

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İNŞAAT MÜHENDİSLİĞİ ANABİLİM DALI**

**SAKARYA HAVZASINA AİT SU KALİTESİ
PARAMETRELERİNİN YAPAY ZEKÂ YÖNTEMLERİ
İLE MODELLENMESİ**

**Hazırlayan
Yusuf ÖZEREN**

**Danışman
Doç. Dr. Hatice ÇITAKOĞLU**

Yüksek Lisans Tezi

**Temmuz 2022
KAYSERİ**

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İNŞAAT MÜHENDİSLİĞİ ANABİLİM DALI**

**SAKARYA HAVZASINA AİT SU KALİTESİ
PARAMETRELERİNİN YAPAY ZEKÂ YÖNTEMLERİ
İLE MODELLENMESİ
(Yüksek Lisans Tezi)**

**Hazırlayan
Yusuf ÖZEREN**

**Danışman
Doç. Dr. Hatice ÇITAKOĞLU**

**Temmuz 2022
KAYSERİ**

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Yusuf ÖZEREN

İmza

“Sakarya Havzasına Ait Su Kalitesi Parametrelerinin Yapay Zekâ Yöntemleri İle Modellenmesi” adlı Yüksek Lisans Tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ ne uygun olarak hazırlanmıştır.

Hazırlayan

Yusuf ÖZEREN

İmza

Danışman

Doç. Dr. Hatice ÇITAKOĞLU

İmza

İnşaat Mühendisliği ABD Başkanı

Prof. Dr. Ahmet Alper ÖNER

İmza

TEŞEKKÜR

Bu tez çalışmamın her aşamasında bana destek olan, bilgi ve deneyimlerini paylaşan, çalışmalarım süresince her türlü yardımı ve fedakârlığı sağlayan, her daim güler yüzle tezimi tamamlamam konusunda beni motive eden değerli danışman hocam Doç. Dr. Hatice ÇITAKOĞLU'na teşekkürlerimi sunuyorum.

Ayrıca; çalışmalarım süresince hiçbir zaman desteğini esirgemeyen ailem ve amcam Mehmet ÖZEREN ile hayatımın her anında olduğu gibi tez çalışmam sürecinde de fedakârca yanımda olan sevgili eşim Ayşe ÖZEREN'e en içten sevgi ve teşekkürlerimi sunarım.

Yusuf ÖZEREN

Temmuz 2022, KAYSERİ

SAKARYA HAVZASINA AİT SU KALİTESİ PARAMETRELERİNİN YAPAY ZEKÂ YÖNTEMLERİ İLE MODELLENMESİ

Yusuf ÖZEREN

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Yüksek Lisans Tezi, Temmuz 2022
Danışman: Doç. Dr. Hatice ÇITAKOĞLU

ÖZET

Yaşam için en önemli kaynaklardan birisi olan su, yeryüzünde kısıtlı miktarda bulunmakta ve ne yazık ki hızlı nüfus artışı, tarımsal ve endüstriyel faaliyetler neticesiyle kirlenmekte ve bu durum telafisi zor durumlara neden olmaktadır. Su kalitesinin izlenmesi çalışmalarında çözünmüş oksijen (ÇO) parametresi önemli bir yere sahiptir.

Birçok alanda kullanılan yapay zekâ teknikleri, su yönetim sürecinde de kullanılmakta ve “su kalitesi” alanındaki çalışmaların sonuçlarında önemli kazanımlar sağlamaktadır. Bu çalışmada, Sakarya Havzası’nda bulunan gözlem istasyonlarına ait 1995–2014 yılları arasındaki; Şubat, Nisan, Haziran, Ağustos ve Kasım aylarında ölçülmüş su kalitesi parametrelerine ait veriler ile ÇO değerinin Yapay Sinir Ağları (YSA), Derin Öğrenme (DÖ), Destek Vektör Makinesi (DVM) ve Gauss Süreç Regresyon (GSR) yöntemleri ile modellenmesi yapılmıştır. Modellerde giriş verisi olarak; biyolojik oksijen ihtiyacı (BOİ), sıcaklık (T), elektriksel iletkenlik (Eİ), aylar ve pH kullanılmış olup, çıkış verisi ÇO parametresidir. Modellerin tahmin performanslarını ölçmek için, Kök Ortalama Karesel Hata (KOKH), Ortalama Mutlak Hata (OMH), determinasyon katsayısı (R^2) ve Nash-Sutcliffe verimlilik katsayısı (NSE) değerleri göz önünde bulundurulmuştur. Ayrıca, Taylor ve Violin diyagramları yardımıyla ölçülmüş veriler ile modellerden elde edilen değerlerin birbirine ne kadar yakın olduğu incelenmiştir. Bu çalışmada, modeller içerisinde en iyi performansa ulaşabilmek için birçok deneme-yanılma yapılmış ve yapılan analizler sonucunda YSA modeli, diğer yöntemlere göre daha iyi sonuç vermiştir. Ayrıca kullanılan yöntemlerin etkinliği Kruskal-Wallis (KW) testi ile kanıtlanmıştır.

Anahtar Kelimeler: Sakarya Havzası, Su Kalitesi Parametreleri, Çözünmüş Oksijen, Yapay Zekâ Yöntemleri, YSA, DÖ, DVM, GSR

MODELING THE WATER QUALITY PARAMETERS OF SAKARYA BASIN WITH ARTIFICIAL INTELLIGENCE METHODS

Yusuf ÖZEREN

Erciyes University, Graduate School of Natural and Applied Sciences

Master Thesis, July 2022

Supervisor: Assoc. Prof. Dr. Hatice ÇITAKOĞLU

ABSTRACT

Water, which is one of the most important resources for life, is found in a limited amount on earth and unfortunately it is polluted as a result of rapid population growth, agricultural and industrial activities, and this situation causes difficult situations. The dissolved oxygen (DO) parameter has an important place in the monitoring of water quality.

Artificial intelligence techniques used in many fields are also used in the water management process and provide significant gains in the results of studies in the field of "water quality". In this study, between the years 1995-2014, the observation stations in the Sakarya Basin; the data of the measured water quality parameters in February, April, June, August and November and the DO value were modeled using Artificial Neural Networks (ANN), Deep Learning (DL), Support Vector Machine (SVM) and Gaussian Process Regression (GPR) methods. As input data in models; biological oxygen demand (BOD), temperature (T), electrical conductivity (EC), months and pH are used and output data is DO parameter. In order to measure the estimation performance of the models, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), coefficient of determination (R^2) and Nash-Sutcliffe coefficient of efficiency (NSE) values were taken into consideration. In addition, it was examined how close the measured data and the values obtained from the models were to each other with the help of Taylor and Violin diagrams. In this study, in order to reach the best performance among the models, many trial and error were made and as a result of the analyzes, the ANN model gave better results than other methods. In addition, the effectiveness of the methods used has been proven by the Kruskal-Wallis (KW) test.

Keywords: Sakarya Basin, Water Quality Parameters, Dissolved Oxygen, Artificial Intelligence Methods, ANN, DL, SVM, GPR

İÇİNDEKİLER

SAKARYA HAVZASINA AİT SU KALİTESİ PARAMETRELERİNİN YAPAY ZEKÂ YÖNTEMLERİ İLE MODELLENMESİ

BİLİMSEL ETİĞE UYGUNLUK	ii
YÖNERGEYE UYGUNLUK.....	iii
KABUL VE ONAY	iv
TEŞEKKÜR.....	v
ÖZET.....	vi
ABSTRACT.....	vii
İÇİNDEKİLER	viii
KISALTMALAR	xii
TABLolar LİSTESİ.....	xiii
ŞEKİLLER LİSTESİ	xiv
GİRİŞ	1

1. BÖLÜM

GENEL BİLGİLER ve LİTERATÜR ÇALIŞMASI

1.1. Suyun Önemi	7
1.2. Yeryüzündeki Su Potansiyeli	7
1.3. Türkiye’deki Su Potansiyeli	8
1.4. Su Kalitesi Parametreleri ve Önemi.....	9
1.4.1. Fiziksel Parametreler.....	9
1.4.1.1. Sıcaklık (T).....	9
1.4.1.2. Elektriksel İletkenlik (Eİ).....	10
1.4.1.3. Askıda Katı Madde (AKM).....	10
1.4.1.4. Renk.....	10
1.4.1.5. Bulanıklık.....	11
1.4.1.6. Suyun Kokusu	11
1.4.1.7. Suyun Tadı.....	12
1.4.1.8. Suyun pH Değeri	12
1.4.2. Kimyasal Parametreler.....	12

1.4.2.1. Çözünmüş Oksijen (ÇO)	12
1.4.2.2. Suyun Sertliği	13
1.4.2.3. Klorür (Cl ⁻)	14
1.4.2.4. Sülfat (SO ₄ ⁻²)	14
1.4.2.5. Biyolojik Oksijen İhtiyacı (BOİ).....	14
1.4.2.6. Kimyasal Oksijen İhtiyacı (KOİ).....	15
1.4.2.7. Azot (N) - Amonyak (NH ₃) - Amonyum (NH ₄).....	15
1.4.2.8. Nitrit (NO ₂) - Nitrat (NO ₃)	16
1.4.2.9. Florür (F)	16
1.4.2.10. Sodyum (Na) - Sodyum Adsorpsiyon Oranı (SAR)	16
1.4.3. Mikrobiyolojik Parametreler	16
1.4.3.1. Koliform Bakteriler	17
1.4.3.2. Toplam Koliform.....	17
1.4.3.3. Fekal Streptokoklar	17
1.4.3.4. Escherichia Coli (Koli Basili (E.koli))	18
1.4.4. Ağır Metaller	18
1.4.4.1. Arsenik (As).....	18
1.4.4.2. Kadmiyum (Cd).....	18
1.4.4.3. Siyanür (CN ⁻)	19
1.4.4.4. Civa (Hg).....	19
1.4.4.5. Kurşun (Pb)	19
1.4.4.6. Selenyum (Se)	20
1.4.4.7. Bakır (Cu)	20
1.4.4.8. Çinko (Zn).....	20
1.4.4.9. Alüminyum (Al).....	21
1.4.4.10. Demir (Fe).....	21
1.4.4.11. Bor (B).....	21
1.5. Su Kalitesi Parametreleri ve Yapay Zekâ Teknikleri.....	21

2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Sakarya Havzası.....	35
2.1.1. Gözlem İstasyonları	39
2.1.2. Alt Havzalar.....	40

2.2. Yapay Sinir Ağları (YSA).....	41
2.2.1. Yapay Sinir Ağının Yapısı.....	43
2.2.1.1. Girdi Katmanı	44
2.2.1.2. Ara (Gizli) Katman	44
2.2.1.3. Çıktı Katmanı	45
2.2.2. Yapay Sinir Hücresinin Bölümleri	45
2.2.2.1. Girdiler.....	45
2.2.2.2. Ağırlıklar.....	45
2.2.2.3. Toplama Fonksiyonu (Birleştirme Fonksiyonu)	46
2.2.2.4. Aktivasyon Fonksiyonu	46
2.2.2.5. Çıktılar	46
2.3. Derin Öğrenme (DÖ)	47
2.3.1. Sinir Ağı Türleri.....	48
2.3.1.1. Konvolüsyonel Sinir Ağları (Convolutional Neural Network).....	48
2.3.1.2. Tekrarlayan Sinir Ağı (Recurrent Neural Network).....	48
2.3.1.3. Uzun Kısa Süreli Hafıza (UKSH) (Long Short-Term Memory).....	48
2.3.1.4. Kısıtlı Boltzmann Makinesi (Restricted Boltzmann Machine)	51
2.3.1.5. Derin İnanç Ağı (DBN-Deep Belief Network)	51
2.3.1.6. Derin Oto-kodlayıcılar (Auto Encoder)	52
2.3.2. DÖ Mimarilerinde Yer Alan Katmanlar	53
2.3.2.1. Konvolüsyon Katmanı	53
2.3.2.2. Aktivasyon (Relu) Katmanı.....	53
2.3.2.3. Havuzlama (MaxPool) Katmanı	53
2.3.2.4. Tam Bağlı (Full-Connected) Katman.....	53
2.3.2.5. Dropout Katmanı	53
2.3.2.6. Normalizasyon (Ölçeklendirme) Katmanı.....	53
2.3.2.7. Entropi (Softmax) Katmanı	54
2.4. Destek Vektör Makineleri (DVM)	54
2.5. Gauss Süreç Regresyon (GSR).....	58
2.6. Materyal.....	63
2.7. Model Performanslarını Karşılaştırma Kriterleri.....	63

3. BÖLÜM

BULGULAR

3.1. YSA Model Çalışması	67
3.2. DÖ Model Çalışması	72
3.3. DVM Model Çalışması.....	74
3.4. GSR Model Çalışması.....	76

4. BÖLÜM

TARTIŞMA-SONUÇ ve ÖNERİLER

4.1.Tartışma	81
4.2. Sonuç ve Öneriler	84
KAYNAKÇA	86
ÖZGEÇMİŞ.....	92

KISALTMALAR

AB	: Avrupa Birliđi
AKM	: Askıda Katı Madde
BOİ	: Biyolojik Oksijen İhtiyacı
ÇO	: Çözünmüş Oksijen
DÖ	: Derin Öğrenme
DSİ	: Devlet Su İşleri
DVM	: Destek Vektör Makinesi
E.koli	: Escherichia Coli
Eİ	: Elektriksel İletkenlik
GSR	: Gauss Süreç Regresyon
KOİ	: Kimyasal Oksijen İhtiyacı
KOKH	: Kök Ortalama Karesel Hata
kPCA	: Kernal Temel Bileşenler Analizi
KW	: Kruskal-Wallis
M5Tree	: M5 Model Tree
MARS	: Çok Değişkenli Adaptif Regresyon Eğrileri
MATLAB	: Matrix Laboratory
NSE	: Nash-Sutcliffe Verimlilik Katsayısı
NTU	: Nephelometric Turbidity Unit
OKH	: Ortalama Karesel Hata
OMH	: Ortalama Mutlak Hata
OMYH	: Ortalama Mutlak Yüzde Hata
pH	: Potentia Hydrogenii
R	: Korelasyon Katsayısı
R²	: Determinasyon Katsayısı
SÇD	: Su Çerçeve Direktifi
T	: Sıcaklık
UKSH	: Uzun Kısa Süreli Hafıza
WHO	: Dünya Sağlık Örgütü
YSA	: Yapay Sinir Ağları

TABLÖLER LİSTESİ

Tablo 1. Sakarya Havzasına ait karakteristik değerler.....	36
Tablo 2. Sakarya Havzası'nda yer alan gözlem istasyonlarına ait bilgiler	39
Tablo 3. Biyolojik sınır sisteminin yapay sınır sistemindeki karşılığı.....	43
Tablo 4. Eğitim ve test değerlerine ait verilerin temel istatistiksel özellikleri.....	65
Tablo 5. ÇO konsantrasyon değerinin YSA modeli ile tahmininde denemiş farklı kombinasyonlar	69
Tablo 6. ÇO parametresi tahmininde modellere ait eğitim ve test verilerinin hata kriterleri	70
Tablo 7. YSA, DÖ, DVM ve GSR Modellere ait eğitim ve test aşamaları için hata kriterleri	78
Tablo 8. KW testinin %5 anlamlılık düzeyinde P değerleri	80
Tablo 9. Literatürdeki benzer çalışmalara ait değerler	82

ŞEKİLLER LİSTESİ

Şekil 1.	Sakarya Havzası ve diğer nehir havzalarının konumu	35
Şekil 2.	Sakarya Havzası'na ait alt havzalar	40
Şekil 3.	Biyolojik sinir ağı yapısı	42
Şekil 4.	Yapay Sinir Ağı yapısı	44
Şekil 5.	Yapay Sinir Hücresinin bölümleri.....	45
Şekil 6.	Tekrarlayan Sinir Ağı yapısı	48
Şekil 7.	Tekrarlayan Sinir Ağı yapısındaki tek katman.....	49
Şekil 8.	UKSH'daki dört etkileşimli katman yapısı	49
Şekil 9.	Kısıtlı Boltzmann Makinesi	51
Şekil 10.	Derin İnanç Ağı yapısı	52
Şekil 11.	Derin Oto-kodlayıcı ağ yapısı	52
Şekil 12.	DVM ile sınıflandırma	55
Şekil 13.	BOİ-Aylar değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	71
Şekil 14.	BOİ-T değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	71
Şekil 15.	BOİ-T-pH değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	71
Şekil 16.	BOİ-T-pH-Eİ değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	72
Şekil 17.	DÖ Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	74
Şekil 18.	DÖ Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	74
Şekil 19.	DVM Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	75
Şekil 20.	DVM Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	76
Şekil 21.	GSR Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	77

Şekil 22.	GSR Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri	77
Şekil 23.	Dört farklı modele ait tahminlerin Taylor diyagramı.....	79
Şekil 24.	Dört farklı modele ait tahminlerin Violin diyagramı	79



GİRİŞ

Dünyanın yaklaşık olarak dörtte üçünün sularla kaplı olması çok fazla miktarda suya sahip olduğumuzu düşündürse de insanların kullanımı için olan tatlı su miktarı ne yazık ki çok sınırlıdır. Yeryüzündeki suların büyük bir kısmı okyanus ve denizlerde bulunmaktadır. Bu sular, içme ya da sulama suyu olarak insanların faydalanabileceği nitelikli sular olmayıp, tuzlu sulardır. Yeryüzünde bu denli kısıtlı miktarda bulunan ve yaşam için hayati önem taşıyan tatlı su ihtiyacı, yüzeysel sular (nehirler, göller vb.) ve yeraltı su kaynaklarından sağlanmaktadır.

Su, canlı yaşamı için büyük bir öneme sahiptir. Ancak son yıllarda ciddi şekilde hissedilmeye başlayan ve daha da hissedilmesi beklenen iklim değişikliği ve çok hızlı değişen nüfus artışı, su kaynakları üzerinde olumsuz sonuçların ortaya çıkmasına neden olabilecektir. Çok hızlı nüfus artışı beraberinde şehirleşmeyi ve sanayileşmeyi getirmiş olup sera gazlarının salınımı ve çevre bilincinin yeterince oluşmaması, su kaynakları üzerindeki bir diğer baskı unsurlarından olmuştur. Azalan su kaynakları yanında artış eğiliminde olan su ihtiyacı, gelecekteki canlı yaşamının sürekliliği ve nitelikli yaşam açısından araştırılması gereken önemli konu başlıklarından birisi olmuştur.

Birçok yaşamsal faaliyette kullanılan ve bu denli büyük bir öneme sahip olan sular, maalesef insanların bilinçsizce yaptıkları faaliyetler neticesinde kirletilmektedir. Su kirliliği, hem günümüzde hem de gelecekte önemli sorunlardan birisi olacaktır. Bilim insanları yapmış oldukları çalışmalarda; su kaynakları üzerinde kirlilik baskısını azaltmak ve su kalitesinin korunmasını sağlamak adına birçok çalışma yapmışlardır.

Dünya üzerinde bulunan sulara oranla tatlı su miktarı çok azdır ve ülkemiz de dünyada su kaynaklarının kısıtlı olduğu coğrafyada bir bölgede yer almaktadır. Ülkemizde su yönetim sürecinin, su kaynaklarındaki bu zorluklar göz önünde bulundurularak daha dikkatli ve sürdürülebilir politikalarla sürdürülmesi gerekmektedir. Su yönetim sürecinde su kalitesi önemli bir konumda bulunmaktadır.

Yaşamımızı sürdürdüğümüz çevre devamlı bir değişim halindedir. Çevremizde yaşanan bu değişimin gelecekte nasıl sonuçları olacağını tahmin edebilmek, ileriye dönük çalışmalarda önemli kazanımlar sağlayacaktır. Birçok bilim dalında yapay zekâ tekniklerinden farklı amaçlar için faydalanılmaktadır. Yapay zekâ tekniklerinin su yönetimi sürecinde kullanılan diğer uygulamalar yanında su kalitesi parametreleri için de kullanım alanı vardır. Su kalitesine yönelik çalışmalarda geçmiş yıllardan elde edilmiş veriler ile tahminde bulunmak gözlem çalışmalarında önemli kolaylıklar sağlayacaktır.

Günümüzde gelişen teknoloji ile beraber bilgiye erişimin kolaylaşması, bilişim teknolojisi alanında yaşanan yenilikler ve yüksek işlem kapasitesine sahip bilgisayarlar ile geçmiş yıllara ait veriler daha kolay işlenebilmektedir. Verilerin toplanması, sınıflandırılması, anlamlı ve işlenebilir hale getirilmesi ile büyük veri kavramı ortaya çıkmıştır. Büyük veri kavramı, sınıflandırma ve tahmin alanında çalışmalar yürüten bilim insanları açısından önemli bir dönüm noktasıdır. Çok sayıda veriyi birlikte ve anlamlı olacak şekilde kullanmak için koordinasyon gereklidir. Bu koordinasyonla yapay zekâ tekniklerinin ve büyük verinin birlikte çalışması ile sağlanacaktır. Yapay zekâ, hayatın birçok alanında karar alma süreçlerinde yardımcı olmaktadır. Yapay zekâ teknikleri ile geçmiş yıllarda yapılmış farklı alanlarda pek çok çalışma bulunmakta ve özellikle hidrolik alanında; akım, debi, yağış, su kalitesi, buharlaşma gibi birçok parametrenin tahminine yönelik çalışmaların olduğu görülecektir.

Su kalitesi parametrelerinin yapay zekâ teknikleriyle modellenmesi ile tahmini, araştırmacıların gelecek ile ilgili planlar yapmalarında önemli kolaylıklar sağlayacaktır. Suyun kalitesi birçok farklı konsantrasyon değeriyle ifade edilmekte olup bu değerlerin her birisinin ayrı ayrı önemi vardır. Ölçümleri yapılan su kalitesi parametrelerinin süreklilik göstermesi, bu verilerin sınıflandırılma ve modellenmesi adına önemlidir.

Bu tez kapsamında, Sakarya Havzası'nda yer alan su kalitesi gözlem noktalarına ait 1995 – 2014 yılları arasındaki aylarda ölçülmüş su kalitesi parametrelerine ait veriler ile ÇO değerinin modellenmesi yapılmıştır. Modelleme verileri; T, Eİ, aylar, BOİ, pH ve ÇO'dur.

Yapay zekâ yöntemleri olan; YSA, DÖ, DVM ve GSR ile Sakarya Havzası'na ait ÇO değerinin diğer su kalitesi parametreleri olan T, Eİ, pH, BOİ ve aylar yardımıyla tahmini amacıyla modeller oluşturulmuş ve model performansları karşılaştırılmıştır.



1. BÖLÜM

GENEL BİLGİLER ve LİTERATÜR ÇALIŞMASI

Doğada bulunan kullanılabilir su miktarında değişme olmamasına rağmen, artan nüfus, tarımsal ve endüstriyel faaliyetlerle aşırı su tüketimi ve günümüzde etkisi artarak hissedilen küresel iklim değişikliği etkin bir su kaynakları yönetimi planlamasını zorunlu kılmaktadır. Su yönetim sürecinin iki önemli unsuru vardır. Bunlar; suyun miktarı ve kalitesidir. Su yönetim sürecinde en önemli amaç, su kaynaklarının optimum faydalı kullanımını sağlamak ve suyun kalitesinin korunmasıdır.

Su yönetiminde, su kalitesi önemli bir konumdadır ve suyun kalitesinin korunması ile sürdürülebilirliğinin sağlanması çok önemli gerekliliklerdendir. Suyun kalitesinin korunması canlı yaşamının devamı açısından hayati önemdedir. İnsan faaliyetleri nedeniyle sular maalesef kirletilmektedir. Su kirliliğine birçok etken sebep olmaktadır. Su kirliliğinin sebeplerinin ve sonuçlarının bilinmesi önceden önlem alınmasını kolaylaştıracaktır.

Uluslararası kuruluşlarca gerçekleştirilen birçok çalışma neticesinde, su kaynaklarının gelecekte karşı karşıya bulunduğu tehlikeler ortaya koyulmuş ve su kıtlığı ile su kirliliği bağlamında su yönetiminin küresel bir sorun olarak değerlendirilmesi amaçlanmıştır. Bu çalışmalar neticesinde ‘bütünleşik su kaynakları yönetimi’ ilkeleri ön plana çıkmıştır [1].

Suların sahip olacağı kalite, yeryüzüne yağmur veya kar olarak düşerken oluşur. Yağmur damlası oluşurken, bir toz partikülünün etrafında başlar ve toprağa düşmeden havada bulunan diğer toz partiküllerini toplar. Toz partiküllerinin yapılarında bakteri bulunmaktadır. Atmosferdeki toz yoğunluğu bakteri yoğunluğuna işaret etmektedir. Kar tanelerinin yüzeyleri geniş olduğu için atmosferdeki asılı partikülleri daha fazla tutabilirler. Yağmur kara göre daha temizdir denilebilir [2]. Doğal kirlenme diyeceğimiz bu süreçte sular hidrolojik çevrimdedir. Endüstriyel faaliyetler (özellikle fosil yakıt

kullanımı ve termik santrallerde filtre sisteminin kullanılmaması) ile oluşan hava kirliliği asit yağmurlarına neden olmaktadır. Yapısında çok fazla miktarda kimyasal olan yağışlar yeryüzüne düşer. Yağışlarla yaşanan doğal kirlenme yanında, insan faaliyetleri olan tarımsal ve endüstriyel faaliyetler ile de su kalitesi bozulmakta, kirlenen su kaynaklarının korunması için ise su kalitesinin izlenmesi gerekmektedir. Bundan dolayı su kalitesi parametrelerinin ölçümlerinin düzenli olarak yapılması gerekmektedir.

Su kalitesi parametrelerinin insan sağlığına etkisi, bunların alt ve üst sınır değerleri ile belirlenir. Ülkemizde; İnsani Tüketim Amaçlı Sular Standardı (TS 266), İçme Suyu Temin Edilen Suların Kalitesi ve Arıtılması Hakkında Yönetmelik, Yerüstü Su Kalitesi Yönetmeliği, Su Kirliliği Kontrolü Yönetmeliği gibi standart ve yönetmeliklerde kriter değerler bulunmaktadır. Yine Avrupa Birliği (AB) ve Dünya Sağlık Örgütü (WHO) gibi devletlerarası kuruluşlarda, özellikle zararlı olabilecek kimyasal ve mikrobiyolojik maddeler için sınır değerler yer almaktadır.

Uluslararası kuruluşların su politikalarına ilişkin çalışmalarında entegre nehir havzası yönetiminin esas alındığı ‘Su Çerçeve Direktifi (SÇD)’ önemli bir çalışma olarak bulunmaktadır. AB mevzuatı içerisinde su kaynaklarının korunması ve yönetimine ilişkin birçok direktif bulunması yanında, su politikalarını gözden geçirerek 23 Ekim 2000 tarihli ve 2000/60/EC sayılı ‘SÇD’ ile havza bazlı yönetim yaklaşımı benimsenmiştir. Direktif, AB sınırları içerisindeki su kaynaklarının kalite ve miktar yönünden korunmasını ve kontrol edilmesini amaçlamaktadır. Su kaynaklarının ortak bir standart kapsamında yönetimi açısından detaylı bir politika ortaya konmuştur. SÇD ile ortak yönetim anlayışı sağlanarak üye ülkelere ait yüzeysel suların (nehir, dere, göl, rezervuar vb.) ekolojik ve kimyasal bakımdan; yeraltı sularının ise miktar ve kalite yönünden iyileştirilmesi amaçlanmaktadır. Bununla havza bazlı çalışmalar ile sağlanması hedeflemektedir. SÇD, su sorununun çözümünde tüm paydaşların aktif olarak katılımında bulunmasını istemektedir. Direktif, suyun ekonomik bir değeri olduğunun kabul edilerek fiyatlandırılmasında gerçekçi ve doğru bir yaklaşımın benimsenmesini istemektedir. Su kaynaklarının sürdürülebilirliğinin sağlanması için suyu kullananın ve kirletenin bedelini ödemesi ilkesi benimsenmiştir [1].

SÇD'nin ana prensipleri olarak şunlar ön plana çıkmaktadır:

- ✓ Sürdürülebilir su kullanımı,
- ✓ Su herkesin konusudur,
- ✓ Suyun adil ücretlendirilmesi,
- ✓ Uluslararası işbirliği ve yeni su birliği,
- ✓ Su hassas bir kaynaktır şeklindedir [3].

AB adaylık süreci devam eden ülkemiz, AB mevzuatını Türk mevzuatına uyumlaştırmaya çalışmakta ve çevre politikaları kapsamında SÇD gereğince mevzuatın uyumlaştırılmasını sağlamaktadır. Ülkemizde su kaynaklarının korunmasında uygulanan, 'Su Kirliliği Kontrolü Yönetmeliği' en önemli mevzuattır. Su Kirliliği Kontrolü Yönetmeliği, su kaynaklarının korunmasına ilişkin esasları ve yasakları, atık su tesisleri ile ilgili esasları ve su kirliliğinin engellenmesi amacıyla yapılacak izleme çalışmaları ve denetleme usul ve esaslarını kapsar. Bu yönetmeliğin amaç ve kapsamının ifade edildiği 1. maddesinde; 'Bu Yönetmeliğin amacı, Ülkenin yeraltı ve yerüstü su kaynakları potansiyelinin korunması ve en iyi bir biçimde kullanımının sağlanması için, su kirlenmesinin önlenmesini sürdürülebilir kalkınma hedefleriyle uyumlu bir şekilde gerçekleştirmek üzere gerekli olan hukuki ve teknik esasları belirlemektir. Bu Yönetmelik su ortamlarının kalite sınıflandırmaları ve kullanım amaçlarını, su kalitesinin korunmasına ilişkin planlama esasları ve yasaklarını, atık suların boşaltım ilkelerini ve boşaltım izni esaslarını, atık su altyapı tesisleri ile ilgili esasları ve su kirliliğinin önlenmesi amacıyla yapılacak izleme ve denetleme usul ve esaslarını kapsar.' denilmektedir [4].

Bu yönetmelik kapsamında, su kirliliğinin engellenmesi amacıyla yapılacak izleme çalışmalarında hangi parametrelerin izleneceği ve bu parametrelere göre su kalitesi sınıflandırması yer almaktadır.

Ülkemiz ve dünyada su kaynakları üzerindeki en önemli tehdit olarak, son yıllarda etkisi her geçen yıl artan iklim değişikliği gösterilebilir. İklim değişikliği ve benzer sorunlar su kaynakları yönetim planlamasında önemli yer tutmaktadır. Akarsu havzalarında akımlarda yaşanabilecek azalma su sorunlarına neden olacaktır. Sıcaklığın artması ve akımlardaki azalmanın su kalitesine etkisi olacaktır. Su kaynaklarının yönetimi açısından

küresel ısınmanın etkilerinin de göz önünde bulundurulduğu politikalar geliştirilmesi gerekecektir.

1.1. Suyun Önemi

Su, yaşamın her alanında en çok ihtiyaç duyulan temel bir gereksinimdir. Birçok kullanım alanının olması ve yerinin doldurulamaması suyun önemini ortaya koymaktadır.

Su ve canlı yaşamı ayrılmaz iki parçadır. Su, sahip olduğu mineraller ve bileşikler ile temel bir besin maddesidir. Vücudumuzda kan dolaşımının sağlanması ve besin maddelerinin taşınması su ile olur.

Su, sucul canlılar için yaşam ortamıdır ve buradaki canlı çeşitliliği açısından önemlidir.

Değerli bir çevresel kaynak olarak su olmazsa yaşam olmaz. Tarihteki birçok büyük uygarlıklar incelendiğinde su kaynaklarının yakınında kurulduğu görülecektir.

1.2. Yeryüzündeki Su Potansiyeli

Dünyamızın dörtte üçünün sularla kaplı olması nedeniyle dünyamız için mavi küre tabiri kullanılır. Büyük bir oranda suya sahip olduğumuz anlaşılsa da tatlı su kaynakları miktarı ne yazık ki çok sınırlıdır. Dünyadaki toplam su miktarının %97.5'ni denizler ve okyanuslarda yer alan tuzlu sular oluşturmaktadır. Geri kalan %2.5'luk kısım tatlı su kaynaklarıdır. Tatlı su kaynaklarının büyük bir kısmı buzullarda ve yeraltı sularında yer alır. Çok az bir kısmı yüzey sularında ve atmosferde bulunur.

Tatlı su kaynakları, toplam su kaynakları içerisinde %2.5 gibi küçük bir kısımdır. Hayatın devamı için çok önemli olan tatlı su ihtiyacı, yüzeysel sular ve yeraltı suları kaynaklarından sağlanmaktadır. Bu miktarlarda çok kısıtlıdır. Çok kısıtlı miktarlarda bulunan tatlı su kaynakları insan faaliyetlerinin etkileri ile giderek azalmaktadır.

Su kaynaklarının dünya üzerinde dengeli bir dağılımının olmaması ve azalan tatlı su kaynakları, tüm canlılar için ülkeleri su ve su kullanımına ilişkin politikalar geliştirirken etkilemektedir.

