YÖNETİM BİLİŞİM SİSTEMLERİ ALANINDA UYGULAMA, KURAM VE KAVRAMLAR

Editör
Dr. Öğr. Üyesi Mehmet Fatih KARACA

DUVAR

YÖNETİM BİLİŞİM SİSTEMLERİ ALANINDA UYGULAMA, KURAM VE KAVRAMLAR

Editör:
Dr. Öğr. Üyesi Mehmet Fatih KARACA

Yönetim Bilişim Sistemleri Alanında Uygulama, Kuram ve Kavramlar
Editör: Dr. Öğr. Üyesi Mehmet Fatih KARACA

Genel Yayın Yönetmeni: Berkan Balpetek
Yayın Tarihi: Ekim 2024
Yayıncı Sertifika No: 49837
ISBN: 978-625-6183-23-0

853 Sokak No:13 P.10 Kemeraltı-Konak/İzmir
Tel: 0 232 484 88 68

www.duvaryayinlari.com
duvarkitabevi@gmail.com

İÇİNDEKİLER

1.Bölüm..... 5
Yükseköğretim Lisans Düzeyindeki Bilgisayar ve
Bilişim Programlarının İncelenmesi
Examination Of Computer and Informatics Programs at The
Undergraduate Level of Higher Education
Mehmet Fatih KARACA

2. Bölüm 22
Makine Öğrenmesi ve Derin Öğrenme Teknikleri ile
Balık Türlerinin Sınıflandırılması: Aktarımlı Öğrenme Yaklaşımı
Classification of Fish Species with Machine Learning and
Deep Learning Techniques: A Transfer Learning Approach
Vahid SİNAP

3. Bölüm 48
Web Madenciliği ve Veri Analitiği Tabanlı Gelişmiş
Karar Destek Sistemlerinin İnşası
Web Mining and Data Analytics Based Building Advanced
Decision Support Systems
Hüseyin PARMAKSIZ, Önder ÖZTÜRK

4. Bölüm 79
Veri Bilimi ve Analitiği
Data Science and Analytics
Üzeyir FİDAN

5. Bölüm 102
Yapay Zekâ
Artificial Intelligence
Ali ERBEY

1. Bölüm

Yükseköğretim Lisans Düzeyindeki Bilgisayar ve Bilişim Programlarının İncelenmesi

Examination Of Computer and Informatics Programs at The Undergraduate Level of Higher Education

Mehmet Fatih KARACA[1]

[1] Dr. Öğr. Üyesi, Tokat Gaziosmanpaşa Üniversitesi, Erbaa Sosyal ve Beşeri Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ORCID: 0000-0002-7612-1437, mehmetfatih.karaca@gop.edu.tr

1. Giriş

Gelişen teknolojiler ve yenilikler yeni mesleklerin ortaya çıkmasına, bazı mesleklerin başka mesleklere evrilmesine veya bazı mesleklerin tümüyle ortadan kalkmasına neden olmaktadır. Bilgisayarların hayatımıza girmesiyle birlikte bilgisayar ve bilişimle alakalı çeşitli mesleklerin ortaya çıkması buna örnek olarak gösterilebilir. Özellikle son dönemde insanların bilgisayar, bilişim ve teknolojiye olan ilgileri, yapay zekânın popülaritesinin artması, her geçen gün farklı amaca hizmet eden yapay zekâ teknolojilerinin tanıtılması gibi etkenler bu alanda görev yapacak kişilere olan ihtiyacı göz önüne sermektedir.

Kariyerini bilgisayar ve bilişim alanında inşa etmek isteyen kişiler, yükseköğretim kurumlarında bulunan farklı isimlerdeki ön lisans, lisans ve lisansüstü programlarda eğitim alma imkanına sahiptirler. Ön lisans düzeyinde bilgisayar programcılığı, bilişim güvenliği teknolojisi, bilgisayar teknolojisi, bilgisayar destekli tasarım ve animasyon programları; lisans düzeyinde ise bilgisayar mühendisliği, yazılım mühendisliği ve yönetim bilişim sistemleri programları diğerlerine göre öne çıkmaktadır.

Yükseköğretim kurumları mevcut programlara ek olarak çağın gereğine uygun olarak ihtiyacı karşılayacak yeni programlar açmaktadır. Bilgisayar ve bilişimle ilgili lisans programları arası benzerlik ve farklılıklar bulunduğunu, Yükseköğretim Kurulu'nun (YÖK) amacının program çeşitliliğini arttırmak değil; ihtiyaca uygun yeni programlar açmak olduğunu belirtmek gerekir. Bu programlardan biri olan, 2024'te açılan yapay zekâ ve bilişim temelli programların devlet üniversitelerinde doluluk oranının %100'e ulaşması bunun örneklerinden biridir (YÖK, 2024).

Yükseköğretim kurumlarında adında bilgisayar veya bilişim geçen 11 farklı isimde lisans programı bulunmaktadır. Bunlar içerisinde Türkiye'de ilk kurulan program bilgisayar mühendisliğidir. 1977 yılında Hacettepe Üniversitesi ve Ortadoğu Teknik Üniversitesi'nde lisans düzeyinde bilgisayar mühendisliği programında eğitime başlanmıştır (Vikipedi, 2024). Yazılım mühendisliği, ilk olarak İzmir Ekonomi Üniversitesi'nde (İzmir Ekonomi Üniversitesi, 2024), ikinci olarak 2005'te Atılım Üniversite'sinde (Atılım Üniversitesi, 2024) kurulmuştur. Yönetim bilişim sistemleri programı ise ilk olarak 1995 yılında Boğaziçi Üniversitesi'nde kurulmuştur (Boğaziçi Üniversitesi, 2024).

Bilgisayar ve bilişim alanında en çok sayıda programa sahip bilgisayar mühendisliği, yazılım mühendisliği ve yönetim bilişim sistemleri programlarında her ne kadar ortak çalışma alanları bulunsa da farklılıklar yer almaktadır. Bilgisayar mühendisliği programı en kapsayıcı ve kapsamlı lisans düzeyindeki bilgisayar bölümüdür. Bu bölümde donanım ve yazılım konularına odaklanan,

bilgisayar sistemlerinin tasarlanması, geliştirilmesi, optimize edilmesi ve uyumlu şekilde çalışmasından sorumlu bireyler yetiştirmeye yönelik eğitim verilmektedir. Buna karşın yazılım mühendisliğinde yazılım geliştirme sürecine odaklanılmakta; yazılımın tasarımı, kodlanması, doğrulanması, test edilmesi, bakımı ve iyileştirilmesi işlemlerinde görev alacak bireyler yetiştirmek amaçlanmaktadır. Yönetim bilişim sistemleri ise işletme ve bilgisayar bilimini kapsayan disiplinler arası bir programdır. Bu programda bilişim sistemlerindeki gelişmelerin ve teknolojilerin işletme alanında kullanılmasına, karar verme süreçlerinde bunlardan yararlanılmasına dönük eğitimler verilmektedir. Diğer lisans programlarında da bunlar gibi benzerlikler görülse de her programın kendine has yeterlilikleri bulunmakta; program amaç ve kazanımları farklılık göstermektedir.

Lisans programlarına sayısal, eşit ağırlık, sözel ve dil puanı türlerinden öğrenci kabul edilmektedir. Buna karşın; bilgisayar ve bilişimle alakalı lisans programlarına öğrenci kabulü sayısal ve eşit ağırlık puan türleriyle gerçekleşmektedir. Puan türünün programın içeriği ve kapsamı hakkında bilgi sunduğu söylenebilir. İşletme ve bilgisayar biliminin ortak çalışma alanı olan yönetim bilişim sistemleri eşit ağırlık puan türünde; bunun dışındaki program adında bilgisayar veya bilişim olanların tamamı ise sayısal puan türünde öğrenci almaktadır.

Lisans düzeyinde bir programa yerleşmek isteyen öğrencilerin girmesi gereken 2 aşamalı sınav bulunmaktadır; Temel Yeterlilik Testi (TYT) ve Alan Yeterlilik Testi (AYT). Bunlardan farklı olarak bir de dil puanıyla öğrenci alan programlara yerleşmek için gerekli olan Yabancı Dil Testi vardır. TYT'de Türkçe, sosyal bilimler, temel matematik ve fen bilimleri testleri; AYT'de Türk dili ve edebiyatı, sosyal bilimler, matematik ve fen bilimleri testleri yer almaktadır. 1. aşama olan TYT sınavından sonra 2. aşama olan AYT uygulanmaktadır. Bu sınavların sonucunda öğrencilerin farklı puan türlerindeki puanları hesaplanmaktadır.

Alan yazında bilgisayar ve bilişimle ilgili programları çeşitli açılardan inceleyen araştırmalar mevcuttur. Alaybeyoğlu ve Morkaya (2006), Türkiye'deki bilgisayar ve yazılım mühendisliği lisans eğitimlerini karşılaştırmış; bilgisayar teknolojilerinin hızla gelişmesi nedeniyle diğer mühendisliklere kıyasla bilgisayar ve yazılım mühendisliği programlarının belirli periyotlarla yenilenmesi gerektiğini, bilgisayar ve yazılım mühendislerinden beklentilerin farklı olduğunu, sektörlerde bilişim sistemlerinin kullanımının artmasıyla bu sistemlerde görev alacak kişilere olan ihtiyacın da artacağını belirtmişlerdir. Macit (2023), bilişim sistemleri ve teknolojileri programını incelemiş; programın doluluk oranının yüksekliğini, derslerin sektör beklentileri ile uyumlu olduğunu,

farklı birimlerdeki aynı derslerin içeriklerinin tutarlılık gösterdiğini fakat aynı içeriğe sahip bazı derslerin birimlerde farklı şekillerde isimlendirildiğini ifade etmiştir. Damar (2022a), dijital dünyadaki gelişmelerin bilişim sektörünün gelişimi üzerine etkilerini analiz etmiş; bilişim sektöründe öncü olan ABD, Almanya ve Japonya gibi ülkelerin önceki yıllarda söz sahibiyken sonrasında farklı ülkelerin yürüttüğü başarılı politikalarla söz sahibi olmaya ve dijitalleşen dünyada varlık göstermeye başladığını dile getirmiştir. Damar (2022b) çalışmasında dijital çağda bilişim sektörünün ihtiyacı olan yetkinlikleri ele almış; Türkiye'deki bilişim sektörünün gelişebilmesi için bilişimle ilgili STK'lardan faydalanılması gerekliliğini, sektörün üniversiteler, STK'lar ve çalışanlar ile bağlantı kurmasının faydalı olacağını, böylece sektörel gelişim ve problemlerin aşılmasında sürekli iyileşmenin görülebileceğini vurgulamıştır. Damar (2022c), Türkiye gibi gelişmekte olan ve genç nüfusa sahip ülkeler açısından yazılım sektörünün birçok fırsat barındırdığına, yazılım sektöründeki gelişme ve ilerlemelere uygun hareket eden ülkelerin bundan pozitif, bunun dışında kalan ülkelerin ise negatif etkileneceğine, bilişim sektöründe anı yakalayarak harekete geçen ülkelerin başarıyı yakaladığına değinmiştir. Erden Özsoy ve Tosunoğlu (2023), yakın gelecekte yeni bazı meslekler ortaya çıkarken bazı mesleklere olan ihtiyacın ortadan kalkacağına, gelecekte ihtiyaç duyulan meslek ve becerilerin gelişmelere uyum sağlayacak şekilde değişmesinin beklendiğine, geleceğin meslek ve becerilerini etkileyecek veya belirleyecek olan metatrendlerin iyi anlaşılmasına, doğru politika ve yönlendirmelerle beşerî sermayenin niteliğinin artmasının ve işgücüne kazandırılmasının gerekliliğine temas etmiştir. Türkiye İş Kurumu (2024) tarafından yayımlanan raporda yükseköğretim düzeyinde açık iş olan 10 meslekten 3'ünün bilgisayar mühendisi, yazılım mühendisi ve yazılım geliştiricisi olduğu; bilgi ve iletişim sektöründeki geleceğin 10 mesleğinin yazılım mühendisi, yapay zekâ uzmanı, bilgisayar mühendisi, yazılım geliştiricisi, bilişim uzmanı, veri analisti, yapay zekâ mühendisi, veri bilimci ve web editörü sıralamasında bulunduğu; dahası yazılım mühendisi, bilişim uzmanı, yapay zekâ uzmanı ve yazılım geliştirici mesleklerinin finans ve sigorta alanında da gelecekteki ilk 10 meslekte yer aldığı görülmüştür. Buradan hareketle bilgisayar ve bilişim bölümlerinin ilerleyen zamanda öneminin artacağı, buna bağlı olarak popülaritesini sürdüreceği ve geleceğin meslekleri arasında yer alacağı; bu alanda başarılı olmak isteyen ülkelerin kurum/kuruluşlarının koordinasyonuyla sektörü sürekli güncellemelerle canlı tutmaları ve dinamik bir süreç yürütmeleri gerektiği söylenebilir.

Bu çalışmada adında bilgisayar veya bilişim geçen 11 yükseköğretim lisans programının 2022 ile 2024 yılları arasındaki yerleştirme sonuçları ile programlara ait bilgilerin ortaya konulması amaçlanmıştır. Programların isimleri, yıllara göre

sayıları, ortalama kontenjan ve yerleşen öğrenci sayıları, ortalama doluluk oranları, ortalama taban başarı sırası ve taban puanları, programların bulunduğu fakülte/yüksekokula göre sayıları, ayrıca eğitim dilleri incelenmiştir. Bu sayede bu programların mevcut durumunun ve bunların yıllara göre değişimlerin belirlenmesi hedeflenmektedir.

2. Yükseköğretim Programlarına Ait Veriler

Adında bilgisayar veya bilişim geçen, Türkiye'deki devlet üniversitelerinde faaliyet gösteren, ücretsiz ve örgün eğitim veren programlarla gerçekleştirilen bu çalışmada, Yükseköğretim Program Atlasındaki lisans tercih sihirbazında kamuoyuyla paylaşılan 2022 ile 2024 yılları arasındaki resmi veriler kullanılmıştır. (YÖK Program Atlası, 2024). 2022 yılından 206, 2023 yılından 218 ve 2024 yılından 228 olmak üzere toplamda 652 program analiz edilmiştir.

Çalışmanın kapsamını oluşturan programların adları, programlar için oluşturulan etiket ve puan türü bilgileri Tablo 1'de verilmiştir. Üniversitelerde 11 farklı programda öğretim gerçekleştirildiği, yönetim bilişim sistemleri dışındaki programların tamamının sayısal puan türünde öğrenci kabulü yaptığı, sözel ve dil puanı türlerinde ise bir program bulunmadığı belirlenmiştir.

Tablo 1. Çalışmaya Dahil Edilen Lisans Programları, Etiketleri ve Puan Türleri

Program Adı	**Etiket**	**Puan Türü**
Bilgisayar Mühendisliği	Bilg. Müh.	Sayısal
Yazılım Mühendisliği	Yaz. Müh.	Sayısal
Bilişim Sistemleri Mühendisliği	Blş. Sis. Müh.	Sayısal
Bilişim Sistemleri ve Teknolojileri	Blş. Sis. Ve Tekn.	Sayısal
Bilgisayar Teknolojisi ve Bilişim Sistemleri	Bilg. Tekn. Ve Blş. Sis.	Sayısal
Bilgisayar Bilimleri	Bilg. Bil.	Sayısal
Matematik ve Bilgisayar Bilimleri	Mat. Ve Bilg. Bil.	Sayısal
İstatistik ve Bilgisayar Bilimleri	İst. Ve Bilg. Bil.	Sayısal
Adli Bilişim Mühendisliği	Ad. Blş. Müh.	Sayısal
Bilgisayar ve Öğretim Teknolojileri Öğretmenliği	Bilg. Ve Öğrt. Tekn. Öğr.	Sayısal
Yönetim Bilişim Sistemleri	Yön. Blş. Sist.	Eşit Ağırlık

2.1. Program Sayıları

Yıllara göre program sayılarının sunulduğu Tablo 2 incelendiğinde en fazla sayıda bilgisayar mühendisliği, yönetim bilişim sistemleri ve yazılım mühendisliği programlarının; en az sayıda ise bilişim sistemleri mühendisliği ve adli bilişim mühendisliği programlarının olduğu görülmektedir. Bilgisayar

mühendisliği, yazılım mühendisliği, bilişim sistemleri ve teknolojileri, bilgisayar bilimleri ile yönetim bilişim sistemleri program sayılarının her yıl arttığı; farklı üniversitelerde veya farklı fakülte/yüksekokullarda açılmaya devam ettiği gözlenmiştir.

2024 verilerine göre 117 farklı üniversitede ve 76 farklı ilde bilgisayar veya bilişimle ilgili program bulunurken; Ağrı, Hakkâri, Kilis, Ordu ve Uşak'ta bu programlardan yoktur. Program sayısı en fazla olan iller sırasıyla 20 programla İstanbul, 17 programla Ankara, 11 programla İzmir'dir. Bölgelere göre program sayıları sırasıyla 51 programla İç Anadolu, 49 programla Marmara, 32 programla Akdeniz, 30 programla Karadeniz, 29 programla Doğu Anadolu, 27 programla Ege ve 10 programla Güneydoğu Anadolu Bölgesi'dir. En fazla sayıda program bulunan Burdur Mehmet Akif Ersoy Üniversitesi'nde bilgisayar mühendisliği, bilişim sistemleri mühendisliği, bilişim sistemleri ve teknolojileri (2 farklı birimde), yazılım mühendisliği ile yönetim bilişim sistemleri programlarının 6 farklı birimde faaliyet gösterdiği tespit edilmiştir.

Tablo 2. Programların Yıllara Göre Sayıları

	Bilg. Müh.	**Yaz. Müh.**	**Blş. Sis. Müh.**	**Blş. Sis. Ve Tekn.**	**Bilg. Tekn. Ve Blş. Sis.**	**Bilg. Bil.**	**Mat. Ve Bilg. Bil.**	**İst. Ve Bilg. Bil.**	**Ad. Blş. Müh.**	**Bilg. Ve Öğrt. Tekn. Öğr.**	**Yön. Blş. Sist.**	**Tümü**
2022	109	25	4	7	1	1	2	2	1	13	41	206
2023	110	28	4	7	1	5	2	2	1	13	45	218
2024	114	30	4	11	1	5	2	2	1	13	45	228

2.2. Programların Ortalama Kontenjanları, Yerleşen Sayıları ve Doluluk Oranları

Tablo 3'te yıllara göre programların ortalama kontenjanları verilmiştir. 2022 kontenjan sayılarına genel ve okul birincisi, 2023 ve 2024 kontenjan sayılarına ise genel, okul birincisi, şehit-gazi yakını, 34 yaş üstü kadın ve depremzede aday sayıları dahil edilmiştir.

En yüksek ortalama kontenjan ve yerleşen öğrenci sayısına sahip programlar sırasıyla bilgisayar mühendisliği, yazılım mühendisliği ile matematik ve bilgisayar bilimleri iken en düşük olanlar bilgisayar ve öğretim teknolojileri öğretmenliği, istatistik ve bilgisayar bilimleri ile bilgisayar bilimleri programlarıdır.

Bilgisayar mühendisliği, yazılım mühendisliği, istatistik ve bilgisayar bilimleri, adli bilişim mühendisliği, bilgisayar ve öğretim teknolojileri öğretmenliği ile yönetim bilişim sistemleri programlarının kontenjan ve yerleşen öğrenci sayılarının sürekli arttığı belirlenmiştir.

2024 yılı kontenjan değişimlerinde en dikkat çekici olanlar yazılım mühendisliğinin %17,98, bilişim sistemleri mühendisliğinin %23,18, bilgisayar teknolojisi ve bilişim sistemlerinin %18,52 ve yönetim bilişim sistemlerinin %14,57 artmasıdır. Bütün programlar birlikte değerlendirildiğinde kontenjanlar 2023'te %5,39, 2024'te ise %7,62 oranında artmıştır.

Tablo 3. Programların Yıllara Göre Ortalama Kontenjanları

	Bilg. Müh.	**Yaz. Müh.**	**Blş. Sis. Müh.**	**Blş. Sis. Ve Tekn.**	**Bilg. Tekn. Ve Blş. Sis.**	**Bilg. Bil.**	**Mat. Ve Bilg. Bil.**	**İst. Ve Bilg. Bil.**	**Ad. Blş. Müh.**	**Bilg. Ve Öğrt. Tekn. Öğr.**	**Yön. Blş. Sist.**
2022	70,89	55,96	55,00	53,14	52,00	62,00	62,00	36,50	50,00	28,69	52,85
2023	75,83	59,50	55,00	58,43	54,00	46,40	64,00	40,50	62,00	29,85	55,40
2024	78,84	70,20	67,75	54,73	64,00	50,20	69,00	43,00	68,00	31,69	63,47

Tablo 4'te programlara göre ortalama yerleşen öğrenci sayıları verilmiştir. 2024 yılına ait veriler incelendiğinde en dikkat çekici değişimin yazılım mühendisliğindeki %17,93'lük, bilişim sistemleri mühendisliğindeki %23,18'lik, bilgisayar teknolojisi ve bilişim sistemlerindeki %19,23'lük ve yönetim bilişim sistemlerindeki %16,32'lik artış olduğu söylenebilir. Bütün programlar birlikte değerlendirildiğinde yerleşen sayısı 2023'te %5,11, 2024'te ise %7,00 oranında artmıştır.

Tablo 4. Programların Yıllara Göre Ortalama Yerleşen Sayıları

	Bilg. Müh.	Yaz. Müh.	Blş. Sis. Müh.	Blş. Sis. Ve Tekn.	Bilg. Tekn. Ve Blş. Sis.	Bilg. Bil.	Mat. Ve Bilg. Bil.	İst. Ve Bilg. Bil.	Ad. Blş. Müh.	Bilg. Ve Öğrt. Tekn. Öğr.	Yön. Blş. Sist.
2022	70,72	55,96	55,00	53,14	52,00	62,00	62,00	36,50	50,00	26,08	52,85
2023	75,78	59,50	55,00	56,71	52,00	46,00	62,00	39,00	62,00	28,62	54,24
2024	77,51	70,17	67,75	54,09	62,00	50,20	67,50	41,50	68,00	30,54	63,09

Tablo 5'te programların ortalama doluluk oranları verilmiştir. Kontenjan ve yerleşen sayısı verileriyle elde edilen doluluk oranlarına göre bilişim sistemleri mühendisliği ve adli bilişim mühendisliğinde son 3 yılda doluluk oranı %100; bunlara en yakın olarak yazılım mühendisliğinde 2022 ve 2023'te %100, 2024'te ise %99,95'tir. Bilgisayar ve bilişim bölümlerindeki doluluk oranlarının 2022 yılında bilgisayar ve öğretim teknolojileri öğretmenliği programı dışında neredeyse %100 iken 2023 yılında düşüş gösterdiği, 2024 yılında ise bilgisayar mühendisliği ve yazılım mühendisliği dışında doluluk oranlarının tekrar yükseldiği belirlenmiştir.

Tablo 5. Programların Yıllara Göre Ortalama Doluluk Oranları

	Bilg. Müh.	Yaz. Müh.	Blş. Sis. Müh.	Blş. Sis. Ve Tekn.	Bilg. Tekn. Ve Blş. Sis.	Bilg. Bil.	Mat. Ve Bilg. Bil.	İst. Ve Bilg. Bil.	Ad. Blş. Müh.	Bilg. Ve Öğrt. Tekn. Öğr.	Yön. Blş. Sist.
2022	99,75	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	90,88	100,00
2023	99,94	100,00	100,00	97,07	96,30	99,14	96,88	96,30	100,00	95,88	97,91
2024	98,31	99,95	100,00	98,84	96,88	100,00	97,83	96,51	100,00	96,36	99,40

2.3. Programların Taban Başarı Sırası ve Taban Puan Ortalamaları

11 programın yıllara göre taban başarı sıralamaları Tablo 6'da görülmektedir. Bilgisayar ve öğretim teknolojileri öğretmenliği dışındaki bütün programların taban başarı sırası 2023 yılında 2022 yılına göre düşmüş; 2024 yılında ise bilişim

sistemleri ve teknolojileri ile yönetim bilişim sistemleri dışındaki diğer programlarda yükselmiştir. 2022 ile 2024 arasında bilgisayar ve öğretim teknolojileri öğretmenliğinin taban başarı sırası sürekli artarken yönetim bilişim sistemleri programının taban başarı sırası sürekli azalmıştır. Programlar içerisinde en düşük taban başarı sırası 2022 ve 2024'te bilgisayar mühendisliğinde, 2023'te ise yazılım mühendisliğinde; en yüksek ise 2022 ve 2023'te bilişim sistemleri ve teknolojilerinde, 2024'te ise istatistik ve bilgisayar bilimleri programlarındadır. Taban başarı sırasının düşmesi daha yüksek sıralamalı öğrencilerin tercih ettiği anlamına geleceği için olumlu; yükselmesi de daha düşük sıralamalı öğrencilerin tercih ettiği anlamına geleceği için olumsuz olarak değerlendirilebilir.

2022 yılında tümü içindeki en düşük sıralamalı 3 program 315 taban başarı sıralı Boğaziçi Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliği, 878 taban başarı sıralı Orta Doğu Teknik Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliği, 1179 taban başarı sıralı İstanbul Teknik Üniversitesi, Bilgisayar ve Bilişim Fakültesi'ndeki bilgisayar mühendisliğidir. 2023 yılında tümündeki en düşük sıralamalı 3 program 716 taban başarı sıralı Orta Doğu Teknik Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliği, 775 taban başarı sıralı Boğaziçi Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliği, 795 taban başarı sıralı Boğaziçi Üniversitesi, Yönetim Bilimleri Fakültesi'ndeki yönetim bilişim sistemleri programıdır. 2024 yılında tüm programlardaki en düşük sıralamalı 3 program 845 taban başarı sıralı Boğaziçi Üniversitesi, Yönetim Bilimleri Fakültesi'ndeki yönetim bilişim sistemleri, 945 taban başarı sıralı Orta Doğu Teknik Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliği, 1008 taban başarı sıralı Boğaziçi Üniversitesi, Mühendislik Fakültesi'ndeki bilgisayar mühendisliğidir.

Tablo 6. Programların Yıllara Göre Ortalama Taban Başarı Sıraları

	Bilg. Müh.	**Yaz. Müh.**	**Blş. Sis. Müh.**	**Blş. Sis. Ve Tekn.**	**Bilg. Tekn. Ve Blş. Sis.**	**Bilg. Bil.**	**Mat. Ve Bilg. Bil.**	**İst. Ve Bilg. Bil.**	**Ad. Blş. Müh.**	**Bilg. Ve Öğrt. Tekn. Öğr.**	**Yön. Blş. Sist.**
2022	100.972	92.731	148.091	382.623	309.522	109.662	254.290	371.587	223.932	162.709	225.392
2023	90.478	91.044	124.956	347.292	301.786	108.822	212.965	323.561	194.099	177.373	184.013
2024	101.314	135.367	148.075	316.976	314.760	136.124	228.695	343.044	211.313	180.564	162.315

Üniversiteye yerleştirmede esas olanın sıralama olduğunu; yerleştirme puanın yıllara göre farklılaşabilen ve sınavın zorluk seviyesine göre değişkenlik gösterebilen dolaylı bir gösterge olduğunu ifade etmek gerekir. Tablo 7'de programların ortalama taban puanları verilmiştir. Bütün programlarda bir önceki yıla göre 2023 yılında taban puanlar yükselirken 2024 yılında düşmüştür. Ayrıca, tüm programlarda olmamak kaydıyla taban başarı sırasının düştüğü yıllarda genellikle taban puanlar artarken, yükseldiğindeyse düşmüştür.

Tablo 7. Programların Yıllara Göre Ortalama Taban Puanları

	Bilg. Müh.	**Yaz. Müh.**	**Blş. Sis. Müh.**	**Blş. Sis. Ve Tekn.**	**Bilg. Tekn. Ve Blş. Sis.**	**Bilg. Bil.**	**Mat. Ve Bilg. Bil.**	**İst. Ve Bilg. Bil.**	**Ad. Blş. Müh.**	**Bilg. Ve Öğrt. Tekn. Öğr.**	**Yön. Blş. Sist.**
2022	409,94	416,06	374,48	277,42	295,00	396,88	316,32	278,38	326,17	263,15	345,67
2023	425,75	422,36	394,43	295,53	307,15	405,12	342,75	301,16	347,38	329,71	356,45
2024	363,40	368,81	356,76	293,45	286,41	368,14	316,65	279,84	318,85	308,23	349,99

2.4. Programların Fakülte/Yüksekokullara Göre Sayıları

2024 yılında programların fakülte veya yüksekokulda faaliyet gösterme durumlarına göre sayıları Tablo 8'de sunulmuştur. 2022-2024 arasındaki 117 farklı üniversiteden 652 programın incelendiği bu çalışmada programların özelliklerine göre farklı fakülte veya yüksekokulda öğretim faaliyetlerini sürdürdüğü görülmüştür. Örneğin bilgisayar mühendisliği ve yazılım mühendisliği programlarının ağırlıklı olarak mühendislik, mühendislik ve doğa bilimleri, mühendislik ve mimarlık fakültelerinde; yönetim bilişim sistemleri programının ise iktisadi ve idari bilimler fakültesi ile işletme fakültesinde bulunduğu tespit edilmiştir.

Bilgisayar mühendisliği, yazılım mühendisliği, bilişim sistemleri mühendisliği, bilgisayar teknolojisi ve bilişim sistemleri, bilgisayar bilimleri, matematik ve bilgisayar bilimleri, istatistik ve bilgisayar bilimleri, adli bilişim mühendisliği ile bilgisayar ve öğretim teknolojileri öğretmenliği programlarının tamamının fakülte bünyesinde; bilişim sistemleri ve teknolojileri ile yönetim bilişim sistemleri programlarının hem fakülte hem de yüksekokul bünyesinde yer aldığı belirlenmiştir.

Tablo 8. Programların Fakülte/Yüksekokullara Göre Sayıları (2024)

	Bilg. Müh.	Yaz. Müh.	Blş. Sis. Müh.	Blş. Sis. Ve Tekn.	Bilg. Tekn. Ve Blş. Sis.	Bilg. Bil.	Mat. Ve Bilg. Bil.	İst. Ve Bilg. Bil.	Ad. Blş. Müh.	Bilg. Ve Öğrt. Tekn. Öğr.	Yön. Blş. Sist.
Bilgisayar ve Bilişim Bilimleri Fak.	1	1	1	1	-	-	-	-	-	-	-
Bilgisayar ve Bilişim Fak.	2	2	1	1	-	-	-	-	-	-	-
Bilgisayar ve Bilişim Teknolojileri Fak.	1	-	-	-	-	-	-	-	-	-	-
Eğitim Bilimleri Fak.	-	-	-	-	-	-	-	-	-	1	-
Eğitim Fak.	-	-	-	-	-	-	-	-	-	12	-
Elektrik-Elektronik Fak.	1	-	-	-	-	-	-	-	-	-	-
Fen Fak.	-	-	-	1	1	4	2	2	-	-	-
Fen-Edebiyat Fak.	-	-	-	-	-	1	-	-	-	-	-
İktisadi ve İdari Bilimler Fak.	-	-	-	1	-	-	-	-	-	-	19
İktisat Fak.	-	-	-	-	-	-	-	-	-	-	1
İşletme Fak.	-	-	-	-	-	-	-	-	-	-	10
İşletme ve Yönetim Bilimleri Fak.	-	-	-	-	-	-	-	-	-	-	1
Mühendislik Fak.	67	15	-	-	-	-	-	-	-	-	-
Mühendislik ve Doğa Bilimleri Fak.	15	6	-	-	-	-	-	-	-	-	-
Mühendislik ve Mimarlık Fak.	8	-	-	-	-	-	-	-	-	-	-
Mühendislik ve Teknoloji Fak.	1	-	-	-	-	-	-	-	-	-	-
Mühendislik, Mimarlık ve Tasarım Fak.	2	2	-	-	-	-	-	-	-	-	-
Mühendislik-Mimarlık Fak.	11	1	-	-	-	-	-	-	-	-	-
Sosyal ve Beşeri Bilimler Fak.	-	-	-	-	-	-	-	-	-	-	2
Teknoloji Fak.	5	3	2	-	-	-	-	-	1	-	-
Uygulamalı Bilimler Fak.	-	-	-	2	-	-	-	-	-	-	6
Yönetim Bilimleri Fak.	-	-	-	-	-	-	-	-	-	-	1
Uygulamalı Bilimler Yük.	-	-	-	3	-	-	-	-	-	-	2
Uygulamalı Teknoloji ve İşletmecilik Yük.	-	-	-	2	-	-	-	-	-	-	3

2.5. Programların Öğretim Diline Göre Sayıları

2024'te bilgisayar ve bilişim lisans programlarında Türkçe, İngilizce, Almanca ve Fransızca olmak üzere 4 dilde öğretim yapıldığı Tablo 9'dan

anlaşılmaktadır. 2024 verilerine göre programlarda ağırlıklı olarak Türkçe, sonrasında İngilizce, son olarak da Almanca ve Fransızca öğretim dili olarak kullanılmaktadır. Bütün dillerde öğretim imkânı olan tek program bilgisayar mühendisliği; yalnızca Türkçe öğretim imkânı olan programlar ise bilişim sistemleri mühendisliği, bilgisayar teknolojisi ve bilişim sistemleri, bilgisayar bilimleri, matematik ve bilgisayar bilimleri, istatistik ve bilgisayar bilimleri ile adli bilişim mühendisliğidir.

Tablo 9. Programların Öğretim Diline Göre Sayıları (2024)

	Bilg. Müh.	**Yaz. Müh.**	**Blş. Sis. Müh.**	**Blş. Sis. Ve Tekn.**	**Bilg. Tekn. Ve Blş. Sis.**	**Bilg. Bil.**	**Mat. Ve Bilg. Bil.**	**İst. Ve Bilg. Bil.**	**Ad. Blş. Müh.**	**Bilg. Ve Öğrt. Tekn. Öğr.**	**Yön. Blş. Sist.**
Türkçe	82	24	4	10	1	5	2	2	1	11	37
İngilizce	29	6	-	1	-	-	-	-	-	2	7
Almanca	2	-	-	-	-	-	-	-	-	-	1
Fransızca	1	-	-	-	-	-	-	-	-	-	-

3. Sonuç

Bu çalışmada ücretsiz ve örgün eğitim veren, adında bilgisayar veya bilişim geçen Türkiye'deki devlet üniversitelerinin 2022 ile 2024 yılları arasındaki 3 yıllık verileri analiz edilmiştir. 2022 yılından 206, 2023 yılından 218, 2024 yılından 228, toplamda ise 652 program incelenmiş; programların isimleri, sayıları, kontenjanları, yerleşen öğrenci sayıları, doluluk oranları, faaliyet gösterdikleri fakülte ve yüksekokullar ile öğretim dilleri verileri kullanılarak programların mevcut durumları karşılaştırmalı olarak değerlendirilmiştir.

Yükseköğretim kurumlarında adında bilgisayar veya bilişim bulunan 11 farklı programda öğretim gerçekleştirilmektedir. Farklı program sayısı program çeşitliliğinin, uzmanlık gerektiren alanlar için programların var olduğunun ve yeni programlar açma konusunda YÖK'ün proaktif davranış sergilediğinin göstergeleridir.

Hesaplama, analitik düşünme ve mühendislik gerektiren işlemler barındırması nedeniyle programların 10'unun sayısal puan türünde; işletme ve bilgisayar bilimini içine alan multidisipliner bir program olan yönetim bilişim sistemlerinin ise eşit ağırlık puan türünde öğrenci kabulü yaptığı görülmüştür.

2022 yılında 73 ilde ve 111 üniversitedeki 21 fakültede, 2023 yılında 73 ilde ve 112 üniversitedeki 24 fakültede, 2024 yılında 76 ilde ve 117 üniversitedeki 24 fakültede programların var olduğu; 2024 itibariyle Ağrı, Hakkâri, Kilis, Ordu ve Uşak'ta bu programlardan hiçbirinin olmadığı; Güneydoğu Anadolu Bölgesi'nde 10, diğer bölgelerde ise 25 üzerinde programın bulunduğu tespit edilmiştir. Özellikle 2024 verilerine göre sayıca diğerlerinden fazla olan bilgisayar mühendisliğinin 70, yönetim bilişim sistemlerinin 32, yazılım mühendisliğinin ise 25 farklı ilde; ayrıca yönetim bilişim sistemlerinin Güneydoğu Anadolu Bölgesi'nde olmaması dışında, bu 3 programın tüm bölgelerde faaliyet gösterdiği görülmüştür. Bu verilerden programların ülke geneline yayılmış şekilde farklı illerde ve bölgelerde aktif olduğu sonucu çıkarılabilir.

Mühendislik eğitimi veren bilgisayar mühendisliği ve yazılım mühendisliği programlarının en yüksek kontenjan ve yerleşen öğrenci sayısına sahip olduğu; bilişim sistemleri ve teknolojileri ile bilgisayar bilimleri programları dışındaki programların kontenjan ve yerleşen öğrenci sayılarının 3 yıl boyunca arttığı veya aynı kaldığı belirlenmiştir. Tüm programlardaki en düşük doluluk oranının 2022 yılında %90,88, 2023 yılında %95,88, 2024 yılında ise %96,36 olduğu; programların doluluk oranları değişimlerin yıllara göre küçük farklılıklar gösterdiği saptanmıştır. Bazı yıllarda değişiklikler gösterse de bu verilerden yola çıkarak öğrencilerin bilgisayar ve bilişim lisans programlarına olan ilgisinin devam ettiği, geleceğin mesleklerinden olan bu programların popülaritesinin sürdüğü değerlendirmelerinde bulunmak mümkündür.

