarafından sunulan ve referans yönetimini bilgi ve bilgi birikimi yönetimi (knowledge management) ile bütünleştiren kapsamlı bir akademik yazılımdır. Citavi, özellikle tez ve doktora düzeyindeki çalışmalar için tasarlanmış olup, klasik referans yöneticilerinden farklı olarak araştırma sürecinin tamamını desteklemeyi amaçlamaktadır (Swiss Academic Software, 2024:1).

Citavi, Avrupa merkezli akademik kurumlarda yaygın biçimde kullanılmakta ve özellikle Almanya, İsviçre ve Avusturya'daki üniversitelerde kurumsal lisanslar aracılığıyla öğrencilere ve akademisyenlere ücretsiz olarak sunulmaktadır. Literatürde Citavi, “tez odaklı referans yönetim yazılımı” olarak tanımlanmakta ve uzun soluklu akademik projelerde sunduğu yapılandırma araçlarıyla öne çıkmaktadır (Becker, 2013:1).

Citavi Teknik Altyapı ve Yazılım Mimarisi

Citavi, temel olarak Windows işletim sistemi için geliştirilmiş masaüstü tabanlı bir yazılımdır ve kapalı kaynak kodlu (proprietary) bir mimariye sahiptir. Yazılım, Microsoft .NET tabanlı teknolojiler kullanılarak geliştirilmiş olup, yerel projeler (offline) ve bulut tabanlı projeler (Citavi Cloud) arasında seçim yapılmasına olanak tanımaktadır (Swiss Academic Software, 2024:1).

Citavi'nin teknik altyapısı, bibliyografik meta verilerin ISBN, DOI ve PMID gibi tanımlayıcılar aracılığıyla otomatik olarak elde edilmesini desteklemektedir. Ayrıca Citavi Picker tarayıcı eklentisi sayesinde Google Scholar, kütüphane katalogları ve akademik veri tabanlarından doğrudan kaynak ekleme imkânı sunulmaktadır. Bu özellikler, Citavi'nin yalnızca bir referans yöneticisi değil, aynı zamanda kapsamlı bir literatür toplama aracı olarak kullanılmasını mümkün kılmaktadır (Becker, 2013:1).

Microsoft Word entegrasyonu, Citavi'nin güçlü yönlerinden biridir. Word eklentisi aracılığıyla metin içi atıflar ve kaynakça listeleri yazım sırasında otomatik olarak oluşturulmakta; APA, IEEE, Chicago ve Vancouver gibi yaygın atıf stilleri desteklenmektedir. Bu entegrasyon, özellikle tez yazım sürecinde biçimsel tutarlılığın korunmasını kolaylaştırmaktadır (Swiss Academic Software, 2024:1).

Citavi Bilgi Yönetimi, Lisanslama ve Akademik Değerlendirme

Citavi'yi benzer referans yönetim yazılımlarından ayıran temel unsur, bilgi organizasyonuna yönelik gelişmiş özellikleridir. Yazılım, kullanıcıların makalelerden elde ettikleri doğrudan alıntılar, özetleri ve yorumları ayrı "knowledge items" olarak saklamasına ve bunları kategori ağaçları altında yapılandırmasına olanak tanımaktadır. Bu yaklaşımın, tez yazım sürecinde argümanların sistematik biçimde geliştirilmesine katkı sağladığı literatürde vurgulanmaktadır (Becker, 2013:1).

Lisanslama açısından Citavi, bireysel kullanıcılar için ücretli bir model benimsemekte; ancak birçok üniversitenin Citavi ile kurumsal lisans anlaşması bulunması nedeniyle öğrenciler ve akademisyenler yazılımı ücretsiz olarak kullanabilmektedir. Bu durum, Citavi'nin bireysel kullanımdan ziyade kurumsal akademik ortamlar için optimize edilmiş bir çözüm olarak konumlanmasına yol açmaktadır (Swiss Academic Software, 2024:1).

Akademik deęerlendirmelerde Citavi, referans ynetimi ile bilgi ynetimini birleřtiren yapısı sayesinde zellikle tez ve doktora alıřmalarında avantajlı bir ara olarak tanımlanmaktadır. Buna karřın, yalnızca Windows platformunu desteklemesi ve bireysel lisans maliyetinin yksek olması, Citavi'nin literatrde dile getirilen bařlıca sınırlılıkları arasında yer almaktadır (Becker, 2013:1). Buna raęmen, sunduęu arařtırma odaklı yaklařım sayesinde Citavi, akademik yazım srecini btncl biimde ele alan ileri dzey bir referans ynetim sistemi olarak kabul edilmektedir.

SONU VE DEęERLENDİRME

Bu alıřmada, akademik arařtırmalarda yaygın olarak kullanılan referans ynetim yazılımları; teknik altyapıları, lisanslama modelleri, entegrasyon yetenekleri ve kullanım amaları aısından karřılařtırmalı olarak incelenmiřtir. Dijitalleřmenin akademik retim srelerine hızla entegre olduęu gnmzde, literatr taraması, kaynak ynetimi ve atıf oluřturma srelerinin manuel yntemlerle yrtlmesi hem zaman kaybına hem de biimsel hatalara yol aabilmektedir.

alıřma kapsamında ele alınan Zotero ve JabRef gibi aık kaynaklı zmler, bte kısıtı bulunan arařtırmacılar iin nemli avantajlar sunarken; EndNote ve RefWorks kurumsal entegrasyon yetenekleriyle byk lekli akademik ortamlarda ne ıkmaktadır. Bookends ve Mendeley ise bireysel ve bulut tabanlı kullanım senaryolarında esnek zmler sunmaktadır.

Citavi, referans ynetimini bilgi organizasyonu ve tez yazım sreciyle btnleřtiren yaklařımı sayesinde zellikle lisansst alıřmalar iin gl bir alternatif olarak deęerlendirilmektedir. Sonu olarak, tek bir yazılımın tm kullanıcı profilleri iin en uygun zm olması beklenmemeli; arařtırmacıların alıřma trne, teknik gereksinimlerine ve bte olanaklarına gre bilinli bir tercih yapmaları gerekmektedir.

Kaynakça/References

Becker, C., Kraxner, M., Plangg, M., & Rauber, A. (2013). Improving decision support for software component selection through systematic cross-referencing and analysis of multiple decision criteria. Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS). Doi: 10.1109/HICSS.2013.263.

Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, 66(11), 2215–2222. Doi: 10.1002/asi.23329.

Clarivate (2024). EndNote: Citation & reference management tool. (15/11/2025 tarihinde <https://endnote.com> adresinden ulařılmıştır).

Elsevier (2024). Mendeley: Free reference manager. (17/11/2025 tarihinde <https://www.elsevier.com/products/mendeley> adresinden ulařılmıştır).

Ex Libris (2024). Recent RefWorks announcements / release information. (20/11/2025 tarihinde <https://capture-knowledge-exlibrisgroup-com-90611502d.getsmartling.com/%40api/deki/pages/114357/pdf/Recent%2BRefWorks%2BAnnouncements.pdf> adresinden ulařılmıştır).

Francese, E. (2013). Usage of reference management software at the University of Torino. *JLIS.it*, 4(2), 145–174. Doi: 10.4403/jlis.it-8679.

Gilmour, R., & Cobus-Kuo, L. (2011). Reference management software: A comparative analysis of four products. *Issues in Science and Technology Librarianship*, (66). Doi: 10.29173/istl1521.

Ivey, C., & Crum, J. (2018). Choosing the right citation management tool: EndNote, Mendeley, RefWorks, or Zotero. *Journal of the*

Medical Library Association, 106(3), 399–403. Doi: 10.5195/jmla.2018.468.

JabRef (t.y.). JabRef license. (23/11/2025 tarihinde <https://docs.jabref.org/advanced/license> adresinden ulařılmıştır).

Kern, M. K., & Hensley, M. K. (2011). Citation management software: Features and futures. *Reference & User Services Quarterly*, 50(3), 204–208.

Paperpile (t.y.). Paperpile pricing information. (26/11/2025 tarihinde <https://paperpile.com/pricing/> adresinden ulařılmıştır).

Perkel, J. M. (2020). Streamline your writing — and collaborations — with these reference managers. *Nature*, 585(7823), 149–150. Doi: 10.1038/d41586-020-02491-2.

ProQuest (2024). RefWorks (learn more). (05/12/2025 tarihinde <https://refworks.proquest.com/learn-more/> adresinden ulařılmıştır).

Puckett, J. (2011). *Zotero: A guide for librarians, researchers, and educators*. Chicago: Association of College and Research Libraries.

Sonny Software (2024). Bookends. (10/12/2025 tarihinde <https://www.sonnysoftware.com/bookends-for-mac/downloads> adresinden ulařılmıştır).

Swiss Academic Software (2024). Citavi (product information). (07/12/2025 tarihinde <https://www.citavi.com/en> adresinden ulařılmıştır).

Zotero (t.y.). Individual file storage. (11/12/2025 tarihinde https://www.zotero.org/support/individual_storage adresinden ulařılmıştır).

BÖLÜM 11

KONUŞMA DUYGU TANIMA İÇİN ÖZELLİK ANALİZİ

Kenan DONUK¹

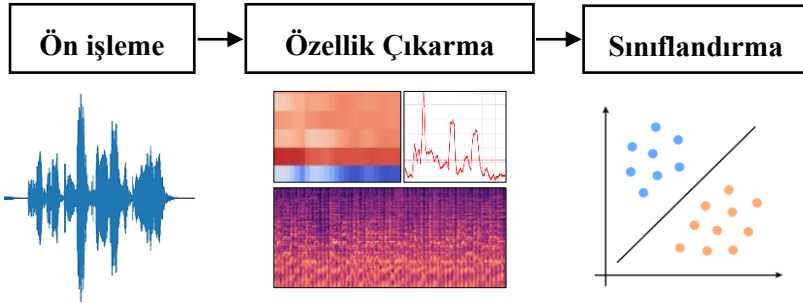
Giriş

Konuşma Duygu Tanıma (KDT), insan konuşmasının altında yatan duygusal durumların zeki sistemler aracılığıyla otomatik olarak algılanmasına odaklanan sinyal işleme ile derin öğrenmeyi birleştiren aktif bir alandır. Konuşma sinyallerinin akustik özellikleri kullanılarak, konuşulan sözcüklerin duygusal içeriği yorumlanır. İnsan-Bilgisayar etkileşiminde zeki sistemlerin kullanıcılarla daha doğru ve empatik bir şekilde etkileşim kurabilmesi için ses sinyallerinden duygusal ipuçlarını tespit edip değerlendirebilmeleri gerekmektedir. KDT teknolojileri, müşteri hizmetlerinde (Raisi & Mourhir, 2024), sesli asistanlarda (Ma & ark., 2025), eğitim sistemlerinde (Xian, 2022), ruh sağlığını izlemede (Jordan & ark., 2025), duygu temelli tehdit değerlendirmesi (Jena & ark., 2025) dahil olmak üzere çeşitli alanlarda kullanılmaktadır. KDT teknolojileri, ses sinyallerindeki zamansal dinamikler ve spektral özellikler de dahil olmak üzere akustik parametreleri analiz ederek ses sinyallerinden duygusal durumları tanımlayan gelişmiş bir

¹ Dr. Öğr. Üyesi, Şırnak Üniversitesi, Bilgisayar Müh., 0000-0002-7421-5587

hesaplama metodolojisi sunar. Metodoloji üç aşamalı bir işlem hattı üzerinden ilerlemektedir: sinyal ön işleme, akustik özellik çıkarma ve akustik tanımlayıcıları duygusal kategorilere eşleyerek sınıflandırmayı gerçekleştiren makine öğrenimi veya derin öğrenme. Şekil 1’de KDT süreci gösterilmiştir.

Şekil 1 KDT Süreci



Konuşma Duygu Tanımda ses sinyallerine uygulanan ön işlem adımlarından sonra gerçekleşen özellik çıkarma, genellikle duygusal ifadenin farklı dinamiklerini yakalayan birden fazla akustik kategoriye içerir. Çağdaş KDT araştırmaları ağırlıklı olarak spektral tabanlı özellikleri kullanmakta ve bu özellikler literatürde (Liu & Yuan, 2023; Singh & ark., 2023) iyi performanslar sergilemektedirler. Bu özellikler önemli başarılar göstermiş olsa da, KDT alanı hala çok sayıda temel sınırlamayla boğuşmaktadır. Başlıca zorluklar arasında sınırlı açıklamalı veri kümeleri, genellemeyi etkileyen kültürel değişkenlik, sınıf dengesizliği sorunları, konuşmacıya bağımlı ve konuşmacıdan bağımsız modeller arasındaki önemli performans farklılıkları yer almaktadır. Bu zorlukların üstesinden gelmek, uygun sınıflandırma mimarilerinin yanı sıra uygun özellikler ve özellik kombinasyonlarının kullanımıyla mümkündür. Bu bölüm, konuşma sinyalindeki spektral ve zamansal kategorilerinden önemli olan akustik özellikleri açıklayan ve analiz eden bir çerçeve sunmaktadır.

Spektral Tanımlayıcılar

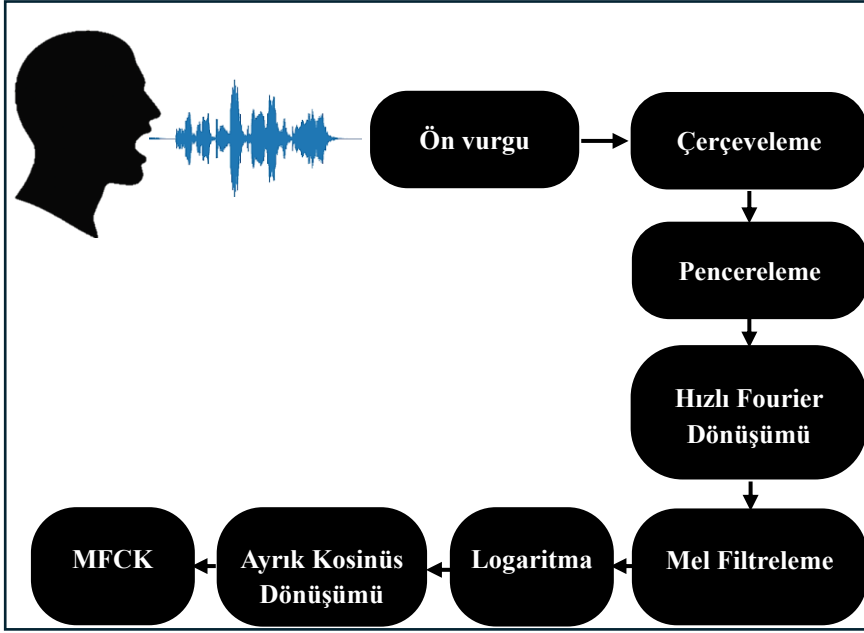
Spektral tanımlayıcılar, farklı duygusal durumlara ait ses enerjisinin dağılımını ve nasıl değiştiğini analiz ederek, insanın ses özelliklerini karakterize eden temel akustik parametrelerdir. Bu parametreler veya bu öznelik vektörleri konuşma analizinden müzik tanımaya kadar birçok uygulamada ses içeriğinin tanımlanmasını ve sınıflandırılmasını sağlar.

- **Mel-Frekans Cepstral Katsayıları**

Davis and Mermelstein tarafından 1980 yılında tanımlanan (Davis & Mermelstein, 1980) Mel-Frekans Cepstral Katsayıları (MFCK) insan kulağının sesi algılama biçimini taklit eden ve ses analizinde kullanılan en yaygın spektral özelliklerden biridir. MFCK çıkarımı, yedi aşamalı bir metodoloji kullanarak konuşma duygu tanımada önemli bir özellik çıkarma yaklaşımını oluşturur. Bu yaklaşımda, ilk adımda isteğe bağlı olarak genellikle yüksek frekanslı bileşenleri belirginleştirmek için ön vurgulama işlemi gerçekleştirilir. Ön vurgulama adımının ardından durağan olmayan ses sinyallerinden anlamlı akustik özellikler çıkarmak için ses sinyaline 20-40 ms'lik çerçevelerle segmentasyon uygulanır. Segmentasyon sonucunda oluşan spektral sızıntıları azaltmak için çerçevelere pencereleme işlevi uygulanır (Hamming, 1983). Kısa aralıklarla segmente edilen ses sinyali bu aralıklarda durağan kabul edilerek Kısa Zamanlı Fourier Dönüşümü (KZFD) uygulanarak bu aralıklarda sinyalin frekans içeriği elde edilir (Gabor, 1946). KZFD den sonra mel ölçekli filtre bankaları (20-40 adet üçgen filtreler), insan işitme özelliklerine dayalı olarak ilgili frekans bilgilerini çıkarırlar. Bu aşamadan sonra insan işitme algısıyla uyumlu bir logaritmik işlem gerçekleşir. Bu işlemin ardından Ayrık Kosinüs Dönüşümü (AKD) ile sinyaldeki özellik korelasyonu ortadan kaldırılır. Daha sonra elde edilen ve genellikle ilk 12-13 katsayı, son MFCK gösterimini oluşturur. Bu gösterim konuşma spektral

özelliklerinin kapsamlı bir karakterizasyonunu sağlar. Şekil 2’de MFCK çıkarma süreci gösterilmiştir.

Şekil 2 MFCK Çıkarma Süreci

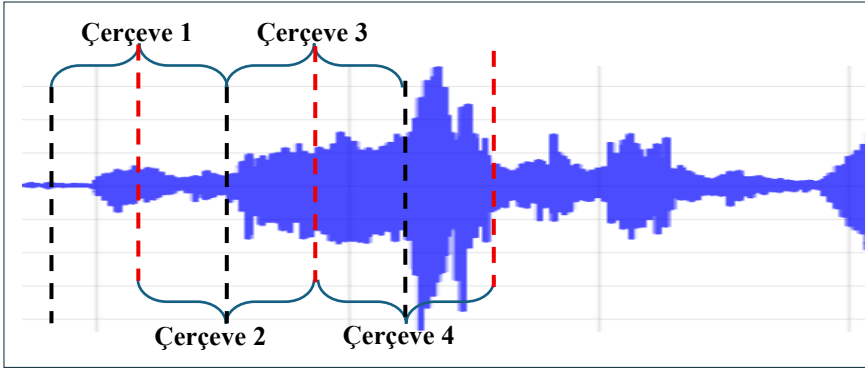


MFCK çıkarma işlemi, yüksek frekanslı bileşenlerin etkisini artırmak için opsiyonel olarak Denklem 1’deki ön-vurgu filtrelemesiyle başlar. Denklemde $y[n]$ parametresi ön-vurgulanmış çıktıyı, $x[n]$ parametresi giriş sinyalin, α parametresi ön-vurgulama katsayısını (0.95-0.97), n parametresi sinyalin örnek indeksini temsil eder.

$$y[n] = x[n] - \alpha \cdot x[n - 1] \quad (1)$$

Ön vurgulamanın ardından ses sinyali, en küçük akustik anlamlı değişimin gerçekleştiği 20-40 ms’lik kısa çerçevelere bölünür. Bu bölünmeler gerçekleşirken bilgi kaybının önüne geçmek için çerçeveler %50 oranında örtüştürülür. Şekil 3’te ses sinyali üzerindeki çerçeveleme süreci ve örtüşme gösterilmiştir.

Şekil 3 Ses Sinyali Çerçeveleme ve Örtüşme



Spektral sızıntıyı azaltmak için Denklem 2'de fonksiyonu verilen Hamming pencereleme, ses sinyalindeki her çerçeveye uygulanır. Çerçevelere uygulanan pencereleme işlemi Denklem 3'te verilmiştir. Denklemlerde yer alan parametreler şunlardır: $w[n]$ parametresi Hamming pencere değerini; N parametresi penceredeki toplam örnek sayısını; n parametresi örnek indeksini; 0.54, 0.46 sabitleri Hamming katsayılarını, $y[n]$ parametresi pencerelenmiş çıkış sinyalini; $x[n]$ parametresi giriş sinyalini veya ön-vurgulanmış giriş sinyalini temsil eder.

$$w[n] = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N-1}\right) \quad (2)$$

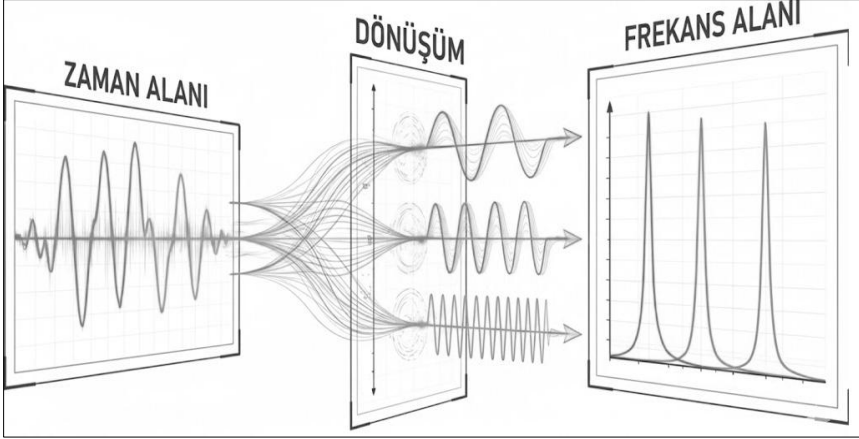
$$y[n] = x[n] \cdot w[n] \quad (3)$$

Pencerelenmiş sinyale Denklem 4'te verilen KZFD işlemi uygulanarak sinyal örneği zaman alanından frekans alanına dönüştürülür. Frekans bileşeninin taşıdığı güç değeri ise Denklem 5 ile hesaplanır. $X[k]$ parametresi KZFD katsayısını, $x[n]$ parametresi pencerelenmiş konuşma sinyal örneğini, N parametresi KZFD uzunluğunu, k parametresi frekans aralığını, j parametresi sanal birimi, $P[k]$ parametresi güç spektrumunu temsil eder. Şekil 4'te ses sinyalinin zaman alanından frekans alanına dönüşümü gösterilmiştir.

