

**T.C.
AKDENİZ ÜNİVERSİTESİ**



**SU KAYIPLARI İÇİN AKILLI SU SAYAÇLARI İLE YAPAY ZEKA TABANLI
OTOMASYON TASARIMI**

Mert Celal NİŞANCI

**FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

ARALIK 2021

ANTALYA

**T.C.
AKDENİZ ÜNİVERSİTESİ**



**SU KAYIPLARI İÇİN AKILLI SU SAYAÇLARI İLE YAPAY ZEKA TABANLI
OTOMASYON TASARIMI**

Mert Celal NİŞANCI

**FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

ARALIK 2021

ANTALYA

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**SU KAYIPLARI İÇİN AKILLI SU SAYAÇLARI İLE YAPAY ZEKA TABANLI
OTOMASYON TASARIMI**

Mert Celal NİŞANCI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ

**Bu tez TUBİTAK ARDEB 1005 - Ulusal Yeni Fikirler ve Ürünler Araştırma
Destek Programı tarafından 118Y10 nolu proje ile desteklenmiştir.**

ARALIK 2021

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**SU KAYIPLARI İÇİN AKILLI SU SAYAÇLARI İLE YAPAY ZEKA TABANLI
OTOMASYON TASARIMI**

Mert Celal NİŞANCI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI
YÜKSEK LİSANS TEZİ

Bu tez 24/12/2021 tarihinde jüri tarafından Oybirliği / Oyçokluğu ile kabul edilmiştir.

Dr. Öğr. Üyesi Yalçın ALBAYRAK (Danışman)
Prof. Dr. Selçuk HELHEL
Doç. Dr. Hilal KAZAN

ÖZET

SU KAYIPLARI İÇİN AKILLI SU SAYAÇLARI İLE YAPAY ZEKA TABANLI OTOMASYON TASARIMI

Mert Celal NİŞANCI

Yüksek Lisans Tezi, Elektrik Elektronik Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Yalçın ALBAYRAK

Aralık 2021; 61 sayfa

Akıllı şehir teknolojileri, kentsel yaşam kalitesini ve verimliliğini iyileştirmek için yüksek potansiyele sahiptir. Belediye idarelerine ve şehirde yaşayanlara kolaylık sağlayan akıllı şehir teknolojileri; ulaşım, trafik, sağlık hizmetleri, güvenlik vb. konularda kullanıcılara yardımcı olmaktadır. Su tüketimindeki verimliliğin artırılması da akıllı şehir teknolojileri sayesinde sağlanabilmektedir.

Proje çalışmasında, su abonelerinin aşırı tüketiminin ve su hattındaki arızalardan kaynaklı sızıntıların tespiti ve yönetilmesi için otomasyon uygulaması geliştirilmiştir. Bu otomasyon uygulaması, su tüketim grafiklerinden faydalanarak, matematiksel formüller ve algoritmalar yardımıyla ve akıllı su sayaçlarıyla ortak kullanılan bir yazılım sayesinde sızıntıların ve aşırı tüketimlerin tespitini, azaltılmasını ve yönetimini gerçekleştirmiştir. Yazılan arayüz yazılımları ile kullanıcı analiz sonuçlarını izleyebilmektedir.

Akıllı su sayaçları ile abonelerin su tüketim verilerinin periyodik olarak alınması, abonelerin aşırı su tüketimlerini ve sızıntılarını tespit etmek üzere otomatik olarak analizlerin yapılması ve bu analiz sonuçlarına göre kullanıcı onayının da alınarak elektronik vana oranlarının ayarlanması sağlanmıştır. Bu sayede su tüketimi sürekli olarak kontrol altında tutulmaktadır. Böylece gerekli su tasarrufu sağlanmaktadır.

ANAHTAR KELİMELER: Akıllı su otomasyonu, Aşırı tüketim analizi, Otomatik su tüketim yönetimi, Sızıntı analizi

JÜRİ: Dr.Öğr.Üyesi Yalçın ALBAYRAK

Prof. Dr. Selçuk HELHEL

Doç. Dr. Hilal KAZAN

ABSTRACT

AI-BASED AUTOMATION DESIGN WITH SMART WATER METERS FOR WATER LOSSES

Mert Celal NİŞANCI

MSc Thesis in Department of Electrical Electronics Engineering

Supervisor: Asst. Prof. Dr. Yalçın ALBAYRAK

December 2021; 61 pages

Smart city technologies have high potential to improve the quality and productivity of the urban life. Smart city technologies which provide convenience to municipal administrations and residents of the city help users in matters such as transportation, traffic, health services, security, etc. Increasing water-use efficiency can also be achieved through smart city technologies.

In the project work, an automation application has been developed to detect and manage over-consumption of water subscribers and leaks caused by faults in the water line. The automation application has realized the detection, reduction and management of leaks and over-consumption through a software that used in common with the smart water meters by using water consumption graphs. The user can monitor the analysis results with the written interface software.

The smart water meters provide to import the water consumption data of the subscribers periodically and to analyze automatically to detect the over-consumption and leakage of subscribers and to set electronic valve ratios with user approval according to the results of the analysis. In this way, water consumption is constantly kept under control. Thus, necessary water savings are achieved.

KEYWORDS: Automatic water consumption management, Intelligent water automation, Leakage analysis, Over-consumption analysis

COMMITTEE: Asst. Prof. Dr. Yalçın ALBAYRAK

Prof. Dr. Selçuk HELHEL

Assoc. Prof. Dr. Hilal KAZAN

ÖNSÖZ

Yüksek Lisans eğitimim boyunca yardım, bilgi ve tecrübeleri ile bana sürekli destek olan, çalışmam boyunca değerli bilgilerini paylaşan, projenin her aşamasında gösterdiği yollarla, hedefe belirlenen sürede ulaşmamı sağlayan saygıdeğer danışman hocam; Dr. Öğr. Üyesi Yalçın Albayrak’a sonsuz teşekkürlerimi sunarım.

Çalışmam boyunca, bilgi ve tecrübeleri ile sürekli destek olan Elektrik Elektronik Mühendisliği Bölümü hocalarıma sonsuz teşekkürlerimi sunarım.

Bu çalışmada, bilgi ve tecrübeleri ile sürekli destek olan ve çevre mühendisliği bölümünde bursiyer olarak çalıştığım projenin yöneticisi Prof. Dr. Ayşe Muhammetoğlu ve proje araştırmacısı Prof. Dr. Habib Muhammetoğlu ’na sonsuz teşekkürlerimi sunarım.

Çevre mühendisliği bölümü ile ilgili olan konularda bilgilerini paylaşan ve emek veren projedeki bursiyer arkadaşlarıma sonsuz teşekkür ederim.

Çalışmalarım boyunca maddi manevi destekleriyle beni hiçbir zaman yalnız bırakmayan aileme sonsuz teşekkür ederim.

Çalışmalarım boyunca manevi desteklerini esirgemeyen tüm arkadaşlarıma sonsuz teşekkür ederim.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
ÖNSÖZ	iii
AKADEMİK BEYAN	vi
SİMGELER VE KISALTMALAR.....	vii
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ	x
1. GİRİŞ	1
2. KAYNAK TARAMASI	4
3. MATERYAL VE METOT	8
3.1. Otomasyon Sistemine Genel Bakış	8
3.2. Otomasyon Arayüzü.....	10
3.3. Akışın Analiz Sonuçlarına Göre Ayarlanması	13
3.3.1. Vana oranlarının belirlenmesi ve veri tabanına yazılması.....	13
3.3.2. Vana oranının fiziksel olarak ayarlanması	13
3.3.3. Kullanıcı onayı için Android uygulama	14
3.4. Sistem Enerjisinin Su Akışından Karşıllanması.....	15
3.5. Veri Setinin Ön Düzenlemesi.....	16
3.5.1. Veri temizleme.....	17
3.5.2. Özellik seçimi	17
3.5.3. Veri dönüşümleri	18
3.5.4. Özellik mühendisliği.....	19
3.5.5. Boyutsal küçülme	19
3.6. Sayaç Verilerinin Düzenlenmesinde Kullanılan Yöntemler	20
3.6.1. Veri ön işleme.....	20
3.6.2. Zaman serisini oluşturma.....	21
3.7. Analiz Yöntemlerine Genel Bakış ve Bazı Terimlerin Tanımı	23
3.8. Tüketim Eğrisinin Düzeltilmesi	23
3.8.1. Hareketli Ortalama (Moving Average).....	24
3.8.2. Günlük maksimuma göre normalizasyon	24
3.9. Sezonsal Analiz	24
3.9.1. Trend.....	24

3.9.2. Mevsimsel dalgalanmalar	25
3.9.3. Rastlantısal hareketler	25
3.10. LSTM	25
3.11. ARIMA	27
3.12. SARIMA	28
3.13. Prophet	29
3.14. Öğrenme Eğrileri	30
3.15. Aykırılık Ölçümü ve Tespiti	35
3.16. Aşırı Tüketim Tespiti	35
3.17. Sızıntı Tespiti	36
4. BULGULAR	38
5. TARTIŞMA	41
6. SONUÇLAR	44
7. KAYNAKLAR	45
ÖZGEÇMİŞ	

AKADEMİK BEYAN

Yüksek Lisans Tezi olarak sunduğum “Su Kayıpları İçin Akıllı Su Sayaçları İle Yapay Zeka Tabanlı Otomasyon Tasarımı” adlı bu çalışmanın, akademik kurallar ve etik değerlere uygun olarak yazıldığını belirtir, bu tez çalışmasında bana ait olmayan tüm bilgilerin kaynağını gösterdiğimi beyan ederim.

24/12/2021

Mert Celal NİŞANCI

SİMGELER VE KISALTMALAR

Simgeler

σ	: Sigmoid fonksiyonu
$\tanh$	: Aktivasyon fonksiyonu
h_{t-1}	: Bir önceki gizli katman
V	: Bir önceki gizli katmanın ağırlığı
x_t	: Mevcut giriş vektörü
U	: Mevcut giriş katmanı ağırlığı
b	: Bias vektörü

Tezde ondalık yazım olarak “.” işareti kullanılmıştır. (Örneğin: 21.01)

Kısaltmalar

Wi-Fi	: Wireless Fidelity (Kablosuz Bağlantı Alanı)
LSTM	: Long Short-Term Memory (Uzun Kısa Süreli Bellek)
RMSE	: Root Mean Squared Error
IoT	: Internet of Things (Nesnelerin İnterneti)
SVR	: Support Vector Regression (Destek Vektör Regresyonu)
ARIMA	: Autoregressive Integrated Moving Average (Birleştirilmiş Otoregresif Hareketli Ortalama Modeli)
MAPE	: Mean Absolute Percentage Error (Ortalama Mutlak Yüzde Hata)
RMSE	: Root Mean Square Error (Kök Ortalama Kare Hata)
RNN	: Recurrent Neural Network (Tekrarlayan Sinir Ağları)
SWaT	: Secure Water Treatment (Güvenli Su Arıtma)
WADI	: Water Distribution (Su Dağıtımı)
GAN	: Generative Adversarial Networks (Çekişmeli Üretici Ağ)

ŞEKİLLER DİZİNİ

Şekil 3.1. a) Su deposu; b) Hidrofor; c) Oransal vana; d) IoT kontrol cihazı; e) Elektromanyetik debimetre; f) Debinin izlenebildiği ekran; g) GSM tabanlı akıllı su sayacı; h) LORA tabanlı akıllı su sayacı; i) NB-IoT tabanlı akıllı su sayacı; j) Mekanik su sayacı; k) Tahliye noktası.....	8
Şekil 3.2. a) Sayaç verileri; b) Sayaç verilerinin server bilgisayardaki veri tabanına iletilmesi; c) Otomasyon programı; d) İnternetteki veri tabanı; e) Android uygulama onayı; f) IOT tabanlı Elektronik Kart ile Oransal Vana Kontrolü	9
Şekil 3.3. Saatlik ve 15 dakikalık tüketim değerlerinin gösterimi	11
Şekil 3.4. Analizlerin yapılması ve sonuçların grafikte gösterimi.....	11
Şekil 3.5. Analiz sonuçlarının belirli tarih arasında filtrelenmesi	12
Şekil 3.6. Belirlenen tarih arasında abonelerin normal ve anormal tüketimlerinin miktarının gösterilmesi.....	12
Şekil 3.7. Abone durumlarının gözlenmesi	12
Şekil 3.8. Vana oran verilerinin kaydedildiği Firebase veri tabanı.....	13
Şekil 3.9. Vana oranının ayarı için elektronik devre	14
Şekil 3.10. Vana oranının ayarlanması için kontrol modülü	14
Şekil 3.11. Android uygulama bildirimi ve onay ekranı.....	15
Şekil 3.12. Su dinamosu ile şarj edilen bir güç besleme modülü	16
Şekil 3.13. Bir akıllı sayaç tarafından sunucu bilgisayara gönderilen verinin ham hali.	21
Şekil 3.14. Ön işlemesi yapılmış veri seti.....	22
Şekil 3.15. Saatlik tüketim verisi	23
Şekil 3.16. a) Zaman Serisi b) Trend c) Sezonallık Grafiği d) Aykırılık Grafiği.....	25
Şekil 3.17. SARIMA tahmini.....	29
Şekil 3.18. Prophet tahmini gösterimi.....	30
Şekil 3.19. Prophet tahmini günlük gösterim.....	30

Şekil 3.20. Yeterli kapasiteye sahip olmayan bir modelin eksik öğrenme halindeki kayıp grafiği	31
Şekil 3.21. Daha fazla eğitim gerektiren bir eksik öğrenme modelini gösteren eğitim öğrenme eğrisi örneği.....	32
Şekil 3.22. Aşırı öğrenme modelini gösteren eğitim ve test öğrenme eğrisi	32
Şekil 3.23. İyi öğrenmeyi gösteren eğitim ve test öğrenme eğrileri örneği	33
Şekil 3.24. Test veri kümesine göre çok küçük olabilecek bir eğitim veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği	33
Şekil 3.25. Eğitim veri kümesine göre az sayıda test veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği	34
Şekil 3.26. Eğitim veri kümesine göre çok küçük olabilecek test veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği	34
Şekil 3.27. Aykırılık grafiği	35
Şekil 3.28. Aşırı tüketim tepiti	35
Şekil 3.29. Basit sızıntı tespiti	36
Şekil 3.30. Aktif tüketim sırasındaki sızıntı tespiti.....	37
Şekil 4.1. LSTM ağ modeli.....	38
Şekil 4.2. Eğitim ve test kayıp grafiği	39
Şekil 4.3. Otomasyon arayüzü	40

ÇİZELGELER DİZİNİ

Çizelge 3.1. Su abonelerinin sınıfları.....	8
Çizelge 4.1. Denenmiş tahmin yöntemleri ve sonuçlar	38



1. GİRİŞ

Su, genellikle sıradan olarak algılsa da, dünyada en çok ilgi çeken maddedir. Günlük yaşantıda yıkanma, içme, yemek pişirme gibi temel eylemlerde sıvı su kullanılır. Vücudumuzun üçte ikisi sudan oluşur ve yaşamımızı devam ettirebilmek için suya ihtiyaç duyarız. Çoğu biyolojik işlevi yerine getiren organizmalar sıvı sudan oluşur. Suyun biyolojik işlevi sadece seyreltici değildir. Bildiğimiz üzere yaşam, su olmadan devam edemez ve onsuz sona erer. Kuraklık kıtlıklara, hastalıklara ve ölümlere neden olur. Önemi nedeniyle su, dünya üzerinde en çok tartışılan maddedir. Suyun öneminin yalnızca bilim insanları tarafından değil, tüm insanlar tarafından benimsenmesi ve anlaşılması gerekmektedir. (Chaplin 2001)

Akıllı şehir teknolojileri, kentsel yaşam kalitesini ve verimliliğini iyileştirmek için yüksek potansiyele sahiptir. Belediye idarelerine ve şehirde yaşayanlara kolaylık sağlayan akıllı şehir teknolojileri; ulaşım, trafik, sağlık hizmetleri, güvenlik vb. konularda kullanıcılara yardımcı olmaktadır. Ayrıca birçok alanda verimliliği artırdığı gibi su tüketimindeki verimliliğin artırılması da akıllı şehir teknolojileri sayesinde sağlanabilmektedir. (Terzi ve Ocakçı 2017)

Bir su hattındaki su tüketim verilerinin izlenmesi ve analizi manuel olarak insan gözü ile yapıldığında anormal tüketim değerlerinin tespiti gecikebilir. Su hattındaki bağlantı noktalarının tahribatı, borulardaki çatlaklar gibi mekanik problemlerin yanı sıra “musluğun açık bırakılması” gibi insan hatalarından kaynaklanan su kayıpları kısa sürede belirlenemez ise su aktif olarak kullanılmadan kaybedilmiş olur. Ayrıca suyun pompalanması sırasındaki tüketilen enerjide de kayıplar artar. Bunun yanında kontrolsüz su kaybı, görülmeyen birçok hasara ve kazaya sebep olabilir. Bu yüzden anormal durumların erken tespiti hem su tüketicisi hem de su tedarikçisi için önemli bir konudur.

Su tedarikçisinin açısından bakıldığında, suyun işlenmesi içinde yer alan suyun kaynaktan alınması, arıtma, iletim ve dağıtım vb. işlemlerde yüksek miktarda iş gücü ve enerji kullanılmaktadır. Su temin sistemlerinde sızıntıların ve aşırı tüketimin azaltılması veya önlenmesi ile enerji daha verimli kullanılır. Su sızıntılarının ve aşırı su tüketimlerinin önlenerek sınırlı olan kaynakların verimli bir şekilde yönetilmesi önemli bir görevdir. Su tüketiminin etkin biçimde yönetilmesi için otomasyon sistemlerinin kullanımı büyük önem taşımaktadır. Su tüketimindeki anormalliklerin tespitini yapmak ve ardından tedarikçiye uyarı vermek, su tüketimine müdahale etmek gibi işlemlerin otomasyon sayesinde otomatik olması tedarikçiye büyük kolaylık sağlar. Böylece sızıntılar ve aşırı tüketimler bir yazılım sayesinde tespit edilip, önlenmesi için gerekli müdahaleler ayrı bir işgücüne gerek kalmadan gerçekleştirilebilir.

Toplu su kullanım alanlarındaki mekanik arızalardan kaynaklanan anormallikler yüzünden su israfı yüksek seviyelere çıkabilir. Ayrıca su basması gibi olumsuz durumlar erken tespit edilmez ise birçok hasara sebep olabilir. Toplu su kullanım alanlarındaki anormal tüketimlerin tespiti ve önlenmesi ile güvenlik önlemleri alınmış olur. Boru patlaması, musluk sızdırması, bağlantı noktalarının tahribatı gibi mekanik arızalardan kaynaklanan problemlerin erken tespiti önemli bir görevdir. Bu problemlerin tespitinden sonra uyarı vermek ve sisteme otomatik müdahale etmek için otomasyon sistemlerinin kullanılması gerekmektedir. Bu otomasyon sayesinde su tüketimindeki

anormallikleri tespit etmek, eğer olumsuz bir durum varsa uyarı vermek ve su tüketimine müdahale etmek gibi işlemlerin otomatik olarak yapılması tüketiciye maddi ve manevi olarak birçok fayda sağlar. Yani, aşırı su maliyeti ve su basması gibi olumsuz durumlar bir yazılım sayesinde öngörülüp, önlenmesi için gerekli müdahaleler otomatik olarak yapılabilir.

Akıllı su ölçüm sistemleri, su idarelerinin kıt su kayıplarını en uygun bir şekilde önlemek için veri analitiğinde saklanabilen ve kullanılacak gerçek zamanlı veri toplamayı kullanmaya başlamasıyla ivme kazanmaya başlamıştır. Bu eğilimi destekleyen en önemli araştırma alanlarından biri, su hizmetleri arasında uzaktan bağlantı sunabilen gelişmiş ölçüm altyapısıdır. Akıllı bir su ölçüm sisteminde, sistemdeki su kaynağını izlemek ve kontrol etmek için belirli eylemlere ilişkin doğru karar almak veya uygun uyarılar vermek için analitik yazılım mimarisinin desteğiyle veriler, akıllı sayaçlar ve su işletmeleri arasında iletişim kurabilir. Tüketicileri uyarır veya tüketimlerini azaltmaları için onlara rehberlik eder. Aynı zamanda su tüketim rejimi de tahmin edebilir.

