

**NESNELERİN İNTERNETİ VE MAKİNE ÖĞRENMESİ İLE JAR
TESTİ SÜRECİNİN OPTİMİZASYONU**

FERDİ AKINCI

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**DANIŞMAN
DR. ÖĞR. ÜYESİ ENVER KÜÇÜKKÜLAHLI**

DÜZCE, 2025

T.C.
DÜZCE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

NESNELERİN İNTERNETİ VE MAKİNE ÖĞRENMESİ İLE JAR
TESTİ SÜRECİNİN OPTİMİZASYONU

Ferdi AKINCI tarafından hazırlanan tez çalışması aşağıdaki jüri tarafından Düzce Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Dr. Öğr. Üyesi Enver KÜÇÜKKÜLAHLI

Düzce Üniversitesi

Jüri Üyeleri

Dr. Öğr. Üyesi Enver KÜÇÜKKÜLAHLI

Düzce Üniversitesi

Prof. Dr. Ali ÇALHAN

Düzce Üniversitesi

Doç. Dr. Nihat DALDAL

Bolu Abant İzzet Baysal Üniversitesi

Tez Savunma Tarihi: 16/06/2025

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün aşamalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

16 Haziran 2025

Ferdi AKINCI

TEŞEKKÜR

Yüksek lisans öğrenimim süresince ve bu tezin hazırlanmasında verdiği kıymetli destek ve rehberlikten dolayı çok değerli danışman hocam Dr. Öğr. Üyesi Enver Küçükkülahlı'ya ve katkılarından dolayı Doç. Dr. Murat Solak'a en içten teşekkürlerimi sunarım.

Çalışma kapsamında bizlere ev sahipliği yapan ve teknik bilgileriyle sürece katkıda bulunan Kdz. Ereğli Belediyesi yönetimine, Su ve Kanalizasyon İşleri Müdürü Sayın Deniz Koca'ya, İçme Suyu Arıtma Tesisi Sorumlusu Sayın Levent Tigin Kalyoncu ve ekibine teşekkür ederim.

Hayatım boyunca desteğini her zaman yanımda hissettiğim aileme, sevgili eşime ve yaşamımıza neşe katan oğlumuz Emir Asaf'a sonsuz şükranlarımı sunarım.

Ayrıca, akademik hayatım boyunca bana yol gösteren tüm hocalarıma ve bu çalışma süresince desteklerini esirgemeyen değerli çalışma arkadaşlarıma teşekkürlerimi sunarım.

16 Haziran 2025

Ferdi AKINCI

İÇİNDEKİLER

Sayfa No



LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

.....	i
ŞEKİL LİSTESİ.....	vii
ÇİZELGE LİSTESİ.....	viii
KISALTMALAR.....	ix
SİMGELER	xi
ÖZET	xii
ABSTRACT	xiii
1. GİRİŞ.....	1
1.1. LİTERATÜR ÖZETİ.....	4
1.2. TEZİN AMACI.....	10
1.3. LİTERATÜRE KATKI.....	11
2. MATERYAL.....	12
2.1. DONANIMLAR.....	12
2.1.1. Nodemcu Esp8266	12
2.1.2. Yağmur Sensörü.....	13
2.1.3. Ph Sensörü	13
2.1.4. Su Sıcaklık Sensörü.....	14
2.1.5. Su Seviye Sensörü.....	15
2.1.6. TDS Sensörü	16
2.1.7. Raspberry Pi.....	16
2.2. FOG BİLİŞİM.....	17
2.2.8.1. Mosquitto	18
2.2.8.2. Node-RED.....	18
2.3. BULUT SİSTEMİ	18
2.3.1.1. Docker.....	19
2.3.1.2. InfluxDB.....	19
2.3.1.3. PostgreSql.....	20
2.3.1.4. Grafana.....	20
2.4. MAKİNE ÖĞRENMESİ.....	21
2.4.1. Veri Ön İşleme.....	21
2.4.2. Model.....	22
2.4.3. Hiperparametre.....	23
2.4.4. Model Değerlendirmeleri.....	23
2.4.4.1. Mean Absolute Error (MAE).....	24
2.4.4.2. Mean Absolute Percentage Error (MAPE)	24
2.4.4.3. Mean Squared Error (MSE).....	24
2.4.4.4. Root Mean Squared Error (RMSE).....	24
2.4.4.5. R-Squared (R^2)	25

3. METOT	26
3.1. IOT MİMARİSİ	27
3.1.1. Sensörlerin Bağlantıları.....	28
3.1.1.1. Bulanıklık ve Su Seviye Sensörü.....	28
3.1.1.2. Ph ve Su Sıcaklığı Sensörü.....	29
3.1.1.3. TDS ve Yağmur Sensörü	30
3.1.2. Fog Katmanı	30
3.1.2.1. Node-RED.....	31
3.1.3. Bulut Katmanı	32
3.1.3.1. Docker.....	32
3.1.3.2. InfluxDB.....	32
3.1.3.3. PostgreSQL.....	32
3.1.3.4. Grafana.....	33
3.1.4. Sensör Verilerinin Değerlendirilmesi.....	33
3.2. IOT DESTEKLİ MAKİNE ÖĞRENMESİ MODELİ.....	34
3.3. MAKİNE ÖĞRENMESİ.....	36
3.3.1. Veri seti	36
3.3.2. Çalışmada Kullanılacak Özellikler.....	36
3.3.3. Veri Ön İşleme.....	38
3.3.4. Verilerin Dağılımı.....	38
3.3.5. Modelin Eğitilmesi	39
3.3.6. Hiperparametre Optimizasyonu.....	39
3.3.7. Modelin Değerlendirilmesi	41
3.4. GELİŞTİRİLEN YAZILIM	41
4. BULGULAR VE TARTIŞMA	45
5. SONUÇ	50
6. KAYNAKLAR.....	52
ÖZGEÇMİŞ.....	57

ŞEKİL LİSTESİ

	<u>Sayfa No</u>
Şekil 2.1. Nodemcu esp8266 [25].....	12
Şekil 2.2. Yağmur sensörü [27].	13
Şekil 2.3. Ph sensörü [29].	13
Şekil 2.4. Su sıcaklık sensörü [31].....	14
Şekil 2.5. Su seviye sensörü [32].....	15
Şekil 2.6. TDS sensörü [33].....	16
Şekil 2.7. Raspberry pi 4 [34].	16
Şekil 2.8. Fog bilişim mimarisi.....	17
Şekil 3.1. IoT, fog, bulut mimarisi.....	27
Şekil 3.2. Bulanıklık ve mesafe sensörü bağlantı şeması.	28
Şekil 3.3. pH ve su sıcaklığı sensörü bağlantı şeması.	29
Şekil 3.4. İletkenlik ve yağmur sensörü bağlantı şeması.	30
Şekil 3.5. Nodered ekran görüntüsü.....	31
Şekil 3.6. Grafana ekran görüntüsü.	33
Şekil 3.7. Akış diyagramı.	34
Şekil 3.8. Korelasyon ısı haritası.	37
Şekil 3.9. Verilerin dağılım grafiği.	39
Şekil 3.10. Yazılım giriş sayfası.	42
Şekil 3.11. Su tesisi anlık veriler sayfası(ana sayfa) ekran görüntüsü.....	42
Şekil 3.12. Yeni kayıt sayfası ekran görüntüsü.	43
Şekil 3.13. Kayıtları listele sayfası ekran görüntüsü.	43
Şekil 3.14. Jar testi sayfası ekran görüntüsü.....	44
Şekil 3.15. Model eğitimi sayfası ekran görüntüsü.	44
Şekil 4.1. IoT ile toplanan verilerin anlık görüntüsü.	45
Şekil 4.2. Yağmur uyarısı mail ekran görüntüsü.	46
Şekil 4.3. Modellerin R ² çıktıları.	48
Şekil 4.4. SVR modeli hata dağılımı grafiği.....	48

ÇİZELGE LİSTESİ

	<u>Sayfa No</u>
Çizelge 2.1. Makine öğrenmesi modelleri.	22
Çizelge 3.1. Bulanıklık ve mesafe sensörü bağlantı pinleri.	28
Çizelge 3.2. pH ve su sıcaklığı sensörü bağlantı pinleri.	29
Çizelge 3.3. İletkenlik ve yağmur sensörü bağlantı pinleri.	30
Çizelge 3.4 Sensörlerin kabul edilebilir değerleri.	33
Çizelge 3.5. Veri seti özellikleri.	36
Çizelge 3.6. Makine öğrenmesinde kullanılacak verilerin veri işleme öncesi istatistiksel sonuçları.	37
Çizelge 3.7. Makine öğrenmesinde kullanılacak verilerin veri işleme sonrası özellikleri.	38
Çizelge 3.8. Hiperparametreler.	40
Çizelge 4.1. Model performans sonuçları.	46

KISALTMALAR

1D-CNN	One-Dimensional Convolutional Neural Network
3D	Three Dimensional
ACID	Atomicity, Consistency, Isolation, Durability
AdaBoost	Adaptive Boosting
ADC	Analog to Digital Converter
Ag/AgCl	Silver/Silver Chloride Electrode
AHP	Analytic Hierarchy Process
Al ₂ (SO ₄) ₃	Aluminium sulfate
ANN	Artificial Neural Network
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BLE	Bluetooth Low Energy
BOD	Biochemical Oxygen Demand
CART	Classification and Regression Trees
CatBoost	Categorical Boosting
CPU	Central Processing Unit
DBPs	Disinfection By-Products
DO	Dissolved Oxygen
DQN	Deep Q-Network
DRAM	Dynamic RAM
DSÖ	Dünya Sağlık Örgütü
DT	Decision Tree
DWQI	Drinking Water Quality Index
E.coli	Escherichia coli
EC	Electrical Conductivity
FDR	False Discovery Rate
FNR	False Negative Rate
GAN	Generative Adversarial Network
GNN	Graph Neural Network
GPIO	General Purpose Input/Output
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
I2C	Inter-Integrated Circuit
I2S	Integrated Interchip Sound
IDE	Integrated Development Environment
IoT	Internet of Things
KB	Kilobyte
KCl	Potassium chloride
KMnO ₄	Potassium permanganate
LCD	Liquid Crystal Display
LightGBM	Light Gradient Boosting Machine
LoraWAN	Long Range Wide Area Network

LPDDR4	Low Power Double Data Rate 4
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MCC	Matthews Correlation Coefficient
MICE	Multivariate Imputation by Chained Equations
ML	Machine Learning
MLP	Multi-Layer Perceptron
MQTT	Message Queuing Telemetry Transport
MSE	Mean Squared Error
NARNet	Nonlinear AutoRegressive Network
NB-IoT	Narrowband Internet of Things
NFS	National Sanitation Foundation
NON-OS	Non-Operating System
ORDBMS	Object-Relational Database Management System
OS	Operating System
PCB	Printed Circuit Board
pH	Hydrogen ion concentration / pH value
PLC	Programmable Logic Controller
PPM	Parts Per Million
PVDF	Polyvinylidene Fluoride
PWM	Pulse Width Modulation
R ²	Coefficient of determination
RAM	Random Access Memory
RF	Random Forest
RMSE	Root Mean Squared Error
RST	Reset
SDK	Software Development Kit
SMS	Short Message Service
SPI	Serial Peripheral Interface
SPIFFS	Serial Peripheral Interface Flash File System
SQL	Structured Query Language
SSL/TLS	Secure Sockets Layer / Transport Layer Security
SVM	Support Vector Machine
SVR	Support Vector Regression
TCP/IP	Transmission Control Protocol/Internet Protocol
TDS	Total Dissolved Solids
TOC	Total Organic Carbon
UART	Universal Asynchronous Receiver-Transmitter
UnionFS	Union File System

SİMGELER

GHz	Gigahertz
H ⁺	Hidrojen iyonu
kg/h	Kilogram/saat
kHz	Kilohertz
kΩ	Kilohm
lt/h	Litre/saat
m ³ /h	Metreküp/saat
Ω	Ohm



ÖZET

NESNELERİN İNTERNETİ VE MAKİNE ÖĞRENMESİ İLE JAR TESTİ SÜRECİNİN OPTİMİZASYONU

Ferdi AKINCI

Düzce Üniversitesi

Lisansüstü Eğitim Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Danışman: Dr. Öğr. Üyesi Enver KÜÇÜKKÜLAHLI

Haziran 2025, 56 sayfa

Dünya yüzeyinin %71'i suyla kaplı olmasına rağmen, bunun yalnızca %2,5'i tatlı sudur. Dünya sağlık örgütü'nün verilerine göre, yaklaşık 844 milyon insan güvenilir içme suyu kaynağına erişememekte, 159 milyon insan ise yüzey sularını kullanmaktadır. Bu durum, içme suyu kalitesinin sürekli izlenmesini ve iyileştirilmesini zorunlu kılmaktadır. İçme suyu kalitesinin değerlendirilmesinde su kalite indeksi yaygın olarak kullanılmakta olup, çeşitli fiziksel ve kimyasal parametrelerin ölçümüne dayanmaktadır. Bu çalışmada, Türkiye'nin Zonguldak ili Ereğli ilçesindeki içme suyu arıtma tesisi örnek alınarak, klasik jar testi sürecinin dijitalleştirilmesi ve optimize edilmesi hedeflenmiştir. Geliştirilen sistem; iot, fog bilişim ve bulut tabanlı makine öğrenmesi altyapılarıyla entegre bir şekilde tasarlanmıştır. IoT katmanında kullanılan sensörler ile giriş suyuna ait bulanıklık, ph, tds, sıcaklık, su seviyesi ve yağmur durumu gibi parametreler 5 dakikalık aralıklarla ölçülmüş ve mqtt protokolü ile nodemcu esp8266 üzerinden raspberry pi cihazına iletilmiştir. Raspberry pi üzerinde çalışan nodered uygulaması ile veriler influxdb bulut veritabanına kaydedilmiş, aynı zamanda grafana arayüzüyle görselleştirilmiştir. Bu yapı sayesinde sistemin uzaktan izlenebilirliği artırılmış, müdahale süreleri azaltılarak operasyonel güvenlik sağlanmıştır. Laboratuvar ortamında yapılan jar testleri, web tabanlı bir arayüz ile dijital ortama aktarılmış ve bu veriler makine öğrenmesi algoritmalarıyla analiz edilmiştir. Özellikle destek vektör regresyonu algoritması, 502 örnek üzerinde yapılan testlerde en yüksek performansı göstermiştir (MAE: 0.29, MAPE: 0.11, MSE: 0.16, RMSE: 0.40, R²: 0.61). Modelin yalnızca 1.27 saniyede sonuç üretmesi, gerçek zamanlı tahmin uygulamaları için yeterli olduğunu göstermektedir. Web tabanlı kullanıcı arayüzü sayesinde operatörler, giriş parametrelerini girerek çıkış bulanıklığını anlık olarak görebilmekte; ayrıca sistemde yer alan “eğitimi başlat” modülüyle model yeni verilerle otomatik olarak güncellenebilmektedir.

Anahtar Sözcükler: Fog Bilişim, Iot, Jar Testi, Makine Öğrenmesi, Su Kalitesi

ABSTRACT

OPTIMIZATION OF THE JAR TESTING PROCESS WITH THE INTERNET OF THINGS AND MACHINE LEARNING

Ferdi AKINCI

Düzce University
Graduate School, Department of Computer Engineering
Master's Thesis
Supervisor: Asst. Prof. Enver KÜÇÜKKÜLAHLI

June 2025, 56 pages

Although 71% of the Earth's surface is covered by water, only 2.5% of it is freshwater. According to the World Health Organization, approximately 844 million people do not have access to a safe drinking water source, while 159 million people use surface water. This situation necessitates continuous monitoring and improvement of drinking water quality. The water quality index is widely used in the assessment of drinking water quality and is based on the measurement of various physical and chemical parameters. In this study, taking the drinking water treatment plant in Ereğli district of Zonguldak province of Turkey as an example, it is aimed to digitize and optimize the classical jar testing process. The system is designed to be integrated with iot, fog computing and cloud-based machine learning infrastructures. With the sensors used in the iot layer, parameters such as turbidity, ph, tds, temperature, water level and rainfall status of the inlet water were measured at 5-minute intervals and transmitted to the raspberry pi device via nodemcu esp8266 via mqtt protocol. With the nodered application running on the raspberry pi, the data was saved to the influxdb cloud database and visualized with the grafana interface. Thanks to this structure, the remote traceability of the system has been increased, intervention times have been reduced and operational security has been ensured. The jar tests performed in the laboratory environment were digitized with a web-based interface and these data were analyzed with machine learning algorithms. In particular, the support vector regression algorithm showed the highest performance in tests on 502 samples (MAE: 0.29, MAPE: 0.11, MSE: 0.16, RMSE: 0.40, R^2 : 0.61). The fact that the model produces results in only 1.27 seconds shows that it is sufficient for real-time forecasting applications. The web-based user interface allows operators to enter input parameters and see the output turbidity instantly, and the “start training” module in the system allows the model to be updated automatically with new data.

Keywords: Fog Computing, Iot, Jar Test, Machine Learning, Water Quality

1. GİRİŞ

Dünya yüzeyinin %71'i suyla kaplı olmasına rağmen, bunun yalnızca %2,5'i tatlı sudur. Dünya Sağlık Örgütü'nün araştırmalarına göre, 844 milyon insan güvenilir içme suyu kaynağından yoksundur ve yalnızca 159 milyon insan yüzey suyuna güvenerek su ihtiyacını karşılamaktadır. Kirli suyun tüketimi insanlarda ishal, dizanteri, kolera, tifo ve çocuk felci gibi hastalıklara neden olmaktadır [1]. İklim değişiklikleri, su kaynağındaki hayvansal kirlilikler ve yağış olayları su kalitesini önemli ölçüde etkilemektedir. Su kirliliğinin artması ile arıtma tesislerinde süreçlerde zorluklar yaşanmaktadır [2].

Su kalitesi tespiti için, Su Kalite İndeksi (WQI-Water Quality Index) hesaplanmaktadır. Su kalite indeksinin hesaplanması; çeşitli parametrelerin ölçülmesi ve ağırlıklarıyla birlikte hesaplanmasıyla mümkün olmaktadır. İndeks sonucuna göre su kalitesi sınıflandırılmaktadır. Su kaynağındaki kalitenin değişimini tespit edebilmek için parametrelerin sürekli bir şekilde toplanıp, kalite indeksinin hesaplanması gerekmektedir. ABD Ulusal Sanitasyon Vakfı Su Kalite İndeksi (NSF-WQI - National Sanitation Foundation-WQI) hesaplamasında; çözünmüş oksijen, fekal koliform, ph, boi, nitrat, sıcaklık, fosfat, bulanıklık ve toplam katı maddeleri ölçülmektedir [3]. Türkiye'deki insani tüketim amaçlı sular hakkında yönetmeliğinde içme suyu kontrolü için izlenmesi gereken parametreler; renk, bulanıklık, koku, tat, iletkenlik, ph, amonyum, alüminyum, demir, c.perfringens(sporlar dahil), e.coli, koliform bakteri, p. aeruginosa, 22 ve 37°C'de koloni sayımı değerleri ölçülmektedir. Parametrelerin kontrolü ile tüketime verilen suyun organoleptik ve mikrobiyolojik kalitesi ve içme-tüketim suyu arıtım sonucunda, arıtma işleminin(dezenfeksiyon) etkili olup olmadığı üzerine düzenli bilgi sağlanmaktadır [4].

İçme suyu arıtma tesislerinde su kalitesini etkileyen en önemli faktörlerden biri, suda askıda kalan katı maddelerdir. Askıda kalan madde miktarı arttıkça suyun bulanıklığı da artmaktadır. Suyun bulanıklığını azaltmak için pıhtılaştırma ve yumaklaştırma işlemi uygulanmaktadır. Pıhtılaştırma için suya katılan malzemeler koagülant (pıhtılaştırıcı) olarak alüminyum sülfat ve flokülant (yumaklaştırıcı) olarak poliakrilamid ve polietilen imin kullanılmaktadır. Katılan bu maddeler askıda kalan maddeleri birleştirerek askıda kalan partiküllerin bir araya gelmesi ve daha büyük parçacıklar oluşturmasını sağlar.

Oluşan büyük taneciklere floklar adı verilir. Floklar, arıtmanın bir sonraki adımında çökeltme işlemi ile sudan ayrıştırılır. Çökeltme işlemi sonrasında suyun bulanıklığında belirgin bir azalma gözlemlenmektedir [5]. Katılacak bu maddelerin dozajının belirlenmesi için su havzasından alınan numuneler 1 litrelik silindirik cam şişeler kullanılarak her model için denemeler yapılarak, optimum pıhtılaştırıcı dozajı bulunur. Bulanıklığın giderilmesi için gerekli maddelerin dozajı yapılan bu test ile belirlenmektedir. Yapılan bu teste jar testi adı verilmektedir [6]. Ancak bu yöntem, suyun anlık izlenmesine olanak tanımadığından, dozajlama işleminin dinamik olarak belirlenmesi mümkün olmamaktadır.

Nesnelerin İnterneti (IoT-Internet of Things), fiziksel nesnelerin dijital ortamda birbirleriyle ve merkezi sistemlerle haberleşerek veri toplaması, iletmesi ve bu veriler doğrultusunda eyleme geçmesini sağlayan bir teknoloji altyapısıdır. Bu paradigma, özellikle çevresel izleme uygulamalarında kritik bir rol üstlenmektedir. Su kaynaklarının kalitesini değerlendirmek amacıyla geliştirilen IoT tabanlı sistemlerde, sıcaklık, pH, bulanıklık ve çözünmüş madde gibi temel su kalite parametrelerini ölçebilen sensörler kullanılarak sürekli ve gerçek zamanlı veri akışı sağlanmaktadır. Böylece, geleneksel yöntemlerin sınırlılıklarını aşarak daha hızlı, doğru ve düşük maliyetli su kalitesi değerlendirmeleri yapılabilmektedir. Bu yaklaşım, yalnızca izleme sürecini değil, aynı zamanda erken uyarı mekanizmalarının oluşturulmasını ve sürdürülebilir su yönetimine yönelik karar alma süreçlerini de güçlendirmektedir [7].