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N} \quad (4)$$

$$P[k] = |X[k]|^2 \quad (5)$$

Şekil 4 Ses Sinyalinin Zaman Alanından Frekans Alanına Dönüşümü



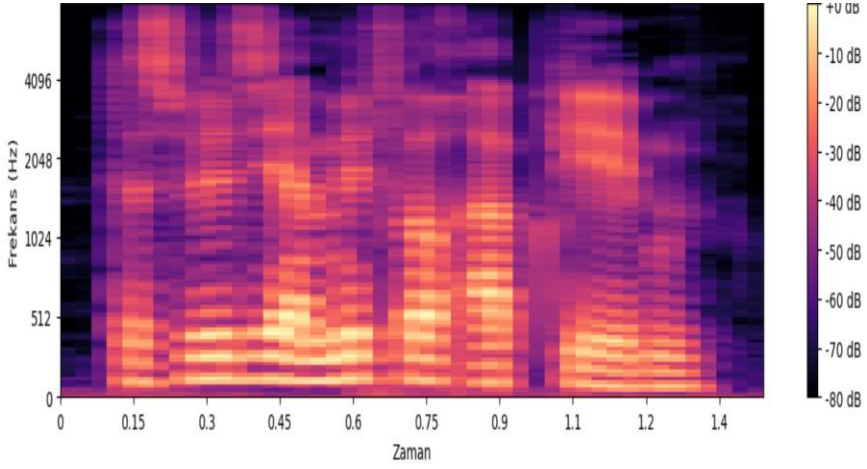
Kaynak: Nano Banana Pro, 2025

Denklem 6 kullanılarak, frekans değerleri insan algısıyla uyumlu mel-frekans değerlerine dönüştürülür. Bu aşamadan sonra Denklem 5’te hesaplanan güç spektrumu, Denklem 7 ile 20-40 üçgen filtreden oluşan bir mel ölçekli filtre bankasından geçirilir. Denklem 6 ve 7’deki parametreler şunlardır: $mel(f)$ parametresi mel frekansını, f parametresi frekansı, 2595 ve 700 sayıları dönüştürme sabitlerini, $S[m]$ parametresi filtre enerjisini, $P[k]$ parametresi spektrumu, $H_m[k]$ parametresi filtre tepkisini, m parametresi filtre indeksini temsil eder. Şekil 5’te mel-ölçekli zaman-frekans gösterimi verilmiştir. Şekildeki farklı renkler ses şiddetinin desibel türünden farklı seviyelerini temsil etmektedir.

$$mel(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \quad (6)$$

$$S[m] = \sum_{k=0}^{N-1} P[k] \cdot H_m[k] \quad (7)$$

Şekil 5 Mel Spektrogram



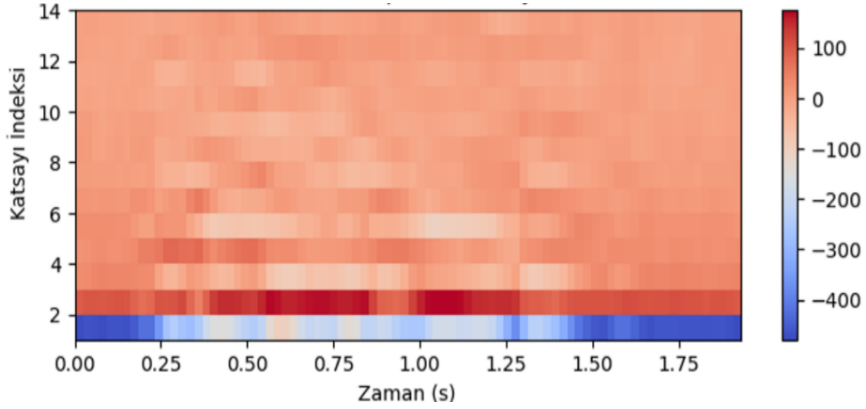
Filtre bankası çıktılarına, logaritmik şekilde çalışan insan ses algısına uygun olarak Denklem 8’de verilen logaritmik sıkıştırma işlemi uygulanır. Logaritmik sıkıştırma işleminden sonra Ayrık Kosinüs Dönüşümü (AKD) uygulanır. Bu dönüşüm ile enerji katsayıları birbirinden bağımsız hâle getirilerek daha az boyuta sahip bir sıkıştırılmış ses özelliği vektörü elde edilir. Denklem 9 ve Denklem 10’da sırasıyla AKD işlemi ve MFCK vektörünü oluşturan 13 adet katsayının gösterimi verilmiştir. Bu denklemlerde yer alan parametreler şunlarıdır: $L[m]$ parametresi log enerjiyi, $S[m]$ parametresi filtre enerjisini, m parametresi filtre indeksini, $C[n]$ parametresi cepstral katsayıları, M parametresi filtre sayısını, n parametresi katsayı indeksini temsil etmektedir. MFCK katsayılarının oluşturduğu görsel Şekil 6’da verilmiştir.

$$L[m] = \ln(S[m]) \quad (8)$$

$$C[n] = \sum_{m=0}^{M-1} L[m] \cdot \cos\left(\frac{\pi n(m+0.5)}{M}\right) \quad (9)$$

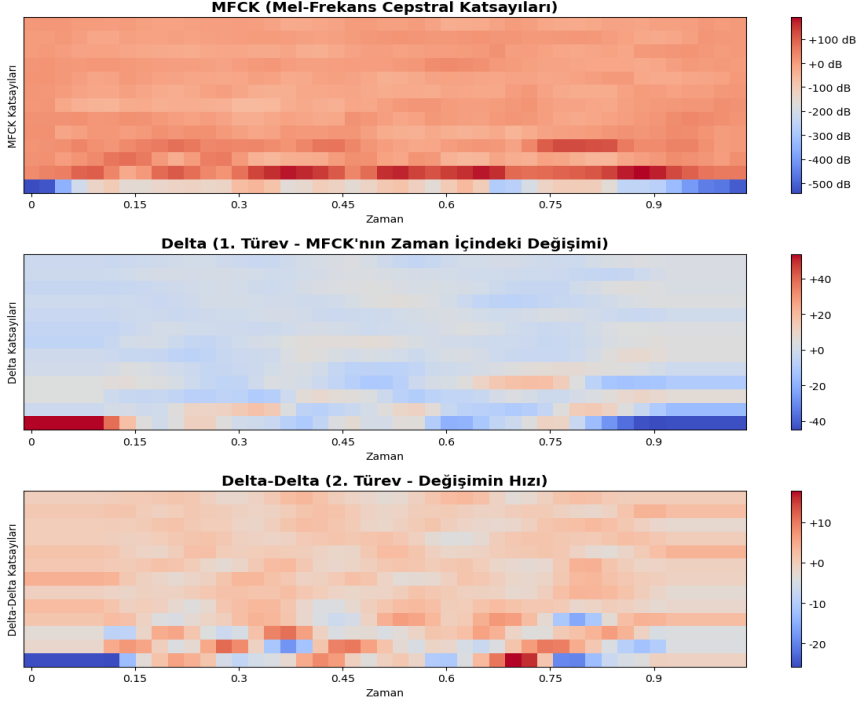
$$\text{MFCK} = [C_1, C_2, C_3, \dots, C_{13}] \quad (10)$$

Şekil 6 MFCK Çerçeve Başına 13 Adet Katsayı Gösterimi



MFCK spektral özelliği, KDT ve birçok farklı alanda kullanılan en yaygın sinyal temsillerinden biridir (Abdul & Al-Talabani, 2022). MFCK, belirli aralıklara bölünmüş ve sinyaldeki her aralık için zaman eksenini boyunca aynı sayıda katsayı içeren bir temsil oluşturmaktadır. Literatürde MFCK özelliğinin kaç adet katsayı ile daha iyi bir temsil ortaya koyacağı aktif bir araştırma konusudur (Rakibul Hasan & ark., 2021). MFCK katsayı adeti uygulama alanına göre değişse de genellikle 12-13 katsayı ile ifade edilmektedir. Literatürde MFCK özelliği insan işitme sistemine yakın bir temsili ifade etmekte olduğundan ve sesin karakterini sağlam bir şekilde yansıttığından KDT alanında önemli bir yere sahiptir. KDT alanındaki algoritmalarda modelden daha yüksek doğruluklar elde etmek için MFCK ile beraber diğer akustik özellikler birleştirilir (Bhangale & Kothandaraman, 2023; Jothimani & Premalatha, 2022). MFCK temsiline yanısıra MFCK'daki spektral zarfın zamansal dinamiklerini yakalamak için MFCK'nın 1. ve 2. türevleride kullanılmaktadır (Aldarmaki & ark., 2022). Şekil 7'de MFCK'nın orijinal formu, 1. ve 2. türevleri gösterilmiştir. 1. türev MFCK'nın zaman içindeki katsayı değişimini ortaya koyarken, 2. türev değişimin hızı hakkında bilgi vermektedir.

Şekil 7 MFCK'nın Orjinal Formu, 1. ve 2. Türevleri



- **Spektral Ağırlık Merkezi**

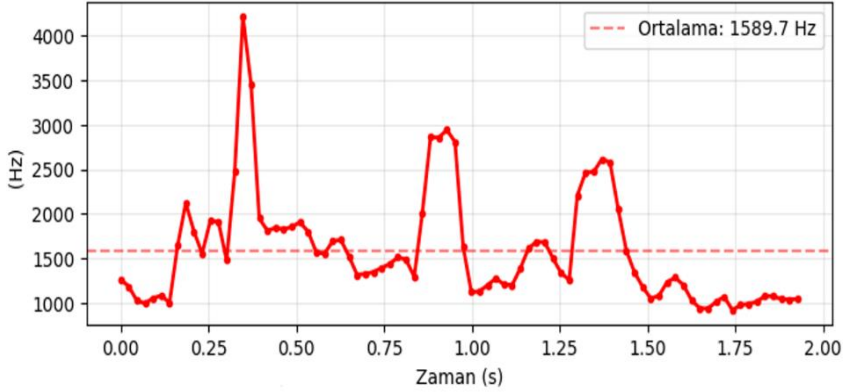
Spektral Ağırlık Merkezi (SAM), ses sinyalindeki frekans spektrumunun ağırlık merkezini işaret eden, sesin parlaklık algısını ölçen ve yaygın olarak kullanılan akustik bir ses özelliğidir. Yüksek SAM değeri, yüksek frekanslı bileşenlerin ağırlığının arttığını düşük SAM değeri ise düşük frekanslı bileşenlerin ağırlığının yüksek olduğunu gösterir (Bhangale & Kothandaraman, 2023).

$$SAM = \frac{\sum_{n=1}^N f(n) \cdot |X(n)|}{\sum_{n=1}^N |X(n)|} \quad (11)$$

Denklem 11'de, $f(n)$ her frekans bölmesinin merkez frekansını, $|X(n)|$ o frekans bandındaki sinyal büyüklüğünü, N toplam bölme sayısını ve n bölme endeksini ifade eder. Bu parametreler, frekansları karşılık gelen enerji değerleriyle

ağırlıklandırarak SAM değerini hesaplar. Bir konuşma sinyalindeki SAM'in zaman-frekans eksenindeki gösterimi Şekil 8'de verilmiştir.

Şekil 8 Spektral Ağırlık Merkezi



SAM, ses sinyalindeki çerçeve başına 1 adet skaler değer ile temsil edilen bir spektral özelliktir. Bu özellikte diğer akustik özellikler ile beraber kullanılarak giriş temsilinin ayırt edici gücünü arttırmaktadır. Literatürde duygusal uyarılma düzeyinin temsilinde ayırt edici bir görev üstlenmektedir (Sievers & ark., 2019). Düşük SAM değerleri sakin, üzgün etiketli duygusal durumlara, yüksek SAM değerleri öfkeli, mutlu etiketli duygusal durumlara işaret etmektedir (Liu & ark., 2018).

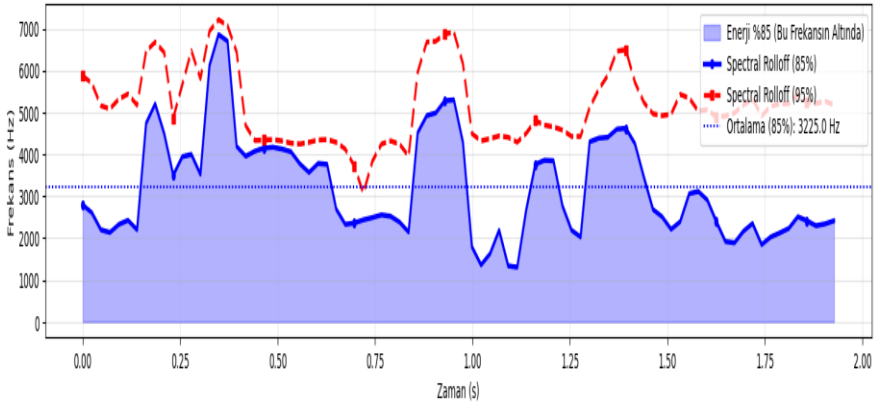
- **Spektral Düşüm Frekansı**

Spektral Düşüm Frekansı (SDF), spektrumun toplam enerjisinin belirli bir yüzdesinin (genellikle %85) hangi frekansın altında kaldığını ölçen (Mashhadi & Osei-Bonsu, 2023) bir akustik özelliktir. Frekanslar arasındaki enerji dağılımını özetler ve ses sinyalinin bant genişliğini karakterize etmek için kullanılır. Düşük frekansta düşüm değerleri düşük frekanslı yoğun sinyalleri, yüksek frekansta düşüm değerleri ise daha yüksek frekanslı yoğun sinyalleri gösterir. Denklem 12'de SDF hesaplaması verilmiştir.

$$\sum_{k=1}^m |X(k)|^2 = \alpha \sum_{k=1}^N |X(k)|^2 \quad (12)$$

Denklem 12'deki parametrelerden m , kümülatif enerjinin eşik değerini karşıladığı noktayı gösteren düşüş indeksini belirtir. N , toplam frekans bölmesi (bin) sayısını temsil eder. $|X(k)|^2$, k . frekanstaki spektral gücü, α düşüş noktasını tanımlayan toplam enerjinin yüzde kaçını belirleyeceği eşik oranıdır (genellikle 0,85 veya 0,95). Denklem 12, eşitliğin sol tarafının sağ tarafa eşit veya daha büyük olduğu ilk m değerini bulur; bu değer, düşüm frekansını belirler. Şekil 9'da bir ses örneği üzerinde SDF'nin gösterimi verilmiştir.

Şekil 9 Spektral Düşüm Frekansı



Frekans spektrumunda toplam enerjinin belirli bir yüzdesinin altında kalan frekansı ifade eden SDF, ses sinyalindeki çerçeve başına 1 adet skaler değer ile temsil edilen bir spektral özelliktir. Farklı spektral dağılımlar farklı duyguları işaret edebileceğinden duygu sınıflandırmasında giriş temsiline ayırt edici günün artırılmasında diğer özelliklerle beraber kullanılır (Mashhadi & Osei-Bonsu, 2023). Örneğin korku ve kızgınlık duygusal durumlarında SDF yüksek iken iğrenme ve üzüntü duygusal durumlarında düşük SDF değeri ölçülmüştür (Chandrasekar & ark., 2014). Başka bir çalışma SDF'nin duygu tanıma doğruluğu üzerinde en iyi ikinci etkiye sahip olduğunu göstermektedir (Van & ark., 2022).

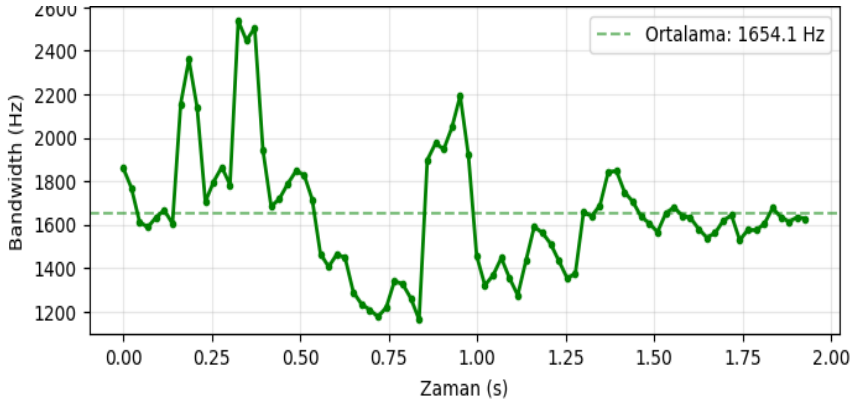
- **Spektral Bant Genişliği**

Spektral Bant Genişliği (SBG), ses sinyallerinde spektral enerjinin spektral ağırlık merkezi etrafındaki dağılımını niceliksel olarak ifade eden bir ölçüttür (Lin & ark., 2014). Denklem 13'te SBG hesaplaması verilmiştir.

$$SBG = \sqrt{\frac{\sum_n S(n) [f(n) - SAM]^2}{\sum_n S(n)}} \quad (13)$$

Bu formülasyonda, $S(n)$ n. frekanstaki spektral büyüklüğü, $f(n)$ n'inci bileşene karşılık gelen frekans değerini, SAM spektral ağırlık merkezini ifade eder. Şekil 10'da SBG'nin bir ses örneği üzerindeki gösterimi verilmiştir.

Şekil 10 Spektral Bant Genişliği



SBG, KDT sistemlerinde yaygın olarak kullanılan tamamlayıcı bir akustik özelliktir. Her çerçeve sinyali için 1 adet değerle temsil sağlar. Bant genişliğinin daraltılması duygu tanıma başarısını olumsuz etkilediği belirtilmiştir (Lech & ark., 2020).

- **Spektral Akı**

Bu özellikte diğer spektral özellikler gibi konuşma sinyalinin önemli özellikleri arasında yer alır. Spektral Akı (SA) bir sinyalin enerji dağılımının zaman içindeki değişim oranı ve konuşmanın

dinamik yapısı hakkında bilgi sağlayan bir özelliktir (Zeng & ark., 2025). Her çerçevenin güç spektrumunun bir önceki çerçevenin güç spektrumuyla karşılaştırılmasıyla hesaplanır . Denklem 14'te SA hesaplaması verilmiştir.

$$SA = \sum_{k=1}^K (|X(n, k)| - |X(n - 1, k)|)^2 \quad (14)$$

Verilen denklemde n çerçeve indeksini, k frekans bölmelerini ve K toplam bölme sayısını belirtir. $X(n, k)$ n . zamandaki, k . frekanstaki spektral büyüklüğe karşılık gelirken, $X(n - 1, k)$ bir önceki zaman çerçevesindeki aynı frekans binine karşılık gelen spektral büyüklüğü ifade eder. Şekil 11'de SA'nın bir ses örneği üzerindeki gösterimi verilmiştir.

Şekil 11 Spektral Akı



KDT alanındaki önemli akustik özelliklerden biri olan SA, konuşma sinyalindeki spektral içeriğin çerçeveler arası hız dinamiğini yakalayarak zamana duyarlı bir temsil ortaya koyar. Çerçeve başına 1 adet özellik değerine karşılık gelen SA, akustik ve prozodik diğer özellik değerleriyle beraber kullanılarak duygu tanıma performansına katkıda bulunur. Yapılan bir araştırmada (Gingras & ark., 2014) yalnızca Spektral Akı ve Spektral Entropinin bir konuşmacının uyarılmışlık derecelerindeki değişkenliğin %65'ni açıkladığını ortaya koymuştur.

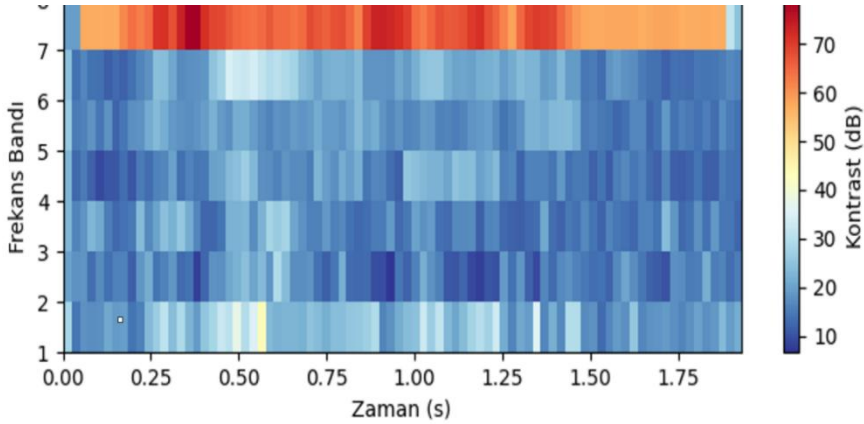
- **Spektral Kontrast**

Spektral Kontrast (SK), frekans spektrumunun yüksek enerji bölgeleri (tepeler) ve düşük enerji bölgeleri (vadiler) arasındaki farkı hesaplar (Tolu & ark., 2024). Denklem 15'te Spektral Kontrast hesaplaması verilmiştir.

$$\text{Skontrast}_k(t) = \log(\text{peak}_k(t)) - \log(\text{valley}_k(t)) = \log\left(\frac{\text{peak}_k(t)}{\text{valley}_k(t)}\right) \quad (15)$$

Denklem 15'deki Spektral Kontrast parametrelerinden k frekans bandı indeksini, t zaman çerçevesini, $\text{peak}_k(t)$ o banttaki spektrumun tepe (maksimum) genlik değerini, $\text{valley}_k(t)$ spektrumun vadi (minimum) genlik değerini temsil eder. Logaritmik ölçekleme ifadesi tepe-vadi oranını daha dengeli bir şekilde temsil eder. Şekil 12'de Spektral Kontrast'ın bir ses örneği üzerindeki gösterimi verilmiştir.

Şekil 12 Spektral Kontrast



KDT sistemlerinin önemli spektral özelliklerden biri olan SK, farklı uyarılma seviyelerine sahip duyuları ayırt etmedeki etkinliği nedeniyle KDT sistemlerindeki tahmin doğruluğu üzerinde önemli katkıları mevcuttur (Kumar & Thiruvengadam, 2021). Yüksek kontrast değerleri belirgin harmonik bileşenlere karşılık gelirken, düşük kontrast değerleri geniş bantlı, gürültülü bileşenlere

karşılık gelmektedir. SK, ortalama spektral bilgiyi içeren ve spektral dağılımı kaybeden MFCK gibi spektral özelliklerden farklı olarak daha ayrıntılı ve göreceli spektral özelliklere sahip olmasıyla müzik türünün ayırt edilmesinde etkilidir (Jiang et al., 2002). MFCK gibi spektral özellikler ile birleştirilerek duygu sınıflandırma başarısında etkili rol oynamaktadır. Örneğin 3 adet konuşma verisetinin kullanıldığı bir çalışmada SK, 7 özellik (MFCK ve Melspektrogramında olduğu) arasından tekil özellik kullanımında doğruluk sıralamasında 3. sırada yer alırken, MFCK ve Melspektrogram ile birleştirildiğinde ise 4 adet kombinasyon arasından en iyi kombinasyon olarak tespit edilmiştir (Dair & ark., 2021). KDT alanındaki çalışmalarda SK temsili için çerçeve başına genellikle 6 veya 7 adet özellik kullanılmaktadır.

Zamansal Tanımlayıcılar

KDT'deki zamansal tanımlayıcı özellikler, konuşma sinyalinin kısa zaman aralıklarındaki zamansal dinamiklerini yansıtır ve duygu tanımda kritik bir rol oynar. Karekök Ortalama Enerji (KOE) değerleri, bir ses sinyalinin güç içeriği hakkında bilgi sağlarken, Sıfır Geçiş Oranı (SGO) sinyalin spektral içeriğini veya sinyal işaretinin ne sıklıkla değiştiğini ölçer. Bu zamansal parametreler, spektral özellikler ile birleştirildiğinde, duygusal durumların daha etkili bir şekilde temsil edilmesini sağlar.

- **Karekök Ortalama Enerji**

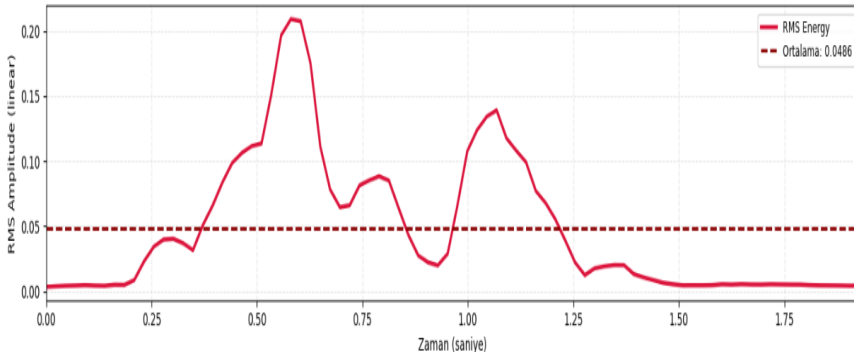
Karekök Ortalama Enerji (KOE), çerçeve seviyesinde bir ses sinyalinin ortalama enerji düzeyini ve sinyalin “güçlü” ya da “zayıf” olduğunu gösteren bir ölçüttür. Konuşma duygu tanımda, 20-40 ms'lik aralıklarda hesaplanan KOE değerleri, duygusal içeriği belirlemek için ses sinyalinin yoğunluğunu ölçer. Denklem 16'da KOE formülü verilmiştir.

$$KOE = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x[n]^2} \quad (16)$$

--227--

KOE denkleminde N toplam örnek sayısını, $x[n]$ her bir zaman noktasındaki sinyal genliğini, n ise örnekleme indeksini temsil eder. Denklem, tüm sinyal örneklerinin karelerinin toplamını alır, örnek sayısına böler ve karekökünü hesaplayarak sinyalin etkin gücünü verir. Bu yöntem, ses sinyalinin ortalama enerjisini tek bir değerle ifade ederek duygu tanımada kullanılır. Şekil 13'te KOE'nin bir ses örneği üzerindeki gösterimi verilmiştir.

Şekil 13 Karekök Ortalama Enerji



Konuşma sinyalinin enerji büyüklüğü hakkında sağladığı bilgiler ve ses sinyalindeki çerçeveler boyunca duygusal içerikteki değişikliklerin göstergesi olması açısından KDT alınındaki önemli özelliklerden biridir (Barhoumi & BenAyed, 2025; Bhangale & Kothandaraman, 2023). KOE her çerçeve başına 1 adet skaler değer ile temsil edilir. Yüksek KOE değerinin mutlu duygularla, düşük KOE değerlerinin ise korku, öfke, üzüntülü duygularla ilişkili olduğu ve olumlu duygusal ifadelerde daha fazla enerji değişimi olduğu ortaya konmuştur (Ma & ark., 2025).

