

**DEĞİŞKEN ŞARTLAR ALTINDA ATIKSU ARITIMINDA
KARBON VE BESİ MADDESİ GİDERİMİNİN MAKİNE
ÖĞRENMESİ YÖNTEMLERİ İLE MODELLENMESİ**

Cem CANTEKİN
191450202

DOKTORA TEZİ

Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Doktora Programı
Danışman: Dr. Öğr. Üyesi Önder ŞAHİNASLAN

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Ocak, 2024

**DEĞİŞKEN ŞARTLAR ALTINDA ATIKSU ARITIMINDA
KARBON VE BESİ MADDESİ GİDERİMİNİN MAKİNE
ÖĞRENMESİ YÖNTEMLERİ İLE MODELLENMESİ**

Cem CANTEKİN

191450202

ORCID: 0000-0003-2891-8318

DOKTORA TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Doktora Programı

Danışman: Dr. Öğr. Üyesi Önder ŞAHİNASLAN

Eş Danışman: Prof. Dr. Nevin YAĞCI

İstanbul

T.C. Maltepe Üniversitesi

Lisansüstü Eğitim Enstitüsü

Ocak, 2024



JÜRİ VE ENSTİTÜ ONAYI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” ile bildirilen 6698 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.



ETİK İLKE VE KURALLARA UYUM BEYANI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” ile bildirilen 6698 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir



TEŞEKKÜR

Bu çalışmadaki rehberliğinden dolayı danışmanım Dr. Öğr. Üyesi Önder ŞAHİNASLAN' a desteğinden dolayı teşekkür ederim. Maltepe Üniversitesi' nde tarafıma sağlanan fırsatlar ve deneyimler bana gelecekteki tüm çalışmalarında başarılı olmaya devam etmem için gerekli araçları, deneyimleri ve bilgiyi sağlamıştır.

Çalışma süresince eş danışmanlığımı yürüten, Prof. Dr. Nevin YAĞCI' ya teşekkürü borç bilirim. Yüksek lisans döneminde de danışmanlığımı yapmış olan Prof. Dr. Nevin YAĞCI, araştırma, akademi ve hayatla ilgili sorularım olduğunda her zaman yanımda olmuştur.

Prof. Dr. A. Mesut RAZBONYALI' ya rehberliğinden ve akademik hayatımda yeni bir sayfa açabilmeme vesile olmasından dolayı özellikle müteşekkir olduğumu belirtmek isterim.

Ayrıca çocukluk arkadaşım ve can dostum Dr. Öğr. Üyesi Can RAZBONYALI' nın dostluğu ve desteği olmadan doktora eğitimimin ve bu çalışmanın tamamlanmasının imkansız olduğu gerçeğini de paylaşmadan geçemem.

Bugüne kadar hayatıma girip çıkan tüm insanlara, varlıkları sayesinde edindiğim yaşam deneyimleri için de ayrıca teşekkür ediyorum.

Çağımızın en önemli aydınlarından biri olan Profesör Richard Dawkins' e de takdirlerimi sunmak isterim. Onun bilimsel düşünce tarzı, akla ve bilime olan bağlılığı, fikirleriyle tanıştığım günden bu yana yol göstericim olmuştur.

En önemlisi de bugün olduğum kişi olmamda en büyük pay sahibi olan iki kişi; hayallerimin peşinden gitmem için beni motive eden sevgili annem Hülya ve anneannem Müjgan' a minnettarım.

Cem CANTEKİN

Ocak, 2024

ÖZET

DEĞİŞKEN ŞARTLAR ALTINDA ATIKSU ARITIMINDA KARBON VE BESİ MADDESİ GİDERİMİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE MODELLENMESİ

Cem Cantekin

Doktora Tezi

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Doktora Programı

Danışman: Dr. Öğr. Üyesi Önder Şahinaslan

Maltepe Üniversitesi Lisansüstü Eğitim Enstitüsü, 2024

Artan kentleşme, sanayileşme ve nüfus artışı eğilimleri, birbirinden çok farklı ölçeklerin varlığı ve eksik veri toplanması gibi faktörlerle birleşerek çevre konularında belirsizlik oluşturmaktadır. Su kıtlığının artışı, emisyonlarda sıkı sınırlamalar, mevcut ortaya çıkan kirleticiler ve hızla gelişen kentleşme, mevcut atıksu arıtımı ve çevre sağlığı üzerinde büyük sorunlara yol açmaktadır. Atıksu sistemlerindeki kirleticilerin ve besli maddelerinin giderilmesiyle ilgili karmaşıklık, yüksek dinamik yapısı ve modellerin kontrol sistemlerinin olasılıklara, varsayımlara ve tahminlere dayanmasıyla daha da zorlaşmaktadır. Atıksu arıtımının işletme maliyeti ve performansı, hidrolik/katı bekletme süresi, geri dönüşüm oranı, havalandırma oranları gibi işletme parametrelerinden oldukça etkilenmektedir.

Makine öğrenmesi, açıkça programlama uygulanmadan verilere dayalı kalıplar öğrenmeyi ve kararlar vermeyi içeren yapay zeka alanıdır. Makine öğrenmesi algoritmalarının eğitimi, girdi verileri ve çıktı verilerinden oluşan bir veri kümesi üzerinde gerçekleştirilir. Algoritma, eğitim verilerindeki kalıplara dayalı olarak girdi verilerini çıktı verileri ile eşleştiren bir işlev öğrenir. Bu öğrenme süreci sonrasında algoritmaya verilen yeni verilere ait çıktı verilerinin öngörüsünde bulunur. Makine öğrenmesinin birçok farklı türü bulunmakta olup; bunlar arasında yönlendirilmiş öğrenme (supervised), yönlendirilmemiş öğrenme (unsupervised) ve pekiştirme öğrenimi (reinforced) bulunur.

Bu alıřmada, karmařık atıksu arıtımı problemlerini ele almak baėlamında  nemli model olan karar aėacı regresyonu (DTR), destek vektr makineleri (SVM) ve ileri beslemeli sinir aėı (FNN) modellerinin uygulanmasına odaklanılmıřtır. Gerek bir atıksu arıtma tesisinden alınan 6 yıllık veriler kullanılarak atıksu arıtma srecine ait ıkıř atıksuyundaki kimyasal oksijen ihtiyaı (KOİ) ile besi maddesi parametrelerinden toplam azot ve toplam fosfor parametrelerinin seilen modeller tarafından ngr performansları incelenmiřtir.

Anahtar Szckler: Makine ėrenmesi, yapay zeka, yapay sinir aėları, atıksu arıtımı



ABSTRACT

MODELING CARBON AND NUTRIENT REMOVAL IN WASTEWATER TREATMENT UNDER VARIABLE CONDITIONS USING MACHINE LEARNING METHODS

Cem Cantekin

PhD Thesis

Computer Engineering Department

Computer Engineering PhD. Programme

Thesis Advisor: Asist. Prof. Önder Şahinaslan

Maltepe University Graduate School, 2024

The increasing trends of urbanization, industrialization, and population growth, combined with factors such as the existence of vastly different scales and incomplete data collection, contribute to uncertainty in environmental issues. The rise in water scarcity, strict emissions limitations, emerging pollutants, and rapid urbanization pose significant threats on existing wastewater treatment and environmental health. The complexity of removing pollutants and nutrients in wastewater systems is further complicated by their high dynamic nature and the reliance of models and control systems on probabilities, assumptions, and predictions. The operational cost and performance of wastewater treatment are highly influenced by operational parameters such as hydraulic/solid retention time, recycling rate, aeration rates, etc.

Machine learning is an artificial intelligence field that involves learning patterns from data and making decisions without explicit programming. Machine learning algorithms are trained on a dataset consisting of input data matching corresponding labels and output data. The algorithm learns a function that maps input data to output data based on patterns in the training data. There are various types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning.

In this study, focusing on addressing complex wastewater treatment problems, three important models; Decision Tree Regression (DTR), Support Vector Machines (SVM), and Feedforward Neural Network (FNN) have been applied. The performance of these models in predicting parameters such as chemical oxygen demand (COD), total nitrogen,

and total phosphorus in the effluent of an actual wastewater treatment plant has been examined using 6 years of data.

Keywords: Machine learning, Artificial neural networks, Artificial intelligence, Wastewater treatment



İÇİNDEKİLER

JÜRİ VE ENSTİTÜ ONAYI	ii
ETİK İLKE VE KURALLARA UYUM BEYANI	iii
TEŞEKKÜR.....	iv
ÖZET	v
ABSTRACT.....	vii
İÇİNDEKİLER	ix
TABLolar LİSTESİ.....	xiii
ŞEKİLLER LİSTESİ	xiv
KISALTMALAR.....	xvi
1. GİRİŞ	1
1.1 Çalışmanın Amacı.....	2
1.2 Çalışmanın Önemi.....	3
1.3 Varsayımlar	4
1.4 Sınırlıklar.....	4
1.5 Tanımlar	4
2. LİTERATÜR ARAŞTIRMASI	6
2.1 Atıksu Arıtımı	6
2.1.1 Atıksu.....	6
2.1.2 Biyolojik atıksu arıtımı	8
2.1.3 Aktif çamur	9
2.1.3.1 Karbon giderimi.....	10
2.1.3.2 Azot giderimi.....	12

2.1.3.3	Fosfor giderimi	15
2.1.4	Atıksu arıtma tesisi	16
2.1.4.1	Giriş ızgaraları	17
2.1.4.2	Ön arıtım/ön çöktürme	18
2.1.4.3	İkincil arıtma	18
2.1.4.4	Üçüncül arıtma	18
2.1.4.5	Çamur arıtımı.....	19
2.2	Makine Öğrenmesi	20
2.2.1	Yönlendirilmiş öğrenme	21
2.2.2	Yönlendirilmemiş öğrenme	21
2.2.3	Pekiştirme öğrenimi	22
2.2.4	Makine öğrenmesi teorisi.....	23
2.2.4.1	Sapma-varyans ödünleşimi.....	23
2.2.4.2	Aşırı uyum	24
2.2.4.3	Genelleştirme.....	24
2.2.4.4	Optimizasyon.....	25
2.2.5	Makine öğrenmesi algoritmaları	25
2.2.5.1	Doğrusal regresyon.....	26
2.2.5.2	Lojistik regresyon	26
2.2.5.3	Karar ağaçları	27
2.2.5.4	Destek vektör makineleri.....	28

2.3	Yapay Sinir Ağları	28
2.3.1	İleri beslemeli sinir ağları	29
2.3.2	Konvolüsyonel sinir ağları.....	30
2.3.3	Geriye dönük sinir ağları	30
2.3.4	Yapay sinir ağlarının eğitimi	31
2.4	Atıksu Arıtımında Tarihsel Süreç Kontrolü	31
2.5	Makine Öğrenmesine Dayalı Literatürde Yapılan Çalışmalar	33
3.	MATERYAL VE YÖNTEM.....	36
3.1	Karar Ağacı Regresyonu	37
3.1.1	Karar ağacı regresyonunun avantajları	37
3.1.2	Karar ağacı regresyonunun sınırlamaları	38
3.2	Destek Vektör Makineleri	39
3.2.1	Destek vektör makinelerinin avantajları	40
3.2.2	Destek vektör makinelerinin dezavantajları	40
3.3	İleri Beslemeli Sinir Ağları	41
3.3.1	İleri beslemeli sinir ağlarının avantajları	42
3.3.2	İleri beslemeli sinir ağlarının sınırlamaları	43
3.4	Veri Seti	43
3.5	Önerilen Yöntem.....	44
3.5.1	Veri ön işleme	45
3.5.2	Veri setinin bölünmesi	46
3.5.3	Model eğitimi.....	46
3.5.4	Hiper-parametre ayarlama	47

3.5.5	Model deęerlendirmesi	48
4.	BULGULAR VE TARTIřMA	50
4.1	Bulgular	50
4.1.1	Karar aęacı regresyonu	50
4.1.2	Destek vektör makineleri	52
4.1.3	İleri beslemeli sinir aęları	54
4.2	Tartıřma	56
5.	SONUÇ ve ÖNERİLER	60
	KAYNAKLAR	62
	ÖZGEÇMİř	70

TABLÖLAR LİSTESİ

Tablo 1. Çalışmada kullanılan veri setine dair örnek tablo.....	44
Tablo 2. Çalışma kapsamında kullanılan modeller ve değerlendirilen parametreler....	45
Tablo 3. Karar Ağacı Regresyonu, Destek Vektör Makineleri ve İleri Beslemeli Sinir Ağları modelleri performansları	57



ŞEKİLLER LİSTESİ

Şekil 1. Atıksu arıtımında azot giderimin basitleştirilmiş gösterimi	13
Şekil 2. Atıksu arıtma tesisi genel akışı.....	17
Şekil 3. Önerilen yönteme ait genel akış şeması	36
Şekil 4. Çalışmada kullanılan veri seti histogramlarına ait örnek görsel	45
Şekil 5. Uygulanan yöntemin akış şeması	49
Şekil 6. Karar Ağacı Regresyonu modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	51
Şekil 7. Karar Ağacı Regresyonu modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	51
Şekil 8. Karar Ağacı Regresyonu modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	52
Şekil 9. Destek Vektör Makineleri modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	53
Şekil 10. Destek Vektör Makineleri modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	53
Şekil 11. Destek Vektör Makineleri modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	54
Şekil 12. İleri Beslemeli Sinir Ağları modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	55
Şekil 13. İleri Beslemeli Sinir Ağları modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	55

Şekil 14. İleri Beslemeli Sinir Ağları modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması	56
Şekil 15. KOİ konsantrasyonlarının gerçek değerler ile karşılaştırılması	58
Şekil 16. Toplam Azot konsantrasyonlarının gerçek değerler ile karşılaştırılması	58
Şekil 17. Toplam Fosfor konsantrasyonlarının gerçek değerler ile karşılaştırılması	59



KISALTMALAR

AAT	: Atıksu arıtma tesisi
KOİ	: Kimyasal oksijen ihtiyacı
BOİ	: Biyolojik oksijen ihtiyacı
DO	: Çözünmüş oksijen
SRT	: Çamur bekletme süresi
HRT	: Hidrolik bekletme süresi
GHG	: Sera gazları
BNR	: Biyolojik besi maddesi giderimi
SVI	: Çamur hacim indeksi
ML	: Makine öğrenmesi
AI	: Yapay zeka
YSA	: Yapay sinir ağları
CNN	: Konvolüsyonel sinir ağları
RNN	: Yineleyici sinir ağları
DTR	: Karar ağacı regresyonu
SVM	: Destek vektör makineleri
FNN	: İleri beslemeli sinir ağları
RMSE	: Kök Ortalama Kare Hata
MAE	: Ortalama Mutlak Hata

1. GİRİŞ

Su kıtlığı ve kirliliği dünyanın farklı bölgelerinde ortak sorunlar haline gelmiştir. Su kaynaklarının yetersizliğinden kaynaklanan baskıyı hafifletmek için atıksu arıtımı yoluyla suyun yeniden kullanımı kentsel alanlar, sanayi bölgesi vb. gibi farklı su talepleri ve sürdürülebilir üretim için etkili bir yol sağlamaktadır.

Artan kentleşme, sanayileşme ve nüfus artışı eğilimleri, dördüncü sanayi devriminde bilginin belirsizliği ve ölçeklerin çokluğu ile çevre konularında belirsizlik yaratmaktadır. Artan su kıtlığı, emisyonlarda katı sınırlamalar, ortaya çıkan kirleticiler ve hızlı kentleşme, mevcut atıksu arıtımı ve çevre sağlığı üzerinde çok fazla kısıtlamaya neden olmaktadır. Atıksu sisteminde ortaya çıkan kirleticilerin ve besinlerin giderilmesinin karmaşıklığı, yüksek dinamik yapısı, modellerin ve kontrol sistemlerinin olasılıklara, varsayımlara ve tahminlere dayanmasını zorlaştırmaktadır. Atıksu arıtımının işletme maliyeti ve performansı, hidrolik/katı bekletme süresi, geri dönüşüm oranı, havalandırma oranları vb. işletme parametrelerinden oldukça etkilenmektedir.

Biyolojik reaksiyon mekanizmaları, fiziksel, kimyasal reaksiyon, kinetik, kataliz, ayırma, taşıma fenomenleri, ortaya çıkan mikro kirleticiler, atıksu arıtma prosesinin çok değişkenli yönleri arasındaki karmaşık etkileşimler nedeniyle (yüksek oranda doğrusal olmayan ve yüksek oranda zamanla değişen) atıksu arıtma tesisi (AAT) süreçlerinin teşhisini karmaşık hale getirmektedir. AAT verilerinin heterojenliği, eksikliği veya belirsizliği AAT'yi diğer çevresel ve biyoteknolojik endüstriyel süreçlere kıyasla benzersiz bir özelliğe sahip kılmakta olup; atıksu arıtma tesisinin uygun olmayan bakımı, çok çeşitli çalışma koşulları, flora ve faunayı etkileyen ciddi halk sağlığı sorunları ve ekolojik sorunları tetikleyebilmektedir (Hossain vd., 2022).

Atıksu parametrelerinin, özellikle besi maddesi parametrelerinin, işletme tarafından arıtmanın başında ve sonunda tespiti ve hesaplamaları çok önemlidir. Gerekli bilgileri elde etmek için, sensörlerden veya numunelerden veri alınarak ve tesisin giriş/çıkış akışı analiz edilerek ham atıksuyun özellikleri belirlenmelidir. Yetersiz arıtılmış atıksular,

yeraltı suyu sistemleri gibi su kütlelerine girerek birçok sağlık sorununa neden olabilmektedir (Bagherzadeh vd., 2021; Sheikh vd., 2023; Hasan vd., 2021).

Su sektörü, güvenilirliği ve sürdürülebilir arzı artırmak için dijital bir devrim geçirmektedir. Atıksu arıtımı ve proses işletimindeki mikrobiyal faaliyetlerin çeşitliliği, işletim ve çevre koşullarından kaynaklanan doğrusal olmayan dinamik davranışlara (proses değişimi) karşı kompanse edilmesi zorunluluğu doğmuştur. Yüksek verimlilik ve düşük maliyetli atıksu arıtımı elde etmek, gelişmekte olan ülkelerde/ekonomilerde yaygın bir zorluktur. Optimizasyon araçları, operasyon ve süreç tasarımı, veriye dayalı sensörlerin kullanımı ve pahalı kontrol sistemlerinin ön şartı olmaksızın optimal strateji ve özel beceriler, su ve atıksu dijitalizasyonunda büyük paydaşları cezbetmektedir (Matheri vd., 2021).

Matematiksel modelleme ve simülasyon araçları biyolojik atıksu arıtımının optimizasyonu için değerli araçlardır (Cortes ve Vapnik, 1995; Guo vd., 2015). Modeller, atıksu arıtma proseslerinin karbon oksidasyonu, denitrifikasyon, nitrifikasyon, fosfor giderimi ve biyokatı zenginleştirme yoluyla dijital dönüşümünü yapabilmektedir (Andrade Cruz vd., 2022). Veri madenciliği, AAT operasyonları için veriye dayalı modellere uygulanan gelişmekte olan teknolojilerden bir tanesidir. Hata kontrol sistemlerinden kaynaklanan eksik veriler (veri kümeleri), veri analizi, tahmin ve modellemede öğrenme ve sınıflandırma doğruluklarını etkilediği bilinmektedir (Gandomi ve Haider, 2015; Hashem vd., 2015; Goodfellow vd., 2016). Uygun yönetim sistemleri, farklı mühendislik, büyük veri, yapay zeka, derin öğrenme (DÖ), makine öğrenimi (ML), veri bilimi/analitiği, veri madenciliği, otomasyon, blok zinciri teknolojisi, biyoteknoloji, mikrobiyoloji, veri akışından, sosyal bilimler ve diğer bilimsel alanlardan çok disiplinli yaklaşımlar gerektirmektedir.

1.1 Çalışmanın Amacı

Makine öğrenmesi (ML), bilgisayar bilimi ve istatistiğin kesiştiği noktada ve yapay zeka (AI) ile veri biliminin merkezinde yer alan en hızlı büyüyen teknik alanlardan biridir. Makine öğrenmesinin etkisi, tüketici hizmetleri, bankacılık, astronomi ve ampirik bilimler gibi yoğun veriler içeren bir dizi endüstride geniş çapta hissedilmektedir. Atıksu arıtma alanında ise veri üretimi, atıksu arıtma tesislerinin (AAT) otomasyonu ile birlikte

ortaya çıkmıştır. Ek olarak, hesaplama ve depolama kapasitesinin artması, su sektöründe farklı kaynaklardan gelen büyük miktarda bilginin depolanmasına olanak sağlamıştır.

Bu çalışmada, aktif çamur bazlı bir AAT'de gerçek zamanlı optimal kontrole odaklanmak hedeflenmektedir. Veriye dayalı yöntemler kullanılarak AAT'nin değişken tahmini ve proses kontrolünün iyileştirilmesine yönelik mevcut yöntemlerin incelenerek performans karşılaştırılmasının yapılması amaçlanmaktadır.

1.2 Çalışmanın Önemi

Artan nüfusun sonucu olarak kentleşmenin ve sanayileşmenin hızlanmasıyla birlikte büyük miktarda atıksu üretilmektedir. Atıksu arıtma tesisleri (AAT'ler), kirleticileri ortadan kaldırmak ve insan faaliyetlerinden kaynaklanan atıksuların çevre kirliliğini azaltmak için tasarlanırlar. Genel olarak atıksuyun arıtılmasında dört aşama vardır: ön arıtma, birincil arıtma, ikincil arıtma ve üçüncül arıtma ve her fazın seçilebilecek çok sayıda teknolojisi bulunmaktadır. Ancak, AAT'lerin inşası ve işletilmesi çok çeşitli kaynakları (tatlı su, enerji, kimyasallar vb.) tüketir, sera gazları (GHG'ler) yayar ve atık çamur üretir. AAT'lerin yüksek doğrusalsızlık ve varyasyon nedeniyle kontrol edilmesi ve optimize edilmesi karmaşıktır.

Veriye dayalı süreç kontrolü ve performans analizindeki son gelişmeler, atıksu arıtma endüstrisine maliyetleri düşürme ve operasyonları iyileştirme imkanı sunmaktadır. Bununla birlikte, atıksu arıtma tesislerinde (AAT) büyük veri, AAT'nin benzersiz özelliklerini tam olarak analiz etmek için gereken arka plan bilgisinin entegre edilememesinden dolayı yaygın olarak yeterince kullanılmamaktadır.

Bu bağlamda veriye dayalı teknikler, atıksu arıtımında değerli süreç bilgileri sağlamaktadır. Yapay zeka ve makine öğrenmesi uygulamaları proaktif karar verme süreçlerine yardımcı olarak süreç verimliliğini artırarak işletme maliyetlerini azaltma ve atıksu çıkış kalitesini artırma açısından büyük önem taşımaktadır. Makine öğrenmesinin adaptasyon yeteneği, atıksu arıtım sistemlerini sürekli olarak geliştirmeye imkan tanıma ve atıksu arıtma endüstrisinde devrim yaratma fırsatı açısında büyük önem taşımaktadır.

1.3 Varsayımlar

Bu çalışma atıksu arıtım süreçlerinin tamamını kapsamamaktadır. Sadece biyolojik besi maddesi giderimi süreci içerisinde Kimyasal Oksijen İhtiyacı, Toplam Azot ve Toplam Fosfor parametreleri ele alınmış olup, seçili yapay zeka modelleri bu süreçten elde edilen veri seti kullanılarak incelenmiştir.

Toplanan verilerin ve kullanılan modellerin gelişime açık olması, daha büyük veri setleri veya hibrit modellerin entegre edilmesi ile modellerin daha etkin hale getirilebileceği göz önünde bulundurulmalıdır.

1.4 Sınırlıklar

Çalışmada kullanılan veri seti besi maddesi giderimi de yapan bir biyolojik atıksu arıtım tesisine ait gerçek verilerdir. Bu veriler arıtım süreci kontrolünün doğası gereği laboratuvar ortamında yapılan deneysel ölçümlere dayanmaktadır. Bu ölçümlerden kaynaklanabilecek olası hataların yanı sıra beklenmedik deşarjlardan kaynaklı olarak verilerde sapmalar olabilmektedir. Bu tarz sapmalar ve ekstrem değerler veri setinin çalışmaya hazırlanması kapsamında gerekli düzenlemeye tabi tutulmuştur.

Veri setine ek olarak makine öğrenmesi aşamasında kullanılacak olan modellerin de bazı sınırlamaları bulunabilmektedir. Aşırı uyum (over-fitting), önyargılı yapılar oluşturma, hesaplama yapısının karmaşıklıkları sebebiyle hesaplama süresinin uzun olabilmesi, hiper-parametre seçimine karşı aşırı duyarlılık gibi sınırlamalar uygulama aşamasında göz önünde bulundurulmalıdır.