Son yıllarda yapılan çalışmalar, IoT temelli sistemlerin su arıtma tesislerinde (WTPs-Water Treatment Plants) gerçek zamanlı veri toplama, uzaktan erişim ve hızlı müdahale gibi olanaklar sağladığını göstermektedir. Özellikle pH, Çözünmüş Oksijen (DO-Dissolved Oxygen), Toplam Çözünmüş Katılar (TDS-Total Dissolved Solids) ve sıcaklık gibi temel su parametrelerini izleyebilen sensörlerin, düşük güçlü Programlanabilir Lojik Denetleyiciler (PLC-Programmable Logic Controllers) aracılığıyla bir ağ içinde çalıştırılması, bu sistemlerin hem enerji verimliliği hem de veri güvenilirliği açısından avantajlı olduğunu ortaya koymaktadır. Elde edilen veriler, kablosuz iletişim protokolleri aracılığıyla bulut tabanlı platformlara aktarılmakta, böylece gerçek zamanlı analiz, alarm sistemleri ve tarihsel veri takibi mümkün olmaktadır. Bu teknolojik entegrasyon, yalnızca operasyonel verimliliği artırmakla kalmamakta, aynı zamanda su kalitesindeki ani değişimlere karşı erken uyarı sistemleri oluşturarak karar destek mekanizmalarını güçlendirmektedir. Bu yönüyle iot, su arıtma süreçlerinin dijitalleşmesinde temel bir yapı

taşı olarak konumlanmakta ve çevre yönetiminde akıllı sistemlerin gelişimine öncülük etmektedir [8].

Kablosuz iletişim yeteneği, bulut entegrasyonu ve gömülü veri işleme kapasitesi sayesinde bu sistemler yalnızca veri toplamakla kalmamakta; aynı zamanda analiz süreçlerini sahada gerçekleştirme ve sonuçları anlık olarak kullanıcıya ulaştırma potansiyeli taşımaktadır. Böylece, geleneksel yöntemlerin ötesine geçilerek, su kaynaklarının daha sürdürülebilir ve etkili biçimde yönetilmesi mümkün hale gelmektedir [9].

İklim değişikliği, su kaynaklarının niteliği ve yönetimi üzerinde giderek artan bir baskı oluşturmaktadır. Özellikle aşırı hava olaylarının sıklaşması, ani kirlilik artışları ve su kalitesinde öngörülemeyen değişimlere neden olarak geleneksel su arıtma sistemlerini zorlamaktadır. İot tabanlı izleme tek başına yeterli kalmayıp, derin öğrenme destekli modelleme yaklaşımları öne çıkmaktadır. İot sensörleri ile sahadan toplanan veriler, kablosuz ağlar aracılığıyla bulut ya da fog bilişim sistemlerine iletilmekte ve burada ön işleme tabi tutulmaktadır. Bu sistemler sayesinde su kaynakları anlık olarak izlenebilmekte, kirlilik parametreleri uzaktan takip edilebilmektedir [10].

Bununla birlikte, yüksek hacimli ve çok değişkenli verilerin analizinde derin öğrenme modellerinin kullanılması, su arıtma süreçlerinin otomasyonu açısından yeni bir dönemi başlatmaktadır. Özellikle ani hava olayları gibi ekstrem koşullarda, koagülant dozajı ve çıkış bulanıklığı gibi kritik parametrelerin doğru tahmini, sistemlerin kesintisiz çalışması ve insan müdahalesinin azaltılması açısından önemlidir . Bu tür yaklaşımlar, yalnızca mevcut durumu izlemekle kalmayıp, aynı zamanda ani değişimlere proaktif yanıt verebilen karar destek sistemlerinin temelini oluşturmaktadır [2].

Öte yandan, elde edilen büyük ve çok boyutlu verilerin işlenmesinde yapay zekâ ve özellikle makine öğrenmesi modelleri önemli avantajlar sağlamaktadır. Makine öğrenmesi algoritmaları, suyun içilebilir olup olmadığını yüksek doğrulukla sınıflandırabilmekte ve karar verme süreçlerinde etkili öngörüler sunmaktadır [11].

Gelişen teknoloji ile birlikte makine öğrenmesi ve otomatik makine öğrenmesi teknikleri kullanılarak suyun içilebilirlik durumunun tahmini ve WQI sınıflandırılması üzerine kapsamlı bir uygulama gerçekleştirilmiştir. Makine öğrenmesi modellerine örnek olarak Naive Bayes, Lojistik Regresyon, Support Vector Machine (SVM), Karar Ağaçları ve Rastgele Orman, XGBoost, Support Vector Regression (SVR), ExtraTrees, AdaBoost,

LightGBM, CatBoost gibi modeller kullanılmaktadır. Bu durum, su kalite analizlerinin otomasyonu ve karar destek sistemlerinin etkinliđi aısından önemli avantajlar sunmaktadır [12].

1.1. LİTERATÜR ÖZETİ

İme suyu kalitesi, toplum sađlığının korunması ve sürdürülebilir kalkınmanın sađlanabilmesi aısından kritik bir öneme sahiptir. İme suyunda bulunan kimyasal ve mikrobiyolojik kirleticiler önemli halk sađlığı sorunlarına yol açmaktadır. Literatürde yapılan alışmalar, bu kirleticilerin su kaynaklarına dođal yollarla veya insan faaliyetleri sonucunda karıştığını ortaya koymaktadır.

Perveen ve Haque (2023) tarafından gerçekleştirilen alışmada, mikrobiyolojik parametreler arasında yer alan *Escherichia coli* (E. coli) ve koliform bakteriler, ime suyunun hijyenik aıdan uygunluđunu deđerlendirmede temel göstergeler olarak tanımlanmaktadır. Bu mikroorganizmalar, özellikle dışkı kaynaklı bulaşının bir göstergesi olup, halk sađlığı aısından ciddi riskler barındırmaktadır. alışmada yer alan veriler, yüzeysel su kaynaklarında bu tür mikrobiyolojik kirleticilerin daha yüksek oranlarda tespit edildiđini ve uygun dezenfeksiyon ile arıtma süreçlerinin uygulanmaması durumunda bu kirleticilerin yayılımının önlenemeyeceđini ortaya koymaktadır [13].

Hasan ve Hassoon (2021) tarafından gerçekleştirilen alışmada, Irak'taki Kerbela Birleşik Su Projesi kapsamında amur geri dönüşümünün su arıtım verimliliđi üzerindeki etkileri incelenmiştir. Bu kapsamda, Euphrates Nehri'nden alınan ham su örnekleri üzerinde kavanoz testi (jar testi) uygulanmış ve %50 oranında geri dönen amur ilavesiyle arıtım verimliliđinde %14'e varan artış sađlandığı gözlemlenmiştir. Ayrıca, kullanılan alüminyum sülfat miktarında yaklaşık %50 oranında tasarruf elde edilmiştir. Bu durum, hem sađlık aısından sistemde daha az alüminyum kalıntısı oluşmasını sađlamış hem de ekonomik aıdan anlamlı bir fayda yaratmıştır. Alüminyumun sülfat ve hidroksit formlarının ökelme mekanizmaları, özellikle bulanıklığın giderilmesinde etkili olmuş ve geri kazanılan amurun köprölleme etkisi ile arıtım sürecine katkı sunduđu deđerlendirilmiştir. Araştırmacılar, geri dönüşüm oranının dikkatle optimize edilmesi gerektiđini, ünkü aşırı miktarların arıtım performansını olumsuz etkileyebileceđini vurgulamışlardır [6].

Benalia ve ark. (2024) tarafından gerçekleştirilen alışmada, dođal bir koagülant olan

Aloe vera kullanılarak jar testi ile arıtım koşulları optimize edilmiştir. Deneyisel süreçte, belediye atıksuyu farklı pH değerlerinde ve Aloe vera sıvı koagülant dozlarında değerlendirilmiştir. Her bir deneme, 1 litre ham su kullanılarak 160 rpm'de 3 dakika hızlı karıştırma, ardından 30 rpm'de 15 dakika yavaş karıştırma ve 25 dakika çökeltme süresiyle tamamlanmıştır. Bulgular, merkezi bileşik tasarım yöntemiyle analiz edilmiş ve pH 12 ile 0.8 mL/L dozajın optimum koşullar olduğu belirlenmiştir. Bu parametrelerde %99.13 bulanıklık, %94.01 askıda katı madde, %56.26 organik madde ve %88.34 aromatik madde giderimi sağlanmıştır. Elde edilen veriler, jar testinin yalnızca etkili koagülant dozajının saptanmasında değil, aynı zamanda çevresel ve ekonomik açıdan sürdürülebilir bir arıtma süreci geliştirilmesinde de kritik rol oynadığını göstermektedir[14].

Şentürk (2023) tarafından geliştirilen kural tabanlı IoT mimarisinde, arduino uno, enc28j60 ethernet lan modülü, acs712 akım sensörü, lcd ekran, jumo type 202630/42-25 klor sensörü, ph sensörü ve sıcaklık sensörleri kullanılmıştır. Sistemin veri kaydı ve izlenmesi amacıyla özel bir yazılım geliştirilmiştir. Önerilen mimaride, havuz suyunun ph değeri, sıcaklığı ve klor seviyesi ilgili sensörler aracılığıyla ölçülmüştür. arduino uno, bilgisayara Evrensel Seri Veri Yolu (USB-Universal Serial Bus) bağlantısı ile bağlanmış, ayrıca ethernet modülü de arduino uno'ya entegre edilmiştir. Sensörlerden elde edilen veriler arduino uno'ya iletilmiş ve buradan USB bağlantısı aracılığıyla bilgisayardaki yazılıma aktarılmıştır. Yazılım, alınan verileri bağlı olduğu veri tabanına kaydetmiştir. Geliştirilen kural tabanlı mimari, ölçülen değerlerin belirlenen eşik değerlerin dışında olması durumunda, arduino uno'nun ethernet modülü aracılığıyla twitter üzerinden bildirim göndermesini sağlamaktadır [15].

Yılmaz (2019) tarafından tuzdan arındırma tesislerinde su kalitesinin anlık olarak izlenmesi ve olası arızaların tespit edilmesi amacıyla bir IoT mimarisi geliştirilmiştir. Sistem, suyun farklı parametrelerini ölçmek için nodemcu esp8266, sım900a modülü, lm75 sıcaklık sensörü, ph sensörü, tur621 bulanıklık sensörü ve su seviye sensörü gibi donanımlardan oluşmaktadır. Önerilen sistemde, sensörlerden her 10 dakikada bir veri alınarak, bu veriler bulut tabanlı bir web sunucusuna kaydedilmektedir. Sensörlerden gelen ölçüm sonuçlarında anomali tespit edilmesi halinde, sistem tarafından kullanıcılara Kısa Mesaj Servisi (SMS-Short Message Service) ile bildirim gönderilmektedir. Bu kapsamda, "su kalitesi düşük", "arıza oluştu" ve "su sızıntısı tespit edildi" şeklindeki uyarı mesajları, belirlenen senaryolar doğrultusunda iletilmiştir [16].

Lakshmikantha ve arkadaşları (2021) tarafından su kalitesinin izlenilmesi ile ilgili IoT mimarisi tasarlanmıştır. Geliştirilen mimari ph, bulanıklık, iletkenlik, sıcaklık, karbondioksit miktarı ve ortam nemi gibi parametreleri ölçen sensörleri, Arduino tabanlı bir mikrodenetleyiciye bağlayarak yapılandırılmıştır. Elde edilen veriler, Kablosuz Bağlantı (WiFi-Wireless Fidelity) aracılığıyla bulut platformuna aktarılmakta ve bu sayede su kalitesi uzaktan izlenebilmektedir. Gerçekleştirilen testlerde, üç farklı su örneği incelenmiş ve yalnızca bir örneğin içilebilir nitelikte olduğu tespit edilmiştir. Bu sistem, düşük maliyetli, taşınabilir ve kullanıcı dostu yapısıyla, IOT tabanlı su kalitesi izleme sistemlerinin pratikteki etkinliğini ve sürdürülebilir su kaynakları yönetimi için sağladığı avantajları gözler önüne sermektedir [17].

Murugan ve çalışma arkadaşları (2023) tarafından, geliştirilen IOT tabanlı su kalitesi izleme mimarisinde suyun bulanıklığı, sıcaklığı ve su seviyesi ölçülerek analiz edilmiştir. Sensörlerin montajı için Üç Boyutlu (3D-Three Dimensional) baskı teknikleri kullanılmıştır. Çalışmada kullanılan sensörler ve modüller arasında bulanıklık, su seviyesi ve sıcaklık sensörleri ile arduino, raspberry pi ve zaman modülü bulunmaktadır. Sensörler arduino'ya bağlanmış, arduino ise raspberry pi ile entegre edilmiştir. Veriler, iki saatlik aralıklarla toplanarak işlenmiştir. Sensörlerden elde edilen veriler önce arduino'ya, ardından raspberry pi'ye aktarılmış ve raspberry pi üzerinden bulutta bulunan veri tabanına kaydedilmiştir. Sistem için gerekli enerji, kurulan güneş enerji sistemi aracılığıyla sağlanmıştır. Sensörlerden alınan veriler ile gerçek ölçümler arasındaki doğruluk oranı $R^2 > 0.97$ olarak tespit edilmiştir [18].

Islam ve Asaduzzaman (2021) tarafından, su kalitesinin izlenmesi ve iyileştirilmesine yönelik IoT ve makine öğrenmesi tabanlı bir sistem geliştirilmiştir. Çalışmada, su deposundaki su kalitesi ph, sıcaklık ve bulanıklık sensörleri aracılığıyla ölçülmüş; elde edilen veriler Dünya Sağlık Örgütü (DSÖ)'nün belirlediği referans aralıklarla karşılaştırılmıştır. Su kalitesinin düşük olduğu durumlarda sistem tarafından uyarı verilmiş ve depo, motora entegre edilmiş fırça mekanizmasıyla otomatik olarak temizlenmiştir. Ayrıca, mesafe ve akış sensörleri yardımıyla su seviyeleri izlenmiş; deponun boş olması durumunda su pompası devreye girerek otomatik dolum gerçekleştirilmiştir. Küresel Konumlama Sistemi (GPS-Global Positioning System) modülü ile konum bilgisi alınmış ve belirlenen kişilere SMS yoluyla bildirim gönderilmiştir. Sistemin geliştirilmesinde nodemcu esp8266, arduino uno r3, sx1278 lora modülü, ds18b20 sıcaklık sensörü, sen0161 ph sensörü, sen0189 bulanıklık sensörü ve

ks0429 Toplam Çözünmüş Katılar (TDS-Total Dissolved Solids) sensörü gibi donanımlar kullanılmıştır. Sensörlerden elde edilen yaklaşık 6000 kayıtlık veri setine ön işleme ve z-skore normalizasyonu uygulanmış, ardından WQI hesaplanmıştır. WQI değerleri üzerinden Su Kalitesi Sınıflandırması (WQC-Water Quality Classification) tahmin edilmiş ve bu süreçte Uzun Kısa Süreli Bellek Ağı (LSTM - Long Short-Term Memory), SVR, Çok Katmanlı Algılayıcı (MLP - Multi-Layer Perceptron) ve Doğrusal Olmayan Otoregresif Sinir Ağı (NARNet - Nonlinear Autoregressive Neural Network) gibi regresyon; SVM, Aşırı Gradyan Artırımı (XGBoost-Extreme Gradient Boosting), Karar Ağacı (DT-Decision Tree) ve Rassal Orman (RF-Random Forest) gibi sınıflandırma algoritmalarından yararlanılmıştır. Model performansları Ortalama Mutlak Hata (MAE-Mean Absolute Error), Ortalama Kare Hata (MSE-Mean Squared Error), Kök Ortalama Kare Hata (RMSE-Root Mean Squared Error), Belirleme Katsayısı (R^2 -Coefficient of Determination), Doğruluk (Accuracy-Accuracy Score), Kesinlik (Precision - Precision Score), Duyarlılık / Geri Çağırma (Recall-Recall Score) ve F1 Skoru (F1-score) gibi metriklerle değerlendirilmiş; geçerlilik analizinde k-fold çapraz doğrulama yöntemi kullanılmıştır [1].

Rajabi ve Moghaddasi (2023) tarafından geliştirilen çalışmada, Derin Q-Ağları (DQN-Deep Q-Network) algoritması ile desteklenen IoT tabanlı bir su kalitesi yönetim sistemi geliştirilmiştir. Sistem; klorür, sülfat, nitrat, nitrit, ağır metaller gibi çeşitli su parametrelerini algılayan sensörlerden oluşmakta ve bu sensörler aracılığıyla toplanan veriler, kısa (Zigbee, BLE- Bluetooth Low Energy) ve Düşük Güçlü Geniş Alan Ağı (LoRaWAN-Long Range Wide Area Network), Dar Bant Nesnelerin İnterneti (NB-IoT-Narrowband Internet of Things) iletişim teknolojileri ile merkezi sisteme iletilmektedir. Toplanan veriler DQN algoritması ile analiz edilerek sistemin mevcut durumu değerlendirilmekte, ardından “uyarı gönderme”, “su akışını kesme” ya da “mevcut durumu sürdürme” gibi aksiyonlar alınmaktadır. Sistem, geçmiş verileri öğrenerek gelecekteki kirlilik durumlarını öngörme ve proaktif müdahalede bulunma yeteneğine sahiptir. 1000 iterasyonluk eğitim süreci sonunda modelin doğruluğu %83, precision %87, recall %80 ve F1 skoru %84 olarak raporlanmıştır. Ortalama yanıt süresi 4.32 birim, kullanılabilirlik skoru ise 0.893 olarak ölçülmüştür. Geliştirilen sistem, dijital dönüşüm sürecinde organizasyonel su güvenliğini artırmaya yönelik yenilikçi ve akıllı bir çözüm sunmaktadır [19].

Yousif ve arkadaşları (2022), yüzey sularında yüksek frekansla ölçüm yapan optik

sensörlerden elde edilen verilerde çevresel etkenlerle oluşan anomalileri (örneğin, buz oluşumu, yaprak veya çöp birikimi) otomatik olarak tespit edip düzeltmek için yenilikçi ve erişilebilir bir yöntem geliştirmiştir. Veriler, Kanada'nın Ontario eyaletinde yürütülen çoklu havza besin maddesi çalışması kapsamında, 2015–2019 yılları arasında 10 farklı tarımsal alanda kurulan sensör istasyonlarından toplanmıştır. Ölçümler 15 dakikalık aralıklarla gerçekleştirilmiş ve bulanıklık ile yağış ve akış verileri de kaydedilmiştir. Geliştirilen yaklaşım üç aşamadan oluşmaktadır: otomatik anomali tespiti ve silme, kullanıcıya açık manuel doğrulama, ve eksik verilerin modellemeyle doldurulması. Otomatik algoritma, zaman serilerindeki otokorelasyon yapısından yararlanarak ani ve gerçek dışı değişimleri tespit ederken, küçük veri boşluklarını yerel doğrusal regresyonla, daha büyük boşlukları ise Çok Değişkenli Zincirleme İmputasyon (MICE-Multiple Imputation by Chained Equations) yöntemiyle ve özellikle Sınıflandırma ve Regresyon Ağacı (CART-Classification and Regression Tree) modeliyle doldurmuştur. Yapay zeka açısından algoritmanın başarısı; Matthews Korelasyon Katsayısı (MCC-Matthews Correlation Coefficient) (0.72–0.94), Yanlış Keşif Oranı (FDR-False Discovery Rate), Yanlış Negatif Oranı (FNR-False Negative Rate) ve Pearson's r gibi metriklerle değerlendirilmiş ve sentetik anomalilerin %99'una kadar başarılı şekilde müdahale edilebildiği görülmüştür. Kullanıcı deneyimleri, uzman kişilerce yapılan manuel doğrulamanın özellikle yanlış pozitif oranını düşürerek sonuçları iyileştirdiğini ortaya koymuştur. Tüm bu süreçler, "Turbidity Cleaner" adlı açık kaynaklı ve kullanıcı dostu bir yazılıma entegre edilmiştir. Sonuçlar, geliştirilen yöntemin yalnızca bulanıklık değil, yüksek frekansta ölçüm yapan diğer çevresel sensörler için de genellenebilecek potansiyele sahip olduğunu göstermektedir [20].

Amador-Castro ve arkadaşları (2024) tarafından yapılan çalışmada, geleneksel su kalitesi izleme yöntemlerinin sınırlılıklarına alternatif olarak sunulan IoT tabanlı sistemler ve vatandaş bilimi uygulamaları kapsamlı biçimde incelenmiştir. IoT sistemlerinin, sıcaklık, pH, bulanıklık, çözünmüş oksijen ve iletkenlik gibi temel parametreleri düşük maliyetle ve gerçek zamanlı olarak izleyebildiği; bazı örneklerde makine öğrenmesi destekli yazılım sensörleriyle Biyokimyasal Oksijen İhtiyacı (BOD-Biochemical Oxygen Demand) ve amonyak gibi doğrudan ölçülemeyen parametrelerin yüksek doğrulukla ($R^2 > 0.78$) tahmin edilebildiği vurgulanmıştır. Ancak sensör kirliliği, düzenli kalibrasyon ihtiyacı, siber güvenlik açıkları ve sınırlı veri doğrulama uygulamaları sistemin yaygınlaşmasını zorlaştırmaktadır. Vatandaş bilimi projelerinde ise gönüllü katılımcıların

basit test kitleri ve taşınabilir sensörlerle ph, nitrat, fosfat ve bulanıklık gibi parametreleri ölçerek geniş alanlarda veri sağladıkları belirtilmiştir. Veri kalitesi, gönüllülerin eğitimi ve uzman doğrulaması ile artırılabilirken; örnekleme önyargısı ve katılımcı bağlılığı gibi sosyal etkenler projelerin sürdürülebilirliğini etkilemektedir. Her iki yaklaşımın da su kalitesine dair zaman ve mekâna yayılmış veriler sağlayarak geleneksel yöntemleri tamamladığı; ayrıca halkın farkındalığını artırarak sürdürülebilir su yönetimi hedeflerine katkı sunduğu sonucuna ulaşılmıştır [21].