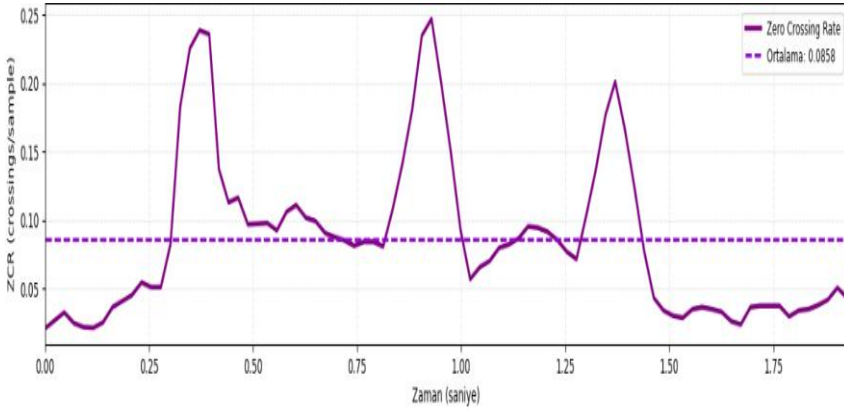
- **Sıfır Geçiş Oranı**

Sıfır Geçiş Oranı (SGO), bir ses sinyalinin sıfır eksenini kaç kez geçtiğini ölçer ve konuşma tanıma, müzik işleme uygulamalarında temel bir zaman alanı özelliği olarak hizmet eder. SGO hesaplaması Denklem 17'de verilmiştir.

$$SGO = \frac{1}{2(N-1)} \sum_{n=1}^{N-1} |\text{sgn}(x[n]) - \text{sgn}(x[n-1])| \quad (17)$$

Bu formülde $x[n]$ ses sinyali örneğini, N toplam örnek sayısını gösterir. $\text{Sgn}(x)$ işaret fonksiyonu, pozitif değerler için +1, negatif için -1 döndürür. Formül, ardışık örnekler arasındaki işaret değişimlerini sayar ve normalize ederek 0-1 arası sıfır geçiş oranını hesaplar. Şekil 14'te SGO'nun bir ses örneği üzerindeki gösterimi verilmiştir.

Şekil 14 Sıfır Geçiş Oranı



Sinyalin baskın frekansının kabaca bir tahmini olan SGO (O'Shaughnessy, 2023), hesaplanması ve sonuçlarıyla sinyaldeki en basit ve en etkili özelliklerden biri sayılabilir. KDT algoritmalarında sınıflandırmada kullanılmak üzere her çerçevede 1 adet özellik ile temsil edilir. SGO, diğer akustik özellikler ile beraber kullanımı anlamlı başarılar ortaya koymaktadır. Yapılan bir KDT çalışmasında 14 akustik özellik ile yapılan karşılaştırmada en iyi sınıflandırma performansında ikinci olmuştur (Shoiynbek & ark., 2019). Örneğin bir konuşma sinyalinin sesli ve sessiz bölümlerinin ayrımında (Bachu & ark., 2010), bir ortamadaki konuşmayı gürültüden ayırmada da (O'Shaughnessy, 2023) kullanılmaktadır.

Sonuç ve Öneriler

Bu çalışmada, Konuşma Duygu Tanımada ses sinyalindeki spektral ve zamansal akustik tanımlayıcılardan yaygın kullanımı olan özellikler incelenmiş, bu özelliklerin matematiksel tanımları, ayırt edici özellikleri ve KDT'deki rolleri ele alınmıştır. Mel-Frekans Cepstral Katsayıları, Spektral Ağırlık Merkezi, Spektral Düşüm Frekansı, Spektral Bant Genişliği, Spektral Akı ve Spektral Kontrast dahil olmak üzere spektral tanımlayıcılar, akustik enerjinin spektrum boyunca nasıl dağıldığını ve spektral zarfların nasıl değiştiğini yakalayarak duygusal durumları ayıran frekans alanı özelliklerini ortaya çıkarır. Zamansal tanımlayıcılardan özellikle KOE ve SGO, duygusal uyarılmayı yansıtan genlik dalgalanmalarını ve sinyal geçişlerini nicelleştirerek zaman alanı dinamiklerini ortaya çıkarır. Analiz, spektral özelliklerin frekans içeriği analizi yoluyla her bir duyguyu karakterize eden akustik özellikleri belirlediğini, zamansal özelliklerin ise bu özelliklerin dinamik ilerleme örüntüleri yoluyla zaman içinde nasıl evrimleştiğini izlediğini göstermektedir. Etkili bir duygu sınıflandırması, her iki tanımlayıcı türünün de entegre edilmesini gerektirmektedir. Spektral ve zamansal tanımlayıcılar arasındaki bu tamamlayıcı ilişki, makine öğrenimi tabanlı duygu sınıflandırma sistemlerinde, sınıflandırma doğruluğunu ve sistem performansını önemli ölçüde artırmaktadır.

Kaynakça

Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review. *IEEE Access*, 10, 122136-122158.

Aldarmaki, H., Ullah, A., Ram, S., & Zaki, N. (2022). Unsupervised Automatic Speech Recognition: A review. *Speech Communication*, 139, 76-91.

Bachu, R. G., Kopparthi, S., Adapa, B., & Barkana, B. D. (2010). Voiced/Unvoiced Decision for Speech Signals Based on Zero-Crossing Rate and Energy. *Advanced Techniques in Computing Sciences and Software Engineering*, 279-282.

Barhoumi, C., & BenAyed, Y. (2025). Real-time speech emotion recognition using deep learning and data augmentation. *Artificial Intelligence Review*, 58(2), 49.

Bhangale, K., & Kothandaraman, M. (2023). Speech Emotion Recognition Based on Multiple Acoustic Features and Deep Convolutional Neural Network. *Electronics* 2023, 12(4), 839.

Chandrasekar, P., Chapaneri, S., Jayaswal, D. (2014). Emotion Recognition from Speech using Discriminative Features. *International Journal of Computer Applications*, 101(16), 975-8887.

Dair, Z., Donovan, R., & O'reilly, R. (2021). Linguistic and Gender Variation in Speech Emotion Recognition Using Spectral Features. *29th Irish Conference on Artificial Intelligence and Cognitive Science*.

Davis, S. B., & Mermelstein, P. (1980). Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), 357-366.

Gabor, D. (1946). Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, 93(26). <https://doi.org/10.1049/JI-3-2.1946.0074>

Gingras, B., Marin, M. M., & Fitch, W. T. (2014). Beyond intensity: Spectral features effectively predict music-induced subjective arousal. *Quarterly Journal of Experimental Psychology*, 67(7), 1428-1446. <https://doi.org/10.1080/17470218.2013.863954>

Hamming, R. W. (1983). *Digital filters*. Prentice-Hall.

Jena, S., Basak, S., Agrawal, H., Saini, B., Gite, S., Kotecha, K., & Alfarhood, S. (2025). Developing a negative speech emotion recognition model for safety systems using deep learning. *Journal of Big Data*, 12(1).

Jiang, D. N., Lu, L., Zhang, H. J., Tao, J. H., & Cai, L. H. (2002). Music type classification by spectral contrast feature. *Proceedings - 2002 IEEE International Conference on Multimedia and Expo, ICME 2002*. <https://doi.org/10.1109/ICME.2002.1035731>

Jordan, E., Terrisse, R., Lucarini, V., Alrahabi, M., Krebs, M. O., Desclés, J., & Lemey, C. (2025). Speech Emotion Recognition in Mental Health: Systematic Review of Voice-Based Applications. *JMIR Mental Health*, 12, e74260. <https://doi.org/10.2196/74260>

Jothimani, S., & Premalatha, K. (2022). MFF-SAUG: Multi feature fusion with spectrogram augmentation of speech emotion recognition using convolution neural network. *Chaos, Solitons & Fractals*, 162, 112512.

Kumar, S., & Thiruvenkadam, S. (2021). An Analysis of the Impact of Spectral Contrast Feature in Speech Emotion Recognition. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, 9(2), 87-95.

Lech, M., Stolar, M., Best, C., & Bolia, R. (2020). Real-Time Speech Emotion Recognition Using a Pre-trained Image Classification Network: Effects of Bandwidth Reduction and Companding. *Frontiers in Computer Science*, 2.

Lin, C. H., Liao, W. K., Hsieh, W. C., Liao, W. J., & Wang, J. C. (2014). Emotion Identification Using Extremely Low Frequency Components of Speech Feature Contours. *The Scientific World Journal*. <https://doi.org/10.1155/2014/757121>

Liu, T., & Yuan, X. (2023). Paralinguistic and spectral feature extraction for speech emotion classification using machine learning techniques. *Eurasip Journal on Audio, Speech, and Music Processing*, 2023(1). <https://doi.org/10.1186/s13636-023-00290-x>

Liu, X., Xu, Y., Alter, K., & Tuomainen, J. (2018). Emotional connotations of musical instrument timbre in comparison with emotional speech prosody: Evidence from acoustics and event-related potentials. *Frontiers in Psychology*, 9.

Ma, Y., Zhang, Y., Fu, D., Zubicueta Portales, S., Kragic, D., & Fjeld, M. (2025). Advancing User-Voice Interaction: Exploring Emotion-Aware Voice Assistants Through a Role-Swapping Approach. *Lecture Notes in Computer Science*, 15802 LNCS, 303-320.

Mashhadi, M. M. R., & Osei-Bonsu, K. (2023). Speech emotion recognition using machine learning techniques: Feature extraction and comparison of convolutional neural network and random forest. *PLOS ONE*, 18(11), e0291500. <https://doi.org/10.1371/JOURNAL.PONE.0291500>

Nano Banana Pro (2025). *Görüntü Oluşturma*. (12.12.2025 tarihinde <https://gemini.google.com/> adresinden ulaşılmıştır).

O'Shaughnessy, D. (2023). Review of analysis methods for speech applications. *Speech Communication*, 151, 64-75.

Raisi, K., & Mourhir, A. (2024). Real-Time Speech Emotion Recognition. *6th International Conference on Intelligent Computing in Data Sciences, ICDS 2024*.

Rakibul Hasan, M., Mahbub Hasan, M., Zakir Hossain, M., & Md Mahbub Hasan, C. (2021). How many Mel-frequency cepstral coefficients to be utilized in speech recognition? A study with the Bengali language. *The Journal of Engineering*, 2021(12), 817-827.

Shoiynbek, A., Kozhakhmet, K., Sultanova, N., & Zhumaliyeva, R. (2019). The robust spectral audio features for speech emotion recognition. *Applied Mathematics and Information Sciences*, 13(5), 867-870. <https://doi.org/10.18576/AMIS/130521>

Sievers, B., Lee, C., Haslett, W., & Wheatley, T. (2019). A multi-sensory code for emotional arousal. *Proceedings of the Royal Society B*, 286(1906). <https://doi.org/10.1098/RSPB.2019.0513>

Singh, J., Saheer, L. B., & Faust, O. (2023). Speech Emotion Recognition Using Attention Model. *International Journal of Environmental Research and Public Health* 2023, 20(6), 5140.

Toleu, A., Tolegen, G., Mussabayev, R., Krassovitskiy, A., & Zhumazhanov, B. (2024). Comparative Analysis of Audio Features for Unsupervised Speaker Change Detection. *Applied Sciences*, 14(24), 12026. <https://doi.org/10.3390/app142412026>

Van, L. T., Le Dao, T. T., Le Xuan, T., & Castelli, E. (2022). Emotional Speech Recognition Using Deep Neural Networks. *Sensors (Basel, Switzerland)*, 22(4), 1414.

Xian, W. (2022). Speech Emotion Recognition Application for Education. *BCP Education & Psychology*, 7, 378-383. <https://doi.org/10.54691/BCPEP.V7I.2691>

Zeng, C., Li, J., & Habibi, A. (2025). Speech emotion recognition based on a stacked autoencoders optimized by PSO based grass fibrous root optimization. *Scientific Reports* 2025, 15(1), 26158. <https://doi.org/10.1038/s41598-025-08703-x>

BÖLÜM 12

SANAL AGORA: SOSYAL SANAL GERÇEKLİK TABANLI ÖĞRENME VE MÜZAKERE ALANLARI

Ebru YILMAZ İNCE¹

Murat İNCE²

Giriş

Çevrimiçi öğrenme ortamları uzun süredir içerik erişimi ve ölçeklenebilirlik avantajları sunsa da, sosyal etkileşim kalitesi, sınıf içi birliktelik hissi ve işbirlikli öğrenme süreçlerini derinleştirme konusunda sınırlılıklar taşır. Son yıllarda sosyal sanal gerçeklik ve metaverse tabanlı ortamlar, kullanıcıların avatarlar aracılığıyla aynı mekânı paylaştığı algısını güçlendirerek sosyal bulunuşluğu artırabilen yeni bir etkileşim katmanı sağlamaktadır. Nitekim pandemi döneminde üniversite seminerlerinde hazır bir sosyal sanal gerçeklik platformunun kullanımı, öğrencilerin video-konferansa kıyasla daha doğal/spontane iletişim, daha güçlü sosyal bulunuşluk ve topluluk hissi yaşadığını göstermiştir (Barreda-Ángeles et al., 2023).

¹ Doç.Dr., Isparta Uygulamalı Bilimler Üniversitesi, Orcid: 0000-0001-9462-0363

² Doç.Dr., Isparta Uygulamalı Bilimler Üniversitesi, Orcid: 0000-0001-5566-5008

Bu bağlamda sanal agora, yalnızca üç boyutlu bir toplantı alanı değil; müzakere, ortak ürün üretimi ve topluluk temelli öğrenme pratiklerini desteklemek üzere tasarlanmış, mekânsal olarak oluşturulmuş sosyal sanal gerçeklik öğrenme ortamıdır. Sanal agora yaklaşımı, çevrimiçi öğrenmenin kuramsal açıklamalarından Community of Inquiry (CoI) çerçevesi ile uyumlu biçimde sosyal, bilişsel ve öğretimsel bulunuşluğu birlikte güçlendirmeyi hedefler (Garrison et al., 1999).

Kuramsal Çerçeve

CoI modeli, çevrimiçi öğrenmede anlamlı öğrenme deneyiminin bilişsel bulunuşluk, sosyal bulunuşluk ve öğretimsel bulunuşluk etkileşimiyle oluştuğunu ileri sürer (Garrison et al., 1999). Sanal agora tasarımında bu üç bileşen, (i) tartışmayı yapılandıran öğretimsel/moderatör desteği, (ii) avatar temelli etkileşimle sosyal bağ kurma, (iii) kanıta dayalı tartışma ve ortak ürün üretimiyle bilişsel süreçlerin derinleştirilmesi üzerinden somutlanır.

Sosyal sanal gerçekliğin eğitimdeki etkisi, daha sürükleyici olmanın ötesinde, etkileşim kalitesi ve bilişsel katılım ile ilişkilidir. Örneğin, işbirlikli sanal gerçeklik öğrenme ortamında sosyal bulunuşluk göstergelerinin (özellikle açık iletişim) aktif-yapıcı-etkileşimli bilişsel katılım modlarıyla anlamlı biçimde ilişkili olduğu rapor edilmiştir (Dunmoye et al, 2024). Bu bulgu, sanal agora tasarımında “söz hakkı, tartışma protokolleri, rol dağılımı, ortak görev” gibi etkileşim düzeneklerinin kritik olduğunu gösterir.

Eğitimsel metaverse literatüründe, öğrenme hedefleriyle uyumlu tasarım, etkileşim/geri bildirim mekanizmaları, teknik erişilebilirlik ve güvenlik boyutları öne çıkmaktadır. Eğitimsel metaverse tasarımına yönelik bir literatür incelemesi, uygulamaların oyun/simülasyon temelli etkileşimlerini değerlendirme boyutlarıyla

birlikte ele alarak tasarım için yol gösterici ilkeler sunmaktadır (Ak et al, 2025).

Sanal Agora Tasarımı: Uygulama Bağlamı ve Bileşenler

Sanal agora, katılımcıların avatarlarla eşzamanlı katıldığı, küçük grup adalarında işbirliği yürüttüğü ve merkez alanda ortak çıktıları paylaştığı sosyal sanal gerçeklik tabanlı bir öğrenme–müzakere ortamı olarak ele alınır. Bu yaklaşım, kolay uygulanabilir sosyal VR kullanımının sosyal bulunuşluğu güçlendirdiğini gösteren bulgularla uyumludur (Barreda-Ángeles et al., 2023).

Sanal agora tasarımında mekân, pedagojik amaçlara göre bölümlenmelidir:

Merkez alan: açılış–kapanış, genel tartışma, sunum ve ortak karar (Şekil 1)

Toplantı alanı: küçük grup problem çözme/argümantasyon (Şekil 2)

Sergileme–paylaşım alanı: poster, kavram haritası, kaynak panosu (Şekil 3)

Moderasyon noktası: söz hakkı, zaman yönetimi, tartışma protokolleri

Bu bölümler, işbirlikli sanal gerçeklik öğrenmede etkileşimi tesadüfi olmaktan çıkarıp yapılandırmaya dönük tasarım önerileriyle tutarlıdır (Dunmoye et al, 2024).

Şekil 1. Sanal agora merkezi alan



Şekil 2. Sanal agora toplantı alanı



Şekil 3. Sanal agora sergileme–paylaşım alanı



Sanal agoranın etkisi, mekân kadar rol–senaryo tasarımına bağlıdır: moderatör, konuşmacı, tartışmacı, raportör gibi roller; tartışma protokolleri; ortak ürün hedefleri (ör. politika notu, proje taslağı, kavram haritası) netleştirilmelidir. Bu yapı, sanal gerçeklik çalışmalarında gözlenen daha doğal iletişim ve topluluk hissi bulgularını pedagojik çıktıya bağlamaya yardımcı olur (Barreda-Ángeles et al., 2023).

Sanal agora değerlendirmesinde tek bir başarı testi yeterli değildir. Önerilen çoklu değerlendirme paketi:

Ürün değerlendirmesi: rubrik (argümantasyon kalitesi, kanıt kullanımı, tutarlılık)

Süreç değerlendirmesi: katılım yoğunluğu, rol üstlenme, konuşma sırası, grup içi etkileşim

Algı ölçekleri: sosyal bulunuşluk, öğretimsel destek, bilişsel süreç göstergeleri

Etik, Mahremiyet ve Yönetişim

Sosyal sanal gerçeklik ortamları; ses, hareket ve yakınlık gibi etkileşimsel veriler nedeniyle mahremiyet riskleri doğurur. Sanal gerçekliğin sorumlu kullanımı için çerçeveler, özellikle rıza, mahremiyet ve zarar risklerini sistematik biçimde ele almayı önerir (Raja et al., 2025). Bu nedenle sanal agora uygulamalarında (i) veri minimizasyonu, (ii) açık bilgilendirme ve rıza, (iii) kayıt/saklama politikası, (iv) davranış kuralları ve ihlal müdahale protokolü standartlaştırılmalıdır.

Sanal agora kavramı, eğitimde işbirliği yanında müzakere kalitesi ve katılımın örgütlenmesi açısından da ele alınabilir. Online deliberasyon tasarımında amaç, hedef kitle, zaman-mekân kurgusu ve moderasyon biçimi gibi tasarım boyutlarının çıktıları etkilediğine dair birikim bulunmaktadır (Muhlberger, 2005).

Sonuç

Sanal agora, sosyal sanal gerçekliğin topluluk hissini artıran etkileşim doğasını, CoI'nin üçlü bulunuşluk bileşenleriyle hizalayarak çevrimiçi öğrenmede katılım ve işbirliğini güçlendirebilecek bir tasarım yaklaşımı sunar (Garrison et al., 1999).

Ancak etkili sonuçlar, yalnızca sanal gerçeklik kullanımı ile değil; rol ve senaryo tasarımı, moderasyon, mekânsal zonlama ve çoklu değerlendirme/analitik bileşenlerinin birlikte kurgulanmasıyla elde edilir. Son olarak, rıza–mahremiyet–zararın önlenmesi odaklı etik çerçeveler, sanal agora uygulamalarının sürdürülebilirliği ve güvenli kullanımı için zorunlu bir yönetim katmanı olarak ele alınmalıdır.

Kaynakça

Ak, O., Şen Akbulut, M., Soydan, S., & Yildirim, E. (2025). Designing educational metaverses: a literature review to determine guidelines. *Interactive Learning Environments*, 33(2), 925-943.

Barreda-Ángeles, M., Horneber, S., & Hartmann, T. (2023). Easily applicable social virtual reality and social presence in online higher education during the covid-19 pandemic: A qualitative study. *Computers & Education: X Reality*, 2, 100024.

Dunmoye, I. D., Rukangu, A., May, D., & Das, R. P. (2024). An exploratory study of social presence and cognitive engagement association in a collaborative virtual reality learning environment. *Computers & Education: X Reality*, 4, 100054.

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *Internet and Higher Education*, 2(2–3), 87–105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)

Muhlberger, P. (2005). The Virtual Agora Project: A research design for studying democratic deliberation. *Journal of Deliberative Democracy*, 1(1).

Raja, U. S., & Al-Baghli, R. (2025). Ethical concerns in contemporary virtual reality and frameworks for pursuing responsible use. *Frontiers in Virtual Reality*, 6, 1451273.

BÖLÜM 13

Yapay Zekâda Veri Kıtlığına Dayanıklı Öğrenme Paradigmaları: Örneksiz Öğrenme, Az Örnekle Öğrenme ve Sentezlenmiş Veri Yaklaşımları

AYŞE GÜL EKER¹

Giriş

Makine öğrenmesi ve derin öğrenme alanındaki klasik yaklaşım, yüksek performanslı modellerin büyük miktarda etiketlenmiş veriye ihtiyaç duyduğu varsayımına dayanmaktadır. Ancak gerçek dünyadaki birçok uygulamada bu varsayım geçerli değildir. Etiketli verinin toplanması maliyetli, zaman alıcı veya etik açıdan sorunlu olabilir, bazı alanlarda ise veri doğası gereği son derece sınırlıdır. Bu durum, literatürde sıklıkla veri kıtlığı problemi olarak ele alınmakta ve yapay zekâ sistemlerinin ölçeklenebilirliğini doğrudan etkilemektedir (Halevy, Norvig, & Pereira, 2009).

Son yıllarda geliştirilen büyük ölçekli ön eğitilmiş modeller, bu veri merkezli yaklaşımı önemli ölçüde dönüştürmüştür. Özellikle dil, görüntü ve çok modlu alanlarda eğitilmiş modeller, göreve özgü büyük veri kümeleri olmadan da anlamlı çıktılar üretebilmektedir.

¹Dr., Kocaeli Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0003-0721-2631
--243--

Bu gelişme, yapay zekâda öğrenmenin yalnızca doğrudan eğitim verisine değil, modelin ön eğitim sürecinde edindiği genel temsil gücüne ve bağlamsal çıkarım yeteneğine dayandığını göstermiştir. Örneksiz (zero-shot) ve az örnekli (few-shot) öğrenme, büyük ölçekli dil modellerinin bağlam içi öğrenme yetenekleri sayesinde ortaya çıkan dönüşümün en görünür örnekleri arasındadır (Brown ve diğ., 2020).

Az ve sıfır örnekle öğrenme yaklaşımları, modelin yeni bir görevi ya hiç örnek görmeden ya da yalnızca birkaç örnek üzerinden çözebilmesini hedefler. Bu tür yaklaşımlar, klasik anlamda yeniden eğitim sürecine dayanmaz; bunun yerine modelin mevcut bilgisini bağlam aracılığıyla kullanmasını esas alır. Literatürde bu mekanizma çoğunlukla bağlama dayalı genelleme (in-context learning) olarak adlandırılmaktadır. Bu durum, öğrenmenin model parametrelerinin güncellenmesinden ziyade, çıkarım sürecinde gerçekleştiği yeni bir paradigmayı işaret etmektedir (Bommasani ve diğ., 2021).

Veri kıtlığına yönelik bir diğer önemli yaklaşım ise sentezlenmiş veri kullanımıdır. Üretici modeller aracılığıyla oluşturulan yapay veriler, gerçek verinin sınırlı olduğu veya erişiminin kısıtlandığı senaryolarda model eğitimi ve değerlendirmesi için alternatif bir kaynak sunmaktadır. Ancak bu yaklaşım, veri gerçekliği, önyargıların çoğaltılması ve güvenilirlik gibi yeni sorunları da beraberinde getirmektedir. Bu nedenle az ve sıfır örnekle öğrenme yaklaşımları ile sentezlenmiş veri tabanlı yöntemler, yalnızca performans açısından değil, sorumlu ve güvenilir yapay zekâ bağlamında da dikkatle ele alınmalıdır.

Bu bölümde, yapay zekâda veri kıtlığına dayanıklı öğrenme paradigmaları bütüncül bir bakış açısıyla incelenmektedir. Örneksiz ve az örnekle öğrenme yaklaşımları, ön eğitilmiş modeller ve bağlama dayalı genelleme çerçevesinde ele alınırken; sentezlenmiş veri kullanımı, bu paradigmanın tamamlayıcı bir bileşeni olarak

tartışılmaktadır. Amaç, veri miktarının sınırlı olduğu koşullarda yapay zekâ sistemlerinin nasıl anlamlı, güvenilir ve uygulanabilir çözümler üretebildiğini ortaya koymaktır.

