

T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**ATIKSU ARITMA TESİSLERİNDE BİYOLOJİK OKSİJEN  
İHTİYACININ (BOİ<sub>5</sub>) KABA KÜMELEME VE MAKİNE  
ÖĞRENMESİ HİBRİT YAKLAŞIMI İLE TAHMİNİ**

**YÜKSEK LİSANS TEZİ**

**Muhammed Alperen ŞERİFOĞLU**

**Enstitü Anabilim Dalı : ENDÜSTRİ MÜHENDİSLİĞİ**  
**Enstitü Bilim Dalı : MÜHENDİSLİK YÖNETİMİ**  
**Tez Danışmanı : Prof. Dr. Orhan TORKUL**

**Haziran 2021**

**T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**ATIKSU ARITMA TESİSLERİNDE BİYOLOJİK OKSİJEN  
İHTİYACININ (BOI<sub>5</sub>) KABA KÜMELEME VE MAKİNE  
ÖĞRENMESİ HİBRİT YAKLAŞIMI İLE TAHMİNİ**

**YÜKSEK LİSANS TEZİ**

**Muhammed Alperen ŞERİFOĞLU**

**Enstitü Anabilim Dalı : ENDÜSTRİ MÜHENDİSLİĞİ**

**Enstitü Bilim Dalı : MÜHENDİSLİK YÖNETİMİ**

**Bu tez 16/06.2021 tarihinde aşağıdaki jüri tarafından oybirliği ile kabul edilmiştir.**

**Jüri Başkanı**

**Üye**

**Üye**

## **BEYAN**

Tez içindeki tüm verilerin akademik kurallar çerçevesinde tarafımdan elde edildiğini, görsel ve yazılı tüm bilgi ve sonuçların akademik ve etik kurallara uygun şekilde sunulduğunu, kullanılan verilerde herhangi bir tahrifat yapılmadığını, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezde yer alan verilerin bu üniversite veya başka bir üniversitede herhangi bir tez çalışmasında kullanılmadığını beyan ederim.

M. Alperen ŞERİFOĞLU

16.06.2021

## TEŐEKKÜR

Yüksek lisans eğitimim boyunca değerli bilgi ve deneyimlerinden yararlandığım, her konuda bilgi ve desteğini almaktan çekinmediğim, araştırmanın planlanmasından yazılmasına kadar tüm aşamalarında yardımlarını esirgemeyen, teşvik eden, aynı titizlikte hem eğitim-öğretim hem de sosyal hayatımda beni yönlendiren değerli danışman hocam Sayın Prof. Dr. Orhan TORKUL'a teşekkür eder, saygılarımı sunarım.

Bilgi ve deneyimlerinden yararlandığım, bu çalışma kapsamında desteklerini esirgemeyen Araştırma Görevlisi Sayın Merve ŐİŐCI'ye, manevi desteklerini benden esirgemeyen eşime, aileme, Arıtma Tesisleri ve Laboratuvar olanakları konusunda anlayış ve yardımlarını esirgemeyen, Sakarya Su ve Kanalizasyon İderesi Genel Müdürlüğü Arıtma Tesisleri Dairesi Başkanı Sayın Yasin ÇAKIR'a, Atıksu Arıtmalar Şube Müdürü Sayın Murat İKİNCİ'ye teşekkür ederim.

# İÇİNDEKİLER

TEŞEKKÜR .....	i
İÇİNDEKİLER .....	ii
SİMGELER VE KISALTMALAR LİSTESİ .....	v
ŞEKİLLER LİSTESİ .....	viii
TABLolar LİSTESİ .....	ix
ÖZET .....	x
SUMMARY .....	xi

## BÖLÜM 1.

GİRİŞ .....	1
-------------	---

## BÖLÜM 2.

KAYNAK ARAŞTIRMASI .....	5
2.1. Atıksu Arıtma İşlemlerinde Sıklıkla Kullanılan Ölçüm Parametreleri ...	5
2.2. Atıksu Arıtma Tesislerinde Uygulanan Arıtma Yöntemleri ve Kimyasal Terimler .....	6
2.2.1. Biyokimyasal Oksijen İhtiyacı (BOİ) arıtma yöntemi .....	6
2.2.2. Kimyasal Oksijen İhtiyacı (KOİ) arıtma yöntemi .....	7
2.2.3. Toplam Askıda Katı Madde (AKM) .....	8
2.2.4. Yağ ve gres arıtma yöntemleri .....	9
2.3. Atıksu Arıtma Tesislerinde PLC ve SCADA Sistemleri ile İlgili Yapılan Çalışmalar .....	9

2.4. Atıksu Arıtma Tesislerinde Yapay Zeka ve Makine Öğrenmesi Teknolojilerinin Kullanımı ile İlgili Gerçekleştirilen Çalışmalar .....	10
2.4.1. Atıksu arıtma tesislerinde performans tahmini ve iyileştirilmesi için kullanılan yapay zeka teknikleri .....	14
2.4.2. Atıksu arıtma tesislerinde parametre tahmini için kullanılan yapay zeka teknikleri .....	18

### BÖLÜM 3.

ATIKSU ARITMA TESİSLERİNDE KULLANILAN ÖLÇÜM VE KONTROLÜ SAĞLAYAN OTOMASYON SİSTEMLERİ .....	24
3.1. Atıksu Arıtma Tesislerinde Kontrol ve Planlamanın Önemi .....	24
3.2. Denetleyici Gözetim ve Veri Toplama için SCADA Otomasyon Sistemleri .....	25
3.3. SCADA Otomasyon Sisteminde Kullanılan PLC .....	27
3.4. SCADA Sisteminin Temel Yapısı .....	29
3.5. SCADA Sisteminin Katmanları .....	31

### BÖLÜM 4.

İLERİ SEVİYE KONTROL VE OTOMASYON SİSTEMİ OLARAK BİR SCADA SİSTEM TASARIMI, PLANLAMA VE UYGULAMASI, ATIKSU ARITMA TESİSİ MODERNİZASYONU ÖRNEĞİ .....	33
4.1. Sakarya'da Faaliyet Gösteren Atıksu Arıtma Tesisleri .....	33
4.2. Atıksu Arıtma Tesisi Modernizasyonu İçin Kullanılan Materyaller .....	34
4.2.1. SCADA sistemi güvenlik kontrol paneli .....	35
4.2.2. Modernizasyonun tanımlanması ve SCADA sisteminin devreye alınması .....	36
4.3. SCADA Otomasyon Sisteminin Tesise Kazandırdığı Avantajlar .....	37
4.3.1. Ekipman açısından SCADA sisteminin faydaları .....	37
4.3.2. Proses açısından SCADA sisteminin faydaları .....	38

4.4. Modernizasyon Amacıyla Yapılan Yatırım Sonrası Maliyet Kıyaslama 41

BÖLÜM 5.

ATIKSU ARITMA TESİSLERİNDE BİYOLOJİK OKSİJEN İHTİYACININ  
(BOİ<sub>5</sub>) TAHMİNİ İÇİN KULLANILAN METOTLAR ..... 45

5.1. Metotlar/ Yöntem ..... 45

5.1.1. Bulanık Hızlı Azaltım algoritması (BHAA) ..... 45

5.1.2. Bulanık-Kaba Örnek Seçimi algoritması (BKÖSA) ..... 48

5.1.3. SmoteR yöntemi ..... 51

5.1.4. Makine öğrenmesi regresyon algoritmaları ..... 55

5.1.4.1. Doğrusal Regresyon (DR) ..... 58

5.1.4.2. Yapay Sinir Ağları (YSA) ..... 59

5.1.4.3. Güçlendirilmiş Karar Ağacı Regresyonu (GKAR) ..... 63

5.1.4.4. Bayesyen Doğrusal Regresyon (BDR) ..... 64

5.1.4.5. Karar Ormanı Regresyonu (KOR) ..... 64

5.2. Metodoloji ve Uygulama ..... 65

5.2.1. Çalışma alanı ve verilerin elde edilmesi ..... 67

5.2.2. Bulanık Hızlı Azaltım algoritması (BHAA) ile nitelik seçimi .... 68

5.2.3. Bulanık-Kaba Örnek Seçimi algoritmasının (BKÖSA)  
uygulanması ..... 70

5.2.4. SmoteR yöntemi ile veri setinin dengelenmesi ..... 71

5.2.5. Makine öğrenmesi algoritmaları ile modellerin oluşturulması .... 72

5.3. Modellerin Performanslarının Karşılaştırılması ..... 73

BÖLÜM 6.

SONUÇ ..... 78

KAYNAKLAR ..... 82

ÖZGEÇMİŞ ..... 92

## SİMGELER VE KISALTMALAR LİSTESİ

AAO	: Anaerobik-Anoksik-Oksik Sistem
AAT	: Atıksu Arıtma Tesisi
ABM	: Ajan Tabanlı Modelleme
AKM	: Askıda Katı Madde
ANFIS	: Uyarlanabilir Nöro-Bulanık Çıkarım Sistemi
ASM	: Aktif Çamur Modeli
BOİ/BOİ <sub>5</sub>	: Biyolojik Oksijen İhtiyacı
BT	: Bilgi Teknolojileri
BDR	: Bayesyen Doğrusal Regresyon
BHAA	: Bulanık Hızlı Azaltım Algoritması
Cd	: Kadminyum
CKM	: Çökebilir Katı Madde
CNN	: Evrişimli Sinir Ağı Modeli
CO <sub>2</sub>	: Karbondioksit
CPU	: Bilgisayar İşlemcisi
Cr	: Krom
Cu	: Bakır
ÇDR	: Klasik Çoklu Doğrusal Regresyon
DR	: Doğrusal Regresyon
Fe	: Demir
FRIS	: Bulanık Kaba Örnek Seçimi Algoritması
GKAR	: Güçlendirilmiş Karar Ağacı Regresyonu
Hg	: Civa



HRT	: Hidrolik Tutma Süresi
HSA	: Hücresel Sinir Ağları
IRR	: Dahili Geri Dönüşüm Oranı
İBGYYSA	: Beslemeli Geri Yayımlı Yapay Sinir Ağları
İBSA	: İleri Beslemeli Sinir Ağı
KHT	: Kareler Hata Toplamı
KNN	: K-en Yakın Komşu
KOH	: Karekök Ortalama Hata
KOİ	: Kimyasal Oksijen İhtiyacı
KOR	: Karar Ormanı Regresyonu
MAE	: Ortalama Mutlak Hata
MAPE	: Ortalama Mutlak Yüzde Hatası
MARS	: Değişkenli Uyarlanabilir Regresyon Uzanımları
MLR	: Çoklu Lineer Regresyon
MPI	: Çok Noktalı Ara Yüz
MTU	: Ana Terminal Birim
NH4	: Amonyum
NO3-	: Nitrat Nitrojen
P	: Fosfor
Pb	: Kurşun
PLC	: Programlanabilir Lojistik Kontrol
Q	: Debi
QRA	: Hızlı Azaltım Algoritması
RAE	: Göreceli Mutlak Hata
RMSE	: Ortalama Karesel Hata
RSE	: Göreceli Kare Hata
RTC	: Gerçek Zamanlı Kontrol Sistemi
RTU	: Uzak Terminal Birim

SCADA	: Merkezi Denetim Ve Veri Toplama Sistemi
SO4	: Sülfat
SRM	: Risk Minimizasyonu
SVM	: Destek Vektör Makinesi
T	: Sıcaklık
TKN	: Toplam Kjeldahl Azotu
TN	: Total Nitrojen
TOC	: Giriş Suyundaki Toplam Organik Karbon
TP	: Toplam Fosfor
YSA	: Yapay Sinir Ağı
YSA-GA	: Yapay Sinir Ağı- Genetik Algoritma
Zn	: Çinko

## ŞEKİLLER LİSTESİ

Şekil 3.1. PLC'nin sistem mimarisi .....	29
Şekil 4.1. Atıksu Arıtma Tesisi Örneği .....	34
Şekil 4.2. Atıksu Arıtma Tesisi SCADA Kontrol Ekranı Görseli .....	35
Şekil 4.3. Atıksu Arıtma Tesisi Network Topolojisi .....	35
Şekil 4.4. Atıksu Arıtma Tesisi Network Topolojisi .....	36
Şekil 4.5. Arıtılan Atıksu Miktarları Karşılaştırılması Debi Artışı .....	38
Şekil 4.6. Birim Enerji Miktarları Karşılaştırılması m3 Başına Azalma %0,7 .....	39
Şekil 4.7. Personel Fazla Mesai Karşılaştırılması %100 Azalma .....	40
Şekil 4.8. Personel Maliyetleri Karşılaştırılması % 45,6 Azalma .....	40
Şekil 4.9. Birim Doğalgaz Miktarları Karşılaştırılması .....	41
Şekil 4.10. Net Büküm Değer Hesap Grafiği .....	44
Şekil 5.1. Kaba Kümelerin Genel İşletim Mantığı .....	48
Şekil 5.2. Bulanık-Kaba Küme Seçim Algoritması Çalışma Mantığı .....	49
Şekil 5.3. SmoteR Algoritmasının Grafikselsel Temsili .....	52
Şekil 5.4. Doğrusal Regresyon Modeli .....	58
Şekil 5.5. Üç Katmanlı İleri Beslemeli YSA .....	60
Şekil 5.6. Yapay sinir ağı nöronu çalışma fonksiyonu .....	61
Şekil 5.7. Çok Katmanlı Yapay Sinir Ağı Görünümü .....	62
Şekil 5.8. Uygulama Akış Şeması .....	66
Şekil 5.9. Karaman Atıksu Arıtma Tesisi .....	67

## TABLolar LİSTESİ

Tablo 4.1. Sakarya’da Bulunan Atıksu Arıtma Tesisleri .....	33
Tablo 4.2. Maliyet Kıyaslama .....	41
Tablo 4.3. Bileşik Faiz Tablosu .....	43
Tablo 4.4. Net Büküm Değer Hesap Tablosu .....	43
Tablo 5.1. Karaman Atıksu Arıtma Tesisi analiz veri setine ait tanımlayıcı istatistikler .....	68
Tablo 5.2. BOİ <sub>5</sub> 'teki Modellerin Veriseti1 Üzerindeki Test Performansları .....	75
Tablo 5.3. BOİ <sub>5</sub> 'teki Modellerin Veriseti2 Üzerindeki Test Performansları .....	75
Tablo 5.4. BOİ <sub>5</sub> 'teki Modellerin Veriseti3 Üzerindeki Test Performansları .....	76
Tablo 5.5. BOİ <sub>5</sub> 'teki Modellerin Veriseti4 Üzerindeki Test Performansları .....	77

## ÖZET

Anahtar Kelimeler: Atıksu Arıtımı, SmoteR, BOİ, Yapay Zeka, Makine Öğrenmesi

Atıksu arıtma tesisi işletmesi sırasında suyun kalitesini gösteren parametreleri hızlı bir şekilde takip edebilmek ve müdahale edebilmek önem arz etmektedir. Atıksu arıtma tesislerine gelen atıksuyun arıtılmasında bakteri bazlı biyolojik yöntemler kullanılmaktadır. Gelen suyun kirlilik yükleri ve oksijen dengesi tesisin içerisindeki yaşamsal döngüyü etkilemektedir. Tesise aniden çok yüksek kirlilik oranına sahip su girdiği zaman operatör hiçbir müdahalede bulunmaz veya yanlış müdahalede bulunur ise bakteriler ölmeye başlar. Dolayısıyla biyolojik denge bozulacağı için arıtma işlemi gerçekleştirilemez. Su kalitesi yönetimi ve planlamasında önemli bir değişken olan Biyolojik Oksijen İhtiyacının (BOİ5) laboratuvar ortamında ölçülmesi diğer parametrelere göre daha uzun sürelerde (5 gün) gerçekleşmektedir. Sakarya da bulunan bir atıksu arıtma tesisinde yapılan bu çalışmada, online ölçüm cihazları ile ölçülebilen parametrelerden alınan verilerin otomasyon ve kontrol sistemlerine aktarılması ve elde edilen bu verilere yapay zeka yöntemleri uygulanarak Biyolojik Oksijen İhtiyacının (BOİ5) tahmin edilmesi amaçlanmıştır. BOİ5'nin tahmin edilmesi amacıyla Bulanık Kaba Kümeleme, SmoteR ve Makine Öğrenmesi Regresyon Algoritmalarını (Güçlendirilmiş Karar Ağacı Regresyonu, Karar Ormanı Regresyonu Yapay Sinir Ağları, Bayesyen Doğrusal Regresyon ve Doğrusal Regresyon) içeren hibrit bir model geliştirilmiştir. Tahmin için kullanılan makine öğrenmesi algoritmalarının performansının artırılması amacıyla Bulanık Kaba Kümeleme tabanlı Nitelik Seçimi ve SmoteR yöntemleri uygulanırken, gereksiz veya gürültülü örneklerin elenmesi için bulanık kaba kümeleme tabanlı Örnek Seçimi Algoritması uygulanmıştır. Elde edilen veri setleri üzerinde makine öğrenmesi regresyon algoritmaları eğitilerek performansları karşılaştırılmıştır. Karşılaştırmalı sonuçlar, Bulanık Kaba Kümeleme tabanlı Nitelik Seçimi, Örnek Seçimi ve özellikle SmoteR yöntemlerinin Makine Öğrenmesi Algoritmalarının performanslarını arttırdığını göstermektedir. Çalışmanın ilk aşamasında (Bulanık Kaba Kümeleme tabanlı Nitelik Seçimi, bulanık kaba kümeleme tabanlı Örnek Seçimi ve SmoteR yöntemlerinin uygulanmadan) en iyi performansı %91,56 R2 değeri ile Bayesyen Doğrusal Regresyon ve %10,35 MAPE değeri ile Doğrusal Regresyon algoritması göstermektedir. Bulanık Kaba Kümeleme tabanlı Nitelik Seçimi, bulanık kaba kümeleme tabanlı Örnek Seçimi ve SmoteR yöntemlerinin uygulanmasıyla yapılan çalışmanın sonucunda makine öğrenmesi algoritmaları arasında en iyi performansı, %97,18 R2 değeri ve %6,07 MAPE değeri ile Güçlendirilmiş Karar Ağacı Regresyon Algoritması göstermiştir.

# **PREDICTION OF BIOCHEMICAL OXYGEN DEMAND (BOD<sub>5</sub>) IN WASTEWATER TREATMENT PLANT WITH ROUGH SET AND MACHINE LEARNING HYBRID APPROACH**

## **SUMMARY**

Keywords: Waste water treatment, SmoteR, BOI, Artificial Intelligence, Machine Learning

During the operation of wastewater treatment in plants, it is important to quickly monitor and intervene the quality of water. Some bacterial-based biological methods are used for the wastewater which is coming to wastewater treatment plants. The pollution and oxygen balance of the incoming water affects the life cycle within the facility. In sudden rises in pollution parameters values at the entrance of the facility, bacteria that provide biodegradation begin to die in the cases where the operator does not intervene or improperly intervenes. Therefore, treatment cannot be carried out as the biological balance will be disrupted. When compared to other parameters, the measurement of Biological Oxygen Demand (BOD), which is an important variable in water quality management and planning, in the laboratory environment takes longer times (5 days). In this study, which was carried out in a wastewater treatment plant in Sakarya, it was aimed to transfer the data obtained from the parameters that can be measured with online measurement devices to the automation and control systems and to estimate the Biological Oxygen Demand (BOD<sub>5</sub>) by applying artificial intelligence methods to these data. In order to predict BOD<sub>5</sub> a hybrid model which includes Fuzzy Rough Set, SmoteR and machine learning Regression algorithms (Boosted Decision Tree Regression, Bayesian Linear Regression, Decision Forest Regression, Neural Networks, Linear Regression) has been developed. While applying Fuzzy Rough Set based feature selection and SmoteR methods in order to increase the performance of machine learning algorithms used for prediction, Fuzzy Rough Set instance selection algorithm has been applied to eliminate unnecessary instances. Machine learning regression algorithms were trained on the obtained data sets and their performances were compared. The comparative results show that Fuzzy Rough Set based feature selection, instance selection and especially SmoteR methods clearly increase the performance of machine learning algorithms. In the first stage of the study (without applying Fuzzy Rough Set based feature selection, Fuzzy Rough Set instance selection, SmoteR methods), Bayesian Linear Regression algorithm with a R<sup>2</sup> value of 91.56% and Linear Regression algorithm with a MAPE value of 10.35% show the best performances. As a result of the study by applying Fuzzy Rough Set based feature selection, Fuzzy Rough Set instance selection, SmoteR methods, the Boosted Decision Tree Regression Algorithm showed the best performance among the machine learning algorithms, with a R<sup>2</sup> value of 97.18% and a MAPE value of 6.07%.

## **BÖLÜM 1. GİRİŞ**

Günümüzde gelinen nokta göz önünde bulundurulduğunda tüketimin hızla artması ve çevre kirliliğinin sürekli olarak devam etmesi su kaynaklarında meydana gelen kirliliğe büyük ölçüde etki etmektedir. Farklı nedenlerle oluşan atıksuların deşarj edildiği ortamlar da yer alan su kaynaklarının, atıksuların oluşturmuş olduğu olumsuz etmenlerden etkilenmemesi amacı ile atıksulara belirli işlemler uygulanmaktadır. Bu işlemler fiziksel, kimyasal, biyolojik ve ileri biyolojik arıtma işlemleridir ve bu işlemlerin uygulandığı sistemlere atıksu arıtma sistemleri denir. Dünyamızdaki hızlı nüfus artışı ve buna paralel hızlı teknolojik gelişmeler ekolojik sistemi olumsuz yönde etkilemektedir. Ekolojik sistemin en önemli parçası olan su kaynakları hızla kirlenmekte, oluşan atıksu miktarları nüfus artışı, plansız kentleşme ve teknolojik gelişmelere paralel olarak artmaktadır.

Su kaynaklarının korunması amacı ile 1983 yılında yürürlüğe giren 2872 sayılı Çevre Kanunu doğrultusunda; 2004 yılında “Su Kirliliği Kontrol Yönetmeliği” ve 2006 yılında “Kentsel Atıksu Arıtma Yönetmeliği” yayımlanmıştır. Yürürlüğe giren yönetmeliklerin amacı, ülkemizde yer altında ve yer üstünde yer alan su kaynaklarının korunarak ilgili su kaynaklarından en etkin şekilde yararlanmanın sağlanmasıdır. Bununla birlikte yürürlüğe giren yönetmelikler kentsel ya da endüstriyel atıksuların toplanarak arıtılmasından sonra, atıksu deşarjının gözlemlenerek raporlanması ve denetlenmesi ile ilgili teknik/idari gerekçeleri kapsamaktadır. Günümüzde, atıksuların deşarj edildiği ortamların çevre ve biyolojik çeşitliliği olumsuz olarak etkilediği bilinmektedir. Atıksu arıtma tesislerinin verdiği olumsuz etkileri en aza indirmek amacıyla alınan klasik önlemler, kapasiteleri giderek artan atıksu arıtma tesislerinin işletme maliyetlerini arttırmakla birlikte, işletme zorluklarını da beraberinde getirmektedir.

Atıksu arıtma tesislerinin uygun şekilde işletilmesi ve kontrolü, çevre sorunları ile ilgili artan endişeler nedeniyle oldukça ilgi görmektedir. Atıksu arıtma tesisinin yanlış işletilmesi ciddi çevresel ve halk sağlığı sorunlarına yol açabilir, çünkü alıcı su kütlesine giden atıklar insanlarda çeşitli hastalıklara neden olabilir veya bunları yayabilir. Atıksu arıtımı, yaşamın her evresinde kullanılan suların yeniden doğaya kazandırılabilmesi için laboratuvar parametreleri dikkatlice kullanılarak yapılan mühendislik çalışmaları ile arıtma verimliliğini artırılacak sistemler ile gerçekleştirilmektedir. Atıksu arıtma, çeşitli kimyasal, fiziksel ve mikrobiyolojik faktörlerden etkilenen karmaşık bir süreçtir. Ayrıca, stokastik tedirginlikler ve etkili değişkenlik, operatörlerin sistem üzerinde uygun operasyonel kontroller gerçekleştirilmesini zorlaştırmaktadır. Klasik yöntem ile, atıksu arıtma tesislerinin işletilmesi, yarı otomatik bir pano ile lokal kontrolünün yapılması ve tüm ekipmanların bu panodan yönetilemediğinden saha kontrollerinin zorunlu hale gelmesi sebebiyle tesiste 7/24 personel bulundurmaya gerekli kılmaktadır. Bu yöntemde, işletmelerde sürekli olarak personelin göz ile ve analiz yaparak verimli bir işletme gerçekleştirilememekle birlikte yetersiz personel ve fazla mesai problemleri ile karşılaşmaktadır.

Günümüz konuşma dillerinde neredeyse her gün sarf edilen teknoloji terimi, içinde bulunduğumuz modern çağın bilimsel ve teknik ileri atılımlarını kapsamakta, modern insanın fazla iş yükünden kurtulmasını, yaşam kalitesini ve konfor düzeylerini arttırmaya yönelik gelişmeleri insanlık hizmetine sunmaktadır. Bu modern çağ, yeni bir sanayi devrimi kabul gören, Endüstri 4.0 hamlesinin konuşulduğu, yapay zekanın ön plana çıktığı, öğrenen makineler, endüstriyel otomasyon sistemleri ve uzman sistemlerin sunduğu avantajlardan uzakta kalmamayı gerekli kılmakta, problemlere yönelik çözümler üreterek kullanabilme fırsatlarını sunmaktadır. Yapay Zeka ile çeşitli disiplinlerdeki zor problemlerin üstesinden gelmek ve gerçek dünyadaki uygulamalarda karmaşık problemleri çözmek için çok sayıda güçlü ve pratik çözüm yolları üretilir. Günümüzde kullanım kolaylığı, bilgiye anında erişmeye olanak sağlayan yüksek hızda çalışması ve bir insan beyni gibi çalışarak problemleri anlayabilmesi ve çözüm üretebilmesi amacıyla bu teknoloji her alanda kullanılmaktadır. Atıksu arıtma tesisleri, giderek daha sıkı analiz değerleri



kısıtlamalarıyla, enerji verimliliği ve kaynak geri dönüşüm yönetmeliklerinden sorumludur. Bu sebep ile araştırmacıların çoğu bu sorunları çözebilmek amacıyla Yapay Zekâ teknolojilerini atıksu arıtma tesislerine de uygulamaya başladılar. Yapay Zekâ teknolojileri, arıtılan Atıksuyun kalitesinin iyileştirilmesi, nehir suyu kalitesi modellemesi, su kaynağı geri dönüşümü, makine arızası teşhisi gibi pratik sorunları çözme becerisi nedeniyle çeşitli mühendislik disiplinlerinde kullanılmıştır.

Bu tez çalışması kapsamında, atıksu arıtma tesisinin işletim maliyetinin düşürülmesi ve ekipmanların daha verimli bir şekilde personelden bağımsız işletilmesi için ekipmanların birbirleri arasında haberleşerek doğru zamanlarda doğru manevralar yapmak üzere, programlanabilir bir endüstriyel otomasyon sistemi olan SCADA sistemi tasarımı uygulanması incelenmiştir. SCADA sisteminden alınan gerçek veriler ile önsezi yeteneği kazandırılmış otomasyon sistemlerinin tesis kazanımlarına katkısı, yeniden programlanabilir PLC'ler açısından SCADA sistemlerinin alarm üretimleri ve uygulanan Makine Öğrenmesi Teknolojileri ile tesis giriş parametresinin önceden tahmin edilmesi amaç edinilmiştir. Laboratuvar parametreleri arasında arıtma tesisi işletimi için büyük öneme sahip olan Biyolojik Oksijen İhtiyacı (BOİ) miktarının analizi ve sonuçlarının elde edilmesi beş (5) gün gibi uzun bir süre almaktadır. Biyolojik oksijen ihtiyacının geç analizi tesis işletimi sırasında yapılacak manevraları da geciktirmekte ve verimliliği olumsuz yönde etkilemektedir. Dolayısıyla bu sorunu ortadan kaldırmak amacıyla BOİ<sub>5</sub> parametresinin, online olarak ölçülebilen diğer parametre sonuçlarından faydalanılarak Yapay Zeka yöntemleri ile daha hızlı bir şekilde tahmin edilmesi amaçlanmıştır. Biyolojik oksijen ihtiyacı tahmini ile ilgili birçok çalışma bulunmasına rağmen bu çalışma kapsamında uyguladığımız Bulanık Kaba Kümeleme ve SmoteR yöntemlerini uygulayan çalışmalara rastlanılmamıştır.

Bu çalışmada; kimyasal oksijen ihtiyacı (KOİ), toplam askıda katı madde (AKM), pH, amonyum azotu (NH<sub>4</sub>-N), nitrat azotu (NO<sub>3</sub>-N), nitrit azotu (NO<sub>2</sub>), toplam azot (TN), toplam fosfor (TP) ve biyolojik oksijen ihtiyacı (BOİ<sub>5</sub>) parametleri kullanılarak eğitim ve test yapılmış ardından girdi parametreleri kullanılarak BOİ<sub>5</sub> parametresi tahmini yapılmıştır. Tahmin işlemi için izlediğimiz yöntemler şu şekilde organize edilmiştir; Verilerin hazırlanması aşamasında eksik veriler temizlenerek, normalize edilmesi

sağlanmıştır. Ardından Bulanık Hızlı Azaltım (Fuzzy QR) ile Nitelik Seçimi algoritmasının uygulanması, Bulanık-Kaba Kaba Önek Seçimi (FRIS) Algoritmasının Uygulanması, Dengesiz Verilerin ortadan kaldırılması için SmoteR yönteminin uygulanması ve Makine Öğrenmesi Modellemesi ile algoritmalar eğitilerek test edilmiş ve performansları karşılaştırılmıştır. Çalışmalar sonucunda Bulanık Kaba Kümeleme ve SmoteR yöntemlerinin makine öğrenmesi regresyon algoritmaları performanslarını önemli ölçüde artırdığı görülmüştür.

## **BÖLÜM 2. KAYNAK ARAŞTIRMASI**

### **2.1. Atıksu Arıtma İşlemlerinde Sıklıkla Kullanılan Ölçüm Parametreleri**

Bir atıksu arıtma tesisi işletmesinde aktif olarak uygulanması gerekli olan çeşitli parametreler bulunmaktadır. Atıksularda sıklıkla karşılaşılan organik maddeler; karbonhidratlar, yağlar, petrol artıkları ve üre olarak sıralanabilir. Bu organik maddelerden başkaca deterjanlar, zirai atıklar, endüstriyel atıklar gibi sentetik katkı maddelerini de barındırmaktadırlar. Orta kirliliğe sahip bir Atıksu, %75 oranında filtre edilebilir askıda katı maddeleri bünyesinde barındırmakla birlikte, bu askıda katı maddelerden yaklaşık %40'ını organik bileşenler oluşturmaktadır (Öztürk ve Timur, 2017). Atıksuların genel karakteristik yapısı ve incelemeleri laboratuvar ortamında yapılır. İncelenmek üzere örnekler belli aralıklarla sıklıkla haftalık olarak laboratuvara teslim edilerek analiz raporları alınır. Bu raporlardan elde edilen verilere yönelik olarak, başlıca elde edilmesi muhtemel parametreler şunlardır:

1. Biyokimyasal Oksijen İhtiyacı (BOİ)
2. Kimyasal Oksijen İhtiyacı (KOİ)
3. Toplam Askıda Katı Maddesi (AKM)
4. Yağ ve Gres
5. Toplam Fosfor (P)
6. Toplam Krom (T.Cr)
7. Krom (Cr)
8. Kurşun (Pb)
9. Toplam Siyanür
10. Kadmiyum (Cd)
11. Demir (Fe)
12. Florür

13. Bakır (Cu)
14. Çinko (Zn)
15. Civa (Hg)
16. Sülfat (SO<sub>4</sub>)
17. Toplam Kjeldal Azotu
18. Suyun Rengi
19. Suyun pH'ı

Belirtilen parametrelere ait parametre değerleri laboratuvar ortamında incelenerek, Su Kirliliği Kontrol Yönetmeliği kapsamında yer alan sınır değerlerine göre, elde edilen ölçüde Atıksuyun sınır değerleri ölçülerek tahliye şekline karar verilmektedir (Türkkak Test Raporu, 2017).

## **2.2. Atıksu Arıtma Tesislerinde Arıtma Tekniklerinin Kullanımı ve Kimyasal Terimler**

Atıksu arıtma tesislerinin temel amacı evsel/endüstriyel kaynaklı olarak oluşan Atıksuların kimyasal, biyolojik ve fiziksel olarak arıtılma işlemine tabi tutularak daha kaliteli çıkış suyunun elde edilmesini sağlamaktır. Arıtılan Atıksu çıkış parametrelerine uygun olması ve böylece nehir, deniz gibi akarsulara deşarj edilebilir hale getirilmesi amacı ile çalışmaktadır. Arıtma yöntemi gelecek olan Atıksuyun karakteri ve çıkış suyu kalitesi hesaplanarak belirlenmelidir. Bu kısımda Atıksu arıtımında sıklıkla kullanılan uygulamalardan bazıları açıklanmaktadır.