Boujelben ve arkadaşları (2023) tarafından yapılan çalışmada, su dağıtım ağlarındaki kaçakları gerçek zamanlı ve yüksek doğrulukla tespit edip konumlandırabilen, IoT ve derin öğrenme tabanlı entegre bir sistem önerilmiştir. Boru yüzeyine yerleştirilen mikrofonlarla toplanan akustik veriler, Arduino tabanlı IoT cihazlarında çalışan hafif Bir Boyutlu Evrişimli Sinir Ağı (1D-CNN-One-Dimensional Convolutional Neural Network) modeliyle analiz edilerek sızıntılar tespit edilmektedir. Sızıntı bilgileri, enerji verimli LoRaWAN teknolojisiyle uzak bir sunucuya iletilmekte ve burada ikinci bir 1D-CNN modeliyle sızıntının tam konumu belirlenmektedir. Sistem, %97 doğrulukla sızıntı tespiti ve X/Y eksenlerinde sırasıyla 0.0006/0.0004 hata payı ile konumlandırma başarısı göstermiştir. Hafifliği ve düşük enerji tüketimi sayesinde sistem, özellikle su kıtlığı yaşayan bölgeler için etkili ve uygulanabilir bir çözüm sunmaktadır [22].

Jin ve arkadaşları (2025) tarafından yapılan çalışmada, içme suyu kalitesinin halk sağlığı üzerindeki etkisini dikkate alarak yenilikçi bir makine öğrenimi tabanlı İçme Suyu Kalite İndeksi (DWQI-Drinking Water Quality Index) modeli geliştirilmiştir. Bu model, sağlık faktörlerini değerlendirme sistemine entegre eden ilk sistemlerden biridir. Araştırmada, su kalitesini etkileyen 16 temel parametre k-means kümeleme algoritması ile belirlenmiş, ağırlık optimizasyonu için dört farklı yöntem test edilmiş ve Analitik Hiyerarşi Prosesi (AHP-Analytic Hierarchy Process) ile RF yöntemlerinin birleşiminden oluşan CWAR yöntemi en etkili çözüm olarak tanımlanmıştır. Model, 2018–2023 yılları arasında Çin'in Zhejiang eyaletindeki içme suyu arıtma tesislerinden elde edilen veriler üzerinde uygulanmış ve DWQI değerlerinin %90,93'ünün Çin'in su kalitesi, kalite sınıfı 2 standardını karşılamıştır. Ayrıca, içme suyu kalitesi ile sosyoekonomik faktörler arasında anlamlı ilişkiler tespit edilmiştir (örneğin, birincil sanayi oranı ile negatif, şehir sakinlerinin gelir seviyesi ile pozitif korelasyon). Gelişmiş arıtma yöntemlerinin geleneksel ve karışık yöntemlere göre daha iyi su kalitesi sağladığı gösterilmiştir. Bu çalışmayla sunulan model, içme suyu yönetimi ve politika geliştirme süreçlerine bilimsel

katkılar sunarak, toplum sađlıđının iyileřtirilmesine y6nelik 6nemli bir ara olarak deđerlendirilmektedir [23].

Li ve arkadaşları (2024) tarafından gerekleřtirilen derleme alıřmasında, ime suyu dađıtım sistemlerinde Makine 6đrenmesi (ML-Machine Learning) uygulamalarının su kalitesi y6netiminde ok boyutlu faydalar sađladıđı ortaya konmuřtur. ML y6ntemlerinin, 6zellikle rezid6el klor, Dezenfeksiyon Yan 6r6nleri (DBPs-Disinfection By-Products), mikrobiyal g6stergeler, ađır metaller ve duyuasal 6zelliklerin tahmininde geleneksel modellere kıyasla daha y6ksek dođruluk sađladıđı vurgulanmıřtır. Anomali tespiti konusunda ise g6zetimli, yarı-g6zetimli ve g6zetimsiz modeller incelenmiř; Yapay Sinir Ađı (ANN-Artificial Neural Network), SVM ve derin 6đrenme (LSTM, 6retici ekiřmeli Ađlar (GAN-Generative Adversarial Network), Graf Sinir Ađı (GNN - Graph Neural Network) gibi) tabanlı modellerin, su kalitesinde ani deđiřimleri ok parametrelili verilerle bařarıyla belirleyebildiđi g6r6lm6řt6r. Ayrıca ML, kontaminasyon kaynaklarının hızlı ve dođru řekilde tespitinde y6ksek bařarı oranı g6stermiř; 6zellikle grafik sinir ađları (GNN) ile 952 d6đ6ml6 bir ađda yalnızca 33 sens6rle %92'nin 6zerinde dođrulukla kaynak yerinin belirlenebildiđi rapor edilmiřtir. Bununla birlikte, veri eksikliđi, model genelleme kapasitesi, aıklanabilirlik ve gerek zamanlı uygulama gibi zorluklar devam etmektedir. alıřma, ML'nin ime suyu dađıtım sistemlerinde erken uyarı, proaktif y6netim ve sistem optimizasyonu aısından 6nemli bir potansiyel tařıdıđını vurgulamakta; gelecekte ok kaynaklı veri b6t6nleřtirme ve transfer 6đrenme gibi stratejilerin 6ne ıkacađını 6ng6rmektedir [24]. Bu yaklařımlar sayesinde, su kalitesindeki ani deđiřimlerin erken tespiti m6mk6n olmakta, b6ylece olası halk sađlıđı tehditlerinin 6n6ne geilebilmektedir.

1.2. TEZİN AMACI

Bu tez alıřmasının temel amacı, ime suyu arıtma tesislerinde yaygın olarak kullanılan jar testi s6recini optimize ederek kimyasal dozajlamada maksimum verimliliđi sađlamaktır. Bu kapsamda, geleneksel y6ntemlerin zaman alıcı ve operat6r bađımlı yapısından kaynaklanan belirsizlikleri ortadan kaldırmak amacıyla makine 6đrenmesi algoritmaları kullanılmıřtır. Tesiste kurulan IoT tabanlı akıllı su analiz sistemi ile pH, bulanıklık, sıcaklık, iletkenlik gibi su kalitesi parametreleri anlık olarak izlenmiřtir. Deđerlerde ani deđiřimler meydana geldiđinde anında bildirim g6ndererek hızlı bir řekilde m6dahale imkanı sunmuřtur. Bu entegre yapı sayesinde, arıtma s6recinin optimizasyonu daha hızlı, daha ekonomik ve s6rd6r6lebilir hale getirilmiřtir.

1.3. LİTERATÜRE KATKI

Literatürde jar testi genellikle laboratuvar ortamında manuel olarak gerçekleştirilen bir işlem olarak ele alınmakta ve bu süreçte deneyime dayalı optimizasyon ön plana çıkmaktadır. Son yıllarda, bu sürecin otomasyonu üzerine bazı çalışmalar yapılmış olsa da, genellikle IoT sensörleriyle veri toplanması ve temel karar destek sistemlerinin oluşturulması çalışmaları yapılmıştır. Yapılan araştırmalarda jar testinin zaman alması nedeniyle testlerin her vardiyada 1 kez olacak şekilde yapıldığı anlaşılmıştır. Jar testi sayısının artırılması ve kimyasal dozajının takip edilip düzenlenmesi sınırlı kaldığı görülmüştür.

Bu çalışma, IoT destekli jar testi sistemini hem fog hem de bulut katmanları üzerinden çalışan, makine öğrenimi ile optimizasyon yapan bütünleşik bir mimariyle sunarak önemli bir boşluğu doldurmayı hedeflemektedir. Fog mimarisi sayesinde karar destek süreçlerinin düşük gecikmeyle yerel olarak çalışması sağlanırken, bulut katmanında sürekli model güncellemeleri ve uzun vadeli veri analizi mümkün olacaktır. Böylece, hem anlık arıtım kararları optimize edilecek, hem de tarihsel verilerle daha doğru ve adaptif modeller geliştirilebilecektir.

Ayrıca, sistemde kullanılacak makine öğrenmesi algoritmaları, geleneksel tek parametrelili eşik değer yaklaşımının ötesine geçerek çok değişkenli analizlerle en uygun kimyasal madde dozajının belirlenmesine katkı sağlayacaktır. Bu yönüyle çalışma, su arıtma süreçlerinde daha az kimyasal madde kullanımını ve daha yüksek arıtma verimliliğini hedefleyerek sürdürülebilirliğe katkı sunacaktır.

Sonuç olarak, önerilen sistemin;

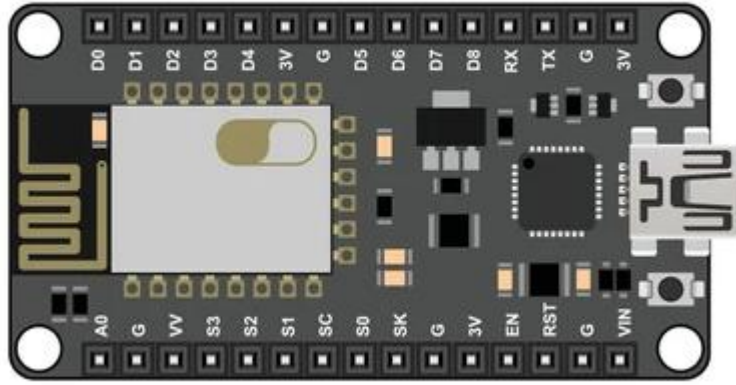
- Gerçek zamanlı ve düşük gecikmeli karar verme kabiliyeti,
- Bulut destekli uzun vadeli öğrenme yeteneği,
- Sürdürülebilir kimyasal kullanım stratejileri

ile literatürdeki mevcut yaklaşımları bir araya getirip önemli bir yenilik sunmuştur.

2. MATERYAL

2.1. DONANIMLAR

2.1.1. Nodemcu Esp8266

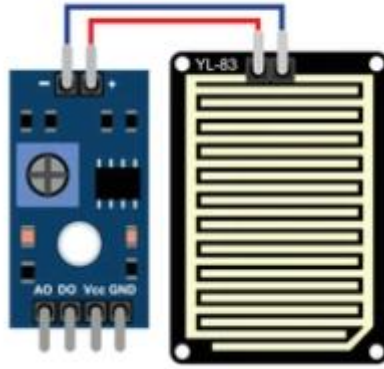


Şekil 2.1. Nodemcu esp8266 [25].

Nodemcu, esp8266 WiFi çipi için geliştirilmiş açık kaynaklı bir firmware ve geliştirme platformudur. LUA programlama diline dayalıdır, ancak Arduino Entegre Geliştirme Ortamı (IDE-Integrated Development Environment) gibi farklı programlama ortamlarında da C/C++ ile kullanılabilir. Geliştirme kartı, Espressif Systems tarafından geliştirilen düşük maliyetli ve İletim Denetim Protokolü / İnternet Protokolü (TCP/IP-Transmission Control Protocol / Internet Protocol) protokolü destekli esp8266 Wi-Fi çipini temel alır. Nodemcu, Seri Çevrimli Flash Dosya Sistemi (SPIFFS-SPI Flash File System) ile dosya yönetimi sağlar ve Espressif'in İşletim Sistemless (NON-OS - Non-Operating System) Yazılım Geliştirme Kiti (SDK-Software Development Kit) üzerine kuruludur. Modül, micro USB portu üzerinden beslenir ve dahili 3.3V regülatör ile çalışır. Teknik olarak 2.5V–3.6V arası çalışır, 80mA akım çeker ve 20µA uyku modu akımı ile enerji verimlidir. Donanım olarak, 32KB RAM, 80KB DRAM, 200KB flash bellek içerir. Nodemcu, toplam 17 Genel Amaçlı Giriş/Çıkış (GPIO-General Purpose Input/Output) pinine, 1 adet 10-bit ADC kanalına, iki UART, SPI, I2C, I2S, PWM desteklerine sahiptir. GPIO pinleri çoklama özelliği ile çok amaçlı kullanılabilir. Ayrıca, sıfırlama (RST) ve

program yükleme (FLASH) düğmeleri ile kullanıcıya programlanabilir bir LED sunar. Yüksek donanım entegrasyonu ve Wi-Fi özellikleri sayesinde, düşük maliyetli ve esnek bir IoT çözümüdür [26].

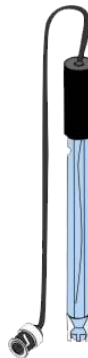
2.1.2. Yağmur Sensörü



Şekil 2.2. Yağmur sensörü [27].

Yağmur damlası sensörü, çevredeki yağmurun varlığını ve yoğunluğunu algılamak için kullanılır. Baskılı Devre Kartı (PCB-Printed Circuit Board) üzerine yerleştirilen çizgili plakaya su değmesi ile birlikte ortamda yağmur olduğu bilgisi alınabilmektedir. Kart üzerindeki suyun yoğunluğuna göre yağmurun yoğunluğu tespit edilebilmektedir [28]. Bu kart hem dijital hem de analog olarak bilgi verebilmektedir. Örneğin yağmur var, yağmur yok şeklinde bilgi alınabilirken, yağmur yoğunluğu 400 şeklinde bilgi alınabilmektedir. Analog sinyalde 0-1024 aralığında ölçüm yapabilmektedir.

2.1.3. Ph Sensörü



Şekil 2.3. Ph sensörü [29].

pH sensör probu, çözeltilerdeki hidrojen iyonu (H^+) konsantrasyonunu ölçen elektrokimyasal bir ölçüm cihazıdır ve genellikle bir cam elektrot ile referans elektrotun birleşiminden oluşur. Propta bulunan ince cam membran, hidrojen iyonlarına duyarlıdır ve çözeltideki iyon derişimine göre bir potansiyel farkı oluşturur. Bu fark, referans elektrotla karşılaştırılarak ölçülür ve pH değeri hesaplanır. pH probu genellikle 0 ile 14 pH aralığında çalışır ve ± 0.01 pH hassasiyetindedir. Standart ölçüm sıcaklığı aralığı $0^{\circ}C$ ila $60^{\circ}C$ olmakla birlikte, bazı endüstriyel modeller daha yüksek sıcaklıklara dayanabilir. Probun içinde yer alan referans elektrot çoğunlukla Ag/AgCl (gümüş/gümüş klorür) türünde olup KCl (potasyum klorür) elektrolitiyle doldurulmuştur. pH problemleri cam, epoksi veya PVDF gibi farklı malzemelerden üretilir; bu da kullanım alanına göre (laboratuvar, çevre analizi, endüstri) değişkenlik gösterir. Bağlantı uçları genellikle BNC tipi olup, bazı modellerde analog yerine dijital çıkışlar da bulunabilir. Kullanım sırasında probun ucunun kuru kalmaması son derece önemlidir; çünkü cam membran kurduğunda ölçüm hassasiyeti düşer. Bu nedenle prob, özel depolama solüsyonlarında muhafaza edilmelidir. Ayrıca ölçüm doğruluğu için belirli aralıklarla pH 4.01, 7.00 ve 10.00 buffer çözeltileriyle kalibrasyon yapılması gerekir [30].

2.1.4. Su Sıcaklık Sensörü

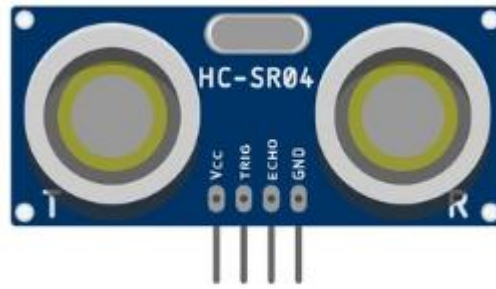


Şekil 2.4. Su sıcaklık sensörü [31].

DS18B20 su geçirmez sıcaklık sensörü, $-55^{\circ}C$ ile $+125^{\circ}C$ arasında sıcaklık ölçümleri yapabilen dijital bir sensördür. Maksimum hassasiyeti, $-10^{\circ}C$ ile $+85^{\circ}C$ arasında $\pm 0.5^{\circ}C$ iken, diğer sıcaklık aralıklarında $\pm 2^{\circ}C$ hassasiyet sunar. Bu sensör, 9 bit ile 12 bit arasında

ayarlanabilen çözünürlük seçeneklerine sahiptir ve 3.0V ile 5.5V arasında bir çalışma voltajı ile enerji sağlar. 1-Wire iletişim protokolü sayesinde, birden fazla sensör aynı veri hattı üzerinden bağlanabilir, bu da sistemlerin daha verimli ve düzenli olmasını sağlar. Su geçirmez metal muhafazası, sensörü zorlu çevre koşullarında, özellikle su altı ve nemli ortamlarda kullanıma uygun hale getirir. Ayrıca, üst ve alt sıcaklık limitlerine göre alarm fonksiyonu da ayarlanabilir. DS18B20 sensörü, hidroponik sistemlerden endüstriyel sıcaklık izleme uygulamalarına kadar geniş bir kullanım alanına sahiptir [31].

2.1.5. Su Seviye Sensörü



Şekil 2.5. Su seviye sensörü [32].

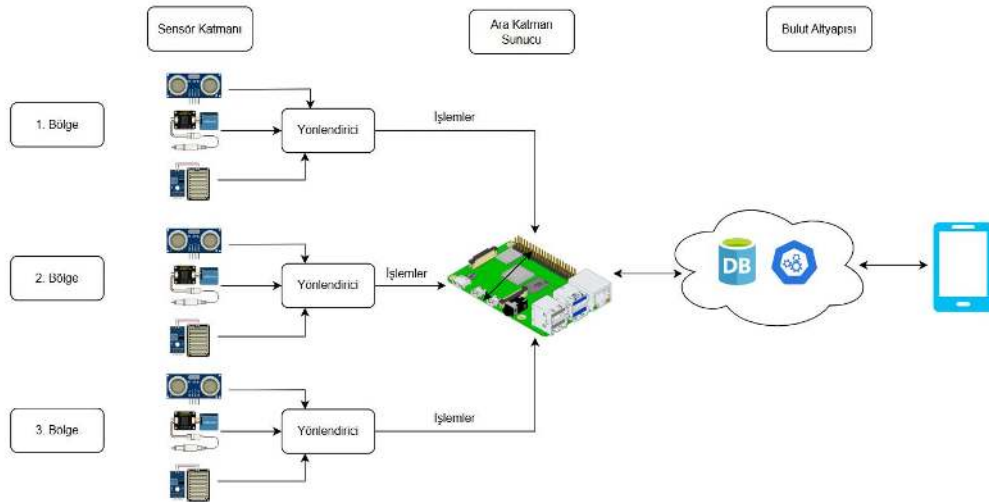
HC-SR04 ultrasonik sensör, ses dalgalarının yankılanma süresini ölçerek mesafe tespiti yapan temasız algılayıcılardır. Bu sensörler, genellikle 20 kHz'in üzerindeki frekansta çalışan piezoelektrik dönüştürücüler vasıtasıyla ses dalgaları üretir ve bu dalgaların nesnelere çarpıp geri dönmesiyle oluşan yankıyı algılar. Gönderilen ve alınan sinyal arasındaki süre, ölçülen mesafe ile doğrudan ilişkilidir. Bu prensip, özellikle sıvı seviyesinin belirlenmesinde, hareketli nesnelerin tespitinde ve yapısal izleme uygulamalarında yaygın olarak kullanılmaktadır.

Son yıllarda IoT destekli akıllı izleme sistemlerinin gelişmesiyle birlikte ultrasonik sensörler, meteorolojik verilerin toplanmasında önemli bileşenler haline gelmiştir. Özellikle yağmur miktarının ölçülmesi, su seviyesi izleme, toprak nemi ve çevresel engellerin tespiti gibi alanlarda düşük maliyetli, enerji verimli ve güvenilir bir çözüm sunmaktadır [28].

Raspberry Pi 4, küçük boyutlu olmasına rağmen güçlü bir işlem kapasitesi sunan, çok amaçlı bir mini bilgisayardır. 1.5 GHz hızında çalışan dört çekirdekli ARM Cortex-A72 işlemcisi, ona yüksek performans sağlar. Farklı RAM seçenekleri (2 GB, 4 GB ve 8 GB LPDDR4) sayesinde hem basit hem de yoğun veri işleme uygulamalarına uygundur. Kablosuz iletişim için 802.11ac Wi-Fi ve Bluetooth 5.0 desteklerken, ağ bağlantısı için Gigabit Ethernet portuna sahiptir. Dört adet USB portu (ikisi USB 3.0, ikisi USB 2.0) ve çift mikro HDMI çıkışı ile çoklu cihaz desteği sunar. Besleme girişi USB Type-C üzerinden sağlanmakta ve 5V/3A güç kaynağı kullanılması önerilmektedir.

Raspberry Pi 4'ün sahip olduğu 40 pinlik GPIO arayüzü, ona fiziksel cihazlarla doğrudan iletişim kurma imkanı verir. Bu pinler üzerinden sensörler, motorlar, ekranlar ve çeşitli diğer bileşenler kolaylıkla bağlanabilir. GPIO üzerinden UART, SPI ve I2C gibi iletişim protokollerinin desteklenmesi, çok sayıda sensör ve modülün eş zamanlı kullanımına olanak tanır [35].

2.2. Fog Bilişim



Şekil 2.8. Fog bilişim mimarisi.