Bir toplu tüketim alanında, su tüketiminin doğru bir şekilde izlenmesini sağlamak için tüketim verileri gerçek zamanlı olarak toplanmalı, saklanmalı ve periyodik olarak analiz edilmelidir. Su tüketim verilerinin gerçek zamanlı olarak toplanmasının, saklanmasının ve periyodik analizlerinin yapılabilmesi için otomasyon sistemine ihtiyaç duyulmaktadır. Bu otomasyon sistemi, su tüketim verilerini bir yazılım ile kontrol etmeli ve anormal durumları tespit edebilmelidir. Böylece mekanik problemlerden veya insan hatalarından kaynaklanan yüksek su kaybı, su basması gibi olumsuz sonuçların yaşanmaması için gerekli önlemler alınabilir. Tüketim verilerinin analiz sonuçlarında anormal bir durum tespit edildiğinde gerekli uyarıların, bilgilendirmelerin ve müdahalelerin otomatik olarak yapılması sağlanmalıdır.

Bir su hattında anormal bir tüketimin olup olmadığını belirlemenin en etkili yöntemlerinden biri, zamana bağlı su tüketim verisinin periyodik olarak analiz edilmesidir. Bu yöntemde, su abonelerinin zamana bağlı tüketim verileri, akıllı su sayaçları ile periyodik olarak ve otomatik bir şekilde toplanması gerekir. Toplanan verilerin işlenebilmesi ve analiz edilebilmesi için yazılımda düzenlenerek zaman serisi haline getirilmesi gerekir. Toplanan verilerden elde edilen bu zaman serisi, sunucuda çalışan bir yazılım ile analiz edilerek, su tüketimindeki anormallikler en kısa zamanda otomatik olarak tespit edilebilir.

Su tüketim verilerinin analizinde, su tüketim değerlerinin tahmini önemli bir veridir. Gelecekteki olması gereken normal tüketim değerlerinin elde edilebilmesi için tüketim tahmini yapılmalıdır. Bu tahmin değerleri sayesinde sızıntı tespiti, tepe tüketim tespiti ve aşırı tüketim tespiti yapılabilmektedir. Su tüketim tahmininin bir yazılım tarafından yapılması için en verimli yöntemlerden biri yapay sinir ağlarının kullanılmasıdır. Su tüketim tahmininin yapılabilmesi için geçmişteki tüketim değerleri, tüketimin yapıldığı zaman dilimi (örneğin; tüketimin haftanın hangi gününde yapıldığı bilgisi) gibi bilgiler kullanılabilir. Belirlenen geçmiş tüketim verilerini eğitim

verisi olarak kullanılmanın uygun olup olmadığını öğrenmek için otokorelasyon grafiği, kutu grafiği ve varyasyon değerlerinden faydalanılabilir. Eğer belirlenen verinin dağılım değerlerinde çok fazla aykırılık varsa veya tekrar eden veriler arasındaki benzerlik düşükse düzenleme veya eleme yapılabilir.

Yapılan çalışmada, akıllı su sayaçlarından toplanan geçmiş zamanlardaki zamana bağlı tüketim verileri ve zaman dilimini temsil etmesi için oluşturulan zaman vektörü kullanılarak yapay sinir ağı eğitilmiştir. Son zamanlardaki gerçek veriler ile tahmin verilerinin karşılaştırılması, gerçek tüketim verileri hakkında matematiksel bir bilgi ortaya çıkarmıştır. Bu matematiksel bilgi, tahmin verisi ve gerçek veri arasındaki farklılıkların bulunarak ortaya çıkan hata değerleridir. Bu hata değerleri kullanılarak, gerçek su tüketiminin normal olup olmadığı konusunda yazılım tarafından bir sonuç elde edilmiştir. Anormal değerler sızıntı, tepe tüketim ve aşırı tüketim olarak üç sınıfta incelenmektedir ve günlük olarak analiz edilmiştir.

2. KAYNAK TARAMASI

Günümüz toplumunda, insan çabalarını azaltmak için ve farklı günlük yaşam görevlerini otomatikleştirmek için çeşitli teknikler kullanıyoruz. Günlük görevleri yerine getirmek için kullanılan otomatik sistemler yüksek verimlilik ve doğruluk gibi birçok avantaj sunar. Günümüzde sadece otomatik sistemlerin oluşturulmasına değil, aynı zamanda akıllı evler, akıllı şehirler gibi akıllı sistemlerin yaratılmasına da odaklanılmaktadır. Bu projeler, nesnelerin interneti, makine öğrenimi, yapay zeka vb. yardımıyla yapılacak birçok görevi içerir. Tüm bu teknolojiler sayesinde, herhangi bir insan çabasına ihtiyaç duyulan bir görev, otomatik olarak tamamlanabilmektedir. Böyle bir teknoloji odaklı toplumda, bir binanın farklı dairelerinde su kullanımını izlemek için kullanılabilecek bir sistem yaratma ihtiyacı doğar. Buradaki temel sorun, mevcuttaki bir binanın farklı dairelerindeki su kullanım grafiğini görmenin bir yolu olmaması ve yeni su sayaçlarının kurulum maliyetinin çok yüksek olması, ayrıca kurulum için çok fazla inşaat işine ihtiyaç duyulmasıdır. Şu anda, birçok aşamada insan çabasına ihtiyaç duyulduğundan, bu tür bir görevi tamamlayan verimli bir sistem yoktur. Böylece ortada, çok az veya hiç yapı işlemi olmadan kolayca kurulabilen bir sistem yaratma ihtiyacı vardır. Ayrıca veri toplayabilmeli ve daha sonra işlemek için sunucuya sürekli olarak tüm verileri aktarmalı ve gerçek zamanlı olarak sonuç üretmelidir. Genel maliyet, ticari olarak kullanılabilecek kadar düşük olmalıdır. Aggarwal vd. (2019), su tüketim verilerini toplamak için kullanılabilecek düşük maliyetli bir akıllı sistem yaratmaya odaklanmıştır.

Akıllı sayaçlardan yayınlanan zaman serisi verilerinin analizi, elektrik tüketim verilerinde yaygınlaşmıştır. Günümüzde, akıllı sayaçlar aracılığıyla toplanan orta çözünürlüklü su tüketimi verilerinin analizi mümkün hale gelmiş ve araştırmalarda, farklı gereksinimlere yanıt vermek için istatistiksel ve makine öğrenimi araçları geliştirilmiştir. Örneğin, su tüketiminin daha iyi anlaşılması, davranışlar ve tüketim tahmini konuları üzerinde çalışma yapılmıştır. Abadi vd. (2017), su tüketimi davranışının dinamiklerini öğrenmek ve çeşitli dışsal değişkenlere dayalı olan gelecekteki günlük tüketim davranışlarını tahmin edebilmek için “Homojen Olmayan Markov Modelleri” ne dayalı yeni bir yaklaşım önermiştir. Burada kullanılan veriler kategorize edilmiş zaman serileridir. Her kategori belirli bir günlük tüketim davranışına karşılık gelir. Deneyler, Fransa'daki bir su idaresi tarafından sağlanan gerçek bir veri seti üzerinde gerçekleştirilmiştir. Önerilen model ile elde edilen tahmin sonuçları, Homojen Markov modelinin sağladığı sonuçlarla karşılaştırılmıştır. Bu tahmine dayalı sınıflandırma, su kaynaklarını daha iyi yönetmek ve tüketici gereksinimlerine yanıt vermek amacıyla su işletmeleri için yararlı olabileceği öngörülmüştür.

Merta ve Fikejz (2019), su sızıntısını tespit etmek için makine öğreniminin kullanımı ile ilgilenmiştir. Veri kaynağı olarak, su tüketiminin izlenmesini sağlayan akıllı su sayacı kullanılmıştır. Bu verilere ve analizlerine dayanarak, modeli oluşturmak için, sadece girdi parametreleri değil, aynı zamanda modelin yapısını da bilmesi gereken sembolik bir regresyon, kullanılmıştır. Modele uygun bir fonksiyon ve standart sapma bulduktan sonra, gerekli hassasiyeti ayarlayarak, izlenen zaman grafiğinde anormal su tüketimlerini tespit etmek mümkündür. Akıllı su sayacında ayrıca bir küresel vana bulunduğundan, ani bir su sızıntısı tespit edilirse, su sayacı ana beslemeyi otonom olarak kapatabilir ve böylece büyük hasarı önleyebilir.

Malinowski ve Povinelli (2020), yalnızca su sayaçları tarafından toplanarak ölçülen akış verilerini kullanarak bir su hizmeti müşterisi popülasyonunu benzerlik ve farklılıklarına göre gruplara ayırma ihtiyacını ele almaktadır. Kümeleme işleminde, gruplar benzer tüketim davranışı modellerine sahip müşterileri temsil eder. Kümeleme, Kuzey Amerika'daki ayrı ayrı ölçülen tüketimler için "normal" ve "olağandışı" müşteri davranış kalıpları hakkında fikir verir. Bu çalışmanın katkıları sadece yeni bir çalışmayı temsil etmekle kalmaz, aynı zamanda kamu hizmeti endüstrisi için pratik bir sorunu da çözer. Bu makale, bir kamu kuruluşunun sınırlı insani, finansal, hesaplamalı ve çevresel kaynaklarına uyum sağlamak için tasarlanmıştır. Bu çalışma, yeniden yapılandırılmış bir faz uzayında Gauss karışım modelleri üzerindeki ölçüm bilgisini kullanan bir hiyerarşik kümeleme yöntemi sunar. Gauss karışım modellerini karşılaştırmak için önerilen ağırlıklı ölçüm varyasyonu, istatistiksel dağılımları büyük varyasyonlu davranışlara göre daha kompakt olan davranışları vurgular ve mevcut karşılaştırma seçeneklerine yeni bir katkı sağlar. Kümeleme sonuçlarının makullüğünü ve tutarlılığını göstermek için hem sentetik hem de gerçek verilerle birkaç deney gerçekleştirilmiştir.

Öneri sistemleri müşterilerin karar vermesine yardımcı olur. Ancak, su endüstrisinde dijital teknolojinin henüz tam benimsenmemesi, evsel su kullanıcıları için böyle bir sistemin olmadığı anlamına geliyor. Su endüstrisi için böyle bir sistem, tüketicilere akıllı su sayaçlarından alınan geçmiş verilerine dayanarak suyu korumanın en etkili yollarını önerecektir. Büyükşehir bölgelerindeki su hizmetleri, yoğun saatlerde düşük basınç veya kuraklık sırasında su kıtlığı gibi sorunların üstesinden gelebilmek için su talebini yönetebilmelidir. Müşteriler için avantajlarından biri para tasarrufudur. Rahim, vd. (2019) tarafından yapılan çalışmada, bir tavsiye sistemi prototipinin yeni bir vizyonunu sunmakta ve hem tüketiciler hem de su hizmeti şirketleri için faydaları tartışılmaktadır. Bu tür bir sistemin başarısı, yararlı ve zamanında tavsiyelerde bulunmak için bir sonraki büyük su kullanımının zamanını tahmin etme yeteneğine bağlıdır. Bu nedenle, prototip, 83 hane için önemli su tüketimi olaylarını (yani duşlar, banyolar, sulama vb.) öngören Uzun Kısa Süreli Belleğe (LSTM) dayalı bir sinir ağına dayanıyor. İlk sonuçlar, LSTM'nin hataların ortalama kare kökü (RMSE) 0.403 olan yararlı bir tahmin yöntemi olduğunu göstermektedir.

Son yıllarda, kıt su kaynakları, insan türünün varlığını ve herhangi bir ulusun ilerlemesini tehlikeye atan temel sorunlardan biri haline geldi. Nasser vd. (2020) tarafından yapılan araştırmada, akıllı su sayaçları Kahire'deki bir bölgesel alanda uygulanmış, veriler belirli aralıklarla toplanıp anında buluta gönderilmiştir. Çözüm modeli olarak, mikro hizmetler ve kapsayıcılara dayalı bir IoT (Nesnelerin İnterneti) uygulaması kullanır. Tasarım, verileri depolamak için gerçek zamanlı akışı izler. Elde edilen verileri analiz etmek için, LSTM (Uzun Kısa Süreli Bellek) kullanarak su tüketimi modellenmiştir. Tasarlanan model, gelecekteki su talebinin tahmininde kullanılmak üzere doğrulanmış ve test edilmiştir. Ayrıca, LSTM'nin üstün olduğu kanıtlanmış karşılaştırmalı bir analiz için, zaman serisi tahmin uygulamalarında yaygın olarak kullanılan SVR (Destek Vektör Regresyonu) ve Rastgele Orman (Random Forest) olmak üzere iki alternatif makine öğrenme yöntemi kullanılmıştır. Sistem öğelerinin uygun şekilde entegrasyonu, önerilen sistem başarısının anahtarıdır. Tasarlanan sistem, başarıya bağlı olarak ulusal ölçekte uygulanabilir. Bu, tüketicilerin talebinin en iyi şekilde yönetilmesini sağlayabilir ve su altyapısının kullanımını iyileştirebilir. Önerilen yöntem, su kaynakları yönetiminde kullanılabilecek çeşitli

senaryolar için bir test ortamı sunar.

İbrahim vd. (2020), kaynak tespiti, depolama, işleme ve dağıtım maliyetlerini düşürmek için gelecekteki talebi tahmin etmeye odaklanan bir su yönetimi stratejisi oluşturma konusunda çalışmıştır. SVR ve ARIMA (Otoregresif Bütünleşik Hareketli Ortalama) kullanarak su talebi tahmini için karşılaştırmalı bir çalışma sunmuştur. Çalışma, Kuveyt Eyaleti günlük su tüketimi üzerine yapılmıştır. Sonuca göre, ARIMA'nın MAPE (Ortalama Mutlak Yüzde Hata) ve RMSE'ye (Kök Ortalama Kare Hata) sahip iken, SVR tahmini su talebinin, gerçek su tüketimine göre sapmasını gösteren MAPE ve RMSE'ye sahiptir.

Mounce ve Boxall (2010), çalışmalarında akış ve basınç izleme tekniklerini kullanarak patlamalardan kaynaklı sızıntıyı tespit eden uygulamayı gerçekleştirmiştir. Uygulama, iletişim yazılımı sayesinde veri tabanına toplanan veriler ve analiz eden bir MATLAB uygulamasından oluşur. Çevrimiçi sistem, saatte bir yapılan veri kümesi analizi ile otomatik uyarılar oluşturur. Analizlerde zaman serisi tahmini ve sınıflandırması kullanılmıştır.

Yıldız ve Aytekin (2019) çalışmalarında Zaman ve Tüketim verileri ile ARIMA, doğrusal regresyon destek vektör regresyonu gibi çeşitli istatistiksel ve makine öğrenmesi gibi yöntemleri kullanarak su talep tahmini yapmıştır.

Zaman serisi analizi, zamana göre sıralı verilerden anlamlı bilgiler çıkarmayı amaçlayan bir dizi görevi ifade eder. Çıkarılan bilgi sadece geçmiş davranışı teşhis etmek için değil, aynı zamanda geleceği tahmin etmek için de kullanılabilir. Zaman serisi analizinin yaygın olarak bilinen örnekleri arasında sınıflandırma, kümeleme, tahmin ve anormallik tespiti yer alır. Verilerden beklenmeyen öğeleri veya olayları belirleme süreci olan anormallik tespiti, birçok araştırmacı ve uygulayıcının ilgi alanı haline gelmiş olup veri madenciliği ve kalite güvencesindeki ana görevlerden biri haline gelmiştir. Çok değişkenli zaman serisi verilerinde anormallik tespiti, zamansal bağımlılıkların ve değişkenler arasındaki ilişkilerin aynı anda dikkate alınmasını gerektirdiğinden özel bir zorluk teşkil eder. Son zamanlarda derin öğrenme temelli çalışmalar bu alanda etkileyici ilerlemeler kaydetmiştir. Choi vd. (2021), zaman serisi verilerinde anormallik tespiti hakkında yapılan son uygulamaları gözden geçirmiştir. Ayrıca, Güvenli Su Arıtma (SWaT) ve Su Dağıtım (WADI) gibi çeşitli veri setleri kullanarak derin öğrenme tabanlı anormallik algılama modellerini karşılaştırmalı olarak analiz etmiştir. Temel olarak F1 puanına bakılıp buna göre yöntemler sıralanmıştır. Sonuç olarak, tüm veri kümeleri için herkese uyan net bir yöntem olmadığını ve LSTM tabanlı yöntemlerin, geçmiş zaman verilerinin özellik haritalarından zamansal bağımlılıkları yakalamada başarılı olduğunu, yani anormallik tespitinde başarılı olduğunu belirtmiştir.

Li vd. (2019), zaman serisi dağılımlarının zamansal korelasyonunu yakalamak için LSTM'yi temel modeller olarak kullanan, Generative Adversarial Networks (GAN)'a dayalı bir denetimsiz çok değişkenli anormallik tespit yöntemi önermiştir. Anormallikleri tespit etmek için DR-skoru adı verilen yeni bir anormallik skoru kullanılarak GAN'dan yararlanılmıştır. Bu denemeler SWaT ve WADI veri setleri üzerinde yapılmıştır Ortaya çıkardıkları yöntem Multivariate Anomaly Detection with GAN (MAD-GAN) adını vermişlerdir. Swat veri seti için MAD-GAN 0.81 ile en

yüksek F1 skorunu elde etmiştir. WADI veri seti için 0.62 F1 skoru ile diğer yöntemlerin önüne geçmiştir. Bunun nedeni karmaşık zaman serilerinin LSTM'nin diğer yöntemlere göre daha iyi öğrenmesidir.

Yapılan literatür taramasındaki incelenen çalışmalarda görüldüğü üzere, su tüketim verilerinin analizinde makine öğrenmesi ve yapay zeka kullanımı son zamanlarda yaygınlaşmıştır. Yapılan çalışmada, geçmişteki çalışmalardan daha verimli analiz yöntemlerini tespit etmek amacıyla farklı bölgelere yerleştirilen akıllı sayaçlardan veriler toplanıp analiz edilmiştir.



3. MATERYAL VE METOT

Yapılan çalışma, su tüketim verilerinin oluşturulması amacıyla TUBİTAK ARDEB 1005 - Ulusal Yeni Fikirler ve Ürünler Araştırma Destek Programı tarafından desteklenen 118Y10 nolu proje kapsamında belirlenen bölgelere Çizelge 3.1’de gösterilen 29 adet akıllı su sayacının montajı ile başlamıştır. Bu sayaçlar Manas Enerji Yönetimi Sanayi Ve Ticaret Anonim Şirketi’nden temin edilmiştir.

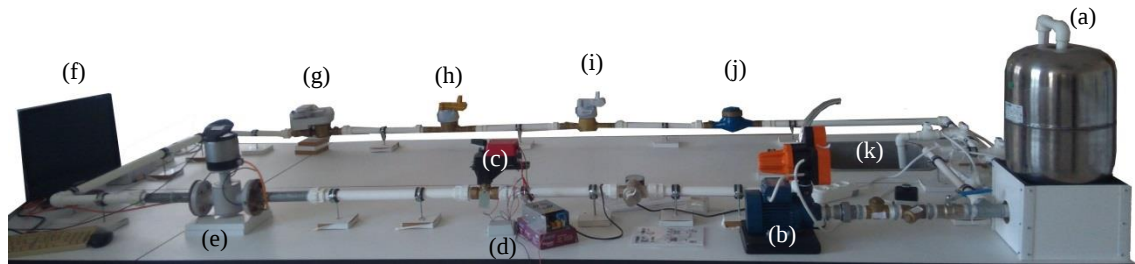
Çizelge 3.1. Su abonelerinin sınıfları

29 Adet Su Abonesi				
Okullar	Camiler	Mezarlıklar	Parklar	Tuvaletler
4 Adet	4 Adet	3 Adet	4 Adet	14 Adet

Su abonelerindeki insan kaynaklı aşırı tüketimlerin ve son tüketici tarafındaki arızalardan kaynaklı sızıntıların tespiti, azaltılması ve yönetimi için akıllı sayaçlarla ortak kullanılabilecek bir otomasyon sistemi geliştirilmiştir. Otomasyonun genel amacı su sayaçlarından alınan tüketim verisindeki aşırı tüketim ve sızıntıları azaltmaktır. Tasarruf için analiz sonuçlarına göre su akış miktarı ayarlanmıştır. Su abonelerinin aşırı su tüketimini ve arıza kaynaklı sızıntıları tespit edebilmek için analiz yapılması ve bu analiz sonuçlarına göre kullanıcı onayının da alınarak elektronik vana oranlarının ayarlanması sağlanmıştır. Sistem, uyarı koşullarına bağlı olarak, kullanıcıyı bilgilendirmektedir. Kullanıcı tarafından (cep telefonu, tablet, bilgisayar ile) alınan onay ile analiz sonucunda bulunan vana oranı, elektronik oransal vanaya uygulanmaktadır. Bunun sonucunda, aşırı tüketim ve sızıntılara karşı önlem olarak, vana kısmen veya tamamen kapatılmış olur. Bu şekilde su tasarrufunun sağlanması için müdahale mümkün olmuştur.