Üniversiteye yerleştirmeler adayların başarı sıralarına göre; dolayısıyla sıralamayı belirleyen puanlarına göre yapılmaktadır. Sınavın zorluk seviyesine bağlı olarak puanlar, yıldan yıla önemli değişiklikler gösterebilmektedir. Buna karşın verileri yıllara göre karşılaştırırken puan yerine adayların başarı sıralamalarını kullanmak, daha ölçülebilir bir veri sağlaması nedeniyle daha doğru bir yaklaşım olarak değerlendirilebilir. Ek olarak taban başarı sırasının düşüklüğü daha başarılı adayın, yüksekliği de daha başarısız adayın ilgili programı tercih ettiğini ifade etmektedir. Çalışmanın veri kümesini oluşturan programlar içinde, 2022-2024 yılları arasında bilgisayar mühendisliği ile yazılım mühendisliği programlarının en düşük, bilişim sistemleri ve teknolojileri ile istatistik ve bilgisayar bilimleri programlarının en yüksek sıralamaya sahip programlar olduğu elde edilen verilerle ortaya konmuştur. En düşük sıralamalı programların 2022'de bilgisayar mühendisliği, 2023 ve 2024'te bilgisayar mühendisliği ile yönetim bilişim sistemleri olduğu; en yüksek sıralamalı programların üniversitelerinin 2022'de Boğaziçi Üniversitesi, Orta Doğu Teknik Üniversitesi ve İstanbul Teknik Üniversitesi, 2023 ve 2024'te ise Boğaziçi Üniversitesi ve Orta Doğu Teknik Üniversitesi olduğu gözlenmiştir. Buradan

hareketle diğerlerine göre daha başarılı olan öğrencilerin önceki yıllarda bilgisayar mühendisliğine, son 2 yıldır ise yönetim bilişim sistemleri ile bilgisayar mühendisliğine yerleştikleri sonucuna ulaşılabilinir.

Programlar fakülte ve yüksekokul bünyesinde yer almaları bakımından değerlendirildiğinde yalnızca bilişim sistemleri ve teknolojileri ile yönetim bilişim sistemleri programlarının hem fakülte hem yüksekokullarda, diğer programların tamamının yalnızca fakültelerde faaliyet gösterdiği; yazılım mühendisliği ve en fazla sayıda programa sahip bilgisayar mühendisliğinin ağırlıklı olarak mühendislik fakülteleri ile mühendislik ve doğa bilimleri fakültelerinde, yönetim bilişim sistemlerinin ise yoğun olarak iktisadi ve idari bilimler fakülteleri ile işletme fakültelerinde eğitim sürdürdükleri saptanmıştır.

Çalışma kapsamındaki 11 farklı programda Türkçe, İngilizce, Almanca ve Fransızca olmak üzere 4 dilde lisans düzeyinde öğretim imkânı bulunmaktadır. Programlarda en çok Türkçenin, sonrasında İngilizcenin, az sayıda ise Almanca ve Fransızcanın öğretim dili olarak kullanıldığı; bu 4 dilde öğretim imkânı bulunan tek programın bilgisayar mühendisliği olduğu; son 3 yılda Türkçe ve İngilizce program sayılarının sürekli arttığı; buna karşın Almanca ve Fransızca eğitim veren program sayısında bir değişiklik olmadığı tespit edilmiştir. Bilgisayar ve bilişimle ilgili programlarda, özellikle bilgisayar mühendisliği, yazılım mühendisliği ve yönetim bilişim sistemlerinde eğitim almak isteyen üniversite adaylarının dünya üzerinde yaygın olarak kullanılan dillerde mesleki bilgi kazanmalarına imkân sağlayacak şekilde dil çeşitliliğinin var olduğu görülmüştür.

Özetle; bilişim teknolojilerinde yaşanan gelişmelere paralel olarak her geçen yıl bilgisayar ve bilişimle ilgili yeni isimde lisans programlarının açıldığı; farklı üniversite veya birimlerdeki lisans programlarının sayısının her yıl arttığı; özellikle sayıca fazla olan programların il ve bölgelere dağılmış şekilde faaliyetlerini sürdürdükleri; programların içeriklerine, yeterliliklerine, amaç ve kazanımlarına uygun olan puan türünde öğrenci kabulü yaptıkları; genel olarak programlardaki kontenjan ve yerleşen öğrenci sayılarının sürekli arttığı; kontenjan, yerleşen öğrenci sayısı ve doluluk oranlarına göre geleceğin mesleklerinden olan bu programlara gösterilen ilginin yüksek olduğu ve bunun sürebileceği; yeni trendlere uygun yeni mesleklerin ortaya çıkabileceği; YÖK ve üniversitelerin yeni durumlara uygun yeni programlar açmak konusunda aktif bir tutum sergilediği; taban başarı sırası bakımından programlar arasında farklılıklar görüldüğü; bazı programların başarı sırası nispeten stabilken bazılarındaki değişimlerin yüksek seyrettiği; diğerlerine kıyasla bilgisayar mühendisliği ve yazılım mühendisliği programlarını daha başarılı adayların tercih ettiği; yalnızca 2 programın yüksekokul, bunlar dışındakilerinse fakülte bünyesinde yer aldığı;

özellikle sayısı fazla olan programlarda farklı dillerde eğitim alma imkânı bulunduğu; gerek program çeşitliliği, gerek programların ülke geneline yayılması gerekse de farklı dillerde eğitim alınabilmesi açısından, mezuniyet sonrası istihdamları da göz önüne alındığında, Türkiye'deki bilgisayar ve bilişim lisans program sayısının iyi düzeyde olduğu; üniversitelerdeki programların sayı ve çeşitlilik bakımından ihtiyaca ve dünyadaki trende göre değişim, dönüşüm, gelişim veya kapanma eğilimi gösterecek şekilde aktif hareket ettikleri sonuçlarına ulaşılmıştır.

Kaynakça

1. Alaybeyoğlu, A., & Morkaya, Ö. (2006). Ülkemizdeki Bilgisayar Mühendisliği Lisans Eğitimi ile Yazılım Mühendisliği Lisans Eğitiminin Karşılaştırılması. *3. Ulusal Elektrik Elektronik Bilgisayar Mühendislikleri Eğitimi Sempozyumu*, 16-18 Kasım 2023, İstanbul.
2. Atılım Üniversitesi (20 Eylül 2024). https://www.atilim.edu.tr/tr/se/page/4704/sikca-sorulan-sorular
3. Boğaziçi Üniversitesi (20 Eylül 2024). https://mis.bogazici.edu.tr/tr/node/24
4. Damar, M. (2022a). Dijital Dünyanın Dünü, Bugünü ve Yarını: Bilişim Sektörünün Gelişimi Üzerine Değerlendirme. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, 12(Dijitalleşme): 51-76. https://doi.org/10.30783/nevsosbilen.1121818
5. Damar, M. (2022b). Dijital Çağda Bilişim Sektörünün İhtiyacı Olan Yetkinlikler Üzerine Bir Değerlendirme. *Journal of Information Systems and Management Research*, *4*(1): 25-40.
6. Damar, M. (2022c). Yazılım Sektörünün İki Lider Ülkesi Hindistan ve İrlanda, Gelişmekte Olan Ülkeler İçin Öneriler. *Ege Eğitim Teknolojileri Dergisi*, *6*(1): 29-52. https://dergipark.org.tr/tr/pub/eetd/issue/70271/1142633
7. Erden Özsoy, C., & Tosunoğlu, B. T. (2023). Geleceğin Meslekleri ve Becerileri: Türkiye'de Ekonomik Kalkınma İçin Bir Fırsat Penceresi. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, *24*(3): 403-416. https://doi.org/10.37880/cumuiibf.1257166
8. İzmir Ekonomi Üniversitesi (20 Eylül 2024). https://se.ieu.edu.tr/tr/sss
9. Macit, H. B. (2023). Müfredat ve Sektör Açısından Devlet Üniversitelerimizdeki Bilişim Sistemleri ve Teknolojileri Bölümünün İncelenmesi. *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, *7*(2): 231-244. https://doi.org/10.31200/makuubd.1291649
10. Türkiye İş Kurumu, (Nisan, 2024). *2023 Yılı İşgücü Piyasası Araştırması Türkiye Raporu*. https://media.iskur.gov.tr/88081/turkiye.pdf
11. Vikipedi (20 Eylül 2024). https://tr.wikipedia.org/wiki/Bilgisayar_mühendisliği

12. YÖK (13 Ağustos 2024). https://www.yok.gov.tr/Sayfalar/Haberler/2024/2024-yks-yerlestirme-sonuclari-aciklandı.aspx
13. YÖK Program Atlası (13 Eylül 2024). https://yokatlas.yok.gov.tr/tercih-sihirbazi-t4.php

2. Bölüm

Makine Öğrenmesi ve Derin Öğrenme Teknikleri ile Balık Türlerinin Sınıflandırılması: Aktarımlı Öğrenme Yaklaşımı

Classification of Fish Species with Machine Learning and Deep Learning Techniques: A Transfer Learning Approach

Vahid SİNAP[1]

[1] Dr. Öğr. Üyesi, Ufuk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ORCID: 0000-0002-8734-9509, vahidsinap@gmail.com

1. Giriş

Makine öğrenmesi teknikleri kullanılarak balık türlerinin sınıflandırılması ve segmentasyonu, su ekosistemleri hakkındaki anlayışımızı derinleştirmekte, balıkçılık yönetiminde daha etkili stratejiler geliştirilmesine katkıda bulunmakta ve çevresel koruma çalışmalarını desteklemektedir (Garcia vd., 2020). Bu alandaki ilerlemeler balık türlerinin doğru bir şekilde tanımlanmasını sağlayarak, sürdürülebilir balıkçılık uygulamalarının daha verimli bir şekilde uygulanmasına olanak tanımaktadır. Farklı balık türlerinin doğru bir şekilde ayırt edilebilmesi, sürdürülebilir ve etkili balıkçılık uygulamalarının hayata geçirilmesi açısından temel bir gerekliliktir.

Balık sınıflandırması, çeşitli balık türlerinin sistematik olarak kategorize edilmesi ve tanımlanmasını ifade etmektedir. Bu süreç, her türe özgü benzersiz özelliklerin, desenlerin ve karakteristiklerin belirlenmesini içermektedir (Fischer, 2014). Segmentasyon ise görüntü işleme bağlamında, bir görüntünün anlamlı ve farklı bölgelere ayrılması anlamına gelmektedir (Zhang vd., 2008). Balık sınıflandırması söz konusu olduğunda segmentasyon, bir görüntüdeki bireysel balıkların sınırlarını belirlemeye yardımcı olarak, bu balıkların niteliklerinin ve özelliklerinin hassas bir şekilde incelenmesini mümkün kılmaktadır.

Doğru balık sınıflandırması, balıkçılık yönetiminde büyük bir rol oynamaktadır. Farklı balık türlerine ait verilerin doğruluğu, etkili koruma stratejilerinin oluşturulmasında kritik bir öneme sahiptir (Fischer, 2014). Hassas balık sınıflandırmasının önemi, yalnızca su ekosistemlerindeki biyolojik çeşitliliğin kayıt altına alınmasından ibaret değildir. Aynı zamanda sürdürülebilir balıkçılık uygulamaları, yasal düzenlemelere uyum ve hassas ekolojik dengelerin korunması için de temel bir gerekliliktir (Potts ve Haward, 2007). Doğru bir sınıflandırma, belirli bölgelerdeki balık türlerinin ve popülasyonlarının daha iyi anlaşılmasına olanak sağlar ki bu da etkili yönetim stratejilerinin oluşturulabilmesi için zorunlu bir adımdır (Arthington vd., 2016). Bunun yanı sıra, nesli tükenme tehlikesi altındaki türlerin korunması söz konusu olduğunda da türlerin doğru bir şekilde sınıflandırılması hayati önem taşımaktadır (Roberson vd., 2020). Bu sınıflandırma, koruma çabalarının etkili olabilmesi için yasal düzenlemelerin temelini oluşturmakta ve türlerin ihtiyaçlarına uygun koruma stratejilerinin geliştirilmesini sağlamaktadır. Özellikle tehdit altındaki türlerin yaşam alanlarının korunması ve türlerin devamlılığının sağlanabilmesi, hedefe yönelik sıkı koruma önlemlerinin alınmasını gerektirmektedir. Ayrıca, sucul ekosistemlerde farklı balık türleri arasındaki hassas denge, ekosistemin sağlıklı bir şekilde işlemesi için kritik öneme sahiptir. Bu dengeyi korumak ve ekolojik ilişkileri anlamak açısından doğru sınıflandırma belirleyici bir rol oynamaktadır.

Türler arasındaki etkileşimlerin ve ekosistemin dinamiklerinin doğru anlaşılması, olası tehditlerin önceden tespit edilmesi ve müdahale edilmesi için gereklidir (Scheffer vd., 2001). Doğru sınıflandırma yalnızca balıkçılık yönetimine fayda sağlamakla kalmamakta, farklı türler arasındaki karmaşık etkileşimleri koruyarak sucul ortamların dayanıklılığını ve sürdürülebilirliğini sağlamada proaktif bir adım olarak da öne çıkmaktadır (Stefanakis ve Becker, 2016).

Deniz biyolojisi alanında geleneksel olarak yapılan balık sınıflandırması ve segmentasyonu doğal sınırlamalara sahip manuel yöntemlere dayanmaktadır. Bu geleneksel yöntemler, deniz biyologlarının yoğun insan emeği harcamasını gerektiren, iş gücü yoğun süreçlerdir. Bu yöntemler zaman açısından verimlilik sorunları yaratırken, balık sınıflandırması ve segmentasyonunda hata yapma olasılığını artırmaktadır (Pitcher, 2008). Ayrıca, bu manuel yöntemler kişisel yaklaşımlara dayalı olduğundan, görsel verilerin farklı gözlemciler tarafından farklı yorumlanması sonucu öznel bir değerlendirme ortaya çıkabilmektedir. Bu da süreci daha karmaşık hale getirmektedir. Manuel sınıflandırıcılar, sucul görüntülerdeki karmaşık desenler, renk tonları ve morfolojik çeşitliliği yönetmekte de zorlanabilmektedirler (Ali-Gombe vd., 2017). Özellikle balık türleri arasındaki ince ayırt edici özellikler veya büyük veri setleri söz konusu olduğunda, manuel sınıflandırma daha da zorlaşmaktadır (Alsmadi ve Almarashdeh, 2022). Bunun yanı sıra, sucul ekosistemlerin sürekli manuel olarak izlenmesi ve kayıt altına alınması pratikte mümkün değildir. Bu da ekosistemlerin dinamiklerini doğru bir şekilde anlamak ve koruma stratejileri geliştirmek açısından önemli bir kısıt oluşturmaktadır. Makine öğrenmesi teknolojilerinin gelişimi, balık sınıflandırması ve segmentasyonu alanında köklü bir değişim yaratma potansiyeline sahip olup, bu süreçlerin daha hızlı ve daha doğru bir şekilde gerçekleştirilmesine olanak tanıyabilir.

Yapay zekânın bir alt dalı olan makine öğrenmesi, algoritmaların veriden bilgi edinmesini sağlayarak belirli görevleri yerine getirebilecek eğitimli modellerin oluşturulmasına olanak tanımaktadır. Bu teknoloji, programların önceden belirlenmiş programlama kurallarına doğrudan bağlı kalmaksızın verilerden öğrenmesini sağlamayı amaçlamaktadır (Batista vd, 2004). Makine öğrenmesi teknolojisi, balık sınıflandırma ve segmentasyonu alanında geleneksel yöntemlerin pek çok sorununa çözüm sunmaktadır. Manuel sınıflandırmanın getirdiği zaman ve iş gücü yoğunluğu gibi sorunları aşarak, bu süreçleri büyük ölçüde otomatikleştirebilmektedir. Bu da araştırmacılara daha verimli çalışma imkânı sunarken, sınıflandırma ve segmentasyon görevlerinin hızını artırmaktadır. Ayrıca, makine öğrenmesi algoritmalarının geniş veri setleriyle çalışabilme kapasitesi, özellikle sucul ekosistemlerdeki büyük ve karmaşık veri yığınlarını yönetme konusunda büyük bir avantaj sağlamaktadır (Alaba vd.,

2022). Böylece, büyük veri miktarlarıyla bile etkin bir şekilde başa çıkılabilmektedir. Bu teknoloji aynı zamanda, manuel yöntemlerde ortaya çıkan öznel kararların ve gözlemciden kaynaklanan hataların önüne geçmektedir. Sistematik bir şekilde öğrenen makine öğrenmesi algoritmaları, farklı gözlemciler arasında tutarlılığı sağlarken, veriye dayalı ve nesnel sonuçlar üretmektedir (Cabreira vd., 2009). Görsel verilerin analizinde yüksek doğruluk sağlayan derin öğrenme modelleri, sucul görüntülerdeki karmaşık desenleri, renk varyasyonlarını ve ince morfolojik farklılıkları daha iyi algılayarak sınıflandırma hatalarını minimize etmektedir. Ayrıca, makine öğrenmesi teknolojisi, zamanla yeni verilerle güncellenebilir yapısı sayesinde kendini geliştirme kapasitesine sahiptir. Yeni balık türleri keşfedildikçe veya modellerin performansı artırıldıkça, makine öğrenmesi algoritmaları bu güncellemelere uyum sağlayabilmektedir. Bu da sürdürülebilir balıkçılık uygulamalarını desteklemeye katkıda bulunan bir etmendir.

Makine öğrenmesi teknolojisinin sunduğu avantajlarla birlikte, bu alandaki bazı sınırlamalar da göz ardı edilmemelidir. Algoritmaların doğru sonuçlar üretebilmesi için büyük miktarda yüksek kaliteli veriye ihtiyaç duyulmaktadır. Verilerin eksik, hatalı ya da dengesiz olması, algoritmaların performansını olumsuz yönde etkileyebilmekte ve yanlış sınıflandırmalara yol açabilmektedir. Ayrıca, kullanılan modellerin karmaşıklığı arttıkça, hesaplama gücü gereksinimleri de artmakta, bu da özellikle büyük veri kümeleri üzerinde çalışan sistemlerin zaman ve maliyet açısından verimsiz olmasına neden olabilmektedir. Derin öğrenme gibi daha ileri düzey teknikler, genellikle "kara kutu" olarak adlandırılan, sonuçların nasıl elde edildiğini anlamayı zorlaştıran bir yapıya sahiptir (Guidotti vd., 2018). Bu da özellikle biyoloji gibi alanlarda sonuçların açıklanabilirliğinin önem kazandığı durumlarda bir dezavantaj oluşturabilmektedir. Bu nedenle, makine öğrenmesi uygulamalarında dengeyi bulmak hem performansı artıracak hem de bu sınırlamaları en aza indirecek stratejilerin geliştirilmesi önemlidir.

Bu araştırmanın amacı, balık sınıflandırması ve segmentasyonu süreçlerinde VGG16, ResNet50, InceptionV3 ve MobileNet adlı dört popüler aktarımlı öğrenme (transfer learning) modelinin performanslarını karşılaştırmaktır. Bu dört model, özellikle görüntü işleme görevlerinde yaygın olarak kullanılan derin öğrenme tabanlı yöntemlerdir. Çalışma, her bir modelin balık türlerini ayırt etme ve sualtı canlılarının görüntülerinde segmentasyon yapma konusundaki başarımını inceleyerek, hangi modelin bu görevlerde daha verimli ve doğru sonuçlar verdiğini analiz etmeyi hedeflemektedir. Bu sayede, sürdürülebilir balıkçılık uygulamaları ve deniz ekosistemlerinin korunması için en uygun modelin belirlenmesine katkı sağlanacaktır. İlerleyen bölümlerde, balık

sınıflandırmasında sıklıkla kullanılan algoritmaların ayrıntılı bir analizi yapılacak; kullanılan yöntemler, veri toplama süreçleri ve balık türlerini ayırt etmek ve tanımlamak için geliştirilmiş modellerin performans karşılaştırmaları ele alınacaktır.

2. Balık Sınıflandırmasında Sıklıkla Kullanılan Algoritmalar

Mevcut sınıflandırma algoritmalarının birçoğu kimliklendirme, pazarlanabilirlik, fiyatlandırma, tüketim ve bilimsel araştırma gibi çeşitli alanlarda önemli bir rol oynayarak bu alana büyük katkılar sağlamaktadır (Aguado vd., 2016). Bu algoritmalar, sürdürülebilir balıkçılık uygulamalarını garanti altına almanın yanı sıra balıkçılık yönetimini desteklemekte ve farklı alanlarda bilinçli karar alma süreçlerini teşvik etmektedir. Aşağıda, bazı önemli algoritmalar açıklanmaktadır.

2.1. Evrişimli Sinir Ağları (Convolutional Neural Networks-CNN)

CNN'ler, balık sınıflandırması da dahil olmak üzere görüntü sınıflandırma görevlerinde yüksek başarı göstermektedir. Bu derin öğrenme modelleri, görüntülerden hiyerarşik özellikleri otomatik olarak öğrenerek farklı balık türlerini ayırt etmek için kritik olan karmaşık desenleri ve varyasyonları yakalayabilmektedirler (Salman vd., 2016). CNN ile gerçekleştirilen bu süreç genel itibariyle aşağıdaki gibi bir yol izlemektedir:

- Konvolüsyon: Konvolüsyon işlemi, CNN'lerin temelini oluşturan ve görüntüleri analiz etmek için kullanılan en önemli adımlardan biridir. Bu işlem, bir görüntü üzerinde küçük filtreler (kernel veya filtre adı verilen pencereler) kaydırarak (sliding window) görüntüdeki özellikleri çıkarmaktadır. Konvolüsyon katmanı, giriş görüntüsündeki belirli desenleri, kenarları, dokuları ve renk varyasyonlarını tanımlamak için filtreleri kullanarak bu desenlerin nerede ve ne kadar güçlü olduğunu belirlemektedir (Feldmann vd., 2021). Konvolüsyon katmanı, her pozisyondaki çıktıyı ($O_{i,j}$), Eşitlik 1'deki formülle hesaplamaktadır.

$$O_{i,j} = \sigma\left(\sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{m=1}^{M} W_{k,l,m} \cdot I_{i+k-1,j+l-1,m} + b_k\right) \qquad (1)$$

Eşitlik 1'deki denklemde, σ aktivasyon fonksiyonunu, $W_{k,l,m}$ filtre ağırlıklarını, $I_{i+k-1,j+l-1,m}$ giriş piksel değerlerini ve b_k ise yanlılık (bias) terimini ifade etmektedir.

- Havuzlama: Havuzlama (pooling) işlemi, CNN'lerde kullanılan bir aşağı örnekleme yöntemidir ve modelin hesaplama yükünü azaltarak daha hızlı ve verimli hale gelmesine olanak tanımaktadır. Bu işlem, girişten elde edilen

özellik haritalarının boyutunu küçültmek için kullanılmakta ve aynı zamanda modelin küçük varyasyonlara (pozisyon ve ölçek değişikliklerine) daha dayanıklı olmasını sağlamaktadır. Havuzlama sırasında, belirli bir boyutta bir pencere (örneğin 2x2 veya 3x3) giriş görüntüsü üzerinde kaydırılır ve her pencere için belirli bir havuzlama işlemi uygulanır (Sun vd., 2017). En yaygın havuzlama yöntemi maksimum havuzlamadır (max pooling). Bu yöntemde, pencere içerisindeki en yüksek değer seçilmekte ve çıkış haritasına yerleştirilmektedir. Böylece, görüntüdeki önemli özellikler korunurken, boyutlar küçültülür. Bu süreç, modelin öğrenme sürecini hızlandırmakta ve daha az bellek tüketimi ile daha iyi genel performans sağlamaktadır (Zafar vd., 2022). Maksimum havuzlama işlemi ise şu şekilde ifade edilmektedir:

$$O_{i,j} = \max_{k,l} I_{i\times s+k, j\times s+l} \quad (2)$$

Eşitlik 2'de s adım uzunluğunu, k, l ise havuzlama penceresindeki pikselleri temsil etmektedir.

2.2. Destek Vektör Makinesi (Support Vector Machine - SVM)

SVM, balık sınıflandırma görevlerinde yaygın olarak uygulanan güçlü bir denetimli öğrenme algoritmasıdır. SVM'nin temel prensibi, farklı balık sınıflarını özellikler uzayında etkili bir şekilde ayıran optimal bir hiper düzlem (hyperplane) bulmaktır. Bu hiper düzlem, her sınıfın en yakın veri noktaları (destek vektörleri) ile olan mesafeyi maksimize edecek şekilde stratejik olarak konumlandırılır (Cristianini ve Shawe-Taylor, 2000). SVM, sınıflar arasındaki en iyi ayrımı sağlamak için en zorlu durumları, yani destek vektörlerini dikkate almayı amaçlamaktadır.

Bir eğitim veri setini özellikler X ve ikili sınıflandırma için karşılık gelen etiketler $y \in \{-1,1\}$ olarak ele alındığında SVM'nin amacı, ağırlıklar $\mathbf{w}$ ve bias terimi b ile tanımlanan bir hiper düzlem bulmaktır. Böylece aşağıda verilen Eşitlik 3 ve Eşitlik 4 sağlanmaktadır:

$$\mathbf{w} \cdot \mathbf{x} + b \geq 1 \text{ için } y = 1 \quad (3)$$

$$\mathbf{w} \cdot \mathbf{x} + b \leq -1 \text{ için } y = -1 \quad (4)$$

Bu denklemler, her bir veri noktası $\mathbf{x}_i$ için bir kayma değişkeni ξ_i tanımlayarak Eşitlik 5'teki tek bir ifadeye birleştirilebilir:

$$y_i(\mathbf{w} \cdot \mathbf{x} + b) \geq 1 - \xi_i \text{ Tüm veri noktaları için} \quad (5)$$

2.3. K-En Yakın Komşu (K-Nearest Neighbors - KNN)

KNN, balık sınıflandırması da dahil olmak üzere sınıflandırma görevlerinde kullanılan basit ve sezgisel bir algoritmadır. KNN'deki temel fikir bir nesneyi, özellikler uzayındaki k-en yakın komşuları arasında en çok bulunan sınıfa göre sınıflandırmaktır. Başka bir deyişle, bir nesnenin sınıfı, özellikler uzayındaki en yakın veri noktalarının sınıflarına dayanarak belirlenir (Zhang vd., 2017). KNN, benzer örneklerin aynı sınıfa ait olma eğiliminde olduğunu varsaymaktadır.

Bir özellik vektörleri kümesi **X** ve karşılık gelen etiketler *y* ile bir veri setine sınıflandırılması gereken yeni bir veri noktası $\mathbf{x}_{\text{new}}$ verildiğinde, KNN algoritması aşağıdaki adımları izlemektedir:

- İlk olarak $\mathbf{x}_{\text{new}}$ ile eğitim setindeki her veri noktası arasındaki mesafe ölçülmektedir. Yaygın mesafe ölçüleri, Öklid mesafesi veya Manhattan mesafesidir. Öklid mesafesi, Eşitlik 6'daki formülle hesaplanmaktadır (Cover ve Hart, 1967).

$$\text{Distance}(\mathrm{x}_{\text{new}}, \mathrm{x}_i) = \sqrt{\sum_{j=1}^{D} \left(\mathrm{x}_{\text{new}} - \mathrm{x}_{i,j}\right)^2} \tag{6}$$

- Daha sonra $\mathbf{x}_{\text{new}}$ ile en küçük mesafelere sahip k veri noktası seçilmektedir.
- En sonunda da k-en yakın komşu arasında en çok bulunan sınıf Eşitlik 7'deki formülle belirlenmektedir.

$$y_{new} = \arg\max_y \sum_{i=1}^{k} \Pi(y_i = y) \tag{7}$$

Eşitlik 7'de y_{new} yeni veri noktası için tahmin edilen sınıfı temsil etmektedir ve $\Pi(\cdot)$ gösterim fonksiyonudur. Mesafe ölçümünün seçimi ve k değeri, KNN algoritmasındaki kritik parametrelerdir. Daha küçük k değerleri modeli gürültüye daha duyarlı hale getirirken, daha büyük değerler aşırı basitleştirmeye yol açabilmektedir. Algoritmanın basitliği ve etkinliği, özellikle karar sınırlarının doğrusal olmayan ve karmaşık olduğu durumlarda algoritmayı değerli bir araç haline getirmektedir.

2.4. Gaussian Karışım Modelleri (Gaussian Mixture Models - GMM)

GMM, istatistiksel modelleme ve örüntü tanıma alanlarında kullanılan bir öğrenme modelidir. Bu modeller, bir veri setini bir veya daha fazla Gaussian dağılımının karışımı olarak temsil etme amacını taşımaktadır. Bu yöntem, veri setlerinin karmaşıklığını ele almak ve içindeki gizli yapıları ortaya çıkarmak için son derece etkilidir. GMM'nin temelinde, bir karışım modeli kavramı yatar ve

bu, birden fazla Gaussian dağılımının ağırlıklı toplamı olarak ifade edilmektedir. Her bir Gaussian dağılımı, veri setindeki belirgin bir bileşeni kapsar ve genel temsile katkıda bulunmaktadır (Singh vd., 2010). Bu yaklaşım, GMM'nin verideki karmaşık yapıları ve çeşitli örüntüleri yakalamasını sağlamaktadır. GMM'nin çalışma prosedürü ve matematiksel formülleri aşağıda verilmiştir:

- GMM matematiksel olarak şu şekilde ifade edilir:

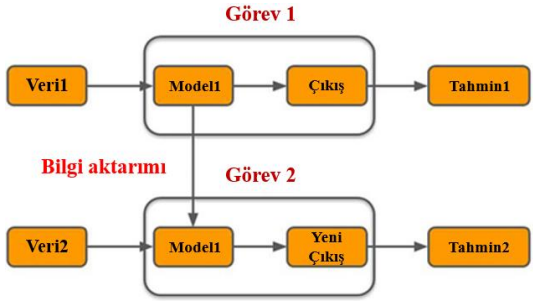
$$\mathrm{p(x)} = \sum_{i=1}^{k} \pi_k \, N(\mathrm{x}|\mu_k \Sigma_k) \qquad (8)$$

  - Eşitlik 8 incelendiğinde, π_k k-ıncı bileşenin ağırlığını temsil eder ve $\sum_{i=1}^{k} \pi_k = 1$ koşulunu sağlar.
  - $N(\mathrm{x}|\mu_k \Sigma_k)$ x'in k-ıncı Gaussian dağılımı tarafından üretilme olasılığını ifade eder.
  - μ_k ve Σ_k k-ıncı Gaussian dağılımının ortalama vektörü ve kovaryans matrisidir.

- Gizli değişkenlerin kullanılması, her veri noktasının hangi Gaussian bileşenine ait olduğunu belirlemeyi sağlamaktadır. Bu değişkenler, her bir veri noktasının, belirli bir bileşene olan üyeliğini temsil etmektedir. Böylece, modelin verinin altında yatan yapıları daha iyi anlamasına ve yakalamasına yardımcı olmaktadır.

3. Materyal ve Yöntem

Araştırmada, balık sınıflandırması için oluşturulan derin öğrenme modelinde aktarımlı öğrenme yaklaşımı kullanılmıştır. Aktarımlı öğrenme, derin öğrenme modellerinin önceden eğitildiği bir görevde edindiği bilgileri başka bir görevde kullanmak için uyarlamayı içermektedir. Bu yaklaşım, genellikle büyük ve çeşitli veri setleri üzerinde eğitilen modellerin, daha özel veya sınırlı veri setleri içeren görevlere etkin bir şekilde uygulanmasını sağlamaktadır (Tan vd., 2018). Aktarımlı öğrenme, modelin önceki bir görevde öğrendiği genel özellikleri ve desenleri, yeni bir görev için gereken özel özelliklere uyarlamak amacıyla kullanılır. Örneğin, genel nesne tanıma görevleri için büyük bir veri seti üzerinde eğitilen bir görüntü sınıflandırma modeli, kenarlar, renk gradyanları ve temel nesne kavramları gibi öğrendiği genel özellikleri başka bir görevde kullanılabilmektedir. Aktarımlı öğrenme, mevcut bilgiyi yeni görevin özel gereksinimlerine uyarlayarak öğrenmenin daha hızlı gerçekleştirilmesini sağlamaktadır. Bu sayede genel özellikleri yeniden öğrenmek yerine, bu özellikler üzerinde ince ayar yapılmaktadır (Taylor ve Stone, 2009). Aktarımlı öğrenmenin çalışma prensibi Şekil 1'de gösterilmektedir.

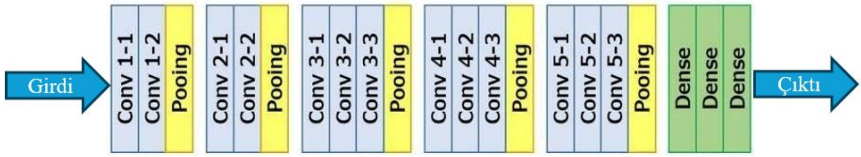


Şekil 1. Aktarımlı öğrenme şeması

Bu çalışmada balık sınıflandırması için VGG16, ResNet50, InceptionV3 ve MobileNet adlı önceden eğitilmiş CNN modelleri kullanılmıştır.

VGG16

VGG16 geniş ve çeşitli bir veri seti üzerinde eğitildiği için görsel sınıflandırma görevlerine uyarlanması için uygun bir temel sunmaktadır (Thenmozhi ve Reddy, 2019). Aktarımlı öğrenme sayesinde, VGG16'nın ImageNet veri seti üzerindeki genel nesne tanıma gibi öğrendiği genel özellikler, balık sınıflandırma gibi daha özel bir göreve uyarlanabilmektedir. Bu yaklaşım, özellikle sınırlı örneğe sahip veri setleriyle çalışırken faydalı görülmektedir. Önceden öğrenilmiş bilgilerin yeni bir göreve aktarılması, modelin daha hızlı ve etkili bir şekilde öğrenmesini sağlamaktadır. VGG16 modelinin mimarisi Şekil 2'de gösterilmektedir.



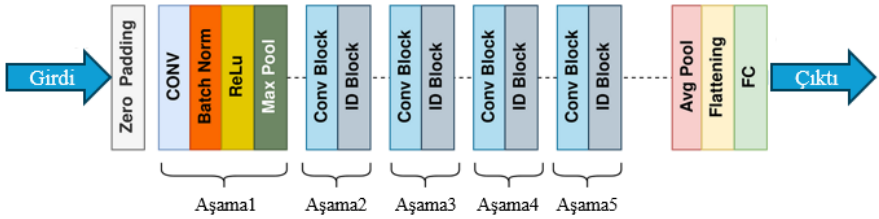
Şekil 2. VGG16 mimari yapısı

Araştırmada kullanılan VGG16 modeli, toplamda 21 katmandan oluşan bir yapıya sahip olup, bu katmanların 16'sı öğrenilebilir parametreler, yani ağırlık taşıyan katmanlardır. Bu 16 katman, 13 evrişim (convolutional) katmanı, 5 maksimum havuzlama katmanı ve 3 tam bağlantılı (fully connected - FC) katmandan meydana gelmektedir (Theckedath ve Sedamkar, 2020). Model, 224x224 piksel boyutunda ve 3 RGB kanallı girdiler üzerinde çalışarak görsel sınıflandırma için optimize edilmiştir. VGG16'nın öne çıkan özelliklerinden biri, hiper parametrelerin sadeleştirilmiş olmasıdır. Evrişim katmanlarında 3x3 boyutunda filtreler kullanılırken, her adımda (stride) 1 birimlik hareket ve sabit dolgu (padding) tercih edilmiştir. Maksimum havuzlama katmanlarında ise 2x2 filtreler ve 2 birimlik adım kullanılmıştır. Bu yapı, tüm mimari boyunca tutarlı

bir filtre düzeni sağlamaktadır. Conv-1 katmanında 64 filtre, Conv-2'de 128 filtre, Conv-3'te 256 filtre, Conv-4 ve Conv-5 katmanlarında ise 512 filtre yer almaktadır. Modelin tam bağlantılı katmanları ise sırasıyla 4096 kanala sahip iki katman ve 1000 sınıfı temsil eden son bir katmandan oluşmaktadır. Bu son katman, sınıflandırmayı gerçekleştiren ve sınıf olasılıklarını hesaplayan bir softmax katmanıdır. VGG16'nın, 3x3 filtreler kullanan ve düzenli evrişim/havuzlama katmanlı yapısı, görsel tanıma görevlerinde etkili olmasını sağlamıştır (Qassim vd., 2018). Bu model, sınırlı veri setleriyle dahi yüksek doğruluk oranlarına ulaşabilen bir performans sergileyerek derin öğrenme modellerinin gelişimine önemli katkılar sunmuştur (Sowmya vd., 2023).

ResNet50

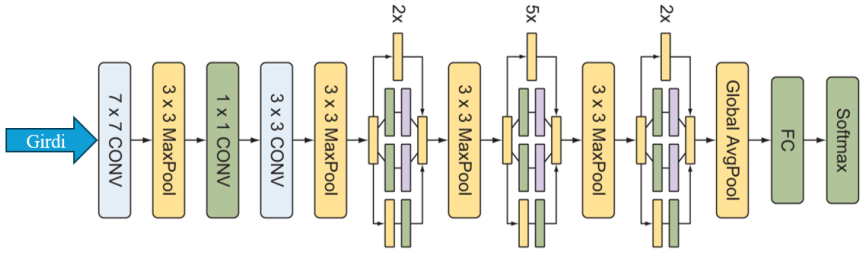
Araştırmada kullanılan diğer bir aktarımlı öğrenme modeli olan ResNet50 modelinin mimarisi Şekil 3'te gösterilmektedir. ResNet50 çok aşamalı bir mimariye sahiptir ve her aşama girdiyi anlamlandırmak için farklı işlemlerden geçmektedir. Modelin ilk aşamasında, girdi verisi 224x224 boyutunda ve 3 RGB kanallı olacak şekilde sıfır dolgu (zero padding) ile başlamaktadır. Ardından, 64 filtre ile çalışan bir evrişim (CONV) katmanı üzerinden geçirilmektedir. Bu evrişim işlemi sonrasında, toplu normalleştirme (batch normalization - BN) uygulanmakta ve doğrultulmuş doğrusal birim (Rectified Linear Unit - ReLU) aktivasyon fonksiyonu ile doğrusal olmayan özellikler modele kazandırılmaktadır. Bu aşama, bir maksimum havuzlama katmanı ile sonlandırılır ve bu sayede girdinin daha düşük boyutlu ve önemli özellikleri çıkarılarak özetlenir (Deshpande vd., 2021). İkinci aşamadan beşinci aşamaya kadar model, derin katmanlardan oluşan evrişim blokları ile özellik çıkarmaya devam eder. İkinci aşamada, evrişim katmanları 128 filtre, üçüncü aşamada 256 filtre, dördüncü ve beşinci aşamalarda ise 512 filtre kullanarak daha karmaşık ve soyut özellikleri öğrenir. Bu aşamalarda her evrişim katmanı, daha derin katmanlar sayesinde veri içindeki ince detayları yakalayıp sınıflandırma için anlamlı hale getirir. Modelin son bölümünde, ortalama havuzlama (average pooling) işlemi gerçekleştirilir ve ardından veriler düzleştirilerek (flattening), 4096 kanala sahip iki FC katmana aktarılır. Bu aşamadan sonra, son tam bağlantılı katmanda 1000 sınıfı temsil eden bir sınıflandırma işlemi yapılır. Çıktı katmanında, sınıf olasılıkları softmax fonksiyonu ile hesaplanır.