Örneksiz ve Az Örnekle Bağlama Dayalı Genelleme

Klasik makine öğrenmesi ve derin öğrenme yaklaşımlarında, yeni bir görevin başarılı bir şekilde çözülmesi genellikle bu göreve ait büyük miktarda etiketlenmiş eğitim verisine dayanır. Ancak etiketlenmiş veri toplamının maliyetli, zaman alıcı ve kimi uygulama alanlarında etik açıdan sınırlı olması, bu varsayımı önemli ölçüde kısıtlamaktadır. Bu veri kıtlığı problemi, yapay zekâ sistemlerinin gerçek dünya uygulamalarında karşılaştığı temel zorluklardan biridir ve bu zorluklara yanıt olarak geliştirilen yöntemler önemli bir araştırma odağı hâline gelmiştir (Ramesh ve diğ., 2025).

Bu bağlamda, örneksiz öğrenme ve az örnekle öğrenme yaklaşımları, eğitim sırasında sınırlı veya hiç örnek verilmemiş yeni görevlere yönelik genelleme becerisini artırmayı hedefler. Bu yöntemler, modellerin yalnızca eğitim verisine değil, önceden öğrenilmiş temsil gücüne ve bağlam içi çıkarım yeteneğine dayanarak karar üretmesini sağlar

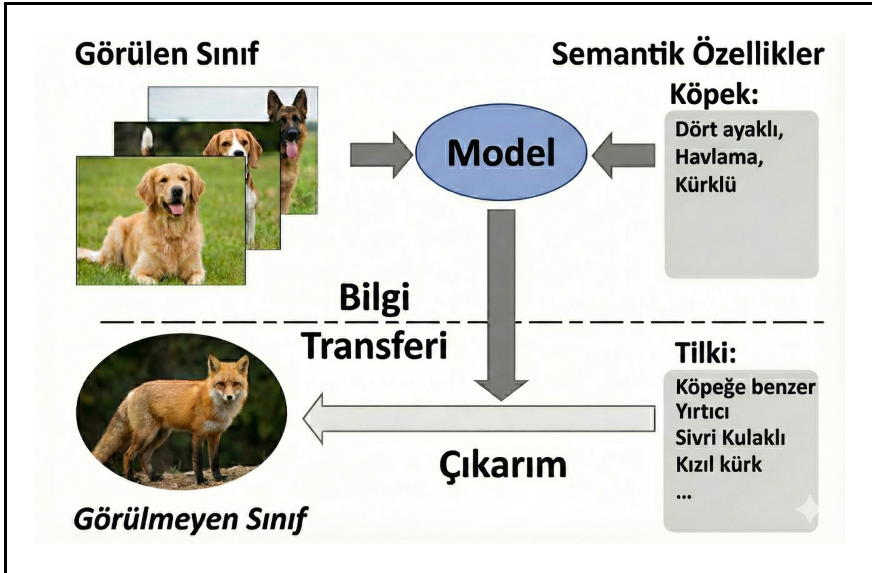
1. Örneksiz Öğrenme

Örneksiz öğrenme, modelin belirli bir görev veya sınıf için hiç etiketlenmiş örnek görmemiş olmasına rağmen bu görevi başarıyla yerine getirebilmesini ifade eder. Bu yaklaşımın temelinde, modelin eğitim aşamasında edindiği genel dil bilgisi, kavramsal temsiller ve semantik ilişkileri daha önce karşılaşmadığı görevler veya sınıflar için genelleyebilme yeteneği yer alır. Özellikle büyük ölçekli ön eğitim süreçlerinden geçen modeller, farklı kavramlar

arasındaki ilişkileri yalnızca yüzeysel istatistikler üzerinden değil, daha soyut anlamsal düzeyde öğrenebilir.

Örneğin, bir metin sınıflandırma modeli eğitim sürecinde “spor” ve “ekonomi” kategorilerine ait haberlerle eğitilmiş olabilir. Model, daha önce hiç “teknoloji” etiketiyle karşılaşmamış olsa bile; yazılım, donanım, yapay zekâ, dijitalleşme gibi kavramların içerdiği anlamsal örüntüleri kullanarak bir haberin teknoloji kategorisine ait olduğunu doğru biçimde tahmin edebilir. Bu durumda model, yeni sınıfı doğrudan öğrenmek yerine, mevcut kavramsal uzayda bu sınıfı konumlandırarak çıkarım yapar.

Resim 1: Örneksiz öğrenme yaklaşımının kavramsal gösterimi



Resim 1’de, örneksiz öğrenme yaklaşımına ait başka bir sezgisel örnek sunulmaktadır. Modelin, eğitim sırasında gördüğü sınıflardan(köpek) edindiği anlamsal temsiller aracılığıyla daha önce hiç örnek görmediği sınıflara(tilki) ilişkin çıkarım yapabildiği gösterilmektedir.

Bu tür bir genelleme yeteneđi, önceden eğitilmiş büyük modellerin bağlama dayalı öğrenme ve doğal dilde verilen görev tanımlarını yorumlayabilme kabiliyeti sayesinde mümkün olmaktadır. Bağlam içine yalnızca görev tanımı veya etiketlerin doğal dil açıklamaları verildiğinde, model eğitim sırasında öğrendiđi temsilleri kullanarak yeni görevlere uyum sağlayabilir. Bu yönüyle örneksiz öğrenme, klasik denetimli öğrenme yaklaşımlarından ayrılır ve özellikle veri etiketlemenin maliyetli veya mümkün olmadığı senaryolarda önemli avantajlar sunar.

2. Az Örnekle Öğrenme

Az örnekle öğrenme, yeni bir görev için modele yalnızca sınırlı sayıda etiketli örnek sunulduđu durumları ifade eder. Bu yaklaşımda model, ilgili göreve özgü bir ince ayar (fine-tuning) sürecinden geçmez; bunun yerine, bağlam içinde verilen küçük örnek setini referans olarak yeni girdiler üzerinde çıkarım yapar. Bu süreçte modelin parametreleri sabit tutulur ve öğrenme, tamamen bağlama eklenen örneklerin yönlendirici etkisiyle gerçekleşir.

Modern büyük dil modellerinde bu mekanizma, bağlama dayalı çıkarım olarak adlandırılmakta ve yaygın biçimde kullanılmaktadır. Model, bağlamda sunulan örneklerdeki girdi çıktı ilişkisini geçici olarak içselleştirerek, benzer yapıya sahip yeni girdiler için uygun çıktılar üretir. Bu durum, klasik makine öğrenmesindeki yeniden eğitim veya ağırlık güncelleme süreçlerinden farklı olarak, çıkarım zamanında gerçekleşen bir uyum mekanizmasını temsil eder.

Örneđin, az örnekle öğrenme kapsamında bir büyük dil modeline film yorumlarının duygu durumunu belirlemeye yönelik olarak yalnızca birkaç etiketli örnek bağlam içinde sunulabilir. Bu örnekler modelin parametrelerinde herhangi bir güncelleme yapılmaksızın, yalnızca bağlama dayalı çıkarım yapmasını sağlar:

Yorum: “Film çok sürükleyiciydi ve izlerken hiç sıkılmadım.”

Duygu: Olumlu

Yorum: “Senaryo son derece sıkıcı ve tahmin edilebilirdi.”

Duygu: Olumsuz

Yorum: “Oyunculuk fena değildi ama beklentimi de karşılamadı.”

Duygu: Nötr

Bu örneklerin ardından modele

“Film gereğinden uzundu ve özgün bir yönü yoktu.”

yorumu verildiğinde, model ek bir eğitim süreci olmadan, bağlamda sunulan örnekleri referans alarak bu yorumu olumsuz duygu sınıfına atayabilir.

Az örnekle öğrenmenin başarısı, modelin önceden eğitildiği geniş ve çeşitli veri dağılımı üzerinden kazandığı genel dil ve dünya bilgisi temsillerini, göreve özgü bağlamla etkili biçimde birleştirebilme yeteneğine dayanır. Bu yaklaşım, klasik denetimli öğrenme yöntemlerine kıyasla çok daha az veri gerektirir ve genellikle basit, sınırlı örneklerle bile anlamlı genelleme yapılmasını mümkün kılar. Bu nedenle, az örnekle öğrenme özellikle hızlı prototipleme, veri kısıtlı senaryolar ve esnek görev tanımları açısından önemli bir paradigma olarak öne çıkmaktadır.

Örneksiz öğrenme ve az örnekle öğrenme, farklı veri yaklaşımlarına dayanıyor gibi görünse de aralarındaki fark, verinin model davranışına nasıl dâhil edildiği ile ilgilidir. Örneksiz öğrenmede model hiç örnek görmezken, az örnekle öğrenmede model çok sınırlı sayıda örneği bağlam içinde kullanır. Her iki durumda da öğrenme, doğrudan parametre güncellemesiyle değil, modelin bağlam içi çıkarım yeteneği ile gerçekleşir. Bu bağlamda bu

yaklaşımlar, veri miktarına değil, önceden öğrenilmiş temsillerin bağlama dayalı genelleme kapasitesine dayanır

3. Uygulama Alanları ve Örnekler

Örneksiz öğrenme ve az örnekle öğrenme paradigmaları, klasik öğrenme yöntemlerinin başarısız kaldığı veya veri elde etmenin zor olduğu birçok gerçek dünya probleminde etkili çözümler sunmaktadır. Bu bölümde, bu öğrenme yaklaşımlarının doğal dil işleme, görsel tanıma ve biyomedikal görüntüleme gibi çeşitli alanlardaki uygulamaları incelenmektedir.

Örneğin doğal dil işleme (NLP) alanında, ön eğitilmiş büyük dil modelleri ile gerçekleştirilmiş uygulamalar, klinik metinlerde bilgi çıkarımı ve ilişki tanıma gibi görevlerde örneksiz ve az örnekle öğrenme stratejilerinin performansını göstermektedir. Klinik bilgi çıkarımı üzerine yapılan deneysel çalışmalarda, farklı istem (prompt) tasarımlarının az örnekle öğrenme bağlamında etkili olduğu ve bu tekniklerin özellikle klinik anlam belirsizliğinin giderimi, ilaç durumu çıkarımı gibi karmaşık NLP görevlerinde yüksek başarı sağladığı raporlanmıştır. Bu sonuçlar, bağlama dayalı öğrenmenin NLP uygulamalarında veri ihtiyacını önemli ölçüde azalttığını göstermektedir (Sivarajkumar ve diğ., 2024).

Görsel tanıma ve bilgisayarlı görü uygulamalarında da az örnekle öğrenme ve örneksiz öğrenme teknikleri, sınırlı etiketlenmiş veriyle yeni nesne kategorilerini tanımada kullanılmaktadır. Örneğin tıbbi görüntü sınıflandırmasında, MRI beyin tümörü gibi nadir görülen durumları tanımak için az örnekle öğrenme yöntemleri kullanılmaktadır. Bir çalışmada, özellikle metrik tabanlı az örnekle öğrenme yaklaşımlarının, geleneksel ince ayarlı modellere kıyasla çok sınırlı örnek sayısı ile dahi daha dayanıklı ve güvenilir sınıflandırma performansı sunduğu gösterilmiştir. Buna karşılık, örneksiz öğrenme yaklaşımlarının esneklik sağlamakla birlikte, tıbbi görüntü alanında performanslarının halen sınırlı olduğu

vurgulanmıştır (Das & Singh, 2025). Benzer biçimde biyomedikal görüntüleme alanında yapılan başka bir çalışmada 3B tıbbi görüntü segmentasyonu için ön eğitilmiş modellerin az örnekle adaptasyon yetenekleri araştırılmıştır. Az örnekle öğrenmeye dayalı, eğitim gerektirmeyen adaptasyon yaklaşımlarının; segmentasyon gibi karmaşık 3B tıbbi görüntüleme görevlerinde, geniş etiketli veri kümelerine ve manuel uzman müdahalesine ihtiyaç duymadan etkili ve kararlı performans sağlayabildiği görülmüştür. Bu yaklaşım, klinik ortamlarda hızlı prototipleme ve otomatik tanı araçlarının geliştirilmesi açısından önemli bir potansiyel sunmaktadır (He ve diğ., 2025).

Az ve örneksiz öğrenmenin önemli uygulama alanlarından bir diğeri ise görsel dil modellerinin (vision language models) desteklediği görevlerdir. Bu modeller, görsel ve metinsel temsilleri birlikte öğrenmeleri sayesinde, bir görseldeki nesnelerin sınıflandırılması veya tanımlanması gibi görevlerde daha önce hiç görülmemiş nesne kategorilerini örneksiz olarak tanıyabilmektedir. Görsel dil entegrasyonuna dayalı bu yaklaşımlar, özellikle büyük ölçekli veri toplamanın maliyetli olduğu senaryolarda esnek ve etkili genelleme yetenekleri sunabilmektedir. Nitekim yakın tarihli bir çalışmada, CLIP tabanlı görsel dil modellerinin uzaktan algılama sahnelerini alana özgü herhangi bir ek eğitim gerektirmeden, örneksiz öğrenme kapsamında yüksek doğrulukla sınıflandırabildiği deneysel olarak gösterilmiştir (Al Rahhal ve diğ., 2023). Benzer şekilde, büyük dil modelleri tarafından üretilen sınıf açıklamalarıyla uyarlanan CLIP benzeri görsel dil modellerinin, daha önce görülmemiş ince ayrımlı nesne kategorilerinde örneksiz öğrenme performansını anlamlı biçimde artırdığı sunulmuştur (Saha ve diğ., 2024). Ayrıca, yalnızca sınıf etiketini kullanarak ve herhangi bir görsel örneğe ihtiyaç duymadan, bir örneğin tek bir hedef sınıfa ait olup olmadığını ayırt edebilen görsel-dil tabanlı örneksiz öğrenmeli

tek sınıflı sınıflandırma yaklaşımının pratik ve yüksek performanslı biçimde uygulanabildiği gösterilmiştir (Bendou ve diğ., 2024).

Bu örnekler, veri kıtlığının doğal bir sınırlama olduğu uygulama alanlarında az örnekle öğrenme ve örneksiz öğrenme yöntemlerinin performansı artırdığını ve veri maliyetini azalttığını göstermektedir. Bununla birlikte, bu paradigmaların etkin biçimde uygulanabilmesi; modelin ön eğitim kapsamı, bağlam tasarımı ve görevle ilişkili anlamsal bilginin dikkatli biçimde kurgulanmasını gerektirir. Aksi hâlde, veri kıtlığını aşma amacıyla kullanılan örneksiz öğrenme ve az örnekle öğrenme yaklaşımları; yetersiz bağlamlandırma, hatalı anlamsal eşleştirme veya uygunsuz yönlendirme nedeniyle performans düşüşlerine yol açabilmekte, bu durum ise sistemin güvenilirliği ve genelleme kapasitesini olumsuz etkilemektedir (Ramesh ve diğ., 2025).

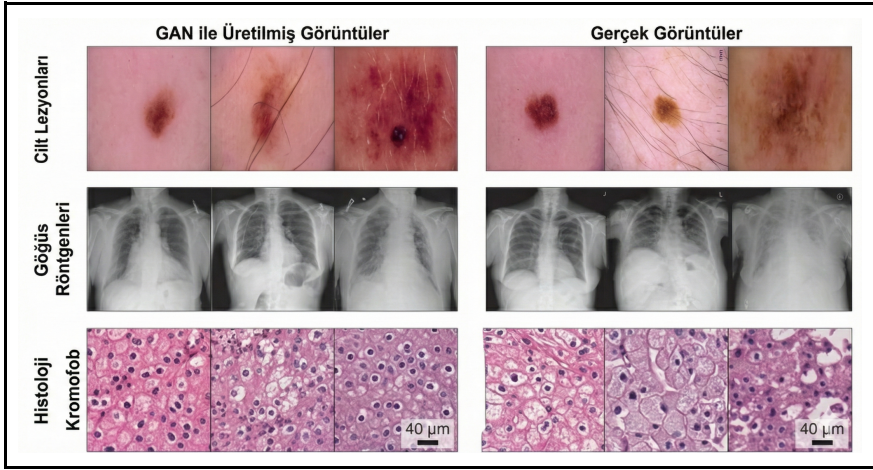
Sentezlenmiş Veri Üretim Yaklaşımları

Veri kıtlığına dayanıklı öğrenme paradigmalarının bir diğer önemli bileşeni, sentezlenmiş veri kullanımıdır. Sentezlenmiş veri, gerçek dünyadan doğrudan toplanmamış; bunun yerine istatistiksel modeller, üretici yapay zekâ sistemleri veya simülasyonlar aracılığıyla oluşturulmuş yapay veri örneklerini ifade eder. Bu yaklaşım, özellikle gerçek veriye erişimin sınırlı, pahalı veya etik açıdan sorunlu olduğu alanlarda, model geliştirme sürecine alternatif bir çözüm sunmaktadır.

Özellikle üretici çekişmeli ağlar (Generative Adversarial Networks, GAN'lar), difüzyon modeller ve büyük dil modelleri, gerçek veri dağılımının istatistiksel özelliklerini öğrenerek yüksek kalitede sentetik örnekler üretebilmektedir. GAN yaklaşımı, üretici ve ayırt edici olmak üzere iki nöral ağın karşılıklı rekabetiyle eğitilmesi esasına dayanarak, eğitim veri setinin altında yatan dağılımı taklit eden yeni örnekler üretmektedir. Bu sayede özellikle

görüntü ve diğer yapısal veri türlerinde gerçekçi sentetik veri üretimi mümkün olmakta; literatürde, bu yaklaşımın makine öğrenimi modellerinde veri yetersizliği problemini azaltmada ve sınıflandırma, tespit gibi görevlerde performansı iyileştirmede etkili olduğu yaygın biçimde rapor edilmektedir (Figueira & Vaz, 2022; Kumar ve diğ., 2025).

Resim 2: GAN ile üretilmiş sentetik medikal görüntü örnekleri



Resim 2’de tıbbi görüntüleme alanında GAN kullanılarak oluşturulan sentetik görüntülerden örnekler sunulmuştur. Farklı medikal modalitelerde üretilen bu sentetik veriler, gerçek görüntülere benzer istatistiksel özellikler taşıyarak veri artırımı ve model geliştirme süreçlerinde destekleyici bir rol üstlenmektedir.

Bunun yanında difüzyon modelleri, özellikle yüksek boyutlu veri üretimi ve örnekleme süreçlerinde doğrudan veri dağılımını modelleme yetenekleri sayesinde gelişmiş sentetik veri elde edebilmektedir; bu modeller, eğitildikleri dağılıma ilişkin örnekler üzerinden rastgele gürültüyü tersine çevirerek yeni ve benzer örnekler üretirler, bu da yüksek kaliteli ve çeşitlilik gösteren veri setleri oluşturulmasına imkân verir. Bu yetenek, son araştırmalarda

difüzyon tabanlı yöntemlerin örnek üretme ve dağılım öğrenme kapasiteleri üzerine yapılan incelemelerde vurgulanmıştır (Zhu, 2024).

Ayrıca büyük dil modelleri, geniş ölçekli ön eğitimleri sayesinde metin, kod ve yapılandırılmış veri türlerinde sentetik veri üretiminde önemli avantajlar sunmaktadır. Bu modeller, gerçek verinin dilsel ve yapısal örüntülerini öğrenebilmektedir. Bu sayede isteme dayalı (prompt-tabanlı) ya da geri kazanım destekli yaklaşımlar kullanılarak görev odaklı sentetik veri üretimi mümkün olmaktadır. Özellikle veri kıtlığının yaşandığı düşük kaynaklı senaryolarda, üretilen bu sentetik veriler gerçek verilerle birlikte kullanıldığında, model performansına anlamlı ve tamamlayıcı katkılar sağlayabilmektedir.

Sentezlenmiş verilerin kullanımı, özellikle sağlık alanında giderek artan bir ilgi odağı hâline gelmiştir. Hasta gizliliğinin korunması, kişisel sağlık bilgilerinin paylaşımına ilişkin yasal ve etik kısıtlamalar ile veri güvenliği gereksinimleri, gerçek tıbbi verilere erişimi çoğu zaman sınırlamaktadır. Bu bağlamda, sentetik hasta kayıtları ve tıbbi görüntüler; makine öğrenmesi modellerinin geliştirilmesi, veri artırımı ve algoritmaların ön değerlendirme süreçlerinde önemli bir alternatif olarak kullanılmaktadır. Yapılan çalışmalar, sentetik verilerle desteklenerek eğitilen modellerin, yalnızca gerçek verilerle eğitilen modellere yakın hatta bazı durumlarda daha yüksek performans sergileyebildiğini göstermektedir. Bununla birlikte, sentetik verilerin tek başına klinik karar destek sistemleri için yeterli olmadığı, klinik uygulamaya geçmeden önce modellerin mutlaka gerçek klinik verilerle doğrulanması ve regülasyonlara uygun biçimde değerlendirilmesi gerektiği vurgulanmaktadır (Chen ve diğ., 2021).

Finans alanında ise sentetik veri, dolandırıcılık tespiti ve risk modellemesi gibi nadir olayların öğrenilmesinde önemli bir rol oynamaktadır. Gerçek dünyada çok az gözlemlenen anomali türleri, sentetik olarak üretilerek modelin bu durumlara karşı daha duyarlı hâle getirilmesi sağlanabilmektedir. Bu yaklaşım, özellikle dengesiz veri setleriyle çalışılan senaryolarda model performansını artırmaktadır (Fiore ve diğ., 2019).

Bu gelişmeler, sentezlenmiş verinin yalnızca destekleyici bir rol oynamasını aşarak, belirli kullanım senaryolarında gerçek veriye yakın nitelikte ve doğrudan model eğitimi, test veya değerlendirme için kullanılabilecek birincil veri kaynağı olabilmesine zemin hazırlamıştır.

Güvenilirlik, Sınırlamalar ve Etik Tartışmalar

Veri kıtlığına dayanıklı öğrenme yaklaşımları ve sentezlenmiş veri üretimi, yapay zekâ sistemlerine esneklik kazandırsa da güvenilirlik ve etik açıdan önemli riskler barındırmaktadır. Özellikle örneksiz öğrenme ve az örnekle öğrenme gibi bağlama dayalı genelleme yöntemlerinde model tahminlerinin başarısı, verilen bağlamanın niteliğine ve ön-eğitim aşamasında öğrenilmiş temsillerin tutarlılığına sıkı şekilde bağlıdır.

Sentezlenmiş veri kullanımında temel bir soru, üretilen verinin gerçek dünya dağılımını ne ölçüde temsil ettiği ve bu temsilin güvenilir bir şekilde model performansına dönüştürülüp dönüştürülemeyeceğidir. Literatürde yayımlanan güncel sistematik incelemeler, GAN'lar gibi üretici yaklaşımların yanı sıra difüzyon modellerinin de yüksek boyutlu sentetik veri üretiminde etkili sonuçlar sunduğunu ortaya koymaktadır. Bununla birlikte, bu modellerin eğitim verilerinden kaynaklanan yapısal önyargıları öğrenerek üretilen sentetik verilere yansıtılma potansiyeline sahip olduğu da vurgulanmaktadır (Liu ve diğ., 2024). Bu durum, sentetik

verilerin özellikle adalet, güvenilirlik ve etik kullanım açısından dikkatle değerlendirilmesini gerekli kılan önemli bir risk unsuru olarak değerlendirilmektedir.

Sentetik verinin yüksek fayda potansiyeline rağmen, gizlilik, etik ve hesap verebilirlik boyutları kritik önemdedir. Yapay verinin özel hayatı ihlal etmeyecek şekilde üretilmesi hedeflense de sentetik verilerde yeniden tanımlanma risklerinin tamamen ortadan kaldırılamadığı literatürde vurgulanmaktadır (Boraschi ve diğ., 2025). Bu nedenle sentetik veri kullanımı, yalnızca model performansına odaklanmak yerine; veri kalitesi, gizlilik, genellenebilirlik ve yasal uyum gibi çoklu kriterlerin birlikte değerlendirildiği yönetim çerçeveleri içinde ele alınmalıdır.

Son olarak, üretici modellerin kendi ürettikleri sentetik verilerle kontrolsüz, naif ve tekrarlı biçimde eğitilmesi durumunda, zaman içinde ciddi kalite sorunları ortaya çıkabilmektedir. Bu tür bir eğitim süreci, literatürde model çöküşü (model collapse) olarak adlandırılan olgulara yol açarak, model çıktılarının hem çeşitliliğini hem de gerçek veri dağılımını temsil etme yeteneğini giderek zayıflatmaktadır. Buna ek olarak, bu süreç önyargıların pekişmesine ve beklenmedik, istenmeyen davranışların ortaya çıkmasına neden olabilmektedir. Son yıllarda bu risklere yönelik ilgi belirgin biçimde artmış; söz konusu olumsuz etkileri azaltmayı, önlemeyi veya tersine çevirmeyi amaçlayan yeni yöntem ve yaklaşımlar önerilmiştir (Alemohammad ve diğ., 2024).