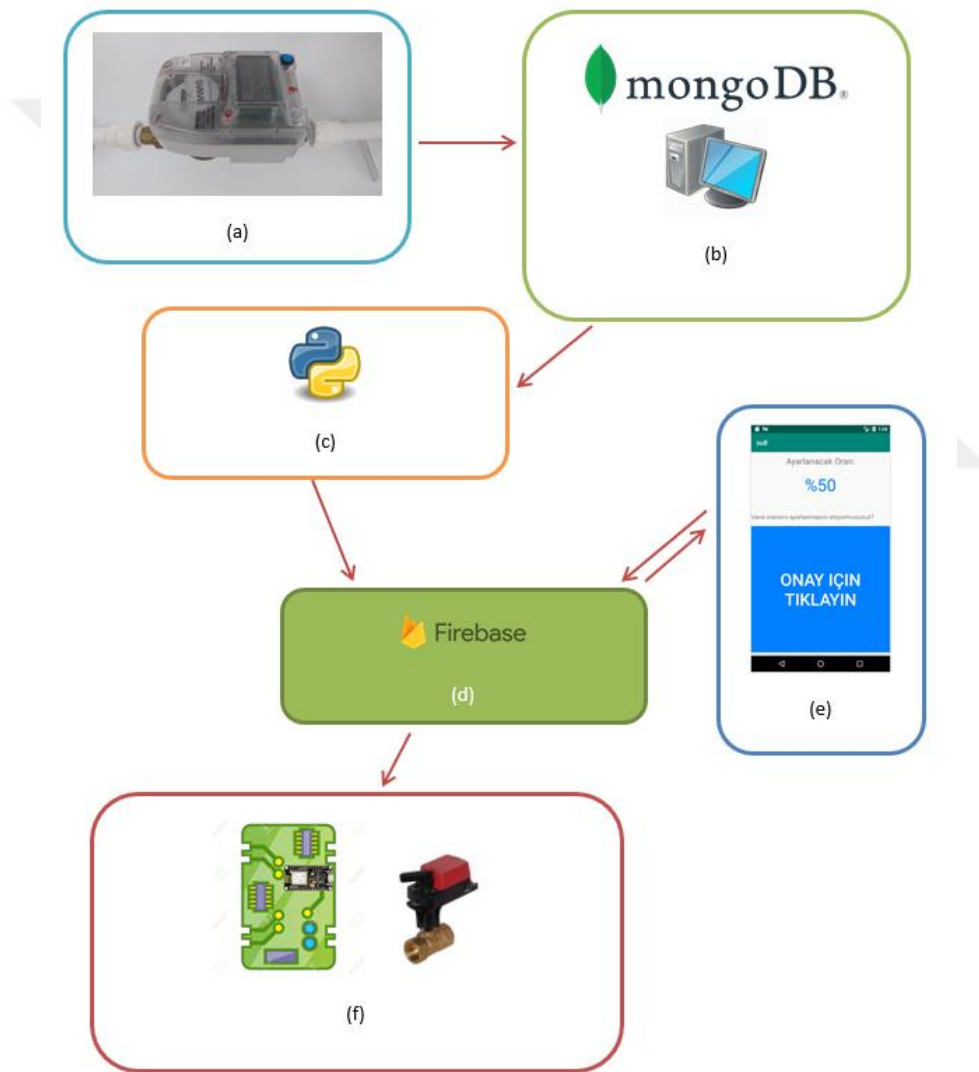
3.1. Otomasyon Sistemine Genel Bakış

Su tüketim analizlerinin sonucuna göre oransal vananın ayarlanmasının son tüketicideki etkisini gözlemleyebilmek için laboratuvar ortamında Şekil 3.1’de görülen su hattı simülasyonu kurulmuştur.



Şekil 3.1. a) Su deposu; b) Hidrofor; c) Oransal vana; d) IoT kontrol cihazı; e) Elektromanyetik debimetre; f) Debinin izlenebildiği ekran; g) GSM tabanlı akıllı su sayacı; h) LORA tabanlı akıllı su sayacı; i) NB-IoT tabanlı akıllı su sayacı; j) Mekanik su sayacı; k) Tahliye noktası

Aynı şehir hattındaki su hattında olduğu gibi, simülasyonda da su basıncının her noktada aynı ve sabit değerde olması gerekir. Bunun için simülasyon hattındaki su basıncı düşüncü, hidrofor depodaki suyu pompalamaya başlar. Su basıncı istenilen seviyeye gelince basınç anahtarı hidroforun enerjisini keser. Böylece hattaki su basıncı sabitlenmiş olur. Akıllı su sayacının seçimi için, çeşitli haberleşme yöntemlerini kullanan sayaçlar laboratuvar koşullarında denenmiştir. Elektromanyetik debimetre ve mekanik su sayacı ile doğrulama işlemi yapılmıştır. Akıllı sayaçlar, sağlıklı veri aktarma yeteneği, ölçüm doğruluğu ve yaygın kullanımı açısından değerlendirilmiştir. Sonuç olarak GSM tabanlı akıllı su sayaçlarının sahaya montajı, uygun görülmüştür. Otomasyonun şematik gösterimi Şekil 3.2’de gösterildiği gibidir.



Şekil 3.2. a) Sayaç verileri; b) Sayaç verilerinin server bilgisayardaki veri tabanına iletilmesi; c) Otomasyon programı; d) İnternetteki veri tabanı; e) Android uygulama onayı; f) IOT tabanlı Elektronik Kart ile Oransal Vana Kontrolü

a $\rightarrow$ b ok yönünde görüldüğü gibi, sahada belirli yerlerde bulunan sayaçların index, tüketim, zaman gibi verilerin server bilgisayara yönlendirilmesi ile değerler periyodik olarak yerel veri tabanına kaydedilmiş olur.

b $\rightarrow$ c ok yönünde görüldüğü gibi, okuyla gösterildiği gibi, yerel veri tabanında bulunan veriler ham haliyle Python ile yazılmış otomasyon programında işlenmek üzere kullanılır.

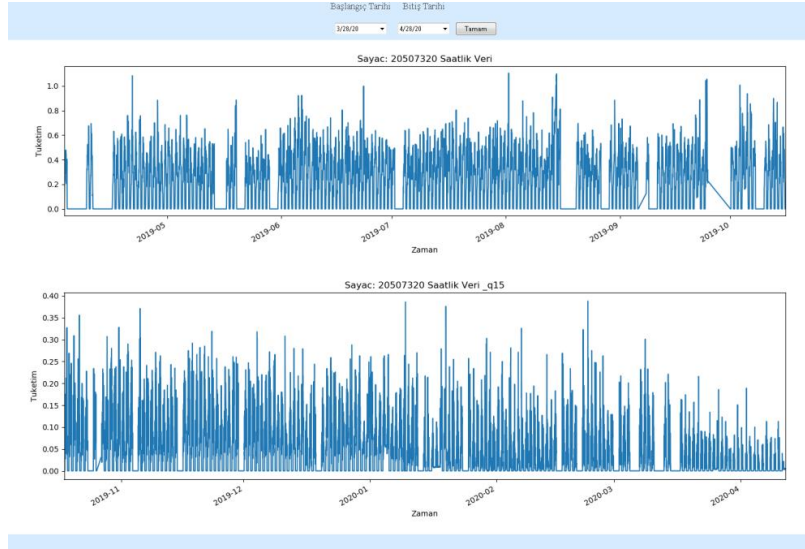
c $\rightarrow$ d ok yönünde görüldüğü gibi, otomasyon programına aktarılan veriler düzenli hale getirildikten sonra analizi yapılır. Analiz sonucu ortaya çıkan sızıntı, aşırı tüketim ve tepe tüketim değerlerinden yola çıkarak vana oran değeri bulunur. Eğer son analiz sonucunda bulunan sızıntı, aşırı tüketim veya tepe tüketim değeri yüksek ise vana oranı küçük bir değer bulunur.

d $\rightarrow$ e ve d $\leftarrow$ e ok yönlerinde görüldüğü gibi, kullanıcıya analiz zamanlarına bağlı olarak bildirim gönderilmektedir. Kullanıcı bu bildirim tıklarsa onay sayfasına yönlendirilir. Bu sayfada, ayarlanacak olan vana oranı gerçek zamanlı olarak veri tabanından alınıp ekranda gösterilmektedir. Onay butonuna tıklanırsa İnternetteki Veri Tabanında bulunan onay ile ilgili alanın değeri “olumlu” olarak değiştirilir.

d $\rightarrow$ f ok yönünde görüldüğü gibi, IoT modülü için geliştirilen gömülü yazılım sayesinde İnternetteki Veri Tabanı gerçek zamanlı olarak kontrol edilir. Onay ile ilgili alan, “olumlu” olarak değiştirilmiş ise vana oran değeri analog voltaj değeri olarak dış dünyaya aktarılır. Elektronik tasarım sayesinde kontrol voltajı oransal vana kontrol edilebileceği değer aralıklarına göre değiştirilir. Voltaj ile kontrol edilebilen oransal vana sayesinde ayarlanmak istenen vana oran değeri elektronik olarak ayarlanmış olur.

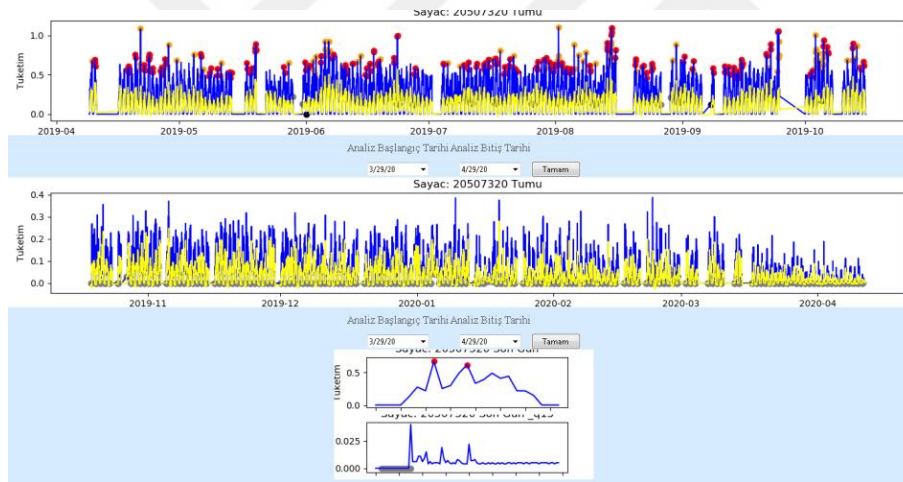
3.2. Otomasyon Arayüzü

Şekil 3.3’te görüldüğü gibi otomasyon sisteminde her abone için tüketim değerleri gözlemlenebilir. Otomasyon sistemi saatlik ve 15 dakikalık olan tüketim değerlerini algılayıp ayırır. Her değerın saatlik ya da 15 dakikalık olduğu bilgisi ham verinin ayrıştırılması sırasında tespit edilir. Grafikte gösterilen bu değerlerde kayıplar olabilir. Çünkü veri tabanında düzensiz ve kayıp veriler elenmiştir.



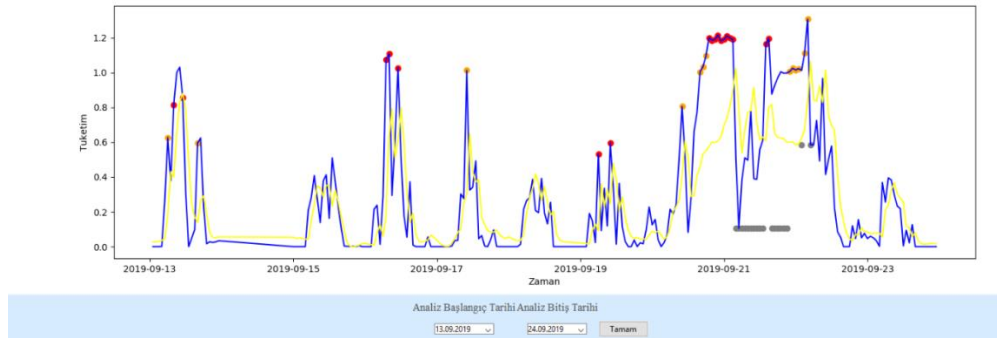
Şekil 3.3. Saatlik ve 15 dakikalık tüketim değerlerinin gösterimi

Şekil 3.4’te görüldüğü gibi otomasyon sistemi verileri analiz edilip analiz sonuçlarını grafikte gösterebilmektedir.



Şekil 3.4. Analizlerin yapılması ve sonuçların grafikte gösterimi

Arayüz, verilerin ve analiz değerlerinin istenilen tarih arasında gözlemlenmesi için filtre imkanı sunmaktadır. Kullanıcı başlangıç ve bitiş tarihlerini girerek Şekil 3.5’teki gibi gözlem yapılabilir.



Şekil 3.5. Analiz sonuçlarının belirli tarih arasında filtrelenmesi

Seçilen tarih aralıkları arasındaki normal ve anormal değerlerin miktarının gözlemlenebilmesi için sızıntı, aşırı, tepe ve normal değer miktarı arayüze aktarılır. Şekil 3.6’da gösterilen bu bilgiler aynı zamanda vana oranını belirler.

Normal: 725.22 m3	Sızıntı: 85.83 m3	Peak: 6.18 m3	Aşırı: 22.15 m3
-------------------	-------------------	---------------	-----------------

Şekil 3.6. Belirlenen tarih arasında abonelerin normal ve anormal tüketimlerinin miktarının gösterilmesi

Analiz sonuçlarına göre durum bilgisinin kolaylıkla gözlemlenebilmesi için renklendirilmesi yapılan abone durumları Şekil 3.7’de görüldüğü gibidir.

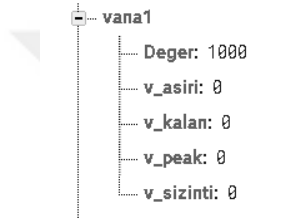
Saatlik Sayaç Seçiniz	Vanalar
20507322 - 6405 - m - arapsuyu me	n - vana - 6
20507214 - 5839 - o - akdeniz üniv	n - vana 1
20507328 - 6411 - c - akdeniz soga	n - vana2-
20507316 - 6399 - c - sezipler cami	n - vana3-
20507320 - 6403 - w - muratpaşa w	s - vana4-
20507324 - 6407 - w - obla bay wc	n - vana5-
20507325 - 6408 - p - karaaahoglu p	n - vana6-
20507330 - 6413 - m - duaci mezarl	n - vana7-
20507318 - 6401 - c - 15 termnu c	n - vana8-
20507323 - 6406 - w - aköla bayan	n - vana9-
20507317 - 6400 - m - fahri mezarl	n - vana10-
20507321 - 6404 - w - yakut wc - C	n - vana11-
20507327 - 6410 - w - karakas wc	s - vana12-
15dk'lık Sayaç Seçünüz	Vanalar
20507329 - 6412 - p - doremealti at	n - vana _
20507319 - 6402 - c - yenayok meri	n - vana _1
20507326 - 6409 - n - null - 0 - n -	n - vana _2
20501728 - 6432 - o - cankaya oku	p - vana _3
20501727 - 6431 - o - bahcesehir ol	s - vana _4
20507333 - 6446 - s - tungut otlan p	s - vana _5
20507335 - 6448 - p - kent orman p	s - vana _6
20507331 - 6444 - w - muh fak 2 -	n - vana _7
20507332 - 6445 - w - muh fak 1 -	s - vana _8
20508242 - 6449 - p - meris park	n - vana _9
20501718 - 6443 - w - muh fak 3 -	n - vana _10
20501714 - 6442 - w - yenikoc wc	s - vana _11
20507334 - 6447 - n - null - 0 - n -	n - vana _12
20501731 - 6433 - w - muh fak zemi	n - vana _13
20508282 - 7433 - w - yakut yeni	s - vana _14
20508283 - 7432 - w - oliibaayan y	n - vana _15

Şekil 3.7. Abone durumlarının gözlenmesi

3.3. Akışın Analiz Sonuçlarına Göre Ayarlanması

3.3.1. Vana oranlarının belirlenmesi ve veri tabanına yazılması

Analiz sonucunda sızıntı için; son gündeki sızıntı değerinin tüm sızıntı değerlerinin en büyük değerine oranı göz önünde bulundurularak vana oranı belirlenmiştir. Örneğin; bir abone için son gündeki saatlik sızıntı değeri ortalama 0.25 m^3 ise ve o abone için tüm analiz sonucundaki en büyük saatlik sızıntı değeri 0.5 m^3 ise sızıntı oranı $0.2/0.6=0.3$ tir. Bu değere göre vana oranı $1-0.3=0.7$, yani %70 açık olarak belirlenmektedir. Sızıntı oranı ile vana açıklık oranı ters orantılıdır. Çünkü son gündeki sızıntı artarsa vana daha az açık olmalıdır. Son günde aşırı tüketimin olduğu abonelerde de aşırı tüketim oranı göz önünde bulundurularak vana oranı belirlenmektedir. Daha sonra Şekil 3.8’de görüldüğü gibi belirlenen vana oranı internete aktarılmaktadır.

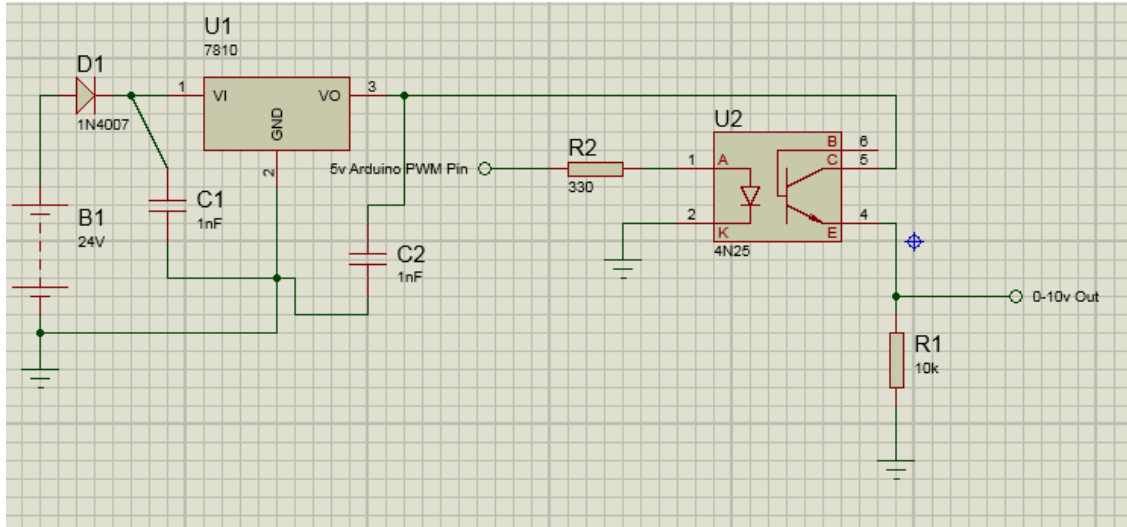


Şekil 3.8. Vana oran verilerinin kaydedildiği Firebase veri tabanı

3.3.2. Vana oranının fiziksel olarak ayarlanması

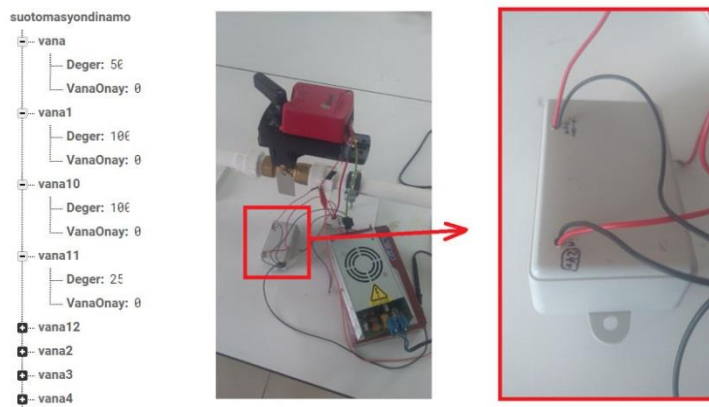
Analiz sonuçlarının ve verilerin belirtilen dönem aralıklarında raporlanması, uyarı koşullarına bağlı olarak belirli bildirimlerin yapılması ve kullanıcı onayı alınarak vana oranlarının değiştirilmesi gibi müdahalelerin yapılması için yazılım geliştirilmiştir. Su tüketim verileri ve analiz sonuçları server bilgisayara ve internetteki veri tabanına da kaydedilmektedir. Veriler kullanıcı tarafından (cep telefonu, tablet, bilgisayar ile) izlenebilmektedir.