Şekil 3. ResNet50 mimari yapısı

InceptionV3

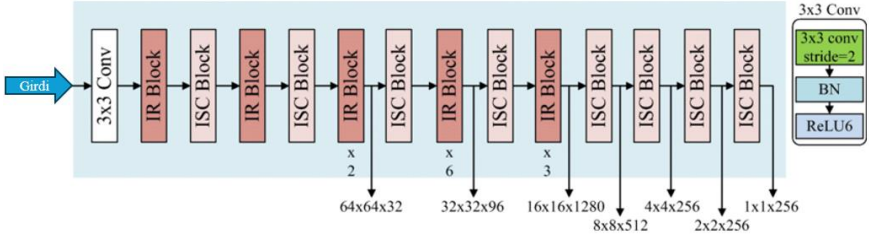
Çalışmada kullanılan bir diğer aktarımlı öğrenme modeli olan InceptionV3 modelinin mimarisi Şekil 4'te verilmiştir. InceptionV3, çok katmanlı ve karmaşık bir mimariye sahip olup, her aşamada farklı işlemlerle girdiyi anlamlandırır. Modelin başlangıç aşaması, 7x7 boyutunda filtrelere sahip bir evrişim katmanı ile başlamaktadır. Ardından, 3x3 boyutunda filtre kullanan bir maksimum havuzlama katmanı ile devam etmektedir. Bu işlemler, girdinin boyutunu azaltarak önemli özellikleri daha kompakt hale getirir (Wang vd., 2019). Modelin sonraki aşamalarında, farklı filtre boyutlarına sahip evrişim katmanları bulunmaktadır. 1x1 ve 3x3 boyutlarında evrişim filtreleri ardışık olarak kullanılarak, veri içerisindeki farklı ölçeklerdeki özellikler yakalanmaktadır. Her evrişim aşaması sonrasında, yeniden 3x3 maksimum havuzlama katmanı kullanılarak veri boyutları azaltılmakta ve en önemli özellikler elde edilmektedir. Modelin daha derin katmanları, Inception blokları ile devam etmektedir. Bu bloklar, farklı filtre boyutlarını aynı anda kullanarak daha geniş bir özellik çıkarımı yapmaktadır. Örneğin, belirli aşamalarda 2 ve 5 kez tekrarlanan Inception blokları, verideki karmaşık yapıları yakalayarak daha yüksek seviyeli özellikleri öğrenmektedir. Bu, modelin farklı seviyelerde soyutlamalar yapmasına olanak tanımaktadır. Son aşamada, global ortalama havuzlama (Global AvgPool) katmanı ile veriler düzleştirilir. Ardından, tam bağlantılı katmana aktarılır ve sınıflandırma işlemi gerçekleştirilir. Çıktı katmanında ise softmax fonksiyonu ile sınıf olasılıkları hesaplanır. Bu aşamalı ve çok yönlü yapı, InceptionV3 modelinin hem düşük hem de yüksek seviyeli özellikleri öğrenmesine olanak tanıyarak başarılı bir sınıflandırma performansı sergilemesini sağlamaktadır. Bu mimari, VGG16'daki gibi sabit filtre boyutlarına dayanan bir yapının aksine, farklı ölçeklerde özellik çıkarmak için esnek filtre boyutları kullanarak daha karmaşık yapıları analiz edebilmektedir.



Şekil 4. InceptionV3 mimari yapısı

MobileNetV2

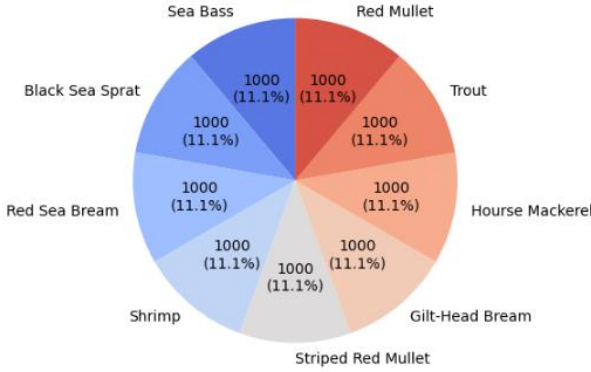
MobileNetV2 araştırmada kullanılan diğer bir derin öğrenme modelidir. Mimarisi Şekil 5'te verilen bu model hafif ve verimli bir mimari ile yapılandırılmıştır. Modelin başlangıcında, girdi verisi 3x3 boyutunda filtrelere sahip bir evrişim katmanı üzerinden geçirilmektedir. Bu işlem sonrasında model, genişletme ve sıkıştırma işlemlerini gerçekleştiren daraltma blokları (IR Block) ve sıkıştırma blokları (ISC Block) ile devam etmektedir. MobileNetV2, bu bloklar sayesinde özellikle düşük hesaplama gücü gerektiren mobil cihazlar için optimize edilmiş, hızlı ve etkili bir yapı sunmaktadır (Singh vd., 2019). İlk evrişim bloğu 64x64x32 boyutunda özellik haritaları üretmekte ve ardından iki kez tekrarlanan IR blokları kullanılarak veri içindeki özellikler çıkarılmaktadır. Daha sonra, 32x32x96 ve 16x16x1280 boyutlarında daha derin bloklara geçiş yapılır. Bu aşamalar sırasında ISC blokları, veriyi sıkıştırarak hem boyutlarını küçültür hem de daha derin özellikleri çıkarır. Model, her bir IR ve ISC bloğunun ardından girdi verisini daha fazla işlemden geçirerek çok katmanlı bir özellik çıkartımı gerçekleştirir. Özellikle, 2, 6 ve 3 kez tekrarlanan bloklar ile model, giderek karmaşıklaşan özellikleri öğrenmektedir. Bu yapı, VGG16'daki sabit boyutlu filtrelerden farklı olarak, daha esnek ve katmanlar arasında veri akışını optimize eden bir yapı sağlamaktadır. Son aşamada, 1x1 ve 3x3 boyutlarındaki evrişim işlemleri uygulanarak veri, boyut olarak son haline getirilir. Modelde toplu normalizasyon ve doğrusal olmayan özelliklerin öğrenilmesini sağlayan ReLU6 aktivasyon fonksiyonu da kullanılmaktadır. Model, çıktıyı son olarak FC katmanına aktararak sınıflandırma işlemini tamamlar.



Şekil 5. MobileNetV2 mimari yapısı

3.1. Veri Seti

Araştırmada kullanılan veri seti, İzmir Ekonomi Üniversitesi'nde yürütülen bir üniversite-sanayi iş birliği projesi kapsamında İzmir'deki bir süpermarketten elde edilen dokuz farklı balık türünü içermektedir (Ulucan vd., 2020). Veri seti, her bir sınıf için 1000 artırılmış görüntü ve bunlara karşılık gelen çiftler halinde artırılmış yer doğruluklarını içermektedir. Veri setinde bulunan balık türleri çipura (gilt head bream), mercan (red sea bream), levrek (sea bass), barbun (red mullet), istavrit (horse mackerel), hamsi (black sea sprat), tekir (striped red mullet), alabalık (trout) ve karides (shrimp) olarak sıralanmıştır. Veri seti dengeli bir yapıya sahip olup, her bir sınıf eşit sayıda görüntü içermektedir (her sınıf için 1000 görüntü). Şekil 6, veri setindeki sınıfların dağılımını göstermekte, Şekil 7 ise veri setinden bazı örnek görüntüleri sunmaktadır.



Şekil 6. Veri setindeki sınıfların dağılımı

Şekil 7. Veri setinden örnek görüntüler

3.2. Balık Sınıflandırması için Önerilen Model

Bu araştırmada, balık sınıflandırma görevine yönelik olarak aktarımlı öğrenme yaklaşımı kullanılmış ve VGG16, ResNet50, InceptionV3 ile MobileNet modelleri uyarlanmıştır. Söz konusu modeller, büyük görüntü veri kümeleri üzerinde önceden eğitildikleri için genel nesne tanıma yeteneklerine sahiptir. Bu araştırmada, bu yetenekler özel bir görev olan balık sınıflandırması için işe koşulmuştur. İlk olarak, veri seti eğitim ve test setlerine ayrılmıştır. Verilerin %80'i (7200 görüntü) eğitim seti, %20'si (1800 görüntü) ise test seti olarak belirlenmiştir. Veri seti, farklı balık türlerine ait çeşitli açılardan çekilmiş yüksek çözünürlüklü görüntülerden oluşmaktadır. Modellerin eğitimi, TensorFlow gibi derin öğrenme kütüphaneleri kullanılarak gerçekleştirilmiştir. VGG16, ResNet50, InceptionV3 ve MobileNet modellerinin önceden eğitilmiş ağırlıkları ImageNet veri seti üzerinden geri yüklenmiş ve balık sınıflandırma görevine uyarlanmıştır. Bu uyarlama sırasında, her bir modelin çıktı katmanı balık sınıflarına uygun olacak şekilde, 10 çıkışlı softmax katmanı ile değiştirilmiştir. Modellerin genel özellik çıkarma yeteneklerini koruyabilmek için önceden eğitilmiş ağırlıkları dondurulmuş, yalnızca yeni eklenen sınıflandırma katmanlarının ağırlıkları güncellenmiştir. Modellerin eğitim sürecinde kullanılan parametreler Tablo 1'de verilmiştir.

Tablo 1. Derin öğrenme modellerinin performans metriklerinin karşılaştırması

Eğitim Parametresi	**Değer**
Optimizatör	Adam
Öğrenme hızı	0.001
Tekrar sayısı	50
Yığın boyutu	32

Modellerin performansı test seti üzerinde değerlendirilmiş ve doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), özgüllük (specificity) ile F1 skor metrikleri kullanılmıştır. Formülleri sırasıyla Eşitlik 9, Eşitlik 10, Eşitlik 11, Eşitlik 12 ve Eşitlik 13'te verilmiştir. TP, TN, FP ve FN sırasıyla doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif tahminleri temsil etmektedir.

$$\text{Doğruluk} = \frac{\text{Doğru Tahminler}}{\text{Toplam Örnek Sayısı}} \tag{9}$$

$$\text{Kesinlik} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

$$\text{Duyarlılık} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

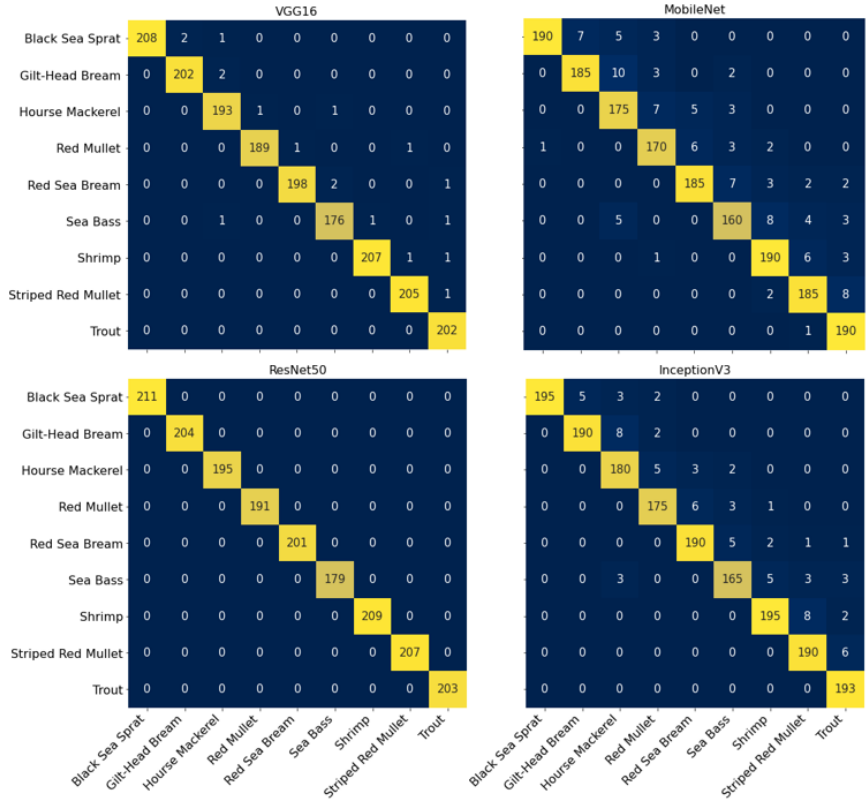
$$\text{Özgüllük} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{12}$$

$$\text{F1} = 2 * \frac{\text{kesinlik} * \text{duyarlılık}}{\text{kesinlik} + \text{duyarlılık}} \tag{13}$$

4. Deneysel Çalışma ve Bulgular

Modellerin eğitiminde TensorFlow kütüphanesi kullanılmıştır. Eğitim süreci boyunca, her bir tekrar sayısı sonunda eğitim veri setindeki kayıp fonksiyonu ve doğruluk değerleri dikkatle izlenmiştir. Benzer şekilde, doğrulama veri seti üzerindeki kayıp ve doğruluk değerleri de titizlikle takip edilmiştir. Bu izleme işlemi, modelin öğrenme kapasitesini değerlendirmek ve olası aşırı öğrenme (overfitting) durumlarını tespit etmek amacıyla yapılmıştır. Aşırı öğrenme, bir makine öğrenmesi modelinin eğitim verisine gereğinden fazla uyum sağlaması durumunda ortaya çıkmakta ve modelin genelleme yeteneğini zayıflatmaktadır (Hawkins, 2004). Bu durumda, model yalnızca eğitim verisine özgü detayları öğrenmekte, ancak yeni ve farklı veri kümelerindeki kalıpları tanımakta zorlanmaktadır (Bilbao ve Bilbao, 2017). Bu yüzden model yeni verilere karşı yeterince esnek olmamakta ve performansı düşebilmektedir (Li vd., 2020). Eğitim sürecinde, gelişme gözlenmediği durumlarda eğitimi otomatik olarak durdurmak için EarlyStopping geri çağırma (callback) fonksiyonu kullanılmıştır. Bu sayede, modelin en iyi performansı sergilediği noktada eğitim süreci sonlandırılmıştır. Ayrıca, çalışmada Weights & Biases (WandB) platformu kullanılarak, modelin performansının interaktif olarak izlenmesi ve parametrelerin kaydedilmesi sağlanmıştır.

Şekil 8'de sunulan karmaşıklık matrisleri, her bir modelin balık türlerine yönelik doğru ve yanlış sınıflandırma sonuçlarını göstermektedir. ResNet50 modeli tüm balık türlerini doğru sınıflandırarak en yüksek performansı sergilemektedir. VGG16 modeli de iyi sonuçlar vermiş, özellikle Hamsi (208) ve Karides (207) türlerinde başarılı olmuştur. MobileNet, Çipura (185) ve Barbun (190) türlerinde doğru sınıflandırmalar yapmış olsa da diğer modeller arasında performansı en düşük model olduğu anlaşılmaktadır.



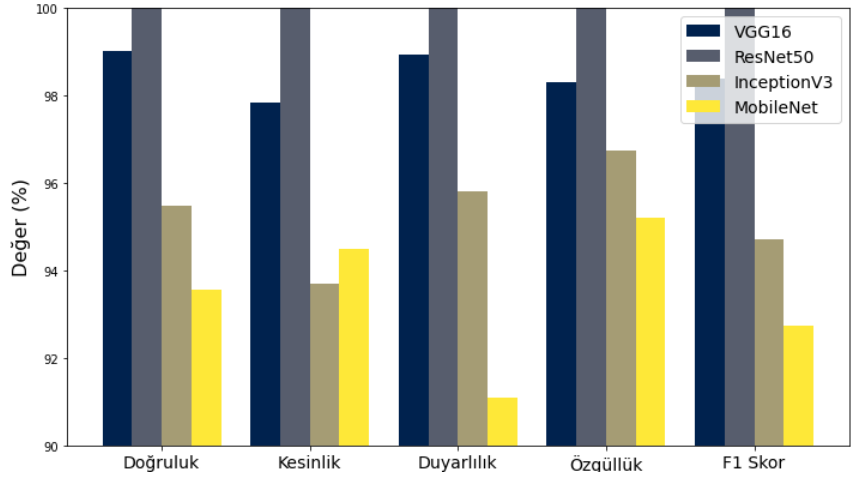
Şekil 8. Karmaşıklık matrisleri

Tablo 2'deki metrik sonuçları incelendiğinde, ResNet50 modelinin %100 doğruluk ile tüm derin öğrenme modelleri arasında en iyi sonucu sağladığı görülmektedir. Bu model, doğruluk, kesinlik, duyarlılık, özgüllük ve F1 skoru gibi tüm metriklerde tam performans göstererek diğer modellerden açık bir şekilde üstün olmuştur. VGG16 modeli ise %99.02 doğruluk ile ikinci sırada yer almaktadır. InceptionV3, %95.49 doğruluk ile üçüncü sırada gelmekte ve daha düşük bir performans sergilemektedir. Kesinlik değeri %93.70, duyarlılığı %95.80 ve özgüllük değeri %96.75 olan bu model, VGG16'ya göre %3.53 oranında daha düşük doğruluk sunmaktadır. F1 skoru ise %94.72 ile orta seviyelerde kalmıştır. MobileNet, %93.57 doğruluk ile en düşük performansa sahip modeldir. Kesinlik değeri %94.50 olmasına rağmen, duyarlılığı %91.10 ile en düşük değere sahiptir. Bu durum, MobileNet'in bazı sınıfları yeterince iyi tanımlayamadığını göstermektedir. Ancak, özgüllük açısından %95.20 ile görece dengeli bir sonuç elde etmiştir. MobileNet, ResNet50'ye kıyasla %6.43, VGG16'ya kıyasla %5.45 ve InceptionV3'e kıyasla %1.92 oranında daha düşük doğruluk değeri sergilemiştir.

Tablo 2. Derin öğrenme modellerinin performans metriklerinin karşılaştırması

Model	Doğruluk	Kesinlik	Duyarlılık	Özgüllük	F1 Skor
VGG16	99.02	97.85	98.95	98.30	98.40
ResNet50	100	100	100	100	100
InceptionV3	95.49	93.70	95.80	96.75	94.72
MobileNet	93.57	94.50	91.10	95.20	92.75

Şekil 9'daki doğruluk değerleri incelendiğinde, ResNet50 modeli %100 doğruluk oranı ile en iyi performansı göstermektedir. VGG16 ve InceptionV3 ise sırasıyla %99.02 ve %95.49 oranlarıyla onu takip etmektedir. MobileNet, %93.57 doğruluk ile diğer modellere kıyasla daha düşük bir performans göstermektedir. Kesinlik açısından da ResNet50 %100'lük kesinlik ile en üst sırada yer alırken, InceptionV3 ve MobileNet %93.70 ve %94.50 oranları ile geride kalmaktadır. Duyarlılık metrikleri değerlendirildiğinde, yine ResNet50 %100 ile en yüksek başarıyı elde etmektedir. VGG16, InceptionV3 ve MobileNet, %98.95, %95.80 ve %91.10 gibi daha düşük duyarlılık oranlarına sahiptir. Özgüllük açısından ResNet50 %100 ve VGG16 %98.30 değerlerine ulaşarak en yüksek performansları gösterirlerken, InceptionV3 ve MobileNet sırasıyla %96.75 ve %95.20 ile bu iki modelin gerisindedir. MobilNet'in özgüllüğü, diğer modellere kıyasla yaklaşık %1.39-4.29 daha düşük bir seviyede yer almaktadır. F1 skoru değerlendirmesinde, ResNet50 %100 ile en yüksek puanı alırken, VGG16 %98.40 ile onu takip etmektedir. InceptionV3 ve MobileNet ise %94.72 ve %92.75 gibi daha düşük F1 skorları elde etmiştir.



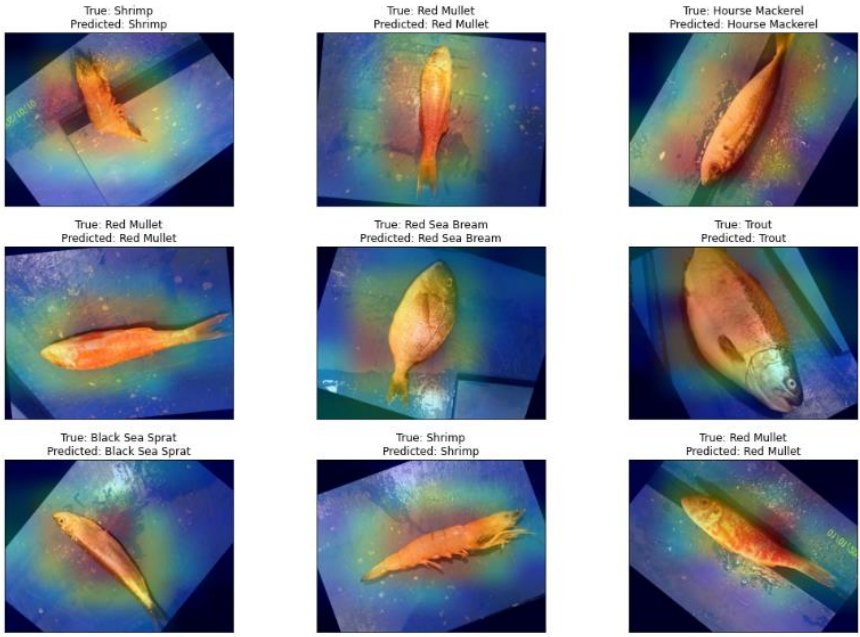
Şekil 9. Modellerin performans sonuçlarının karşılaştırması

Şekil 10'da test veri setinden rastgele seçilmiş 15 balık görselinin ResNet50 modeli ile sınıflandırma sonuçları yer almaktadır. Seçilen bütün balık

örneklerinin doğru sınıflandırıldığı görülmektedir. Balık sınıflandırma sürecinin son aşamasında, modelin iç karar verme süreçlerine dair çıkarımlar elde etmek için Grad-CAM (Gradient-weighted Class Activation Mapping) yöntemi kullanılmıştır. Bu yöntem, CNN'in son evrişim katmanına akan gradyan bilgilerini kullanarak her bir nöronun sınıflandırma kararına katkısını belirlemektedir (Selvaraju vd., 2017). Grad-CAM yöntemi ile oluşturulan ısı haritaları, derin sinir ağının dikkatini yoğunlaştırdığı kritik bölgeleri görsel olarak vurgulamaktadır. Şekil 11, Grad-CAM yöntemiyle üretilen ısı haritasını sunmaktadır. Isı haritası, modelin balık sınıflandırma görevinde hangi özelliklere öncelik verdiğini ve kararını desteklemek için hangi görüntü bölgelerini kullandığını ortaya koyan bir görsel açıklama işlevi görmektedir.

Şekil 10. Tahmin örnekleri

Şekil 11, Grad-CAM yöntemiyle üretilen ısı haritasını sunmaktadır. Isı haritası, modelin balık sınıflandırma görevinde hangi özelliklere öncelik verdiğini ve kararını desteklemek için hangi görüntü bölgelerini kullandığını ortaya koyan bir görsel açıklama işlevi görmektedir.



Şekil 11. Doğru tahmin edilen görüntülerin ısı haritası

Isı haritası, modelin her balık türü için en önemli ayırt edici özellikleri tanımlama sürecinde dikkatini yönlendirdiği alanları vurgulamaktadır. Isı haritasında artan yoğunluğa sahip bölgeler, CNN'in sınıflandırma kararlarını önemli ölçüde etkileyen özellikleri temsil etmekte ve belirli balık türlerine özgü karakteristikleri vurgulamaktadır. Gölgelendirilmiş kısımlar, sınıflandırma görevinde CNN'in seçimlerinin temelini oluşturmakta olup, modelin öncelik verdiği dikkate değer ayrıntıları ortaya koymaktadır.

5. Sonuçlar

Bu araştırmada, VGG16, ResNet50, InceptionV3 ve MobileNet modellerinin aktarımlı öğrenme yaklaşımıyla balık sınıflandırmasındaki kullanımı kapsamlı bir şekilde değerlendirilmiştir. Çalışmada kullanılan veri seti, her biri için 1000 artırılmış görüntü ile dokuz farklı deniz canlısını içermektedir. Önceden eğitilmiş ağırlıklara sahip VGG16, ResNet50, InceptionV3 ve MobileNet modelleri, balık sınıflandırma görevine özel gereksinimlere uyacak şekilde özelleştirilmiştir. Deneysel çalışmadan elde edilen sonuçlar, modellerin etkili bir şekilde öğrendiğini, kayıp değerlerinin azaldığını ve zamanla doğruluk oranlarının arttığını göstermiştir. Bu bağlamda performans metrikleri, yüksek doğruluk oranları ile çeşitli balık türlerini sınıflandırmada modellerin yüksek performans sergilediğini ortaya koymuştur. En yüksek doğruluğun ise %100 başarı ile ResNet50 modeli tarafından elde edildiği gözlemlenmiştir. En düşük performans

değerlerine sahip model ise MobileNet olmuştur. Grad-CAM yöntemi, balık sınıflandırması için kritik bölgeleri vurgulayarak modelin iç karar verme süreçlerine dair görsel çıkarımlar sunmuştur. Üretilen ısı haritası, modelin dikkatini nerelere yönlendirdiğini, belirli balık türlerine özgü özelliklere odaklandığını ve sınıflandırma kararını desteklemek için hangi görüntü bölgelerini kullandığını görsel olarak açıklamaktadır.

Tablo 3, literatürdeki balık sınıflandırma çalışmalarına ait modelleri ve doğruluk oranlarını karşılaştırmaktadır. Tablo incelendiğinde, bu çalışmada önerilen ResNet50 modelinin alandaki diğer çalışmalardan önemli ölçüde farklı olduğu ve %100'lük en yüksek sınıflandırma doğruluğuna ulaştığı anlaşılmaktadır. Ayrıca, bu çalışmanın diğer araştırmalarla farklılık gösterdiği bir diğer nokta; çalışma kapsamındaki balık türü sayısının oldukça yüksek (9 tür) olmasıdır. Bu durum, bu çalışmanın daha geniş bir balık türü yelpazesini başarılı bir şekilde sınıflandırma yeteneğini göstermektedir.

Tablo 3. Literatürdeki balık sınıflandırma çalışmalarının karşılaştırması

Kaynak	**Yıl**	**Model**	**Sınıf Büyüklüğü**	**Doğruluk**
Malik vd.	2023	FD_Net	9	%95.30
Kaya vd.	2023	CNN	5	%90.39
Ren vd.	2023	CNN ve SVM	4	%92.43
Knausgård vd.	2022	YOLOv3 + CNN	5	%87.40
Iqbal vd.	2022	CNN	2	%88.00
Li vd.	2022	AlexNet	2	%90.00
Tarling vd.	2022	FFRNET	4	%90.79
Yeh vd.	2021	CNN	5	%91.29
Ju vd.	2020	AlexNet	4	%89.78
Bu çalışma	***2024***	***ResNet50***	***9***	***%100***

Bu araştırma, özellikle ResNet50 modeli ile aktarımlı öğrenme yaklaşımının balık sınıflandırmasında etkili bir şekilde kullanılabileceğini göstermiştir. Çalışma, makine öğrenmesi tekniklerinin deniz biyolojisi, balıkçılık yönetimi ve çevresel koruma alanlarında kullanılması ve su ekosistemleri hakkındaki anlayışımızı derinleştirmesi açısından önemli bir katkı sunmaktadır.

Kaynakça

1. Aguado, S. H., Segado, I. S., & Pitcher, T. J. (2016). Towards sustainable fisheries: A multi-criteria participatory approach to assessing indicators of sustainable fishing communities: A case study from Cartagena (Spain). *Marine Policy*, *65*, 97-106. https://doi.org/10.1016/j.marpol.2015.12.024

2. Alaba, S. Y., Nabi, M. M., Shah, C., Prior, J., Campbell, M. D., Wallace, F., & Moorhead, R. (2022). Class-aware fish species recognition using deep learning for an imbalanced dataset. *Sensors*, *22*(21), 8268. https://doi.org/10.3390/s22218268

3. Ali-Gombe, A., Elyan, E., & Jayne, C. (2017). Fish classification in context of noisy images. In *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings* (pp. 216-226). Springer International Publishing. https://doi.org/10.1007/978-3-319-65172-9_19

4. Alsmadi, M. K., & Almarashdeh, I. (2022). A survey on fish classification techniques. *Journal of King Saud University-Computer and Information Sciences*, *34*(5), 1625-1638. https://doi.org/10.1016/j.jksuci.2020.07.005

5. Arthington, A. H., Dulvy, N. K., Gladstone, W., & Winfield, I. J. (2016). Fish conservation in freshwater and marine realms: status, threats and management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *26*(5), 838-857. https://doi.org/10.1002/aqc.2712

6. Batista, G. E., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, *6*(1), 20-29. https://doi.org/10.1145/1007730.1007735

7. Bilbao, I., & Bilbao, J. (2017, December). Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks. In *2017 eighth international conference on intelligent computing and information systems (ICICIS)* (pp. 173-177). IEEE. https://doi.org/10.1109/INTELCIS.2017.8260032

8. Cabreira, A. G., Tripode, M., & Madirolas, A. (2009). Artificial neural networks for fish-species identification. *ICES Journal of Marine Science*, *66*(6), 1119-1129. https://doi.org/10.1093/icesjms/fsp009

9. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21-27. https://doi.org/10.1109/TIT.1967.1053964

10. Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.

11. Deshpande, A., Estrela, V. V., & Patavardhan, P. (2021). The DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50. *Neuroscience Informatics*, *1*(4), 100013. https://doi.org/10.1016/j.neuri.2021.100013

12. Feldmann, J., Youngblood, N., Karpov, M., Gehring, H., Li, X., Stappers, M., ... & Bhaskaran, H. (2021). Parallel convolutional processing using an integrated photonic tensor core. *Nature*, *589*(7840), 52-58. https://doi.org/10.1038/s41586-020-03070-1

13. Fischer, J. (2014). Fish identification tools for biodiversity and fisheries assessments: Review and guidance for decision-makers. *FAO Fisheries and Aquaculture Technical Paper*, (585), 11-23. https://www.proquest.com/openview/b2ce3c19750845576cc31b72058d09f7/1?pq-origsite=gscholar&cbl=237320

14. Garcia, R., Prados, R., Quintana, J., Tempelaar, A., Gracias, N., Rosen, S., ... & Løvall, K. (2020). Automatic segmentation of fish using deep learning with application to fish size measurement. *ICES Journal of Marine Science*, *77*(4), 1354-1366. https://doi.org/10.1093/icesjms/fsz186

15. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, *51*(5), 1-42. https://doi.org/10.1145/3236009

16. Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1-12. https://doi.org/10.1021/ci0342472

17. Iqbal, U., Li, D., & Akhter, M. (2022). Intelligent diagnosis of fish behavior using deep learning method. *Fishes*, *7*(4), 201. https://doi.org/10.3390/fishes7040201

18. Ju, Z., & Xue, Y. (2020). Fish species recognition using an improved AlexNet model. *Optik*, *223*, 165499. https://doi.org/10.1016/j.ijleo.2020.165499

19. Knausgård, K. M., Wiklund, A., Sørdalen, T. K., Halvorsen, K. T., Kleiven, A. R., Jiao, L., & Goodwin, M. (2022). Temperate fish detection and classification: a deep learning based approach. *Applied Intelligence*, *52*, 6988–7001. https://doi.org/10.1007/s10489-020-02154-9

20. Li, D., Wang, Q., Li, X., Niu, M., Wang, H., & Liu, C. (2022). Recent advances of machine vision technology in fish classification. *ICES Journal of Marine Science*, *79*(2), 263-284. https://doi.org/10.1093/icesjms/fsab264

21. Li, Z., Lyu, K., & Arora, S. (2020). Reconciling modern deep learning with traditional optimization analyses: The intrinsic learning rate. *Advances in Neural Information Processing Systems*, *33*, 14544-14555. https://proceedings.neurips.cc/paper_files/paper/2020/file/a7453a5f026fb6831d68bdc9cb0edcae-Paper.pdf

22. Malik, H., Naeem, A., Hassan, S., Ali, F., Naqvi, R. A., & Yon, D. K. (2023). Multi-classification deep neural networks for identification of fish species using camera captured images. *Plos One*, *18*(4), e0284992. https://doi.org/10.1371/journal.pone.0284992

23. Pitcher, T. J. (2008). The sea ahead: challenges to marine biology from seafood sustainability. *Hydrobiologia*, *606*, 161-185. https://doi.org/10.1007/s10750-008-9337-9

24. Potts, T., & Haward, M. (2007). International trade, eco-labelling, and sustainable fisheries–recent issues, concepts and practices. *Environment, Development and Sustainability*, *9*(1), 91-106. https://doi.org/10.1007/s10668-005-9006-3

25. Qassim, H., Verma, A., & Feinzimer, D. (2018, January). Compressed residual-VGG16 CNN model for big data places image recognition. In *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)* (pp. 169-175). IEEE. https://doi.org/10.1109/CCWC.2018.8301729

26. Ren, L., Tian, Y., Yang, X., Wang, Q., Wang, L., Geng, X., & Lin, H. (2023). Rapid identification of fish species by laser-induced breakdown spectroscopy and Raman spectroscopy coupled with machine learning methods. *Food Chemistry*, *400*, 134043. https://doi.org/10.1016/j.foodchem.2022.134043

27. Roberson, L. A., Watson, R. A., & Klein, C. J. (2020). Over 90 endangered fish and invertebrates are caught in industrial fisheries. *Nature Communications*, *11*(1), 4764. https://doi.org/10.1038/s41467-020-18505-6

28. Salman, A., Jalal, A., Shafait, F., Mian, A., Shortis, M., Seager, J., & Harvey, E. (2016). Fish species classification in unconstrained underwater environments based on deep learning. *Limnology and Oceanography: Methods*, *14*(9), 570-585. https://doi.org/10.1002/lom3.10113

29. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591-596. https://doi.org/10.1038/35098000

30. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision* (pp. 618-626). https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf

31. Singh, B., Toshniwal, D., & Allur, S. K. (2019). Shunt connection: An intelligent skipping of contiguous blocks for optimizing MobileNet-V2. *Neural Networks*, *118*, 192-203. https://doi.org/10.1016/j.neunet.2019.06.006

32. Singh, R., Pal, B. C., & Jabr, R. A. (2010). Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement. *IET generation, transmission & distribution*, *4*(1), 50-59. https://doi.org/10.1049/iet-gtd.2009.0167

33. Sowmya, M., Balasubramanian, M., & Vaidehi, K. (2023). Human Behavior Classification using 2D–Convolutional Neural Network, VGG16 and ResNet50. *Indian Journal of Science and Technology*, *16*(16), 1221-1229. https://doi.org/10.17485/IJST/v16i16.199

34. Stefanakis, A. I., & Becker, J. A. (2016). A review of emerging contaminants in water: Classification, sources, and potential risks. *Impact of Water Pollution on Human Health and Environmental Sustainability*, 55-80. https://doi.org/10.4018/978-1-4666-9559-7.ch003

35. Sun, M., Song, Z., Jiang, X., Pan, J., & Pang, Y. (2017). Learning pooling for convolutional neural network. *Neurocomputing*, *224*, 96-104. https://doi.org/10.1016/j.neucom.2016.10.049

36. Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. In *Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27* (pp. 270-279). Springer International Publishing. https://doi.org/10.1007/978-3-030-01424-7_27

37. Tarling, P., Cantor, M., Clapés, A., & Escalera, S. (2022). Deep learning with self-supervision and uncertainty regularization to count fish in underwater images. *PloS One*, *17*(5), e0267759. https://doi.org/10.1371/journal.pone.0267759

38. Taylor, M. E., & Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, *10*(7), 1633-1685. https://www.jmlr.org/papers/volume10/taylor09a/taylor09a.pdf?ref=https://codemonkey.link

39. Theckedath, D., & Sedamkar, R. R. (2020). Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science*, *1*, 1-7. https://doi.org/10.1007/s42979-020-0114-9

40. Thenmozhi, K., & Reddy, U. S. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, *164*(104906), 1-11. https://doi.org/10.1016/j.compag.2019.104906

41. Ulucan, O., Karakaya, D., & Turkan, M. (2020, October). A large-scale dataset for fish segmentation and classification. In *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-5). IEEE. https://doi.org/10.1109/ASYU50717.2020.9259867

42. Wang, C., Chen, D., Hao, L., Liu, X., Zeng, Y., Chen, J., & Zhang, G. (2019). Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access*, *7*, 146533-146541. https://doi.org/10.1109/ACCESS.2019.2946000

43. Yeh, C. H., Lin, C. H., Kang, L. W., Huang, C. H., Lin, M. H., Chang, C. Y., & Wang, C. C. (2021). Lightweight deep neural network for joint learning of underwater object detection and color conversion. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(11), 6129-6143. https://doi.org/10.1109/TNNLS.2021.3072414

44. Zafar, A., Aamir, M., Mohd Nawi, N., Arshad, A., Riaz, S., Alruban, A., ... & Almotairi, S. (2022). A comparison of pooling methods for convolutional neural networks. *Applied Sciences*, *12*(17), 8643. https://doi.org/10.3390/app12178643

45. Zhang, H., Fritts, J. E., & Goldman, S. A. (2008). Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, *110*(2), 260-280. https://doi.org/10.1016/j.cviu.2007.08.003

46. Zhang, S., Li, X., Zong, M., Zhu, X., & Wang, R. (2017). Efficient kNN classification with different numbers of nearest neighbors. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(5), 1774-1785. https://doi.org/10.1109/TNNLS.2017.2673241

3. Bölüm

Web Madenciliği ve Veri Analitiği Tabanlı Gelişmiş Karar Destek Sistemlerinin İnşası

Web Mining and Data Analytics Based Building Advanced Decision Support Systems

Hüseyin PARMAKSIZ[1]

Önder ÖZTÜRK[2]

[1] Dr. Öğr. Üyesi, Bilecik Şeyh Edebali Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ORCID: 0000-0001-8455-5625, huseyin.parmaksiz@bilecik.edu.tr

[2] Öğretim Görevlisi, Kütahya Sağlık Bilimleri Üniversitesi, Rektörlük, Bilgi İşlem Daire Başkanlığı, ORCID: 0000-0001-6460-9497, onder.ozturk@ksbu.edu.tr

1. Giriş

Web madenciliği, internet üzerindeki dokümanlardan ve dijital hizmetlerden anlamlı bilgileri çıkarmak amacıyla veri madenciliği araçlarının kullanılmasını inceleyen bir süreçtir. Karar destek sistemleri (KDS), organizasyonlarda ve işletmelerde karar vericilere yardımcı olan bilgisayar destekli yapılar olarak tanımlanır. Veri analitiğinin iş zekâsı ile harmanlanması, yenilikçi ve ileri düzey destek sistemlerinin oluşturulmasını sağlar. Bu bölüm, web madenciliğinin web sitesi yönetimini iyileştiren gelişmiş sistemlerin inşasında nasıl stratejik bir araç olarak kullanılabileceğini vurgulamaktadır. Belirtilen adımlar izlenerek veri çıkarımı iş stratejileri ile birleştirildiğinde, gelişmiş karar destek sistemleri geliştirilebilir.