### **2.2.1. Biyokimyasal Oksijen İhtiyacı (BOİ) arıtma yöntemi**

Biyokimyasal oksijen ihtiyacı Atıksulardaki kirlilik yükünün (organik maddelerin) ölçüsü olarak tanımlanmaktadır. Aerobik bakteriler tarafından parçalanabilen organik maddelerin biyokimyasal oksidasyon (karbon içeren maddelerin oksitlenmesi olayı) sırasında harcanan oksijen miktarı esas alınarak hesaplanan değere BOİ adı verilmiştir. BOİ<sub>5</sub> testinde minimum beş (5) gün süre ile 20 °C sıcaklıkta bekletilerek sonuca ulaşılır. Suyun içinde bir yanma olayı olan, biyokimyasal oksidasyon yöntemi

Atıksuyun içinde çözülmüş halde bulunan oksijen miktarından yararlanılarak yapılan bir yakma yöntemidir. Bu yöntemde Atıksuda bulunan karbon miktarının tamamen karbondioksit (CO<sub>2</sub>) olarak sudan ayrıştırılmasını ve tahliyesini amaçlayan arıtma biçimidir. Bu yöntem Atıksuda bulunan tüm karbonlu maddelerin tükenmesine kadar ardı ardına uygulanır. Gerçekleştirilen analiz biyolojik şekilde arıtılabilme derecesinin belirlenmesinin yanı sıra bir arıtma tesisinin sahip olduğu performansın değerlendirilmesinde kullanılan etkin bir biyoanaliz yöntemidir. Bu test evsel ve endüstriyel atıkların kirlilik derecesini belirleyen temel testtir. Oksitlenme olayının gerçekleşme aşamasında oksijen tüketimi miktarında meydana gelen artışın sebebi Atıksuda var olan fazla organik madde miktarıdır. Bu durumda Atıksuda var olan organik madde miktarı ile oksijen tüketimi arasında doğru bir orantı gözlemlenmektedir.

Arıtma yönteminin seçiminde, aktif çamur ve damlatmalı filtre ünitelerinin boyutlandırılmasında, BOİ önemli bir faktördür. Alıcı su ortamlarında ne kadar BOİ harcanacağı veya oksidasyon hızının belirlenmesinde kullanılan, biyolojik olarak oksitlenebilecek organik madde miktarının ölçümünü sağlayan tek testtir. Arıtma yönteminin seçiminde, aktif çamur ve damlatmalı filtre ünitelerinin boyutlandırılmasında, BOİ önemli bir faktördür (Nas, 2017).

### **2.2.2. Kimyasal Oksijen İhtiyacı (KOİ) arıtma yöntemi**

Kimyasal oksijen ihtiyacı organik maddelerin, kimyasal olarak oksitlenmesi esnasında gerekli olan oksijen ihtiyacı olarak tanımlanmaktadır. Arıtma sistemlerindeki BOİ ayrıştırması, zor ayrışabilen (inert) maddelerin meydana gelmesine sebep olur. KOİ bileşenleri asidik ortamda kuvvetli oksitleyicilerden potasyum gibi kuvvetli bir oksitleyici ile ölçülerek zor ayrıştırılabilen KOİ, reaktörde biyolojik ayrıştırılmadan etkilenmemektedir. Bu sebeple parça halindeki inert KOİ, havuzlama sisteminde bulundurulmaktadır. Bu doğrultuda biyolojik çamurda biriktirilen KOİ, çamurla beraber sistemden atılmaktadır. Buradaki mikroorganizmatik popülasyon barındıran inert KOİ, organik maddenin hidroliz aşamasından geçtikten sonra kolay ayrışan substrat'a dönüştürülmüş olur (Turovskiy ve Mathai, 2006 ; Öztürk ve Timur, 2017).

BOİ ile sadece mikroorganizmalar tarafından kullanılabilen organik maddeler ölçülürken, Kimyasal Oksijen ihtiyacı ile bütün organik maddeler ölçülebilir. Bu sebep ile Atıksuların verimli olarak arıtılabilmesi BOİ/KOİ oranına bakılarak karar verilir. KOİ deney sonuçları kısa zamanda elde edilebilir ve bu sayede tesislerde işletme yönünden oluşabilecek hatalar aynı gün içerisinde düzeltilebilir. BOİ deneyi ile birlikte kullanıldığında KOİ deneyi toksik şartların ve biyolojik olarak dirençli organik maddelerin belirlenmesinde oldukça faydalıdır.

### **2.2.3. Toplam Askıda Katı Madde (AKM)**

Toplam Askıda Katı Madde (AKM) bir atıksu arıtma tesisinde arıtma işlemine tabi tutulan suyun işlem sonrasındaki kalitesinin analiz edilmesi için kullanılan, su kalitesine ilişkin bir parametredir. AKM tayini Atıksuların kirlenme derecesini değerlendirmek için kullanılır. Ön çöktürme ünitesinde çökebilen katıların giderilmesinden sonra kalan askıda katıların miktarının tayininde önemlidir. Böylece geriye kalan miktar yani ikincil biyolojik arıtma birimlerine gidecek olan yük tayin edilebilir. Büyük arıtma tesislerinde askıda katı madde tayinleri, arıtma birimlerinin veriminin bir ölçüsü olarak rutin olarak kullanılmaktadır (Nas, 2017).

Toplam Askıda Katı maddelerin atıksuda bulunan granüllerin filtrelerden geçirilerek, bir nevi süzülerek arıtılması biçimidir. Bu bölümde çıkış suyundaki toplam AKM'yi ayrıştırmak için çeşitli filtrasyonlar uygulanmaktadır. En bilinen ve yaygın olan çeşitleri granüler filtrasyon, kimyasal ilaveli filtrasyonlar, elektrokoagülasyonel filtrasyonlardır. Atıksuyun, deşarj edilmesinden önce uygulandığı gibi ayrıca biyolojik arıtmadan kalan partikülleri gidermede, çökertme işlemleri sonrasında da uygulanmaktadır. Atıksu dönüştürüldükten sonra insan teması her aşamada göz önünde bulundurularak arıtma işlemi yapılmaktadır. Bunun için filtrasyon birkaç aşamada uygulanmaktadır. Filtreleme işlemleri için filtrasyon malzemelerinin temizlenmesi ve yeniden kullanımı hususu dikkatle yapılmalıdır (Öztürk ve Timur, 2017).

#### **2.2.4. Yağ ve gres arıtma yöntemleri**

Konutlar, lokantalar, benzin ve akaryakıt istasyonları gibi ortamlardan elde edilen ve yağ içeren Atıksuların arıtılmasını sağlayan yöntem yağ ve gres tutucunun kullanılmasıdır. Yağ ve Gres katılarak suyun üzerine çıkar ve Atıksudan sıyrılarak temizlenir. Arıtma ünitesi perdeler ile birbirinden ayrılmış bölmeler halinde bir iç tasarıma sahiptir. Septik tank içerisinde hidrololik çökme sağlandıktan sonra üstte kalan köpük ve çürümüş çamur birlikte kanallardan toplanarak başka bir arıtım ünitesine veya uygun bir araziye deşarj edilir. Tankın dibinde kalan maddeler ise anaerobik mikroorganizmalar tarafından parçalanıp karbondioksit, metan ve hidrojen sülfür olarak tahliye edilmektedir (Hager, 1997 ; Trovskiy ve Mathai, 2006).

### **2.3. Atıksu Arıtma Tesislerinde PLC ve SCADA Sistemleri ile İlgili Yapılan Çalışmalar**

SCADA sistemlerinin, atıksu arıtma tesislerinde kullanıldığı ve literatürde yer alan çalışmalardan bazıları şu şekildedir: SCADA sistemleri ile atıksu arıtma tesislerinde kullanılması düşüncesini yararlı bularak SCADA sistemleri için bir yazılım programı önermişlerdir (Nascu, 2018). Atıksu arıtma tesislerinde SCADA sisteminin uygulanması adı altında bir makale ile Houston Şehri'nde bir atıksu arıtma tesisinin içerisinde yer aldığı bir projeden ve projenin uygulanmasından söz etmiş ve diğer tesislerde de uygulanması ile bu maliyetli işe alternatif olarak SCADA kullanımını önermiştir (Dieu, 2001).

Çanakkale ili, Bayramiç kazasında yer alan ve daha önce analog kontrol kumanda tarzında projelendirilmiş bir atıksu arıtma tesisindeki (Bayramiç Kazası (Çanakkale) Atıksu Arıtma Tesisi) proses işleyişinin PLC (Programmable Logic Controllers) ve SCADA (Supervisory Control and Data Acquisition) sistemine dönüştürülmesi sürecini incelemişlerdir (Güzelbeyoğlu, 2002). Atıksu arıtma tesisinde yapılan bir çalışmada SCADA sistemin kullanımı ile %17-21 oranında enerji tasarrufu sağlanacağını savunmuştur (Gülsoy, 2017). İstanbul Paşaköy Atıksu Toplama Havzası'nda, bir otomatik kontrol sistemi kurulması konusunda öneri getirilmiş

geliştirilen önerinin maliyet boyutu hesaplanmıştır (Çakmakçı, 2001). Atıksu arıtma tesislerinde yapılan çalışma sonucunda belediyelerde SCADA sistemlerin kurulmasının tasarruflu olacağı sonucuna varmışlardır (Pitas ve ark, 2010).

Dınış ve Popa 2014 yılında atıksu arıtma tesisinde Kullanılan Scada Sisteminde Ölçümler üzerine yaptıkları çalışmada Brad Belediyesi Atıksu Arıtma Tesisi teknolojik akışını incelmış ve Atıksu Arıtmalar için bir yazılım platformunun tasarımı ve uygulanmasını rapor etmiştir. Platform, karmaşık kontrol sistemlerinin uygulanması, test edilmesi ve tesis genelinde operasyon için veri kalitesi analizi için araçlar sağlamak amacıyla mevcut bir kontrol sistemine, yani tesis SCADA sistemine entegre edilebilir. Sistem bir kamu kuruluşunun atıksu arıtma tesisinde uygulanmakta ve test edilmektedir (Dinis ve Popa, 2014).

#### **2.4. Atıksu Arıtma Tesislerinde Yapay Zeka ve Makine Öğrenmesi Teknolojilerinin Kullanımı ile İlgili Gerçekleştirilen Çalışmalar**

Atıksu arıtma tesislerinde oldukça karmaşık ve zaman alan parametre verileri Yapay Zeka yöntemleri kullanarak tahmin edilebilmektedir. Bu tahminler sayesinde hem ekonomik, hem kolay, hem de hızlı yönetimin önünü açmaktadır. “Zhao ve arkadaşları 2020 yılında yaptıkları çalışmada Atıksu arıtımını, kirletici maddenin azaltılması ve su ortamının teşvik edilmesi için önemli bir su kalitesi göstergesi olarak tanımlamışlardır. Doğal koşulların karmaşıklığı, giriş şoku ve atıksu arıtma teknolojisi atıksu arıtma sisteminde belirsizlikler söz konusu olduğundan yapay zeka teknolojisinin uygulanmasındaki eğilimlerle su kalitesinin arttırılacağını öne sürmüşlerdir. Araştırmacılar çalışmalarında, büyük ölçekli bir bibliyometrik analizi atıksu arıtma tesisinde yapay zekanın, hangi işlemleri tanıyacağını ve işlemi nasıl gerçekleştirmesi gerektiğini PLC kodlarıyla sağlamışlardır. Ayrıca, uygulamanın dört yönünün sistematik bir incelemesini çalışmalarında sunmuşlardır (Zhao ve ark., 2020).

Arıtma proseslerinin kontrol ve gözlem aşamasında çoklu istatistiksel yöntemler ve yapay zekanın işlevselliğinin ve uyarlanabilirlik potansiyelinin araştırılması için gerçekleştirilen bir çalışmada KOİ, BOİ, pH, İletkenlik, Bulanıklık, Alkalinite, Toplam



Azot Miktarı (TN), Giriş Suyu Debisi (Q), Askıda Katı Madde (AKM), Çözünmüş Oksijen (ÇO), Sıcaklık (t) parametreleri Aktif Çamur Modelleri (ASM1 ve ASM3), Temel Bileşenler Analizi, Çoklu Lineer Regresyon Analizi, Bulanık Mantık (FL), YSA, Adaptif Ağ Temelli Bulanık Çıkarım Sistemi tahmin yöntemleri kullanılarak analiz edilmiştir. Elde edilen çıktılar göz önünde bulundurulduğunda Yapay Zeka yöntemlerinden biri olan YSA ve Adaftif Ağ Temelli Bulanık Çıkarım Sistemi'nin işlevselliklerinin ve uygulanabilirliklerinin diğer yöntemlere göre daha yüksek olduğu sonucuna varılmıştır (Civelekoğlu, 2006).

Atıksularda bulunan ağır metal ve boyar maddelerin giderilmesi için adsorban olarak atıksu arıtma çamurunun kullanılabilirliğinin Yapay Sinir Ağları modellemesi ile araştırılması için yapılan bir çalışmada kirleticilerin maksimum giderim veriminin daha az deneyle belirlenmesi hedefi doğrultusunda YSA'nın oldukça etkili bir tahmin yöntemi olduğu ortaya konulmuştur (Öztürk, 2014).

Aktif çamur prosesinin sistem davranışının ve değişkenlerin kendi aralarındaki ilişkilerinin analizi için yapay sinir ağlarının kullanıldığı bir çalışmada Kohonen Özöğütlemeli Harita Ağı Metodu kullanılmış ve Kohonen Özöğütlemeli Harita Ağı Metodu'nun çıkış parametrelerinin ve oluşabilecek problemlerin tahmininde kolaylık sağladığı ortaya konulmuştur (Çınar ve Yılmaz 2005).

Atıksu arıtma tesisinin işletilmesi esnasında tipik bir karmaşık kontrol problemini ele alan araştırmacılar, doğrusal olmayan dinamikler ve değişkenler arasındaki kuplaj etkileri ile gerçek zamanlı optimum kontroller incelenmiştir. Bu çalışmada araştırmacılar, bir Q-öğrenme 2d klavuzlu (ASM2d klavuzlu) ödül ayarlı (entegre bir ASM2d-QL algoritması) önermişler ve yaygın olarak uygulanan anaerobik-anoksik-oksik (AAO) sistemi araştırma paradigması olarak seçmişlerdir. AAO'nun kontrol stratejilerini optimize etmek için değişen akış yükleri altında sistem, Q matrisleri hem HRT'ler hem de IRR optimizasyonu için oluşturmuşlardır. Akıllı modelleme ve kararlı atıksu arıtımı sırasında dalgalanan giriş yükleri altında optimal kontrol stratejileri ele alınmıştır. Atıksu arıtma sürecini optimize etme görevi, optimum AAT'nin işletme

koşullarının biyolojik, fiziksel ve biyolojik olması nedeniyle kontrol edilmesinde zorlanmışlardır (Wang ve ark., 2019).

Atıksu Arıtma süreçleri fiziksel, biyolojik ve kimyasal arıtma proseseleri ile operasyonel açıdan karmaşıklık oluştuğu için genellikle programlanabilir mantık denetleyicileri kullanmışlardır. Yaygın olarak kullanılan stratejiler çoğunlukla belirli operasyonları dikkate almayan iş kurallarına ve uzman yönergelerine dayanır. Bu koşullar aşırı veya yetersiz havalandırmaya yol açabilir ve bu çok maliyetli olabilir. Havalandırma sürecine odaklanan bu çalışmada, makine öğrenimi yaklaşımı uygulanarak atıksu arıtma süreçlerini optimize etmeyi amaçlamışlardır. Hedefler blower'ı (havalandırıcıyı) uyumlu hale getirmek, gereksiz dalgalanmaları en aza indirmek ve makine aracılığıyla anlamak için günlük operasyonel kararların nasıl ele alındığını vurgulamışlardır (Mercier ve ark., 2019).

AAT'nin biyolojik sürecini oldukça karmaşık, doğrusal olmayan, dinamik davranışı ve mikrobiyal bileşenlerin çeşitliliği nedeniyle dokunulmaz iki makine öğrenimi tekniği ile Yapay Sinir Ağları desteklemiş ve bu karmaşık sistemi modellemek için vektör regresyonu kullanmıştır. Sonraki verilerin kullanılması nesil dizileme, mikrobiyal topluluk kompozisyonu ortaya çıkarmış olup bu veriler, makine öğrenimi modellerinin bir seçimini tahmin etmek için girdi olarak kullanılmış ve işlem performans parametreleri her iki modelde tatmin edici sonuçların ötesinde sonuçlar göstermiştir. Modellenen her sürecin hassasiyetini analiz ederek her mikroorganizmanın parametresi, mikrobiyal etkisinin bir göstergesi olup süreç performansı ile ilgili yapı oluşturmuşlardır (Cornelissen, 2018).

Ulsan'da (Kore) gıda atığı karıştırılan bir atıksu arıtma tesisinde TN konsantrasyonunu tahmin edebilmek amacıyla YSA ve Destek Vektör Makinesi (DVM) yöntemlerini kullanmışlar ve her iki modelde de iyi performans yakalmışlar ortaya koymuştur. Bu çalışma, entegre gıda atığı ve Atıksu arıtma sürecinin su kalitesinin erken tahmini için verimli ve sağlam doğrusal olmayan zaman serisi modelleme yöntemini önermektedir (Guo ve ark., 2015). Gerçekleştirilen bir çalışmada Nöral makine öğrenimi sistemine bağlı bir mikrobiyal diziyi hızlı BOİ değerlendirmesi yapmak amacı ile Özdevimli

Öğrenme algoritması ve mikrobik bakterilerin birlikte kullanılması ile Biyolojik Oksijen İhtiyacı (BOİ<sub>5</sub>)'nin ölçülme süresinin kısaltılması amaçlamıştır. Çoklu Lineer Regresyon (MLR), Kısmi En Küçük Kareler Regresyonu (PLS), Özdevimli Öğrenme algoritması (ML), Çok Katmanlı Yapay Sinir Ağları yöntemleri kullanılmış ve çalışmada geliştirilen metot ile referans metot arasındaki güçlü ilişki ortaya konulmuştur. Geliştirilen metodun atıksuyun gözlemini ve kontrolünü kolaylaştırdığı ve arıtım süresini kısalttığı gözlemlenmiştir (Jouanneau ve ark., 2019).

Sensör Füzyonu ve Makine Öğrenimi Sistemi ile atıksu kalitesinin izlenmesi için gerçekleştirilen bir çalışmada UV/Vis Spektrokopi ve Türbidimetri kullanarak atıksu kalitesinin izlenmesinde kullanılacak bir sistem geliştirilmiştir. Araştırmacılar UV/Vis Spektroskopi, Kısmi En Küçük Kareler Regresyonu (PLS) metodları kullanılarak tahmin edilen değerler ile analitik değerlerin yüksek uyum içinde olduğu göstererek ve geliştirilen sistemin etkinliği ortaya koymuştur (Qin ve ark., 2012).

Atıksu Kalite Göstergeleri tahmini için Makine Öğrenimi Algoritmalarının kullanıldığı bir çalışmada drenaj havzasının bazı özellikleri baz alınarak Atıksu kalitesinin analiz edilmesinde dolaylı bir metodoloji elde edilmesi için (BOİ<sub>5</sub>), (KOİ), (AKM), Toplam Çözünmüş Katı parametrelerine ilişkin veriler Destek Vektör Regresyonu (DVR), Karar Ağaçları (Regression Trees, RT), En Küçük Kareler Regresyonu (LS) kullanılarak analiz edilmiştir. Oluşturulan modeller ile sağlamlık, güvenilirlik ve yüksek genelleme yeteneği potansiyeli ortaya koymuştur (Granata ve ark., 2017).

Kuzey Avrupa'da bulunan atıksu arıtma tesislerinden alınan veriler ile proses parametreleri ve enerji tüketimi arasındaki ilişkileri tanımlamak için Makine Öğrenmesi Maliyet Modellemesi (MLCM) metodolojisini kullanılmıştır. Bu çalışmada ise ilk defa enerji fiyatı model parametresi olarak kullanılmış ve önemi değerlendirilmiştir. Yapılan çalışma ile MLCM algoritmaları ile model performans göstergelerinin genellikle literatürde bulunanlardan daha iyi olduğu gösterilmiştir (Güçlü ve Dursun, 2007). Tam ölçekli kentsel atıksu arıtma tesislerinin bilgisayar programı kullanılarak modellenmesi ve arıtma performanslarının incelenmesi için

gerçekleştirilen bir çalışmada atıksu arıtma tesislerinin proses kontrolü ve izlemesi aşamasında yapay zeka ve hibrit modellemelerinin uygulanabilirlikleri belirlenmeye çalışılmış ve Hibrit Model'in, performans tahmininde Aktif Çamur Modeli (ASM1) ve Yapay Sinir Ağları Modeli (YSA)'dan daha başarılı olduğu ortaya konulmuştur (Güçlü, 2007).

#### **2.4.1. Atıksu arıtma tesislerinde performans tahmini ve iyileştirilmesi için kullanılan yapay zeka teknikleri**

Xiaofei He 2014 yılında yaptığı çalışmasında, modelleme ve endüstriyel sistemlerin optimizasyonunu iyileştirmeyi amaçlayan aktif bir araştırma alanı olmuştur. Bu tür verilerde çok değerli bilgiler yer alsa da verilerin endüstriyel sistemlerin modellenmesinde kullanımı gecikmektedir. Veri madenciliği yeni bir bilimdir, karmaşık sistemleri ve süreçleri modellemek için bir platform ve teknik sağlamaktadır. Veri madenciliği sistemlerin hedef ve operasyonel maliyeti arasındaki ilişkiyi açıklayan modelleri üretmek için kullanılır ve sistem kontrol değişkenleri operasyonel maliyeti en aza indirmek için bir optimizasyon modeli oluşturmuştur. Araştırmacının amacı pompa sistemi ve HVAC sistemi yönetimini optimize ederek enerji tasarrufu sağlamaktır. Veri madenciliği yöntemini kullanarak pompalar üzerine yapılan ilk araştırma su dağıtım sistemlerinin tasarımına odaklanmak ve pompa kontrolü olmuştur. Tasarım araştırması, maliyetlerin en aza indirilmesine de vurgu yapmıştır (Zhang ve ark., 2016).

Mukherjee ve arkadaşlarının 2013 yılında yayımladığı çalışmada; ABD Atıksuyunun değişkenliğini teşhis etmek için hiyerarşik bir modelleme yaklaşımı kullanılır. Dönem akış hızı, kapasite kullanımı, mevsimsellik ve uyumluluk geçmişi gibi özellikler ele alınmış ve Kriging kullanılarak mekânsal bir analize tabi tutulmuştur. Mekânsal değişkenliği modellemek için hiyerarşi seviyesi amonyak, mevsimsellik, izin verilen atık amonyak konsantrasyonlarını etkilemez ve limitler genellikle mevsimseldir. Amonyak deşarj limitlerine uygunluk seviyesi önemlidir ve önceki atık konsantrasyonundan büyük ölçüde etkilenmiştir. Hiyerarşi ile anormal uyumluluğun coğrafi bölgelerini tanımlar ve modelin deşarj izin sınırlarına yakın olan bir tesis alt

kümesi için tahmin performansı olarak atıksu arıtma tesis tasarımı ve işletilmesini hedeflemiştir (Suchetana ve ark., 2013).

Najafzadeh, Zeinolabedini 2019 yılında yaptıkları çalışma ile atıksu arıtma tesisinde (AAT) girişte bulunan aşırı safsızlıkları veya kirleticileri istenen seviyeye kadar ortadan kaldıracak uygun bir sistem sağlamaya çalışmışlardır. Bu şekilde, günlük akış hızları, Atıksu arıtma proseslerini ve tesis birimlerini tasarlamak için katkıda bulunulması gereken en önemli bileşenlerden biridir. Bu nedenle, mevcut araştırma çalışmasında günlük akış oranlarının tahmini için, İleri Geri Besleme Sinir Ağı, Radyal Temel İşlevli Sinir Ağı, Uyarlanabilir Nöro-Bulanık Çıkarım Sistemi (ANFIS) ve Destek Vektör Makinesi (DVM) kullanılmıştır. Önerilen modellerin performansı istatistiksel hata göstergeleri kullanılarak eğitim ve test aşamaları için değerlendirilmiştir. Önceki araştırmalar, deneysel çalışmalar yapmanın zaman aldığını ve laboratuvar çalışmaları için yapılan harcamaların arttığını göstermiştir. Dolayısıyla, son yıllarda atıksu arıtma tesis süreçlerinin farklı çevresel yönlerini analiz etmek için Yapay Sinir Ağları (YSA), Uyarlanabilir Nöro-Bulanık Girişim Sistemi (ANFIS), Gen İfade Programlama (GEP), Evrimsel Polinom Regresyon (EPR) ve Destek Vektör Makinesi (DVM) kullanılmıştır. Mantıksal bir bakış açısından, önerilen modellerin performansı, bir Gauss çekirdeği tarafından tasarlanan DVM'nin, diğer yapay zeka tekniklerine kıyasla akış hızlarını tahmin etmede en verimli yaklaşımlardan geçtiğini göstermiştir. Başka bir açıdan bakıldığında, önerilen modeller, arıtma süreçlerinin karmaşıklığı ile beraber AAT'nin en önemli bileşenlerinden birinde başarılı bir şekilde yapılmıştır (Najafzadeh ve Zeinolabedini, 2019).

Hamed ve arkadaşları 2004 yılında bir atıksu arıtma tesisinin (AAT) performansını geçmiş bilgilere dayanarak tahmin etmek için Yapay Sinir Ağları (YSA) modelleri ile çalışmışlardır. Bu çalışmada kullanılan veriler, Mısır'ın Büyük Kahire bölgesindeki El Gabal El Asfar isimli en büyük konvansiyonel atıksu arıtma tesisinden ortalama 1 milyon m<sup>3</sup>/gün debiyle elde edilmiştir. Arıtma sürecinin çeşitli aşamalarında 10 ay boyunca Biyokimyasal Oksijen İhtiyacı (BOİ) ve Askıda Katı Madde (AKM) konsantrasyonlarının günlük kayıtları tesis laboratuvarından elde edilerek, verilerdeki ilişkileri tespit etmek ve veri bağımlılığını değerlendirmek için keşif veri analizi

kullanılmıştır. Tesis Atıksularındaki BOİ ve Askıda Katı Madde (AKM) konsantrasyonlarının tahmini için YSA tabanlı iki model sunulmuştur. Sinir ağı modellerinin uygun mimarisi, modellerin eğitiminin ve test edilmesinin birkaç adımıyla belirlenmiştir. YSA tabanlı modellerin AAT performansını tahmin etmede etkili ve sağlam bir araç sağladığı sonucu elde edilmiştir. Geliştirilen modellerin, değişken doğruluk ve girdi verilerinin boyutları karşısında tutarlı bir şekilde iyi performans gösterdiği belirlenmiştir (Hamed ve ark., 2004).

Yapay Sinir Ağları, atıksu arıtma tesislerinin performans tahmininde dikkate alınmaya değer bir araçtır. Yapay Zeka yöntemleri ile atıksu arıtma tesislerinde yapılan çalışmalar neticesinde birçok iyileştirme ve verimlilik artırıcı yöntemler bulunmuş ve uygulanmıştır. “Atıksu arıtma tesisi performansının tahmini için YSA uygulamalarını ele alan çalışmada su kalitesinin belirlenmesine yönelik olarak kullanılan 3 ana parametrenin tahminde Yapay Sinir Ağları modeli ve regresyon analizi yöntemlerinin kullanımının karşılaştırılması amaçlanmıştır. Su kalitesinin tahmini için BOİ, Sıcaklık, pH, KOİ, AKM parametrelerine ilişkin değerler ele alınarak Kök Ortalama Kare Sapması (RMSE), Ortalama Mutlak Yüzde Hatası (MAPE) gibi hata performans ölçüm metrikleri kullanılmıştır. Çalışma sonucunda BOİ, KOİ, AKM parametrelerinin tahmininde Yapay Sinir Ağları yöntemlerinin Çoklu Lineer Regresyon yöntemine göre daha etkin sonuç verdiği kanıtlanmıştır (Hamada ve ark., 2018).

Atıksu performansının yapay zeka yöntemlerinin görüntü işleme metotları ile birlikte kullanılarak gözlemlenmesini amaçlayan bir çalışmada BOİ, AKM, KOİ, Uçucu Askıda Katı Madde (UAKM), Çökebilir Katı Madde (CKM), Sıcaklığı (t), pH, Debi (Q), Toplam Kjeldahl Azotu (TKN), Toplam Fosfor (TP) parametreleri Yapay Sinir Ağları (YSA), Geri Yayılım Algoritması, Levenberg-Marquardt Öğrenme Metodu, Genetik Algoritma, Yapay Arı Kolonisi Algoritması ve Hücresel Sinir Ağları (HSA) tahmin yöntemleri ile değerlendirilerek atıksu arıtma tesisinin performansının ve aktif çamur sisteminin izlenmesinde görüntü işleme tekniklerinin kullanılabilirliği ortaya konulmuştur. Ayrıca bu tekniklerin aktif çamur prosesinde ortaya çıkabilecek sorunların tahmini noktasında yararlılıkları ortaya konulmuştur (Boztoprak, 2014). Yapay Sinir Ağı ile atıksu arıtma performansının modellenmesi için gerçekleştirilen

bir çalışmada Yapay Sinir Ağı modeli ile atıksu arıtma performansı belirlenmiş ve birden fazla model girdisi kullanmanın model performansını tek model girdisi kullanmaktan çok daha fazla arttırdığı ortaya konulmuştur (Ersü ve Subaşı, 2010). Atıksu arıtma tesisi performansının Yapay Sinir Ağları kullanılarak tahmin edilmesini sağlamak için gerçekleştirilen bir çalışmada Geliştirilen iki tip Yapay Sinir Ağı modeli, Atıksuda bulunan BOİ ve AKM miktarlarının tahmininde kullanılmış ve daha fazla veri ve ek parametre ölçümleri ile modellerin etkinliklerinin arttırılabileceği ortaya konulmuştur (Moreno-Alfonso ve Redondo, 2018).

Nourani, Elkiran ve Abba 2018 yılında Lefkoşa Atıksu Arıtma Tesisi'nde (NWWTP) İleri Beslemeli Sinir Ağı (İBSA), Uyarlanabilir Bulanık Mantığa Dayalı Yapay Sinir Ağı (BMDYSA), Destek Vektör Makinesi (DVM) yaklaşımları ve Klasik Çoklu Doğrusal Regresyon (ÇDR) yöntemlerini kullanarak tesis performansının incelemişlerdir. Üç model tekniğini oluşturmak ve uygulamak için belirtilen tekli modellerin çıktılarını kullanarak tahminin genel verimliliğini arttırmak amacıyla, ağırlıklı olarak basit doğrusal ortalama, doğrusal ortalama alma ve doğrusal olmayan sinir topluluğu yöntemlerin çıktılarının birleştirilmesini önermişlerdir (Nourani ve ark., 2018).

Atıksu kalitesinin tahmini ve atıksu mühendisliğinde ortak makine öğrenimi modellerinin performans değerlendirilmesi için yapılan bir çalışmada Askıda Katı Madde (AKM), (KOİ), (TN), (TP), (t), pH parametrelerine ilişkin değerler kullanılmıştır. Yapay Sinir Ağları (YSA), Destek Vektör Regresyonu (SVR), Çok Katmanlı Yapay Sinir Ağları, Öğrenme Algoritması (SVM), Yapısal Risk Minimizasyonu (SRM) metodları kullanılarak analiz edilen çalışma ile kullanılan yöntemlerin uygulanabilirliklerinin ve etkinliklerinin yüksek olduğu ortaya konulmuştur (Guo ve ark., 2015). Kuzey Portekiz'de bulunan bir atıksu arıtma tesisinin performansının tahmini için Destek Vektör Makinelerinin kullanıldığı bir çalışmada makine öğrenmesi tekniklerinin atıksu arıtma tesisi performansının belirlenmesinde kullanımının etkililiğini ölçülmesi amaçlanmıştır. Çalışmada DVM Öğrenme Algoritması, Veri madenciliği Teknikleri, Sıralı Minimal Optimizasyon Algoritması (SMO) yöntemleri kullanılmış ve özellikle BOİ'nin tahmininde tahmine dayalı

modellerin başarı ile uygulanabileceği ortaya konulmuştur. Çalışmada kullanılan SMO ve DVM yöntemlerinin oldukça benzer sonuçlar ortaya koyduğu gözlemlenmiştir (Ribeiro ve ark., 2013).

#### **2.4.2. Atıksu arıtma tesislerinde parametre tahmini için kullanılan yapay zeka teknikleri**

Atıksu arıtma tesisi performansının modellenmesi için YSA tekniklerinin uygulanmasını öneren çalışmada İleri Beslemeli Geri Yayılı Yapay Sinir Ağları (İBGYYSA) metodunun günlük BOİ giriş miktarının tahmininde Çoklu Regresyon Analizine kıyasla çok daha etkin bir model olduğu kanıtlanmıştır (Yılmaz ve ark., 2008). Atıksu arıtma tesislerinin (AAT) yönetimi ve işletilmesinde parametreler arasındaki düzensizlikler genellikle üstesinden gelinemeyen yönetim sorunlarına yol açmaktadır. Yapılan çalışmalar ile bu sorunlara basit yöntemler ile çözümler aranmıştır. Biyokimyasal Oksijen İhtiyacını (BOİ) tahmin edebilmek amacıyla güvenilir tahmin modelleri seçilerek, atıksu sıcaklığı, pH, kimyasal oksijen talebi, askıda katı madde, toplam azot, toplam fosfor, elektriksel iletkenlik ve giriş debisi gibi parametrelerden yararlanılmıştır. Atıksu arıtma tesisi modelinin oluşturulmasında Klasik Regresyon Analizi, Çok Değişkenli Uyarlamalı Regresyon Spline'lar (MARS), Yapay Arı Kolonisi ve Öğretme-Öğrenme tabanlı optimizasyon kullanılmışlardır. Kök ortalama kare hatası ve ortalama mutlak hata, her biri için performans kriterlerini değerlendirmek amacıyla kullanılmış ve yapılan analizlerin sonuçları birbiriyle karşılaştırıldığında, MARS yöntemi, çalışmada kullanılan diğer yöntemlere göre daha iyi tahmin sonuçları vermiştir. Sonuç olarak, MARS yönteminin, BOİ tahmininde kabul edilebilir sonuçlar ürettiği görülmüş ve ele alınmıştır (Bakia ve ark., 2019).