1.5 Tanımlar

Atıksu: Kalitesi insan faaliyetleri tarafından olumsuz etkilenen su kütlesi.

Atıksu arıtımı: Çeşitli insan faaliyetleri sonucu kirletici yüklere maruz kalan suyun belirli kimyasal, fiziksel ve biyolojik süreçlerin uygulanması sonucu çevreye ve insan sağlığına tehlike arz etmeyecek karakterizasyona getirilmesi işlemidir (Mackenzie, 2010).

Makine öğrenmesi: Bilgisayarların açık bir programlama uygulanmadan verilerden öğrenmeleri ve tahminler yapmalarını sağlayan algoritmaların kullanıldığı yapay zeka alanıdır (Bishop, 2006).

Yapay Zeka: Canlılarda olan biyolojik zekaya özgü olan algılama, öğrenme, çözüm üretme, iletişim kurma, çıkarım yapma gibi fonksiyonları otonom olarak yerine getirebilen yapay yazılımlar (Poole vd., 1998).

Algoritma: Matematikte ve bilgisayar biliminde bir işi veya hesaplamayı yapmak için tanımlanan, bir başlangıç durumundan başladığında, açıkça belirlenmiş bir son durumunda sonlanan, net olarak tanımlanmış sonlu işlemler bütünü.



2. LİTERATÜR ARAŞTIRMASI

Atıksu arıtma tesislerinden elde edilen geniş veri setlerini inceleyen makine öğrenmesi modelleri, deşarj özelliklerini önceden tahmin edebilir, bu da proaktif karar alma süreçlerine katkı sağlar. Bu modeller, çevresel standartlara uyumu sağlamak ve işletme maliyetlerini azaltmak konusunda etkiye sahiptir. Ayrıca, hızlı anormallik tespiti yaparak sorunları hızlı bir şekilde saptamayı kolaylaştırabilirler. Makine öğrenmesi uygulamaları atıksu arıtma tesislerinin işletme süreçlerini optimize etmelerine ve mevcut kaynakları daha etkili bir şekilde kullanmalarına olanak tanır. Ayrıca, çevresel sürdürülebilirlik çabalarına katkıda bulunarak su kaynaklarının korunmasına yardımcı olur. Dolayısıyla, makine öğrenmesi, atıksu arıtımındaki operasyonel etkinlik, çevresel uyum ve sürdürülebilirlik açısından önemli bir araçtır, atıksu arıtma süreçlerine önemli avantajlar sunar.

2.1 Atıksu Arıtımı

Belediye atıksu arıtma tesisleri (AAT), besin maddeleri, patojenler ve organik ve inorganik kirleticiler nedeniyle su yollarının kirlenmesini önlemek için hayati öneme sahiptir. Atıksu çevreye deşarj edilmeden önce arıtılmazsa hastalıklar hızla yayılabilir (Feachem vd., 1983); çürüyen organik bileşikler, sucul yaşamı sürdürmek için gereken çözülmüş oksijeni (DO) tüketebilir (Bunzel vd., 2013); metaller ciddi akut ve kronik toksikolojik etkilere sahip olabilir (Akpör, 2014); fazla besin maddesi, ötrofikasyona veya DO tüketen ve potansiyel olarak toksik bileşikler üreten alglerin aşırı büyümesine neden olabilir (Chislock vd., 2013); ve farmakolojik olarak aktif bileşikler, düşük konsantrasyonlarda bile, doğal bakteri popülasyonlarında direnç katkıda bulunabilir ve balık ve sucul hayvanların endokrin fonksiyonlarını bozabilir (Hernando vd., 2006). Gelişmiş ülkelerin çoğunda, evsel atıksu merkezi bir tesise iletilir ve bu tesislerde arıtım işlemi gerçekleştirilir.

2.1.1 Atıksu

Genel olarak atıksu, kalitesi insan faaliyetleri tarafından olumsuz etkilenen su olarak tanımlanabilir. Bu tanım suyun, geniş bir madde yelpazesi ile kontamine olabileceği

anlamına gelmektedir. Olası kirleticilerin çeşitliliğine rağmen, tipik biyolojik arıtma süreci nitrifikasyon ve denitrifikasyon kullanarak karbon, azot ve fosforla birlikte ana kirliliğin giderimine odaklanır. Buna ek olarak, farklı mekanik adımlar vasıtasıyla partikül maddeler atıksudan giderilir.

Evsel atıksu temelde haneler veya benzeri tesisler tarafından kullanılan su olup; iki gruba ayrılabilir: 1. Başlıca olarak insan veya hayvanlardan gelen dışkı ve idrar tarafından etkilenen siyah su. 2. Yıkama, yıkama, pişirme ve benzeri faaliyetlerden kirlenen gri su. Tüm durumların çoğunda siyah ve gri su ayrılmaz ve karışım halinde kanalizasyon sistemine ulaşır. Arıtma süreci için önemli olan şey, bu faaliyetlerin atıksuya belirli bir kirletici / besi maddesi bileşimi kazandırması gerçeğidir. Belediye atıksuyu terimi, bir şehirden gelen atıksuyu ifade eder ve evsel ve endüstriyel atıksuyun bir karışımı anlamına gelir.

Endüstriyel atıksu, üretim, temizleme veya soğutma gibi endüstriyel süreçler sırasında kirlenen sudur. Endüstriyel atıksuyun bileşimi evsel atıksudan tamamen farklı olabilir ve deşarj eden endüstri türüne büyük ölçüde bağlıdır. Genel olarak endüstriyel atıksular yüksek KOİ değerlerine sahip olmakla birlikte endüstri türüne göre yüksek miktarda azot ve diğer kirletici yükler de içerirler. Üretim süreci parti şeklinde düzenlenmişse, atıksu karakterizasyonu günden güne önemli ölçüde değişebilir. Bu belirsizlikler, kontrol ve optimizasyon tekniklerinin uygulanmasını karmaşıktırır.

Tarımsal atıksu, tarımsal alanlarda ve çiftliklerden kaynaklanan gübreler, pestisitler, dışkılar vb. Kirletici yüklerini içeren atıksulardır. Arıtma tesisleri için, genellikle canlı üreme tesisleri gibi nokta kaynaklardan gelen tarımsal atıksu ilgi odağıdır. Tarlalardaki kirli su genellikle toplanmaz ve doğrudan yeraltı suyuna veya su kütlesine akar.

Sızıntı suları özel bir atıksu türüdür. Sızıntı suları, genellikle kanalizasyon sistemine sızan su (genellikle yeraltı suyu) anlamına gelir. Kanalizasyonda diğer atıksu türleriyle karıştırılır ve atıksu arıtma tesislerine taşınır. Atıksu arıtma tesislerine girmesi sebebiyle genellikle diğer atıksu türlerinin seyrelmesine yol açar. Optimizasyon veya kontrolör geliştirme sırasında, yıl boyunca miktarı değişir ve yeraltı su seviyesi ile değişir, bu nedenle dikkate alınmalıdır.

Yüzeysel atıksular, sokaklardan, otoparklardan ve temel olarak kanalizasyon sistemine bitişik tüm sızdırmaz yüzeylerden gelen kentsel akış suyudur. Bu nedenle, yağ, yakıt, lastik, egzoz, yol aşınması vb. ile kontamine olabilir. Arıtma tesisi ve kontrol stratejisi için yağış olayları sırasında önemli hale gelir. Kombinasyonlu bir kanalizasyona bağlandığında, arıtma tesislerine giren atıksu miktarını önemli ölçüde artar.

2.1.2 Biyolojik atıksu arıtımı

Atıksu arıtma tesislerinde biyolojik işlemler genellikle karbon, azot ve fosforun giderilmesine odaklanır. Süreçler ayrı ayrı ele alınabilir olsa da, birbiriyle yakından ilişkili olup ve birbirini etkilemektedir. Belirtilen kirletici yüklerden farklı olarak, arıtma süreci boyunca kısmen giderilebilir olmalarına rağmen farmakolojik bileşikler veya ağır metaller gibi diğer maddeler özellikle hedeflenmemektedir. Genel konsept olarak giderim daha çok maddelerin çamura bağlanması ve bu şekilde bertaraf edilmesi şeklinde tanımlanabilir (Henze vd., 2008).

Doğal su kaynaklarına arıtılmamış atıksu deşarjı ciddi çevre kirliliğine neden olabilir. Atıksu, özellikle organik madde, azot (N) ve fosfor (P) yüksek konsantrasyonları içerir ve bunlar alıcı su kaynaklarında ötrofikasyona neden olabilir (Metcalf ve Eddy, 2003). Ötrofikasyon, alg ve su bitkilerinin aşırı büyümesine yol açarak, düşük oksijen seviyelerine, balık ölümlerine ve su ekosistemlerinin tahribine neden olabilir. Bu nedenle, atıksudan besi maddelerinin giderilmesi, insan ve çevre sağlığını korumak için kritik öneme sahiptir.

Biyolojik besi maddesi giderimi (BNR), azot ve fosforun atıksudan giderimi için etkili ve sürdürülebilir bir yöntemdir. BNR işlemi, nitrifikasyon ve denitrifikasyon süreçleriyle organik madde ve besin maddelerinin giderimini içerir. BNR prosesi iki adımdan oluşur: organik madde ve besi maddesi giderimi. İlk adımda, mikroorganizmalar atıksudaki organik maddeleri tüketerek karbondioksit ve suya dönüştürürler. İkinci adımda, mikroorganizmalar nitrifikasyon ve denitrifikasyon süreçleri yoluyla atıksudan azot ve fosfor giderilirler.

Bir biyolojik atıksu arıtma prosesinde pH, sıcaklık, çözülmüş oksijen (DO), hidrolik bekletme süresi (HRT) ve çamur bekletme süresi (SRT) gibi çeşitli temel karakteristik

parametreler yer almaktadır. pH seviyesi mikroorganizma aktivitesini etkiler ve BNR için optimal pH aralığı 7,0 ile 8,0 arasındadır. Sıcaklık da BNR işleminde önemli bir rol oynar ve optimal sıcaklık aralığı 20°C ile 30°C arasındadır.

Çözünmüş oksijen (DO), nitrifikasyondan sorumlu olan aerobik mikroorganizmaların büyümesi ve aktivitesi için önemlidir. Nitrifikasyon aşamasında DO seviyesi 2-3 mg/L seviyesinde tutulmalıdır. Denitrifikasyon sırasında DO seviyesi düşük tutulmalıdır, 0,5 mg/L'nin altında olmalıdır.

Hidrolik bekletme süresi (HRT), atıksuyun arıtma sisteminde kalma süresidir. Daha uzun bir HRT, mikroorganizmaların atıksuyla daha uzun süre maruz kalmasına izin verir ve daha yüksek giderim verimlilikleri sağlar. Çamur bekletme süresi (SRT), mikroorganizmaların arıtma sisteminde kalma süresidir. Daha uzun bir SRT, sağlam bir mikrobiyal topluluğun gelişmesine olanak tanır ve arıtma sisteminin kararlılığını ve performansını artırabilir (Mackenzie, 2010).

2.1.3 Aktif çamur

Aktif çamur, atıksu arıtma tesislerinde yaygın olarak kullanılan bir tekniktir. Bu teknik kullanılarak atıksuyun mikroorganizmalarla oluşturulan biyokütle ile birleştirilerek çökelti oluşturulması prensibine dayanmaktadır.

Tüm aktif çamur tesislerinde, çamurun son arıtma işleminde sudan ayrılması; çöktürme işlemi gerçekleştirilmektedir. Ayrılan biyokütle daha sonra belirli oranlarda biyoreaktörlere geri pompalanır. Biyokütlenin sürekli taze atıksu ile beslenmesi nedeniyle biyokütle artar. Bu, geri pompalanan çamurun bir kısmının, yani fazla çamurun ayrılması ve çamur arıtma işlemine yönlendirilmesini gerektirir (Henze vd., 2008).

Avantajlarının yanı sıra, arıtma süreci sırasında farklı işletme sorunları ortaya çıkabilmektedir. Tipik sorunlar, şişen çamur ve yüzen çamurdur. Her iki durumda da arıtılan atıksuyun biyokütleden ayrılması düzgün gerçekleşmez. Şişen Çamur, filamentli bakterilerin, yavaş çökme koşullarına yol açan büyüme avantajı elde ettiği durumlarda ortaya çıkar. Çamur, Çamur Hacim İndeksi (SVI)' nin 150 ml/g'den büyük olduğu durumlarda şişmiş çamur olarak kabul edilir (Lee ve Lin, 2007). Yükselen Çamur/Yüzen Çamur, son arıtmada denitrifikasyon meydana geldiğinde ortaya çıkar, bu da

karbondioksit gazı üretir ve *Nocardia* gibi belirli bakterilerin büyümesi sonucu daha etkin hale gelmektedir. Sonuç olarak, çökeltme tankında yüzen bir çamur tabakası oluşur. Şişen çamur ve yükselen çamur sorunları her ikisi de çamur çıkışına ve deşarj değerlerinin ihlaline yol açabilmektedir (Henze vd., 2008).

2.1.3.1 Karbon giderimi

Tarihsel olarak, karbon atıksu arıtma tesislerinde hedeflenen ilk parametredir ve 1970'lerin ortasına kadar tesisler yalnızca organik madde giderimi için inşa edilmiştir. Bu tesisler çamur istikrar kazanana kadar sürekli havalandırılmıştır. Bu işlem, yüksek miktarda enerji gerektirmektedir ancak zamanın düşük enerji maliyetleri nedeniyle uygulanmasında sakınca görülmemiştir (Ludzack ve Ettinger, 1962). Karbon giderimi Biyolojik Oksijen İhtiyacı (BOİ) ve Kimyasal Oksijen İhtiyacı (KOİ) parametreleri üzerinden takip edilmektedir.

Kimyasal Oksijen İhtiyacı, su içindeki organik ve inorganik maddelerin kimyasal oksidasyonu için gereken oksijen miktarını ölçen, sıkça KOİ olarak kısaltılan, nicel bir ölçüdür. Kimyasal oksijen ihtiyacı (KOİ), suda bulunan bir organik bileşiğin tamamen oksidasyonu için gereken oksijen kütlesi olarak tanımlanır. Glikozun (CH_2O)₆ oksidasyonu için temsili bir hidrokarbon olan 1 mol O_2 / mol C gereklidir (Wiesmann vd., 2007). Atıksu bağlamında, KOİ özellikle organik kirlilik yükünü değerlendirmek için temel bir metrik olarak hizmet etmektedir. Bu ölçüm, karbonhidratlar, yağlar, proteinler ve diğer biyolojik olarak parçalanabilir maddeler de dahil olmak üzere geniş bir organik bileşen yelpazesini kapsamaktadır.

Atıksu içindeki yüksek KOİ seviyeleri, karmaşık organik bileşiklerin varlığını belirtir. Bu bileşikler, endüstriyel deşarjlar, evsel kanalizasyon ve doğal süreçler kaynaklı olarak atıksudaki genel organik yükü artırır. KOİ seviyelerinin yönetilmesi ve azaltılması, organik kirliliğin etkilerini kontrol etmek açısından önemlidir.

KOİ, Biyokimyasal Oksijen İhtiyacı (BOİ) ile yakından ilişkilidir. BOİ, organik madde biyolojik olarak parçalandığında mikroorganizmalar tarafından tüketilen oksijen miktarını gösterir. KOİ, toplam oksijen talebini ölçerken, BOİ, organik kirleticilerin biyolojik olarak parçalanabilir fraksiyonu hakkında bilgi verir. Yüksek KOİ seviyeleri

genellikle yüksek BOİ'yi ima eder, bu da sucul ortamlarda oksijen tükenmesi potansiyeli açısından daha büyük bir risk anlamına gelmektedir.

Dünya genelinde çevresel düzenleyici kuruluşlar, atıksu arıtma tesislerinden deşarj edilen suyun KOİ seviyeleri için sıkı standartlar belirlemektedir. Bu standartlara uyum, yalnızca yasal bir gereklilik değil, aynı zamanda deşarj edilen suyun çevreye bir tehdit oluşturmamasını sağlamaktadır. Etkili karbon giderimi, bu düzenleyici hedeflere ulaşmak için önemlidir.

Biyolojik karbon giderimi, atıksu arıtımının temel taşlarından biridir ve etkili giderim stratejileri tasarlamak için teorik temellerin anlaşılması önemlidir. Bu süreç, özellikle atıksudaki organik karbon bileşiklerini metabolize etmek ve parçalamak için mikroorganizmaların kullanılmasını içerir. Bakteri ve mantarları içeren mikroorganizmalar, biyolojik karbon gideriminde merkezi bir rol oynar. Bu mikroorganizmalar, karmaşık organik bileşikleri basit formlara çevirmek için gerekli olan enzimatik mekanizmalara sahiptir. Bu mikroorganizmaların metabolik faaliyetleri, organik karbonu mikrobiyal biyokütle ve karbondioksit gibi son ürünlere dönüştürür. Atıksu, etkili karbon giderimi için çeşitli organik bileşikler içerir, bu da etkili bir mikrobiyal topluluğun varlığını gerektirir. Farklı mikrobiyal türler, belirli organik bileşiklerin parçalanmasına uzmanlaşmıştır. Mikrobiyal konsorsiyum, geniş bir organik kirleticiler yelpazesinin parçalanmasını sağlamak için işbirliği yapar.

Biyolojik karbon giderimi, hem aerobik (oksijenli) hem de anaerobik (oksijensiz) koşullar altında gerçekleşir. Aerobik koşullarda mikroorganizmalar, organik karbonu tamamen karbondioksit ve suya oksidize etmek için oksijeni kullanır. Anaerobik koşullarda ise organik karbon metana ve karbondioksite dönüşür. Aerobik ve anaerobik koşullar arasında seçim, atıksu özellikleri ve arıtma hedefleri gibi faktörlere bağlıdır.

Azot giderimi uygulayan tesislerde, denitrifikasyon için yeterli konsantrasyonlarda karbon bulunması gerekmektedir, ancak arıtılmış su tesislerden ayrılmadan önce yeterince giderilmelidir. Bu, tesisin işletme mühendisliğini ve kontrolünü karmaşıktırır. Süreçte karbon özellikle iki süreç için kullanılır:

- Heterotrofik biyokütle büyümesi.
- Çamurun endojenik solunumu.

Tesisin kontrolü için iki husus dikkate alınmalıdır:

- Denitrifikasyon süreci sırasında yeterli KOİ'nin olması.
- Deşarjı standartlarını sağlayabilmek için yeterli KOİ gideriminin sağlanması (Mackenzie, 2010; Metcalf ve Eddy, 2003).

Etkili biyolojik karbon giderimi, organik kirleticilerin suya salınmasını engeller. Kontrolsüz organik kirlilik, sucul ekosistemlerde oksijen tükenmesine yol açarak balıklar ve diğer sucul organizmaların sağlığını tehlikeye atabilir. Biyolojik karbon giderimi, ekosistemlerin dengesini korumaya yardımcı olur.

Organik karbon giderimi, sucul ortamlardaki besin yükünün azaltılmasıyla yakından ilişkilidir. Organik maddenin parçalanması, azot ve fosfor bileşiklerinin giderilmesini içerir ve ötrofikasyon riskini azaltır. Aşırı besinler, alg patlamalarına neden olabilir ve su sistemlerinin ekolojik dengesini bozabilir.

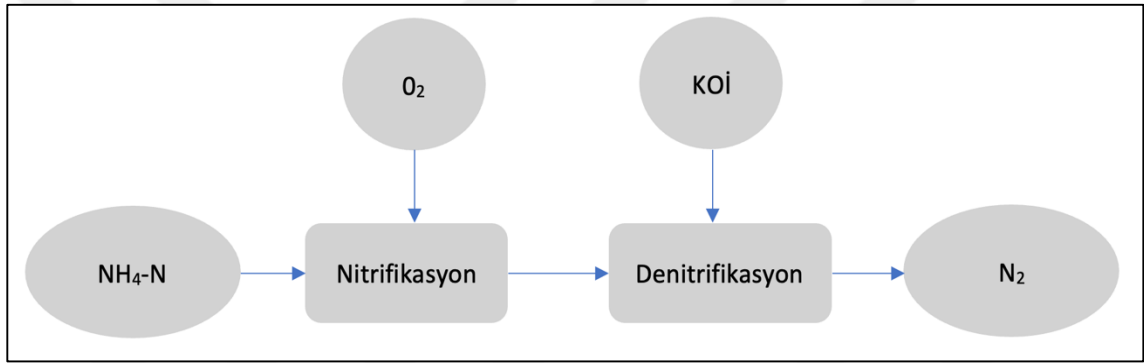
Yüksek seviyede organik kirleticiler içeren atıksu, halk sağlığı için bir risk oluşturur. Patojenler ve zararlı mikroorganizmalar bu ortamlarda gelişebilir. Etkili biyolojik karbon giderimi, deşarj edilen suyun zararlı kirleticilerden arındığından emin olur, böylece halk sağlığını korur.

Çevre düzenlemelerine uyum sağlamak, atıksu arıtma tesisleri için KOİ standartlarını karşılamak önemlidir. Uyumluluk, yalnızca yasal bir gereklilik değil, aynı zamanda sürdürülebilir uygulamalarla uyumlu olmayı gerektirir. Etkili biyolojik karbon giderimi, su arıtma süreçlerinin genel sürdürülebilirliğine katkıda bulunur.

2.1.3.2 Azot giderimi

Azot, su kalitesi yönetimindeki ana parametrelerden biridir. Amonyum ($\text{NH}_4\text{-N}$) su bitkilerinin temel besin maddesidir ve aşırı konsantrasyonları ötrofikasyona neden olur. İyonize olmayan amonyum (NH_3) su ortamı için toksiktir ve bakteriyel nitrifikasyon süreçlerinden dolayı çözünmüş oksijen talebi artar. Alıcı ortamlarda ötrofikasyon,

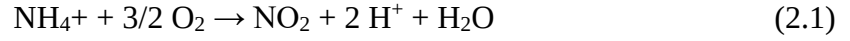
çözünmüş oksijen tükenmesi ve toksisite sorunları, azot parametresinin kontrol edilmesiyle yönetilir. Günümüzde azot giderimi biyolojik işlem süreçleriyle başarıyla gerçekleştirilmektedir (Orhon and Artan, 1997). İki temel adımdan oluşur: 1. Nitrifikasyon ve 2. Denitrifikasyon. Nitrifikasyon, amonyumun aerobik koşullar altında nitrata dönüştürüldüğü bir süreç olup; denitrifikasyon ise nitratin anoksik koşullar altında azot gazına dönüştürüldüğü bir süreçtir. Şekil 1’de görülebileceği gibi bu sürecin işlemesi, idrardaki başlıca maddelerden biri olan ürenin, önceki bir adımda üre hidrolizi yoluyla amonyuma dönüştürülmesini gerektirmektedir. Bu hidroliz süreci genellikle atıksuyun arıtma tesisine ulaşmasından önce kanalizasyon sistemlerinde gerçekleşir (Mackenzie, 2010).



Şekil 1. Atıksu arıtımında azot giderimin basitleştirilmiş gösterimi

Denitrifikasyon, atıksu içeriğinin çoğunlukla nitrat formunda olduğu durumlarda uygulanabilir. Bununla birlikte, evsel atıksu genellikle oksitlenmiş azot içermez. Bu nedenle, belediye uygulamalarında azot kaldırılmadan önce atıksudaki organik ve amonyak azotunun öncelikle nitrit (NO_2^-) veya nitrat (NO_3^-) formuna oksitlenmesi gerekmektedir. Süreç yapılandırması açısından, bu, biyolojik denitrifikasyonun nitrifikasyon ile çiftleştirilmesini gerektirir (Orhon and Artan, 1997).

Nitrifikasyon, iki aşamalı bir süreç olup, sırasıyla amonyumun nitrite ve nitrata dönüşümüdür. Aerobik koşullarda, *Nitrosomonas*, *Nitrosococcus*, *Nitrosopira*, *Nitrosovibrio* ve *Nitrosolobus* gibi amonyak oksitleyen nitrikler, amonyumu nitrite (NO_2^-) oksitleyebilirler (Liu, 2008). Amonyumun nitrite oksidasyonu;



ve nitritin nitrata oksidasyonu;



Nitrobacter, nitrit oksidasyonundan sorumlu bakteriler arasında en ünlü cinstir. Nitrifikasyon bakterileri ototrof, kemolitotrof ve zorunlu aeroblardır. Ototrof olduklarından, nitrifikasyonun yüksek enerji gerektiren bu sürecinden dolayı nitrifikasyonun çok daha küçük bir biyokütle verimi (Y) vardır (Rittmann and McCarty, 2001).