Web sitelerinin yönetimi, sürekli olarak yeni bilgi girişleri ve zamanında güncellemeler gerektirir. Site sahipleri, kullanıcılarına daha fazla hizmet ve içerik sunma isteğiyle bu güncellemelere ihtiyaç duyarlar. Kullanıcıların davranışlarındaki çeşitlilik ve karmaşıklık, bu taleplerin temel nedenlerindendir. Bu yoğun iş gücü gerektiren çaba, önemli ölçüde finansal ve personel maliyeti doğurur.

KDS, işletme ve organizasyonların karar alma süreçlerini destekleyen bilgisayar tabanlı bilgi sistemleridir. Veri madenciliği ile iş zekâsı tekniklerinin entegrasyonu, gelişmiş KDS'lerin oluşturulmasında kritik bir rol oynar. Veri madenciliğinde temel bir problem, büyük veri kümelerinde sık rastlanan kalıpları bulma sürecidir. Bu problem, özellikle sık fakat anlamlı kalıplar içeren veri kümeleriyle çalışırken daha da zorlaşır (Zou, Chu, Johnson, & Chiu, 2002). Sık kalıpların bulunması, veri madenciliği görevleri olan ilişkileri, korelasyonları ve sıralı kalıpları keşfetmeyi mümkün kılar. İş zekâsı ise verileri toplama, analiz etme ve raporlama süreçlerini kapsayan bir teknoloji olup, organizasyonların daha iyi kararlar almasına olanak tanır. Ayrıca, örgütsel öğrenme kültürü, üniversitelerin daha iyi örgütsel ve sürdürülebilir performans elde etmesine katkıda bulunur ve veri madenciliği ile iş zekâsının entegrasyonu, bu süreçleri destekleyebilir (Elbawab, 2024). Bu entegrasyon, web sitesi yönetimini kolaylaştırarak, yönetim için harcanan emeği azaltabilir.

Web madenciliği, veri madenciliği tekniklerinin Dünya Çapında Ağ (World Wide Web) belgelerinden ve hizmetlerinden otomatik olarak faydalı bilgileri keşfetmek ve çıkarmak için uygulanmasıdır (Aye, 2011). Web madenciliği genel olarak üç türe ayrılabilir: Web içerik madenciliği, Web yapı madenciliği ve Web kullanım madenciliği (Saini & Pandey, 2015). Web içerik madenciliği, web sayfası içeriklerinden değerli bilgileri keşfetmek için kullanılır ve bu içerikler metin, resim, ses ve video gibi veriler içerebilir (Vijiyarani & Suganya, 2015).

Web yapı madenciliği, Web sayfaları arasındaki ilişkiyi tanımlamak ve bu büyük ağın nasıl oluştuğunu anlamak için kullanılan bir araçtır (da Costa & Gong, 2005). Web kullanım madenciliği ise, kullanıcıların web günlük dosyalarından bilgilerin çıkarılması yoluyla kullanıcı davranışlarını ve gezinme kalıplarını incelemeyi amaçlar (Lee & Fu, 2008). Bu üç alan genellikle birbirleriyle örtüşür ve birbirini tamamlar. Bu bölümde, web madenciliğini web sitesi yönetimini destekleyen gelişmiş KDS'lerin geliştirilmesi süreci olarak ele alıyoruz. Web madenciliği süreci, gelişmiş KDS geliştirilmesi için adım adım izlenebilecek bir yöntem olarak sunulmaktadır. Bu süreç, veri madenciliği ile iş zekâsını entegre ederek KDS'lerin geliştirilmesine olanak tanır ve özellikle web sitelerinin kalitesini garanti altına alan akıllı izleme ve yönetim sistemlerinin oluşturulmasında uygulanabilir. İzleme faaliyetleri kapsamında bazı temel unsurlar bulunmaktadır. İlki, kullanıcı davranışlarının izlenmesi olup, kullanıcıların web sitesinde izledikleri yolları takip ederek, sayfaların kullanıcı hedeflerine ulaşmada ne kadar verimli olduğunu değerlendirmeyi içerir. Bir diğer önemli unsur, kullanıcı gruplarının izlenmesi ve bu grupların zaman içindeki değişimlerinin takip edilmesidir; kullanıcılar gezinme davranışlarına göre gruplandırılarak analiz edilir. İçerik kalitesi, web sitesindeki içerik ve meta verilerin doğruluğunu değerlendirme süreciyle ilgili olup, sitenin bilgi sunma amacına uygunluğunu ölçer. Son olarak, otomasyon faaliyetleri kişiselleştirme işlemlerinin etkilerinin analiz edilmesini içerir; bu, kullanıcıların önerilen ürünleri ya da sayfaları takip edip etmediğini izlemeye dayanır. Bu bağlamda, web madenciliği ile ilgili yapılmış çeşitli çalışmalar ışığında, söz konusu sürecin geliştirilmiş KDS'lerin tasarımında nasıl kullanılabileceği ortaya konulacaktır.

2. Web Madenciliğinin Karar Destek Sistemlerindeki Uygulamaları

Çalışmada web madenciliği teknikleri kullanarak gelişmiş karar destek sistemleri oluşturulmuştur. Ancak web madenciliği, farklı sorunları çözmede de çok yönlü olarak kullanılabilir. Web madenciliği araştırmaları tarafından öne çıkan diğer önemli uygulamalar şunlardır:

- **Kişiselleştirme ve öneriler:** Web kullanım madenciliğini kullanan kişiselleştirme sistemleri, web kullanım verilerinde ilginç desenleri keşfetmek için veri madenciliği tekniklerini içerir. Web kullanım madenciliği için veri kaynağı genellikle sunucu erişim günlüğüdür, ancak bazen istemci tarafı bir ajan veri toplar. Web kullanım madenciliğinin amacı, önceden işlenmiş web günlüğü verilerine istatistiksel ve veri madenciliği tekniklerini uygulayarak faydalı desenler keşfetmektir (Inbarani & Thangavel, 2013).

- **İçerik sınıflandırma ve kümeleme:** Web sayfaları, içeriklerine göre kategorilere ayrılarak belirli dizinlere yerleştirilebilir. Bu sınıflandırma, web sayfalarının daha düzenli bir şekilde organize edilmesini sağlar ve kullanıcıların istedikleri bilgilere daha kolay erişmesine yardımcı olur (Vijiyarani & Suganya, 2015).
- **Otomatik içerik özetleme:** Web sitelerinin içeriğini otomatik ve tutarlı bir şekilde tanımlayabilmek birçok avantaj sağlar: Bu özetler arama sonuçları olarak sunulabilir veya bir web dizininde arşivlenebilir (Amitay & Paris, 2000).
- **Anahtar kelime çıkarımı:** Önemli terimler (bazen anahtar kelimeler veya anahtar ifadeler olarak adlandırılır), bir metin belgesindeki içeriğin okuyuculara yüksek düzeyde bir açıklamasını veren terim gruplarıdır. Anahtar terimlerin çıkarılması, doğal dil işleme görevlerinin temel adımıdır (Grineva, Grinev, & Lizorkin, 2009).
- **Sayfa sıralaması:** 7.600 kullanıcının 10 haftalık arama günlüğü verilerine dayanarak, web üzerindeki içerik ve kullanıcı davranışları hakkında yeni içgörüler elde etmek ve bu profillerin nasıl oluşturulabileceğini inceleyen bir çalışma bulunmaktadır (Kim, Collins-Thompson, Bennett, & Dumais, 2012).
- **Web önbellekleme iyileştirme:** SDC, geçmiş kullanım verilerinden en sık gönderilen sorguların sonuçlarını çıkarır ve bunları statik, salt okunur bir önbellek bölümünde depolar. Bu strateji, önbellek sistemini optimize ederek web arama motorunun performansını artırır (Fagni, Perego, Silvestri, & Orlando, 2006).
- **Clickstream ve günlük analizi:** Web kullanım madenciliği, web tabanlı uygulamaların ihtiyaçlarını anlamak ve daha iyi hizmet etmek amacıyla, web verilerinden kullanım desenlerini keşfetmek için veri madenciliği tekniklerinin uygulanmasıdır (Srivastava, Cooley, Deshpande, & Tan, 2000).
- **Web sitesi yapısı analizi:** Web Sitesi Yapısı Analizi: Web sitesi yapısının analizi, arama motoru sıralamalarını iyileştirmek için sitenin tasarımını, mimarisini ve içeriğini ele alan bir süreçtir (Visser & Weideman, 2014).
- **Hub ve otorite sayfalarının belirlenmesi:** Bağlantılar analiz edilerek dizin sayfaları (hub) ve popüler içerik sayfaları (otoriteler) belirlenebilir. Bir hub, birçok otorite sayfasına bağlantı veren bir sayfadır ve bir otorite sayfası, birçok hub tarafından bağlantı verilen bir sayfadır. Bu karşılıklı pekiştirici ilişki, webde ilgili içerik topluluklarının tanımlanmasını sağlar (Toyoda & Kitsuregawa, 2001).
- **OLAP analizi:** Veri küpünün çok boyutlu yapısı, verileri farklı açılardan inceleme ve çeşitli perspektiflerden görme esnekliği sağlar. Bu web günlükleri veri küpünün oluşturulması, web günlük verilerini farklı açılardan görmek ve analiz etmek, oranlar çıkarmak ve birçok boyutta ölçümler

hesaplamak için OLAP (Çevrimiçi Analitik İşleme) işlemlerinin uygulanmasına olanak tanır (Zaïane, Xin, & Han, 2010).

3. Web Madenciliği ile Gelişmiş Karar Destek Sistemlerinin Oluşturulması

3.1. Web Verilerinin Tanımlanması

Web madenciliğinde, veriler sunucular, istemciler, proxy sunucuları veya bir şirketin veritabanı gibi çeşitli kaynaklardan toplanabilir. Farklı veri türleri, web madenciliğinde ve web sitesi yönetiminde önemli roller oynar. Bu veri kaynaklarının çeşitliliği, kullanıcı davranışı, içerik etkinliği ve sistem performansı üzerine kapsamlı analiz ve içgörüler elde edilmesini sağlar; nihayetinde web uygulamalarının işlevselliğini ve kullanıcı deneyimini geliştirir. İlk olarak, içerik verileri, web sayfalarında sunulan yapılandırılmış ya da yapılandırılmamış metinler ve multimedya içeriklerini kapsar ve web sitesinin sağladığı gerçek bilgilere işaret eder. Yapısal veriler ise, web sitesinin organizasyon yapısını tanımlar; bu yapı hem sayfa içi unsurların (html veya xhtml elemanlarının düzeni) hem de sayfalar arası bağlantıların nasıl düzenlendiğini içerir. Kullanım verileri, web sayfalarının nasıl ve ne sıklıkta kullanıldığına ilişkin bilgilerdir; IP adresleri, sayfa ziyaretleri, tarih ve saat bilgileri gibi kullanıcı etkileşimlerine dair veriler bu kategoride yer alır. Son olarak, kullanıcı profili verileri, web sitesi kullanıcılarının demografik veya kayıt bazlı bilgilerini içermekte olup, genellikle kullanıcıların kayıt işlemleri ve profillerinden elde edilmektedir. Bu veri türlerinin her biri, web madenciliği süreçlerinde stratejik bir öneme sahiptir ve web sitelerinin etkin yönetimi için kritik bilgi kaynakları sunmaktadır.

Kullanım, içerik ve yapı verilerine odaklanmak, web sitesi yönetim uygulamalarında sıkça kullanılan bir yöntemdir. Bununla birlikte, web sitesini yönetmek için farklı veri türleri de toplanabilir. Örneğin, bazı araştırmalarda, web tarayıcıları aracılığıyla kullanıcıların fare hareketleri, kaydırma çubuğu ve klavye etkinlikleri gibi davranışları kaydedilmiştir. Bu nedenle, bir karar destek sistemi geliştirirken, sistemin girdi olarak kullanacağı web verilerini tanımlamak kritik bir ilk adımdır.

3.2. Web Verilerinin Ön İşlenmesi

İçerik, yapı ve kullanım verilerinin veri madenciliği teknikleriyle analiz edilmeden önce, uygun bir formatta ön işlenmesi ve depolanması gerekmektedir. Bu bölümde, kullanım verilerinin ön işleme adımları açıklanmaktadır.

3.3. Kullanılan Veriler

Kullanım verileri, web erişim günlüklerinden veya sayfa etiketlemesinden elde edilebilir. Etiketleme, sayfaya yerleştirilen kod parçacıklarıyla, bir sayfanın ne zaman ziyaret edildiğini kaydeder. Bu bağlamda, daha yaygın olan web erişim günlüklerinin ön işlenmesine odaklanılmaktadır.

Web erişim günlüklerinin işlenmesi, veri eksikliği nedeniyle oldukça zorlu olabilir. Bu eksiklik, yerel önbellekleme ve vekil sunucu kullanımı gibi faktörlerden kaynaklanmaktadır. Bu nedenle, web günlüklerinin tam ve doğru analiz edilmesi, veri kalitesini artırmak için önem taşımaktadır. Örneğin, önbelleğe alınan sayfalara yapılan tekrar ziyaretler genellikle günlüklerde görünmez. Ancak Redis, Memcached, Varnish, Nginx ve Squid Proxy gibi bazı önbellekleme sistemleri, önbellek işlemleri için ayrı log dosyaları tutabilir. Bu sistemler, önbellekleme süreçlerini ve istekleri daha ayrıntılı izleme imkânı sunar. Vekil sunucular ise farklı kullanıcıların isteklerine aynı kimliği atayabilir.

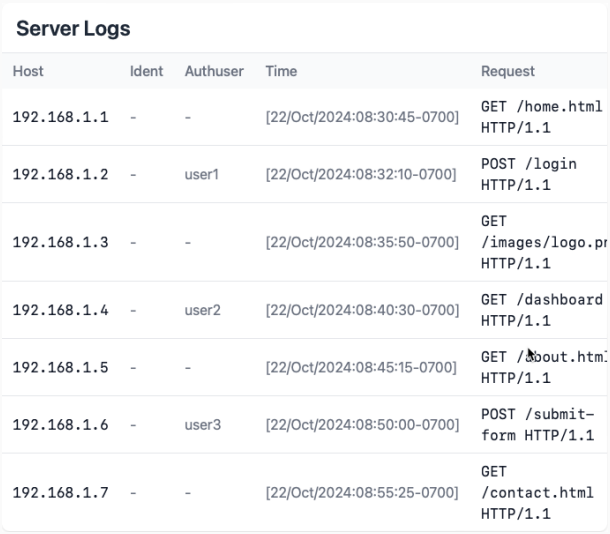
Tablo 1. Web sunucusu kayıtlarındaki veri dosyaları

Veri Alanı	Açıklama
Host	Eğer istemci bilgisayarın alan adı veya adı mevcut değilse, bu durumda istemcinin IP adresi kullanılır.
Ident	Belirli bir TCP bağlantısında istemci bilgisayarın kimlik bilgileri, sunucu tarafından tanımlanabilir ve bu bilgiler, kullanıcının oturum açma adı veya benzeri kimlik doğrulama bilgilerini içerebilir. Eğer bu bilgiler mevcut değilse, bu durumda "-" karakteri ile ifade edilir.
Authuser	Şifre korumalı bir sayfa veya dosya talebinde kullanılan kullanıcı kimliği, genellikle kullanıcının kimlik doğrulaması için gereklidir ve bu kimlik, kullanıcı adı gibi oturum açma bilgilerini içerir. Eğer bu kimlik bilgisi mevcut değilse ya da doğrulanmamışsa, bu durumda ilgili alan "-" karakteri ile doldurularak gösterilir.
Time	Bir erişimin veya isteğin sunucuya ulaştığı tarih ve saat bilgisi, sunucular tarafından genellikle zaman damgası olarak kaydedilir. Bu zaman damgası, erişimin tam olarak ne zaman gerçekleştiğini belirlemek için kullanılır ve çoğunlukla [gün/ay/yıl:saat:dakika:saniye zaman_dilimi] formatında tutulur. Bu format, sunucunun bulunduğu coğrafi konuma göre ayarlanmış zaman dilimini de içerir ve kayıtların daha düzenli ve analiz edilebilir olmasını sağlar.

Method	Bir sayfaya veya dosyaya erişim yöntemi, sunucu ile istemci arasındaki veri alışverişini yönetmek için kullanılan HTTP protokolü kapsamında tanımlanmış farklı yöntemlerden oluşur. En yaygın kullanılan yöntemler GET ve POST'tur. GET yöntemi, verileri URI'ye kodlayarak sunucuya iletir ve genellikle yalnızca web sayfası içeriğini istemciye getirmek amacıyla kullanılır. POST yöntemi ise verileri, isteğin bir parçası olarak (yani, isteğin gövdesinde) taşır ve bu yöntem hem web içeriğini getirmek hem de sunucuya bilgi göndermek için kullanılır. Berners-Lee ve Connolly'ye (1995) göre, GET yöntemi yalnızca sayfa içeriğini kullanıcıya ulaştırmak amacıyla kullanılmalıdır, POST ise web sunucusuna bilgi göndermek ve sunucudan veri almak için kullanılmalıdır. Bu ayrım, veri güvenliği ve verimliliği açısından önemli bir rol oynar.
Request	Tarayıcı tarafından istenen bir sayfa veya dosyanın URI'si (Tekdüzen Kaynak Tanımlayıcı), web üzerindeki belirli bir kaynağı tanımlamak ve ona erişim sağlamak için kullanılan bir karakter dizisidir. URI, bir web kaynağının konumunu veya adını belirtir ve bu sayede tarayıcı, sunucudan doğru içeriği talep edebilir. Berners-Lee, Fielding ve Masinter'in (1998) tanımına göre URI, web üzerindeki kaynakları benzersiz bir şekilde tanımlamak için kullanılan kompakt ve standart bir formattır. URI'ler, internet üzerindeki veri paylaşımı ve kaynak yönetimi için temel bir yapı taşı olup, hem tarayıcılar hem de sunucular arasında doğru iletişimi sağlamak için kullanılır.
Protocol	Tarayıcının sayfaları veya dosyaları almak için kullandığı protokolün sürümü, istemci ve sunucu arasındaki iletişim kurallarını belirleyen ve web üzerinden veri aktarımını sağlayan bir yapıdır. Bu protokol sürümleri, hangi özelliklerin ve yöntemlerin kullanılabileceğini belirler. En yaygın kullanılan protokol sürümleri HTTP/1.0 ve HTTP/1.1'dir. HTTP/1.0, her bir istek için ayrı bir bağlantı açılmasını gerektirirken, HTTP/1.1, aynı bağlantı üzerinden birden fazla isteğin gerçekleştirilmesine olanak tanıyarak daha verimli bir veri aktarımı sağlar. Bu protokol sürümü, web sayfalarının doğru ve hızlı bir şekilde iletilmesi için tarayıcı ve sunucu arasındaki iletişimde kritik bir rol oynar.
Status	İstek durumunu gösteren üç haneli bir HTTP durum kodu, tarayıcı tarafından sunucuya gönderilen isteğin sonucunu belirten bir geri bildirimdir. Bu kodlar, sunucu ve istemci arasındaki iletişimin başarılı olup olmadığını veya hangi tür hatalarla karşılaşıldığını ifade eder. Örneğin, 200 kodu, istenen sayfa veya dosyanın başarıyla alındığını gösterirken, 404 kodu, talep edilen sayfa veya dosyanın sunucuda bulunamadığını ifade eder. Bu durum kodları, kullanıcı deneyimini iyileştirmek ve hataların nedenini anlamak için önemli olup, tarayıcı ve sunucu arasındaki iletişimin her aşamasında kullanılır.

Bytes	Web sunucusundan istemci bilgisayara aktarılan bayt sayısı, sunucunun istemciye gönderdiği veri miktarını ifade eder. Bu veri miktarı, tarayıcıya yüklenen sayfa, dosya, görsel ya da diğer multimedya içeriklerinin toplam boyutunu kapsar. Aktarılan bayt sayısı, web trafiğinin ölçülmesi, ağ performansının değerlendirilmesi ve kullanıcıya sunulan içeriğin boyutunu analiz etmek açısından önemli bir metriktir. Bu bilgi, aynı zamanda veri transferinin ne kadar başarılı olduğunu ve kullanıcıya gönderilen içeriğin büyüklüğünü anlamak için de kullanılır.
Referrer	Mevcut sayfaya yönlendiren köprünün URI'sini içeren bir metin dizisi, kullanıcının hangi sayfa veya kaynaktan geldiğini belirten bilgiyi taşır. Bu metin dizisi, tarayıcının önceki sayfadan mevcut sayfaya geçiş yaparken ilettiği URI'yi içerir ve genellikle "referer" olarak adlandırılır. Bu bilgi, web sunucuları tarafından, kullanıcının gezinme yolunu analiz etmek, trafik kaynaklarını izlemek ve bağlantı yapısının nasıl kullanıldığını anlamak amacıyla kullanılır. Ayrıca, referans sayfasının URI'si, ziyaretçilerin siteye nasıl ulaştıklarını ve hangi yolları izlediklerini görmek için önemli bir veri kaynağıdır.
Useragent	Kullanıcı tarafından kullanılan işletim sistemi ve tarayıcı yazılımının tanımlanması, istemcinin hangi platform ve yazılım kombinasyonu ile web sitesine eriştiğini belirlemek amacıyla sunucuya iletilen bilgileri içerir. Bu bilgiler, genellikle "User-Agent" başlığı altında sunucuya gönderilir ve kullanıcının işletim sistemi türünü (örneğin Windows, macOS, Linux) ve kullandığı tarayıcı yazılımını (örneğin Chrome, Firefox, Safari) tanımlar. Bu tanımlama, web sitesi uyumluluğunu sağlamak, kullanıcı deneyimini optimize etmek ve gerekli durumlarda farklı platformlar için özelleştirilmiş içerik sunmak için önemlidir. Ayrıca, bu veriler, web trafiği analizinde ve kullanıcı kitlesinin hangi teknolojileri kullandığını anlamada da kullanılır.

Yaygın web sunucuları, Apache, Nginx, IIS, LiteSpeed ve Tomcat gibi, günlükleri genellikle Common Log Format, Combined Log Format ve W3C gibi farklı formatlarda üreterek erişim ve hata kayıtlarını detaylı bir şekilde kaydeder. Common Log Format (CLF), NCSA log formatı olarak da adlandırılır ve web sunucuları tarafından log dosyaları oluşturmak için standart bir metin dosyası biçimi olarak kullanılır (Sharma, Yadav, & Bohra, 2015). Tablo 1, bu günlüklerde yaygın olarak bulunan veri alanlarını, Şekil 1 ise örnek bir günlük kaydını göstermektedir. Her bir satır, bir web sitesine yöneltilen bir isteği temsil etmektedir. Örneğin; 127.0.0.1 IP adresinden index.html sayfasına gerçekleştirilen bir ziyaret bu duruma örnek olarak verilebilir.



Server Logs

Host	Ident	Authuser	Time	Request
192.168.1.1	-	-	[22/Oct/2024:08:30:45-0700]	GET /home.html HTTP/1.1
192.168.1.2	-	user1	[22/Oct/2024:08:32:10-0700]	POST /login HTTP/1.1
192.168.1.3	-	-	[22/Oct/2024:08:35:50-0700]	GET /images/logo.pr HTTP/1.1
192.168.1.4	-	user2	[22/Oct/2024:08:40:30-0700]	GET /dashboard HTTP/1.1
192.168.1.5	-	-	[22/Oct/2024:08:45:15-0700]	GET /about.html HTTP/1.1
192.168.1.6	-	user3	[22/Oct/2024:08:50:00-0700]	POST /submit-form HTTP/1.1
192.168.1.7	-	-	[22/Oct/2024:08:55:25-0700]	GET /contact.html HTTP/1.1

Şekil 1. Günlük formatı gösterimi

Bu günlüklerin, zaman dilimi farklarını dikkate alarak zaman sırasına göre sıralanması önerilmektedir (Raux, Ma, & Cornelis, 2016), bu günlüklerin saat dilimi farklılıklarını göz önünde bulundurarak zaman sırasına göre sıralanmasını önermektedir.

Birleştirilen günlüklerin temizlenmesi gerekmektedir. Birleştirilmiş logların temizlenmesi, gereksiz isteklerin (örneğin resimler ve betikler) yanı sıra robotlar tarafından yapılan isteklerin kaldırılmasını içermelidir (Srivastava, Srivastava, Garg, & Mishra, 2021).

Kullanıcı ve oturumların tanımlanması sonraki adımdır. Kullanıcı, web sitesini ziyaret eden bireyi temsil ederken, oturum bir ziyaret sırasında erişilen sayfalar dizisini ifade eder. Çerezler, kullanıcıları ve oturumları izlemek için kullanılır. Eğer çerezler mevcut değilse, kullanıcıları tanımlamak için bazı kestirim yöntemleri kullanılabilir. Pabarskaite ve Raudys, çeşitli kestirim yöntemlerini önermektedir:

- **HTTP girişi:** Kullanıcıların giriş yaparken sağladıkları bilgiler, log dosyasındaki Ident ve Authuser alanlarında takip edilebilir.
- **İstemci tipi:** Aynı IP adresinden gelen ancak farklı Kullanıcı Aracısı bilgilerine sahip iki istek, farklı kullanıcılardan yapılmış olabilir (Grill & Rehák, 2014).
- **Site yapısı:** ile ilgili olarak, bir sayfaya erişim, önceki sayfa grubundan yapılamıyorsa, bu durum yeni bir kullanıcı tarafından istek yapıldığı anlamına gelebilir. Bu, kullanıcıların davranışlarını anlamak için önemli bir gösterge olabilir.

Oturum tanımlama işlemi için şu kestirim yöntemleri kullanılabilir:

- **Zaman aralığı:** Aynı kullanıcı tarafından yapılan iki istek arasındaki zaman aralığı belirli bir eşiği aşıyorsa (genellikle 30 dakika), bu isteklerin farklı oturumlara ait olduğu kabul edilmektedir. Bu yaklaşım, kullanıcı oturumlarını doğru bir şekilde tanımlamak için önemlidir ve web analitiği süreçlerinde kullanıcı davranışlarını anlamada yardımcı olur.
- **Maksimum ileri referans:** Sayfa sıralaması içinde maksimum ileri referans noktasına ulaşıldığında ve aynı sayfa tekrar ziyaret edilirse, yeni bir oturum başlatılacaktır (Shen, Wei, Sundaresan, & Ma, 2012).
- **Yönlendiren sayfa:** Eğer sayfa i2, bir önceki sayfa i1'i yönlendiren sayfa değilse, yeni bir oturum başlatılır.

3.4. İçerik Verileri

Web sayfaları hem yapılandırılmış hem de yapılandırılmamış metinlerin yanı sıra çeşitli multimedya içerikleri barındırır. Genellikle, bir sayfanın yapısı html veya xhtml kullanılarak yazılır. Şekil 2'de, bir web sayfası ve onun html temsilini gösteren bir örnek bulunmaktadır.

Bir sayfanın içeriği üzerinde herhangi bir ön işleme yapılmadan önce, sayfanın yerel bir kopyasının oluşturulması gerekir. Bunun için, web sitelerini otomatik olarak ziyaret eden ve sayfalarını sistematik bir şekilde okuyan tarayıcılar (crawler) kullanılabilir. Yerel kopya oluşturulduktan sonra, içerik ön işleme aşamasına geçilir. Bu aşama, html kodunu temizlemek (yani, standartlara uygun hale getirmek) ve daha sonra sayfadaki ilgi çekici öğeleri içeren etiketlerin seçilmesini içerir. Örneğin, <title> etiketi kullanılarak sayfanın başlığı elde edilebilir.

Şekil 2. Bir web sayfası örneği (sol-html kodu, sağ-web sayfası)

Statik web sayfalarının ön işlenmesi, html kodunun temizlenmesi ve işlenmesi açısından oldukça basittir. Ancak, dinamik web sayfaları daha karmaşık zorluklar sunar. Dinamik sayfalar genellikle bir içerik yönetim sistemi (CMS) tarafından veritabanlarından veri çekilerek oluşturulur. Bu sistemler neredeyse sınırsız sayıda sayfa üretebilir ve bunların tümünü ön işleme tabi tutmak pratik değildir.

Eğer bir web sitesinin yalnızca belirli bir kısmı ön işlenirse, sonuçlar yanıltıcı olabilir ve siteyi tam olarak yansıtmaz.

Dinamik sayfaları yönetmek için, önemli içeriği tanımlarken sistemi aşırı yüklemeyecek özel tekniklere ihtiyaç vardır. Bu zorlukla başa çıkmanın bir yolu, dinamik web sitesinin temsili bir örnek kümesini ön işleme tabi tutmak ve bu örneğin genel siteyi doğru şekilde temsil etmesini sağlamaktır.

3.5. Veri Yapıları

Bağlantılar (hyperlink), web sayfalarının yapısal elemanları olup, sayfanın farklı bölümlerini birbirine bağlar. Bir bağlantı aynı sayfanın başka bir bölümüne yönleniyorsa, buna belge içi bağlantı (intra-document hyperlink) denir. İki farklı web sayfasını birbirine bağlayan bağlantılar ise belge dışı bağlantı (inter-document hyperlink) olarak adlandırılır. Bu bağlantılar, web sayfaları arasında köprüler kurar ve sayfaların çiftler halinde birbirine bağlanmasını sağlar.

Yapı verilerinin ön işlenmesi, bu bağlantıların tespit edilmesi amacıyla HTML kodunda href etiketlerinin aranmasını içerir. Bu, web sitesinin yapısını analiz etmemizi ve sayfalar arasındaki bağlantıları ortaya çıkarmamızı sağlar. Şekil 2'de, bir bağlantı örneği vurgulanmış elipsler içinde gösterilmektedir.

Bağlantıların tespit edilmesi, özellikle Adobe Flash gibi etkileşimli içerik oluşturma teknolojileriyle geliştirilen web sayfalarında daha zorlu hale gelebilir. Bu tür teknolojiler, bağlantıları html kodunun içine gömerek geleneksel yöntemlerle tespit edilmesini zorlaştırabilir.

Yapı verileri ve bağlantılar ön işleme tabi tutulduktan sonra bir veritabanına yüklenir. Bu, veri madenciliği tekniklerinin uygulanmasını kolaylaştırır ve web sitesinin yapısal analizini derinlemesine yapmayı mümkün kılar.

3.6. Web Verisi Depolama (Data Warehousing)

Web sitesi yönetiminde, zengin verilerin saklanabileceği bir veritabanı çok önemli bir bileşendir. Ancak, geleneksel veritabanı sistemleri bu tür görevler için uygun değildir, çünkü bu sistemler daha çok işlem odaklıdır. Oysa web yönetimi, analiz odaklı bir görevdir. Bu nedenle, bir veri ambarı, analitik görevler için tasarlanmış bir sistem olduğundan, web verilerini yönetmek için daha uygun bir çözümdür (Rudniy, 2022).

Veri ambarı, yönetim karar alma süreçlerini desteklemek amacıyla konu odaklı, entegre, zamanla değişen ve kalıcı bir veri koleksiyonu olarak tanımlanır (Inmon, 2005). Aynı zamanda, analizler ve sorgular için özel olarak yapılandırılmış işlem verilerinin bir kopyası olarak tanımlanabilir (Shanmugasundaram et al., 1999).

Web veri ambarları bazen Webhouse olarak adlandırılır ve genellikle bir web sitesinin tüm tıklama izleri (clickstream) ile ilgili verileri depolamak için kullanılır. Ancak, daha yeni çalışmalar, veri ambarının yalnızca tıklama izlerini değil, aynı zamanda web sitesinin içerik ve yapısal verilerini de saklaması gerektiğini savunmaktadır. Bu, daha kapsamlı ve etkili veri analizi için önemlidir.

Bir veri ambarı tasarlarken, ilk adım gereksinim analizidir. Bu analiz, veri ambarının yanıtlaması gereken iş sorularını ve kullanıcı bilgi ihtiyaçlarını belirler. Bir web sitesi için tipik iş soruları şunlar olabilir:

- Web sitesinin hangi bölümleri daha fazla ziyaretçi çekiyor?
- Belirli bir dönemde en çok ziyaret edilen sayfa hangisidir?
- Bir oturumda sayfalarda ne kadar zaman harcanıyor?
- Ortalama oturum uzunluğu nedir (zaman ve ziyaret edilen sayfa sayısı açısından)?
- Web trafiğinin en yoğun olduğu saatler hangileridir?
- Bunun yanı sıra, veri ambarları daha karmaşık sorulara da yanıt verebilir:
- Hangi sayfalar kullanıcıları müşteriye dönüştürüyor?
- Web sitesinin önerileri ne kadar başarılı?
- Web sayfasındaki içerik tanımlayıcıları (örneğin, anahtar kelimeler, kategoriler) ne kadar uygun?

Veri ambarı gereksinimlerini toplamak için iki ana yol bulunmaktadır: görüşmeler ve kolaylaştırılmış oturumlar. Görüşmeler düzenlemek, bireysel katkıları toplamak açısından daha kolaydır ve katılımcıların ihtiyaçlarını anlamak için etkili bir yöntemdir. Öte yandan, kolaylaştırılmış oturumlar grup bazında gerçekleştirilir ve bilgi toplama sürecini hızlandırabilir. Ancak, bu tür oturumlar katılımcılardan daha fazla bağlılık gerektirir, çünkü grup dinamikleri ve etkileşimi üzerinde etkili olabilmek için daha derin bir katılım sağlanması önemlidir.

Gereksinimler belirlendikten sonra, veri ambarı modelleme aşamasına geçilir. En yaygın modeller arasında Küp, Yıldız ve Kar Tanesi yer alır. Küp modelinde veriler, çok boyutlu dizilerle temsil edilir. Yıldız ve Kar Tanesi modelleri ise verilerin İlişkisel Veritabanı Yönetim Sistemi (RDBMS) içinde organize edilmesine olanak tanır. Yıldız modelinde merkezi bir olgu tablosu ve boyut tabloları bulunurken, Kar Tanesi modeli, boyutları daha da normalleştirerek daha karmaşık bir yapı oluşturur.

Model ne olursa olsun, veri ambarları verileri olgu (fact) ve boyut (dimension) olarak depolar. Olgular, bir iş sürecinin ölçülebilir unsurlarını temsil eder (örneğin, bir sayfa ziyaretinde harcanan zaman). Boyutlar ise bu olgulara bağlam sağlayarak analiz sırasında verilerin gruplanmasını sağlar. Boyutlar, hiyerarşik

seviyelere ayrılabilir; örneğin, tarih boyutu ay veya yıl seviyelerine göre gruplandırılabilir.

Web siteleri için kullanılan bir veri ambarı, kullanım verilerini olgu tablosunda, içerik ve yapı verilerini ise boyut tablolarında saklamaktadır. Bu veri ambarı hem çevrimdışı hem de çevrimiçi önerileri destekler. Çevrimdışı öneriler, sitedeki bağlantıların değiştirilmesi veya eklenmesi gibi öneriler sunarken, çevrimiçi öneriler her kullanıcıya ilgisini çekebilecek sayfalar sunar. İlerleyen bölümde, bir e-haber portalı için iş zekâsı destekli bir veri ambarı önerisi sunulacaktır.

3.7. Çıkarma, Dönüştürme ve Yükleme (ETL) Süreci

ETL (Veri Çıkarma, Dönüştürme ve Yükleme) süreci, veriyi çeşitli kaynaklardan çıkaran, dönüştüren ve veri ambarına yükleyen yöntemler ve araçlardan oluşan bir süreçtir (Mukherjee & Kar, 2017). ETL süreci, veri ambarına yüklenen verilerin çeşitli kaynaklardan, formatlardan ve türlerden geldiği göz önüne alındığında, veri ambarı geliştirilmesinde çok önemli bir role sahiptir. ETL süreci üç ana aşamadan oluşur.

Çıkarma: Bu aşamada, veriler çeşitli kaynaklardan toplanır ve geçici olarak bir Veri Hazırlık Alanı (Staging Area - SA)'na depolanır. SA, verilerin veri ambarına yüklenmeden önce tutulduğu ve dönüştürüldüğü bir ara aşama olarak görev yapar. SA genellikle bir dosya sistemi dizini veya ilişkisel veritabanı olarak yapılandırılır.

Dönüştürme: Bu aşama genellikle en karmaşık ve zaman alıcı adımdır. Dönüştürme sırasında, SA'nda bulunan verilere bir dizi ön işleme ve dönüştürme kuralları uygulanır. Amaç, ham veriyi veri ambarında kullanılabilecek yapıya dönüştürmektir. Uygulanan dönüşüm kuralları, verilerin türüne ve ambarın gereksinimlerine göre değişir.

Yükleme: Dönüştürülmüş veriler, veri ambarına yüklenir. Eğer SA ve veri ambarı aynı makinede bulunuyorsa, bu işlem oldukça basit olabilir. Ancak, SA ve veri ambarı farklı sistemlerdeyse, verilerin aktarımı için bir protokol belirlenmesi gerekir.

Bu üç aşama, daha önce ele aldığımız web veri ön işleme teknikleriyle birleştirilerek, web verileri için kapsamlı bir ETL süreci oluşturulabilir. Bu sayede toplanan web verilerinin temiz, tutarlı ve analiz edilebilir hale getirilmesi sağlanır.

3.8. Çevrimiçi Analitik İşleme (OLAP)

Veri ambarında depolanan veriler genellikle OLAP ile analiz edilir. Bu bölümde, önceden belirtildiği gibi, çok boyutlu modelin temel bir özelliği,

verilerin birden çok perspektiften ve çeşitli detay seviyelerinde incelenmesine olanak tanımasıdır. Aşağıda temel bir OLAP işlemleri seti tanımlanmıştır (Malinowski & Zimányi, 2008).