Fog bilişim, nesnelerin interneti (IoT) cihazları tarafından üretilen verilerin, doğrudan merkezi bulut sistemlerine gönderilmeden önce, ağın uç noktalarına yakın yerel sunucular, gateway'ler veya edge cihazlar üzerinde işlenmesini, analiz edilmesini ve geçici olarak depolanmasını sağlayan mimaridir. Bu yapı, özellikle zaman duyarlı uygulamalarda ağ gecikmesini azaltmak, bant genişliği kullanımını optimize etmek ve veri güvenliğini artırmak amacıyla geliştirilmiştir. Verinin doğrudan buluta iletilmesi

yerine yerel ortamda işlenmesine imkan tanıyan fog bilişim, gerçek zamanlı tepki kabiliyeti, sistem otonomisi ve yüksek ölçeklenebilirlik sunarak kritik uygulamalarda hızlı ve güvenilir çözümler sağlar. Akıllı su yönetimi, akıllı şehir altyapıları, endüstriyel otomasyon sistemleri, sağlık hizmetleri ve ulaşım ağları gibi birçok alanda etkin bir şekilde kullanılan fog bilişim, aynı zamanda bulut bilişim altyapılarını tamamlayıcı bir katman olarak uçtan uca bilgi işlem süreçlerinde hız, esneklik ve güvenlik avantajı sağlamaktadır[36].

2.2.8.1. Mosquitto

Mosquitto, açık kaynaklı bir MQTT (Message Queuing Telemetry Transport) broker yazılımıdır. Bu sürüm, IoT (Nesnelerin İnterneti) gibi düşük bant genişliği ve yüksek gecikme süresi gereksinimleri olan ortamlarda veri iletimi için tasarlanmıştır. Mosquitto, MQTT protokolü üzerinden cihazlar arasında publish/subscribe (yayınla/abone ol) modelini kullanarak, veri iletimini güvenli ve verimli bir şekilde gerçekleştirir. Mosquitto 1.4.8, SSL/TLS şifrelemesi ile kimlik doğrulama ve yetkilendirme gibi güvenlik özellikleri sunar ve aynı zamanda sınırlı kaynaklara sahip cihazlarda bile yüksek performans sağlar. Bu özellikleriyle, IoT uygulamalarında ve düşük maliyetli cihazların iletişiminde sıklıkla tercih edilir [37].

2.2.8.2. Node-RED

Node-RED, açık kaynaklı bir görsel programlama platformudur ve özellikle nesnelerin interneti (IoT) projelerinde veri akışlarını görsel olarak yönetmeye olanak tanır. Kullanıcılar, sürükle ve bırak yöntemiyle çeşitli cihazlardan ve servislerden gelen verileri işleyebilir, analiz edebilir ve yönlendirebilir. JavaScript tabanlı olarak geliştirilmiş olan Node-RED, kullanıcıların yazılım geliştirmeden de akışlar oluşturmalarını sağlayan basit ve sezgisel bir arayüz sunar. Bu platform, MQTT, HTTP, WebSocket gibi popüler iletişim protokollerini destekler ve farklı cihazlar arasındaki veri iletimini kolaylaştırır. Genişletilebilir yapısı sayesinde kullanıcılar, ihtiyaçlarına göre özel düğümler ve işlevler ekleyebilir. Node-RED, IoT uygulamaları, veri entegrasyonu, otomasyon ve bulut tabanlı sistemlerde yaygın olarak kullanılır [38].

2.3. Bulut Sistemi

Bulut bilişim (cloud computing), donanım ve yazılım kaynaklarının internet üzerinden isteğe bağlı olarak sağlandığı, esnek ve ölçeklenebilir bir bilgi işlem modelidir. Bu

yaklaşım, kullanıcıların fiziksel altyapıya sahip olmadan yüksek işlem gücüne, veri depolama alanına ve yazılım hizmetlerine erişmesini mümkün kılar. Bulut bilişim yapısı dağıtık sistem bileşenlerinin merkezi olarak yönetilmesine, verilerin coğrafi sınır olmaksızın erişilebilir hale gelmesine ve IoT cihazlarının ölçeklenebilirlik sorunlarının çözülmesine olanak tanımaktadır [36].

Bu bağlamda bulut bilişim, yalnızca donanım merkezileştirme değil; aynı zamanda IoT sistemlerinin işlevselliğini artıran, veri sürekliliği ve bütünlüğünü destekleyen, merkezi analiz ve yönetim altyapısı olarak konumlanmaktadır.

Bulut sisteminde kurulu olan sistemler; Docker, InfluxDB, PostgreSQL ve Grafana'dır.

2.3.1.1. Docker

Docker, uygulama dağıtım süreçlerini otomatikleştirmek amacıyla geliştirilen, Linux tabanlı konteyner teknolojisini esas alan açık kaynaklı bir platformdur. Modern yazılım geliştirme ve dağıtım sistemlerinde yaygın olarak kullanılan Docker, uygulamaları ve ilgili tüm bağımlılıklarını tek bir kapsayıcı (container) içerisinde çalıştırarak taşınabilir, tutarlı ve izole bir çalışma ortamı sağlar. Bu sayede uygulamalar, farklı donanım yapılandırmalarında ya da işletim sistemlerinde herhangi bir yapılandırma değişikliğine gerek duyulmaksızın aynı şekilde çalıştırılabilir. Docker, sanal makinelerden farklı olarak, yalnızca uygulama katmanını ve işletim sistemi düzeyindeki gerekli bileşenleri izole eder; bu da hem başlatma süresini azaltmakta hem de kaynak tüketimini minimize etmektedir. Linux çekirdeğinin sunduğu control groups (cgroups) ve namespaces gibi özellikleri kullanarak sistem kaynaklarını sınırlar ve her konteynerin birbirinden izole biçimde çalışmasını sağlar. Konteynerler, salt okunur katmanlardan oluşan ve Union File System (UnionFS) yapısını kullanan Docker imajlarından (images) türetilir. Bu yapı, yazılım bileşenlerinin katmanlı biçimde yeniden kullanılmasına ve güncellemelerin yalnızca gerekli katmanlar üzerinden yapılmasına olanak tanır [39].

2.3.1.2. InfluxDB

InfluxDB, özellikle zaman serisi veri yönetimi için optimize edilmiş, açık kaynaklı bir veritabanı sistemidir. Sensör verileri, sistem metrikleri, uygulama performans izlemeleri ve olay günlükleri gibi zaman damgası içeren verilerin yüksek hızda saklanması, işlenmesi ve sorgulanması amacıyla geliştirilmiştir. InfluxDB 1.8 sürümü, verilerin yazım ve sorgulama süreçlerinde düşük gecikme süresi ve yüksek verimlilik sunarak,

büyük hacimli zaman serisi verileriyle çalışmaya olanak tanır. SQL benzeri bir sorgulama dili olan InfluxQL kullanarak veri analizi yapılmasını sağlar. Veri noktaları ölçüm (measurement), etiketler (tags) ve alanlar (fields) yapısında organize edilir; bu da hızlı indeksleme ve esnek sorgulama imkanları sunar. Özellikle IoT uygulamaları, enerji izleme sistemleri, finansal veri analizi ve altyapı performans takibi gibi alanlarda geniş kullanım alanı bulmuştur [40].

2.3.1.3. *PostgreSql*

PostgreSQL, nesne-ilişkisel bir veritabanı yönetim sistemi (ORDBMS) olarak, geleneksel ilişkisel veritabanı modellerini genişleten açık kaynak kodlu ve güçlü bir çözümdür. ANSI SQL standardına tam uyumlu olmasının yanı sıra, gelişmiş veri türleri, kullanıcı tanımlı fonksiyonlar, tetikleyiciler (triggers) ve prosedürel uzantılar gibi özelliklerle fonksiyonel genişletilebilirlik sağlar. Transaction yönetimi açısından ACID (Atomicity, Consistency, Isolation, Durability) ilkelerini tam olarak desteklemesi, onu yüksek güvenilirlik ve veri bütünlüğü gerektiren sistemler için uygun hâle getirmektedir. PostgreSQL'in en belirgin teknik avantajlarından biri, Write-Ahead Logging (WAL) mekanizmasıdır. Bu yapı sayesinde, veritabanı üzerindeki tüm değişiklikler kalıcı hâle getirilmeden önce loglanır ve sistemin tutarlılığı güvence altına alınır. Ayrıca, PostgreSQL, logical replication ve physical replication olmak üzere iki farklı veri çoğaltma (replication) yöntemini destekleyerek hem ölçeklenebilirlik hem de yüksek erişilebilirlik sağlar. Çoklu tablo yapıları, karmaşık sorgu yetenekleri ve geniş uzantı (extension) desteği sayesinde PostgreSQL, sadece geleneksel veri tabanı çözümleri değil; aynı zamanda gerçek zamanlı, dağıtık ve olay güdümlü veri mimarilerinde de etkili bir bileşen olarak kullanılmaktadır [41].

2.3.1.4. *Grafana*

Grafana, açık kaynaklı bir veri görselleştirme ve izleme platformudur. Kullanıcıların farklı veri kaynaklarından (örneğin InfluxDB, Prometheus, PostgreSQL gibi) gerçek zamanlı verileri çekerek, interaktif grafikler, tablolar ve panolar oluşturmaya olanak tanır. 10.3.1 sürümü, dashboard izin yönetimi, veri kaynağı bağlantıları ve sorgu editörü üzerinde yapılan geliştirmelerle kullanıcı deneyimini iyileştirmiştir. Grafana, özellikle IT altyapı izleme, IoT cihaz verisi analizi ve sistem performansı takibi gibi uygulamalarda yaygın olarak kullanılmaktadır [42].

2.4. MAKİNE ÖĞRENMESİ

Makine öğrenmesi, veri setlerinden öğrenerek tahmin ve sınıflandırma gibi görevleri otomatikleştiren, istatistiksel temelli bir yapay zekâ alt dalıdır. Bu yaklaşımda sistemler, açık biçimde programlanmaksızın veriler içerisindeki örüntüleri tanıyarak karar alma süreçlerini yürütür. Makine öğrenmesi için ilk olarak veriseti belirlenir ardından veri analiz edilir, aykırı değer yada boş değerler varsa bunlar temizlenir. Bu bölüm veri ön işleme olarak adlandırılır. Veri ön işleme tamamlandıktan sonra korelasyon analizi ve özellik önem sıralaması (feature importance) teknikleri sayesinde tahmin sürecine en çok katkı sunan değişkenler belirlenebilir. Sonrasında veriseti eğitim ve test olmak üzere ikiye bölünür. Eğitim veriseti üzerinde model eğitilir, test verisetiyle sonuçlar eğitimdeki sonuç değerleriyle karşılaştırılır. Bu karşılaştırma sonucu makine öğrenmesi modelinin başarı metrikleri hesaplanır [11].

Sınıflandırma problemlerinde model başarıları doğruluk (accuracy), AUC-ROC ve F1-score gibi metriklerle değerlendirilir [11]. Regresyon problemlerinde ise hata oranlarını ölçmek amacıyla ortalama mutlak hata (MAE), ortalama mutlak yüzde hatası (MAPE), ortalama kare hata (MSE), kök ortalama kare hata (RMSE) ve determinasyon katsayısı (R^2) gibi istatistiksel ölçütler kullanılmaktadır [7]. Bu metrikler, modelin tahmin gücünü nesnel biçimde değerlendirmeye olanak sağlar. Parametre ayarlarının model performansı üzerindeki etkisini optimize etmek için yaygın olarak GridSearchCV ve benzeri hiperparametre optimizasyon yöntemleri uygulanır.

Bu bütünsel yaklaşım sayesinde, hem sınıflandırma hem de regresyon problemleri için güçlü ve güvenilir tahmin modelleri elde edilebilmekte; makine öğrenmesi tabanlı yöntemler su kalitesi değerlendirme süreçlerinde etkin dijital çözümler olarak değerlendirilmektedir.

2.4.1. Veri Ön İşleme

Makine öğrenmesi modelinde, ilk işlem olarak veri ön işleme yapılmalıdır. Veri ön işleme, veri setindeki aykırı değerlerin, tekrarlanan kayıtların, eksik kayıtların ve boş kayıtların veri setinden temizlenmesi işlemidir. Ön işleme sonrasında makine öğrenmesi modelinin performansı ve tahmin doğruluğu artmaktadır [43].

2.4.2. Model

Makine öğrenmesi, regresyon ve sınıflandırma problemlerinin çözümünde çeşitli algoritmaların uygulanmasıyla önemli başarılar elde eden bir alandır. Bu tez kapsamında, farklı teorik temellere ve öğrenme yaklaşımlarına sahip çeşitli makine öğrenmesi modelleri değerlendirilmiştir. Bu modeller, doğrusal ve doğrusal olmayan regresyon yöntemlerinden başlayarak topluluk (ensemble) yöntemlerine kadar geniş bir yelpazeyi kapsamaktadır. Her modelin farklı veri yapılarında ve problem tiplerinde gösterdiği performans değişiklik göstermektedir. Bu nedenle, model çeşitliliği, eldeki verinin yapısal özelliklerine en uygun yöntemi belirlemek açısından önemlidir.

Kullanılan model listesi Çizelge 2.1’de belirtilmiştir.

Çizelge 2.1. Makine öğrenmesi modelleri.

Kategori	Model Adı
Doğrusal Regresyon Modelleri	Linear Regression
	Ridge Regression
	Lasso Regression
	ElasticNet
	Bayesian Ridge
	Huber Regressor
	SGD Regressor
	Passive Aggressive Regressor
	TheilSen Regressor
	RANSAC Regressor
Destek Vektör Regresyonu (SVR)	SVR
	NuSVR
	LinearSVR
Karar Ağaçları ve Ensemble Yöntemler	Decision Tree Regressor
	Random Forest Regressor
	Extra Trees Regressor
	AdaBoost Regressor
	Bagging Regressor
	Gradient Boosting Regressor
Modern Boosting Yöntemleri	HistGradientBoosting Regressor
	XGBoost
	LightGBM
	CatBoost

2.4.3. Hiperparametre

Makine öğrenmesinde hiperparametreler, bir modelin eğitimi başlamadan önce dışarıdan belirlenen ve modelin genel performansını doğrudan etkileyen ayar değişkenleridir. Bu parametreler, modelin öğrenme sürecini kontrol eder ve eğitim verisi üzerinden öğrenilmezler. Örneğin, destek vektör makinelerinde (SVM) kullanılan C ve gamma değerleri, karar ağaçlarında max_depth ya da sinir ağlarında katman ve nöron sayısı gibi parametreler hiperparametre olarak değerlendirilir. Hiperparametrelerin uygun bir şekilde seçilmesi, modelin hem eğitim verisine uyumunu hem de yeni verilere karşı genellenebilirliğini belirler. Bu nedenle, hiperparametre ayarlaması (tuning) aşaması, modelin başarısı açısından kritik bir rol oynar. Bu süreç genellikle grid search, random search, bayes optimizasyonu gibi tekniklerle gerçekleştirilir. Grid search tüm parametre kombinasyonlarını denerken, random search rastgele seçimle parametreleri test eder ve bazı çalışmalarda daha verimli olduğu görülmüştür [44].

2.4.4. Model Değerlendirmeleri

Regresyon modellerinin performansını değerlendirmek için çeşitli hata metrikleri kullanılır. MAE (Mean Absolute Error - Ortalama Mutlak Hata), tahmin edilen değerlerle gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasını verir ve modelin ortalama hata büyüklüğünü gösterir. MASE (Mean Absolute Scaled Error - Ortalama Mutlak Ölçeklenmiş Hata), farklı veri setleri arasında karşılaştırmayı mümkün kılmak için hatayı ölçeklendirir ve özellikle zaman serisi analizlerinde tercih edilir. MAPE (Mean Absolute Percentage Error - Ortalama Mutlak Yüzde Hatası), hatayı yüzde cinsinden ifade eder ancak gerçek değerlerin sıfıra yakın olduğu durumlarda yanıltıcı olabilir. MSE (Mean Squared Error - Ortalama Kare Hata), hataların karelerinin ortalamasını hesaplayarak büyük hatalara daha fazla ceza verir. RMSE (Root Mean Squared Error - Kök Ortalama Kare Hata) ise MSE'nin karekökü alınarak hata biriminin orijinal veriyle aynı olması sağlanır ve daha yorumlanabilir hale gelir. Son olarak, R^2 (R-Squared - Belirlilik Katsayısı), modelin açıklayabildiği toplam varyans oranını gösterir ve 1'e yaklaştıkça modelin veriyi daha iyi açıkladığı kabul edilir. Bu metriklerin tamamı birlikte değerlendirildiğinde, modelin hata büyüklüğü, hata dağılımı ve açıklayıcılığı hakkında kapsamlı bir anlayış sağlar ve performans karşılaştırmalarına olanak tanır [1], [11], [45], [46].

Parametrelerde tahmin ile gerçek değer arasındaki farkı bulmak için aşağıdaki formül kullanılmaktadır.

$$\text{Error} = \text{Predicted}_i - \text{Actual}_i$$

2.4.4.1. Mean Absolute Error (MAE)

MAE, Denklem 1'de gösterildiği gibi tahmin ve gerçek değer arasındaki farkın mutlak değeri alınarak her bir hata hesaplanır. Daha sonra, bu hata değerleri toplam kayıt sayısına bölünerek ortalaması alınır. Bu işlem, tüm hata değerlerinin toplanıp kayıt sayısına bölünmesiyle gerçekleştirilir. MAE değeri küçüldükçe, ortalama tahmin hatasının düşük olduğu anlamına gelmektedir [47].

$$MAE = \sum_{i=0}^n \frac{|\text{Error}_i|}{n} \quad (2.1)$$

2.4.4.2. Mean Absolute Percentage Error (MAPE)

MAPE, Denklem 5'te gösterildiği gibi gerçek değer ile tahmin edilen değerlerin hata oranını hesaplamaktadır. Bu değer 0'a yakın oldukça başarı oranı artmaktadır [2].

$$MAPE(\%) = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - y_p}{y_i} \right| \quad (2.2)$$

2.4.4.3. Mean Squared Error (MSE)

MSE, Denklem 2'de gösterildiği gibi tahmin ve gerçek değer arasındaki farkın karesi alınarak her bir hata hesaplanır. Daha sonra, bu hata değerleri toplam kayıt sayısına bölünerek ortalaması alınır. Bu işlem, tüm hata karelerinin toplanıp kayıt sayısına bölünmesiyle gerçekleştirilir. MSE değeri küçüldükçe, modelin tahmin doğruluğu artar ve güvenilirliği yükselir [41].

$$MSE = \sum_{i=1}^n \frac{(\text{Error}_i)^2}{n} \quad (2.3)$$

2.4.4.4. Root Mean Squared Error (RMSE)

RMSE, Denklem 3'te gösterildiği gibi MSE değerinin kareköklü sonucudur. Değer düştükçe tahmin edilen ve gerçek değer arasındaki sonuçların uyumlu olduğunu göstermektedir [42].

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\text{Error}_i)^2}{n}} \quad (2.4)$$

2.4.4.5. R-Squared (R^2)

R^2 , Denklem 4'te gösterildiği gibi bağımlı değişken ile bağımsız değişkenin ilişkisini göstermektedir. R^2 değeri 0 ile 1 arasındadır. 1'e doğru yaklaştıkça başarı oranı artmaktadır [43].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_p - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.5)$$

3. METOT

Bu çalışmada, içme suyu arıtma tesislerinde akıllı izleme ve kontrol süreçlerini gerçekleştirmek üzere çok katmanlı, IoT, fog bilişim (sis bilişim), bulut bilişim ve makine öğrenmesi mimarilerine dayalı bir sistem geliştirilmiştir. Sistem, üç adet Nodemcu ESP8266 mikrodenetleyici üzerine kurulmuş olup, her biri farklı çevresel ve operasyonel verileri izlemek üzere yapılandırılmıştır. Birinci ESP modülü, pH ve su sıcaklığı sensörü ile donatılmıştır. İkinci modül, bulanıklık ve ultrasonik (su seviyesi) sensörü ile donatılmıştır. Üçüncü modül ise yağmur ve TDS sensörü ile donatılmıştır.

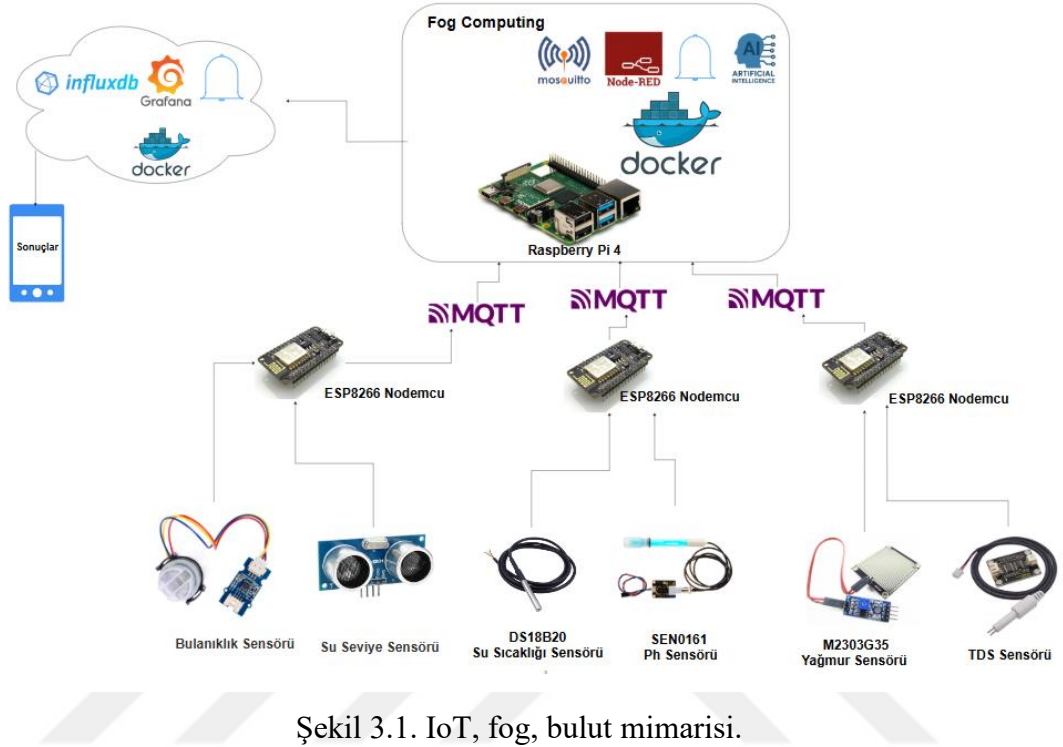
Bu ESP modülleri, Raspberry Pi ile iletişim kurmaktadır. Raspberry Pi, gelen verileri internet aracılığıyla bulut veritabanına aktarımını sağlamaktadır. Raspberry Pi aynı zamanda fog node (sis düğümü) olarak konumlandırılmış; makine öğrenmesi için kullanılan yöntemlerde verilerin ön işlenmesi, filtrelenmesi, anlık kararların verilmesi ve bildirimlerin oluşturulması gibi görevleri gerçekleştirmektedir. Böylece, özellikle performans gerektiren işlemlerin bulut yerine yerel ağ düzeyinde gerçekleştirilmesiyle hem gecikme süreleri azaltılmış hem de sistemin gerçek zamanlı aksiyon alınmasını sağlamıştır.