1.3. Türkiye’deki Su Potansiyeli

Dünya’da tatlı su kaynaklarının yeterli miktarda olmamasının yanında, ülkemiz dünyada su kaynaklarının kısıtlı olduğu bir coğrafyada yer almaktadır. Ülkemizin yağış rejimi bölgesel ve mevsimsel koşullara bağlı olarak çok büyük farklılık göstermektedir.

Türkiye’de yıllık ortalama yağış yaklaşık olarak 574 mm’dir. Yılda ortalama 450 milyar m³ civarında suya karşılık gelmektedir. Günümüz teknik ve ekonomik imkânlarıyla farklı amaçlar için kullanılabilecek yerüstü suyu potansiyeli yılda ortalama toplam 94 milyar m³’tür. Yeraltı suyu potansiyeli de yaklaşık 18 milyar m³ olarak belirlenmiştir. Ülkemizin tüketilebilir yerüstü ve yeraltı su potansiyeli yılda yaklaşık toplam 112 milyar m³ olup bunun 57 milyar m³’ü kullanılmaktadır. Ülkemizde kişi başına düşen kullanılabilir yıllık su miktarı 2000 yılında 1652 m³, 2009 yılında 1544 m³, 2020 yılında ise 1346 m³ olmuştur [5].

Ülkemizde, kişi başına kullanılabilir su miktarında yıllar içerisinde önemli düşüşler olmuştur. Kişi başına kullanılabilir su miktarındaki bu düşüş ülkemizin gelecekte su sıkıntısı yaşayabileceğini göstermektedir.

Ülkemiz geneli yarı kurak bir iklime sahip olmakla birlikte kimi bölgeleri arasında aşırı sıcaklık farklılıkları vardır. İç Anadolu ve Trakya bölgelerinde sulanabilir alanlara göre yetersiz yağış nedeniyle su miktarı kısıtlıyken, Doğu Karadeniz bölgesinde tersi bir durum yani fazla yağış olmasına rağmen daha az sulanabilir alan mevcuttur. Yağışların bölgelere göre değişmesi ve düzensiz olmasının yanında su miktarı ile nüfusun oransız dağılımı, su ihtiyacında değişkenlik olmasına neden olacaktır.

Su kaynaklarının planlaması ve yönetimi başlı başına zor bir konudur. Su yönetimi; çok boyutlu, birçok paydaşın bir arada bulunması gereken bir organizasyonu gerektirir. Ülkemizde, mevcut su kaynaklarının korunarak alternatif su kaynaklarının geliştirilmesi, değişen şartlara uygun ekolojik çevre ile uyumlu teknolojilerin kullanılması ve yaygınlaştırılması, sürdürülebilir su politikalarının oluşturulması gerekmektedir [1].

1.4. Su Kalitesi Parametreleri ve Önemi

Tarımsal ve endüstriyel faaliyetler neticesinde su kaynakları kirlenmektedir. Su ortamlarındaki kirlenme, suda yaşayan canlıların biyolojik çeşitliliği ve su kalitesi üzerinde etkili olmaktadır. Doğal su kaynaklarının korunması ve sürdürülebilirliğin sağlanması için su kalitesi parametrelerinin düzenli bir şekilde izlenmesi ve yürürlükteki ilgili kriter değerlere göre değerlendirilmesi büyük önem arz etmektedir. Ulusal ve uluslararası standart ve yönetmeliklerde, suyun kullanım amacı ve su kaynaklarına göre farklı kriter değerler belirlenmiştir. Suyun kullanım amaçlarına, içme suyu ve sulama suyu örnek verilebilir. Kaynaklarına göre sular da, yüzeysel sular ve yeraltı suyu olarak sınıflandırılabilir. Literatürde su kalitesi parametreleri için genel olarak fiziksel, kimyasal, biyolojik parametreler şeklinde bir sınıflandırma yapılmaktadır. Aşağıda bu parametreler ile ilgili detaylı açıklamalar yer almaktadır.

1.4.1. Fiziksel Parametreler

1.4.1.1. Sıcaklık (T)

İçilebilir suyun sıcaklığının 15 °C'nin altında olmaması ve en uygun sıcaklık aralığı değerinin 10-12 °C civarında olması istenilmektedir [6]. 20 °C'nin üzerindeki sular bulantı verici ve lezzetsiz olabilir. Yani bu parametre bir yerde damak lezzetine hitap eder. Su sıcaklığı parametresi, endüstriyel amaçlı kullanılan sularda çok dikkat edilmese de içme suyu amaçlı sularda oldukça önemlidir.

Su kaynakları, suyun özelliği ve iklimsel koşullar gibi nedenlerden dolayı farklı sıcaklıklarda olabilir. T parametresi arttıkça sudaki reaksiyonların hızı da beraberinde artacaktır. T parametresi, biyolojik aktiviteyi yani mikroorganizma gelişim hızını etkiler.

T parametresi, sucul yaşam için de çok önemli bir yere sahiptir. Sıcaklığın artması sudaki ÇO miktarının azalmasına neden olacaktır. Su içerisindeki katı maddelerin çökelme hızı artacak ve çok ciddi su kalitesi sorunlarının oluşumuna neden olacaktır.

T parametresi; suyun çözünürlüğü ve yoğunluğunu etkileyebilmesi ihtimali dolayısıyla, su yapılarının (su arıtma tesisleri vb.) inşası açısından da ayrıca önemli bir parametredir.

Su sıcaklığı termometreler ile ölçülür. Derin sular için özel olarak geliştirilmiş termometreler kullanılmaktadır.

1.4.1.2. Elektriksel İletkenlik (Eİ)

Bu parametre, suyun elektrik akımını iletebilme kabiliyetini ifade eder. İletkenlik, suda çözülmüş maddelere işarettir. İçme sularında Eİ artışı, suyun kirlendiği veya tuz miktarındaki artışın, deniz suyu karışımı olasılığını gösterir.

Ülkemizde yürürlükte olan "İnsani Tüketim Amaçlı Sular Hakkında Yönetmeliğe" göre izin verilen en üst iletkenlik değeri 2500 microS/cm'dir [7].

Doğal sular sahip oldukları tuz çözeltileri sayesinde elektriği iletirler. Tuz miktarının artması iletkenliğin artmasına sebep olacaktır. Elektriksel iletkenliği yüksek olan sular korozyon özellik gösterirler. Bu sular, aşındırıcı etkiye sahiptirler. Bundan dolayı, işletmeler (Su deposu, borular vb.) açısından önemli ve takibi gereken bir parametredir.

Eİ kondüktometre ile ölçülür. İletkenlik birimi de Siemens/cm'dir.

1.4.1.3. Askıda Katı Madde (AKM)

Suda, askıda maddelerin yüzer halde ya da çökme halinde bulunmasına 'askıda katı madde' denilir. Erozyon ile madde taşınması, organik madde parçacıklarının varlığı ve planktonlardan kaynaklanan AKM'ler vardır.

AKM'ler sularda bulanıklığa neden olabilirler ve bulanıklık ışık geçişini engeller. Ayrıca, AKM'ler sularda doğal ve geçici renk oluşumuna neden olabilirler.

1.4.1.4. Renk

İçilebilir nitelikte olması istenilen sular renksiz olmalıdır. Suyun rengi, içinde bulunması muhtemel kirleticilerin göstergesi, yani organik bileşik varlığına işaret olabilir. Suyun renginin oluşumunda birçok etken etkilidir. Bunları, organik ve inorganik maddeler olarak sınıflandırabiliriz. Organik maddeler, topraktaki organik maddeler ile bitkisel varlıklardır. Organik maddelerden kaynaklı suyun rengi, gerçek renktir. İnorganik maddeler ise suda çözülmüş halde bulunan mangan, krom, demir vb. maddelerdir. Daha çok evsel ve endüstriyel kirlenme ile ortaya çıkar.

Suyu tüketecek kişiler açısından suyun rengi önemli bir faktördür. Tüketici, içerisinde sağlığa zararlı bir madde bulunmasa bile renkli bir su yerine, renksiz ama sağlık yönünden güvenli olmayan bir suyu tercih edebilir. Suyun renginin önemi, estetik açıdan ön plana çıkıyor gibi görülse de sudaki ışık geçirgenliğini etkilemesi nedeniyle de önemli bir parametredir.

1.4.1.5. Bulanıklık

Sular için ışık geçirgenliğinin ölçüsü olması bakımından bulanıklık parametresi önemlidir. İçme ve günlük hayatta kullandığımız sularının berrak olması istenir. Sudaki bulanıklığa askıdaki katı maddeler neden olmaktadır. Organik ve inorganik maddeler nedeniyle kirlenme oluşacak, bu sebeple bakteri oluşumu hızlanacak ve bakteri oluşumu nedeniyle de suda bulanıklık artacaktır.

Suda bulanıklık parametresinin ölçümü bulanıklık ölçüm cihazı ile sağlanır. Bulanıklık ölçüm deneyi şu şekilde gerçekleştirilir. Su numunesi cihaz içerisine yerleştirilir ve cihazın içindeki ışık kaynağı ile aydınlatılır. Su içerisinde bulunan parçacıklar gelen ışınların bir kısmını kırar ve dağıtırlar. Işık kaynağı yönüne dik açıda olan kırılmış ve yansıtılmış ışınların yoğunluğu detektörler vasıtası ile ölçülür ve deney tamamlanır. Su analizlerinde bulanıklık birimi olarak ‘NTU’ (Nephelometric Turbidity Unit) kullanılmaktadır. 5 NTU veya daha yüksek miktarda olan bulanıklık göz ile fark edilir. Ayrıca, içme suyunda bulanıklık 1 NTU’dan fazla olmamalıdır [8].

Bulanıklık, estetik açıdan suda istenmeyen görüntü oluşturur. Su içerisindeki askıda katı maddeler çökeltme ve filtrasyon ile sudan uzaklaştırılarak bulanıklık giderilebilir.

1.4.1.6. Suyun Kokusu

İçme ve kullanma suları kokusuz olmalıdır. Koku estetik açıdan önemli bir parametredir. Suların kokularındaki değişimin fark edilmesi muhtemel kirleticiler için önlem alınmasını sağlayabilir. Sudaki kokunun giderilmesinde havalandırma ve daha farklı yöntemler tercih edilmektedir.

1.4.1.7. Suyun Tadı

Suya tadını, içerisinde erimiş halde bulunan karbondioksit ve oksijen gazları verir. Örneğin kaynamış suyun tadı bu gazların eksikliğinden dolayı iyi değildir. Suyun tadı, damak lezzeti yanında sucul ortamlarda anaerobik ortamın oluştuğuna da işaret edebilir. Çürümüş bitkiler, algler vb. suyun tadını olumsuz etkileyebilir.

1.4.1.8. Suyun pH Değeri

Latince ‘potentia hydrogenii’ kelimesi baş harflerinin kısaltması olan pH kavramı, çözeltideki hidrojen iyonu miktarını gösterir. Yani, suların asidik veya bazik durumunu gösteren logaritmik ölçünün adına ‘pH’ denilir. Ölçüm aralığı 0 ile 14 arasında değişir. pH değerinin; 7’nin altında olması asidik, 7’nin üstünde olması bazik ve 7 olması hali nötr durumdur.

Suların geçtiği jeolojik tabakalar suyun pH'ını etkiler. Organik maddelerin su içerisinde bozulması da suyun pH'ını değiştirebilir. pH değeri ‘pH metre’ cihazı ile ölçülür. TS 266 Standardında (İnsani Tüketim Amaçlı Sular) pH değer aralığının 6,5 ile 9,5 arasında olması istenilmektedir.

1.4.2. Kimyasal Parametreler

1.4.2.1. Çözünmüş Oksijen (ÇO)

ÇO, suda çözünmüş halde bulunan oksijen konsantrasyonudur. Oksijenin çözünürlüğüne birçok faktör etki edebilir. ÇO konsantrasyonunda basıncın belirleyici bir etkisi vardır. Yüksek basıncın oksijen çözünürlüğünü artıracak ve basıncın azaltılması durumunda azaltılma oranı kadar gaz çıkışı olacağından oksijenin çözünürlüğü azalacaktır.

Saf sular, yüksek mineral içerikli sulara göre daha çok oksijene sahiptirler. Suyun içerdiği mineralin miktarı, oksijeni çözme yeteneğini etkiler. Damıtık sular, yüksek mineral içerikli sulara göre daha çok oksijeni absorbe edebilir.

ÇO suya, atmosferden doğrudan ya da rüzgâr, dalga ve havalandırma sonucu girer. Sığ sularda yüzeydeki dalgaları oluşturan rüzgâr ÇO parametresini artıracaktır.

ÇO parametresi diğer su kalitesi parametrelerinden etkilenmekte ve etkilemektedir. T parametresinin ÇO miktarı üzerinde etkisi çok fazladır. Sıcaklığın artması ÇO miktarını azaltacaktır. Su içerisinde anaerobik çürüme olacak ve suda kötü koku sorunu olacaktır. Ayrıca Eİ parametresi de ÇO konsantrasyonunu etkilemektedir. Düşük Eİ değeri ÇO konsantrasyonu değerini artıracaktır.

ÇO miktarı sucul yaşam içinde çok önemlidir. Suda organik kirleticilerin bulunması, zamanla suda oksijensizlik durumunun oluşmasına neden olacaktır. Bu durum suda canlı yaşamını olumsuz etkileyecektir. Suda, ÇO'nun az olması yanında fazla olmasının da olumsuz etkileri vardır. Düzenli takibinin yapılması ile oluşacak olumsuzluklar engellenebilir.

ÇO'nun korozyon döngüsünde hızlandırıcı etkisi vardır. Bundan dolayı metal yapıların zarar görmesinin takibi açısından da ayrıca önem arz eder.

Su kalitesinin izlenmesinde, su kirliliğinin belirlenmesinde ve atık su arıtma faaliyetlerinde ÇO parametresi birçok yönden önemlidir. ÇO parametresinin belirlenmesinde iki ayrı yöntem kullanılmaktadır. Bunlar; winkler veya iyodometrik metot ile elektrometrik metottur. Winkler titrasyon yöntemi, çözünmüş oksijenin oksitleme özelliğine dayanan manuel bir titrasyon yöntemidir. Membran elektrot işlemi ise moleküler oksijenin membrana karşı difüzyon hızına dayanan bir yöntemdir. Kimyasal ve cam malzeme kullanılmadan multimetre üzerine takılan bir prop ile gerçekleştirilir. Membran elektrot işlemi hızlı sonuç elde edilmesi nedeniyle iyodometrik metoda göre daha çok tercih edilir.

1.4.2.2. Suyun Sertliği

Suyun sertliği, sahip olduğu çözünmüş kalsiyum (Ca^{+2}) ve magnezyum (Mg^{+2}) iyonlarından gelir. Suyun sertliği, magnezyum ve kalsiyum iyonlarının kalsiyum karbonat cinsinden toplam konsantrasyonları olarak da ifade edilebilir. Sertlik, geçici ve kalıcı sertlik olarak ifade edilebilir. Geçici sertlik (karbonat sertliği), kalsiyum ve magnezyum bikarbonatlarıdır ve suyun kaynatılması ile giderilebilir. Kalıcı sertlik (karbonat olmayan sertlik); sülfat, nitrat, klorür, fosfat ve silikatlarıdır ve arıtma ile giderilir. Sertlik, çeşitli sertlik dereceleri (fransız sertliği, ingiliz sertliği vd.) ile ifade edilir.

Sert sular; insan sađılıđı aısından deride tahriše neden olabilirler. Suyun sertlik derecesi, insan sađılıđına etkileri yanında, daha ok ekonomik aıdan ve lezzet bakımından onemlidir. Yumuřak sulara gre sert suların temizlik amalı kullanımda daha fazla sabun tketilmektedir. Sert sular, fazla sabun tketilmesine neden olması ve ekonomik olmamalarından dolayı endstriyel kullanımda genellikle tercih edilmezler.

Dokuma endstrisinde, sert suların kullanılması halinde boyalardan istenilen verim alınamayabilir. Ayrıca, sert sular ile yıkanan sebzelerde katılařması ve renk deđiřimi gzlenebilir.

1.4.2.3. Klorr (Cl⁻)

Klorr, dođada dođal veya kullanılmıř sulara yaygın olarak bulunan bir elementtir. NaCl (Sodyum klorr), yaygın olarak bilinen adı ile tuz maddesini suda zdđmzde klorr iyonu oluřur. Sulara, topraktan znme ile ya da tuzlu su (deniz suyu) karıřımı nedeniyle bulunabileceđi gibi atıksu karıřımı da etken olabilir.

Topraktan znme ile karıřan klorr sađlık aısından sorun teřkil etmez. Klorr'n insan sađılıđı aısından dřk ya da fazla deđil, dengeli olması istenilen durumdur.

1.4.2.4. Slfat (SO₄⁻²)

Slfat, dođal sulara deđiřen oranlarda bulunur. Bir tuz bileřiđi olan slfatlar, jips kkenlidir. Atık su karıřımı nedeniyle dođal sulara slfat miktarı artabilir. Sulara yksek miktarda slfat bulunması; yksek tuzluluk ve yksek sertliđi ifade eder. Damıtma iřlemi ile sudaki slfat giderilebilir.

1.4.2.5. Biyolojik Oksijen İhtiyacı (BOİ)

Sudaki oksijenin, sudaki mikroorganizmalar tarafından ne kadar hızlı kullanıldığını tespit eden biyolojik bir prosedrdr. Bir bařka deyiře, sudaki organik maddelerin, suda mevcut bulunan mikroorganizmalar tarafından paralanması iin gerekli oksijen miktarıdır. Suda BOİ deđerinin yksek olması suyun kirlilik dzeyinin yksek olduđunu gsterir.

BOİ deđeri, BOİ₅ testi ile tespit edilir. Bu test, 20  C sıcaklık ve beř gn standart inkbasyon sresinde suda organik maddelerin oksitlenmesi iin kullanılacak oksijen

miktardır. BOİ değeri atık su arıtma tesislerinde nasıl arıtma yönteminin kullanılacağı yönünden gereksinim duyulan bir değerdir.

1.4.2.6. Kimyasal Oksijen İhtiyacı (KOİ)

Evsel ve endüstriyel atıkların organik kirlilik derecesinin belirlenmesinde kullanılan bir parametredir. KOİ, su içerisinde yer alan yükseltgenebilir özellikteki maddelerin oksitlenmeleri için gerekli oksijen miktarıdır.

Organik atık maddelerle kirlenmiş olan su numunelerinin oksijen tükenme kapasitesinin belirlenmesi için ‘KOİ Tayini’ gerçekleştirilmektedir. KOİ temelde sudaki organik maddelerin asidik bir ortamda, yüksek sıcaklık altında kuvvetli oksitleyici kimyasallar ile oksitlendirilmesidir. Bu test ile birlikte, karbonlu organik maddeler CO_2 ve suya, azotlu organik maddeler de NH_3 (amonyağa) dönüşür.

KOİ, su kalitesinin belirlenmesinde kullanılan önemli bir parametredir. Genellikle BOİ’nin tahmin edilebilmesi için gerçekleştirilmektedir. KOİ analizinin kısa sürede sonuç vermesi BOİ analizi yerine tercih sebebi olarak görülebilir.

1.4.2.7. Azot (N) - Amonyak (NH_3) - Amonyum (NH_4)

Azot (N), sularda sık görülen bir parametredir. Sularda farklı oksidasyon kademelerinde bulunur ve bakteriler tarafından besi kaynağı olarak kullanılır. Azot türleri sularda; Amonyak Azot ($\text{NH}_3\text{-N}$), Organik Azot (Org-N), Nitrik Azot ($\text{NO}_2\text{-N}$) ve Nitrat Azot ($\text{NO}_3\text{-N}$) şeklinde bulunur. Alıcı ortamda bu azot türlerinin aşırı miktarda bulunması halinde organizmalar tarafından kullanılarak ötrofikasyona neden olacaktır. Sulardaki azotu gidermek için; havalandırma, klorlama, iyon değiştirme gibi arıtma yöntemleri kullanılabilir.

Amonyak (NH_3), hayvan yemi üretiminde ve gübre olarak kullanılır. Amonyak ve amonyum tuzlarına genellikle temizlik maddelerinde rastlanır. Amonyum fazlalığı sularda tat ve koku problemi oluşturur ve insan sağlığı üzerinde olumsuz etkisi vardır. Sularda mikroorganizma faaliyetleri sonucu oluşan amonyak sağlığa zararlıdır.

1.4.2.8. Nitrit (NO₂) - Nitrat (NO₃)

Nitrit, suda mikrobiyolojik kirlenmeye işaret eder. Sularda bulunması kan zehirlenmesine sebep olur.

Nitrat, sularda kolay çözüldüğü için tehlikelidir. Sulardaki nitratın kaynağı tarımdaki gübreler ve atıklardır. İnsan sağlığı açısından özellikle bebeklerde mavi hastalığa neden olur. WHO tarafından içme sularında 45 mg/L üstünde olması istenmemektedir.

1.4.2.9. Florür (F)

Florür, sudaki miktarına göre faydalı veya zararlı olabilir. Yüksek miktarda florür içeren sular diş ve kemik sağlığı açısından zararlıdır. TS 266 Standardına (İnsani Tüketim Amaçlı Sular) göre içme sularında florür miktarı 1,5 mg/L'den fazla olmamalıdır. İçme suyu amaçlı tüketilecek sular için genellikle 1 mg/L'den fazla olmaması istenilir. İçme sularında düşük konsantrasyon halinde sulara florid katılmalı, fazla olması halinde arıtma işlemi yapılmalıdır.

1.4.2.10. Sodyum (Na) - Sodyum Adsorpsiyon Oranı (SAR)

Sodyum elementinin içme sularında 200 mg/L'nin üzerinde olması istenmemektedir. Fazla olması halinde tuzluluk hissinden dolayı hoş olamayan tada neden olur.

Tarımsal sulamada sodyum fazlalığı önemli sorunlardandır. Sulama suyunda sodyum fazlalığı, toprak yapısının bozulmasına neden olacak ve suyun toprağa nüfuz etmesini engelleyecektir. Sulama suyunda sodyum kriteri, "sodyum adsorpsiyon oranı (SAR)" olarak ifade edilir. Yüksek SAR değerine sahip suyun sürekli kullanımı toprağın fiziksel yapısında bozulmaya sebep olacaktır. Karbonat ve bikarbonat konsantrasyonu yüksek olan sular, kalsiyum karbonata ve magnezyum karbonata çökelme eğiliminde olduğundan SAR değerini yükseltecektir.

1.4.3. Mikrobiyolojik Parametreler

Sanayi bölgelerine yakın bulunan su kaynakları, yerleşim alanlarının yakınında bulunan nehir, baraj, göl vb. sular ile su arıtma ve kanalizasyon sistemi sorunları olan bölgelerdeki sular, insan sağlığı için zararlı olabilen mikroskobik canlıların üreme ve çoğalmaları için uygun ortamı sağlar. Suyun kalitesinin bozulmasına yol açabilecek kanalizasyon

atıklarının varlığının tespiti ve hastalık yapıcı canlıların (patojenlerin) tespitinde koliform bakterilerinin varlığı aranır. Patojenlerin sayısının az olması zararsız olacaklarını düşündürse de uygun ortamı bulmaları halinde çok hızlı çoğalabilirler.

Suların mikrobiyolojik kalitelerinin korunması için yapılması gereken öncelikli adım, suyun kirlenmekten korunması olmalıdır. Su kirliliğinin erken belirlenmesi için izleme çalışmaları düzenli olarak yapılmalıdır.

Sularda bulunabilecek patojenlerin varlığı ve miktarı mikrobiyolojik analizler ile tespit edilir.

1.4.3.1. Koliform Bakteriler

Kanalizasyon suları ya da dışkı karışması neticesinde kirlenmiş sular, bulaşıcı bağırsak hastalıklarının taşınmasına neden olur. Hastalıklı kişilerin dışkılarında bulunan patojen mikroorganizmalar arasında virüsler, parazitler, bakteriler ve protozoalar yer alır [6]. Su kaynaklarında dışkı ile kirlenmenin olup olmadığı patojen olmayan koliform bakteriler ile tespit edilir.

1.4.3.2. Toplam Koliform

Su içerisine kanalizasyon karışabileceğinin bir göstergesidir. Koliform grubu içindeki patojenlerin göstergesidir.

İçme suyunun kalitesinin belirlenmesinde, toplam koliform bakteri tayini önemlidir. İçme suyunda toplam koliform bakteri bulunmaz ise suyun kullanılabilir olduğu söylenebilir. Toplam koliform bakteri tespit edilmesi halinde ise suyun patojen varlığı nedeniyle kullanılmasının risk taşıdığı kabul edilir.

1.4.3.3. Fekal Streptokoklar

Fekal streptokoklar, çoğunlukla insan ve hayvanların dışkısında bulunurlar. Kirli sularda genellikle çoğalamazlar fakat koliform bakteriler ve koli basili (e.koli)'den daha dayanıklıdır. Su kalitesinin kontrolü çalışmalarında arıtma işlemlerinin verimliliğinin ölçülmesinde önem arz ederler.

1.4.3.4. Escherichia Coli (Koli Basili (E.koli))

E.koli bakterisi, memeli canlıların kalın bağırsağında yaşar. E.kolinin çevresel sularda varlığı zararlı mikroorganizmaların bulunduğu belirtilir. Arıtma işlemleri ile sulardan uzaklaştırılarak hastalık oluşması engellenir.

1.4.4. Ağır Metaller

1.4.4.1. Arsenik (As)

Doğada sülfür ve oksit halinde bulunan arsenik çok etkili bir zehirdir. Doğal suda nadir olarak bulunan arsenik, akiferlerin jeolojik ve jeokimyasal özelliklerine bağlı olarak minerallerin çözünmesi yanında endüstriyel faaliyetler ve pestisitlerden kirlilik olarak da suya karışabilmektedir [2]. Su içerisinde arsenik olması halinde; doku bozulmaları, dolaşım sistemi rahatsızlıkları, zehirlenmeler ve kanser riski gibi problemler ortaya çıkabilir.

Arsenik, içme suları ve su ürünleri açısından önemli olup, arıtımı için birçok yöntem kullanılabilir. Kireçle yumuşatma yöntemi, oldukça ekonomik ve basit bir yöntem olarak önemli oranda arsenik arıtımı sağlar. WHO içme ve kullanma amacıyla tüketilecek sularda 10 µg/l (ppb)'ye kadar arsenik bulunabileceği, 50 µg/l (ppb)'den daha yüksek miktarlarda arsenik bulunması halinde bu suların hiç kullanılmaması gerektiğini ifade etmiştir.

1.4.4.2. Kadmiyum (Cd)

Kadmiyum, birçok endüstriyel faaliyetlerde kullanılmakta ve bu faaliyetler neticesinde yer üstü ve yeraltı sularına karışabilmektedir.

WHO, içme sularında kadmiyum seviyesinin tolere edilebilir değerinin 0,005 µg/L olabileceğini belirlemiştir. İçme suyu için çok düşük konsantrasyondaki kadmiyumun herhangi bir yan etkisi görülmemekle birlikte insanlarda toksisitesi kanser riski problemlerindendir.

1.4.4.3. Siyanür (CN⁻)

Siyanür, çok etkili bir zehirdir. Birçok farklı endüstri kolunda kullanılmaktadır. Bunların en önemlileri kimyasal ve plastik üretim endüstrileridir. Ayrıca madencilikte, altın ve gümüş üretiminde kullanılmaktadır.

Sudaki siyanür canlı yaşam için çok tehlikelidir. Çeşitli endüstriyel faaliyetler neticesinde sulara siyanür karışabilmektedir. Az ya da yüksek miktarda siyanüre maruz kalma durumunda ölüm bile gerçekleşebilmektedir. WHO tarafından, sağlık açısından siyanüre kısa süreli (5 günlük süre) maruziyette limit değer olarak 0,5 mg/L hesaplanmıştır.

1.4.4.4. Civa (Hg)

Civa, gümüş renkli olup ağır metal sınıfında yer alan bir elementtir. Oda sıcaklığında (25 °C) sıvı halde bulunur. Sularda genellikle endüstriyel faaliyetler neticesinde rastlanır. Ayrıca, besin zincirinde biriken civa suda yaşayan canlılara daha sonra bunlarla beslenen insanlara geçebilir. İnsan vücudunda depolandığı yerlere göre yarılanma ömürleri değişmekte olup beyinde depolanan civanın yarılanma ömrü uzun olabilir. Civa kirlenmesi yoksa dahi, taze sularda eser miktarda civaya rastlanır. Arıtım işlemleri ile birlikte su içerisindeki miktarı daha da azalacaktır.

Civa bileşiklerinin insanlarda toksik etkisi daha çok böbreklerde olmakla birlikte dış sağlığı açısından da olumsuz etkileri vardır. Civa bileşiklerinin biyolojik yarılanma ömürleri çok uzun olduğu için vücuttan atılması zor ve uzun zaman alır.

1.4.4.5. Kurşun (Pb)

Kurşun, toprağın doğal elementlerindendir. Doğada yaygın olarak bulunması ve kolay işlenebilmesinden dolayı birçok endüstriyel faaliyet alanında kullanılmakta ve bu faaliyetler neticesinde su kaynaklarında kirlenmeler meydana gelmektedir. İçme suyundaki kirlilik arıtma ile giderilebilir. Kurşun metalinin insan sağlığına etkisi (zehirlenme) daha çok sinir sistemine olur. Beyin ve böbreklerde de hasara neden olabilir.

1.4.4.6. Selenyum (Se)

Selenyum, sülfür grubunun kimyasal elementidir. Cam, kimyasal, lastik gibi birçok endüstriyel faaliyet alanında kullanılmaktadır. Selenyum bileşiklerine, doğada az rastlanır ve doğal sularda nadiren bulunurlar.

Selenyum, insan vücudunun gereksinim duyduğu minerallerdendir. Eksikliği halinde, çok odaklı kalp kası iltihabı hastalığının oluşumuna neden olabilir [9]. Yüksek miktarda alınması halinde cilt, saç ve dişte çeşitli rahatsızlıklara neden olabilir. Ayrıca, vücutta birikerek zehir etkisi gösterebilir.

1.4.4.7. Bakır (Cu)

Bakır elementi, doğada birçok yerde bulunması ve iyi bir iletken olması nedeniyle endüstride çok sık kullanılmaktadır. Bakır, doğal su kaynaklarında ve endüstriyel kirlenme sonucu sularda bulunabilir.

Bakır, besin olarak ihtiyaç duyulan bir elementtir. Eksikliği halinde sinir sistemi bozukluğu ve kansızlığa neden olabileceği gibi, yüksek miktarda alımı sindirim sistemi rahatsızlıkları ile böbrek rahatsızlıklarına neden olabilir. WHO tarafından içme suları için açıklanan sınır değeri 2 mg/L'dir [10].

1.4.4.8. Çinko (Zn)

Çinko, dünyada yaygın olarak bulunan elementlerdendir. Çelik gibi metallerin korozyondan korunması amacıyla galvanize edilmesinde, pillerin gövdelerinde, otomotiv endüstrisi gibi birçok sanayi dalında geniş kullanım alanı vardır. Su iletim hatlarında galvanizli borulardan suya geçebilir.

Birçok besinde çinko minerali vardır ve bağışıklık sistemi için önemlidir. Ulusal ve uluslararası kuruluşlar, çinko miktarının sınır değeri için limit değerler belirlememişlerdir. WHO, 3 mg/L'yi geçen çinko değerinin suda bulanıklık ve tatla değişikliğe neden olabileceğini belirtmiştir.

1.4.4.9. Alüminyum (Al)

Alüminyum metali, yerkürede bol miktarda bulunur. Suda bulunmasında iki faktörün etkisi vardır. Bunlar; suyun kaynağında ya da su arıtımındaki alüminyum tuzlarıdır. Fazlalığı halinde suyun rengi bulanık mavimsiye görünümüne alır. Başta sinir sistemi olmak üzere, böbreklerde tahribat yapar [6].

1.4.4.10. Demir (Fe)

Demir elementi yeryüzünde yaygın olarak bulunmakla birlikte doğada daha çok oksitlenmiş halde bulunur. Suda demirin iki şekline rastlanır. Suda çözünebilen demir içeren ferro demir (Fe^{2+}), oksijenle reaksiyonu sonucu suda çözünemeyen demir içerikli ferri demire (Fe^{3+}) dönüşür. Suda çözünemez haldeki demir çökelir ve suyun rengini de değiştirir.

Demir metali sulara kayaların ayrışması sonucu veya kirlenmeyle karışabilir. Suyun rengi ve tadında değişime neden olur. WHO, tatlı su kaynaklarındaki demir konsantrasyonunun 0,5-50 mg/L aralığında bulunabileceği belirtmiştir. Demir insan sağlığı açısından önemli ve gerekli bir element olduğu için dengeli bir şekilde tüketilmelidir.