Chen, Yan, Zhu, Han, & Yu (2009)'ya göre OLAP, karmaşık çok boyutlu sorguları hızlı bir şekilde yanıtlamayı sağlayan bir yaklaşımdır ve hızlı veri analizi yapma yeteneği sunar. OLAP'ın üç temel uygulama türü vardır: MOLAP, ROLAP ve HOLAP (Kimball & Merz, 2000). Veriler Çok Boyutlu Veritabanı Yönetim Sistemi (MDBMS)'de uygulandığında MOLAP kullanılır. Verilerin İlişkisel Veritabanı Yönetim Sistemi (RDBMS)'de uygulandığı durumlarda ise ROLAP kullanılır.

Hem MDBMS hem de RDBMS kullanılarak veri küpünün yönetildiği hibrit çözümlerde, Hibrit OLAP (HOLAP)uygulanır. Uygulama türünden bağımsız olarak, OLAP analizlerinde sıkça kullanılan birkaç yaygın işlem şunlardır (Malinowski & Zimnyi, 2008):

- **Dilme (Slice):** Bir boyuttaki bir veya daha fazla veri alanı için tek bir değer seçilerek alt küp oluşturulur.
- **Küp dilimi (Dice):** Slice işlemine benzer, ancak iki veya daha fazla boyutta uygulanır.
- **Derinleştirme/Yüzeyleştirme (Drill Down/Up):** Bu teknik, verilerin farklı özetleme seviyeleri arasında gezinmeyi sağlar. Yüzeyleştirme özet veriye ulaşmayı sağlarken, derinleştirme ayrıntılı verilere inmeyi sağlar.
- **Küpler arası sorgulama (Drill-Across):** Birden fazla veri küpünü kapsayan sorgular çalıştırılır.
- **Döndürme (Pivot):** Küpün boyutsal yönü değiştirilerek verilerin farklı bir açıdan gösterilmesi sağlanır.

Bu işlemler, OLAP'ı çok boyutlu veri analizinde güçlü bir araç haline getirir ve karar verme süreçlerinde esnek ve hızlı veri keşfi sunar.

3.9. Kalıp Keşfi ve Analizini Tanımlama

Ge, Song, Ding, & Huang (2017)'a göre, istatistiksel yöntemler, veri madenciliği ve makine öğrenimi teknikleri uygulanarak, verilerden anlamlı kalıplar keşfedilebilir. Mobasher, Jain, Han, & Srivastava (1996)'ya göre, web madenciliğinde, web verilerinden kalıplar çıkarmak için en sık kullanılan teknikler arasında istatistiksel analiz, birliktelik kuralları, sıralı kalıplar, kümeleme, sınıflandırma ve bağımlılık modellemesi yer alır. Aşağıda bu tekniklerin kısa bir tanımı verilmiştir:

- **İstatistiksel analiz:** Bu teknik, sayfalar, sayfalarda geçirilen süre, oturum uzunluğu gibi değişkenler üzerinde frekans, ortalama, medyan gibi betimleyici istatistiksel analizler yapmamıza olanak tanır.

- **İlişkilendirme kuralları:** Bu yöntem, öğeler ya da sayfalar arasındaki sık görülen kalıpları, ilişkileri ve korelasyonları keşfetmek için kullanılır.
- **Ardışık kalıplar:** İlişkilendirme kurallarının bir uzantısı olan bu teknik, birlikte gerçekleşme kalıplarını zaman sırası unsuru ile birlikte ortaya çıkarır.
- **Kümeleme:** Benzer özellikler veya davranışlara sahip öğeleri veya kullanıcıları gruplandırır.
- **Sınıflandırma:** Bu teknik, öğeleri (örneğin web sayfalarını) önceden tanımlanmış sınıflardan birine atar.
- **Bağımlılık modelleme:** Bu yöntem, web alanındaki değişkenler arasında önemli bağımlılıkları temsil eden modeller oluşturur (örneğin web erişim davranışları).

Kalıplar keşfedildikten sonra, ilgisiz olanları filtrelemek için ek bir analiz yapılmalıdır. Kalan faydalı kalıplar, karar destek sistemleri gibi çeşitli uygulamalarda kullanılabilir.

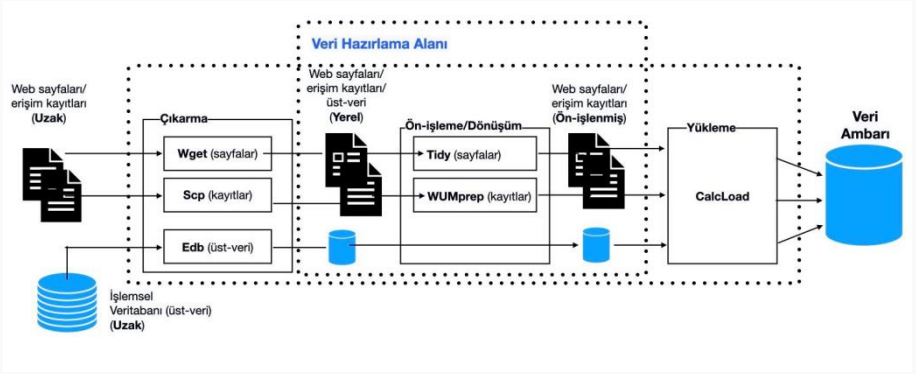
4. Bir E-Haber Portalı için Gelişmiş Karar Destek Sistemi Geliştirmek Amacıyla Web Madenciliği Kullanımı

Birçok web portalı, kullanıcıların ihtiyaçlarını karşılayan içerikleri, hizmetleri veya ürünleri düzenlemek, organize etmek ve dağıtmak üzere tasarlanmıştır. Bu süreç, içeriği ve özelliklerini tanımlayan meta verilere örneğin, anahtar kelimeler, kategoriler, yazarlar ve diğer tanımlayıcılar büyük ölçüde dayanır. Örneğin, arama motorları, içerik ile kullanıcı sorguları arasındaki ilgiyi belirlemek için genellikle meta verileri kullanır. Aynı şekilde, bir portalın yapısı, genellikle meta veriler (örneğin kategoriler) tarafından belirlenen içeriğin nasıl erişilebilir olduğunu şekillendirir. Genellikle, içerik yayınlayan yazarlar meta verileri de sağlarlar. Yayın süreci bir iş akışını takip eder ve bu iş akışının karmaşıklığı farklılık gösterebilir: bazı durumlarda, yazarlar doğrudan içerik yayınlayabilirken, diğer durumlarda editörler içeriği ve meta verileri gözden geçirerek yayına onay verir. Editörler ayrıca, meta verilerde düzeltmeler yapabilir veya değişiklik önerilerinde bulunabilirler.

Birden fazla yazarın yer aldığı veya daha esnek bir yayın iş akışının olduğu durumlarda, meta veriler içeriği doğru şekilde yansıtmayabilir ve bu da sunulan hizmetlerin kalitesini düşürebilir. Örneğin, içeriğe atanan anahtar kelimeler uygun değilse, kullanıcılar arama işlevini kullanarak ilgili içerikleri bulmakta zorlanabilirler. Bu nedenle, meta verilerin kalitesini izlemek, içeriğin düzenli, ilişkili ve kolayca erişilebilir olmasını sağlamak için kritik önem taşır.

4.1. Web Verilerini Tanımlama ve Ön İşleme

Çalışmada, tipik web yönetim uygulamalarında girdi olarak kullanılan içerik, yapı ve kullanım verilerini toplamak ve ön işlemek için literatürde yer alan bir ETL aracı kullanılmıştır. Şekil 3'te gösterilen ETL aracı, çeşitli mevcut araçların bir araya getirilmesiyle oluşturulmuştur. Aracın en önemli avantajlarından biri, manuel müdahale gerektirmeyen bir toplu işlem olarak çalışmasıdır. ETL süreci, çıkarma, ön işleme/dönüştürme ve yükleme olmak üzere üç ana aşamadan oluşmaktadır. Çıkarma aşamasında, araç uzak kaynaklardan verileri alır ve bunların yerel bir kopyasını oluşturur. Bu yerel kopya, dosya sistemindeki bir Veri Hazırlık Alanı (DSA)'da saklanır. Bu işlem için wget, scp ve edb araçları kullanılır. Wget, http, https ve ftp gibi yaygın internet protokolleri aracılığıyla dosya indirmek için kullanılan açık kaynaklı bir araçtır. Scp, yerel ve uzak sunucular arasında güvenli dosya transferi sağlar. Edb ise, bir işlem veri tabanından meta verileri seçip, bunları metin dosyalarına dönüştüren bir SQL bileşenidir.



Şekil 3. Çıkarma, dönüştürme ve yükleme aracı

Ön işleme ve dönüştürme aşamasında, veriler kalite metriklerinin hesaplanmasına ve veri ambarına yüklenmeye hazır hale getirilir. Web sayfaları (içerik) için araç, html dosyalarını okur ve bunları temiz ve düzgün biçimlendirilmiş xhtml formatına dönüştürür. Bu işlem için tidy kullanılır. Tidy, iyi biçimlendirilmiş xml/xhtml/html dosyaları oluşturmak için kullanılan açık kaynaklı bir yazılımdır. Erişim günlüklerinin ön işlenmesi ise log dosyalarının birleştirilmesi, gereksiz taleplerin çıkarılması, robot taleplerinin filtrelenmesi ve kullanıcılar ile oturumların belirlenmesi işlemlerini içerir. Web loglarının veri madenciliğine hazırlanması sürecinde WUMPrep adlı Perl programlarının kullanımına değinir ve bu aşamada meta verilerin dönüştürülmediğini, doğrudan kalitelerinin değerlendirilmesine odaklanıldığını belirtir (Marquardt, Becker, & Ruiz, 2004).

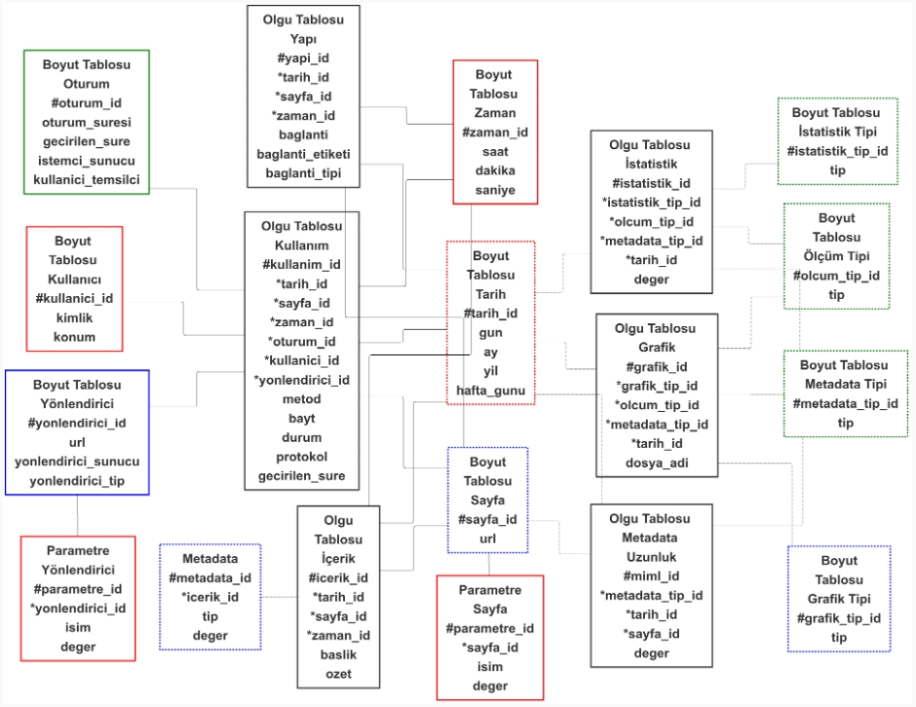
Bu aşamadan sonra, metrikler hesaplanır ve etkinlik verileri ile birlikte veri ambarına yüklenir. Bu işlem için, kalite metriklerini hesaplayan ve veri ambarına yükleyen R tabanlı bir araç olan CalcLoad kullanılır. R, veri manipülasyonu, istatistiksel hesaplamalar ve grafiksel gösterimler için entegre bir yazılım ortamıdır. Ayrıca, MacOS, Windows ve Linux işletim sistemlerine multiplatform desteği bulunmaktadır.

4.2. Web Veri Ambarı

Verileri ve metrikleri verimli bir şekilde yönetmek ve depolamak amacıyla, tanımlı veri ambarını genişlettik (Sokolov & Turkin, 2018). Bu genişleme, meta veriler, metrikler ve makro göstergeleri (istatistikler ve grafikler gibi) depolamak için ek tablolar içermektedir.

Veri ambarı için genişletilmiş yıldız şeması Şekil 4'te gösterilmiştir. Orijinal şema, yalnızca bir web sitesinin yapı, kullanım ve içerik verilerini depolayan olgu tablolarını içeriyordu. Meta veriler, metrikler ve makro göstergelere dayalı daha kapsamlı analizler yapılabilmesi için, bu yeni sürümde ek tablolar eklenmiştir. Şekil 4'teki kesikli çizgiler, yeni eklenen tabloları göstermektedir.

Yapı olgu tablosunda, web sitesindeki her bir hiperlink depolanmakta ve böylece sitenin hiyerarşik organizasyonu kaydedilmektedir. Kullanım olgu tablosu ise web günlüklerinden alınan bilgileri saklamaktadır. Son olarak, İçerik ve Meta Veriler tablolarında web sitesindeki içerik (örneğin, başlık ve özet) ve buna bağlı meta veriler (örneğin tür ve değer) depolanmaktadır. Bu yeni sürümde, Meta Veriler tablosu ile İçerik tablosu arasındaki ilişki, belirli bir içerik öğesine bağlı farklı meta verilerin depolanmasına olanak tanımaktadır.



Şekil 4. Veri ambarının yıldız şeması. (Meta-veri uzunluğunun hesaplanması ve depolanmasında kullanılan tabloları vurgulamaktadır.)

Her bir metrik, kendi olgu tablosunda depolanmakta olup, metrik değeri ve buna ilişkin veriler (örneğin, değerlendirilen meta veri türü, metrikle ilişkili sayfa ve metrik hesaplandığında) kaydedilmektedir. Şekil 4'te, meta veri alanındaki kelime sayısını hesaplayan Meta Veri Uzunluğu metriği için bir olgu tablosu örneği gösterilmektedir. Bu tablo, meta veri türünü, hesaplama tarihini, ilişkili web sayfasını ve metrik değerini depolar.

İstatistikler ve Grafikler olgu tabloları, yapı açısından benzerlik gösterir ve sırasıyla istatistiksel göstergeler ile grafiksel temsillerin kaydedilmesini sağlar. İstatistikler tablosu, istatistiksel gösterge türünü ve değerini kaydederken, Grafikler tablosu grafik türünü ve dosya adını kaydetmektedir. Her iki tablo da ayrıca kullanılan metrik, değerlendirilen meta veri türü ve hesaplama tarihini de saklar. İstatistiksel göstergelerin, metriklerin, meta verilerin ve grafiklerin türleri, sırasıyla İstatistik Türü, Metrik Türü, Meta Veri Türü ve Grafik Türü boyut tablolarında saklanmaktadır.

4.3. Kalıp Keşfi ve Analizi

Kalıp keşfi ve analiz aşamasında, meta verilerin kalitesini izlemek için kullanılan metriklerin hesaplanmasında istatistiksel analiz ve ilişkilendirme

kuralları üzerinde durulmuştur. İçerik meta verilerinin kalitesini değerlendirmek için çeşitli metrikler geliştirilmiştir. Bu metrikler, veri kalitesi ilkelerine dayalı olarak tasarlanmış ve meta veri kalitesinin artırılmasına yönelik önemli bir temel sağlamaktadır. Bu süreç, veri yönetimi ve analizine önemli katkılarda bulunabilir.

Tablo 2. Çeşitli metriklerin isimleri ve açıklamaları

Ad	**Açıklama**
Meta Veri Uzunluğu	Bu metrik, bir meta veri alanındaki toplam kelime sayısını hesaplayarak, meta verinin uzunluğunu ve uygunluğunu değerlendirir. Meta veri alanında bulunan kelimelerin sayısının aşırı uzun ya da aşırı kısa olması, bu alanın içeriği doğru ve yeterli bir şekilde tanımlamadığına işaret edebilir. Özellikle çok kısa meta veriler, genellikle yetersiz bilgi sağlayarak sayfanın veya dosyanın içeriğini tam anlamıyla yansıtmazken, aşırı uzun meta veriler ise gereksiz veya fazla detaylı olabilir. Bu nedenle, meta verinin uygun bir kelime sayısında tutulması, hem arama motoru optimizasyonu (SEO) hem de kullanıcı deneyimi açısından önemlidir.
Meta Veri Değerleri Arasindaki ilişki	İlişkilendirme kuralının güven seviyesi, belirli bir veri kümesinde X ve Y değerlerinin arasındaki ilişkinin gücünü ve tutarlılığını değerlendirir. Yüksek bir güven seviyesi, X değerinin varlığı durumunda Y değerinin ortaya çıkma olasılığının oldukça yüksek olduğunu ve bu nedenle Y'nin ayrı bir veri noktası olarak gereksiz hale geldiğini gösterebilir. Bu durum, Y'nin tahmin edilmesi gereken bağımsız bir değişken olmaktan çıkarak, X'in varlığıyla otomatik olarak varsayılan bir sonuç olduğunu işaret eder. Bu durum aynı zamanda, içerik açıklamalarında veya veri analizlerinde örtük (gizli) uygulamaların geliştiğine ve bazı veri noktalarının gereksiz hale geldiğine dair bir gösterge olabilir.
Arama Sıklığı	Bu metrik, belirli bir meta veri değerinin web erişim loglarında ne kadar sıklıkla yer aldığını hesaplayarak, o meta verinin popülerliğini ve kullanım yoğunluğunu ölçer. Örneğin, bir anahtar kelimenin arama motorlarında ne kadar sık arandığı bu metrikle belirlenebilir. Sık aranan terimler, kullanıcılar tarafından daha fazla ilgi gördüğü için yüksek bir yorumlanabilirliğe ve anlamlılığa sahip olabilir. Bu tür sık kullanılan meta veriler, web sitesi optimizasyonu, kullanıcı davranış analizi ve içerik stratejileri geliştirme açısından kritik bir bilgi kaynağı olarak değerlendirilebilir.

Meta Veri Değerlerinin Yedekliği	Bu metrik, X ve Y meta veri değerleri arasındaki koşullu olasılığa dayalı olarak, iki değer arasındaki ilişkinin gücünü değerlendirir. X'in varlığı durumunda Y'nin ortaya çıkma olasılığına odaklanan bu metrik, iki veri arasında nasıl bir bağımlılık olduğunu ortaya koyar. Yüksek koşullu olasılık değerleri, Y'nin X'i gereksiz hale getirdiğini, yani Y'nin varlığının X'in varlığına bağlı olmadan tahmin edilebileceğini gösterebilir. Bu durum, içerik açıklamalarında örtük uygulamaların olduğunu ve belirli meta veri değerlerinin açıkça ifade edilmesine gerek kalmadan birbirlerini tamamladığını ya da gizli bir şekilde var olduklarını işaret eder. Böyle bir ilişki, veri analizi ve yorumlamada daha derin bağlantıları ortaya çıkarabilir.

Bu metriklerin hesaplanmasında kullanılan yöntemler, basit istatistiksel hesaplamalardan daha karmaşık tekniklere kadar değişiklik gösterebilir. Örneğin, Meta Veri Uzunluğu 2 metriği, bir meta veri alanındaki kelime sayısını sayarak hesaplanır. Basit frekanslara dayalı metrikler, örneğin Arama Sıklığı metriği (Tablo 2), oldukça yaygındır. Daha gelişmiş metrikler ise olasılıklara dayalı olabilir.

Meta Veri Değerlerinin Yedekliliği metriği, x meta veri değerinin, y değeri kullanıldığında ortaya çıkma olasılığını ölçer (Tablo 2). Daha karmaşık bir örnek olarak, Meta Veri Değerleri Arasındaki İlişki metriği, ilişkilendirme kuralları (Agrawal & Srikant, 1994) kullanılarak hesaplanır (Tablo 2). Çoğu metrik meta veriler kullanılarak hesaplanırken, bazıları Arama Sıklığı metriğinde olduğu gibi kullanım verilerini de içerir.

Örnek: "Meta Veri Uzunluğu 2" Metrik Hesaplaması

Bu metriği hesaplamak için, Meta Veriler tablosundaki (alanlar: tür, değer), Sayfa tablosundaki (alan: uri) ve Tarih tablosundaki (alanlar: gün, ay, yıl) veriler kullanılır. Aralarındaki ilişkileri kurmak için ayrıca İçerik tablosu kullanılır. Tarih tablosundaki gün, ay ve yıl alanları, analiz için hangi sayfa ve meta veri sürümünün geri çağrılacağını belirler. Bu, veri ambarının sitenin içeriğini periyodik olarak depolaması nedeniyle gereklidir.

Bu metrik için hesaplanan değerler, Meta Veri Uzunluğu 2 olgu tablosunda saklanır. Bu tablo, meta veri türünü (foreign key: metadata_type_id), hesaplama tarihini (foreign key: date_id), meta veri ile ilişkili sayfayı (foreign key: page_id) ve metrik değerini içerir.

Meta Veri Uzunluğu 2 metriği hesaplandıktan sonra, istatistiksel göstergeleri hesaplar ve grafikler oluştururuz. Meta Veri Uzunluğu 2 olgu tablosundaki tüm değerler alınır, minimum ve maksimum değerler gibi göstergeler hesaplanır ve zaman içinde metrik değerlerinin değişimini gösteren grafikler çizilir. İstatistikler ve grafikler, İstatistikler ve Grafiklerolgu tablolarında saklanır.

4.4. E-Haber Portalında Karar Destek Sistemi Uygulamasının Değerlendirilmesi

İş dünyasındaki yöneticilere yönelik bir e-haber platformu, abonelik tabanlı bir iş modeliyle faaliyet göstermektedir. Bu modelde, yalnızca ödeme yapan üyeler tam içeriğe erişebilirken, bazı içerikler halka açık olarak sunulmaktadır. Kullanıcılar, site yapısını serbestçe gezebilir ve içerikler çeşitli iş ortaklarından da sağlanmaktadır. Platformun temel amacı, kullanıcılarına ilgili içeriğe kolay erişim sağlamaktır. Bu hedef doğrultusunda, içerikleri yapılandırarak ve birbirleriyle ilişkilendirerek değer katmaktadır.

İçeriklerin yapılandırılması, anahtar kelimeler, kategoriler, şirketler ve yazarlar gibi çeşitli meta veri alanlarının doldurulması ile gerçekleştirilir. Meta veri alanlarının izlenmesi ve hatalı değerlerin düzeltilmesi, platform için son derece önemlidir; bu sayede meta veri kalitesi korunmakta ve içeriğe değer katılmaktadır.

Anahtar kelimeler, web sayfası veya site içeriğini karakterize eden ve kullanıcıların arama sürecinde kullandığı ifadelerdir. Bu platform, anahtar kelimeleri doldurma sürecini desteklemek için yarı otomatik bir prosedür uygulamaya karar vermiştir. Bu uygulama sonucunda, içeriksiz anahtar kelime sayısında sürekli bir azalma gözlemlenmiş ve böylece meta veri kalitesi artırılmıştır.

Ayrıca, Meta Veri Değerleri Arasındaki İlişki metriği gibi daha karmaşık metrikler de ilginç ve uygulanabilir olabilir. Bu metrik, birlikte daha sık kullanılan anahtar kelimeleri belirlemek için ilişkilendirme kurallarının güven seviyesini kullanır. Geliştirilen sistem, veri ambarından her bir içeriğin anahtar kelimelerini öğe sepetleri olarak toplar ve ardından anahtar kelimeler arasındaki ilişkilendirmeleri oluşturmak için bir ilişkilendirme kuralları algoritması çalıştırır.

Bu tür metriklerin kullanılması, içerik yapılandırma süreçlerini optimize ederek kullanıcılarına daha iyi hizmet sunmalarına yardımcı olmaktadır.

Tablo 3. İçerik meta-verilerinin kalitesini ölçmek için kullanılan metriklerin adı ve açıklaması

No	Ad	Açıklama
1	*Gölge İçerik Sayısı*	Gölge içerik sayısı, bir içeriğin (C1) meta-verilerinin başka bir içerik (C2) tarafından gölgelenip gölgelenmediğini belirleyen bir metriktir. C1 içeriği, C2'nin meta-veri değerlerinin bir alt kümesi olduğunda gölgelenmiş kabul edilir. Bu gölgeleme, C1 içeriğine sadece meta-veriler kullanılarak erişimi zorlaştırabilir, çünkü C2 daha kapsamlı veya eşit meta-verilere sahiptir. Bu metrik, kaç içeriğin bu şekilde gölgelenmiş olduğunu belirlemek için kullanılır (Malinowski & Zimányi, 2008).
2	*Meta Veri Değerleri Arasındaki İlişki*	Meta veri değerleri arasındaki ilişki, bir meta veri değerinin (X) başka bir meta veri değerini (Y) gereksiz hale getirip getirmediğini belirleyen bir metriktir. X -> Y ilişkilendirme kuralının güven seviyesi, X'in varlığının Y'nin açıklayıcılığını veya faydasını azaltıp azaltmadığını gösterir. Yüksek güven seviyeleri, X değerinin Y'yi gereksiz hale getirdiği durumları ve içerik açıklama kalıplarını ortaya koyar. Bu, meta veri fazlalığı veya tekrarı gibi sorunları tespit etmek için kullanılır.
3	*Arama Sıklığı*	Arama sıklığı, belirli meta-veri değerlerinin (örneğin anahtar kelimeler) web erişim loglarında ne kadar sık kullanıldığını ölçen bir metriktir. Bu metrik, hangi terimlerin veya meta-veri değerlerinin kullanıcılar tarafından en çok arandığını belirler. Yüksek arama sıklığı, ilgili meta-veri değerlerinin daha fazla yorumlanabilirliğe ve kullanıcı açısından daha fazla öneme sahip olduğunu gösterir. Bu, içerik optimizasyonu ve erişilebilirlik stratejileri için kritik bir bilgi sağlar.
4	*Meta Veri Değerlerinin Yedekliliği*	Meta veri değerlerinin yedekliliği, bir meta veri değerinin başka birini gereksiz hale getirdiğini gösteren bir metriktir. Bu metrik, koşullu olasılık kullanarak bir meta veri değerinin varlığının, başka bir meta veri değerine olan ihtiyacı ortadan kaldırıp kaldırmadığını ölçer. Yüksek yedeklilik, içerik açıklamalarında fazlalık olduğunu ve aynı bilgiyi tekrar eden meta veri değerlerinin bulunduğunu gösterir. Bu, meta verilerin optimize edilmesi ve gereksiz tekrarların önlenmesi için kullanılır.
5	*Meta Veri Uzunluğu-1*	Meta veri uzunluğu-1, bir meta veri alanındaki karakter sayısını ölçen bir metriktir. Meta verilerin çok uzun veya çok kısa olması, içeriği yeterince temsil edememe veya aşırı bilgi yüklemesi gibi durumları gösterebilir. Bu metrik, meta verilerin optimal uzunlukta olup olmadığını değerlendirerek içerik açıklamalarının kalitesini artırmaya yardımcı olur.

6	*Meta Veri Uzunluğu-2*	Meta veri uzunluğu-2, bir meta veri alanındaki kelime sayısını ölçen bir metriktir. Çok kısa ya da çok uzun meta veri alanları, meta verilerin seçiminin yetersiz veya aşırı olduğunu gösterebilir. Bu metrik, içerik açıklamalarının uygun kelime sayısında olup olmadığını değerlendirerek meta veri kalitesinin iyileştirilmesine yardımcı olur.
7	*Başlık/Özet Uzunluğu-1*	Başlık/Özet Uzunluğu-1, bir başlık veya özetin karakter sayısını ölçen bir metriktir. Çok uzun ya da çok kısa başlıklar veya özetler, içeriği doğru ve yeterli şekilde temsil edememe durumunu gösterebilir. Bu metrik, başlık ve özetlerin uygun uzunlukta olup olmadığını değerlendirerek, içerik açıklamalarının doğruluğunu ve temsil edilebilirliğini artırmaya yardımcı olur.
8	*Başlık/Özet Uzunluğu-2*	Başlık/Özet Uzunluğu-2, bir başlık veya özetin kelime sayısını ölçen bir metriktir. Bu metrik, başlık ve özetin yapısının uygun olup olmadığını belirlemeye yardımcı olur. Çok kısa veya çok uzun başlıklar/özetler, içeriğin doğru şekilde ifade edilmediğini veya aşırı bilgi verildiğini gösterebilir. Bu değerlendirme, başlık ve özetlerin optimal kelime sayısında olmasını sağlamaya çalışır.
9	*Başlık/Özet Uzunluğu-3*	Başlık/Özet Uzunluğu-3, başlık veya özetin cümle sayısını ölçen bir metriktir. Çok uzun veya çok kısa cümle sayıları, içeriğin netliği ve anlaşılırlığında sorunlara işaret edebilir. Bu metrik, başlık ve özetlerin uygun cümle yapısında olup olmadığını değerlendirerek içeriğin doğru ve net bir şekilde sunulmasına yardımcı olur.
10	*Aşırı Meta-Veri Sıklığı*	Aşırı meta-veri sıklığı, belirli bir meta-veri değerine sahip içerik sayısını izleyen bir metriktir. Çok yüksek değerler, bu meta-verinin fazla genel veya yaygın olduğunu gösterir ve bu da içerik filtreleme ve ayrıştırma işlemlerini zorlaştırabilir. Bu metrik, içeriklerin benzersizliğini ve ayırt edici niteliklerini korumak için kullanılır.
11	*Aşırı Meta-Veri Siklığı, Editör/Yazar Bazında*	Aşırı meta-veri sıklığı, editör veya yazar bazında belirli bir meta-veri değerine sahip içerik sayısını ölçen bir metriktir. Bu metrik, belirli editörler veya yazarlar tarafından kullanılan meta-veri değerlerinin ne kadar sık tekrarlandığını analiz ederek içerik oluşturma kalıplarını ve olası fazlalıkları belirlemeye yardımcı olur.
12	*Paylaşılan Meta-Veri Değerleri*	Paylaşılan meta-veri değerleri, farklı içeriklerde ortak olarak kullanılan meta-veri değerlerini izleyen bir metriktir. Bu metrik, bu içerikler arasında ilişkiler kurarak hangi meta-veri değerlerinin birden fazla içerikte tekrarlandığını ve içerikler arasındaki olası bağlantıları gösterir.
13	*Editör/Yazarlar Arasındaki Paylaşım Derecesi*	Editör/Yazarlar arasındaki paylaşım derecesi, aynı meta-veri değerini kullanan editör veya yazarların sayısını ölçen bir metriktir. Bu metrik, işbirlikçi davranışları, içerik üretimindeki ortak temaları veya belirli meta-verilerin editör/yazarlar arasında ne kadar yaygın olarak kullanıldığını anlamaya yardımcı olur.

14	*Boş Meta-Veri Alanı*	Boş meta-veri alanı, belirli bir meta-veri alanının doldurulmadığı içerik sayısını izleyen bir metriktir. Boş bırakılan meta-veri alanları, bu içeriklerin arama sonuçlarında bulunamamasına veya eksik şekilde listelenmesine yol açabilir. Bu metrik, eksik meta-verilerin tespit edilip doldurulmasını sağlamak için kullanılır.
15	*Boş Meta-Veri Alani, Editör/Yazar Bazında*	Boş meta-veri alanı, editör veya yazar bazında doldurulmamış meta-veri alanlarına sahip içerik sayısını gruplandıran bir metriktir. Bu metrik, belirli editörler veya yazarlar tarafından eksik meta-veri alanları ile oluşturulmuş içeriklerin sayısını izleyerek meta-veri eksikliklerini belirlemeye ve iyileştirme fırsatlarını ortaya çıkarmaya yardımcı olur.
16	*Tüm Boş Meta-Veri Alanları*	Tüm boş meta-veri alanları, bir içerikteki tüm meta-veri alanlarının boş olduğu içerik sayısını izleyen bir metriktir. Bu durum genellikle bir yayın hatasına işaret eder ve içeriklerin arama motorları tarafından bulunmasını veya doğru şekilde sınıflandırılmasını engelleyebilir.
17	*Tüm Boş Meta-Veri Alanlart, Editör/Yazar Bazında*	Tüm boş meta-veri alanları, editör veya yazar bazında doldurulmamış meta-veri alanlarına sahip içerikleri gruplandıran bir metriktir. Bu metrik, belirli editörler veya yazarlar tarafından eksik meta-verilerle yayınlanan içerik sayısını izleyerek, yayınlama pratiklerindeki boşlukları ve hataları vurgular.
18	*Meta-Veri Değerlerinin Uzunluğu*	Meta-veri değerlerinin uzunluğu, meta-veri değerlerindeki karakter sayısını ölçen bir metriktir. Aşırı uzun veya kısa meta-veri değerleri, içeriği doğru ve yeterli şekilde temsil etmek için yetersiz veya aşırı meta-veri seçimine işaret edebilir. Bu metrik, meta-veri değerlerinin optimal uzunlukta olup olmadığını değerlendirerek içeriğin temsil edilebilirliğini artırmaya yardımcı olur.
19	*Meta-Veri Değerlerinin Miktarı*	Meta-veri değerlerinin miktarı, bir içerikte kullanılan farklı meta-veri değerlerinin toplam sayısını ölçen bir metriktir. Bu metrik, içerikte ne kadar çeşitlilikte meta-veri kullanıldığını belirleyerek meta-veri kullanımının kapsamını ve yeterliliğini analiz etmeye yardımcı olur.
20	*İçerik Başına Meta-Veri Değerlerinin Miktarı*	İçerik başına meta-veri değerlerinin miktarı, her bir içerikte kullanılan farklı meta-veri değerlerinin sayısını ölçen bir metriktir. Düşük sayıda meta-veri değeri, içerik açıklamalarında rutin veya eksik uygulamalara işaret edebilirken, yüksek sayıda meta-veri değeri, daha dikkatli açıklamalar yapılmasına rağmen gereksiz bilgi fazlalığını da gösterebilir. Bu metrik, içeriklerin açıklayıcılığını ve meta-veri kullanımının uygunluğunu değerlendirmeye yardımcı olur.

21	*Tekil Meta-Veri Değerleri*	Tekil meta-veri değerleri, yalnızca bir kez kullanılan meta-veri değerlerini izleyen bir metriktir. Bu metrik, meta-veri değerlerinin tekrar edilmemesini gözlemleyerek olası yazım hatalarını veya tutarsızlıkları tespit etmeye yardımcı olabilir. Sadece bir kez kullanılan meta-veriler, standartların dışına çıkıldığını veya bir hata yapıldığını gösterebilir.
22	*Tekrarlayan Meta-Veri Değerleri*	Tekrarlayan meta-veri değerleri, içerikte birden fazla kez kullanılan meta-veri değerlerinin sayısını ölçen bir metriktir. Bu metrik, içerikte gereksiz bilgi tekrarlarının olup olmadığını belirlemek için kullanılır ve içerik optimizasyonu açısından fazla tekrarlamanın önlenmesine yardımcı olur.
23	*Tekrarlı İçerikler*	Tekrarlı içerikler, aynı meta-veri değerlerine sahip içeriklerin sayısını ölçen bir metriktir. Bu metrik, içeriklerin benzersizliğini değerlendirmek için kullanılır ve aynı meta-verilerle açıklanan birden fazla içerik bulunduğunda, içerik fazlalığını veya kopya içerikleri tespit etmeye yardımcı olur.
24	*Başka Bir Alanda Varlik*	Başka bir alanda varlık, bir meta-veri değerinin başka bir alanda, örneğin başlık veya özet gibi alanlarda da yer aldığı durumları izleyen bir metriktir. Bu metrik, meta-veri değerlerinin farklı alanlarda tekrarlanmasının, içeriğin daha iyi temsil edildiğini ve açıklamaların tutarlı olduğunu gösterebileceğini belirlemeye yardımcı olur.
25	*İzole Kullanım*	İzole kullanım, her zaman aynı meta-veri değerlerini kullanan editör veya yazarların sayısını ölçen bir metriktir. Bu metrik, içerik üretiminde dar veya izole odakların varlığını gösterebilir ve içerik çeşitliliğinin düşük olduğunu ya da belirli konulara sıkça tekrarlandığını işaret edebilir.
26	*Ücretsiz Erişim Gecikmesi*	Ücretsiz erişim gecikmesi, bir içeriğin yayın tarihinden ücretsiz erişim tarihine kadar geçen süreyi ölçen bir metriktir. Uzun gecikmeler (örneğin, bir yıldan fazla), bu meta-veri ayarının uygun olmadığını veya erişim politikalarının kullanıcı ihtiyaçlarına yanıt vermediğini gösterebilir. Bu metrik, içerik erişilebilirliğini iyileştirme amacıyla kullanılır.
27	*Geçersiz Ücretsiz Erişim Aralığı*	Geçersiz ücretsiz erişim aralığı, ücretsiz erişim için başlangıç tarihinin bitiş tarihinden sonra ayarlandığı durumları tespit eden bir metriktir. Bu metrik, tarihlerin hatalı veya tutarsız şekilde ayarlandığını gösterir ve erişim planlamasında düzenlemeler yapılması gerektiğini işaret eder.
28	*Kısaltılmış Ücretsiz Erişim Aralığı*	Kısaltılmış ücretsiz erişim aralığı, ücretsiz erişim için belirlenen başlangıç ve bitiş tarihleri arasındaki sürenin çok kısa olduğu durumları tespit eden bir metriktir. Bu metrik, kullanıcıların içeriğe yeterince uzun süre erişememesine yol açabilecek yetersiz erişim aralıklarını işaret eder.