Nitrifikasyon, karbon gideriminden daha zor gerçekleşen bir süreçtir, çünkü ototrofik biyokütle heterotrofik biyokütleden yaklaşık beş kat daha yavaş büyümektedir. Aynı nedenle karbon amonyaktan daha hızlı tüketilir. Bu sebepten ötürü atıksu arıtma tesisinin düzgün işletilmesinde atıksuyun doğru C:N:P oranına sahip olması önem taşımaktadır (Wiesmann vd., 2007; Orhon ve Artan, 1997).

Denitrifikasyonda nitrat (NO_3^-) veya nitrit (NO_2^-), anoksik koşullarda mikrobiyal solunumda çözülmüş oksijenin yerine geçer. Denitrifikasyon ve karbon giderimi, farklı elektron alıcıları kullansalar da aynı heterotrofik biyokütle üzerine dayanır. Azot giderimi için işletilen tek çamur aktif çamur sistemlerinde, heterotrofik biyokütle, anoksik koşullar altında nitrat tüketmeye başlayacak şekilde kendini adapte etmesi beklenir (Artan ve Orhon, 2005).

Azot giderimine ek olarak, denitrifikasyon oksijen geri kazanımında da fayda sağlamaktadır. Bu durum, nitrifikasyonda nitrat içinde bulunan oksijenin bir kısmının karbon gideriminde kullanılabileceği anlamına gelmekte olup, uygulamada yaklaşık %50 geri kazanım oranı elde edilebilmektedir (Henze vd., 2008).

Bu süreç, anaerobik/aerobik koşullar, organik substratlar ve uygun pH ortamını gerektirir (Wiesmann vd., 2007). Çamur atık olmaması, aktif çamur sürecinde yavaş büyüyen nitrifikatörlerin ve denitrifikatörlerin aktivitesi için çamurun en aza indirilmesinin mümkün olacağı anlamına gelir. Mikroorganizmalar süreç boyunca canlı kaldığı sürece çamurun en aza indirilmesi mümkün olur. Geleneksel çamur atık modu bu yavaş büyüyenler için popülasyon kaybına neden olabilir. Öte yandan sonsuz SRT'yi başarabilecek çamur azaltma, bu bakterilerin büyümesi için bir çözüm olabilir (Liu, 2008).

Denitrifikasyon prosesinde karbon yükü kısıtlayıcı bir parametre olabilmekte olup, bu sorunu farklı tesis tasarımları ile veya ek karbon kaynağı kullanarak aşmak mümkün olmaktadır (Mackenzie, 2010). Mükemmel bir verimli sistem, KOİ/N oranı 4.2 olan azotu giderebilir, ancak uygulamada genellikle 5-10 arasındaki oranlar gereklidir (Henze, 1991; Artan ve Orhon, 2005). Gelen atık akımları her zaman kabul edilebilir azot giderimine ulaşmak için yeterli karbon sağlamaz ve KOİ kolayca giderilebildiğinden, denitrifikasyon için kullanılabilir konsantrasyon, girişten önemli ölçüde daha düşük olabilir.

2.1.3.3 Fosfor giderimi

Atıksularda fosfor üç formda bulunur; ortofosfat, polifosfat ve organik bağlı fosfor. Basit ortofosfatlar ve polifosfatlar toplam anorganik fosforu oluşturur. Polifosfatlar, asit hidrolizi yoluyla kolayca ortofosfata dönüştürülebilirken, organik maddeyi tamamen oksidatif olarak parçalamak, organik fosforun ortofosfat olarak serbest kalması için gereklidir. Ortofosfatlar mikroorganizmalar tarafından kolayca asimile edilir. Diğer fosfor formları, sadece ortofosfatlara hidroliz sonucunda mikroorganizmalar tarafından erişilebilir hale gelir (Wiesmann vd., 2007; Artan and Orhon, 2005).

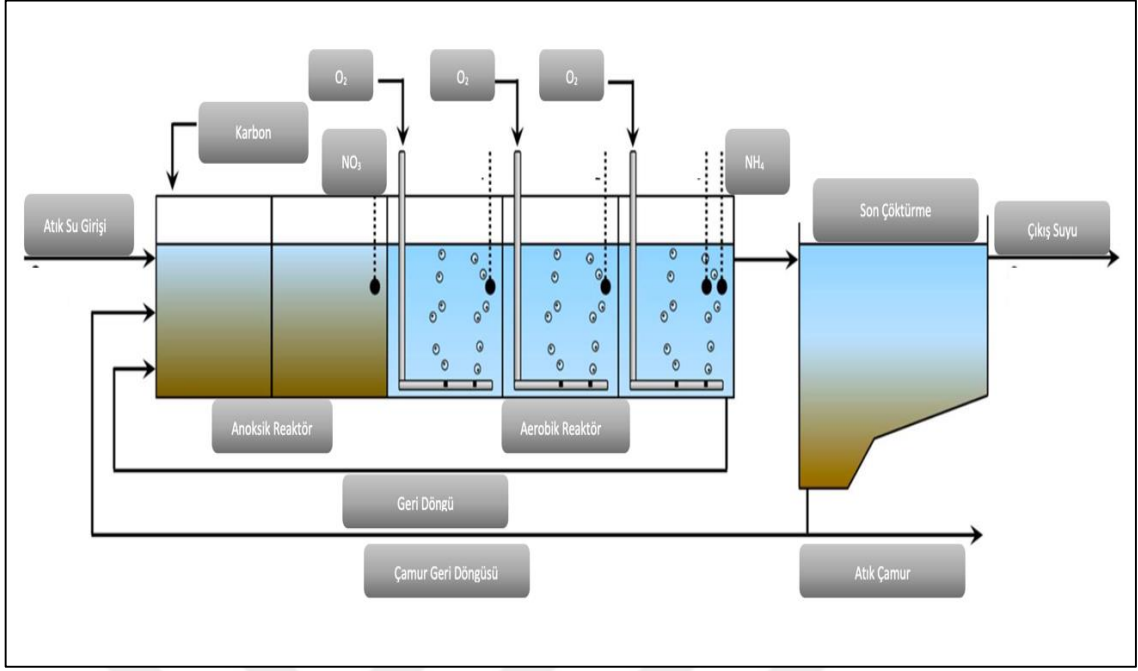
Geçmişte, fosfatın çoğu zaman kimyasal işlemler kullanılarak giderildiği dönemler olmuştur. Bu işlem, atıksuya demir veya alüminyum tuzları gibi maddelerin eklenmesini içerir, örneğin $FeSO_4$ veya $AlCl_3$ gibi. Bu proseste çözünmüş fosfat bileşikleri çözünmeyen fosfatlara dönüştürülür ve daha sonra çamur ile birlikte tesislerden giderilmektedir. En yaygın kullanılan işlem eş zamanlı çökeltmedir, bu işlemde çökeltici havalandırma tanklarına dozlanır. Alternatifler arasında ön çökeltme ve çökelticinin ön işlem amacıyla giriş akışına dozlandığı ön çökeltme ve son çökeltmeden sonra ek bir

reaksiyon tankının gerektiği çöktürme sonrası işlemleri bulunmaktadır (Metcalf ve Eddy, 2003). Günümüzde biyolojik fosfat giderimi de kullanılmaktadır, bu nedenle fosfat eliminasyonu genellikle kimyasal ve biyolojik çöktürme süreçlerinin kombinasyonu ile gerçekleştirilmektedir (Mackenzie, 2010).

Geliştirilmiş Biyolojik Fosfor Giderimi (EBPR) aktif çamur sistemi anaerobik/aerobik bir dizi ile işletildiğinde, polifosfat biriktiren organizmaların (PAOs) üstünlüğü sonucunda ortaya çıkar. Anaerobik ortamda PAO'lar, bu dönemde fermante edilebilir substrat kullanan normal heterotrofların çoğunlukla asetat gibi uçucu yağ asitlerini (VFA) depolayabilir. VFA'ların polihidroksi-alkanoatlar (PHA) olarak depolanması, iç polifosfat (poli-P) havuzunu enerji kaynağı olarak kullanarak ortofosfatın salınmasına yol açar. Sonraki aerobik dönemde PAO'lar depolanmış PHA üzerinde büyür ve poli-P rezervlerinin yenilenmesi için ortofosfat alır, bu da çamurun fosfor içeriğini artırır. Normal heterotrofların bir elektron alıcısının eksikliği nedeniyle asetat gibi substratları alamaması, anaerobik koşulların aktivasyon çamurunda PAO'ların zenginleşmesini destekler (Mackenzie, 2010; Wiesmann vd., 2007).

2.1.4 Atıksu arıtma tesisi

Arıtma tesisleri, atıksu arıtma sistemlerinin temel bileşenleridir. Bu tesislerde gerçek arıtma süreci aşamalar halinde gerçekleştirilmektedir. Bu tesisler mekanik, biyolojik ve kimyasal arıtma süreçlerini kullanmaktadır. Çeşitli mekanik arıtma adımlarının ardından karbon, azot ve fosforun biyolojik arıtımı ana reaktörlerde gerçekleştirilir. Bu süreçleri aktif çamurun çöktürme yoluyla atıksudan mekanik olarak ayrıldığı son işlem takip eder. Şekil 2'de atıksu biyolojik arıtma tesisine ait genel süreç akışı verilmiştir (Newhart, 2020).



Şekil 2. Atıksu arıtma tesisi genel akışı.

2.1.4.1 Giriş ızgaraları

Giriş ızgaraları, bir yandan mekanik hasarı önlemek, diğer yandan katı madde içeriğini azaltmak için arıtma tesisinden partikül maddeleri dışarıda tutma görevini üstlenmektedir. Iızgaralar kaba ve ince ızgaralar olarak ikiye ayrılır. Kaba ızgaraların açıklıkları 10 mm ile 50 mm arasındadır, ince ızgaralardaki açıklıklar ise 2 mm ile 10 mm arasındadır. Günümüzde genellikle otomatik ızgaralar, örneğin adım ızgaraları kullanılmaktadır. Bu sistemler, ızgaraların ne kadar tıkanmış olduğunu tespit eder, bunları otomatik olarak temizler ve biriken atığı doğrudan bir konteynere aktarır.

Iızgaraların ne kadar tıkanmış olduğunu belirlemede geleneksel yöntem, ızgaraların önündeki su seviyesini ölçmektir. Su seviyesi belirli bir seviyeyi aştığında, otomatik temizleme süreci başlatılır. İyileştirilmiş sistemler, ızgaranın önündeki ve arkasındaki su seviyesini ölçer ve aradaki farkı belirleyici kriter olarak kullanır. Yeni bir yaklaşım ise, ızgaranın ne kadar tıkanmış olduğunu değerlendirmek için görüntü tanıma yöntemlerinin kullanılmasıdır (Mackenzie, 2010; Metcalf ve Eddy, 2003).

2.1.4.2 Ön arıtım/ön çöktürme

Birincil arıtma tankı veya ön çöktürme tankı, mekanik arıtma sürecinin bir parçasıdır. Amaç, önceki aşamalarda giderilemeyen askıda katı maddelerden atıksuyu arındırmaktır. Atıksuyun tank içinde yavaşça akmasını sağlayacak şekilde inşa edilmiştir. Dikdörtgen tanklarda, genellikle tankın altında ve yüzeyinde biriken veya yüzen malzemeyi temizlemek için kazıyıcılar kullanılır. Eski tesislerin amacı mümkün olduğunca fazla maddenin çökmesini sağlamak olsa da, günümüzde çöktürme tankı süresi, denitrifikasyon süreci için gerekli miktarda karbonun korunması amacıyla daha kısa tutulmaktadır (Mackenzie, 2010; Metcalf ve Eddy, 2003).

2.1.4.3 İkincil arıtma

İkincil arıtma, biyokimyasal arıtma aşamasıdır. Aktifleştirilmiş çamur tesislerinde tipik olarak kullanılan çeşitli türlerde reaktör düzeni veya yapılandırması bulunmaktadır. Her reaktör düzeninin belirli özellikleri, avantajları ve dezavantajları vardır. Uygulamada, tesislerin karmaşıklığı genellikle artar çünkü özellikle orta ve büyük ölçekli tesislerin birçok kanalı bulunmaktadır.

İkincil arıtma, birincil arıtma sonrasında atıksuda kalan çözünmüş ve askıda kalmış organik bileşiklerin parçalanmasını içerir. Yaygın olarak kullanılan yöntem, mikroorganizmaların kalan organik maddeyi tüketmesi olan aktif çamur prosesidir, bu da biyokimyasal oksijen ihtiyacı (BOİ) ve kimyasal oksijen ihtiyacı (KOİ) seviyelerini önemli ölçüde azaltır. Bu süreç, çevreye geri salınan suyun zararlı organik maddelerden nispeten arınmış olmasını sağlamakta olup, sucul yaşamı ve genel ekosistem sağlığını korumaktadır (Mackenzie, 2010; Metcalf ve Eddy, 2003).

2.1.4.4 Üçüncül arıtma

Üçüncül arıtma, özellikle besin maddeleri ve patojenler olmak üzere kalan kirlenmeleri hedefleyen aşama olarak hizmet vermektedir. Bu aşamada, azot ve fosforun azaltılması da dahil olmak üzere besin maddesi giderimi, üçüncül arıtmada kritik bir rol oynar. Deşarj edilen suda yüksek konsantrasyonlarda azot ve fosfor bulunması, ötrofikasyona neden olarak sucul ekosistemleri bozabilir. Biyolojik nitrifikasyon ve denitrifikasyon gibi

süreçlerin yanı sıra kimyasal çökelme, bu besi maddelerinin etkili bir şekilde giderilmesi için kullanılmaktadır.

Ayrıca, filtrasyon, dezenfeksiyon ve besin maddesi giderimi gibi çeşitli yöntemler istenen su kalite standartlarına ulaşmak için kullanılmaktadır. Kum ve membran filtrasyonu dahil olmak üzere filtrasyon süreçleri, kalan askıda katı maddeleri etkili bir şekilde giderirken, klorlama veya ultraviyole (UV) ışınları gibi dezenfeksiyon yöntemleri kalan patojenleri etkisiz hale getirerek suyun çevresel deşarj veya potansiyel yeniden kullanım için güvenli olmasını sağlar (Mackenzie, 2010; Metcalf ve Eddy, 2003).

2.1.4.5 Çamur arıtımı

Çamur işlemi, atıksu arıtma alanındaki başka bir önemli unsurdur. Atıksu arıtma sırasında çamur farklı işlem adımlarında oluşturulur. Birincil çamur, birincil arıtmada atıksudan mekanik olarak ayrılır ve yağ içeriği nedeniyle yüksek KOİ konsantrasyonu ile karakterize edilir. İkinci kaynak ise aktif çamur prosesinden çıkan fazla çamurdur. Biyokütle sürekli beslendiği için sürekli olarak artar. Bu nedenle, atıksu arıtma tesisi için çamur bir atık bertarafı sorunu haline gelmektedir. Stabilize edilmesi gerekmektedir, yani organik içerik azaltılmalı ve biyolojik aktivite minimuma indirilerek çamurun depolanabilir hale gelmesini sağlamak için yapılması gereken bir işlemdir.

Çamur stabilizasyonu için farklı yöntemler kullanılmaktadır:

Aerobik çamur stabilizasyonu

Bu yöntem genellikle işletme maliyetleri nedeniyle çürütücüye sahip olmayan küçük ölçekli atıksu arıtma tesislerinde kullanılır. Temel olarak, çamur mikroorganizmalarının sürekli açlık durumunda olmaları için havalandırılır, böylece açlık durumundan kaçınmak için tüm mevcut kaynakları kullanmaya zorlanırlar. Bu yöntemin başlıca dezavantajı havalandırma için çok fazla enerjiye ihtiyaç duyulmasıdır. Yanı sıra yalnızca hafif yüklü olan tesislerde nitrifikasyonla aynı tankta mümkündür.

Anaerobik çamur sindirimi

Çamurun Anaerobik Sindirimi (AD) (1) Karbon ve nitrat giderimi ve (2) enerji üretimi olmak üzere iki temel amaca hizmet etmektedir. Bu hedeflere ulaşmak için çamur,

biyogaz üretmek için anaerobik bir ortamda çürütücülerde işlenir; bu biyogaz çoğunlukla metan (CH₄), karbondioksit (CO₂), hidrojen (H₂) ve hidrojen sülfür (H₂S) içermektedir. Biyogaz daha sonra kombine ısı ve güç ünitelerinde yanarak elektrik ve termal enerji üretmek üzere kullanılır veya doğal gaz kalitesine yükseltilir ve ardından gaz şebekesine sağlanır. Arıtma tesisi sindiricilerinde AD süreci genellikle 35 ila 42°C arasında sabit sıcaklıkta ve %5 ila 10 toplam katı konsantrasyonunda mesofilik aralıkta çalıştırılır ve işletilir. Her ne kadar işlem verimliliğini artırmak için çoklu çürütücüye sahip AD tasarımları olsa da, daha düşük maliyetlerden dolayı tek reaktör tasarımı atıksu arıtma tesislerinde en yaygın olanıdır (Mackenzie, 2010).

Genel olarak, çamur işleminin atıksu arıtımının ana işletme maliyetlerinden biri olduğu söylenebilir. Bu sebeple tesis optimizasyonu veya kontrol sistemi geliştirme sırasında çamur işlemi göz önünde bulundurulmalıdır.

2.2 Makine Öğrenmesi

Makine öğrenmesi ve yapay zeka, 20. yüzyılın ortalarına kadar uzanan bir tarihe sahiptir. "Yapay zeka" terimi 1956 yılında kullanılmaya başlanmış olup, o zamandan beri araştırmacılar insan zekasını taklit eden bilgisayar programları geliştirmek için çalışmaktadırlar. Herhangi bir problemi çözmede yapay zekâ tekniklerinin özü, veriler arasındaki işlevsel ilişkilerin yakalanması için kendilerine sunulan veri girdileri ve çıktıları örnekleriyle öğrenmektir. Bu nedenle, yapay zeka modelleri, bir sistemi yöneten yapı ve parametreleri belirlemek için veriye dayalı modellerdir. Makine öğrenmesi, yapay zekanın bir alt kümesi olarak, bilgisayarların açıkça programlanmadan verilerden öğrenmelerini ve tahminler yapmalarını sağlayan algoritmaların geliştirilmesine odaklanır. Teknolojinin ilerlemesi ve büyük veriye erişim, makine öğrenimi ve yapay zekanın sağlık, finans ve ulaşım gibi alanlarda giderek daha önemli hale gelmesini sağlamıştır.

Makine öğrenmesi, açıkça programlanmadan verilere dayalı kalıplar öğrenmeyi ve kararlar vermeyi içeren yapay zeka alanıdır (Bishop, 2006; Mohri vd., 2018). Bu, arama motorları (Brin ve Page, 1998), konuşma tanıma (Hinton vd., 2012) ve otonom araçlar gibi pek çok pratik uygulamanın arkasındaki önemli bir teknolojidir.

Makine öğrenmesinin amacı, yeni verilere genelleştirme yeteneğine sahip modeller oluşturmaktır, yani önceden görülmemiş verilere yönelik doğru tahminler veya kararlar verebilme yeteneğine sahip olmaktır (Vapnik, 1995). Bu amaçla, makine öğrenmesi algoritmaları, girdi verileri ile eşleşen etiketler ve çıktı verilerinden oluşan bir veri kümesi üzerinde eğitilir. Algoritma, eğitim verilerindeki kalıplara dayalı olarak girdi verilerini çıktı verilerine eşleyen bir işlev öğrenir (Murphy, 2012).

Makine öğrenmesinin birçok farklı türü vardır, bunlar arasında yönlendirilmiş öğrenme (supervised), yönlendirilmemiş öğrenme (unsupervised) ve pekiştirme öğrenimi (reinforced) bulunur (Sutton ve Barto, 2018).

2.2.1 Yönlendirilmiş öğrenme

Yönlendirilmiş öğrenme için, algoritma, etiketli eğitim verileriyle birlikte verilir, yani her girdi için doğru çıktı verilir (Hastie vd., 2009). Algoritma, bu eğitim verilerini kullanarak girdileri çıktılara eşleyen bir işlev öğrenir. Yönlendirilmiş öğrenme, sınıflandırma ve regresyon gibi görevler için kullanılır, bu görevlerde amaç, sırasıyla belirli bir sınıfı veya sürekli bir değeri tahmin etmektir (Murphy, 2012). Örneğin, tıbbi teşhislerde, e-posta spam filtrelerinde veya finansal tahminlerde, algoritmanın doğru bir şekilde sınıflandırma yapabilmesi için önceden etiketlenmiş veri setleri kullanılır. Yönlendirilmiş öğrenme, bilgisayar sistemlerine belirli bir girdiye dayanarak doğru çıktıları üretme yeteneği kazandırır. Sınıflandırma görevlerinde bu, bir girdinin belirli bir kategoriye ait olup olmadığını belirleme anlamına gelmektedir. Örneğin, bir e-postanın spam olup olmadığını belirleme veya bir görüntünün içeriğini tanımlama gibi uygulamalarda bu tür öğrenme yaklaşımı büyük önem taşımaktadır. Yönlendirilmiş öğrenme algoritmalarının başlıca amacı, eğitim verilerine dayalı olarak genelleme yeteneğine sahip modeller oluşturmaktır. Bu modeller, önceden görülmemiş verilere karşı doğru tahminlerde bulunma yeteneğiyle öne çıkar. Bu, algoritmaların gerçek dünya uygulamalarında güvenilir ve etkili sonuçlar üretebilmesini sağlar.

2.2.2 Yönlendirilmemiş öğrenme

Yönlendirilmemiş öğrenme için, algoritma etiketli eğitim verileriyle birlikte verilmez (Frey ve Dueck, 2007). Bunun yerine, veri üzerinde kendi başına kalıplar ve yapılar

bulmalıdır. Yönlendirilmemiş öğrenme, veri noktalarını kümelere gruplama veya verinin daha düşük boyutsal gösterimlerini bulma gibi görevler için kullanılır (Bishop, 2006). Yönlendirilmemiş öğrenme, veri setindeki benzerlikleri, gruplamaları veya desenleri belirleyerek verinin doğal yapısını ortaya çıkarmayı amaçlar. Kümeleme (clustering) ve boyut indirgeme (dimensionality reduction) gibi tekniklerle sıklıkla uygulanır. Örneğin, bir pazarlama uzmanı, müşteri segmentasyonunu anlamak için yönlendirilmemiş öğrenme algoritmalarını kullanabilir. Algoritma, müşteriler arasındaki benzerlikleri belirleyebilir ve bu bilgiyi kullanarak özelleştirilmiş pazarlama stratejileri oluşturabilir. Bu öğrenme türü, özellikle büyük ve karmaşık veri setlerinde gizli desenleri keşfetme kapasitesi nedeniyle önemlidir. Örneğin, genetik veri analizi, nörolojik hastalıkların teşhisi veya doğal dil işleme uygulamalarında, algoritmaların verinin içsel yapısını anlamaları ve öğrenmeleri için kullanılır. Yönlendirilmemiş öğrenme, makine öğrenmesinin geniş bir yelpazesine uygulanabilen çok yönlü bir yaklaşımdır. Bu, veri setinin doğasına bağlı olarak farklı yöntemlerin ve algoritmaların kullanılabilmesine olanak tanır. Bu özellikleri nedeniyle, bu öğrenme türü, veri bilimi ve yapay zeka alanlarında geniş bir uygulama yelpazesi bulmuştur.

2.2.3 Pekiştirme öğrenimi

Pekiştirme öğrenimi, bir çevrede kararlar vermeyi öğrenen algoritma eğitimini içerir (Sutton ve Barto, 2018). Algoritma, deneme ve yanılma yoluyla öğrenir ve ödülleri aldığında eylemlerini ayarlar. Pekiştirme öğrenimi, oyunlar oynama (Mnih vd., 2015) ve robotları kontrol etme (Kober vd., 2013) gibi görevler için kullanılır. Model, belirli bir aksiyonu gerçekleştirdiğinde ödüllendirilir veya cezalandırılır. Bu ödüllendirme veya cezalandırma mekanizması, modelin gelecekteki kararlarını optimize etmeye yönlendirir. Pekiştirme öğrenme, ajanın çevresiyle etkileşime girerek deneme yanılma yoluyla en iyi stratejiyi bulmasına odaklanmaktadır. Pekiştirme öğrenimi, özellikle oyun teorisi, robotik kontrol, finansal uygulamalar ve yapay zeka tabanlı karar destek sistemleri gibi alanlarda yaygın olarak kullanılır. Algoritmanın hedefi, çeşitli durumlarla başa çıkma ve ödüllendirme sinyallerini doğru bir şekilde değerlendirme yeteneğini geliştirmektir. Bu öğrenme türü, ajanın öğrenme sürecindeki deneyimlerinden çıkarımlar yaparak belirli bir görevde optimal performans elde etmesine olanak tanır.