Sonuç ve Gelecek Çalışmalar

Bu bölümde, yapay zekâ sistemlerinin veri kıtlığı koşullarında nasıl etkili biçimde öğrenebildiği; örneksiz öğrenme, az örnekle öğrenme ve sentezlenmiş veri yaklaşımları çerçevesinde ele alınmıştır. Büyük ölçekli ön eğitilmiş modellerin sunduğu bağlama dayalı genelleme yetenekleri, öğrenmenin yalnızca doğrudan eğitim

verisine bağılı olmadığını; bunun yerine önceden edinilmiş temsil gücü ve çıkarım mekanizmalarının belirleyici hâle geldiğini göstermektedir. Bu durum, klasik veri merkezli öğrenme anlayışından daha esnek ve görev odaklı bir öğrenme paradigmasına geçişı işaret etmektedir.

İncelenen çalışmalar, örneksiz ve az örnekle öğrenme yaklaşımlarının özellikle etiketli verinin sınırlı olduđu doğal dil işleme, bilgisayarlı görü ve biyomedikal uygulamalarda anlamlı performans kazanımları sağladığını ortaya koymaktadır. Bununla birlikte, bu yaklaşımların başarısının büyük ölçüde bağlam tasarımına, ön eğitim kapsamına ve görevle ilişkili anlamsal bilginin doğruluğuna bağılı olduđu görülmektedir. Benzer şekilde, sentezlenmiş veri kullanımı veri erişim maliyetlerini azaltma ve gizliliği koruma açısından önemli avantajlar sunsa da, üretilen verinin temsil gücü, önyargı aktarımı ve güvenilirlik boyutları dikkatle değerlendirilmediği sürece sınırlı kalabilmektedir.

Gelecek çalışmalar açısından, veri olmadan veya sınırlı veriyle öğrenme yaklaşımlarının sistematik değerlendirme çerçevelerinin geliştirilmesi önemli bir araştırma alanı olarak öne çıkmaktadır. Özellikle bağlama dayalı öğrenmenin sınırlarının daha iyi anlaşılması, sentetik verinin kalite ve güvenilirliğinin nicel ölçütlerle değerlendirilmesi ve bu yöntemlerin sorumlu yapay zekâ ilkeleriyle uyumlu biçimde bütünleştirilmesi gerekmektedir. Bu doğrultuda, veri kıtlığına dayanıklı öğrenme paradigmalarının yalnızca performans artırıcı teknikler olarak değil, aynı zamanda güvenilir, şeffaf ve etik yapay zekâ sistemlerinin temel bileşenleri olarak ele alınması büyük önem taşımaktadır.

Kaynakça

Al Rahhal, M. M., Bazi, Y., Elgibreen, H., & Zuair, M. (2023). Vision-language models for zero-shot classification of remote sensing images. *Applied Sciences*, 13(22), 12462.

Alemohammad, S., Humayun, A. I., Agarwal, S., Collomosse, J., & Baraniuk, R. (2024). Self-improving diffusion models with synthetic data. *arXiv preprint arXiv:2408.16333*.

Bendou, Y., Lioi, G., Padeloup, B., Mauch, L., Cardinaux, F., Boukli Hacene, G., & Gripon, V. (2024). LLM meets vision-language models for zero-shot one-class classification. *arXiv*. <https://arxiv.org/abs/2404.00675>

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.

Bommasani, R. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Boraschi, D., van der Schaar, M., Costa, A., & Milne, R. (2025). Governing synthetic data in medical research: the time is now. *The Lancet Digital Health*, 7(4), e233-e234.

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, 5(6), 493–497.

Das, A., & Singh, S. (2025). Few-shot and zero-shot learning for MRI brain tumor classification using CLIP and vision transformers. *Sensors*, 25(23), 7341.

Figueira, A., & Vaz, B. (2022). Survey on synthetic data generation, evaluation methods and GANs. *Mathematics*, 10(15), Article 2733.

Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479, 448-455.

Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE intelligent systems*, 24(2), 8-12.

He, X., Hu, Y., Zhou, Z., Jarraya, M., & Liu, F. (2025). Few-shot adaptation of training-free foundation model for 3D medical image segmentation. *arXiv*.

Jun Zhu, Synthetic data generation by diffusion models, *National Science Review*, Volume 11, Issue 8, August 2024, nwae276, <https://doi.org/10.1093/nsr/nwae276>

Kumar, S. V., Kumar, K. S., Kotteeswaran, C., & Manoj, S. (2025, March). Synthetic Data Generation Using GANS for Improved Machine Learning Models. In *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2024*, Volume 1 (Vol. 886, p. 325). Springer Nature.

Liu, Y., Acharya, U. R., & Tan, J. H. (2025). Preserving privacy in healthcare: A systematic review of deep learning approaches for synthetic data generation. *Computer Methods and Programs in Biomedicine*, 260, 108571.

Ramesh, G., Sahil, M., Palan, S. A., Bhandary, D., Ashok, T. A., Shreyas, J., & Sowjanya, N. (2025). A review on NLP zero-shot and few-shot learning: methods and applications. *Discover Applied Sciences*, 7(9), 1-22.

Saha, O., Van Horn, G., & Maji, S. (2024). Improved zero-shot classification by adapting vision-language models with text descriptions. arXiv.

Sivarajkumar, S., Kelley, M., Samolyk-Mazzanti, A., Visweswaran, S., & Wang, Y. (2024). An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing: Algorithm development and validation study. *JMIR Medical Informatics*, 12, e55318.

Ye, Q., Lin, B. Y., & Ren, X. (2021). CrossFit: A few-shot learning challenge for cross-task generalization in NLP. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), 7163–7189. Association for Computational Linguistics

BÖLÜM 14

DİJİTAL İKİZ TEKNOLOJİSİ: KAVRAM, MİMARİ VE DİSİPLİNLERARASI UYGULAMALAR

RÜYA AKINCI¹

Giriş

Son yıllarda siber-fiziksel sistemlerdeki gelişmeler, fiziksel sistemlerin ya da nesnelerin yalnızca izlemesinin yanı sıra bu yapıları analiz ederek öngöründe bulunabilen ve aynı zamanda sistemi optimize edebilen bir yapı oluşturmaktadır. Bu yapının en güncel ve etkili sistemlerinden biri dijital ikiz teknolojisidir.

Dijital ikiz fiziksel bir nesnenin, varlığın veya sistemin veri etkileşimiyle beslenen sanal ve dinamik kopyasıdır. Dijital ikiz, fiziksel model ya da nesne ile gerçek zamanlı olarak etkileşim halindedir. Bu sayede gerçek sistem ile çift yönlü olarak çalışmaktadır (Fuller, Fan, Day, & Barlow, 2020).

Dijital ikiz kavramı, ilk zamanlarda üretim ve endüstriyel bakım süreçlerinde kullanılmıştır. Ancak yapay zeka teknolojisinin ilerlemesiyle birlikte kendine yeni kullanım alanları bulmuştur.

¹ Öğr. Gör., Harran Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Orcid: 0000-0001-6272-1178

Günümüzde dijital ikizler; akıllı üretim sistemleri, otonom robotlar, enerji yönetim sistemleri, sağlık ve biyomedikal uygulamalarda kullanılmaktadır. Bu bağlamda dijital ikiz kavramı sadece simülasyon modeli değil aynı zamanda veri, model ve yapay zeka katmanlarının bütünleştiği dinamik bir sistem olarak çalışmaktadır (Singh, ve diğerleri, 2022).

Özellikle sağlık ve rehabilitasyon alanında dijital ikiz teknolojisi, bireye özgü fiziksel yapıların modellenmesinde önemli rol oynamaktadır. Rehabilitasyon alanında özellikle kas ve iskelet sisteminin modellenmesi açısından oldukça önemlidir. Bu alanda dijital ikiz, eklem hareketleri ve kas aktivasyonları gibi parametrelerin dijital ortamda modellenenebilmesi ve bu sayede kişiye özel ve dinamik bir yapı oluşturmaktadır. Elektromiyografi (EMG) gibi biyosinyallerin dijital ikiz modellerine entegre edilmesiyle, insan-makine arayüzleri ile yapı daha doğal ve adaptif bir şekilde kontrol edilmektedir.

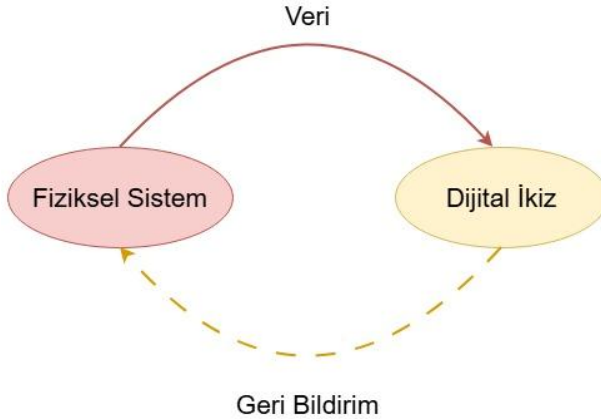
Bütün bunlara ek olarak, Sanal Gerçeklik (VR) ve Arttırılmış Gerçeklik (AR) teknolojileriyle desteklenen dijital ikizler, sadece görselleştirme sunmayıp aynı zamanda etkileşimli olarak geribildirim sunmaktadır. Bu sayede terapilerin ya da uygulamaların daha motive edici olmaktadır. Bu tür sistemler, rehabilitasyon robotları, dış iskeletlerle kullanılarak terapiyi daha eğlenceli hale getirerek hasta motivasyonunu sağlamaktadır (Rashid & Ghiasi, 2025).

Bu çalışmada dijital ikiz teknolojisinin kavram, mimari ve disiplinlerarası uygulamaları hakkında bilgi verilerek çalışma prensipleri sistematik bir şekilde incelenmiştir. Ayrıca robotik ve sağlık odaklı yaklaşımlarda dijital ikizin sunduğu olanaklar tartışılmakta ve gelecekteki çalışmalara yönelik değerlendirmeler sunulmaktadır.

1. Dijital İkiz Kavramı

Dijital ikiz kavramı, fiziksel bir nesnenin, sistemin ya da varlığın fiziksel, işlevsel ve yapısal özelliklerini dijital ortamda temsil eden ve gerçek zamanlı olarak veriler ile güncellenerek fiziksel sistemlerle etkileşim kurabilen sanal modeldir (VanDerHorn & Mahadevan, 2021). Bu yapı sadece mevcut fiziksel durumu yansıtmakla kalmaz aynı zamanda simülasyon, optimizasyon, öngörü ve karar destek sistemi görevlerini yerine getirir. Bütün bu yönleriyle dijital ikiz, klasik simülasyon yaklaşımlarından ayrılarak dinamik ve gerçek zamanlı olarak çalışabilen bir sistem olmaktadır. Literatürde fiziksel nesneyi doğrudan temsil ettiği için dijital kopya olarak da tanımlanmaktadır (Duman, 2021).

Şekil 1 Dijital İkiz



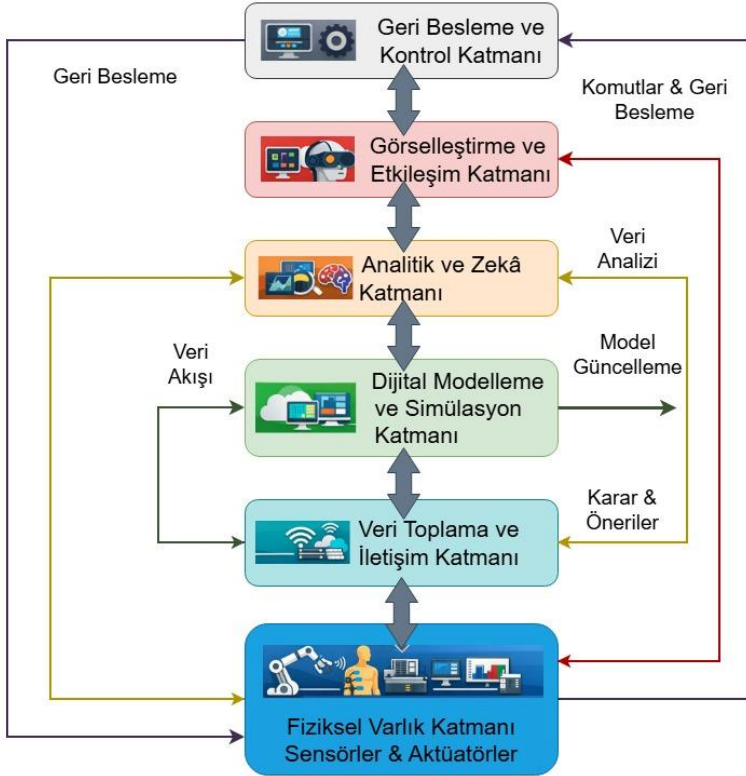
2. Dijital İkiz Mimarisi ve Sistem Bileşenleri

Dijital ikiz mimarisi, fiziksel sistem ile onu temsil eden sanal model arasındaki veri akışını ve karar verme süreçlerini katmanlı bir yapı içerisinde sunarak gerçek zamanlı olarak kontrol sağlar (Dihan, ve diğerleri, 2024).

2.1. Dijital İkiz Katmanları

Dijital ikiz katmanları, fiziksel sistem ile sanal sistem arasında kesintisiz, güvenilir ve gerçek zamanlı olarak etkileşim kurmasını sağlar (Dihan, ve diğerleri, 2024). Şekil 2 de dijital ikiz katmanları yer almaktadır.

Şekil 2 Dijital İkiz Katmanlar



2.2.1. Fiziksel Varlık Katmanı

Fiziksel varlık katmanı, dijital ikizin temsil ettiği gerçek sistemden oluşur. Bu sistem, robot, insan, üretim hattı, güç santrali vs. olabilir. Bu katman, dijital ikizin referans olarak kullandığı katmandır. Bu katmandan sensör, aktüatör ve kontrol elemanlarından

verileri alarak konumunu ve hareketini oluşturur (Qian, ve diğerleri, 2024).

2.2.2. Veri Toplama ve İletişim Katmanı

Veri toplama ve iletişim katmanı, fiziksel katmanından aldığı verileri doğru ve gerçek zamanlı olarak alır. Alınan veriler yine bu katmanda ön işleme yapılarak gecikme olmadan sanal ortama aktarılmaktadır (Rahman, ve diğerleri, 2024). Bu katmanda, alan veri yolları, endüstriyel Ethernet, kablosuz ağlar, 5G gibi endüstriyel iletişim sistemlerinin gerçek zamanlı olarak iletilmesini sağlar (Mistry, Tanwar, Tyagi, & Kumar, 2020).

2.2.3. Dijital Modelleme ve Simülasyon Katmanı

Dijital modelleme ve simülasyon katmanında, fiziksel modelin ya da varlığın sanal kopyasının oluşturulduğu bölümdür. Burada katman, fiziksel dünyadan aldığı veriler ile gerçek zamanlı olarak çalışabilmektedir. Katman özellikle karar destek, optimizasyon ve maliyet düşürme gibi aşamalar için kritik rol oynamaktadır (Segovia & Garcia-Alfaro, 2022).

2.2.4. Analitik ve Zekâ Katmanı

Analitik ve zeka katmanı en önemli katmanlardan biridir. Bu katman gerçek zamanlı sensör verilerinden elde edilen bilgileri kullanarak analiz, tahmin ve optimizasyon yapan üst katmandır. Bu katmanda çoğunlukla yapay zeka, makine öğrenmesi ve derin öğrenme teknikleri kullanılır (Bozkaya-Aras, Onel, Eriskin, & Karatas, 2025). Bu katmanın dijital ikizin beyni gibi çalışarak sensörlerden alınan verileri bilgiye, bilgiyi de tahmin ve optimizasyona dönüştürmektedir (Liu, Xi, Quan, & Cheng, 2024).

2.2.5. Görselleştirme ve Etkileşim Katmanı

Görselleştirme ve etkileşim katmanı, sanal ortamda dijital ikizin ürettiği verilerin kullanıcıya anlamlı bir şekilde sunmaktadır.

Aynı zamanda bu katmanda kullanıcı ile etkileşim kurarak insan ve dijital ikiz arasında köprü görevini üstlenmektedir (Zhang, Xu, Liu, & Chai, 2024). Bu katman, dijital ikizin görünen yüzüdür. Burada tasarlanan yapı doğru bir şekilde tasarlandığında sistemin bilişsel yükü azaltılarak karar kalitesi artar ve kullanıcı daha verimli bir şekilde sisteme dahil edilebilir (Fan, ve diğerleri, 2021).

2.2.6. Geri Besleme ve Kontrol Katmanı

Geri besleme ve kontrol katmanı, dijital ikizin sadece görsel olarak değil aynı zamanda fiziksel sistemi aktif bir şekilde analiz eden ve yönlendiren bir yapıya dönüştürmektedir. Bu katman sayesinde dijital ikiz, analiz ve zeka katmanından aldığı çıktıları fiziksel dünyaya aktararak kapalı çevrim bir kontrol yapısı sunmaktadır. Bu sayede fiziksel ortamdan yeni veri geldiğinde sistem kendini güncelleyerek adaptif bir kontrol sağlamış olacaktır. Dolayısıyla sistem bu katman sayesinde fiziksel sistemi sürekli kontrol eden, yeni gelen verileri öğrenen ve bu veriler ışığında sistemi tekrar güncelleyebilen bir yapı haline gelmektedir (Sjöberg, ve diğerleri, 2023).

2.2.7. Dijital İkiz Mimari Yapısının Önemi

Dijital ikiz sistemlerinin mimari tasarımı ölçeklenebilir, güvenilirlik ve sürdürülebilirlik açısından oldukça önemlidir. Mimaride, katmanlar arasında görev dağılımının net olması sistemin geliştirme sürecini kolaylaştırmakla beraber farklı uygulamalarda kullanılabilirliğini arttırmaktadır. Bu nedenle, dijital ikiz bütüncül bir sistemi temsil etmektedir (Zhuang, ve diğerleri, 2017).

3. Dijital İkiz: Disiplinlerarası Uygulamalar

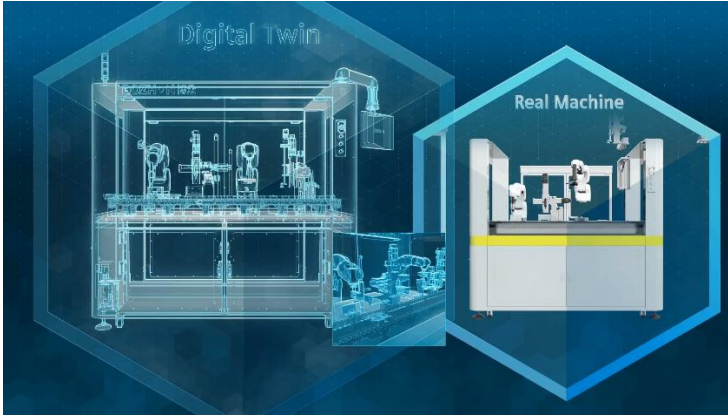
Dijital ikiz teknolojisi, disiplinler arası bir yaklaşıma sahiptir. Sensör teknolojileri, veri bilimi, kontrol teorisi, yapay zeka, insan-bilgisayar etkileşimi ve farklı alanlardaki mühendislik bilgilerinin bir araya getirilerek farklı disiplinlerde ortak bir dijital ortam

sunmaktadır. Bu yönüyle dijital ikiz teknolojisi, tek bir alana özgü değildir. Çok sayıda disiplini ortak bir çatı altında toplayarak bütüncül bir sistem sunmaktadır (Emmert-Streib, Tripathi, & Dehmer, 2023).

3.1. Endüstri ve Üretim Sistemleri

Dijital ikiz teknolojisi, özellikle endüstri ve üretim hatlarının yer aldığı fabrikalarda çok kullanılmaktadır. Dijital ikiz, fabrika sistemleri için sürekli güncellenerek gerçek zamanlı olarak makinaların ya da üretim hatlarının kontrolünü gerçekleştirmektedir. Özellikle, kestirimci bakım (Iheanacho, Ozurumba, Amajoh, & Igwe, 2025), verimlilik artışı, enerji optimizasyonu ve akıllı fabrika uygulamalarında tercih edilmektedir (Karic, Schilberg, & Arockiadoss, 2025).

Şekil 3 Dijital İkiz ve Gerçek Makine



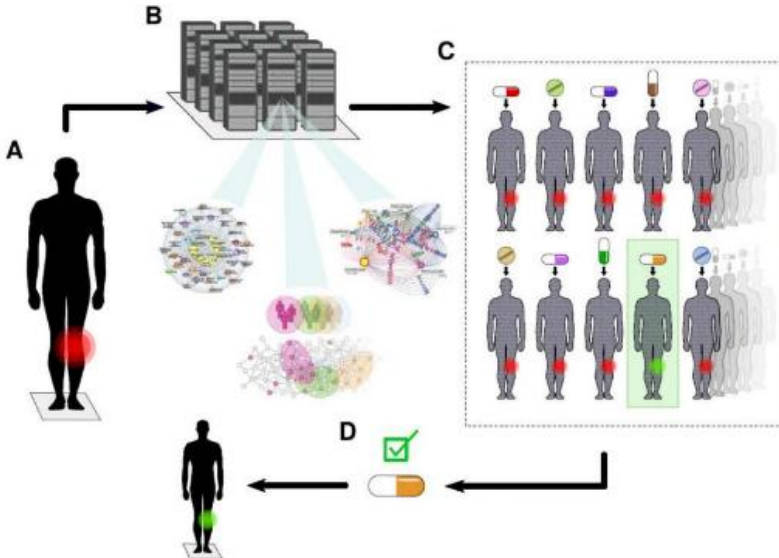
(Selvaraj, 2020)

3.2. Sağlık ve Biyomedikal Uygulamalar

Dijital ikiz teknolojisi, sağlık ve biyomedikal alanda kişiye özel, veri temelli adaptif çözümler sunar. Dijital ikiz, tanı, tedavi ve rehabilitasyon süreçlerinde önemli rol oynamaktadır. Burada fiziksel sistemlerin (insan vücudu, organ veya biyomekanik yapıların vs.)

sanal ortamda modeli oluşturularak fiziksel sistem gerçek zamanlı olarak temsil edilmektedir. Dijital ikiz bu aşamada yapıyı sayısal verilere dayandırarak tedavi süreçlerinin sürdürülmesi ve optimize edilmesine olanak sağlamaktadır. Sağlık ve biyomedikal alanda çoğunlukla kullanılan dijital ikizler; hastaya özel dijital ikiz, kas ve iskelet sisteminin modellenmesi, rehabilitasyon ve robot destekli terapilerdir. Şekil 4 de kişiselleştirilmiş tıp için dijital ikiz kavramı yer almaktadır. Burada, a: bireysel bir hastanın lokal bir hastalık belirtisi vardır (kırmızı). b: Bu hastanın dijital ikizi, binlerce hastalıkla ilgili değişkenin hesaplamalı ağ modellerine dayalı olarak sınırsız sayıda kopyada oluşturulur. c: Her ikiz, binlerce ilaçtan biri veya daha fazlasıyla bilgisayar ortamında tedavi edilir. Bu, bir hastanın dijital olarak iyileşmesiyle sonuçlanır (yeşil). d: Dijital ikiz üzerinde en iyi etkiyi gösteren ilaç, hastanın tedavisi için seçilir (Björnsson, ve diğerleri, 2019).

Şekil 4 Kişiselleştirilmiş Tıp İçin Dijital İkiz Kavramı



(Björnsson, ve diğerleri, 2019)

3.3. Robotik ve Otonom Sistemler

Dijital ikiz teknolojisi, robotik ve otonom sistemlerinde özellikle tasarım, geliştirme ve test süreçlerinde oldukça önemli rol oynamaktadır. Fiziksel robot ya da otonom sistemin sanal ortamda kopyası yani dijital ikizi oluşturularak sistemin performans analizi etkin ve gerçek zamanlı bir şekilde gözlemlenebilmektedir. Bu yapı sayesinde sistem öğrenen, öngören ve karar verebilen tümleşik bir kontrol ve analiz sistemini oluşturmaktadır. Robotik sistemlerde, dijital ikiz robotun mekanik yapısını, sensörlerini, aktüatörlerini ve kontrol algoritmalarını sanal ortamda kopyasının oluşturulmuş halidir. Bu model yine gerçek zamanlı olarak kontrol edilebilmekte ve fiziksel robot ile senkronize olarak çalışabilmektedir (Ramasubramanian, Mathew, Kelly, Hargaden, & Papakostas, 2022). Otonom sistemlerde ise otonom sistemin çevresini algılama ve karar verme süreçlerinde oldukça önemli bir parametreyi oluşturmaktadır. Bu sistemde, otonom sistemler, gerçek dünyada karşılaşılabileceği karmaşık problemleri güvenli bir şekilde öğrenebilecektir (Almeaibed, Al-Rubaye, Tsourdos, & Avdelidis, 2021). Şekil 5 de fabrikada yer alan bir fiziksel robot kol ve robot kolun dijital ikizi yer almaktadır.