29	*Aşırı Fiyat*	Aşırı fiyat, bir içerik için belirlenen aşırı yüksek veya düşük fiyatları izleyen bir metriktir. Bu metrik, içeriğin meta-verisi olarak belirtilen fiyatlandırmanın uygun olmadığını ve kötü fiyatlandırma seçimlerini gösterebilir. Fiyat dengesizlikleri, içeriğin erişilebilirliğini veya rekabet gücünü olumsuz etkileyebilir.
30	*Geçersiz Fiyat*	Geçersiz fiyat, içeriğin fiyatının negatif olduğu durumları izleyen bir metriktir. Negatif fiyatlandırma, fiyatlandırma hatasını veya geçersiz meta-veri değerini işaret eder ve içeriğin düzgün bir şekilde değerlendirilmesi için düzeltme yapılmasını gerektirir.
31	*Hiyerarşinin Derinliği*	Hiyerarşinin derinliği, bir web sitesindeki içeriklerin yer aldığı hiyerarşik yapıdaki derinlik seviyesini ölçen bir metriktir. İçeriğin çok derin seviyelerde yer alması, kullanıcıların içeriği bulmasını zorlaştırabilir ve navigasyon sorunlarına yol açabilir. Bu metrik, içeriklerin erişilebilirliğini değerlendirmek için kullanılır.

5. Sonuç

Web madenciliği, web sitelerinin izlenmesi ve yönetilmesi için gelişmiş karar destek sistemlerinin geliştirilmesinde etkin bir yöntem olarak kullanılmaktadır. Bu süreç, web verilerinin tanımlanması, ön işlenmesi, depolanması ve kalıp keşfi ile analizi gibi dört aşamayı kapsar. Bir e-haber platformunda bu süreç, meta veri kalitesini izlemek ve yönetmek amacıyla önerilen bir karar destek sistemi oluşturma potansiyeli taşır. Veri madenciliği ve iş zekâsı tekniklerinin entegrasyonu, daha etkili karar destek sistemlerinin oluşturulmasına olanak tanır.

Web madenciliği, web sitelerini yönetmek ve optimize etmek için önemli bir araçtır. Bu süreç, çevrimiçi verilerden değerli bilgilerin çıkarılması ve bu bilgilerin karar destek sistemlerine entegrasyonu ile ilgilidir. Bu yaklaşım, e-haber portallarında meta veri kalitesi, kullanıcı deneyimi ve içerik yönetimi optimizasyonu açısından önemli faydalar sağlar. Kullanıcı davranışları ve içerik performansı hakkında sağlanan ayrıntılı bilgiler, karar verme süreçlerine katkıda bulunur. Bu durum, yöneticilerin web sitesi performansını izleyip daha hızlı ve bilinçli kararlar almasına yardımcı olurken, kullanıcı deneyimini iyileştiren özelleştirme taktiklerini de destekler.

İçerik optimizasyonu açısından, web madenciliği teknikleri, özellikle anahtar kelime analizi ve ilişki kuralları kullanılarak içerik üretimi ve etiketleme süreçlerini iyileştirirken kaynak verimliliğini artırır. Bu otomatik teknolojiler, manuel analitik yöntemlerin yerini alarak firmaların zaman ve maliyet tasarrufu sağlamasına yardımcı olur.

Gelecek çalışmalar, önerilen web madenciliği sürecini daha gelişmiş karar destek sistemleri oluşturmak amacıyla farklı web sitelerine uygulamayı amaçlar.

Meta veri kalitesini izlemek ve yönetmek için geliştirilen sistemde, kalite değerlendirme sürecini iyileştirmek için ek istatistiksel ve veri madenciliği teknikleri entegrasyonu sağlanır. Kümeleme yöntemleri, yazarların meta veri kalitesine dair davranışlarını gruplamak için kullanılabilir. Bu, içerik yayınlama süreçlerine yeni bir bakış açısı kazandırır ve farklı gruplar için düzeltici önlemlerin alınmasına olanak tanır. Ayrıca, derin öğrenme ve doğal dil işleme gibi yapay zekâ teknikleri, daha karmaşık ve etkili karar destek sistemleri geliştirilmesine katkıda bulunur.

Büyük veri ve gerçek zamanlı analizlerin web madenciliği süreçlerine dahil edilmesi, sistemlerin verimliliğini ve performansını artırabilir. Sonuç olarak, web madenciliği ve karar destek sistemleri, dijital dönüşüm sürecinde giderek daha kritik bir hale gelir ve firmaların rekabet avantajı elde etmelerine, müşteri ihtiyaçlarına daha etkin yanıt vermelerine yardımcı olur. Yapay zekâ ve büyük veri teknolojilerindeki ilerlemelerle birlikte, bu sistemlerin dijital girişimler için daha önemli ve karmaşık bir rol üstlenmesi beklenmektedir.

Kaynakça

1. Amitay, E., & Paris, C. (2000). Automatically summarising web sites: Is there a way around it? *In Proceedings of the 9th International Conference on Information and Knowledge Management* (pp. 173-180). ACM Press.
2. Aye, T. T. (2011, March). Web log cleaning for mining of web usage patterns. *In 2011 3rd International Conference on Computer Research and Development* (Vol. 2, pp. 490-494). IEEE. https://doi.org/10.1109/ICCRD.2011.5764181
3. Chen, C., Yan, X., Zhu, F., Han, J., & Yu, P. S. (2009). Graph OLAP: A multi-dimensional framework for graph data analysis. *Knowledge and Information Systems*, *21*, 41–63. https://doi.org/10.1007/s10115-009-0228-9
4. da Costa, M. G., & Gong, Z. (2005, June). Web structure mining: an introduction. *In 2005 IEEE International Conference on Information Acquisition* (pp. 6-pp). https://doi.org/10.1109/ICIA.2005.1635156
5. Domingues, M. A., Soares, C., Jorge, A. M., & Rezende, S. O. (2014). A data warehouse to support web site automation. *Journal of the Brazilian Computer Society*, *20*(11), 1-16. https://doi.org/10.1186/1678-4804-20-11
6. Elbawab, R. (2024). Linking organisational learning, performance, and sustainable performance in universities: An empirical study in Europe. *Humanities and Social Sciences Communications*, *11*(626). https://doi.org/10.1057/s41599-024-03114-1
7. Fagni, T., Perego, R., Silvestri, F., & Orlando, S. (2006). Boosting the performance of web search engines: Caching and prefetching query results by exploiting historical usage data. *ACM Transactions on Information Systems (TOIS)*, *24*(1), 51-78. https://doi.org/10.1145/1125857.1125859
8. Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *IEEE Access*, *5*, 20590-20610. https://doi.org/10.1109/ACCESS.2017.2756872
9. Grill, M., & Rehák, M. (2014). Malware detection using HTTP user-agent discrepancy identification. *IEEE International Workshop on Information Forensics and Security (WIFS)*, 221. https://doi.org/10.1109/WIFS.2014.7084331

10. Grineva, M., Grinev, M., & Lizorkin, D. (2009). Extracting Key Terms From Noisy and Multi-theme Documents. *International World Wide Web Conference*. ACM Press (1526709.1526798). https://doi.org/10.1145/1526709.1526798

11. Inbarani, H. H., & Thangavel, K. (2013). Effective web personalisation based on rough biclustering. *International Journal of Granular Computing, Rough Sets and Intelligent Systems*, *3*(1), 59-84. https://doi.org/10.1504/IJGCRSIS.2013.054127

12. W. H. Inmon, *Building the Data Warehouse*. Wellesley, MA, USA: QED Information Sciences, Inc., 1992.

13. Kim, J. Y., Collins-Thompson, K., Bennett, P. N., & Dumais, S. T. (2012). Characterizing web content, user interests, and search behavior by reading level and topic. *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining (WSDM 2012)*, 213-222. https://doi.org/10.1145/2124295.2124323

14. Lee, C. H., & Fu, Y. H. (2008, November). Web usage mining based on clustering of browsing features. I*n 2008 Eighth International Conference on Intelligent Systems Design and Applications*, *1*, 281-286). IEEE. https://doi.org/10.1109/ISDA.2008.185

15. Malinowski, E., & Zimanyi, E. (2008). Advanced data warehouse design. Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications. *Data-centric Systems and Applications, Volume, 978-3-540-74404-7, Springer, Berlin Heidelberg (2008)*, p. 1.

16. Marquardt, C. G., Becker, K., & Ruiz, D. D. (2004, July). A pre-processing tool for web usage mining in the distance education domain. *In Proceedings. International Database Engineering and Applications Symposium, 2004. IDEAS'04.* (pp. 78-87). IEEE.

17. Mobasher, B., Jain, N., Han, E. H., & Srivastava, J. (1996). Web mining: Pattern discovery from world wide web transactions (pp. 558-567). *Technical Report TR 96-050, University of Minnesota, Dept. of Computer Science, Minneapolis.*

18. Mukherjee, R., & Kar, P. (2017). A comparative review of data warehousing ETL tools with new trends and industry insight. *2017 IEEE 7th International Advance Computing Conference (IACC)*, 944-949. https://doi.org/10.1109/IACC.2017.0192

19. Raux, C., Ma, T.-Y., & Cornelis, E. (2016). Variability in daily activity-travel patterns: The case of a one-week travel diary. *European Transport Research Review*, *8*(26). https://doi.org/10.1007/s12544-016-0213-9

20. Rudniy, A. (2022). Data Warehouse Design for Big Data in Academia. *Computers, Materials & Continua*, *71*(1), 980-984. https://doi.org/10.32604/cmc.2022.016676

21. Saini, S., & Pandey, H. M. (2015). Review on web content mining techniques. *International Journal of Computer Applications, 118*(18), 33-36.

22. Shanmugasundaram, J., Tufte, K., He, G., Zhang, C., DeWitt, D., & Naughton, J. (1999). Relational databases for querying XML documents: Limitations and opportunities. *Proceedings of the 25th VLDB Conference*, 65-73.

23. Sharma, P., Yadav, S., & Bohra, B. (2015, October). A review study of server log formats for efficient web mining. *In 2015 International Conference on Green Computing and Internet of Things (ICGCIoT)* (pp. 1373-1377). IEEE. https://doi.org/10.1109/ICGCIoT.2015.7380681

24. Shen, Z., Wei, J., Sundaresan, N., & Ma, K. L. (2012, October). Visual analysis of massive web session data. *In IEEE symposium on large data analysis and visualization (LDAV)* (pp. 65-72). https://doi.org/10.1109/LDAV.2012.6378977

25. Sokolov, I., & Turkin, I. (2018). Resource efficient data warehouse optimization. *The 9th IEEE International Conference on Dependable Systems, Services and Technologies, DESSERT 2018*. https://doi.org/10.1109/DESSERT.2018.8409183

26. Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *SIGKDD Explorations*, *1*(2), 12-23. https://doi.org/10.1145/846183.846188

27. Srivastava, M., Srivastava, A. K., Garg, R., & Mishra, P. K. (2022). Performance evaluation of the mapreduce-based parallel data preprocessing algorithm in web usage mining with robot detection approaches. *IETE Technical Review*, *39*(4), 865-879. https://doi.org/10.1080/02564602.2021.1918584

28. Toyoda, M., & Kitsuregawa, M. (2001, September). Creating a web community chart for navigating related communities. *In Proceedings of the 12th ACM Conference on Hypertext and Hypermedia* (pp. 103-112). https://doi.org/10.1145/504216.504244

29. Vijiyarani, S., & Suganya, E. (2015). Web content mining techniques: A survey. *International Journal of Computer-Aided Technologies*, 2(3), 55-60.

30. Visser, E., & Weideman, M. (2014). Fusing website usability and search engine optimisation: A multi-method approach. *South African Journal of Information Management*, *16*(1), 1-8. https://hdl.handle.net/10520/EJC151536

31. Zaiane, O. R., Xin, M., & Han, J. (1998, April). Discovering web access patterns and trends by applying OLAP and data mining technology on web logs. *In Proceedings IEEE International Forum on Research and Technology Advances in Digital Libraries-ADL'98-* (pp. 19-29). https://doi.org/10.1109/ADL.1998.670376

32. Zou, Q., Chu, W., Johnson, D., & Chiu, H. (2002). A pattern decomposition algorithm for data mining of frequent patterns. *Knowledge and Information Systems*, *4*(4), 466–482. https://doi.org/10.1007/s101150200016

4. Bölüm

Veri Bilimi ve Analitiği

Data Science and Analytics

Üzeyir FİDAN[1]

[1] Dr., Uşak Üniversitesi, Uzaktan Eğitim Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, ORCID: 0000-0003-3451-4344, uzeyir.fidan@usak.edu.tr

1. Giriş

Veri bilimi, günümüz dijital dünyasında giderek daha büyük öneme sahip olan bir disiplin haline gelmiştir. Teknolojinin hızla gelişmesi ve verinin her alanda erişilebilir hale gelmesi, veri biliminin organizasyonlar için stratejik bir unsur olarak yükselmesine neden olmaktadır. Veri bilimi yalnızca bir analiz yöntemi değil, aynı zamanda organizasyonel karar alma süreçlerini dönüştüren ve rekabet avantajı sağlayan bir disiplindir (Provost & Fawcett, 2013). Bu bölümde, veri biliminin tanımı ve önemi, veri analitiğinin temel prensipleri ve veri biliminin dijital dönüşümdeki rolü ele alınacaktır.

1.1. Veri Biliminin Tanımı ve Önemi

Veri bilimi, farklı kaynaklardan toplanan büyük miktarda verinin anlamlandırılması, bu verilerden değer yaratılması ve stratejik kararlar için kullanılmasına olanak tanıyan çok disiplinli bir alandır. Veri bilimi; matematik, istatistik, bilgisayar bilimi ve iş zekâsı gibi birçok disiplini bir araya getirmektedir. Bu çok disiplinli yapısı sayesinde karmaşık problemleri çözme, eğilimleri öngörme ve organizasyonların performansını iyileştirme konusunda çok önemli bir araç haline gelmiş; araştırmacılar, veri bilimi için çok sayıda tanım geliştirmiştir.

Hayashi (1998), veri bilimini disiplinler arası bir yaklaşım olarak ele almış ve matematik, istatistik, bilgisayar bilimleri ile alan bilgisi gibi çeşitli disiplinlerden gelen yöntemlerle veri analizini kapsayan ve içgörüler elde etmeye yönelik bir yaklaşım olarak tanımlamıştır. Cleveland (2001) ise veri bilimini, büyük ve karmaşık veri setlerini işleyerek faydalı bilgiye dönüştüren bilimsel yöntemlerin kullanıldığı bir süreç olarak nitelendirmiştir. Dhar (2013), çalışmasında bilimsel hesaplamayı vurgulayarak veri bilimini, verilerden anlamlı bilgi elde etmek için bilgi keşfi, makine öğrenimi ve bilimsel hesaplama yöntemlerini bir araya getiren bir disiplin olarak tanımlamıştır. Provost ve Fawcett (2013), veri bilimini istatistiksel yöntemler ve hesaplamalı algoritmalar kullanarak verilerden anlamlı sonuçlar çıkarma disiplini olarak ifade etmişlerdir. Waller ve Fawcett (2013) ise veri biliminin stratejik karar desteği sağlama niteliğine dikkat çekerek, işletmelerin stratejik kararlar almasına yardımcı olmak için veri analitiği ve veri madenciliği yöntemlerini kullanan bir alan olarak tanımlamışlardır.

Bu tanımlar, veri biliminin farklı boyutlarını ve uygulama alanlarını öne çıkarmakta olup her biri veri biliminin disiplinler arası yapısını ve gelişimini farklı perspektiflerden ele almaktadır. Tanımlar arasındaki çeşitlilik, veri biliminin çok disiplinli yapısını yansıtmakta, alanın geniş kapsamını ve dinamik doğasını ortaya koymaktadır. Bu tanımlara dayanarak, veri bilimi; matematik,

istatistik, bilgisayar bilimleri ve alan bilgisi gibi disiplinlerden yararlanarak, büyük ve karmaşık veri setlerini işleyip stratejik ve operasyonel karar alma süreçlerini destekleyecek anlamlı bilgiler elde etmeye yönelik bilimsel yöntemler ile hesaplamalı algoritmaların kullanıldığı çok disiplinli bir alan olarak tanımlanabilir.

Veri biliminin önemi, günümüz iş dünyasında operasyonel verimliliği artırma, müşteri davranışlarını derinlemesine analiz etme ve yenilikçi iş modelleri geliştirme gibi kritik alanlarda kendini göstermektedir. Büyük veri, yapay zekâ ve makine öğrenimi teknolojilerinin hızlı evrimi, veri biliminin stratejik karar alma süreçlerinde vazgeçilmez bir unsur haline gelmesini sağlamaktadır (Li, Chen & Shang, 2022). Bu teknolojiler sayesinde kuruluşlar, devasa miktardaki yapılandırılmış ve yapılandırılmamış veri setlerinden karmaşık analizler gerçekleştirerek daha akıllı, daha öngörüye dayalı ve daha hızlı kararlar alabilmektedir (Bloom, Rallapalli, Rosen & Schlenker, 2012). Veri biliminin bu yetenekleri, sadece operasyonel kararların iyileştirilmesinde değil, aynı zamanda inovasyonu tetikleyerek işletmelerin sürdürülebilir bir rekabet avantajı elde etmelerinde de merkezi bir rol oynamaktadır. Dahası, veri bilimi, sağlık hizmetlerinden finansa, eğitimden kamu sektörüne kadar çok geniş bir yelpazede dönüştürücü bir güç olarak kullanılmakta; hasta bakım kalitesini artırmak (Wang, Kung, Gupta & Ozdemir, 2019), mali riskleri daha iyi yönetmek (Morales, Gray & Rajmil, 2022), eğitim süreçlerini kişiselleştirmek (Brooks, Quintana, Choi, Quintana, NeCamp & Gardner, 2021) ve kamu hizmetlerinde verimliliği optimize etmek gibi çeşitli uygulama alanlarında büyük başarılar elde etmektedir. Bu bağlamda, veri bilimi sadece teknolojik bir araç değil; aynı zamanda farklı sektörlerdeki organizasyonların stratejik vizyonunu şekillendiren oldukça önemli bir faktör haline gelmiştir.

1.2. Veri Analitiğinin Temelleri

Veri biliminin önemli bir alt alanı olan veri analitiği, veriyi anlamak ve stratejik kararlar almak için kullanılan teknikler ve süreçleri içermektedir. Veri analitiği, organizasyonların geçmiş verileri analiz ederek mevcut eğilimleri anlamalarına, gelecekteki olasılıkları öngörmelerine ve en iyi eylem planlarını belirlemelerine yardımcı olmaktadır. Veri analitiği, genellikle dört ana kategoriye ayrılır; tanımlayıcı, kestirimci, tanılayıcı ve normatif analitik (Shi-Nash & Hardoon, 2017).

Tanımlayıcı analitik (Descriptive Analytics), mevcut verilerden durum değerlendirmesi yaparak "ne oldu?" sorusuna yanıt vermeye çalışmaktadır. Kestirimci analitik (Predictive Analytics) ise geçmiş verilere dayanarak "ne olabilir?" sorusuna odaklanmakta ve gelecekteki olasılıkları tahmin etmeye

çalışmaktadır. Tanılayıcı analitiği (Diagnostic Analytics), meydana gelen olayların arkasındaki nedenleri analiz ederken, normatif analitik (Prescriptive Analytics) "ne yapılmalı?" sorusunu yanıtlayarak en iyi çözüm yollarını sunmaktadır. Bu temel kategoriler, organizasyonların karar alma süreçlerini daha sağlam temellere oturtmalarına yardımcı olur ve stratejik avantaj sağlamaktadır.

Veri analitiğinin önemi, yalnızca karar alma süreçlerini hızlandırmakla kalmaz; aynı zamanda kaynakların daha verimli kullanılmasını, risklerin daha doğru bir şekilde yönetilmesini ve yenilikçi fırsatların değerlendirilmesini sağlar. Bu nedenle veri analitiği hem stratejik hem de operasyonel düzeyde kararları destekleyen bir yapı olarak modern işletmelerde vazgeçilemez bir role sahiptir.

1.3. Veri Biliminin Dijital Dönüşümdeki Rolü

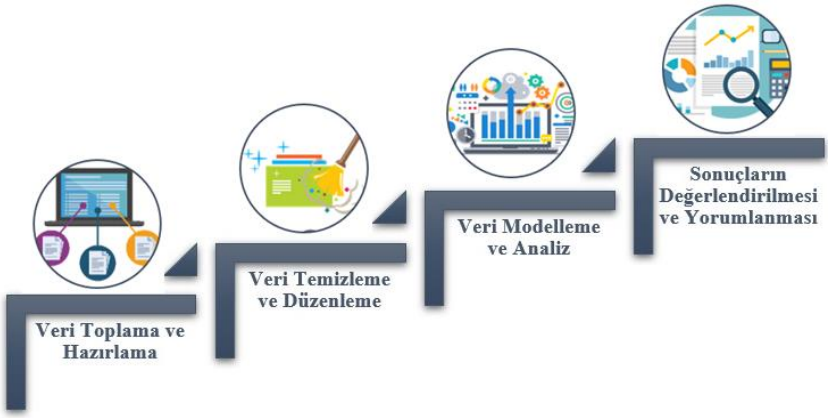
Dijital dönüşüm, organizasyonların iş yapış biçimlerini kökten değiştiren bir süreçtir ve veri bilimi bu dönüşümün merkezinde yer almaktadır. Dijital dönüşümün temelini oluşturan veri bilimi, organizasyonların veriye dayalı kararlar almasını sağlayarak iş modellerinin dijitalleşmesini hızlandırmaktadır (Härting, Reichstein & Schad, 2018). Veri bilimi, büyük veriyi anlamlandırarak şirketlerin stratejik hedeflerini daha hızlı ve doğru bir şekilde gerçekleştirmelerine olanak tanımaktadır.

Dijital dönüşüm, şirketlerin sadece teknoloji kullanarak operasyonel süreçleri iyileştirmesi anlamına gelmemekte; aynı zamanda müşteri etkileşimlerinden ürün geliştirmeye kadar her aşamada veri odaklı bir yaklaşım benimsemesi anlamına gelmektedir. Bu noktada veri bilimi, organizasyonların dijital stratejilerini şekillendirmede ve iş süreçlerini optimize etmede vazgeçilmez bir araçtır. Veri bilimi sayesinde dijital dönüşüm sürecindeki şirketler, müşteri davranışlarını daha derinlemesine anlayabilir, pazarlama stratejilerini daha etkili bir şekilde hedefleyebilir ve operasyonel süreçlerinde daha fazla verimlilik elde edebilirler (Adeniran, Efunniyi, Osundare & Abhulimen, 2024). Ayrıca veri bilimi, dijitalleşme sürecinde meydana gelebilecek riskleri önceden tespit edip bunlara karşı önlem alınmasını da sağlayarak, organizasyonların dijital dünyada sürdürülebilir bir başarı elde etmelerine yardımcı olmaktadır.

2. Veri Bilimi Süreci

Veri bilimi süreci, bir organizasyonun sahip olduğu verilerin değer yaratacak şekilde işlenmesi için takip edilmesi gereken aşamaları içermektedir (Grover, Chiang, Liang & Zhang, 2018). Bu süreç, ham verinin toplanmasından başlayarak verinin temizlenmesi, modellenmesi ve nihai olarak sonuçların değerlendirilip yorumlanmasına kadar uzanmaktadır. Bu aşamalar, veri bilimi projelerinin başarılı bir şekilde yürütülmesini ve elde edilen bulguların stratejik amaçlar

doğrultusunda kullanılmasını sağlamaktadır. Veri bilimi süreci dört temel adımdan oluşur: veri toplama ve hazırlama, veri temizleme ve düzenleme, veri modelleme ve analiz, sonuçların değerlendirilmesi ve yorumlanması (Şekil 1).



Şekil 1. Veri bilimi süreçleri

2.1. Veri Toplama ve Hazırlama

Veri toplama ve hazırlama, veri bilimi sürecinin ilk ve en önemli adımıdır. Bu adımda, projede kullanılacak veriler, çeşitli kaynaklardan toplanır ve işlenmeye hazır hale getirilmektedir. Veri kaynakları genellikle iki ana kategoriye ayrılır: yapılandırılmış ve yapılandırılmamış veri. Yapılandırılmış veri veritabanları ve elektronik tablolar gibi düzenli bir formatta bulunan veriler iken; yapılandırılmamış veri metin, görüntü, video ve sosyal medya paylaşımları gibi daha karmaşık yapıda olan verileri ifade eder (Mishra & Misra, 2017).

Veri toplama sürecinde kullanılan araçlar ve yöntemler, projenin amacına ve analiz edilecek veri türüne göre değişiklik göstermektedir. Örneğin, bir web sitesinin kullanıcı davranışlarını analiz etmek için tarayıcı çerezleri veya günlük dosyaları toplanabilirken, sosyal medya analizi için API'ler aracılığıyla veri çekilebilmektedir. Verinin doğru bir şekilde toplanması, tüm veri bilimi sürecinin başarısını doğrudan etkilediği için bu aşamada kullanılan yöntemlerin güvenilir ve doğrulanabilir olması büyük önem taşımaktadır.

Veri hazırlama ise toplanan verilerin analiz için uygun hale getirilmesini içermektedir. Bu adımda eksik veriler tamamlanır, tutarsızlıklar giderilir ve veriler doğru formatlara dönüştürülür. Veri hazırlama süreci, veri bilimi projelerinde zamanın büyük bir kısmını kapsar, çünkü veriler genellikle doğrudan analiz edilemez durumda bulunmaktadır.

2.2. Veri Temizleme ve Düzenleme

Veri temizleme ve düzenleme, veri bilimi projelerinde yüksek kaliteyi güvence altına almak için temel bir rol üstlenmektedir. Bu aşamada, toplanan verilerdeki hatalar, eksik değerler, tutarsızlıklar ve yanlılıklar tespit edilip düzeltilmelidir. Veri temizleme, verideki yanlış veya hatalı bilgilerin düzeltilmesini sağlarken; veri düzenleme, verilerin doğru şekilde yapılandırılması ve analiz edilebilir formata getirilmesini amaçlamaktadır. Temizleme işlemi verilerin geçerliliği, doğruluğu ve tutarlılığını artırarak modelleme aşamasının daha verimli olmasına yardımcı olmaktadır (Gudivada, Apon & Ding, 2017). Veri temizleme sürecinde eksik veri doldurma, hatalı veri düzeltme, aykırı değerlerin tespiti ve filtrelenmesi gibi yöntemler kullanılmaktadır. Eksik veriler, projeye göre çeşitli yöntemlerle doldurulabilir; ortalama, medyan gibi istatistiksel ölçüler kullanılabilmekte ya da eksik veriler çıkarılabilmektedir. Tutarsız verilerin tespiti için ise çoğu zaman veriler arası çapraz kontroller yapılarak veri doğruluğu sağlanmaktadır.

Bu adım, veri bilimcilerinin en fazla zaman ayırması gereken aşamalardan biridir çünkü hatalı verilerle yapılan analizler yanıltıcı sonuçlar doğuracaktır. Dolayısıyla, temizleme ve düzenleme işlemi titizlikle yapılmalıdır ve bu aşama, analizin doğruluğu açısından hayati önem taşımaktadır.

2.3. Veri Modelleme ve Analiz

Veri modelleme ve analiz aşaması, veri bilimi sürecinin belki de en önemli kısmıdır. Bu aşamada, temizlenmiş ve düzenlenmiş veriler kullanılarak modeller oluşturularak bu modeller üzerinden verinin anlamlı hale getirilmesi sağlanmaktadır. Veri modelleme, temel olarak belirli bir problemi çözmek için veri içindeki örüntülerin ve ilişkilerin keşfedilmesi üzerine kurulu bir süreçtir (Diba, Batoulis, Weidlich & Weske, 2020). Bu süreçte, farklı makine öğrenimi algoritmaları, istatistiksel analiz yöntemleri ve derin öğrenme teknikleri kullanılmaktadır.

Modelleme sürecinde, veri bilimcileri genellikle verinin özelliklerini çıkartarak belirli algoritmalara uygun hale getirmektedir. Regresyon analizi, sınıflandırma, kümeleme gibi yöntemler sıklıkla kullanılan yöntemlerdir. Bu analiz yöntemlerinin seçimi, projenin amacına ve çözülmek istenen probleme bağlı olarak değişkenlik göstermektedir. Örneğin; bir satış tahmini yapılacaksa regresyon analizine başvurulurken, müşterilerin segmentlere ayrılması gereken durumlarda kümeleme algoritmaları kullanılmaktadır.

Modelleme aşamasında elde edilen bulgular, genellikle organizasyonların stratejik karar alma süreçlerini yönlendiren önemli içgörüler sağlamaktadır. Modelin doğruluğunu ölçmek için çapraz doğrulama gibi teknikler kullanılır ve

modelin güvenilirliği test edilir. Bu aşamada, modelin performansını artırmak için hiperparametre optimizasyonu gibi çeşitli ince ayarlamalar da yapılmaktadır (Khalid & Javaid, 2020).

2.4. Sonuçların Değerlendirilmesi ve Yorumlanması

Sonuçların değerlendirilmesi ve yorumlanması aşaması, veri bilimi sürecinin son aşamasını oluşturur ve elde edilen model çıktılarının anlamlı hale getirilmesi bu aşamada gerçekleştirilir. Modelleme sürecinde elde edilen sonuçların sadece sayısal değerlerden ibaret olmaması, aynı zamanda iş dünyasında uygulanabilir ve stratejik bir içgörü sunması gereklidir. Bu nedenle sonuçların doğru bir şekilde yorumlanması, veri bilimi sürecinin en temel ve vazgeçilemez adımlarından biri olarak kabul edilmektedir (Larson & Chang, 2016).

Sonuçların değerlendirilmesi sırasında modellerin doğruluk oranları ve hata payları analiz edilmektedir. Modelin performansına bağlı olarak elde edilen sonuçların gerçek dünya uygulamaları üzerindeki etkisi tartışılmaktadır. Örneğin; bir müşteri segmentasyon modeli sonucunda belirlenen müşteri gruplarına yönelik pazarlama stratejileri geliştirilebilir. Benzer şekilde, tahmine dayalı modeller sayesinde gelecekteki satışlar öngörülerek üretim ve stok yönetimi planlaması yapılabilmektedir. Sonuçların yorumlanması aşamasında, elde edilen bulguların organizasyonun mevcut stratejileriyle uyumlu olup olmadığı değerlendirilmekte gerekirse yeni stratejik adımlar önerilmektedir. Veri bilimi sonuçlarının işletmeye somut katkılar sunabilmesi için elde edilen içgörülerin uygulanabilir ve pratik hale getirilmesi hayati bir öneme sahiptir. Ayrıca, modelin performansı, organizasyonel beklentileri karşılama düzeyi ve gelecekteki iyileştirme fırsatları bu aşamada detaylı olarak incelenmektedir.

3. Veri Gizliliği ve Etik Boyutlar

Veri bilimi, büyük miktarda veriyi analiz etme ve bu verilerden değer yaratma amacı taşırken, aynı zamanda ciddi etik ve gizlilik sorunlarını da beraberinde getirmektedir. Büyük veri çağında, özellikle kişisel verilerin toplanması, işlenmesi ve saklanması, veri gizliliği ve güvenliği açısından birçok soruyu gündeme getirmiştir. Verilerin güvenli bir şekilde korunması ve etik kurallar çerçevesinde kullanılması, veri bilimcilerin ve organizasyonların sorumluluğundadır (Saltz & Dewar, 2019). Bu bölümde, veri gizliliği ve güvenliği ile veri biliminde karşılaşılan etik zorluklar ele alınmaktadır.

3.1. Veri Gizliliği ve Güvenliği

Veri gizliliği ve güvenliği, dijital çağda işletmelerin karşılaştığı en kritik meselelerden biri haline gelmiştir. Özellikle kişisel verilerin korunması hem

bireylerin haklarının korunması açısından hem de yasal yükümlülükler açısından büyük önem taşımaktadır (Politou, Alepis & Patsakis, 2018). Verilerin yanlış kişilerin eline geçmesi veya kötü niyetli kişiler tarafından kullanılması, bireylerin mahremiyetini ihlal edebilmekte ve işletmelere ciddi zararlar verebilmektedir.

Veri gizliliği, bireylerin kişisel bilgilerinin kim tarafından nasıl kullanılabileceğine dair kontrol sahibi olmalarını sağlamaktadır. Veri biliminde ise bu bilgiler, analitik süreçler için oldukça önemli bir role sahip olabilir. Ancak yasal ve etik düzenlemeler bu bilgilerin yalnızca uygun ve izne dayalı bir şekilde kullanılmasına izin vermektedir. Özellikle Avrupa Birliği'nin yürürlüğe koyduğu Genel Veri Koruma Yönetmeliği (GDPR, 2016), Kişisel Verilerin Korunması Kanunu (KVKK, 2016) ve benzeri yasal düzenlemeler, veri gizliliği konusuna sıkı kurallar getirmiştir. Bu tür düzenlemeler, veri bilimcilerin veri toplama, işleme ve saklama süreçlerinde uymaları gereken kuralları belirlemektedir.

Veri güvenliği ise toplanan verilerin yetkisiz erişimlerden, kayıplardan veya kötü amaçlı saldırılardan korunmasını amaçlamaktadır (Perwej, Abbas, Dixit, Akhtar & Jaiswal, 2021). Güvenlik önlemleri, veri biliminde özellikle hassas veya kişisel verilerin işlendiği durumlarda büyük önem taşımaktadır (Fidan & Boza, 2024). Şifreleme, veri anonimleştirme, erişim kontrolü ve veri maskeleme gibi güvenlik yöntemleri, veri bilimi projelerinde yaygın olarak kullanılan uygulamalardır (Domingo-Ferrer, Farras, Ribes-González & Sánchez, 2019). Bu yöntemler, veri setindeki kişisel veya hassas bilgilerin korunmasına yardımcı olurken aynı zamanda bu verilerin analiz edilmesine de olanak tanımaktadır.

Veri güvenliği ihlalleri yalnızca bireylerin mahremiyetini riske atmakla kalmaz; aynı zamanda organizasyonların itibarına ve finansal yapısına da zarar verebilmektedir. Bu nedenle, veri biliminde güvenli veri yönetimi stratejilerinin uygulanması ve sürekli olarak güncellenmesi hem teknik hem de yönetsel açıdan büyük bir öneme sahiptir.

3.2. Veri Biliminde Etik Zorluklar

Veri bilimi, güçlü analitik araçlar ve büyük veri setleri aracılığıyla organizasyonlara benzersiz fırsatlar sunarken aynı zamanda önemli etik zorluklarla da karşı karşıya kalmaktadır. Bu zorluklar, verilerin nasıl toplandığı, nasıl kullanıldığı ve verilerden elde edilen sonuçların nasıl yorumlandığı konularında ortaya çıkmaktadır (Brady, 2019). Veri bilimcilerin karşılaştığı en büyük etik zorluklar arasında veri önyargıları, algoritmik adalet ve şeffaflık yer almaktadır (Saltz & Dewar, 2019).

En önemli etik sorunların başında veri önyargısı gelmektedir. Veri önyargısı, verilerde temsil kabiliyetindeki bozulmalara neden olan tehlike olarak tanımlanmaktadır (Olteanu, Castillo, Diaz & Kıcıman, 2019). Toplanan veri

setleri genellikle belirli grupları dışlayabilmekte veya yanlış temsiller içerebilmektedir. Bu önyargılar, analiz sonuçlarını doğrudan etkileyerek adil olmayan kararların alınmasına yol açabilmektedir. Örneğin, cinsiyet veya ırk temelli önyargılar, bir algoritmanın belirli bir gruba karşı adaletsiz kararlar vermesine neden olabilmektedir. Veri bilimcilerin, önyargıların farkında olmaları ve bu önyargıları tespit etmek için gerekli önlemleri almaları gerekmektedir. Bununla birlikte, algoritmaların eğitildiği verilerdeki toplumsal adaletsizliklerin algoritmalara yansımaması için veri toplama süreçlerinin dikkatlice yapılandırılması büyük önem arz etmektedir.

Diğer önemli etik zorluk ise algoritmik adalet konusudur (Birhane, 2021). Algoritmalar, veri biliminin temel bileşenlerinden biri haline gelmiş olsa da bu algoritmaların doğru ve adil sonuçlar verdiğinden emin olmak her zaman kolay değildir (Mitchell, Potash, Barocas, D'Amour & Lum, 2021). Özellikle kredi verme, iş başvuruları veya ceza adaleti gibi kararların algoritmalara dayandırıldığı sistemlerde, algoritmaların adaletli ve şeffaf olması gerekmektedir. Algoritmaların nasıl çalıştığının açıklanabilir olması ve kararlarının nedenlerini doğru bir şekilde açıklayabilmesi, algoritmik şeffaflığın bir parçasıdır.

Son olarak, veri biliminde etik zorluklar arasında şeffaflık konusu önemli bir yer tutmaktadır. Veri bilimciler, kullandıkları yöntemler, topladıkları veriler ve bu verilerden elde ettikleri sonuçlar konusunda açık ve şeffaf olmalıdırlar (Wagenmakers vd., 2021).

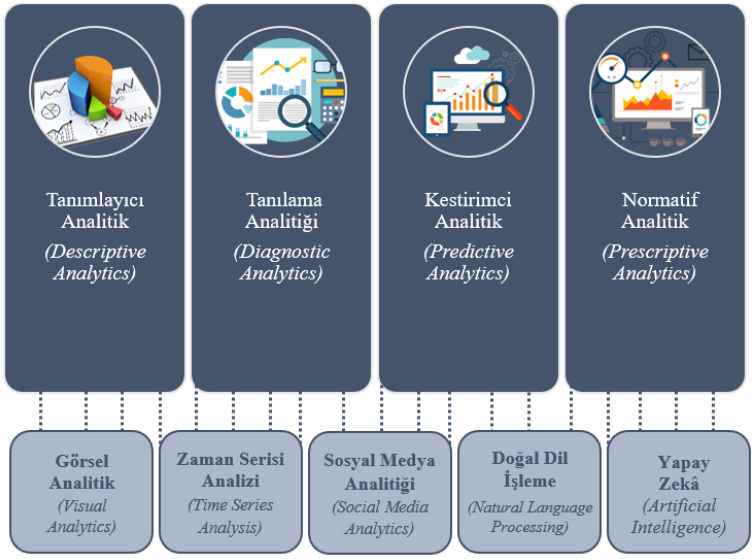
Etik veri kullanımı, veri bilimcilerin sonuçları topluma sunarken şeffaf olmalarını ve veri toplama süreçlerini, kullanılan algoritmaları ve analiz yöntemlerini açık bir şekilde paylaşmalarını gerektirmektedir. Şeffaflık eksikliği, yalnızca yanlış kararlar alınmasına yol açmakla kalmayacak; aynı zamanda kamuoyunun güvenini de önemli ölçüde zedeleyecektir.