1.4.4.11. Bor (B)

Su kaynaklarında bor elementine, borat ve borosilikat içeren kayalar ile toprakların sulara karışması sonucunda ve cam, sabun, deterjan, kozmetik ve ilaç endüstrisi faaliyetleri neticesinde rastlanır. Ülkemizde bor madenlerinin yoğun olduğu yerlerde yeraltı suyunda görülebileceğinden takip edilmesi önemlidir.

WHO, içme sularındaki bor konsantrasyonunun 0,5 mg/L'nin altında olmasını istemektedir. Fazla olması halinde insan sağlığına etkisi, sinir ve sindirim sistemine olabilir.

1.5. Su Kalitesi Parametreleri ve Yapay Zekâ Teknikleri

Dijitalleşme modern dünyamızın vazgeçilmezi olmuştur ve birçok bilim dalı da bu dijitalleşme sürecinden etkilenmektedir. Bu dijitalleşme süreci önemli kazanımlar sağlamaktadır. Mühendislik ve doğa olaylarında kullanım alanı son yıllarda önemli

şekilde artan yapay zekâ teknikleri, su yönetim sürecinde de önemli kolaylıklar sağlamaktadır.

Yapay zekâ tekniklerinin yağış, akım, taşkın, kuraklık gibi birçok su yönetimi çalışmalarının yanında su kalitesi çalışmaları alanında da kullanım alanı vardır. Su kalitesine ait geçmiş yıllardan elde edilmiş veriler ile model oluşturarak tahminde bulunmak, muhtemel sorunların çözümünde önemli kolaylıklar sağlayacaktır.

Son yıllarda veri analitiği ve akıllı sensörlerde yaşanan gelişmeler neticesinde yapay zekâ tabanlı çözümlerler hız kazanmıştır. Özellikle gelişmiş ülkelerde su yönetimi için yapay zekâ alanında önemli yatırımlar yapılmaktadır. Bunun en önemli nedeni olarak, yapay zekâyâ yönelik optimizasyon araçlarının standartların karşılanmasında sağladığı kolaylık ile verimlilik artışı ve işletme maliyetlerini düşürüyor olmasıdır [11].

Yapay zekâ teknikleriyle;

- ✓ Suyu olan talep ile tüketimin tahmini
- ✓ Su kalitesinin izlenmesi ve değerlendirilmesi
- ✓ Su altyapılarının fiziki durumunun tahmini
- ✓ Su kaynaklı doğal afetlerin oluşmasının engellenmesi için gerekli tedbirlerin alınması sağlanabilir.

Ayrıca, yapay zekâ tekniklerinin kullanımı ile ileride karşılaşılması muhtemel doğal afetlerin (taşkın, kuraklık vb.) zararı asgari düzeye indirilebilir.

Su kaynakları üzerinde kirlenmeye bağlı oluşan baskı, etkili bir su kalitesi yönetimini zorunlu kılmaktadır. Bu yönetim sürecinde, geleneksel istatistiksel yöntemlerin yanında yapay zekâ tekniklerinden de faydalanılabilmektedir.

Yapay zekâ yöntemleri, su kaynaklarına dair çalışmalarda geleneksel istatistiksel yöntemlerin karşılaştığı bazı zorlukları aşarak alternatif yöntemler olarak kullanılmaktadır. Bu da yapay zekâ tekniklerini geleneksel istatistiksel yöntemlere önce tamamlayıcı hale getirmiş, daha sonra da alternatif bir yöntem haline getirmiştir. Yapay zekâ tekniklerinin bağımlı ve bağımsız değişkenler arasında isteğe bağlı lineer olmayan ilişkiler kurabilmesi, bağımlı değişkenler arasındaki bütün muhtemel etkileşimlere izin vermesi, tek değişkenli yapılardan çok değişkenli yapılara doğru kolaylıkla

geniřletilebilmesi alternatif bir yöntem olmasına imkân sağlamaktadır. İstatistiksel yöntemler ve yapay zekâ teknikleri arasında küçük bir etkileřim bulunması yanında, her iki yöntemde hemen hemen bağımsızdırlar [12].

Tek deęiřkenli istatistiksel analiz çalışmalarında, deęiřkenler tek tek incelenir. Deęiřkenler sayısal veya kategorik olabilir. Çok deęiřkenli istatistiksel analiz çalışmalarında ise iki ya da daha fazla deęiřken aynı anda analiz edilir ve birbiriyle iliřki halinde çok sayıda deęiřken söz konusudur. Gerçek hayatın temel alındığı bilimsel çalışmalarda deęiřken sayısı birden fazla olup bu deęiřkenlerin birbirleriyle etkileřimleri de söz konusudur. Çok deęiřkenli istatistiksel analiz çalışmaları bu ihtiyaçlar doğrultusunda ortaya çıkmıřtır. Tek deęiřkenli istatistiksel analiz çalışmalarında veri olarak kabul edilen birçok etken, çok deęiřkenli istatistiksel analiz çalışmalarında deęiřken olarak yer alır. Çok deęiřkenli istatistiksel analiz yöntemleri ile çalışma yapılan konuya iliřkin bir deęiřken incelenirken bu deęiřken ile iliřkili dięer deęiřkenlerde sistem içerisine dahil edilerek incelenmelidir. Özellikle, mühendislik, eğitim, saęlık alanlarında kümeleme ve sınıflandırma yapma, veri indirgeme gibi çalışma alanlarında kullanılmaktadır. Çok deęiřkenli istatistiksel analiz yöntemleri uygulamada, MATLAB (Matrix Laboratory), Minitab, Maple, Statistica, SAS, SPSS gibi paket programlardan faydalanılmaktadır.

Yapay zekâ tekniklerinin su kalitesi izleme çalışmalarında potansiyel bir kullanım alanı vardır. Önümüzdeki yıllarda izleme çalışmalarının daha düşük maliyetli ve daha kolay erişebilir olarak yapılması durumunda yapay zekâ uygulamalarının kullanımı daha da artacaktır. Su politikalarına iliřkin konular genellikle yerelden küresel düzeye uzanmaktadır. Bu da politika geliřtirmede ve uygulamada işbirliğini gerektirmektedir. Yapay zekâ tekniklerinin kullanımında birlikteliğin saęlanması çalışmalardan saęlanacak verimi artıracaktır. Yapay zekâ tekniklerinden; özellikle sınıflandırma yapma, verileri eşleřtirme ve yorumlama gibi birçok alanda faydalanılabilir. Tahmin problemlerinde saęladığı doğru sonuçlar yapay zekâ tekniklerinin öne çıkan özelliklerindendir.

İnsan beyin yapısının çalışmasından etkilenen YSA, yapay zekâ yöntemleri içerisinde sıklıkla kullanılmaktadır. İnsan beyninin tecrübeler ile saęladığı kazanımlar, YSA'da eğitime aşamasıyla saęlanmaya çalışılır. Eğitime aşaması daha çok örnekler yardımıyla

öğrenmenin gerçekleştirilmesi olup, girdi ve çıktı elemanları sisteme tanıtılarak öğrenme sağlanır ve sonrasında eğitime işlemi gerçekleştirilir [13].

Veri madenciliğinin önem kazanması ile birlikte makine öğrenimi de önem kazanmıştır. Makine öğrenmesinin bir alt dalı olan DÖ, birden fazla katmana sahip mevcut veriler ile sonuçları tahmin etmeye yarayan bir öğrenme yöntemidir.

Makine öğrenmesi ve YSA'dan yararlanarak DVM ağ yöntemi oluşturulur. DVM, makine öğrenmesinin önemli yöntemlerinden biri olan kernel yöntemlerinin örneklerindendir. DVM yönteminden özellikle doğrusal olmayan sınıflandırma ve tahmin çalışmalarında yararlanılır.

Regresyon analizi, bağımlı ve bağımsız değişkenlerin arasındaki sayısal ilişkinin incelendiği bir yöntemdir. GSR modeli, karmaşık fonksiyon haritalamasının parametrik olmayan denetimli bir öğrenme yaklaşımı olup çekirdek tabanlı olasılık modelidir.

Bu tez kapsamında, su kalitesi parametrelerinin tahmin edilmesinde farklı yapay zekâ yöntemlerinin performansları karşılaştırılmıştır. Literatürde de su kalitesinde yapay zekâ tekniklerinin kullanıldığı birçok çalışma da göstermektedir ki, su kalitesinde yapay zekâ tekniklerinin önemli bir kullanım alanı vardır.

Modern dünyamızda yaşanan teknolojik yenilikler hayatımızın her alanında görülmekte ve bu yeniliklerden faydalanılmaktadır. İnsanlar gibi öğrenen, karar verebilen ve tahmin yapabilen makineler yapay zekâ kavramının temelini oluşturmuştur. Bilgisayar teknolojisinde son zamanlarda yaşanan gelişmeler, yapay zekâ yöntemlerine olan ilginin artmasını sağlamıştır. Birçok bilim dalında kullanım alanı olan yapay zekâ yöntemleri, verimliliğin ve otomasyonun artmasını sağlayarak önemli kazanımlar sağlamaktadır.

Ülkemizin önemli nüfus yoğunluğunu barındıran ve başkentimiz Ankara ile birlikte birçok sanayi şehrinin yer aldığı Sakarya Havzasında, su kalitesi verilerinin yapay zekâ teknikleri ile modellenmesi önemli kazanımlar sağlayacaktır. Su yönetim sürecinin birçok aşamasında yapay zekâ tekniklerinden yararlanılmaktadır. Literatürde; akım, debi, taşkın, kuraklık vb. birçok alanda olduğu gibi su kalitesine dönükte yapay zekâ tekniklerinin kullanıldığı birçok çalışma vardır. Literatürde su kalitesi parametrelerinin belirlenmesine yönelik incelenen bazı çalışmalar şu şekildedir.

Sengorur et al. (2005), Nehirlerin su kalitesinin belirlenmesinde, ÇO'nun önemli bir yere sahip olduğunu ortaya koydukları çalışmalarında, nehir su kalitesi kontrolü için ÇO miktarının tahmini için YSA'lerden yararlanmışlardır. Ölçümü yapılmış ve modelleme giriş verileri olarak; $\text{NO}_2\text{-N}$ (Nitrit Azotu), $\text{NO}_3\text{-N}$ (Amonyak Azotu), BOİ , debi ve T seçilmiştir. YSA modeli ileri beslemeli olarak oluşturulmuş ve modellerin performansı, Ortalama Karesel Hata (OKH) ve R^2 hata değerleri ile ölçülmüştür. ÇO değerinin, tahmin değeri ile gerçek değerleri kıyaslanmıştır. Çalışma neticesinde 1 gizli katman ile 3 nöron sayısına sahip ileri beslemeli YSA'nın, 0.9186 R^2 değeri ile en iyi tahmini sağlandığı yapılan çalışmayla ortaya konulmuştur. YSA'lerden, nehir suyu kalitesi parametrelerinin tahminde yararlanılabileceği sonucuna varılmıştır [14].

Terzi ve Baykal (2012), Su kaynakları mühendisliğinin birçok alanında kullanım alanı olan YSA ile katı madde tahmini çalışması yapmışlardır. Çalışma alanı olarak ülkemizin en büyük nehri olan Kızılırmak Nehri seçilmiş ve Yamula gözlem istasyonuna ait 1973-2003 yılları arasında ölçülmüş akım değerlerinden yararlanılmıştır. Akım değerlerinin giriş verisini, katı madde miktarı değerlerinin çıkış verisini oluşturacağı şekilde YSA modelleri geliştirilmiştir. Modelleme çalışmalarında akım ile katı madde miktarı değerleri boyutsuzlaştırılmıştır. Verilerin; 1973-1997 yılları arasına ait olanların % 80'ni eğitmede, kalan % 20'si ile 1998-2003 yılları arasındaki veriler test verisi olarak kullanılmıştır. Modellerin performanslarını kıyaslanırken KOKH ve R^2 değerleri kullanılmıştır. Geliştirilen modeller karşılaştırıldığında en iyi modelin, 1 girdi, 4 gizli ve 1 çıktı katmana sahip ve aktivasyon fonksiyonu logaritmik sigmoid olan YSA modeli olduğu görülmüştür. Yapılan çalışma neticesinde; son yıllarda geniş kullanım alanına sahip YSA'nın katı madde miktarının tahmininde kullanılabileceği, ayrıca ölçüm yapılamayan durumlar ya da kayıp veriler durumu ile eksik verilerin de YSA ile tamamlanabileceği ifade edilmiştir [15].

Solanki et al. (2015), göller ve rezervuarların canlı yaşamı için önemini belirttikleri çalışmalarında su kaynaklarında yaşanması muhtemel kirlenme tahmininin su kaynaklarının korunması konusunda çok önemli olduğunu belirtmişlerdir. Çalışma alanı olarak, Hindistan, Maharashtra, Nasik yakınlarındaki Chaskaman Nehri seçilmiştir. DÖ modeliyle, tahmine dayalı analiz teknikleri kullanılmıştır. Su kalitesi parametrelerinden; ÇO, pH ve bulanıklık konsantrasyon değerlerinin farklı modeller sonucu elde edilen gerçek ve tahmini değerleri karşılaştırılmıştır. Model performansları; OMH ve OKH

değerlerine göre değerlendirilmiştir. Denetimsiz DÖ ağ modelinin denetimli DÖ ağ modeline kıyasla daha doğru sonuçlar elde ettiği belirtilmiştir. ÇO değerinde yaz aylarında değişim gözlemlendiği bu durumun sıcaklık kaynaklı olabileceği değerlendirilmiştir. Ayrıca, pH değerinin ÇO ve bulanıklığa göre daha kararlı olduğu, bulanıklık değerinin pH ve ÇO'ya göre daha değişken olduğu bu durumun yağış kaynaklı olabileceği ifade edilmiştir [16].

Kisi ve Parmar (2015), Nehir su kirliliğinin tahminine yönelik modelleme çalışmalarında; DVM, çok değişkenli adaptif regresyon eğrileri (MARS) ve M5 model tree (M5Tree)'nin tahmin başarısı karşılaştırılmıştır. Çalışma alanı, Hindistan Delhi'deki Yamuna Nehridir. Modellerde giriş verileri; serbest amonyak, toplam kjeldahl azotu, T, toplam koliform, fekal koliform ve pH olup, çıkış verileri KOİ'dir. Veriler üç alt gruba bölünerek kullanılmıştır. Model performansları, KOKH, OMH ve R^2 ile zirve yüzde eşik istatistikleri değerlerine göre kıyaslanmıştır. DVM ve MARS modellerinin tahmin başarısı birbirine yakın ve M5Tree'ye göre daha yüksektir. DVM ve M5Tree modellerinin KOKH değerleri, MARS modeli kullanılarak sırasıyla %1,47 ve %19,1 azaltılmıştır. KOİ değeri tahmininde; pH ve fekal koliform değerlerinin eklenmesi tahmin başarısını düşürmüş olup, T değerinin eklenmesi de tahmin başarısının artmasını sağlamamıştır. Serbest amonyak, toplam kjeldahl azotu ve T parametreleri kullanılarak nehir su kirliliğinin, DVM ve MARS modelleri ile tahmin edilebileceği ortaya konulmuştur [17].

Mohammadpour et al. (2015), Malezya Bilim Üniversitesi'ndeki serbest yüzeyli sulak alanda yaptıkları çalışmalarında; DVM ve iki YSA (ileri geri yayımlı ve radyal esaslı) ile su kalitesi endeksi tahmin edilmeye çalışılmıştır. Çalışma alanında; 17 noktada izleme çalışması yapılmış, 14 aylık (Ekim 2010-Aralık 2011) zaman aralığında ayda iki kez veriler alınmıştır. 11 adet su kalitesi değişkenine ait konsantrasyon değerleri izlenmiştir. Bunlar; T, pH, ÇO, Eİ, AKM, nitrit, nitrat, amonyak nitrojen, BOİ, KOİ ve fosfattır. Toplanan veriler normalize edilmiştir. Üç ayrı oluşturulan modellerin performansları; R^2 , OMH, KOKH'ya göre kıyaslanmıştır. İleri Geri Yayımlı YSA, R^2 değeri 0.9988 ve OMH değeri 0.0044 ile, diğer modellere ait performans değerlerinden daha iyidir. DVM modeline ait değerler İleri Geri Yayımlı YSA'ya çok yakın olup, R^2 değeri 0.9984 ve OMH değeri 0.0052'dir. Radyal Esaslı YSA, R^2 değeri 0.9960 ve OMH değeri 0.0080 ile tahmin performansı, diğer iki model performansından çok az düşüktür. Sulak alanlarda,

su kalitesi endeksinin öneminden dolayı İleri Geri Yayılımlı YSA ve DVM modellerinin kullanılması ile güvenilir sonuçlar elde edilebileceği ifade edilmiştir [18].

Zhang et al. (2016) Çin'deki üç rüzgâr çiftliğinden toplanan rüzgâr hızı verilerini kullanarak GSR tabanlı hibrit yaklaşım ile YSA ve DVM modellerini karşılaştırmıştır. Rüzgâr enerjisinin güç sistemine entegrasyonunun artmasıyla birlikte, kısa vadeli rüzgâr hızı tahmini için doğru ve güvenilir yöntemlere acilen ihtiyaç duyulduğu belirtilen çalışmada kısa vadeli rüzgâr hızı tahmini için otoregresif (OR) model ile GSR'ye dayalı hibrit bir model önermişlerdir. Önerilen modelle, rüzgâr hızı serilerinden genel yapıyı yakalamak için otoregresif modeli kullanılmış ve yerel yapıyı çıkarmak için GSR benimsenmiştir. Tahmin sonuçları, önerilen yöntemin diğer yöntemlerle karşılaştırıldığında yalnızca nokta tahminlerini iyileştirmekle kalmayıp, aynı zamanda tatmin edici tahmin aralıkları da ürettiğini göstermektedir. Model sonuçları kıyaslandığında KOKH, OMH ve Ortalama Mutlak Yüzde Hata (OMYH) performans sonuçları açısından hibrit modelin en iyi tahmin performansını verdiğini ifade etmişlerdir [19].

Ozel et al. (2017), Çevre kirliliğinin önemli ölçüde nehir ekosistemlerini etkilediği ve sürdürülebilir bir su yönetimi için nehirlerin kirliliğinin izlenmesi gerektiğini belirttikleri çalışmalarında, Bartın Nehri'nde (2012 Aralık - 2013 Aralık) tarihleri arasında beş farklı gözlem istasyonunda aylık olarak ölçülen BOİ, KOİ, AKM, pH, Eİ ve T değerleri ile Çoklu Doğrusal Regresyon (ÇDR), Radyal Temelli Sinir Ağı (RTSA) ve Çok Katmanlı Algılayıcı Sinir Ağları (ÇKASA) modelleri ile su kalitesi tahmini için çalışma yapılmıştır. Bu modellerde, giriş verisi olarak; T, pH, KOİ, AKM, Eİ parametreleri kullanılarak BOİ değeri tahmin edilmiştir. Gözlem istasyonlarına ait 41 ölçüm verisi eğitmede, diğer 18 ölçüm verisi test sürecinde kullanılmıştır. Elde edilen sonuçlara göre YSA modelleri çoklu doğrusal regresyon modelinden daha iyi sonuçlar sağlamıştır. Kurulan modellere kıyasla en iyi performans değerlerine radyal tabanlı bir YSA modeli ile ulaşılmıştır. Bu modelde OMH, KOKH ve R^2 değerleri sırasıyla 0.998, 1.230 ve 0.890 olarak elde edilmiştir. Mevcut çalışma sonuçlarına göre, Bartın Nehri'ndeki aylık BOİ değerleri için YSA modelleri ile en başarılı tahmin elde edilmiştir [20].

Talesh et al. (2019), DVM yardımıyla su kalitesi endeksinin değişimini etkileyen faktörleri tanımlamak ve su kalitesi endeksi tahmin edilmeye çalışılmıştır. İran'ın kuzey kesiminde yer alan Sefidrud havzasında, Aralık 2007'den Kasım 2008'e kadar aylık olarak su numuneleri alınmış olup; bu numunelere ait ÇO, pH, toplam katı miktarı, T, nitrat, fosfat, BOİ, bulanıklık ve koliform değerleri model değişken verileri olarak kullanılmıştır. Verilerin; %80'ni eğitme verisi olarak, %20'si test verisi olarak kullanılmıştır. MATLAB yazılım programında DWM modeli, dokuz adet su kalitesi parametresi giriş verileri, su kalite endeksi değeri çıkış verisi olarak kullanılmıştır. Dokuz ay boyunca havzada belirlenmiş beş istasyona ait su kalite endeksi değerleri tahmin edilen değerler ile kıyaslanmıştır. R^2 ve KOKH 0.87 ile 0.061 olarak bulunmuştur. Ayrıca nitrat parametresinin su kalitesi endeksinin etkileyen önemli bir değer olduğu belirtilmiştir. Yapılan çalışma neticesinde DVM modelinin su kalitesi endeksi tahmininde başarı sağlayabildiği ve su kalitesi yönetim sürecine önemli katkılar sağlayabileceği ifade edilmiştir [21].

Liu et al. (2019), Yangzhou City'deki Guazhou su kaynağının otomatik su kalitesi izleme istasyonu tarafından ölçülen içme suyu kalitesi verilerini tahmin etmek için DÖ modelinden yararlanmışlardır. Model giriş verileri olarak, 2016–2018 tarihleri arasındaki T, pH, ÇO, Eİ, bulanıklık, KOİ konsantrasyonu değerlerini kullanmışlardır. Çalışmalarının sonucunda, DÖ modeliyle tahmin edilen değer ile ölçülmüş değerlerin birbirine yakın olduğu değerlendirilmiştir. DÖ modelinin su kalitesi parametrelerinin tahmin edilmesinde kullanılabileceği ifade edilmiştir [22].

Ma et al. (2019), New York City gibi kıyı kentleri için liman suyu kalitesinin kontrolü ve yönetiminin öneminden bahsedilen çalışmalarında, DÖ yöntemi ile liman suyundaki BOİ konsantrasyon değerinin tahmini amaçlanmıştır. BOİ konsantrasyon değeri, tespit edilmesi çok zaman alan ve su kalitesi açısından önemli bir değerdir. BOİ değerinin tespitinde işletme ve insan kaynaklı hatalar nedeniyle ortaya çıkan seyrek matris sorununun çözümü için Derin Matris Faktörizasyonu (DMF) ve Derin Sinir Ağı'nı (DSA) birleştiren DÖ tabanlı bir model çalışması yapılmıştır. Sonuçlar incelendiğinde kurulan modelin klasik matris tamamlama yöntemleri ve geleneksel makine öğrenimi algoritmalarına göre daha doğru tahmin yapabildiği görülmektedir. Model performans kriteri olarak KOKH değerinin; geleneksel matris tamamlama yöntemlerine göre 11,54'e 17,23, geleneksel makine öğrenimi algoritmalarına göre ise %19,20'ye 25,16 daha düşük

olarak daha iyi tahmin performansı elde edebileceğini göstermiştir. Seyrek matris problemini DÖ yöntemi ile ilk inceleyen çalışma olan bu çalışmada geleneksel makine öğrenimine göre DÖ yönteminin tahmin yeteneğinin daha iyi olduğu ifade edilmiştir [23].

Barzegar et al. (2020), Makine öğrenme yöntemlerini kullanılarak Yunanistan'ın Küçük Prespa Gölü'ne ait su kalitesi için önemli iki parametre olan ÇO ile klorofil-a (Chl-a) değerlerini tahmin etmişlerdir. Çalışmasında DÖ, DVR ve karar ağacı (KA) modellerini YSA ile birleştirerek tahmin modelleri geliştirmişlerdir. Modellerde giriş değişkeni olarak; pH, El, T, oksidasyon indirgeme potansiyeli (OİP) değerleri kullanmışlardır. Verilerin %70'i eğitmede, %30'u test verisi olarak kullanılmıştır. Model performansları karşılaştırılırken korelasyon katsayısı (R), KOKH, OMH ve normalleştirilmiş yüzde olarak ifade edilen eşdeğerler (NOMH), sapma yüzdesi (SY), NSE ve Willmott's Index (WI) istatistiksel hata kriterlerinden yararlanmışlardır. Çalışmalarının sonucunda, DÖ'nün, ÇO tahmini için YSA modelinden daha iyi performans gösterdiğini, bağımsız DÖ modellerinin Chl-a tahmini için benzer performanslar verdiğini belirtmişlerdir. Ayrıca, Uzun Kısa Süreli Hafıza (UKSH) ve YSA modelleri birleştirilerek oluşturulan hibrit modelin, özellikle ÇO konsantrasyonları için su kalitesi değişkenlerinin hem düşük hem de yüksek seviyelerini başarıyla tahmin ettiğini ifade etmişlerdir [24].

Nghe et al. (2020), Küresel iklim değişikliği ve su kirliliği nedeniyle balık ve karides ölümlerinin bu işi yapan çiftçiler açısından önemli bir sorun olduğu nedeniyle, yaptıkları çalışmalarında bu sorunla mücadele için suyun kalitesinin izlenmesi ve yönetilmesinin öneminden bahsetmişlerdir. Su ürünleri yetiştiriciliği ve balıkçılıkta su kalitesini takip etmek için tahmin modelli bir IoT sistemini (Nesnelerin İnterneti) öneriyorlar. Gerçek zamanlı sensör veri değerleri (tuzluluk, T, pH ve ÇO) takip edilerek erken uyarı almaları ile suyun kalitesi yönetebilir olacaktır. DVM'nin regresyon sürümü ile UKSH algoritması karşılaştırılmıştır. Model performans kriteri olan KOKH değeri UKSH algoritmasında daha iyi sonuçlar vermiştir. Su kalitesi değerlerinin tahmini için UKSH algoritması ile DÖ modeli önerilmiştir [25].

Zhang et al. (2020) yapmış oldukları çalışmalarında, Avustralya'da yer alan Burnett Nehrine ait su kalitesi verileri ile ÇO değerinin tahmini için dört farklı model oluşturmuşlardır. Günümüzde su kalitesinin tahmininin, tarımsal faaliyetlerin yönetimi ve çevrenin korunması için çok önemli olduğu, su kalitesi verilerinin modern veri

toplayıcılar ile toplanması, su kalitesindeki eğilimleri tahmin etmek için veriye dayalı modellerin kullanımına olan ilgiyi artırdığı ifade edilmiştir. ÇO değerinin tahmini için oluşturulmuş dört farklı model; Kernal Temel Bileşenler Analizi (kPCA) ve tekrarlayan sinir ağı kombinasyonuna dayanan bir su kalitesi tahmin modeli, ileri beslemeli sinir ağı, destek vektör regresyonu ve genel regresyon sinir ağıdır. Model performansları kıyaslandığında, kPCA-tekrarlayan sinir ağı modelinin 1, 2 ve 3 saat içinde ÇO konsantrasyonu tahmininde R^2 değerleri; 0.908, 0.823 ve 0.671 olup kPCA- tekrarlayan sinir ağı modelinin tahmine dayalı doğruluğunun diğer üç modellerden (ileri beslemeli sinir ağı, destek vektör regresyonu ve genel regresyon sinir ağı) en az % 8, % 17 ve % 12 daha iyidir. Ayrıca kPCA-tekrarlayan sinir ağı, destek vektör regresyonu ve genel regresyon sinir ağı modelinin aynı anda birden fazla su değişkenini tahmin etmeyi sağlayabileceği de ifade edilmiştir [26].

Cıtakoglu et al. (2021), Sakarya Havzası'nda ÇO değişimini belirlemek için L–Momentleri yöntemi ile bölgeselleştirme çalışması yapmışlardır. Sakarya Havzası'nda bulunan gözlem istasyonlarının ÇO verileri 1995-2014 yılları arasında değerlendirilmiştir. Yapılan bu çalışmada, homojenlik kriteri (h_1) istatistiksel sonuçlarına göre, çalışma alanının homojen olmadığı belirlenmiştir. L–Momentleri yönteminin uygulanabilmesi için, Sakarya Havzası topoğrafik özellikleri göz önünde bulundurularak beş homojen bölgeye ayrılmıştır. Wakeby dağılımının parametre değerleri yardımıyla, istatistiksel boyutsuz ÇO içerik değerleri elde edilmiş ve L–Momentleri yöntemi ile tahmin edilen ÇO içerik verileri görsel olarak Ters mesafe ağırlıklı enterpolasyon yöntemi (IDW) ile ifade edilmiştir. Yapılan bu çalışma sonucunda, Ters mesafe ağırlıklı enterpolasyon yöntemiyle (IDW) elde edilen tematik ÇO tahmin haritalarının daha iyi sonuç verebilmesi için Sakarya Havzası'ndaki istasyonların artırılması ve ayrıca ülke genelinde su kalitesi izleme faaliyetlerinin artırılması belirtilmiştir [27].

Dehghani et al. (2021), Dünya'daki yaşam için en önemli kaynağın su ve oksijen olduğu, ÇO parametresini tahmin için bu çalışmalarında, T ve akış değişkenlerinden yararlanmışlardır. Çalışma alanı, Amerika Birleşik Devletleri'nin güneyindeki Cumberland Nehri olup, 2008'den 2018'e kadar aylık veriler kullanılmıştır. ÇO parametresinin, hem bağımsız hem de hibrit formlarda tahmin edilmesi için destek vektör regresyonu (SVR) kullanılmıştır. En iyi girdi kombinasyonunu belirlemek için Pearson korelasyon katsayısı kullanılmış ve sonuçların yorumlanmasında kutu grafikleri ve

Taylor diyagramlarından yararlanılmıştır. Yapılan çalışma neticesinde, dört hibrit modelin hepsinin bağımsız destek vektör makinesi (SVM) modelinden daha iyi sonuçlar elde ettiği belirtilmiştir. Yapılan bu çalışmada, hibrit modellerin tek destek vektör makinesi (SVM) modelinin doğruluğunu %6.52-1.75 oranında artırdığı ifade edilmiştir [28].

Şener vd. (2022), Şehir şebeke sularının kalite seviyesinin makine öğrenme yöntemleri ve topluluk öğrenme algoritmalarıyla ölçülmesi için çalışma gerçekleştirmişlerdir. Veri kümesi kentsel su kalitesini ölçmeye ilişkin verilerden oluşmakta ve toplam 7999 örnek numune değerine sahiptir. Veri kümesinin %20'si test verisi, %80'i eğitim verisi olarak kullanılmıştır. Veri kümesi önce, makine öğrenme yöntemleri (Lojistik regresyon, K-En yakın komşu, Karar ağacı, Rastgele Orman) ile daha sonra topluluk öğrenme algoritmaları (Uyarlanabilir artırma (AdaBoost), Aşırı gradyan artırma (XGBoost)) kullanılarak analizler gerçekleştirilmiştir. Yöntemler içerisinde en iyi performansı topluluk öğrenme algoritmalarında yer alan XGBoost algoritması vermiştir. Bu algoritmanın genel doğruluk başarısı %96,44 olarak elde edilmiştir. Yapılan çalışma ile kentsel içme suyu şebekelerinde suyun kalite seviyesinin belirlenmesinde yapay zekâ tabanlı yöntemler ile doğruluk oranı yüksek sonuçlar elde edilebileceği ifade edilmiştir [29].

Kumar et al. (2022), Hindistan'da yer alan Chennai şehrine 18 yıldan fazla süredir içme suyu sağlayan 990 dönümlük bir alana sahip Korattur Gölü'nün Chennai bölümünde yaptıkları çalışmalarında, 2009'dan 2019'a kadar su örneklerine ait bulanıklık, toplam çözünmüş katı madde miktarı, KOİ, pH, fosfat, demir, nitrat, sodyum ve klorür parametre değerlerine sahip 5000 veriyi kullanmışlardır. DÖ modeline veriler tanımlanırken 4000 âdeti eğitim seti, 1000 âdeti test seti olarak kullanılmıştır. Su kalite indeksi belirlemek için parametreler ve bunların su kalitesi üzerindeki etki derecelerine göre ağırlıklar belirlenmiştir. Su kalitesi indeksine uygun suyun kalitesini öngörmek için derin öğrenme teknikleri kullanılmıştır. CNN (Konvolüsyonel Sinir Ağları), LSTM (Long Short-Term Memory) ve AutoKeras (Derin Öğrenme Modelini Otomatikleştirme) modelleri, su kalitesi tahmini için kullanılmıştır. CNN ve LSTM manuel derin öğrenme tekniği olup AutoKeras ise bir AutoDL kütüphanesidir. AutoDL, derin öğrenme işlemlerini otomatikleştirmek için yeni bir alandır ve minimum kodla model oluşturma ve yorumlamayı sağlar. AutoDL, henüz başlangıç aşamasındadır. Yapılan bu çalışmada, su

kütlelerini farklı sınıflara ayıran ‘Su Kalitesi İndeksi’ne dayalı suyun kalitesini öngörmek için AutoDL ile geleneksel modeller arasında bir karşılaştırma yapılmış ve su kalitesi değerlendirmesi için AutoDL'nin benimsenmesinin uygunluğunu araştırılmıştır. Çalışma neticesinde, geleneksel DL'nin doğruluğu, ikili sınıf su verileri için AutoDL'nin doğruluğundan %1,8 daha yüksek olduğu, geleneksel DL'nin doğruluğu, çok sınıflı su verileri için AutoDL'nin doğruluğundan %1 daha yüksektir. Geleneksel modelin doğruluğu ~%98 ila ~%99 iken, AutoDL yöntem ~%96 ila ~%98 arasında sonuç vermiştir. Bu çalışma sonucunda; AutoDL modelinin uygun DL modelini bulma da önemli kazanımlar sağladığı ifade edilmiştir [30].