Şekil 5 Fiziksel Robot Kol (solda) ve Dijital ikizi (sağda).



(Liang, Mcgee, Menassa, & Kamat, 2020)

3.4. Sanal Gerçeklik (VR), Artırılmış Gerçeklik (AR) ve Eğitim

Dijital ikiz, sanal gerçeklik ve artırılmış gerçeklik ile çalıştığında özellikle eğitim ve öğrenme süreçlerinde oldukça faydalı

bir yapı sunmaktadır. Dijital ikiz, sanal ya da artırılmış ortamlarda kullanıcıların karmaşık süreçleri ya da adımları deneyimleyerek öğrenmesini sağlamaktadır. Bu yapı eğitimde deneyim temelli yaklaşımı güçlendirmektedir. Burada VR, kullanımı tamamen sanal ortamda olduğu yapıyı temsil eder. AR ise dijital ikizin nesneler üzerine bindirilmesi ile sistem üzerinde öğrenme olanağı sunmaktadır (Al-Ansi, Jaboob, Garad, & Al-Ansi, 2023).

3.5. Akıllı Şehirler ve Altyapı Sistemleri

Dijital ikiz teknolojisi, akıllı şehirler ve altyapı sistemlerinde özellikle karmaşıklığı gidererek çok paydaşlı kontrol yapılarını tek bir çatı altında toplayarak yönetmektedir. Şehir ölçeğinde oluşturulmuş olan dijital ikizler, fiziksel altyapı, insan hareketliliği ve çevresel verilerin bir araya getirilerek gerçek zamanlı bir sistem sunmaktadır. Akıllı şehir dijital ikizleri; ulaşım, enerji su ve iletişim altyapılarının modellenmesi, ulaşım ve trafik gibi adımlarda aktif rol oynamaktadır (Marai, Taleb, & Song, 2020) . Şekil 6 da yol altyapısının dijital ikizi oluşturulmuştur.

Şekil 6 Yol Altyapısının Dijital İkizi

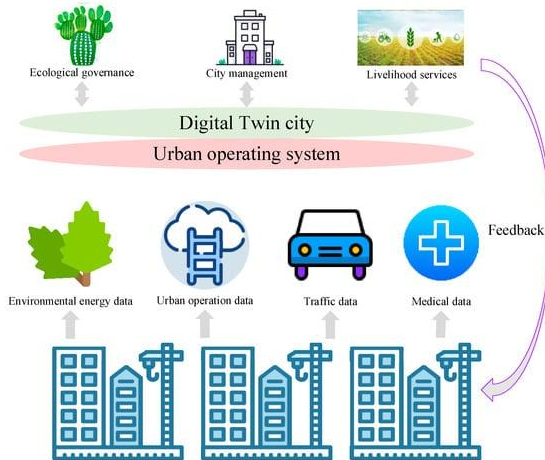


(Marai, Taleb, & Song, 2020)

3.6. Enerji, Çevre ve Sürdürülebilirlik

Dijital ikiz teknolojisi, enerji sistemleri, çevresel izleme ve sürdürülebilirlik alanlarında oldukça önemli hale gelmiştir. Dijital ikizin, fiziksel sistemden gerçek zamanlı veri entegrasyonu sayesinde, öngörüye dayalı analiz, ve optimizasyon yetenekleri sayesinde sistemin kaynakları verimli kullanması, çevresel atıkların azaltılması gibi sürdürülebilir bir sistemin tasarlanması açısından oldukça verimli bir yapı oluşturmaktadır. Enerji sistemlerinde dijital ikizler, üretimden tüketime kadar bütün bu sistemlerin izlenmesi ve optimize edilmesini sağlar. Özellikle elektrik üretimlerinde sanal güç santrali ikizi oluşturularak üretim ve tüketim dengesi analizi yapılmakta ve optimize edilmektedir. Çevresel faktörlerin izlenmesinde ise çevresel faktörler sanal ortamla bağlantılı olarak izlenerek çevresel risklerin tespiti yapılmaktadır (Yu, Patros, Young, Klinac, & Walmsley, 2022). Şekil 7 de şehrin çeşitli alanlarındaki bilgi kaynakları çeşitli çevresel faktörler yer almaktadır. Bu görselde akıllı şehir modeli ile dijital ikiz çevre sürdürülebilirlik alanı ile entegre bir şekilde çalışmaktadır. Fiziksel şehir model ve gerçek veriler sanal ortamda dijital ikiz teknolojisi ile birlikte çalışmaktadır (Yang, Lv, & Wang, 2022).

Şekil 7 Dijital İkiz Şehri



(Yang, Lv, & Wang, 2022)

Sonuç

Bu çalışma, dijital ikiz teknolojisinin kavramsal çerçevesi, sistem mimarisi ve disiplinlerarası uygulama alanlarından bahsedilmiştir. Dijital ikizin, fiziksel sistem, veri altyapısı, dijital modelleme, analitik zeka, görselleştirme ve geri besleme katmanlarının etkileşimiyle çalışan dinamik ve gerçek zamanlı olarak çalışan bütüneşik bir yapı olduğu ortaya konmuştur. Bu yapı, dijital ikizin klasik simülasyon ve analiz yaklaşımlarının aksine gerçek zamanlı olarak çalışabilen, öğrenen, karar veren adaptif bir yapı olduğu görülmektedir.

Dijital ikiz yapısı gereği zaten disiplinlerarası uygulamalarda kullanılmaktadır. Robotik ve otonom sistemlerde, güvenli test ortamları, sistemin adaptif kontrolü gibi sistemin güvenli bir şekilde kontrolünü sağlamaktadır. Aynı zamanda dijital ikiz, robotik ve otonom sistemlerinin geliştirme süreçlerini hızlandırarak sistemin güvenilirliğini arttırmaktadır. Sağlık ve biyomedikal uygulamalarında ise hastaya özel dijital ikiz yaklaşımları geliştirilmektedir. Özellikle kişiselleştirilmiş tedavi ve rehabilitasyon çalışmalarının önü açılarak sağlık ekosisteminde önemli bir devrim yaratmaktadır. Sanal ve artırılmış gerçeklik ortamlarında kullanılarak da eğitim ve öğretim faaliyetlerinde yine kişiye uyarlanabilir eğitim modeli oluşturabilmektedir. Akıllı şehirler, enerji ve çevre uygulamalarında ise dijital ikizler, kaynak verimliliği, karmaşık altyapı problemlerinin çözümü açısından önemli rol oynamaktadır.

Elde edilen değerlendirmeler, dijital ikiz teknolojisinin başarı ve gelişiminin yalnızca teknik doğruluğa değil aynı zamanda disiplinler arası entegrasyon, insan merkezli yaklaşım ve veri kalitesine bağlı olduğu görülmektedir. Gelecekte yapay zeka ve veri analitiğindeki ilerlemelerle birlikte dijital ikizlerin daha otonom kendi kendine öğrenebilen ve geliştirebilen sistemlere dönüşmesi öngörülmektedir. Bu bağlamda dijital ikiz teknolojisi mühendislik,

sağlık, eğitim ve şehircilik gibi pek çok alanda dijital dönüşümün temel yapı taşlarından biri olmaya adaydır.

Kaynakça

- Al-Ansi, A. M., Jaboob, M., Garad, A., & Al-Ansi, A. (2023). Analyzing augmented reality (AR) and virtual reality (VR) recent development in education. *Social Sciences & Humanities Open*, 100532.
- Almeaibed, S., Al-Rubaye, S., Tsourdos, A., & Avdelidis, N. P. (2021). Digital Twin Analysis to Promote Safety and Security in Autonomous Vehicles. *IEEE Communications Standards Magazine*, 40-46.
- Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D. R., Gustafsson, M., . . . Benson, M. (2019). Digital twins to personalize medicine. *Genome medicine*, 12(1), 4.
- Bozkaya-Aras, E., Onel, T., Eriskin, L., & Karatas, M. (2025). Intelligent human activity recognition for healthcare digital twin. *Internet of Things*, 30, 101497.
- Dihan, M. S., Akash, A. I., Tasneem, Z., Das, P., Das, S. K., Islam, M. R., . . . Hasan, M. M. (2024). Digital twin: Data exploration, architecture, implementation and future. *Heliyon*, 10(5).
- Duman, M. Ç. (2021). İŞLETMELER İÇİN YENİ BİR VERİMLİLİK TEKNOLOJİSİ: DİJİTAL İKİZ. *Verimlilik Dergisi*.
- Emmert-Streib, F., Tripathi, S., & Dehmer, M. (2023). Analyzing the Scholarly Literature of Digital Twin Research: Trends, Topics and Structure. *IEEE Access*, 69649-69666.
- Fan, Y., Yang, J., Chen, J., Hu, P., Wang, X., Xu, J., & Zhou, B. (2021). A digital-twin visualized architecture for Flexible

- Manufacturing System. *Journal of Manufacturing Systems*, 176-201.
- Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*, 108952 - 108971.
- Iheanacho, C. C., Ozurumba, E., Amajoh, N., & Igwe, E. (2025). Enhancing predictive maintenance in lean manufacturing for continuous process improvement using digital twin technology. *World Journal of Advanced Research and Reviews*.
- Karic, H., Schilberg, D., & Arockiadoss, A. S. (2025). Smart Factory and Industry 4.0: A Survey on Advancements, Technologies, Methods and Perspectives of Digital Transformation in Manufacturing. *AHFE International*.
- Liang, C.-J., Mcgee, W., Menassa, C., & Kamat, V. R. (2020). Bi-directional communication bridge for state synchronization between digital twin simulations and physical construction robots. *International Symposium on Automation and Robotics in Construction (ISARC 2020)* (s. 1480-1487). Proceedings of the International Symposium on Automation and Robotics in Construction (IAARC).
- Liu, Z., Xi, L., Quan, Y., & Cheng, C. (2024). Brain-Inspired Collaborative Power Dynamic Dispatch for the Digital Twin Smart Generation System. *IEEE Transactions on Industry Applications*, 1420-1430.
- Marai, O. E., Taleb, T., & Song, J. (2020). Roads Infrastructure Digital Twin: A Step Toward Smarter Cities Realization. *IEEE network*, 136-143.

- Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G-enabled IoT for industrial automation: A systematic review, solutions, and challenges. *Mechanical Systems and Signal Processing*.
- Qian, C., Guo, Y., Hussaini, A., Musa, A., Sai, A., & Yu, W. (2024). A New Layer Structure of Cyber-Physical Systems under the Era of Digital Twin. *ACM Transactions on Internet Technology*.
- Rahman, H., Shah, U. M., Riaz, S. M., Kifayat, K., Moqurrah, S. A., & Yoo, J. (2024). Digital twin framework for smart greenhouse management using next-gen mobile networks and machine learning. *Future Generation Computer Systems*, 156, 285-300.
- Ramasubramanian, A. K., Mathew, R., Kelly, M., Hargaden, V., & Papakostas, N. (2022). Digital Twin for Human–Robot Collaboration in Manufacturing: Review and Outlook. *Applied Sciences*, 4811.
- Rashid, S. M., & Ghiasi, A. R. (2025). Adaptive digital twin integration with multilevel inverter control for energy efficient smart rehabilitation systems. *Scientific Reports*, 15(1), 8511.
- Segovia, M., & Garcia-Alfaro, J. (2022). Design, Modeling and Implementation of Digital Twins. *Sensors*, 22.
- Selvaraj, R. R. (2020, 6 29). *Embrace the digital twin in manufacturing: Deliver performance, reliability and efficiency*. Siemens:
<https://blogs.sw.siemens.com/simcenter/embrace-the-digital-twin-in-manufacturing> adresinden alındı

- Singh, M., Srivastava, R., Fuenmayor, E., Qiao, V. K., Murray, N., & Devine, D. M. (2022). Applications of digital twin across industries: A review. *Applied Sciences*, 5727.
- Sjöberg, A., Önnheim, M., Frost, O., Cronrath, C., Gustavsson, E., Lennartson, B., . . . Jirstrand, M. (2023). Online geometry assurance in individualized production by feedback control and model calibration of digital twins. *Journal of Manufacturing Systems*, 71-81.
- VanDerHorn, E., & Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, 145, 113524.
- Yang, B., Lv, Z., & Wang, F. (2022). Digital Twins for Intelligent Green Buildings. *Buildings*, 856.
- Yu, W., Patros, P., Young, B., Klinac, E., & Walmsley, T. G. (2022). Energy digital twin technology for industrial energy management: Classification, challenges and future. *Renewable and Sustainable Energy Reviews*, 112407.
- Zhang, K., Xu, Q., Liu, C., & Chai, T. (2024). Intelligent decision-making system for mineral processing production indices based on digital twin interactive visualization. *Journal of Visualization*, 421-436.
- Zhuang, C., Liu, J., Xiong, H., Ding, X., Liu, S., & Weng, G. (2017). Connotation, architecture and trends of product digital twin. *Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems*, 753-768.

BÖLÜM 15

HASTA YORUMLARINI KULLANARAK DUYGU ANALİZİ

İpek AKPINAR¹
Ali Mete ÇIPLAK²
Ediz ENGİN³
Enes ÇAKMAK⁴
Mahmutcan SAKINCI⁵
Mustafa TETİK⁶
Nail KOCABAY⁷
Ozan DESTE⁸
Ayşe Berna ALTINEL⁹

¹ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0009-4474-1189, ipekakpinar@marun.edu.tr

² Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0007-9324-3157, ali.mete@marun.edu.tr

³ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0006-2075-0222, edizengin@marun.edu.tr

⁴ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0001-2998-5320, enescakmak@marun.edu.tr

⁵ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0004-5289-3258, mahmutcansakinci@marun.edu.tr

⁶ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0003-8638-9730, mustafatetik@marun.edu.tr

⁷ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0001-1677-3206, nailkocabay@marun.edu.tr

⁸ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, İstanbul/Türkiye,
Orcid: 0009-0002-1473-346X, ozandeste@marun.edu.tr

⁹ Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul/Türkiye,
Orchid: 0000-0001-5544-0925, berna.altinel@marmara.edu.tr

Giriş

Dijitalleşmenin hızla arttığı günümüzde, sağlık hizmetleri yalnızca fiziksel tedavi süreçlerinden ibaret olmaktan çıkmış ve hasta memnuniyeti, hizmet kalitesi ve hasta deneyimi gibi unsurlar sağlık sisteminin değerlendirilmesinde kritik ölçütler haline gelmiştir. Özellikle sağlık kuruluşları ve çalışanları hakkında yapılan çevrimiçi yorumlar bu sistemin güçlü ve zayıf yönlerini ortaya koymak açısından büyük bir veri kaynağı sunmaktadır. Bu bağlamda hasta yorumlarının analizi yalnızca bireysel deneyimlerin değerlendirilmesiyle kalmamakta, aynı zamanda sağlık sistemlerinin iyileştirilmesi adına önemli yapısal verilere ulaşılmasını da sağlamaktadır.

Bu çalışmada çevrimiçi platformlarda yer alan hasta yorumları incelenerek sağlık kuruluşlarının ve çalışanlarının hizmet kalitelerinin veri madenciliği teknikleri ile değerlendirilmesi amaçlanmaktadır. Özellikle tekrar eden şikayet, öneri ve olumlu geri bildirimlerin belirlenmesiyle birlikte hizmet sunucularının gelişime açık yönleri daha görünür hale getirilmektedir. Projenin temel çıktısı, her yorumu pozitif, nötr veya negatif olarak sınıflandıran bir sistem oluşturarak sağlık kurumlarının stratejik planlamalarında kullanılabilir somut sonuçlar üretmektir. Bu sistem, hasta ile sağlık kurumu arasındaki dolaylı iletişimi analiz edilebilir hale getirerek hizmet kalitesine yönelik doğrudan ve anlamlı bir geri bildirim mekanizması sunmaktadır.

Bu amaç doğrultusunda, veri madenciliği ve doğal dil işleme (NLP) tekniklerinden yararlanılmıştır. Araştırmanın ilk sürecinde K-En Yakın Komşu (KNN) algoritması, Destek Vektör Makineleri (SVM) ve çift yönlü uzun kısa süreli bellek (bi-LSTM) modelleri kullanılmış ve çeşitli sonuçlar elde edilmiştir. Son aşamasında ise bu yöntemlerin yanı sıra Evrişimsel Sinir Ağı (CNN), BERT ve türevi derin öğrenme modelleri ile boosting algoritmaları (XGBoost, CatBoost gibi) entegre edilerek çok yönlü bir değerlendirme gerçekleştirilmiştir. Yöntem seçimleri, hem literatürdeki benzer çalışmalardan yararlanılarak hem de proje kapsamında test edilen model performanslarına göre belirlenmiştir.

Literatür taramaları, hasta deneyimlerine odaklanan duygu analizlerinde özellikle son yıllarda Transformer tabanlı modellerin öne çıktığını göstermektedir. BERT, XLNet ve RoBERTa gibi modellerin yorum sınıflandırmasında üstün başarı sağladığı ve geleneksel yöntemlerin ise belirli veri türlerinde hâlâ geçerliliğini koruduğu görülmüştür. Bu kapsamda proje boyunca incelenen makaleler, kullanılan veri setleri, yöntem tercihleri ve deney sonuçları bir bütün olarak değerlendirilmiş; sağlık alanında veri odaklı çözüm önerileri geliştirilmiştir.

Çalışmanın kalanı ise aşağıdaki şekilde düzenlenmiştir. 2. bölümde ilgili çalışmalar ve literatür özetlerine yer verilmiştir. 3. bölümde projede kullanılan veri setleri ve yöntemler detaylandırılmıştır. 4. bölümde deneysel çalışmaların sonuçları karşılaştırmalı biçimde sunulmuştur. Son bölümde ise genel değerlendirme yapılmış ve gelecekte yapılabilecek çalışmalara dair öneriler paylaşılmıştır.

İlgili Çalışmalar

Hasta Geri Bildirimleri Üzerine Otomatik Metin Analizi

Alexander ve arkadaşları(Alexander, Bahja ve Butt, 2022), NHS hasta geri bildirim verilerini analiz ederek sağlık hizmet kalitesi üzerine içgörüler sunan bir sistem geliştirmiştir. Çalışmada öncelikle metinler ön işleme tabi tutulmuş, ardından Latent Dirichlet Allocation (LDA) ile konu modelleme yapılmıştır. Ayrıca metinler Naive Bayes sınıflandırıcısı ile duygu kategorilerine ayrılmıştır.

Elde edilen sonuçlarda, olumsuz yorumların büyük çoğunluğunun randevu sistemine ve personel ilgisine yönelik olduğu gözlemlenmiştir. Sağlık sisteminde sıkça tekrar eden sorunları tespit edebilmek için makine öğrenmesi tabanlı metin analizi yöntemlerinin etkili olduğu belirtilmiştir.

Sağlık Platformlarında Duygu ve Değerlendirme Analizi

Bu çalışmada(Tang ve diğerleri, 2023), Healthgrades adlı sağlık değerlendirme platformundaki metin yorumları incelenmiştir.

Yorumlara VADER duygu analiz algoritması uygulanmış ve yıldızlı değerlendirme puanları ile karşılaştırılmıştır. Ayrıca kelime frekans analizi yapılarak kullanıcıların öne çıkardığı temalar belirlenmiştir.

Çalışmanın dikkat çekici sonucu, metinlerdeki duygu skorları ile yıldız puanlamalarının yüksek korelasyon göstermesidir. Yorumların içeriği, yıldız puanlarından bağımsız olarak hizmet kalitesini daha ayrıntılı yansıtabilmektedir.

BERT Tabanlı Çin Sağlık Platformu Analizi

Zhao ve ekibi (Zhao, Li, Yuan ve Deng, 2024), Çin merkezli haodf.com platformundaki hasta yorumlarını sınıflandırmak için BERT tabanlı bir sistem geliştirmiştir. BERT modeline ek olarak CNN ve Bi-GRU yapıları da kullanılarak performans karşılaştırması yapılmıştır.

Deneyler sonucunda, BERT modelinin %98,36 doğruluk oranı ile en iyi sonuç verdiği görülmüştür. Ayrıca F2 skoru da dikkate alınarak olumlu, olumsuz ve nötr sınıflandırmasında BERT'in diğer modellere üstün geldiği kanıtlanmıştır.

Duygu Sınıflandırmasında Derin Öğrenme Modelleri

Suresh ve arkadaşları(Sagarika Suresh Thimmanayakanapalya, Pavankumar Mulgund, ve Raj Sharman, 2022), hasta yorumlarını duygu kategorilerine ayırmak amacıyla Bi-LSTM, LSTM ve BERT modellerini karşılaştırmıştır. Harvard n2c2 veriseti ve RateMDs platformu kullanılmıştır.

Deneyler sonucunda BERT modeli %99,7 gibi oldukça yüksek bir doğruluk oranına ulaşarak en başarılı model olmuştur. Bu sonuç, transformer tabanlı modellerin geleneksel RNN tabanlı modellere göre anlamlı bir üstünlük sağladığını göstermektedir.

Portekizce Hasta Yorumlarında BERTopic Kullanımı

Bu çalışmada(Osório ve Fachada, 2024), Google Review üzerinden toplanan Portekizce sağlık yorumları analiz edilmiştir. LDA, LSA ve BERTopic yöntemleri karşılaştırılarak en uyumlu konu modelleme algoritması araştırılmıştır.

Sonuçlara göre BERTopic modeli, hem konu tutarlılığı hem de duygu-konu uyumu açısından en yüksek performansı göstermiştir. Bu modelin özellikle küçük ve çok çeşitli metinlerde avantaj sağladığı belirtilmiştir.

Macar Hasta Deneyimlerinin Negatifliğinin Ölçülmesi

Macaristan'dan derlenen hasta yorumları, huBERT, HIL-SBERT ve HDBSCAN yöntemleri ile analiz edilmiştir. Bu modeller kullanılarak, yorumların tematik kümeleme ve duygu sınıflandırması yapılmıştır.

Çalışma(Osváth, Yang ve Kósa, 2023), özellikle devlet hastanelerine yönelik yorumların %94,4 oranında negatif duygu içerdiğini ortaya koymuştur. Bu durum, sağlık politikaları açısından dikkate alınması gereken çarpıcı bir geribildirim sunmaktadır.

Hasta Deneyimi Yorumlarında BERTopic Kullanımı

Bu çalışmada(Steele ve diğerleri, 2025), hasta deneyim anketlerindeki yorumlar analiz edilmiştir. Anketler, Kanada'da hastanelerde hizmet alan yetişkin ve çocuk hastalardan toplanmış ve BERTopic ile konu modellemesi yapılmıştır. LDA gibi klasik yöntemler yerine, önceden eğitilmiş dil modeli tabanlı BERTopic tercih edilmiştir.

Sonuçlara göre BERTopic, az ayarla bile yorumlardan anlamlı başlıklar çıkarabilmiş ve mevcut anketlerde yer almayan yeni bakım temalarını tanımlamıştır. Yorumlar genel olarak olumlu bulunmuştur. Bu modelin, kişi merkezli bakım, hasta güvenliği ve kalite iyileştirme açısından katkı sunduğu belirtilmiştir.

GPT-4 Turbo ile Hasta Yorumlarının Çok Etiketli Sınıflandırması

Sakai ve ekibi(Sakai, Lam, Mikaeili, Bosire ve Jovin, 2025), HCAHPS anketinden elde edilen 1.089 hasta yorumunu çok etiketli olarak sınıflandırmak amacıyla GPT-4 Turbo modelini kullanmıştır. Sıfır örnekli (zero-shot) öğrenme yöntemiyle %76.12 örnek bazlı F1

skoru ve %73.61 ağırlıklı F1 skoru elde edilmiştir. Bu başarı oranları, geleneksel makine öğrenimi yöntemleri ve diğer dil modellerine kıyasla anlamlı düzeyde yüksektir. Ayrıca, yorumlardan kişisel sağlık bilgilerini ayıklamak üzere bir PHI tespit çerçevesi geliştirilmiştir.

Duygu analizi ile ilgili literatürdeki bazı çalışmalarla ilgili yöntem, veri seti, kullanılan metrik ve alınan sonuçlara ait bilgiler Tablo 1’de gösterilmektedir.