Analog sinyal ile kontrol edilen oransal vananın kontrol voltajı 0 ile 10 Volt arasındadır. Fakat IoT modülünün analog çıkışı 0 ile 5 Volt arasında değişebilmektedir. Bu yüzden, IoT modülü ile oransal vananın kontrol edilmesi için 0-5 Voltluk sinyal voltajı ile 0-10 Voltluk sinyal voltajını kontrol eden bir elektronik devre tasarlanmıştır.(Şekil 3.9)



Şekil 3.9. Vana oranının ayarı için elektronik devre

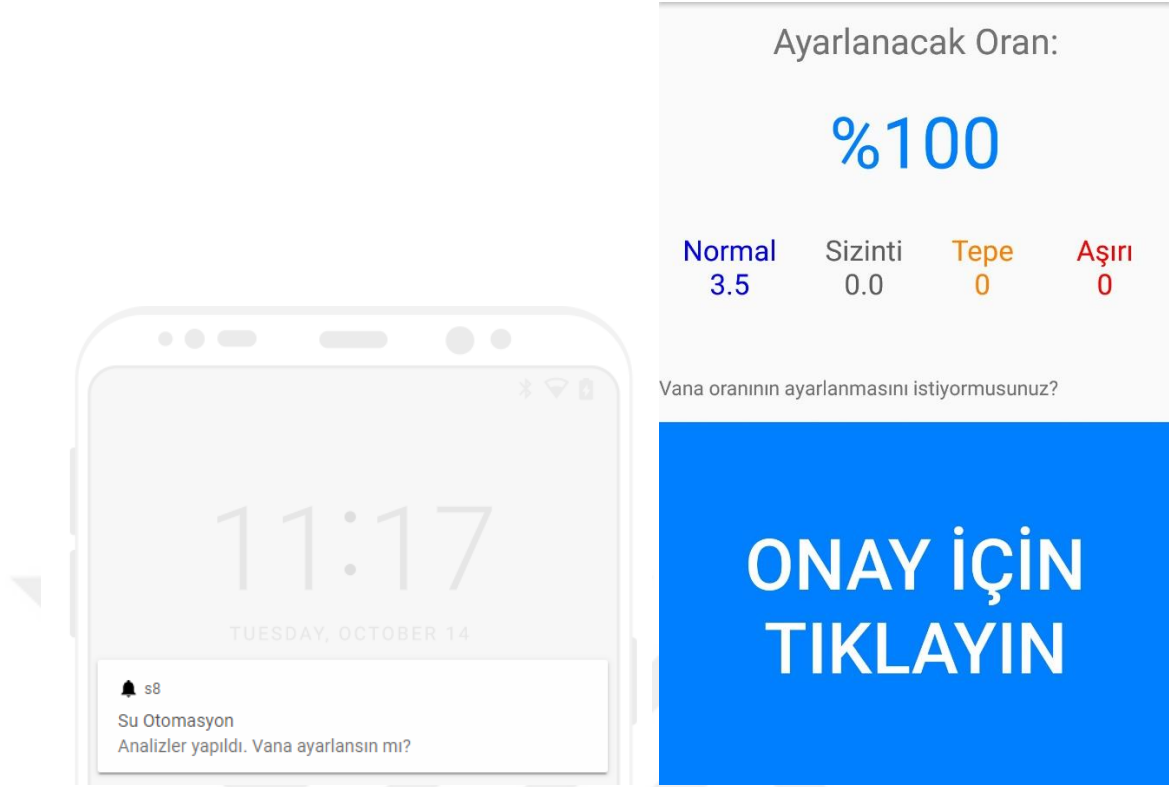
Elektronik devre, Şekil 3.10 da görüldüğü gibi su geçirmeyen devre kutusunun içine yerleştirilmiştir. Bu kutu sayesinde elektronik devrenin çevre koşullarından etkilenmemesi için su yalıtımı sağlanmıştır.



Şekil 3.10. Vana oranının ayarlanması için kontrol modülü

3.3.3. Kullanıcı onayı için Android uygulama

Geliştirilen Android yazılım sayesinde kullanıcıya her gün Android cihazı tarafından analizlerin tamamlandığı bildirilmektedir. Bildirime tıklanınca kullanıcı onayı için Şekil 3.11’de görülen Android arayüzü açılmaktadır.



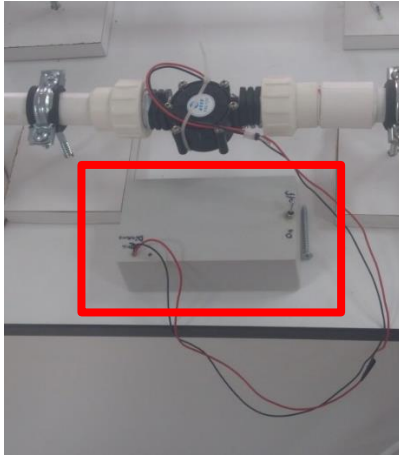
Şekil 3.11. Android uygulama bildirimi ve onay ekranı

Analiz sonucunda ortaya çıkan vana oranının değeri server bilgisayardaki arayüz yazılımı tarafından internetteki veri tabanına yazılmıştır. Bu oran Android arayüzüne alınıp kullanıcı ayarlanacak olan vana oranı konusunda bilgilendirilmiştir. Onay butonuna tıklanması ile internetteki veri tabanının gerekli alanı değiştirilir. Bu değişimi gömülü yazılım barındıran elektronik tasarım gerçek zamanlı olarak takip ettiği için vana oranının ayarlanması için gerekli işlemler yapılmış olur.

3.4. Sistem Enerjisinin Su Akışından Karşılanması

Çalışmaların sonucunda ortaya çıkan kontrol modülünün enerji ihtiyacının karşılanması için her abone başına bir elektrik hattı çekilmesi gerekseydi, sayaç sisteminin dönüşümü için maliyet ve iş yükü oldukça yükselirdi. Eğer şarj edilemeyen pil takılması tasarlansaydı, bu pillerin periyodik olarak değişimi de iş gücü ihtiyacını ve sarf malzeme tüketimini artırırdı. Bu tür problemlerin çözümü için bir güç modülü tasarımına ihtiyaç duyulmuştur. IoT modülü ve vana kontrol devresini içinde barındıran kontrol modülünün enerjisini karşılamak için Şekil 3.12’de görülen modül tasarlanmıştır. Bu güç modülü, su dinamosu, şarj devresi ve bataryalardan oluşmaktadır. Su hattına bağlı olan su dinamosu 12 Volt 0.5 Amper enerji üretme kapasitesine sahiptir. Bu enerji Li-Ion (Lityum İyon) pillerin şarj edilmesi için kullanılmaktadır. Voltaj değeri 12 Volt olan su dinamosunun bağlantı uçları şarj devresine girmektedir. Şarj sistemi için 12 Volt giriş ile 3.7 Volt’luk bir Li-Ion bataryayı doldurma görevini üstlenen elektronik devre tasarlanmıştır. Bu elektronik devre, pil şarj modülleri ve voltaj yükseltici modülleri barındırmaktadır. Voltaj yükseltici modüllerin kullanım amacı, pil voltajı ile kontrol modülünü beslemektir. Kontrol modülü 24 Volt ile çalıştığı için pil

besleme geriliminin yükseltilmesi gerekir. Pil voltajının 3.7 Volt'tan 24 Volta yükseltilmesi için DC Voltaj yükseltici modül kullanılmıştır.



Şekil 3.12. Su dinamosu ile şarj edilen bir güç besleme modülü

3.5. Veri Setinin Ön Düzenlemesi

Manas Enerji Yönetimi Sanayi Ve Ticaret Anonim Şirketi'nden temin edilen 29 adet sayaçtan 13 tanesi saatlik çözünürlüklü, 16 tanesi de 15 dakikalık çözünürlüklü veri sağlamaktadır. Verilerin oluşturulması, bölgesel veri izleme noktalarına yerleştirilen saatlik ve 15 dakikalık çözünürlüklü 2 farklı türdeki akıllı su sayaçları ile yapılan tüketim ölçümleri ile sağlanmaktadır. Veri kümesinin oluşturulması için belirli bir süre içinde tüketim verileri toplanmıştır.

Makine öğrenimi algoritmaları için ham veriler değerlendirilememektedir. Bu yüzden, bireysel makine öğrenimi algoritmalarının gereksinimlerini karşılamak için veriler dönüştürülmelidir. Dahası, bir tahmine dayalı modelleme projesinde mevcut kaynaklar göz önüne alındığında en iyi performansı elde etmek için, tahmin probleminin bilinmeyen temel yapısını, öğrenme algoritmalarına en iyi şekilde gösteren veriler için bir temsil seçilmelidir. Tahmine dayalı modelleme projeleri, verilerden öğrenmeyi içerir. Veriler, çözmek istenilen sorunu karakterize eden etki alanındaki örneklerle veya vakalara atıfta bulunur. Denetimli öğrenmede veriler, her örneğin bir modele sağlanacak bir girdi unsuruna ve modelin tahmin etmesi beklenen bir çıktı veya hedef unsuruna sahip olduğu örneklerden oluşur. Bir modele giriş için kullanılan sütunlara giriş değişkenleri, tahmin edilecek hedefi içeren sütuna ise çıkış değişkeni denir. Bir modeli eğitmek için kullanılan verilere eğitim veri seti, modeli değerlendirmek için kullanılan verilere ise test veri seti denir. Girdi değişkenleri, bir tahmin yapmak için bir modele sağlanan veri kümesindeki sütunlardır. Çıkış değişkenleri, bir model tarafından tahmin edilecek veri kümesindeki sütunlardır. Tahmine dayalı makine öğrenimi projeleri her zaman bir tür veri hazırlığı içerir. Bir veri kümesi için gereken özel veri hazırlığı, değişken türleri gibi veri özelliklerine ve ayrıca bunları modellemek için kullanılacak algoritmalara bağlıdır. Bununla birlikte, yapılandırılmış verilere uygulanabilecek bir dizi standart veri hazırlama algoritması vardır. Bu veri hazırlama algoritmaları, belirli bir proje için teknikleri karşılaştırırken ve seçerken yardımcı olabilecek bir çerçeveye göre düzenlenebilir veya gruplandırılabilir.

Veri hazırlamayı, ham verilerin modellemeye daha uygun bir forma dönüştürülmesi olarak tanımlanabilir. Bununla birlikte, bir tahmine dayalı modelleme projesinde, veri hazırlama adımından önce ve sonra önemli olan ve yapılacak veri hazırlığı adımları vardır. Herhangi bir proje için adımlar arasında istenilen sıralama yapılabilir. Ancak tüm projeler aynı genel adımlara sahiptir.

- Adım 1: Problemi Tanımlama.
- Adım 2: Verileri Hazırlama.
- Adım 3: Modelleri Değerlendirme.
- Adım 4: Modeli Sonlandırma.

Bir makine öğrenimi projesinde veri hazırlama adımı sırasında kullanabilecek ortak veya standart görevler vardır. Gerçekleştirilen veri hazırlama türleri, verilere bağlıdır. Bu görevler şunları içerir:

- Veri Temizleme: Verilerdeki hataları veya hataları belirleme ve düzeltme.
- Özellik Seçimi: Görevle en ilgili olan girdi değişkenlerini belirleme.
- Veri Dönüşümleri: Değişkenlerin ölçeğini veya dağılımını değiştirme.
- Özellik Mühendisliği: Mevcut verilerden yeni değişkenler türetme.
- Boyut Azaltma: Verileri kompakt hale getirme.

Bu, belirli bir proje üzerinde düşünülebilecek farklı veri hazırlama algoritmaları hakkında düşünmek ve bu algoritmalarda gezinmek için kullanabilecek kaba bir çerçeve sağlar.

3.5.1. Veri temizleme

Veri temizleme, genellikle diğer veri hazırlama işlemlerinden önce gerçekleştirilen bir işlemdir.

Dağınık, gürültülü, bozuk veya hatalı veriler belirlenip sonra ele alınabilir. Bu, bir satırın veya sütunun kaldırılmasını içerebilir. Alternatif olarak, gözlemlerin yeni değerlerle değiştirilmesini içerebilir. Bununla birlikte, gerçekleştirilebilecek genel veri temizleme işlemleri vardır. Bunlar:

- Normal verileri tanımlamak ve aykırı değerleri belirlemek için istatistikleri kullanma.
- Aynı değere sahip olan veya varyansı olmayan sütunların belirlenmesi ve kaldırılması.
- Yinelenen veri satırlarını belirleme ve bunları kaldırma.
- Boş değerleri eksik olarak işaretleme.
- İstatistikleri veya öğrenilmiş bir modeli kullanarak eksik değerleri yüklemek.

3.5.2. Özellik seçimi

Özellik seçimi, tahmin edilmekte olan hedef değişkenle en alakalı girdi özelliklerinin bir alt kümesinin seçiminde yardımcı olur.

Alakasız ve gereksiz girdi değişkenleri, öğrenme algoritmalarının dikkatini

dağıtabileceği veya yanlış yönlendirebileceği ve muhtemelen daha düşük tahmin performansına yol açabileceği için bu önemlidir. Ek olarak, örneğin mümkün olan en basit iyi performans gösteren modeli tercih etmek için yalnızca bir tahmin yapmak için gerekli olan verileri kullanarak modellerin geliştirilmesi arzu edilir.

3.5.3. Veri dönüşümleri

Veri dönüşümleri, veri değişkenlerinin türünü veya dağılımını değiştirmek için kullanılır. Bu, farklı tekniklerin geniş bir şemsiyesidir. Girdi ve çıktı değişkenlerine aynı kolaylıkla uygulanabilir.

Veri türleri kategorilere ayrılmaktadır. Bunlar:

- Sayısal veri türleri:
 - Tamsayı (Integer): Kesirli kısmı olmayan tam sayılar.
 - Reel: Kesirli sayılar.
- Kategorik veri türleri:
 - Sıralı (Ordinal): Sıralama sırasına sahip etiketler.
 - Nominal: Sıralama içermeyen etiketler.
 - Boole: Doğru ve Yanlış değerleri.

Ayrıklaştırma (Discretization) adı verilen bir süreçte sayısal bir değişken sıralı bir değişkene dönüştürmek istenilebilir. Alternatif olarak, çoğu sınıflandırma görevinde gerekli olan kategorik bir değişken tamsayılar veya boole değişkenleri olarak kodlanabilir.

- Ayrıklaştırma Dönüşümü (Discretization Transform): Sayısal bir değişkenin sıra değişkeni olarak kodlanması.
- Sıralı Dönüşüm (Ordinal Transform): Kategorik bir değişkenin bir tamsayı değişkenine kodlanması.
- One-Hot Dönüşümü (One-Hot Transform): Kategorik bir değişkenin ikili değişkenlere kodlanması.

Reel değerli sayısal değişkenler için, bir bilgisayarda temsil edilme biçimleri, 0 ile 1 aralığındadır. Veri türünün fazla çözünürlüğe sahip olduğu anlamına gelir. Bu nedenle, değişkenler normalleştirme adı verilen bu aralığa ölçeklendirmek istenebilir. Verinin bir Gauss olasılık dağılımı varsa, verileri ortalama sıfır ve standart sapması bir olan standart bir Gauss'a kaydırmak daha yararlı olabilir.

- Normalleştirme Dönüşümü (Normalization Transform): Bir değişkenin 0 ve 1 aralığında ölçeklendirilmesi.
- Standardizasyon Dönüşümü (Standardization Transform): Bir değişkenin standart bir Gauss'a ölçeklendirilmesi.

Örneğin, dağılım Gauss'a yakınsa, ancak çarpık veya kaydırılmışsa, bir güç dönüşümü kullanılarak daha uygun hale getirilebilir. Alternatif olarak, olağandışı bir doğal dağılıma sahip bir değişken üzerinde tek tip veya Gauss gibi bir olasılık dağılımını zorlamak için nicel dönüşümler kullanılabilir.

- Güç Dönüşümü (Power Transform): Bir değişkenin dağılımının Gauss'a daha uygun olacak şekilde değiştirilmesi.
- Kuantil Dönüşüm (Quantile Transform): Tek tip veya Gauss gibi bir olasılık dağılımının empoze edilmesi.

Veri dönüşümleriyle ilgili önemli bir husus, işlemlerin genellikle her değişken için ayrı ayrı gerçekleştirilmesidir. Hal böyle olunca farklı değişken türleri üzerinde farklı işlemler yapmak istenebilir.

3.5.4. Özellik mühendisliği

Özellik mühendisliği, mevcut verilerden yeni girdi değişkenleri oluşturma sürecini ifade eder.

Mühendislik özellikleri, verilerle ve veri türleriyle oldukça ilgilidir. Bu nedenle, verilerden oluşturulabilecek yeni özelliklerin belirlenmesine yardımcı olmak için genellikle bir uzmanının işbirliğini gerektirir. Bu işbirliği, genel yöntemlerle genelleme yapmayı zor bir konu haline getirmektedir. Bununla birlikte, yeniden kullanılabilir bazı teknikler vardır. Bazı durumlar için bir boole bayrağı değişkeni ekleme, ortalama gibi bir grup veya genel özet istatistik ekleme, bileşik değişkenin her bileşeni için tarih-saat gibi yeni değişkenler ekleme teknikleri buna örnektir.

İstatistiklerden elde edilen popüler bir yaklaşım, onları bir güce yükseltmek veya polinom özellikleri olarak adlandırılan diğer girdi değişkenleriyle çarpmak gibi basit bir matematiksel işlemle değiştirilen sayısal girdi değişkenlerinin kopyalarını oluşturmaktır.

Özellik mühendisliğinin teması, tek bir gözleme daha geniş bağlam eklemek veya her ikisi de girdi verileri üzerinde daha basit bir bakış açısı sağlamak amacıyla karmaşık bir değişkeni ayrıştırmaktır.

3.5.5. Boyutsal küçülme

Bir veri kümesi için girdi özelliklerinin sayısı, verinin büyüklüğü olarak kabul edilebilir. Örneğin, iki girdi değişkeni birlikte, her veri satırının o uzayda bir nokta tanımladığı iki boyutlu bir alan tanımlayabilir. Bu fikir daha sonra büyük çok boyutlu hiper hacimler oluşturmak için herhangi bir sayıda girdi değişkenine ölçeklenebilir. Sorun şu ki, bu uzay ne kadar çok boyuta sahipse (örneğin ne kadar çok girdi değişkeni), veri kümesinin o uzayın çok seyrek ve muhtemelen temsili olmayan bir örnekleme temsili etmesi o kadar olasıdır. Buna boyutsallığın laneti denir. Bu özellik seçimini motive eder. Ancak özellik seçimine bir alternatif olarak orijinal verilerin en önemli özelliklerini koruyan daha düşük boyutlu bir alan seçilebilir. Bu, genel olarak boyut indirgeme olarak adlandırılır ve özellik seçiminin bir alternatifidir. Özellik seçiminden farklı olarak, yansıtılan verilerdeki değişkenler, orijinal girdi değişkenleriyle doğrudan ilişkili değildir. Ve bu da verinin yorumlanmasını zorlaştırır.

Boyut indirgemeye yönelik en yaygın yaklaşım, bir matris çarpanlara ayırma tekniği kullanmaktır:

- PCA: Principal Component Analysis (Temel Bileşen Analizi)
- SVD: Singular Value Decomposition (Tekil Değer Ayrışımı)

Bu tekniklerin ana etkisi, girdi değişkenleri, örneğin ilişkili değişkenler arasındaki doğrusal bağımlılıkları ortadan kaldırmalarıdır. Daha düşük boyuta azaltmayı sağlayan başka yaklaşımlar da mevcuttur. Bunlar LDA ve belki de otomatik kodlayıcılar gibi model tabanlı yöntemler olarak adlandırılabilir.

- LDA: Linear Discriminant Analysis (Lineer Diskriminant Analizi)

3.6. Sayaç Verilerinin Düzenlenmesinde Kullanılan Yöntemler

Yapılan ön çalışmada kullanılan veri seti Mart 2019'dan Ekim 2019'a kadar saatlik, Kasım 2019'dan Mayıs 2020'ye kadar 15 dakikalık periyotlarla kaydedilmiştir. Tüm analizler her iki periyot için ayrı kurallar ile yapılmıştır. Abonelerin sınıfı (cami, mezarlık, tuvalet vb.) belirli olduğu için manuel olarak girilmektedir. Bununla birlikte, veri toplama noktasında, gürültü ve diğer faktörler nedeniyle, veri işlemenin etkinliğini büyük ölçüde etkilenir. Veri değerinin eksik olması vb. durumlardan kaynaklı anormal veriler ortaya çıkabilir. Bu nedenle, başarılı bir veri analizi için etkili tanımlama ve ön düzenleme gereklidir.