2.2.4 Makine öğrenmesi teorisi

Makine öğrenmesi teorisi, makine öğrenmesi algoritmalarının matematiksel temellerini ve ilkelerini inceleyen bir alandır (Vapnik, 1995). Bu, algoritmaların nasıl ve neden çalıştığını anlamaya ve yeni algoritmalar ve teknikler geliştirmeye yöneliktir (Mohri vd., 2018). Makine öğrenimi teorisi için önemli konular arasında sapma-varyans ödünleşimi (bias-variance tradeoff), aşırı uyum (overfitting), genelleştirme ve optimizasyon bulunur (Hastie vd., 2009).

2.2.4.1 Sapma-varyans ödünleşimi

Sapma-varyans ödünleşimi, modelin varsayımları tarafından sağlanan hata (sapma) ve verinin değişkenliği tarafından sağlanan hata (varyans) arasındaki dengeyi ifade eder (Geman vd., 1992). Sapma, bir modelin öğrenme sürecinde yaptığı basitleştirmelerin veya yanlılıkların bir ölçüsüdür, yani modelin gerçek dünyayı temsil etme yeteneğini etkiler. Varyans ise modelin farklı eğitim veri setlerine tepkisindeki dalgalanmayı temsil eder, yani modelin esnekliğini ve genelleme yeteneğini etkiler. Yüksek sapmaya sahip bir model, veriyi aşırı basitleştirebilir ve yeni verilerde iyi performans gösteremeyebilirken yüksek varyansa sahip bir model, eğitim verilerine aşırı uyum gösterebilir ve yeni verilerde iyi performans gösteremeyebilir (Hastie vd., 2009). Yüksek sapma, modelin aşırı basitleştirildiği ve gerçek dünyayı karmaşıklıkla temsil edemediği anlamına gelir. Yüksek varyans ise modelin eğitim verilerine çok fazla uyum sağladığı, ancak yeni veri setlerine kötü genelleme yaptığı durumu ifade eder. Sapma-varyans ödünleşimi, bu iki faktör arasında bir denge bulunması gerektiğini vurgular. Bir modelde ideal sapma-varyans ödünleşimi, hem eğitim verilerinde hem de yeni veri setlerinde iyi performans gösterme yeteneğidir. Bu dengeyi bulmak için, model karmaşıklığını ayarlamak ve uygun düzenleme tekniklerini kullanmak önemlidir. Aşırı uyarlamayı (overfitting) önlemek ve genel performansı artırmak için çapraz doğrulama gibi teknikler de sapma-varyans ödünleşimi üzerinde olumlu etkiler sağlar. Bu dengeyi kurmak, bir makine öğrenmesi modelinin gerçek dünyada güvenilir ve genel sonuçlar üretebilmesi için kritiktir (Bishop, 2006).

2.2.4.2 Aşırı uyum

Aşırı uyum (overfitting), makine öğrenimi modellerinin eğitim verilerine aşırı derecede uyması ve bu nedenle yeni ve görülmemiş veri setlerinde düşük performans göstermesi durumunu ifade eder. Bu durum, modelin eğitim verilerindeki gürültüyü veya tesadüfi örüntüleri öğrenmesi sonucu ortaya çıkar ve modelin genelleme yeteneğini zayıflatır (Hastie vd., 2009). Aşırı uyum, modelin eğitim verilerine aşırı adapte olduğunda meydana gelir, bu da modelin eğitim verilerindeki her detayı hatırlamasına ve bu detaylara aşırı hassasiyet göstermesine yol açar. Bu durum modelin yeni veri setlerinde önceki başarısını sürdürmeyeceği anlamına gelmektedir. Pratikte, aşırı uyumun önlenmesi ve kontrol edilmesi önemlidir çünkü makine öğrenimi modellerinin gerçek dünyadaki genelleme yeteneği, görülmemiş durumlarla başa çıkma kabiliyetiyle ölçülür. Aşırı uyum, modelin karmaşıklığının dengelemesi, düzenleme tekniklerinin uygulanması ve uygun ölçüm metriklerinin kullanılması gibi stratejilerle önlenabilir. Ayrıca, modelin eğitim sürecinde çapraz doğrulama (cross-validation) gibi tekniklerin kullanılması, modelin gerçek dünyadaki performansını daha iyi yansıtan bir değerlendirme sağlar. Aşırı uyumun kontrol altına alınması, makine öğrenimi modellerinin güvenilir ve genel geçer sonuçlar üretmelerini sağlayarak uygulamalarının gerçek dünyadaki etkinliğini artırır (Kohavi, 1995).

2.2.4.3 Genelleştirme

Genelleştirme, makine öğrenimi modellerinin eğitim verileri üzerinde elde ettiği başarıyı, yeni ve görülmemiş veri setlerine uygulama yeteneğini ifade eder (Hastie vd., 2009). Bir modelin genelleme yeteneği, eğitim sürecinde öğrendiklerini sadece eğitim verileriyle sınırlı kalmadan, genel bir probleme uygulayabilme becerisine dayanır. Bu, modelin öğrendiği örüntüleri, kuralları ve ilişkileri yeni ve çeşitli veri setlerine de doğru bir şekilde uygulayabilmesini içerir. İdeal bir genelleme, modelin hem eğitim verilerinde hem de yeni veri setlerinde başarılı bir performans sergilemesini içerir. Ancak genelleme yeteneği, aşırı uyum (overfitting) ve yetersiz uyarlamayı (underfitting) dengeleme sürecine dayanır. Aşırı uyum, modelin eğitim verilerine fazlasıyla uyması ve yeni veri setlerinde başarısız olması durumunu ifade ederken, yetersiz uyarlama ise modelin eğitim verilerini yeterince öğrenememesi ve genelleme yeteneğinin düşük olması durumunu ifade eder. Genelleme yeteneği, modelin karmaşıklığının ayarlanması, doğru özelliklerin

seçilmesi, uygun düzenleme tekniklerinin uygulanması ve çapraz doğrulama gibi değerlendirme stratejilerinin kullanılması ile geliştirilebilir. Genelleme yeteneği arttıkça, makine öğrenimi modelleri daha çeşitli uygulamalarda başarılı bir şekilde kullanılabilir ve gerçek dünyadaki değişkenliklere daha iyi uyum sağlayabilir.

2.2.4.4 Optimizasyon

Optimizasyon, makine öğrenmesi modelinin parametrelerini bir kayıp fonksiyonunu minimize etmek için ayarlamak sürecidir (Murphy, 2012). Yapay zekâ teknolojileri, iş süreçlerinde verimlilik artışı ve iyileştirmelere fırsat sunmanın yanında yapay zekâ tabanlı analitik araçlarla iş süreçleri optimizasyonu için strateji geliştirmeye de katkı sunar(Şahinaslan, E. 2023). Bu süreç, modelin eğitim verilerinde ve yeni veri setlerinde daha iyi bir performans sergilemesini sağlamak için parametrelerin en iyi değerlerini bulmayı amaçlamaktadır. Makine öğrenimi modellerinin başarısı, modelin parametrelerinin bir kayıp fonksiyonunu minimize ederek belirlenir. Kayıp fonksiyonları, tahmin edilen çıktı ile gerçek çıktı arasındaki farkı ölçer ve optimizasyonun amacı, en düşük kaybı sağlayan parametre kümesini bulmaktır. Optimizasyon süreci, bu kayıp fonksiyonunu en aza indirerek modelin doğruluğunu artırmayı hedefler. Farklı optimizasyon algoritmaları vardır, bunlar arasında “gradient descent”, “stochastic gradient descent” ve “Adam” bulunur (Bottou vd., 2015). Optimizasyon süreci, modelin hiperparametrelerini belirleyerek, aşırı uyum (overfitting) veya yetersiz uyarılama (underfitting) gibi sorunları önleyerek ve genel performansı iyileştirerek makine öğrenimi modellerinin daha etkili ve güvenilir olmasını sağlar. Bu, modelin gerçek dünyada geniş bir uygulama yelpazesinde başarılı bir şekilde çalışabilmesini sağlar.

2.2.5 Makine öğrenmesi algoritmaları

Makine öğrenmesi, veri biliminin önemli bir parçasıdır ve veri toplama, temizleme, ön işleme ve özelleştirme gibi aşamaları içerir. Bu aşamalar, verinin makine öğrenimi algoritmaları tarafından kullanılabilir hale getirilmesi için yapılır.

Makine öğrenimi, derin öğrenme algoritmaları kullanarak insan zekâsını taklit eden bir teknolojidir(Şahinaslan vd, 2023). Makine öğrenmesi algoritmaları, makine öğrenmesi sistemlerinin öğrenmeyi ve tahminler yapmayı veya kararlar vermeyi sağlayan yapı

taşlarıdır (Hastie vd., 2009). Birçok farklı türde makine öğrenmesi algoritmaları vardır, bunlar arasında doğrusal regresyon, lojistik regresyon, karar ağaçları ve destek vektör makineleri (Hastie vd., 2009) bulunur. Bir görev için doğru algoritmayı seçmek, verinin özelliklerine ve istenen sonuca bağlıdır (Murphy, 2012).

2.2.5.1 Doğrusal regresyon

Doğrusal regresyon, sürekli değerleri tahmin etmek için kullanılan bir yönlendirilmiş (supervised) öğrenme algoritmasıdır (Murphy, 2012). Temel amacı, bağımlı bir değişkenin (çıktı) bağımsız değişkenlerle olan ilişkisini tanımlayan bir doğrusal denklem oluşturarak bu ilişkiyi modellemektir. Verilere bir çizgi sığdırarak çalışır, böylece tahmin edilen değerler ile gerçek değerler arasındaki karelerin hata toplamı en aza indirilir (Bishop, 2006). Bu regresyon tekniği, sürekli değerlere sahip çıktı değişkenleri üzerinde çalışır. Model, bağımsız değişkenlerin (girdilerin) doğrusal bir kombinasyonunu kullanarak çıktı değişkenini tahmin eder. Doğrusal regresyonun temel varsayımı, bağımsız değişkenler ile çıktı arasındaki ilişkinin doğrusal olduğudur. Doğrusal regresyon, istatistikte yaygın olarak kullanılan en basit regresyon tekniğidir ve genellikle regresyon analizlerine giriş olarak öğretilir. Modelin performansını değerlendirmek için kullanılan metrikler arasında R-kare (R-squared), ortalama kare hatası (Mean Squared Error - MSE) ve katsayılarının anlamlılığı gibi ölçütler bulunur. Bu regresyon tekniği, ekonomik tahminlerden finansal analize, pazarlama stratejilerinden tıbbi araştırmalara kadar birçok alanda uygulanabilir. Doğrusal regresyon, veri setindeki ilişkileri analiz etmek ve gelecekteki olayları tahmin etmek için güçlü bir araç olabilmektedir.

2.2.5.2 Lojistik regresyon

Lojistik regresyon, sınıflandırma görevleri için kullanılan başka bir yönlendirilmiş (supervised) öğrenme algoritmasıdır (Hastie vd., 2009). Bu yöntem, doğrusal regresyonun bir genişlemesi olarak kabul edilir ve sınıflandırma problemlerinde kullanılır. Temel amacı, bağımlı bir kategorik değişkenin (genellikle iki sınıfa ait) bağımsız değişkenlerle olan ilişkisini modellemektir. Verilere bir lojistik eğri sığdırarak çalışır, böylece farklı sınıfların tahmin edilen olasılıkları en yüksek seviyeye çıkarılır (Murphy, 2012). Lojistik regresyon, logit fonksiyonunu kullanarak olasılık değerlerini tahmin eder ve bu olasılıkları belirli bir eşik değeri üzerinden sınıflara dönüştürür. Bu

sayede, bir girdi seti verildiğinde, ilgili sınıfa ait olma olasılığı hesaplanabilir. Modelin eğitiminde genellikle maksimum olabilirlik tahmini kullanılır. Lojistik regresyon, tıp alanında hastalık teşhisinden pazarlama stratejilerine kadar birçok uygulama alanında kullanılır. Örneğin, bir tıbbi araştırmada, belirli semptomların varlığına göre bir hastalığın varlığını tahmin etmek için kullanılabilir. Pazarlamada ise, müşterinin belirli bir ürünü satın alma olasılığını tahmin etmek amacıyla kullanılabilir. Bu regresyon tekniğinde, modelin performansını değerlendirmek için kullanılan metrikler arasında doğruluk oranı (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 puanı gibi ölçütleri bulunmaktadır. Lojistik regresyon, sınıflandırma problemlerinde etkili bir çözüm sunar ve birçok endüstriyel uygulamada başarıyla kullanılmaktadır.

2.2.5.3 Karar ağaçları

Karar ağaçları, sınıflandırma ve regresyon görevleri için kullanılan bir yönlendirilmiş (supervised) öğrenme algoritması türüdür (Breiman vd., 1984). Özellik değerlerine göre veriyi daha küçük alt kümelerine sınıflandırarak çalışır, böylece alt kümelerin saflığı en yüksek seviyeye çıkarılır (Hastie vd., 2009). Karar ağaçları, veri setindeki karmaşık ilişkileri anlama ve açıklama konusunda etkili bir araçtır. Her düğüm, bir özellik üzerinde bir testi temsil eder ve bu test sonucuna göre ağaç dallara ayrılır. Bu dallar, sonuçları temsil eden yapraklara kadar devam eder. Model, veri setindeki desenleri öğrenerek ve genelleme yeteneği kazanarak yeni veri noktalarını sınıflandırabilir veya tahmin edebilir. Karar ağaçları, özellikle sınıflandırma problemlerinde kategorik ve sayısal verilerle çalışmak için uygundur. Ayrıca, ağacın genişletilmiş versiyonları, ensefalopati teşhisi koymaktan finansal risk tahminine kadar çeşitli uygulamalarda kullanılır. Bir avantajı da modelin sonuçları anlaşılır bir şekilde görselleştirebilmesidir. Bu yöntem, hiperparametre ayarlaması ve veri seti büyüklüğüne duyarlılık açısından bir dengesizlik oluşturabilir. Ancak, karar ağaçları, diğer modellerle birleştirilerek veya topluluk öğrenme teknikleriyle birleştirilerek bu zorlukları aşabilir. Karar ağaçları, veri madenciliği ve karar destek sistemleri gibi birçok alanda kullanılan güçlü ve esnek bir makine öğrenimi aracıdır.

2.2.5.4 Destek vektör makineleri

Destek vektör makineleri (SVM), sınıflandırma ve regresyon görevleri için kullanılan bir yönlendirilmiş (supervised) öğrenme algoritması türüdür (Cortes ve Vapnik, 1995). Verideki farklı sınıfları en iyi şekilde ayıran bir hiperdüzlem bulunarak çalışır (Hastie vd., 2009). SVM, özellikle sınıflar arasındaki ayrım zor olduğunda veya veri seti gürültülü olduğunda etkilidir. Çalışma prensibi, farklı sınıflara ait örnekleri en iyi şekilde ayıran bir hiperdüzlemi bulmaktır. Bu hiperdüzlem, veri noktalarının sınıflara göre en uzak mesafede olmasını sağlar, bu mesafe "margin" olarak adlandırılır. SVM'nin avantajlarından biri, yüksek boyutlu veri setlerinde ve karmaşık özellik uzaylarında etkili olabilmesidir. Hem sınıflandırma hem de regresyon görevlerinde kullanılabilir. Ayrıca, çekirdek fonksiyonları aracılığıyla non-lineer ilişkileri modelleyebilir, bu da özellikle veri seti lineer olmayan bir şekilde ayrılabilirse avantaj sağlar. Ancak, SVM'nin bazı dezavantajları da vardır. Özellikle büyük veri setlerinde eğitim süresi uzun olabilir ve hiperparametre ayarlaması gerektirebilir. Ayrıca, çok sayıda özellik çalışırken performansı düşebilir. Destek vektör makineleri, biyomedikal mühendislikten finansal analize kadar birçok alanda başarıyla kullanılmaktadır. Güçlü sınıflandırma yetenekleri ve genelleme yeteneği, SVM'yi makine öğrenimi uygulamalarında önemli bir araç haline getirir.

2.3 Yapay Sinir Ağları

Yapay sinir ağları (YSA) insan beyninin yapısı ve işlevinden ilham alan bir makine öğrenmesi algoritmasıdır. Yapay sinir ağları, bilgisayar sistemlerine öğrenme ve karar verme yeteneği kazandırmak amacıyla tasarlanmış matematiksel modellerdir. YSA'lar, bilgi işleme ve iletimine yönelik "nöronlar" denen birbirine bağlı düğümlerden oluşur. Bu ağlar, girdi verisine dayalı olarak öğrenme ve uyum sağlama yeteneğine sahiptir ve kalıp tanıma, sınıflandırma ve tahmin gibi çeşitli görevler için kullanılabilir (LeCun vd., 2015; Shrestha ve Mahmood, 2019).

Yapay sinir ağlarının temel yapı taşı nöronlardır. Bu yapay nöronlar, giriş verilerini alır, bu verileri işler ve bir çıkış üretirler. Yapay sinir ağları, genellikle birbirine bağlı çok sayıda nöron içeren katmanlardan oluşur. Giriş katmanı, veriyi ağına tanıtan birinci

katmandır. Gizli katmanlar, giriş verisini işleyen ve çeşitli desenleri öğrenen katmanlardır. Çıkış katmanı ise ağı sonu ıktısını üretir.

Yapay sinir ağıları, öğrenme süreçlerini gerçekleştirebilmek için genellikle iki temel öğrenme türünü kullanır: denetimli öğrenme ve denetimsiz öğrenme. Denetimli öğrenme, ağı eğitildiğı bir süreçtir ve giriş verileri ile bu verilere karşılık gelen doğru ıkışlar kullanılarak gerçekleştirilir. Denetimsiz öğrenme ise doğru ıkışlar kullanılmadan, ağı giriş verileri arasında gizli desenleri ve yapıları keşfetmesine odaklanır.

YSA'ların önemli özelliklerinden biri, gürültülü veya eksik olabilen büyük veri miktarlarını işleme yeteneğidir (Goodfellow vd., 2016). Bu, özellikle girdi verisi karmaşık ve çeşitli olabilecek görüntü ve konuşma tanıma gibi görevler için önemli bir avantaj sağlar. YSA'ların farklı türleri vardır, bunlar arasında ileri besleme ağıları (FNN), konvolüsyonel sinir ağıları (CNN) ve yineleyici sinir ağıları (RNN) bulunur (Shrestha ve Mahmood, 2019).

2.3.1 İleri beslemeli sinir ağıları

İleri beslemeli sinir ağıları (FNN), en temel tür olan YSA'dır ve girdi katmanı, bir veya daha fazla gizli katman ve ıktı katmanından oluşur (Bishop, 2006). Girdi katmanı girdi verisini alır ve ıktı katmanı ıktıyı üretir. Gizli katmanlar, ağırlıklı bağlantılar aracılığıyla nöronlar arasında veriyi işler. Bu bağlamda bir katmandaki her nöronun bir sonraki katmandaki her nöronla bağı olmaktadır (LeCun vd., 2015). İleri Beslemeli Sinir Ağıları genellikle birbirine tamamen bağı (fully connected) katmanlardan oluşur. Bu, bir katmandaki her bir nöronun, bir önceki ve bir sonraki katmandaki tüm nöronlarla bağı olduğu anlamına gelir. Giriş katmanı, ağı dış veriyi tanıtır, gizli katmanlar, veriyi işleyip öğrenmeyi sağlar, ve ıkış katmanı, sonu ıktısını üretir. Bu nöronların bağlantılarındaki ağırlıklar, öğrenme süreci boyunca uyarlanır. İleri beslemeli sinir ağıları, genellikle sınıflandırma ve regresyon gibi görevlerde kullanılır. Örneğın, görüntü tanıma, dil işleme, ses analizi gibi birçok alanda başarıyla uygulanmışlardır. İleri Beslemeli Sinir Ağıları, basit ve düzenli bir yapıya sahip oldukları için eğitim süreci genellikle hızlıdır. Ancak, daha karmaşık ve büyük veri setleriyle çalışan uygulamalarda derin öğrenme modellerine kıyasla sınırlamalara sahip olabilirler. Bu nedenle, ileri beslemeli sinir ağıları, giriş

verilerini basit bir şekilde işlemek ve çıktı üretmek için kullanılan temel bir yapay sinir ağı mimarisidir.

2.3.2 Konvolüsyonel sinir ağları

Konvolüsyonel sinir ağları (CNN), özellikle resim gibi ızgara yapısı olan veriyi işlemek için tasarlanmış bir besleme yapılı sinir ağı türüdür (LeCun vd., 1998). Bu ağlar, veri içindeki özellikleri belirlemek ve öğrenmek için kullanılan özel bir mimariye sahiptir. Konvolüsyon katmanları kullanılır, bu katmanlar veriye bir filtre kümesi uygular ve özellikleri çıkarır ve havuzlama katmanları, verinin boyutunu azaltmak için veriyi indirgemektedir (Goodfellow vd., 2016). Konvolüsyonel sinir ağları (CNN), genellikle katmanlar halinde düzenlenir. İlk katman, giriş verisi üzerinde belirli özellikleri vurgulamak ve çıkartmak için konvolüsyon işlemlerini içerir. Bu sayede, özellikler hiyerarşik bir şekilde tanımlanabilir. Ardından, tipik olarak öğrenilmiş özellik haritalarını düzleştiren bir veya daha fazla tam bağlantılı katman gelir. Bu ağlar, özellikle görüntü sınıflandırma görevlerinde kullanılır, çünkü konvolüsyon işlemleri, özellikleri lokal olarak tanımlayarak nesne tanıma ve tespiti gibi görevlerde etkili olabilmektedir. Ayrıca, dil işleme ve video analizi gibi alanlarda da kullanılmaktadır. Konvolüsyonel sinir ağlarının (CNN) başarısının bir nedeni, özellikle büyük veri setleri üzerinde eğitildiklerinde, öğrenilen özelliklerin genelleme yetenekleridir. Bu ağlar, önceki katmanlardan elde edilen özellikleri bir sonraki katmanlarda kullanarak karmaşık ilişkileri öğrenebilir ve bu da genel performanslarını artırabilir. Sonuç olarak, konvolüsyonel sinir ağları, özellikle görsel veri üzerinde çalışan ve özellik çıkarma süreçlerinde etkili olan güçlü bir sinir ağı mimarisidir.

2.3.3 Geriye dönük sinir ağları

Geriye dönük sinir ağları (RNN), zaman serisi veya doğal dil gibi sıralı veriyi işlemek için tasarlanmış bir YSA türüdür (Goodfellow vd., 2016). Geri besleme bağlantıları vardır, bu bağlantılar nöronların bir zaman adımından diğerine bilgi aktarabilmesine olanak verir ve ağıın geçmişle ilgili bilgileri saklama özelliğine sahiptir (Bishop, 2006). Geriye dönük sinir ağlarının (RNN) temel özelliği, önceki zaman adımlarındaki bilgileri mevcut adıma dahil edebilmeleridir. Bu özellik, metin verilerinde veya zaman serilerindeki bağımlılıkları anlamak ve gelecekteki durumları tahmin etmek için büyük

önem taşır. Ancak, basit geriye dönük sinir ağı yapıları, uzun vadeli bağımlılıkları iyi öğrenememe ve gradient kaybı sorunlarıyla karşılaşabilme sorunlarına sahiptir. Bu nedenle, LSTM (Long Short-Term Memory) ve GRU (Gated Recurrent Unit) gibi gelişmiş RNN varyasyonları geliştirilmiştir. Bu yapılar, uzun vadeli bağımlılıkları daha etkili bir şekilde öğrenir ve gradient kaybı sorunlarına daha iyi bir çözüm sunar. LSTM, özellikle uzun vadeli bağımlılıkları daha etkili bir şekilde koruyarak, metin anlama, dil modelleme ve müzik oluşturma gibi uygulamalarda başarılı olmaktadır. Geriye dönük sinir ağıları, genellikle doğal dil işleme, konuşma tanıma, hisse senedi fiyat tahmini gibi zaman serileri analizinde, ya da video analizi, el yazısı tanıma gibi sıralı verilerin işlenmesi gereken görevlerde kullanılmaktadır. Bu ağılar, önceki adımlardan gelen bağlantılar sayesinde verideki zaman ve sıra içindeki desenleri daha iyi anlama yeteneği ile bilinmektedir.

2.3.4 Yapay sinir ağlarının eğitimi

YSA'ların eğitimi, ağı büyük bir veri kümesi sunarak ve nöronlar arasındaki bağlantıların güçlerini, ağı tahminlerinin doğruluğuna göre ayarlama işlemine dayanır (LeCun vd., 2015). Bu işlem, “backpropagation” olarak bilinir ve gradient inişi gibi bir optimizasyon algoritmasının ağırlıkları ayarlamak için kullanılmasını içerir (Goodfellow vd., 2016).