Biyokimyasal Oksijen İhtiyacı (BOİ) tahmininde kullanılacak basit ve güvenilir bir tahmin modeli belirlemek amacıyla gerçekleştirilen bir çalışmada tahmin parametreleri olarak atıksu sıcaklığı, pH, kimyasal oksijen ihtiyacı, askıya katı madde miktarı, toplam nitrojen miktarı, toplam fosfor miktarı, elektriksel iletkenlik baz alınmıştır. Ele alınan parametrelere ilişkin veriler Klasik Regresyon Analizi, Çok Değişkenli Uyarlanabilir Regresyon Uzanımları (MARS), Yapay Arı Kolonisi



Algoritması, Öğretme- Öğrenme Esaslı Optimizasyon tahmin yöntemleri ile analiz edilmiştir. Çalışmanın sonucunda Biyokimyasal Oksijen İhtiyacı (BOİ) tahminde kullanılacak en güvenilir yöntem Çok Değişkenli Uyarlanabilir Regresyon Uzanımları (MARS) Yöntemi olarak belirlenmiştir (Bakia ve ark., 2019).

Bilgisayar ortamında modellemenin, BOİ parametresinin tahmin edilmesine olanak tanınması ile Yapay Sinir Ağları kullanılarak oluşturulan modellerin değerlendirilmesini sağlamak için gerçekleştirilen bir çalışmada DVM, Öğrenme Algoritması ve Pasif Agresif Regresör Öğrenme Algoritmalarının BOİ<sub>5</sub> parametresinin tahmin edilmesinde Atıksu arıtma tesislerinde kullanılabileceği sonucuna varılmıştır (Baki ve Aras, 2018).

Biyolojik karbon ve azot giderme proseslerinin temsilinde kullanılacak en iyi aktivasyon fonksiyonunun, farklı aktivasyon fonksiyonları arasından Geri Beslemeli Yapay Sinir Ağı (GBYSA) kullanılarak belirlenmesi için gerçekleştirilen bir çalışmada GBYSA kullanımının analiz maliyetleri ve suresi bakımından daha kullanışlı olduğu anlaşılmıştır. Ayrıca KOİ ve TKN giderimlerinin tahmininde GBYSA'nin Sinc fonksiyonu daha işlevsel iken TN gideriminin tahmininde Lojistik ve Sinc fonksiyonlarının daha işlevsel olduğu gözlemlenmiştir (Manav ve Demir, 2017).

Atıksu arıtma tesisinde Kimyasal Oksijen İhtiyacı (KOİ)'nin tahmininde ve tesisin kalite ve performansının izlenmesinde Yapay Sinir Ağı modelinin performans ve yeterliliğinin belirlenmesi amacı ile yapılan bir çalışmada Biyolojik Oksijen İhtiyacı (BOİ), Kimyasal Oksijen İhtiyacı (KOİ), pH, İletkenlik, Toplam Azot (TN), Toplam Fosfor (TP), Çözünmüş Oksijen (ÇO), Askıda Katı Madde (AKM) parametrelerine ilişkin veriler kullanılmıştır. Elde edilen veriler Yapay Sinir Ağı (YSA), Çoklu Lineer Regresyon (MLR), İleri Beslemeli Yapay Sinir Ağları (İBGYSA), Geri Yayılma Eğitim Algoritması (BP) kullanılarak analiz edilmiş ve Yapay Sinir Ağı modelinin KOİ tahmininde Çoklu Lineer Regresyon yöntemine kıyasla çok daha iyi sonuçlar verdiği gözlemlenmiştir (Abba ve Elkiran, 2017).

Atıksu arıtma tesisinin günlük Biyolojik Oksijen İhtiyacı (BOİ)'nin belirlenmesinde kullanılacak bir Yapay Sinir Ağı Modelinin geliştirilmesi için yapılan bir çalışmada

Yapay Sinir Ağı modelinin BOİ tahmininde etkili bir yöntem olduğu kanıtlanmıştır ve Çoklu Lineer Regresyon (MLR) yöntemiyle kıyaslandığında YSA'nın üstünlük sağladığı ortaya konulmuştur (Parthasarathy ve Narayanan, 2014). 3 Farklı Yapay Sinir Ağı metodu kullanılarak Biyolojik Oksijen İhtiyacı (BOİ)'nin belirlenmesi için gerçekleştirilen çalışmada Çok Katmanlı Yapay Sinir Ağı, Öğretme-Öğrenme Esaslı Algoritma, Yapay Arı Kolonisi Algoritması, kullanılmış ve Çok Katmanlı Yapay Sinir Ağı modelinin en iyi yöntem olduğu gözlemlenmiştir (Baki ve Aras, 2018).

Gerçekleştirilen bir çalışmada atıksu arıtma tesisinde oksijen ihtiyacının Yapay Sinir Ağları kullanılarak tahmin edilmesi için atıksu arıtma tesisinin koagülatör malzemelerinin modellenmesi ve optimizasyonunda kullanılacak bir yapay sinir ağı modelinin oluşturulması amacı ile Kimyasal Oksijen İhtiyacı (KOİ), Debi (Q), Biyolojik Oksijen İhtiyacı (BOİ<sub>5</sub>), pH, Askıda Katı Madde (AKM) ilişkin veriler Yapay Sinir Ağı (YSA), Çoklu Lineer Regresyon (MLR), Çok nüfuslu Genetik Algoritması yöntemleri ile analiz edilmiştir. Elde edilen sonuçlar model üzerinde etkisi olabilecek bütün faktörler ele alındığında YSA'nın atıksu arıtma tesisinin performans analizinde ve çıkış deşarj suyunda bulunan KOİ miktarının tahmininde yeterli bir araç olduğu ortaya koymuştur (Gheytaspour ve Bigdarvish, 2018).

Atıksu arıtma tesisinde kirlilik parametresinin tahmini için yapay sinir ağlarını kullanmayı öneren bir çalışmada atıksu arıtma tesisinde Geri Yayılım Algoritması kullanılarak oluşturulan YSA modeli ile tesise ait çıkış parametrelerinin giriş parametreleri üzerinden tahmin edilmesi amaçlanmıştır. Bu doğrultuda Sıcaklık (t), Debi, pH, AKM, KOİ parametreleri Tesis Elektrik Sarfiyatı, Yapay Sinir Ağları (YSA), Geri Yayılım Algoritması kullanılarak analiz edilmiştir. Geri Yayılım Yapay Sinir Ağı ile yapılan çıkış parametrelerinin tahmini için yüksek oranda doğru sonuç verdiği gözlemlenmiştir (Öztemel ve Dügenci, 2016). Yapay Sinir Ağları modelinin ön arıtma ile biyolojik arıtma işlemlerinin gerçekleşmesi sürecinde çıkış parametrelerinin tahmininde, giriş değerleri baz alınarak yapılan çalışmada daha verimli sonuçlar elde etmek için Yapay Sinir Ağlarının iki farklı modeli olan YSA-1 ve YSA-2'nin birlikte kullanılmasının gerekliliği ortaya konulmuştur (Tezel ve ark., 2010).

Kentsel atıksu arıtma tesisi Biyolojik Oksijen İhtiyacının (BOİ<sub>5</sub>) makine öğrenmesi yöntemleri ile tahmin edilmesi amacı ile gerçekleştirilen bir çalışmada Biyolojik Oksijen İhtiyacı (BOİ<sub>5</sub>) parametresinin tahmin edilmesinde 3 farklı makine öğrenmesi algoritması kullanılarak oluşturulan tahmin modelleri değerlendirilmiştir. Debi (Q) , Sıcaklık (T) , pH, İletkenlik, Kimyasal Oksijen İhtiyacı (KOİ), Askıda Katı Madde (AKM) parametreleri SVM Öğrenme Algoritması, SGD Regresör Öğrenme Algoritması, Pasif Agresif Regresör Öğrenme Algoritması öğrenme algoritması ile analiz edilmiştir. Çalışma sonucunda SVM Öğrenme Algoritması ve Pasif Agresif Regresör Öğrenme Algoritması BOİ<sub>5</sub> parametresinin tahmin edilmesinde atıksu arıtma tesislerinde kullanılabilceği anlaşılmıştır (Sakarya ve ark., 2019). Almanya’da bir atıksu arıtma tesisinde yürütülen çalışmada, arıtma tesisi girişinde, havalandırma tankında ve çıkışta pH, iletkenlik, bulanıklık, redoks potansiyeli, KOİ, NH<sub>4</sub>- N ve NO<sub>3</sub>-N parametreleri ölçülmüştür. Giriş KOİ, NH<sub>4</sub>-N konsantrasyonları ile havalandırma tankı NH<sub>4</sub>-N ve NO<sub>3</sub>-N konsantrasyonları MLP modeli ile tahmin edilmiş ve aynı zaman da alınan eksik verilerinin Yapay Sinir Ağları (YSA) kullanarak tamamlanabileceği gösterilmiştir (Hack ve Köhne, 1996).

Kentsel atıksu arıtma tesisine gelen atıksuyun giriş parametreleri baz alınarak çıkış parametreleri olan Kimyasal Oksijen İhtiyacı, Biyolojik Oksijen İhtiyacı ve Askıda Katı Madde değerlerinin tahmini ile ilgili yapılan çalışmada çok katmanlı algılayıcı kullanılmış ve arıtma verimine etki eden parametrelerin bilinmesi ile YSA yöntemi kullanılarak verim değerleri elde edilmiştir. Gerçekleştirilen çalışmada R<sup>2</sup> değeri 0.845 bulunmuş ve çok katmanlı algılayıcı modelin gerçek verilere karşı oldukça iyi bir sonuçla uyum sağladığını öne sürülmüştür (Keleş ve ark., 1999). Baeza ve arkadaşları 2002 yılında atıksu arıtma tesisinde giriş parametrelerinden debi, biyokütle konsantrasyonunu ve sıcaklık değerlerini kullanarak Kimyasal Oksijen İhtiyacı parametresini tahmin etmek amacıyla Yapay Sinir Ağları modelini kullanmışlardır. Yapılan çalışmada Yapay Sinir Ağları modelleri kullanılarak KOİ değerini tahmin edilebileceğini göstermişlerdir (Baeza ve ark., 2002). Oliveira-Esquerrel vd. 2002 yılında bir atıksu arıtma tesisinde giriş BOİ değeri ve çıkış BOİ değerlerini geliştirdiği model yardımıyla tahmin etmeye çalışmış ve en iyi sonucu MLP ile elde etmiştir. Çalışmanın sonucunda giriş BOİ için R<sup>2</sup>=0.813 ve çıkış BOİ için ise R<sup>2</sup>=0.697

sonuçları ile Biyolojik Oksijen İhtiyacı parametrelerinin modellenebileceğini ortaya koymuştur (Olivera- Esquerrel ve ark., 2002). Açıkalın 2005 yılında Adapazarı Atıksu Arıtma Tesisi çıkış kirlilik parametrelerinden BOİ, KOİ ve AKM değerlerini tahmin etmek için YSA metodu kullanmıştır. Yapılan bu çalışmada çok katmanlı algılayıcı modelinin gerçek verilere göre oldukça iyi bir sonuç verdiği görülmüştür (Açıkalın, 2005).

Düğenci 2007 yılında, Adapazarı Karaman Atıksu Arıtma Tesisinde yaptığı çalışmada tesis giriş parametrelerinden sıcaklık, pH, debi, askıda katı madde, KOİ ve havalandırma havuzunda bakterilerin yaşamsal döngüsü için kullanılan elektrik tüketimleri değerleri kullanılarak çıkış parametre değerlerinin tahmini sağlanmıştır. Yapay Sinir Ağı modelinin eğitimleri için klasik geri yayılım metodu, genetik algoritma ve arı algoritması kullanılmıştır. Çalışma neticesinde arı algoritmasının yapay sinir ağı eğitiminde başarılı bir şekilde uygulanması ve elde edilen model ile tesis kontrolünde faydalanılmasının mümkün olabileceği ortaya konmuştur (Düğenci, 2007).

Sezer 2007 yılında yaptığı çalışma ile 2 farklı biyolojik atıksu arıtma tesisinden almış olduğu KOİ ve BOİ değerlerini YSA yöntemiyle tahmin ederek YSA'da elde edilen sonuçların regresyon analizinden elde edilen sonuçlar kadar iyi olduğunu ileri sürmüştür (Sezer, 2007).

Doğan ve arkadaşları 2008 yılında yaptığı çalışmada Adapazarı Atıksu Arıtma Tesisinde giriş parametrelerinden biyolojik oksijen ihtiyacı tahmini için Yapay Sinir Ağları ve çoklu regresyon modeli yöntemlerini kullanmışlardır. Yapılan bu çalışma da tahmin için kullanılan giriş parametreleri (KOİ, debi, AKM, TN, TP) ile kurulan modellerin sonucunda YSA modeli MLR modeline göre daha başarılı olmuştur (Doğan ve ark., 2008). Yılmaz 2009 yılında Adapazarı Karaman Atıksu Arıtma Tesisinden alınan KOİ, AKM, P, N, Q parametrelerine ait veriler ile BOİ parametresini tahmin edebilmek amacıyla YSA modelini kullanmıştır. Yapılan çalışma sonuçlarına göre YSA modelinin Çoklu Regresyon Analizi'nden çok daha etkili bir model olduğu belirtmiştir (Yılmaz, 2009). Lee ve arkadaşları 2011 yılında bir atıksu arıtma tesisinde giriş

parametrelerinden KOİ, AKM, toplam azot (tN) parametrelerini kullanarak BOİ parametre tahmini için yaptıkları çalışmada yapay sinir ağlarında 4 adet çoklu bağımsız modelleme yapmışlar ve bu modellemenin kullanılabilir bir model olduğunu ortaya koymuşlardır (Lee ve ark., 2011). Heddam ve arkadaşları 2016 yılında, Sidi Marouane (Cezayir) Atıksu Arıtma Tesisinde BOİ tahmini için yaptıkları bir çalışmada KOİ, AKM, elektriksel iletkenlik, sıcaklık ve pH parametrelerine ait veriler ile Genelleştirilmiş Regresyon Yapay Sinir Ağı (GR-YSA) yöntemini kullanmışlardır. Bu model ile elde edilen korelasyon katsayısı  $r=0.922$ , standart çoklu regresyon analizinden (SÇRA) elde edilen ise  $r=0.897$  bulunmuş ve GR-YSA'nın SÇRA'ya göre daha güvenilir tahmin kapasitesi olduğunu ileri sürmüşlerdir (Heddam ve ark., 2016).

## **BÖLÜM 3. ATIKSU ARITMA TESİSLERİNDE KULLANILAN ÖLÇÜM VE KONTROLÜ SAĞLAYAN OTOMASYON SİSTEMLERİ**

### **3.1. Atıksu Arıtma Tesislerinde Kontrol ve Planlamanın Önemi**

Atıksu arıtma tesislerinde uygulanması ve uyulması gereken kurallar, yasalar kapsamında çerçevesi belirlenmiş düzenlemeler ile halk sağlığı ve güvenliği yükümlülüklerini yerine getirme görevleri, ilçe ve bağlı oldukları il belediyelerine verilmiştir. Türk mevzuatında, 21 Mayıs 1991 tarih ve 91/271/EEC sayılı Kentsel Atıksu Arıtımı Konsey Direktifine (KAAD) dayalı olarak kentsel Atıksu deşarjını düzenleyen iki temel dayanak bulunmaktadır. Bunlar;

1. Kentsel Atıksuların Kentsel Atıksu Arıtımı Tesislerinde (KAAT) toplanması, arıtılması ve deşarjını düzenleyen Kentsel Atıksu Arıtımı Yönetmeliğı (KAAY).
2. Tüm meskenlerin ve sanayilerin yüzey sularına yaptıkları deşarjlar nedeniyle meydana gelebilecek su kirliliğini düzenlemek amaçlı, Su Kirliliğı Kontrolü Yönetmeliğı'dir.

Denetim mekanizması ise halk sağlığı kuruluşları aracılığı ile Sağlık Bakanlığı, çevre kirliliğı kontrolü aracılığı ile de Çevre ve Şehircilik Bakanlığı ana kuruluşlar olarak belirlenmiştir. Bu yasal çerçeve kapsamında belediyeler, KAAT'lerin ve yüzey sularına yapılan deşarjlarla ilgili mevzuatı uygulamak ve denetlemek üzere Atıksularını Arıtma Tesisleri kurmak ve mevzuat kapsamında işletmekle yükümlüdür. Aşağıdaki şematik gösterimde atıksu arıtma tesislerinde uygulanan yasal çerçevede her iki yönetmeliğın de ilgili maddelerine yer verilmiştir. Atıksu arıtma tesislerinde planlamanın başlıca etmeni yasal çerçeveye uyma sorumluluğudur.

### 3.2. Denetleyici Gözetim ve Veri Toplama için SCADA Otomasyon Sistemleri

SCADA sistemleri, kullanıcıların fiziksel aygıtları, işlemleri ve olayları uzaktan ve gerçek zamanlı olarak izlemelerine ve denetlemelerine olanak tanır. Bu sistemler endüstriyel prosesler için kritik önem taşıdığından, genellikle oldukça güvenilir ve özel bir donanımda çalıştırılmaktadırlar.

Denetimde yer alan SCADA sistemleri, kontrol ve veri toplama sistemleri, endüstriyel kontrol sistemleri, ICS adındaki süreçlerin yönetimi için kullanılan sistem grubuna dâhildir. Bu sebeple yapısal olarak ortaklıkları mevcuttur ve buna ilişkin örnek Şekil.9'da yer almaktadır. Binlerce kilometrekarelik sahada, coğrafi olarak dağılmış alanları yönetmek için kullanılmaktadır. SCADA sistemlerde kontrollü sürece ve sisteme bağlı olarak değişen karmaşıklık dereceleri bulunmaktadır. SCADA sistemlerde özel uygulama SCADA sistemleri, giderek daha karmaşık sistemlerin incelendiği bilimsel araştırma süreci alanında yaygın olarak kullanılmaktadır. Bu alandaki Ar-Ge, ölçüme zorluklar ve özel gereksinimler ekler, kurulum testleri altında kullanılan izleme ve kontrol sistemi programlanır.

Atıksu arıtma tesisleri (AAT), işlenmiş hammadde kütlesi açısından en büyük endüstri olarak kabul edilebilir. Dağıtım şebekesi SCADA sadece veri toplama ve denetleme kontrolü modellerini ve yöntemlerini değil, aynı zamanda güç şebekesi topolojisi modellerini ve yöntemlerini de içermektedir. Güç ağının genişlemesi ve piyasaya yönelik çalışmasıyla, çalışma ve arıza modları artmaya başlar, bu nedenle izlenmesi ve analiz edilmesi gereken veriler artar. Özellikle cihazlar arızalanırken veya acil durum altındayken, analiz ve karar verme görevleri o kadar acil ve ağırdır ki, operatör genellikle bunları mükemmel bir şekilde başaramaz. Bu nedenle, etkin SCADA'ya bağımlı olmaları gerekir (Stojkovic ve Vukasovic, 2009 ; Guozhen ve ark., 2009 ; Mathew ve Shankar, 2010 ; Dinis ve Popa, 2014).

Gerçek zamanlı verileri toplamak ve analiz etmek için bir bilgisayar sistemi olan SCADA platformları, telekomünikasyon, su ve atık kontrolü, enerji, petrol, gaz arıtma ve taşıma gibi endüstrilerdeki bir tesisi veya ekipmanı izlemek ve yönetmek için kullanılır (Amin ve ark., 2012).

Bir SCADA sistemi bir dizi işlemde bilgi toplar, bilgileri merkezi bir siteye geri aktarır, bir olay meydana geldiğinde ev istasyonunu uyarır, olayın kritik olup olmadığını belirleme ve bilgileri görüntüleme gibi gerekli analiz ve kontrolleri mantıklı ve organize bir şekilde (Haijing ve ark., 2006 ; Hadžiosmanović ve ark., 2012). Aslında, çoğu kontrol işlemi otomatik olarak uzak terminal birimi (RTU) veya programlanabilir lojik kontrolörler (PLC) tarafından gerçekleştirilir. Örneğin, bir PLC endüstriyel bir işlemin bir parçası aracılığıyla soğutma suyu akışını kontrol edebilir, ancak SCADA sistemi operatörlerin akış için ayar noktalarını değiştirmesine ve akış kaybı ve yüksek sıcaklık gibi alarm koşullarını etkinleştirmesine, görüntülenmesine ve kaydetmesine izin verebilir (Lakhoua ve Jbira, 2012). Proses endüstrisinde SCADA sistemlerinin devreye alınması, yüzlerce hatta binlerce kontrol döngüsü ile proseslerin kontrolünü ve izlenmesini kolaylaştırmıştır. Bununla birlikte, yalnızca son on yılda, Atıksu kalite standartları gittikçe katılaşmaktadır, büyük ölçekli kentsel atıksu arıtma tesislerinde benzer denetim kontrol sistemleri bulunmaktadır (Sanchez ve ark., 2003). Su ve Atıksu uygulamalarında, SCADA genellikle uzak ve/veya yerel algılama birimlerini izleme verilerine koordineli yanıtların başlatılabileceği merkezi bir izleme konumuna bağlamak için uygulanır.

Ancak, SCADA sistemlerde saldırı tespiti için çeşitli teknikler ve algoritma çalışmaları derin öğrenme teknikleri araştırma sahası olarak görülmektedir (Kor, 1974 ; Libhaber ve Orozco-Jamamillo, 2012). Saldırı tespit algoritmalarının büyük bir ailesi, kural tabanlı algoritmalarıdır (Topacık ve Koyuncu, 2006). Ancak gerçek uygulamalarda, anormal durumlarda, sistemin davranışı tahmin edilemez ve bilinen herhangi bir model veya kuralı izlemez. Bu özellik, kural tabanlı algoritmalarda izinsiz girişi tespit edememektedir. Genel olarak, anormallik tespiti destek vektör makineleri, K-en yakın komşu (KNN) ve kabul edilebilir ikili sınıflandırma problemi olarak ve dolayısıyla sinir ağlarını, gizli anormallikleri tespit etmek için kullanılan birçok sınıflandırma algoritması olarak Markov modeli de kullanılabilir. Bununla birlikte, kesin olarak konuşursak, saldırı tespitinin temel nesnesini saptıran ne tür bir anormallik beklediğini bilmeyi gerektirdiklerinden, bunlar saldırı tespit algoritmaları değildir. Bu nedenle SCADA sistemlerde güvenlik riski tespit edilmekte zorlanan bir açıklığı beraberinde



getirmektedir. Derin öğrenme sistemleri genel anlamda eğitilmiş modellerden oluşturulur (Turovskiy ve Mathai, 2005).

Atıksu arıtma tesisleri (AAT), yönetilmesi zor olan ve verimli çalışma için uzun vadeli uzmanlık ve sürekli izleme gerektiren karmaşık dinamik sistemlerdir. Çalışmalar kapsamından da anlaşılacağı üzere, SCADA sistemlerinin atıksu arıtma tesislerinde uygulanmasına öncülük eden kuruluşlar genelde belediyeler ve belediyeler bünyesinde operasyon idarelerini yürüten bağıl su ve kanalizasyon operasyonlarını yürüten birimlerdir. Bu çıkarım genel halk sağlığının yanı sıra, halkın maliyet analizleri ve enerji tasarrufunun belediye bütçesi ile bütünleşik olduğunu ve otomasyon sistemlerinin kullanılması ile halkın hem sağlık hem de gelir kaynaklarının akıllıca idare edilmesinde yükümlülüğünü emanet ettiği belediye kuruluşlarının çağa ayak uydurma zorunluluğunu gözler önüne sermektedir. SCADA sistemlerin kullanım alanları farklılık göstermektedir. Bu sistemler çoğunlukla elektrik sistemlerinde denetleyici olarak konumlandırılrsa da buzlanma çalışmaları, santrallerin kontrolü, endüstriyel denetim ve önceki bölümlerde de bahsedildiği üzere Atıksu arıtma sistemlerinde kendilerine yer bulurlar.

SCADA sistemin atıksu arıtma tesislerinde kullanılması sistemin verimliliği ve çok yönlülüğüyle birlikte mevzuat kapsamına uygunluğu yönüyle de önemlidir. Aynı zamanda bu sistemler sayesinde toplanan büyük verileri işleyebilmek ve o verileri yapay zeka yöntemleri ile yapılacak çalışmalarda kullanabilme açısından oldukça önemlidir. Bu bağlamda bir sonraki bölümde SCADA sistemin atıksu arıtma tesislerindeki kullanımına esas çalışma prensiplerine yer verilecektir.

### **3.3. SCADA Otomasyon Sisteminde Kullanılan PLC**

1969 yılında ilk uygulaması elektrik rölelerinde kumanda görevi görmek üzere tasarlanan PLC, Kontrol teknolojisindeki en yaygın donanım platformu programlanabilir lojistik kontrolör olarak SCADA otomasyon sisteminin yazılımı uygulama kısmıdır. Bir PC güç kaynağı, bir kontrol işlemci CPU, bazı merkezi dijital giriş ve çıkış modülleri ve dâhili bir veri yolu sisteminden oluşur. Gerektiğinde analog

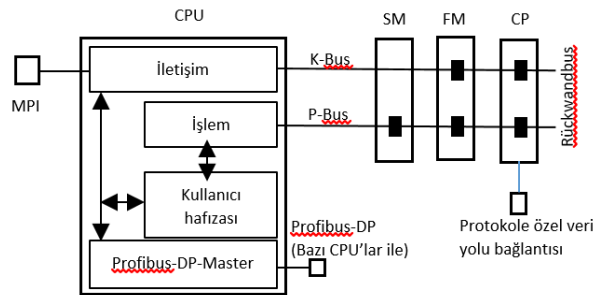
değer işleme veya kontrolörler ve konumlandırma gibi özel işlevler için modüller de eklenebilir. Çevresel modüller ve programlama dilleri kontrol teknolojisinin ihtiyaçlarına yöneliktir.

PLC kontrolü PC işletim sisteminden bağımsızdır ve PC, süreç kontrolüne ek olarak, tipik bir veri toplama veya veri işleme görevi (süreç görselleştirme, veri tabanı, üretim verilerinin depolanması vb.) sunduğu düşük maliyet ile tercih edilebilirliğini arttırmaktadır. PC tabanlı kontroller işlem verilerini bağlı bir fieldbus sistemi (örn. PROFIBUS-DP) üzerinden alır (Miklovic ve Liptak, 2002). Bir programlama cihazı olan PC yansıtma, PLC' nin donanım modülleri için konfigürasyon verileri ile kod blokları ve veri bloklarında bulunan kontrol komutları ve kontrol verilerinden oluşan bir kullanıcı programı oluşturabilmeyi sağlar.

Ayrıca, sembol bağlantıları ve program hakkında yorumlar ve iletişim bağlantıları hakkında bilgiler de bu programlama sırasında PLC ara yüzüne işlenir. CPU bilgisayar işletim sisteminin güçlendirilmesi için kullanılan bir tür cihazdır ve makine öğrenmesi, derin öğrenme, YSA, ESA gibi tüm makine öğrenmesi alt dallarında olduğu gibi PLC'de programlanabilmesi için PC işletim sisteminin güçlendirilmesine ihtiyaç duyulmaktadır (Sağiroğlu ve ark., 2003).

PLC sistem ara yüzü, kullanıcı belleği, sistem belleği ve kontrol ünitesi birlikte CPU merkezi modülünün işlem bölümünü oluşturur. İşlem kısmı, bir PLC'nin kullanıcı programını işleme, çevre modüllerine erişme ve tüm program yürütme sistemini izleme / yönetme gibi klasik işleme görevlerini üstlenir. MPI programlama ara yüzünün çalışması ve bir akıllı otomasyon cihazı PLC, içindeki akıllı, iletişim yeteneğine sahip çevresel modüllerle veri trafiği için ek bir iletişim parçası sağlanmaktadır ve bu duruma ilişkin örnek Şekil 3.1.'de gösterilmektedir. Çevresel veri yolu (p veri yolu), arka duvar veri yolunun bir parçasıdır ve CPU ile sinyal modülleri arasındaki veri trafiğini yönetir. P-bus mono bir ana veri yoludur, yalnızca CPU veri trafiğini başlatabilir (Bailey ve Wright, 2003).

İletişim veri yolu (K-büst), arka panel veri yolunun bir parçasıdır ve iletişim kapasitesine sahip FM modülleri (hızlı sayılar, kontrol, konumlandırma için fonksiyon modülleri) ve CP'ye (fieldbus sistemlerinin bağlantısı için iletişim modülleri) veri trafiğini devralır. Böylelikle sistemdeki iletişim veri akışı daha fazla miktarda veri alışverişini için optimize edilmiş olur (Brodsky ve Radvonovsky, 2013).



Şekil 3.1. PLC'nin sistem mimarisini (Bailey ve Wright, 2003)

MPI = Çok Noktalı Ara yüz, CPU = merkezi modül, SM = sinyal modülü, FM = fonksiyon modülü, CP = iletişim modülü, K-Bus = iletişim veri yolu, P-Bus = çevresel veri yolu PLC, OPC, RTU ünite ve teknolojilerini barındıran SCADA Sistemleri; Çeşitli alanlarda tamamen veya kısmen otomatik hale getirilebilen bir üretim sürecinin mevcut durumu hakkında uç nokta cihazları olarak bilinmektedir. Süreç kontrol ortamlarındaki tipik BT sistemleri, üretim süreci hakkında bilgi toplar. PCS, DCS, SCADA ve benzerleri, otomatik mantık alarmları ve operatör etkileşimi gerektiren olaylara dayalı olarak değerleri okur ve etkileşime girer.

PLC bu anlamda yeniden programlanabilir olması açısından uzman sistemlerin güçlü yönlerinden birini temsil etmektedir. Kullanımı son derece yaygın olan bu sistemlerden PLC ve SCADA sistemleri atıksu arıtma tesislerinde sıklıkla kullanılan sistemlerdir.

### 3.4. SCADA Sisteminin Temel Yapısı

Genel anlamda SCADA mimarisini bir haberleşme kanalı yoluyla uzak ekipmanı kontrol etmeyi ve izlemeyi amaçlayan uygulamaları açıklar (Stamp ve ark., 2003).

Genelleştirilmiş bir SCADA çerçevesi ve mimarisi oluşturma girişimleri olmuştur (Nader ve ark., 2016).

SCADA sistemlerde, veriler bir ağ üzerinden olayları işleyen bir kontrol sunucusuna aktarılır. Sensör verilerinin belirlenen sınırları aşması durumunda SCADA sistemi üzerinden kullanıcıyı bilgilendirir. Bunun aksine, SCADA çerçevesi oluşturan bileşenleri tanımlayan derinlemesine bir standart tanımlamaktadır. IEA standardına göre, sistem uzak bir sahaya ve ana istasyona ayrılmıştır. SCADA sistemlerde, internet, bir cihaz sunucusuna bağlı saha cihazlarından oluşan bir veri tabanı oluşturur. Toplanan bilgiler, ana istasyondaki gerçek zamanlı ve geçmiş veritabanında saklanır. Ana istasyonun bileşenleri arasındaki iletişimde dahili bir iletişim protokolü kullanılır. Ana Terminal Birimi (MTU), veritabanlarında depolanan verilerle etkileşime giren araçlar, gerçek zamanlı veritabanındaki değişikliklere tepki veren bir olay işleyicisi olarak bilinirler (Bailey ve Wright, 2003).

SCADA saha cihazlarının davranışını değiştirebilen bir cihaz yöneticisi, kullanıcıların izleme kurallarını ayarlamasına ve kuralların çiğnenmesi durumunda kullanıcıyı bilgilendirmesine olanak tanıyan bir alarm yöneticisi, saklanan verilerin analizini sağlayan bir veri sisteminde depolanmasına ve uzak sitenin grafik sunumunu sağlayan bir GUI veya yeniden programlanmasına imkân tanır. Her SCADA sisteminde çok sayıda saha cihazı kullanılmaktadır. Bir saha cihazında sensörler, anahtarlar ve aktüatörlerden verileri anlar ve toplar. Bu nedenle saha cihazları, Programlanabilir Mantık Denetleyici (PLC) veya Uzak Terminal Birimi (RTU) adı verilen bir hesaplama cihazından yararlanır.

Bir ağ arayüzü veya bir İnsan Makine Arayüzü (HMI) aracılığıyla, bağlı sensörlere, anahtarlara veya aktüatörlere yapılandırma ve erişime izin verir. Dahili bir bilgisayar, bağlı sensörlerden, anahtarlardan veya aktüatörlerden gelen sinyalleri dijital verilere veya tam tersi şekilde dönüştüren algoritmik bir kodun çalışmasına yol açar.

Saha cihazlarında çalışan kod iki kategoriye ayrılır: Örnekleme ortalamayı içerebilecek sensörler için izleme döngüleri ve durumu kontrol eden durum

diyagramları aktüatörler gibi çıkış cihazlarının kod, bağlı saha cihazlarını doğrudan izlediği ve kontrol ettiği için gerçek zamanlı gereksinimlere sahiptir. Saha cihazları güvenilir olacak şekilde tasarlanmıştır ve genellikle yedek G / Ç kartları biçiminde yedek güç sistemleri gibi yedek ekipmanları da içeren bir çalışma biçimini benimsemektedir. Böylelikle arızalanan bir G / Ç kartı olduğunda, yedek G/Ç bağlıysa, aygıt otomatik olarak yedek G / Ç'ye bağlanır ve devreden çıkmaz aksi durumda, bir giriş kartı arızalanırsa, kart değiştirilene kadar sinyaller kaybolur. SCADA ve Saha Cihazları kullanımıyla, büyük ölçekli sistemleri çok büyük mesafeleri kapsayan gaz boru hatları gibi ucuz ve verimli bir şekilde izlemek ve kontrol etmek mümkündür. Bir saha cihazı ve SCADA sistemi arasında veri aktarmak için iletişim protokolleri kullanılır. Protokoller yoklama veya tanımlanmış –eğer.... – ise..... olacak biçimde algoritmik kodlamada yer alan senaryodaki olaya dayalıdır. Yoklama protokolleri, her aktarım arasındaki süreye yoklama aralığı adı verilen zamanlanmış bir döngüde veri aktarır. Olay güdümlü protokoller sensör değişimindeki verileri aktarır. Yaygın olarak kullanılan yoklama protokolleri arasında Modbus ve Profibus bulunmaktadır (Huising ve ark., 2008).