Makine öğrenmesi modeli, Raspberry Pi üzerinde çalışacak şekilde optimize edilmiş ve toplam 23 farklı regresyon modeli (örneğin: SVR, Random Forest, CatBoost, LightGBM, XGBoost vb.) test edilmiştir. Bu modellerden en yüksek doğruluk ve en düşük hata değerlerine ulaşan model seçilerek gerçek zamanlı kimyasal dozaj tahmini ve anomali tespiti amacıyla sisteme entegre edilmiştir.

Veri yönetimi ve kayıt işlemleri bulut sunucuda yapılandırılmıştır. Aynı zamanda, web uygulama arayüzü sayesinde operatörler su kalitesi verilerini anlık olarak izleyebilmekte ve sistemden gelen otomatik uyarılarla anormal durumlara hızla müdahale edilebilmektedir. Bu çok katmanlı yapı (IoT-Fog-Bulut mimarisi), arıtma tesisinin daha akıllı, esnek ve verimli bir şekilde yönetilmesini mümkün kılmaktadır.

3.1. IOT MİMARİSİ

Fog ve bulut tabanlı geliştirilen IoT mimarisi Şekil 3.1’de gösterilmektedir.

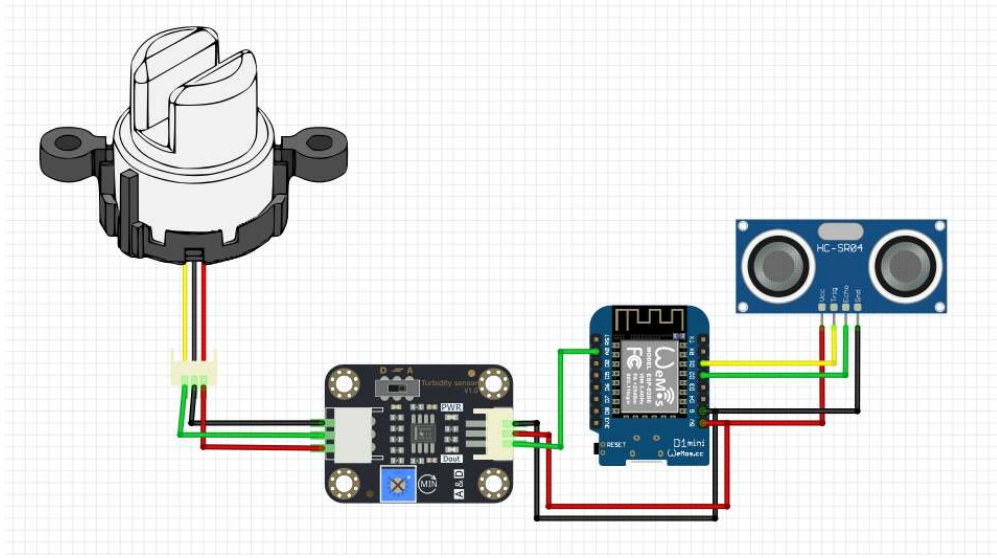


Şekil 3.1. IoT, fog, bulut mimarisi.

Şekil 3.1’de gösterilmekte olan IoT mimarisinde kullanılan ESP8266 ortamda bulunan önceden belirlenmiş wifi ağına bağlanacak şekilde programlanmıştır. Her 5 dakikada 1 çalışacak şekilde öncelikle bağlı sensörlerden anlık değerleri alıp MQTT iletişim protokülü ile fog katmanı olan raspberry pi’ye iletir. Raspberry pi üzerinde kurulu olan node-red üzerinde hazırlanan akıştaki işlemler yapıldıktan sonra veriler buluttaki sunucuda bulunan influxdb’ye kaydedilmektedir. Veriler, belirtilen referans değerleri dışında ise operatöre mail yoluyla uyarı vermektedir.

3.1.1. Sensörlerin Bağlantıları

3.1.1.1. Bulanıklık ve Su Seviye Sensörü



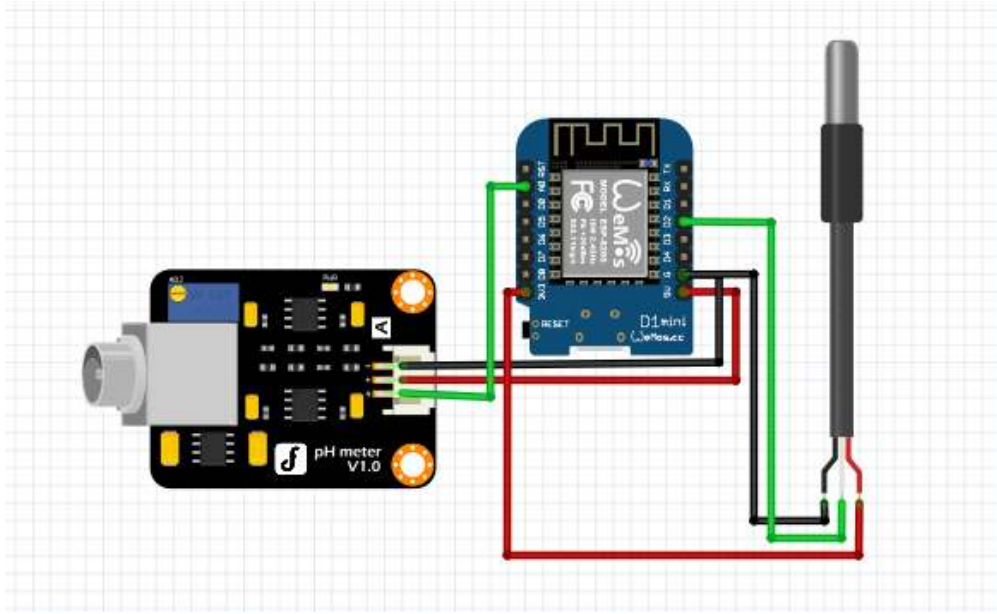
Şekil 3.2. Bulanıklık ve mesafe sensörü bağlantı şeması.

Şekil 3.2’de gösterilmekte olan devre tasarımı tesis giriş suyunun bulanıklık sensörü ile suyun kirlilik düzeyini, ultrasonik sensör ile deponun su seviyesini ölçmektedir. Bu veriler nodemcu esp8266 üzerinde toplanıp ve fog katmanındaki raspberry pi’ye mqtt iletişim protokolü üzerinden kablosuz olarak iletilir. Çizelge 3.1’de sensör bağlantı pinleri gösterilmiştir.

Çizelge 3.1. Bulanıklık ve mesafe sensörü bağlantı pinleri.

Sensör Adı	Sensör Pin	Nodemcu ESP8266 Pin
Bulanıklık Sensörü	VCC	5V
	GND	GND
	A0	A0
Mesafe Sensörü	VCC	5V
	GND	GND
	Trig	D5 (GPIO14)
	Echo	D6 (GPIO12)

3.1.1.2. Ph ve Su Sıcaklığı Sensörü



Şekil 3.3. pH ve su sıcaklığı sensörü bağlantı şeması.

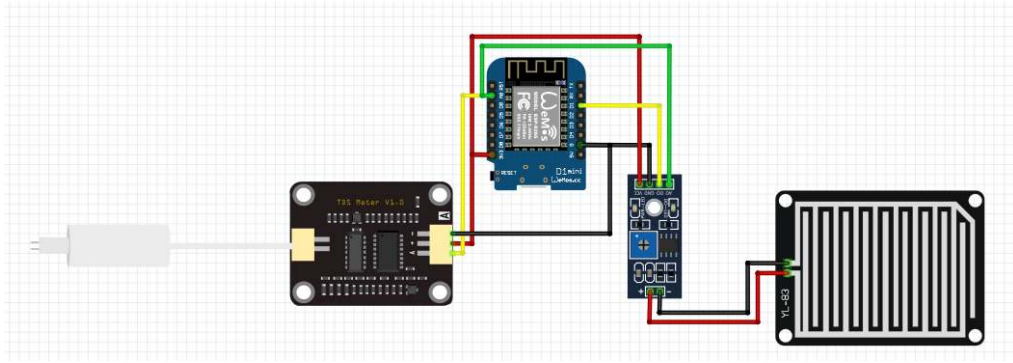
Şekil 3.3'te gösterilmekte olan devre tasarımı tesis giriş suyunun ph sensörü ile suyun ph değerini, su sıcaklık sensörü ile suyun sıcaklığını ölçmektedir. Bu veriler nodemcu esp8266 üzerinde toplanıp ve fog katmanındaki raspberry pi'ye mqtt iletişim protokolü üzerinden kablosuz olarak iletilir. Çizelge 3.2'de sensör bağlantı pinleri gösterilmiştir.

Çizelge 3.2. pH ve su sıcaklığı sensörü bağlantı pinleri.

Sensör Adı	Sensör Pin	Nodemcu ESP8266 Pin
pH Sensörü	VCC	5V
	GND	GND
	AOUT	A0
Sıcaklık Sensörü	VCC	3V3
	GND	GND
	Data	D4 (GPIO2)

Verilerin düzenli alınması için sıcaklık sensörü data pinine 4.7k Ω 'luk bir direnç bağlanmıştır.

3.1.1.3. TDS ve Yağmur Sensörü



Şekil 3.4. İletkenlik ve yağmur sensörü bağlantı şeması.

Şekil 3.4’de gösterilmekte olan devre tasarımı tesis giriş suyunun tds sensörü ile suyun çözünmüş madde yoğunluğunu, hc-sr04 sensörü ile yağmur durumunu ölçmektedir. Bu veriler nodemcu esp8266 üzerinde toplanıp ve fog katmanındaki raspberry pi’ye mqtt iletişim protokolü üzerinden kablosuz olarak iletilir. Çizelge 3.3’te sensör bağlantı pinleri gösterilmiştir.

Çizelge 3.3. İletkenlik ve yağmur sensörü bağlantı pinleri.

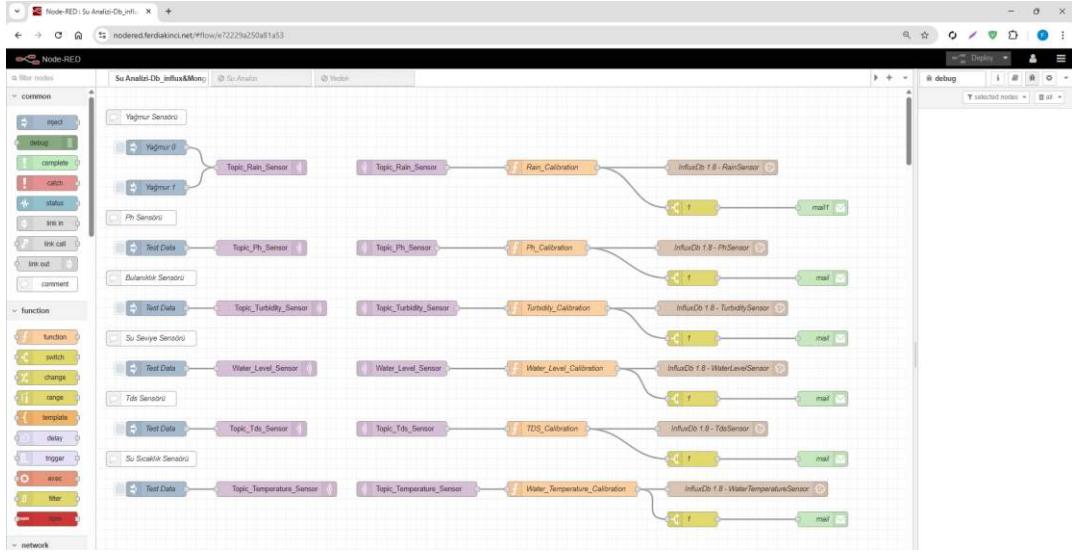
Sensör Adı	Sensör Pin	Nodemcu ESP8266 Pin
İletkenlik Sensörü	VCC	5V
	GND	GND
	AOUT	A0
Yağmur Sensörü	VCC	3V3
	GND	GND
	D0	D3 (GPIO0)

3.1.2. Fog Katmanı

Fog katmanında raspberry pi 4 kullanılmıştır. Raspberry pi üzerine raspberry pi os (64 bit) kurulumu yapılmıştır. İşletim sistemi üzerine docker kurulumu yapılmıştır. Tüm uygulamalar docker üzerinde çalıştırılmıştır. Docker üzerine, raspberry pi ile esp’lerin haberleşebilmesi için mqtt protokolünü destekleyen mosquitto kurulumu yapılmıştır. Sensörlerden gelecek verilerin ön işleme, kontrol ve buluttaki sunucuda bulunan veritabanına kayıt işlemlerinin yürütülebileceği node-red uygulaması kurulmuştur. Node-red üzerinde, sensörlerden gelen veri analiz edilmektedir. Değerler buluttaki kurulu olan influxdb’ye kaydedilmektedir.

3.1.2.1. Node-RED

Projede node-red 3.1.3 versiyonu kullanılmıştır. esp↔raspberry pi haberleşmesi için MQTT Broker oluşturulmuştur. Şekil 3.5’te nodered’de hazırlanan akış diyagramı gösterilmektedir.



Şekil 3.5. Nodered ekran görüntüsü.

Şekil 3.5’te yer alan node-red tabanlı akış diyagramı, su kalitesi izleme sistemine entegre edilmiş sensörlerden elde edilen verilerin uçtan uca işleme sürecini temsil etmektedir. Her bir sensör (yağmur, pH, bulanıklık, su seviyesi, TDS ve sıcaklık) için bağımsız bir MQTT “topic” tanımlanmış ve bu sayede sensör verilerinin ayrıştırılması ve yönetilmesi sağlanmıştır. Bu topic’ler üzerinden gelen veriler doğrudan kalibrasyon işlev düğümlerine yönlendirilmiştir. Örneğin, Topic_Rain_Sensor üzerinden alınan yağmur verileri Rain_Calibration işleviyle kalibre edilmekte, benzer şekilde pH, bulanıklık, su seviyesi, TDS ve sıcaklık verileri de sırasıyla kendi kalibrasyon blokları (Ph_Calibration, Turbidity_Calibration, vb.) aracılığıyla işlenmektedir.

Bu kalibrasyon adımı, her sensörün fiziksel koşullardan ve cihaz hassasiyetlerinden kaynaklanan ölçüm sapmalarını gidermek için kritik bir rol oynamaktadır. Böylece sensörlerden elde edilen veriler, daha güvenilir ve standartlaştırılmış hale getirilmektedir.

Kalibrasyonu tamamlanan veriler, InfluxDB üzerinde tanımlanan tablolara kaydedilmektedir. Sistemin bir diğer önemli bileşeni ise uyarı mekanizmasıdır. Her bir sensör verisi, belirli eşik değerlerin aşılp aşılmadığını kontrol eden switch düğümleri ile işlenmektedir. Eğer tanımlanan kritik sınırlar aşılsa, sistem tarafından tanımlanmış e-

posta bildirimi (mail düğümü) devreye girerek kullanıcılara anlık uyarı iletilmektedir.

3.1.3. Bulut Katmanı

Bulut sistemi için sunucu hizmeti sağlayan bir firmadan 4 Core CPU, 6 GB Ram, 50 GB Disk altyapısına sahip bir sunucu kiralanmıştır. Sunucuya açık kaynak kodlu işletim sistemi olan ubuntu 22.04 versionu kurulmuştur. Ubuntu üzerine docker kurulumu yapılmıştır. Influxdb, postgresdb, grafana ve geliştirilen web tabanlı uygulama docker üzerinde containerda çalıştırılmıştır.

3.1.3.1. Docker

Bu projede docker 27.1.1 versionu kurulmuştur. Fog sisteminde kullanılan Raspberry Pi ve bulut sistemindeki sunucuda kullanılan işletim sistemleri üzerine docker kurulumu yapılarak tüm uygulamaların docker üzerinde çalışması sağlanmıştır.

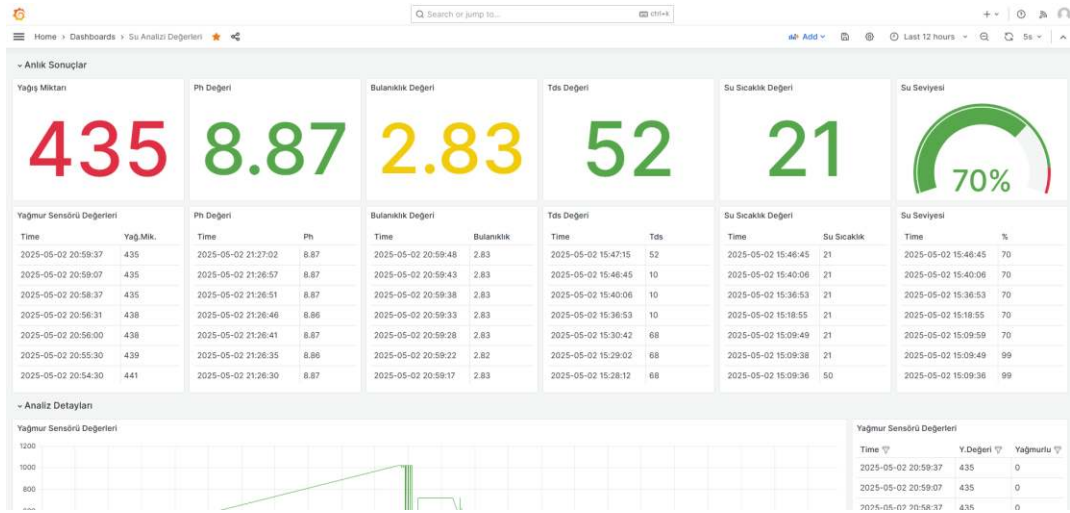
3.1.3.2. InfluxDB

Geliştirilen mimaride “WaterAnalysis” adında bir veritabanı oluşturulmuştur. Veritabanı içerisinde 6 farklı ölçüm parametresi için WaterTemperatureSensor, TdsSensor, WaterLevelSensor, TurbiditySensor, PhSensor ve RainSensor ölçüm alanları oluşturulmuştur. Her bir ölçüm sonucu kendi ölçüm alanında tutulmaktadır.

3.1.3.3. PostgreSql

Bu projede PostgreSQL 15 versionu kurulmuştur. PostgreSQL kullanım amacı; makine öğrenmesi için gerekli olan günlük ölçümlerin düzenli ve güvenli bir şekilde saklanması amacıyla kullanılmıştır.

3.1.3.4. Grafana



Şekil 3.6. Grafana ekran görüntüsü.

Bu projede grafana, influxdb'ye bağlantı yapılmıştır. Grafana veritabanındaki tüm metrikleri sürekli izleyerek görselleştirilmiş arayüzüyle kullanıcıya sunmaktadır. Grafana üzerinde geçmişe dönük verilerin de gösterilmesiyle birlikte, karşılaştırmalar yapılabilmektedir. Bu sayede geçmiş verilerle birlikte analiz ve değerlendirme yapılabilmektedir.

3.1.4. Sensör Verilerinin Değerlendirilmesi

Sensörlerden gelen veriler fog katmanında kalibre edilmektedir. Sensörlerden gelen verilerde ph, bulanıklık, tds ve su sıcaklık değeri İnsani Tüketim Amaçlı Sular Hakkında Yönetmelik'te belirtilen değerlere göre karşılaştırılmakta ve bu değerlere göre eşik değerleri belirlenmekte, uyarı mekanizmaları oluşturulmaktadır. Yağış miktarı ve su seviyesi ise yapılmış sensör kalibrasyonları ve testleri sonucu belirlenen eşik değerlerine göre işlem görmektedir. Çizelge 3.4'te sensörlerin kabul edilebilir aralıkları gösterilmiştir.

Çizelge 3.4 Sensörlerin kabul edilebilir değerleri.

Parametre	Ölçüm Birimi	Normal Değerler
Yağış Miktarı	Analog Değer	<100
pH Değeri	pH Birimleri	6,5-9,5
Bulanıklık Değeri	Nephelometric Turbidity Unit	<1

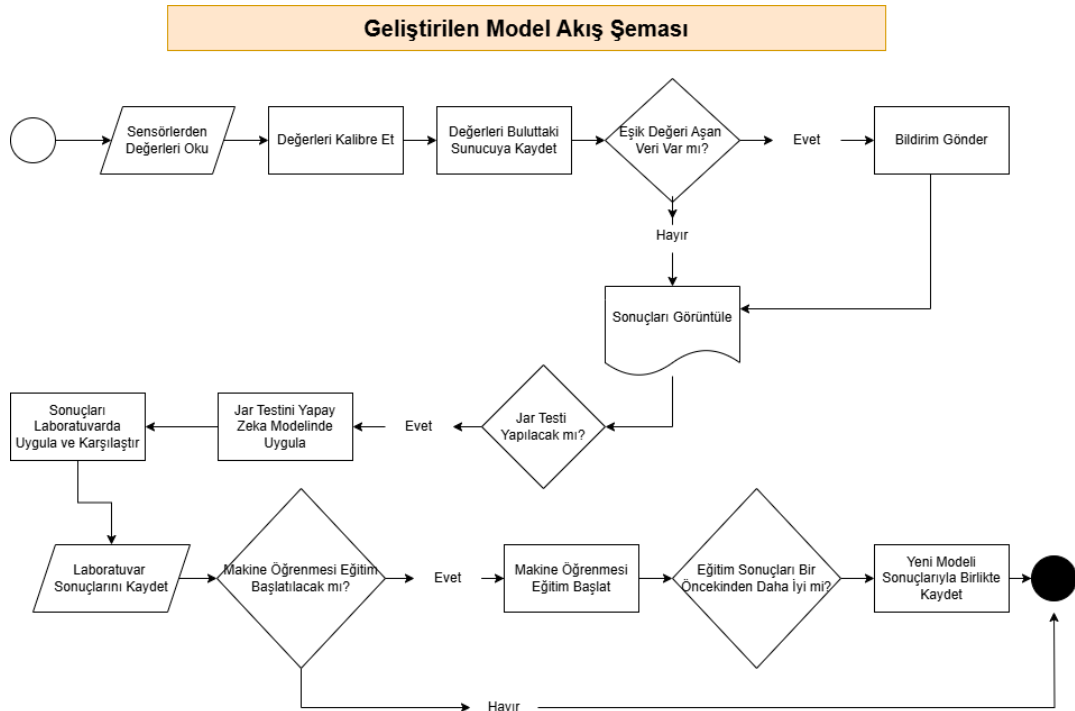
Çizelge 3.4 (devam). Sensörlerin kabul edilebilir değerleri.