Nacar vd. (2022), Gümüşhane’de bulunan ve şehrin atıksularının etkisi altında kalan Harşit Çayı’nda ÇO konsantrasyon değerinin, çok değişkenli uyarlanabilir regresyon eğrileri (MARS) ve TreeNet gradyan arttırma makinesi (TreeNet) isimli regresyon tabanlı yöntemler ile modelleme çalışması yapmışlardır. Harşit Çayı üzerinde, 6 adet su kalitesi gözlem noktasında 15 gün aralıklarla ve 24 kez yerinde gerçekleştirilen izleme çalışmasında ÇO, T, pH ve Eİ ölçümleri ile akarsudan alınan numuneler üzerinde laboratuvar da gerçekleştirilen sertlik deneyleri sonucunda elde edilen veriler kullanılmıştır. Veri setinin %80’i modellerin eğitilmesinde geriye kalan %20’si ise modellerin test edilmesinde kullanılmıştır. Modellerin performanslarını değerlendirmede ortalama karesel hatanın karekökü (OKHK), ortalama mutlak hata (OMH), ortalama rölatif hata (ORH) ve determinasyon katsayısı (R²) performans istatistikleri kullanılmıştır. TreeNet yönteminin eğitim veri seti için MARS yöntemine göre daha doğru sonuçlar verdiği, test veri seti için MARS yönteminin TreeNet’e göre daha iyi performans sonuçları verdiği belirtilmiştir. Yapılan çalışma sonucunda, MARS yöntemi temel olarak veri setine özel taban fonksiyonları ürettiği ve bağımsız değişkenlerin tüm serisine bir katsayı üretmek yerine daha esnek küçük doğrusal fonksiyonlarla tüm seri için bir denklem elde edildiğinden test veri setinde diğer regresyon yöntemlerinden daha başarılı sonuçlar üretebildiği ifade edilmiştir. Bu çalışmada kullanılan MARS yönteminin diğer regresyon yöntemlerinden daha iyi performans değerleri elde etmesinde kullanım kolaylığı, daha az bilgisayar gücü ve zaman gerektirmesi gibi üstünlükleri bulunduğu belirtilmiştir. Bu çalışma sonucunda, dört farklı su kalitesi değişkeni kullanılarak yapılan ÇO konsantrasyonu değerinin modellemesinde MARS yönteminin tercih edilebileceği ifade edilmiştir [31].

Başakın and Özger (2022), Güneş radyasyonu verilerinin ölçümünde eksik olan değerlerin tamamlanması amacıyla Derin Sinir Ağı (DNN) yöntemini kullanmışlardır. Eksik verilerin tamamlanmasında literatürde sıklıkla kullanılan, Ortalama Atama (MI), Doğrusal İnterpolasyon (LI) ve Spline İnterpolasyon ile karşılaştırma yapmışlardır. Çalışma yapılan istasyon İç Anadolu Bölgesinde yer almakta ve karasal iklim bölgesidir. 2016-2021 yılları arasındaki 1857 veri örneği kullanılmıştır. Derin Sinir Ağı (DNN) yönteminde verilerin %80'ni eğitmede, %20'si test verisi olarak kullanılmıştır. Girdi verisi olarak meteorolojik veriler, çıktı olarak günlük toplam güneş radyasyonu verileri kullanılmıştır. Güneş radyasyonu veri setinde rastgele boşluklar oluşturulmuştur. Modellerin performansları KOKH, Ortalama mutlak yüzde hatası (OMYH) ve NSE değerlerine göre değerlendirilmiştir. Derin Sinir Ağı (DNN) yönteminin diğer klasik yöntemlere kıyasla daha iyi sonuçlar verdiği belirtilmiştir. Bu çalışmayla, Derin Sinir Ağı (DNN) yönteminin bu alanda yapılan çalışmalara önemli katkılar sağlayacağı ifade edilmiştir [32].

Alqahtani et al. (2022), Asya'nın en büyük nehirlerinden birisi olan İndus Nehri'nde yaptıkları çalışmalarında, Tarbela Barajı'nın da yer aldığı Yukarı İndus Havzasındaki Bisham Qilla gözlem noktasına ait su kalitesi verilerini kullanmışlardır. 1975-2005 yılları arasında ölçülmüş 9 su kalitesi parametresi değerleri kullanılmıştır. Bu parametreler; Bikarbonat (HCO_3^-), Kalsiyum, Sülfat, pH, Magnezyum, Klorür, Sodyum, Toplam Çözünmüş Katı ve Eİ'dir. Girdi ve çıktı değerleri arasındaki ilişkiyi belirlemek için veriler istatistiksel yöntemlerle analiz edilmiştir. Nehir suyu tuzluluğunun tahmini amacıyla Eİ ve Toplam Çözünmüş Katı parametreleri tahmin edilmeye çalışılmıştır. Verilerin %70'i eğitmede %30'u test verisi olarak kullanılmıştır. Denetim tabanlı programlar olan Gene Expression Programming (GEP) ve YSA ile eğitim aşamasında çok sayıda karar ağacı oluşturarak problemin tipine göre sınıf veya sayı tahmini yapan bir toplu öğrenme yöntemi olan Random Forest (RF) kullanılmış ve model performansları karşılaştırılmıştır. RF modelinin optimizasyonu için 20 alt model üretilmiştir. R^2 değerine göre model sonuçları kıyaslandığında, GEP modeli 0.96, RF modeli 0.98 ve YSA modeli ile 0.92 değeri elde edilmiştir. Bu elde edilen sonuçlar ile girdiler, model çıktıları ve sonuçlar arasında güçlü bir ilişki olduğu belirtilmiştir. Yapılan çalışma sonucunda; en iyi sonucu RF modelinin verdiği, YSA modelinin performansının düşük olduğu, en etkili olan değişkenin Bikarbonat olduğu ve diğer etkili değişkenlerin sırasıyla Klor ve Sülfat

olduđu belirtilmiřtir. Ayrıca seilen parametrelere sahip RF modelinin su kalitesi deđerlendirmesi iin diđer modellere gre nceliklendirilebileceđi ifade edilmiřtir [33].

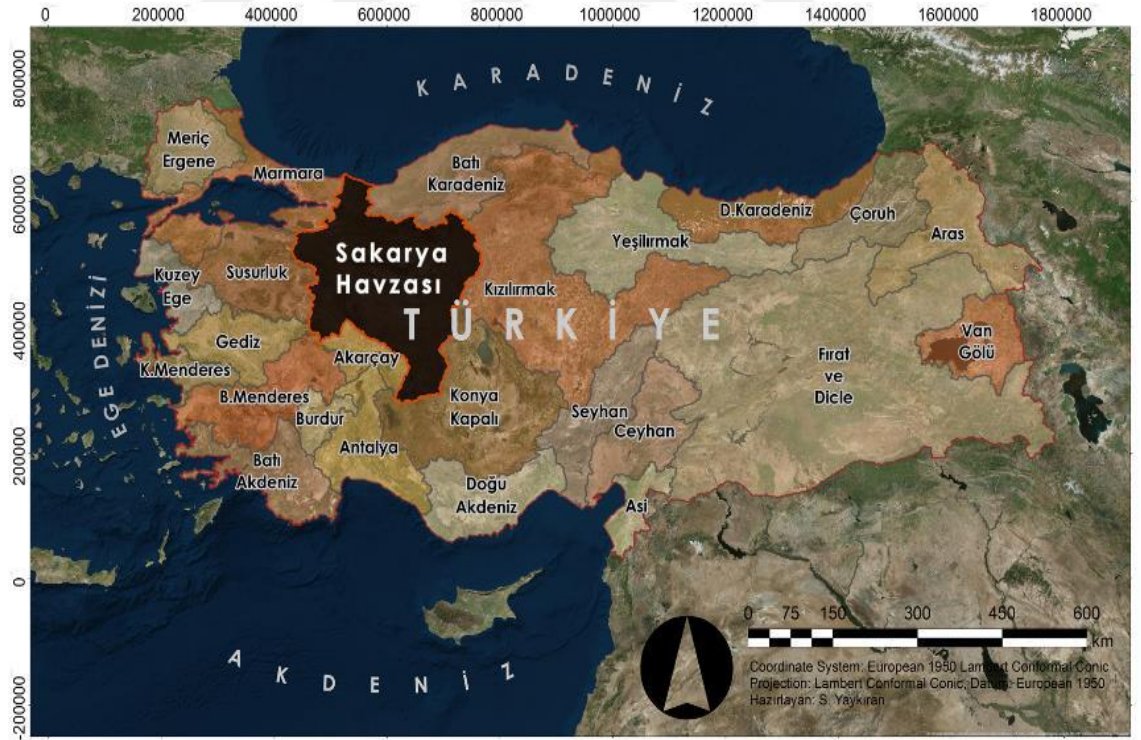


2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Sakarya Havzası

Sakarya Havzası, ülkemizin kuzeybatısında bulunmakta olup havzanın komşuları, batısında Susurluk, güneyinde Akarçay ve Konya Kapalı Havzaları ile doğusunda Kızılırmak ve Batı Karadeniz Havzaları ile çevrilidir. Şekil 1’de Sakarya Havzası ve diğer nehir havzalarının konumu yer almaktadır.



Şekil 1. Sakarya Havzası ve diğer nehir havzalarının konumu [34]

Ülkemizin, 25 hidrolojik havzasının 13 tanesi müstakil nehir havzasıdır. Sakarya Havzası da müstakil nehir havzalarındandır. Havza’nın ana kolu olan Sakarya Nehri; Eskişehir’in güneyinde yer alan Çifteler Sakaryabaşı kaynaklarından doğarak; Ankara Çayı, Porsuk

Çayı, Göksu Çayı, Karasu Çayı, Çarksuyu, Mudurnu gibi yan kollar ve birçok derelerle beslenerek, Sakarya ili Karasu ilçesi yakınlarından Karadeniz'e dökülmektedir [5]. Havzaya ait karakteristik değerler Tablo 1'de verilmiştir [34].

Tablo 1. Sakarya Havzasına ait karakteristik değerler

Karakteristik	Değer	Birim
Drenaj Alanı	58.160	km ²
Nüfus (2012)	7.588.968	kişi
Nüfus Yoğunluğu	130	kişi/ km ²
Yıllık Ortalama Akış	6.40	km ³
Yıllık Ortalama Akış	193	m ³ /s
Ortalama Yıllık Akış Yüksekliği	88.90	mm
Türkiye Su Potansiyeline Katkısı	3.4	%
Ortalama Yıllık Verim	3.6	L/s/ km ²
Ortalama Yıllık Yağış Yüksekliği	479	mm
Ortalama Yıllık Yağış	32×109	m ³
Ortalama Rakım	965	m
Ortalama Sıcaklık	10.6	°C
Drenaj Alanı	58.160	km ²

Sakarya Nehri, yaklaşık olarak 58.000 km²'lik drenaj alanı ve 720 km'lik uzunluğu ile ülkemizdeki hidrolojik havzaların alansal büyüklük açısından 3.'sünü oluşturan ve taşıdığı su ile de havzasının en önemli kaynaklarından. Sakarya Nehri, bölgedeki diğer dere ve çayların birleştiği ana akarsu olması nedeniyle havzanın ana kolu olarak bilinir ve hemen hemen her mevsimde debisi yüksek ve ılıktır. Nehrin debisi ilkbaharda yükselmeye başlamakta, mart, nisan ve mayıs aylarında debi yüksek kalmaya devam etmektedir. Debinin en düşük olduğu aylar da temmuz, ağustos ve eylül gibi yaz sonu aylarıdır. Sakarya Nehri'nin suları yerleşim alanlarının içme ve temel ihtiyaçlarını gidermek için gerekli olan suyu karşılamının yanında, tarımsal ve sanayi kuruluşlarında kullanılan su ihtiyacını gidermek, rekreasyon ve atık su deşarjı gibi birçok farklı amaç için de kullanılmaktadır.

Sakarya Havzasının ortalama kotu 965 m, yıllık ortalama yağış miktarı 479 mm'dir. Havza yıllık ortalama yağış miktarı olan 479 mm değeri ile yıllık ortalama yağış miktarı 500 mm'den az olan ülkemizdeki dört havzadan biridir.

Sakarya Havzasında arazi kullanımı incelendiği zaman arazilerin yaklaşık olarak % 50'ye yakın kısmı tarımsal alan, % 45'lik kısmı ormanlık ve doğal alan, %2.5'luk kısmı şehir vb. insan yaşamının olduğu ve doğal yapısının bozulduğu yapay alanlardır. Havzada sulak alanlar ile yüzey suları yaklaşık % 0.5'lik alanı kaplamaktadır.

Sakarya Havzası, ülkemiz yüzölçümünün yaklaşık % 7'sini kaplar ve ortalama yüksekliği 508 m'dir. Sakarya Havzası, sınırlarının uç noktaları arasındaki mesafe; doğu-batı doğrultusunda yaklaşık 340 km, kuzey-güneydoğu doğrultusunda yaklaşık 378 km olduğundan dolayı geniş bir alanı kaplar. Havza sahip olduğu geniş alan sebebiyle çeşitli iklimlerin etkisi altındadır. Havzanın kuzey kısımlarında Karadeniz bölgesinde görülen ılıman ve yağışlı iklim hâkimken, orta ve güney kısımlarda İç Anadolu bölgesinde görülen yazların sıcak ve az yağışlı, kışlar ise soğuk ve karlı geçtiği iklim etkilidir.

Sakarya Havzasının en önemli nehirleri; Sakarya Nehri ve kolları olan Porsuk Çayı, Karasu Çayı, Ankara Çayı, Çubuk Çayı, Koca Çay, Kirmir Çayı, Çark Suyu ve Mudurnu Çayıdır. Sapanca, Mogan ve Eymir Gölleri de havzada yer almakta olup Porsuk, Sarıyar, Gökçekaya, Kurtboğazi ve Çamlıdere Baraj Gölleri de bulunmaktadır.

Havzada tarımın önemli bir yer tutmasından dolayı, sulama amaçlı birçok baraj ve gölet ile sulamada kullanılan tesisler yer alır. Ayrıca ülkemizin başkenti Ankara ile beraber önemli kent merkezlerinin içme ve kullanma suyunu sağlayan Çubuk I, Çubuk II, Çamlıdere, Kavşakkaya, Bayındır, Eğrekkaya, Kurtboğazi, Akyar ve Porsuk Barajları havzada yer alır.

Sakarya Havzası'nda birçok sulak alan bulunmaktadır. Bu sulak alanlar içerisinde uluslararası önemde olanlar, havzanın yukarisından aşağı doğru olacak şekilde;

- ✓ Eskişehir'de Sakarya Nehri'nin yukarı kesiminde bulunan Balıkdami,
- ✓ Konya ilinin Yunak ilçesinde yer alan Akgöl,
- ✓ Konya'nın Ilgın ilçesinde bulunan Ilgın (Çavuşçu) Gölü,
- ✓ Ankara ilinde Mogan-Eymir Gölleri ve çevresi,
- ✓ Eskişehir-Ankara il sınırında Sarıyar Baraj Gölü,

- ✓ Havzadaki en önemli göl olan Sakarya ilinde bulunan Sapanca Gölü ve çevresidir [34].

Sakarya Havzası, Ankara ve İstanbul gibi büyük kentler arasında olması nedeniyle ekonomik ve kültürel faaliyetler, ulaşım ve sağlık imkânları yönünden Türkiye'nin diğer bölgelerine nazaran daha gelişmiş bölgesi durumdadır. Özellikle Eskişehir, Ankara, Sakarya, Bilecik ve Kütahya illeri ağırlıklı olmak üzere havzada geniş bir alanda endüstriyel faaliyet sürdürülmektedir. Diğer taraftan, havzada su kirliliği sorununun nüfus ve endüstrileşmeyle birlikte hızlı bir şekilde arttığı, nehrin bazı kollarının evsel ve endüstriyel atık sular ile tarımsal faaliyetlerin etkileri sonucunda kirlendiği bilinmektedir.

Sakarya Havzası, Türkiye'de endüstriyel faaliyetlerin yoğun olduğu havzalardan birisidir. Havzada birçok önemli kent bulunmakta ve bu kentlerde birçok Organize Sanayi Bölgesi vardır. Bu sanayi bölgelerinde birçok farklı endüstriyel faaliyet yapılmaktadır.

Sakarya Havzasında endüstriyel faaliyetlerin yanında tarım ve hayvancılık faaliyetleride önemli yer tutar. Hayvancılık ve tarımsal faaliyetler havzanın birçok yerinde yoğun bir şekilde yapılmaktadır. Ülkemizin tahıl ambarlarından birisi olarak tanımlayabileceğimiz Ankara'da bulunan Polatlı Ovası, sebze yetiştiriciliğinin yoğun olduğu Ankara'nın Beypazarı İlçesi, Bursa'da bulunan Yenişehir Ovası, ülkemizin meyve ve sebze yetiştiriciliğinde önemli yer tutan Pamukova ve tarımsal faaliyetlerin yoğun olduğu Adapazarı Ovası havzada yer alan ve tarımsal faaliyetlerin yoğun bir şekilde yapıldığı alanlardır.

Tarımsal faaliyetlerde bilinçsiz gübre kullanımı birçok soruna neden olabilmektedir. Azotlu, fosforlu ya da potasyumlu tuzlar içeren kimyevi gübrelerin ihtiyaç fazlası kullanılması halinde, zamanla toprakta birikmesinden dolayı topraklar tuzlanacak ve toprağın kalitesi bozulacaktır. Zamanla toprakta biriken gübreler su kaynaklarına karışarak kirlenmelerine sebep olabilecektir.

Havzadaki su kaynakları, yürütülen birçok farklı faaliyetten kaynaklı baskı unsurlarının tehdidi altındadır [34].

Su kalitesi üzerinde etkili olabilecek baskı unsurları olarak;

- ✓ Tarımsal ve hayvancılık faaliyetleri sonucunda oluşması muhtemel kirlenme,

- ✓ Evsel ve endüstriyel atık suların arıtılmadan deşarj edilmesi,
- ✓ Nehir yataklarında malzeme alımından kaynaklı bozulmalar örnek gösterilebilir.

Sakarya Havzasında birçok gelişmiş şehir bulunmakta ve ülkemiz nüfusunun önemli bir kısmı bu şehirlerde yaşamaktadır. Kentleşme ve belediyeçilik çalışmaları neticesinde, atık su arıtma ve kanalizasyon konusunda önemli mesafe alınmıştır.

2.1.1. Gözlem İstasyonları

Su kaynaklarının kullanım alanlarına göre kalitesinin takibinin düzenli bir şekilde yapılması gerekmekte ve önlem alınmasını gerektirecek durumlar ortaya çıktığında da hızlı bir şekilde karar alınmasını sağlanmak gerekmektedir. Devlet Su İşleri (DSİ) Genel Müdürlükleri tarafından Sakarya Havzasında su kalitesi izleme programı yürütülmekte olup Tablo 2’de Sakarya Havzası’nda yer alan gözlem istasyonlarına ait bilgiler yer almaktadır.

Tablo 2. Sakarya Havzası’nda yer alan gözlem istasyonlarına ait bilgiler

Sıra No	İstasyon Adı	İstasyon No	Havza
1	Seydisuyu-Kozyaka	12-03-00-008	Sakarya
2	Sakarya Nehri-Kavuncu	12-03-00-009	Sakarya
3	Porsuk Çayı-Ağaçköy	12-03-00-010	Sakarya
4	Porsuk Çayı-Çalça	12-03-00-011	Sakarya
5	Porsuk Çayı-Başdeğirmen	12-03-00-012	Sakarya
6	Porsuk Çayı-Benzinlik	12-03-00-013	Sakarya
7	Porsuk Çayı-Şeker Çiftliği	12-03-00-015	Sakarya
8	Porsuk Çayı-Beylikova	12-03-00-017	Sakarya
9	Porsuk Çayı-Yunus Emre	12-03-00-018	Sakarya
10	Porsuk Çayı-Sazılar (İlören)	12-03-00-019	Sakarya
11	Sakarya Nehri-Karacaahmet	12-03-00-020	Sakarya
12	Ankara Çayı-Meşelik	12-03-00-029	Sakarya
13	Sakarya Nehri-Dümrek	12-03-00-030	Sakarya
14	Sakarya Nehri-Yenice	12-03-00-031	Sakarya
15	Sakarya Nehri-Doğançay	12-03-00-034	Sakarya
16	Çarksuyu-Beşköprüler	12-03-00-035	Sakarya
17	Çarksuyu-Söğütlü	12-03-00-039	Sakarya
18	Kocadere-Karasu Öncesi	12-03-00-056	Sakarya
19	Porsuk Çayı-Alpu	12-03-00-072	Sakarya

2.1.2. Alt Havzalar

Sakarya Havzası, nehrin doğduğu yerden Karadeniz'e döküldüğü yere kadar 6 alt havzaya bölünmek suretiyle incelenmiştir [34]. Şekil 2'de Sakarya Havzası'na ait alt havzalar yer almakta olup, bu alt havzalar;

- ✓ Yukarı Sakarya Alt Havzası
- ✓ Porsuk Çayı Alt Havzası
- ✓ Ankara Çayı Alt Havzası
- ✓ Orta Sakarya Alt Havzası
- ✓ Göksu-Karasu Çayları Alt Havzası
- ✓ Aşağı Sakarya Alt Havzası şeklindedir.



Şekil 2. Sakarya Havzası'na ait alt havzalar [34]

2.2. Yapay Sinir Ağları (YSA)

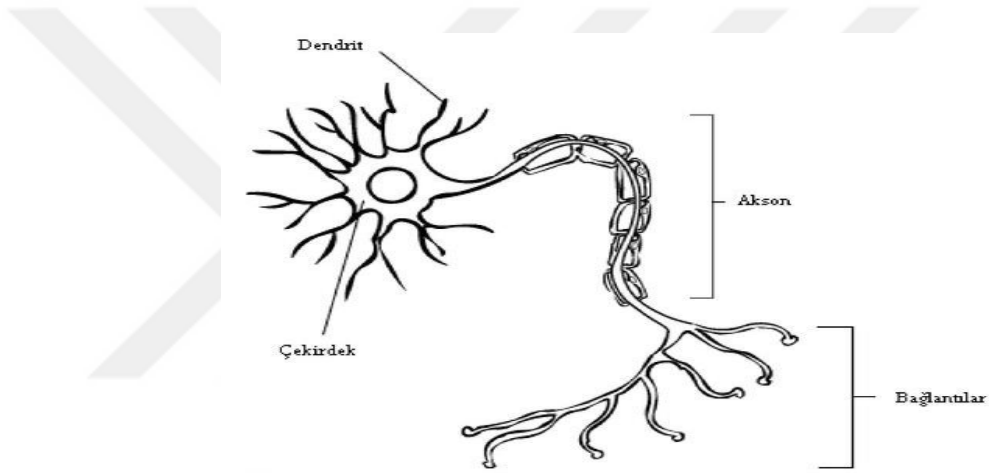
Modern dünyada gelişen teknoloji yapay zekâ yöntemleri kavramının önemini daha fazla artırmıştır. İnsan beyni gibi düşünebilme ve öğrenebilme özelliklerinin bilgisayar programlarına uyarlanması ile bilgisayarlara insan beynine özgü düşünebilen ve davranabilen sistemlerin kazandırılması, yapay zekâ çalışmalarının temelini oluşturmaktadır. Yapay zekâyâ yönelik çalışmalarda yapay sinir ağları konusu önemli yer tutar. Bilim insanları, insan beyninin çalışma özellikleri üzerinde araştırma ve sorgulama yapmışlar ve beynin çalışma işlevinden esinlenerek matematiksel model oluşturmaya çalışmışlardır. Model oluşturma çalışmalarında, beynin sahip olduğu tüm bileşenlerin doğru olarak yer alması önemli olduğundan birçok ağ modeli oluşturulmuştur. Bu şekilde model oluşturma çalışmaları ile YSA ortaya çıkmıştır. YSA, günümüz bilgisayarlarının sahip olduğu algoritmik hesaplama yönteminden farklı ve yeni bir bilim dalı olarak ortaya çıkmıştır [35].

YSA model çalışmalarında esas olarak beynin nörofiziksel yapısından esinlenerek matematiksel model oluşturulmaya çalışılır. İnsan beyninin en önemli özellikleri olan öğrenme, verileri işleyebilme ve depolama ile genelleme yapma özellikleri YSA ile taklit edilmeye çalışılır. YSA ile bilgisayarlar, insan beynini taklit ederek öğrenme ve düşünme özelliklerini programlamaya çalışırlar. YSA insan beyninin en önemli özelliği olan öğrenme fonksiyonunu sağlayan bilgisayar sistemleridir. Örnekler yardımı ile öğrenme işlemini gerçekleştirirler. Öğrenme işlemi bazı kurallar tanımlayarak, giriş çıkış bilgileri verilerek gerçekleştirilir.

YSA alanında yapılan çalışmaların tarihçesi incelendiğinde, her ne kadar 1940'lı yıllarda çalışmalarına başlamış olsalar da, W.A. Pitts ve W.S. McCulloch'ın 1943 yılında yayınladıkları makaleleri YSA modellenen hesaplama yönteminin temellerini oluşturur. Daha sonraki yıllarda B.G. Farley ve W.A. Clark tarafından 1954 yılında oluşturulmuş ağ model çalışmasında, ağ içerisinde uyarılara tepki veren uyarılara adapte olabilen model geliştirilmiştir. 1960 yılına geldiğimizde ise ilk neural bilgisayarın ortaya çıktığı yıldır. 1970 ve 1980'lerde modellerdeki eksikliklerin fak edilmesi ve giderilmesine yönelik çalışmalar hız kazanmıştır. YSA ile ilgili araştırmaların ve tanınırlığın arttığı 1985 yılı dönüm noktası sayılabilecek bir yıldır. YSA günümüzde, nesneleri veya görüntüleri ayırma ve tanıma, sistem tanımlama, ses ve görüntü tanıma gibi birçok farklı alanda

kullanılmaktadır. YSA'dan birçok problemin çözümünde faydalanılmakta ve önemli kolaylıklar sağlamaktadır. YSA modelleri sayısal veri esasına göre çalıştığından bazı problemlerin çözümünde sayısal veriye erişememek veya dönüştürülememesi YSA'nın dezavantajı olarak görülebilir.

Biyolojik sinir ağlarının en temel bileşeni biyolojik sinir hücresidir. İnsan beyinde yaklaşık 1011 sinir hücresi bulunur. Bu sinir hücreleri 1000-10000 arasında değişen başka hücrelerle ilişki içerisinde bulunarak bu sinirlerden uyarı alır. Şekil 3'de biyolojik sinir ağ yapısı yer almakta ve sinir hücresinin temel bileşenleri olan çekirdek, dendrit, bağlantılar ve akson bulunmaktadır.



Şekil 3. Biyolojik sinir ağ yapısı [36]

Şekil 3'de hücre gövdesi ve çekirdek, akson, dendrit (sinir ucu) ile sinir hücresinin sinir ucu arasında kalan ince bağlantılar (sinaps) olmak üzere dört bölümden oluşan biyolojik sinir hücresi yer almaktadır. Bu bileşenleri inceleyecek olursak dendritler, gelen sinyallerin çekirdeğe iletildiği kısımdır. Çekirdek bölümünde gelen sinyaller toplanır ve aksona iletir. Aksonda toplanan bu sinyaller daha sonra işlenerek sinapslara gönderilir ve sinapslarda yeni üretilen bu sinyalleri diğer sinir hücrelerine iletirler.

Biyolojik sinir ağlarının yapısında gördüğümüz sinir hücresinin yapay sinir ağlarında karşılığı olarak yapay sinir hücrelerini görebiliriz. Yapay sinir hücreleri yerine bazen proses elemanları kavramı da kullanılmaktadır. Tablo 3'de biyolojik sinir ağ yapısının yapay sinir ağındaki karşılıkları yer almaktadır.

Tablo 3. Biyolojik sinir sisteminin yapay sinir sistemindeki karşılığı

Biyolojik Sinir Sistemi	Yapay Sinir Sistemi
Sinir	İşlemci eleman
Dendrit	Toplama fonksiyonu
Hücre gövdesi	Transfer fonksiyonu
Aksonlar	Yapay nöron çıkışı
Sinapslar	Bağlantı ağırlıkları

YSA’lar, giriş veri seti olarak kendisine verilen bilgilere karşılık çıkış verisi üretebilme özelliğine sahiptir. YSA modeli, yapay sinir hücrelerinin birbirleriyle farklı şekillerde bağlanmasından ve farklı katmanların bir araya gelmesinden oluşur. Donanım olarak genellikle farklı türde yazılımlar geliştirilir. YSA temel çalışma prensibi insan beynini taklit etmek olduğu için öğrenme aşamasından sonra beynin bilgi işlemesine benzer bilgiyi saklar ve genelleme işlemi yapar. Genelleme işlemini yapabilmesi YSA’nın en büyük özelliği olarak görülür. Kendisine sunulan girdi parametreleri ile eğitime işlemi gerçekleştirilir ve genelleme yapar, bu genellemenin sonrasında tanımlanan girdi parametrelerine karşılık çıktı parametresini oluşturur.

2.2.1. Yapay Sinir Ağının Yapısı

YSA’da, sinir hücrelerinin grup şeklinde işlev gördükleri durum ‘ağ’ olarak adlandırılır ve bu şekildeki bir grupta binlerce nöron bulunabilir. Bu şekildeki yapay nöronların birbirlerine bağlanarak bir araya gelmeleri ‘yapay sinir ağı’ oluşturmaktadır. Burada amaçlanan biyolojik sinir ağının bir modelinin yapay sinir ağıyla oluşturulmak istenmesidir. Aynı doğrultu üzerinde nöronların bir araya gelmesiyle katmanlar oluşur ve katmanların değişik şekilde birbirleriyle bağlanmaları değişik ağ mimarilerini oluşturur [12]. YSA girdi, ara ve çıktı olmak üzere üç katmadan oluşur.

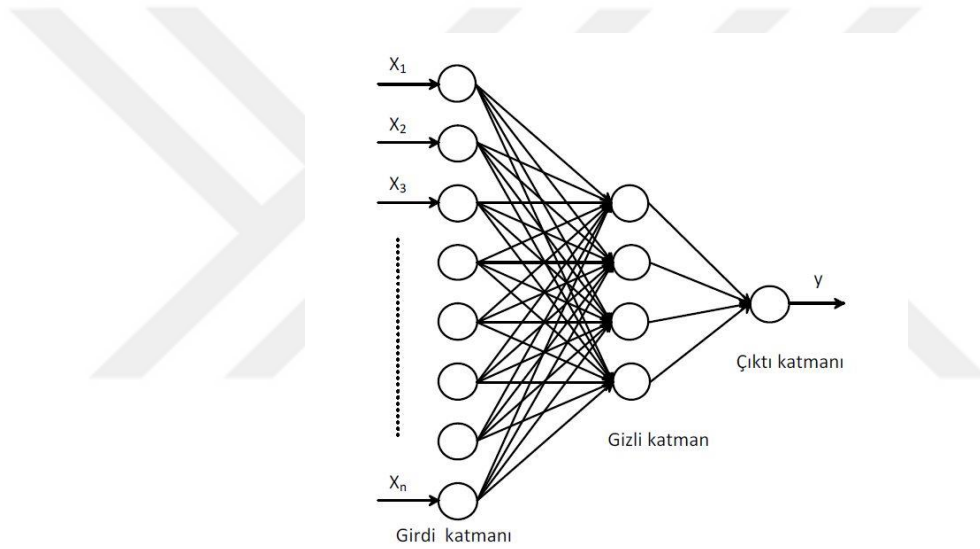
$$u_k = w_{k1} \cdot x_1 + w_{k2} \cdot x_2 + \dots + w_{km} \cdot x_m \quad (2.1)$$

$$u_k = \sum_{j=1}^m w_{kj} \cdot x_j \quad (2.2)$$

$$y_k = u_k \cdot \varphi \quad (2.3)$$

Eşitlik 2.1, 2.2, ve 2.3'deki formülleri incelediğimiz zaman, $x_1, x_2, \dots, x_n$ nöron girişleridir. Her bir nöron girişi toplama işlemine girmeden önce $w_{k1}, w_{k2}, \dots, w_{kn}$ ağırlıkları ile çarpılmaktadır. Toplam olarak, girişler X vektörü ile, ağırlıklar ise W vektörü ile ifade edilmiştir. Formüllerde yer alan; u_k , lineer toplama çıkışı, ϕ , aktivasyon fonksiyonu ve y_k , çıkış nöronudur.