Analiz yöntemlerinden biri olan su tüketim tahminin amacı istenilen zaman aralığı için normal bir tüketim örneğinin üretilmesidir. Su tüketim tahmini için eğitimde gürültülü verinin kullanılması, tahmin verisinde gürültülerin ortaya çıkmasına neden olur. Yani, veri gürültüsünü ortadan kaldırmak veri tahminini büyük ölçüde etkiler.

3.6.1. Veri ön işleme

Bir metin bütünü olarak alınan veri, bir metin düzenleme algoritması ile ayrıştırılıp sadeleştirilmiştir. Ayrıştırmadan sonra Şekil 3.13'te görüldüğü gibi "meterId" değeri koşul olarak konulup değerler alınmıştır. Böylece istenilen sayacın tarih ve tüketim verisi metin olarak elde edilmiştir.

verP9x için meterId değerini girin:

```
6403
{'_id': ObjectId('5ca47ffe4b0cb2cd8b882e7'), '_class': 'com.manas.metermanager.library.model.vendor.manas.gprs.VerP9x',
'packetSize': 257, 'packetCount': 3, 'creditValidation': 50, 'dataSaveDateTime': datetime.datetime(1999, 12, 31, 22, 0),
'lastGprsError': 255, 'lastGprsConnectionReason': 1, 'currentPressure': 0, 'highPressureTreshold': 0, 'lowPressureTreshold': 0,
'highConsumptionTreshold': 0, 'hourlyLog': [{'pressure': 0, 'consumption': 0.375, 'reverseConsumption': 0.0, 'index': 173.323,
'date': datetime.datetime(2019, 4, 3, 4, 0)}], {'pressure': 0, 'consumption': 0.135, 'reverseConsumption': 0.0, 'index': 172.948,
'date': datetime.datetime(2019, 4, 3, 3, 0)}], {'pressure': 0, 'consumption': 0.0, 'reverseConsumption': 0.0, 'index':
172.81300000000002, 'date': datetime.datetime(2019, 4, 3, 2, 0)}], {'pressure': 0, 'consumption': 0.0, 'reverseConsumption': 0.0,
'index': 172.81300000000002, 'date': datetime.datetime(2019, 4, 3, 1, 0)}], {'pressure': 0, 'consumption': 0.0,
'reverseConsumption': 0.0, 'index': 172.81300000000002, 'date': datetime.datetime(2019, 4, 3, 0, 0)}], {'pressure': 0,
'consumption': 0.0, 'reverseConsumption': 0.0, 'index': 172.81300000000002, 'date': datetime.datetime(2019, 4, 2, 23, 0)}],
{'pressure': 0, 'consumption': 0.0, 'reverseConsumption': 0.0, 'index': 172.81300000000002, 'date': datetime.datetime(2019, 4, 2, 22, 0)}], {'pressure': 0, 'consumption': 0.0, 'reverseConsumption': 0.0, 'index': 172.81300000000002, 'date': datetime.datetime(2019, 4, 2, 21, 0)}], {'pressure': 0, 'consumption': 0.0, 'reverseConsumption': 0.0, 'index': 172.81300000000002, 'date':
datetime.datetime(2019, 4, 2, 20, 0)}], {'pressure': 0, 'consumption': 0.192, 'reverseConsumption': 0.0, 'index':
172.81300000000002, 'date': datetime.datetime(2019, 4, 2, 19, 0)}], {'pressure': 0, 'consumption': 0.314, 'reverseConsumption':
0.0, 'index': 172.621, 'date': datetime.datetime(2019, 4, 2, 18, 0)}], {'pressure': 0, 'consumption': 0.347, 'reverseConsumption':
0.0, 'index': 172.30700000000002, 'date': datetime.datetime(2019, 4, 2, 17, 0)}], {'pressure': 0, 'consumption': 0.429,
'reverseConsumption': 0.0, 'index': 171.96, 'date': datetime.datetime(2019, 4, 2, 16, 0)}], {'pressure': 0, 'consumption': 0.425,
'reverseConsumption': 0.0, 'index': 171.531, 'date': datetime.datetime(2019, 4, 2, 15, 0)}], {'pressure': 0, 'consumption': 0.513,
'reverseConsumption': 0.0, 'index': 171.106, 'date': datetime.datetime(2019, 4, 2, 14, 0)}], {'pressure': 0, 'consumption': 0.439,
'reverseConsumption': 0.0, 'index': 170.593, 'date': datetime.datetime(2019, 4, 2, 13, 0)}], {'pressure': 0, 'consumption': 0.582,
'reverseConsumption': 0.0, 'index': 170.154, 'date': datetime.datetime(2019, 4, 2, 12, 0)}], {'pressure': 0, 'consumption': 0.558,
'reverseConsumption': 0.0, 'index': 169.572, 'date': datetime.datetime(2019, 4, 2, 11, 0)}], {'pressure': 0, 'consumption': 0.625,
'reverseConsumption': 0.0, 'index': 169.014, 'date': datetime.datetime(2019, 4, 2, 10, 0)}], {'pressure': 0, 'consumption': 0.423,
'reverseConsumption': 0.0, 'index': 168.389, 'date': datetime.datetime(2019, 4, 2, 9, 0)}], {'pressure': 0, 'consumption': 0.339,
'reverseConsumption': 0.0, 'index': 167.966, 'date': datetime.datetime(2019, 4, 2, 8, 0)}], {'pressure': 0, 'consumption': 0.466,
'reverseConsumption': 0.0, 'index': 167.627, 'date': datetime.datetime(2019, 4, 2, 7, 0)}], {'pressure': 0, 'consumption': 0.376,
'reverseConsumption': 0.0, 'index': 167.161, 'date': datetime.datetime(2019, 4, 2, 6, 0)}], {'pressure': 0, 'consumption': 0.242,
'reverseConsumption': 0.0, 'index': 166.785, 'date': datetime.datetime(2019, 4, 2, 5, 0)}], 'dailyLog': [{'consumption':
6.78000000000001, 'reverseConsumption': 0.0, 'index': 173.323, 'date': datetime.datetime(2019, 4, 2, 5, 0)}], 'monthlyLog':
[{'consumption': 159.742, 'reverseConsumption': 0.107, 'index': 159.635, 'date': datetime.datetime(2019, 2, 28, 21, 0)},
{'consumption': 0.0, 'reverseConsumption': 0.0, 'index': 0.0, 'date': datetime.datetime(2019, 1, 31, 21, 0)}], 'meterType': 'S',
'signalQuality': 15.0, 'versionNr': 13.0, 'version': 'P13', 'imei': '353656100211871', 'ccId': '8990011103250078121',
'communicationUnitDateTime': datetime.datetime(2019, 4, 3, 5, 43, 6), 'penaltyDate': datetime.datetime(2019, 3, 10, 21, 0),
'resetDateTime': datetime.datetime(1999, 12, 31, 22, 0), 'removeDateTime': datetime.datetime(1999, 12, 31, 22, 0),
'magnetDetectionDateTime': datetime.datetime(1999, 12, 31, 22, 0), 'failureDate': datetime.datetime(1999, 12, 31, 22, 0),
'lastCreditRecordDate': datetime.datetime(2019, 4, 2, 21, 0), 'holiday1Start': datetime.datetime(2008, 12, 31, 21, 0),
'holiday1End': datetime.datetime(2008, 12, 31, 21, 0), 'holiday2Start': datetime.datetime(2008, 12, 31, 21, 0), 'holiday2End':
datetime.datetime(2008, 12, 31, 21, 0), 'pulseDataSaveDateTime': datetime.datetime(2019, 4, 2, 5, 32, 6), 'valveOpenDateTime':
datetime.datetime(2019, 2, 27, 8, 53), 'valveCloseDateTime': datetime.datetime(2019, 3, 11, 10, 18), 'dayOfWeek': 3,
'communicationUnitSerialNr': 20507320, 'totalConsumption': 173.605, 'fastDay': -1, 'issue': 'DU', 'totalReverseConsumption':
0.107, 'flag': 0, 'flag2': 16, 'setting': 136, 'decimalCount': 3, 'gprsConnectionTryCount': 45, 'hourlyDataLogIndex': 24,
'dailyDataLogIndex': -1, 'monthlyDataLogIndex': 2, 'timeSpanDataLogIndex': -1, 'q01Distance': 1920.0, 'q12Distance': 480.0,
'q23Distance': 96.0, 'q34Distance': 38.0, 'q45Distance': 19.0, 'q0Consumption': 0.041, 'q1Consumption': 6.113, 'q2Consumption':
77.066, 'q3Consumption': 89.957, 'q4Consumption': 0.34, 'q5Consumption': 0.058, 'qStartConsumption': 0.0, 'qMinConsumption': 0.0,
'qTConsumption': 0.0, 'qNConsumption': 0.0, 'qMaxConsumption': 0.0, 'remainingCredit': 999826.288, 'expendedCredit': 173.712,
'criticalCredit': 10.0, 'maximumFlow': 1320, 'testAmount': 0, 'valveOpenCount': 0, 'valveCloseCount': 2, 'hourlyLogConsumption':
0.175, 'hourlyLogReverseConsumption': 0.0, 'dailyLogConsumption': 0.0, 'dailyLogReverseConsumption': 0.0,
'monthlyLogConsumption': 13.863, 'monthlyLogReverseConsumption': 0.0, 'mechanicMeterNr': '205073', 'meterSize': 20,
'connectionPeriod': 1477, 'valveCloseStart': 8, 'valveCloseEnd': 8, 'errorPulseFactor': 200, 'batteryEndLimit': 150,
'batteryLowLimit': 157, 'displaySettings1': 71, 'displaySettings2': 129, 'valveCheckHour': 0, 'valveCheckDay': 0,
'valveCheckPeriod': 35, 'mainBatteryLevel': 194.0, 'backupBatteryLevel': 7.0, 'customerNr': '205073', 'customerType': 0,
'multipleConnectionMinute': -1, 'epicNo': 0, 'irrigation1Start': datetime.datetime(1999, 12, 31, 22, 0), 'irrigation1End':
datetime.datetime(1999, 12, 31, 22, 0), 'irrigation1Limit': -1.0, 'irrigation2Start': datetime.datetime(1999, 12, 31, 22, 0),
'irrigation2End': datetime.datetime(1999, 12, 31, 22, 0), 'irrigation2Limit': -1.0, 'creditLoadDateTime1': datetime.datetime(2019, 3, 4, 21, 0), 'creditLoadDateTime2': datetime.datetime(2019, 2, 26, 21, 0), 'creditLoadDateTime3': datetime.datetime(2013, 12, 31, 22, 0), 'creditLoadAmount1': -6633.96, 'creditLoadAmount2': 0.0, 'creditLoadAmount3': 0.0, 'flags': [{'flagType': 'INFO',
'name': 'ValveOpened', 'activationDate': datetime.datetime(2019, 4, 3, 5, 43, 6), 'attributes': []}], 'readDate':
datetime.datetime(2019, 4, 3, 5, 43, 6), 'meterId': 6403}
```

Şekil 3.13. Bir akıllı sayaç tarafından sunucu bilgisayara gönderilen verinin ham hali

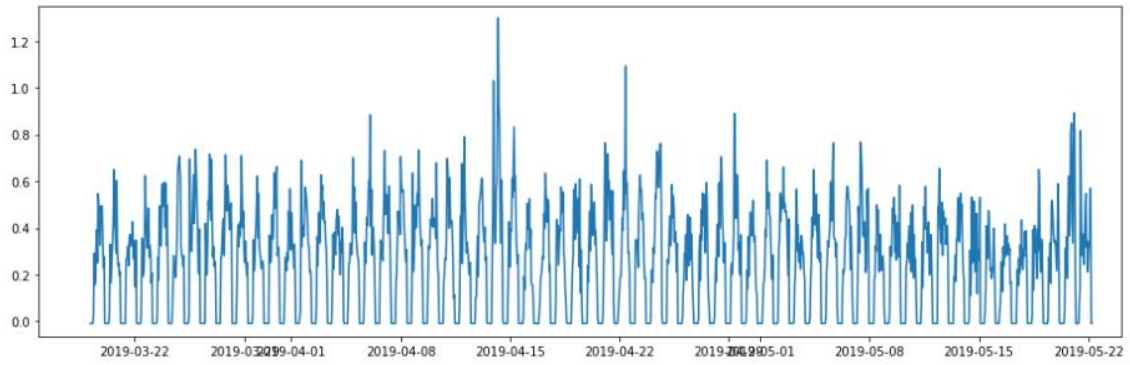
3.6.2. Zaman serisini oluşturma

Belirli kurallara göre birbiri ardınca yer alan veri bütününe seri, zaman vasfının şıklarını dikkate alarak düzenlenen serilere de zaman serisi denir. Zaman serisini oluşturmak için düzensiz olan verilerden boş alanlar silinmiş, tarih ve tüketim verileri yazılımda işlem yapılacak hale getirilmek üzere düzenlenmiştir. “Zaman” ve “Tüketim(m3)” alanları olan bir veri seti haline getirilmiştir. Ayrıca veri tabanından veriler alındığında tarih sıralamasında bir düzensizlik gözlemlenmiştir. Tarih verisinin sıralı olması için her tarihe denk düşen tüketim değerinin de tarihle beraber yerini değiştirmek üzere sıralama işlemi gerçekleştirilmiştir.(Şekil 3.14)

	Zaman	Tüketim(m3)
0	2019-03-19 01:00:00	0.000
1	2019-03-19 02:00:00	0.000
2	2019-03-19 03:00:00	0.000
3	2019-03-19 04:00:00	0.000
4	2019-03-19 05:00:00	0.000
5	2019-03-19 06:00:00	0.031
6	2019-03-19 07:00:00	0.295
7	2019-03-19 08:00:00	0.163
8	2019-03-19 09:00:00	0.183
9	2019-03-19 10:00:00	0.330
10	2019-03-19 11:00:00	0.394
11	2019-03-19 12:00:00	0.256
12	2019-03-19 13:00:00	0.546
13	2019-03-19 14:00:00	0.491
14	2019-03-19 15:00:00	0.403
15	2019-03-19 16:00:00	0.329
16	2019-03-19 17:00:00	0.494
17	2019-03-19 18:00:00	0.466
18	2019-03-19 19:00:00	0.493
19	2019-03-19 20:00:00	0.342
20	2019-03-19 21:00:00	0.234
21	2019-03-19 22:00:00	0.281
22	2019-03-19 23:00:00	0.000
23	2019-03-19 00:00:00	0.000

Şekil 3.14. Ön işlemesi yapılmış veri seti

Düzenlenmiş veri seti belirlenen periyotlarda bir dosyaya yedeklenmiştir. Bunun sebebi, denenen analiz yöntemlerinde kullanılan verilerde farklılık olmamasının istenmesidir. Böylece aynı veri ile farklı analiz sonuçları gözlemlenebilir. Fakat analiz yöntemi belirlendiğinde bu yedeklemeye ihtiyaç kalmaz ve veriler gerçek zamanlı olarak analiz edilir. Analiz edilecek verinin grafikteki gösterimi Şekil 3.15'deki gibidir.



Şekil 3.15. Saatlik tüketim verisi

3.7. Analiz Yöntemlerine Genel Bakış ve Bazı Terimlerin Tanımı

Bazı su idareleri otomatik okuma sayaçları ile uzaktan erişimle su sayaçlarını okumaktadır. Ancak bu sayaçlar, aşırı su tüketimlerinin ve arıza kaynaklı sızıntılarının otomatik tespiti ve yönetimine yönelik kapasiteye, yazılıma ve kullanıcıya özgü uyarı koşullarına ve müdahaleye sahip değildir. Mevcut teknolojilerden ve uygulamalardan farklı olarak, proje kapsamında geliştirilecek yazılım ile su tüketim grafiklerinin algoritmalar ve matematiksel yöntemler kullanılarak analizlerin yapılması sağlanacaktır. Analizin yüksek başarı ile yapılması için birçok algoritma, kütüphane ve analiz yöntemi denenmiştir. Bu analiz yöntemlerinin ortak özelliği, analiz sonucunda elde edilen simülasyon verileri ile gerçek verilerin karşılaştırılmasıdır. Böylece normal olarak tanımlanan veriler ile gerçek veriler arasındaki farklar analiz edilerek, anormal tüketim verileri de tespit edilmiş olur. Toplanan su tüketimi verilerinin analizi için tanımlanabilir anormallikler, eksik değer, sıfır değer, büyük değer vb. içerir. Bu anormalliklerden, sıfır değer ve eksik değer, doğrudan tespit edilip elenebilir. Bu eleme sonucunda örnek bir normal değer ortaya çıkar. Örnek olarak düzenlenen simülasyon verilerinin yaklaşık olarak normal dağılıma uyduğu varsayılarak, rastgele hatalar içeren veriler test edilip hata bölgesinin belirlenmesi gerekir. Böylece anormal değer miktarı belirlenebilir. Su tüketim verilerinde sıfır değeri ve eksik değer çıkarıldıktan sonra veri seti oluşur. Doğrudan tespit edilen anormal veri değeri elde edildikten sonra, normal aralığı elde etmek için verilerin kurtarılması gerekir. Ardından, daha önce yukarıdaki saptama yöntemi kullanılarak tanımlanmış olan sıfır değeri, eksik değer, anormal büyük değer ve anormal küçük değeri içeren aykırı değerleri kaldırmak için veri kümesinin ortalama değerini hesaplamak için ortalama doldurma yöntemi kullanılabilir. Doğrudan tespit edilen anormal değerlerden sonra bile, su tüketim verileri kaçınılmaz olarak hatalar ve gürültüler barındırır. Bu yüzden tüketim verisi detaylı bir şekilde ve farklı yöntemler ile analiz edilmelidir. Denenmiş analiz yöntemlerinde kullanılan yöntemler ve algoritmalar, Hareketli Ortalama (Moving Average), ARIMA, SARIMA, Prophet, Sezonsal Ayrıştırma (Seasonal Decomposition), LSTM'dir.

3.8. Tüketim Eğrisinin Düzeltilmesi

Sistemi eğitmek için, düzenli veri seti içerisinde bulunan aşırı tüketim ve tepe tüketim verileri, yumuşatma işlemi için Hareketli Ortalama (Moving Average) formülü kullanılmıştır. Ayrıca her 24 saatlik verinin en büyük değerine göre normalizasyon yapılmıştır. Normalizasyon işlemi gerçek veri grafiğinde tepe noktaları aynı düzeye

getirmek amacıyla yapılmaktadır. Böylece sanal bir normal veri oluşturulmuştur. Bu işlem gerçek anormal değerlerin tespiti için gereklidir.

3.8.1. Hareketli Ortalama (Moving Average)

Hareketli Ortalama, zaman serisi tahmininde etkili bir tekniktir. Veri hazırlama, özellik mühendisliği ve hatta doğrudan tahmin yapmak için kullanılabilir. Yumuşatma (Smoothing), zaman adımları arasındaki küçük varyasyonları ortadan kaldırmak için zaman serilerine uygulanan bir tekniktir. Düzleştirme umudu, gürültüyü ortadan kaldırmak ve gerekli olan sinyali daha iyi ortaya çıkarmaktır. Hareketli Ortalamalar, zaman serisi analizinde ve zaman serisi tahmininde kullanılan basit ve yaygın bir yumuşatma türüdür. Hareketli Ortalama'nın hesaplanması, değerlerin orijinal zaman serisindeki ham verilerin ortalamasından oluşan yeni bir serinin oluşturulması ile yapılır. Hareketli Ortalama için, pencere genişliği adı verilen bir boyutun belirlenmesini gerektirir. Bu, Hareketli Ortalama değerini hesaplamak için kullanılan ham verilerin sayısını tanımlar. Hareketli Ortalama ismindeki "Hareketli" kavramı, yeni serideki ortalama değerleri hesaplamak için pencere genişliği ile tanımlanan pencerenin zaman serisi boyunca kaydırılmasını ifade eder. Günlük tüketim eğrilerinin eğitimde kullanılmak üzere Hareketli Ortalama formülü kullanılmıştır. Veri hazırlamada etkili bir yöntem olduğu için bu formül seçilmiştir. Eğrinin daha yumuşak bir halde olması için bu formül kullanılmıştır.