### **3.5. SCADA Sisteminin Katmanları**

Bağlamanın üç temel başlıkta ayırt edilebilir ve yönetimin nasıl uygulanacağı açısından farklı birimlerin birlikte kullanımı da mümkündür. Genel anlamda, yerel kontrol, merkezi kontrol, dağıtılmış kontrol birimleri olarak adlandırılan denetleme sistemlerinde; yerel kontrol tipi, kontrolör işleminin yerel olarak yapılır. Bununla birlikte sensörleri ve aktüatörlerin yönetimini sağlar (Galitsky, 2019). Basit SCADA sistemlerde kontrol parametrelerini ayarlamak ve işlemin çeşitli parametrelerini izlemek için kullanılabilen bir insan makine ara yüzü (HMI veya PC ekranı), süpervizör seviyesine bağlantı, bir parametreyi başlatmak ve bitirmek veya değiştirmek için basit komutlar ve durum bilgisi ve alarmları göndermek için kullanım parametrelerinin ayarlanmasını sağlamaktadır (Cheung, 2007). İkinci tip denetleyici sistemlerde merkezi kontroller için, sensörler ve aktüatörler, MTU, HMI ve diğer tüm ünitelerin bulunduğu kontrol odasına bağlantı sağlayan sistemler olarak tanımlanır. Burada makine öğrenmesi mimarilerinin avantajı, operatörün tüm sistemi tek bir

yerden izleyebilmesi ve yönetebilmesi ve acil durumlarda hızlı tepki verebilmesine imkan tanır (Annie, 2012). Dezavantajı ise programlanın yapılmasındaki yüksek fiyattır, çünkü yönetilen süreçteki sensörlere ve aktüatörlere olan bağlantıların çoğaltılması ve ana kontrolör sistemin kazalara karşı sürdürülebilirliği için öğelerinin dikkate alınmasını esas kılar. Bu PLC yönetim biçimleri geçmişte kullanılmıştır, ancak şimdi en çok başvurulan sistemler daha gelişmiş seviyede önsezi yeteneği sunan derin öğrenme yöntemiyle elde edilmiş model ile programlamadır. Bu sistem bir çeşit uzman sistem yöntemi olarak SCADA sistemlerde uygulanmaktadır (Moris ve Gao, 2014).

Dağıtılmış kontrol yöntemi sunan derin öğrenme mimarileri diğer iki sistemin avantajlarını birleştirdiği için en yaygın kullanılan sistem biçimlerini temsil etmektedir. Denetim gerçekleştiren denetleyiciler yönetilen sürece yakındır, ancak bir ağ tarafından bir merkeze bağlanır (Shawartz ve ark., 2014). Sistem içindeki birleşenler derin öğrenme mimarilerinin kullanılması ile önsezi yeteneğinin kullanımına olanak sağlamaktadır. Atıksu arıtma sistemlerinde, elektrik sistemlerinde kullanılan SCADA sistemler kadar geniş bir siber saldırı düzenlemesi olmasa da sistem programlandığı zararlı etkiye karşı alarm üreterek kullanıcıyı uyarılmaktadır. Şimdiye dek makine insan ara yüzü olarak ifade edilen ekranlar bir PC ekranını tanımlamaktadır. Makine insan ara yüzü, diğer bir deyişle HMI kırmızı ekran ile üretilen alarmı karşılamaktadır (Adamsky ve ark., 2018).

## BÖLÜM 4. İLERİ SEVİYE KONTROL VE OTOMASYON SİSTEMİ OLARAK BİR SCADA SİSTEM TASARIMI, PLANLAMA VE UYGULAMASI, ATIKSU ARITMA TESİSİ MODERNİZASYONU ÖRNEĞİ

### 4.1. Sakarya'da Faaliyet Gösteren Atıksu Arıtma Tesisleri

SAKARYA - SASKİ Genel Müdürlüğüne bağlı olarak faaliyet gösteren 6 adet (Karaman, Akyazı, Hendek, Geyve, Karasu ve Kocaali) Atık Su Arıtma Tesisine ilişkin veriler Tablo 4.1'de gösterilmiştir. Atıksu arıtma tesislerine ait görsele Şekil 4.1'de yer verilmiştir. Merkezde bulunan Atıksu Arıtma Tesisi'nin kapasitesi 198.800 m<sup>3</sup>/G olup 269.000 m<sup>2</sup> alana kurulmuştur. 1 m<sup>3</sup> Atıksuyun arıtılması için ortalama 0,28 kW enerji tüketilmiş ve arıtım maliyeti ise ortalama 0,44 ₺ olarak hesaplanmıştır.

Tablo 4.1. Sakarya'da bulunan Atıksu Arıtma Tesisleri

TESİS	Karaman AAT	Akyazı AAT	Hendek AAT	Geyve AAT	Karasu AAT	Kocaali AAT
Eşdeğer Nüfus (kişi)	1.000.000	60.000	65.000	66.000	100.000	30.000
Kapasite (m <sup>3</sup> /G)	198.800	13.970	12.970	12.513	17.160	2.080
Kurulu Alan (m <sup>2</sup> )	269.000	16.150	51.000	39.900	38.360	13.266
Arıtma Sistemi	Uzun Havalandırmalı Aktif Çamur	Klasik Aktif Çamur	Klasik Aktif Çamur	Uzun Havalandırmalı İleri Biyolojik Aktif Çamur Sistemi	Uzun Havalandırmalı İleri Biyolojik Aktif Çamur Sistemi	Uzun Havalandırmalı İleri Biyolojik Aktif Çamur Sistemi

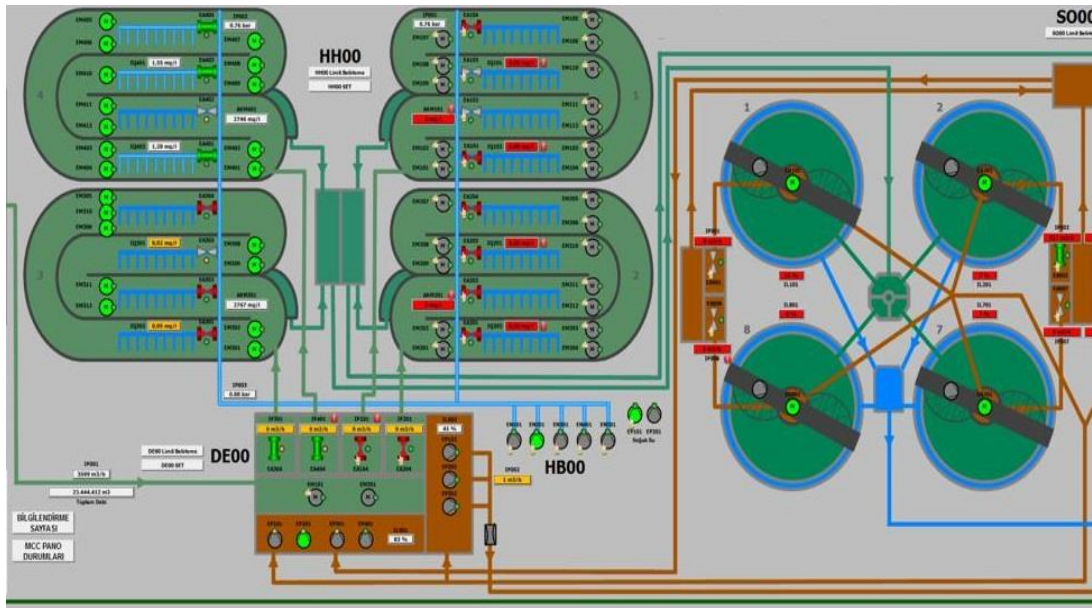


Şekil 4.1. Atıksu Arıtma Tesisi Örneği

#### 4.2. Atıksu Arıtma Tesisi Modernizasyonunda Kullanılan Materyaller

SCADA programında ana ekran üzerinde hava sağlayıcılar, giriş terfi, geri devir ve fazla çamur pompaları, fiziksel arıtma ekipmanları, çöktürme havuzları köprüleri ve teleskopik vanaları, Atıksu Giriş İzleme Sistemi verileri ve kapalı ortam gaz değerleri gibi ekipmanların izlenmesi ve uzaktan kontrol edilmesine olanak sunmaktadır. Ekranda ekipmanlara ait açık-kapalı durumlarının animasyonla sunulmasının yanı sıra renksel uyarı yapılarak operatörün dikkatinin yoğunlaştırılması sağlanmaktadır. Fiziksel arıtma üniteleri, tesise gelen Atıksuyun seviyesine göre, zamana göre ve manuel olarak çalışmaktadır. Seviye Sensörlerinde yaşanabilecek arızalara karşı kendinden şamandıralı pompalar kullanılmaktadır. Sensörlerden alınan kanal ve haznedeki Atıksu doluluk oranları SCADA ana ekranına sayısal olarak sunulmaktadır. Ayrıca, kapalı ortamlardaki  $H_2S$  (Hidrojen Sülfür) ve  $CH_4$  (Metan) gazları konsantrasyonları SCADA (Uzaktan İzleme ve Kontrol Sistemi) ekranları üzerinden görülebilmektedir. Herhangi bir ekipmanın arıza veya bakım durumlarında sistem üzerindeki bakım kilidi vasıtasıyla sahadan ve SCADA (Uzaktan İzleme ve Kontrol Sistemi) ekranı üzerinden çalışması engellenebilmektedir (Şekil 4.2.).

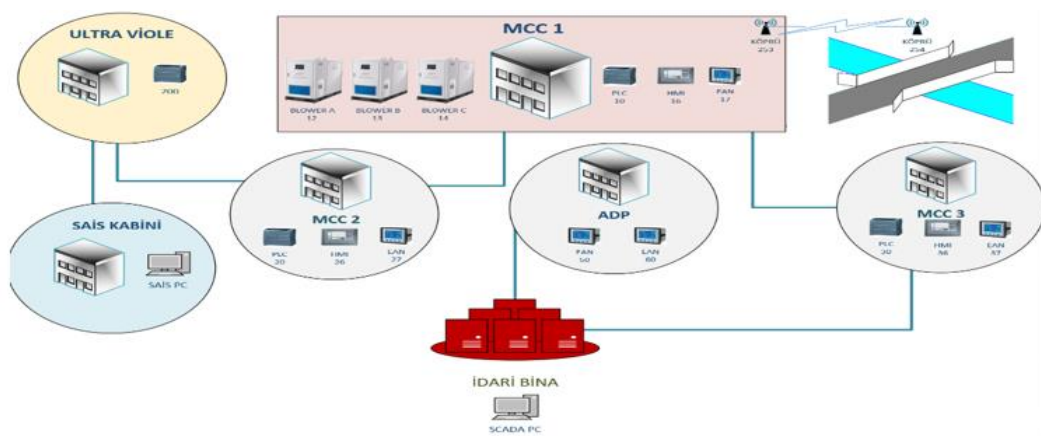




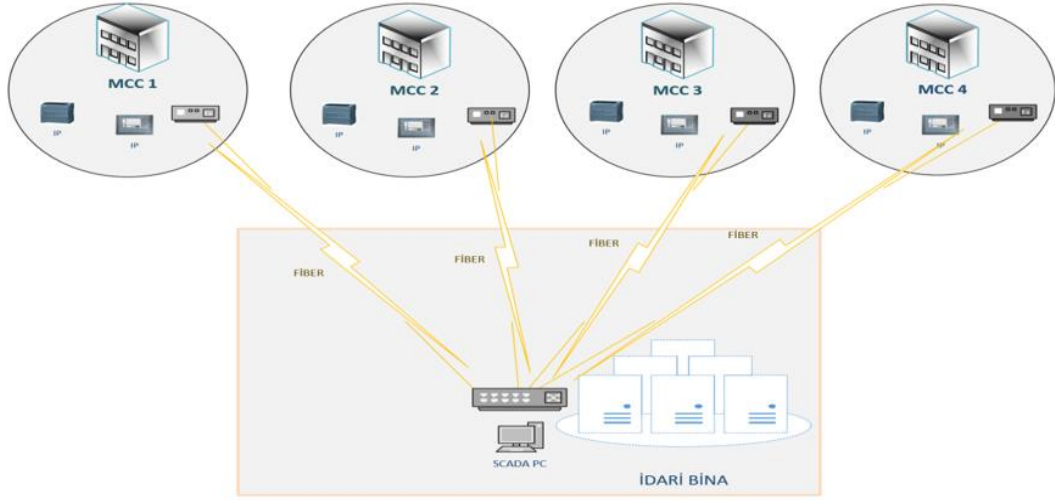
Şekil 4.2. Atıksu Arıtma Tesisi SCADA Kontrol Ekranı Görşeli

#### 4.2.1. SCADA sistemi güvenlik kontrol paneli

Ağ iletişim alt yapısında Ağ Erişim Kontrolü (Network Access Control) yazılımı kullanılmakta olup tesislerde bulunan panolarda dışarıdan fiziksel olarak yapılan girişleri engellemek amacıyla panolar arasında akıllı switchler takılmaktadır. Bu switchler, tanımlı olmayan ve dışarıdan giriş yapan tüm yabancı cihazların verilerini engelleyerek merkeze uyarı bilgisi geçmektedir.



Şekil 4.3. Atıksu Arıtma Tesisi Network Topolojisi



Şekil 4.4. Atıksu Arıtma Tesisi Network Topolojisi

Önceden tanımlanmış ve yetkilendirilmiş olan kullanıcılar sisteme kayıt edilmektedir. Dışarıdan izinsiz bağlantı yapılma durumunda rapor ekranında cihazlara bağlantı yapan veya erişmeye çalışan tüm cihazlar görüntülenerek kayıt altına alınır.

İletişim ağı ana dağıtım istasyonu noktadan noktaya iletişim kurmakta olup şehrin birçok bölümünü kapsamaktadır. Bu sistem sayesinde hiçbir operatör bağlantısı olmadığı için dışarıya internet çıkışı sağlanmamakta ve dışarıdan yapılacak olan erişimler engellenmektedir (Şekil 4.3.), (Şekilde 4.4.).

Ayrıca yaşanabilecek herhangi bir siber saldırıya karşı makineler yedeklenmektedir. Herhangi bir olumsuzluk yaşanması durumunda sistem kapatılıp yedek makine ile değiştirilmekte olup eski makine ve giriş yapılan kullanıcılar incelemeye alınmaktadır. Yetkilendirilmiş personeller sadece izleme yapabilme veya müdahale edebilme olarak iki sınıfa ayrılmaktadır.

#### 4.2.2. Modernizasyonun tamamlanması ve SCADA sistemin devreye alınması

Yapılan çalışmalarda Atıksu arıtma tesislerinde ekipmanların PLC (Programlanabilir Lojik Kontrol) kontrollü ölçüm cihazları ve seviye sensörlerinden alınan verilere göre çalışması için SCADA (Uzaktan Kontrol ve Gözleme Sistemi) yazılımı belirlenen senaryolara göre tamamlanarak sistem devreye alınmıştır. Gerçekleştirilen çalışma

tesislere; prosesin optimum şartlarda işletilmesi, sistemin kontrol fonksiyonlarının artırılması ve tesislerin işletme maliyetlerinin düşürülmesi gibi birçok avantaj sağlamıştır.

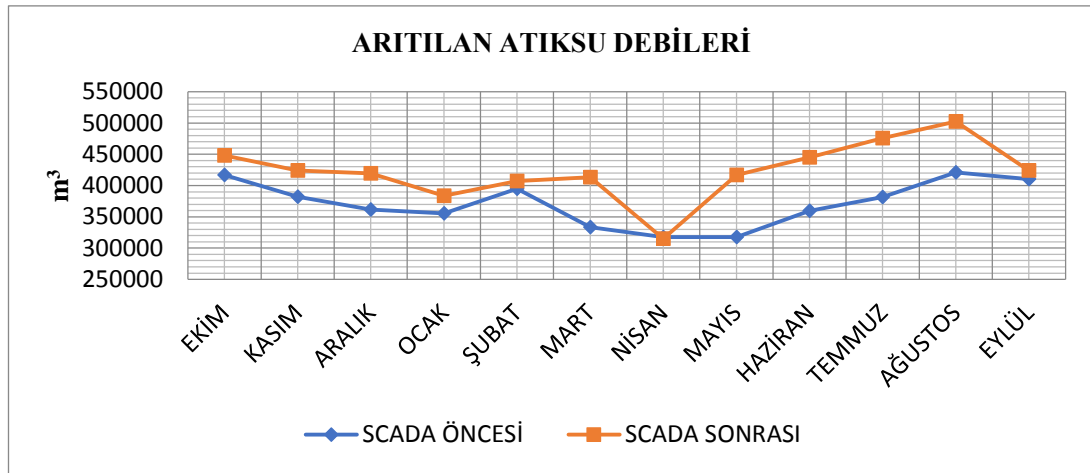
Ayrıca sistem üzerinden arızaların algılanabilmesi, yerlerinin tespiti ve uzaktan kumandalı olarak belli bir merkezden giderilebilmesi, sistemle ilgili alarm sinyallerinin operatörleri uyaracak şekilde oluşturulması ve görüntülenmesi, çeşitli verilerin geçmişe dönük olarak veri tabanı şeklinde saklanabilmesi, gece görüşlü kamera sistemi ile tesislerin giriş ve çıkış sularının yirmidört (24) saat görüntülenebilmesi ile daha etkin ve daha ekonomik işletme gerçekleştirilmiştir.

### **4.3. SCADA Otomasyon Sisteminin Tesise Kazandırdığı Avantajlar**

SCADA sistemine geçiş ile birlikte işgücü, makine, donanım, enerji ve zamanın doğru kullanımıyla verimlilik artmakta ve bununla doğru orantılı olarak maliyetler düşmektedir. Ayrıca yapılacak performans çalışmaları, Makine Öğrenmesi çalışmalarımız için yol gösterici olacaktır. Bu sistem ile kişilere bağlı kalımsızın sistem hakkında toplanan verilere dayalı ayrıntılı bilgi edinilmesi sağlanmıştır.

#### **4.3.1. Ekipman açısından SCADA sisteminin faydaları**

Atıksu arıtma tesislerinde, nüfus artışlarına bağlı olarak arıtılan atıksu miktarlarında artış söz konusudur. Atıksu arıtma tesisinin SCADA sistemi sonraki dönemde önceki döneme oranla arıtılan atıksu miktarında %14 lük bir artış meydana gelmiş olup arıtılan atıksuda artış olmasına rağmen işletme giderlerinde düşüş meydana gelmiştir (Şekil 4.5.).



Şekil 4.5. Arıtılan Atıksu Miktarları Karşılaştırılması Debi Artışı

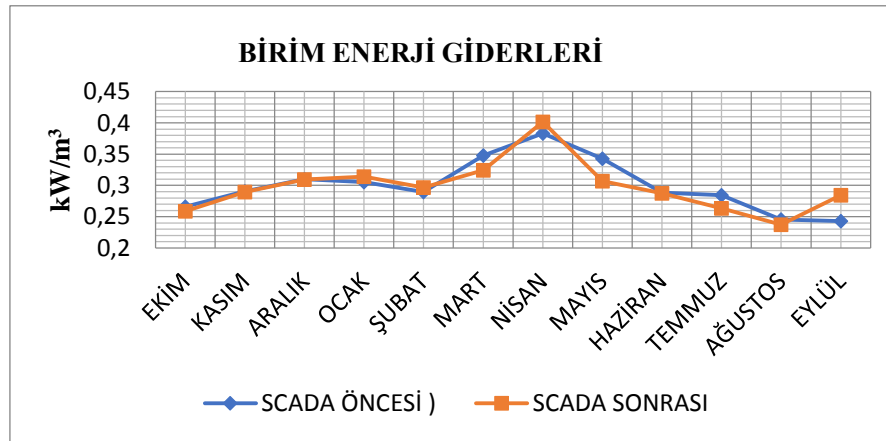
Yedekli çalışan ekipmanlara eş yaşlandırma algoritması uygulanarak, motor ömrünün uzatılması ve bakım periyodunun daha da uzaması sağlanmıştır. Yedekli çalışan ekipmanların ve hava sağlayıcıların havuzdaki oksijen seviyelerine ve belirlenen limit değerlerine göre frekans konvertör üzerinden hızlarının kademeli olarak artırılıp azaltılarak bu ekipmanların elektriksel ve mekanik aksamalarının korunması ve enerji tasarrufu sağlanmıştır.

Mekanik ekipmanların çalışma zamanlarının SCADA sistemi üzerinden görülebilmesi ile ekipmanların bakımlarının zamanında yapılması sağlanarak bakım eksikliklerinden kaynaklanacak arızaların önüne geçilmiştir. Fiziksel arıtma ekipman ve enstrümanlarının gelen Atıksu seviyesine göre çalışması sağlanarak ekipmanların sürekli çalışmasının önüne geçilmiş, bu sayede ekipmanların kullanım ömürleri uzatılmış ve enerji tasarrufu sağlanmıştır.

#### 4.3.2. Proses açısından SCADA sistemin faydaları

Ekipmanların SCADA sisteminde belirlenen limit değerlere göre frekans konvertör yardımıyla çalışması ile enerji giderlerinde düşüş meydana gelmiştir. Online ölçüm cihazları [pH (Hidrojen Konsantrasyonu), iletkenlik, AKM (Askıda Katı Madde), debi, seviye, TOC, TN, TP, RTC (Gerçek Zamanlı Kontrol Sistemi) vb.] ile sahada ölçülen değerler baz alınarak frekans konvertörler yardımıyla motor hızlarının ayarlanması ve

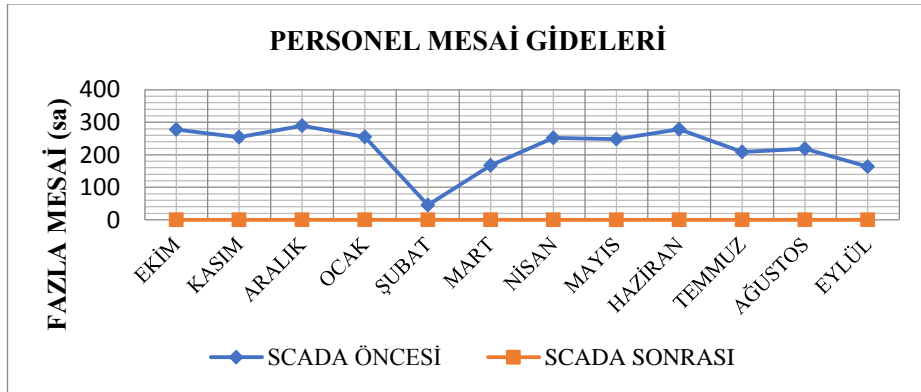
kontrol fonksiyonlarının arttırılması ile proses kontrolünün optimum şartlara getirilmesi sağlanmıştır. Giriş suyundaki Toplam Organik Karbon (TOC) değerini ölçen TOC analizöründe ölçülen değerlere göre tesislerin giriş suyu kirlilik değerleri online olarak ölçülmektedir. Sisteme girilen senaryolara göre pik kirlilik geldiğinde hava sağlayıcıların kapasitesi arttırılarak prosesin güvenliği sağlanmaktadır. Giriş suyunda belirlenen limit değerler aşıldığında operatöre gerekli bildirimler sms/mail yoluyla yapılmaktadır. Havalandırma havuzlarına montajı yapılan RTC (Gerçek Zamanlı Kontrol) sistemi ile biyolojik arıtmanın nitrifikasyon – denitrifikasyon süreleri bu problemlerin ölçtüğü  $\text{NH}_4\text{-N}$  (Amonyum Azotu),  $\text{NO}_2\text{-N}$  (Nitrit Azotu),  $\text{NO}_3\text{-N}$  (Nitrat Azotu) değerlerine göre otomatik olarak belirlenmektedir. Bu ölçüm sonuçları neticesinde hava sağlayıcıların kapasitesi sistem üzerinden ayarlanabilmektedir. Bu sayede gereksiz enerji sarfiyatı engellenerek prosesin verimliliği arttırılmaktadır. Şekil 4.6.'da görüldüğü üzere SCADA sistemine geçmeden önce ile SCADA sistemine geçtikten sonra yapılan karşılaştırmalarda  $\text{m}^3$  başına %0.7'lik kW azalması gerçekleşmiştir.



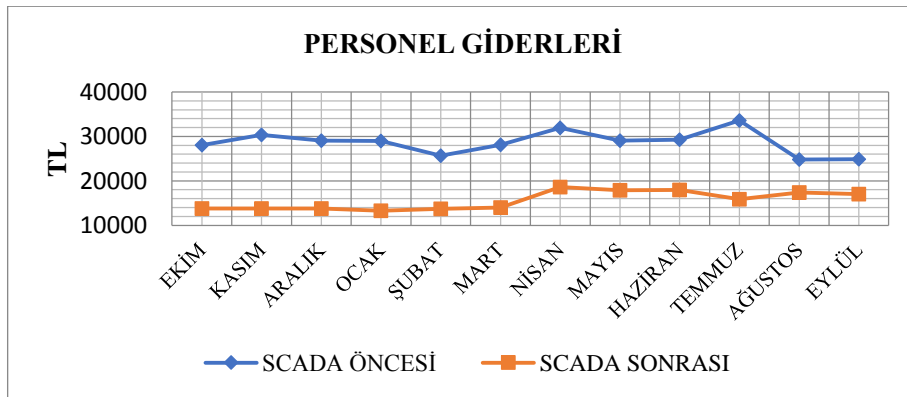
Şekil 4.6. Birim Enerji Miktarları Karşılaştırılması  $\text{m}^3$  Başına Azalma %0,7

SCADA sistemi ile tesislerin tek bir ekrandan kontrol edilebilir olması, ekipmanların çalışma, durma, manuel, otomatik ve arıza durumları, vana ve kapak açıklık durumları, seviye, sıcaklık ve sistem durumlarının tek bir ekran üzerinden gözlemlenebilmesi, gerekli müdahaleleri uzaktan yapabilmesi, sistem üzerinden belirlenen limit değerler aşıldığında bildirim mesajı gelmesi, sistemi kontrol eden operatörler için kolaylık sağlayacaktır. Ayrıca SCADA kontrol ekranında ekipmanlar üzerinde yer alan bakım

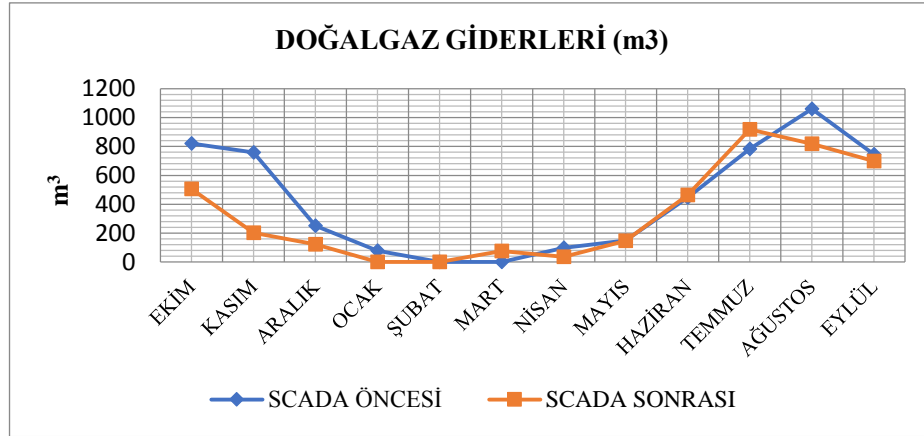
kilidi butonu vasıtasıyla ekipmanın bakım ya da arıza durumlarında SCADA sistemi üzerinden çalışması engellenmiş olup ekipmanlara ait hareket, yanıp sönme, uzak veya yakın çalışmaya göre renk değişimi gibi operatörlerin ilgisini çekip kullanım için kolaylık sağlayabilecek nitelikleri barındırmaktadır. Kapalı ortamlardaki zehirleyici ve boğucu özelliklere sahip hidrojen sülfür ve metan gazlarını ölçen sabit tip gaz dedektörlerinin ölçtüğü değerlerin SCADA ekranında görülebilmesi, bu ortamlarda yapılacak çalışmalar öncesinde ortamda bulunan gaz konsantrasyonlarına bağlı olarak gerekli iş güvenliği önlemlerinin alınması sağlanmıştır. SCADA sistemi ile tesislerin uzaktan kontrol ve müdahalesinin sağlanması ile tesislerde hafta sonu ve gece vardiyası çalışmaları ile resmi tatil günlerindeki çalışmalar kaldırılmıştır. SCADA sistemi öncesi ile SCADA sistemi sonrasında Şekil 4.7.'de de görüldüğü üzere personel mesai karşılaştırmalarında %100 azalma gerçekleştirilmiş ve Şekil 4.8.'de görüldüğü üzere personel maliyetlerinde ise %45,6'lık azalma gerçekleştirilmiştir.



Şekil 4.7. Personel Fazla Mesai Karşılaştırılması %100 Azalma



Şekil 4.8. Personel Maliyetleri Karşılaştırılması %45,6 Azalma



Şekil 4.9. Birim Doğalgaz Miktarları Karşılaştırılması

Doğalgaz birim m<sup>3</sup> giderleri şekil 4.9’da gösterilmiştir. Tesis işletme değerlerinin tarihsel ve gerçek zamanlı olarak sistem yazılımında tutulması, geçmişe dönük verilerin istenen zaman aralıklarında raporlanması açısından kolaylık sağlamıştır.

#### 4.4. Modernizasyon Amacıyla Yapılan Yatırım Sonrası Maliyet Kıyaslama

Tablo 4.2. Maliyet Kıyaslama

GİDER	BİRİM	SCADA Öncesi (11 Ay)	SCADA Sonrası (11 Ay)	Yıllık Kazanım (₺)
Cihazların Bakım ve Kimyasal Giderleri	₺	*	-15,000	-15,000
Enerji	₺	396,989.00	422,566.00	3,382.096*
Enerji	kW	1.318.218	1.491.145	
Personel	₺	344,024	187,187	156,837
Doğalgaz	₺	5.942	4,426	1,516
Arıtılan Atıksu	m <sup>3</sup>	4.450.845	5.074.969	-
kW/m <sup>3</sup>		0.296	0.294	-
₺/kW		0.301	0.284	
<b>TOPLAM</b>	<b>₺</b>	<b>746,955.00</b>	<b>599,179.00</b>	<b>146,735.10</b>

SCADA Sonrası arıtılan Atıksu miktarı : 5.074.969 m<sup>3</sup>

SCADA Öncesi m<sup>3</sup> başına enerji tüketim : 0,296 kW

SCADA Sonrası m<sup>3</sup> başına enerji tüketim : 0,294 kW/m<sup>3</sup>

SCADA Sonrası m<sup>3</sup> başına enerji bedeli (yaklaşık) : 0,284 tl/kW

Modernizasyon amacıyla yapılan yatırım sonrası ve yatırım öncesi giderler üzerinden maliyet kıyaslama yapılmış ve Tablo 4.2.'de gösterilmiştir. SCADA sonrası arıtılan Atıksu miktarını artırmak için gerekli olan SCADA öncesi olması tahmin edilen enerji tüketimini hesaplarken SCADA sonrası arıtılan Atıksu miktarı ile SCADA öncesi kW/m<sup>3</sup> değeri ile çarpılır.

$$5.074.969 * 0,296 = 1.503.066,381 \text{ kW/yıl}$$

Eğer SCADA Sistemine geçilmeseydi arıtılan 5.074.969 m<sup>3</sup> Atıksu için 1.503.066,381 kW enerji harcanacaktı. Fakat SCADA sonrası arıtılan Atıksu miktarına göre tüketilen enerji miktarı: 1.491.145 Kw'dır.