Aynı doğrultu üzerindeki yapay nöronlar katmanları oluştururlar ve katmanların farklı şekillerde bağlanmasıyla YSA oluşur. YSA'da, yapay nöronların sınıflandırılmasında tabakalar ve birbirleriyle etkileşimi önemlidir. Şekil 4'de yapay sinir hücresi katmanları gösterilmektedir. Girdi, ara (gizli) ve çıktı katmanı YSA'nın temel katmanlarını oluşturur.



Şekil 4. Yapay Sinir Ağının yapısı [37]

2.2.1.1. Girdi Katmanı

Girdi katmanı ($x_1, x_2, x_3, \dots, x_n$), gelen bilgilerin alındığı katmandır. Buradaki veriler ara katmana aktarılır. Her nöronun sadece bir çıkış verisi olması yanında sınırsız sayıda veri girişi alabilir.

2.2.1.2. Ara (Gizli) Katman

Ara ya da gizli katman olarak ifade edilen bu bölümde girdi katmanından gelen veriler işlenerek çıktı katmanına gönderilir. Birden fazla ara katman bulunabilir.

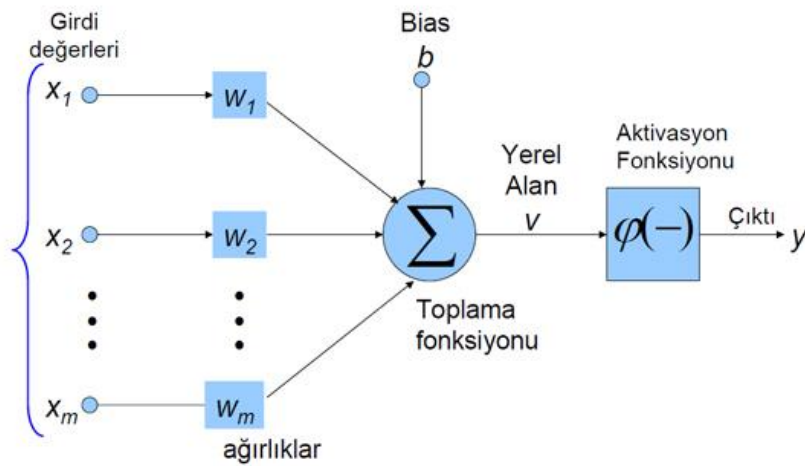
2.2.1.3. Çıktı Katmanı

Girdi katmanındaki veriler için üretmesi gereken çıktıyı, ara katmandan gelen bilgileri işleyerek üretirler.

Tek bir katmandan oluşan YSA model yapıları tek katmanlı sinir ağı, çok sayıda ara (gizli) katmandan oluşan model yapıları çok katmanlı sinir ağı şeklinde ifade edilir.

2.2.2. Yapay Sinir Hücresinin Bölümleri

Yapay sinir ağındaki katmanlarda yapılan işlevleri gerçekleştiren elemanlara ‘yapay sinir hücresinin bölümleri’ denir. Şekil 5’de gösterildiği gibi bu bölümler; girdi değerleri, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktılarıdır.



Şekil 5. Yapay Sinir Hücresinin bölümleri [38]

2.2.2.1. Girdiler

Yapay sinir hücresine dış ortamdan gelen bilgilerdir. Bu gelen bilgiler biyolojik sinir hücrelerindeki gibi benzer şekilde toplanmak amacıyla nöron çekirdeğine gönderilir.

2.2.2.2. Ağırlıklar

Yapay sinir hücresine gelen veriler ilk olarak girdiler bölümünde toplanır ve çekirdeğe bağlantıların ağırlığıyla çarpılarak iletilir. Bu şekilde üretilen çıktı değerleri girdilerin belirlenecek ağırlık değeriyle ilişkilidir.

2.2.2.3. Toplama Fonksiyonu (Birleştirme Fonksiyonu)

Toplama fonksiyonu (birleştirme fonksiyonu), yapay sinir hücresinde ağırlıklarla çarpılarak gelen bilgilerin toplandığı ve o hücreye ait net girdinin hesaplandığı bölümdür. Hesaplama sonrasında toplamı transfer işlevine aktaran fonksiyondur.

2.2.2.4. Aktivasyon Fonksiyonu

Toplama fonksiyonunda girdilerin ağırlıklı toplamı bu fonksiyona aktarılır ve bu bölümde bir çıkış değeri (genellikle doğrusal olmayan) üretilir. Üretilen çıkış değeri bir sonraki katmana aktarılır. ReLU, tanjant, sigmoid vb. yaygın olarak kullanılan aktivasyon işlevleridir.

2.2.2.5. Çıkışlar

Aktivasyon fonksiyonundan üretilen çıkış değeri yapay sinir hücrenin çıktı değeri olmaktadır. Yapay sinir hücreleri birden fazla girdiye sahip olsa da bir tek çıktısı olur.

YSA modelleri yapısal olarak (doğrusal durumu) incelediği zaman, genel olarak ileri beslemeli ve geri beslemeli olarak ikiye ayrılır. İleri Beslemeli YSA'da yer alan nöronlar girişten çıkışa doğru düzenli katmanlar şeklinde yer alır. Girdi katmanına gelen bilgiler herhangi bir değişime uğratılmadan ara (gizli) katmandaki hücelere iletilir. Sonrasında çıkış katmanından işlenerek çıktı değeri dış ortama aktarılır. Geri Beslemeli YSA'da ileri beslemeli ağlardaki gibi bir nöronun çıktısı sadece kendinden sonra gelen nöron katmanına girdi olarak iletilmez. İleri beslemeli ağlardan farklı olarak nörona ait çıktı verisi kendinden önceki katmanda veya kendi katmanında bulunan herhangi bir nörona girdi verisi olarak bağlanabilir. Bu özelliğinden dolayı geri beslemeli YSA'nın doğrusal olmayan dinamik bir davranış gösterdiği belirtilmiştir [39].

YSA'ların en önemli özelliklerinden birisi olarak öğrenme yeteneği ön plana çıkmaktadır. Öğrenme işlemi, eğitme, deneme ve yanıtlar neticesinde sağlanır. YSA hücrelerinin arasındaki bağlantılarda kullanılan ağırlıkların belirlenmesi açısından öğrenme işlemi önemlidir. Farklı öğrenme algoritmalarına göre YSA'lar; danışmanlı, danışmansız ve pekiştirmeli öğrenme olarak sınıflandırılabilir.

Danışmanlı öğrenmede, YSA kullanılmaya başlamadan önce eğitilir. Eğitim aşamasında sisteme hem girdi hem de o girdi değerlerine ait çıktı değerleri verilir. Ağın ürettiği çıktı ile istenen çıktı karşılaştırılarak hata hesaplaması ve ağırlıkların değerleri belirlenir.

Danışmansız öğrenmede, öğrenmeye yardımcı olacak danışman olmadığı için sisteme girdiler verilir ve örnekler yardımıyla sistemin öğrenmeyi kendisinin sağlaması beklenir.

Pekiştirmeli öğrenmede, giriş verileri uygulandıktan sonra sonuç danışman tarafından değerlendirilir. Ödül-ceza yöntemiyle ağın ağırlıkları güncellenir. Danışmanlı öğrenmeye benzerliği olsa da sisteme çıkış değeri verilmemesi ayırt edici yanındır.

2.3. Derin Öğrenme (DÖ)

Makine öğrenmesi, 1980 yıllarda ortaya çıkmış ve sisteme verilen bir veri kümesini işleyerek tahmin veya sınıflandırma yapmak amacıyla kullanılan bir yöntemdir. Yapay zekâ gibi makine öğreniminde de, uzun yıllar boyunca önemli bir gelişim olmamıştır. Veri madenciliğinde 1990'lı yıllarda yaşanan gelişim ile makine öğrenmesinin önemi artmış ve 2000'li yılların başında DÖ alanında da önemli gelişimler olmuştur. Yaşanan teknolojik gelişmeler ile bilgisayarların performansları arttı ve veri setleri ile algoritmalar geliştirilerek DÖ alanında iyileştirilme sağlandı.

DÖ, bir makine öğrenmesidir. Makine öğrenmesi, yazılım ve donanım alanında yaşanan gelişmelerle birlikte önemi daha fazla artmış bir yöntem olup, büyük veri setlerini işleyen bilgisayarların kendi başlarına öğrenmesidir. DÖ ile verilen bir veri kümesi ile çıktıların tahmin edilmesi amaçlanır. Yine öğrenme işlemi eğitime ile gerçekleştirilecek ve eğitmek için hem denetimli hem de denetimsiz öğrenme kullanılabilecektir.

YSA'nın yapısında insan beyninin taklidi amacıyla nöronlar bulunur. Tüm nöronlar birbirine bağlıdır ve sistemin çıktısını etkilemektedir. Girdi, gizli ve çıktı olmak üzere üç katman vardır. DÖ'deki, 'derin' isimlendirmesi birden fazla gizli katmana sahip olmaktan gelmektedir.

Problemlerin türüne ve verilerin özelliklerine göre çözümde kullanılacak YSA modelleri farklılık gösterecektir.

2.3.1. Sinir Ağı Türleri

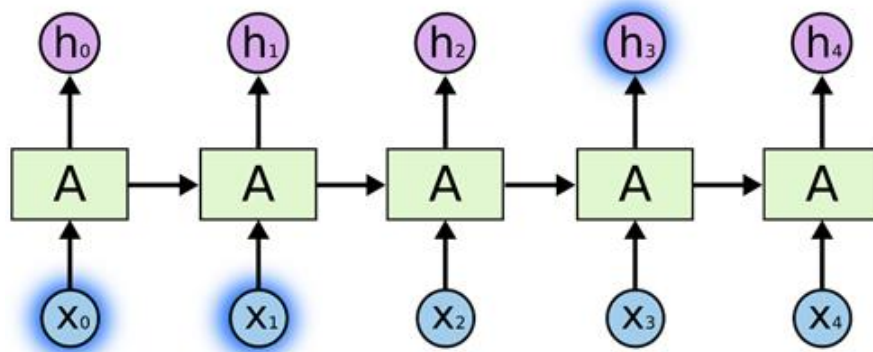
2.3.1.1. Konvolüsyonel Sinir Ağları (Convolutional Neural Network)

Konvolüsyonel sinir ağı, çok katmanlı ileri beslemeli bir YSA'dır. DÖ yapıları içerisinde en çok kullanılan sinir ağı olup, görüntü analizlerinin yapılmasında önemli kolaylık sağlar.

Çok katmanlı bir sinir ağı türü olup içerisinde fazla sayıda konvolüsyon katmanı, aktivasyon katmanı, sınıflandırıcı katman, tam bağlı katman, havuzlama katmanı ve ilave ek katmanlar yer almaktadır. Her katman kendisine ait işlevleri yürütür ve son olarak sınıflandırıcı katmanda sonuç verilir [40].

2.3.1.2. Tekrarlayan Sinir Ağı (Recurrent Neural Network)

Tekrarlayan sinir ağları, sadece sisteme verilen giriş verilerini değil daha önce sisteme verilmiş zaman serisi giriş verilerini de alır. Sisteme giren veriler ardıl şekilde kullanılır ve sistemdeki çıktı verileri önceki hesaplamalara bağlıdır. Bu sinir ağı arka arkaya işlemlerden sonra gelecek işlemin tahmininde kullanılır. Finans, dil çevrimleri vb. çalışmalar olabilir. Zaman dizeleri aralarında boşluk olması tahmin performansını etkiler. Şekil 6'da tekrarlayan sinir ağı yapısı yer almaktadır.



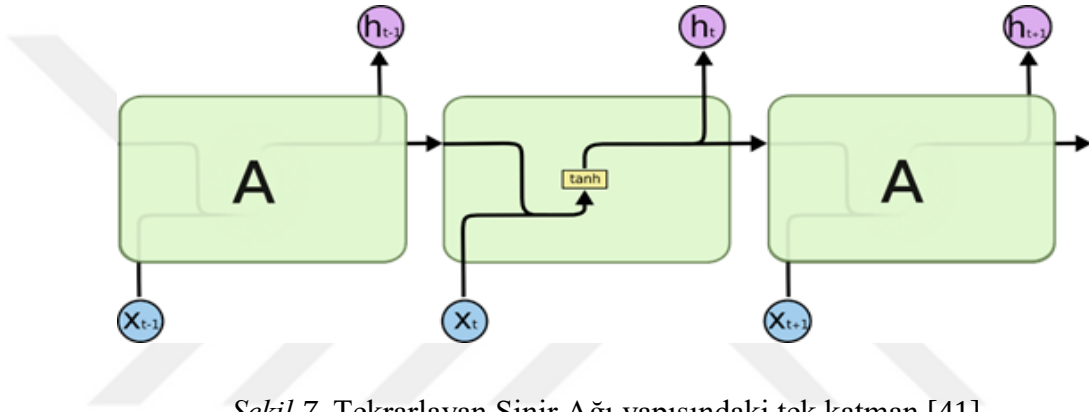
Şekil 6. Tekrarlayan Sinir Ağı yapısı [41]

2.3.1.3. Uzun Kısa Süreli Hafıza (UKSH) (Long Short-Term Memory (LSTM))

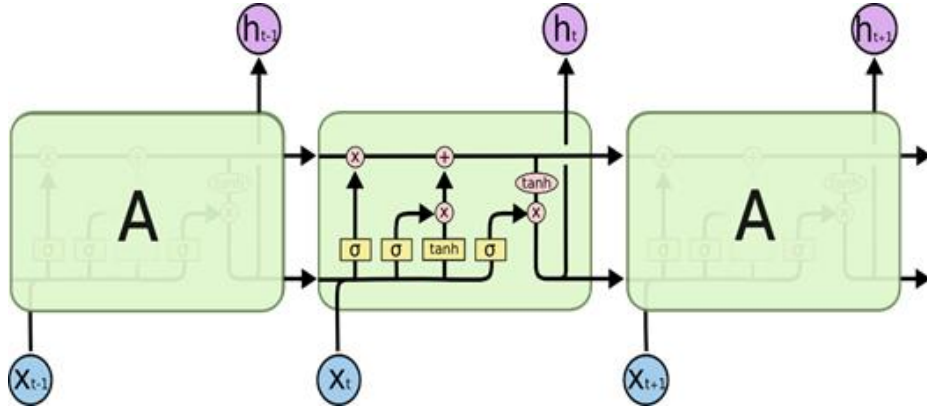
UKSH, DÖ alanında kullanılan bir tekrarlayan sinir ağı mimarisine sahip makine öğrenmesidir. Bu ağ mimarisi Jürgen Schmidhuber ve Sepp Hochreiter tarafından 1997

yılında oluşturulmuş ve UKSH'nın ilk sürümü hücreler, giriş ve çıkış kapılarından oluşmaktadır. Daha sonra unut kapısı da sisteme eklenmiştir. Unut kapısı ile hücrenin değişken uzunlukta zaman aralıklarındaki değerleri hatırlaması sağlanır. Giriş, unut ve çıkış kapılarında hücreye giren ve çıkan veri akışı düzenlenir.

Tekrarlayan sinir ağı yapısının yer aldığı Şekil 7 de görüleceği gibi tekrarlayan kısımda (X_t) tek bir tanjant katmanı (tanh) bulunur. Şekil 8'de UKSH ağ mimarisi görülmektedir. Şekil 8'de UKSH'daki tek bir sinir katmanı yerine dört etkileşimli katman yapısı bulunmaktadır.



Şekil 7. Tekrarlayan Sinir Ağı yapısındaki tek katman [41]



Şekil 8. UKSH'daki dört etkileşimli katman yapısı [41]

UKSH ağlarının tekrarlayan sinir ağlarından en belirgin farklılıkları olarak hafıza hücreleri ön plana çıkmaktadır. Hafıza hücrelerinde gizli durumunun hesabı yapılır. Hafıza hücreleri hangi bilgilerin saklanacağına ya da silineceğine karar verir. Daha sonraki aşamalara geçildiğinde önceki duruma ait veriler ile giriş verileri birleştirilir. Bu

yaklaşım uzun vadeli bağımlılıkların ortadan kaldırılması ve veri dizilerinin devam ettirilmesini mümkün kılar [40].

UKSH'nın klasik ileri beslemeli sinir ağlarından farkı olarak geri bildirim bağlantıları bulunur. Bu özelliği bu ağ yapılarının uzun vadeli bağımlılıkları öğrenebilmesini sağlar. UKSH ağ yapılarında temelde yer alan kapılar şu şekildedir;

- ✓ Giriş Kapısı
- ✓ Unutma Kapısı
- ✓ Çıkış Kapısı

UKSH ağ yapısında ilk işlem aşaması, unutma kapısı ile yapılır ve burada hücre durumunda olan bilginin, saklanıp saklanmayacağına karar verilir. Bu aşamada girdi ile beslenen hücrelere, 0 ile 1 arasında bir sayı verilir. Hücre, eğer 1 çıkarsa tamamen bunu tut anlamında ve 0 çıkarsa da bunu tutmana gerek yok anlamındadır.

UKSH ağ mimarisinde unutma kapısında yapılan işlemden sonra, giriş kapısı ile yapılan hangi bilginin hücrede tutulacağını belirlemenin aşamasına geçilir. Giriş kapısı sigmoid sinir ağı ve tanjant katmanı olmak üzere olmak üzere iki katmandan oluşur. Hangi değerlerin güncelleneceğine sigmoid sinir katmanında karar verilir. Tanjant katmanda ise saklanması gereken yeni vektör oluşturulur. Bu işlemler sonucunda iki farklı değer birleşerek giriş katmanında saklanacak olan yeni değeri oluşturur. UKSH ağ mimarisinin çıkış katmanı da tanjant ve sigmoid katmanı olmak üzere iki katmandan oluşmaktadır. Tanjant katmanında saklanan bilginin ne kadarının kullanılacağı hesaplanmaktadır. Tanjant katmanında yer alan sonuçların değerleri -1 ile 1 arasında değişmektedir. Sigmoid katmanında ise yeni gelen bilginin kullanılıp kullanılmayacağına karar verilir. Bu katmanda sonuç çıktısı 0 ile 1 arasında olur. Sigmoid ve tanjant katmanlarından gelen değerlerin çarpılmasıyla çıktı değeri oluşur ve çıkış katmanı sonucudur. [41].

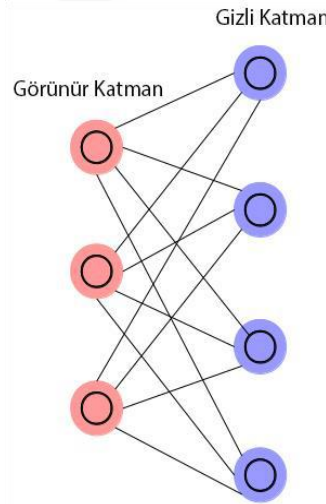
Zaman içerisinde geleneksel UKSH ağ mimarilerine eklemeler yapılarak geliştirilmiştir. Hücrede saklanan verinin, giriş ve unutma kapısına da etki etmesini sağlamak amacıyla gözetleme deliği bağlantıları eklenmiştir. Bu sayede verinin saklanıp saklanmayacağına, değiştirilip değiştirilmeyeceğine ve çıkış bilgisinde, yeni gelen bilginin olup olmayacağına etki edilebilecektir. UKSH ağ mimarilerinde yapılan bir diğer iyileştirmeye unutma ve giriş kapıları tek bir kapı üzerinde birleştirilmiş ve güncelleme

kapısı adı verilen Gated Recurrent Unit (GRU) oluşturulmuştur. Elde edilen bu yeni model, geleneksel UKSH modelinden, daha basit ve işlevseldir.

2.3.1.4. Kısıtlı Boltzmann Makinesi (Restricted Boltzmann Machine)

Kısıtlı Boltzmann Makinesi, giriş veri seti üzerinde olasılıksal dağılımları öğrenebilen ve sınıflandırma, regresyon ve özellik öğrenimi işlemlerini yapabilen bir sinir ağıdır. Girdi (görünür) ve gizli katman olmak üzere iki katmanlı bir yapıya sahiptir.

Her bir düğümün bir nöron olduğu Kısıtlı Boltzmann Makinesinde, hesaplamalar bu düğümlerde yapılır. Şekil 9’da gösterildiği şekilde, görünür katmandaki her düğüm gizli katmanda yer alan bir başka düğümler (nöron) ile bağlanır. Aynı katmanda yer alan düğümler birbirleri ile bağlanmazlar bu durum katmanlar arası iletişim olmadığını gösterir. Aynı katmandaki düğümler arası iletişimin kısıtlı olmasından dolayı kısıtlı boltzman makineleri olarak adlandırılırlar. Kısıtlı Boltzmann Makinesinde görünür katmanda hesaplanan girdilerin bir sonraki düğüme iletilip ileilmeyeceği rastgele olarak belirlenir [40].

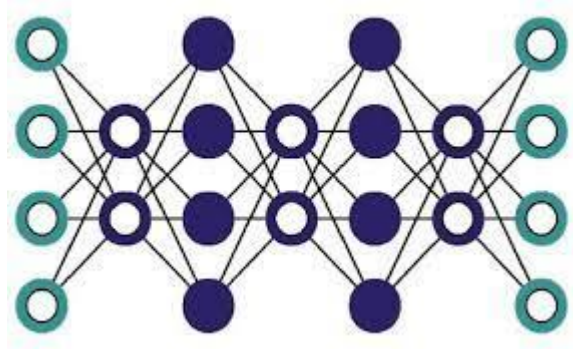


Şekil 9. Kısıtlı Boltzmann Makinesi [40]

2.3.1.5. Derin İnanç Ağı (DBN-Deep Belief Network)

Derin inanç ağları, Kısıtlı Boltzman Makinelerinin yığınlarıyla oluşturulur. Kısıtlı Boltzman Makinelerinin eğitilmesi ve ardından öğrenilmesiyle gerçekleşir. Derin inanç

ağı modeli girdi verisi ile gizli katman arasındaki olasılıksal dağılımdır. Şekil 10’da derin inanç ağı yapısı yer almaktadır.



Şekil 10. Derin İnanç Ağı yapısı [40]

2.3.1.6. Derin Oto-kodlayıcılar (Auto Encoder)

Derin oto-kodlayıcılar denetimsiz öğrenme için kullanılan bir özel YSA’dır. Bu model girdi verisini çıktı katmanında tekrar oluşturur. İleri beslemeli bir sinir ağı olup, giriş veri setinin sıkıştırılmasıyla öğrenme boyutu artırılmaya çalışılır.



Şekil 11. Derin Oto-kodlayıcı ağ yapısı [42]

Şekil 11’de gösterildiği gibi derin oto-kodlayıcıların ağ mimarisinde girdi katmanı, gizli katman ve çıktı katmanı olmak üzere üç katman bulunur. Gizli katmandaki nöron sayısı girdi ve çıktı katmanına göre değişkenlik göstermektedir. Giriş ve çıkış katmanında yer alan nöronların sayısı gizli katmandan fazla olduğu zaman veri kümesi sıkıştırılır. Girdi verisinin sıkıştırılması daha az verinin ağ içerisinde yer almasını ve en iyi özelliklerin ağ tarafından öğrenilmesini sağlayacağından çalışma performansını artacaktır.

2.3.2. DÖ Mimarilerinde Yer Alan Katmanlar

DÖ yapıları çok katmanlı bir yapıya sahip olduğundan her katmanda ayrı bir işlev yerine getirilir ve bir sonraki katmana veriler aktarılır. DÖ mimarilerinde yer alan katmanlar ve işlevleri aşağıda ayrı ayrı yer almaktadır.

2.3.2.1. Konvolüsyon Katmanı

Konvolüsyon katmanında filtreleme işlemi gerçekleştirilerek tanımlamak istenilen özellikler ağa daha iyi bir şekilde eklenir. Hangi filtrenin tercih edileceği ağın başarısını etkileyeceğinden filtre seçimi de önemlidir.

2.3.2.2. Aktivasyon (Relu) Katmanı

Aktivasyon katmanı, konvolüsyon katmanından sonra gelir ve aktivasyon fonksiyonunun gerçekleştiği katmandır.

2.3.2.3. Havuzlama (MaxPool) Katmanı

Havuzlama katmanı genellikle aktivasyon katmanı sonrasında yer alır ve verilerin indirgenmesi ile giriş boyutu azaltılır. Veri miktarındaki azalma ağın hızını artıracaktır.

2.3.2.4. Tam Bağlı (Full-Connected) Katman

Bu katmanda nöronlar tam bağlı haldedir ve her nöron kendinden sonraki nöronla bağlanır. Bu yüzden tam bağlı katman olarak bilinir.

2.3.2.5. Dropout Katmanı

DÖ ağ yapısının aşırı öğrenme ve ezber yapması bağlantılarını ortadan kaldırma işlemini yaparak ağın performansını artırır.

2.3.2.6. Normalizasyon (Ölçeklendirme) Katmanı

Normalizasyon katmanı, diğer katmanlardan gelen verileri düzenli hale getirmesi ile ağın performansını artıracaktır.

2.3.2.7. Entropi (Softmax) Katmanı

Olasılıksal bir girdi verisinin sınıflandırmasının yapıldığı katmandır. Bu işlemi kendisinden önce gelen tam bağlı katmanın girdi verisi ile yapar.

Son zamanlarda DÖ algoritmalarını, birçok farklı bilim dalında yapılan çalışmalarda görebiliriz. Bunun yanında DÖ günlük yaşamımızdaki birçok teknolojik unsurun içerisinde kullanılmaktadır. DÖ algoritmaları ile yapılan çalışmalarda yüksek başarımların elde edilmesi bu alana olan eğilimi artırmıştır [43]. Özellikle sınıflandırma, görüntü işleme, tanıma ve gelecek tahmini için kullanılmaktadır.

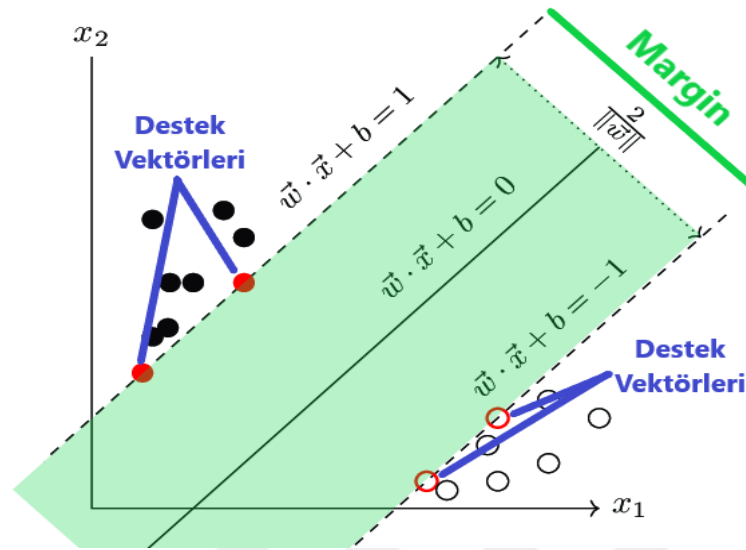
2.4. Destek Vektör Makineleri (DVM)

DVM, istatistiksel öğrenme teorisi ile yapısal riski en aza indirmeyi amaçlayan, özellikle sınıflandırma ve regresyon problemlerinin çözümü amacıyla Vapnik tarafından ileri sürülmüş bir makine öğrenme yöntemidir [44].

DVM yöntemi ilk olarak ortaya çıktığı zaman sensör verilerini yorumlamanın bir yolu olarak geliştirilmiştir ve iki katmanlı bir yapı sahiptir. Bu yapının ilk katmanı, giriş değişken serilerine sahip destek vektörleri üzerindeki ağırlıksız doğrusal olmayan bir çekirdektir. İkinci katman, çekirdek çıktılarının ağırlıklı bir toplamının yer aldığı katmandır. DVM model yapıları, uygun çekirdek filtreleri ve destek vektörleri belirlendikten sonra YSA yöntemlerine göre daha verimli olabilir [45].

DVM ile sınıflandırma veya regresyon problemleri, karesel programlama problemine dönüştürülerek daha hızlı çözümleme başarısı elde edilebilir. Bu özellik, DVM'nin diğer yöntemlere göre üstünlüklerinden olarak ifade edilebilir. DVM'lerin diğer yöntemlere göre bir diğer en önemli üstünlüğü olarak yüksek genelleme yapabilme yeteneğine de ön plana çıkmaktadır.

DVM çoğunlukla sınıflandırma problemlerinde kullanılmakta olup denetimli bir makine öğrenmesi algoritmasıdır. Sınıflandırma problemlerinin çözümü için iki sınıfı ayıracak doğru çizilir. Yani DVM, düzlem üzerinde bulunan veri setleri için bir doğru çizer. Bu doğrunun bu iki sınıfa ait noktalardan en fazla uzaklıkta olmasını amaçlar. Küçük ve orta ölçekteki karmaşık veri setleri için kullanılabilir [46].



Şekil 12. DVM ile sınıflandırma [47]

Şekil 12’de, siyahlar ve beyazlar olarak ayrılmış iki farklı veri sınıfı vardır. Sınıflandırma problemlerindeki en önemli amacımız verilerin hangi sınıfta yer alacağını belirlemektir. Şekil 12’de görüleceği üzere sınıflandırma işlemini yapabilmek için iki sınıfı ayıran bir doğru çizilir. Bu doğrunun ± 1 ’lik kısmının arasında kalan bölgeye (yeşil renkli) ‘Margin’ adı verilir. Margin şeklinde isimlendirdiğimiz bu alanın çok fazla geniş olması sınıflandırmanın başarısını artıracaktır.

$$\hat{y} = \begin{cases} 0 & \text{if } \mathbf{w}^T \cdot \mathbf{x} + \mathbf{b} < 0, \\ 1 & \text{if } \mathbf{w}^T \cdot \mathbf{x} + \mathbf{b} \geq 0 \end{cases} \quad (2.4)$$

Eşitlik 2.4’de yer alan lineer fonksiyona ait formülü incelediğimizde;

- ✓ $\mathbf{x}$, girdi vektörünü,
- ✓ $\mathbf{w}$, ağırlık vektörünü ve
- ✓ $\mathbf{b}$, sapmayı ifade eder.

Şekil 12’de iki sınıfı ayıran doğrunun 0’dan küçük değerler için olan beyaz noktalara yaklaştığı, 0’dan büyük değerler için olan siyah noktalara yaklaştığı görülmektedir. Yeni bir değer için yapılacak sınıflandırmada çıkan sonuca göre beyaz noktalara mı yoksa siyah noktalara mı yakın olacağına karar verilir.

DVM, veri setinin doğrusal olarak ayrılıp ayrılamamasına göre; Doğrusal DVM ve Doğrusal Olmayan DVM olarak ikiye ayrılır. Doğrusal DVM’nin en önemli amacı veri

setinin bir hiper düzlem ile ayrılmasıdır. Bu ayırma hiper düzlemini farklı sınıflara ait destek vektörleri arasındaki uzaklığı maksimize ederek sağlar. Doğrusal Olmayan DVM’lerde veri seti doğrusal bir fonksiyonla ayrılamaz. Doğrusal Olmayan DVM’lere veri setinin hiper bir düzlem ile ayrılamadığı gerçek yaşam problemleri örnek verilebilir. Bu problemlerde veri setinin doğrusal olarak ayrılması mümkün değildir. Sınıfları ayırma işlemini gerçekleştirmede ayırma eğrisinin tahmin edilmesi gerekir. Uygulamada eğrinin tahmin edilmesinin zor olduğu söylenebilir [48].

Veri setinin, birbirinden düz bir doğru (düzlem) ile ayrılamadığı durumlarda bu tip veri setlerinin başka bir yöntemle ayrılması gerekmektedir. DVM’lerde bu tip ayırma problemlerinin çözümü için ‘çekirdek’ kavramı kullanılmaktadır. DVM’lerin yüksek başarı göstermelerinde çekirdek yöntemlerinin kullanılmasının önemli bir rolü vardır. Bir düzlem ile ayrılmayan veri setlerindeki grupların birbirinden ayrılmasını sağlayan çekirdeğin çalışma prensibi, iç çarpım hesabını daha yüksek boyutlu bir uzaya taşıyarak (yüksek boyutlu uzaydan kastın 2 boyutluyu 3 boyuta, 3 boyutluyu 4 boyuta taşımak gibi) iki vektör arasında benzerliği araştırır. Çekirdek yöntemleri sayesinde bir düzlem ile ayrılamayan veri sınıflarının boyutunun artırılması bir düzlem vasıtası ile ayrılmasını sağlar [49].

Farklı tipte kullanılan çekirdek fonksiyonlarından bazıları şöyledir;

- ✓ Doğrusal,
- ✓ Gauss,
- ✓ Polinom,
- ✓ Anova,
- ✓ Laplace,
- ✓ Bessel,
- ✓ Sigmoid (Hiperbolik Tanjant) çekirdek fonksiyonlarıdır.