3.8.2. Günlük maksimuma göre normalizasyon

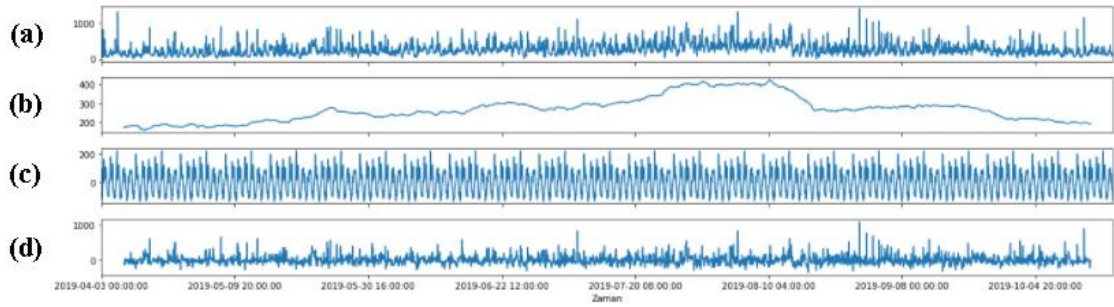
Günlük en büyük tüketim değerleri her gün için sırasıyla listelendiğinde bu değerler arasındaki varyasyonun yüksek olduğu gözlemlenebilir. Bu durumda günlük tepe noktaları aynı hizaya getirilerek eğitim için daha uygun bir veri elde edilir. Bu yöntem tepe noktalarının eşitlenmesi anlamına gelir. Tepe noktalarının ortalamaları alınır. Daha sonra bu oran her tepe noktası ile oranlanarak katsayılar belirlenir. Bu katsayı her bir gün için özeldir. Bu katsayılar gün içindeki tüm değerler ile çarpılır. Böylece tepe noktaları aynı değerde olan fakat karakteri değişmeyen bir tüketim serisi ortaya çıkar. Bu işlem ile veriler eğitim için daha etkin kullanılabilir.

3.9. Sezonsal Analiz

Zaman serisini birçok değişken etkileyebilir. Değişkenlerin etkilerine bağlı bir şekilde, zaman serisi grafiği farklı değerler alabilir. Su tüketiminin zaman serisindeki dalgalanmaların 3 bileşeni vardır. Bunlar; trend, mevsimsel dalgalanmalar, rastlantısal hareketlerdir.

3.9.1. Trend

Trend, zaman serisindeki değerlerin, geniş bir zaman dilimindeki yükselişi veya düşüşünü gösterir. Böylece su tüketimindeki trend verisi, aşırı tüketim ve sızıntı tespitinde işe yarar bir veridir. Şekil 3.16'da görüldüğü gibi trend ve sezonsallık grafiği su tüketimindeki aykırılıkların tespiti için kullanılmıştır.



Şekil 3.16. a) Zaman Serisi b) Trend c) Sezonsallık Grafiği d) Aykırılık Grafiği

3.9.2. Mevsimsel dalgalanmalar

Zaman serilerinde kısa dönem aralıklarındaki değer oynamalarına mevsimsel dalgalanmalar denir. Mevsimsel dalgalanmalar periyodiktir. Dalga uzunlukları aynıdır ve bir yıldır. Doğa koşulları, sosyolojik adetler ve kutlamalar gibi olayların bir yıl içindeki dağılımlarından meydana gelir. Sıcaklık derecesi, gibi meteorolojik nedenler bir yıl içerisinde her yıl yaklaşık aynı tarihlerde gerçekleşen dalgalanmalar mevsimsel dalgalanmalardır. Mevsimsel dalgalanmalar yıl içinde farklılık da gösterebilir. Mevsimsel dalgalanmaların belirlenmesinde ve ölçülmesinde yaygın olarak kullanılan yöntem hareketli ortalama yöntemidir.

3.9.3. Rastlantısal hareketler

Su tüketiminde etkili olan ve sistematik olmayan hareketlerden biri rastlantısal hareketlerdir. Arıza, deprem vb. olaylar buna örnek gösterilebilir. Bu etkiler sistematik olmadığı için ile önceden tahmin edilemezler. Rastlantısal hareketlerin etkileri doğal ve sosyolojik olaylardır ve belirlenemez.

3.10. LSTM

Yararlı önerilerde bulunmak için sistemin, simüle edilecek normal su verilerini yüksek doğrulukla tahmin edebilmesi gerekir. Bu çalışmada, LSTM kullanılarak normal su akışını tahmin etmek için önerilen yaklaşımı sunar. LSTM ağı, geri besleme kullanılarak eğitilmiş tekrarlayan bir sinir ağıdır. LSTM, optimum sonuçları elde etmek için makine öğrenimindeki zor dizi problemlerini çözmek için kullanılır. LSTM ağlarında nöronlar yerine bellek blok katmanları bulunur. Bir blok, son diziler için bir bellekle birleştirilmiş geleneksel bir nörondan daha iyi performans göstermesini sağlayan bileşenlerden oluşur. Blok, hem bloğun durumunu hem de çıktıyı yöneten kapılara sahiptir. Bir giriş dizisi üzerinde çalışır ve bir blok içindeki her kapı, tetikleme durumunu kontrol etmek için aktivasyon birimlerini kullanır. Blok yoluyla durum değişikliğini ve bilgi eklenmesini koşullu hale getirir. LSTM, bir LSTM katmanından yeterli öğrenme elde edebilir. LSTM derin öğrenme alanında kullanılan bir RNN (Tekrarlayan Sinir Ağı) mimarisidir. Geleneksel sinir ağlarından farklı olarak LSTM'nin (Uzun-kısa süreli bellek) geri beslemeleri vardır. İnsan düşünmeye başlarken, hafızasındaki bazı bilgileri kullanır. Yani sıfırdan düşünmeye başlamazlar.

LSTM'ler, Hochreiter ve Schmidhuber (1997) tarafından tanıtılan özel bir RNN türüdür. LSTM'ler uzun vadeli bağımlılıkları öğrenme yeteneğine sahip olduklarından ve çeşitli sorunları çözebildiklerinden, zaman serisi tahmini dahil olmak üzere farklı alanlarda yaygın olarak popüler hale geldiler. Bir LSTM modeli için ortak bir mimari, hücre olarak da adlandırılan bir bellek bileşeninden ve LSTM modeline bilgi akışını kontrol etmek için kapılar olarak bilinen üç "düzenleyiciden" oluşur.

Giriş geçidinin girdisi ilk olarak -1 ve 1 arasında sıkıştırılır ve

$$g = \tanh(b^g + x_t U^g + h_{t-1} V^g) \quad (3.1)$$

olarak ifade edilen bir tanh etkinleştirme işlevi burada giriş önyargısıdır ve sırasıyla giriş ve önceki hücrenin çıktısının ağırlıklarıdır. Daha sonra, sıkıştırılmış girdi, 0 ile 1 arasında bir değer olarak, gizli bir sigmoid-aktif düğüm katmanından oluşan girdi geçidinin çıktısı ile eleman bazında çarpılır. Eleman bazında çarpıldığında, kapı hangi girdilerin alınacağına karar verir veya atar. Kısacası, giriş geçidi, girişler için bir filtre görevi görür. Giriş kapısı

$$i = \sigma(b^i + x_t U^i + h_{t-1} V^i) \quad (3.2)$$

olarak formüle edilmiştir. LSTM hücresinin bir sonraki aşaması, hücreden hangi bilgilerin çıkarılması gerektiğini belirleyen unutma kapısıdır. Bu görev, aşağıdaki denklem kullanılarak "unutma katmanı" adı verilen bir sigmoid katmanla gerçekleştirilir:

$$f = \sigma(b^f + x_t U^f + h_{t-1} V^f) \quad (3.3)$$

Unutma geçidinden gelen çıktı, zamana göre ayrılan girdiler arasındaki ilişkiyi öğrenmek için dahili bir tekraralama döngüsü sağlamak için S_{t-1} ile eleman bazında çarpılır. Çarpmanın ürünü, önceki hangi adımların hatırlanması ve hangilerinin unutulması gerektiğini belirler.

Burada;

h_{t-1} : bir önceki gizli katman

V : bir önceki gizli katmanın ağırlığı

x_t : mevcut giriş vektörü

U : mevcut giriş katmanı ağırlığı

b : bias vektörü

LSTM'ler, bu üç formülü kullanarak, , bilgileri uzun süre zorluk çekmeden hatırlayacak şekilde tasarlanmıştır. Bu yeteneği nedeniyle, LSTM'ler, su tüketimi olaylarıyla ilgili tahmin problemlerini çözmek için iyi bir seçimdir. LSTM'lerin önemli

su tüketimi olaylarını tahmin etmek için kullanılıp kullanılamayacağını ve nasıl kullanılabileceğini araştırmak için bir ön çalışma olarak, 13 tane saatlik okuma periyodu, 16 tane 15 dakikalık okuma periyodu ile toplam 29 hanenin her biri için basit bir LSTM ağı kurulmuştur. LSTM'ler, TensorFlow kütüphanesinin (Abadi vd. 2016) üzerinde çalışan üst düzey bir sinir ağı API'si olan Keras (Chollet 2018) kullanılarak uygulanmıştır. Verinin LSTM'ye uyması için eğitim ve test seti gerekir. İyi uyum modeli durumunda, eğitim ve test kaybı azalır ve bir noktada neredeyse sabitlenir.

3.11. ARIMA

ARIMA, zaman serisi tahmini için popüler ve yaygın olarak kullanılan bir istatistiksel yöntem modelidir. ARIMA, Auto Regressive Integrated Moving Average anlamına gelen bir kısaltmadır. Zaman serisi verilerinde bir dizi farklı standart zamansal yapı yakalayan bir model sınıfıdır. Bir ARIMA modeli analizi ve tahmini zaman serisi verileri için bir istatistiksel model sınıfıdır. Zaman serisi verilerinde bir dizi standart yapıyı besler. Bu nedenle ustaca zaman serisi tahminleri yapmak için basit ama güçlü bir yöntem sağlar.

ARIMA kısaltması, modelin kendisinin temel özelliklerini kapsar. Bunlar:

- AR: Autoregression (Oto regresyon). Bir gözlem ile bazı gecikmeli gözlemler arasındaki bağımlı ilişkiyi kullanan bir model.
- I: Integrated (Fark). Zaman serisini durağan hale getirmek için ham gözlemlerin farkının kullanılması
- MA: Moving Average (Hareketli Ortalama). Gecikmeli gözlemlere uygulanan hareketli ortalama modelinden bir gözlem ile artık bir hata arasındaki bağımlılığı kullanan bir model.

Bu bileşenlerin her biri modelde bir parametre olarak açıkça belirtilir. Kullanılan belirli ARIMA modelini hızlı bir şekilde belirtmek için parametrelerin tamsayı değerleriyle değiştirildiği gösterilen

$$\text{ARIMA}(p, d, q) \quad (3.4)$$

biçiminde standart bir gösterim kullanılmaktadır ARIMA modelinin parametreleri aşağıdaki gibi tanımlanır:

- p: Modele dahil edilen gecikme gözlemlerinin sayısı, gecikme sırası olarak da adlandırılır.
- d: Ham gözlemlerin fark edilme sayısı, aynı zamanda fark derecesi olarak da adlandırılır.
- q: Hareketli ortalamasının sırası olarak da adlandırılan hareketli ortalama penceresinin boyutu.

Belirlenen terim sayısı ve tipini içeren lineer bir regresyon modeli oluşturulmakta ve veriler, regresyon modelini olumsuz etkileyen trend ve mevsimsel yapıları ortadan kaldırmak, durağan hale getirmek için bir dereceye kadar farklılaştırılarak hazırlanmaktadır.

Modelin o ögesinin kullanılmayacağını belirten bir parametre için 0 değeri kullanılabilir. Bu şekilde, ARIMA modeli bir ARMA modelinin ve hatta basit bir AR, I veya MA modelinin işlevini yerine getirecek şekilde yapılandırılabilir. Bir zaman serisi için bir ARIMA modelinin benimsenmesi, gözlemleri oluşturan temel sürecin bir ARIMA süreci olduğunu varsayar. Bu açıkça görünebilir, ancak ham gözlemlerde ve modelden kalan tahmin hatalarında modelin varsayımlarını doğrulama ihtiyacını motive etmeye yardımcı olur. ARIMA modeli, gelecekteki zaman adımlarını tahmin etmek için kullanılabilir. Eğitim veri setini eğitim ve test setlerine bölünüp, eğitim setini modele uyacak şekilde kullanabilir ve test setindeki her eleman için bir tahmin üretilebilir. Fark ve AR modeli için önceki zaman adımlarındaki gözlemlere bağımlılık göz önüne alındığında, yuvarlanan bir tahmin gereklidir. Bu yuvarlanan tahmini gerçekleştirmenin kaba bir yolu, her yeni gözlem alındıktan sonra ARIMA modelini yeniden oluşturmaktır. Bir ARIMA modeline uydurmak için klasik yaklaşım, Box-Jenkins Metodolojisini takip etmektir. Bu, ARIMA modeli için iyi parametreleri keşfetmek için zaman serisi analizini ve tanılamayı kullanan bir süreçtir. Özetle, bu sürecin aşamaları şu şekildedir:

- Model Tanımlama: Gerekli olacak farkın miktarı ve gecikmenin boyutu hakkında bir fikir edinmek için eğilimleri, mevsimselliği ve otoregresyon öğelerini belirlemek için grafikleri ve özet istatistikleri kullanılır.
- Parametre Tahmini: Regresyon modelinin katsayılarını bulmak için uygun bir prosedür kullanılır.
- Model Kontrolü: Model tarafından yakalanmayan zamansal yapının miktarını ve türünü belirlemek için artık hataların grafiklerini ve istatistiksel testlerini kullanılır.

İşlem, örnek içi veya örnek dışı gözlemlerde (örneğin, eğitim veya test veri kümeleri) istenen bir uyum düzeyine ulaşılan kadar tekrarlanır.

3.12. SARIMA

ARIMA, tek değişkenli zaman serisi veri tahmini için en yaygın kullanılan tahmin yöntemlerinden biridir. Yöntem bir trendle verileri işleyebilmesine rağmen, mevsimsel bileşenli zaman serilerini desteklemez. Serinin mevsimsel bileşeninin doğrudan modellenmesini destekleyen ARIMA'nın bir uzantısı SARIMA olarak adlandırılır. ARIMA'da, entegre eleman, yöntemin bir trend ile zaman serisi verilerini desteklemesine izin veren farklılığa atıfta bulunur. ARIMA mevsimsel olmayan veya mevsimsel bileşeni kaldırılmış, örneğin mevsimsel fark gibi yöntemlerle mevsimsellikten arındırılmış veriler beklemektedir. SARIMA serinin mevsimsel bileşeni için otoregresyon (AR), fark (I) ve hareketli ortalamayı (MA) belirtmek için üç yeni hiper parametrenin yanı sıra mevsimsellik dönemi için ek bir parametre ekler. Bir SARIMA'nın yapılandırılması, serinin hem trend hem de mevsimsel öğeleri için hiperparametrelerin seçilmesini gerektirir. Yapılandırma gerektiren üç trend ögesi vardır. ARIMA modeli ile ayırdırlar. Bunlar:

- p : Trend otomatik regresyon sırası.
- d : Trend farkı sırası.
- q : Trend hareketli ortalama sırası.

ARIMA'nın parçası olmayan ve yapılandırılması gereken dört mevsimsel öge vardır. Bunlar:

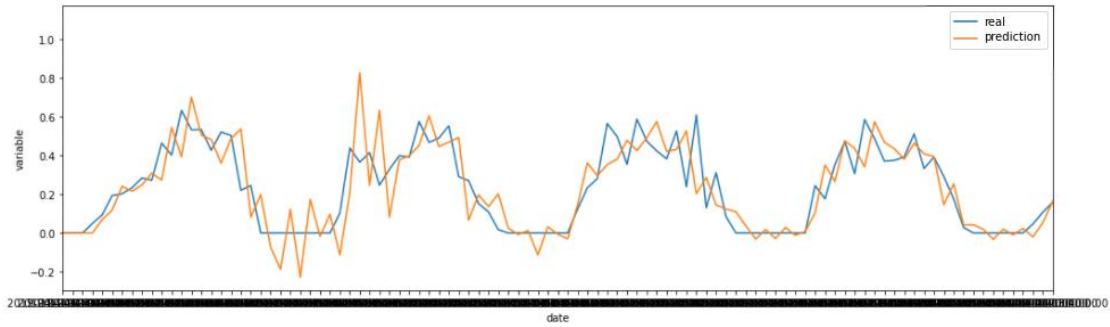
- P : Mevsimsel otoregresif düzen.
- D : Mevsimsel fark sayısı.
- S : Mevsimsel hareketli ortalama düzeni.
- m : Tek bir mevsimsel dönem için zaman adımı sayısı.

SARIMA modelinin gösterimi şu şekilde belirtilir:

$$\text{SARIMA}(p, d, q)(P, D, Q)m \quad (3.5)$$

Bu gösterimde, m parametresi P , D ve Q parametrelerini etkiler.

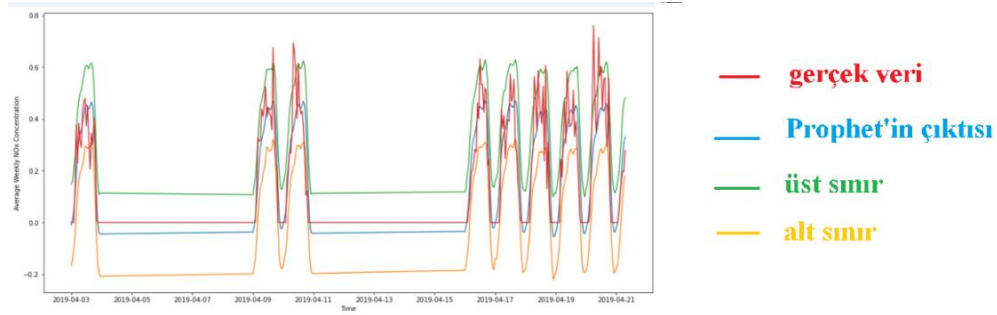
Şekil 3.17'de rastgele seçilmiş 4 günlük su verisini tahmin eden bir SARIMA tahmini gösterilmiştir.



Şekil 3.17. SARIMA tahmini

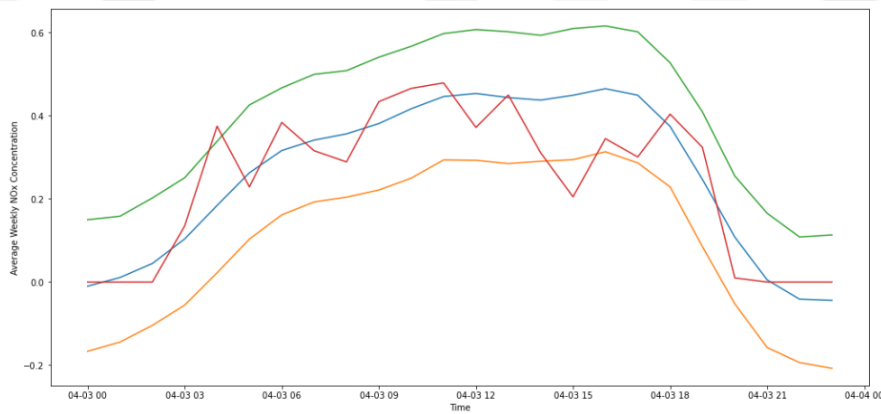
3.13. Prophet

Prophet kütüphanesi, tek değişkenli zaman serisi veri kümeleri için tahminler yapmak için tasarlanmış açık kaynaklı bir kütüphanedir. Kullanımı kolaydır ve varsayılan olarak trendler ve mevsimsel yapıya sahip veriler için ustaca tahminler yapmak amacıyla model için iyi bir hiperparametre kümesini otomatik olarak bulmak üzere tasarlanmıştır. Prophet, eklemeli zaman serisi tahmin modeli olarak adlandırdıkları şeyi uygular ve uygulama trendleri, mevsimselliği ve tatilleri destekler. Şekil 3.18'de görüldüğü gibi Prophet girdi verisini kullanarak 3 adet veri sunmuştur. Bunlar Prophet tahmini, üst sınır ve alt sınırdır.



Şekil 3.18. Prophet tahmini gösterimi

Şekil 3.19'da kırmızı ile gösterilen eğri; gerçek veri, mavi ile gösterilen eğri; tahmin verisi, yeşil ile gösterilen eğri; üst sınır, turuncu ile gösterilen eğri; alt sınırdır. Tahmin değeri alt ve üst sınırın dışında kalan noktalar anormal olarak nitelendirilebilir.