SCADA öncesi tahmin edilen Tüketim - SCADA Sonrası Tüketim

$$1.503.066,381 - 1.491.145 = 11.921,381 \text{ kW'dır.}$$

Elde edilen bu rakamı SCADA sonrası ortalama tl/kw değeri (0,284) ile çarpar isek yıllık olarak enerjiden elde edilen kar bulunmuş olur.

$$11.921,381 * 0,284 = 3.382,096 \text{ TL kar elde edilir.}$$

Sonuç olarak 25.577,00TL fazla enerji tüketilmiş gibi görünse bile arıtılan Atıksu miktarının artması sebebi ile SCADA öncesiyle oranladığımızda yıllık kazancımız 3.382,096 TL kar elde etmiş oluruz. SCADA sistemi öncesi ve sonrasındaki maliyetler kıyaslandığında ise yıllık giderler ortalama olarak çıkarıldığında toplamda yıllık olarak 146.735,10 ₺' lik bir kazanım görülmektedir. Otomasyona geçiş için yapılan yatırımların toplamı ₺486,106.00' dir. Başabaş Noktası (Amorti süresi) olarak  $486.106,00 / 146.735,10 = 3,31$  yıl (39,72 ay) hesaplanmıştır. Tablo 4.3.'te bileşik faiz tablosu ile Denklem 4.1 ve Denklem 4.2'de net büküm değer hesapları ve formülleri görülmektedir.

$$\text{Net Büküm Değer (NBD) Hesabı; } P=A \left[ p = A \left[ \frac{(1+i)^n - 1}{i \cdot (1+i)} \right] \right] = P=A(p/A; \%i; n) \quad (4.1)$$

P: Şimdiki Değer,

A: Eşit Değer,



İ: faiz,

N: Dönem

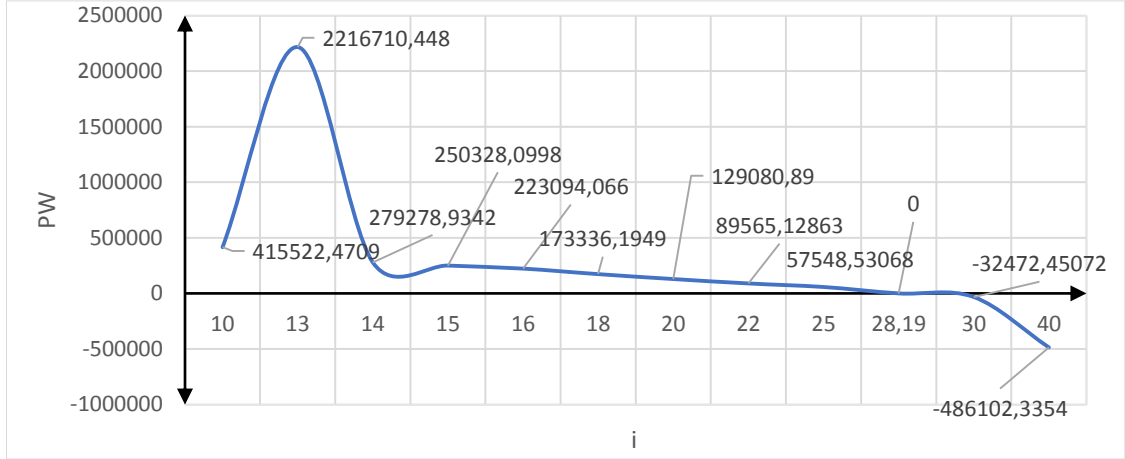
$$\text{NBD (Net Büküm Değer)} = -486106 + A(P/A; i; n) \quad (4.2)$$

Tablo 4.3. Bileşik Faiz Tablosu

(n, i) değeri	(P/A) değeri
n=10 yıl, İ=%10 için	6.1446
n=10 yıl, İ=%13 için	18.4197
n=10 yıl, İ=%14 için	5.2161
n=10 yıl, İ=%15 için	5.0188
n=10 yıl, İ=%16 için	4.8332
n=10 yıl, i=%18 için	4.4941
n=10 yıl, i=%20 için	4.1925
n=10 yıl, i=%22 için	3.9232
n=10 yıl, i=%25 için	3.705
n=10 yıl, i=%30 için	3.0915
n=10 yıl, i=%40 için	2.4136

Tablo 4.4. Net Büküm Değer Hesap Tablosu

		Net Büküm Değer (NBD)					
n=10 yıl, İ=%10 için	-486,106.00	+	₺146,735.10	*	6.1446	=	415,522.4709
n=10 yıl, İ=%13 için	-486,106.00	+	₺146,735.10	*	18.4197	=	2,216,710.448
n=10 yıl, İ=%14 için	-486,106.00	+	₺146,735.10	*	5.2161	=	279,278.9342
n=10 yıl, İ=%15 için	-486,106.00	+	₺146,735.10	*	5.0188	=	250,328.0998
n=10 yıl, İ=%16 için	-486,106.00	+	₺146,735.10	*	4.8332	=	223,094.066
n=10 yıl, i=%18 için	-486,106.00	+	₺146,735.10	*	4.4941	=	173,336.1949
n=10 yıl, i=%20 için	-486,106.00	+	₺146,735.10	*	4.1925	=	129,080.89
n=10 yıl, i=%22 için	-486,106.00	+	₺146,735.10	*	3.9232	=	89,565.12863
n=10 yıl, i=%25 için	-486,105.00	+	₺146,735.10	*	3.705	=	57,548.53068
n=10 yıl, i=%30 için	-486,104.00	+	₺146,735.10	*	3.0915	=	-32,472.45072
n=10 yıl, i=%40 için	-486,103.00	+	₺0.28	*	2.4136	=	-486,102.3354



Şekil 4.10. Net Büküm Değer Hesap Grafiği

Çalışmada birleşik faiz hesabı ve net büküm değer hesabı kullanılarak ulaşılan sonuçlar neticesinde başabaş noktası hesaplanmış ve ardından enterpolasyon hesabı yapılmıştır (Tablo 4.4. ve Şekil 4.10.).

Yapılan yatırımda kullanılan sermayenin başka yatırım araçlarından biri olan banka faiz oranlarına göre kıyaslaması yapılmış, dönemin en yüksek faizini veren bankanın faiz oranı dikkate alınmış ve yapılan bu yatırımın diğer yatırım araçlarına göre daha avantajlı olduğu görülmüştür (Denklem 4.3).

$$\frac{(\%i - 25)}{(35 - 25)} = \frac{(0 - 57548,53068)}{(-32472,4507 - 57548,53068)} = \%28,19 \quad (4.3)$$

$$\text{İRR } i = \%28,19 > \text{MARR } (\%12,25)$$

Yapılmış olan yatırım ile işletme kolaylığı, veri havuzu, anlık takip ve sürdürülebilir faydaları olması dışında ekonomik olarak değerlendirmeye alındığı zaman da kabul edilebilir yatırım olarak görülmektedir.

Yapılan çalışmalar sonucunda Atıksu arıtma tesislerinde iyileştirme işlemi gerçekleştirilmiş, motivasyon ve verimlilik artışı elde edilmiş, minimum kaynak ile darboğazlar, arızalar, hatalar ve problemler tespit edilmeye başlanmıştır.

## **BÖLÜM 5. ATIKSU ARITMA TESİSLERİNDE BİYOLOJİK OKSİJEN İHTİYACININ (BOİ<sub>5</sub>) TAHMİNİ İÇİN KULLANILAN METOTLAR**

### **5.1. Metod/Yöntem**

#### **5.1.1. Bulanık Hızlı Azaltım algoritması (BHAA)**

Sürekli ve artan bir hızla oluşan veri kümeleri elde edilen verilerin etkin şekilde kullanılmasını gerektirmiştir. Günümüzde karar alma amacı ile ihtiyaç duyulan verilerin elde edilmesinde büyük veri tabanları önemli araçlar olarak kullanılmaktadır. Bu durumda karmaşıklık ve belirsizlik içeren büyüklükler ile karşılaşmaktadır. Veri tabanları üzerinden elde edilen verilerin analiz edilerek karar destek sistemlerinde kullanılması herhangi bir araç kullanma gerekliliğini meydana getirmiştir. Bu sebeple akıllı veri işleme yöntemleri ortaya çıkarılmıştır. Bulanık mantık fazla miktarda elde edilen verilerin karmaşıklık ve belirsizliği göz önünde bulundurulduğunda verilerin işlenmesi için matematiksel bir niceliği ifade etmektedir. Bu durumda bulanık işlemler belirsiz ve karmaşık durumların çözümü için kullanılan matematiksel işlemlerdir. Lofty A. Zadeh bulanık mantığı; herşeyin ve doğrunun bir derece meselesi olduğunu belirterek akıl yürütmek amacı ile geliştirilmiş bir model olarak tanımlamış ve bulanık mantık kavramının temelinde sözcük hesaplama anlamının yer aldığını vurgulamıştır (Şevki, 2008). Bulanık mantık kavramının en belirgin özelliklerinden biri elde edilen sonuçların geçerliliğinin kesin olmamakla beraber yaklaşık olarak çıkarım kurallarına sahip olmasıdır. Bir önerme bulanık mantık kavramı ile  $\mu(a) = 1 - \mu(\sim a)$ , ve  $\mu(\sim a) = 1 - \mu(a)$  olarak ifade edilmektedir. Bulanık mantık kavramının temel elemanı üyelik fonksiyonları ile ifade edilen bulanık kümelerdir. Bulanık küme (fuzzy sets) farklı üyelik ya da ait olma derecelerine sahip elemanlar barından bir kümedir. Bulanık küme içerisinde bulundurduğu elemanların her birine 0 ve 1 arasında değerler atayan ve bir üyelik fonksiyonu ile karakterize edilen küme türüdür. Bu doğrultuda kümeye dahil

olmayan elemanların değeri 0 olarak atanırken, kümeye tam dahil olan elemanlara atanan değer 1 olmaktadır ve kümeye dahil olma durumu belirli olmayan elemanlara ise belirsizlik durumuna göre 0 ile 1 arasında değer atanmaktadır (Telçeken ve Doğan, 2004). Buna göre özellikle bulanık kümeye dahil olma durumu belirsizlik gösteren eleman sayısının ve niteliklerinin fazla olması verilerin analizi için farklı araçlar kullanılması gerekliliğini beraberinde getirmiştir. Günümüzde veri tabanlarında saklanan bilginin hızla artış gösteriyor olması büyük veri setlerinin analiz edilmesini gerektirmektedir. Bulanık modelleme teknikleri büyük veri setlerinin analiz edilmesinde yetersiz kalmaktadır. Bulanık küme tekniklerinin etkinliğini arttırmak amacı ile kaba küme yaklaşımı temelli algoritmalar kullanılmaktadır. Kaba kümeler (rough sets); eksik, yetersiz veya belirsiz verileri düzenleyerek veri analizi için yeterli hale getirilmesini sağlayan araçlardır (Riza ve ark., 2014). Bulanık-Kaba küme teorisi verileri analiz etmek için ek parametreler kullanılmasını gerektirmemektedir ve özellikle su kalitesinin analiz edilmesinde kullanım uygunluğu göstermektedir (Jensen ve Shen, 2002). Bulanık- Kaba küme teorisi bir veri setindeki bulanık küme yaklaşımından türetilen, kaba bir kümenin genellemesidir ve “nesnelerin belirlenen parametrelere göre hangi kümelere yerleştirilmesi gerektiği denklem 5.1 ve 5.2’deki gibi tanımlanır;

$$\mu_{RX}([x]R) = \inf\{\mu_X(x)|x \in [x]R\} \quad (5.1)$$

$$\mu_{RX}([x]R) = \sup\{\mu_X(x)|x \in [x]R\} \quad (5.2)$$

$\mu_X(x)$ , x’in bulanıklığa ait olduğu dereceyi belirtmektedir.  $\langle RX, RX \rangle$  kaba bulanık küme olarak adlandırılmaktadır ve belirtilen tanımlar alt ve üst yaklaşımlar ile özdeş hale gelmektedir (Jensen ve Shen, 2002). Bulanık-kaba kümeler tüm eşdeğerlik sınıflarının bulanık olabildiği durumlarda kullanılabilir ve bu durum karar değerleri ile koşullu değerlerin tümünün bulanıklık özelliği göstermesi gerektiği anlamına gelir. Bu durumda alt ve üst yaklaşımlar denklem 5.3 ve 5.4’teki eşitlikle ifade edilebilir;

$$\mu_X(F_i) = \inf_x \max\{1 - \mu_{F_i}(x), \mu_X(x)\} \quad \forall i \quad (5.3)$$

$$\mu_X(F_i) = \sup_x \min\{\mu_{F_i}(x), \mu_X(x)\} \quad (5.4)$$

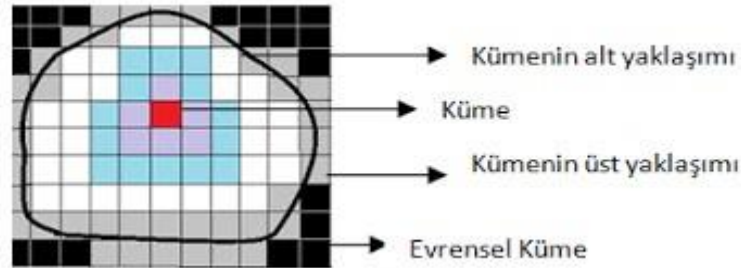
$F_i$ , tek bir bulanık eşdeğerlik sınıfını belirtmektedir ve  $\langle X, X \rangle$  belirsiz-kaba küme olarak adlandırılır (Chouchoulas, ve Shen, 2001). Verilerin analizinde kaba küme uygulamalarının kullanımı söz konusu olduğunda belirsiz çok fazla verinin azaltılması gerekmektedir. Bu durumda azaltılmış hesaplamanın gerçekleştirilmesi amacı ile Hızlı Azaltım Algoritması (QRA) kullanılmaktadır. Hesaplamalarda yaşanan karmaşıklık ve hafıza gereksinimleri göz önünde bulundurulduğunda ele alınan bir veri kümesinin tüm olası alt kümelerinin hesaplanmasında zorluklar yaşanmaktadır. Bu sorunun çözümü için mevcut bilgi sisteminde yer alan veri setini temsil edecek minimum nitelikteki indirgenmiş nitelikler kümesinin elde edilmesine ihtiyaç duyulmaktadır. QRA algoritması işleme mekanizması doğrultusunda veri kümelerine ait alt kümeler bir ağaç geçişi üzerinde ele alınır ve ağacın her bir düğümü bir koşullu niteliği belirtmektedir. QRA algoritması burada indirgeme işlemi için kullanılmaktadır. QRA algoritmasının işleyişi şu şekilde gerçekleşmektedir; Q alt kümeler için boş bir öznitelik sunmaktadır. R bağımlılıktaki değişiklikleri temsil etmektedir (Sai ve Rao, 2009). QRA algoritmasının işleme mantığı şu şekildedir; kaba kümelerde yer alan DT, U bir dizi nesneden oluşan karar tablosunu, C koşullu öznitelikler kümesi ve D karar öznitelikleri kümesini belirtmektedir. Belirli bir kavram için  $X \subseteq U$ , BX daha düşük ve üst yaklaşımı göstermektedir.  $B \subseteq C$  nitelikleri kümesine göre (D) B pozitif bölgeyi belirtirken,  $R \subseteq C$  kümesi için (D) ,  $R\gamma$  bağımlılık ölçüsünü göstermektedir. “QRA algoritması (C, D);  $R = \Phi$  do  $T = R$  for each  $x \in (C - R)$  if  $( ) ( ) \gamma R U \{x\} D > \gamma T D$   $T = R U \{x\}$  endfor  $R = T$  until (D) (D)  $R \subseteq C \gamma = \gamma$  (10) return şeklinde işlem gerçekleştirmektedir (Mafarja ve Mirjalili, 2019). Kaba-küme algoritması farklı algoritmalarla hibrit şekilde kullanılarak gerçekleştirilen uygulamanın etkinliğinin artırılmasına olanak tanımaktadır. Bu durumla ilgili olarak gerçekleştirilen bir çalışmada ilgili veri setinden olabilecek en az bilgi kaybı ile belirgin özelliklerin çıkarımı için etkili bir algoritma elde etmek amacı ile bulanık-kaba küme algoritması, oluşturulan hibrit sistem için kullanılan diğer algoritmayı (CEBARKCC) destekleyici olarak işlem yapmaktadır. Bulanık-küme algoritmasının oluşturulacak hibrit sistem için kullanılmasının sebebi özellik çıkarım işleminin daha hızlı gerçekleşmesini sağlamaktır. Önerilen hibrit yaklaşım sınıflandırma problemleri için özellik çıkarımında kullanılmıştır ve sonuçlar bulanık-kaba küme temelli hibrit algoritmanın

sınıflandırma doğruluğu açısından daha iyi performans gösterdiğini ortaya koymuştur (Pawlak ve Skowron, 2007).

Bu çalışmada oluşturulan veri kümeleri üzerinden optimum sonuç verecek verilerin, elde edilmesi ve veri kümelerinin boyutlarının azaltılması amacı ile Fuzzy OuickReduct Algorithm kullanılarak veri kümelerine ait alt kümelere indirgenme işlemi uygulanmıştır.

### 5.1.2. Bulanık–Kaba Örnek Seçimi (FRIS) algoritması (BKÖSA)

Kaba Küme (Rough Sets) Teorisi ilk kez, Zdzislaw I. Pawlak tarafından öne sürülmüştür (Turgay ve Torkul, 2018). Kaba kümeler alttan ve üstten yaklaşan biçimsel olarak tanımlanabilen kümeleri ifade etmektedir (Şekil 5.1.).



Şekil 5.1. Kaba Kümelerin Genel İşletim Mantığı (Turgay ve Torkul, 2018)

Kaba Kümeleme teorisinin biçimsel tanımı şu şekildedir;  $I = (U, A)$ ,  $B = \{b \mid b \in A\}$ ,  $IND_I(B) = \{(x, x') \in U^2 \mid \forall a \in B, a(x) = a(x')\}$ ,  $\underline{B}X = \{x \mid [x]B \subseteq X\}$ ,  $\overline{B}X = \{x \mid [x]B \cap X \neq \emptyset\}$ ,  $BR_B(X) = \overline{B}X - \underline{B}X$  (Güleç, 2013). Kaba Kümeleme'nin biçimsel tanımı doğrultusunda I: bilgi sistemini, U: Boş küme olmayan sınırlı nesne kümesini, A: Boş küme olmayan sınırlı özellik kümesini, B: Boş küme olmayan sınırlı özellik kümesi-koşul özelliklerini, X: Tanımlanmak istenen kümeyi,  $IND_I$ : Bilgi kümesi üzerindeki ayırt edilemezliği,  $\underline{B}X$ : X'i B özellik alt kümesi ile alttan tanımlayan kümeyi,  $\overline{B}X$ : X'i B özellik alt kümesi ile üstten tanımlayan kümeyi ve  $BR_B(X)$ : X'in B özellik alt kümesi ile tanımlanan sınır kümesini temsil etmektedir.

Kaba kümeleme yönteminin kullanılmasındaki temel amaç baz alınan veri kümesi üzerinden bilgi çıkarımının sağlanmasıdır. Buna göre kaba küme teorisi fazla sayıda belirsiz verinin matematiksel olarak çözümlenmesi ile verimli şekilde kullanılabilir hale getirilmesinde kullanılmaktadır. Ancak belirsiz durumların matematiksel olarak çözümlenmesinde kaba kümeleme teorisinin kullanımının yetersiz kaldığı durumlar için yeni modellerin keşfine ihtiyaç duyulmaktadır. Bulanık-Kaba Küme teorisi bu durum üzerine geliştirilmiştir ve belirsiz veri kümelerinden optimum fayda sağlanacak alt veri kümelerinin elde edilmesinde kullanılmaktadır. Bulanık ve kaba kümelerin entegre şekilde kullanılması amacı ile alttan ve üstten yaklaşan veri kümelerine ait formüllerin bulanıklaştırılmasına odaklanılmıştır. Bu durumda A kümesinin X'teki bir bulanık kümede genelleştirilmesi ile nesnelerin belirli bir kavrama ait olması ve nesnelere temsil eden yaklaşık eşitlik değerlerinin ölçülmesi olmak üzere 2 temel ilke baz alınmıştır (Huang ve ark., 2013).

Bulanık-Kaba Örnek Seçimi (FRIS) Algoritmasının temel çalışma mantığı şu şekildedir; veri setinde yer alan indirgenecek nesnelere a parametresi olan bulanık benzerlik ölçüsünde kullanılmaktadır ve ihtiyaç duyulan verilerden fazlasını elemek için  $\tau$  parametresi kullanılmaktadır. FRIS Algoritması verilere ait üyelik derecelerini X pozitif bölgeye göre değerlendirerek verilerin elenmesi işlemini gerçekleştirmektedir. Böylelikle hangi sınıfa ait olduğu belirsiz olan nesnelere için değerlendirme işlemi ile ihtiyaç duyulan veri seti elde edilmektedir (Şekil 5.2.) (Jensen ve Cornelis, 2010).

FRIS (S, a,  $\tau$ )

S: Azaltılacak Nesnelere Kümesi

a: taneciklik Parametresi

$\tau$ : seçim eşiği

1.  $Y \leftarrow S$
2. foreach  $x \in S$
3. if ( $P_{OS\alpha, S A}(x) < \tau$ )
4.  $Y \leftarrow Y - \{x\}$
5. return Y

Şekil 5.2. Bulanık-Kaba Küme Seçim Algoritması Çalışma Mantığı (Verbiest, 2014).

Kaba küme teorisini bulanık küme teorisinden ayıran en önemli fark veri setine ilişkin eksik bilgiler olması durumunda modellemeye olanak tanınmasıdır. Daha açık bir ifade ile bulanık küme teorisi belirsiz durumlar için modelleme uygulamalarının gerçekleştirilmesinde kullanılırken, kaba küme teorisi eksik bilgiler ışığında da modelleme uygulamaları için kullanılabilir. Bulanık kaba model alt ve üst tanımların bulanıklaştırılması ile edilmektedir. Bir  $X \in U$  elemanı için koşulun net yaklaşımı  $\forall y \in U : (x, y) \in R \rightarrow y \in A$  şeklinde ifade edilmektedir. Bu durumda eşdeğerlik bağıntısı  $R$  bulanık bir bağıntıya dönüşmektedir ve  $A$  artık bulanık bir kümedir. Değerler  $R(x, y)$  ve  $A(y)$  bulanık bir çıkarım elde edilmesi adına  $I$  ile birbirine bağlıdır. Bu sebeple  $I(R(x, y), A(y))$   $x$  ile benzerlik gösteren elemanların ne ölçüde  $A$  kümesine ait olduğunu belirtmektedir. Bu doğrultuda  $X \in U$  elemanının üyelik değeri düşük yaklaşım göz önünde bulundurulduğunda yüksek ise  $I(R(x, y), A(y))$  tüm  $y \in U$  için yüksek bir değer belirtecektir ve bu durum  $\forall x \in U : (R \downarrow A)(x) = \min_{y \in U} I(R(x, y), A(y))$  bağıntısı ile ifade edilmektedir. Bulanık-Kaba küme teorisi karar sistemlerinin oluşturulmasında özellikle kullanılmaktadır. Sistem  $(U, A \cup \{d\})$  olmak üzere,  $U$  evreni örneklem kümesini temsil etmektedir. Ayırt edilemezlik ilişkisi  $R$  koşullu öznelilikler aracılığı ile modellenmektedir.  $A$ : koşullu özneliliklerin yüksek oranda ilişkili olduğu  $x$  ve  $y$  örnekleri, yüksek üyelik derecesi  $R(x, y)$  ya da  $R$  için eşdeğer durumdadır. Bu doğrultuda karar nitelikleri göz önünde bulundurularak kavramlar alttan ve üstten yaklaştırılacak karar sınıflarına karşılık gelmektedir (Verbiest, 2014). Buna göre bulanık-kaba küme teorisi karar sistemlerinin modellenmesi için belirsiz ve eksik bilgileri barındıran veri setleri üzerinden işlem yapılmasını sağlamaktadır. Bulanık-kaba küme teorisinin uygulanması sırasında sınıflandırma veya regresyon problemlerinin çözümü için K-En Yakın Komşu Algoritması (KNN) kullanılmaktadır. KNN algoritması uygulanması kolay ve gözetimli makine öğrenmesi algoritmalarından biridir. KNN algoritması `CNNClassif(form, dat, dist = "Euclidean", p = 2, Cl = "smaller")` şeklinde kullanılmaktadır (Aydilek, 2018). Burada form tahmin problemi için açıklayıcı formül, dat veri setini içeren veri çerçevesi, dis alttan ve üstten yaklaşım için mesafe ölçüsünü kullanan bir karakter dizisi, p 'P-norm mesafesinin seçilmesi durumunda p'nin değerini belirten sayı ve Cl en önemli sınıfları belirten karakter vektörüdür.

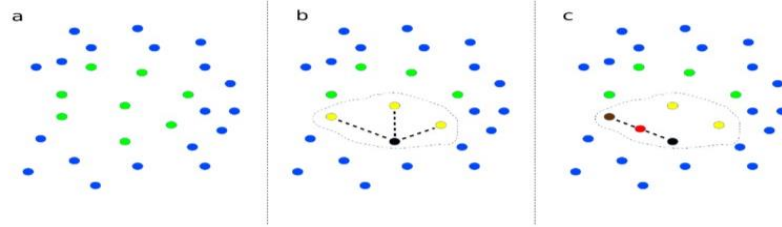


KNN algoritması, sınıfları belirli olan bir veri setindeki verileri kullanarak işlem yapmaktadır. Veri setine katılacak olan yeni verinin, mevcut verilere alttan ve üstten uzaklığını hesaplayarak, k sayısına göre yakın komşuluğunun hesaplanmasını sağlar. Bulanık-kaba kümelerde sınıflandırma problemlerinin çözümü için KNN algoritmasının kullanıldığı bir çalışmada; alt ve üst yaklaşımların örnekleme dahil edilecek elemanların hangi kümeye dahil olacağına ilişkin üyelik olasılığı belirlenirken daha kesin sonuçları verdiğini ortaya koyulmuştur (Tsang ve ark., 2016). Bulanık- Kaba modellerin uygulanması durumunda tahmin yöntemleri, sınıflandırma algoritmaları, en yakın komşu sınıflandırması algoritmaları, karar ağaçları, doğrusal regresyon ve olasılık modelleri kullanılmaktadır. Aynı zamanda sıralama tabanlı ve OWA tabanlı olmak üzere bulanık-kaba küme modelleri oluşturulmuştur (Vluymans, 2019). Bu çalışmada verilere ait özelliklerin bağıllığının tespit edilmesi ile verilerin etkin şekilde tahmin edilerek daha iyi sonuçlar elde edilmesi amacı ile gürültülü örneklerin elenmesi için Bulanık-Kaba Küme Örnek Seçimi (FRIS) Algoritması kullanılmıştır.

### 5.1.3. SmoteR yöntemi

Sınıflandırma işlemi veri setlerinin analiz edilmesinde temel görevlerden biri olarak karşımıza çıkmaktadır. Ve sınıflandırma işleminin gerçekleştirilmesinde en fazla karşılaşılan sorun dengesiz veri setleri üzerindeki işlem yapılmasıdır. Teoriden bağımsız olarak gerçek yaşamda dengesiz veri setleri ile karşılaşılması daha olası bir durumdur. Bu sebeple dengesiz veri setlerinin analiz edilmesinde makine öğrenmesi yöntemleri kullanılmaktadır. Dengesiz veri seti (İmbalanced Data Set); sınıf dağılımlarının birbirlerine yakın olmadığı durumlarda görülmektedir ve sorun problemin fazla olduğu sınıfın, daha az problemlili sınıfı domine etmesinden kaynaklanmaktadır. Bu sebeple oluşturulan modelin çoğunluğa sahip sınıfa yakınlık göstermesinden kaynaklı olarak azınlık sınıf kötü şekilde sınıflandırılmaktadır. Dengesiz veri setleri ile karşılaşılması durumunda doğru analiz yapılması ve dengenin sağlanması amacı ile uygulanabilecek çeşitli yöntemler vardır. SMOTER yöntemi regresyon tahmini için kullanılan bir yöntemdir ve SmoteR Algoritması dengeli olmayan veri setleri içerisinde azınlıkta olan sınıf değerine ait kayıt sayısının sentetik olarak artırılması işlemini

gerçekleştirerek dengeli veri kümelerinin oluşturulmasında kullanılmaktadır (Aydilek, 2018).



Şekil 5.3. SmoteR Algoritmasının Grafiksel Temsili (Schubach ve ark., 2017)

Smoter algoritmasının grafiksel gösterimi Şekil 5.3.'te verilmiştir. SmoteR Algoritması sentetik azınlık örnekler üretme işlemini gerçekleştiren bir aşırı örnekleme metodudur. SmoteR Algoritmasının çalışma prensibi şu şekildedir;

1. Azınlık sınıfına ait her gözlemin k yakın komşusunun aranması işlemi gerçekleştirilir.
2. Azınlık sınıfına ait gözlem ile k yakın komşusu (kNN) olan gözlem arasındaki farkın hesaplanması işlemi gerçekleştirilir.
3. (0,1) arasında rastgele bir sayı ( $\alpha$ ) seçilerek, azınlık sınıfına ait gözlem ile k yakın komşu (kNN) gözlem arasındaki fark ile çarpılması işlemi gerçekleştirilir.
4. Eşitlik 2.15 kullanılarak yeni sentetik gözlem elde edilir.  $X_{yeni} = x_i + (x_j - x_i) * \alpha$
5. İstenilen sayıda sentetik gözlem oluşturulması için 1. ve 4. Adımın tekrarlanması işlemi gerçekleştirilir (Haklı, 2018).

Smote algoritması dengesiz sınıf dağılımları ile karşılaşıldığı durumlarda sınıflandırma problemlerinin çözümünde kullanılır, SmoteR algoritması ise regresyon problemleri için kullanılan bir örnekleme yöntemidir. Temel SmoteR Algoritması şu şekildedir;

```

function SmoteR(D, tE, o, u, k)
// D - A data set
// tE - The threshold for relevance of the target variable values
// %o,%u - Percentages of over- and under-sampling
// k - The number of neighbours used in case generation
rareL ← {hx, yi ∈ D : φ(y) > tE ∧ y < ỹ} // ỹ is the median of the target Y
newCasesL ← genSynthCases(rareL, %o, k) // generate synthetic cases for rareL
rareH ← {hx, yi ∈ D : φ(y) > tE ∧ y > ỹ}
newCasesH ← genSynthCases(rareH, %o, k) // generate synthetic cases for
rareH
newCases ← newCasesL S newCasesH
nrNorm ← %u of |newCases|
normCases ← sample of nrNorm cases ∈ D \ {rareL S rareH} // under-sampling
return newCases S normCases
end function (Torgo ve ark., 2013).

```

SmoteR yönteminin temel özelliği azınlıkta veri kümelerine eklenen elemanları ve çoğunlukta olan veri kümelerini dengelemektedir. Bu durum SmoteR yönteminin diğer yöntemlere göre avantajlı olduğunu göstermektedir. SmoteR algoritması veri kümelerine ait nadir rastlanan hedef bir değer için sentetik sonuçlar oluşturulması adına aşırı bir örnekleme stratejisi kullanmaktadır. Bu strateji her durum için nadir değerlere gözlem kümesine en yakın elemanın seçilmesidir. SmoteR algoritmasının regresyon problemlerinde uygulanması için üç temel bileşen vardır. Bunlardan ilki hangi durumun ilgili gözlemler için normal olduğunun belirlenmesi, ikincisi yeni sentetik örneklerin hangi yolla oluşturulacağını belirlemek ve üçüncüsü sentetik örneklerin değerlerinin belirlenmesi adına hedef değişkene karar verilmesidir (Branco ve ark., 2013). Gerçekleştirilen bir çalışmada çok dengesiz dağılımlara sahip veri setlerini kullanarak tahmine dayalı modeller oluşturma uygulamalarının geliştirilmesinde regresyon problemlerinin çözümü için SmoteR yönteminin kullanımı önerilmiştir. İlgili çalışmada SmoteR algoritması regresyon görevlerine uyarlanmıştır. SmoteR algoritmasının regresyon problemlerine uyarlanmasının temel nedeni üç temel bileşenin çözüm için optimum sonuç vereceği beklentisidir. Bu doğrultuda belirli

durum için alakalı olma doğrultusunda işlev ve kullanıcı tanımlı setlerin belirlenmesi için eşik kullanılmıştır. Setlerin birine ait gözlemler aşırı örneklenmiş ve diğer sete ait gözlemler yetersiz örneklenmiştir. Bu sorunun çözümlenmesi adına SmoteR algoritması kullanılmıştır (Branco ve ark., 2015).

Tahmine dayalı modellerin oluşturulmasında regresyon için kullanılacak analitik yöntemlerin önerildiği bir çalışmada oluşturulan modellerden elde edilen tahmin sonuçlarının doğruluğunun artırılması için denenen algoritmalarından biri SmoteR algoritmasıdır. SmoteR algoritmasının ilgili çalışmada kullanılmasının nedeni dengesiz veri setlerinde regresyon problemlerinin çözümünde daha doğru sonuçlar almaya yönelik olarak gerçekleştirilmesidir. Çalışmada hangi elemanın hangi veri kümesine dahil edileceğinin belirlenmesi amacı ile SmoteR algoritmasının bileşenleri göz önünde bulundurularak bir alaka düzeyi ile alaka eşiği belirlenmiştir. Bu işlem ile veri kümelerine dahil edilecek elemanlara ilişkin ait olacakları veri kümesi doğrultusunda önem dereceleri belirlenmiştir. Böylece SmoteR algoritması kullanılarak hedef değişken değerlerinin ağırlıklı ortalamalarının hesaplanmasıyla gerçekleştirilmiştir (Branco ve ark., 2016).

Dengesiz veri kümelerinde regresyon işlemi için bir ön işleme yaklaşımı önerilen bir diğer çalışmada hedef değişkenin süreklilik gösterdiği regresyon değerleri için bir ön işleme stratejisi olarak SmoteR algoritması önerilmiştir. SmoteR algoritması genel olarak rastgele az örnekleme işlemi gerçekleştirilmesine imkan tanımaktadır. Bunun yanı sıra nadir rastlanan durum değişkenleri için yeni sentetik örnekler de SmoteR algoritmaları ile oluşturulabilir ve sentetik durumlar enterpolasyon stratejisi aracılığı ile elde edilmektedir. İlgili strateji rastlanan iki farklı nadir durum için kullanılır ve durumlar ilişkin alt ve üst yaklaşımlar baz alınarak sonuç üretilmektedir. İki duruma ait özellikler, hedef değişken değerlerinin ağırlıklı ortalamalarının hesaplanması ile belirlenmektedir. Son olarak tüm durumlara ait örnekler sıralanır ve aşırı /yetersiz örnekleme yüzdeleri tanımlanır (Branco ve ark., 2017).