Tablo 1: Duygu analizi ile ilgili literatürdeki çalışmalar

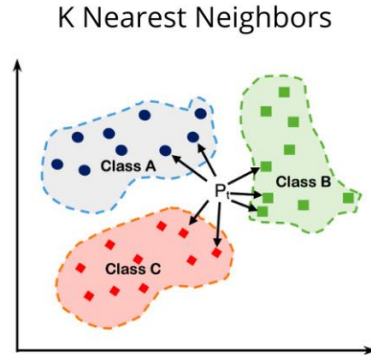
Yazar	Yıl	Yöntem	Veri Seti	Metrikler	Sonuç
Alexander et al.	2022	LDA, Naive Bayes	NHS Free Text Dataset	F1	Negatif yorum oranı en yüksek; randevu/personel şikayetleri önde
Tang et al.	2023	VADER, Linear Regression	Healthgrades	VADER score, frekans	Yorum duygu skoru yıldızla ilişkili, frekans analizi temaları belirledi
Zhao et al.	2024	BERT, CNN, Bi-GRU	haodf.com	Accuracy, F2	BERT %98.36 doğruluk ile üstün performans sağladı
Suresh et al.	2022	BERT, Bi-LSTM, LSTM	Harvard n2c2, RateMD	Accuracy	BERT %99.7 doğruluk ile en iyi sonucu verdi
Osorio & Fachada	2024	BERTopic, LDA, LSA	Google Reviews (PT)	Accuracy, uyum skoru	BERTopic konuyu bakımından en güçlü sonuçları verdi
Osváth et al.	2023	huBERT, HIL-SBERT, HDBSCAN	praxis.blog.hu, yorumlar	Sentiment Accuracy	%94.4 oranında hasta deneyimi negatif bulundu

Steele et al.	2025	BERTopic, Duygu Analizi	Kanada hasta deneyim anketleri (2016–2020)	Sentiment Score, Comment Rate, Comment Lenght	BERTopic, yeni ve anlamlı hasta deneyimi temaları belirleyerek klasik anketleri tamamladı
Sakai et al.	2024	GPT-4 Turbo (0-shot)	HCAHPS	F1 (76.12%), Weighted F1 (73.61%)	GPT-4 Turbo en yüksek sınıflandırma başarımını gösterdi

Yöntemler

Kullanılan Makine ve Derin Öğrenme Modelleri

KNN: KNN, sınıflandırma yaparken basit mantık ile hareket eder. Yeni bir veri geldiğinde, daha önce sınıfı bilinen verilere bakar ve ona en çok benzeyen K tanesini bulur. Bu benzer verilerin çoğunlukta olan sınıfı neyse, yeni veriye de o sınıfı atar Geeks For Geeks (2025) K-Nearest Neighbor(KNN) Algorithm,) (30.06.2025 tarihinde <https://www.geeksforgeeks.org/machine-learning/k-nearest-neighbours/> adresinden ulaşılmıştır).



Şekil 1: KNN temsili(Sachinsoni, 2023)K Nearest Neighbours — Introduction to Machine Learning Algorithms (01.07.2025 tarihinde <https://medium.com/@sachinsoni600517/k-nearest->

neighbours-introduction-to-machine-learning-algorithms-9dbc9d9fb3b2 adresinden ulaşılmıştır).

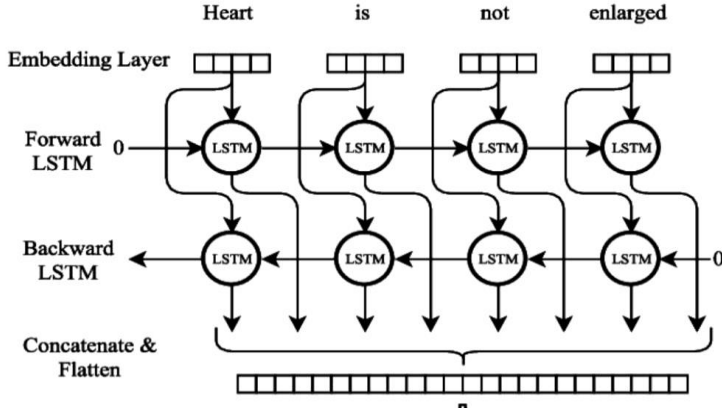
Veriler arasındaki benzerlik, genellikle iki nokta arasındaki uzaklık ölçülerek bulunur. Bu ölçüm için Öklid uzaklığı, cosine benzerliği veya Hamming uzaklığı gibi yöntemler tercih edilir.

Öklid (Euclidean) Uzaklığı, iki nokta arasındaki düz çizgi mesafesini hesaplar. Değerler ne kadar yakınsa, veriler o kadar benzer kabul edilir.

Cosine Benzerliği, vektörler arasındaki açısal benzerliği ölçer. Burada amaç, iki vektörün aynı yöne ne kadar baktığını anlamaktır. Aynı değerlere sahip olmasalar bile, benzer yönelimleri varsa yüksek benzerlik puanı elde edilir.

Hamming Uzaklığı, iki vektör arasında kaç elemanın farklı olduğunu sayar. Eğer veriler 0 ve 1 gibi ikili biçimlerdeyse, bu yöntem veriler arasındaki farkları basit ama etkili bir şekilde hesaplamak için uygundur

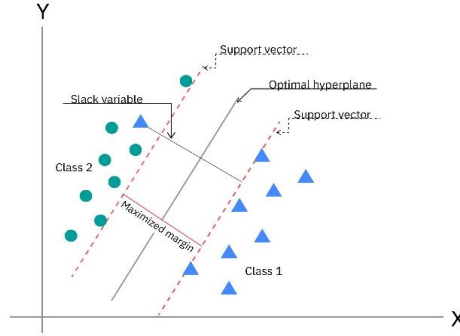
Bi-LSTM: Metin gibi verileri işlerken hem geçmiş hem de gelecek kelimeleri aynı anda göz önüne alan bir derin öğrenme modelidir. Normal LSTM modelleri sadece metni soldan sağa okurken, Bi-LSTM iki yönlü olarak çalışır. Bu sayede bir kelimenin tam anlamını çevresindeki tüm bağlamı kullanılarak daha iyi anlar(*Anishnama, 2023) Understanding Bidirectional LSTM for Sequential Data Processing* (30.06.2025 tarihinde <https://medium.com/@anishnama20/understanding-bidirectional-lstm-for-sequential-data-processing-b83d6283befc> adresinden ulaşılmıştır).



Şekil 2: Bi-LSTM şeması(Preserving Privacy in Arabic Judgments: AI-Powered Anonymization For Enhanced Legal Data Privacy”, t.y.)

Bi-LSTM özellikle duygu analizi gibi cümle yapısının önemli olduğu görevlerde başarılı sonuçlar verir. Bu çift yönlü yapı, modelin bağlamı daha derinlemesine kavramasını sağlar. Uzun ve karmaşık cümlelerde bile anlamı yakalayabilir.

SVM : Destek Vektör Makineleri (SVM), denetimli öğrenme kapsamında sınıflandırma, regresyon ve anormal veri noktalarını saptama gibi görevlerde kullanılan bir metottur. Bu yöntem, bir veri kümesindeki iki veya daha fazla nokta grubunu birbirinden ayırmak için uzayda hiperdüzlemler (doğrular) oluşturur. SVM'in temel prensibi, oluşturulan bu hiperdüzlemin, ayrılacak iki kümedeki en yakın noktalara olan mesafesini (marjını) maksimuma çıkarmaktır (Türkoğlu, 2021) Support Vector Machine (SVM) Algoritması-Makine Öğrenmesi- (30.06.2025 tarihinde <https://mfatih.to.medium.com/support-vector-machine-algoritması-makine-öğrenmesi-8020176898d8> adresinden ulaşılmıştır).



Şekil 3: SVM şeması (“What Is Support Vector Machine?”, 2023)

Bu yaklaşım, modelin yeni verileri daha doğru sınıflandırmasını sağlar ve öğrenilen sınırın genellenebilirliğini artırır.

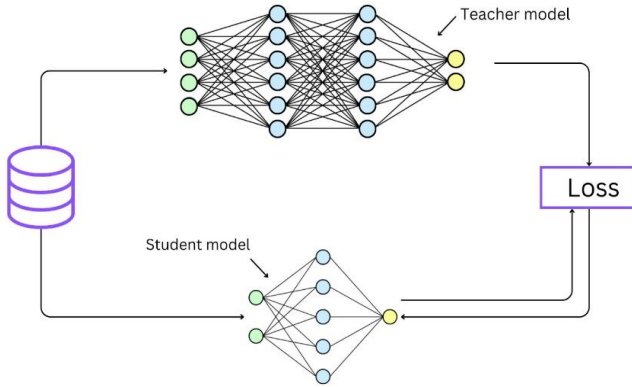
CNN : Convolutional Neural Network (CNN), özellikle görüntü tanıma ve doğal dil işleme görevlerinde yaygın olarak kullanılan bir mimaridir. Özelliklerin çıkarımı için konvolüsyon filtreleri kullanır, bu sayede işlenmemiş veriden anlamlı sonuçlar üretebilir. Konvolüsyon Formülü(Awaits, 2020) Denklem (1)’de verilmiştir.

$$G[i, j] = \sum_{u=-k}^k \sum_{v=-k}^k H[u, v] F[i - u, j - v] \quad (1)$$

Konvolüsyon İşlemi Formülü, CNN’in temel yapı taşı olan konvolüsyon işlemi tanımlar. X, giriş görüntüsü veya özelliğini, K ise filtreyi temsil eder. K filtresi X üzerinde kaydırılarak çarpım ve toplam işlemleri uygulanır. Bu sayede filtre, giriş verisindeki belirli bir özelliği tespit etmeye çalışır ve filtreler farklı özellikleri öğrenerek anlamlı temsiller oluşturur.

RoBERTa : RoBERTa, Facebook AI arařtırmacıları tarafından geliştirilen, BERT modelinin gelişmiş bir versiyonudur. Doğal dil işleme alanında kullanılan güçlü bir derin öğrenme modelidir. Transformer mimarisi üzerine kurulu olan bu model, metinleri anlamak ve bağlama göre yorumlamak için kelimeler arasındaki ilişkileri öğrenir.

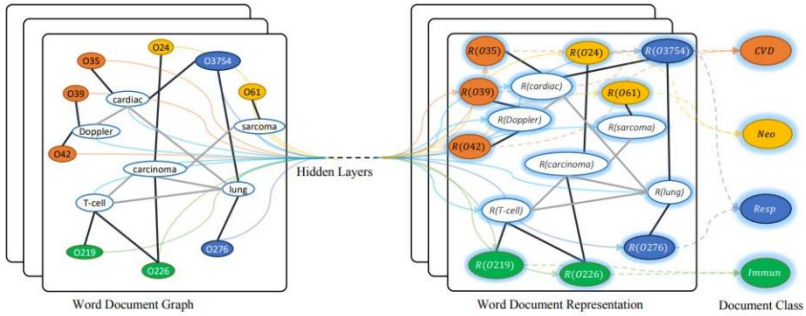
DistilBert : DistilBERT, BERT'in daha küçük, daha hızlı ve daha verimli bir versiyonudur. Hugging Face tarafından geliştirilen bu model, BERT'in doğruluk seviyesini büyük ölçüde korurken, %40 daha az parametre ve %60 daha hızlı çalışma gibi önemli performans özellikleriyle öne çıkar.



Şekil 4: Öğretmen-öğrenci modelinde bilgi aktarımı(“Distilbert: A Smaller, Faster, and Distilled BERT - Zilliz Learn”, t.y.)

textRNN : Text RNN, sıralı olan verilerle çalışmak için tasarlanmış bir yapay sinir ağı türüdür. Bir metindeki kelimeleri sırayla işler ve her kelimenin önceki kelimelerle olan bağlamını öğrenmeye çalışır. Bu sayede metin içinde anlam ilişkilerini kurabilir. Geçmiş bilgileri belleğinde tutar ve bu belleği her adımda güncelleyerek sonraki kelime işlenirken kullanır ancak uzun cümlelerde eski bilgileri unutabilme gibi sorunları vardır.

textGCN : Metin sınıflandırma için geliştirilmiş graf tabanlı bir derin öğrenme yöntemidir. Metinleri sırayla değil, graf yapısı olarak ele alır; belgeler ve kelimeler düğüm, aralarındaki ilişkiler ise kenarlarla temsil edilir. Kelime eş-oluşumları ve belge-kelime bağlantıları bu graf üzerinden modellenir.

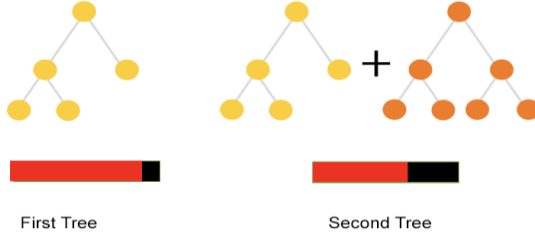


Şekil 5: Metin GCN'nin Şematiği(Yao, Mao ve Luo, 2018)

Model, Graph Convolutional Network (GCN) işlemleriyle her düğüm için bağlamsal temsiller üretir. Böylece sıralı yapıya ihtiyaç duymadan metin içi ilişkiler öğrenilir. Özellikle etiketli verinin az olduğu senaryolarda etkili sonuçlar verir.

XGBoost : Gradient Boosting olarak bilinen algoritmanın geliştirilmiş halidir. Karar ağaçlarının birbirlerinin hatalarını giderecek şekilde ard arda dizilmesi prensibine dayanır. Öncelikle bir karar ağacı eldeki veri ile eğitilir. Daha sonra bir sonraki karar ağacı ilk karar ağacının hataları ile eğitilir ve onu düzeltmeye çalışır. Üçüncü karar ağacı ikinci karar ağacının hatalarını düzeltmeye çalışır. Bu şekilde karar ağaçları eklenerek süreç devam eder. En sonunda bütün ağaçların sonuçları toplanır.

CatBoost : CatBoost, karar ağaçlarını ardı ardına kullanarak hataların minimize edildiği ve tahmin oranının iyileştirilmesini sağlayan bir gradient boosting tekniğidir.



Şekil 6: Birinci ve ikinci ağaçlar genelde yüksek hata içerir(“CatBoost—Open-source gradient boosting library”, t.y.)

Bu teknik çalıştırılırken önce karar ağacı oluşturulur sonrasında tahmin içerisinde ne kadarlık hata oranı var hesaplanır. Her bir oluşturulan karar ağacının amacı bir önceki karar ağacı tarafından meydana gelen hataların düzeltilip, stabil hâle getirilmesini sağlamaktır ve bu bir döngü halinde çalışmaya devam eder.

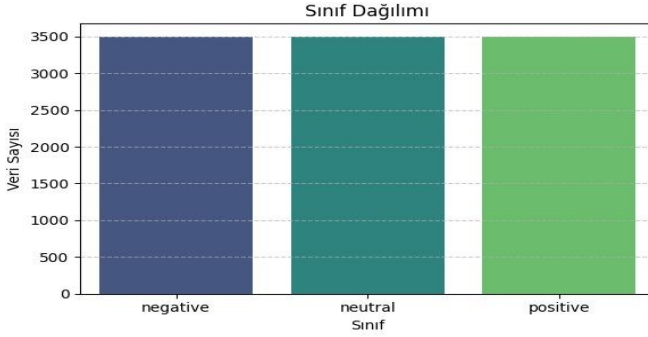
Deney Ortamı, Sonuçlar ve Değerlendirme

Veri Seti

NHS Verisi

National Health Service (NHS)(“NHS website developer portal”, t.y.), Birleşik Krallık’ta 1948 yılında kurulmuş olan kamuya ait sağlık hizmetleri sistemidir. NHS resmî sitesinde hastaların yorumları ve verdikleri puanların (Rating) verilerini çekebilmemizi sağlayan bir api hizmeti mevcuttur [kaynak1]. Rating etiketinin ortalamaları alınarak 0 ile 5 arasında değerler elde edilmiştir. Bu değerler 0 – 2.5 arası negatif, 2.5 – 3.5 arası nötr ve 3.5 – 5 arası pozitif olarak etiketlenmiş ve 3 sınıf elde edilmiştir. Bu şekilde veri seti daha sade ve analiz edilebilir bir hale getirilmiş, sınıflandırma modelleri için uygun bir yapı oluşturulmuştur. Toplam 10500 veri

bulunun bu veri setinde her sınıfın veri sayısı eşitlenmiştir. Her sınıfta 3500 veri bulunmaktadır.

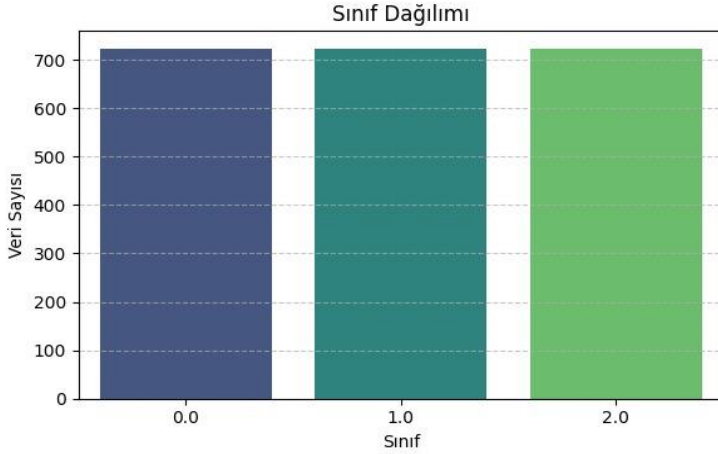


Şekil 7: NHS veri kümesi sınıf dağılımı

RateMD Verisi

Avijit Thawani'nin "Are Online Reviews of Physicians Biased Against Female Providers?" adlı makalesinde kullandığı verisetidir (Thawani, Paul, Sarkar ve Wallace, 2019). RateMDs.com sitesinden açık kaynaklı bir kod kullanılarak elde edilmiştir. Veri setinde doktorlar, hastalar tarafından dört farklı başlık altında 1 ile 5 arasında puanlanarak değerlendirilmiştir. Her başlık için ayrı ayrı puanlama yapılmış olup bu çoklu değerlendirmeler sadeleştirilerek tek bir sınıfa (label) dönüştürülmüştür. Dört başlıktan elde edilen puanlar bir araya getirilerek genel bir sınıflandırma yapılmış, bu sınıflandırma sürecinde 0 negatif, 1 nötr ve 2 pozitif olmak üzere üç farklı sınıf tanımlanmıştır. Etiketleme aşamasında, Yüksek puanlardan oluşan yorumlar pozitif sınıfa dahil edilmiştir. Aynı şekilde, düşük puanlardan oluşan yorumlar negatif, düşük ve yüksek puanlardan oluşan çoklu değerlendirmeler ise nötr olarak etiketlenmiştir. Bu şekilde veri seti daha sade ve analiz edilebilir bir hale getirilmiş, sınıflandırma modelleri için uygun bir yapı oluşturulmuştur. Bu veride de bütün sınıfların veri sayısı eşitlenmiştir. Nötr verilerin sayısı görece az olduğu için diğer

sınıflardaki veri sayıları azaltılmıştır. Toplam 2169 veri vardır ve her sınıf 723 veriye sahiptir.



Şekil 8: RateMD veri kümesi sınıf dağılımı

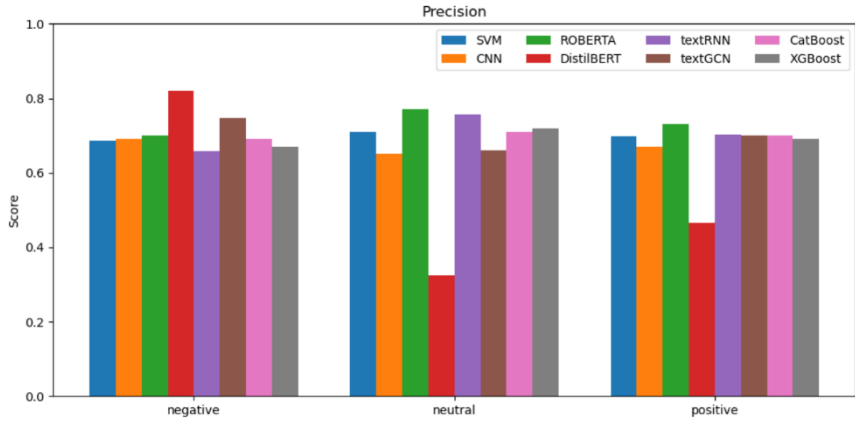
Sonuçlar ve Değerlendirme

NHS Veri Kümesi

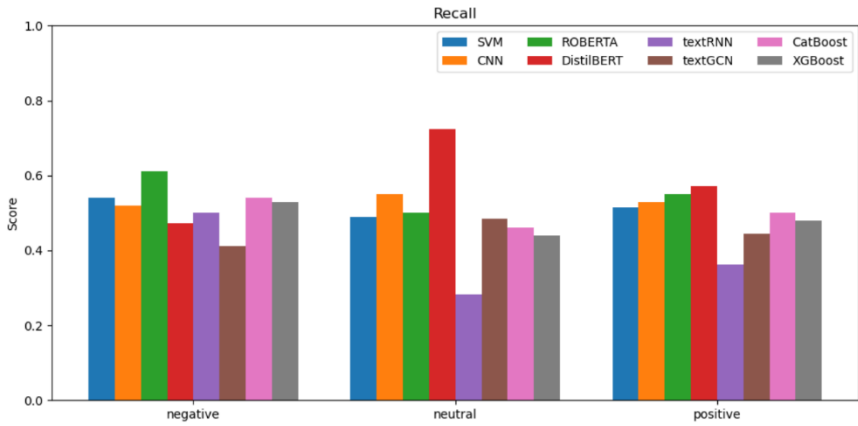
Tablo 2: NHS veri kümesi

Algoritma Adı	%70 Eğitim %30 Test	%80 Eğitim %20 Test
SVM	Accuracy: 0.6804 F1 score: 0. 6762 Recall: 0.6804 Precision: 0.6736	Accuracy: 0.6838 F1 score: 0.6785 Recall: 0.6838 Precision: 0.6755
CNN	Accuracy: 0.66 F1 score: 0.66 Recall: 0.66 Precision: 0.67	Accuracy: 0.65 F1 score: 0.65 Recall: 0.65 Precision: 0.65

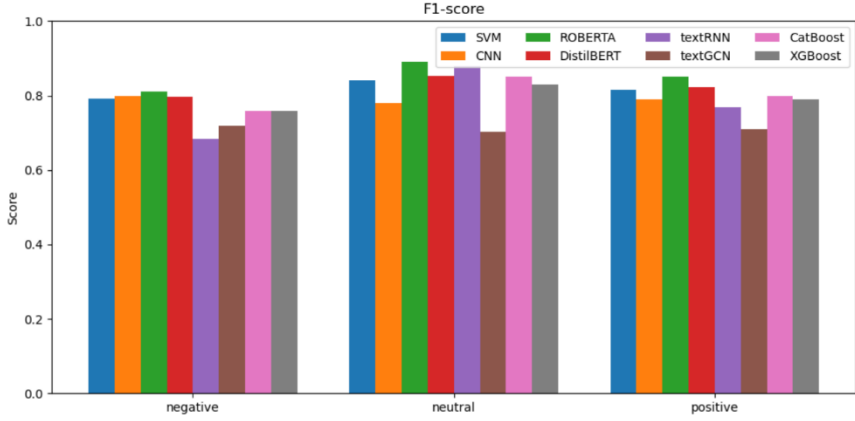
RoBERTa	Accuracy: 0.72 F1 score: 0.71 Recall: 0.72 Precision: 0.71	Accuracy: 0.73 F1 score: 0.72 Recall: 0.73 Precision: 0.72
DistilBERT	Accuracy: 0.6341 F1 score: 0.6203 Recall: 0.6340 Precision: 0.6963	Accuracy: 0.6890 F1 score: 0.6843 Recall: 0.6890 Precision: 0.6816
textRNN	Accuracy: 0.6390 F1 score: 0.6111 Recall: 0.6390 Precision: 0.6138	Accuracy: 0.6467 F1 score: 0.6420 Recall: 0.6467 Precision: 0.6388
textGCN	Accuracy: 0.6252 F1 score: 0.6186 Recall: 0.6252 Precision: 0.6159	Accuracy: 0.6452 F1 score: 0.6378 Recall: 0.6455 Precision: 0.6380
XGBoost	Accuracy: 0.66 F1 score: 0.66 Recall: 0.66 Precision: 0.65	Accuracy: 0.67 F1 score: 0.66 Recall: 0.67 Precision: 0.66
CatBoost	Accuracy: 0.6737 F1 score: 0.6669 Recall: 0.6737 Precision: 0.6639	Accuracy: 0.6771 F1 score: 0.6706 Recall: 0.6771 Precision: 0.6678