DVM’ler makine öğrenmesinin bir alt dalı olarak son yıllarda önemli aşama kaydetmişler ve makine öğrenmesi, istatistik ve sinir ağlarından çeşitli teknikleri kombine etmektedirler [50].

DVM’ler sınıflandırma problemlerinde oldukça başarılı sonuçlar verirler. DVM’lerin sınıflandırma stratejileri, sadece istatistiksel bir kriterden ziyade marj tabanlı bir

geometrik kriteri kullanmalarına bağlıdır. Başka bir ifadeyle, DVM'ler sınıflandırma görevini gerçekleştirmek için sınıfların istatistiksel dağılımlarının bir tahminine ihtiyaç duymazlar ve marj maksimizasyonu kavramından yararlanarak sınıflandırma modelini tanımlarlar. DVM'lerin birçok alanda kullanılmasına yönelik ciddi bir ilgi vardır. Bu ilgi şu nedenlerden kaynaklanmaktadır:

- ✓ İçsel etkililik, bu da beraberinde yüksek sınıflandırma doğruluklarını ve çok iyi genelleme yeteneklerini sağlar,
- ✓ Mimari tasarım için gerekli olan daha az çalışmayı, yani birkaç kontrol parametresi içerirler,
- ✓ Doğrusal kısıtlı kuadratik programlama yöntemlerine göre öğrenme problemini çözme olasılığıdır (bilimsel literatürde yoğun olarak çalışılan).

Bunlar yanında DVM'lerin önemli bir dezavantajı da teorik bir bakış açısından, ikili sınıflandırma problemlerini çözmek için geliştirilmiş olmaları şeklinde ortaya çıkmaktadır. Bu dezavantaj, hiperspektral sensörlerden elde edilen verilerle (Havacılık ve uzay çalışmaları kaynaklı sensörler ile elde edilen uzaktan algılama görüntülerine ait veri seti) çalışma yapılırken ortaya çıkmaktadır [51].

DVM'ler coğrafi amaçlı çevresel veri analizi ve bu analizlerin modelleme çalışmalarında sıkça kullanılmaktadır. Son dönemde sıkça kullanılan ve önemli faydaları olan uydu görüntüleri üzerinden haritalama çalışmalarında DVM algoritmaları da sıkça kullanılmaktadır. Son yıllarda, tematik haritalamanın öneminin artışı ve birçok alanda kullanıldığı bilinmektedir. Uydu görüntülerinin sınıflandırılmasıyla elde edilen tematik haritalama çalışmalarında DVM'lerden yararlanılmaktadır.

DVM'nin birçok bilimsel araştırma alanında uygulanabilmesi yanında mühendislik problemleri içinde kullanım alanı vardır. Hidroloji ve su kalitesi alanında kullanılabilmekte ve önemli yararlar sağlamaktadır. DVM, yağış - akış kavramlarına yönelik uygun model yapıları geliştirilerek, taşkın tahmini amacıyla geleceğe yönelik tahminlerin geliştirilebilmesi için geçmişten öğrenilen bilgileri ve geçmiş olayları izleme potansiyeline sahiptir [52].

DVM, ilk ortaya çıktığında doğrusal vektörler yardımıyla ikili veri setini kolaylıkla sınıflara ayırabilmekteydi. Daha sonra geliştirilen DVM modelleri, çekirdek

fonksiyonların kullanılmasıyla iç içe geçmiş birden fazla sınıfı da kolayca ayırabilecek seviyeye ulaşılmıştır. DVM modellerindeki bu gelişme birçok alanda kullanılmasına yönelik tercih sebebidir. DVM yönteminin çok geniş alanda uygulanabilmesi mümkündür. Tanıma, sınıflandırma ve tahmin alanında birçok çalışma vardır.

DVM'nin kullanım alanlarının birkaçı şöyledir.

- ✓ Yüz algılama çalışmalarında, piksel değerleri ile sınıflandırma ve etiketleme yapılır. Daha sonra bu değerler eğitim verilerini oluşturur. Sınıflandırma, piksel parlaklığına göre sınırlar oluşturularak yapılır.
- ✓ Tıpta biyolojik diziler (gen dizilimi) arasında tanımlama yaparken sıklıkla kullanılabilir. Hastaların gen dizilimine göre veya diğer birçok farklı biyolojik parametrelere göre sınıflandırılması için kullanılmaktadır.
- ✓ Belgeler üzerindeki imzalar ile elle yazılmış karakterleri tanımak için de kullanılabilir.
- ✓ Ayrıca son yıllarda sıklıkla duyduğumuz ses ve nesne tanıma ile metin sınıflandırma alanlarında da DVM yöntemlerinin kullanım alanı vardır.

2.5. Gauss Süreç Regresyon (GSR)

Günümüzde artan teknolojik gelişmeler üretilen veri miktarlarının artmasını sağlamıştır. Artan veri miktarları da büyük veri analizi kavramının oluşmasını sağlamıştır. Büyük veri kavramının en önemli özelliği hacimsel veri büyüklüğüdür. Günümüzde geleneksel veri işleme yöntemleri ile analizi yapılamayan büyük veriler, donanım ve yazılım alanındaki gelişmeler sayesinde detaylı olarak incelenebilmektedir. Büyük verilerin analizinde, yapay zekâ teknikleri sıkça kullanılmaktadır.

Yapay zekânın bir alt dalı olan makine öğrenmesi, istatistiksel ve matematiksel yöntemlerin kullanıldığı ve veriler yardımıyla kestirmeler yapan algoritmalarla oluşan bir modelleme yöntemidir. Yapay zekâ tekniklerinin günümüzde en popüler alt dalı olan makine öğrenmesinin, 1980'li yıllarda temelleri atılmış olup birçok çalışma alanında uygulamaları vardır. Makine öğrenmesi ile çalışma yaparken yüksek kapasiteli belleklere ve yüksek hızlı grafik işlem birimine (GPU) sahip bilgisayarlara ihtiyaç vardır.

Büyük veri analiziyle, çalışma konusuna ilişkin bilgi sahibi olmak ve tahminde bulunmak amacıyla farklı makine öğrenmesi algoritmaları kullanılabilir. Bu algoritmalar,

kestirimde bulunmak için öğrenme yollarına göre denetimsiz öğrenme ve denetimli öğrenme olarak ikiye ayrılmaktadır.

Denetimsiz öğrenme, verilerdeki önceden bilinmeyen bağlantıları bulmak için kullanılan yöntemdir. Bu yöntemin amacı, veri içindeki benzer grupları bulmayı ve veriler hakkında daha fazla bilgi sahibi olmak için verilerdeki temel yapıyı veya dağılımı modellemektir.

Denetimli öğrenme süreci, eğitime verilerine ihtiyaç duyan bir makine öğrenme modelidir. Bu modelde eğitim verilerine ait algoritma, veri setinin önemli bir bölümünden oluşur ve öğrenme süreci yine veri setine ait test verisi ile denetlenir. Denetimli makine öğrenmesinin sahip olduğu öğrenme süreci tahmin problemlerinde başarısı artırır.

Denetimli makine öğrenmesinde sınıflandırma ve regresyon konuları önemli konu başlıkları olarak ön plana çıkmaktadır. Sınıflandırma denetimli bir öğrenme çeşidi olarak, nitel değişkenlerin modellenmesi ve tahmin edilmesini amaçlanır ve bunun için bu kategorilere gözlemler atar. Regresyon, iki ya da daha fazla nicel değişken arasındaki ilişkiyi tahmin için kullanılan denetimli bir öğrenme yöntemidir [53].

Regresyon kelimesi ilk olarak, Francis Galton tarafından ‘Family Likeness in Stature (Boy Bakımından Aile Benzerliği)’ isimli çalışmasında kullanılmıştır. 1805 yılında Adrien Marie Legendre tarafından ortaya konulan ‘En Küçük Kareler Yöntemi’ regresyon yönteminin ilk hali olarak görülebilir. Daha sonra aynı yöntemi 1809 yılında C.F. Gauss tarafından açıklamıştır. Regresyon analizi değişkenler arasındaki ilişkiyi incelemek ve modellemek için kullanılan istatistiksel bir yöntemdir. Bu yöntemde bağımlı değişken, bağımsız değişkenden etkilenen değişkendir.

Regresyon analizinde; bağımlı ve bağımsız değişken sayısına göre analiz yöntemi belirlenmektedir. Bir bağımlı ve bir bağımsız değişken olması durumunda basit regresyon analizi, bir bağımlı ve birden fazla bağımsız değişken olması durumunda çoklu regresyon analizi, birden fazla bağımlı değişkenin olması durumunda çok değişkenli regresyon analizi yöntemleri uygulanır. Değişkenler arasındaki ilişki doğrusal ise doğrusal regresyon analizi, değilse eğrisel regresyon analizi olarak adlandırılır [54].

Basit bir regresyon modelini inceleyecek olursak;

$$y = \beta_0 + \beta_1 X + \varepsilon \quad (2.5)$$

Eşitlik 2.5'deki denklem şeklinde ifade edilir.

- ✓ Y; Bağımlı (açıklanan) değişkendir.
- ✓ X; Bağımsız (açıklayıcı) değişkendir.
- ✓ β_0 ; Sabit değerdir. (X=0 olduğunda Y'nin aldığı değerdir)
- ✓ β_1 ; Regresyon katsayısıdır. (Bağımsız değişkendeki 1 birim değişime karşılık bağımlı değişkende meydana gelecek değişmeyi ifade eder)
- ✓ ε ; Rassal hata terimidir. (Bağımlı değişkenin belli bir hata içerdiği varsayılır. Bağımsız değişkende hata yoktur)

Gauss süreçleri kavramı, Gauss dağılımı (normal dağılım) kavramına dayandığı için Carl Friedrich Gauss'un adını almıştır. Gauss süreçleri, çok değişkenli normal dağılımların sonsuz boyutlu bir genellemesi olarak görülebilir. Gauss dağılımı doğal olaylara ilişkin uygulanan çalışmalarda oldukça uyumlu sonuçlar vermiştir. Dağılıma uyumlu, uygun olduğu için gauss dağılımı adıyla birlikte normal dağılım ismi de kullanılmaktadır.

Gauss süreçleri ile yapılan çalışmalarda normal dağılımın sahip olduğu özelliklerden yararlanılır ve istatistiksel modellemede önemli kazanımlar sağlanır. Gauss süreci ile modellenen rastgele bir süreçte niceliklerin dağılımları elde edilebilir. Gauss süresinde belirlenen zaman aralığının işleme ait ortalaması ile belirlenmiş daha küçük zaman dizisinde numune değerleri yardımıyla ortalama değer tahmin edilmesindeki hataya sahiptir.

Gauss sürecinde veri miktarının artması zayıf şekilde ölçeklenmeye neden olacak ve hesaplama süresini düşürecektir. Bu sayede model doğruluğunun arttığı çoklu yaklaşım metodları geliştirilmiştir.

GSR modelleri, değişken sayısının fazla olduğu rastgele dağılıma sahip parametrik olmayan çekirdek tabanlı bir olasılık modelidir.

Gauss süreçleri istatistiksel modelleme, çoklu hedef değerlere regresyon ve haritalamanın daha yüksek boyutlarda analiz edilmesinde kullanılır [19].

Gauss süreçleri, Bayes öğrenmesini çekirdek makineleriyle birleştirerek, regresyon için ilkeli ve olasılıklı bir yaklaşım sağlar [55].

Gauss süreci, gauss dağılımına sahip sonlu sayıda rastgele değişkenlerin toplamıdır. Eğitim verilerinden oluşan küme, $D = \{(x_i, f_i), i = 1, \dots, n\}$ olmak üzere, $f_i = f(x_i)$, x_i değerinin ilgili fonksiyondaki aldığı değere karşılık gelmektedir [56].

Gauss süreci;

$$f(x) \sim GP \left(m(x_i), k(x_i, x_j) \right) \quad (2.6)$$

Eşitlik 2.6 ile ifade edilebilir. Burada x , ortalamaya ve (kx_i, x_j) , kovaryans fonksiyonuna karşılık gelmektedir. Kovaryans matrisinin elemanları, rastgele değişkenler arasındaki benzerliği ifade eder ve bu benzerlikler üstel fonksiyon aracılığı ile belirlenebilmektedir [56].

$$k(x_i, x_j) = \exp \left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right) \quad (2.7)$$

Eşitlik 2.7'yi incelediğimiz zaman, σ parametresi çekirdek genişliği parametresidir.

Örneklem dışı X_* verisi için GSR yönteminin vereceği gömüleme fonksiyonu Eşitlik 2.8'de ifade edildiği şekilde olacaktır.

$$p(f_* | X_*, X, f) = N(f_* | \mu_*, \sigma_*) \quad (2.8)$$

Eşitlik 2.8'deki μ_* ve σ_* değerleri, Eşitlik 2.9 ve 2.10'da verilen formüllerle belirlenebilir.

$$\mu_* = \underline{k_*^T} [K(X, X) + \sigma^2 I]^{-1} \underline{y} \quad (2.9)$$

$$\sigma_x^2 = k(X_*, X_*) - k^T [K(X, X) + \sigma_n^2 I]^{-1} \underline{k_*} \quad (2.10)$$

Burada k_* değişkeni, X eğitim verileri ve X^* örneklem dışı verisi arasındaki kovaryans değerlerini içeren vektöre karşılık gelmektedir. σ_x^2 ise gürültü varyansına karşılık gelmektedir. σ_x^2 değişkeni de girdi ve çıktı verileri arasında oluşturulan modelin güven (confidence) ölçütüne karşılık gelmektedir [56].

Gauss süreçleri, karmaşık veri kaynaklarını modellemek için bilimlerde ve endüstride yaygın olarak kullanılan esnek, parametrik olmayan Bayes araçları sınıfıdır. Gauss süreç modellerini uygulamanın anahtarı, birçok programlama dilinde mevcut olan iyi geliştirilmiş açık kaynaklı yazılımın mevcudiyetidir. Gauss süreçleri, verileri modellemek için esnek, parametrik olmayan bir araç sağlayan stokastik süreçler ailesidir. En temel durumda, bir Gauss süreci, sonlu bir gözlemler kümesine dayalı olarak gizli bir işlevi modeller. Gauss süreci, çok değişkenli bir Gauss dağılımının sonsuz sayıdaki boyuta bir uzantısı olarak görülebilir; burada boyutların herhangi bir sonlu kombinasyonu, tamamen ortalama ve kovaryans fonksiyonları tarafından belirtilen çok değişkenli bir Gauss dağılımı ile sonuçlanacaktır. Çekirdek olarak da bilinen ortalama ve kovaryans fonksiyonunun seçimi, ilgilenilen gizli fonksiyon üzerinde düzgünlük varsayımları uygular ve ilgili girdi veri noktaları x arasındaki Öklid mesafesinin bir fonksiyonu olarak çıktı gözlemleri y arasındaki korelasyonu belirler. [57].

GSR, son yıllarda makine öğrenmesi alanında çokça adını duyduğumuz ve büyük ilgi gören bir yöntemdir. GSR, belirsizlik tahminlerindeki avantajları yanında, doğrusal olmayan, karmaşık sınıflandırma ve regresyon sorunlarını ele alma konusunda önemli faydalar sağlar. YSA ile kıyaslandığında, GSR pratikte anlama ve uygulama açısından daha kolaydır. Gauss süreci, rastgele bir süreç olup herhangi bir sonlu sayısı ortak Gauss dağılımına sahip rastgele değişkenler dizisi olarak tanımlanır [58].

Gauss süreçlerinin önemli bir gerçeği, onların ikinci dereceden istatistiklerle tamamen tanımlanabilmeleridir. Bu nedenle, bir Gauss sürecinin ortalama sifra sahip olduğu varsayılırsa, kovaryans fonksiyonunun tanımlanması sürecin davranışını tanımlar. Önemli olarak, bu fonksiyonun negatif olmayan kesinliği, Karhunen-Loève açılımını kullanarak spektral ayrışmasını sağlar. Kovaryans fonksiyonu aracılığıyla tanımlanabilecek temel yönler, sürecin durağanlığı, izotropisi, düzgünlüğü ve periyodikliğidir [55].

Gauss süreçleri için birçok kovaryans fonksiyonu vardır. Bunlar;

- ✓ Constant,
- ✓ Linear,
- ✓ SE (Kare üstel),
- ✓ Matern,
- ✓ RQ (Rasyonel ikinci dereceden),
- ✓ Periyodik vb. şeklindedir.

Gauss süreçleri, jeostatistikte uzaysal verileri modellemek için uzun süredir kullanılmaktadır. Makine öğrenmesine artan ilgiyle birlikte GSR ile yapılan çalışmalarda da artış olmuştur. Tahmin ve sınıflandırma konusunda birçok bilim alanında faydalanılmaktadır.

2.6. Materyal

Sakarya Havzası'nda yer alan gözlem istasyonlarına ait 1995 - 2014 yılları arasındaki Şubat, Nisan, Haziran, Ağustos ve Kasım aylarında DSİ tarafından ölçülmüş su kalitesi verileri ile ÇO değerinin yapay zekâ yöntemleri olan; YSA, DÖ, DVM ve GSR ile modellenmesi yapılmış ve oluşturulan modeller karşılaştırılmıştır. Modellerde kullanılan bağımsız değişkenler; pH, BOİ, T, Eİ, aylar ve ÇO değerleridir.

2.7. Model Performanslarını Karşılaştırma Kriterleri

ÇO konsantrasyonunun tahminine yönelik, YSA, DÖ, DVM ve GSR modellerinden elde edilen sonuçların karşılaştırılmasında; OMH, KOKH, R^2 ve NSE kriterleri kullanılmıştır.

OMH, regresyon ve zaman serisi problemlerinde sıkça kullanılmakta olup iki sürekli değişken arasındaki farka ait ölçüdür. OMH değeri 0'dan ∞ 'a kadar değişebilir. KOKH, tahmin hatalarının standart sapması olup değeri yine 0 ile ∞ arasında değişir. OMH ve KOKH değerlerinin sıfır olması modelin mükemmel sonuçlar elde ettiği yani hiç hata yapmadığı anlamına gelir. R^2 , deney ya da model çalışması ile elde edilen verilerin doğrusal bir eğriye ne kadar iyi uyduğunun ölçütüdür. Çok fazla veri noktası olması değerlerin güvenilirliğini artırır ve R^2 değerinin 1 olması doğrusal eğrinin kusursuz olduğu anlamına gelir. NSE hata kriteri değeri $-\infty$ ile 1 arasında değişmektedir. NSE değeri 1 bulunması halinde tahmin başarısı yüzde yüz olarak ifade edilir. Literatürde NSE değeri

0.3–0.5 arasında olması durumu için düşük tahmin yeteneği olduğu, 0.5–0.7 arasında olması durumu için kabul edilebilir başarı yeteneği olduğu, 0.7–0.9 arasında olması durumu için yüksek tahmin başarısı yeteneği olduğu, 0.9–1 arası durumun ise mükemmel tahmin başarısını gösterdiği ifade edilir. OMH, KOKH, R^2 ve NSE değerleri eşitlik 2.11, 2.12, 2.13 ve 2.14’de yer alan denklemler ile gösterilmiştir.

$$OMH = \frac{1}{n} \sum_{i=1}^n |\zeta O_{T,i} - \zeta O_{ölçülen,i}| \quad (2.11)$$

$$KOKH = \frac{1}{n} \sum_{i=1}^n \sqrt{(\zeta O_{T,i} - \zeta O_{ölçülen,i})^2} \quad (2.12)$$

$$R^2 = 1 - \sum_i^n \frac{(R_{i\ gerçek} - R_{i\ tahmin})^2}{(R_{i\ gerçek} - R_{ort})^2} \quad (2.13)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (\zeta O_{ölçülen,i} - \zeta O_{T,i})^2}{\sum_{i=1}^n (\zeta O_{ölçülen,i} - \overline{\zeta O_{ölçülen}})^2} \quad (2.14)$$

Denklemlerde yer alan;

- ✓ ζO_T , modeller ile tahmin edilen çözünmüş oksijen değeri,
- ✓ $\zeta O_{ölçülen}$, DSI tarafından ölçülen çözünmüş oksijen değeri,
- ✓ $\overline{\zeta O_{ölçülen}}$, DSI tarafından gözlem noktalarından ölçülen çözünmüş oksijen değerlerinin ortalaması,
- ✓ n, serinin uzunluğunu gösterir.

3. BÖLÜM

BULGULAR

Yapay zekâ yöntemlerinin hidroloji biliminin diğer alanlarıyla birlikte su kalitesi parametreleri için de kullanım alanı vardır ve sağladığı kolaylıklar nedeniyle önemi sürekli artmaktadır. Bu tez kapsamında, ülkemizin 25 nehir havzasından biri olan Sakarya Havzası'nda yer alan gözlem istasyonlarına ait 1995 – 2014 yılları arasındaki Şubat, Nisan, Haziran, Ağustos ve Kasım aylarında ölçülmüş su kalitesi verileri ile ÇO değerinin YSA, DÖ, DVM ve GSR ile modellemesi yapılmıştır. Model giriş verileri; T, Eİ, pH, BOİ ve aylar, çıkış verisi ÇO olup MATLAB programı yardımıyla veriler işlenmiştir. 1995 – 2014 (20 yıl) yılları arasındaki toplam verisi sayısı 1388'dir. Bu verilerin, %80'ni (1107 adet) eğitmede, %20'si (281 adet) test verisi olarak kullanılmıştır. Tablo 4'de eğitme ve test değerlerine ait verilerin temel istatistiksel özellikleri verilmiştir.

Tablo 4. Eğitme ve test değerlerine ait verilerin temel istatistiksel özellikleri

<i>Eğitme Verisi</i>	<i>X_{ort}</i>	<i>SS</i>	<i>C_v</i>	<i>Min</i>	<i>Mak</i>	<i>C_{sx}</i>	<i>C_k</i>
T (°C)	14.524	5.698	39.23	0.00	29.5	0.11	-0.75
pH	7.778	0.346	4.45	4.16	9.4	-1.09	11.22
Eİ (mikromhos/cm)	836.90	425.30	50.82	193	6680	3.58	39.88
BOİ (mg/L)	22.37	57.12	255.32	0.27	626	6.48	50.46
ÇO (mg O₂/l)	7.108	3.179	44.73	0.00	14	-0.31	-0.81
<i>Test Verisi</i>	<i>X_{ort}</i>	<i>SS</i>	<i>C_v</i>	<i>Min</i>	<i>Mak</i>	<i>C_{sx}</i>	<i>C_k</i>
T (°C)	15.083	6.202	41.12	2	30	0.08	-0.85
pH	7.8657	0.4745	6.03	6.2	9.3	-0.01	0.74
Eİ (mikromhos/cm)	841.8	387	45.98	212	2670	0.98	2.2
BOİ (mg/L)	9.899	16.264	164.3	0.6	160	4.96	34.49
ÇO (mg O₂/l)	7.79	2.709	34.77	0.4	13.5	-0.52	-0.23

Tablo 4’de yer alan istatistiksel parametreler;

- ✓ X_{ort} , istatistik bilim dalında ve veri analizinde kullanılan bir veri dizisinin orta konumunu yani merkezi eğilim ölçüsü ifade eder.
- ✓ SS (Standart sapma), istatistik ve olasılık gibi farklı alanlarda kullanılan ve değişken veri değerlerine ait yayılımı özetlemek için kullanılan bir ölçüttür.
- ✓ C_v (Varyasyon katsayısı), bir olasılık dağılımı için normalize edilmiş istatistiksel yayılma ölçüsüdür.
- ✓ C_{sx} (Çarpıklık (Skewness) katsayısı), dağılımın ortalama etrafındaki simetriden ne kadar saptığını gösterir
- ✓ C_k (Basıklık (Kurtosis) katsayısı), verilerin tepe noktalarının durumunun normalden sapması (basık veya sivri olması) hakkında bilgi verir.

Tablo 4’de yer alan parametreler incelendiği zaman, pH parametresinin C_v değeri diğer parametrelerden daha küçük değere sahip olması nedeniyle en az değişken özelliğe sahiptir. C_{sx} katsayısının sıfırdan küçük olması durumunun sola çarpık, sıfırdan büyük olması durumunun sağa çarpık olduğundan hareketle ÇO ve pH değişkenlerinin hem eğitme hem de test verilerinin C_{sx} katsayıları sıfırdan küçük olduğu için sola çarpık dağılıma sahip oldukları, diğer değişkenlerin ise C_{sx} katsayıları sıfırdan büyük oldukları için sağa çarpık dağılıma sahip oldukları söylenebilir. BOİ, Eİ ve pH verilerinin değerleri C_k katsayısı yani basıklık katsayısının sıfırdan büyük olması sebebiyle olasılık yoğunluk dağılımında sivri dağılım özelliği göstermektedir. Diğer değişkenler ise C_k katsayısının sıfırdan küçük olması sebebiyle basık dağılım özelliği göstermektedir. C_{sx} ve C_k katsayılarının sıfır olması durumuna normal dağılım denir. Tablo 4’deki bu değerlerin sıfırdan farklı olmasından dolayı bu çalışmada kullanılan giriş-çıkış verileri normal dağılım özelliği göstermemektedir.

ÇO konsantrasyonu özellikle sucul yaşam için önemli bir su kalitesi parametresidir. Su kalitesi parametrelerinin birçoğu birbirlerini etkilemekte ve birbirlerinden etkilenmektedir. ÇO konsantrasyonda meydana gelecek değişiklikte T parametresi doğrudan etkili olmaktadır. ÇO, su içinde çözünmüş halde bulunan oksijen konsantrasyonudur. Sabit sıcaklıkta, sıvı içinde çözünen gaz miktarının doğrudan basınç miktarına bağlı olmasına Henry Kanunu denir. Bundan dolayı ÇO değeri doğrudan doğruya kısmi basınçla ilgilidir. Yüksek basınç altında, çözünen oksijen miktarı

artacaktır. Basınç azaldığında, azaltılma oranı kadar gaz çıkışı olur. Aylık olarak ölçümü yapılmış parametrelerden ÇO konsantrasyonunu etkileyecek T, Eİ, pH ve BOİ değerlerinden yararlanılmıştır.

ÇO konsantrasyonunun tahminine yönelik model çalışmalarında giriş verileri; aylık olarak ölçülmüş T, Eİ, pH, BOİ değerleridir. Modellemeye başlamadan önce veriler rastgele eğitim ve test gruplarına ayrıldı. Eğitim veri seti, toplam verinin %80'ini içerirken, test veri seti %20'sini içerir. Bu tahmin çalışması için 4 farklı yöntem kullanılmış ve model performansları karşılaştırılmıştır. Bu modellere ait çalışmalar aşağıda verilmiştir.

3.1. YSA Model Çalışması

ÇO konsantrasyon değerini, YSA modeli ile en iyi YSA mimari yapıyı belirlemek için deneme-yanılma yapılmış olup YSA en iyi mimari yapılar Tablo 5'de verilmiştir. Tablo 5'de farklı kombinasyonlara ait en iyi YSA mimari yapıları görülmektedir. Bu farklı kombinasyonların tamamında çok katmanlı YSA modellerinden yararlanılmıştır. YSA model mimarilerinin giriş ve çıkış katmanındaki aktivasyon fonksiyonu olarak logaritmik sigmoid transfer fonksiyonu (logsig) ve tanjant sigmoid transfer fonksiyonu (tansig) deneme-yanılma sonucunda belirlenmiştir. Modeller, 1 gizli katmandan oluşmaktadır. Nöron sayısı 1-10 aralığında ve iterasyon sayısı 1-100 aralığında değişmektedir. Öğrenme algoritması olarak hızlı sonuç vermesinden dolayı Levenberg-Marquardt optimizasyon algoritması kullanılmıştır.

YSA modelinde farklı kombinasyon denemelerinin ilk adımı olarak bağımsız değişkenlerimiz olan T, BOİ, Eİ, pH ve aylar modele ayrı ayrı 1 girişli veri olarak tanımlanmıştır. 1 veri girişli analizlerin determinasyon (R^2) katsayılarına göre değerlendirilmesi sonucunda, 5 değişken içerisinde BOİ değişkeni diğer değişkenlere göre daha iyi sonuç vermiştir. Tablo 5'de gösterildiği gibi BOİ değerinin giriş verisi olarak tanımlandığı modelde giriş ve çıkış katmandaki aktivasyon fonksiyonu tanjant sigmoid transfer fonksiyonu (tansig), gizli katmandaki nöron sayısı 2 ve iterasyon sayısı 94'dür. R^2 değeri 0.493 olup, bu değer 1'e yakın olmamasına rağmen diğer değişkenlerin R^2 değerlerine göre daha iyidir. Bir sonraki aşamada, 1 girişli analizlerde diğer değişkenlere göre iyi sonuç veren BOİ diğer değişkenlerle ayrı ayrı tanımlanarak 2 girişli modeller oluşturulmuş ve analiz sonuçları alınmıştır. Tablo 5'de de görüleceği üzere 2

girişli modeller; BOİ-T, BOİ-Aylar, BOİ-Eİ ve BOİ-pH'dır. Modellerde giriş ve çıkış katmanında farklı aktivasyon fonksiyonları, gizli katmanda farklı nöron sayıları ve farklı iterasyon denemeleri yapılmıştır. 2 girişli bu 4 ayrı modelin R^2 değerleri incelendiği zaman en iyi sonuç 0.620 ile BOİ-Aylar değişkenleri modeline ait olup, R^2 değeri 0.611 olan BOİ-T modeli ikinci iyi sonuçtur. BOİ-T modelinin R^2 değeri BOİ-Aylar modelinin R^2 değerine çok yakındır. BOİ-T modelinin, giriş ve çıkış katmanındaki aktivasyon fonksiyonu tansig, gizli katmandaki nöron sayısı 3 ve iterasyon sayısı 26'dır. BOİ-Aylar modelinin, giriş katmanındaki aktivasyon fonksiyonu logsig ve çıkış katmanındaki aktivasyon fonksiyonu tansig olup gizli katmandaki nöron sayısı 6 ve iterasyon sayısı 70'dir. Tablo 5'deki değerler incelendiği zaman, 2 girişli modeller ile elde edilen R^2 değerlerinin 1 girişli modeller ile elde edilen sonuçlara göre daha iyi olduğu görülmektedir. Bir sonraki aşamada giriş sayısı 3'e çıkarılarak R^2 değerinin 1'e yaklaşması hedeflenmiştir. Bunun için 3 girişli; BOİ-T-pH, BOİ-Aylar-T, BOİ-T-Aylar ve BOİ-T-Eİ modelleri farklı aktivasyon fonksiyonları, nöron sayıları ve iterasyon sayıları ile oluşturulmuştur. Modellere ait R^2 değerleri 0.64 ile 0.69 arasında değişmekte ve biraz daha 1 değerine yaklaşmıştır. 3 girişli 4 ayrı modelin R^2 değerlerinin, 2 girişli modellere göre daha iyi olduğu görülmüş olup en iyi sonuç BOİ-T-pH modelinden elde edilmiştir. 3 girişli 4 ayrı modelde giriş katmanındaki aktivasyon fonksiyonu logsig seçilmiştir. En iyi sonucun elde edildiği BOİ-T-pH modelinde çıkış katmanındaki aktivasyon fonksiyonu logsig olup, gizli katmandaki nöron sayısı 4 ve iterasyon sayısı 19'dur. BOİ-T-pH modelinde R^2 değeri 0.686 olarak bulunmuştur. Bir sonraki adımda giriş sayısı değişkenleri 4'e çıkarılmış ve R^2 değerindeki iyileşme gözlenmiştir. 4 girişli modeller, BOİ-T-pH-Eİ ve BOİ-Aylar-T-pH değişkenleri ile oluşturulmuştur. Bu modellere ait sonuçlar hata kriteri olan R^2 değerine göre değerlendirilmiş ve 4 girişli analizler neticesinde en iyi sonuçları, BOİ-T-pH-Eİ değişkenlerine ait model vermiştir. Bu modelde, giriş fonksiyonu olarak logaritmik sigmoid ve çıkış fonksiyonu olarak da tanjant sigmoid kullanılmıştır. Modelin gizli düğüm numarası 4 olup iterasyon sayısı da 55'tir.