SmoteR yöntemi dengesiz zaman serilerinin tahmin edilmesinde yeniden örnekleme stratejilerinden biridir. Zaman serisi tahminlerinde ortaya çıkan önemli sorunlardan biri hedef değişkenin dengesiz dağılım göstermesidir. Hedef değişkenlere ilişkin

değerler kullanıcı için önemli olduğundan, ilgili değerlerin tahmin için yeterli doğrulukta olmaması sorun oluşturmaktadır. Zaman serilerinin tahmininde hedef değişkenlerine ilişkin nadir değerlerin belirlenmesi adına SmoteR yönteminin kullanıldığı bir çalışmada dengesiz zaman serisi verilerini kullanarak tahmin doğruluğunu arttırmak amacı ile yeniden örnekleme stratejisi olarak SmoteR algoritması önerilmiştir. 24 zaman serisine ait veri seti üzerinde SmoteR ve diğer standart yöntemler kullanılmış ve alınan sonuçlar SmoteR algoritmalarının tahmin doğruluğunda önemli bir artış olduğunu göstermiştir (Moniz ve ark., 2017). Smote algoritması bu çalışmada regresyon problemlerinde dengeli olmayan veri kümelerinden dengeli veri kümeleri elde etmek amacı ile kullanılmıştır.

#### **5.1.4. Makine öğrenmesi regresyon algoritmaları**

Makine Öğrenmesi, çözüm bulunması amaç edinilen bir probleme ilişkin belli algoritmalar ile bir model tasarlamak ve problem üzerinde test edilmesi sağlanarak, sorunun çözüme kavuşması noktasında sisteme hız kazandıran, sistemi otomatikleştiren, bilgisayar yazılımı ile kodlama sistemlerinin ortak edinimini birlikte sunan bir yapay zeka teknolojisidir. Makine öğrenmesi bir sistemin performansını yükseltmek veya tanımlanan sorun ile ilgili doğru tahminlerde bulunmak için tecrübelerden yararlanabilen hesaplama yöntemleri olarak geniş kapsamda tanımlanabilir. Makine öğrenmesi, tanımlanan soruna ilişkin çözümü çıktı olarak vermesi için çözümü amaçlanan sorunun, oluşturulmuş hazır veri kümesindeki bilgiler ile tasarlanıp eğitilen bir model ile tarayarak çözüme ilişkin bilgileri çıktı olarak sunan yapay zekanın bir alt birimi olarak kabul görür. Makine öğrenmesi, soruna ilişkin oluşturulmuş veri kümesindeki sıralı örüntüleri tespit ederek, en akıllıca kararı vermek amacı ile istatistiksel bilgi, mantık ve algoritmik hesaplamayı bütünleşik kullanabilen çözümlenmesi amaç edinilen konuda tasarlanarak eğitilebilen tanımlanmış algoritmaya göre karar vererek soruna ilişkin çözümü çıktı olarak sunan yöntemdir (Bishop, 2006).

Yapay zekâ ve makine öğrenmesi sistemleri kısaca bir modelin eğitilmesiyle istenen eylemi gerçekleştirmek üzere programlanmış sistemleri sunar. Programlanmış

sistemler saldırı tespit veya anormallik durumlarının tespitinde alarm üreten sistemlerdir. Normal olmayan durumların tespit edilmesi amacı ile Yapay Sinir Ağları (YSA) ve derin öğrenme modelleri ile eğitilen sistemler bir arada kullanılmaktadır (Yang ve ark., 2019). Bu sistemlerle günümüzde sıklıkla karşılaşılan büyük veri setlerinin oluşturulması, veri madenciliği ve ihtiyaca yönelik hizmet sunmak için bilginin nitelikli kullanımını sağlamaktadır. İstendiğinde erişilebilir ve doğru ayrımla kişilerin ihtiyaçlarını karşılayabilen algılama ve karar verme yeteneğine sahip algoritmik bilimsel sistemlerdir (Jain, 2016). Bilginin nitelikli olarak yönetilmesi, ayrılması, karmaşıklıktan kurtulmasını amaçlayan bu hızlı ve deneyimli makine öğrenmesi ve yapay zekâ algoritmalarını kullanarak öğrenilmiş karar verme yeteneği nedeniyle uzman sistem olarak adlandırılır (Zhang, 1994).

Makine öğrenmesi denetimli, denetimsiz ve yarı denetimli makine öğrenmesi biçimindedir. Denetimli öğrenme sınıflandırma yöntemi ve regresyon yöntemi olarak ikiye ayrılır. Regresyon analiz yöntemi, istenilen en az iki veya daha fazla değişken arasındaki ilişkilerin incelenmesini sağlayan güçlü bir istatistiksel yöntemdir. Bir regresyon gerçekleştirme süreci, hangi faktörlerin en önemli olduğunu, hangi faktörlerin göz ardı edilebileceğini ve bu faktörlerin birbirlerini nasıl etkilediğini güvenle belirlenmesini sağlar (Selvi, 2019).

Denetimli öğrenme içerisinde regresyon modellerinde kullanılan 8 adet yöntem vardır. Bu yöntemler; Lineer Regresyon, Lojistik Regresyon, Çoklu Doğrusal Regresyon, Karar Ağaçları, Destek Vektör Makinası, Polinom Regresyonu, Olasılıksal Dereceli Azalma ve Pasif Agresif Regresör'tür. Bu çalışmada kullandığımız denetimli makine öğrenme süreci, harici bir öğrenmeyi içermektedir. "Giriş ve çıkışı belli olan veriler ile öğrenme oluşturulur ve ardından yeni verilere yanıt için uygun tahminler oluşturmak üzere bir model eğitilir. Denetimli öğrenmeyle ilgili önemli bir konu, hata yakınsaması sorunudur, yani istenen ve hesaplanan veri değerleri arasındaki hatanın en aza indirilmesidir (Guo ve ark., 2015). Bu kavramların anlamlarına değinmeden önce makine öğrenmesini incelemek yerinde olacaktır. Makine öğrenmesi temelde, makinenin öğrenmesi istenilen konu hakkında bir veri oluşturarak girdi ve çıktıları belirlenmiş ve aradaki öğrenen ağların hangi girdide hangi sonuçları vereceğine karar

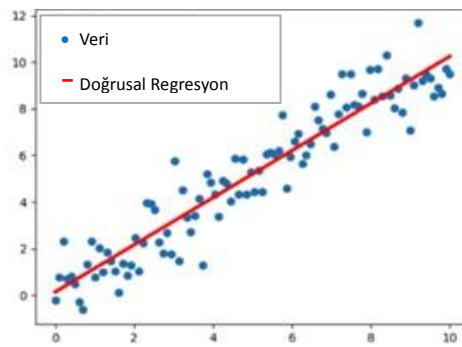
vermesi gerektiği öğretilen sistem biçimini ifade etmektedir (Lantz, 2015). Örneğin biz bilgisayarda bir sınıflandırma yapmak istiyorsak modelimizi oluşturduktan sonra yani algoritmalarla kodlamayı sağladıktan sonra oluşturduğumuz veri seti üzerinde eğitebilir, daha sonra eğitimini tamamlayıp yeni bir görüntüyü sisteme yüklediğimizde sistemin bunu tanıyarak ayırım yapmasını sağlamış yani karar vermesini gerçekleştirmiş oluruz (Jihyun ve ark., 2016). Aynı durum; seslerin ayrılması, haritaların tanınması, sosyal medya kullanımı, internetten yapılan alışveriş ve incelemeler için de daha büyük bir veri ile mümkün hale gelmektedir (Hoque ve ark., 2014).

Yapay zekânın öğrenme biçimlerinden biri olan makine öğrenmesi, daha sonraları insan beyninde yer alan nöronlara benzer bir şekilde tanımlanarak oluşturulmuş derin öğrenme yapay sinir ağlarının oluşmasına dönüşerek, bugün tanımlanmış bir önsezi ile hareket eden otonom sistemlerin oluşmasını sağlamıştır (Feng ve ark., 2017). Bu teknolojik öğrenme biçimi günümüzde en çok kullanılan sistemleri tanımlamaktadır (Pratt, 1994). Örneğin derin öğrenme ile tasarlanan ve dünya satranç şampiyonu Kasparov'u yenen IBM'in Deep Blue isimli cihazı bir satranç uzman sistemidir ve dakikada onlarca hamleyi hesaplayabilmektedir. Aynı şekilde Google tarafından Tenserflow Kütüphanesi kullanılarak tasarlanan Alpha Go isimli oyun da Uzakdoğu'da çokça oynanan Go isimli oyunda başarı sağlayan Go oyununun uzman sistemidir. Oyun haricinde bilgisayarlı görme ve sınıflandırma işlemlerini sağlayan Yapay zekâ tıp alanından, mühendisliğe, insansız hava araçlarına, uydu cihazlarına akıllı otomasyon sistemlerine dek hemen her alanda kullanılmaktadır.

Bu çalışmada Makine Öğrenmesi Regresyon Algoritmaları içerisinde Doğrusal Regresyon, Yapay Sinir Ağları, Bayesyen Doğrusal Regresyon ve Karar Ormanı Regresyonu kullanılmıştır. Bu kısımda kullanılan algoritmalar hakkında bilgi verilmektedir.

### 5.1.4.1. Doğrusal Regresyon (DR)

Çok değişkenli veri kümelerinin oluşturulmasından sonra gerçekleştirilmesi gereken işlem veriler arasındaki ilişkilerin belirlenmesidir. Regresyon bir veya birden fazla değişken arasındaki ilişkiyi belirtir ve bu ilişkinin seviyesi ile kapsamının ölçülmesi işlemi regresyon analizidir. Regresyon teknikleri veri kümelerinin birbiri ile ilişkisinin ortaya konmasını sağlayan yöntemler sunmaktadır ve bu yöntemlerden biri Doğrusal Regresyon (DR) yöntemidir. İki veya daha fazla değişken arasındaki ilişkiyi modellemede kullanılan yöntemlerden en yaygın olanı regresyon analizidir. Regresyon analizi tekli ve çoklu lineer regresyon olarak 2 farklı türde karşımıza çıkmaktadır. Bu doğrultuda “bağımlı değişkeni tahmin etmek için geliştirilen model de giriş olarak bir bağımsız değişken kullanılıyorsa bu tekli lineer regresyon, birden fazla bağımsız değişken kullanılıyorsa çoklu lineer regresyon analizi gerçekleştirilmektedir (Gabrali ve Aslan, 2020). Doğrusal Regresyon kullanımının amacı parametreler arasında fonksiyon olarak elde edilen ilişkinin belirlenmesini sağlamaktır ve doğrusal regresyon en temel makine öğrenmesi modelidir.



Şekil 5.4. Doğrusal Regresyon Modeli

Şekil 5.4.’te görülen “Doğrusal regresyon modeli; bağımlı değişken (y), bağımsız değişken (x), kesişim noktası ( $\beta_0$ ), doğrunun eğimi ( $\beta_1$ )  $\varepsilon$  olmak üzere  $y = \beta_0 + \beta_1 x + \varepsilon$  olarak ifade edilmektedir (Gök, 2017)”. Doğrusal regresyon modelindeki bağımlı değişkenin Y olduğu varsayılırsa, Y ile ifade edilen bağımsız değişkenler arasındaki ilişki için doğrusal bir denklem oluşturulur.  $X_1, X_2, \dots, X_n$ . Bu denklem, tahmin edilmesi istenen bağımlı değişken Y'nin regresyon denklemidir ve Denklem 5.5'deki



gibi ifade edilir (Çatal ve ark., 2019). Bu denklemden,  $b_1, b_2, \dots, b_n$  değerler regresyon katsayılarını temsil eder ve  $\varepsilon$  hata terimidir ..

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n + \varepsilon \quad (5.5)$$

Bu çalışmada parametre tahmininde kolaylık sağlanması, lineer olmayan parametrelerin standart hale dönüştürülmesi ve regresyon tahmini amacı ile Doğrusal (Lineer) Regresyon modeli kullanılmıştır.

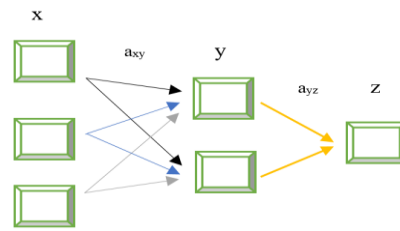
#### 5.1.4.2. Yapay Sinir Ağları (YSA)

Yapay Sinir Ağları (YSA), yapay zeka teknolojilerini oluşturan algoritmaların biridir. YSA teknolojisi, insan beynine ait tekniklerinin taklit edilmesi yolu ile beynin öğrenme, hatırlama, genelleme yapma yolu ile topladığı verilerden yeni veri üretebilme gibi temel işlevlerin gerçekleştirildiği bilgisayar yazılımlarıdır (Öztürk ve ark., 2018). YSA eğitim verileri üzerinden öğrenen; simülasyon, tahmin ve optimizasyon için kullanılabilen veri noktaları arasındaki ilişkileri analiz etmeye olanak tanır. YSA'lar, insan beynine benzeyen bir tür bilgi işleme sistemidir (Ataseven, 2013). Uygun bir eğitim algoritması ve doğru miktarda veri verildiğinde, YSA'lar çok değişkenli doğrusal olmayan problemleri çözebilir (Ayşe ve Berberler, 2017). YSA'lar ayrıca deneysel tasarımlarda su ve Atıksu arıtımı sırasında kirleticileri gidermek için sıklıkla kullanılır (Küçükpehlivan ve ark., 2017).

YSA doğrusal olmayan, paralel şekilde çalışabilen, öğrenme yeteneği olan, hatalara karşı esneklik gösteren, eksik verilerle çalışmasını sağlayan, çok sayıda parametre ve değişken ile işlem yapılabilmesine imkan tanıyan algoritmalarıdır. YSA uygulamaları yaygın olarak tahmin, verilerin sınıflandırılması, veriler arasındaki ilişkilerin belirlenmesi, verilerin yorumlanması ve filtrelenmesi işlemlerinde kullanılmaktadır. YSA'lar, sistemlerin kara kutu temsillerini oluşturmak için değişken ağırlıktaki bağlantılar ile birbirine bağlanan birçok işleme ögesinden, yapay nöronlardan oluşan oldukça basitleştirilmiş modeller kullanır. Her nöron diğer nöronlardan giriş sinyallerini alır, işler ve sonraki nöronlara gönderir (Yurtoğlu, 2005).

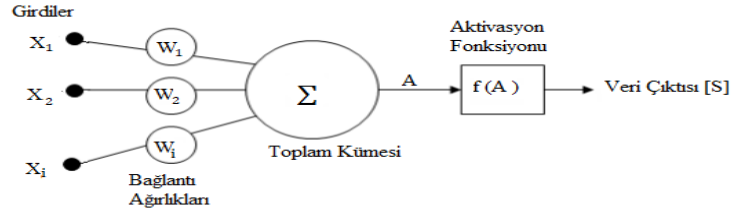
YSA fazla sayıda işleme unsurundan oluşmaktadır. YSA'nın temelinde biyolojik sinir ağlarına benzer şekilde oluşturulmuş hesaplama sistemleri ile işlem yapmaktadır. YSA üç temel bileşen ile karakterize edilmektedir ve bunlar; düğüm, ağırlıklar (bağlantı gücü) ve bir aktivasyon (transfer) işlemidir. YSA modellemesi doğrusal olmayan istatistiksel bir teknikler bütünüdür ve geleneksel istatistiksel ve matematiksel yöntemler ile çözülemeyen problemlerin çözümlenmesinde kullanılmaktadır (ASCE Task Committee, 2000).

YSA modelinde yer alan temel birim düğümdür ve düğümler birbiri ile bağlantılıdır. Bu bağlantılara sinaps adı verilmektedir ve her bir sinaps ile ilişkili bir ağırlık faktörü mevcuttur. YSA genellikle belirli bir girdi kümesi üretecek doğrultuda eğitilmiştir. En yaygın kullanılan YSA üç katmanlı ileri beslemeli yapay sinir ağlarıdır. İleri beslemeli yapay sinir ağlarında her katmanda düğümler bulunur. Giriş katmanında ve iç katmanlarda yer alan her düğüm girdileri kaydeder ve elde edilen değerler bir sonraki katmana aktararak işlem gerçekleştirilir. Bu işlemler ağırlıklara göre yapılır ve ağırlık iki düğüm arasındaki bağlantı gücünü belirtmektedir. Giriş katmanındaki ve çıktıdaki nöron sayısı giriş değerlerine göre belirlenerek sırasıyla çıktı parametreleri oluşturulur (Eren ve ark., 2008). Bu çalışmada verilerin gerçek zamanlı olarak işlenebilmesi ve kalite analizi gerçekleştirilmesi amacı ile BOİ<sub>5</sub> tahmini için YSA algoritması kullanılmıştır.



Şekil 5.5. Üç Katmanlı İleri Beslemeli YSA

Şekil 5.5.'te x, y, z sırasıyla düğüm giriş katmanını, gizli katmanı ve çıktı katmanını belirtir. a düğümlerin ağırlığıdır ve a simgesinin alt simgeleri düğümler arası bağlantıları belirtir.  $a_{xy}$ , x ve y düğümleri arasındaki ağırlıktır. Basit bir sinir ağı yapısı Şekil 5.6.'te yer almaktadır.



Şekil 5.6.Yapay sinir ağı nöronu çalışma fonksiyonu (F. Rosenblatt'ın 1958 yılında tanıttığı model)

Bu sinir ağı tek katmanlı bir sinir ağı olduğundan basit bir perseptron yapı olarak bilinir. Perseptron beynimizdeki tek nöronun genel yapısını ifade etmektedir. Sinir ağında girdi ve çıktı bilgileri önemli olmayıp asıl önemli nota orta katmandaki toplam kümesidir. Bu katman öğrenmenin gerçekleştiği alandır. Bu basit sinir ağının fonksiyonları Denklem 5.6 ve 5.7'teki gibi algoritmik fonksiyonlara sahiptir (O' Shea ve Hoydis, 2017).

$$S = f(A) = f(X_1 \cdot W_1 + X_2 \cdot W_2 + \dots + X_i \cdot W_i) = f\left(\sum_{p=1}^i X_i W_i\right) \quad (5.6)$$

$$S = f(A) = f(X_1 \cdot W_1 + X_2 \cdot W_2 + \dots + X_i \cdot W_i) = f\left(\sum_{p=1}^i X_i W_i + b\right) \quad (5.7)$$

Formülde;

S= Sistemden İstenen Çıktıyı,

A = Aktivasyon Nöronunu,

f(A) = Aktivasyon Fonksiyonunu,

X<sub>1</sub>, X<sub>2</sub>,.....X<sub>i</sub> = Sisteme Girilen Veri Kümesini,

W<sub>1</sub>, W<sub>2</sub> .....W<sub>i</sub> = Fonksiyonun Vektörel Ağırlıklarını

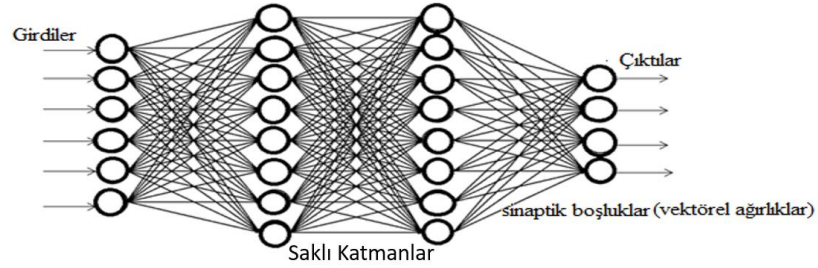
X<sub>p</sub> = (p=1,....,i) Veri Kümesinin P Sırasında Bulunan Elemanını,

W<sub>p</sub> = (P=1,....,i) Vektörel Ağırlığın P Sırasındaki Elemanını,

b = Bias önyargı değerini, ifade etmektedir.

Şekil.5.6'daki bağlantı çizgileri insan beyni nöronlarından esinlenen ağırlık vektörleridir. "Ağırlık vektörünün ataması düzgün şekilde atanmamış ise bilgi, taşıma tam olarak sağlanamayacağından, çok katmanlı sinir ağında kendisinden sonraki katmana eksik olarak geçebileceği gibi, taşınmaması riskini beraberinde getirmektedir

(Ying ve ark., 2019). Perseptronların bir araya gelerek oluşturdukları çok katmanlı yapay sinir ağı Şekil.6'da yer almaktadır.



Şekil 5.7. Çok Katmanlı Yapay Sinir Ağı Görünümü

Şekil 5.7.'da yapay sinir ağı nöronu olan perseptronların bir araya gelerek sinaptik boşluk veya diğer adıyla vektörel ağırlıklar vasıtasıyla, bilgiyi kendisinden sonraki katmana taşıdığı ve öğrenmenin sağlandığı çok katmanlı bir sinir ağı görüntüsü yer almaktadır. YSA için ilk katmanın girdi katmanı olduğunu ve bu katmanda herhangi bir öğrenmenin gerçekleşmediği vektörel ağırlıkların daha sonraki orta katmanlara geçişte kullanıldığı şekilden de anlaşılmaktadır. Bu nedenle, girdi ve çıktı katmanlarında taşıma kendinden sonraki katmana aktarılmadığından, öğrenmenin bu katmanlarda sağlanmaması ve orta katmanlarda öğrenmenin gerçekleştiği görülür olmaktadır. Bu durumda, ağırlık vektörünün ataması doğru şekilde yapıldığında bilgi ancak taşınabilir ve kendisinden sonraki katmana ulaşabilir olmaktadır. Çıktılar ise, modelin öğrenmesi sırasında baştan belirlenmekte ve orta katmanlar öğrenmesini sağlamaktadır.

Bilgiyi bu şekilde öğrenen YSA veri setindeki girdilerle tanımlanmış çıktıları benimseyerek öğrenip, ayırt edebilir hale gelmektedir. YSA model eğitiminde orta katman sayısı arttıkça sistem daha çok öğrenebilir olmakta veri setinin büyütülmesiyle daha etkili bir tahmin yapabilir duruma gelmektedir. Ancak, uzman sistemlerde, uzmanlık alanı olarak düşünülen alanın dışında başkaca verilerle ağı çalışması, bir süre sonra ağı ezberlemesine neden olacağından tüm şeyleri bilen ve ayırt edebilen insanüstü bir varlığın oluşturulmasından henüz çok uzakta olduğumuzu belirtmek ve bu noktada, ağı büyük veri sorunu ile karşılaşmasına neden olacağını vurgulamak elzemdir. Büyük veri sorunu kısaca, bir doktorun mühendislik çalışmalarını sadece

okumalar yaparak tüm hatlarıyla bilmesi ve uygulamada kullanması gibi çelişki yaratan sorun olarak tanımlanabilir.

#### 5.1.4.3. Güçlendirilmiş Karar Ağacı Regresyonu (GKAR)

Ağaç tabanlı öğrenme algoritmaları yaygın olarak kullanılan denetimli öğrenme yöntemlerinden birisi olmasının sebebi yüksek doğruluk, kararlılık ve yorumlama kolaylığına alt yapı sunmasıdır. Karar Ağacı öğrenmesi ise genel olarak çıkarım için kullanılmaktadır. Güçlendirilmiş Karar Ağacı Regresyonu (GKAR) tahmin modellerinin oluşturulmasında kullanılmaktadır. Bunun sebebi öğrenme görevlerinin daha kolay yerine getirilmesi ve kesine yakın sonuçlara ulaşılmasıdır. Nihai tahmin değerinin elde edilmesi için Güçlendirilmiş Karar Ağacı Regresyonu ile oluşturulan tahmin modelinde yer alan değerlerin ağırlıklı ortalaması yinelemeli olarak hesaplanmaktadır. Her adımda ek bir sınıflandırıcının çağırılması ile tüm veri setinin performansı ölçülmektedir. "Tahminde dayalı model için GKAR algoritması şu şekildedir;

Input: A set of data points  $(x_i, y_i)$  from the given dataset

Output: A regression tree  $T$

Initialisation: the weak learners to an individual list  $T$  the current regression tree with low values

Initialisation:  $iter \leftarrow$  number of iterations

1: for  $i = 1$  to  $iter$  do

2: Calculate new weights of examples  $(x_0, y_0)$  to  $(x_i, y_i)$  by evaluation on incremental examples with low prediction accuracy by  $T$

3: project new weak classifier  $h_i$  on pre-weighted examples

4: compute weight  $\beta_i$  of new weak classifier

5: add the pair  $(h_i, \beta_i)$  to  $T$

6: end for

7: return  $T$  (Manna ve ark., 2018)".

Bu çalışmada daha doğru sonuçlara ulaşılması için tahmin modellerinin oluşturulmasında Güçlendirilmiş Karar Ağacı Regresyonu kullanılmıştır.

#### **5.1.4.4. Bayesyen Doğrusal Regresyon (BDR)**

Gelişmiş ve yaygın olarak kullanılan olasılık yöntemlerinden biri olan Bayes Teoremi, yeterli verinin elde edilememesinden kaynaklı olarak karar vermede zorluk yaşanan durumlarda koşullu olasılıklarının kullanılarak işlem yapılmasına olanak tanımakta ve olasılık içeren bütün alanlarda kullanılmaktadır. Bayes teoremi rassal bir değişkene ait olasılık dağılımı içerisinde yer alan koşullu olasılıklar ve marjinal olasılıklar arasındaki ilişkiyi belirlenmesini sağlar ve teoreme göre incelenen bir olay doğrultusunda B olayına koşullu bir A olayına ait olasılık değeri ile A olayına koşullu olarak B olayına ait olasılık değeri birbirinden farklıdır (Yüksel, 2019). Bayesyen Doğrusal Regresyon yönteminin temelini Doğrusal Regresyon oluşturmaktadır. Doğrusal (Lineer) Regresyon yöntemlerinin kullanılması durumunda nokta tahminlerinden yararlanmaktadır. Bundan farklı olarak Bayesian Linear Regression yöntemi ile olasılık dağılımlarını kullanmak suretiyle işlem gerçekleştirilmektedir. Bayesian Linear Regression yöntemi doğrultusunda iki aşama izlenmektedir. “İlk aşama modele ait parametrelerin dağılımının belirlenerek Lineer Regresyon fonksiyonu ile çarpılması ve ikinci aşama tahmin dağılımının hesaplanmasıdır. Böylece tahmin dağılımının oluşturulan model ile ilişkisi belirlenerek yeni parametreler ve tahmin aralıkları elde edilmektedir (Panni ve ark., 2016).

Bu çalışmada Lineer Regresyon modelinden elde edilen sonuçların kesinliğinin ölçülmesi ve Lineer Regresyon modelinde oluşan hataların belirlenmesi amacı ile Bayesian Linear Regression yöntemi kullanılmıştır.

#### **5.1.4.5. Karar Ormanı Regresyonu (KOR)**

Karar ağacı yaklaşımı veri bilimlerinde yaygın olarak kullanılmaktadır. Karar Ormanı tahmin modellerinde karar ağaçlarının kullanılması durumunda karar ağacının gerçekleştirmiş olduğu tahmin performansının arttırılması adına geliştirilmiş bir

yöntemdir. Karar Ormanı Regresyonu birden fazla karar ağacının kullanılmasıyla modele ilişkin daha doğru tahmin sonuçlarının elde edilmesini sağlayan regresyon modelidir.

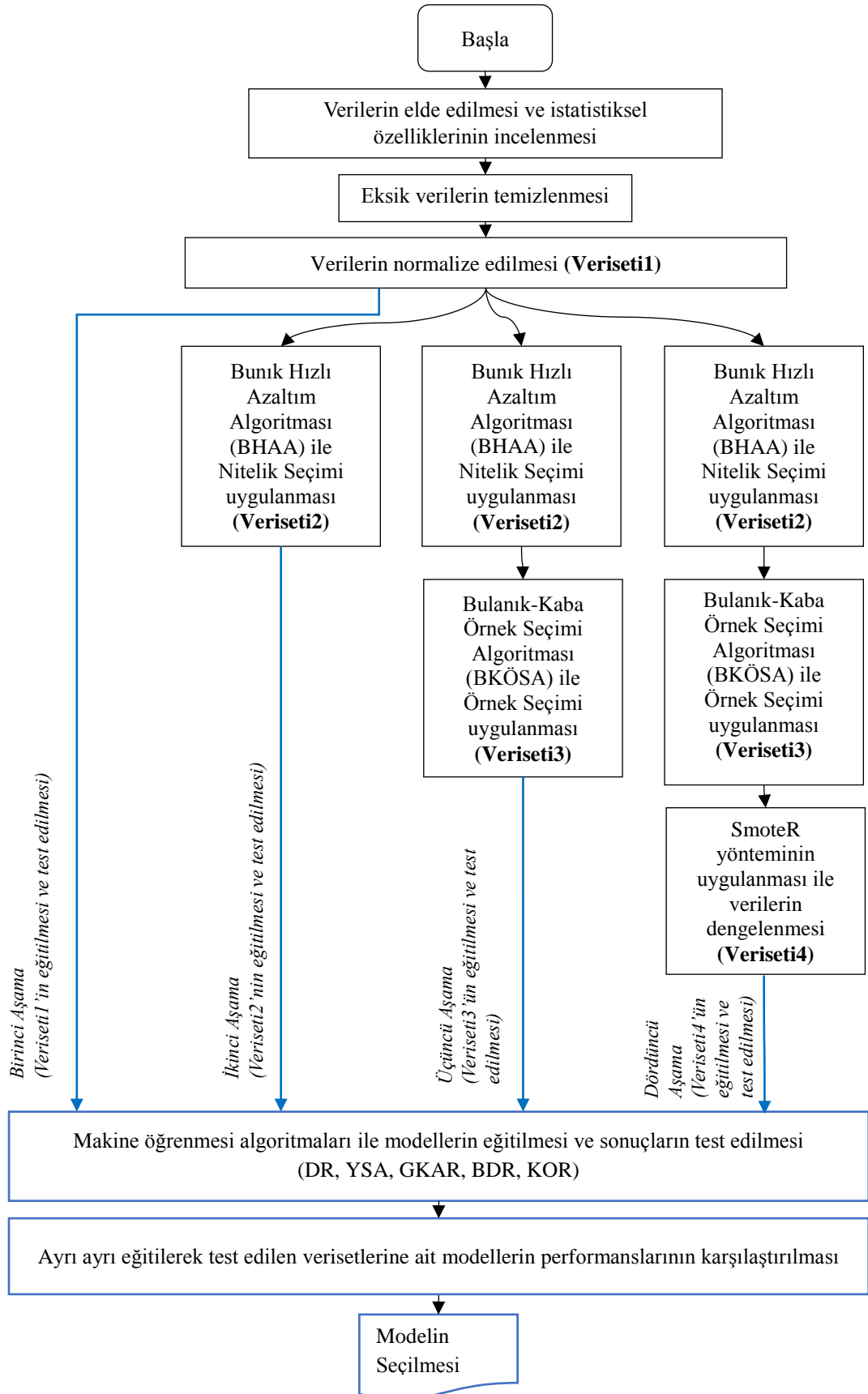
Karar Ormanı Regresyonu yönteminde temel ilke birden fazla karar ağacına ilişkin sonucun ölçülerek sonuçların birleştirilmesi ile karar alınmaktadır. Böylece birkaç doğrusal regresyon modeli birleştirilerek tahmine dayalı yeni bir model oluşturulmaktadır. İlgili modele göre her bir ağaçtan elde edilen tahmin değeri bir diğer ağaçtan daha iyi sonuç veriyorsa ağaçlara ilişkin bütün çıktılarının birleştirilmesi potansiyel olarak daha doğru bir tahminleme yapılmasına olanak tanımaktadır (Criminisi, 2011). Bu çalışmada verilerden elde edilen bilgilerin doğruluk ve kesinliğinin artırılması amacı ile regresyon tekniklerinin etkin şekilde kullanılmasına imkan tanıyan Karar Ormanı Regresyon yöntemi kullanılmıştır.

## 5.2. Metodoloji ve Uygulama

Atıksu arıtma tesisinde BOİ<sub>5</sub> tahmini çalışması için uygulama akış şeması Şekil 5.8.'de verilmektedir. İlk aşamada, makine öğrenmesi çalışmaları için veriler elde edilerek eksik veriler temizlenmiştir. Daha sonra veriler normalize edilerek Veriseti1 elde edilmiştir. İkinci aşamada gereksiz ve ilgisiz niteliklerin elimine edilmesi için elde edilen veri setine Bulanık Hızlı Azaltım algoritması uygulanarak nitelik seçimi gerçekleştirilmiştir. Böylece Veriseti2 elde edilmiştir.

Üçüncü aşamada Veriseti3 için, Bulanık-Kaba Örnek Seçimi Algoritması (BKÖSA) uygulanarak veri setinden tutarsız ve gürültülü örnekler çıkartılmıştır. Dördüncü aşamada verilerin dengelenmesi için veri setine SmoteR algoritması uygulanarak Veriseti4 elde edilmiştir.

Elde edilen her bir verisetinden sonra makine öğrenmesi regresyon algoritmaları ayrı ayrı eğitilerek test edilmiş ve her verisetine ait modellerin performansları hata ölçüm yöntemlerine göre karşılaştırılmıştır.



Şekil 5.8. Uygulama Akış Şeması



### 5.2.1. Çalışma alanı ve verilerin elde edilmesi

Bu çalışma, Sakarya'da bulunan Karaman Atıksu Arıtma Merkezinde gerçekleştirilmiştir. Sakarya merkezde bulunan Şekil 5.9.'daki Karaman Atıksu Arıtma Tesisi 2003 yılında faaliyete alınmış olup 1.000.000 kişi eşdeğer nüfusa hizmet edebilecek büyüklüktedir. Adapazarı ilçesinin Karaman mevkiinde kurulan Atıksu Arıtma Tesisi 296.000 m<sup>2</sup> alana kurulmuş ve 198.800 m<sup>3</sup>/gün kapasiteye sahiptir. Tesiste Sakarya İline bağlı olan Adapazarı, Erenler, Serdivan, Arifiye ve Sapanca ilçelerinden gelen evsel nitelikli atıksular arıtılmaktadır.