YSA'ların çok sayıda yeteneklerine rağmen, sınırlamaları da vardır. Örneğin, YSA'lar eğitim verisine iyi bir performans gösterebilir ancak yeni, görülmemiş veri üzerinde düşük bir performans gösterme eğilimine sahip olabilecekleri aşırı uyumlama (overfitting) problemi yaşayabilirler (Goodfellow vd., 2016). Ayrıca, YSA'lar hesaplama yoğun olabilir ve eğitim için önemli bir işlem gücü ve zaman gerektirir (LeCun vd., 2015).

Sonuç olarak, yapay sinir ağıları makine öğrenmesi için güçlü bir araçtır ve çeşitli uygulamalarda yaygın olarak kullanılır. Sınırlamaları olmasına rağmen, bu alandaki araştırmalar ve gelişmeler bu algoritmaların performansını ve yeteneklerini geliştirmeye yardımcı olmaktadır.

2.4 Atıksu Arıtımında Tarihsel Süreç Kontrolü

Veriye dayalı süreç kontrolü, tarihsel olarak atıksu arıtımında fazla tercih edilmemiştir (Metcalf ve Eddy, 2003; O'Day, 2004). 1920'lerden itibaren histogramlar ve kontrol

grafikleri gibi istatistiksel araçlar, resmi olmayan teşhisler için kullanılmıştır. 1960'lardan önce dijital kontrolün bir seçenek olmamasında ötürü bu dönemde kontrol, manuel ayarlamalara ve gözlemlere dayanmaktadır. Bilgisayar ve enstrümantasyon maliyeti yüksek; arıtma dinamikleri iyi anlaşılmamış olup; tesisler ek esneklikle tasarlanmamıştır ve kontrol teorisi yeterince gelişmemiştir (Olsson, 2012). 1980'lerde, uygun maliyetli bilgi işlem gücü, basit ilk prensip modellerine olanak tanınmıştır, ancak karmaşıklıkları ve güvenilirlik eksiklikleri nedeniyle bunlar zayıf danışmanlık sistemleri olmuştur (Olsson vd., 1998). 2000'li yılların başlangıcıyla birlikte dijital devrim atıksu arıtımına ulaşmış ve çoğu atıksu arıtma tesisi, programlanabilir mantık denetleyicileri (PLC) ve denetimli kontrol ve veri toplama (SCADA) sistemleri şeklinde kendi sürümünü sürekli izleme sürecine entegre etmiştir.

Atıksu arıtma tesisi verileri tarafından ortaya konulan benzersiz zorluklara rağmen, veriye dayalı sistem otomasyonu ve gerçek zamanlı kontrol, modern tesis işletmeciliğinin ayrılmaz bir parçasıdır. En yaygın işlem kontrol uygulaması, çevrimiçi sensör okumalarını ve geri besleme kontrolünü kullanarak arıtma sürecini korumaktır. Örneğin, çözünmüş oksijen konsantrasyonları, hava üfleyici hızını (yani havalandırma yoğunluğunu) ayarlayarak kontrol edilebilir. Tek bir üfleyici hızında çalışmak yerine, çevrimiçi ölçümler, çözünmüş oksijen konsantrasyonu gibi ölçülen bir değere göre üfleyici hızının artırılması, azaltılması veya sabit kalması gerekip gerekmediğini belirleyen kontrol sistemine geri bildirim sağlar. Kirleticinin çöktürme sürecini artırmak için kimyasal dozlama veya biyolojik süreçler için ek karbon gibi in-situ besin konsantrasyonlarına dayalı geri besleme kontrolü başka bir örnektir. Bu işlem kontrol yöntemi, sürekli, pürüzsüz ve otomatik bir şekilde işletme koşullarını değiştirmek için değişken frekans sürücüsüne sahip bir denetleyici ile genellikle başarılmaktadır.

İşlem değişkenleri için hedef değerler belirleme ("set-point") en basit kontrol yöntemlerinden biridir ve atıksu arıtma tesislerinde yaygın bir uygulamadır. Bu tek değişkenli izleme paradigması, çoğu modern arıtma tesisinde hata tespiti için temel oluşturmaktadır. Bu yaklaşımın düşük yanlış alarm oranı olmasına rağmen (örneğin, bir akış hızı ölçümü bir set noktanın altında ise, sistemde kesinlikle bilinmeyen bir nedenle hata oluşmuştur), hataları tespit etmek için çok yavaş olabilir, gelecekteki değerleri öngörmez ve değişkenler arasındaki ilişkileri hesaba katmaz. Tesis operatörleri, ekipman

hasarını veya sistem arızasını önlemek için hızlı bir şekilde bir sistem hatasına yanıt vermek için hazırda bulunmalıdır, bu da ekipmana ve personele ek stres ve işletme maliyeti getirmektedir. Hata tespiti ve öngörüye yönelik proaktif ve kapsamlı yaklaşımlar geliştirilmesi, atıksu arıtma tesisi sistemlerinin maliyetini azaltmaya ve verimliliğini artırmaya yardımcı olacaktır (Capizzi ve Masarotto, 2017; Jiang vd., 2012; Kazor vd., 2016; Odom vd., 2018; Wang ve Jiang, 2009).

2.5 Makine Öğrenmesine Dayalı Literatürde Yapılan Çalışmalar

Son yıllarda, makine öğrenmesi atıksu arıtma tesisi (AAT) araştırmalarında çeşitli su kalitesi parametrelerini tahmin ve kontrol etmek için etkili bir araç olarak kullanılmaktadır. Zhao (2009) tarafından yürütülen çalışmada, kümeleme ve karar ağacı gibi makine öğrenmesi teknikleri AAT'deki çamur kabarması olaylarını tespit etmek ve analiz etmek için kullanılmıştır. Sonuçlar, makine öğrenmesi algoritmalarının çamur kabarması olaylarını etkili bir şekilde tespit edebileceğini ve teşhis edebileceğini göstermiştir.

Matheri vd. (2021) yapay sinir ağları (YSA) kullanılarak AAT'lerdeki iz elementlerinin ve kimyasal oksijen ihtiyacının (KOİ) performansı tahmin edilmiştir. Bu çalışmada iz elementlerinin ve KOİ'nin giderim verimliliklerini tahmin etmek için çeşitli veri kümeleri kullanılarak YSA modellerini eğitilmiş ve test edilmiştir. Sonuçlar, YSA modellerinin AAT'lerde iz elementleri ve KOİ'nin performansını etkili bir şekilde tahmin edebileceğini göstermiştir.

Güçlü ve Dursun (2010) yapay sinir ağları (YSA) kullanarak çıkış kimyasal oksijen ihtiyacı (KOİ), askıda katı madde ve karışık sıvı askıda katı madde (MLVSS) konsantrasyonlarını tahmin etmeye çalışmışlardır. Atıksu modelleri kullanılarak sentetik veri üzerine yapılan bu çalışmada yüksek doğrulukta performans elde edilmiştir. Kashani ve Shahhosseini (2010) dinamik yapay zeka nöral ağları (DANNs) kullanarak kimyasal oksijen ihtiyacı (KOİ) ve uçucu askıda katı (VSS) konsantrasyonlarını tahmin etmeye çalışmışlardır.

Bagherzadeh vd. (2021), çeşitli özellik seçme yöntemleri ve makine öğrenmesi algoritmaları kullanarak AAT'lerdeki toplam azotun tahmin edilmesi üzerine

karşılaştırmalı bir çalışma yürütmüştür. Toplam azot tahmini için en ilgili özellikleri seçmek için temel bileşen analizi (TBA), korelasyon temelli özellik seçimi (KTS) ve karşılıklı bilgi özellik seçimi (KBÖS) gibi özellik seçme yöntemleri kullanılmıştır. Daha sonra, destek vektör makineleri (DVM'ler), rastgele orman (RO) ve K-en yakın komşu (KYK) gibi çeşitli makine öğrenmesi algoritmaları kullanılarak tahmin modelleri geliştirilmiştir. Çalışmanın sonuçları, özellik seçme yöntemlerinin toplam azotun tahmin doğruluğunu önemli ölçüde artırabileceğini göstermiştir.

Başka bir çalışmada Hvala ve Kocijan (2021) tarafından MÖ algoritmaları, çamur kabarma kontrolü için girdi değişkeni seçiminde kullanılmıştır. Morris yöntemi, Sobol yöntemi ve genişletilmiş Fourier Amplitüd Hassasiyet Testi (eFAST) dahil çeşitli global hassasiyet analizi (GSA) yöntemlerini kullanarak en etkili girdi değişkenlerini belirlemek için çalışmışlardır. Daha sonra, tanımlanan girdi değişkenlerini kullanarak SVM, KNN ve yapay sinir ağları (YSA) dahil çeşitli makine öğrenmesi algoritmalarını eğitmek ve test etmek için kullanılmıştır. Sonuçlar, tanımlanan girdi değişkenleri kullanılarak geliştirilen MÖ modellerinin tüm girdi değişkenlerini kullanan modellerden daha iyi performans gösterdiğini göstermiştir.

Asami vd. (2021), atıksu arıtma tesislerinde biyolojik ve kimyasal oksijen ihtiyacı ile toplam askıda katı madde simülasyonu yapmak için bir veri madenciliği yaklaşımı kullanmıştır. Çalışmada gerçek bir atıksu arıtma tesisi tarafından toplanan işletme verilerine dayanarak bu değişkenlerin davranışını tahmin etmek için makine öğrenimi tekniklerini kullanmışlardır. Sonuçlar, veri madenciliği yaklaşımının değişkenlerin dinamiklerinin doğru bir temsili sağladığını ve tesisin performansını optimize etmek için kullanışlı olabileceğini göstermiştir. Liu vd. (2020), kağıt endüstrisi atıksuyu arıtımında ELM metodları kullanarak çıkış COD parametresinin öngörü doğruluğunu incelemişlerdir.

Guo vd. (2015), makine öğrenmesi modellerini kullanarak bir atıksu arıtma tesisi'nde deşarj konsantrasyonunu tahmin etmişlerdir. Yapay sinir ağları, karar ağaçları ve destek vektör makineleri de dahil olmak üzere çeşitli makine öğrenmesi algoritmalarını kullanarak, gerçek bir atıksu arıtma tesisinden toplanan operasyonel verilere dayanarak deşarj konsantrasyonunu tahmin edebilen modeller geliştirilmiştir. Sonuçlar, makine öğrenmesi modellerinin deşarj konsantrasyonunu doğru bir şekilde tahmin ettiğini ve bu

durumun tesisin işletmesini optimize etmek ve çevresel etkisini azaltmak için faydalı olabileceğini göstermektedir.

Yang ve diğerleri (2021) atıksu arıtma süreçlerinde deşarj kalitesi için bir uyarlamalı dinamik tahmin modeli geliştirmişlerdir. Gerçek bir atıksu arıtma tesisinden toplanan giriş verilerine dayanarak etkili bir tahmin yapmak için kısmi en küçük kareler (PLS) ile ilişki vektör makinesini (İVM) kullanmışlardır. Çalışma, önerilen modelin, deşarj kalitesinin doğru ve sağlam bir tahminini sağladığını göstermiştir.

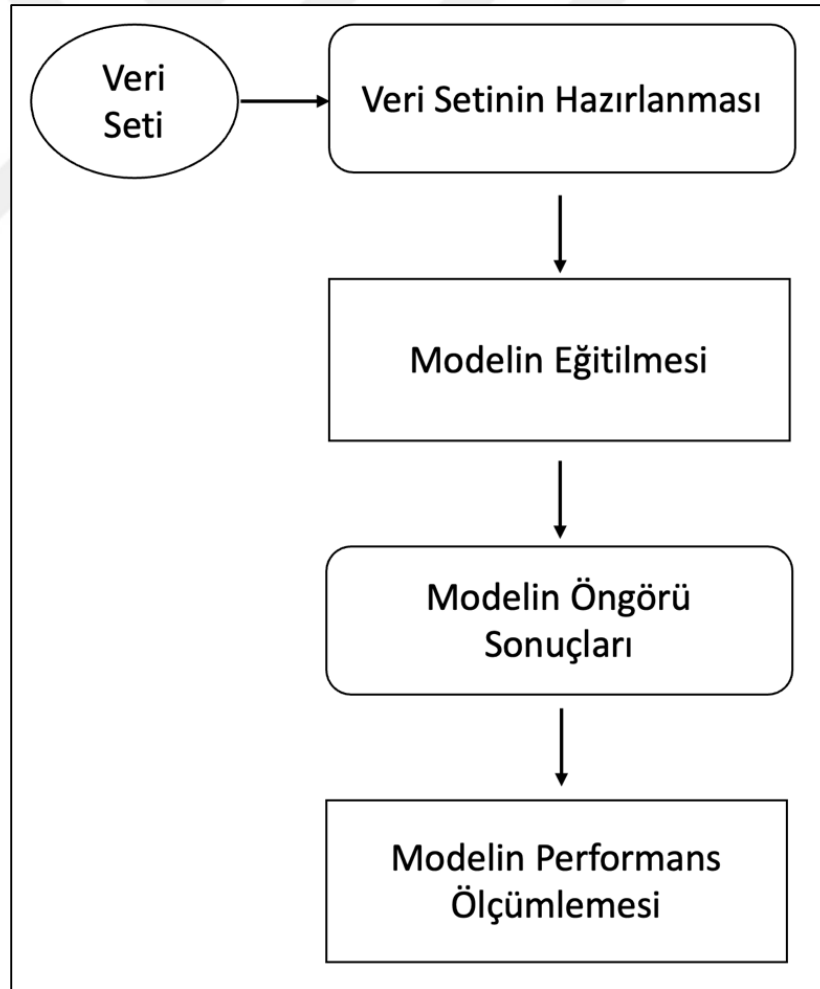
Nadiri vd. (2018), bir atıksu arıtma tesisinde deşarj kalite parametrelerini öngörmek için bir bulanık mantık modeli geliştirmiştir. Gerçek bir atıksu arıtma tesisi tarafından toplanan operasyonel verileri kullanarak, kimyasal oksijen ihtiyacı, biyokimyasal oksijen ihtiyacı, toplam askıda katı madde ve toplam azot dahil deşarj kalite parametrelerini öngörebilen bir model geliştirmişlerdir. Önerilen modelin, tesisin işletmesini optimize etmek ve çevresel etkisini azaltmak için faydalı olabilecek bir deşarj kalite parametreleri öngörüsü sağladığını göstermektedir.

Sonuç olarak, literatürde yapılan bu çalışmalar, atıksu arıtma tesislerinin işletmesini optimize etmek için makine öğrenmesi tekniklerinin kullanımını araştırmışlardır. Bu teknikler, işletme verilerine dayanarak deşarj kalitesini ve diğer değişkenleri öngörmek için güçlü bir araç sağlamakta olup, atıksu arıtma tesislerinin performansını artırarak çevresel etkisini azaltmaya yardımcı olmaktadır. Ayrıca, bu çalışmalar, makine öğrenmesi ve veri madenciliği tekniklerinin diğer çevresel sorunlara uygulanma potansiyelini göstererek, karmaşık sistemlerin dinamikleri hakkında değerli içgörüler sağlayarak daha verimli ve sürdürülebilir çözümler geliştirmeye yardımcı olmaktadır.

3. MATERYAL VE YÖNTEM

Bu çalışmada, karmaşık regresyon problemlerini ele almak bağlamında üç önemli model olan karar ağacı regresyonu (DTR), destek vektör makineleri (SVM) ve ileri beslemeli sinir ağı (FNN) modellerinin uygulanmasına odaklanılmıştır. Bu modeller, veri setleri içerisindeki karmaşık, doğrusal olmayan ilişkilerle başa çıkmada kanıtlanmış etkinlikleri nedeniyle seçilmiştir.

Atıksu arıtımının önemli parametreleri olan karbon, azot ve fosfora ait tesis çıkış değerlerinin öngörülmesinde bu üç modelin performansı incelenmiş, uygulanabilirliği değerlendirilmiş ve karşılaştırılmıştır. Önerilen yönteme ait genel akış şeması Şekil 3’ te verilmiştir.



Şekil 3. Önerilen yönteme ait genel akış şeması

3.1 Karar Ağacı Regresyonu

Karar ağacı regresyonu, öngörüs el modellemede ve veri analizinde g   l  bir ara  olarak ortaya  ıkmı  olup, karma ık veri k meleri ile ba a  ıkma ve yorumlanabilirlik  zellikleri sunmaktadır.

Karar ağacı regresyonu,  zellik uzayını belirli  zniteliklere g re tekrar tekrar b lerek ağa  benzeri hiyerar ik bir yapı olu turma s recini i erir. Ağa taki her d ğ m bir  zniteliđi temsil eder ve dallar,  zniteliđin alabileceđi olası deđerleri g sterir. D ğ mlerin b l nmesi, genellikle bilgi kazancı veya Gini indeksi gibi belirli kriterlere dayalı olarak ger ekle tirilir.

Bilgi kazancı, Quinlan (1993) tarafından tanıtılan, veri k mesini sınıflandırmada bir  zniteliđin etkinliđini deđerlendirmek i in temel bir  l  d r. Bir b l nme sonrasında belirsizliđin veya entropinin azalmasını nicelendirerek, veri k mesini b lmek i in en bilgilendirici  zniteliđi belirlemede yardımcı olur. Diđer taraftan, Breiman vd. (1984) tarafından ilk kez  nerilen Gini indeksi, karar ağa ındaki bir d ğ m n saflıđını  l er. Ağa  yapım s recinde Gini indeksinin en aza indirilmesi, homojen d ğ mlerin olu turulmasını sađlayarak modelin  ng r sel dođruluđunu arttırır.

3.1.1 Karar ağacı regresyonunun avantajları

Karar ağacı regresyonunun temel avantajlarından biri, i sel i leyi in anla ılabilir olmasıdır. Karar ağa ının hiyerar ik yapısı, ara tırmacıların ve uygulayıcıların girdi  znitelikleri ile hedef deđi ken arasındaki ili kileri anlamalarını sađlar. Bu yorumlanabilirlik, modelin i sel i leyi ini anlamının karar alma s re lerinde hayati  neme sahip olduđu alanlarda  zellikle deđerli olmaktadır.

Karar ağacı regresyonu, sayısal, kategorik ve sıralı gibi  e itli veri t rlerini etkili bir  ekilde i leme esnekliđi g sterir. Bir ok karma ık algoritmanın aksine, karar ağa ları eksik deđerleri ve aykırı deđerleri etkili bir  ekilde ele alabilir, veri manip lasyonu ihtiya ını azaltarak modelin sađlamlıđını sađlar.

Bir ok sofistike makine  đrenimi tekniklerine kıyasla, karar ağacı regresyonu d   k d zeyde veri  ni leme gerektirir. Kategorik ve sayısal verileri kapsayabilmesi, geni 

kapsamlı normalizasyon veya dönüşüm gerektirmemesi, veri önışleme için sınırlı kaynaklara ve zamana sahip senaryolarda çekici bir seçenek oluşturur.

Karar ağacı regresyonu, her bir düğüm içindeki yerel bilgiye dayanarak aykırılara karşı dayanıklılık gösterir. Verinin belirli bir bölgesindeki aykırı değerlerin genel ağaç yapısını önemli ölçüde etkilemesi olası değildir, bu da modelin öngörüşel performansının gürültülü verilerin varlığında dahi sabit kalmasını sağlar.

Karar ağacı regresyonunun esnekliğı, hem regresyon hem de sınıflandırma problemlerine uygulanabilmesini sağlar. Her düğümdeki karar kriterini değıştirerek, karar ağacı regresyonu çeşitli öngörüşel modelleme görevlerini ele alacak şekilde özelleştirilebilir, bu da veri bilimciler için çok yönlü ve geniş uygulanabilir bir araç olmasını sağlar (Breiman vd., 1984).

3.1.2 Karar ağacı regresyonunun sınırlamaları

Karar ağacı regresyonunun başlıca sınırlamalarından biri, gürültülü veya karmaşık veri kümelerinin varlığında aşırı öğrenmeye karşı duyarlı olmasıdır. Karar ağaçları, kısıtlamalar olmaksızın büyüdüğünde genellikle eğitim verilerini ezberleyerek, görünmeyen verilere kötü genelleme yapar. Aşırı öğrenme, modelin öngörüşel performansını ciddi şekilde etkiler ve gerçek dünya senaryolarında uygulanabilirliğini sınırlar.

Karar ağacı regresyonu, eğitim verisindeki küçük değışikliklere duyarlıdır ve bu da sonuçta elde edilen karar ağacı yapısında önemli değışikliklere yol açabilir. Veri kümesindeki hafif bir değışiklik bile modelin öngörüşel performansını önemli ölçüde etkileyebilir, bu da model oluşturma sürecinde tekrarlanabilirliğı ve sağlamlığı sağlamak için dikkatli düşünce gerektirir.

Karar ağacı regresyonu, dengesiz veri kümeleri üzerinde eğitildiğinde genellikle çoğunluk sınıfının dağılımını yansıtan önyargılı ağaçlar oluşturur. Bu durum, azınlık sınıfının önemli olduğı senaryolarda önyargılı tahminlere yol açabilir. Önyargılı ağaçların sorununu ele almak, modelin adil ve öngörüşel doğruluğunu korumak için uygun örnekleme stratejileri ve topluluk yöntemlerinin uygulanmasını gerektirir.

Karar ağacı regresyonu, yorumlanabilirlik, çeşitli veri türlerine uyum sağlama yeteneği ve aykırılara karşı dayanıklılık gibi özellikler sunarak öngörüsöl modelleme alanında değerli bir araç olarak karşımıza çıkmaktadır. Bununla birlikte, aşırı öğrenme ve küçük değişikliklere duyarlılık gibi zayıflıkları nedeniyle, genelleme yeteneğini ve öngörüsöl doğruluğunu artırmak için uygun düzenleme tekniklerinin ve topluluk yöntemlerinin uygulanmasını gerektirir (Breiman vd., 1984).

3.2 Destek Vektör Makineleri

Destek Vektör Makineleri (SVM'ler), sınıflandırma ve regresyon analizi için geniş ölçüde kullanılan bir denetimli öğrenme algoritması sınıfını temsil eder. Erken 1990'ların başında Vapnik ve meslektaşları tarafından tanıtılan SVM'ler, karmaşık veri sınıflandırma görevlerinde sağlam performansları nedeniyle önemli dikkat çekmiştir (Vapnik, 1995). İstatistiksel öğrenme teorisi prensiplerini kullanarak SVM'ler, veri noktalarını farklı sınıflara ayıran optimal bir hiperdüzlemi belirlemede uzmanırlar (Cortes ve Vapnik, 1995).

Destek Vektör Makineleri, farklı veri noktaları arasındaki marjı maksimize eden optimal bir hiperdüzlemi belirlemeye dayanır. Matematiksel olarak, $x_1, y_1, x_2, y_2, \dots, x_n, y_n$ gibi bir eğitim veri kümesi verildiğinde, x_i giriş özelliklerini temsil ederken; y_i karşılık gelen sınıf etiketlerini gösterir. Bir SVM'nin hedefi şu denklemle tanımlanan hiperdüzlemi belirlemektir:

$$w \cdot x - b = 0 \quad (3.1)$$

Burada w ağırlık vektörünü, b sapmayı ve x giriş vektörünü temsil eder. Optimal hiperdüzlem, aşağıdaki optimizasyon problemini çözerek elde edilir:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (3.2)$$

Şu kısıtlamalara tabidir:

$$y_i (w \cdot x_i - b) \geq 1 \quad (3.3)$$

$i = 1, 2, \dots, n$ için. Bu formülasyon, iki sınıf arasındaki marjın maksimize edildiğinden emin olur ve hiperdüzlemdeki en yakın veri noktalarının pozisyonunu belirleyen destek vektörleri olarak hareket eder (Boser vd., 1992).

3.2.1 Destek vektör makinelerinin avantajları

Destek Vektör Makineleri birkaç dikkate değer avantaj sunar, bu da onların farklı alanlarda yaygın olarak benimsenmesine katkıda bulunur. İlk olarak, SVM'ler yüksek boyutlu uzaylarda yüksek performans sergiler, bu da karmaşık veri kümeleri ve karmaşık sınıf sınırları içeren görevler için uygun hale getirir (Cortes ve Vapnik, 1995). Ayrıca, çekirdek hilesinin kullanımı, SVM'lerin doğrusal olmayan ayrılabilir verileri etkili bir şekilde işlemesini sağlar, böylece daha yüksek boyutlu bir uzaya dolaylı olarak haritalar ve böylece gelişmiş sınıflandırma doğruluğunu kolaylaştırır (Boser vd., 1992). Dahası, SVM'ler boyutluluğun laneti karşısında nispeten bağımsızlık gösterir, bu da sınırlı sayıda eğitim örneği içeren görevler için uygun hale getirir (Vapnik, 1995).