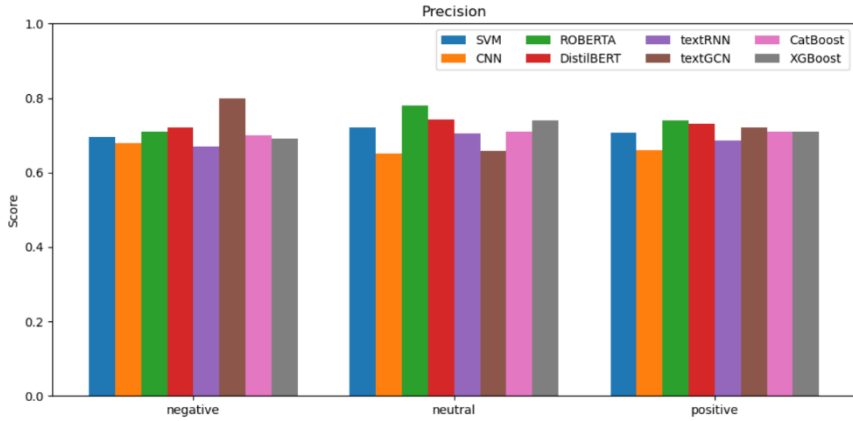
Şekil 9: NHS veri kümesi için yüzde 70 precision sonucu



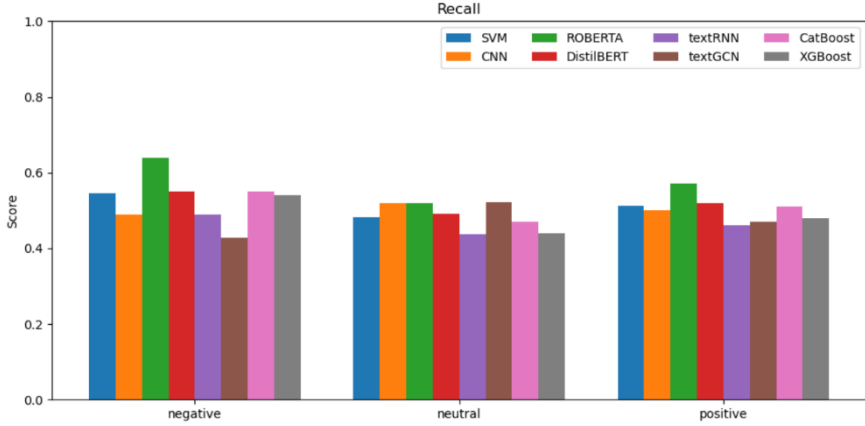
Şekil 10: NHS veri kümesi için yüzde 70 recall sonucu



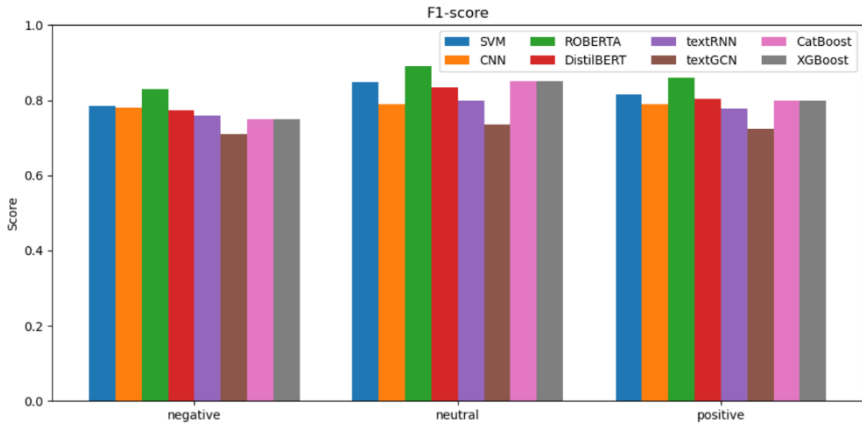
Şekil 11: NHS veri kümesi için yüzde 70 F1-score sonucu



Şekil 12: NHS veri kümesi için yüzde 80 precision sonucu



Şekil 13: NHS veri kümesi için yüzde 80 recall sonucu



Şekil 14: NHS veri kümesi için yüzde 80 F1-score sonucu

NHS veri kümesi göz önünde bulundurulduğunda en başarılı sonuçları veren algoritmanın hem %70 eğitim – %30 test hem de %80 eğitim – %20 test oranlarında RoBERTa olduğu görülmektedir. Eğitim yüzdesi arttıkça performansı da artan RoBERTa modeli özellikle yüksek veriyle daha güçlü genelleme yeteneği sergilemektedir. Benzer şekilde DistilBERT de düşük eğitim

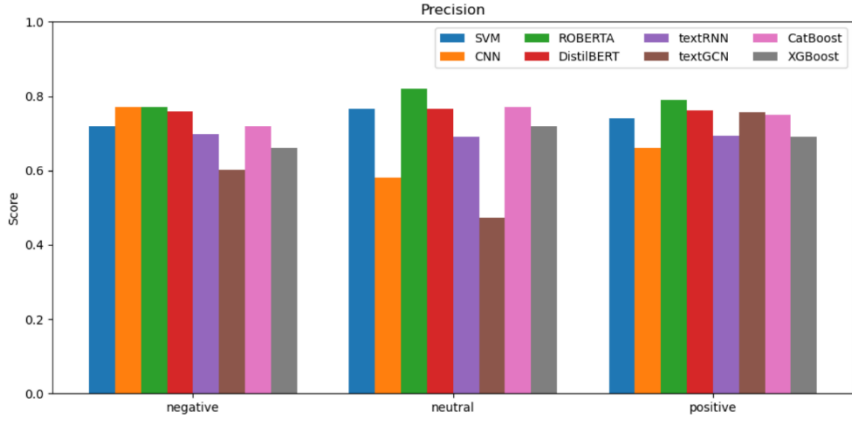
oranlarında nispeten zayıf sonuçlar vermesine rağmen eğitim oranı arttığında önemli bir performans artışı göstererek daha rekabetçi hale gelmektedir. CatBoost algoritması da benzer şekilde yüksek eğitim oranlarında daha iyi sonuç vermekte ve istikrarlı bir başarı profili çizmektedir. SVM modeli ise düşük ve yüksek eğitim oranlarında oldukça benzer sonuçlar vermekte, dolayısıyla eğitim oranından çok fazla etkilenmeyen dengeli bir algoritma olarak öne çıkmaktadır. Ancak bu modelin başarı düzeyi, RoBERTa gibi daha gelişmiş modellerin gerisinde kalmaktadır. Öte yandan CNN modeli dikkat çekici bir şekilde eğitim oranı arttıkça daha kötü sonuçlar vermektedir; bu da modelin yüksek veriyle overfitting ya da yeterli öğrenememe gibi sorunlar yaşadığını göstermektedir. textRNN ve textGCN algoritmaları genel olarak zayıf sonuçlar verirken, eğitim oranı artsa bile diğer modellerle aralarındaki farkı kapatamamaktadır. XGBoost ise ortalama bir başarı sergilemekte ve eğitim yüzdesi değişiminden çok fazla etkilenmemektedir.

RateMD Veri Kümesi

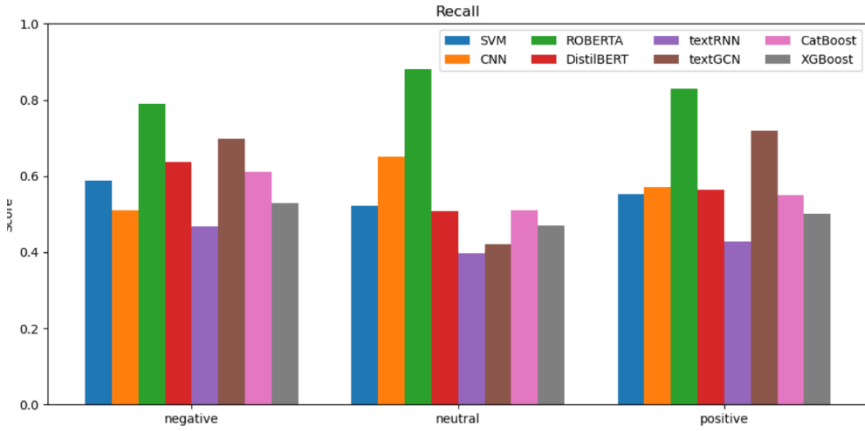
Tablo 2: RateMD veri kümesi

Algoritma Adı	%70 Eğitim %30 Test	%80 Eğitim %20 Test
SVM	Accuracy: 0.6897 F1 score: 0.6859 Recall: 0.6897 Precision: 0.6843	Accuracy: 0.7074 F1 score: 0.7028 Recall: 0.7076 Precision: 0.7017
CNN	Accuracy: 0.65 F1 score: 0.66 Recall: 0.65 Precision: 0.68	Accuracy: 0.63 F1 score: 0.64 Recall: 0.63 Precision: 0.66

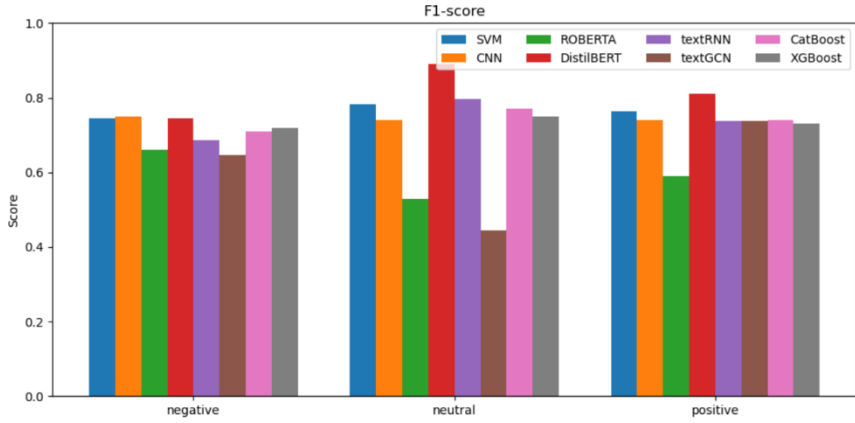
RoBERTa	Accuracy: 0.75 F1 score: 0.74 Recall: 0.75 Precision: 0.74	Accuracy: 0.76 F1 score: 0.75 Recall: 0.76 Precision: 0.75
DistilBERT	Accuracy: 0.7204 F1 score: 0.7122 Recall: 0.7204 Precision: 0.7130	Accuracy: 0.7488 F1 score: 0.7430 Recall: 0.7489 Precision: 0.7421
textRNN	Accuracy: 0.6283 F1 score: 0.6204 Recall: 0.6283 Precision: 0.6172	Accuracy: 0.6221 F1 score: 0.6066 Recall: 0.6221 Precision: 0.6154
textGCN	Accuracy: 0.6125 F1 score: 0.6095 Recall: 0.6125 Precision: 0.6102	Accuracy: 0.6205 F1 score: 0.6179 Recall: 0.6205 Precision: 0.6183
XGBoost	Accuracy: 0.64 F1 score: 0.63 Recall: 0.64 Precision: 0.63	Accuracy: 0.65 F1 score: 0.64 Recall: 0.65 Precision: 0.64
CatBoost	Accuracy: 0.6851 F1 score: 0.6798 Recall: 0.6851 Precision: 0.6793	Accuracy: 0.6959 F1 score: 0.6895 Recall: 0.6961 Precision: 0.6882



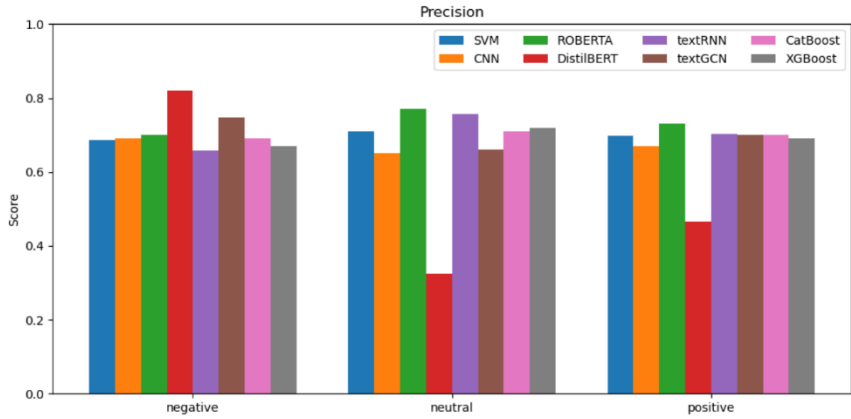
Şekil 15: RateMD veri kümesi için yüzde 70 precision sonucu



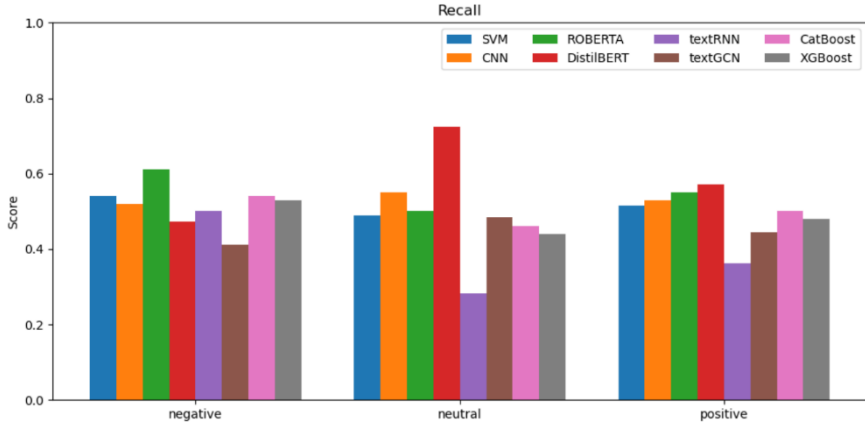
Şekil 16: RateMD veri kümesi için yüzde 70 recall sonucu



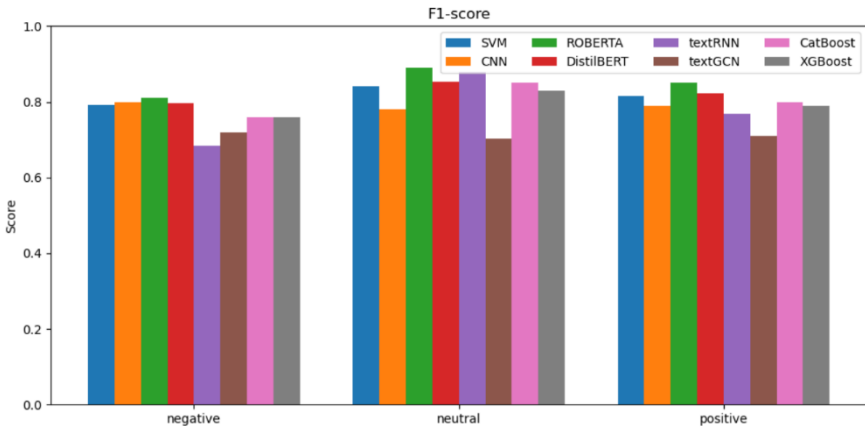
Şekil 17: RateMD veri kümesi için yüzde 70 F1-score sonucu



Şekil 18: RateMD veri kümesi için yüzde 80 precision sonucu



Şekil 19: RateMD veri kümesi için yüzde 80 recall sonucu



Şekil 20: RateMD veri kümesi için yüzde 80 F1-score sonucu

RateMD veri kümesi göz önünde bulundurulduğunda ise NHS veri kümesine benzer bir tablo görülmektedir. En yüksek başarıyı yine RoBERTa sağlamaktadır ve eğitim oranı yükseldikçe performansı artmaya devam etmektedir. DistilBERT modeli, NHS veri kümesinde olduğu gibi RateMD veri kümesinde de eğitim yüzdesi arttıkça belirgin bir başarı yükselişi göstermekte ve RoBERTa'ya yaklaşmaktadır. Bu durum, DistilBERT'in daha fazla

veriye eriştiğinde öğrenme kapasitesinin güçlendiğini ortaya koymaktadır. CatBoost yine güvenilir ve istikrarlı bir model olarak dikkat çekmekte ve eğitim oranı arttıkça küçük ama pozitif değişimler göstermektedir. SVM, NHS veri kümesindeki gibi burada da dengeli bir model olarak yer almakta ve eğitim oranı yükseldiğinde olumlu bir performans artışı göstermektedir. Ancak bu artış sınırlı kalmakta ve üst düzey modellerin gerisinde kalmaktadır. CNN algoritması NHS veri kümesindeki gibi burada da benzer bir şekilde düşük eğitim oranında daha iyi sonuç verirken yüksek eğitim oranında performans kaybı yaşamaktadır. Bu durum, CNN'in bu tür metin verilerine uygun olmadığını ya da aşırı öğrenme sorunları olduğunu göstermektedir. textRNN modeli ise eğitim oranı yükseldikçe performansı düşen nadir modellerden biridir. Bu, modelin eğitim verisini verimli kullanmadığını göstermektedir. textGCN ve XGBoost da genel olarak ortalama veya düşük başarılar sergilemiş, eğitim oranı değişimleri performanslarına belirgin bir etki yapmamıştır.

Sonuç

Bu çalışmada, dijitalleşmenin artmasıyla birlikte sağlık sistemine dair önemli bir geri bildirim kaynağı haline gelen çevrimiçi hasta yorumlarının duygu analizi gerçekleştirilmiştir. Temel problem, hasta deneyimlerinin sağlık hizmeti kalitesine yönelik anlamlı içgörülere dönüştürülmesidir. Bu doğrultuda yorumlar, doğal dil işleme ve makine öğrenmesi algoritmaları aracılığıyla pozitif, nötr ve negatif olarak sınıflandırılmıştır.

Çalışmada iki farklı sağlık değerlendirme platformundan alınan NHS ve RateMD veri kümeleri kullanılmış; KNN, SVM, CNN, textRNN, textGCN, XGBoost ve CatBoost gibi geleneksel makine öğrenmesi algoritmalarının yanı sıra RoBERTa, DistilBERT gibi Transformer tabanlı derin öğrenme modelleri de değerlendirmeye alınmıştır. Elde edilen sonuçlar göstermektedir ki Transformer mimarili modeller, özellikle RoBERTa, her iki veri kümesinde de en yüksek doğruluk, F1, precision ve recall değerlerine ulaşarak diğer algoritmaların önüne geçmiştir. Eğitim

yüzdesi arttıkça RoBERTa ve DistilBERT gibi modellerin performansının belirgin şekilde iyileştiği gözlemlenmiştir. CatBoost ve SVM ise istikrarlı fakat görece olarak daha sınırlı başarı sağlamıştır. CNN gibi bazı modellerin ise eğitim verisi arttıkça performansında düşüş yaşadığı görülmüştür.

Gelecekte yapılacak çalışmalarda veri seti çeşitliliğinin artırılması, Türkçe gibi farklı dillerdeki yorumların da analiz edilmesi, yorumların daha ayrıntılı kategori etiketlerine ayrılması ve GPT benzeri büyük dil modelleriyle sıfır örnekleli sınıflandırma yöntemlerinin entegrasyonu planlanmaktadır. Ayrıca, yorumlardan yalnızca duygu değil; aynı zamanda öneri, şikâyet ve memnuniyet temalarının otomatik çıkarımı üzerine yoğunlaşılması hedeflenmektedir. Bu doğrultuda geliştirilecek sistemler, sağlık politikalarının şekillenmesine ve hasta merkezli hizmet iyileştirmelerine önemli katkılar sunabilir.

Kaynakça / References

Alexander, G., Bahja, M. ve Butt, G. F. (2022). Automating Large-scale Health Care Service Feedback Analysis: Sentiment Analysis and Topic Modeling Study. JMIR Medical Informatics, 10(4), e29385. <https://doi.org/10.2196/29385>

Anishnama. (2023, 18 Mayıs). Understanding Bidirectional LSTM for Sequential Data Processing. Medium. <https://medium.com/@anishnama20/understanding-bidirectional-lstm-for-sequential-data-processing-b83d6283befc>

Awaits. (2020, 17 Mart). [DL] 8. CNN 1(Convolutional Neural Network Basics). Learning. <https://medium.com/jun94-devpblog/dl-8-cnn-1-convolutional-neural-network-basics-2f60c93d7d22>

CatBoost—Open-source gradient boosting library. (t.y.). 2 Temmuz 2025 tarihinde <https://catboost.ai/news/catboost-enables-fast-gradient-boosting-on-decision-trees-using-gpus> adresinden erişildi.

Distilbert: A Smaller, Faster, and Distilled BERT - Zilliz Learn. (t.y.). 2 Temmuz 2025 tarihinde <https://zilliz.com/learn/distilbert-distilled-version-of-bert> adresinden erişildi.

GeeksforGeeks. (t.y.). K-Nearest Neighbor(KNN) Algorithm. 2 Temmuz 2025 tarihinde <https://www.geeksforgeeks.org/machine-learning/k-nearest-neighbours/> adresinden erişildi.

NHS website developer portal. (t.y.). 2 Temmuz 2025 tarihinde <https://www.nhs.uk> adresinden erişildi.

Osório, L. ve Fachada, N. (2024). Patient-Centered Healthcare: 15th Advanced Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2024. L. M. Camarinha-Matos ve F. Ferrada (Ed.), DoCEIS, IFIP Advances in Information and Communication Technology (ss. 152–163) içinde. Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-63851-0_10

Osváth, M., Yang, Z. G. ve Kósa, K. (2023). Analyzing Narratives of Patient Experiences: A BERT Topic Modeling Approach. Acta Polytechnica Hungarica, 20(7), 153–171. <https://doi.org/10.12700/APH.20.7.2023.7.9>

(PDF) Preserving Privacy in Arabic Judgments: AI-Powered Anonymization For Enhanced Legal Data Privacy. (t.y.). 2 Temmuz 2025 tarihinde https://www.researchgate.net/publication/374691308_Preserving_Privacy_in_Arabic_Judgments_AI-Powered_Anonymization_For_Enhanced_Legal_Data_Privacy adresinden erişildi.

Sachinsoni. (2023, 11 Haziran). K Nearest Neighbours—Introduction to Machine Learning Algorithms. Medium. <https://medium.com/@sachinsoni600517/k-nearest-neighbours-introduction-to-machine-learning-algorithms-9dbc9d9fb3b2>

Sagarika Suresh Thimmanayakanapalya, Pavankumar Mulgund, ve Raj Sharman. (2022). Development of an Automated

Physician Review Classification System: A Hybrid Machine Learning Approach. In International Conference of Information Systems. ResearchGate.
https://www.researchgate.net/publication/365960526_Development_of_an_Automated_Physician_Review_Classification_System_A_Hybrid_Machine_Learning_Approach

Sakai, H., Lam, S. S., Mikaeili, M., Bosire, J. ve Jovin, F. (2025, 20 Şubat). Large Language Models for Patient Comments Multi-Label Classification. arXiv.
<https://doi.org/10.48550/arXiv.2410.23528>

Steele, B., Fairie, P., Kemp, K., D'Souza, A. G., Wilms, M. ve Santana, M. J. (2025). Identifying Patient-Reported Care Experiences in Free-Text Survey Comments: Topic Modeling Study. JMIR Medical Informatics, 13(1), e63466.
<https://doi.org/10.2196/63466>

Tang, J. E., Arvind, V., Dominy, C., White, C. A., Cho, S. K. ve Kim, J. S. (2023). How Are Patients Reviewing Spine Surgeons Online? A Sentiment Analysis of Physician Review Website Written Comments. Global Spine Journal, 13(8), 2107–2114.
<https://doi.org/10.1177/21925682211069933>

Thawani, A., Paul, M. J., Sarkar, U. ve Wallace, B. C. (2019). Are Online Reviews of Physicians Biased Against Female Providers? Proceedings of the 4th Machine Learning for Healthcare Conference (ss. 406–423). PMLR.
<https://proceedings.mlr.press/v106/thawani19a.html>

Türkoğlu, M. F. (2021, 14 Temmuz). Support Vector Machine Algoritması-Makine Öğrenmesi-. Medium. 2 Temmuz 2025 tarihinde <https://mfatihto.medium.com/support-vector-machine-algoritmas%C4%B1-makine-%C3%B6%C4%9Frenmesi-8020176898d8> adresinden erişildi.

What Is Support Vector Machine? | IBM. (2023, 12 Aralık). 2 Temmuz 2025 tarihinde <https://www.ibm.com/think/topics/support-vector-machine> adresinden erişildi.

Yao, L., Mao, C. ve Luo, Y. (2018, 13 Kasım). Graph Convolutional Networks for Text Classification. arXiv. <https://doi.org/10.48550/arXiv.1809.05679>

Zhao, Y., Li, T., Yuan, Q. ve Deng, S. (2024). How to detect fake online physician reviews: A deep learning approach. DIGITAL HEALTH, 10, 20552076241277171. <https://doi.org/10.1177/20552076241277171>

GEÇİCİ KAPAK

*Kapak tasarımı
devam ediyor.*