Tds Değeri	20 °C’de $\mu\text{S}/\text{cm}-1$	<2500
Su Sıcaklık Değeri	°C	20
Su Seviyesi	%	<90

İnsani tüketim amaçlı sular hakkında yönetmelikte pH değeri 6,5–9,5 aralığında, bulanıklık değeri 1 NTU altında, TDS 2500 altında ve su sıcaklığı 20°C eşit olması gerekmektedir [4]. Yağmur sensörü analog ve dijital pinleri üzerinden değerleri okuyabilmektedir. PCB kart üzerine gelen su damlaları ile analog değerlerde 0-1024 aralığında çalışmaktadır. Dijital değerde 0-1 olacak şekilde çalışmaktadır. Yağmurun artması ile PCB üzerine düşen su damlacıkları artarak değerini arttırdığı görülmüştür. Yapılan testler ile 100’ün üstünde yağmur var olduğu belirlenmiştir. Su seviyesi deponun üstüne konulan ultrasonik sensörle ölçülmüştür. Sensör değeri kalibre edildikten sonra 90% ve üstünde su seviyesi riskli seviyeye yükseldiği belirtilmiştir.

3.2. IOT DESTEKLİ MAKİNE ÖĞRENMESİ MODELİ

IoT destekli makine öğrenmesi modelinin akış diyagramı Şekil 3.7’de gösterilmiştir.



Şekil 3.7. Akış diyagramı.

Şekil 3.7’de, geliştirilen “nesnelerin interneti ve makine öğrenmesi tabanlı jar testi süreci”ne ilişkin ayrıntılı akış şeması gösterilmektedir.

Akış şeması aşağıdaki adımlardan oluşmaktadır.

1. "Sensörlerden Değerleri Oku" adımıyla başlamaktadır.
2. Okunan değerler "Değerleri Kalibre Et" adımıyla kalibre edilmektedir.
3. Kalibre edilen değerler "Değerleri Buluttaki Sunucuya Kaydet" adımıyla InfluxDB veritabanına kaydedilmektedir.
4. Sonra "Eşik Değeri Aşan Veri Var mı?" karar noktası gelmektedir.
 - Eğer "Evet" ise, "Bildirim Gönder" adımıyla operatöre bildirim gönderilmektedir.
 - Eğer "Hayır" ise ya da bildirim gönderildikten sonra "Sonuçları Görüntüle" adımıyla geçilmektedir.
5. Sonrasında "Jar Testi Yapılacak mı?" karar noktası gelmektedir.
 - Eğer "Evet" ise, "Jar Testini Yapay Zeka Modelinde Uygula" ve sonra "Sonuçları Laboratuvarı Uygula ve Karşılaştır" adımları gerçekleştirilmektedir.
6. Laboratuvar sonuçları kaydedildikten sonra "Makine Öğrenmesi Eğitimi Başlatılacak mı?" karar noktası gelmektedir.
 - Eğer "Evet" ise, "Makine Öğrenmesi Eğitimi Başlat" adımıyla geçilmektedir.
 - Eğer "Hayır" ise, doğrudan "Yeni Modeli Sonuçlarıyla Birlikte Kaydet" adımıyla geçilmektedir.
7. Eğitim sonrasında "Eğitim Sonuçları Bir Öncekinden Daha İyi mi?" kararı verilmektedir.
 - İyileşme var ise "Yeni Modeli Sonuçlarıyla Birlikte Kaydet" adımıyla geçilmektedir.
8. Süreç tamamlandıktan sonra başa dönülerek döngü devam etmektedir.

Bu akış şeması, sensörlerden toplanan verilerin otomatik analizi, eşik değerleri aşıldığında bildirim sistemi, laboratuvar ölçümleriyle karşılaştırma ve sürekli makine öğrenmesi iyileştirmesi içeren akıllı bir sistem mimarisini temsil etmektedir. Su kalitesi izleme, içme suyu arıtma tesisindeki jar testi sürecinin gelişen teknoloji ile birlikte optimizasyonu gerçekleştiriliyor.

3.3. MAKİNE ÖĞRENMESİ

3.3.1. Veri seti

Türkiye'nin Zonguldak ilinin Ereğli belediyesindeki içme suyu arıtma tesisi incelenmiştir. Tesiste günlük su analizi yapılmaktadır, analiz sonuçları Excel dosyalarında kayıt altına alınmaktadır. Veri seti Ereğli belediyesi içme suyu arıtma tesisinin 2022 Ocak ayı ile 2024 Mart ayı aralığındaki verilerden oluşmaktadır. Veri setinde günlük hava durumu, ham su bilgileri, dozlama bilgileri, dozlanmış su bilgileri, durulmuş su bilgileri, filtrelenmiş su bilgileri, temiz su bilgileri yer almaktadır. Veri setinde toplamda 791 kayıt bulunmaktadır. Veri seti özellikleri Çizelge 3.5'te gösterilmektedir.

Çizelge 3.5. Veri seti özellikleri.

Hava Durumu	a1_HavaSicakligi a2_Yagis_Miktari	a3_Nem a4_Ruzgar
Ham Su Bilgileri	b1_Ham_Su_Debi_(m3/h) b10_Ham_Su_Mangan b2_Ham_Su_Debi_m3/Gun b3_Ham_Su_Giris_Bulaniklik b4_Ham_Su_Sicaklik b5_Ham_Su_pH b6_Ham_Su_Iletkenlik b7_Ham_Su_Alkalinite b8_Ham_Su_PermanganatIndeksi	b9_Ham_Su_Demir c1_Ham_Su_Haftalik_Nitrit c2_Ham_Su_Haftalik_Nitrat c3_Ham_Su_Haftalik_Amonyum c4_Ham_Su_Haftalik_Aluminyum c5_Ham_Su_Haftalik_TOC c6_Ham_Su_Haftalik_Silikat c7_Ham_Su_Haftalik_Cozunmus_O2
Dozlama Bilgileri	d1_Dozlama_SodyumHidroksit_(lt) d2_Dozlama_Aktif_Karbon_(kg) d3_Dozlama_Katyonik_Poli_(lt/h) d4_Dozlama_Anyonik_Poli_(lt/h)	d5_Dozlama_Klor_(kg/h) d6_Dozlama_On_Klor_(kg/h) d7_Dozlama_Koagulant_(ppm) d8_Dozlama_KMnO4
Dozlanmış Su Bilgileri	e1_Dozlanmış_Su_Aluminyum e2_Dozlanmış_Su_pH e3_Dozlanmış_Su_Cozunmus_O2	
Durulmuş Su Bilgileri	f1_Durulmuş_Su_DurultucuBulaniklik f2_Durulmuş_Su_Demir	f3_Durulmuş_Su_Mangan f4_Durulmuş_Su_Aluminyum
Filtrelenmiş Su Bilgileri	g1_Filtrelenmiş_Su_Filtre_CikisiBulaniklik g2_Filtrelenmiş_Su_Aluminyum g3_Filtrelenmiş_Su_Klor g4_Filtrelenmiş_Su_Kum_Filtre1 g5_Filtrelenmiş_Su_Kum_Filtre2	g6_Filtrelenmiş_Su_Kum_Filtre3 g7_Filtrelenmiş_Su_Kum_Filtre4 g8_Filtrelenmiş_Su_Kum_Filtre5 g9_Filtrelenmiş_Su_Kum_Filtre6
Temiz Su Bilgileri	h1_Temiz_Su_Bulaniklik h2_Temiz_Su_pH h3_Temiz_Su_Bakiye_Klor h4_Temiz_Su_Demir	h5_Temiz_Su_Mangan h6_Temiz_Su_Amonyum h7_Temiz_Su_Aluminyum

3.3.2. Çalışmada Kullanılacak Özellikler

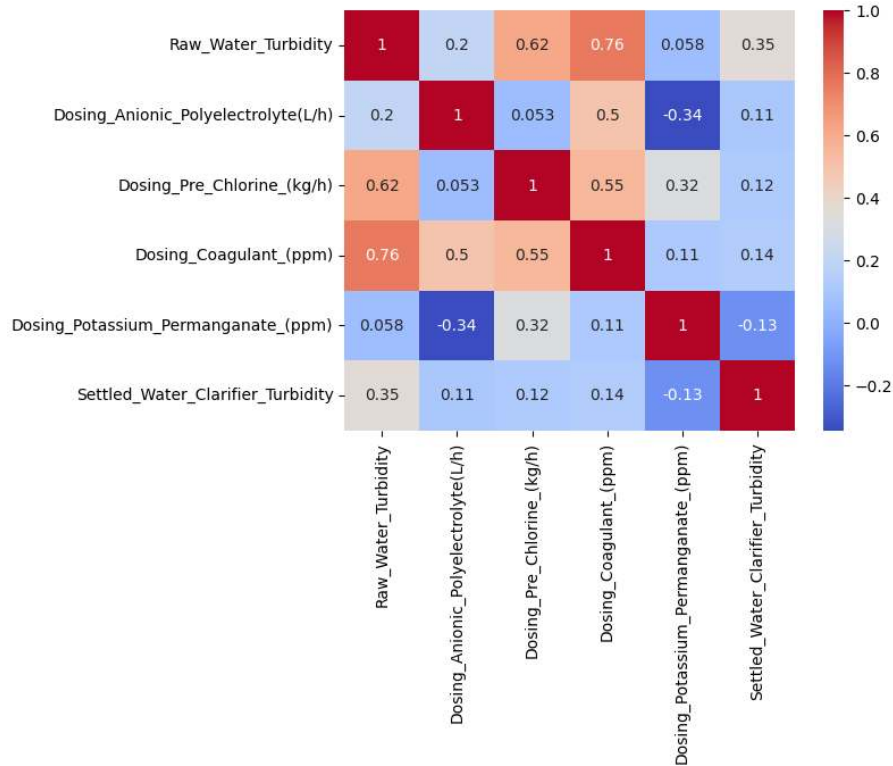
Bu çalışmada kullanılan özellikler; ham su giriş bulanıklığı, suya dozlanan anyonik polielektrolit, ön klor ve koagülant miktarları ile suya maddelerin dozajlanmasının

ardından işlem gören suyun bulanıklık değerlerinden oluşmaktadır. Çalışmada kullanılacak özellikler ve verinin istatistiksel sonuçları Çizelge 3.6’te gösterilmektedir.

Çizelge 3.6. Makine öğrenmesinde kullanılacak verilerin veri işleme öncesi istatistiksel sonuçları.

index	count	mean	std	min	25%	50%	75%	max
Raw Water Turbidity	781,00	66,00	90,22	9,70	20,30	30,60	66,20	810,00
Dosing Anionic Polyelectrolyte(L/h)	781,00	219,12	72,54	54,70	150,00	225,00	290,00	385,00
Dosing Pre Chlorine (kg/h)	781,00	2,97	1,61	-	2,07	2,82	4,60	7,17
Dosing Coagulant (ppm)	781,00	68,20	33,24	25,00	43,90	62,00	80,00	270,00
Dosing Potassium Permanganate (ppm)	781,00	14,65	34,59	-	-	-	-	222,20
Settled Water Clarifier Turbidity	781,00	2,85	0,71	1,12	2,38	2,82	3,26	7,18

Veri setindeki değerlerin ilişkileri Şekil 3.8’de korelasyon ısı haritasında gösterilmektedir.



Şekil 3.8. Korelasyon ısı haritası.

Korelasyon ısı haritası incelendiğinde, ham su giriş bulanıklığına göre eklenmesi gereken madde miktarlarının; anyonik polielektrolit için 0.2, ön klor için 0.62, koagülant için 0.76,

potasyum permanganat için 0.058 ve durultucu su bulanıklığı için 0.35 oranında olduğu görülmüştür. Sonuçlar incelendiğinde, potasyum permanganatın modele anlamlı bir etkisinin olmadığı, en yüksek etkiye sahip değişkenlerin ise koagülant ve ön klor olduğu görülmektedir.

3.3.3. Veri Ön İşleme

Makine öğrenmesi modelinde, ilk işlem olarak veri ön işleme yapılmalıdır. Veri ön işleme, veri setindeki aykırı değerlerin, tekrarlanan kayıtların, eksik kayıtların ve boş kayıtların veri setinden temizlenmesi işlemidir. Ön işleme sonrasında makine öğrenmesi modelinin performansı ve tahmin doğruluğu artmaktadır [43]. Ereğli veri setinde ön işleme sonrası 791 kayıttan 289 kayıt silinmiş olup son halinde 502 kayıt kalmıştır. Makine öğrenmesinde kullanılacak verilerin istatistiksel sonuçları Çizelge 3.7’de gösterilmektedir.

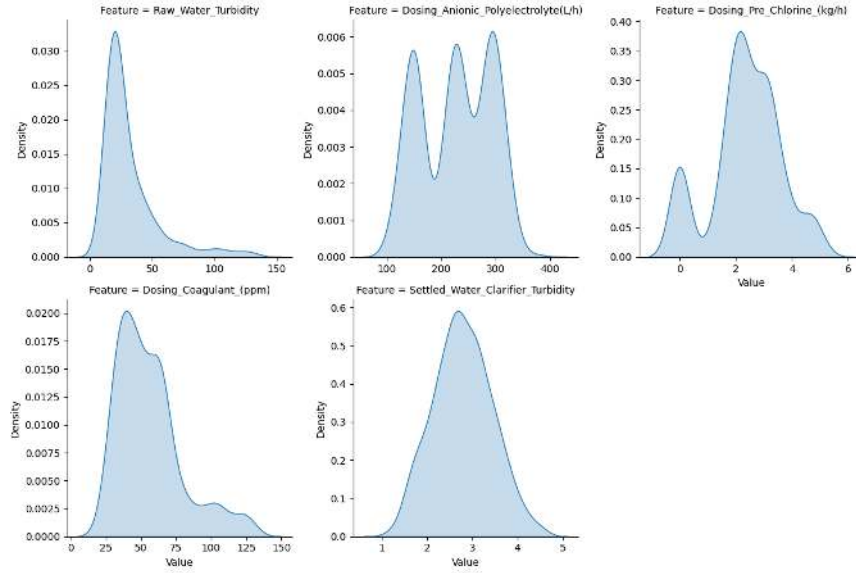
Çizelge 3.7. Makine öğrenmesinde kullanılacak verilerin veri işleme sonrası özellikleri.

index	count	mean	std	min	25%	50%	75%	max
Raw Water Turbidity	502,00	31,49	23,07	9,70	18,13	23,00	37,28	134,00
Dosing Anionic Polyelectrolyte (L/h)	502,00	226,69	64,10	112,00	155,00	230,00	290,00	377,30
Dosing Pre Chlorine (kg/h)	502,00	2,35	1,25	-	1,84	2,37	3,13	5,12
Dosing Coagulant (ppm)	502,00	56,08	23,54	25,00	38,00	50,00	65,00	130,00
Dosing Potassium Permanganate (ppm)	502,00	-	-	-	-	-	-	-
Settled Water Clarifier Turbidity	502,00	2,79	0,65	1,15	2,32	2,77	3,22	4,55

Veri ön işleme sonrası aykırı değerler temizlenmiştir. Potasyum permanganatın veri ön işleme sonrası değerleri sıfırlanmıştır. Bunun ana sebebi incelendiğinde potasyum permanganat’ın tesiste kullanımının az olması ile veri setinde de sıfır değerinin oldukça fazla olmasından kaynaklanmaktadır. Sonuçlara istinaden tesisle tekrar görüşme yapılmıştır. Potasyum permanganat’ın daha çok olağanüstü durumlarda kullanıldığı öğrenilmiştir. Potasyum permanganat, veri ön işleme esnasında değerlerinin sıfırlanmasından kaynaklı modelden çıkartılmıştır.

3.3.4. Verilerin Dağılımı

Modelde kullanılacak verilerin dağılımları Şekil 3.9’da gösterilmektedir.



Şekil 3.9. Verilerin dağılım grafiği.

3.3.5. Modelin Eğitilmesi

Makine öğrenmesi modelleri için korelasyonlar incelenmiştir. Veri seti incelendiğinde problemin regresyon problemi olduğu görülmüştür. Regresyon modeli algoritmaları ile çalışmalar yapılmıştır. Yapılan incelemeler sonucunda suyun giriş bulanıklığı ve bulanıklığa göre suyun arıtılması için gerekli olan aktif karbon, anyonik polielektrolit, klor, koagülant, potasyum permanganat katılmasıyla çıkış bulanıklığının tahmini ile ilgili regresyon modelleri oluşturulmuştur. Kullanılan regresyon modelleri; Linear Regression, Ridge Regression, Lasso (Least Absolute Shrinkage and Selection Operator), ElasticNet, Bayesian Ridge, Huber Regressor, SGDRegressor (Stochastic Gradient Descent Regressor), Passive-Aggressive Regressor, Theil-Sen Regressor, RANSACRegressor (RANDOM SAMple Consensus Regressor), SVR (Support Vector Regressor), NuSVR (Nu Support Vector Regressor), LinearSVR (Linear Support Vector Regressor), DecisionTreeRegressor (Decision Tree Regressor), RF (Random Forest), Gradient Boosting, Extra Trees Regressor, AdaBoostRegressor (Adaptive Boosting Regressor), BaggingRegressor (Bagging Regressor), HistGradientBoostingRegressor (Histogram-based Gradient Boosting Regressor), XGBRegressor (Extreme Gradient Boosting Regressor), LGBMRegressor (Light Gradient Boosting Machine Regressor), CatBoostRegressor (Categorical Boosting Regressor).

3.3.6. Hiperparametre Optimizasyonu

Modelin eğitilmesinde tahmin hatasını minimuma düşürmek ve doğruluk değerlerini

yükseltmek için hiperparametre uygulanmaktadır. Hiperparametre ile en iyi kombinasyon ve parametreler bulunur, eğitim bu parametreler ile yapılır. Bununla birlikte eğitimdeki başarı oranı da artmaktadır. Tüm modeller için çeşitli hiperparametre optimizasyon yöntemleri vardır [16]. Modelde, en iyi performansı bulabilmek için farklı hiperparametre kombinasyonları, gridsearchcv parametreleriyle farklı kombinasyonlar çalıştırılmış olup en son başarılı çıktıya ulaşılmıştır. Optimizasyonda 5 katlı çapraz doğrulama(k-fold) kullanılarak model performansı değerlendirilmiştir. Sonuçlar R² değerine göre yüksekten aza doğru sıralanarak tüm modellerin çıktıları excel dosyasına kaydedilmiştir. En iyi sonucu veren model olan SVR eğitim modeli uygulamada kullanılmak üzere kaydedilmiştir. Kullanılan hiperparametreler Çizelge 3.8’de gösterilmektedir.

Çizelge 3.8. Hiperparametreler.

Model	Hiperparametreler
LinearRegression	fit_intercept: [True, False]
Ridge	alpha: [1e-4, 1e-3, 1e-2, 0.1, 1, 10, 100]
Lasso	alpha: [1e-4, 1e-3, 1e-2, 0.1, 1, 10]
ElasticNet	alpha: [1e-4, 1e-3, 1e-2, 0.1, 1, 10], l1_ratio: [0.1, 0.3, 0.5, 0.7, 0.9]
BayesianRidge	alpha_1: [1e-6, 1e-5, 1e-4, 1e-3]
HuberRegressor	epsilon: [1.1, 1.2, 1.35, 1.5, 1.75]
SGDRegressor	alpha: [1e-5, 1e-4, 1e-3, 1e-2], penalty: ["l2", "l1", "elasticnet"]
PassiveAggressiveRegressor	C: [0.001, 0.01, 0.1, 1, 10]
TheilSenRegressor	max_subpopulation: [None, 1000, 10000], random_state: [0, 42]
RANSACRegressor	min_samples: [0.05, 0.1, 0.3, 0.5, 0.9], residual_threshold: [None, 0.5, 1.0, 5.0]
SVR	C: [0.01, 0.1, 1, 10, 100], kernel: ["linear", "rbf", "poly"], gamma: ["scale", "auto"]
NuSVR	C: [0.01, 0.1, 1, 10, 100], nu: [0.1, 0.3, 0.5, 0.7, 0.9]
LinearSVR	C: [0.01, 0.1, 1, 10, 100]
DecisionTree	max_depth: [None, 5, 10, 20, 50], min_samples_split: [2, 5, 10, 20]
RandomForest	n_estimators: [100, 200, 300], max_depth: [None, 5, 10, 20, 50]
GradientBoosting	n_estimators: [100, 200, 300], learning_rate: [0.001, 0.01, 0.1, 0.2]
ExtraTrees	n_estimators: [100, 200, 300], max_depth: [None, 5, 10, 20, 50]
AdaBoost	n_estimators: [50, 100, 200], learning_rate: [0.001, 0.01, 0.1, 1]
Bagging	n_estimators: [10, 50, 100, 200]
HistGradientBoosting	max_iter: [100, 200, 300], learning_rate: [0.001, 0.01, 0.1, 0.2]
XGBoost	n_estimators: [100, 200, 300], learning_rate: [0.001, 0.01, 0.1, 0.2], max_depth: [3, 6, 9]
LightGBM	n_estimators: [100, 200, 300], learning_rate: [0.001, 0.01, 0.1, 0.2], num_leaves: [10, 31, 50]
CatBoost	iterations: [100, 200, 300], learning_rate: [0.001, 0.01, 0.1, 0.2], depth: [4, 6, 8]

3.3.7. Modelin Değerlendirilmesi

Makine öğrenmesinde uygulanan modellerin başarı oranları, model performanslarının ölçümleri için materyal kısmında belirtilen değerlendirme metrikleri kullanılmıştır.