Ayrıca, SVM'ler istatistiksel öğrenme teorisi prensiplerine dayanan sağlam bir teorik temel sunar, bu da genelleme yeteneklerine katkıda bulunarak model seçimi ve düzenlileştirmeye prensip sahibi bir yaklaşım sağlar (Vapnik, 1995). Bu, aşırı uyumlanmaya karşı dayanıklılığı, SVM'leri sınırlı eğitim verisi içeren görevler için özellikle tercih edilebilir kılar ve gürültü veya aykırı değerler nedeniyle yanlış sınıflandırmayı önler.

3.2.2 Destek vektör makinelerinin dezavantajları

Birçok avantajına rağmen, Destek Vektör Makinelerinin de kısıtlamaları vardır. Temel dezavantajlardan biri, özellikle büyük ölçekli veri kümeleri söz konusu olduğunda, hesaplama karmaşıklıklarıdır, eğitim süresi ve bellek gereksinimleri önemli olabilir (Cortes ve Vapnik, 1995). Ayrıca, doğrusal olmayan verilerde uygun bir çekirdek fonksiyonunun seçimi zor olabilir, sıklıkla optimal sınıflandırma performansını elde etmek için uzmanlık ve dikkatli deney gerektirebilir (Boser vd., 1992). Bu çekirdek

fonksiyonunun doğru seçimiyle ilgili karmaşıklık, yorumlanabilirlik ve basitliğin ön planda olduğu senaryolarda SVM'lerin pratik uygulamasını engelleyebilir.

Ek olarak, SVM'ler hiperparametre seçimine duyarlı olabilir ve bu parametrelerdeki küçük değişiklikler sınıflandırma performansını önemli ölçüde etkileyebilir, dikkatli bir şekilde ayarlanmazlarsa alt-optimal sonuçlara yol açabilir (Vapnik, 1995). Dahası, SVM'ler doğrudan olasılık tahminleri sağlamaz, güvenilir olasılık tahminleri elde etmek için ek kalibrasyon teknikleri gerektirir, bu da uygulamada zahmetli ve hesaplama açısından yoğun olabilir (Cortes ve Vapnik, 1995).

Destek Vektör Makineleri, makine öğrenme alanında güçlü bir araç oluşturur ve karmaşık veri sınıflandırma görevlerine prensipli bir yaklaşım sunar. Yüksek boyutlu verilerle başa çıkma yetenekleri, aşırı uydurmaya karşı direnç göstermeleri ve güçlü teorik temelleri, farklı alanlarda kullanımlarını vurgular. Bununla birlikte, hesaplama karmaşıklığı, hiperparametrelere duyarlılık ve çekirdek seçimi ile ilgili zorluklar, SVM'leri gerçek dünya problemlerine uygularken dikkatli düşünce gerektirir. Bu kısıtlamalara rağmen, SVM'ler, sürekli olarak yeni tekniklerin ve yöntemlerin geliştirilmesiyle makine öğrenmesi alanında temel bir unsur olmaya devam etmektedir.

3.3 İleri Beslemeli Sinir Ağları

İleri Beslemeli Sinir Ağları (FNN) yapay sinir ağları (ANN) dünyasına temel mimari örneklerden biridir. İlk olarak biyolojik sinir ağlarının basitleştirilmiş modelleri olarak önerilen FNN'ler, desen tanıma, sınıflandırma ve regresyon analizi gibi çeşitli görevler için güçlü araçlar haline gelmiştir.

FNN'lerin teorik çerçevesi, birbirine bağlı düğümlerden oluşan katmanlar içeren çok katmanlı algılayıcıların (MLP'ler) prensiplerine dayanır. Bu nöronlar, giriş katmanı, gizli katmanlar ve çıkış katmanı gibi farklı katmanlara örgütlenmiştir. Bir katmandaki her bir nöron, takip eden katmandaki her nörona bağlıdır, bu da giriş katmanından gizli katmanlara ve çıkış katmanına doğru tek yönlü bir bilgi akışı oluşturur. Bu geri besleme mekanizması olmaksızın tek yönlü akış, bu ağların ileri besleme özelliğini karakterize eder.

Matematiksel olarak, bir FNN'nin fonksiyonu aşağıdaki gibi ifade edilebilir:

$$y = f(\sum_{i=1}^n w_i x_i + b) \quad (3.4)$$

burada y nöronun çıktısını temsil eder, f aktivasyon fonksiyonudur, w_i giriş nöronlarıyla ilişkili ağırlıkları gösterir, x_i önceki katmandan gelen girişi temsil eder ve b biası simgeler. Yaygın olarak kullanılan aktivasyon fonksiyonları arasında sigmoid fonksiyonu, hiperbolik tanjant fonksiyonu ve doğrusal doğru birim (ReLU) bulunur, her biri ağıın davranışı ve eğitim dinamiklerine farklı özellikler katar (Goodfellow vd., 2016).

Bir FNN'nin eğitimi, öngörülen çıktı ile gerçek çıktı arasındaki farkı en aza indirmek için ağıın ağırlıklarını ve biasını ayarlamayı kolaylaştıran geri yayılım gibi algoritmaların kullanımını içerir. Bu süreç, ağıın veri içindeki karmaşık desenleri ve ilişkileri öğrenmesini sağlamak için hayati önem taşır (Rumelhart vd., 1986).

3.3.1 İleri beslemeli sinir ağlarının avantajları

FNN'lerin temel avantajlarından biri, karmaşık, doğrusal olmayan fonksiyonları yaklaşabilme yetenekleridir, bu da onları karmaşık veri ilişkileri içeren görevler için uygun hale getirir. Hiyerarşik yapıları, ham verilerden karmaşık özelliklerin çıkarılmasına izin verir, etkili özellik temsili ve öğrenme sağlar. Ayrıca, FNN'ler, paralel işleme yetenekleri ve grafik işlemcileri gibi sağlam hesaplama çerçevelerinin bulunabilirliği sayesinde büyük ölçekli veri kümeleriyle verimli bir şekilde başa çıkabilirler (Nielsen, 2015). Ek olarak, FNN'ler, görünmeyen verilerde iyi performans gösterme yetenekleri sayesinde yüksek bir genelleme derecesine sahiptir, bu da onların çeşitli uygulamalardaki sağlamlığını gösterir.

Ayrıca, FNN'lerin yorumlanabilirliği, katman katman önem yayılımı (LRP) ve belirginlik haritaları gibi tekniklerin geliştirilmesiyle önemli ölçüde gelişmiştir, bu da kullanıcıların ağıın karar verme sürecine içgörü kazanmalarını sağlar. Bu yorumlanabilirlik, tıbbi teşhis veya finansal tahmin gibi kararların gerekçelendirilmesinin önemli olduğu alanlarda güven ve şeffaflığı teşvik eder (Montavon vd., 2018).

3.3.2 İleri beslemeli sinir ağlarının sınırlamaları

Dikkate değer yeteneklerine rağmen, FNN'lerin de sınırlamaları vardır. Özellikle küçük veri kümeleri veya karmaşık modellerle uğraşırken aşırı uyum sorununa yatkınlıkları dikkat çekicidir. Aşırı uyum, ağıın eğitim verilerine fazlasıyla uyum sağlaması sonucunda görünmeyen verilerde kötü genelleme performansına yol açar. Bu sorunu hafifletmek için düzenleme teknikleri, erken durdurma ve çapraz doğrulama gibi stratejiler kullanılmıştır. Ancak, model karmaşıklığı ile genelleme arasında optimal bir dengeye ulaşmak sürekli bir zorluk oluşturur (Goodfellow vd., 2016).

Ayrıca, FNN'ler, derin mimariler ve büyük ölçekli veri kümeleri ile uğraşırken hesaplama yoğunluğuna sahiptir, bu durum uzun süren eğitim süreleri ve artan kaynak gereksinimleri ile sonuçlanır. Bu kısıtlama, yüksek performanslı bilgi işlem altyapısına sınırlı erişimi olan araştırmacılar ve organizasyonlar için önemli bir engel oluşturur (Nielsen, 2015).

Yanı sıra, FNN'lerin doğuştan bellek eksikliği ve geri besleme mekanizmasının olmaması, zaman içindeki bağımlılıkları ve dinamik desenleri dizisel veriler içine entegre etme yeteneklerini kısıtlar. Bu kısıtlama, zaman serisi analizi, doğal dil işleme ve zamansal dinamiklerin kritik bir rol oynadığı diğer alanlar gibi görevler için FNN'leri daha az uygun hale getirir (Goodfellow vd., 2016).

3.4 Veri Seti

Bu çalışmada kullanılan veri seti, İstanbul'daki gerçek bir atıksu arıtma tesisinden 2017-2022 yıllarını kapsayan altı yıllık günlük operasyonel verileri kapsamaktadır ve giriş ve çıkış atıksu karakteristikleri ile ilgili kapsamlı bir parametre yelpazesini içermektedir. Bu parametreler arasında KOİ konsantrasyonları, azot ve fosfor seviyeleri, askıda katılar, sıcaklık ve çeşitli diğer operasyonel değişkenler bulunmaktadır.

Tablo 1. Çalışmada kullanılan veri setine dair örnek tablo

Date	Temp	COD inf	TN inf	NH4N inf	TP inf	SS inf	TN eff	NH4N eff	TP eff	SS eff	COD eff
1.06.2017	22,3	946,0	36,8	30,4	15,5	3168,0	4,6	1,6	0,5	22,0	41,0
2.06.2017	23,3	258,0	38,9	27,5	1,1	1054,0	5,4	1,6	0,2	23,0	26,0
3.06.2017	24,7	521,0	36,4	15,5	2,0	988,0	8,3	1,5	0,2	12,0	32,0
4.06.2017	22,6	395,0	34,2	26,1	4,0	726,0	8,8	3,5	0,2	27,0	47,0
5.06.2017	21,7	556,0	33,0	23,5	11,0	1200,0	8,8	0,1	0,2	26,0	20,0
...	...	...	...	...	...	...	...	...	...	...	...

3.5 Önerilen Yöntem

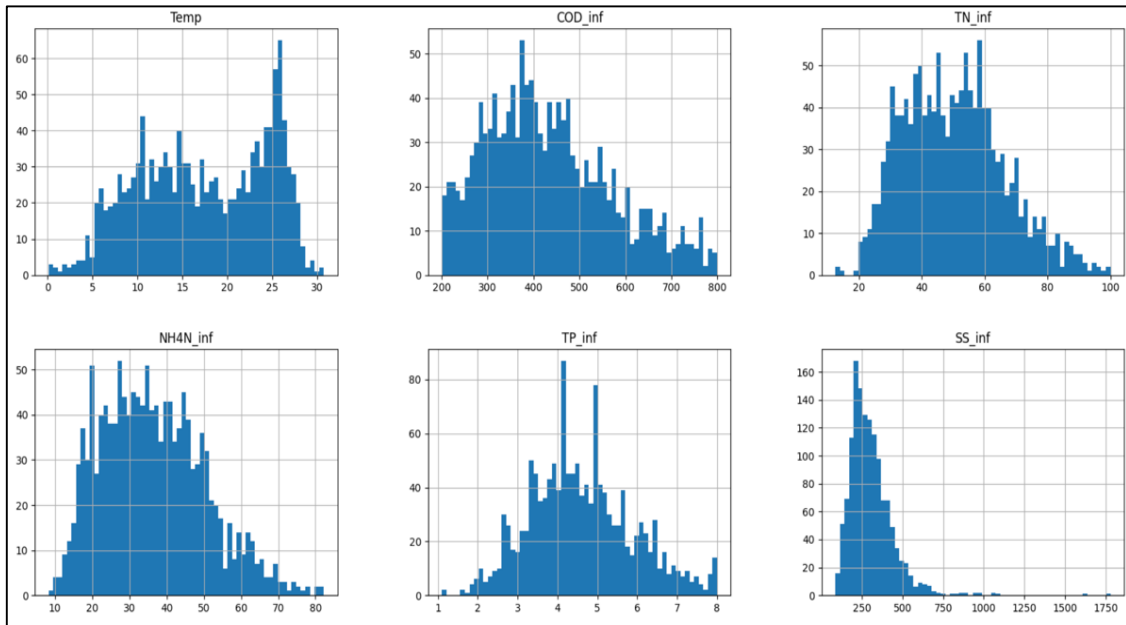
Atıksu arıtma tesislerinde deşarj kalitesi önemli bir husustur ve Kimyasal Oksijen İhtiyacı (KOİ), Toplam Azot ve Toplam Fosfor gibi temel parametrelerin konsantrasyonlarının doğru bir şekilde tahmin edilmesi, etkili tesis işletmesi ve çevresel düzenlemelere uyum sağlamak için kritiktir. Bu çalışmada gerçek bir Atıksu Arıtma Tesisi'nden elde edilen altı yıl boyunca günlük veriler üzerinden üç makine öğrenmesi modeli; Karar Ağacı Regresyonu, Destek Vektör Makineleri (SVM) ve İleri Beslemeli Sinir Ağları kullanılarak deşarj konsantrasyonlarının öngörülmesine odaklanılmıştır. Seçilen modeller Python programlama dili üzerinde, ilgili Scikit-Learn ve Pandas kütüphaneleri kullanılarak, Google Colab ortamında geliştirilmiştir. Google Colab, internet tarayıcısı aracılığı ile Google tarafından sağlanan bilgisayarlarda Jupyter Notebook üzerinde Python kodu çalıştırmaya olanak sağlayan bir platformdur. Veri seti, dikkatlice eğitim ve test setlerine ayrılmıştır; eğitim seti model eğitimi ve hiperparametre ayarlamayı sağlarken, test seti tahmin performansını değerlendirmek üzere kullanılmaktadır. Eğitim seti ile eğitilen modeller, sonrasında test setinin kullanılması ile performans değerlendirmesine tabi tutulmuştur.

Tablo 2. Çalışma kapsamında kullanılan modeller ve değerlendirilen parametreler

Model	DTR	SVM	FNN
	KOİ	KOİ	KOİ
Parametre	Toplam Azot	Toplam Azot	Toplam Azot
	Toplam Fosfor	Toplam Fosfor	Toplam Fosfor

3.5.1 Veri ön işleme

İlk adım, veri setinin kalitesini ve güvenilirliğini sağlamak için titiz bir veri ön işleme işlemidir. Bu işlem, eksik verilerle başa çıkmayı, aykırı değerleri ele almayı ve veriyi normalize etmeyi içerir. Veri setinin günlük doğasından kaynaklanan özelliklere özel dikkat gösterilmektedir, böylece doğru model eğitimi ve güvenilir tahminler elde edilmesi sağlanmaktadır. Veri setinde bulunması olası sapma değerler, hatalar vb. faktörleri belirlemenin en etkili yöntemlerinden biri histogramlardır. Her bir parametreye ait histogramlar oluşturularak parametrelere ait verilerin durumu yakından incelenir ve gerekli aksiyonlar alınır. Oluşturulan histogramlara ait örnek görsel Şekil 4’te verilmiştir.



Şekil 4. Çalışmada kullanılan veri seti histogramlarına ait örnek görsel

Veri Temizleme: İlk adım, eksik veya hatalı değerlerin belirlendiği ve uygun bir şekilde ele alındığı kapsamlı bir veri temizleme sürecini içerir. Eksik veri noktaları, veri seti bütünlüğünü korumak için dikkatli bir şekilde yerine konur (Hashem vd., 2015).

Özellik Seçimi: Daha sonra, seçilen deşarj konsantrasyonlarını tahmin etmek için en uygun girdi değişkenlerini belirlemek için özellik seçimi metodolojileri uygulanır. Bu özellik seçimi süreci çift katmanlı bir amaç hizmet eder: boyutu en aza indirerek hesaplama verimliliğini optimize eder ve en etkili faktörlere odaklanarak model yorumlanabilirliğini artırır (Hatcher ve Yu, 2018).

Veri Bölme: Veri seti dikkatlice üç farklı alt kümeye bölünür: eğitim, doğrulama ve test setleri. Bu bölme, makine öğrenmesi iş akışı için temel bir adımdır, çünkü model geliştirmeyi, hiperparametre ayarlama işlemlerini ve sağlam bir değerlendirme yapmayı kolaylaştırır (Hatcher ve Yu, 2018).

3.5.2 Veri setinin bölünmesi

Altı yıllık veri seti, eğitim ve test setlerine bölünmüştür. Eğitim seti, verinin çoğunluğunu oluşturarak model eğitime ve hiperparametre ayarlama olanağı sağlar. Test seti ise daha sonraki bir zaman dilimini temsil eder ve modellerin eğitim dönemini aşan verilere ne kadar iyi genelleme yapabildiğini değerlendirmek için kullanılır.

3.5.3 Model eğitimi

Eğitim süreci, her bir modelin eğitim veri setine beslenmesini, hiperparametrelerin ayarlanmasını ve modellerin tahmin performansı için optimize edilmesini içerir. Hiperparametre ayarlama sağlaması sağlamak için çapraz doğrulama (cross-validation) gibi çeşitli yöntemler kullanılabilir. Amaç, her bir modelin girdi özellikleri (örneğin, atıksu giriş karakteristikleri, işletme parametreleri) ile deşarj konsantrasyonları arasındaki temel ilişkileri öğrenmesine izin vermektir.

Üç farklı makine öğrenimi modeli, birbirini tamamlayan avantajları nedeniyle seçilmiştir. Seçilen modeller Python programlama dili üzerinde, ilgili Scikit-Learn ve Pandas kütüphaneleri kullanılarak, Google Colab ortamında geliştirilmiştir.

Karar Ağacı Regresyonu (DTR), yorumlanabilirliği ve verideki doğrusal olmayan ilişkileri ele alma yeteneği nedeniyle seçilmiştir. Karar ağaçları, karmaşık karar sınırlarını yakalamada özellikle başarılıdır ve aykırı değerlere karşı dirençlidir. Karar Ağacı Regresyonu (DTR) modelinde, her bir düğümün bir kararı temsil ettiği bir ağaç benzeri bir yapı oluşturmak için eğitilir. Ağaç, bilgi kazancına dayanarak genişletilir ve aşırı uyuma karşı önlem almak için budama uygulanabilir.

Destek Vektör Makineleri (SVM) hem doğrusal hem de doğrusal olmayan ilişkilerle başa çıkma yetenekleri nedeniyle seçilmiştir. Çekirdek hilesi, SVM'leri girdi verilerini daha yüksek boyutlu uzaylara eşlemek için kullanır, karmaşık desenleri yakalamada etkili bir araç sağlar. Destek Vektör Makineleri (SVM) modelinde veriyi deşarj konsantrasyonlarına karşılık gelen sınıflara en iyi şekilde ayıran hiperplanı bulmak üzere eğitilir. Modelin çekirdek parametreleri, sınıflar arasındaki marjı maksimize etmek için ayarlanır.

İleri Beslemeli Sinir Ağları (FNN), özellikle büyük veri setlerinde karmaşık desenleri yakalamada etkili olan hiyerarşik temsiller öğrenme yetenekleri nedeniyle seçilmiştir. İleri Beslemeli Sinir Ağları (FNN) modelinde sinir ağı, geriye yayılım kullanılarak eğitilir; ağırlıkları ve sapmaları, tahmin edilen ve gerçek deşarj konsantrasyonları arasındaki hatayı minimize etmek için ayarlar. Mimarisi, katman sayısı, katman başına düğüm sayısı ve aktivasyon fonksiyonları gibi unsurlar bu süreçte ayarlanır.

3.5.4 Hiper-parametre ayarlama

Her modelin hiperparametreleri, eğitim veri setinde performanslarını optimize etmek için ayarlanır. “Grid Search” veya “Random Search” gibi teknikler, hiperparametre uzayını etkili bir şekilde keşfetmek için kullanılabilir. Amaç, aşırı uyumdan kaçınılarak eğitim verisine en iyi tahmin performansını sağlayacak hiperparametre kombinasyonunu bulmaktır.

Karar Ağacı Regresyonu (DTR) modelinde “ağaç derinliği”, “düğüm ayırma büyüklüğü” ve “yaprak büyüklüğü”; Destek Vektör Makineleri (SVM) modelinde “kernel”, “C” ve “gamma” gibi hiper-parametreler çapraz doğrulama (cross-validation) ile ayarlanarak uygulanmıştır.

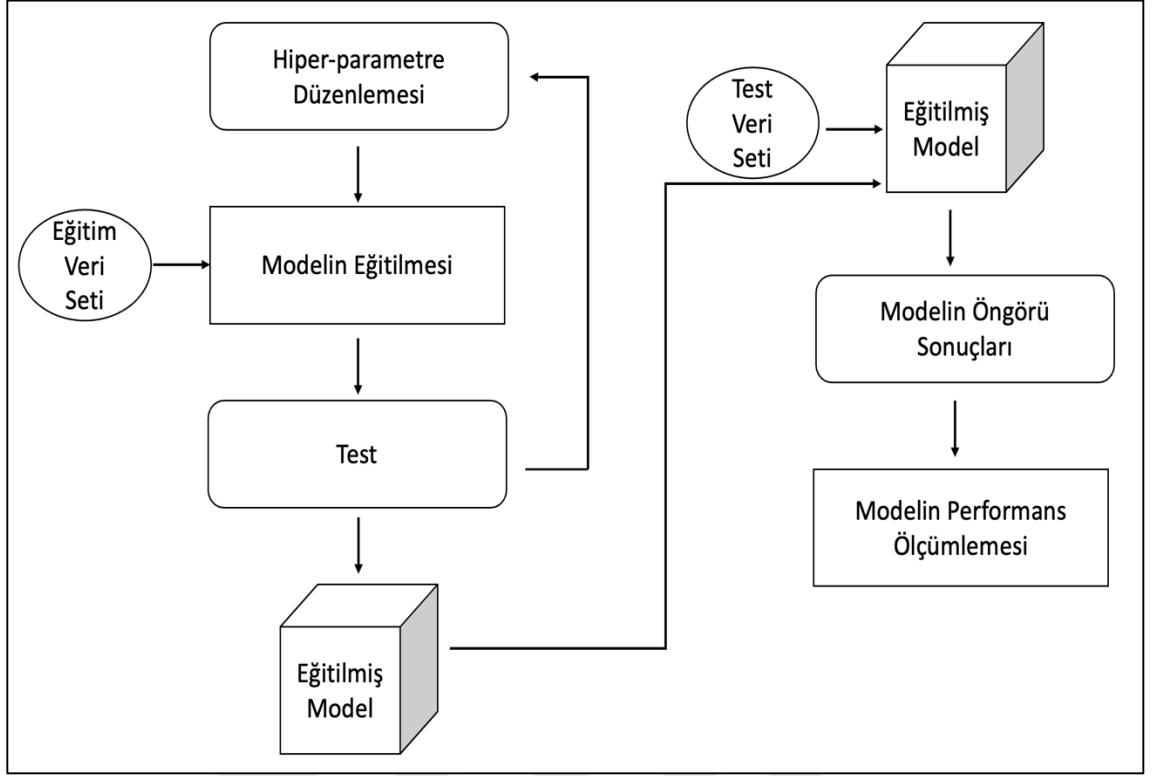
3.5.5 Model deęerlendirmesi

Eęitim ařamasını takiben modeller, tahmin yeteneklerini grnmeyen veri zerinde deęerlendirmek iin test veri setinde deęerlendirilir. Alınan sonular zerinde uygulanan iki temel performans metrięi, Kk Ortalama Kare Hata (RMSE) ve Ortalama Mutlak Hata (MAE), modellerin doęruluęunu lmek iin kullanılır. Bu metrikler, hem tahmin hatalarının byklęn hem de ynn dikkate alarak modellerin deęarj konsantrasyonlarını tahmin etme yetenekleri hakkında kapsamlı bir anlayıř saęlar.

Her bir modelin performansı, deęerlendirme metriklerine dayanarak sistematik bir řekilde karřılařtırılır. Ama, her bir modelin deęarj konsantrasyonlarını tahmin etmedeki gl ve zayıf ynlerini belirlemektir. İnterpretasyon kabiliyeti, hesaplama verimlilięi ve tahmin doęruluęu gibi faktrler, karřılařtırmalı analizde dikkate alınır. RMSE ve MAE gibi deęerlendirme metriklerinin seimi, deęerlendirme srecine farklı doęruluk ynlerinden bakarak modellerin performansı hakkında igrler sunar.

RMSE ve MAE dahil edilen model deęerlendirmelerinden elde edilen sonular, Karar Aęacı Regresyonu, Destek Vektr Makineleri ve İleri Beslemeli Sinir Aęları'nın atıksu arıtma tesisi deęarj konsantrasyonlarını tahmin etme konusundaki kullanılabilirlik ve gvenilirlięi hakkında detaylı bir řekilde yorumlanır.