Şekil 5.9. Karaman Atıksu Arıtma Tesisi

Çalışma kapsamında kullanılan veriler, Karaman Atıksu Arıtma Tesisinin girişinden 2019-2020 yılları arasında elde edilen 177 örnekten oluşmaktadır. Her bir örneğe ait sekiz girdi (KOİ, AKM, ph, (NH<sub>4</sub>-N), (NO<sub>3</sub>-N), (NO<sub>2</sub>-N), (TN), (TP)) ve bir çıktı (BOİ<sub>5</sub>) parametresine ilişkin değerler laboratuvar ortamında ölçülmüştür. 18 örnekte ölçüm değerlerine ilişkin eksik nitelikler bulunduğundan bu 18 örnek silinerek çalışmaya 159 örnek üzerinden devam edilmiştir. Veri setine ait tanımlayıcı istatistikler Tablo.8'de verilmektedir. Tabloda sırasıyla ortalama ( $X_{ort}$ ), medyan ( $X_{med}$ ), minimum ( $X_{min}$ ), maksimum ( $X_{maks}$ ), standart sapma ( $S_x$ ), çarpıklık katsayısı ( $C_{sx}$ ) ve basıklık katsayısı ( $C_k$ ) değerleri verilmektedir.

Tablo 5.1. Karaman Atıksu Arıtma Tesisi analiz veri setine ait tanımlayıcı istatistikler

Parametre	$X_{ort}$	$X_{med}$	$X_{min}$	$X_{maks}$	$S_x$	$C_{sx}$	$C_k$
KOİ	356,584906	334	47	970	167,0476435	1,09138853	1,667487982
AKM	263,915157	236	13	926	140,6228017	1,200308039	2,704852919
Ph	7,348239	7,37	6,52	8,1	0,304105941	-0,08220624	-0,24112730
NH <sub>4</sub> -N	15,87434	14,1	3,3	39,3	6,313935314	0,691276485	0,393578754
NO <sub>3</sub> -N	0,723836	0,38	0	10,67	1,209172643	4,924289673	33,25449557
NO <sub>2</sub> -N	0,134969	0,06	0	1,71	0,207119712	3,987172499	23,16423281
TN	32,451321	31,81	3,36	66	10,66365708	0,275046065	0,538700732
TP	6,851761	6,3	0,01	24,99	3,688161961	1,339575352	3,897922155
BOİ <sub>5</sub>	113,264151	110	14	300	52,526359	0,891839254	1,322693424

Tablo 5.1. incelendiğinde, deneysel veri setinin çok farklı ölçekteki sayısal değerlere sahip olan niteliklerden oluştuğu görülmektedir. Bu durum, modelleme sırasında deneye potansiyel olarak hata ekleyebilir veya makine öğrenmesi algoritmalarının performansını olumsuz yönde etkileyebilir. Değerleri 0-1 aralığında ortak bir ölçekte olacak şekilde dönüştürmek için girdi değişkenlerine Denklem 5.8 kullanılarak MinMax normalizasyon işlemi uygulanmıştır. Bu şekilde Veriseti1 oluşturulmuştur.

$$\tilde{X}^i = \frac{X^i - X_{min}}{X_{max} - X_{min}} \quad (5.8)$$

Burada  $x^i$  özelliğin değeri,  $\tilde{x}^i$  ölçeklendirilmiş özellik değeridir ve  $X_{min}$  ve  $X_{max}$  sırasıyla veri kümesinin minimum ve maksimum değeridir.

### 5.2.2. Bulanık Hızlı Azaltım Algoritması (BHAA) ile nitelik seçimi

Bu çalışmada, verisetinden gereksiz ve ilgisiz niteliklerin kaldırılarak optimal bir alt küme veriseti elde etmek amacıyla bulanık kaba kümeleme tabanlı nitelik seçimi için Bulanık Hızlı Azaltım Algoritması (BHAA) kullanılmıştır. BHAA uygulanması için 2015 yılında Rıza vd. tarafından R programlama dili için geliştirilen 'RoughSets' paketinden yararlanılmıştır. "RoughSets" paketi, kaba kümeler ve bulanık kaba kümeler teorisine dayalı algoritma uygulamalarını içeren bir yazılım paketidir.

Paketteki yöntemler sayesinde, temel kavramlar, ayrıklaştırma, özellik seçimi, örnek (veri) seçimi, kural çıkarma ve tahmin/sınıflandırma gibi uygulamalar gerçekleştirilebilmektedir. BHAA adımları R Studio programında FS.quickreduct.FRST() fonksiyonu aracılığıyla gerçekleştirilmiştir. Fonksiyonun uygulanabilmesi öncelikle install.packages("RoughSets") 'RoughSets' paketinin R Studio ortamına yüklenmesi gerekmektedir. Paket daha sonra library(RoughSets) kodu ile etkinleştirilmiştir. Analiz edilen veri setinde bulanık kaba kümeleme tabanlı nitelik seçimi için R Studio programında geliştirilen kodlar şu şekildedir.

```
install.packages("RoughSets")
library(RoughSets)
karar.tablosu <- SF.read.DecisionTable (file.choose(), decision.attr = 9, sep=",")
colnames(karar.tablosu) <- c("KOI", "AKM", "Ph", "NH4-N", " NO3-N ", "
NO2-N", " TN", "TP", "BOI5")
kontrol <- list(alpha = 0.9, q.some = c(0.1, 0.6), q.most = c(0.2, 1),
type.aggregation = c("t.tnorm", "lukasiewicz"))
nitsec <- FS.quickreduct.FRST(karar.tablosu, type.method = "vqrs", type.QR =
"fuzzy.QR", control = kontrol)
karar.tablosu2 <- SF.applyDecTable(karar.tablosu, nitsec)
```

Uygulamada alt ve üst yaklaşımlar için yöntem olarak 2008 yılında Cornelis ve Jensen (Baydilli, 2020), tarafından geliştirilen belirsiz nicelikli kaba küme (VQRS) yaklaşımı kullanılmıştır. 0 ile 1 arasında bir eşik değerini ifade eden gerçek bir sayı veya durdurma kriteri olan 'alpha' 0.9 olarak belirlenmiştir. VQRS'nin alfa ve beta parametresi için bir çift sayısal değer olan bazı niceleyici (q.some) 0.1 ve 0.6 olarak, en çok niceleyici (q.most) 0.2 ve 1 olarak belirlenmiştir. Toplama (aggregation) operatörü olarak üçgen (triangular) norm operatörü olan t.tnorm operatörü ve lukasiewicz'in t-normu (lukasiewicz's t-norm) kullanılmıştır. Atıksu arıtma tesisinden elde edilen KOİ, AKM, ph, NH<sub>4</sub>-N, NO<sub>3</sub>-N, NO<sub>2</sub>-N, TN, TP, BOI<sub>5</sub> olmak üzere dokuz niteliğe sahip veri seti (Veriseti1) BHAA uygulanmasıyla 4 niteliğe indirgenmiş ve Veriseti2 elde edilmiştir. Veriseti2'de bulunan nitelikler sırasıyla, KOİ, NO<sub>2</sub>-N, TN ve BOİ<sub>5</sub>'dir.

### 5.2.3. Bulanık-Kaba Örnek Seçimi (BKÖSA) algoritmasının uygulanması

Bu aşamada Veriseti2’de bulunan tutarlı örnekleri korurken, gürültülü, gereksiz veya tutarsız örnekleri kaldırmak amacıyla Bulanık-Kaba Örnek Seçimi (BKÖSA/FRIS) algoritması kullanılmıştır. Algoritmanın uygulanması için bulanık kaba kümeleme tabanlı nitelik seçimi uygulamasında olduğu gibi R Studio programı aracılığıyla ‘RoughSets’ paketinden yararlanılmıştır. Bulanık-kaba örnek seçimi algoritması adımları IS.FRIS.FRST () fonksiyonu aracılığıyla gerçekleştirilmiştir. Bulanık kaba kümeleme tabanlı örnek seçimi uygulaması R Studio kodları şu şekildedir.

```
library(RoughSets)
karar.tablosu3 <- SF.read.DecisionTable (file.choose(), decision.attr = 4,
sep=",")
colnames(karar.tablosu3) <- c("KOI", " NO2-N", " TN", "BOI5")
res.1 <- IS.FRIS.FRST(decision.table = karar.tablosu3, control =
list(threshold.tau = 0.9, alpha = 0.8, type.aggregation = c("t.tnorm",
"lukasiewicz"), t.implicator = "lukasiewicz"))
karar.tablosu4<- SF.applyDecTable(karar.tablosu3, res.1)
```

Kodlarda görülen threshold.tau bir nesnenin kaldırılıp kaldırılamayacağını belirleyen bir eşik değeridir. Eşik değerinden düşükse nesne kaldırılabilir (Criminisi, 2012). alpha ise pozitif değerlere sahip ( $> = 0$ ) bulanık benzerlik ölçüsünün ayrıntı düzeyini belirleyen bir parametredir. Bu çalışmada threshold.tau değeri 0.9, alpha değeri ise 0.8 olarak belirlenmiştir. Toplama (aggregation) operatörü olarak, bulanık hızlı azaltım algoritmasında olduğu gibi üçgen (triangular) norm operatörü olan t.tnorm operatörü ve lukasiewicz’in t-normu (lukasiewicz’s t-norm) kullanılmıştır. Algoritmanın uygulanmasıyla 5 örnek elenerek 154 örneğe sahip Veriseti3 elde edilmiştir.

### 5.2.4. SmoteR yöntemi ile veri setinin dengelenmesi

Bu çalışmada, nadir hedef değişkeni durumları ile en sık karşılaşılan değerler arasındaki dengesizlik sorununu azaltmak için SmoteR yöntemi kullanılmıştır.

Yöntem adımlarının uygulanabilmesi için 2017 yılında Branco vd. tarafından R programlama dili için geliştirilen 'UBL' paketinden yararlanılmıştır. Paket, hem regresyon hem de sınıflandırma çalışmalarında kullanılmak üzere maliyete duyarlı ve maliyet/fayda görevlerinde daha iyi tahmin performansı elde etmek için kullanılabilir bir dizi fonksiyon sağlar. Orijinal veri kümesini kullanıcı tercihlerine göre yönlendiren yeniden örnekleme yaklaşımlarını içerir (Orçanlı, 2019). SmoteR algoritması açık kaynak kodlu R ortamında SmoteRegress() fonksiyonu yardımıyla uygulanabilmektedir. 'UBL' paketinin R Studio ortamına yüklenebilmesi için `install.packages("UBL")` komutu, paketin aktifleştirilmesi için ise `library(UBL)` kullanılmaktadır. SmoteR algoritmasının R Studio ortamında uygulanması için gerekli kod bloğu aşağıdaki gibidir.

```
install.packages("UBL")
library(UBL)
karar.tablosu5 <- SF.read.DecisionTable (file.choose(), decision.attr = 4,
sep=",")
colnames(karar.tablosu5) <- c("KOI", " NO2-N", " TN", "BOI5")
smote.kt <- SmoteRegress(BOI5~, karar.tablosu5, dist = "HEOM",
C.perc=list(1.5,2.5))
```

SmoteR algoritmasının uygulanmasında K en yakın komşu belirlenirken hem nominal hem de sayısal (numeric) nitelikler(feature) ile çalışabilen HEOM mesafesi (distance) kullanılmıştır. Yeni sentetik örnekler oluşturmak için kullanılan komşuların sayısı 5 olarak belirlenmiştir. İlgililik fonksiyonu (relevance function) otomatik (auto) olarak ayarlanmış, üstünde bir vakanın nadir "sınıfa" ait olduğu kabul edilen ilgililik eşliğini (relevance threshold) gösteren sayı 0.5 olarak belirlenmiştir. C.perc parametresi eşikle elde edilen her bir "sınıf"a (tümsek) uygulanacak aşırı örnekleme ve yetersiz örnekleme miktarını kontrol eder. Bu uygulamada aşırı örnekleme yapılacağından bu liste 1.5 ve 2.5 olarak belirlenmiştir. SmoteR algoritmasının uygulanması ile 244 örneğe sahip Veriseti4 elde edilmiştir.

### 5.2.5. Makine öğrenmesi algoritmaları ile modellerin oluşturulması

Biyolojik Oksijen İhtiyacı tahmini için makine öğrenmesi algoritmaları ile regresyon modellerinin eğitilmesi ve test edilmesi çalışmaları Microsoft Azure Machine Learning Studio ortamında gerçekleştirilmiştir. Modellerin geliştirilmesi için veriseti eğitim veriseti ve test veriseti olmak üzere ikiye ayrılmıştır. Verisetinin %70'i eğitim için kullanılırken, %30'u test için kullanılmak üzere rassal bir şekilde seçilmiştir. Bu işlem Azure Machine Learning Studio programında "Split Data" modülü kullanılarak yapılmıştır.

Bu çalışmada kullanılan makine öğrenmesi algoritmaları GKAR, KOR, YSA, DR'dur. Kullanılan makine öğrenmesi algoritmalarından GKAR, KOR ve YSA algoritmalarının parametreleri, Azure ML Studio'da "Tune Model Hyperparameters" modülü kullanılarak tüm ızgara (entire grid) seçeneğiyle belirlenmiştir. Tüm ızgara seçeneği ile modül, sistem tarafından önceden tanımlanmış bir kılavuz üzerinde döngü gerçekleştirerek farklı birleşimler denemektedir ve en iyi öğrenme parametrelerini belirlemektedir.

GKAR algoritması ile oluşturulan modellerde, ağaç başına maksimum yaprak sayısı 8, yaprak düğüm başına minimum örnek sayısı 1, öğrenme oranı 0.1 ve toplulukta oluşturulacak toplam karar ağacı sayısı 500 olarak belirlenmiştir. KOR algoritmasının parametreleri ise yaprak düğüm başına Minimum örnek sayısı 1, düğüm başına rastgele bölme sayısı 128, karar ağaçlarının maksimum derinliği 16, toplulukta oluşturulacak toplam karar ağacı sayısı 32 olarak bulunmuştur.

YSA'nın yapısı bir girdi katmanı, bir saklı katman ve bir çıktı katmanı olmak üzere üç katmandan oluşmaktadır. Saklı katmandaki nöron sayısı 100 olarak belirlenirken, öğrenme oranı 0.1 olarak belirlenmiştir. Kayıp fonksiyonu olarak karesel hata fonksiyonu kullanılmıştır. Öğrenme iterasyon sayısı ise 80 olarak belirlenmiştir. Doğrusal regresyonda hatayı ölçmek ve regresyon çizgisini oluşturmak için çözüm yöntemi olarak Sıradan En Küçük Kareler yöntemi kullanılmıştır. Modelin aşırı öğrenme sorununu çözmek için kullanılan L2 düzenleme ağırlığı 0.001 olarak

belirlenmiştir. BOİ<sub>5</sub> tahmini için doğrusal regresyon yöntemi ile niteliklerin ağırlıkları KOİ, NO<sub>2</sub>-N ve TN için sırasıyla 268.663, 25.9467 ve 17.3191 olarak bulunurken, bias 12.2639 olarak bulunmuştur. BDR algoritmasında gürültünün kesinliğinden önceki ağırlık hassasiyetinin oranını temsil eden düzenleme ağırlığı sabiti 1 olarak belirlenmiştir.

### 5.3. Modellerin Performanslarının Karşılaştırılması

Makine öğrenmesi algoritmaları ile eğitilen modellerin performansları (Denklem 5.5-5.10) kullanılarak altı farklı performans ölçütüne göre değerlendirilmiştir ve bunlar ortalama mutlak hata (MAE), ortalama karesel hata (RMSE), göreceli mutlak hata (RAE), göreceli kare hata (RSE), ortalama mutlak yüzde hatası (MAPE) ve belirleme katsayısı ( $R^2$ )'dir (Gültepe, 2019).

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - \bar{Y}_i| \quad (5.5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{Y}_i - Y_i)^2} \quad (5.6)$$

$$RAE = \frac{\sum_{i=1}^N |\hat{Y}_i - Y_i|}{\sum_{i=1}^N |\bar{Y} - Y_i|^2} \quad (5.7)$$

$$RSE = \frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)^2}{\sum_{i=1}^N (\bar{Y} - Y_i)^2} \quad (5.8)$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|Y_i - \bar{Y}_i|}{Y_i} \quad (5.9)$$

$$R^2 = \left( \frac{\sum_{i=1}^N (Y_i - \bar{Y}_i) - \sum_{i=1}^N (\hat{Y}_i - \bar{Y}_i)}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y}_i)^2} \sqrt{\sum_{i=1}^N (\hat{Y}_i - \bar{Y}_i)^2}} \right)^2 \quad (5.10)$$

$Y_i$  ve  $\hat{Y}_i$  sırasıyla gözlemlenen ve tahmin edilen değerlerin ortalamasıdır. N, test veri kümesinin örnek numarasını gösterir. Performans ölçütleri olan MAE, RMSE, RAE, RSE ve MAPE değerlerinin sıfıra yakın olacak kadar düşük olması kullanılan modellerin tahmin doğruluğu için iyi performans sergilediğini göstermektedir.  $R^2$

değeri ise 0 ile 1 değerleri arasında olmalı ve 1'e yaklaştıkça en iyi performansı göstermektedir.

BOİs tahmini için elde ettiğimiz her verisini microsoft azure machine learning studio programı üzerinden makine öğrenmesi algoritmaları (Güçlendirilmiş Karar Ağacı Regresyonu (GKAR), Bayesyen Doğrusal Regresyon (BDR), Karar Ormanı Regresyonu (KOR), Doğrusal Regresyon (DR), Yapay Sinir Ağları (YSA)) ile eğiterek gerçekleştirdiğimiz tüm modellerin performans endekslerinin sonuçları Tablo 5.2.'de verilmiştir.

Atıksu arıtma tesisinden toplanan veriler Microsoft Azure Machine Learning Studio programında "Clean Missing Data" modülü kullanılarak eksik veriler temizlenmiş ve 177 örnek içerisinde 18 örnek silinerek çalışmaya 159 örnek üzerinden devam edilmiştir.

Eksik veri temizleme işleminden sonra verilerin çok farklı ölçekteki sayısal değerlere sahip olduğu görülmektedir. Örneğin giriş parametrelerinden KOİ parametresinin minimum ve maximum değerli 47-970 mg/Lt, AKM parametresine ait veriler 13-926 mg/Lt arasında, pH 6,5-8,1 arasında değiştiğinden dolayı büyük ölçekte bulunan parametrelere daha fazla ağırlık verilmektedir. Bu sorunu ortadan kaldırabilmek için parametrelere ait minimum ve maximum değerlerinin 0-1 aralığına dağıtılmasını sağlamak amacıyla "Normalize Data" modülü uygulanarak veriler normalize edilmiş ve Veriseti1 oluşturulmuştur.

Tablo 5.2.'de görüldüğü üzere elde edilen Veriseti1, Microsoft Azure Machine Learning Studio programı üzerinden makine öğrenmesi algoritmaları uygulanarak her bir model ayrı ayrı eğitilmiş ve test edilmiştir. Yapılan performans testleri sonucunda MAE, RAE ve MAPE için en düşük değeri DR modelinin, RMSE ve RSE için en düşük değeri BDR modelinin verdiği görülmektedir. R2 için en yüksek değeri ise %91,56 ile BDR algoritması vermiştir. MAPE sonuçlarına göre en iyi performansı %10,35 hata ile DR algoritması göstermiştir.



Tablo 5.2. BOİs'teki Modellerin Veriseti1 Üzerindeki Test Performansları

Model	MAE	RMSE	RAE	RSE	MAPE	R <sup>2</sup>
<b>Veriseti1</b>						
GKAR	15.250172	22.172688	0.357368	0.175684	13,43187352	0.824316
BDR	12.341851	<b>15.362177</b>	0.289215	<b>0.084333</b>	13,19429062	<b>0.915667</b>
KOR	14.214974	19.474741	0.333109	0.135531	14,01442928	0.864469
DR	<b>11.824379</b>	16.191638	<b>0.277089</b>	0.093686	<b>10,35002566</b>	0.906314
YSA	14.092879	17.163436	0.330248	0.10527	17,36872068	0.89473

Veriseti1 eğitilerek performans testleri yapıldıktan sonra çalışmanın sonucunun iyileştirilmesi amacıyla verisetinden gereksiz ve ilgisiz niteliklerin kaldırılarak optimal bir alt küme veriseti elde etmek amacıyla bulanık kaba kümeleme tabanlı nitelik seçimi için Bulanık Hızlı Azaltım Algoritması (BHAA) kullanılmıştır. R Studio programı üzerinden yapılan çalışma ile 9 nitelik arasından 4 niteliğe indirgenerek Veriseti2 elde edilmiştir.

Tablo 5.3'te görüldüğü üzere elde edilen Veriseti2, Microsoft Azure Machine Learning Studio programı üzerinden makine öğrenmesi algoritmaları uygulanarak her bir model ayrı ayrı eğitilmiş ve test edilmiştir. Yapılan performans testleri sonucunda BHAA uygulanmasıyla neredeyse tüm algoritmaların tahmin performansının arttığı görülmektedir. Sadece BDR algoritmasının RMSE ve RSE değerleri biraz artarak, R<sup>2</sup> değerinde düşüş yaşanmıştır. MAPE sonuçlarına göre en iyi performansı %9,79 hata ile DR algoritması göstermiş ve Veriseti1 uygulamasına göre daha iyi bir performans sergilemiştir.

Tablo 5.3. BOİs'teki Modellerin Veriseti2 Üzerindeki Test Performansları

Model	MAE	RMSE	RAE	RSE	MAPE	R <sup>2</sup>
<b>Veriseti2</b>						
GKAR	13.58617	17.440139	0.318374	0.108691	13,0753301	0.891309
BDR	11.887472	<b>15.627701</b>	0.278567	<b>0.087274</b>	12,95957943	<b>0.912726</b>
KOR	13.503373	18.53018	0.316434	0.122703	13,07717945	0.877297
DR	<b>11.216992</b>	15.710536	<b>0.262855</b>	0.088202	<b>9,793055196</b>	0.911798
YSA	12.991923	17.305579	0.304449	0.10702	14,73066599	0.89298

Veriseti2 eğitilerek performans testleri yapıldıktan sonra çalışmanın iyileştirilmesi amacıyla R Studio programı üzerinden Bulanık-Kaba Örnek Seçimi Algoritması (BKÖSA) kullanılarak gürültülü, gereksiz veya tutarsız örnekleri kaldırmış ve 159 örnek içerisinde tutarsız olan 5 örnek elenerek 154 örneğe sahip Veriseti3 oluşturulmuştur.

Tablo 5.4'te görüldüğü üzere elde edilen Veriseti3, Microsoft Azure Machine Learning Studio programı üzerinden makine öğrenmesi algoritmaları uygulanarak her bir model ayrı ayrı eğitilmiş ve test edilmiştir. Yapılan performans testleri sonucunda Örnek Seçimi olan BKÖSA uygulanması ile tahmin performanslarının Veriseti2'ye göre daha fazla arttığı görülmüştür. Veriseti3'e uygulanan makine öğrenmesi algoritmaları ile DR modelinin performans testi en iyi sonuçları göstermektedir.

Tablo 5.4. BOİs'teki Modellerin Veriseti3 Üzerindeki Test Performansları

Model	MAE	RMSE	RAE	RSE	MAPE	R <sup>2</sup>
<b>Veriseti3</b>						
GKAR	9.978678	15.876328	0.259116	0.095777	12,47114204	0.904223
BDR	12.582426	17.00973	0.326728	0.10994	14,46653209	0.89006
KOR	9.552253	13.191693	0.248043	0.066124	11,9367435	0.933876
DR	<b>9.160436</b>	<b>12.71608</b>	<b>0.237869</b>	<b>0.061442</b>	<b>9,968557594</b>	<b>0.938558</b>
YSA	11.040238	14.715567	0.286682	0.082284	12,1710765	0.917716

Veriseti3 eğitilerek performans testleri yapıldıktan sonra veri setinde dengesiz veriler olduğu tespit edilmiştir. Tesise gelen atıksuların BOİ değerleri genellikle 14-300mg/lit arasında değişmektedir. Ancak 250-300mg/lit kirlilik yüküne sahip olan atıksular verisetinde nadir olarak bulunmakta olup verilerin yoğunlaştığı değerler arasında dengesizlik oluşmaktadır. Oluşan bu sorunun çözümü için R Studio programı üzerinden dengesiz verilerin dengelenmesi (artırılması) amacıyla SmoteR yöntemi uygulanmış ve 154 örneğe sahip olan veriseti 244 örneğe yükseltilerek Veriseti4 elde edilmiştir.

Tablo 5.5'te görüldüğü üzere elde edilen Veriseti4, Microsoft Azure Machine Learning Studio programı üzerinden makine öğrenmesi algoritmaları uygulanarak her bir model

ayrı ayrı eğitilmiş ve test edilmiştir. Yapılan performans testleri hata sonuçları incelendiğinde tüm modellerimizdeki algoritma değerlerinde ciddi azalışlar görülürken,  $R^2$  değerinde artış görülmüştür. MAE, RMSE, RAE, RSE ve MAPE sonuçlarına göre en iyi performansın Güçlendirilmiş Karar Ağacı Regresyonu (GKAR) algoritmasında olduğu görülmektedir. Yapılan son çalışma ile MAPE sonuçlarına göre en iyi performansı veren GKAR modeli %6,075287884 hataya sahip ve %97,1805  $R^2$  değeri ile en iyi tahmin performansı elde edilmiştir.

Tablo 5.5. BOİs'teki Modellerin Veriseti4 Üzerindeki Test Performansları

Model	MAE	RMSE	RAE	RSE	MAPE	$R^2$
<b>Veriseti4</b>						
GKAR	<b>7.000193</b>	<b>9.321163</b>	<b>0.158623</b>	<b>0.028195</b>	<b>6,075287884</b>	<b>0.971805</b>
BDR	9.430184	12.661962	0.213686	0.052027	9,260435488	0.947973
KOR	7.617682	11.125451	0.172615	0.040166	6,853280763	0.959834
DR	8.671448	11.850078	0.196493	0.045569	7,635109023	0.954431
YSA	9.91696	13.029563	0.224716	0.055092	9,676174387	0.944908

Özellikle SmoteR yönetimi ile veri dengelemesi yaparak  $R^2$  dışındaki modellere uygulanan tüm algoritmaların hata değerlerinin düşmesi ve  $R^2$  değerinin artması, veriseti uygulamalarımız sayesinde tahmin performanslarının her aşamada yükseldiğini ve daha iyi sonuçlar verdiğini göstermektedir.

## BÖLÜM 6. SONUÇ

Günümüzde çevre ve halk sağlığı dikkate alındığında, atıksu arıtma tesislerinin işletmesinde atıksuyun kirlilik parametrelerinin hızlı bir şekilde izlenmesi ve atıksuya müdahale edilebilmesi önem arz etmektedir. Elde edilen veriler ile izlenmesi önem arz eden parametrelerin makine öğrenmesi algoritmaları ile tahmin edilmesi ile tesislerin işletiminde zaman ve maliyet açısından fayda sağlayacağı beklenmektedir. Bu çalışmada laboratuvar şartlarında 5 (beş) gün içerisinde gerçekleşen Biyolojik Oksijen İhtiyacı (BOİ<sub>5</sub>) testi, yapay zeka yöntemi olan makine öğrenmesi algoritmaları ile daha kısa sürede ve laboratuvar ölçeğinde tahmin edilmesine yönelik çalışmalar yapılmıştır.

Atıksu arıtma tesisinin giriş suyunda bakılan Kimyasal Oksijen İhtiyacı (KOİ), Askıda Katı Madde (AKM), pH, Amonyum Azotu (NH<sub>4</sub>-N), Nitrat (NO<sub>3</sub>-N), Nitrit (NO<sub>2</sub>), Toplam Azot (TN), Toplam Fosfor (TP) ve BOİ<sub>5</sub> parametreleri bu çalışma kapsamında kullanılmıştır. BOİ parametresinin tespiti için laboratuvar şartlarında yapılan analizin yüksek maliyetlere sebep olması ve uzun zamanda sonuçlanması tesisin verimli işletilebilmesi açısından problem oluşturmaktadır. Bu sebep ile giriş suyu değerleri ile BOİ<sub>5</sub> tahmini için geçmiş yıllara ait veriler kullanılarak tahmin performansı hesaplanmıştır.

BOİ tahmini için makine öğrenmesi algoritmaları ile regresyon modellerinin eğitilmesi ve test edilmesi çalışmaları, Microsoft Azure Machine Learning Studio ortamında gerçekleştirilmiştir. Modellerin geliştirilmesi için elde edilen verisetleri, eğitim veriseti ve test veriseti olmak üzere iki aşamada değerlendirilmiştir. Verisetinin %70'i eğitim için kullanılırken, %30'u test için kullanılmak üzere rassal bir şekilde seçilmiştir. Bu çalışmada kullanılan makine öğrenmesi algoritmaları, Güçlendirilmiş Karar Ağacı Regresyonu, Karar Ormanı Regresyonu, Yapay Sinir Ağları, Doğrusal Regresyon ve Bayesyen Doğrusal Regresyon'larıdır.

Eđitilen modellerin performansları 6 farklı performans ölçütüne göre deęerlendirilmiř ve bunlar ortalama mutlak hata (MAE), ortalama karesel hata (RMSE), greceli mutlak hata (RAE), greceli kare hata (RSE), ortalama mutlak yzde hatası (MAPE) ve belirleme katsayısı ( $R^2$ )'dir.

Yapılan alıřmada kullanılan veri setleri 2019-2020 yılları arasında elde edilen 177 rnekten oluřmaktadır. Her bir rneęe ait 8 girdi (KOİ, AKM, ph, (NH<sub>4</sub>-N), (NO<sub>3</sub>-N), (NO<sub>2</sub>-N), (TN), (TP)) ve 1 ıktı (BOİ<sub>5</sub>) parametresine iliřkin toplam 9 paramatreden deęerler toplanmıřtır. 18 rnekte lm deęerlerine iliřkin eksik nitelikler bulunduęundan bu 18 rnek silinerek alıřmaya 159 rnek zerinden devam edilmiřtir. Verilerin hepsi farklı lekler arasında deęiřtięinden dolayı algoritmaların daha iyi performans gstermesi amacıyla normalize iřlemi uygulanmıř ve Veriseti1 elde edilmiřtir. Elde edilen Veriseti1 zerinden algoritmalar eđitilerek performans testleri gerekleřtirilmiř ve MAE, RAE ve MAPE iin en dřk hata deęeri Doęrusal Regresyon modelinin, RMSE ve RSE iin en dřk deęeri Bayesyen Doęrusal Regresyon modelinin verdięi grlmektedir.  $R^2$  iin en iyi performansı %91,5667 ile Bayesyen Doęrusal Regresyon MAPE sonularına gre ise en iyi performansı %10,35002566 hata ile Doęrusal Regresyon algoritması gstermektedir.

Regresyon tahmini iin makine ęrenmesi algoritmalarının etkinlięini arttırmak amacı ile bulanık kaba kme yaklařımı temelli algoritmalar kullanılmıř ve verilerin analizinde belirsiz ok fazla verinin azaltılması gerekmiřtir. Bu sebeple Bulanık Hızlı Azaltım Algoritması (BHAA) ile nitelik seimi gerekleřtirilerek toplam 9 nitelikten oluřan parametreler 4 nitelięe indirgenmiř ve Veriseti2 elde edilmiřtir. Veriseti2'de bulunan nitelikler sırasıyla, KOİ, NO<sub>2</sub>-N, TN ve BOİ<sub>5</sub>'dir. Yapılan bu alıřmanın ardından makine ęrenmesi modelleri ile algoritmalar eđitilerek performans testleri yapılmıřtır. Birinci ařamaya gre neredeyse tm algoritmaların tahmin performansının arttıęı grlmektedir. Ancak sadece Bayesyen Doęrusal Regresyon algoritmasının RMSE ve RSE deęerleri biraz artarak,  $R^2$  deęerinde dřř yařanmıřtır. Veriseti2 zerinde yapılan performans testleri incelendięinde MAPE sonularına gre en iyi

performansı %9,79 hata ile Doğrusal Regresyon algoritması göstermiş ve Veriseti1 uygulamasına göre daha iyi bir performans sergilediği görülmüştür.

Nitelik seçimi amacıyla uygulanan Bulanık Hızlı Azaltım Algoritması'nın ardından tutarsız ve gürültülü örneklerin elenmesi amacı ile Bulanık-Kaba Örnek Seçimi Algoritması (BKÖSA) kullanılmıştır. Bu aşama veriseti2'de bulunan tutarlı örnekleri korurken, gürültülü, gereksiz veya tutarsız örnekleri kaldırmak amacıyla 159 örnek içerisinde 5 örnek elenerek 154 örneğe sahip Veriseti3 elde edilmiştir. Elde edilen Veriseti3 üzerinde algoritmalar eğitilerek performans testleri yapılmıştır. Yapılan makine öğrenmesi çalışmaları ile tahmin performanslarının Veriseti2'ye göre daha iyi sonuç verdiği görülmüştür. Veriseti3 ile eğitilen algoritmalar içerisinde MAPE sonuçlarına göre %9,968557594 hata payı ve %93,3876 R<sup>2</sup> tahmin performansları ile Doğrusal Regresyon en iyi sonucu göstermiştir.

Dengesiz veri setleri ile karşılaşıldığı için veri setleri içerisinde azınlıkta olan sınıf değerlerine ait kayıt sayısı sentetik olarak artırılarak doğru analiz yapılması ve dengenin sağlanması amacı ile SmoteR yöntemi ile veri artırma işlemi uygulanmış ve nadir hedef değişkeni durumları ile en sık karşılaşılan değerler arasındaki dengesizlik sorunu azaltılmıştır. Yöntem adımlarının uygulanabilmesi için R programlama dili için geliştirilen 'UBL' paketinden yararlanılmıştır. SmoteR algoritmasının uygulanması ile 244 örneğe sahip Veriseti4 elde edilmiştir.