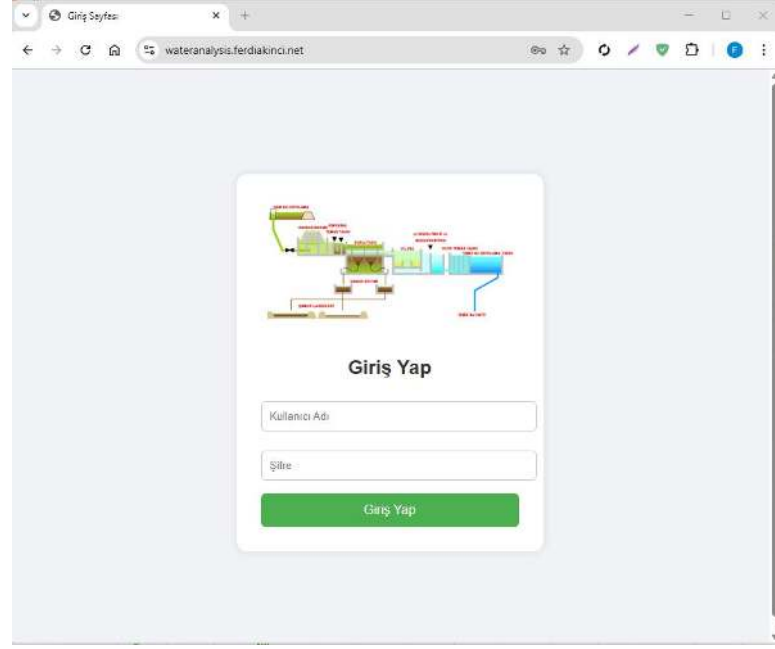
Model başarıları için Belirlilik Katsayısı (R^2) metriği, hata metrikleri için Ortalama Mutlak Hata (MAE), Ortalama Mutlak Yüzde Hatası (MAPE), Ortalama Kare Hata (MSE), Kök Ortalama Kare Hata (RMSE) kullanılmıştır.

3.4. GELİŞTİRİLEN YAZILIM

Tüm IoT sistemin anlık izlenebilmesi, yeni analizlerin kaydedilebilmesi ve makine öğrenmesi algoritmalarının tekrarlı bir şekilde çalıştırılabilmesi için donanım ve ortam bağımsız ve kurulumu gerek kalmadan çalıştırabilmesi için web tabanlı bir yazılım geliştirilmiştir. Yazılım, Python dilinde geliştirilmiştir ve buluttaki sunucuda docker altyapısında bir container içinde çalışmaktadır. Geliştirilen bu yazılım ile, tesisteki su değerlerinin anlık olarak izlenebilmesi, laboratuvarı yapılan Jar testi sonuçları ve değerlendirmelerin dijital ortamda saklanabilmesi ve makine öğrenmesi modelinin toplanan verilerle birlikte tekrar çalıştırılabilmesi sağlanmıştır.

IoT ile kurulan sistemde tesisten alınan veriler Grafana üzerinde gösterilmiştir. Grafana üzerinde hazırlanan pano (dashboard) ekranının ve oluşturulan makine öğrenmesi modelinin yazılıma entegre edilmesiyle birlikte tüm işlemlerin tek bir ekran üzerinden takip edilmesi ve işlenmesi sağlanmıştır.

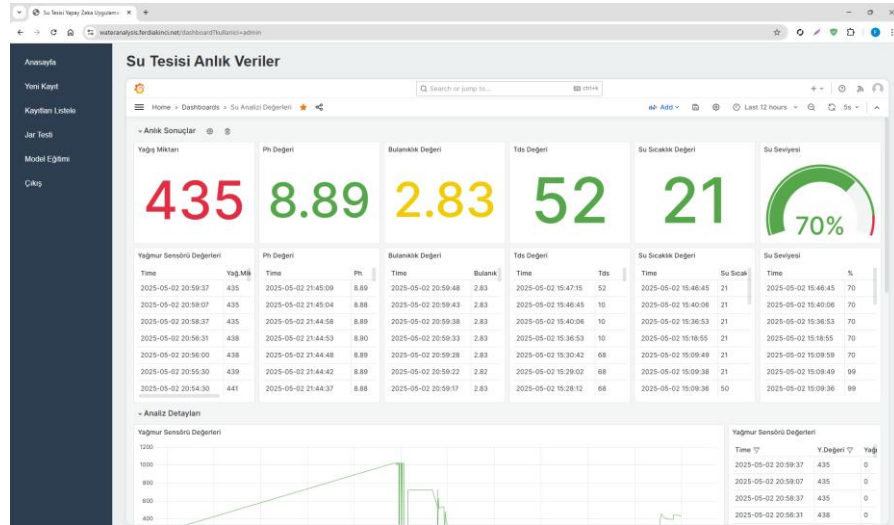
Yazılımın ilk sayfası Şekil 3.10'da gösterilmiştir.



Şekil 3.10. Yazılım giriş sayfası.

Yazılımın web tabanlı olmasıyla birlikte güvenlik altyapısının sağlanması için girişte kullanıcı doğrulaması bulunmaktadır. Doğrulama işlemi yapıldıktan sonra uygulamaya giriş yapılmaktadır. Uygulama içerisinde sol tarafta menüler ve sağ tarafta içerikler yer almaktadır.

Giriş yapıldıktan sonra açılan Ana sayfada su tesisindeki anlık veriler gösterilmektedir. Ana sayfa ekran görüntüsü Şekil 3.11’de gösterilmiştir.



Şekil 3.11. Su tesisi anlık veriler sayfası(ana sayfa) ekran görüntüsü.

Anasayfa içerisinde grafana dashboard ekranı açılmaktadır. Buradan değerler anlık olarak

izlenebilmektedir. Yazılımda sol tarafta yer alan “Yeni Kayıt” sayfasında laboratuvarlarda uygulanan Jar testi işlem sonuçları kaydedilmektedir. Yeni Kayıt sayfası ekran görüntüsü Şekil 3.12’de gösterilmiştir.

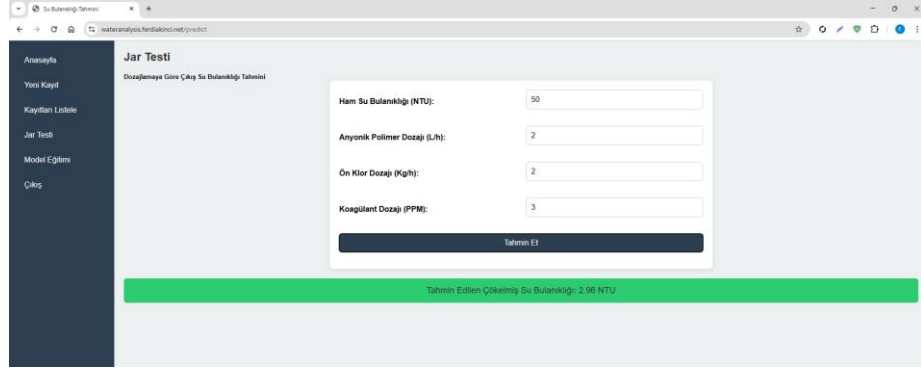
Şekil 3.12. Yeni kayıt sayfası ekran görüntüsü.

Yeni kayıt sayfasında, laboratuvarlarda testi yapılan gerçek değerler sisteme kaydedilmektedir. Kaydedilen veriler; konum, hava sıcaklığı, su sıcaklığı, yağış miktarı, giriş su bulanıklığı, anyonik polielektrolit dozajı, ön klor dozajı, koagülant dozajı, potasyum permanganat dozajı ve çıkış bulanıklığıdır. Kayıtları Listele sayfası ekran görüntüsü Şekil 3.13’te gösterilmiştir.

Id	Kayıt Tarihi	Konum	Hava Sıcaklığı	Su Sıcaklığı	Yağış Miktarı	Giriş Su Bulanıklığı	Anyonik Polielektrolit	Ön Klor	Koagülant	Potasyum Permanganat	Çıkış Bulanıklığı
4	2025-05-02 17:49:42.889620+00:00	1	1.00	1.00	1	1.00	1.00	1.00	1.00	1.00	1.00
3	2025-05-02 14:29:05.914198+00:00	2	2.00	2.00	2	2.00	2.00	2.00	2.00	2.00	2.00
2	2025-05-02 14:12:21.744720+00:00	1	1.00	1.00	1	1.00	1.00	1.00	1.00	1.00	1.00
1	2025-05-02 07:04:21.190237+00:00	1	1.00	1.00	1	1.00	1.00	1.00	1.00	1.00	1.00

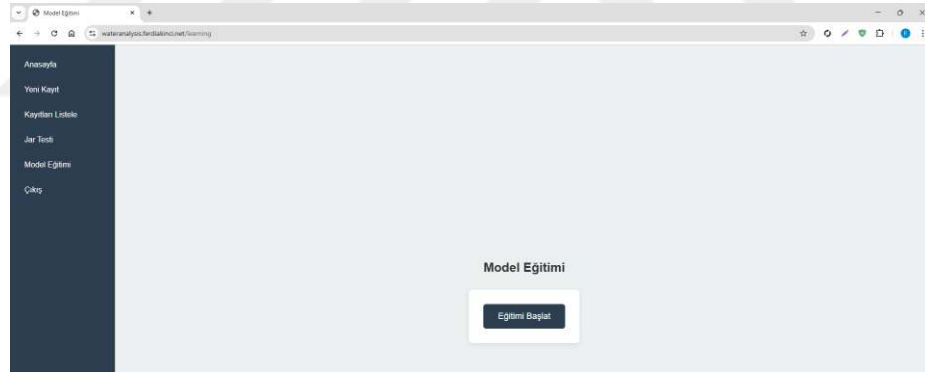
Şekil 3.13. Kayıtları listele sayfası ekran görüntüsü.

“Kayıtları Listele” sayfasında yeni kayıt sayfasında girilen değerler sondan başa (Z-A) olacak şekilde sıralanarak gösterilmiştir. Burada sondan başa tüm kayıtlar görüntülenebilmektedir. “Jar Testi” sayfası ekran görüntüsü Şekil 3.14’te gösterilmiştir.



Şekil 3.14. Jar testi sayfası ekran görüntüsü.

Jar Testi sayfasında eğitimi yapılan makine öğrenmesi modeli kullanılarak, giriş su bulanıklığına göre suyun arıtılması için suya katılması gereken madde miktarlarının ekrana girilmesiyle suyun çıkış bulanıklığı tahmin edilmektedir. Bu sayede Jar testi süreci hızlı bir şekilde sonuçlanıp son bulunan değerler ile Jar testi laboratuvarında gerçekleştirilip gerçek sonuçlar ile kıyaslayıp suya dozajlama yapılabilir. Buradaki makine öğrenmesi modülü Jar testi sürecini hızlandırmıştır. “Model Eğitimi” sayfası ekran görüntüsü Şekil 3.15’te gösterilmiştir.

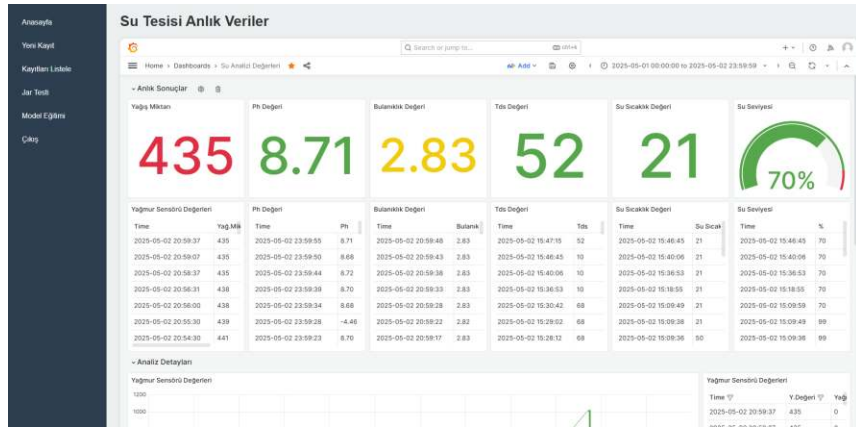


Şekil 3.15. Model eğitimi sayfası ekran görüntüsü.

Model eğitimi sayfasında yer alan “Eğitimi Başlat” butonuyla uygulamada kaydedilen verilerle makine öğrenmesi algoritmaları tekrardan çalıştırılmaktadır. Alınan sonuçlarla bir önceki sonuçlar karşılaştırılmaktadır. Eğer yeni sonuç bir önceki sonuçtan daha iyi ise önceki model dosyası silinir. Yeni model dosyası kaydedilir. Sonuçlar bir önceki sonuçtan daha iyi değilse hiçbir işlem yapılmadan işlem sonlanır. Sayfa üzerinde bilgilendirme yapılır.

4. BULGULAR VE TARTIŞMA

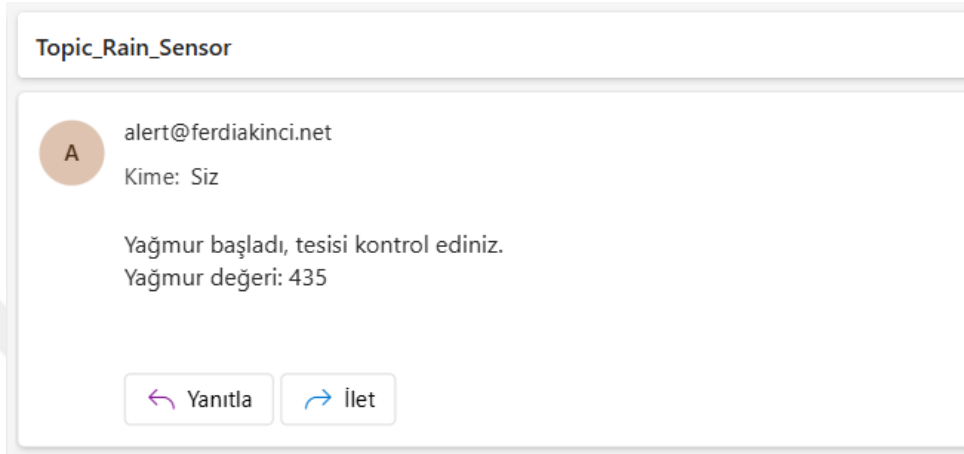
Gerçekleştirilen projede, IoT ile su değerleri anlık olarak izlenmiştir. Nodemcu esp8266 devre kartlarına bağlı olan sensörler ile 5 dakikada 1 değerler alınıp raspberry pi'ye gönderilmiştir. Veri akışları raspberry pi üzerinde kurulu olan docker üzerindeki nodered uygulaması içerisinde yapılmıştır. Esp devre kartından gelen veriler analiz edilmiştir. Gelen değerler kalibre edilip ve buluttaki sunucuya kurulan docker üzerinde kurulu olan influxdb'ye kayıt edilmiştir. Kaydedilen veriler anlık olarak Grafana arayüzünde izlenmiş olup, izlenen parametrelerin eşik değeri dışında veri geldiğinde otomatik olarak uyarı vermektedir. Uyarı hem mail yoluyla hemde grafana dashboarddaki değerler kırmızı renkli olacak şekilde belirtilmiştir. Sensörler ile sunucu arasında kurulan Fog katmanı sunucudaki iş yükünü almış olup systemin sürdürülebilirliğinin sağlanmasında destek olmuştur. Parametrelerde eşik değerleri dışı bir veri alındığında gelen uyarı ekranı Şekil 24'de ve uyarı mail ekran görüntüsü Şekil 4.1'de gösterilmiştir.



Şekil 4.1. IoT ile toplanan verilerin anlık görüntüsü.

Şekil 4.1'de gösterilen su tesisi anlık verilerdeki üst tarafta bulunan 6 parametre değerlerinden kırmızı renkli olanlar aykırı bir durum olduğu ve kontrol edilmesi gerektiğini bildirmektedir. Sarı renkli olan değerler izlenmesi, kontrol altında tutulması gerektiğini ve yeşil renkli değerler su kalitesinin iyi olduğu değerleri göstermektedir. Şekil 24'de gösterilen ekran görüntüsünde ise yağmur yağdığını, su ph 8.71 olduğunu, bulanıklık değerinin 2.83 olduğunu, tds değerinin 52 olduğu, su sıcaklığının 21 derece

olduğu ve su seviyesinin 70%'lere ulaştığı görülmektedir. Bu durumda dikkat edilmesi gereken durum ise yağmur yağdığı için su seviyesi takip edilmeli, bulanıklık değerinin takip edilmesi gerektiği yorumlanabilmektedir. Diğer parametrelerin normal olduğu gözlemlenmektedir. Yağmurun başladığı mail ile bildirilmiş olup ekran görüntüsü Şekil 4.2'de gösterilmiştir.



Şekil 4.2. Yağmur uyarısı mail ekran görüntüsü.

Yağmur sensörü 0-1024 aralığında analog değer ve 0-1 dijital olarak ölçüm yapabilmektedir. Yağmurun yoğunluğu 100 den büyük olduğu durumlarda yağmur yağıyor olarak belirlenmiştir. Şekil 24'deki ekran görüntüsünde sensörden gelen değer 435 olduğu için mail bildirimi sağlanmıştır. Değer kırmızı renge dönüşmüştür. IoT ve değerlerin takibi bu şekilde bu ekran üzerinden izlenebilmektedir. Ayrıca içme suyu arıtma tesisinde uygulanmak olan Jar testi için geliştirilen makine öğrenmesi modeli jar testini optimize etmiş olup zamandan ve katılacak madde miktarından tasarruf sağlamıştır.

Makine öğrenmesi için 23 farklı algoritma ile regresyon modeli eğitilmiş ve testleri yapılmıştır. Model performansını değerlendirmek için MAE, MAPE, MSE, RMSE, R^2 parametreleri hesaplanmıştır. Eğitim sonuçları Çizelge 4.1'de gösterilmiştir.

Çizelge 4.1. Model performans sonuçları.

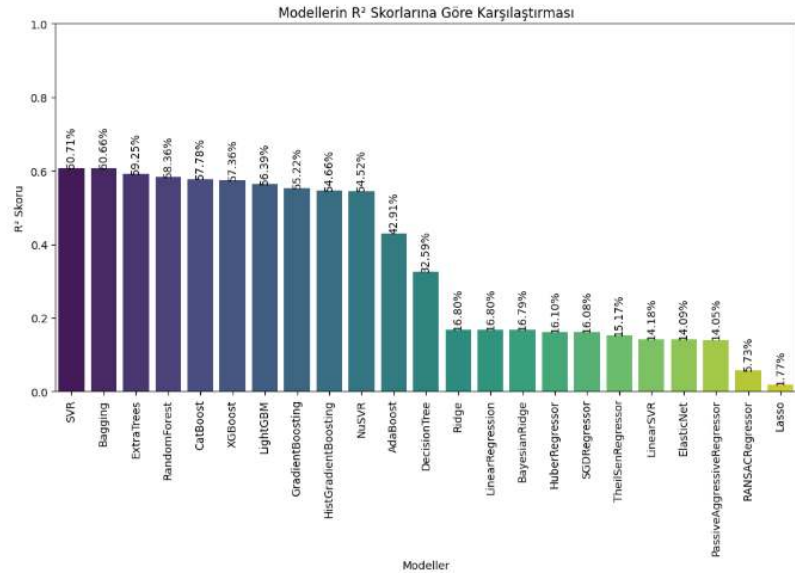
Model	MAE ↓	MAPE ↓	MSE ↓	RMSE ↓	R^2 ↑	Training Duration (s) ↓
SVR	0,294	0,109	0,161	0,402	0,607	1,271
RandomForest	0,316	0,121	0,162	0,403	0,605	23,862
ExtraTrees	0,314	0,119	0,165	0,406	0,599	11,013

Çizelge 4.1 (devam). Model performans sonuçları.