Tablo 5. ÇO konsantrasyon değerinin YSA modeli ile tahmininde denemiş farklı kombinasyonlar

<i>Değişkenler</i>	<i>Gizli katmandaki aktivasyon fonksiyonu</i>	<i>Çıkış katmandaki aktivasyon fonksiyonu</i>	<i>Gizli katmandaki nöron sayısı</i>	<i>İterasyon</i>	<i>R²</i>
BOİ	Tansig	Tansig	2	94	0.493
Aylar	Logsig	Logsig	4	35	0.224
T	Logsig	Logsig	2	50	0.221
Eİ	Logsig	Logsig	2	23	0.105
pH	Tansig	Logsig	4	75	0.100
BOİ-T	Tansig	Tansig	3	26	0.611
BOİ-Aylar	Logsig	Tansig	6	70	0.620
BOİ-Eİ	Tansig	Logsig	2	28	0.536
BOİ-pH	Logsig	Tansig	4	27	0.472
BOİ-T-pH	Logsig	Logsig	4	19	0.686
BOİ-Aylar-T	Logsig	Tansig	7	78	0.648
BOİ-T-Aylar	Logsig	Logsig	6	16	0.641
BOİ-T-Eİ	Logsig	Logsig	4	19	0.639
BOİ-T-pH-Eİ	Logsig	Tansig	4	55	0.722

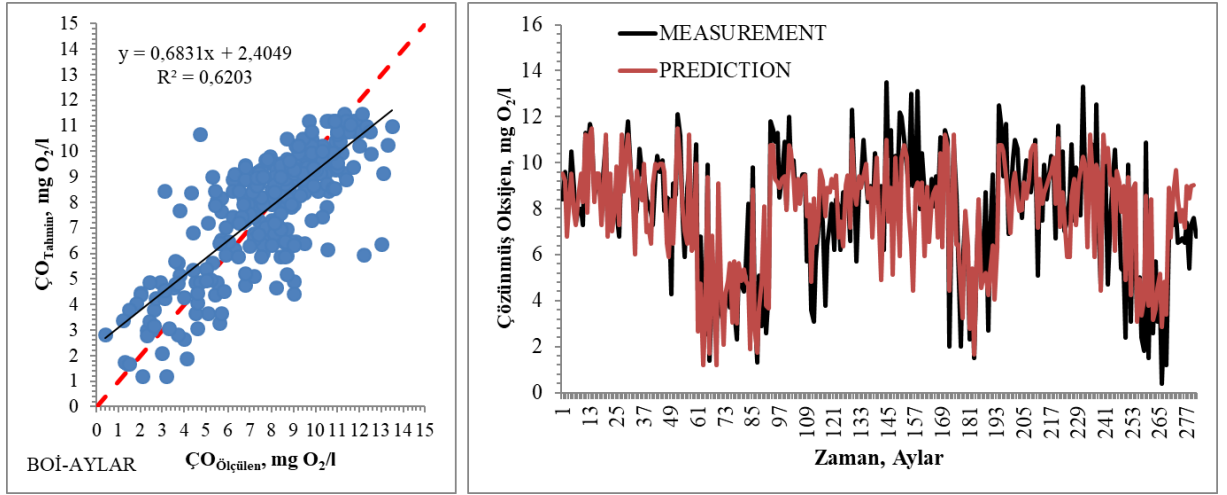
Sonuç olarak YSA model çalışmalarında Tablo 5’de de gösterildiği gibi deneme-yanılma yöntemiyle R² değerinin 1’e en fazla yaklaştığı model tespit edilmiştir. Daha sonra model performans kriterleri yardımıyla model doğruluğu değerlendirilmiştir. Model performanslarının değerlendirilmesinde OKH, OMH, OMBH ve NSE hata kriterlerinden yararlanılmıştır. Tablo 6’da ÇO parametresi tahmininde deneme-yanılma ile elde edilen modellerin performanslarının değerlendirmesine ait hata kriterleri yer almaktadır. OKH ve OMH hata değerlerinin sıfıra yaklaşması modellerden elde edilen ÇO değerinin, ölçülen ÇO değerlerine daha çok yaklaştığı anlamına gelmektedir. Tablo 6’da yer alan OKH değerlerini incelediğimizde BOİ-T modeli hariç diğer modellerde test verileri eğitme verilerinden daha düşük yani daha iyi sonuç vermiştir. En yüksek OKH değeri, 3.64 ile BOİ-T test modelinde, en düşük 2.09 ile BOİ-T-pH-Eİ test modelinde yer alır. Tablo 6’daki OMH değerlerini incelediğimizde yine OKH değerleri gibi, BOİ-T modeli hariç diğer modellerde test verileri eğitme verilerinden daha düşük yani daha iyi sonuç

vermiştir. En yüksek OMH değeri, 1.47 ile BOİ-T test modelinde, en düşük 1.10 ile BOİ-T-pH-Eİ test modelinde yer alır. NSE hata kriteri değerleri, 0.608 ile 0.748 arasında değişmektedir. NSE değerinin 1 olması yüzde yüz başarıyı temsil eder. 0.5–0.7 arası kabul edilebilir başarı ve 0.7–0.9 arası yüksek tahmin başarısı olarak ifade edilir. NSE hata değerlerini incelediğimizde, BOİ-Aylar ve BOİ-T modellerinin eğitme ve test verileri ile BOİ-T-pH modelinin test verileri 0.5–0.7 arasında olduğu için kabul edilebilir başarıyı, BOİ-T-pH modelinin eğitme verisi ile BOİ-T-pH-Eİ modelinin eğitme ve test verileri 0.7–0.9 arasında olduğu için yüksek tahmin başarısını temsil etmektedir.

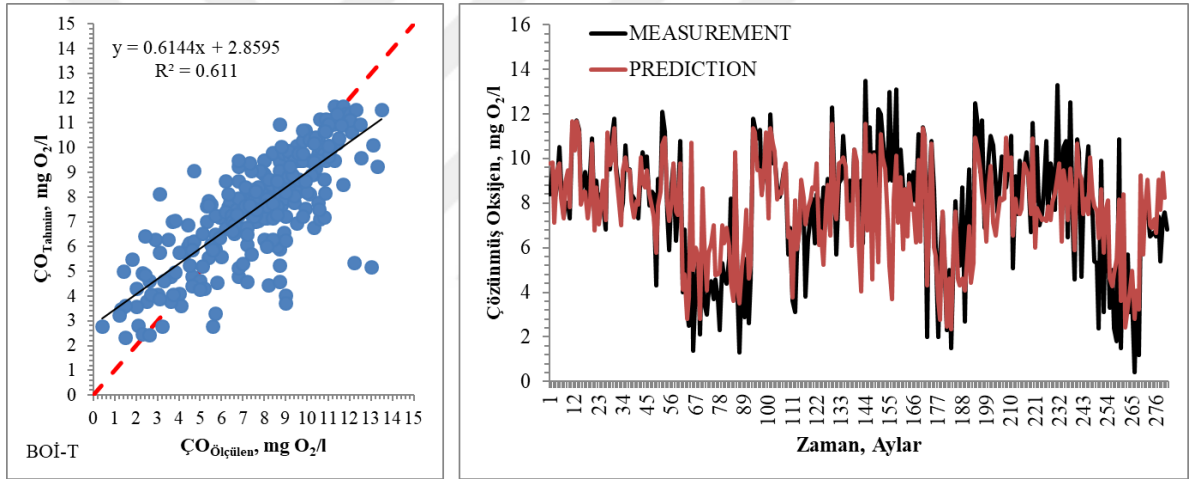
Tablo 6. ÇO parametresi tahmininde modellere ait eğitme ve test verilerinin hata kriterleri

<i>Model</i>	<i>Veri Grubu</i>	<i>OKH</i>	<i>OMH</i>	<i>OMBH</i>	<i>NSE</i>
BOİ-Aylar	Eğitme	3.30	1.40	35.32	0.672
	Test	2.87	1.30	23.41	0.613
BOİ-T	Eğitme	2.91	1.25	24.22	0.608
	Test	3.64	1.47	38.07	0.639
BOİ-T-pH	Eğitme	2.92	1.36	35.64	0.710
	Test	2.40	1.19	22.01	0.677
BOİ-T-pH-Eİ	Eğitme	2.54	1.25	30.55	0.748
	Test	2.09	1.10	20.03	0.719

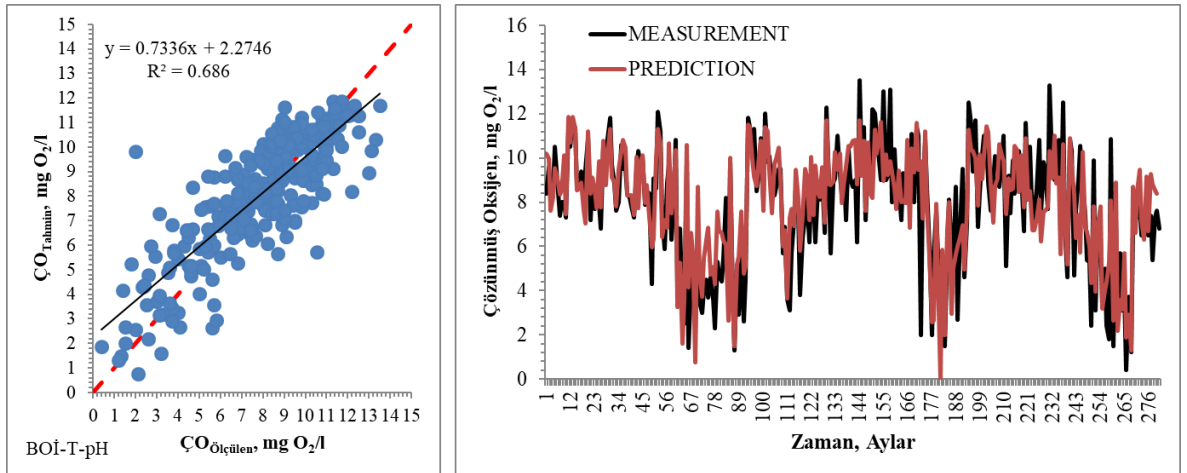
ÇO konsantrasyon değerinin tahminine yönelik çok katmanlı YSA modellerine ait saçılma ve zaman serisi grafikleri şekil 13, 14, 15 ve 16’da yer almaktadır. $\text{ÇO}_{\text{Tahmin}}$ – $\text{ÇO}_{\text{Ölçülen}}$ saçılma grafikleri incelendiğinde, şekil 13, 14 ve 15’e kıyasla şekil 16’ya ait grafiğin alttaki uç noktalar haricinde ÇO konsantrasyonu tahmini değerleri ile ÇO konsantrasyonu ölçülen değerleri arasında oldukça büyük uyumun olduğu görülmektedir. Ayrıca şekil 16’ya ait saçılma grafiğinde değerler fazla saçılma göstermemektedir.



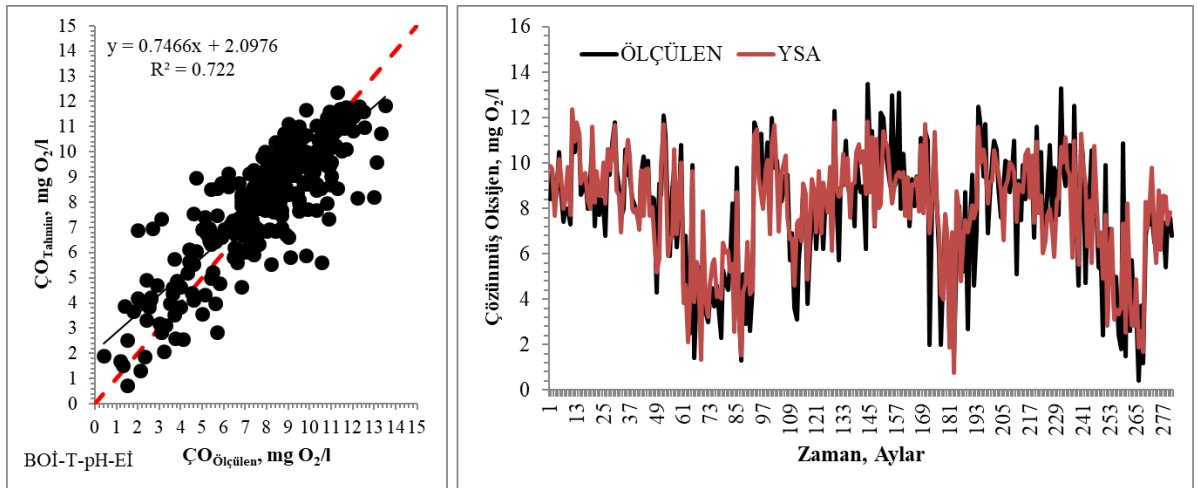
Şekil 13. BOİ-Aylar değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



Şekil 14. BOİ-T değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



Şekil 15. BOİ-T-pH değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



Şekil 16. BOİ-T-pH-Eİ değişkenlerine ait YSA modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri

ÇO konsantrasyonunun YSA modelleri ile tahmini amacıyla deneme-yanılma yapılarak birçok model oluşturulmuştur. 4 girişli BOİ-T-pH-Eİ modeli, Tablo 6’da yer alan model performans kriterleri incelendiğinde diğer modellere göre daha iyi sonuçlar vermiştir.

YSA model çalışmalarında en iyi sonucu, 4 veri girişli BOİ-T-pH-Eİ modeli vermiştir. Diğer makine öğrenmesi yöntemleri olan; DÖ, DVM ve GSR model çalışmaları da, 4 veri girişli BOİ-T-pH-Eİ ile oluşturulmuştur. YSA, DÖ, DVM ve GSR modelleri ile ÇO konsantrasyonunun tahminine yönelik elde edilen sonuçlar KOKH, OMH, R^2 ve NSE hata kriteri değerlerine göre değerlendirilmiştir.

3.2. DÖ Model Çalışması

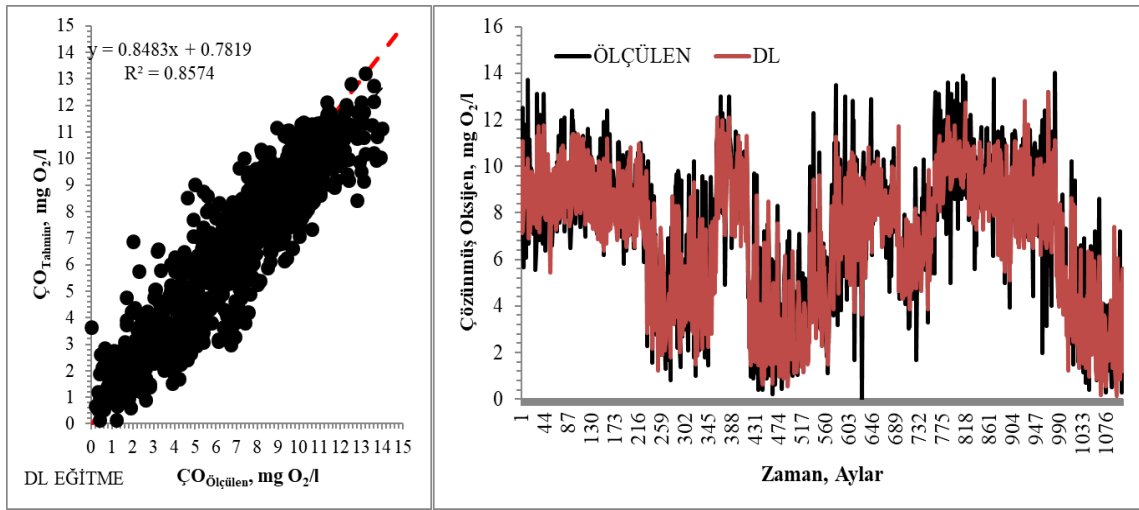
Çalışmanın ilk aşamasında, ÇO konsantrasyonunun YSA modelleri ile tahmini amacıyla deneme-yanılma yapılarak 4 girişli BOİ-T-pH-Eİ modeline ait sonuçların diğer modellere göre daha iyi sonuçlar verdiği görülmüş olup, çalışmanın ikinci aşamasında 4 girişli BOİ-T-pH-Eİ, DÖ modeli ile ÇO konsantrasyonunun tahmini amaçlanmıştır.

DÖ modellerinde “State Activation Function” olarak tanh ve “Gate Activation Function” olarak ise sigmoid kullanmıştır. Ayrıca, DÖ modelinde ağı eğitmek için stochastic gradient descent with momentum (SGDM) optimizasyon tekniği kullanılmıştır. DÖ modeli yapısında 1 gizli katmanda nöron sayısı ise 5 ile 20 arası ve iterasyon sayısı ise

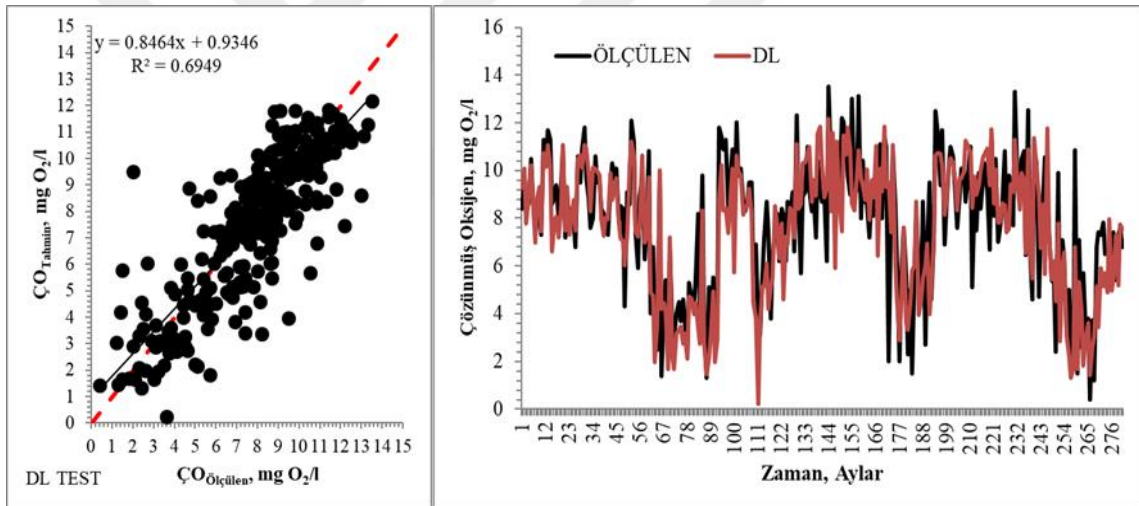
50 ile 200 denenmiştir. DÖ modelinde diğer parametreler ise ilk öğrenme oranı katsayısı 0.05, öğrenme oranını düşürme faktörü 0.2 ve öğrenme oranı düşme süresi 125 alınmıştır.

DÖ modellerinin geliştirme aşamasında gizli katmandaki nöron sayısı değiştirilerek çok sayıda deneme gerçekleştirilmiştir. DÖ modellemesi aşamasında yapılan denemeler sonucunda elde edilen sonuçlar Tablo 7'de verilmiştir. Tablo 7'de de görüleceği üzere, eğitme ve test verileri için sırasıyla, KOKH değeri 1.24 ile 1.61, OMH değeri 0.94 ile 1.18, R^2 değeri 0.857 ile 0.695 ve NSE değeri 0.849 ile 0.653'tür. Eğitme verilerine ait sonuçların hata kriterlerinin tamamı test verilerine ait sonuçların hata kriterlerinden daha iyi sonuçlar vermiştir. KOKH ve OMH değerleri, test veri grubunda eğitme veri grubuna göre sıfırdan daha fazla uzaklaşmıştır. R^2 değeri eğitme grubunda 0.857 ile test verilerine ait 0.695'e göre daha yüksek tahmin gücüne sahiptir. NSE hata değerlerini incelediğimizde; eğitme verisinde 0.849 ile yüksek tahmin başarısını, test verisinde 0.653 ile kabul edilebilir tahmin başarısını temsil etmektedir.

Şekil 17'de DÖ Eğitme modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri ile şekil 18'de DÖ Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri yer almaktadır. Şekil 17 ve 18'deki saçılma grafikleri incelendiğinde, eğitme grubuna ait değerlerin saçılma grafiği çıkış değişkenleri ile ölçülen çıkış değişkenleri arasında önemli oranda büyük bir uyum vardır. Test veri grubuna ait saçılma grafiği daha dağınık ve uyumsuzluğun varlığı söz konusudur.



Şekil 17. DÖ Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



Şekil 18. DÖ Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri

3.3. DVM Model Çalışması

Çalışmanın üçüncü aşamasında 4 girişli BOİ-T-pH-Eİ, DVM modeli ile ÇO konsantrasyonunun tahmini amaçlanmıştır.

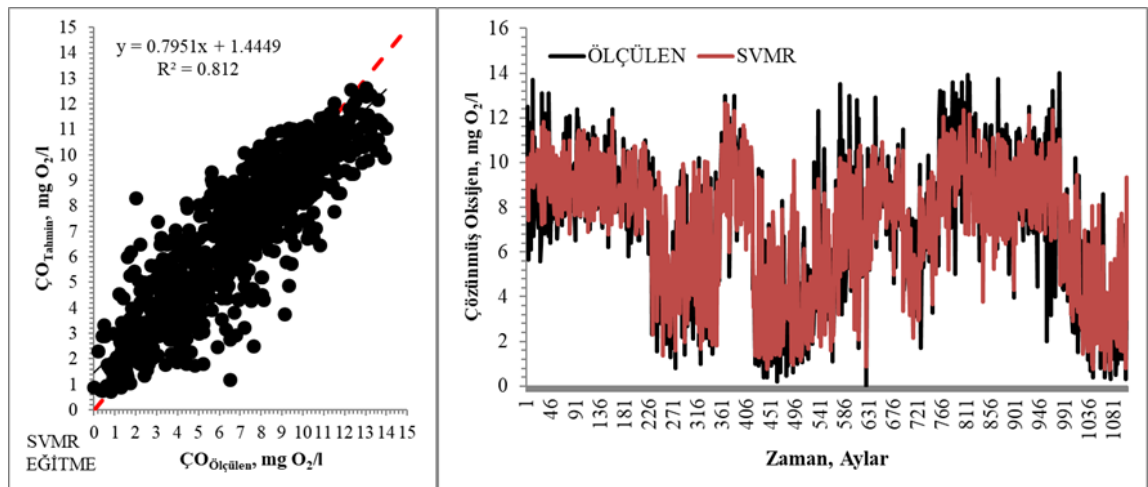
DVM model çalışmasında Kernel fonksiyonu kullanılarak tahmin modelleri geliştirilmiştir. Deneme-yanılması sonucunda diğer Kernel işlevlerine kıyasla daha iyi performans göstermesi nedeniyle bu çalışmada Radial basis function (RBF) doğrusal olmayan kernel işlevi belirlenmiştir. İki Lagrange çarpanı arasındaki fark olan $\alpha_i - \alpha_i^*$ ve bias (b) parametrelerinin en küçük değerleri Sequential minimal optimization

(SMO) ile elde edilmiştir. ÇO konsantrasyonunun modellenmesinde kullanılan bias (b) parametresi -0.1497'dir.

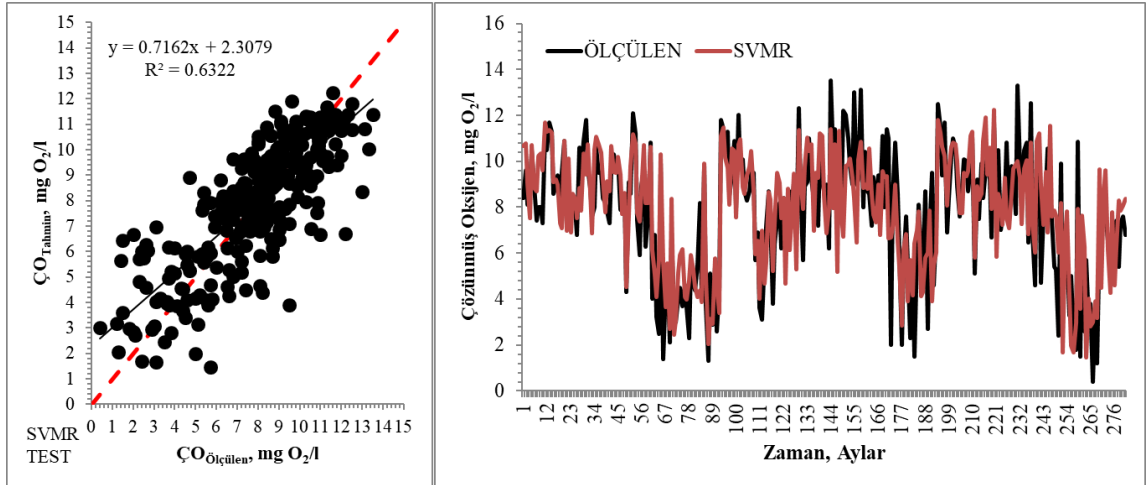
Tablo 7'de literatürde yaygın olarak kullanılan hata kriterleri kullanılmakta olup, DVM modeline ait eğitim ve test verileri için sırasıyla, KOKH değeri 1.38 ile 1.68, OMH değeri 1.01 ile 1.30, R^2 değeri 0.812 ile 0.632 ve NSE değeri 0.812 ile 0.620'dir. Eğitim verilerine ait sonuçların hata kriterlerinin tamamı test verilerine ait sonuçların hata kriterlerinden daha iyi sonuçlar vermiştir.

Eğitim verilerine ait R^2 ve NSE değerleri aynı değerler olduğu için DVM modelinin tahmin gücü yüksektir. Fakat Test verilerine ait R^2 ve NSE değerlerinin düşük olmasından dolayı DVM modelinin bağımlı ve bağımsız değişkenler arasındaki doğrusal olmayan ilişkilerin yorumlanmasındaki kapasitesinin iyi performansa sahip olduğunu söylemek oldukça zordur.

Şekil 19'da DVM Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri ile Şekil 20'de DVM Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri yer almaktadır. Şekil 19 ve 20'deki saçılma grafikleri incelendiğinde, eğitim grubuna ait değerlerin saçılma grafiği çıkış değişkenleri ile ölçülen çıkış değişkenleri arasında önemli oranda büyük bir uyum vardır. Test veri grubuna ait saçılma grafiği daha dağınık ve uyumsuzluğun varlığı söz konusudur.



Şekil 19. DVM Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



Şekil 20. DVM Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri

3.4. GSR Model Çalışması

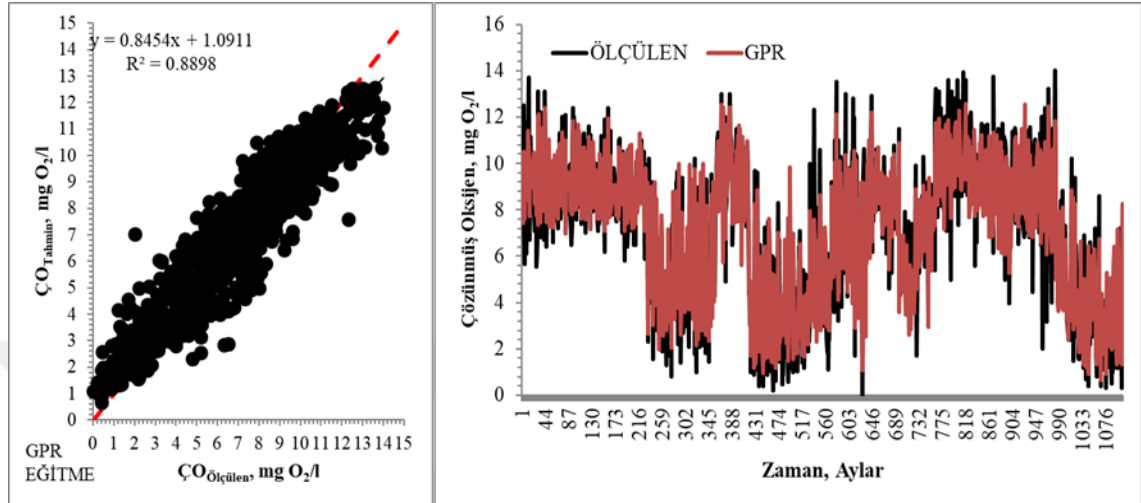
Çalışmanın dördüncü aşamasında 4 girişli BOİ-T-pH-Eİ, GSR modeli ile ÇO konsantrasyonunun tahmini amaçlanmıştır.

GSR yaklaşımında Kernel fonksiyonu kullanılarak tahmin modelleri geliştirilmiştir. Deneme–yanılması sonucunda diğer Kernel işlevlerine kıyasla daha iyi performans göstermesi nedeniyle bu çalışmada Squared Exponential Kernel (ardsquaredexponential) kovaryans (covariance) işlevi belirlenmiştir. GSR yaklaşımında kullanılan beta ve sigma parametrelerini belirlemek için “Subset of regressors approximation” ve “Fully independent conditional approximation” kullanılmıştır. Ayrıca bu çalışmada iterasyon sayısı 50 alınmıştır.

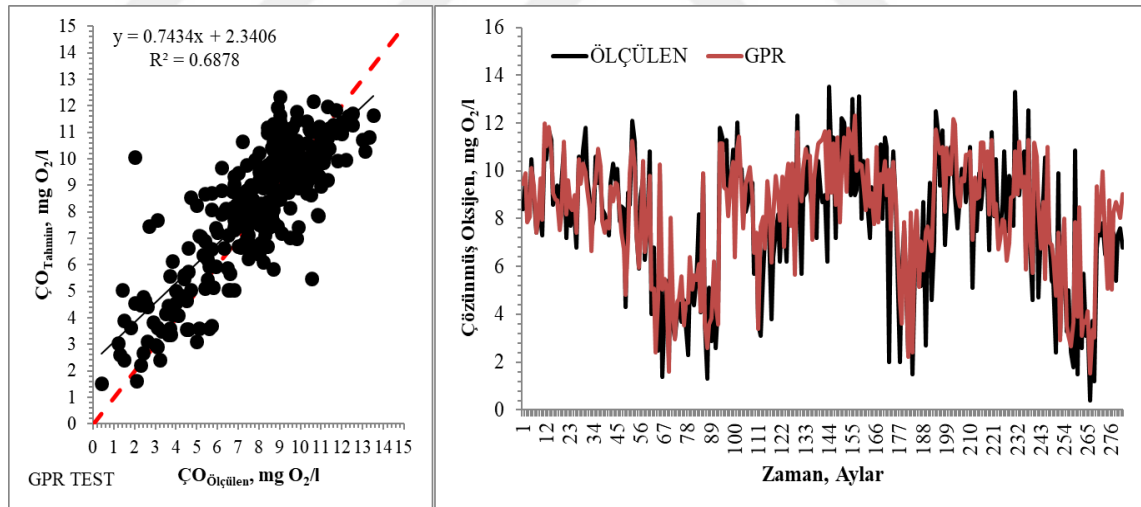
Tablo 7’de literatürde yaygın olarak kullanılan hata kriterleri kullanılmakta olup, GSR modeline ait eğitim ve test verileri için sırasıyla, KOKH değeri 1.06 ile 1.57, OMH değeri 0.82 ile 1.21, R^2 değeri 0.890 ile 0.688 ve NSE değeri 0.888 ile 0.668’dir. Eğitim verilerine ait sonuçların hata kriterlerinin tamamı test verilerine ait sonuçların hata kriterlerinden daha iyi sonuçlar vermiştir.

Şekil 21’de GSR Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri ile şekil 22’de GSR Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri yer almaktadır. Şekil 21 ve 22’deki saçılma grafikleri incelendiğinde, eğitim grubuna ait değerlerin saçılma grafiği

çıkış değişkenleri ile ölçülen çıkış değişkenleri arasında önemli oranda büyük bir uyum vardır. Test veri grubuna ait saçılma grafiği daha dağınık ve uyumsuzluğun varlığı söz konusudur.



Şekil 21. GSR Eğitim modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri



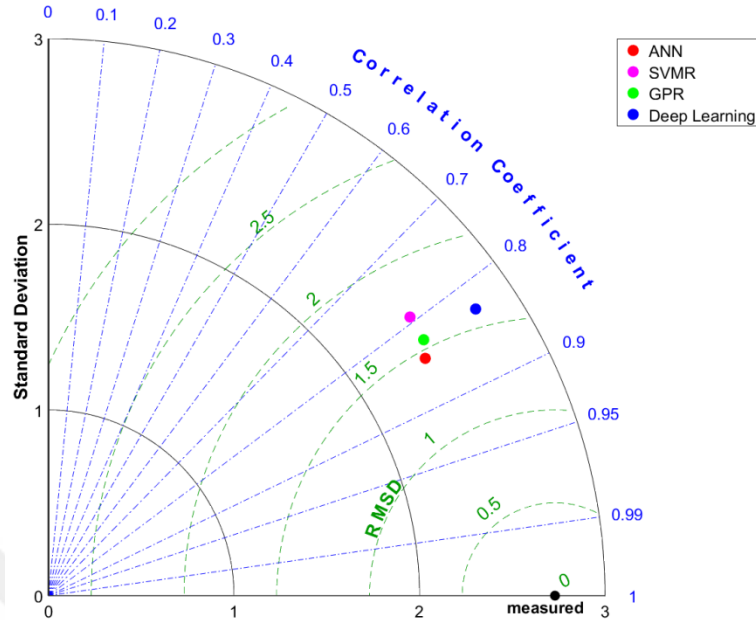
Şekil 22. GSR Test modeli ile tahmin edilen ÇO konsantrasyon değerinin saçılma ve zaman serisi grafikleri

Tablo 7'de görüleceği üzere performans kriterlerine göre 4 makine öğrenmesi modellerinden en iyi tahmini; DÖ, DVM ile GSR modellerine kıyasla YSA modeli vermiştir.