Oluşturulan her verisetinde makine öğrenmesi regreyon algoritmaları eğitilerek test edilmiş ve elde edilen modellerin performansları hata ölçüm yöntemlerine göre karşılaştırılmıştır. Yapılan bu çalışmalar sonucunda makine öğrenmesi algoritmaları arasında, topluluk makine öğrenmesi algoritmalarından olan Güçlendirilmiş Karar Ağacı Regresyon (GKAR) algoritması %6,0753 hata payı olan MAPE değeri ve %97,18 R<sup>2</sup> değeri ile en iyi tahmin performansını gösteren algoritma olmuş ve tahmin çalışmalarında uygulanan metotların başarılı olduğu görülmüştür.

Karşılaştırmalı sonuçlar, bulanık kaba küme tabanlı nitelik seçimi ve örnek seçim algoritmalarının ve özellikle SMOTER yöntemlerinin makine öğrenmesi

algoritmalarının performansını artırdığını göstermektedir. Bu sayede BOİ parametresine ait analizlerin uzun ve maliyetli sonuçlarını beklemeden online ölçülebilen değerler ile veya laboratuvarında kısa zamanda ölçüm sonuçlarına ulaşılabilen parametreler ile anlık olarak tespit edilmesi sağlanmaktadır. Bu çalışmada görüldüğü üzere makine öğrenmesi yöntemleri ile birçok karmaşık olgunun çözüme ulaştırılarak hem zamandan hem maliyetten tasarruf sağlayabileceği söylenebilir. Bu sebep ile birçok parametrik değere sahip işletmelerde makine öğrenmesi teknolojileri ile daha hızlı sonuçlar elde edilerek işletmelerin performansları artırılabilir. Gelecekteki araştırmalar, farklı özellik seçimi ve örnek seçim algoritmalarının kaba kümelere ve bulanık kaba kümelere dayanan etkilerinin araştırılmasına odaklanmalıdır.

## KAYNAKLAR

- Abba, S. I., & Elkiran, G. (2017). Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application. *Procedia Computer Science*, 120(July 2019), 156–163.
- Açıklalın, S., 2005. Atıksu Arıtma Tesisi Veriminin Yapay Sinir Ağları ile Tahmin Edilmesi, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.
- Adamsky, F., Aubigny, M., Battisti, F., Carli, M., Cimorelli, F., Cruz, T., Giorgio, A.D., Foglietta, C., Galli, A., Giuseppi, A., Liberati, F., Neri, A., Panzieri, S., Pascucci, F., Proença, J., Pucci, P., Rosa, L. and Soua., R., (2018). Integrated Protection of Industrial Control Systems from Cyberattacks: the ATENA Approach. *International Journal of Critical Infrastructure Protection*, 21(72) 82.
- Annie, G., 2012. Anomaly Detection Based on Machine Learning: Dimensionality Reduction using PCA and Classification using SVM. *International Journal of Computer Applications*, 47(21).
- ASCE Task Committee.(2000). Artificial neural networks in hydrology. I. Preliminary concepts. *Journal of Hydrologic Engineering ASCE*, 5, 115–123.
- Ataseven, B. (2013). Yapay Sinir Ağları ile öngörü modellemesi. *Öneri Dergisi*, 10(39), 101-115.
- Atıksu Arıtma Tesislerinde Çalışan Teknik Personele İlişkin Eğitim ve Sertifika Programlarının Araştırılması ve Ülkemize Özgü Modelin Geliştirilmesi Projesi, Çevre ve Şehircilik Bakanlığı, 2020, Ankara.
- AYDIN HAKLI, D. (2018). Sınıf Dengesizliği Sorununu Çözmek İçin Kullanılan Algoritmaların Farklı Sınıflandırma Yöntemlerinde Performanslarının Karşılaştırılması.
- Aydilek, İ. B. (2018). Analyzing and improving information gain of metrics used in software defect prediction in decision trees. *Pamukkale University Journal of Engineering Sciences*, 24(5), 906–914.
- Ayşe, A. R. I., & Berberler, M. E. (2017). Yapay Sinir Ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı. *Acta Infologica*, 1(2), 55-73.
- B. Stojkovic, M. Vukasovic, A new SCADA System design in the Power System of Montenegro –ICCP/TASE.2 and Web- based real- time electricity demand metering extensions, Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES, 29 Oct.- 1 Nov., Atlanta, Georgia, pp.2194- 2199.



- Baeza, J.A., Gabriel, D. ve Lafuente, J., 2002. In-line fast OUR (Oxygen Uptake Rate) Measurements for Monitoring and Control of WWTP. *Water Science and Technology*, 45, 4–5, 19–28.
- Bailey, D. Wright, E., (2003).” Practical Scada For Industry” , Elsevier Isbn: 9780750658058, Pp. 203-247.
- Baki, O. T., & Aras, E. (2018). Estimation of BOD in wastewater treatment plant by using different ANN algorithms. *Membrane Water Treatment*, 9(6), 455–462.
- Bakia, O. T., Arasb, E., Akdemirc, U. O., & Yilmaza, B. (2019). Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models. *Desalin. Water Treat*, 157, 79-89.
- Baydilli, Y. Y. (2020). Derin öğrenme yöntemleri ile periferik yayma görüntülerinin analizi ve sınıflandırılması (Doctoral Dissertation).
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.
- Boztoprak, H. (2014). Görüntü işleme teknikleri ve yapay zeka yöntemleri kullanarak Atıksu arıtmada performans analizlerinin incelenmesi (Doctoral dissertation, Selçuk Üniversitesi Fen Bilimleri Enstitüsü).
- Branco, P., Ribeiro, R. P., & Torgo, L. (2016). UBL: an R package for Utilitybased Learning. <http://arxiv.org/abs/1604.08079>
- Branco, P., Ribeiro, R. P., Torgo, L., Krawczyk, B., & Moniz, N. (2017). SMOGN: a Pre-processing Approach for Imbalanced Regression. *Proceedings of Machine Learning Research*, 74, 36–50.
- Branco, P., Torgo, L., & Ribeiro, R. (2015). A Survey of Predictive Modelling under Imbalanced Distributions. 1–48. <http://arxiv.org/abs/1505.01658>
- Brodsky, J., & Radvanovsky, R. (2013). Sociological and cultural aspects. *Handbook of SCADA/Control Systems Security*, 15-28.
- Bulut, F. (2016). Sınıflandırıcı Topluluklarının Dengesiz Veri Kümeleri Üzerindeki Performans Analizleri. *Bilişim Teknolojileri Dergisi*, 9(2), 153.
- Cheung,S., Dutertre,B. Fong, M., Ulf Lindqvist, Keith Skinner, and Alfonso Valdes. Using model-based intrusion detection for scada networks. In *Proceedings of the SCADA security scientific symposium*, volume 46, pages 1–12. Citeseer, 2007.
- Chouchoulas, A., & Shen, Q. (2001). Rough set-aided keyword reduction for text categorization. *Applied Artificial Intelligence*, 15(9), 843–873. <https://doi.org/10.1080/088395101753210773>
- Civelekoğlu, G. 2006. Arıtma Proseslerinin Yapay Zeka Ve Çok İstatistiksel Yöntemler İle Modellenmesi. Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü Çevre Mühendisliği Bölümü, Doktora Tezi.
- Cornelissen, E. (2018). Prediction of wastewater treatment plants process performance parameters based on microbial communities using machine learning techniques (Doctoral dissertation).

- Criminisi, A., Shotton, J., & Konukoglu, E. (2012). Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning.
- Criminisi, A., Shotton, J., & Konukoglu, E. (2011). Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning. *Microsoft Research Cambridge, Tech. Rep. MSRTR-2011-114*, 5(6), 12.
- Çakmakçı, M. (2001). Atıksu toplama şebekeleri için otomatik kontrol sistemi geliştirme esasları (Doctoral dissertation, Fen Bilimleri Enstitüsü).
- Çatal, C, Ece, K, Arslan, B, Akbulut, A. (2019). Satış Tahmini için Regresyon Algoritmalarının ve Zaman Serisi Analiz Tekniklerinin Kıyaslanması. *Balkan Elektrik ve Bilgisayar Mühendisliği Dergisi*, 7 (1), 20-26. DOI: 10.17694 / bajece.494920.
- Çinar, Ö., & Yılmaz, A. S. Yapay Sinir Ağlarının Atıksu Arıtma Tesisi İşletimine Uygulanması: Bir Örnek Çalışma.
- D. Torregrossa, U. Leopold, F. Hernández-Sancho, ve J. Hansen, "Machine learning for energy cost modelling in wastewater treatment plants", *J. Environ. Manage.*, c. 223, ss. 1061-1067, Eki. 2018.
- Dieu, B. (2001). Application of the SCADA system in wastewater treatment plants. *ISA transactions*, 40(3), 267-281.
- Dinis, C. M., & Popa, G. N. (2014). Measurements in SCADA system used at a wastewater treatment plant. *Annals of the Faculty of Engineering Hunedoara*, 12(4), 207.
- Doğan, E., Ateş, A., Yılmaz, E.C. ve Eren, B., 2008. Application of Artificial Neural Networks to Estimate Wastewater Treatment Plant Inlet Biochemical Oxygen
- Düğenci, M., 2007. Arı Algoritması'nın Yapay Sinir Ağı Öğrenmesi İçin Kullanımı ve Atıksu Arıtma Tesis Kontrolü Uygulaması, Doktora Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü. Sakarya.
- Eren, B., Ates, A., Yılmaz, E, C., Dogan, E. (2008). Application of Artificial Neural Networks to Estimate Wastewater Treatment Plant Inlet Biochemical Oxygen Demand. *Environmental Progress* 27(4):439-446
- Fatih TOPTAŞ (ÇOB) Ülke Raporu, 03.03.2008
- Feng, C., Li, T., Zhu, Z., & Chana, D. (2017). A deep learning-based framework for conducting stealthy attacks in industrial control systems. arXiv preprint arXiv:1709.06397.
- Gabrali, D., & Aslan, Z. (2020). Güneş Enerjisi Potansiyelinin Çoklu Lineer Regresyon ve Yapay Sinir Ağları ile Modellenmesi. *AURUM Mühendislik Sistemleri ve Mimarlık Dergisi*, 4(1), 23-36.
- Galitsky, B. (2019). Developing enterprise chatbots: Learning linguistic structures. *Developing Enterprise Chatbots: Learning Linguistic Structures* (pp. 1–559). Springer International Publishing.
- Gheytaspour, M., & Bigdarvish, O. H. (2018). Forecasting Oxygen Demand in Treatment Plant Using Artificial Neural Networks. March.

- Gök, M. (2017). Makine Öğrenmesi Yöntemleri ile Akademik Başarının Tahmin Edilmesi. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 5(3), 139–148. <https://dergipark.org.tr/gujsc/issue/31140/311082>
- Granata, F., Papirio, S., Esposito, G., Gargano, R., & De Marinis, G. (2017). Machine learning algorithms for the forecasting of wastewater quality indicators. *Water*, 9(2), 105.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. P., ... & Cho, K. H. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences*, 32, 90-101.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J., Kim, J. H., & Cho, K. H. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models ScienceDirect Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *JES*, 32(December 2017), 90–101.
- Güçlü, D. (2007). Tam ölçekli kentsel atık su arıtma tesislerinin bilgisayar programı kullanılarak modellenmesi ve arıtma performanslarının incelenmesi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Doktora Tezi.
- Gülsoy, İ. (2017). Kentsel Atıksu Arıtma Tesisinin online izlenmesi ile enerji etüt ve optimizasyonu.
- Gültepe, Y. (2019). Makine Öğrenmesi Algoritmaları ile Hava Kirliliği Tahmini Üzerine Karşılaştırmalı Bir Değerlendirme. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 8-15.
- Güzelbeyoğlu, N. (2002). Bayramiç Kazası Manuel Atıksu Tesisinin Scada Otomasyonu (Doctoral dissertation, Fen Bilimleri Enstitüsü).
- Häck, M. ve Köhne, M. 1996. Estimation of Wastewater Process Parameters Using Neural Networks. *Water Science and Technology*, 33, 1, 101-115.
- Hadžiosmanović, D., Bolzoni, D., & Hartel, P. H. (2012). A log mining approach for process monitoring in SCADA. *International Journal of Information Security*, 11(4), 231-251.
- Hager, W. H. (1999). *Wastewater Hydraulics: Theory And Practice*. Berlin: Springer.
- Hamada, M., Zaqoot, H. A., & Jreiban, A. A. (2018). Application of artificial neural networks for the prediction of Gaza wastewater treatment plant performance-Gaza strip. *Journal of Applied Research in Water and Wastewater*, 9(1), 399–406.
- Hamed, M. M., Khalafallah, M. G., & Hassanien, E. A. (2004). Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling & Software*, 19(10), 919-928.
- Heddami, S., Lamda, H. ve Filali, S., 2016. Predicting Effluent Biochemical Oxygen Demand in a Wastewater Treatment Plant Using Generalized Regression Neural Network Based Approach: a Comparative Study. *Environmental Processes*, 16, 129-3.

- Hoque, N., Bhuyan, M. H., Baishya, R. C., Bhattacharyya, D. K., & Kalita, J. K. (2014). Network attacks: Taxonomy, tools and systems. *Journal of Network and Computer Applications*, 40, 307-324.
- Hu Guozhen, Cai Tao, Chen Changsong, Duan Shanxu, Solutions for SCADA system Communication Reliability in Photovoltaic Power Plants, IEEE 6th International Conference Power Electronics and Motion Control IPEMC'09, 17- 20 May 2009, Wuhan, China, pp. 2482- 2485.
- Huang, C. C., Tseng, T. L. B., Fan, Y. N., & Hsu, C. H. (2013). Alternative rule induction methods based on incremental object using rough set theory. *Applied Soft Computing*, 13(1), 372-389.
- Huitsing, P., Chandia, R., Papa, M., and Shenoi, S., Attack Taxonomies for the Modbus Protocols. *International Journal of Critical Infrastructure Protection*, 1:37-44, 2008.
- J. Mathew, S.S. Shankar et al., Implementation of high availability SCADA system for superheater steam temperature control in a 210 MW thermal power plant, IEEE International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCT), 2014 , 6 – 7 Jan., Bangalore, India, pp.1- 6.
- Jain, A. (2016). Learning From Natural Human Interactions For Assistive Robots. Proquest Dissertations And Theses. Retrieved From [Http://0-Search.Proquest.Com.Mylibrary.Qu.Edu.Qa/Docview/1844397898](http://0-Search.Proquest.Com.Mylibrary.Qu.Edu.Qa/Docview/1844397898)  
<http://asheshjain.org/pdfs/AsheshJainPhdThesis.pdf>
- Jain, A. (2016). Learning From Natural Human Interactions For Assistive Robots. Proquest Dissertations And Theses. Retrieved From [Http://0-Search.Proquest.Com.Mylibrary.Qu.Edu.Qa/Docview/1844397898](http://0-Search.Proquest.Com.Mylibrary.Qu.Edu.Qa/Docview/1844397898)
- Jensen, R., & Cornelis, C. (2010). Fuzzy-rough instance selection. 2010 IEEE World Congress on Computational Intelligence, WCCI 2010, August. <https://doi.org/10.1109/FUZZY.2010.5584791>
- Jensen, R., & Shen, Q. (2002). Fuzzy-rough sets for descriptive dimensionality reduction. *IEEE International Conference on Fuzzy Systems*, 1(June), 29-34. <https://doi.org/10.1109/fuzz.2002.1004954>
- Jihyun Kim, Jaehyun Kim, Huong Le Thi Thu, and Howon Kim. Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection. In *Proceedings of the International Conference on Platform Technology and Service (PlatCon)*, pages 1-5. IEEE, 2016.
- Jouanneau, S., Grangé, E., Durand, M. J., & Thouand, G. (2019). Rapid BOD assessment with a microbial array coupled to a neural machine learning system. *Water research*, 166, 115079.
- Keleş, R., Açıkalın, S. ve İleri, R., (1999). Adapazarı Kentsel Atıksu Arıtma Tesisi Çıkış Suyu Parametreleri ve Verim Değerlerinin Yapay Sinir Ağları İle Tahmin Edilmesi. Sakarya Üniversitesi, Müh. Fak., Çevre Müh. Bölümü, Esentepe Kampusu, Sakarya.
- Kor, N (1974). Çevre Sağlığı Ve Teknolojisi, 1. Cilt. İstanbul: Teknik Üniversitesi Matbaası

- Küçükpelvan, H., Yarimtepe, C. C., & Öz, N. A. (2017). Deri Atıksuyunun Arıtım Metotları. Çanakkale Onsekiz Mart Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 3(1), 59-96.
- Lakhoua, M. N., & Jbira, M. K. (2012). Project Management Phases of a SCADA System for Automation of Electrical Distribution Networks. *International Journal of Computer Science Issues (IJCSI)*, 9(2), 157.
- Lantz, B., (2015). *Machine Learning With R*, Packt Publishing, Birmingham.
- Lee, J.W., Suh, C., Hong, Y.S.T. ve Shin, H.S., 2011. Sequential Modelling of a Full-Scale Wastewater Treatment Plant Using an Artificial Neural Network. *Bioprocess Biosyst. Eng.*, 34, 963–973.
- Libhaber, M., Orozco-Jamamillo, A. (2012). *Sustainable Treatment And Reuse Of Municipal Wastewater*. London: Iwa.
- Mafarja, M. M., & Mirjalili, S. (2019). Hybrid binary ant lion optimizer with rough set and approximate entropy reducts for feature selection. *Soft Computing*, 23(15), 6249–6265. <https://doi.org/10.1007/s00500-018-3282-y>
- Manav, N., & Orcid, D. E. M. İ. R. (2017). Use Of Artificial Neural Networks As A Tool To Predict Carbon And Nitrogen Removal Efficiencies In Biological Wastewater Treatment Plants. 6(2), 375–386.
- Manna, S., Biswas, S., Kundu, R., Rakshit, S., Gupta, P., & Barman, S. (2018). A statistical approach to predict flight delay using gradient boosted decision tree. *ICCIDS 2017 - International Conference on Computational Intelligence in Data Science*,
- Mehmet Güleç, F. (2013). *Kaba Küme Teorisinin Literatür Tabanlı Bilgi Keşfine Uygulanması*.
- Mercier, T., Dembele, A., Denoeux, T., & Blanc, P. (2019). Machine Learning As A Decision Support Tool For Wastewater Treatment Plant Operation. *Wit Transactions On Ecology And The Environment*, 229, 103-107.
- Miklovic, D., & Liptak, B. G. (2002). Integration of DCS, PLC, HMI and SCADA systems. In *Process Software and Digital Networks*. CRC Press.
- Moniz, N., Branco, P., & Torgo, L. (2017). Resampling strategies for imbalanced time series forecasting. *International Journal of Data Science and Analytics*, 3(3), 161–181. <https://doi.org/10.1007/s41060-017-0044-3>
- Moreno-alfonso, N., & Redondo, C. F. (2018). Intelligent waste-water treatment with neural-networks. 7017(March). [https://doi.org/10.1016/S13667017\(01\)00071-X](https://doi.org/10.1016/S13667017(01)00071-X)
- Morris, T. and Gao, W. Industrial Control System Traffic Data Sets for Intrusion Detection Research. *Advances in Information and Communication Technology Critical Infrastructure Protection VIII*, pages 65–78, 2014.
- Nader, P., Honeine, P., & Beuseroy, P. (2016, April). Detection of cyberattacks in a water distribution system using machine learning techniques. In *2016 Sixth International Conference on Digital Information Processing and Communications (ICDIPC)* (pp. 25-30). IEEE.

- Najafzadeh, M., & Zeinolabedini, M. (2019). Prognostication of waste water treatment plant performance using efficient soft computing models: an environmental evaluation. *Measurement*, 138, 690-701.
- Nascu, I. (2018). Hierarchical predictive control of Wastewater Treatment Plants. In *MATEC Web of Conferences* (Vol. 210, p. 02002). EDP Sciences.
- Nourani, V., Elkiran, G., & Abba, S. I. (2018). Wastewater treatment plant performance analysis using artificial intelligence—an ensemble approach. *Water Science and Technology*, 78(10), 2064-2076.
- O'Shea, T. J., & Hoydis, J. (2017). An introduction to machine learning communications systems. arXiv preprint arXiv:1702.00832.
- Oliveira-Esquerre, K.P., Mori, M. ve Bruns, R.E., 2002. Simulation of an Industrial Wastewater Treatment Plant Using Artificial Neural Networks and Principal Components Analysis. *Brazilian Journal of Chemical Engineering*, 19, 365-370.
- ORÇANLI, K. (2019). Kalite Kontrol Grafiklerinde R Programlama Dilinin Kullanımı ile İlgili İçerik Analizi. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 13(19), 1390-1429.
- Öztemel, E., & Düğencİ, M. (2016). Atıksu Arıtma Tesis Kontrolde Yapay Sinir Ağı ile Kirlilik Parametre Tahmini. November, 4–6.
- Öztürk, K., & Şahin, M. E. (2018). Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış. *Takvim-i Vekayi*, 6(2), 25–36. <http://www.sloi.org/sloi-name-of-thisarticle>
- Öztürk, N. (2014). İçme suyu arıtma tesisi atık çamurunun sulu çözültiden bazı kirletici türlerin adsorpsiyonla gideriminde kullanılabilirliğinin araştırılması ve Yapay Sinir Ağları ile modellenmesi. Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Kimya Anabilim Dalı, Doktra Tezi.
- Öztürk, İ. Ve Timur,H., ( 2017).“Atıksu Mühendisliği” Teknik Kitaplar Serisi, İSKİ, İstanbul.
- Panni, O., West, G. M., Catterson, V. M., McArthur, S. D. J., Shi, D., & Mogridge, I. (2016). Implementation Of A Bayesian Linear Regression Framework For Nuclear Prognostics. Third European Conference of the Prognostics and Health Management Society, July.
- Parthasarathy, P., & Narayanan, S. K. (2014). Effect of Hydrothermal Carbonization Reaction Parameters on. *Environmental Progress & Sustainable Energy*, 33(3), 676–680.
- Pawlak, Z., & Skowron, A. (2007). Rudiments of rough sets. *Information Sciences*, 177(1), 3–27. <https://doi.org/10.1016/j.ins.2006.06.003>
- Pitas, V., Fazekas, B., Banyai, Z., & Karpati, A. (2010). Energy efficiency of the municipal wastewater treatment. *Journal of Biotechnology*, (150), 163-164.
- Pratt, L., Y., (1994), Experiments on the transfer of knowledge between neural networks. In S. Hanson, G. Drastal, and R. Rivest, editors, *Computational Learning Theory and Natural Learning Systems, Constraints and Prospects*, volume 1, chapter 19, pages MIT Press, franklinite.mines.colorado.

- Qin, X., Gao, F., & Chen, G. (2012). Wastewater quality monitoring system using sensor fusion and machine learning techniques. *Water research*, 46(4), 1133-1144.
- Ribeiro, D., Sanfins, A., & Belo, O. (2013, July). Wastewater treatment plant performance prediction with support vector machines. In *Industrial Conference on Data Mining* (pp. 99-111). Springer, Berlin, Heidelberg.
- Riza, L. S., Janusz, A., Bergmeir, C., Cornelis, C., Herrera, F., Ślęzak, D., & Benítez, J. M. (2014). Implementing algorithms of rough set theory and fuzzy rough set theory in the R package “roughSets.” *Information Sciences*, 287, 68–89. <https://doi.org/10.1016/j.ins.2014.07.029>
- Sağiroğlu, Ş., Erler, M., & Beşdok, E., (2003). *Mühendislikte Yapay Zeka Uygulamaları - I : Yapay Sinir Ağları* . Kayseri: Ufuk Kitabevi.
- Sai Prasad, P. S. V. S., & Rao, C. R. (2009). IQuickReduct: An improvement to quick reduct algorithm. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5908 LNAI, 152–159. [https://doi.org/10.1007/978-3-642-10646-0\\_18](https://doi.org/10.1007/978-3-642-10646-0_18)
- Sakarya, C., Selv, E., Dal, A., Dan, T., & Eren, B. (2019). Kentsel atıksu arıtma tesisi biyolojik oksijen ihtiyacının (boi 5 ) makina öğrenmesi yöntemleri ile tahmin edilmesi yüksek lisans tezi.
- Sanchez, A., Katebi, MR ve Johnson, MA (2003, Haziran). Atıksu endüstrisinde gelişmiş kontrol sistemlerinin ve veri kalitesi yönetiminin test edilmesi için gerçek zamanlı bir kontrol platformunun tasarımı ve uygulaması. In *Kontrol ve Otomasyon Proceedings 2003 4. Uluslararası Konferansı* (s. 431-435). IEEE.
- Saurabh Amin, Xavier Litrico, S. Shankar Sastry and Alexandre M. Bayen, *Cyber Security of Water SCADA Systems: (I) Analysis and Experimentation of Stealthy Deception Attacks*, *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY*, VOL. XX, NO. X, 2012, pp.1- 8.
- Schubach, M., Re, M., Robinson, P. N., & Valentini, G. (2017). Imbalance-Aware Machine Learning for Predicting Rare and Common Disease-Associated NonCoding Variants. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-01703011-5>
- Selvi, E. (2019). Kentsel Atıksu Arıtma Tesisi biyolojik oksijen ihtiyacının (BOİ5) makina öğrenmesi yöntemleri ile tahmin edilmesi (Master's thesis, Sakarya Üniversitesi).
- Selvi, E. (2019). Kentsel atıksu arıtma tesisi biyolojik oksijen ihtiyacının (boi5) makina öğrenmesi yöntemleri ile tahmin edilmesi (Master's thesis, Sakarya Üniversitesi).
- Sezer, M., 2007. *Yapay Sinir Ağları (YSA) Kullanılarak KOİ Parametresinden BOİ Parametresinin Tahmin Edilmesi*, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.
- Shalev-Shwartz, S., Ben-David, S., *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

- Stamp, J., Campbell, P., DePoy, J., Dillinger, J. ve Young, W. (2003). Altyapı SCADA için sürdürülebilir güvenlik. Sandia Ulusal Laboratuvarları, Albuquerque, New Mexico (www. Sandia. Gov / scada / belgeler / SustainableSecurity. Pdf) .
- Subaşı, H., & Ersü, Ç. B. (2010). Yapay Sinir Ağı ile Atıksu Arıtma Performansının Modellenmesi (Doctoral dissertation, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Çukurova Üniversitesi).
- Suchetana, M., Aninda, M., Sudha, G., & Datta, A. K. (2013). Pollination events in *Nigella sativa* L.(Black cumin). *International Journal of Research in Ayurveda and Pharmacy (IJRAP)*, 4(3), 342-344.
- Şevki, I. (2008). Bulanık Mantık ve Bulanık Teknolojiler. Araştırma Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Felsefe Bölümü Dergisi (Yayınlanmıyor), 19(0), 105–126. Üniversitesi, K. T. (1999). Bulanık Mantık : Bulanıklılık Kavramı, 80–85.
- Telçeken, S., & Doğan, M. (2004) Kaba Kümeler Teorisi Yardımı İle Büyük Veri Topluluklarının Analizi.
- Tezel, G., Yel, E., & Sinan, R. K. (2010, May). Artificial neural network (Ann) model for domestic wastewater treatment plant control. In The fourth scientific conference on water observation and information system for decision support (BALWOIS), Ohrid, Republic of Macedonia.
- Topacık, D., Koyuncu, İ. (2006). Arıtma Tesislerinin Hidroliği. İstanbul: Su Vakfı.
- Torgo, L., Ribeiro, R. P., Pfahringer, B., & Branco, P. (2013). SMOTE for regression. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8154 LNAI(September), 378–389. [https://doi.org/10.1007/978-3-642-40669-0\\_33](https://doi.org/10.1007/978-3-642-40669-0_33)
- Tsang, E. C. C., Hu, Q., & Chen, D. (2016). Feature and instance reduction for PNN classifiers based on fuzzy rough sets. *International Journal of Machine Learning and Cybernetics*, 7(1), 1–11. <https://doi.org/10.1007/s13042-014-0232-6>
- Turgay, S., Torkul, O. (2017) KABA KÜMELEME, Mühendislikte yapay zeka ve uygulamaları Sakarya Üniversitesi Kütüphanesi Yayınevi
- Turovskiy, I. S., & Mathai, P. K. (2006). Wastewater sludge processing. John Wiley & Sons.
- Turovsky, I.S., Mathai, P.K. (2005). Wastewater Sludge Processing ,John Wiley & Sons, Inc. (Sayfa: 173-212).
- TÜRKAK Test Raporu/2017
- Türkiye'deki Küçük ve Orta Ölçekli Belediyelerde Atıksu Arıtımı için Rehber El Kitabı, Ekim 2010.
- Verbiest, N. (2014). Fuzzy Rough and Evolutionary Approaches to Instance Selection. March.
- Vluymans, S. (2019). Dealing with imbalanced and weakly labelled data in machine learning using fuzzy and rough set methods. Switzerland: Springer International Publishing.



- Wang, K., Pang, Y., He, C., Li, P., Xiao, S., Sun, Y., ... & He, D. (2019). Xiangxi Körfezi'ndeki çözünmüş organik maddenin optik ve moleküler imzaları ve Three Gorges Reservoir, Çin'in ana akımı: Mekansal varyasyonlar ve çevresel etkiler. *Toplam Çevre Bilimi*, 657, 1274-1284.
- Yang Haijing, Yang Yihan, Zhang Dongying, The structure and application of flexible SCADA, Power Engineering Society General Meeting, 2006. IEEE, Montreal, Quebec, Canada, pp.1- 8
- Yang, H., Cheng, L., & Chuah, M. C. (2019). Deep-Learning-Based Network Intrusion Detection For Scada Systems. In 2019 Ieee Conference On Communications And Network Security, Cns 2019. Institute Of Electrical And Electronics Engineers Inc.
- Yılmaz, E.C., 2009. Bir Atıksu Arıtma Tesisinin Girişindeki Biyolojik Oksijen İhtiyacının Yapay Sinir Ağları Kullanılarak Modellenmesi. Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Çevre Mühendisliği Bölümü, Yüksek Lisans Tezi.
- Yılmaz, E. C., & Dogan, E. (2008). Artificial Neural Networks Application for Modelling of Wastewater Treatment Plant Performance. *Electronic Letters on Science and Engineering*, 4(2), 1-9.
- Ying, C., Klein, A., Christiansen, E., Real, E., Murphy, K., & Hutter, F. (2019, Mayıs). Nas-bench-101: Yeniden üretilebilir sinir mimarisi araştırmasına doğru. In *Makine Öğrenmesi Konferansı* (s. 7105-7114). PMLR.
- Yuksel, S. (2019). Araştırmalarda karar süreci ve bayes teorimi. December 2018.
- Yurtoğlu, H. (2005). Yapay Sinir Ağları metodolojisi ile öngörü modellemesi: bazı makroekonomik değişkenler için Türkiye örneği. DPT.
- Zhang, Y. (1994). A Foundation For The Design And Analysis Of Robotic Systems And Behaviors. Proquest Dissertations And Theses. Retrieved From [Http://Login.Proxy.Library.Vanderbilt.Edu/Login?Url=Http://Search.Proquest.Com/Docview/275928607](http://Login.Proxy.Library.Vanderbilt.Edu/Login?Url=Http://Search.Proquest.Com/Docview/275928607).
- Zhang, Z., Kusiak, A., Zeng, Y., & Wei, X. (2016). Modeling and optimization of a wastewater pumping system with data-mining methods. *Applied energy*, 164, 303-311.
- Zhao, L., Dai, T., Qiao, Z., Sun, P., Hao, J., & Yang, Y. (2020). Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. *Process Safety and Environmental Protection*, 133, 169-182.

## ÖZGEÇMİŞ

**Adı Soyadı** : Muhammed Alperen ŞERİFOĞLU

### ÖĞRENİM DURUMU

Derece	Eğitim Birimi	Mezuniyet Yılı
Yüksek Lisans	Sakarya Üniversitesi / Fen Bilimleri Enstitüsü / Endüstri Mühendisliği	2021
Lisans	Sakarya Üniversitesi / Mühendislik Fakültesi / Çevre Mühendisliği	2015
Lise	Muğla Anadolu Öğretmen Lisesi	2008

### İŞ DENEYİMİ

Yıl	Yer	Görev
2019-Halen	Sakarya Su ve Kanalizasyon İdaresi Genel Müdürlüğü / Atıksu Arıtmalar Şb. Md./ Çevre ve Kalite Yönetim Birimi	Mühendis
2016-2019	Sakarya Su ve Kanalizasyon İdaresi Genel Müdürlüğü / İçmesuyu Arıtmalar Şb. Md.	Arıtma Tesisleri ve Dezenfeksiyon Birim Sorumlusu
2015-2016	Abacı Group/ Abacı Kimya / Abacı Su ve Çevre Teknolojileri	Bölge Müdürü
2015-2016	Sosyobyte Dijital Medya Ajansı	Şirket Ortağı
2015-2016	Şerifoğlu Mühendislik Ltd. Şti.	Satış Planlama Sorumlusu

### YABANCI DİL

İngilizce

**ESERLER (makale, bildiri, proje vb.)**

1. İçmesuyu Arıtma Tesislerinin Personelsiz Yönetilmesi
2. Atıksu Arıtma Tesislerinin Personelsiz Yönetilmesi
3. Tavuk Gübresini Termik Yakma İle Yakarak Isı, Enerji Edesi ve Bertarafı
4. Enerji Bitkilerini (C4 vb bitkiler) Termik Yakma İle Yakarak Elektrik Üretimi
5. Endüstriyel Tesislerden Oluşan Hava Kirliliği, Baca Gazı Arıtma Tesisleri

**HOBİLER**

Proje/organizasyon/etkinlik yönetimi, ekip yönetimi, at biniciliği, yüzme, basketbol, sosyal sorumluluk projeleri, sivil toplum kuruluşlarında faaliyetlerde bulunmak.