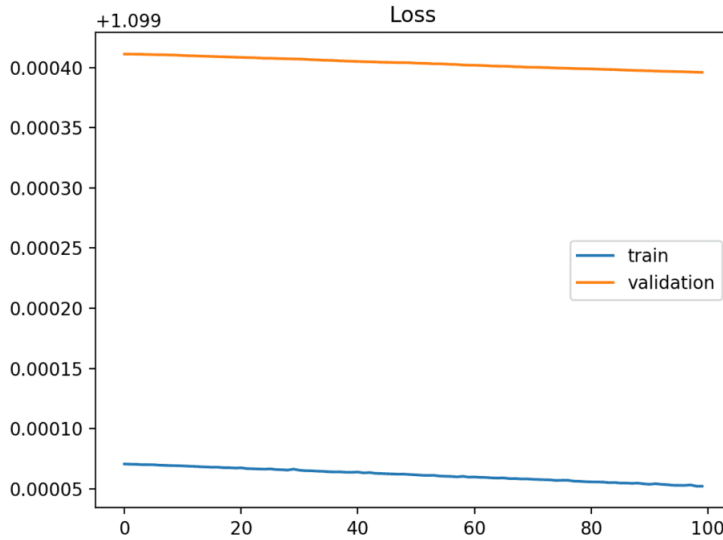
Şekil 3.19. Prophet tahmini günlük gösterim

3.14. Öğrenme Eğrileri

Öğrenme eğrileri, deneyim açısından zaman içinde öğrenme performansındaki değişiklikleri gösteren grafiklerdir. Eğitimdeki model performansının öğrenme eğrileri ve doğrulama veri kümeleri, eksik, fazla veya tam uyumlu bir modeli teşhis etmek için kullanılabilir. Model performansının öğrenme eğrileri, eğitim veya test veri kümelerinin problem alanını nispeten temsil edip etmediğini teşhis etmek için kullanılabilir. Öğrenme eğrileri, bir eğitim veri kümesinden aşamalı olarak öğrenen algoritmalar için makine öğreniminde yaygın olarak kullanılan bir tanı aracıdır. Model, eğitim sırasındaki her güncellemeden sonra eğitim veri kümesinde ve bir bekleme doğrulama veri kümesinde değerlendirilebilir ve öğrenme eğrilerini göstermek için ölçülen performansın grafikleri oluşturulabilir. Eğitim sırasında, modellerin öğrenme eğrilerinin gözden geçirilmesi, eksik veya fazla öğrenen model gibi öğrenme ile ilgili sorunları, eğitim ve doğrulama veri kümelerinin uygun şekilde temsil edilip edilmediğini teşhis etmek için kullanılabilir. Genel olarak, bir öğrenme eğrisi, x ekseninde zamanı ve y ekseninde öğrenmeyi veya gelişmeyi gösteren bir çizimdir (Anzanello ve Fogliatto 2011). Öğrenme eğrileri, derin öğrenme sinir ağları gibi zamanla aşamalı olarak öğrenen (iç parametrelerini optimize eden) algoritmalar için makine öğreniminde yaygın olarak kullanılır. Öğrenmeyi değerlendirmek için iyi puanlar (daha büyük

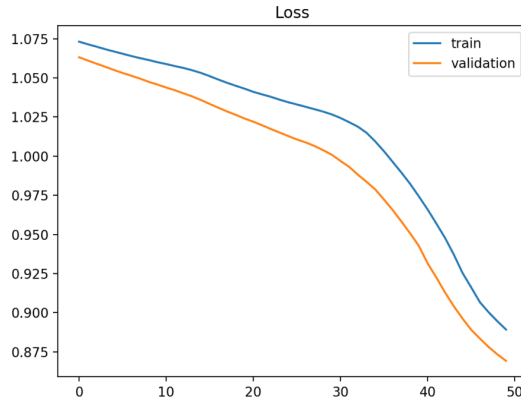
sayılar) daha fazla öğrenmeyi gösterir. Kayıp veya hatayı en aza indiren bir puanın kullanılması daha verimlidir. Bu sayede daha iyi puanlar (daha küçük sayılar) daha fazla öğrenmeyi gösterir ve 0,0 değeri eğitim veri kümesinin mükemmel şekilde öğrenildiğini ve hata yapılmadığını gösterir. Bir makine öğrenmesi modelinin eğitimi sırasında, eğitim algoritmasının her adımında modelin mevcut durumu değerlendirilebilir. Modelin öğrenme konusunda ne kadar iyi olduğuna dair bir fikir vermek için eğitim veri seti de değerlendirilebilir. Ayrıca, eğitim veri kümesinin parçası olmayan test veri kümesinde de değerlendirilebilir. Test veri kümesi üzerindeki değerlendirme, modelin ne kadar iyi genelleştirildiği hakkında bir fikir verir. Hem eğitim hem de test veri kümelerinde eğitim sırasında bir makine öğrenimi modeli için ikili öğrenme eğrileri oluşturmak yaygındır. Bir öğrenme eğrisinin şekli ve dinamikleri, bir makine öğrenimi modelinin davranışını teşhis etmek için kullanılabilir ve belki de, öğrenmeyi ve/veya performansı iyileştirmek için yapılabilecek yapılandırma değişikliklerinin türüne ilişkin önerilerde bulunabilir. Öğrenme eğrilerinde muhtemelen gözlemleyeceğimiz üç ortak dinamik vardır. Bunlar; aşırı öğrenme, eksik öğrenme ve iyi öğrenmedir.

Eksik öğrenme, eğitim veri kümesini öğrenemeyen bir modeli ifade eder. Eksik uyum modeli, yalnızca eğitim kaybının öğrenme eğrisinden tanımlanabilir. Modelin eğitim veri setini hiç öğrenemediğini gösteren düz bir çizgi veya nispeten yüksek kayıplı gürültülü değerler gösterebilir. Bunun bir örneği Şekil 3.20’de verilmiştir. Eksik öğrenme, veri kümesinin karmaşık olduğu durumlarda ve uygun bir yapıya sahip olmadığında gözlemlenir.



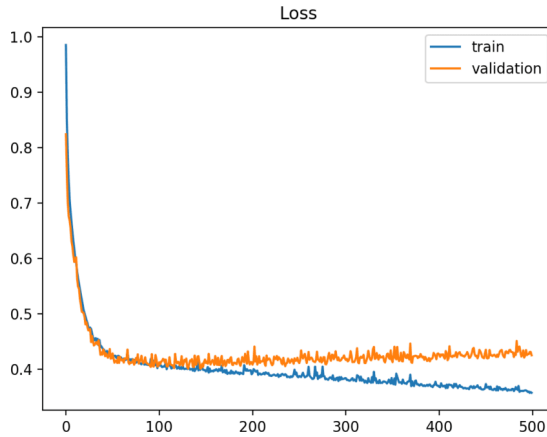
Şekil 3.20. Yeterli kapasiteye sahip olmayan bir modelin eksik öğrenme halindeki kayıp grafiği

Bir eksik öğrenme modeli, Şekil 3.21’de görüldüğü gibi sonlara doğru azalan ve azalmaya devam eden bir eğitim kaybı ile de tanımlanabilir. Bu, modelin daha fazla öğrenme ve olası daha fazla iyileştirme kapasitesine sahip olduğunu ve eğitim sürecinin zamanından önce durdurulduğunu gösterir.



Şekil 3.21. Daha fazla eğitim gerektiren bir eksik öğrenme modelini gösteren eğitim öğrenme eğrisi örneği

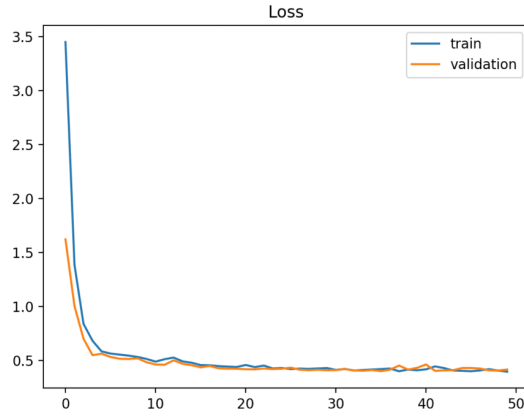
Aşırı öğrenme, eğitim veri kümesindeki istatistiksel gürültü veya rastgele dalgalanmalar dahil olmak üzere, eğitim veri kümesini çok iyi öğrenmiş bir modeli ifade eder. Aşırı öğrenme ile ilgili sorun, model eğitim verilerine ne kadar çok özelleşirse, yeni verilere o kadar az genelleme yapabilir ve bu da genelleme hatasında bir artışa neden olur. Bu istenmeyen bir durumdur. Genellikle, modelin sorun için gerekenden daha fazla kapasiteye ve dolayısıyla çok fazla esnekliğe sahip olması durumunda ortaya çıkar. Model çok uzun süre eğitilirse de oluşabilir. Genelleme hatasındaki bu artış, modelin test veri setindeki performansı ile ölçülebilir. Bir öğrenme eğrisinde Şekil 3.22'deki gibi test kaybı grafiği bir noktaya inip tekrar artarsa aşırı öğrenmeyi gösterir. Bu yüzden artmaya başlayan noktada eğitim durdurulabilir.



Şekil 3.22. Aşırı öğrenme modelini gösteren eğitim ve test öğrenme eğrisi

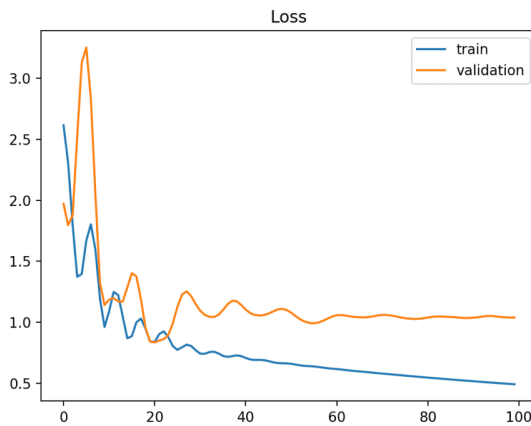
İyi öğrenme, algoritmanın hedefidir. Aşırı öğrenme ve eksik öğrenmenin arasında bulunur. İyi öğrenme, iki nihai kayıp değeri arasında minimum bir boşluk ile bir kararlılık noktasına düşen bir eğitim ve test kaybı ile tanımlanır. Eğitim kaybı grafiğinin bir denge noktasına düşüşü iyi öğrenme olduğunu gösterir (Şekil 3.23). Modelin kaybı, eğitim veri setinde neredeyse her zaman test veri setinden daha düşük

olacaktır. Bu, eğitim ve test kaybı öğrenme eğrileri arasında bir miktar boşluk beklememiz gerektiği anlamına gelir. Bu boşluk, "genelleme boşluğu" olarak adlandırılır. İyi bir uyum için devam eden eğitim, muhtemelen aşırı uyuma neden olacaktır.



Şekil 3.23. İyi öğrenmeyi gösteren eğitim ve test öğrenme eğrileri örneği

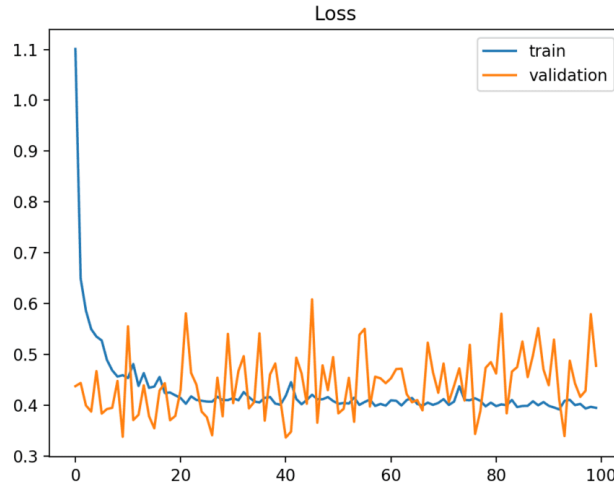
Yetersiz örneğe sahip bir eğitim veri kümesi, onu değerlendirmek için kullanılan test veri kümesine göre çok az örnek barındırıyorsa, eğitim veri kümesinin sorunu öğrenmek için yeterli bilgi sağlamadığı anlamına gelir. Bu durum, iyileşme gösteren eğitim kaybı için bir öğrenme eğrisi ve benzer şekilde iyileşme gösteren test kaybı için bir öğrenme eğrisi ile tanımlanabilir, ancak her iki eğri arasında Şekil 3.24'te görüldüğü gibi büyük boşluk kalır.



Şekil 3.24. Test veri kümesine göre çok küçük olabilecek bir eğitim veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği

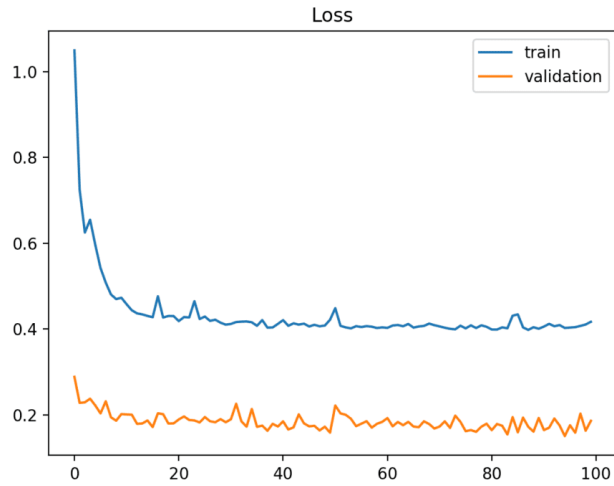
Yetersiz örneğe sahip bir test veri kümesi, doğrulama veri kümesinin, modelin genelleme yeteneğini değerlendirmek için yeterli bilgi sağlamadığı anlamına gelir. Bu, doğrulama veri kümesinde eğitim veri kümesine kıyasla çok az örnek varsa ortaya çıkabilir. Bu durum, Şekil 3.25'te görüldüğü gibi iyi bir uyuma sahip gibi görünen

eğitim kaybı ve eğitim kaybı etrafında gürültülü hareketler gösteren test kaybı gözlemlenen bir öğrenme eğrisi ile tanımlanabilir.



Şekil 3.25. Eğitim veri kümesine göre az sayıda test veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği

Şekil 3.26’da görüldüğü gibi eğitim kaybindan daha düşük bir test kaybı da olabilir. Bu durum, model için test veri setinin eğitim veri setinden daha kolay tahmin edilebileceğini gösterir.



Şekil 3.26. Eğitim veri kümesine göre çok küçük olabilecek test veri kümesini gösteren eğitim ve test öğrenme eğrileri örneği

3.15. Aykırılık Ölçümü ve Tespiti

Gerçek veri ile analiz verisi arasındaki farklılıkların analiz edilmesi için aykırılık grafiği kullanılmıştır. Aykırılık grafiği hataların zaman serisinde grafiğe dökülmesi ile oluşturulmuştur. Şekil 3.27'deki aykırılık grafiği,

$$e[n] = r[n] - a[n] \quad (3.6)$$

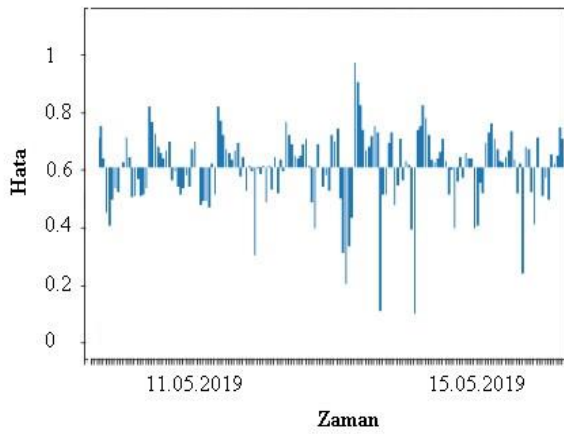
formülü ile oluşturulmuştur.

n = belirlenen tarih verisi

$r[n]$ = n tarihindeki gerçek veri

$a[n]$ = n tarihindeki hareketli ortalama

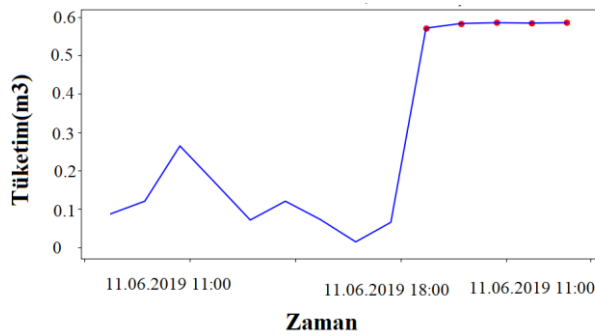
$e[n]$ = n tarihindeki hata



Şekil 3.27. Aykırılık grafiği

3.16. Aşırı Tüketim Tespiti

Analiz sonucunda bulunan tepe tüketimlerin tekrar sayısına bakılarak bulunan aşırı tüketimler Şekil 3.28'de görüldüğü gibi gerçek veri ile birlikte grafikte gösterilmektedir.



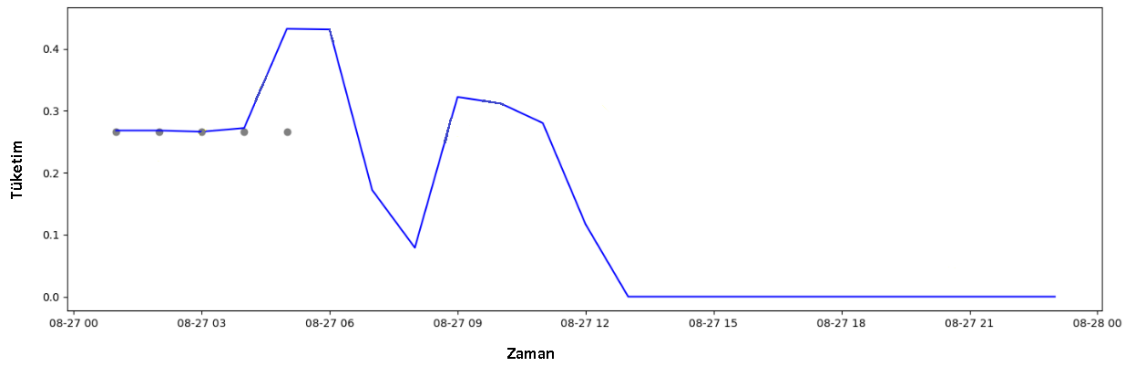
Şekil 3.28. Aşırı tüketim tespiti

3.17. Sızıntı Tespiti

Gece saatlerindeki tüketimin ortalamasına ve bir gün içindeki aktif kullanımdaki birbirine yakın değerlerin sayısına bakılarak sızıntı bulunmaktadır. Gece saatlerindeki değerlerin ortalaması eşik değeri aşıyorsa o geceye sızıntı tespiti konulmaktadır. Aktif kullanım saatlerinde birbirine yakın değerlerin sayısı eşik değeri geçiyorsa birbirine yakın değerler sızıntı olarak değerlendirilmektedir. Verilerin birbirine yakınlık hassasiyeti de parametreler ile ayarlanabilmektedir.

Örnek 3.1. Günlük verinin belirtilen saatler arasında (Örneğin; 01:00:00'den 05:00:00'e kadar) tüketim değeri ortalaması, kullanıcı tarafından belirlenen eşik değerden (Örneğin; 0.1 m³) büyükse sızıntı tespiti konulmuştur. Eşik değerler abonelerin özelliklerine göre kullanıcı tarafından değiştirilebilmektedir. Belirtilen zaman değerleri arasında ortalama eşik değerin üstünde ise o zaman değerleri arasındaki minimum değer o aralık boyunca sızıntı olarak belirlenmiştir. Örneğin; 01:00:00 ile 05:00:00 saatleri arasında ortalamanın 0.1 m³ ü aştığı gözlemlendikten sonra bu saatler arasındaki minimum değer 0.26 m³ ise sızıntı bu 5 saat boyunca saatte 0.26 m³ olarak belirlenmiştir.

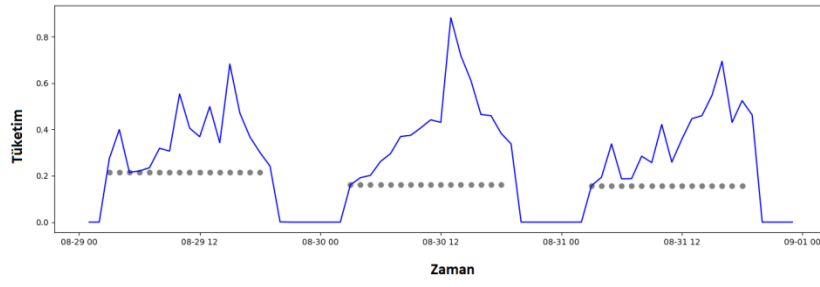
Şekil 3.29'da görüldüğü gibi aktif tüketimdeki sızıntı gün içinde birbirine yakın art arda gelen tüketim değerlerinin sayısı hesaba katılarak tespit edilmektedir. Art arda gelen birbirine yakın tüketim verilerinin sayısı eşik değerin üzerinde ise bu değerlere sızıntı tespiti konulmaktadır.



Şekil 3.29. Basit sızıntı tespiti

Günlük aktif tüketim verisinin en küçüğünün tekrar sayısına bakılmaktadır. Yani günlük en küçük tüketim verisi belirli dalgalanma aralıklarındaki tekrar sayısı kullanıcı tarafından belirlenmiş eşik değeri aşıyorsa sızıntı tespiti konulmuştur. Bu sızıntı Şekil 3.30'da görüldüğü gibi günlük aktif kullanım süresi kadar devam etmektedir.