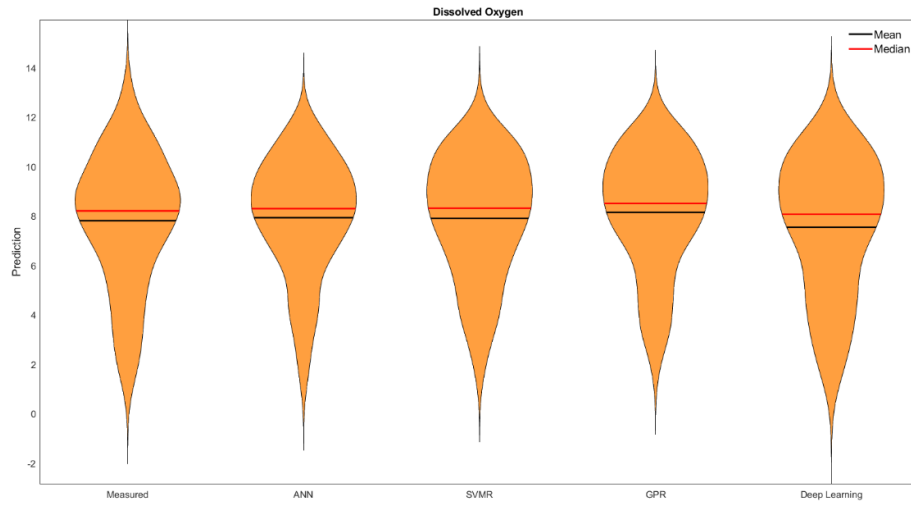
Tablo 7. YSA, DÖ, DVM ve GSR Modellere ait eğitme ve test aşamaları için hata kriterleri

Modeller	Eğitme Verisi				Test Verisi			
	KOKH	OMH	R ²	NSE	KOKH	OMH	R ²	NSE
YSA	1.59	1.25	0.748	0.748	1.45	1.10	0.719	0.719
DVM	1.38	1.01	0.812	0.812	1.68	1.30	0.632	0.620
GSR	1.06	0.82	0.890	0.888	1.57	1.21	0.688	0.668
DÖ	1.24	0.94	0.857	0.849	1.61	1.18	0.695	0.653

Bu çalışmada, YSA, DÖ, DVM ve GSR modellerinden elde sonuçları karşılaştırmak amacıyla klasik performans kriterlerine ek olarak Taylor diyagramından yararlanılmıştır. Taylor diyagramı, modeller ile elde edilen sonuçların gözlemler ile ne kadar yakın olduğunun grafik özeti olarak gösterilmesini sağlar. İki model arasındaki benzerliği, korelasyonları, merkezci ortalama karekökleri ve değişimlerinin büyüklükleri ile değerlendirmektedir [59]. Modellerin korelasyonu ile karesel ortalama hataların referans veri kümesi ile olan ortalamalar kullanılarak karşılaştırmaları yapılmıştır. Ayrıca, hata dağılımlarına dayalı, Taylor diyagramı ile model doğrulukları araştırılmaktadır. Bu çalışmada, Taylor diyagramı, oluşturulan modellerin karşılıklı karşılaştırmasını sağlamak için şekil 23’de sunulmuştur. Şekil 23’de görüleceği üzere 4 farklı makine öğrenmesi modelleri içerisinde ÇO değerini en iyi YSA modeli tahmin etmektedir. Ayrıca modellerin Violin diyagramları da çizilmiştir. Şekil 24’de her bir modele ait violin diyagramları da verilmiştir. Şekil 24’de görüleceği üzere modeller birbirine benzemektedir. Şekil 24’e göre ÇO değişkeninde YSA modeli ölçüm verilerine daha çok benzemektedir. Şekil 24’de görüleceği üzere ölçülen ÇO konsantrasyonunun yaklaşık 15’den büyük değerleri için 4 makine öğrenmesi modellerinin tahmin değerleri düşük sonuçlar vermişlerdir. Ayrıca DÖ tahmin modelinin ortalaması ölçülen ÇO konsantrasyonunun ortalamasından daha düşük, GSR modelinin ortalaması ölçülen ÇO konsantrasyonunun ortalamasından biraz daha yüksek sonuçlar verdiği görülmektedir.



Şekil 23. Dört farklı modele ait tahminlerin Taylor diyagramı



Şekil 24. Dört farklı modele ait tahminlerin Violin diyagramı

Çalışmanın son aşamasında ise istatistiksel anlamlılık testi olan KW testi yapılmıştır. Modellere ait sonuçların doğruluğunu kontrol etmek için KW testinden yararlanılmıştır. KW testi yardımıyla, tahminle ölçülen ÇO değerlerinin ortalamalarının aynı olup olmadığı belirlendi.

Tablo 8. KW testinin %5 anlamlılık düzeyinde P değerleri

	YSA	DVM	GSR	DÖ
p-value	0.7000	0.7519	0.1245	0.3946
*H ₀	RED	RED	RED	RED

*H₀: Ortalama beklenen ve gerçek değerler arasında bir farklılık var.

Tablo 8’de KW testinin sonuçları verilmiştir. H₀ hipotezi, Tablo 8’de gösterildiği gibi reddedilir. Yani, beklenen ve ölçülen ÇO değerlerinin ortalamalarının önemli ölçüde farklı olmadığını göstermektedir. Başka bir deyişle, H₀ hipotezini reddeder ve öngörülen değerlerin doğru olduğunu göstermektedir.

4. BÖLÜM

TARTIŞMA-SONUÇ VE ÖNERİLER

4.1.Tartışma

Yaşamın devamı için suyun değeri tartışılmaz önemdedir. Günümüzde artarak devam eden yapay zeka alanındaki teknolojik gelişmeler hidroloji biliminin birçok alanında karar alma süreçlerinde yardımcı olmaktadır. Su kalitesi parametreleri içerisinde ÇO konsantrasyonu parametresi dikkatli ve sürekli bir şekilde izlenmesi gereken bir parametredir. Bu çalışmada Sakarya Havzası'na ait ÇO değerinin diğer su kalitesi parametreleri olan T, Eİ, pH, BOİ ve aylar yardımıyla tahmini amaçlanmıştır. ÇO tahmini için YSA, DÖ, DVM ve GSR gibi dört farklı makine öğrenmesi modelleri oluşturulmuş ve bu modellerin performansları karşılaştırılmıştır.

Bu çalışmada yapay zekâ yöntemlerinin su kalitesi parametreleri için kullanımı, dört farklı makine öğrenmesi modelleri ile incelenmiştir. Bu modeller içerisinde en iyi performansa ulaşabilmek için birçok deneme-yanılma yapılmış ve YSA modeli diğer yöntemlere göre daha iyi sonuç vermiştir.

Bu çalışma sonucunda, Sakarya Havzası'na ait su kalitesi parametrelerinde yapay zekâ yöntemlerinin kullanılabileceği ortaya konulmuştur. Dört farklı makine öğrenmesi olan YSA, DÖ, DVM ve GSR modelleri kullanılmıştır. Suyun kalitesinin belirlenmesinde ve canlı yaşam için önemli bir parametre olan ÇO konsantrasyonu parametresi bu modeller ile tahmin edilmiştir.

Tablo 9. Literatürdeki benzer çalışmalara ait değerler

Makale	Çalışma Alanı	Girdi Parametreleri	Çıktı Parametreleri	Yöntem	Performans Kriterleri
Sengorur et al. (2005)		NO ₂ -N, NO ₃ -N, BOİ, debi, T	ÇO	YSA	OKH, R ²
Solanki et al. (2015)	Hindistan	ÇO, pH, bulanıklık	ÇO, pH, bulanıklık	DÖ, YSA, LR	OMH, OKH
Kisi ve Parmar (2015)	Hindistan	Serbest amonyak azotu, T, toplam koliform, fekal koliform, pH	KOİ	M5Tree, DVM, MARS	KOKH, OMH, R ²
Mohammadpour et al. (2015)	Malezya	T, pH, ÇO, Eİ, AKM, nitrit, nitrat, amonyak nitrojen, BOİ, KOİ ve fosfat	Su kalitesi indeksi	DVM ve YSA	R ² , OMH, KOKH
Ozel et al. (2017)	Bartın	T, pH, KOİ, AKM, Eİ	BOİ	ÇDR, RTSA, ÇKASA	OMH, KOKH ve R ²
Talesh et al. (2019)	İran	ÇO, pH, toplam katı miktarı, T, nitrat, fosfat, BOİ, bulanıklık ve koliform	Su kalitesi indeksi	DVM	R ² ve KOKH
Liu et al. (2019)	Çin	T, pH, ÇO, Eİ, bulanıklık, KOİ	T, pH, ÇO, Eİ, bulanıklık, KOİ	DÖ	Bayesian information criterion (BIC)
Ma et al. (2019)	ABD	T, pH, ÇO	BOİ	DÖ	KOKH, OMH,
Barzegar et al. (2020)	Yunanistan	pH, Eİ, T, OİP	ÇO, Chl-a	DÖ, DVR, KA-YSA	R, KOKH, OMH, NOMH, SY, NSE, Wİ
Zhang et al. (2020)	Avustralya	T, Eİ, pH, bulanıklık, Chl-a	ÇO	kPCA-YSA, DVM	KOKH, OMH, R ²
Dehghani et al. (2021)	ABD	T, akış	ÇO	CSO-DVM, SSD-DVM, BWO-DVM, AIG-DVM	KOKH, OMH, R ² , NSE, BIAS
Kumar et al. (2022)	Hindistan	Bulanıklık, toplam çözülmüş katı madde miktarı, KOİ, pH, fosfat, demir, nitrat, sodyum ve klorür	Su kalitesi indeksi	CNN, LSTM	R, KOKH, OMH, NOMH, SY, NSE
Nacar vd. (2022)	Gümüşhane	T, pH ve Eİ	ÇO	MARS	OKHK, OMH, ORH, R ²
Alqahtani et al. (2022)	Asya	HCO ₃ ⁻ , Kalsiyum, Sülfat, pH, Magnezyum, Klorür, Sodyum, Toplam Çözülmüş Katı ve Eİ	Çözülmüş katı madde	GEP, RF, YSA	R ²

Su kalitesi parametrelerinin yapay zeka teknikleri ile modellenmesiyle ilgili literatür incelendiğinde, Sengorur et al. (2005), Solanki et al. (2015), Zhang et al. (2020), Dehghani et al. (2021), Nacar vd. (2022) gibi çalışmalar başta olmak üzere ÇO konsantrasyon parametresinin tahmini üzerine oldukça az çalışma mevcuttur. Literatürde akarsuya ve atıksuya ait su kalitesi parametrelerinin, makine öğrenmesi ile tahmini üzerine pek çok çalışma mevcuttur.

Literatürdeki mevcut çalışmalar ile bu çalışmadaki giriş verileri benzer olup, giriş verilerinin ÇO'yu tahmin etmedeki duyarlılıkları incelenmemiştir. Yapılan bu çalışmada her bir giriş değişkeni ile çıkış değişkeni arasındaki duyarlılık YSA ile incelenmiştir. Değişkenlerin başarı sırasına göre, YSA'ya giriş değişkenleri verilmiştir.

Literatürdeki çalışmalarda makine öğrenmesi yöntemlerinden DVM, DÖ ve YSA yöntemleri bu tezde de kullanılmış olup, çok yaygın kullanımı olmayan GSR metoduna da yer verilmiştir. YSA, DVM, DÖ ve GSR modellerinin tahmin başarısını değerlendirmek amacıyla KOKH ve OMH kullanılmıştır. KOKH ve OMH kriterlerine göre bu dört yöntemin başarılı tahminler elde edildiği söylenebilir. Sengorur et al. (2005), Ozel et al. (2017), Talesh et al. (2019) çalışmalarında R^2 kriterinde iyileştirme gözlemlense de, bu tezdeki yöntemler ile R^2 değerinde başarılı sonuçlar elde edilememiştir. Bu tez çalışmasında Sakarya Nehri'nin bazı akarsu kollarındaki istasyonlardan yararlanılmıştır. R^2 değerinde bu sebeple iyileşmenin olmadığı düşünülmektedir.

Solanki et al. (2015) ve Liu et al. (2019) çalışmalarında, tek bir istasyona ait değerleri kullanan zaman serisinden yararlanmışlardır. Çalışmalarında DÖ yöntemini tek bir istasyon için geleceğe yönelik tahminlerde bulunmak için kullanmışlardır. Bu çalışma da ise Sakarya Havzası'ndaki 19 adet gözlem istasyonunun BOİ, T, pH ve EI değişkenlerini giriş verisi olarak, DÖ yöntemi ile regresyon modellemesi yapılmıştır.

Literatürdeki çalışmalardan farklı olarak klasik performans kriterlerinin yanı sıra son zamanlarda popülaritesi artan ve aynı zamanda yöntemleri görsel karşılaştırmaya elverişli olan Taylor diyagramından yararlanılmıştır. Taylor diyagramı, ölçülmüş ÇO değişkenin standart sapma-korelasyon değerleri ile modellerin standart sapma-korelasyon değerleri arasındaki mesafeyi görsel olarak sunmaktadır. Böylece ÇO değerini en iyi tahmin eden YSA modeli tespit edilmiştir.

Tez çalışmasının sonunda ise, literatürde su kalitesi parametresinin tahminin de kullanılmayan, fakat diğer tahmin modellerinde sık kullanılan KW testi yapılmıştır. KW testi YSA, DVM, DÖ ve GSR modellerinden elde edilen sonuçların doğruluğunu araştırmıştır.

4.2. Sonuç ve Öneriler

Sakarya Havzası'nda yer alan gözlem istasyonlarına ait 1995 - 2014 yılları arasındaki Şubat, Nisan, Haziran, Ağustos ve Kasım aylarında ölçülmüş su kalitesi verileri ile ÇO değerinin yapay zekâ yöntemleri olan; YSA, DÖ, DVM ve GSR ile modellenmesi yapılmıştır. Bu çalışmada ilk olarak YSA modelinde farklı kombinasyonlar denenmiş ve BOİ, Eİ, pH ve aylar modele ayrı ayrı 1 girişli veri olarak tanımlanmıştır. R^2 değeri sonucuna göre BOİ daha iyi sonuç vermiştir. 2 girişli model BOİ yanında diğer parametrelerle oluşturulmuş ve R^2 değeri sonucuna göre BOİ-Aylar daha iyi sonuç vermiştir. 3 girişli modeller içerisinde R^2 değeri sonucuna göre BOİ-T-pH diğerlerinden daha iyi sonuç vermiştir. Son olarak 4 girişli analizler gerçekleştirilmiş ve en iyi sonucu BOİ-T-pH-Eİ değişkenlerine ait model vermiştir. Daha sonra DÖ, DVM ve GSR makine öğrenmesi modelleri ile, BOİ-T-pH-Eİ değişkenlerine ait 4 girişli analizler gerçekleştirilmiştir. Bu modellere ait sonuçlar Tablo 7'de yer alan hata kriterlerine göre karşılaştırılmıştır. Modellerde kullanılan bağımsız değişkenler; T, Eİ, pH, BOİ ve aylardır. Çalışma sonucunda, hata kriterleri ile Taylor ve Violin diyagramlarına göre en iyi tahmin sonucunu YSA modeli vermiştir. Karşılaştırmada önerilen dört yöntemin etkinliği KW testi ile kanıtlanmıştır.

Ayrıca, 'Machine Learning' temelli DÖ modeli tahmin yöntemi başarılı sonuçlar verdiği tespit edilmiştir.

ÇO parametresini etkileyebilecek T, Eİ, pH ve BOİ parametreleri giriş verisi olarak tercih edilmiştir. YSA modelde ÇO parametresinin tahmininde en etkili değişkenin BOİ değişkeni olduğu R^2 değerlerine göre görülmektedir.

Bu çalışmanın temel sınırlılıkları; ülkemizdeki 25 hidrolojik havzadan birisi olan ve ülke yüzölçümünün %7'sini kapsayan Sakarya Havzasında yapılması, Sakarya Havzası'nda yer alan gözlem istasyonlarına ait 1995 - 2014 (20 yıl) yılları arasındaki Şubat, Nisan, Haziran, Ağustos ve Kasım aylarında ölçülmüş su kalitesi verilerinin kullanılması, ÇO

konsantrasyonu parametresinin tahmininde en fazla dört girişli analizlerin yapılması ve dört farklı makine öğrenmesi yönteminin kullanılmasıdır. Yapılan bu çalışma, ülkemizin birçok sanayi şehri ile birlikte başkentimiz Ankara'nın da yer aldığı Sakarya Havzasında su kalitesinin izlenmesi çalışmalarında karar vericilere fayda sağlayacaktır.

Çalışmada dört farklı makine öğrenmesi yöntemi kullanılmıştır. YSA, DÖ, DVM ve GSR yöntemleri diğer su kalitesi parametrelerinin tahminine yönelik çalışmalarda kullanılabilir.

Su kalitesi parametreleri içerisinde ÇO parametresi su kirliliği ve sucul ekosistem için önemli bir parametredir. Bunun için düzenli izlenmesi önemlidir. İzleme çalışmalarında bazen personel eksikliği bazen de hava koşullarına bağlı olarak aksamalar olabilmektedir. Eksik olan verilerin istatistiksel yöntemler ile tamamlanması yanında yapay zekâ yöntemleri ile de tamamlanması mümkündür.

Su kalitesi parametreleri, su yapıları inşa planlaması çalışmalarında önemli bir altlık oluşturmaktadır. Gelecek yıllar tahmini verileri su yapıları inşa sürecine de önemli katkılar sunabilir.

ÇO parametresinin insan sağlığı üzerinde doğrudan bir etkisi olmamakla birlikte içme sularında düşük konsantrasyonlarda olması boru ve tesisatın korozyonuna neden olabilecektir. ÇO parametresinin takibi ve mümkün değilse tahminine yönelik çalışmalar oluşabilecek sağlık sorunlarını en aza indirebilir.

Bu çalışma, Sakarya Havzası'nda başarılı sonuç vermiş olup diğer havzalara örnek teşkil etmekte ve diğer havzalar içinde uygulanabilir. YSA, DÖ, DVM ve GSR gibi 4 farklı makine öğrenmesi modeli kullanılmış olup diğer makine öğrenmelerinin performansı incelenebilir. Ayrıca YSA, DÖ, DVM ve GSR gibi makine öğrenmelerinin performans kriterlerini iyileştirici hibrit modellerin (Genetic-YSA, parçacık sürü optimizasyonu-YSA gibi) kullanılması ileriki çalışmalar için önerilebilir.

KAYNAKÇA

1. T. C. Kalkınma Bakanlığı, 2018. On birinci kalkınma planı 2019-2023: su kaynakları yönetimi ve güvenliği özel ihtisas komisyonu raporu. (Web sayfası: [https://www.sbb.gov.tr/wp-content/uploads/2020/04/SuKaynaklari-Yonetimi ve GuvenligiOzelIhtisasKomisyonuRaporu.pdf](https://www.sbb.gov.tr/wp-content/uploads/2020/04/SuKaynaklari-Yonetimi-ve-GuvenligiOzelIhtisasKomisyonuRaporu.pdf)), (Erişim tarihi: Şubat 2022).
2. Güler, Ç., 1997. Su Kalitesi. Çevre Sağlığı Temel Kaynak Dizisi: 43. Ankara, 95 s.
3. Akkaya, C., Efeoğlu, A., Yeşil, N., 2006. Avrupa Birliği Su Çerçeve Direktifi ve Türkiye’de Uygulanabilirliği, s.195-204. *TMMOB Su Politikaları Kongresi*, 21-23 Mart 2006, Ankara, TMMOB, 350 s.
4. T. C. Cumhurbaşkanlığı Mevzuat Bilgi Sistemi, 2004. Su kirliliği kontrolü yönetmeliği. (Web sayfası: <https://www.mevzuat.gov.tr/mevzuat?Mevzuat-No=7221&MevzuatTur=7&MevzuatTertip=5>), (Erişim tarihi: Nisan 2022).
5. DSİ, 2020. Devlet su işleri genel müdürlüğü toprak su kaynakları. (Web sayfası: <https://www.dsi.gov.tr/Sayfa/Detay/754>), (Erişim tarihi: Şubat 2022).
6. T. C. Milli Eğitim Bakanlığı, Mesleki Eğitim ve Öğretim Sisteminin Güçlendirilmesi Projesi (MEGEP) Yayınları, 2011. Çevre Sağlığı, Suların Analiz Parametreleri, 850CK0011. Ankara, 64 s.
7. T. C. Cumhurbaşkanlığı Mevzuat Bilgi Sistemi, 2005. İnsani tüketim amaçlı sular hakkında yönetmelik. (Web sayfası: <https://www.mevzuat.gov.tr/-mevzuat?MevzuatNo=7510&MevzuatTur=7&MevzuatTertip=5>), (Erişim tarihi: Şubat 2022).
8. Akgiray, Ö., 2003. İçme suyu kalitesi parametreleri. **Tesisat Dergisi**, Mayıs, 2003, 89.
9. Oğuz, T. C., 2015. İçme Suyu Arıtımında Yaygın Olarak Karşılaşılan Su Kalite Problemleri ve Arıtımı İçin Çözüm Önerileri. T. C. Orman ve Su İşleri Bakanlığı, Uzmanlık Tezi, 92 s.

10. World Health Organization (WHO), 1996. Trace elements in human nutrition and health. Geneva, 343 s.
11. Yıldız, D., Özgüler, H., 2020. Yapay zeka ve su yönetimi. **Su Politikaları Derneği, Rapor No:30**, 35 s.
12. Köklü, R., 2010. Melen Nehri Su Kalitesinin İstatistiksel Analiz Yöntemleri ve Yapay Zekâ Teknikleri Kullanılarak Değerlendirilmesi. Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Sakarya, 299 s.
13. Öztemel, E., 2012. Yapay Sinir Ağları. Papatya Yayıncılık, İstanbul, 233 s.
14. Sengorur, B., Dogan, E., Koklu, R., Samandar, A., 2005. Dissolved oxygen estimation using artificial neural network for water quality control. **Electronic Letters on Science & Engineering**, 1 (2): 13-16.
15. Terzi, Ö., Baykal, T., 2012. Akarsulardaki katı madde miktarının yapay sinir ağları ile tahmini: kızılırmak nehri örneği. **SDU International Journal of Technologic Sciences**, 4 (3): 8-14.
16. Solanki, A., Agrawal, H., Khare, K., 2015. Predictive analysis of water quality parameters using deep learning. **International Journal of Computer Applications**, 125 (9): 0975-8887.
17. Kisi, O., Parmar, K. S., 2015. Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. **Journal of Hydrology**, 534: 104-112.
18. Mohammadpour, R., Shaharuddin, S., Chang, C. K., Zakaria, N. A., 2015. Prediction of water quality index in constructed wetlands using support vector machine. **Environmental Science and Pollution Research**, 22 (8): 6208-6219.
19. Zhang, N., Xiong, J., Zhong, J., Leatham, K., 2018. Gaussian process regression method for classification for high-dimensional data with limited samples, pp. 358-363. *8th International Conference on Information Science and Technology*, June 30-July 6, 2018, Granada, Cordoba, and Seville, Spain, IEEE, 414 s.
20. Ozel, H. U., Gemici, B. T., Ozel, H. B., Gemici, E., 2017. Determination of water quality and estimation of monthly biological oxygen demand (BOD) using by

- different artificial neural networks models in the bartin river. **Fresenius Environmental Bulletin**, **26** (8): 5465-5476.
21. Kamyab-Talesh, F., Mousavi, S-F., Khaledian, M., Yousefi-Falakdehi, O., Norouzi-Masir, M., 2019. Prediction of water quality index by support vector machine: a case study in the sefidrud basin, Northern Iran. **Water Resources**, **46** (1): 112-116.
 22. Liu, P., Wang, J., Sangaiah A. K., Xie, Y., Yin, X., 2019. Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. **Sustainability**, **11** (7): 1-14.
 23. Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., Xu, Z., 2019. Soft detection of 5-day BOD with sparse matrix in city harbor water using deep learning techniques. **Water Resources**, **170**: 115350.
 24. Barzegar, R., Aalami, M. T., Adamowski, J., 2020. Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. **Stochastic Environmental Research and Risk Assessment**, **34**: 415–433.
 25. Thai-Nghe, N., Thanh-Hai, N., Ngon, N. C., 2020. Deep learning approach for forecasting water quality in IoT systems. **International Journal of Advanced Computer Science and Applications (IJACSA)**, **11** (8): 686-693.
 26. Zhang, Y., Fitch, P., Thorburn, P. J., 2020. Predicting the trend of dissolved oxygen based on the kPCA-RNN model. **Water**, **12** (2): 585.
 27. Citakoglu, H., Gemici, B. T., Demir, A., 2021. Regionalization and mapping of dissolved oxygen concentration of sakarya basin by l–moments method. **Mühendislik Bilimleri ve Tasarım Dergisi**, **9** (2): 495-510.
 28. Dehghani, R., Torabi Poudeh, H., Izadi, Z., 2021. Dissolved oxygen concentration predictions for running waters with using hybrid machine learning techniques. **Modeling Earth Systems and Environment**, **8**: 2599–2613.
 29. Şener, A., Ergen, B., Mızrak, H., 2022. Kentsel içme suyu şebekelerinde suyun kalite seviyesinin yapay zekâ yaklaşımları kullanılarak belirlenmesi. **Fırat Üniversitesi Fen Bilimleri Dergisi**, **34** (1): 49-57.

30. Kumar, P. S., Prasad, D. V. V., Venkataramana, L. Y., Prasannamedha, G., Harshana, S., Srividya, S. J., Harrinei, K., Indraganti, S., 2022. Analysis and prediction of water quality using deep learning and auto deep learning techniques. **Science of The Total Environment**, **821**: 153311.
31. Nacar, S., Baki, O. T., Bayram, A., 2022. Akarsularda çözünmüş oksijen konsantrasyonunun regresyon tabanlı yöntemlerle modellenmesi: harşit çayı örneği. **Uludağ Üniversitesi Mühendislik Fakültesi Dergisi**, **27** (1): 309-324.
32. Başakın, E. E., Özger, M., 2022. Missing data imputation for solar radiation by deep neural network . **Avrupa Bilim ve Teknoloji Dergisi**, **35**: 548-555.
33. Alqahtani, A., Shah, M. I., Aldrees, A., Javed, M. F., 2022. Comparative assessment of individual and ensemble machine learning models for efficient analysis of river water quality. **Sustainability**, **14**: 1183.
34. Yaykiran, S., 2016. Sakarya Havzası'nın Yüksek Çözünürlüklü Hidrolojik Modelinin Yapılandırılması. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, 178 s.
35. Ataseven, B., 2013. Yapay sinir ağları ile öngörü modellemesi. **Öneri Dergisi**, **10** (39): 101-115.
36. Orhan, U., 2011. Makine öğrenmesi (yapay sinir ağlarına giriş, tek katmanlı ysa'lar). (Web sayfası: <https://ceng.cu.edu.tr/uorhan/DersNotu/Ders06.pdf>), (Erişim tarihi: Şubat 2022).
37. Esen, B. A., 2017. Yapay sinir ağı modeli. (Web sayfası: <https://www.akanesen.com/2017/09/yapay-sinir-ag-modeli.html>), (Erişim tarihi: Şubat 2022).
38. Çınar U. K., 2018. Yapay sinir ağları ve R programıyla uygulama (artificial neural network). (Web sayfası: <https://www.veribilimiokulu.com/yapay-sinir-aglari/>), (Erişim tarihi: Şubat 2022).
39. Öztürk, K., Şahin, M. E., 2018. Yapay sinir ağları ve yapay zekâya genel bir bakış. **Takvim-i Vekayi**, **6** (2): 25-36.

40. Doğan, F., Türkoğlu, İ., 2018. Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme. **Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi**, 10 (2): 409-445.
41. Alp, E. C., 2018. Recurrent neural network. (Web sayfası: <http://elitcenkalp.blogspot.com/2018/04/recurrent-neural-network.html>), (Erişim tarihi: Şubat 2022).
42. Gürbüz F. B., 2020. Derin oto kodlayıcılar (denoising autoencoders)(DAE) ve oto-kodlayıcılar (autoencoders)(AE). (Web sayfası: <https://medium.com/@batincangurbuz/derin-oto-kodlayicilar-denoising-autoencoders-dae-ve-oto-kodlayicilar-autoencoders-ae-bbd3b1389d60>), (Erişim tarihi: Şubat 2022).
43. Doğan, F., Türkoğlu, İ., 2018. Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması. **Sakarya University Journal of Computer and Information Sciences**, 1 (1): 10-21.
44. Vapnik, V., 1998. The support vector method of function estimation, pp. 55-85. *In: Nonlinear modeling* (Eds: J. A. K. Suykens, J. Vandewalle). Springer, Boston, MA.
45. Demirci, M., 2019. Destek vektör makineleri ve M5 karar ağacı yöntemleri kullanılarak yağış akış ilişkisinin tahmini. **Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi**, 10 (3): 1113-1124.
46. Yang, J., Awan, A. J., Vall-Llosera, G., 2019. Support vector machines on noisy intermediate scale quantum computers. **arXiv preprint arXiv:1909.11988v1**.
47. Akca, M. F., 2020. Nedir bu destek vektör makineleri? (makine öğrenmesi serisi-2). (Web sayfası: <https://medium.com/deep-learning-turkiye/nedir-bu-destek-vekt%C3%B6r-makineleri-makine-%C3%B6%C4%9Frenmesi-serisi-2-94e576e4223e>), (Erişim tarihi: Mart 2022).
48. Ayhan, S., Erdoğan, Ş., 2014. Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. **Eskişehir Osmangazi Üniversitesi İİBF Dergisi**, 9 (1): 175-198.

49. Metlek, S., Kayaalp, K., 2020. Makine Öğrenmesinde, Teoriden Örnek Matlab Uygulamalarına Kadar Destek Vektör Makineleri. İksad Yayınevi, Ankara, 99 s.
50. Tolun, S., 2008. Destek Vektör Makineleri: Banka Başarısızlığının Tahmini Üzerine Bir Uygulama. İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi, İstanbul, 265 s.
51. Melgani, F., Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. **IEEE Transactions on Geoscience and Remote Sensing**, **42** (8): 1778-1790.
52. Bray, M. T. J., Han, D., 2004. Identification of support vector machines for runoff modelling. **Journal of Hydroinformatics**, **6** (4): 265-280.
53. Çağlayan Akay, E., 2018. Ekonometride yeni bir ufuk: büyük veri ve makine öğrenmesi. **Social Sciences Research Journal**, **7** (2): 41-53.
54. Büyüköztürk, Ş., 2012. Veri Analizi El Kitabı (İstatistik, Araştırma Deseni SPSS Uygulamaları Ve Yorum). Pegem Akademi, Ankara, 201 s.
55. Rasmussen, C. E., Williams, K. I., 2006. Gaussian Processes For Machine Learning. The MIT Press, London, England, 248 pp.
56. Taşkın, G., 2022. Manifold öğrenme yöntemlerinin ileri seviye regresyon yöntemleri ile genelleştirilmesi. **Journal of the Faculty of Engineering and Architecture of Gazi University**, **37** (1): 485-495.
57. Fairbrother, J., Nemeth, C., Rischard, M., Brea, J., Pinder, T., 2018. GaussianProcesses.jl: a nonparametric bayes package for the julia language. **arXiv preprint arXiv:1812.09064**.
58. Yergök, D., Acı, M., 2019. Toplu yemek üretiminde günlük talep tahmini için alternatif bir yaklaşım: öğrenci regresyon. **European Journal of Science and Technology, (Özel Sayı)**, 64-73.
59. Cıtakoglu, H., 2021. Comparison of multiple learning artificial intelligence models for estimation of long-term monthly temperatures in Turkey. **Arabian Journal of Geosciences**, **14** (20): 2131.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı: Yusuf ÖZEREN
Uyruğu: Türkiye (T.C)

EĞİTİM

Derece	Kurum	Mezuniyet Tarihi
Lisans	Hacettepe Üniversitesi, Mühendislik Fakültesi, Hidrojeoloji Mühendisliği	2014
Lise	İbni Sina Anadolu Lisesi, Ankara	2003

İŞ DENEYİMLERİ

Yıl	Kurum	Görev
2021-Halen	DSİ 12. Bölge Müdürlüğü	Şube Müdürü
2015-2021	DSİ 12. Bölge Müdürlüğü	Mühendis

YABANCI DİL

İngilizce

YAYINLAR

1. Çıtakoğlu, H., Özeren, Y., Çoşkun, Ö., 2019. Short time drought estimation of sakarya basin station with wavelet model-adaptive neuro-fuzzy inference system, s. 190-197 *WW'2019: International Conference on Image Processing, Wavelet and Application* 20 October 2019, Kocaeli, Turkey.
2. Çıtakoğlu, H., Özeren, Y., 2021. Sakarya basin water quality parameters modeling with artificial neural networks. **Avrupa Bilim ve Teknoloji Dergisi**, 24 (2): 10-17.