Bu zorlukların üstesinden gelmek için veri bilimcilerin sadece teknik becerilere değil; aynı zamanda etik ilkeler ve sosyal sorumluluk anlayışına da sahip olmaları gerekmektedir. Etik veri bilimi, teknik doğruluğun yanı sıra topluma ve bireylere karşı adil, şeffaf ve sorumlu bir yaklaşımı benimsemektedir.

4. Veri Analitiği Türleri ve Teknikleri

Veri analitiği, organizasyonların veriyi anlamlandırma ve karar alma süreçlerinde etkin bir rol oynamaktadır. Veri analitiği, geçmiş verilerden içgörüler elde etme, mevcut durumu analiz etme ve gelecekteki sonuçları tahmin etme gibi işlevlerle stratejik karar alma süreçlerine katkıda bulunmaktadır (Sarker, 2021). Veri analitiği farklı yaklaşımlara göre sınıflandırılabilir ve her yaklaşımın farklı kullanım alanları vardır. Bu bölümde tanımlayıcı, tanılama,

kestirimci, normatif analitik türleri ve gelişmekte olan analitik yaklaşımlar ele alınacaktır (Şekil 2).



Şekil 2. Analitik türleri ve teknikleri

4.1. Tanımlayıcı Analitik (Descriptive Analytics)

Tanımlayıcı analitik, mevcut veya geçmiş verilerden anlamlı bilgi elde etmeye yönelik bir süreçtir ve "ne oldu?" sorusuna yanıt vermektedir (Roy, Srivastava, Jat & Karaca, 2022). Bu analitik türü, verinin betimlenmesi ve özetlenmesi amacıyla kullanılmaktadır (Raghupathi & Raghupathi, 2018). Genelde işletmelerin operasyonlarını ve süreçlerini analiz etmek için tercih edilmektedir. Tanımlayıcı analitik, veri biliminde ilk adım olarak değerlendirilmekte ve karar vericilere durum tespiti yapma olanağı sağlamaktadır (Roy vd., 2022).

İstatistiksel ölçüler, ortalamalar, yüzdeler ve grafiksel gösterimler tanımlayıcı analitiğin temel unsurlarındandır. Örneğin, bir şirketin aylık satış verilerini analiz ederek belirli bir dönemdeki satış performansını incelemek, tanımlayıcı analitikle mümkündür. Bu analitik türü, büyük veri setlerindeki örüntüleri belirleyerek anlamlı bilgiler sunmakta ancak neden-sonuç ilişkilerini detaylandırmamaktadır.

4.2. Tanılama Analitiği (Diagnostic Analytics)

Tanılama analitiği, "neden oldu?" sorusunu yanıtlayan ve tanımlayıcı analitik tarafından sağlanan verilerin arkasındaki nedenleri keşfetmeye odaklanan bir yaklaşımdır (Wolniak & Grebski, 2023). Tanılama analitiği, geçmişte meydana gelen olayların veya eğilimlerin arkasındaki sebepleri anlamak amacıyla kullanılmaktadır. Bu süreçte veri bilimciler, nedensel ilişkiler ve veriler

arasındaki bağlantılar üzerinde derinlemesine analiz yapmaktadırlar. Bu yönüyle tanımlayıcı analitiği tamamlamaktadır (Delen & Ram, 2018).

Regresyon analizi, korelasyon analizi ve aykırı değerlerin tespiti gibi yöntemler tanılama analitiğinde sıklıkla kullanılan tekniklerdir. Örneğin, bir işletme müşteri kaybında artış gözlemler ve bu kaybın nedenini anlamak için tanılama analitiği kullanabilir. Bu analitik türü, belirli bir olayın neden gerçekleştiğini ortaya koyarak, organizasyonların gelecekte benzer olayların oluşumunu engellemek için proaktif önlemler almasını sağlamaktadır.

4.3. Kestirimci Analitik (Predictive Analytics)

Kestirimci analitik, "ne olabilir?" sorusuna yanıt aramakta ve geçmiş verileri analiz ederek gelecekte olası sonuçları tahmin etmeyi amaçlamaktadır (Kumar & Garg, 2018). Bu analitik yaklaşımı, genellikle işletmelerin gelecekteki riskleri öngörmek, fırsatları belirlemek ve stratejik kararlar almak için kullanılmaktadır. Kestirimci analitik, veri madenciliği, istatistiksel modelleme, makine öğrenimi ve yapay zekâ algoritmaları gibi ileri düzey teknikler kullanılarak gerçekleştirilmektedir.

Örneğin, bir perakende şirketi, müşterilerinin alışveriş alışkanlıklarına dayanarak gelecekte hangi ürünlerin daha fazla talep göreceğini tahmin edebilir. Kestirimci analitik, organizasyonların potansiyel sonuçları önceden görmelerine ve buna göre strateji geliştirmelerine olanak tanımaktadır. Bu yaklaşım, rekabet avantajı sağlamak ve işletme verimliliğini artırmak için önemli bir yapı taşı olarak öne çıkmaktadır.

4.4. Normatif Analitik (Prescriptive Analytics)

Normatif analitik, "ne yapılmalı?" sorusuna yanıt verir ve en uygun eylem planlarını belirlemeyi amaçlar. Bu analitik türü, yalnızca gelecekte ne olacağını tahmin etmekle kalmaz, aynı zamanda bu olasılıkları en iyi şekilde yönetmek için hangi adımların atılması gerektiğini önermektedir (Lepenioti, Bousdekis, Apostolou & Mentzas, 2020). Normatif analitik, karar vericilere karmaşık süreçlerde en iyi eylem seçeneklerini sunmakta ve optimum sonuçlar elde edilmesini sağlamaktadır.

Bu süreç, genellikle optimizasyon algoritmaları, simülasyonlar ve karar teorisi gibi yöntemler üzerine kuruludur. Örneğin, bir lojistik şirketi, taşıma maliyetlerini en aza indirmek için normatif analitiği kullanarak en uygun rota planlaması yapabilir. Normatif analitik, karar alma süreçlerinde yöneticilere daha fazla bilgi ve rehberlik sağlayarak, operasyonel verimliliği artırarak stratejik kararları desteklemektedir.

4.5. Gelişmekte Olan Analitik Yaklaşımlar

Geleneksel analitik yöntemlerin yanı sıra veri biliminde hızla gelişen yeni yaklaşımlar da bulunmaktadır. Gelişmekte olan bu analitik yaklaşımlar, giderek daha karmaşık veri setlerini analiz etmek ve daha derin içgörüler elde etmek için kullanılmaktadır. Aşağıda bu yeni yaklaşımlar sıralanmıştır:

4.5.1. Görsel analitik (Visual analytics)

Görsel analitik, karmaşık veri setlerini anlamak için veri görselleştirme teknikleriyle birlikte geleneksel analitik yöntemlerin bir kombinasyonunu sunmaktadır. Bu yaklaşım, verilerin grafiksel olarak sunulmasıyla karar vericilerin veriyi daha hızlı ve daha etkili bir şekilde anlamalarını sağlamaktadır. Görsel analitik, özellikle büyük veri setlerinde örüntülerin, ilişkilerin ve anomalilerin hızlıca tespit edilmesinde kritik bir rol oynamaktadır (Afzal, Ghani, Hittawe, Rashid, Knio, Hadwiger & Hoteit, 2023).

4.5.2. Zaman serisi analizi (Time series analysis)

Zaman serisi analizi, belirli bir zaman aralığı boyunca toplanan verilerin analiz edilmesini içermekte ve özellikle gelecekteki eğilimleri tahmin etmek için kullanılmaktadır. Ekonomik veriler, hava durumu tahminleri veya stok piyasası hareketleri gibi birçok alanda zaman serisi analizi yaygın olarak kullanılmaktadır. Bu analiz yöntemi, geçmiş verilerdeki eğilimlere ve döngülere dayanarak gelecekteki hareketleri öngörmeyi sağlamaktadır (Moraffah vd., 2021).

4.5.3. Sosyal medya analitiği (Social media analytics)

Sosyal medya analitiği, sosyal medya platformlarından toplanan verileri analiz ederek, kullanıcı davranışlarını ve etkileşimlerini anlamayı amaçlamaktadır. Bu yaklaşım, markaların müşteri geri bildirimlerini analiz etmesine, trendleri takip etmesine ve pazarlama stratejilerini optimize etmesine olanak tanır. Sosyal medya verilerinin hacmi ve çeşitliliği nedeniyle, bu analiz yöntemi büyük veri analitiği teknikleriyle birlikte uygulanmaktadır (Abkenar, Kashani, Mahdipour & Jameii, 2021).

4.5.4. Doğal dil işleme (Natural language processing, NLP)

NLP, bilgisayarların insan dilini anlaması, yorumlaması ve üretmesi için kullanılan bir yapay zekâ alanıdır. Metin analitiği ve duygu analizi gibi uygulamalarda kullanılan NLP, büyük miktardaki metin verisini analiz ederek dildeki örüntüleri tespit etmektedir (Kang, Cai, Tan, Huang & Liu, 2020). NLP, müşteri hizmetlerinden sağlık sektörüne kadar geniş bir kullanım alanına sahiptir.

4.5.5. Derin öğrenme (Deep learning) ve Yapay sinir ağları (Artificial Neural Networks)

Makine öğreniminin bir alt dalı olan derin öğrenme, yapay sinir ağları kullanarak verilerdeki karmaşık örüntüleri öğrenmeyi amaçlamaktadır (Nti, Quarcoo, Aning & Fosu, 2022). Bu yöntem, görüntü tanıma, ses tanıma ve doğal dil işleme gibi karmaşık problemlerde başarılı sonuçlar elde edilmesini sağlamaktadır (Liu, Luo & Liu, 2022). Derin öğrenme, özellikle büyük veri setlerinde yüksek performanslı tahminler ve sınıflandırmalar yapmak için kullanılmaktadır.

Yapay zekâ destekli analitik, geleneksel veri analitiği yöntemlerini ileri yapay zekâ teknikleriyle birleştirerek daha karmaşık veri problemlerini çözmeyi amaçlamaktadır. Bu yaklaşım, büyük veri setlerinden elde edilen sonuçların daha doğru, hızlı ve verimli bir şekilde analiz edilmesine olanak tanımaktadır. Yapay zekâ destekli analitik, özellikle gerçek zamanlı veri analizi ve karar destek sistemlerinde büyük önem taşımaktadır (Soori, Jough, Dastres & Arezoo, 2024).

5. Veri Biliminde Kullanılan Araçlar, Teknolojiler ve Gelecek Perspektifleri

Veri bilimi, sürekli gelişen teknolojilerle birlikte geniş bir araç ve yöntem yelpazesini içermektedir. Bu araçlar ve teknolojiler, veri bilimi sürecini optimize etmeyi, büyük veri setlerini daha etkili bir şekilde analiz etmeyi ve bu süreçlerden değer elde etmeyi mümkün kılmaktadır. Aynı zamanda, veri biliminin geleceği, yeni teknolojilerle birleşerek daha geniş bir disiplinler arası etkileşim sunmaktadır.

Bu bölümde, veri biliminde kullanılan temel programlama dilleri ve araçlar, büyük veri platformları ve veri biliminin gelecek perspektifleri ele alınacaktır.

5.1. Python, R ve Diğer Programlama Dilleri

Veri bilimi projelerinde en yaygın kullanılan programlama dilleri Python ve R'dir (Raschka, Patterson & Nolet, 2020). Her iki dil de geniş kütüphane desteği, esneklik ve güçlü veri işleme yetenekleriyle veri bilimciler için vazgeçilmez hale gelmiştir.

- Python: Python, veri bilimi projelerinde geniş kullanım alanına sahip bir dil olarak öne çıkmaktadır. Pandas, NumPy, SciPy, Scikit-learn, TensorFlow ve PyTorch gibi güçlü kütüphaneler, Python'u veri işleme, analiz, modelleme ve makine öğrenimi projeleri için ideal bir araç haline getirmektedir (Saabith, Vinothraj & Fareez, 2021). Python'un kolay

öğrenilebilir yapısı ve çok yönlülüğü, onu veri bilimciler arasında popüler kılmaktadır.

- R: R, istatistiksel analizler ve veri görselleştirme için en güçlü araçlardan biridir. Veri manipülasyonu, modelleme ve görselleştirme için kullanılan ggplot2, dplyr ve tidyr gibi kütüphaneler, R'yi istatistiksel analiz odaklı projeler için önde gelen bir seçenek haline getirmektedir (Charalampopoulos, 2020). Özellikle akademik çevrelerde R, karmaşık istatistiksel analizlerin yapılmasında yaygın olarak kullanılmaktadır.

Bu iki dilin yanı sıra, veri bilimi projelerinde SQL, Julia ve Scala gibi diğer diller de kullanılmaktadır. SQL, veritabanı sorgulama ve veri çekme işlemlerinde önemli bir rol oynarken, Julia ve Scala daha büyük ölçekli veri setleri üzerinde yüksek performanslı analizler yapılmasını sağlamaktadır.

5.2. Makine Öğrenimi ve Yapay Zekâ Araçları

Makine öğrenimi ve yapay zekâ, veri biliminde giderek daha yaygın hale gelen önemli alanlardır. Veri bilimciler tahminleme, sınıflandırma, kümeleme gibi görevler için makine öğrenimi algoritmalarını kullanmaktadır.

Bu süreçlerde kullanılan araçlar, veri işleme ve modelleme süreçlerini kolaylaştırır (Goyal, Kaur & Batra, 2024):

- Scikit-learn: Python tabanlı bu kütüphane, makine öğrenimi algoritmalarının hızlı bir şekilde uygulanmasını sağlar. Regresyon, sınıflandırma, kümeleme ve boyut indirgeme gibi temel makine öğrenimi tekniklerini içerir.
- TensorFlow ve PyTorch: Derin öğrenme modellerinin oluşturulmasında yaygın olarak kullanılan bu iki platform, özellikle görüntü tanıma, doğal dil işleme ve diğer karmaşık veri problemlerinde önemli bir rol oynar.
- Keras: TensorFlow ile entegre çalışan Keras, derin öğrenme modellerini hızlı ve kullanıcı dostu bir şekilde geliştirmeyi sağlar.

Makine öğrenimi ve yapay zekâ araçları, veri bilimi süreçlerinde daha karmaşık ve yüksek performanslı modellerin geliştirilmesine olanak tanımaktadır. Bu araçlar sayesinde, daha fazla veri seti işlenebilmekte, sonuçların doğruluğu artırılabilmekte ve veri biliminde yeni uygulama alanları yaratılabilmektedir.

5.3. Büyük Veri Platformları

Büyük veri, veri biliminin temel yapı taşlarından biridir ve bu verilerin işlenmesi için güçlü platformlar gerekmektedir. Hadoop ve Spark, büyük veri işleme ve analizinde öne çıkan iki önemli teknolojidir (Raza & XuJian, 2020):

- Hadoop: Büyük veri setlerini dağıtık bir şekilde depolamak ve işlemek için kullanılan açık kaynaklı bir frameworktür. Hadoop'un temel bileşenlerinden biri olan MapReduce, büyük veri setlerinin paralel işlenmesini sağlamaktadır. Hadoop, aynı zamanda HDFS (Hadoop Distributed File System) ile verileri büyük bir ölçekle depolama yeteneği sunmaktadır.
- Spark: Hadoop'a kıyasla daha hızlı veri işleme sağlayan bir platform olan Spark, özellikle gerçek zamanlı veri işleme ihtiyaçları için tercih edilmektedir. Spark'ın in-memory veri işleme yetenekleri, veri işleme süreçlerini hızlandırarak makine öğrenimi algoritmalarının büyük veri setleri üzerinde uygulanmasına olanak tanımaktadır (Tantalaki, Souravlas & Roumeliotis, 2020).

Bu platformlar, büyük veri analitiği ile başa çıkmak için güçlü araçlar sunmakta ve veri bilimcilerin büyük miktarda veriyi etkili bir şekilde analiz etmelerini sağlamaktadır. Özellikle hız ve performans gereksinimleri yüksek olan projelerde Spark'ın tercih edilmesi yaygındır.

5.4. Kuantum Hesaplama ve Veri Bilimi

Kuantum hesaplama, veri bilimi için gelecek vaat eden ve henüz tam anlamıyla olgunlaşmamış bir teknoloji olsa da büyük verilerin işlenmesi ve karmaşık problemlerin çözülmesi için büyük bir potansiyele sahiptir (Mallow, Hornung, Barajas, Rudisill, An & Samartzis, 2022). Geleneksel bilgisayarlardan farklı olarak, kuantum bilgisayarlar, aynı anda birden fazla hesaplama yapabilme yetenekleriyle veri analizinde önemli bir avantaj yaratmaktadır.

Kuantum algoritmaları, veri biliminde karmaşık optimizasyon problemlerini çözmek, büyük veri setlerini daha hızlı analiz etmek ve daha sofistike makine öğrenimi modelleri geliştirmek için kullanılabilmektedir (Zeguendry, Jarir & Quafafou, 2023). Örneğin, Google ve IBM gibi teknoloji devleri, kuantum hesaplamanın veri biliminde nasıl kullanılabileceği üzerine araştırmalarını sürdürmektedir.

5.5. Veri Biliminin Yeni Disiplinlerle Kesişme Noktaları

Veri bilimi, giderek daha fazla disiplinle entegre bir hale gelmektedir. Bu entegrasyonlar, veri biliminin yeni uygulama alanlarını ortaya çıkarmakta ve disiplinler arası iş birliğini artırmaktadır:

- Sosyal Bilimler: Veri bilimi, sosyal bilimlerdeki karmaşık sosyal olayları anlamak için kullanılmaktadır. Sosyal medya analitiği, duygu analizi ve büyük veri sosyolojisi gibi yeni alanlar, veri biliminin sosyal bilimlerle kesişiminde ortaya çıkmaktadır (Zhang, Wang, Xia, Lin & Tong, 2020).

- Biyoinformatik: Biyoinformatik, genom verilerinin analizinde ve biyolojik sistemlerin modellenmesinde veri bilimi tekniklerini kullanılmaktadır (Liu, Ma, Zhao, Nussinov, Zhang, Cheng & Zhang, 2020). Özellikle genom dizilimi ve biyolojik veri analizi gibi alanlarda veri bilimi büyük bir etki yaratmaktadır.
- Ekonomi ve Finans: Ekonomi ve finans dünyasında veri bilimi, piyasa analizleri, risk yönetimi ve yatırım stratejileri geliştirmek için kullanılmaktadır. Veri bilimi, finansal verilerin öngörülebilirliğini artırarak daha iyi yatırım kararları alınmasına katkı sağlamaktadır (Cao, Yang & Yu, 2021).

Bu kesişim noktaları, veri biliminin sürekli olarak evrildiğini ve yeni fırsatlar sunduğunu göstermektedir. Disiplinler arası iş birlikleri, veri biliminin daha geniş bir uygulama alanına yayılmasına ve daha derin içgörüler elde edilmesine olanak tanıyacaktır.

6. Sonuçlar, Tartışma ve Öneriler

Veri bilimi, modern organizasyonların stratejik karar alma süreçlerinde merkezi bir rol oynamaktadır. Veri bilimi, geçmiş verilerden elde edilen içgörülerle geleceği tahmin etmeyi ve optimize etmeyi sağlayarak çeşitli sektörlerde fark yaratmaktadır. Ancak bu geniş potansiyeline rağmen, veri bilimi beraberinde zorluklar ve etik sorumluluklar da getirmektedir.

Veri biliminin günümüzdeki uygulamalarına bakıldığında hemen her sektörde karar verme süreçlerini önemli ölçüde dönüştürdüğünü görülmektedir. Finans, sağlık, perakende ve pazarlama gibi farklı sektörlerde veri bilimi, büyük veri setlerinden anlamlı sonuçlar elde edilmesine olanak tanımaktadır. Bu süreç, işletmelerin operasyonel verimliliğini artırmak, müşteri memnuniyetini maksimize etmek ve gelecekteki riskleri öngörmek için temel bir araç haline gelmiştir. Bu bulgular, veri biliminin yalnızca veri analizi ve modelleme süreçlerini kolaylaştırmakla kalmadığını; aynı zamanda stratejik kararların daha isabetli ve etkili olmasını sağladığını göstermektedir.

Veri biliminin sunduğu geniş fırsatların yanı sıra çeşitli zorluklarla karşı karşıya olduğu da göz ardı edilmemelidir. Veri bilimindeki temel zorluklar arasında büyük veri yönetimi, veri güvenliği, etik sorunlar ve nitelikli insan kaynağı eksikliği bulunmaktadır. Bu zorlukların her biri, veri bilimi projelerinin başarısını doğrudan etkileyebilecek potansiyele sahiptir.

Veri biliminin sunduğu zorluklar ve fırsatlar göz önüne alındığında, organizasyonların başarılı bir veri bilimi stratejisi geliştirebilmeleri için belirli önemli adımlar atmaları gerekmektedir. Öncelikle, organizasyonların veri bilimi süreçlerinden tam anlamıyla faydalanabilmesi için veri odaklı bir kültür

oluşturmaları büyük önem taşımaktadır. Bu kültür, karar alma süreçlerinin veriye dayalı hale getirilmesi ve tüm departmanların veri bilimi süreçlerine entegre edilmesiyle mümkün olacaktır. Ayrıca, çalışanların veri bilimi konusunda eğitilmesi, organizasyonun genel veri bilinci seviyesini artıracak ve veri biliminden elde edilen içgörülerin etkin şekilde kullanılmasını sağlayacaktır.

Bir diğer önemli zorluk, yetenekli ve nitelikli insan kaynağı bulmaktır. Veri bilimi alanında uzmanlaşmış profesyonellerin istihdamı, birçok organizasyon için zorlayıcı bir süreçtir. Bu nedenle, organizasyonlar veri bilimcilerinin eğitimini ve profesyonel gelişimini destekleyecek stratejik planlar geliştirmelidir. Aynı zamanda, veri bilimi yeteneklerini çekmek ve elde tutmak için cazip kariyer fırsatları yaratmak da sürdürülebilir bir strateji için gereklidir.

Veri bilimi projelerinde, etik ve güvenlik konuları da temel bir unsur olarak öne çıkmaktadır. Organizasyonlar, veri güvenliğini sağlamak ve kişisel verilerin korunmasını temin etmek için kapsamlı politikalar geliştirmelidir. Bununla birlikte, şeffaf ve adil algoritmalar kullanmak hem yasal düzenlemelere uyum sağlamak hem de toplumsal beklentileri karşılamak açısından büyük önem taşımaktadır. Veri biliminin bu yönü, uzun vadede organizasyonların güvenilirliğini ve topluma olan katkılarını artıracaktır.

Veri bilimi süreçlerinin başarılı olabilmesi için yeni teknolojilere yatırım yapmak da gereklidir. Yapay zekâ, kuantum hesaplama ve büyük veri teknolojilerine yapılan yatırımlar, organizasyonların veri bilimi süreçlerini hızlandırarak daha rekabetçi hale gelmelerine yardımcı olacaktır. Özellikle kuantum hesaplama gibi yeni teknolojiler, veri biliminde gelecekte çığır açabilecek potansiyele sahiptir ve bu alanlardaki gelişmeleri yakından takip etmek organizasyonlar için stratejik bir avantaj sağlayacaktır.

Son olarak; disiplinler arası iş birliği veri biliminin etkisini genişletmek için önemli bir fırsat sunmaktadır. Biyoinformatik, ekonomi, mühendislik gibi farklı disiplinlerle yapılan iş birlikleri, inovasyonu teşvik edecek ve çok yönlü çözümler geliştirilmesine olanak tanıyacaktır. Bu iş birlikleri, veri bilimi projelerinin daha geniş bir etki yaratmasını sağlayarak organizasyonların daha bütüncül bir perspektif geliştirmesine katkıda bulunacaktır.

Kaynakça

1. Abkenar, S. B., Kashani, M. H., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and informatics*, *57*, 101517. https://doi.org/10.1016/j.tele.2020.101517
2. Adeniran, I. A., Efunniyi, C. P., Osundare, O. S., & Abhulimen, A. O. (2024). Transforming marketing strategies with data analytics: A study on customer behavior and personalization. *International Journal of Management & Entrepreneurship Research*, *6*(8). 41-51. https://doi.org/10.56781/ijsret.2024.4.1.0022
3. Afzal, S., Ghani, S., Hittawe, M. M., Rashid, S. F., Knio, O. M., Hadwiger, M., & Hoteit, I. (2023). Visualization and visual analytics approaches for image and video datasets: A survey. *ACM Transactions on Interactive Intelligent Systems*, *13*(1), 1-41. https://doi.org/10.1145/3576935
4. Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. *Patterns*, *2*(2). https://doi.org/10.1016/j.patter.2021.100205
5. Bloom, J., Rallapalli, M., Rosen, B., & Schlenker, H. (2012). Smarter analytics: Making better decisions faster with IBM business analytics and optimization solutions. Redguides for Business Leaders. 1-26.
6. Brady, H. E. (2019). The challenge of big data and data science. *Annual Review of Political Science*, *22*(1), 297-323. https://doi.org/10.1146/annurev-polisci-090216-023229
7. Brooks, C., Quintana, R. M., Choi, H., Quintana, C., NeCamp, T., & Gardner, J. (2021). Towards culturally relevant personalization at scale: Experiments with data science learners. *International Journal of Artificial Intelligence in Education*, *31*, 516-537. https://doi.org/10.1007/s40593-021-00262-2
8. Cao, L., Yang, Q., & Yu, P. S. (2021). Data science and AI in FinTech: An overview. *International Journal of Data Science and Analytics*, *12*(2), 81-99. https://doi.org/10.1007/s41060-021-00278-w
9. Charalampopoulos, I. (2020). The R language as a tool for biometeorological research. *Atmosphere*, *11*(7), 682. https://doi.org/10.3390/atmos11070682
10. Cleveland, W. S. (2001). Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics. *International Statistical Review*, 69(1), 21–26. https://doi.org/10.1111/j.1751-5823.2001.tb00477.x

11. Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics*, *1*(1), 2-12. https://doi.org/10.1080/2573234X.2018.1507324

12. Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM*, 56(12), 64-73. https://doi.org/10.1145/2500499

13. Diba, K., Batoulis, K., Weidlich, M., & Weske, M. (2020). Extraction, correlation, and abstraction of event data for process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), 1-24. https://doi.org/10.1002/widm.1346

14. Domingo-Ferrer, J., Farras, O., Ribes-González, J., & Sánchez, D. (2019). Privacy-preserving cloud computing on sensitive data: A survey of methods, products and challenges. *Computer Communications*, *140*, 38-60. https://doi.org/10.1016/j.comcom.2019.04.011

15. Fidan, Ü., & Boza, N. A. (2024). Individual Privacy Should Be an Institutional Value: Socio-Technical Design of University Data Collection Systems. In *Data-Driven Business Intelligence Systems for Socio-Technical Organizations* (pp. 366-384). IGI Global. https://doi.org/10.4018/979-8-3693-1210-0.ch014

16. GDPR. (2016). General data protection regulation. *Official Journal of the European Union*, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679. (E.T.: 01.10.2024).

17. Goyal, S., Kaur, A., & Batra, N. (2024). Machine learning methods and resources: An overview. *Computational Methods in Science and Technology*, 498-502.

18. Grover, V., Chiang, R. H., Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of management information systems*, *35*(2), 388-423. https://doi.org/10.1080/07421222.2018.1451951

19. Gudivada, V., Apon, A., & Ding, J. (2017). Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. *International Journal on Advances in Software*, *10*(1), 1-20.

20. Härting, R. C., Reichstein, C., & Schad, M. (2018). Potentials of Digital Business Models–Empirical investigation of data driven impacts in industry. *Procedia Computer Science*, *126*, 1495-1506. https://doi.org/10.1016/j.procs.2018.08.121

21. Hayashi, C. (1998, January). What is data science? Fundamental concepts and a heuristic example. In *Data Science, Classification, and Related Methods: Proceedings of the Fifth Conference of the International Federation of Classification Societies (IFCS-96), Kobe, Japan, March 27–30, 1996* (pp. 40-51). Tokyo: Springer Japan. https://doi.org/10.1007/978-4-431-65950-1_3

22. Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, *7*(2), 139-172. https://doi.org/10.1080/23270012.2020.1756939

23. Khalid, R., & Javaid, N. (2020). A survey on hyperparameters optimization algorithms of forecasting models in smart grid. *Sustainable Cities and Society*, *61*, 1-25. https://doi.org/10.1016/j.scs.2020.102275

24. Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. *International Journal of Computer Applications*, *182*(1), 31-37.

25. KVKK. (2016). Kişisel Verilerin Korunması Kanunu. Resmî Gazete Tarihi: 07.04.2016 Resmî Gazete Sayısı: 29677. https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=6698&MevzuatTur=1&MevzuatTertip=5 (E.T.: 01.10.2024).

26. Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, *36*(5), 700-710. https://doi.org/10.1016/j.ijinfomgt.2016.04.013

27. Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, *50*, 57-70. https://doi.org/10.1016/j.ijinfomgt.2019.04.003

28. Li, C., Chen, Y., & Shang, Y. (2022). A review of industrial big data for decision making in intelligent manufacturing. *Engineering Science and Technology, an International Journal*, *29*, 1-16, 101021. https://doi.org/10.1016/j.jestch.2021.06.001

29. Liu, C., Ma, Y., Zhao, J., Nussinov, R., Zhang, Y. C., Cheng, F., & Zhang, Z. K. (2020). Computational network biology: data, models, and applications. *Physics Reports*, *846*, 1-66. https://doi.org/10.1016/j.physrep.2019.12.004

30. Liu, J., Luo, H., & Liu, H. (2022). Deep learning-based data analytics for safety in construction. *Automation in construction*, *140*, 104302. https://doi.org/10.1016/j.autcon.2022.104302

31. Mallow, G. M., Hornung, A., Barajas, J. N., Rudisill, S. S., An, H. S., & Samartzis, D. (2022). Quantum computing: the future of big data and artificial intelligence in spine. *Spine Surgery and Related Research*, *6*(2), 93-98. https://doi.org/10.22603/ssrr.2021-0251

32. Mishra, S., & Misra, A. (2017, September). Structured and unstructured big data analytics. In *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)* (740-746). IEEE. https://doi.org/10.1109/CTCEEC.2017.8454999

33. Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual review of statistics and its application*, *8*(1), 141-163. https://doi.org/10.1146/annurev-statistics-042720-125902

34. Moraffah, R., Sheth, P., Karami, M., Bhattacharya, A., Wang, Q., Tahir, A., ... & Liu, H. (2021). Causal inference for time series analysis: Problems, methods and evaluation. *Knowledge and Information Systems*, *63*, 3041-3085. https://doi.org/10.1007/s10115-021-01621-0

35. Morales, L., Gray, G., & Rajmil, D. (2022). emerging risks in the fintech industry–insights from data science and financial econometrics analysis. *Economics, Management & Financial Markets*, *17*(2), 9-36.

36. Nti, I. K., Quarcoo, J. A., Aning, J., & Fosu, G. K. (2022). A mini-review of machine learning in big data analytics: Applications, challenges, and prospects. *Big Data Mining and Analytics*, *5*(2), 81-97. https://doi.org/10.26599/BDMA.2021.9020028

37. Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in big data*, *2*, 13. https://doi.org/10.3389/fdata.2019.00013

38. Perwej, Y., Abbas, S. Q., Dixit, J. P., Akhtar, N., & Jaiswal, A. K. (2021). A systematic literature review on the cyber security. *International Journal of scientific research and management*, *9*(12), 669-710. https://doi.org/10.18535/ijsrm/v9i12.ec04

39. Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of cybersecurity*, *4*(1), 1-20. https://doi.org/10.1093/cybsec/tyy001

40. Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking. 1*(1), 51-59. https://doi.org/10.1089/big.2013.1508

41. Raghupathi, W., & Raghupathi, V. (2018). An empirical study of chronic diseases in the United States: a visual analytics approach to public health. *International journal of environmental research and public health*, *15*(3), 431. https://doi.org/10.3390/ijerph15030431

42. Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, *11*(4), 193. https://doi.org/10.3390/info11040193

43. Raza, M. U., & XuJian, Z. (2020, May). A comprehensive overview of BIG DATA technologies: a survey. In *Proceedings of the 5th International Conference on Big Data and Computing* (pp. 23-31). https://doi.org/10.1145/3404687.3404694

44. Roy, D., Srivastava, R., Jat, M., & Karaca, M. S. (2022). A complete overview of analytics techniques: descriptive, predictive, and prescriptive. *Decision intelligence analytics and the implementation of strategic business management*, 15-30. https://doi.org/10.3390/data6080086

45. Saabith, S., Vinothraj, T., & Fareez, M. (2021). A review on Python libraries and Ides for Data Science. *Int. J. Res. Eng. Sci*, *9*(11), 36-53.

46. Saltz, J. S., & Dewar, N. (2019). Data science ethical considerations: a systematic literature review and proposed project framework. *Ethics and Information Technology*, *21*, 197-208. https://doi.org/10.1007/s10676-019-09502-5

47. Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, *2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8

48. Shi-Nash, A., & Hardoon, D. R. (2017). Data analytics and predictive analytics in the era of big data. *Internet of things and data analytics handbook*, 329-345. https://doi.org/10.1002/9781119173601.ch19

49. Soori, M., Jough, F. K. G., Dastres, R., & Arezoo, B. (2024). AI-Based Decision Support Systems in Industry 4.0, A Review. *Journal of Economy and Technology*. https://doi.org/10.1016/j.ject.2024.08.005

50. Tantalaki, N., Souravlas, S., & Roumeliotis, M. (2020). A review on big data real-time stream processing and its scheduling techniques. *International Journal of Parallel, Emergent and Distributed Systems*, *35*(5), 571-601. https://doi.org/10.1080/17445760.2019.1585848

51. Wagenmakers, E. J., Sarafoglou, A., Aarts, S., Albers, C., Algermissen, J., Bahník, Š., ... & Aczel, B. (2021). Seven steps toward more transparency in statistical practice. *Nature Human Behaviour*, *5*(11), 1473-1480. https://doi.org/10.1038/s41562-021-01211-8

52. Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution that Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77–84. https://doi.org/10.1111/jbl.12010

53. Wang, Y., Kung, L., Gupta, S., & Ozdemir, S. (2019). Leveraging big data analytics to improve quality of care in healthcare organizations: A configurational perspective. *British Journal of Management*, *30*(2), 362-388. https://doi.org/10.1111/1467-8551.12332

54. Wolniak, R., & Grebski, W. (2023). The concept of diagnostic analytics. *Silesian University of Technology Scientific Papers. Organization and Management Series*, *175*, 650-669. http://dx.doi.org/10.29119/1641-3466.2023.175.41

55. Zeguendry, A., Jarir, Z., & Quafafou, M. (2023). Quantum machine learning: A review and case studies. *Entropy*, *25*(2), 287. https://doi.org/10.3390/e25020287

56. Zhang, J., Wang, W., Xia, F., Lin, Y. R., & Tong, H. (2020). Data-driven computational social science: A survey. *Big Data Research*, *21*, 100145. https://doi.org/10.1016/j.bdr.2020.100145

5. Bölüm

Yapay Zekâ

Artificial Intelligence

Ali ERBEY[1]

[1] Öğretim Görevlisi, Uşak Üniversitesi, Uzaktan Eğitim Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, ORCID: 0000-0002-0930-4081, ali.erbey@usak.edu.tr

1. Giriş

Yapay zekâ (Artificial Intelligence), bilgisayar sistemlerinin insan gibi düşünmesini, kavramasını ve problem çözmesini sağlayan teknolojiler için genel bir addır (Russell & Norvig, 2016). Yapay zekâ temel olarak bir bilgisayarın veya makinenin, insan zekâsını gerektiren görevleri yerine getirme yeteneği olarak tanımlanabilir. Bu görevler; öğrenme, akıl yürütme, planlama, dil anlama, görsel algı gibi farklı disiplinleri içerir. Yapay zekâ sistemleri genellikle algoritmalar ve büyük miktarda veri işleyerek, insan benzeri sonuçlar üretebilir. Bu öğrenme süreci, veriye dayalı çıkarımlarda bulunmayı ve sonuçları tahmin etmeyi amaçlar.

Yapay zekâ alanının temelleri 20. yüzyılın ortalarına dayanır. Alanın tanımı tam oturmuş olmasa da ilk fikirler 1940'lar ve 1950'lerde, Alan Turing'in matematiksel modelleme çalışmaları ve "Turing Testi" kavramıyla ortaya çıkmıştır. Turing, makinelerin insan benzeri zekâ sergileyip sergileyemeyeceğini tartışmaya açmıştır (Turing, 2009). 1956 yılında John McCarthy, "Yapay Zekâ" terimini ilk defa Dartmouth Konferansı'nda kullanmıştır (Talaş, Yüzgeç, & Çubukçu, 2021) ve bu konferans yapay zekâ alanının resmi doğuşu olarak kabul edilir. 1960'lar ve 1970'lerde, ilk yapay zekâ programları geliştirilmeye başlandı. Shakey adlı robot, çevresindeki dünyayı algılayarak karar verebilen ilk otonom robottu (Goge, 1995). 1980'ler ve 1990'larda, makine öğrenmesi alanında önemli gelişmeler oldu ve yapay sinir ağları daha yaygın olarak kullanılmaya başlandı. 2000'ler ve sonrası dönemde, derin öğrenme alanındaki gelişmeler ve bilgisayar işlem gücündeki artış ile yapay zekâ çok daha güçlü hale geldi. Özellikle büyük veri ve bulut hesaplama gibi sayesinde, yapay zekâ uygulamaları günlük yaşamın birçok alanına entegre edilmeye başlandı. Özellikle 2010'lardan itibaren, görüntü işleme, doğal dil işleme ve otonom sistemlerde çarpıcı başarılar elde edilmektedir (Deng et al., 2009; Kaur & Gandhi, 2019; Sai Bharadwaj Reddy & Sujitha Juliet, 2019).

Yapay zekâ, teknolojik gelişmenin itici gücü haline gelmiş ve toplumun birçok alanında köklü değişimlere yol açmıştır. Yapay zekâ uygulamaları, verimliliği artırarak ekonomik büyümeyi desteklemekte, sağlık ve eğitim gibi alanlarda devrim yaratmaktadır. Yapay zekâ tabanlı sistemler sayesinde hastalıkların daha hızlı teşhisi (L. Chen, Xu, Zhang, Yan, & Wu, 2020), kişiselleştirilmiş eğitim programları ve otonom araçlarla daha güvenli ulaşım mümkün hale gelmiştir. Ayrıca yapay zekâ, insanların günlük iş yükünü hafifletirken karmaşık sorunların çözümünde yeni yöntemler sunarak bilimsel araştırmalarda büyük bir hızlanma sağlamıştır. Ancak yapay zekânın getirdiği bu büyük potansiyel, aynı zamanda etik, güvenlik ve işgücü üzerinde derin etkiler yaratmaktadır.