Veri setinin hazırlanması, seilen modellerin eęitimi ve performanslarını artırmak iin uygulanan hiper-parametre ayarlamaları, eęitim sonucu oluřturulan modellerin test veri seti ile ngr performanslarının deęerlendirilmesinden oluřan yntemin akıř řeması řekil 5'te verilmiřtir;



řekil 5. Uygulanan y ntemin akıř řeması

4. BULGULAR VE TARTIŞMA

Gelişmiş makine öğrenmesi ve yapay zeka algoritmalarının uygulanması atıksu arıtım sistemlerinin tahmin edilmesi ve optimizasyonu için değerli araçlar sunmaktadır. Bu bölümde çalışma kapsamında kullanılan Karar Ağacı Regresyonu (DTR), Destek Vektör Makineleri (SVM) ve İleri Beslemeli Sinir Ağları (FNN) modellerinin atıksu arıtma sürecinden çıkan Kimyasal Oksijen İhtiyacı (KOİ), Toplam Azot ve Toplam Fosfor parametrelerine ait konsantrasyonlara dair performans bulguları paylaşılmış olup, bulguların değerlendirmesi yapılmıştır.

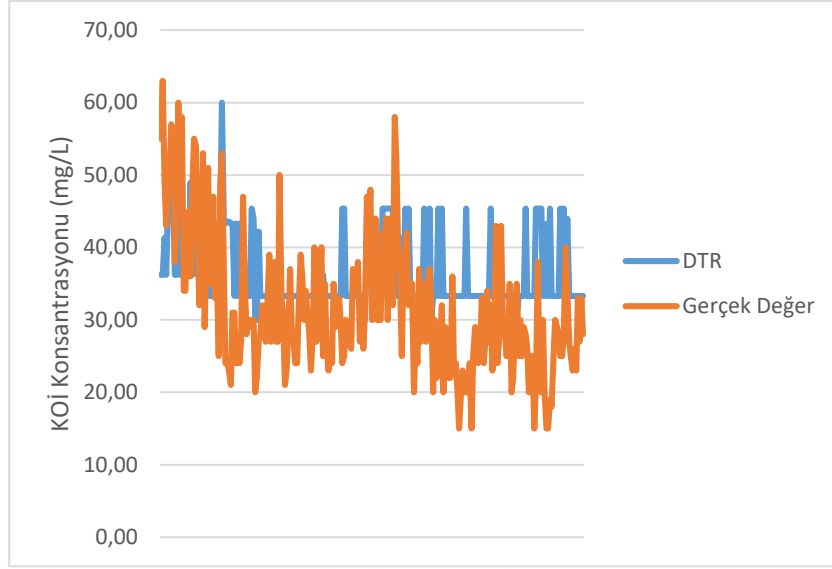
4.1 Bulgular

Bu bölümde çalışma kapsamında kullanılan Kara Ağacı Regresyonu (DTR), Destek Vektör Makineleri (SVM) ve İleri Beslemeli Sinir Ağları (FNN) modellerinin atıksu arıtma sürecinden çıkan Kimyasal Oksijen İhtiyacı (KOİ), Toplam Azot ve Toplam Fosfor parametrelerine ait konsantrasyonlara dair performans bulguları paylaşılmıştır.

4.1.1 Karar ağacı regresyonu

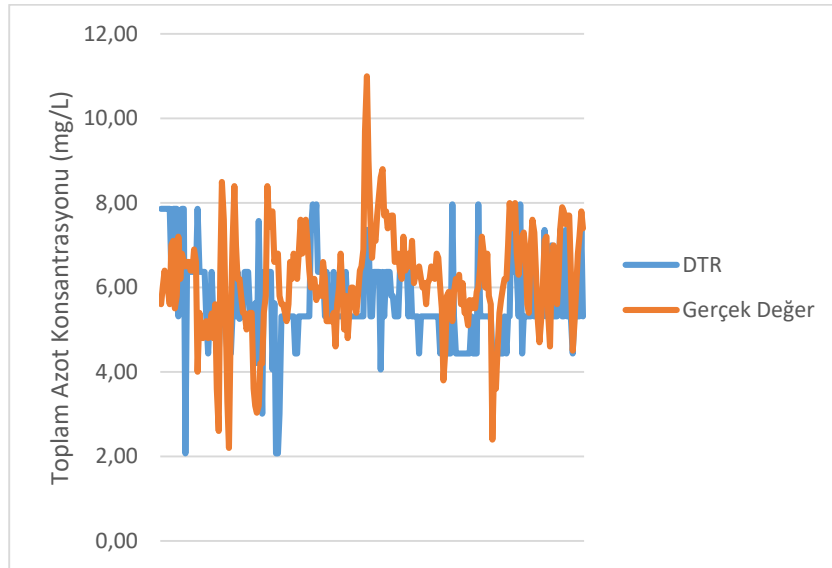
Atıksu arıtma tesislerinde çıkış suyuna ait kimyasal oksijen ihtiyacı (KOİ), toplam azot ve toplam fosfor konsantrasyonlarını tahmin etmede Karar Ağacı Regresyonunun performansı, kök ortalama karesel hata (RMSE) ve ortalama mutlak hata (MAE) metrikleri kullanılarak değerlendirilmiştir.

Seçilen parametrelerin atıksu konsantrasyonlarını tahmin etmede Karar Ağacı Regresyonu (DTR) modelinin analizi, etkinliği hakkında değerli bilgiler sunmaktadır. KOİ parametresi için RMSE, modelin tahminlerinin ortalama olarak gerçek değerlerden 10.27 mg/L sapma gösterdiğini gösterirken, 8.36 mg/L'lik MAE, tahminlerdeki hataların ortalama büyüklüğünü belirtmektedir. DTR modelinin KOİ konsantrasyonlarını tahmin etme performansı, atıksu arıtma süreçlerinin dinamik yapısıyla ilişkili karmaşıklıklar göz önüne alındığında özellikle dikkate değerdir. Karar Ağacı Regresyonu modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 6'da görülmektedir.



Şekil 6. Karar Ağacı Regresyonu modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

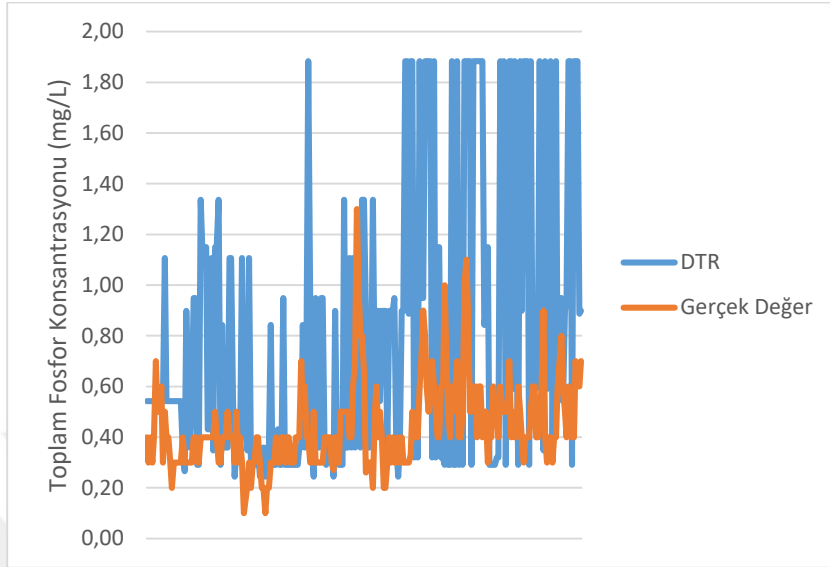
Benzer şekilde, Toplam Azot parametresi için DTR modeli 1.52 mg/L RMSE ve 1.21 mg/L MAE sergilemektedir. Karar Ağacı Regresyonu modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 7’de verilmiştir.



Şekil 7. Karar Ağacı Regresyonu modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

Yanı sıra, Toplam Fosfor parametresi için DTR modeli 0.66 mg/L RMSE ve 0.47 mg/L MAE sonuçlarını vermiştir. Karar Ağacı Regresyonu modelinin çıkış toplam fosfor

konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 8’de görülmektedir.

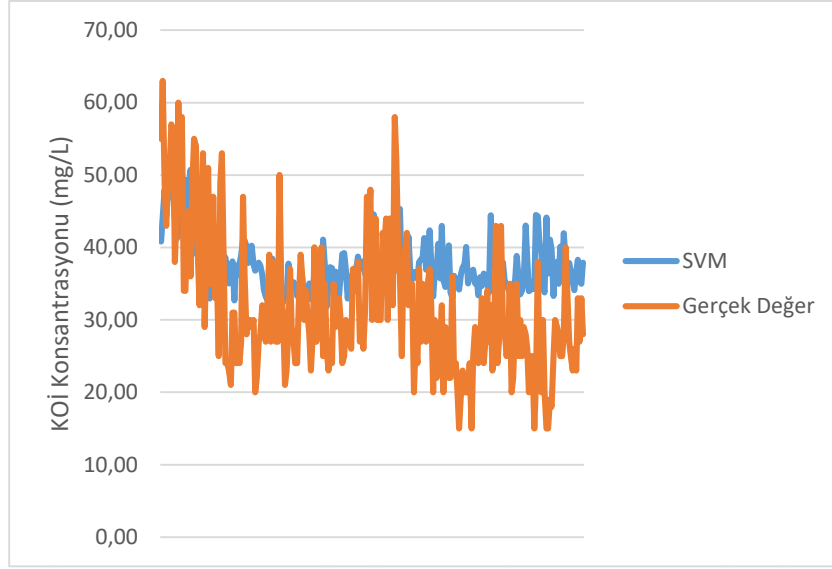


Şekil 8. Karar Ağacı Regresyonu modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

4.1.2 Destek vektör makineleri

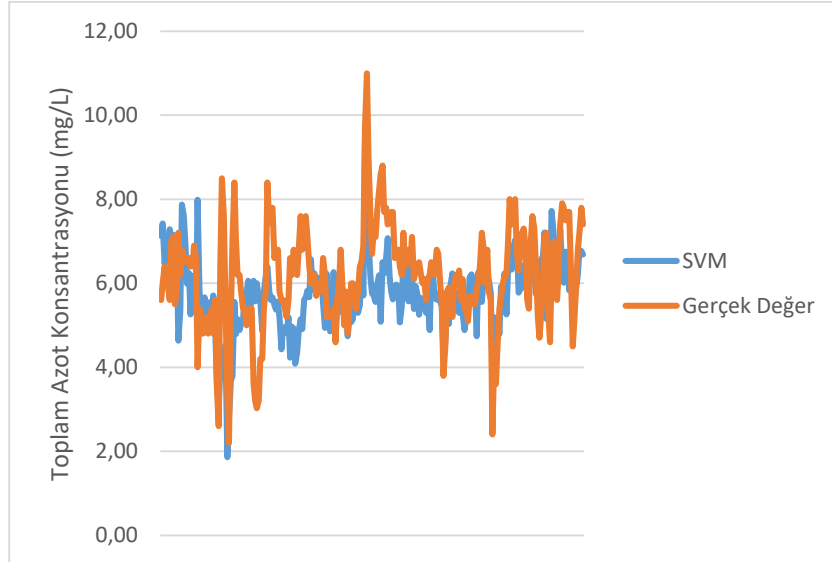
Atıksu arıtma tesislerinde çıkış suyuna ait kimyasal oksijen ihtiyacı (KOİ), toplam azot ve toplam fosfor konsantrasyonlarını tahmin etmede Destek Vektör Makinelerinin performansı, kök ortalama karesel hata (RMSE) ve ortalama mutlak hata (MAE) metrikleri kullanılarak değerlendirilmiştir.

Kimyasal Oksijen İhtiyacı (KOİ), Toplam Azot ve Toplam Fosforun atıksu konsantrasyonlarını tahmin etmede Destek Vektör Makineleri (SVM) algoritmasının performans analizi, tahmin yetenekleri hakkında değerli içgörüler sunmaktadır. KOİ parametresi için, SVM modeli 10.33 mg/L RMSE ve 8.77 mg/L MAE sergilemiş olup; DTR modeli ile karşılaştırılabilir bir performans göstermiştir. Destek Vektör Makineleri modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 9’da paylaşılmıştır.



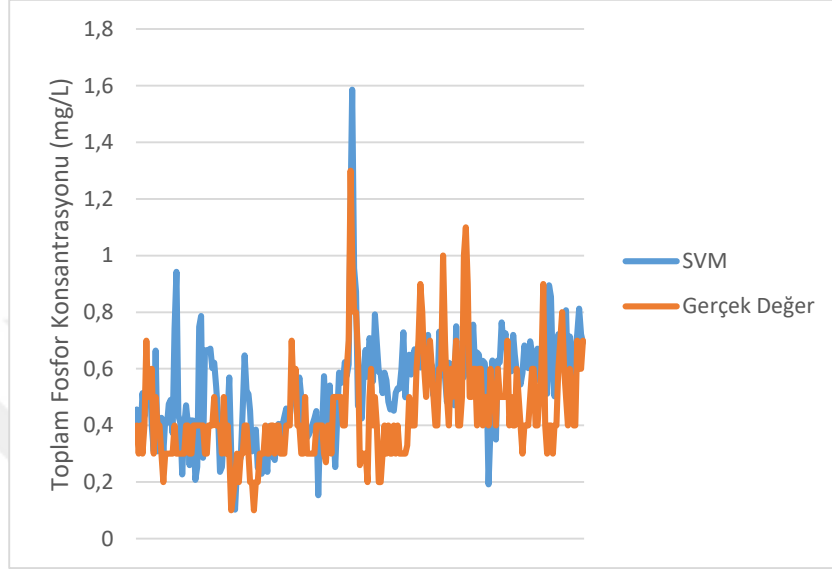
Şekil 9. Destek Vektör Makineleri modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

Toplam Azot parametresi için ise, SVM modeli 1.22 mg/L RMSE ve 0.95 mg/L MAE sonuçlarını vermiştir. DTR modeline kıyasla göreceli olarak iyileştirilmiş bir tahmin doğruluğunu işaret etmektedir. Destek Vektör Makineleri modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 10'da görülmektedir.



Şekil 10. Destek Vektör Makineleri modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

Çıkış Toplam Fosfor konsantrasyonu tahmini durumunda, Destek Vektör Makineleri yaklaşımı RMSE için 0.23 mg/L ve MAE için 0.17 mg/L sonuçlarını vermiştir. Destek Vektör Makineleri modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 11’de verilmiştir.

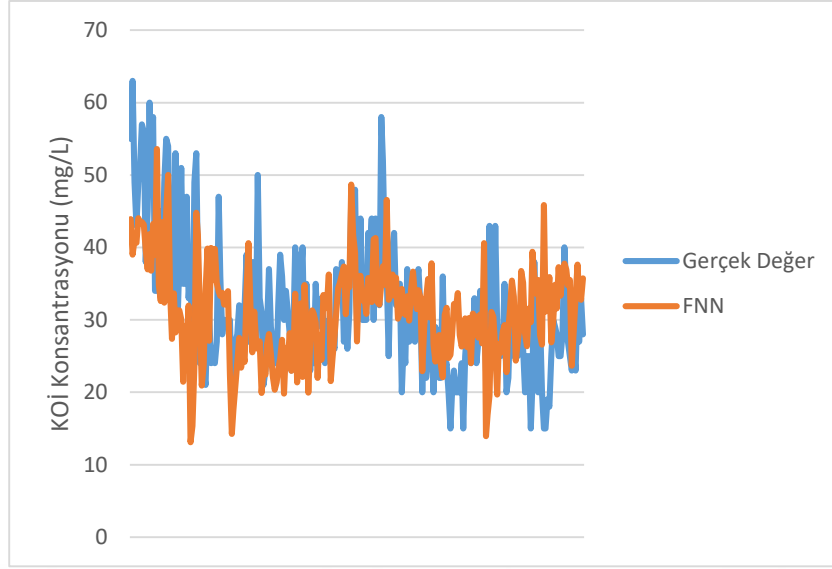


Şekil 11. Destek Vektör Makineleri modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

4.1.3 İleri beslemeli sinir ağları

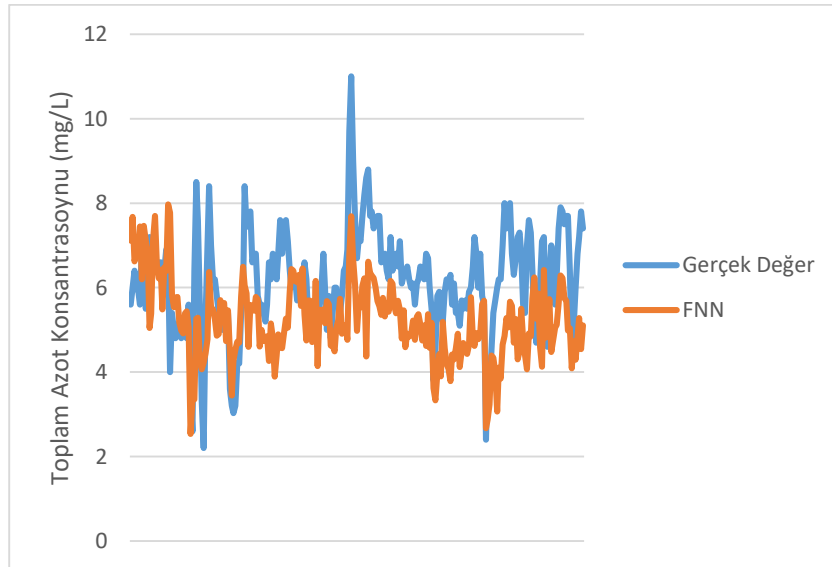
Atıksu arıtma tesislerinde çıkış suyuna ait kimyasal oksijen ihtiyacı (KOİ), toplam azot ve toplam fosfor konsantrasyonlarını tahmin etmede İleri Beslemeli Sinir Ağlarının performansı, kök ortalama karesel hata (RMSE) ve ortalama mutlak hata (MAE) metrikleri kullanılarak değerlendirilmiştir.

KOİ parametresi için, İleri Beslemeli Sinir Ağları modeli 9.03 mg/L RMSE ve 7.08 mg/L MAE göstermektedir ve hem DTR hem de SVM modellerine kıyasla iyileştirilmiş bir tahmin doğruluğunu işaret etmektedir. FNN modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 12’de görülmektedir.



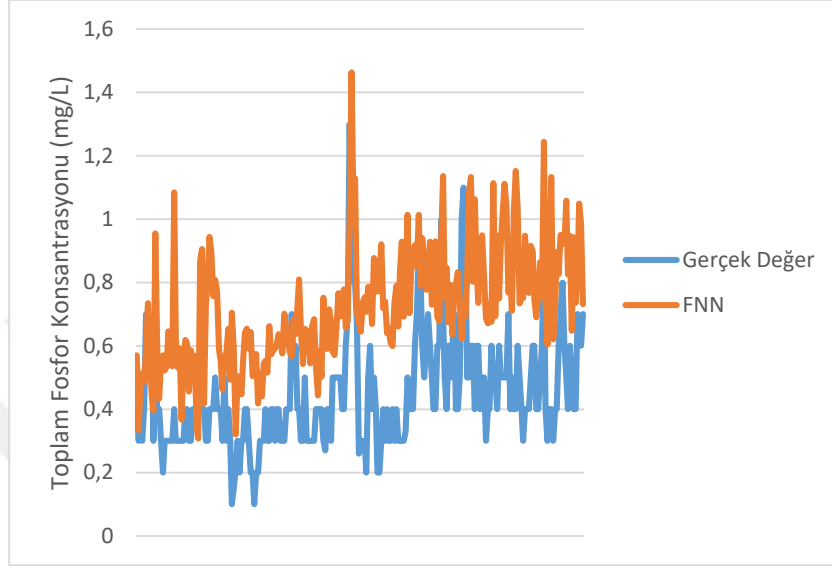
Şekil 12. İleri Beslemeli Sinir Ağları modelinin çıkış KOİ konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

Çıkış Toplam Azot konsantrasyonu tahmini durumunda, İleri Beslemeli Sinir Ağları yaklaşımı RMSE için 1.52 mg/L ve MAE için 1.27 mg/L sonuçlarını vermiştir. İleri Beslemeli Sinir Ağları modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 13'te verilmiştir.



Şekil 13. İleri Beslemeli Sinir Ağları modelinin çıkış toplam azot konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

Çıkış Toplam Fosfor parametresi için, İleri Beslemeli Sinir Ağları yaklaşımı RMSE için 0.37 mg/L ve MAE için 0.31 mg/L sonuçlarını vermiştir. İleri Beslemeli Sinir Ağları modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması Şekil 14’te paylaşılmıştır.



Şekil 14. İleri Beslemeli Sinir Ağları modelinin çıkış toplam fosfor konsantrasyonları için yaptığı tahminlerin gerçek değerlerle olan karşılaştırması

4.2 Tartışma

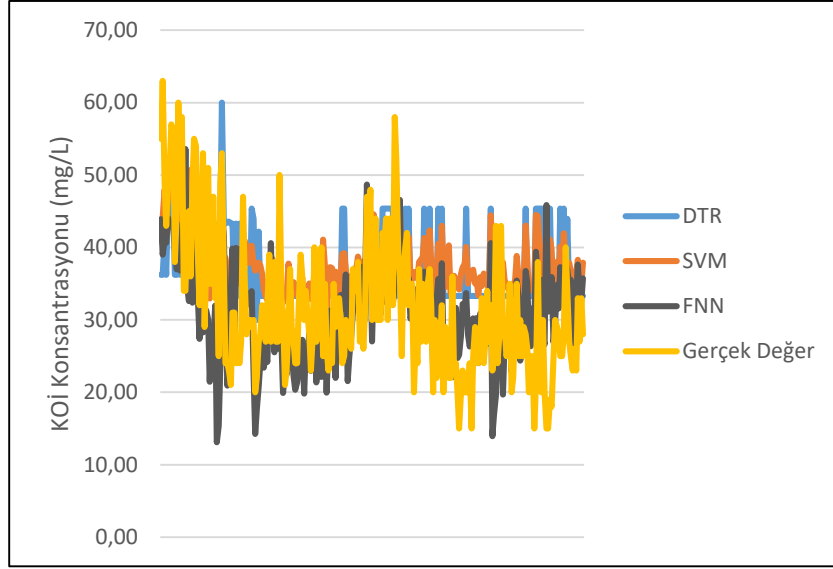
Üç makine öğrenimi algoritmasının, yani Karar Ağacı Regresyonu (DTR), Destek Vektör Makineleri (SVM) ve İleri Beslemeli Sinir Ağları (FNN) modellerinin, KOİ, Toplam Azot ve Toplam Fosforun çıkış atıksu konsantrasyonlarını tahmin etmedeki performanslarının karşılaştırmalı analizi, ilgili güçlü yönleri ve kısıtlamaları hakkında değerli içgörüler sunmaktadır.

Tablo 3. Karar Ağacı Regresyonu, Destek Vektör Makineleri ve İleri Beslemeli Sinir Ağları modelleri performansları

	DTR			SVM			FNN		
	KOİ	TA	TF	KOİ	TA	TF	KOİ	TA	TF
RMSE (mg/L)	10.27	1.52	0.66	10.33	1.22	0.23	9.03	1.52	0.37
MAE (mg/L)	8.36	1.21	0.47	8.77	0.95	0.17	7.08	1.27	0.31

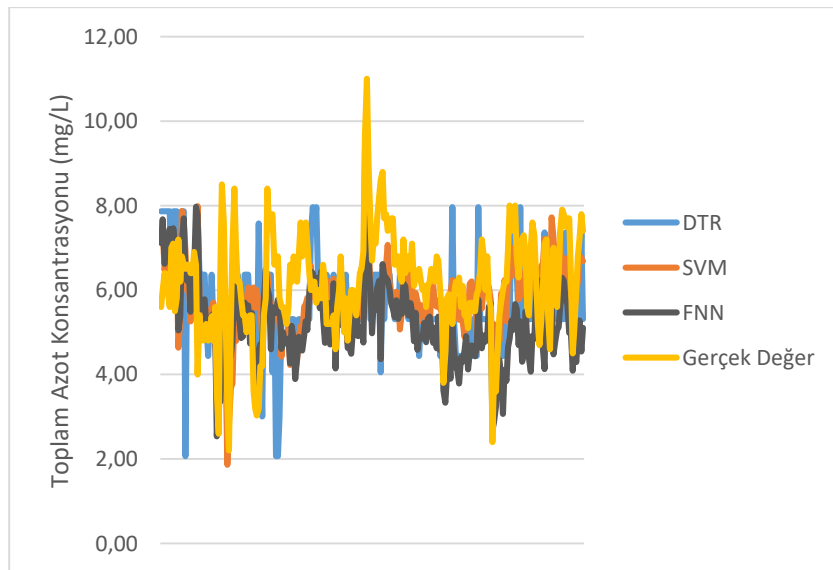
SVM modeli, üç parametrenin tümünü tahmin etmede orta ila yüksek doğruluk gösterirken, performansı özellikle KOİ durumunda üstün tahmin doğruluğu sergileyerek, atıksu arıtımı stratejileri için potansiyelini göstermektedir. Öte yandan, FNN modeli, KOİ tahmini için rekabetçi bir performans sergilerken, Toplam Azot için performansı hem DTR hem de SVM modelleriyle karşılaştırıldığında göreceli olarak daha az doğru olmaktadır. Bununla birlikte, FNN modeli Toplam Fosfor için SVM modeliyle benzer performans sergilerken DTR modelinden hafif derecede daha iyi performans göstererek, bu parametrenin doğru tahmin edilmesi konusundaki potansiyelini vurgulamaktadır.