Bagging	0,317	0,120	0,165	0,407	0,597	2,437
CatBoost	0,315	0,119	0,173	0,416	0,578	4,074
XGBoost	0,325	0,126	0,175	0,419	0,574	2,600
LightGBM	0,320	0,121	0,179	0,423	0,564	1,072
GradientBoosting	0,316	0,122	0,184	0,429	0,553	5,994
HistGradientBoosting	0,323	0,123	0,186	0,432	0,547	3,992
NuSVR	0,328	0,124	0,187	0,432	0,545	1,177
AdaBoost	0,403	0,155	0,234	0,484	0,430	4,062
DecisionTree	0,392	0,146	0,279	0,528	0,322	0,406
Ridge	0,463	0,182	0,342	0,585	0,168	0,201
LinearRegression	0,462	0,182	0,342	0,585	0,168	0,112
BayesianRidge	0,463	0,182	0,342	0,585	0,168	0,267
HuberRegressor	0,462	0,181	0,345	0,587	0,161	0,503
SGDRegressor	0,465	0,183	0,345	0,587	0,160	0,241
TheilSenRegressor	0,461	0,180	0,346	0,588	0,158	5,799
LinearSVR	0,465	0,182	0,353	0,594	0,142	0,227
ElasticNet	0,467	0,185	0,353	0,594	0,141	0,502
PassiveAggressiveRegressor	0,475	0,182	0,380	0,616	0,076	0,364
Lasso	0,494	0,198	0,404	0,635	0,018	0,335
RANSACRegressor	0,588	0,232	0,585	0,765	0,424	1,255

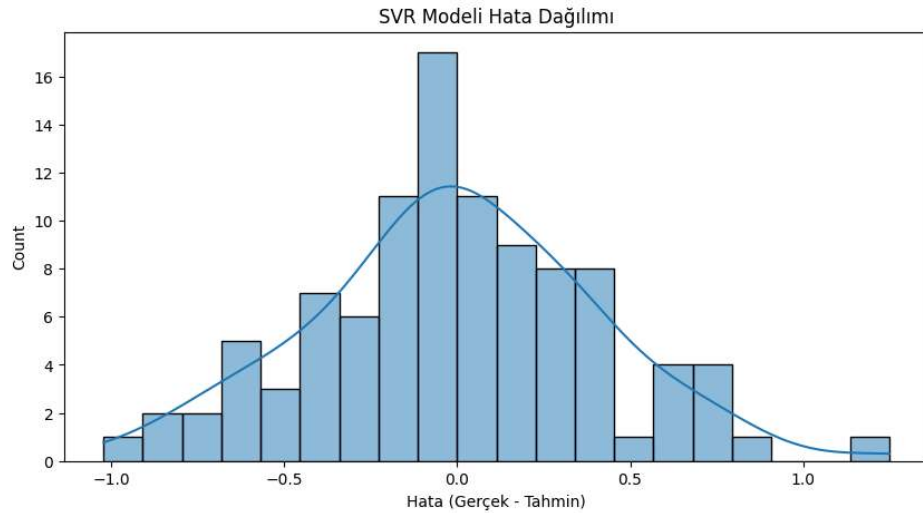
Uygulanan makine öğrenmesi algoritmaları eğitim sonuçları karşılaştırılmış ve en yüksek performans MAE: 0.29, MAPE: 0.11, MSE: 0.16, RMSE: 0.40, R^2 : 0.61, çalışma süresi: 1,27 saniye ile SVR modeliyle elde edilmiştir. SVR modeli, düşük hata oranları ve kısa eğitim süresiyle hem doğruluk hem verimlilik açısından öne çıkmıştır. Toplu öğrenme ve boosting temelli yöntemler benzer performanslar sergilemiş, ancak bazıları yüksek eğitim süresi nedeniyle uygulama açısından sınırlı kalmıştır. Lineer modeller ise düşük açıklayıcılık değerleriyle yetersiz bulunmuştur. Modellerin genel olarak düşük R^2 ve yüksek MAE, MAPE, MSE, RMSE skorlarına sahip olması, veri setinin küçük olmasından kaynaklanmaktadır. Veri seti küçük ve dağılımlar da eşit olmadığı için smote yöntemlerinden de sonuç alınamamıştır. Ancak geliştirilen sistem, veri toplama altyapısı da sunarak ileride daha kapsamlı ve etkili modellerin geliştirilmesine olanak sağlayacaktır.

Eğitim ve test çıktılarına göre hesaplanan tüm modellerin R^2 çıktıları Şekil 4.3'te gösterilmektedir.



Şekil 4.3. Modellerin R² çıktıları.

En iyi model olan SVR modelinin hata dağılımları Şekil 4.4’de gösterilmektedir.



Şekil 4.4. SVR modeli hata dağılımı grafiği.

Grafik, gerçek değerler ile tahmin edilen değerler arasındaki farkları (hataları) göstermektedir. Hata dağılımı yaklaşık olarak normal dağılım özelliği göstermekte olup, sıfır merkez etrafında simetrik bir yapı sergilemektedir.

Hataların çoğunluğu -0.3 ile +0.3 aralığında yoğunlaşmıştır, bu da modelin tahminlerinin gerçek değerlere oldukça yakın olduğunu göstermektedir. Pozitif hatalar modelin gerçek değerlerden daha düşük tahminler yaptığını, negatif hatalar ise modelin gerçek değerlerden daha yüksek tahminler yaptığını ifade etmektedir.

Bazı aykırı deęerler gözlemlense de (-1.0 ve +1.0 civarında), bunların sayısı oldukça azdır. Bu durum, modelin genel performansının tutarlı olduğunu ve aşırı tahmin hatalarının minimum düzeyde olduğunu göstermektedir.

Bu hata dağılımı, SVR modelinin veri setindeki temel yapıyı makul bir doğrulukla yakaladığını ve tahminlerde sistematik bir yanlılık olmadığını göstermektedir.

Geliştirilen python tabanlı web uygulamasında IoT ile toplanan veriler için grafanada hazırlanan dashboard ekranı, uygulamaya entegre edilmiştir. Ayrıca uygulama içerisinde bulunan laboratuvar testi kaydı, laboratuvar kayıt sonuçları, jar testi uygulaması ve makine öğrenmesi modeli geliştirme modülleri de bulunmaktadır. Bu sayede tek bir uygulama üzerinden tüm işlemler yapılabilmektedir. Uygulama üzerindeki makine öğrenmesi destekli jar testi sonucunun gerçek sonuca yakın deęerleri tahmin ettięi görülmüştür.

Makine öğrenmesi sonuçlarının iyileştirilebilmesi için daha fazla veri setine ihtiyaç duyulmaktadır. Gelecekte veri setinin genişletilerek tekrar üzerinde çalışma yapılabilir ayrıca geliştirilen bu modelde veri toplama işlemi de gerçekleştirildiğinden zamanla uygulama içerisinde de model eğitimleri gerçekleştirilip, başarı oranının artırması ön görülmektedir.

5. SONUÇ

Gelişen teknolojiler sayesinde su kalitesi analizlerinin IoT tabanlı sistemlerle ve makine öğrenmesi algoritmalarıyla entegre biçimde yürütülmesi mümkün hale gelmiştir. Bu çalışmada, içme suyu arıtma tesislerinde uygulanan klasik jar testi süreci, iot, fog bilişim ve bulut tabanlı makine öğrenmesi altyapıları ile dijitalleştirilerek optimize edilmiştir. Bu çalışmanın literatüre katkısı içme suyu arıtma tesisinde uygulanan jar testi sürecinin iot, fog ve bulut altyapısı ile destekli makine öğrenmesi entegrasyonu ile optimizasyonu yapılmasıdır.

Geliştirilen sistemde, iot katmanında yer alan sensörler aracılığıyla tesisin giriş suyuna ait bulanıklık, ph, tds, su sıcaklığı, su seviyesi ve yağmur durumu değerleri 5 dakikalık aralıklarla ölçülmüş ve mqtt protokolü kullanılarak nodemcu esp8266 üzerinden raspberry pi cihazına iletilmiştir. Raspberry pi üzerinde çalışan nodered uygulaması sayesinde sensör verileri, influxdb bulut veritabanına kaydedilmiş ve olası anormal durumlar için sistem otomatik bildirim mekanizmalarıyla desteklenmiştir. Ayrıca veriler, kullanıcıya anlık olarak sunulmak üzere grafana arayüzü ile görselleştirilmiştir. Bu yapı, hem sistemin uzaktan izlenebilirliğini artırmış hem de operatörlerin müdahale süresini azaltarak süreç güvenliğini güçlendirmiştir.

Laboratuvar ortamında gerçekleştirilen jar testleri, web tabanlı bir uygulama aracılığıyla dijital olarak kayıt altına alınmıştır. Bu sayede elde edilen veriler, makine öğrenmesi modellerinin eğitilmesi için uygun bir veri yapısına dönüştürülmüş ve sürdürülebilir bir tahmin sistemi altyapısı oluşturulmuştur. Sisteme entegre edilen fog bilişim katmanı olarak kullanılan raspberry pi cihazı, verinin doğrudan bulut yerine ara birim üzerinde işlenmesini sağlayarak gecikme sürelerini düşürmüş ve sistem performansını anlamlı ölçüde artırmıştır.

Makine öğrenmesi tarafında, tesisten elde edilen ve ön işleme süreçlerinden geçirilen 502 örnek üzerinde farklı algoritmalar eğitilmiş, en iyi performans metriklerini destek vektör regresyonu (SVR) algoritması vermiştir. Modelin R^2 : 0.61, MAE: 0.29, MAPE: 0.11, MSE: 0.16 ve RMSE: 0.40 sonuçlar üretmesi, makine öğrenmesi ile su çıkış bulanıklığının başarılı şekilde tahmin edilebildiğini göstermektedir. Ayrıca, modelin sadece 1.27 saniyelik çalışma süresiyle yanıt vermesi, gerçek zamanlı tahmin

uygulamaları için yeterli performansa sahip olduğunu ortaya koymaktadır.

Geliştirilen web tabanlı uygulama kullanıcı dostu arayüz sayesinde operatörler, giriş suyu bulanıklığı ve uygun görülen kimyasal dozaj miktarlarını sisteme girdikten sonra beklenen çıkış bulanıklığını anlık olarak görebilmektedir. Ayrıca, uygulama içerisinde yer alan “eğitimi başlat” modülü sayesinde yeni verilerle model yeniden eğitilmekte ve daha yüksek doğruluk sağlayan yeni modeller otomatik olarak sisteme entegre edilmektedir. Bu yapı, modelin zamana bağlı olarak gelişmesine olanak tanıyarak sürekli iyileştirme süreci sağlamaktadır.

Sonuç olarak, bu çalışma ile içme suyu arıtma tesislerinde uygulanan jar testi süreci, iot ve yapay zeka teknolojileri aracılığıyla dijitalleşmiş, operasyonel verimlilik artırılmış ve manuel iş gücü gereksinimi azaltılmıştır. Elde edilen bulgular, bu yaklaşımın sadece zaman ve maliyet açısından değil, aynı zamanda çevresel sürdürülebilirlik, veri doğruluğu ve proses güvenliği açısından da önemli katkılar sunduğunu ortaya koymaktadır. [48]



6. KAYNAKLAR

- [1] Md. J. Islam ve Asaduzzaman, "Smart Water Quality Monitoring and Controlling System," içinde *2021 5th International Conference on Electrical Information and Communication Technology (EICT)*, IEEE, Ara. 2021, ss. 1-6.
- [2] J. Kim, C. Hua, S. Lin, S. Kang, J.-H. Kang, ve M.-H. Park, "Deep learning-based coagulant dosage prediction for extreme events leveraging large-scale data", *Journal of Water Process Engineering*, c. 66, s. 105934, Eyl. 2024.
- [3] A. Yavuz, "Su kalite indeksinin makine öğrenme yöntemleri ile tahmini," *Research Journal of Biology Sciences*, c. 15, sayı 2, ss. 39–49, 2022.
- [4] İnsani tüketim amaçlı sular hakkında yönetmelik. (17 Şubat 2005). *T.C. Resmi Gazete*, Sayı: 25730.
- [5] R. Thorburn, A. Margheri, ve F. Paci, "Towards an integrated privacy protection framework for IoT: contextualising regulatory requirements with industry best practices," *IET Cyber-Physical Systems: Theory & Applications*, Avustralya, Aralık 2019, doi: 10.1049/cp.2019.0170
- [6] A. A. Hasan ve A. F. Hassoon, "The Efficacy of the Return Sludge in Removing Undesirable Materials from the Karbala Unified Water Project", içinde *2021 4th International Iraqi Conference on Engineering Technology and Their Applications (IICETA)*, IEEE, Eyl. 2021, ss. 230-235.
- [7] M. A. Rahu, A. F. Chandio, K. Aurangzeb, S. Karim, M. Alhussein, ve M. S. Anwar, "Toward Design of Internet of Things and Machine Learning-Enabled Frameworks for Analysis and Prediction of Water Quality", *IEEE Access*, c. 11, ss. 101055-101086, 2023.
- [8] H. M. Forhad vd., "IoT based real-time water quality monitoring system in water treatment plants (WTPs)", *Heliyon*, c. 10, sy 23, Ara. 2024.
- [9] V. Garrido-Momparler ve M. Peris, "Smart sensors in environmental/water quality monitoring using IoT and cloud services", Eyl. 2022, *Elsevier B.V.* doi: 10.1016/j.teac.2022.e00173.
- [10] N. Tewari ve M. Joshi, "Water Quality Prediction System (WQPS) and Method using Fog of Thing (FoT)", içinde *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, IEEE, Ara. 2020, ss. 158-162.

- [11] T. Turan vd., “Makine öğrenmesi algoritmaları ile su kalitesi ve içilebilirlik tahmini,” *Uluborlu Mesleki Bilimler Dergisi (UMBD)*, 2023.
- [12] D. Venkata Vara Prasad vd., “Automating water quality analysis using ML and auto ML techniques”, *Environ Res*, c. 202, s. 111720, Kas. 2021.
- [13] S. Perveen ve Amar-Ul-Haque, “Drinking water quality monitoring, assessment and management in Pakistan: A review”, *Heliyon*, c. 9, sy 3, s. e13872, Mar. 2023.
- [14] A. Benalia vd., “The use of central composite design (CCD) to optimize and model the coagulation-flocculation process using a natural coagulant: Application in jar test and semi-industrial scale”, *Journal of Water Process Engineering*, c. 57, s. 104704, Oca. 2024.
- [15] H. Şentürk, “Nesnelerin interneti ile endüstriyel su depolarının ve kimyasal üretim yapan tesislerin yağmur suyu kalitesinin uzaktan izlenmesi, analizi ve kontrolü,” Yüksek lisans tezi, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Fen Bilimleri Enstitüsü, Manisa Celal Bayar Üniversitesi, Manisa, Türkiye, 2023.
- [16] S. M. Yılmaz, “Nesnelerin interneti cihazları için kural tabanlı kontrol sistemi geliştirilmesi ve uygulaması,” Yüksek lisans tezi, Bilgisayar Mühendisliği Anabilim Dalı, Fen Bilimleri Enstitüsü, Düzce Üniversitesi, Düzce, Türkiye, 2019.
- [17] V. Lakshmikantha, A. Hiriyannagowda, A. Manjunath, A. Patted, J. Basavaiah, ve A. A. Anthony, “IoT based smart water quality monitoring system”, *Global Transitions Proceedings*, c. 2, sy 2, ss. 181-186, Kas. 2021.
- [18] T. M. Murugan, R. Kiruba Shankar, P. Shivkumar, S. Raja Kumar, K. Gayathri, ve A. Jeyam, “Monitoring and controlling the desalination plant using IoT”, *Measurement: Sensors*, c. 27, s. 100720, Haz. 2023.
- [19] S. Rajabi ve K. Moghaddasi, “IoT-Driven Water Quality Management System using Deep Q-Network”, içinde *2023 14th International Conference on Information and Knowledge Technology (IKT)*, IEEE, Ara. 2023, ss. 1-6.
- [20] M. Yousif vd., “An innovative approach to correct data from in-situ turbidity sensors for surface water monitoring”, *Environmental Modelling & Software*, c. 155, s. 105461, Eyl. 2022.
- [21] F. Amador-Castro vd., “Internet of Things and citizen science as alternative water quality monitoring approaches and the importance of effective water quality communication”, *J Environ Manage*, c. 352, s. 119959, Şub. 2024.

- [22] M. Boujelben, Z. Benmessaoud, M. Abid, ve M. Elleuchi, “An efficient system for water leak detection and localization based on IoT and lightweight deep learning”, *Internet of Things*, c. 24, s. 100995, Ara. 2023.
- [23] S. Jin, K. Zhang, T. Zhang, Y. Yang, Y. Shuai, ve T. Hu, “Establishment of an innovative machine learning-driven drinking water quality assessment model with health considerations”, *Journal of Environmental Sciences*, Mar. 2025, doi: 10.1016/j.jes.2025.03.019.
- [24] Z. Li vd., “Applications of machine learning in drinking water quality management: A critical review on water distribution system”, *J Clean Prod*, c. 481, s. 144171, Kas. 2024.
- [25] Shutterstock. (2025 08 Mayıs). *NodeMCU – Shutterstock Görselleri*. [Online]. Erişim: <https://www.shutterstock.com/tr/search/node-mcu>
- [26] Y. Singh Parihar ve Y. S. Parihar. (2025 08 Mayıs). *Internet of Things and Nodemcu A review of use of Nodemcu ESP8266 in IoT products*. [Online]. Erişim: www.jetir.org
- [27] Shutterstock. (2025 08 Mayıs). *Top View Illustration of a Rain Drop Sensor*. [Online]. Erişim: <https://www.shutterstock.com/tr/image-vector/top-view-illustration-rain-drop-sensor-1578117640>
- [28] T. L. Narayana vd., “Advances in real time smart monitoring of environmental parameters using IoT and sensors”, *Heliyon*, c. 10, sy 7, s. e28195, Nis. 2024.
- [29] PNGFind. (2025 08 Mayıs). *This Free Icons PNG Design of pH Probe*. [Online]. Erişim: https://www.pngfind.com/mpng/hiJbwhJ_this-free-icons-png-design-of-ph-probe/
- [30] Atlas Scientific LLC. (2025 08 Mayıs). *Lab Grade pH Probe – Technical Documentation*. [Online]. Erişim: <https://www.atlas-scientific.com>
- [31] Trudyo. (2025 08 Mayıs). *Su Geçirmez DS18B20 Sıcaklık Sensörü*. [Online]. Erişim: <https://trudyo.com/magaza/su-gecirmez-ds18b20-sicaklik-sensoru/>
- [32] Shutterstock. (2025 08 Mayıs). *Ultrasonic Ranging Module HC-SR04 for Arduino–Illustration*. [Online]. Erişim: <https://www.shutterstock.com/tr/image-illustration/ultrasonic-ranging-module-hc-sr04-arduino-2085289012>

- [33] OpenAI. (2025 08 Mayıs). *Yapay Zekâ Destekli Çizim: TDS Sensörü*. [Online]. Erişim: <https://chat.openai.com>
- [34] Shutterstock. (2025 08 Mayıs). *Mega Board Microprocessor Interfaces, LEDs, Connectors*. [Online]. Erişim: <https://www.shutterstock.com>
- [35] D. Wang, E. Regentova, V. Muthukumar, M. Berli, ve F. C. Harris, “A machine learning framework to measure Water Drop Penetration Time (WDPT) for soil water repellency analysis”, *Machine Learning with Applications*, c. 18, s. 100595, Ara. 2024.
- [36] A. Jenifer J., G. J. Leelipushpam Paulraj, I. J. Jebadurai ve J. Jebadurai, “IoT-based smart water tank supply management system using MQTT protocol,” *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Bhimtal, India, Kasım 2022, ss. 647–652.
- [37] Eclipse Mosquitto. (2025 08 Mayıs). *Mosquitto v1.4.8 Documentation*. [Online]. Erişim: <https://mosquitto.org/>
- [38] OpenJS Foundation. (2025 08 Mayıs). *Node-RED Documentation*. [Online]. Erişim: <https://nodered.org/>
- [39] N. Nguyen ve D. Bein, “Distributed MPI Cluster with Docker Swarm Mode,” *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA, USA, 2017, ss. 3583–3592.
- [40] InfluxData. (2025 08 Mayıs). *InfluxDB 1.8 Documentation*. [Online]. Erişim: <https://docs.influxdata.com/influxdb/v1.8/>
- [41] A. Sayar, Ş. Arslan, T. Çakar, S. Ertuğrul, ve A. Akçay, “High-Performance Real-Time Data Processing: Managing Data Using Debezium, Postgres, Kafka, and Redis”, içinde *2023 Innovations in Intelligent Systems and Applications Conference (ASYU)*, IEEE, Eki. 2023, ss. 1-4.
- [42] Grafana Labs. (2025 08 Mayıs). *What’s New in Grafana v10.3*. [Online]. Erişim: <https://grafana.com/docs/grafana/latest/whatsnew/whats-new-in-v10-3/>
- [43] Y. J. Wong vd., “Toward industrial revolution 4.0: Development, validation, and application of 3D-printed IoT-based water quality monitoring system”, *J Clean Prod*, c. 324, s. 129230, Kas. 2021.
- [44] J. Bergstra ve Y. Bengio, “Random Search for Hyper-Parameter Optimization”,

Journal of Machine Learning Research, c. 13, ss. 281-305, 2012.

[45] J. Brownlee. (2025 08 Mayıs). *Regression metrics for machine learning*. [Online]. Erişim: <https://www.machinelearningmastery.com/regression-metrics-for-machine-learning/>

[46] R. J. Hyndman ve A. B. Koehler, “Another look at measures of forecast accuracy”, *Int J Forecast*, c. 22, sy 4, ss. 679-688, 2006.

[47] Y. Zhao, W. Zhang, ve X. Liu, “Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting”, *Appl Soft Comput*, c. 154, s. 111362, Mar. 2024.

[48] S. Lin, J. Kim, C. Hua, M. H. Park, ve S. Kang, “Coagulant dosage determination using deep learning-based graph attention multivariate time series forecasting model”, *Water Res*, c. 232, s.119665, Nis. 2023.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Ferdi AKINCI

Yabancı Dili : İngilizce

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Bilgisayar Mühendisliği	Düzce Üniversitesi	2025
Lisans	Bilgisayar Mühendisliği	Düzce Üniversitesi	2021
Lisans	Yönetim Bilişim Sistemleri	Anadolu Üniversitesi	2020
Ön Lisans	Bilgi Güvenliği Teknolojisi	Bülent Ecevit Üniversitesi	2016
Lise	Ağ İşletmenliği	Kdz. Ereğli Teknik ve Endüstri Meslek Lisesi	2014

TEZDEN ÇIKAN YAYIN

F. Akıncı ve E. Küçükkülahlı (2025). Optimization Of Water Purification Process With Machine Learning. *10th International Azerbaijan Congress on Life, Engineering, Mathematical, and Applied Sciences* (ss. 251-63), Bakü, Azerbaijan.

DİĞER YAYINLAR

F. Akıncı, F. Kayaalp, ve E. Küçükkülahlı (2024). Parkinson Hastalığı El Titremesi Analizi. *5. Uluslararası Mühendislik Araştırmaları Sempozyumu* (ss. 90-98), Düzce, Türkiye