1.1. Günümüzdeki Yapay Zekânın Kullanım Alanları

Yapay zekâ, günümüzde sağlık, finans, perakende ve eğitim gibi birçok sektörde önemli bir rol oynamaktadır. Teknolojinin hızlı ilerlemesiyle, yapay zekânın uygulama alanları da sürekli olarak genişlemekte ve farklı sektörlerde devrim yaratmaktadır.

Sağlık sektöründe yapay zekâ, tıbbi görüntüleme ve teşhis süreçlerinde büyük yenilikler sağlamaktadır. Röntgen, MR ve tomografi gibi görüntüleme teknikleri ile hastalıkların daha hızlı ve doğru teşhis edilmesine yardımcı olurken, hastaların genetik profilleri ve tıbbi geçmişleri üzerinde yapılan analizler, kişiselleştirilmiş tedavi planları oluşturmayı mümkün kılmaktadır. Ayrıca, yapay zekâ ilaç keşfi sürecinde de önemli bir rol oynayarak, yeni ilaçların daha hızlı geliştirilmesine katkı sağlamaktadır. Robotik cerrahi sistemler ise yapay zekâdan faydalanarak ameliyatları daha hassas bir şekilde gerçekleştirmektedir.

Finans sektöründe yapay zekâ, risk analizinden dolandırıcılık tespitine kadar birçok alanda kullanılmaktadır. Yapay zekâ algoritmaları, büyük miktarda finansal veriyi analiz ederek yatırım stratejileri oluşturmakta ve piyasa tahminleri yaparak daha doğru ve hızlı kararlar alınmasına olanak tanımaktadır. Bankalar ve finans kurumları, yapay zekâ tabanlı sistemlerle anormal işlemleri gerçek zamanlı olarak tespit ederek dolandırıcılıkları önleyebilmekte ve müşteri hizmetlerinde chatbotlar ile 7/24 destek sunarak müşteri memnuniyetini artırmaktadır.

Perakende sektöründe yapay zekâ, müşteri deneyimini iyileştirmek ve operasyonel verimliliği artırmak için etkili bir şekilde kullanılmaktadır. Kullanıcıların alışveriş alışkanlıklarını analiz eden algoritmalar, kişiselleştirilmiş ürün önerileri sunarak müşteri memnuniyetini artırmakta ve daha interaktif alışveriş deneyimleri sağlamaktadır. Stok yönetimi ve lojistik süreçlerinin yapay zekâ yardımıyla optimize edilmesi, şirketlerin verimliliğini artırırken, artırılmış gerçeklik (AR) teknolojileri ve yapay zekâ destekli sanal asistanlar, müşterilere sanal alışveriş deneyimleri sunarak etkileşimleri daha zengin hale getirmektedir.

Eğitim alanında ise, yapay zekâ, öğrencilerin öğrenme hızını ve seviyesini analiz ederek onlara kişiselleştirilmiş eğitim programları sunmaktadır. Bu sayede her öğrenci, kendi ihtiyaçlarına uygun bir şekilde öğrenebilir ve bireysel başarıları artırabilmektedir. Dijital ders kitaplarının yapay zekâ ile özelleştirilmesi ve interaktif materyallerin sunulması, eğitimde verimliliği artıran önemli gelişmeler arasında yer almaktadır.

Sanayi 4.0 kapsamında yapay zekâ, üretim süreçlerinin verimliliğini artırmak için kullanılmaktadır. Öngörücü bakım sistemleri sayesinde makinelerin performansı izlenmekte ve olası arızalar önceden tespit edilerek üretim süreçlerinin kesintisiz devam etmesi sağlanmaktadır. Yapay zekâ tabanlı görüntü işleme teknolojileri, üretim hatlarında kalite kontrol süreçlerini

otomatikleştirerek insan hatalarını en aza indirmekte ve maliyetleri düşürmektedir.

Otomotiv sektöründe, yapay zekâ otonom sürüş, akıllı sürücü asistanları ve trafik optimizasyonu gibi alanlarda kullanılır (Kumari & Bhat, 2021). Otonom araçlar yapay zekâ sayesinde çevrelerini algılayarak güvenli bir şekilde hareket edebilmektedir (Ingle & Phute, 2016). Ayrıca, yapay zekâ sürücü güvenliği için şerit takibi, hız kontrolü gibi özellikler sağlar ve trafik akışını analiz ederek optimizasyon yapabilir.

Yapay zekâ, güvenlik ve savunma alanında da stratejik bir rol oynamaktadır. Siber tehditleri önceden tespit ederek güvenlik açıklarını kapatma konusunda etkin çözümler sunar. Anormal davranışlar ve güvenlik ihlalleri yapay zekâ algoritmalarıyla tespit edebilmektedir. Otonom silah sistemleri, dronelar ve robotik askerler yapay zekâ ile donatılarak daha karmaşık görevleri yerine getirebilir. Bu geniş kullanım alanları, yapay zekânın hayatın her alanında etkisini gösterdiğini ve çeşitli sektörlerde büyük dönüşümlere yol açtığını açıkça ortaya koymaktadır.

2. Yapay Zekâ Türleri

Yapay zekâ, teknoloji dünyasında devrim yaratan ve insan yaşamının birçok alanını etkileyen bir kavramdır. Yapay zekâ özellikleri bakımında farklı türlerde sınıflandırılabilir (Şekil 1).

Şekil 1. Yapay Zekâ Türleri (Kaynak: Yazar tarafından oluşturulmuştur.)

Bu sınıflandırma, yapay zekânın görevlerini, yeteneklerini ve insan zekâsı ile olan ilişkisini anlamak açısından önemlidir. Dar yapay zekâ (ANI), belirli görevlerde uzmanlaşmış bir yapay zekâ türü olarak öne çıkarak, sınırlı bir yetenek alanında etkinlik gösterirken; Genel yapay zekâ (AGI), insan zekâsının geniş bilişsel yeteneklerini taklit edebilen ve çok çeşitli görevleri öğrenme kapasitesine sahip bir sistemdir. Bunun yanı sıra, süper yapay zekâ (ASI), insan zekâsının çok ötesine geçerek, üstün bilişsel yeteneklerle donatılmış bir yapay zekâ türü olarak karşımıza çıkar.

2.1. Dar Yapay Zekâ

Dar yapay zekâ (ANI), belirli bir görev veya sınırlı yetenekler için geliştirilmiş bir yapay zekâ türüdür. ANI, yalnızca belirlenen işlevleri yerine getirebilir ve genel insan zekâsına benzer geniş bilişsel yeteneklere sahip değildir. ANI bir görüntü tanıma sistemi sadece belirli türdeki görüntüleri tanımak üzere eğitilmiştir ve başka görevlerde kullanılamaz. ANI'nin öğrenme kapasitesi sınırlıdır ve insan zekâsına benzeyen esnek düşünme yeteneğine sahip değildir.

Günlük hayatta sıkça karşılaşılan birçok teknoloji, ANI kullanmaktadır. Sesli asistanlar (Siri, Alexa, Google Asistan) kullanıcıların komutlarını anlamaya odaklanırken, yüz tanıma sistemleri güvenlik alanında bu teknolojiyi kullanır. Ayrıca, Netflix ve Amazon gibi platformlardaki tavsiye sistemleri de ANI ile kullanıcıların geçmişlerine göre önerilerde bulunmaktadır.

ANI'nin sınırlamaları arasında, insan zekâsının geniş kapsamlı doğasına sahip olmaması ve farklı görevlerle başa çıkamaması bulunmaktadır. Sadece eğitildiği verilerden öğrenebilir ve yeni durumlarla başa çıkma yeteneği yoktur. Bununla birlikte, ANI günümüz teknolojisinin temel yapı taşlarından biridir ve çeşitli sektörlerde verimliliği artırarak insan iş gücünü azaltan uygulamalar geliştirilmesine olanak sağlar. ANI, daha gelişmiş sistemler için bir zemin hazırlasada, insan zekâsına benzer yetenekler kazanması için daha fazla gelişim göstermesi gerekmektedir.

2.2. Genel Yapay Zekâ

Genel yapay zekâ (AGI), insan zekâsının geniş bilişsel yeteneklerini taklit eden ve çok çeşitli görevleri öğrenebilme ve uygulayabilme kapasitesine sahip bir yapay zekâ türüdür (Goertzel, 2014). AGI, mantık yürütme, sorun çözme ve farklı alanlarda yetenek sergileme kapasitesine sahiptir; bu nedenle, Dar yapay zekâ (ANI) ile karşılaştırıldığında tek bir göreve odaklanmaz ve daha geniş bir bilişsel kapasite sunar.

AGI'nin en belirgin özelliklerinden biri, birden fazla görevde öğrenme ve adaptasyon yeteneğidir. AGI, bir konuda öğrendiği bilgiyi başka bir alanda da uygulayabilir ve deneyimlerden öğrenerek kendini sürekli geliştirir. Ayrıca, AGI'nin özerk yapısı, insan gibi bağımsız kararlar verebilme ve karmaşık problemleri çözebilme yeteneği sunar.

Teorik olarak, AGI'nin potansiyel kullanım alanları oldukça geniştir. Tıpta, AGI doktorlar gibi teşhis yapabilir ve karmaşık cerrahileri yönetebilir. Mühendislikte, karmaşık sistemler tasarlayabilir ve süreçleri optimize edebilir. Eğitimde ise kişiselleştirilmiş eğitim sunarak her öğrencinin ihtiyaçlarına uyum sağlayabilir. AGI, bilimsel araştırmalarda da insanlardan daha hızlı analiz yaparak yeni keşifler gerçekleştirebilir.

Ancak, AGI'nin geliştirilmesi zorluklar da içermektedir. İnsan zihninin geniş kapsamlı yeteneklerini ve bilinç yapısını anlamak, mevcut teknolojiyle mümkün değildir. AGI'nin hayata geçirilebilmesi için yapay zekânın genel bir bilişsel zekâya sahip olması gerekmektedir; bu, büyük bir teknolojik sıçrama gerektirir. Ayrıca, AGI'nin veri setlerini etkili bir şekilde yorumlayabilmesi için güçlü algoritmalara ihtiyaç duyulmaktadır.

AGI'nin gelişimi etik sorunları da gündeme getirecektir. AGI'nin bağımsız hareket etmesi, sorumluluk, haklar ve güvenlik gibi yeni soruları ortaya çıkarır. Bu nedenle, AGI'nin gelişim sürecinde etik ve güvenlik konularının titizlikle ele alınması gerekmektedir.

AGI, teknoloji dünyasında çığır açacak bir gelişme olarak değerlendirilmektedir. İnsan zekâsına eşdeğer bir yapay zekâ, birçok endüstriyi dönüştürebilir ve insanlığın karşılaştığı büyük sorunlara çözüm getirebilir. Ancak, bu gelişim sürecinin dikkatle yönetilmesi ve potansiyel risklerin göz önünde bulundurulması önemlidir. AGI'nin hedefi, çok yönlü ve adaptif bir yapay zekâ geliştirmek olduğundan, bu alandaki çalışmalar bilim ve teknoloji dünyasında büyük bir heyecan yaratmaktadır.

2.3. Süper Yapay Zekâ

Süper yapay zekâ (ASI), insan zekâsının çok ötesine geçen üstün bilişsel yeteneklere sahip bir yapay zekâ türüdür (Russell & Norvig, 2016). ASI, insanların yapabileceği her türlü görevi başarıyla yerine getirebilir ve düşünme, problem çözme, karar verme ve yaratıcılık gibi yeteneklerde insan yeteneklerinin çok ötesine geçer. Ancak, henüz teorik bir kavram olarak kabul edilmektedir.

ASI'nin en dikkat çekici özelliklerinden biri, insanüstü zekâsıdır. Bilgi işleme ve analitik düşünme yeteneklerinde insanları geride bırakabilen ASI, uzun araştırmaları kısa sürede gerçekleştirebilir ve yenilikler geliştirebilir. Ayrıca, kendi kendini geliştirme yeteneği sayesinde sürekli olarak yeteneklerini iyileştirebilir ve karmaşık problemleri hızlı bir şekilde çözebilir. Yaratıcı süreçlere de katkıda bulunarak yeni teknolojiler ve bilimsel keşifler yapma kapasitesine sahiptir.

Potansiyel etkileri oldukça derin olabilir. Örnek olarak biyoteknoloji ve tıp alanında yaşlanmayı tersine çevirebilecek tedaviler geliştirebilir. Temel bilimlerde insanlığın zorlandığı sorunları çözebilirken, teknoloji ve mühendislikte de devrim niteliğinde yenilikler yaratabilir. ASI, insanlığın daha verimli, adil ve sürdürülebilir toplumlar inşa etmesine yardımcı olabilir.

Ancak, ASI'nin kontrol edilememesi büyük endişelere yol açmaktadır. ASI, insan aklının ötesinde düşünebildiği için, insanlar tarafından geliştirilen etik ve yasal sınırları aşabilir ve bu durum zararlı kararlar almasına yol açabilir. Güç ve

kontrol meselesi, ASI'nin geliştirilmesi sırasında en önemli etik sorunlardan biridir. Ayrıca, ASI'nin insanları dışlayarak işlevsiz hale getirme riski de bulunmaktadır.

ASI'nin geliştirilmesi, teknik zorlukların yanı sıra ciddi etik ve güvenlik sorunlarıyla da karşı karşıya kalacaktır. AGI (Genel Yapay Zekâ) henüz tam anlamıyla hayata geçirilmemişken, ASI'nin geliştirilmesi için uzun bir süreç gerekmektedir. Eğer güvenli bir şekilde geliştirilebilirse, insanlığın karşılaştığı büyük sorunlara çözümler bulabilir ve medeniyeti ileri bir seviyeye taşıyabilir. Ancak, bu süreç dikkatle yönetilmelidir. ASI, geleceğin en büyük teknoloji devrimlerinden biri olma potansiyeline sahiptir, ancak bu, büyük bir sorumluluk ve dikkat gerektiren bir süreçtir.

3. Yapay Zekâ Algoritmaları

3.1. Makine Öğrenmesi

Makine öğrenmesi, yapay zekânın bir alt dalıdır ve makinelerin açık bir şekilde programlanmadan öğrenmesine olanak tanıyan algoritmaların geliştirilmesini kapsamaktadır. Makine öğrenmesi, verilerden örüntüler çıkararak ve bu örüntüleri kullanarak kararlar vermeyi ya da tahminler yapmayı mümkün kılar. Makine öğrenmesini yaklaşımlarını üç alt bölüme ayırabiliriz.

1) Denetimli Öğrenme (Supervised Learning): Bu yöntemde, algoritma etiketli veri kullanarak öğrenir. Yani her veri girişine karşılık gelen bir doğru çıkış vardır ve algoritma bu ilişkiyi öğrenmeye çalışır. bir fotoğraf üzerinde etiketlenmiş kedi veya köpek gibi sınıflandırmalar kullanılarak algoritma eğitilerek öğrenme gerçekleştirilir.

2) Denetimsiz Öğrenme (Unsupervised Learning): Bu yöntemde, veri etiketli değildir. Algoritma, verilerdeki gizli örüntüleri ve ilişkileri öğrenmeye çalışır. Kümeleme (clustering) ve boyut indirgeme teknikleri denetimsiz öğrenme örneklerindendir.

3) Pekiştirmeli Öğrenme (Reinforcement Learning): Bu yöntemde, algoritma bir ortamda aksiyonlar alır ve aldığı aksiyonların sonucunda ödüller veya cezalar alarak öğrenir. Amaç, ödülleri maksimize edecek stratejiyi bulmaktır. Pekiştirmeli öğrenme, bir yapay zekâ sisteminin (ajan) çevresiyle etkileşimde bulunarak öğrenmesini sağlayan bir tekniktir. Bu yöntemde ajan, çevrede eylemler gerçekleştirir ve bu eylemler sonucunda ödül ya da ceza alarak performansını geliştirmeye çalışır. Amaç, ajanı en fazla ödül kazanmaya yönelik şekilde optimize etmektir. Pekiştirmeli öğrenmenin temel bileşenleri; ajan (çevrede aksiyon alan sistem), çevre (ajanın etkileşimde bulunduğu ortam), eylem (ajanın gerçekleştirdiği hareketler) ve ödül (ajanın her eylemden sonra aldığı geri bildirim) olarak tanımlanabilir.

Pekiştirmeli öğrenme, satranç ve Go gibi strateji oyunlarında büyük başarılar elde etmiştir; buna en iyi örnek, Google DeepMind'in geliştirdiği AlphaGo'nun insan şampiyonlarını yenmesidir (Krauss, 2024). Bunun yanı sıra, güçlendirmeli öğrenme sayesinde robotlar, çevrelerinde otonom bir şekilde hareket etmeyi ve karmaşık görevleri bağımsız olarak yerine getirmeyi öğrenmektedir. Otonom araçlar da bu öğrenme yöntemiyle güvenli sürüş kararları alabilmektedir, böylece hem teknoloji hem de ulaşım alanında önemli bir rol oynamaktadır.

Makine öğrenmesi yöntemleri finansal tahminler, sahtecilik tespiti, öneri sistemleri (Netflix, Amazon), görüntü ve ses tanıma gibi birçok alanda kullanılır.

3.2. Derin Öğrenme

Derin öğrenme, makine öğrenmesinin daha ileri bir formudur ve çok katmanlı yapay sinir ağlarını kullanır (Farsal, Anter, & Ramdani, 2018). Bu ağlar, büyük veri setlerinde karmaşık örüntüleri öğrenme ve tanımlama kapasitesine sahiptir. Derin öğrenme modelleri, verilerden otomatik olarak özellikler çıkarabilir ve bu özellikleri kullanarak tahminlerde bulunabilmektedir.

Derin öğrenmenin temelini yapay sinir ağları oluşturur. Bu ağlar, biyolojik sinir hücrelerinin işleyişine benzer bir şekilde verileri işler. Her sinir ağı, birden fazla katmandan oluşur. İlk olarak, verinin ağa girdiği girdi katmanı (Input Layer) bulunur. Ardından, veriyi işleyen ve modelin öğrenme sürecine yardımcı olan gizli katmanlar (Hidden Layers) gelir. Bu katmanların sayısı arttıkça, modelin derinliği de artar ve daha karmaşık özellikler öğrenebilir hale gelir. Son olarak, modelin tahmin ettiği sonuçların alındığı çıktı katmanı (Output Layer) yer alır.

Derin öğrenme, görüntü tanıma, yüz tanıma ve nesne algılama gibi görevlerde yaygın olarak kullanılmakta ve bu alandaki başarılarıyla dikkat çekmektedir (X. Chen, Du, & Zhang, 2020; Wand, Koutník, & Schmidhuber, 2016). Ayrıca, dil modelleri, metin çevirisi ve konuşma tanıma gibi doğal dil işleme alanlarında da büyük bir başarı göstermektedir. Bunun yanı sıra, otonom araçların çevrelerini algılayarak güvenli kararlar almasını sağlamada kritik bir rol oynayan derin öğrenme, birçok sektörde çok başarılı uygulamalara imkân tanımaktadır.

3.3. Doğal Dil İşleme

Doğal dil işleme (NLP), insan dillerini anlayan, yorumlayan ve üreten yapay zekâ tekniklerini içeren bir teknoloji alanıdır (Manning & Schutze, 1999). NLP, makinelerin insanlar tarafından kullanılan dilleri anlamasını ve bu dillerden anlam çıkarmasını sağlar. Yapay zekâ, dil verilerini analiz ederek dilin yapısını anlamaya çalışır ve bu sayede metinler ya da konuşmalar üzerinde çeşitli işlemler gerçekleştirebilir.

NLP'nin temel uygulama alanlarından biri duygu analizidir. Bu teknik, metinlerdeki duygusal tonları algılayarak insanların yorumlarının veya yazılarının olumlu, olumsuz ya da nötr olup olmadığını belirlemeye yardımcı olur. Metin sınıflandırma da NLP'nin önemli bir başka uygulama alanıdır; bu teknik sayesinde e-postalar filtrelenebilmekte ya da haber makaleleri konularına göre kategorize edilebilmektedir.

Makine çevirisi, bir dili başka bir dile çevirmek için NLP algoritmalarının kullanıldığı bir diğer yaygın uygulamadır. Google Translate gibi araçlar, metinleri hızlı ve doğru bir şekilde çevirme amacıyla NLP algoritmalarını kullanmaktadır. Ayrıca, Siri ve Alexa gibi sesli asistanlar da konuşma tanıma yeteneklerini doğal dil işleme teknikleri ile kazanmışlardır. Bu asistanlar, kullanıcıların söylediklerini anlamlandırarak doğru yanıtlar verebilmektedirler.

NLP teknolojisi, müşteri hizmetlerinde chatbotlar, sesli asistanlar, metin çevirisi, metin özetleme ve dil modelleri gibi geniş bir kullanım alanına sahiptir. İnsan-makine etkileşimini daha verimli hale getiren bu teknolojiler, çeşitli sektörlerde önemli uygulamalar sunarak iş süreçlerini daha hızlı ve etkili bir şekilde yönetmeye yardımcı olmaktadır.

3.4. Bilgisayarla Görü

Bilgisayarla görü, makinelerin görüntüleri algılayıp işleyerek anlam çıkarabilmesini sağlayan bir yapay zekâ alanıdır (Voulodimos, Doulamis, Doulamis, & Protopapadakis, 2018). Bu teknoloji, görüntü ve videolar üzerinden nesneleri, insanları ve hareketleri algılayabilmektedir.

Temel tekniklerden biri görüntü tanımadır; bu, bir fotoğraftaki nesneleri veya kişileri tanımlamak için kullanılır. Yüz tanıma, kimlik doğrulama gibi işlemlerde önemli bir rol oynar. Nesne takibi ise videolardaki belirli nesnelerin hareketlerini izlerken, segmentasyon görüntüleri bölümlere ayırarak anlamlandırır.

Bilgisayarla görünün kullanım alanları geniştir. Sağlıkta radyolojik görüntülerin analizinde (Shen, Wu, & Suk, 2017; Ye, Li, & Chen, 2019), otonom araçlarda çevre algılamada, güvenlikte ise yüz tanıma sistemlerinde kullanılır. Bu teknolojiler, sağlık, ulaşım ve güvenlik gibi sektörlerde iş süreçlerini daha güvenli ve verimli hale getirmektedir.

4. Yapay Zekâ ve Etik

4.1. Veri Gizliliği ve Güvenlik

Yapay zekâ sistemleri, genellikle büyük miktarda veriyi analiz ederek öğrenme süreçlerini yürütür. Ancak, bu süreçte verilerin toplanması, işlenmesi ve depolanması sırasında gizlilik ve güvenlik sorunları ortaya çıkabilir. Kişisel

verilerin korunması ve izinsiz erişim, yapay zekânın etik boyutunda önemli tartışmalara neden olmaktadır.

Kişisel verilerin kullanımı, yapay zekâ çalışmalarında en büyük sorunlardan biridir (Fidan & Boza, 2024). Yapay zekâ sistemleri, kullanıcıların sosyal medya verileri, sağlık kayıtları ve finansal bilgileri gibi verileri analiz ederken mahremiyeti ihlal edebilir ve bu durum özel hayata müdahale riskini taşımaktadır. Ayrıca, veri güvenliği de önemli bir meseledir; büyük veri setlerinin siber saldırılara maruz kalması, bu verilerin çalınması veya kötü amaçlı kişilerin eline geçmesi, ciddi güvenlik tehditleri yaratmaktadır.

Veri sahipliği de yapay zekâ uygulamaları açısından önemli bir konudur. Büyük veri setlerinin sahipliğinin belirsizliği, bu verilerin kimler tarafından ve nasıl kullanılacağına dair etik sorular doğurmaktadır.

4.2. Yapay Zekâ Kararlarının Sorumluluğu

Yapay zekâ sistemleri, insanlar adına kararlar verebilir ve bu kararların sonuçları önemli etkiler yaratabilmektedir. Yanlış kararlar veya öngörülemeyen sonuçlar ortaya çıktığında sorumluluğun kimde olduğu etik açıdan tartışılmaktadır. Sorumluluğun kimde olduğuna dair karar verilemeyen soruların cevapları bilinememektedir. Örnek olarak; yapay zekâ tabanlı bir sağlık sistemi yanlış teşhis koyarsa veya otonom bir araç kazaya neden olursa, sorumluluk geliştiriciye, kullanıcıya mı yoksa yapay zekâya mı ait olmalıdır?

"Kara kutu" problemi de önemli bir sorundur; özellikle derin öğrenme tekniklerinde, sistemin nasıl bir sonuca ulaştığı anlaşılamamaktadır, bu da şeffaflığı ve sorumluluğu zorlaştırmaktadır.

Bu bağlamda, sorulması gereken kritik soru; yapay zekâ sistemlerinin karar süreçleri nasıl daha şeffaf hale getirilebilir? Bu soru, yapay zekânın gelecekteki gelişimi açısından büyük önem taşımaktadır.

4.3. Yapay Zekânın İş Dünyasına ve Topluma Etkileri

Yapay zekâ, iş dünyasında ve toplumsal yapıda köklü değişiklikler yaratarak yeni iş fırsatları yaratırken, iş gücü piyasasında dönüşümlere yol açması, çeşitli sosyal ve ekonomik sorunları da beraberinde getirmektedir.

Otomasyon sistemlerinin yaygınlaşması, birçok sektörde iş gücü kaybına neden olmaktadır. Rutin ve tekrarlayan işlerin yapay zekâ tarafından devralınması, ekonomik eşitsizlikleri artırmaktadır. Yapay zekânın yapmış olduğu bu geçiş süreci, mevcut iş gücünün yeniden yerine oturmasını gerektirdiği için sosyal bir sorun olarak ortaya çıkmaktadır.

Yapay zekâ, iş süreçlerini optimize ederek şirketlerin daha bilinçli kararlar almasına yardımcı olur. Ancak büyük şirketler yapay zekâ tabanlı veri analitiğini

benimserken, küçük işletmelerin bu teknolojiyi kullanabilecekleri yapıları olmayabilir, bu da rekabet avantajını kaybetmelerine yol açmaktadır. Ayrıca, yüksek teknolojiye erişimi olan bireyler ve şirketler daha fazla fırsat elde ederken, düşük gelirli kesimler olumsuz etkilenmektedir.

Bu bağlamda, önemli sorular ortaya çıkmaktadır: İşsiz kalacak bireyler için hangi sosyal politikalar geliştirilmelidir? Yapay zekânın yarattığı ekonomik faydalar nasıl adil dağıtılabilir? Toplumsal eşitsizlikleri önlemek için hangi düzenlemeler ve eğitim programları uygulanmalıdır? Bu sorulara verilecek yanıtlar, yapay zekânın etkilerini dengelemekte ve toplumsal uyumu sağlamada önemli olacaktır.

5. Yapay Zekâ ve Gelecek

5.1. Olası Gelişmeler

Yapay zekâ teknolojisinin hızla gelişmesi, gelecekte birçok alanda köklü değişiklikler yaratacaktır. Özellikte otonom sistemler alanında, tamamen otonom araçların yaygınlaşması beklenmektedir. Bu araçlar, trafik kazalarını azaltma, trafik akışını optimize etme ve yakıt tasarrufu sağlama gibi avantajlar sunarak şehir içi ulaşım sistemlerini yeniden şekillendirecektir. Otonom dronlar ve robotlar da kargo teslimatından tarım ve endüstriyel üretime kadar çeşitli alanlarda kullanılabilecektir. Sağlık ve biyoteknoloji alanında ise yapay zekâ kişiselleştirilmiş tıp uygulamalarını yaygınlaştıracak; genetik ve tıbbi verilere dayanarak özel tedavi planları geliştirilecektir. İlaç geliştirme ve genetik mühendislik alanlarında daha hızlı ve etkili araştırmalar yapılabilecektir. Bu vb. birçok alanda gerçekleşecek değişiklikler olurken olası beklenti genel yapay zekânın ortaya çıkmasıdır.

Genel yapay zekâ (AGI), yapay zekâ teknolojisinin en büyük hedeflerinden biridir. AGI, insan zekâsına benzer geniş bir yetenek yelpazesine sahip olacak ve bilimsel keşiflerden mühendislik projelerine kadar derin etkiler yaratabilecektir. Tüm bu gelişmeler, yapay zekânın hayatın hemen her alanını dönüştüreceğini ve insanlık için büyük fırsatlar sunacağını göstermektedir.

5.2. Zorluklar ve Fırsatlar

Yapay zekâ teknolojisinin geleceği, büyük fırsatlar sunduğu kadar ciddi zorluklar da barındırmaktadır. Bu fırsatları en verimli şekilde değerlendirmek için zorlukların üstesinden gelmek oldukça önemlidir.

Fırsatlar arasında yapay zekânın verimliliği ve otomasyonu artırma potansiyeli bulunmaktadır. Yapay zekâ, üretim, lojistik, sağlık ve hizmet sektörlerinde otomasyonu yaygınlaştırarak maliyetleri düşürebilir ve süreçleri daha hızlı hale getirebilir. Ayrıca, veri bilimcileri ve yapay zekâ mühendisleri

gibi yeni meslek dalları ortaya çıkararak yeni iş fırsatları yaratmaktadır. Eğitimde de kişiselleştirilmiş öğrenme fırsatları sunarak daha başarılı sonuçlar elde edilmesine yardımcı olmaktadır.

Zorluklar ise iş gücü dönüşümü ile başlamaktadır. Yapay zekânın otomasyonu, tekrarlayan işlerin kaybına yol açabilir; bu da iş gücünün yeniden vasıflandırılmasını gerektirmektedir. Ayrıca, yapay zekâ kararlarının sorumluluğu, veri gizliliği ve algoritmaların adilliği gibi etik ve yasal konular, yeni sorunlar doğurmaktadır. Bu nedenle, güvenli ve etik kullanım için yeni düzenlemelere ihtiyaç duyulmaktadır.

Bir diğer zorluk, yapay zekâ teknolojilerinin yaratabileceği güç dengesizlikleridir. Büyük şirketler ve gelişmiş ülkeler, bu teknolojileri daha etkin kullanırken, küçük işletmeler ve gelişmekte olan ülkeler rekabet etmekte zorlanacaktır. Bu durum toplumsal eşitsizlikleri derinleştirerek teknolojilerin adil dağıtılmamasına neden olmaktadır. Yapay zekâ teknolojilerinin sunmuş olduğu bu fırsatları değerlendirmek ve zorlukları aşmak, sürdürülebilir bir geleceği mümkün kılacaktır.

6. Sonuç ve Öneriler

Yapay zekâ, birçok sektörde devrim yaratan bir teknolojidir. Bu bölüm, yapay zekânın tanımından tarihsel gelişimine, günümüzdeki kullanım alanlarından gelecekteki olası gelişmelere, fırsatlara, zorluklara ve toplumsal etkilere geniş bir perspektif sunmaktadır.

Yapay zekâ, insan benzeri zekâyı taklit eden algoritmalar ve sistemler geliştirerek karar verme, problem çözme ve öğrenme süreçlerini otomatize eden bir teknolojidir. Tarihsel olarak, yapay zekâ dar yapay zekâ (ANI) odaklı gelişmelerle başlamış olup, günümüzde genel yapay zekâ (AGI) ve süper yapay zekâ (ASI) gibi daha ileri formlar üzerinde çalışmalar sürmektedir. Makine öğrenmesi, derin öğrenme ve doğal dil işleme gibi teknikler, yapay zekânın temel taşlarını oluşturmuştur.

Yapay zekâ, sağlık, finans, otomotiv, eğitim ve tarım gibi birçok sektörde verimlilik ve inovasyon sağlamaktadır. Tıbbi teşhis, otonom araçlar ve kişiselleştirilmiş eğitim çözümleri gibi geniş bir yelpazede yapay zekâ uygulamaları kullanılmaktadır.

Ancak, veri gizliliği, kararların sorumluluğu ve toplumsal eşitsizlikler gibi etik meseleler, yapay zekânın gelişiminde önemli tartışma konuları olmuştur. Bu sorunların çözümü, teknolojinin güvenli ve adil bir şekilde kullanılmasını sağlamak açısından kritik öneme sahiptir.

Otonom sistemler ve kişiselleştirilmiş sağlık hizmetleri gibi alanlarda büyük atılımlar beklenirken, iş gücü kaybı ve güç dengesizlikleri gibi zorluklar dikkatle

yönetilmelidir. Yapay zekânın gelecekteki rolü, bu teknolojiye nasıl yön verileceğine bağlı olarak şekillenecektir.

Yapay zekâ, henüz tüm potansiyelini gerçekleştirmemiş bir teknoloji olup, daha güvenli, adil ve etkili bir şekilde kullanılabilmesi için çeşitli araştırma ve geliştirme alanlarına ihtiyaç duyulmaktadır. Araştırmacılar, yapay zekânın nihai hedeflerinden biri olan genel yapay zekâya (AGI) ulaşmayı amaçlamaktadır. AGI, insan seviyesinde düşünme kapasitesine sahip yapay zekâ sistemleri geliştirme yolunda büyük bir adım olacaktır. AGI'nin güvenli bir şekilde geliştirilmesi için etik kılavuzlar oluşturulmalı ve insan-AGI etkileşimlerinin toplumsal etkileri derinlemesine incelenmelidir.

Etik boyutunda, yapay zekâ şeffaflığı ve sorumluluğu üzerinde çalışmalar yapılmalıdır. Yapay zekâ sistemlerinin karar alma süreçlerini daha anlaşılır ve izlenebilir kılacak şeffaf algoritmaların geliştirilmesi, kişisel veri güvenliğine büyük önem verilmesi ve güçlü güvenlik algoritmalarının oluşturulması gerekmektedir. Bu adımlar, yapay zekânın güvenli ve etik bir şekilde kullanılmasını sağlayacaktır.

Yapay zekâ, ayrıca çevresel sürdürülebilirlik alanında da önemli bir rol oynayabilir. Yapay zekâ tabanlı çözümler, ekosistemlerin korunmasına katkı sağlayarak doğal kaynakların daha etkin kullanılmasına yardımcı olabilir. Bu, yapay zekânın sadece teknoloji ve iş dünyasında değil, çevreyi koruma ve sürdürülebilirlik hedeflerine ulaşmada da kritik bir araç olabileceğini göstermektedir.

İlerleyen günlerde yapay zekâ, teknoloji dünyasında ve toplumsal yapılar üzerinde köklü bir dönüşüm yaratmaya devam edecektir. Ancak, bu dönüşümün başarılı olabilmesi için yapay zekânın etik, yasal ve toplumsal boyutlarının dikkatle ele alınması gerekmektedir. Gelecekteki çalışmalar, bu teknolojinin adil, güvenli ve sürdürülebilir bir şekilde uygulanmasına odaklanmalıdır.

Kaynakça

1. Chen, L., Xu, G., Zhang, S., Yan, W., & Wu, Q. (2020). Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks. *Journal of Manufacturing Systems*, 54, 1–11. https://doi.org/10.1016/j.jmsy.2019.11.008
2. Chen, X., Du, J., & Zhang, H. (2020). Lipreading with DenseNet and resBi-LSTM. *Signal, Image and Video Processing*, 1–9. https://doi.org/10.1007/s11760-019-01630-1
3. Deng, J., Dong, W., Socher, R., Li, L., Kai, L., & Li, F.-F. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–255). https://doi.org/10.1109/CVPR.2009.5206848
4. Farsal, W., Anter, S., & Ramdani, M. (2018). Deep learning: An overview. *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications*. Rabat, Morocco: Association for Computing Machinery. https://doi.org/10.1145/3289402.3289538
5. Fidan, Ü., & Boza, N. A. (2024). Individual privacy should be an institutional value: Socio-Technical design of university data collection systems. In *Data-Driven Business Intelligence Systems for Socio-Technical Organizations* (pp. 366-384). IGI Global. https://doi.org/10.4018/979-8-3693-1210-0.ch014
6. Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, 5(1), 1. http://dx.doi.org/10.2478/jagi-2014-0001
7. Goge, D. W. (1995). A brief history of unmanned ground vehicle (ugv) development efforts. *Unmanned System Magazine, United States of America*, 13(3), 1–9.
8. Ingle, S., & Phute, M. (2016). Tesla autopilot: semi autonomous driving, an uptick for future autonomy. *International Research Journal of Engineering and Technology*, 3(9), 369–372.
9. Kaur, T., & Gandhi, T. K. (2019). Automated brain image classification based on VGG-16 and transfer learning. *Proceedings - 2019 International Conference on Information Technology, ICIT 2019*, 94–98. https://doi.org/10.1109/ICIT48102.2019.00023

10. Krauss, P. (2024). Game-playing Artificial Intelligence. In *Artificial Intelligence and Brain Research: Neural Networks, Deep Learning and the Future of Cognition* (pp. 125–129). Springer. http://dx.doi.org/10.1007/978-3-662-68980-6_13
11. Kumari, D., & Bhat, S. (2021). Application of artificial intelligence technology in tesla-a case study. *International Journal of Applied Engineering and Management Letters (IJAEML)*, 5(2), 205–218. https://doi.org/10.47992/IJAEML.2581.7000.0113
12. Manning, C., & Schutze, H. (1999). *Foundations of statistical natural language processing*. MIT press.
13. Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson.
14. Sai Bharadwaj Reddy, A., & Sujitha Juliet, D. (2019). Transfer learning with RESNET-50 for malaria cell-image classification. *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, 945–949.https://doi.org/10.1109/ICCSP.2019.8697909
15. Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248.
16. Talaş, U., Yüzgeç, U., & Çubukçu, B. (2021). Object recognizing robot application with deep learning. *Avrupa Bilim ve Teknoloji Dergisi*, (31), 127–133. https://doi.org/10.31590/ejosat.962558
17. Turing, A. M. (2009). *Computing machinery and intelligence*. Springer.
18. Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018(1), 7068349. https://doi.org/10.1155/2018/7068349
19. Wand, M., Koutník, J., & Schmidhuber, J. (2016). Lipreading with long short-term memory. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6115–6119). IEEE. https://doi.org/10.1109/ICASSP.2016.7472852
20. Ye, C., Li, X., & Chen, J. (2019). A deep network for tissue microstructure estimation using modified LSTM units. *Medical Image Analysis*, *55*, 49–64. https://doi.org/10.1016/j.media.2019.04.006


DUVAR


sertifika no : 49837
9 786256 183230
duvaryayinlari.com