KOİ parametresinin deneysel ölçümünde kabul edilen hata payı olan ~10 mg/L göz önünde bulundurulduğunda her üç modelin de bu parametre için güvenilir performans gösterdiği görülmektedir. Bu durum da işletme süreçlerinin daha hassas izlenmesini ve kontrol edilmesini kolaylaştırmaktadır.



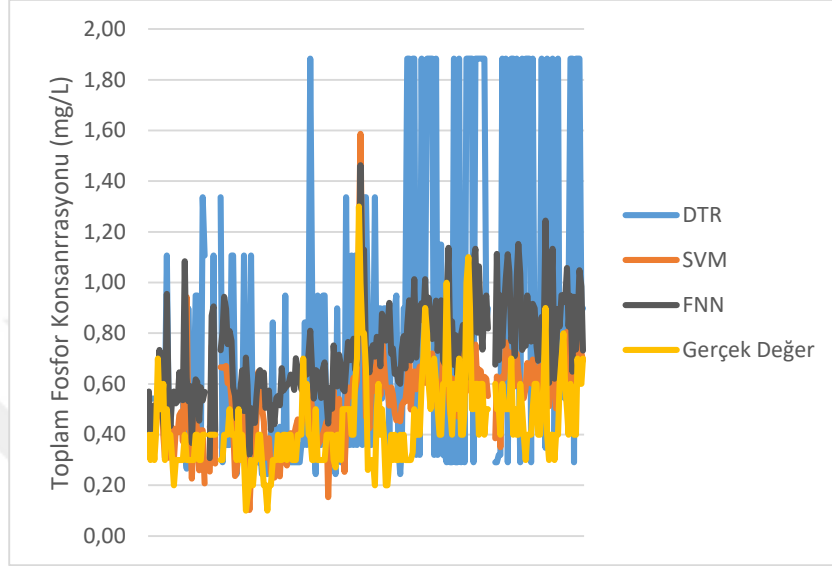
Şekil 15. KOİ konsantrasyonlarının gerçek değerler ile karşılaştırılması

Toplam Azot parametresinin deşarj standartları (10 mg/L) ele alındığında ise SVM modelinin başarılı olduğu görülürken, DTR ve FNN modelleri başta olmak üzere SVM modelinin de operasyonel şartlar sebebiyle deşarj konsantrasyonlarının 10 mg/L seviyelerine yaklaşması durumunda işletme sorunlarına yol açma ihtimalleri bulunmaktadır. Azot parametresi söz konusu olduğunda özel parametre performansını artırmak için sinir ağı mimarisinin daha fazla optimize edilmesi veya ayarlanması ihtiyacını yada hibrit modellerin kullanılmasının gerekebileceğini ortaya koymaktadır.



Şekil 16. Toplam Azot konsantrasyonlarının gerçek değerler ile karşılaştırılması

Toplam Fosfor parametresi incelendiğinde ise; deşarj standartları (1 mg/L) göz önünde bulundurulduğunda DTR ve FNN modellerinin hata paylarının tesis işletmesi açısından güvenilir gözükmediği, SVM modelinin ise işletme şartlarının deşarj standardına yaklaşmadığı sürece kabul edilebilir performans sergilediği gözlemlenmiştir.



Şekil 17. Toplam Fosfor konsantrasyonlarının gerçek değerler ile karşılaştırılması

Sonuç olarak, atıksu arıtımında KOİ, Toplam Azot ve Toplam Fosforun atıksu konsantrasyonlarını tahmin etmede DTR, SVM ve FNN gibi üç önemli makine öğrenimi algoritmasının performansları kapsamlı bir şekilde değerlendirilmiştir. RMSE ve MAE metriklerini kullanarak gerçekleştirilen analiz, seçilen parametrelerin tahmin edilmesinde her bir algoritmanın belirgin güçlü ve zayıf yönlerini ortaya koymaktadır. SVM modeli tüm parametreler için orta ila yüksek tahmin doğruluğu gösterirken, FNN modeli, KOİ için rekabetçi tahmin doğruluğu göstermektedir.

5. SONUÇ ve ÖNERİLER

Bu çalışma, atıksu konsantrasyonlarını tahmin etmede her makine öğrenimi algoritmasının belirli güçlü ve zayıf yönlerini vurgulayarak, atıksu arıtım süreçlerinde tahmin modellemesi yaklaşımlarının ilerlemesine katkıda bulunmaktadır. Modellerin tahmin performanslarındaki belirlenen farklılıklar, her bir parametrenin belirli gereksinimlerine dayalı olarak uygun algoritmaların seçilmesinin önemini vurgulamaktadır. Ayrıca, bulgular, makine öğrenimi modellerinin tahmin yeteneklerini artırmak için daha fazla optimizasyon ve ayarlama olanak sağladığını vurgulayarak, atıksu arıtma sistemlerinin daha hassas ve etkili bir şekilde yönetilmesini mümkün kılmaktadır.

Atıksu arıtımında çıkış kimyasal oksijen ihtiyacı, toplam azot ve toplam fosfor konsantrasyonu tahmininin kritik bir konusuna, özellikle Karar Ağacı Regresyonu (DTR), Destek Vektör Makineleri (SVM) ve İleri Beslemeli Sinir Ağları (FNN) gibi makine öğrenimi algoritmalarının gücünden yararlanarak çözüm bulmaya çalışılmıştır. Biyolojik besi maddesi giderimi yapan gerçek bir atıksu arıtma tesisine ait kapsamlı altı yıllık günlük atıksu verilerini kullanarak, bu çalışma atıksu arıtma süreçlerinin optimizasyonuna katkıda bulunmayı hedeflemiştir. RMSE ve MAE metriklerini kullanarak gerçekleştirilen analiz, seçilen parametrelerin tahmin edilmesinde her bir algoritmanın belirgin güçlü ve zayıf yönlerini ortaya koymuştur.

Bu çalışmanın bulguları, atıksu arıtım süreçlerinin verimliliğini ve sürdürülebilirliğini artırmak için makine öğrenimi algoritmalarını kullanmanın önemini vurgulamaktadır. Modellerin tahmin performanslarındaki belirlenen farklılıklar, atıksu arıtımındaki her bir parametreyle ilişkili belirli gereksinimlerin dikkatlice düşünülmesi gerektiğini vurgulamaktadır. Bu bulguların sonuçları, makine öğrenmesi modellerinin optimizasyonu ve ayarlanmasına kadar uzanan etkileri, tahmin yeteneklerinin artırılmasını sağlayarak atıksu arıtma sistemlerinin daha hassas ve etkili bir şekilde yönetilmesine yardımcı olmaktadır. İlerleyen süreçte, atıksu arıtımında tahmin modellemesi alanında daha fazla araştırma ve geliştirme, makine öğrenimi

algoritmalarının yeteneklerini iletirmek ve sürdürülebilir çevre uygulamalarını teşvik etmek için gereklidir.

Çevresel açıdan, bu sonuçların etkilerinin önemli olacağı düşünülmektedir. Bu modellerin doğru tahmin performansları, atıksu arıtma tesislerinin çevresel etkisini önemli ölçüde azaltmaya katkıda bulunacaktır. Yetersiz şekilde arıtılmış atıksu deşarjını en aza indirerek, bu modeller sucul ekosistemlerin korunmasına ve sıkı çevre düzenlemelerine uyum sağlanmasına destek olacaktır. Ayrıca, KOİ, toplam azot ve toplam fosfor konsantrasyonlarını hassas bir şekilde tahmin etme yeteneği, daha verimli ve sürdürülebilir atıksu arıtımını sağlayarak, yüksek enerji gereksinimine yol açabilen aşırı arıtmadan kaçınmayı ve çevreye yönelik riskleri azaltmayı sağlayacaktır.

Çevresel endişeler arttıkça ve düzenlemeler sıkılaştıkça, gelişmiş makine öğrenimi tekniklerinin entegrasyonu, atıksu arıtımını devrimleştirmeye ve hem çevresel sorumluluk hem de operasyonel verimlilik sağlamaya potansiyel sunmaktadır.

Kalıntı ve tahmin verileri detaylı bir şekilde incelendiğinde değişen hata varyansı ve model yetersizliği (modelin önemli değişkenleri gözden kaırması veya verilerde belirli desenleri yakalayamaması) işaretleri görülebilmektedir. Her üç modelin tahminleri, hedef değişkenin belirli aralıklarında daha yüksek veya daha düşük doğruluk sağlayabilmekte, bu da veri setinin belirli özelliklerine bağlı olarak tahmin performansında değişikliğe yol açabilmektedir.

Ancak, çevresel modelleme ve atıksu yönetimi alanı dinamik bir alandır ve bu çalışmaların üzerine inşa edilebilecek model iyileştirmesi, veri artırma, hibrit modeller ve gerçek zamanlı izleme gibi gelecekteki çalışma ve araştırma alanları için fırsatlar bulunmaktadır.

Sonuç olarak, çevresel modelleme ve atıksu yönetimi alanı, yapay zeka ve makine öğrenmesi gibi uygulamaların entegrasyonu ile yapılacak gelecekteki çalışmalar, yenilikler ve disiplinler arası iş birlikleri için birçok fırsat sunmaktadır. Bu araştırmanın temelleri üzerine inşa edilerek gerçekleştirilecek gelecek çalışmalar makine öğrenmesi uygulamalarının getireceği katkılar ile daha sürdürülebilir ve bilinçli atıksu arıtma uygulamalarına ve azaltılmış çevresel etkiye katkıda bulunabilecektir.

KAYNAKLAR

- Akpor, O., (2014). Heavy metal pollutants in wastewater effluents: Sources, effects and remediation. *Adv. Biosci. Bioeng.* 2, 37. DOI: 10.11648/j.abb.20140204.11
- Andrade Cruz, I., Chuenchart, W., Long, F., Surendra, K.C., Renata Santos Andrade, L., Bilal, M., Liu, H. vd. (2022). Application of machine learning in anaerobic digestion: perspectives and challenges. *Bioresource Technology*, 345, 126433.
- Artan, N., Orhon, D. (2005). Mechanism and Design of Sequencing Batch Reactors for Nutrient Removal, *Scientific and Technical Report No.19*, IWA Publishing
- Asami H., Golabi M., Albaji M., (2021). Simulation of the biochemical and chemical oxygen demand and total suspended solids in wastewater treatment plants: Data-mining approach. *Journal of Cleaner Production*, 296, 2021.
- Bagherzadeh, F., Mehrani, M.J., Basirifard, M., Roostaei, J., (2021). Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance, *Journal of Water Process Engineering*, 41.
- Bishop, C. M. (2006). *Pattern recognition and machine learning (1st ed.)*. New York: Springer.
- Boser, B. E., Guyon, I. M., ve Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *In Proceedings of the Fifth Annual Workshop on Computational Learning theory*, 144-152.
- Bottou, L., Curtis, F.E., Nocedal, J. (2015). Optimization methods for large-scale machine learning. *SIAM Review*, 60(2), 223-311.
- Breiman, L., Friedman, J., Olshen, R. A., Stone, C.J, (1984). *Classification and regression trees*. CRC.
- Brin, S., Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117.

- Bunzel, K., Kattwinkel, M., Liess, M., (2013). Effects of organic pollutants from wastewater treatment plants on aquatic invertebrate communities. *Water Res.* 47, 597–606. <https://doi.org/10.1016/j.watres.2012.10.031>
- Capizzi, G., Masarotto, G., (2017). Phase I distribution-free analysis of multivariate data. *Technometrics*, 59, 484–495.
- Chislock, M.F., Doster, E., Zitomer, R.A., Wilson, A.E., (2013). Eutrophication: Causes, consequences, and controls in aquatic ecosystems. *Nat. Educ. Knowl.*, 4, 10.
- Cortes, C., Vapnik, V., (1995). Support-vector networks. *Machine Learning*, 20, 273–297.
- Feachem, R.G., Bradley, D.J., Garelick, H., Mara, D.D., (1983). Sanitation and disease: health aspects of excreta and wastewater management. *Sanit. Dis. Health Asp. Excreta Wastewater Manag.*
- Frey, B. J., Dueck, D. (2007). Clustering by passing messages between data points. *Science*, 315(5814), 972-976.
- Gandomi, A., ve Haider, M. (2015). Beyond the hype: big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Geman, S., Bienenstock, E., Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural computation*, 4(1), 1-58.
- Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep learning (vol. 1)*. Cambridge: MIT press.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y., M., Park, J. vd. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal Of Environmental Sciences*, 32, 90-101

- Güçlü, D., Dursun, Ş., (2010). Artificial neural network modelling of a large-scale wastewater treatment plant operation. *Bioprocess and Biosystems Engineering*, 33, 1051–1058.
- Hasan, M., N., Altaf, M., M., Khan, N., A., Khan, A., H., Khan, A., A., Ahmed, S. vd. (2021). Recent technologies for nutrient removal and recovery from wastewaters: a review. *Chemosphere*, 277.
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., ve Khan, S. U. (2015). The rise of “big data” on cloud computing: review and open research issues. *Information Systems*, 47, 98-115.
- Hastie, T., Tibshirani, R., Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction (2nd ed.)*. New York: Springer.
- Hatcher, W., G., Yu, W., (2018). A survey of deep learning: Platforms, applications and emerging research trends, *IEEE Access*, 6, 24411 – 24432.
- Henze, M. (1991). Capabilities of biological nitrogen removal processes from wastewater. *Water Science and Technology*, 23, 669–679.
- Henze, M., Ekama, G.A., von Loosdrecht, M.C.M. Brdjanovic, D., (2008). *Biological wastewater treatment: principles, modelling and design*. London: IWA Pub.
- Hernando, M.D., Mezcuca, M., Fernández-Alba, A.R., Barceló, D., (2006). Environmental risk assessment of pharmaceutical residues in wastewater effluents, surface waters and sediments. *Talanta, 1st Swift-WFD workshop on validation of Robustness of sensors and bioassays for Screening Pollutants*, 69, 334–342. DOI: 10.1016/j.talanta.2005.09.037
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N. vd. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6), 82-97.
- Hossain, S.M.Z., Sultana, N., Jassim, M.S., Coskuner, G., Hazin, L.M., Razzak, S.A. VD. (2022). Soft-computing modeling and multiresponse optimization for nutrient

- removal process from municipal wastewater using microalgae. *J. Water Process Eng.*, 45, 102490.
- Hvala N., Kocijan J., (2021). Input variable selection using machine learning and global sensitivity methods for the control of sludge bulking in a wastewater treatment plant. *Computers and Chemical Engineering*, 154.
- Jiang, W., Wang, K.B., Tsung, F., (2012). A variable selection-based multivariate EWMA chart for process monitoring and diagnosis. *J. Qual. Technol.*, 44, 209–230.
- Kashani, M.N., Shahhosseini, S., (2010). A methodology for modeling batch reactors using generalized dynamic neural networks. *Chemical Engineering Journal*, 159, 195–202.
- Kazor, K., Holloway, R.W., Cath, T.Y., Hering, A.S., (2016). Comparison of linear and nonlinear dimension reduction techniques for automated process monitoring of a decentralized wastewater treatment facility. *Stoch. Environ. Res. Risk Assess.*, 30, 1527–1544.
- Kober, J., Bagnell, J.A., Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11), 1238-1274.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*, 14, 1137-1145.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Lee, C.C., Lin, S. (Eds.), (2007). *Handbook of environmental engineering calculations*, 2nd ed., McGraw-Hill handbooks. New York: McGraw Hill,

- Liu H., Zhang Y., Zhang H., (2020). Prediction of effluent quality in papermaking wastewater treatment processes using dynamic kernel-based extreme learning machine. *Process Biochemistry*, 97, 72-79.
- Liu, Y. (2008). *Simultaneous Sludge Minimization and Nutrients Removal in Activated Sludge Process*, M.Sc. Thesis, The University of Utah, USA
- Ludzack, F.J., Ettinger, M.B., (1962). Controlling Operation to minimize activated sludge effluent nitrogen. *Journal of the Water Pollution Control Federation*, 34 (9), 920-931.
- Mackenzie L. Davis (2010). *Water And Wastewater Engineering: Design Principles And Practice*. New York: McGraw Hill.
- Matheri A.N., Ntuli F., Ngila Jane C., Seodigeng T., Zvinowanda C., (2021). Performance prediction of trace metals and cod in wastewater treatment using artificial neural network. *Computers and Chemical Engineering*, 149.
- Metcalf, W., Eddy, C., (2003). *Wastewater Engineering: Treatment and Reuse*. New York: McGraw Hill,
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J. vd. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533.
- Mohri, M., Rostamizadeh, A., Talwalkar, A. (2018). *Foundations of machine learning (2nd ed.)*. Cambridge, MA: The MIT Press.
- Montavon, G., Samek, W., ve Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73, 1-15.
- Murphy, K. P. (2012). *Machine learning: A probabilistic perspective (1st ed.)*. Cambridge, MA: The MIT Press.
- Nadiri A.A., Shokri S., Tsai F.T., Moghaddam A.A., (2018). Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model. *Journal of Cleaner Production*, 539-549, 180.

- Newhart, K., (2020). *Data-Driven Process Monitoring And Control In Municipal Wastewater Treatment*, PhD. Thesis, Department of Civil and Environmental Engineering, Colorado School of Mines 18 Eylül 2023 tarihinde <https://scholar.google.com/citations?user=IgFxIPYAAAAJ&hl=en> adresinden alındı.
- Nielsen, M. A. (2015). Neural networks and deep learning. Determination Press. *Journal of Software Engineering and Applications*, Vol.11 No.2,
- O'Day, D.K., (2004). Statistical Process Control Can Help Treatment Plants. *Water World* 20.
- Odom, G.J., Newhart, K.B., Cath, T.Y., Hering, A.S., (2018). Multistate multivariate statistical process control. *Appl. Stoch. Models Bus. Ind.*, 34, 880–892.
- Olsson, G., (2012). ICA and me – A subjective review. *Water Res.* 46, 1585–1624.
- Olsson, G., Aspegren, H., Nielsen, M.K., (1998). Operation and control of wastewater treatment - a Scandinavian perspective over 20 years. *Water Sci. Technol.* 37, 1–13.
- Orhon, D., Artan, N., (1997). *Modeling of Activated Sludge Systems*. Florida: CRC Press.
- Poole, D., Mackworth, A., Goebel, R., (1998). *Computational Intelligence: A Logical Approach*, Oxford: Oxford University Press
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. New York: Morgan Kaufmann.
- Rittmann, B. E., McCarty, P. L. (2001). *Environmental Biotechnology: Principles and Applications*, New York: McGraw Hill
- Rumelhart, D. E., Hinton, G. E., ve Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- Sheikh, M., Harami, H., R., Rezakazemi, M., Cortina, J., L., Aminabhavi, T., M., Valderrama, C., (2023). Towards a sustainable transformation of municipal

- wastewater treatment plants into biofactories using advanced $\text{NH}_3\text{-N}$ recovery technologies: a review. *Science of the Total Environment*, 904.
- Shrestha, A., Mahmood, A., (2019). Review of Deep Learning Algorithms and Architectures. *IEEE Access*, 7, 53040—53065.
- Sutton, R. S., Barto, A. G. (2018). *Reinforcement learning: An introduction (2nd ed.)*. Cambridge, MA: The MIT Press.
- Şahinaslan, Ö., Dalyan, H (2021). Sosyal Medya Uygulamalarında Yapay Zekânın Yeri. *Yapay Zeka Uygulamalarında Güncel Konular ve Araştırmalar* içinde (ss. 24-40). Konya, Çizgi Kitabevi Yayınları
- Şahinaslan, E. (2023). İş Süreci Optimizasyonu: Yöntem, Teknoloji, Riskler Ve Fırsatlar. *Akademik İzdüşüm*, 8(2): 570-604. ISSN: 2547-9725
- Şahinaslan, E., Günerkan, M., & Şahinaslan, Ö. (2023). Makine Öğrenmesinde Kategorik Veri Kodlama Tekniğinin Kullanımına Alternatif Bir Çözüm Yöntemi. *Journal of Intelligent Systems: Theory and Applications*, 6(1), 1-11. <https://doi.org/10.38016/jista.1140499>
- Vapnik, V. (1995). *The nature of statistical learning theory (1st ed.)*. New York: Springer.
- Wang, K.B., Jiang, W., (2009). High-dimensional process monitoring and fault isolation via variable selection. *J. Qual. Technol.*, 41, 247–258.
- Wiesmann, U., Choi, I. S., Dombrowski, E., (2007). *Fundamentals of Biological Wastewater Treatment*, Wiley-VCH Verlag GmbH ve Co. KGaA, Weinheim, Almanya
- Yang C., Zhang Y., Huang M., Liu H., (2021). Adaptive dynamic prediction of effluent quality in wastewater treatment processes using partial least squares embedded with relevance vector machine. *Journal of Cleaner Production*, 314.

Zhao, Y., (2009). *Detection and Analysis of Sludge Bulking Events Using Data Mining and Machine Learning Approach*, MSc. Thesis, Marquette University, 2009. 18 Eylül 2023 tarihinde https://epublications.marquette.edu/theses_open/143/ adresinden alındı.



ÖZGEÇMİŞ

Cem CANTEKİN

Eğitim

Derece	Yıl	Üniversite, Enstitü, Anabilim/Anasanat Dalı
Yüksek Lisans	2014	İ.T.Ü, Fen Bilimleri Enstitüsü, Çevre Biyoteknolojisi
Lisans	2010	Anadolu Üniversitesi, Müh. Fak., Çevre Mühendisliği

İş/İstihdam (Varsa)

Yıl	Görev
2020-....	Ar. Gör. / Maltepe Üniversitesi Bilgisayar Müh. Böl.
2017-2019	Technical SEO Analyst ve Chatbot Developer / Ingoduo Digital Marketing
2016-2017	Teknik Danışman / TÜBİTAK MAM Melen Baraj Havzası Projesi
2016	Danışmanlık / Ironman Consulting
2011-2012	OHSAS Mühendisi / DETAM Şirketler Grubu

Yayınlar ve Diğer Bilimsel/Sanatsal Faaliyetler

Makale

Cubukcu Cerasi Ceren, Cantekin Cem (2021). Using a combined Fuzzy-AHP and TOPSIS Decision Model for Selecting the Best Firewall Alternative. Journal of Fuzzy Extension and Applications

Cantekin Cem, Taybuga Emine Simge, Yağcı Nevin, Orhon Derin (2019). Potential For Simultaneous Nitrogen Removal And Sludge Reduction Of The Oxic-Settling-Anaerobic Process Operated As A Dual Fed Sequencing Batch Reactor. JOURNAL OF ENVIRONMENTAL MANAGEMENT

Sunum

Cantekin Cem, Şahinaslan Önder, Yağcı Nevin (2023). Prediction Performances of Decision Tree Regression and Support Vector Machines (SVM) Algorithms on Effluent Chemical Oxygen Demand (COD) Concentrations of a Wastewater Treatment Plant. International Conference of Mathematical Sciences (ICMS 2023)

Cubukcu Ceren, Cantekin Cem (2021). Using a Fuzzy-AHP Decision Model for Selecting the Best Firewall Alternative. Intelligent and Fuzzy Systems (INFUS 2021)

Cantekin Cem, Karlikanovaite Agne, Yagci Nevin (2017). Simultaneous Nutrient Removal and Solids Reduction in OSA Process. International Conference on Civil and Environmental Engineering (ICOCEE)

Kitap

Contemporary Approaches in Social Science Research, Chapter Name: DIGITAL TRANSFORMATION IN EDUCATION: MALTEPE UNIVERSITY EXAMPLE (2022)., CANTEKİN CEM, ÇUBUKÇU ÇERASİ CEREN, SRA Academic Publishing Press, ISBN:978-625-7148-42-9