Örnek 3.2. Saat 20:00 ile 03:00 saatleri arasında ana vananın kapatıldığı bir abonede aktif vakitlerde minimum saatlik tüketim 0.2 m³ ise ve gün içinde 0.01 m³ dalgalanma farkı ile tekrar sayısı 2'yi aşıyorsa sızıntı tespiti konulmuştur.



Şekil 3.30. Aktif tüketim sırasındaki sızıntı tespiti

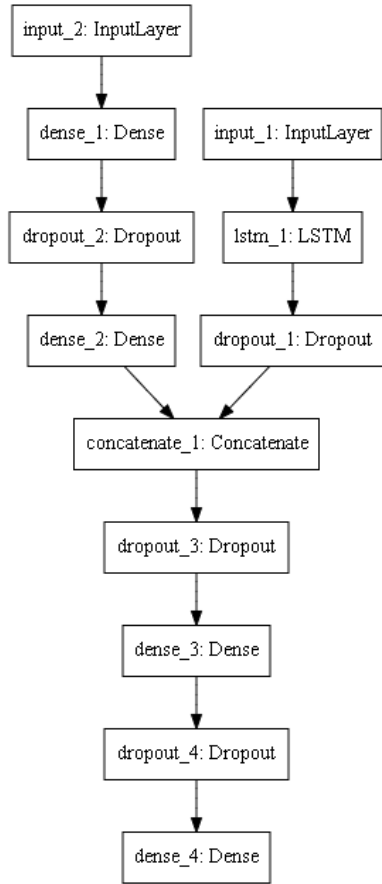
4. BULGULAR

Tahmin modellerinin uygunluğunu doğrulamak için ölçüt değerleri Çizelge 4.1’de gösterilmiştir. Bu tabloda gösterilen değerler her model için en başarılı değerlerdir.

Çizelge 4.1. Denenmiş tahmin yöntemleri ve sonuçlar

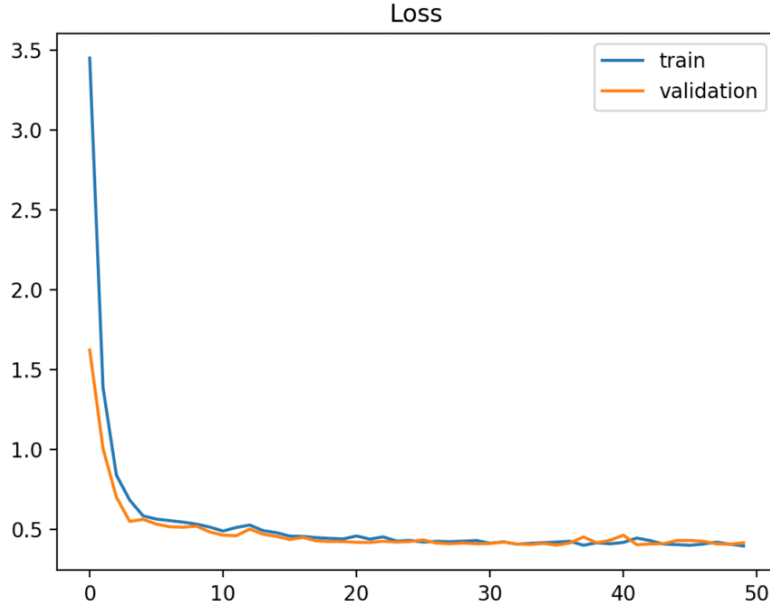
Yöntem	LSTM	SARIMA	Prophet
R² Değeri	0.82	0.74	0.72
Eğitim Kayıp Eğrisi	İyi	İyi	İyi
Test Kayıp Eğrisi	İyi	İyi	Kötü

Analizde kullanılan yöntemlerin başarı oranları göz önünde bulundurularak en verimli tercih edilmiştir. Seçilen model en verimli model olan, LSTM modelidir ve Şekil 4.1’deki gibi oluşturulmuştur.



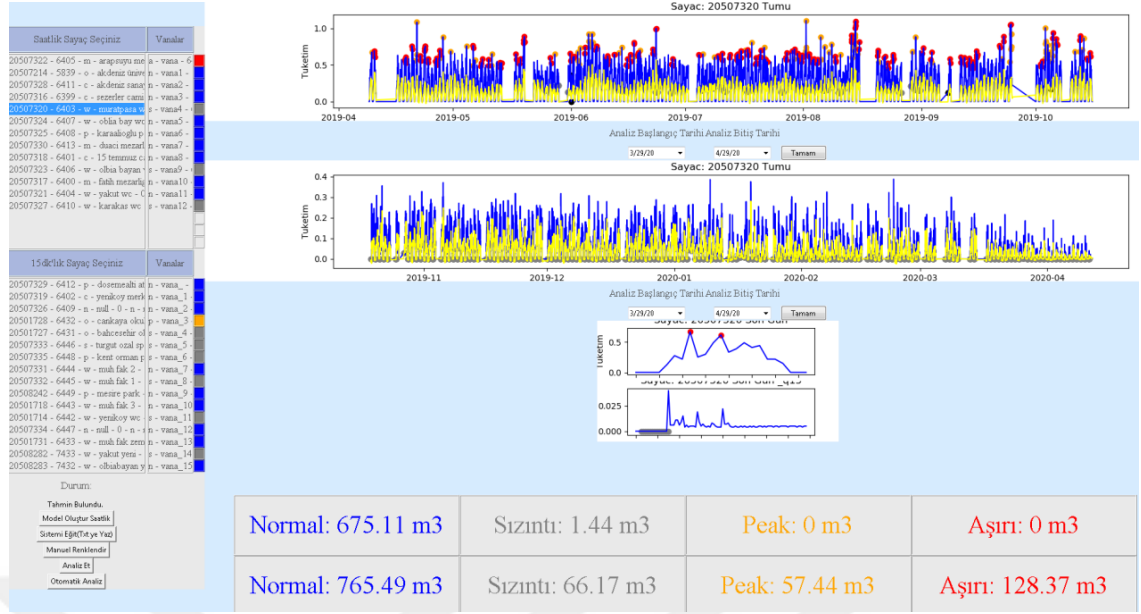
Şekil 4.31. LSTM ağ modeli

Su tüketim tahmininde LSTM modeli kullanılmıştır. Eğitim ve test kayıp verileri görselleştirilmiş olup Şekil 4.2'deki grafik elde edilmiştir. Bu grafik eğitim ve testin başarılı olduğunu gösterir.



Şekil 4.32. Eğitim ve test kayıp grafiği

Otomasyon arayüz tasarımının son hali Şekil 4.3'deki gibidir. Arayüzün sol tarafında abone bilgileri, vana numaraları ve abonelerin tüketim durumları yer almaktadır. Tüketim durumları 4 farklı renk ile gösterilmektedir. Mavi renk son günün normal olduğunu, gri renk son günde sızıntı bulunduğunu, turuncu renk son günde tepe tüketim olduğunu, kırmızı renk de son günde aşırı tüketim bulunduğunu belirtmektedir. Bu bölüm arayüz kullanıcılarına, abonelere genel bakış olanağı sağlamaktadır. Abone bilgilerinin ve durumlarının gözlemlendiği bölgenin alt tarafında model eğitimi ve manuel analiz ile ilgili butonlar bulunmaktadır. Model eğitim butonlarına basılarak aynı türdeki diğer abone verileri ile model eğitilir. Manuel analiz butonuna basılarak yazılımın çıkardığı analiz sonuçları grafiğe aktarılır. Arayüzün üst ve orta kısmında seçilen abonenin tüketim grafikleri ve analiz sonuçları bulunmaktadır. Bu grafiklerde saatlik ve 15 dakikalık veriler ayrı ayrı gösterilmektedir. Grafiklerde, gerçek tüketim verileri mavi, tahmin verisi sarı, sızıntı değerleri gri, tepe tüketim değerleri turuncu, aşırı tüketim değerleri kırmızı ile gösterilmektedir. Sol bölümde bulunan bir aboneye tıklandığında ilk olarak abonenin saatlik tüm verileri grafiğe aktarılmaktadır. Kullanıcı verileri iki tarih arasında filtreleyebilmektedir. İkinci grafikte 15 dakikalık veriler grafikte gösterilmektedir. Üçüncü grafikte saatlik son gün grafiği, dördüncü grafikte de 15 dakikalık son gün grafiği bulunmaktadır. Arayüzün alt kısmında seçilen tarih aralığındaki normal, sızıntı, tepe ve aşırı tüketim miktarları bulunmaktadır. Bu veriler saatlik ve 15 dakikalık veriler için gösterilmiştir.



Şekil 4.33. Otomasyon arayüzü

Uyarıların %70'i, gerçek patlama ve arızalar ile ilişkilendirilmiştir. Diğer uyarıların çoğunluğu insan kaynaklı aşırı tüketimdir. Yapılan çalışmadan elde edilen sonuçlar, otomasyon sisteminin küçük ila orta büyüklükteki patlamaların ve sızıntılarının gerçek zamanlı tanımlanması için yararlı bir araç olarak kullanıma potansiyeline sahip olduğunu göstermiştir.

5. TARTIŞMA

Bir su abonesinin tüketim grafiğinin gerçek zamanlı izlenmesi ve analiz edilmesi ile aşırı tüketim, tepe tüketim ve sızıntılar tespit edilebilmektedir. Bu yöntemlerin en uygununu seçebilmek için ölçüm parametreleri mevcuttur. Bu çalışma, sürdürülebilir kentsel su yönetimi amacıyla bir otomasyon sistemi için yeni bir konsept sunmuştur. Bu sistemin hem tüketiciler hem de su işletmeleri için potansiyel faydaları büyüktür. Bu çalışmada, zaman serisinin analiz edilmesi için birçok yöntem denenmiştir.

Uygulanan IoT sistemi aracılığıyla veri toplamanın temel amacı, daha sonraki “makine öğrenimi” işleme aşamaları için yeterli veriyi elde etmektir. Ölçülen veriler bir büyük veri kontrolcü sistem tarafından değerlendirilmektedir. Bu araştırmadaki veriler, tüm abonelerden izin alınarak toplanmıştır. Antalya’da bulunan belirli mahallelere akıllı su sayaçları kurulmuştur. Elde edilen veriler, 29 aboneden toplanan içme suyu tüketim ölçümlerini içermektedir. Ölçülen veriler, akıllı su sayaçlarının kurulduğu dağıtılmış konumlardan su tüketimi bilgilerini içermektedir. Veri setlerinden hatalı ölçümlerin çıkarılması için denetimin ilk aşamasında kullanılmak üzere 3 aylık veri toplanmıştır. Sonuç olarak, ölçümlerde herhangi bir tutarsızlıktan kaçınacak şekilde yaklaşık 1 yıl boyunca alınan 29 veri seti vardı. Verilerin örnekleme periyodu önce saatlik, sonra 15 dakikalık olarak tutulmuştur. Her iki örnekleme periyodu için ayrı analizler yapılmıştır. Bu nedenle, elde edilen veriler, katılımcı hanelerde hacimsel su tüketimini ortaya çıkaran 1 yıllık bir süreyi kapsamaktadır. Ayrıca, toplu veri seti, su talebi yönetimi için gerekli olan diğer özellikleri ortaya çıkarmak için kullanılabilir. Veri toplama aşamasına ne kadar çok hane katılırsa, modelin başarısı o kadar artar. Elde edilen veri kümelerinde bazı beklenmedik korelasyonları bulmak için başarılı olan ve makine öğrenimi ve toplu işleme için bunları eşzamanlı olarak aktarabilen akıllı sayaçlar kullanılmıştır.

Akış verileriyle bir vaka çalışması için Antalya’da tam ölçekli bir su hattındaki aboneler akıllı sayaçlar ile izlendi. Yaklaşık 1 yıllık bir süre boyunca, 29 akıllı su sayacı için otomatik veri iletimi ve bu verileri sürekli olarak analiz edip zamanında uyarılar sağlayan çevrimiçi bir sistem çalışmaktadır. Sistem, akış olayının anormalliğinin yüzdesi ve sızıntı ölçeğini sunar. Bir yıllık dönem için sistem sonuçlarının raporu alınabilir. Görünür sızıntının ve aşırı tüketimlerin raporları gibi, sistemin birçok anormal olayı başarılı bir şekilde belirlediği görüldü. Bu çalışma ile, sızıntıların ve aşırı tüketimlerin kötü sonuçlar doğurmadan önce tespit edilebileceğini kanıtlamak mümkün oldu.

Yapılan uygulamada, su tüketimindeki anormallikler, “tepe tüketim”, “aşırı tüketim” ve “sızıntı” olarak üç kategoriye ayrılmıştır. Analiz değerleri ile gerçek değerler arasındaki farklılıklar analiz edilerek anormal tüketimler tespit edilmiştir. Anormal durum analizi yapılan mevcut uygulamalarda verimli olduğu gözlemlenen analiz yöntemleri kullanılmıştır. Su tüketim anormalliklerinin tespiti için yapılan analizlerin başarısının yüksek olabilmesi için veriler düzenlenerek işlenmiştir. Doğru sonuçlar elde edebilmek amacıyla analiz için düzensiz ve eksik verilerden arındırılmış veri topluluğu kullanılmıştır. Yazılımda her su abonesi için uyarı koşulları ve parametreler belirlenip, normal su tüketimi dışındaki aşırı tüketim ve sızıntılar algoritmalar ve matematiksel formüller yardımı ile tespit edilmektedir.

Tahmini yapılacak olan abonenin sınıfı eğitim için önemli bir bilgidir. Daha önce manuel olarak girilen abone sınıfı bilgisi eğitim verisi oluştururken kullanılmaktadır. Hangi abone için tahmin yapılırsa o abone türünün tüketim verilerinin tümü eğitimde kullanılmaktadır. Örneğin; bir mezarlık için tahmin yapmak gerekirse diğer tüm mezarlık verileri eğitimde kullanılır. Böylece daha önce manuel olarak girilen sınıf bilgisi eğitim sürecinde işe yarar. Eğitim ve test setleri girdi ve çıktı değişkenlerine bölünmüştür. Girdiler LSTM'lerin gerektirdiği üç boyutlu formatta yeniden şekillendirilmiştir. Tahmin modeli için abone sınıfındaki tüm tüketim verilerini kullanmak üzere ilk gizli katmanda 64 nöron ve çıkış katmanında bir nöron içeren basit bir LSTM ağı tanımlanmıştır. Gün içinde bazı zaman aralıklarında su tüketimi artarken bazı zaman aralıklarında tüketim olmayabilir. Tatil günleri ve çalışma günleri arasındaki tüketim farklılık göstermektedir. Haftalık veriye bakıldığında, hafta içi ve hafta sonu tüketim grafiği farklı olabilmektedir. Bu yüzden zaman kavramı tahmini etkileyen önemli bir faktördür. Saatlik, günlük, haftalık, aylık olarak oluşturulan zaman vektörü modele ayrı bir girdi olarak alınmıştır. Zaman vektörünün eğitimde kullanılmasının sebebi, yapay sinir ağının “tatil günleri”, “aktif tüketim saatleri” gibi kavramları da göz önünde bulundurarak tahmin yapmasını sağlamaktır. LSTM'ler, özellikle sigmoid veya tanh aktivasyon fonksiyonları kullanıldığında, giriş verilerinin ölçeğine duyarlı olduğu için veriler eğitilmeden ve test edilmeden önce 0'dan 1'e aralığına yeniden ölçeklendirilmiştir. Zaman serisi verilerinde, değerlerin sırası önemlidir. Bu nedenle, veri kümesi, modeli eğitmek için kullanılan gözlemlerin %70'i ile eğitim veri kümelerine bölünmüş ve kalan %30'u model testi için bırakılmıştır. Modellerin uygunluğunu doğrulamak için her eğitim ve doğrulama dönemi için kayıp puanları hesaplanıp kaydedilmiştir. Bu değerlerin çizilmesi, modelin uygunluğunun anlaşılmasını sağlamaktadır. Bir model eğitim veri setinde daha iyi performans gösteriyorsa ve test veri setinde zayıfsa, o zaman uygun olmayan model olarak kabul edilir. Bu, eğitim kaybının test kaybından daha düşük olduğu anlamına gelir. Aşırı öğrenme gösteren bir model söz konusu olduğunda, model eğitim seti için iyidir ve gelişmeye devam eder. Ancak test seti için bir noktadan sonra performans düşmeye başlar. İyi uyum modeli durumunda, eğitim ve test kaybı azalır ve bir noktada neredeyse sabitlenir. Sistemin başarısı büyük ölçüde sistemin tüketiciler tarafından su tüketimi davranışını tahmin etme yeteneğine bağlıdır. Özellikle küvet, duş, sulama gibi önemli su tüketimi olayları tahmini yüksek oranda etkilemektedir. Bu tüketim olayları her abone sınıfında farklılık gösterebilir. Ayrıca aynı sınıftaki aboneler arası tüketim eğrilerinde benzerlik gözlemlenebilir. Kişiselleştirilmiş bir öneri sistemi gerçekleştirmeye yönelik bir adım olarak, 29 abone için yüksek su tüketimi olaylarını tahmin etmek amacıyla basit bir LSTM ağı kurulmuştur. Ön deneyimler, LSTM'nin bu tür bir sorunu çözmek için iyi bir aday olduğunu göstermektedir. Ağırlık güncellenmeden önce ağı gösterilecek numune sayısını sınırladığı için optimal parti boyutu göreve bağlıdır. Bu soruna bir çözüm, çevrimiçi öğrenmeyi kullanan modeldir. Bu, her eğitim örneğinden sonra ağı ağırlıklarını güncellerken parti boyutunu küçük tutularak ayarlayarak başarılabilir. Bu, daha hızlı öğrenme etkisine sahip olabilir, ancak ağırlıklar her partide büyük ölçüde değiştiğinden öğrenme sürecine istikrarsızlık da ekleyebilir. Bu nedenle, hem gizli katmandaki nöron sayısını hem de ilk eğitim aşamasında toplu iş boyutu optimize edilmiştir. Model, eğitim veri kümesi kullanılarak eğitildikten sonra, modelin performansı, karşılaştırmaya uygun metrikler verecek şekilde tahmin edilebilir. Birleştirilmiş hanelerin veri kümelerinin sayısı ne kadar fazla olursa, LSTM modeli tarafından yakalanan periyodikliğin o kadar iyi olduğu fark

edilmiştir. Model, su tüketimi tahmini ve saat cinsinden zaman ilişkisini yakalayarak daha iyi bir iş çıkarmıştır. Uygun doğruluk ölçütleri kullanılarak ayrıca değerlendirilmiştir. LSTM her zaman uzun vadeli bağımlılıkları dikkate aldığı ve tüm seri modelini anladıktan sonra yeni değeri değerlendirdiği için, büyük boyutlu veri kümelerini kullanırken derin öğrenme yeteneklerinde üstündür. Tahmin sonuçlarının başarısının ölçülmesi için R^2 ölçütü kullanılmıştır. LSTM'ler stokastik olduğu için, yani değişken sonuçlar verdiği için her çalıştırma farklı bir R^2 değeri ile sonuçlanabilir. Bu nedenle, modeller her abone için beş kez eğitilmiştir ve değerlendirilmiştir. Her değerlendirme için R^2 değeri hesaplanmıştır. R^2 metriği, bir dizi tahminin gerçek değerlere uygunluğunun bir göstergesidir. İstatistik literatüründe bu ölçü, belirleme katsayısı olarak adlandırılır. Bu ölçü, uyumsuzluk ve mükemmel uyum için sırasıyla 0 ile 1 arasında bir değerdir. Su tüketim tahmini sayesinde normal değerler tahmin edilip gerçek değerler ile karşılaştırılmıştır. Gerçek değerlerin içinde anormal değerlerde olduğu için eğitimde bu değerlerin arındırılmış hali kullanılmıştır. Yani anormal değerler manuel olarak silinmiştir. Eğitilen yapay sinir ağı sayesinde, son zamanlardaki su tüketim verileri, zamana bağlı olarak gerçeğe yaklaşık %82 benzerlik oranı ile tahmin edilmiştir. Daha düzgün bir tüketim eğrisi elde etmek için günlük maksimum değere göre normalleştirme yapılarak, hareketli ortalama formülü uygulanmıştır. Bu ön çalışmadan, LSTM'nin aşırı tüketim olaylarını tahmin etmek için iyi bir aday olduğu sonucuna varıyoruz.

Sızıntı tespiti, aktif kullanım saatleri dışında kalan zamanlardaki (genellikle gece saatlerindeki) tüketim ortalaması esas alınarak yapılmaktadır. Fakat, “kullanıcının gece saatlerinde ana vanayı kapatması” gibi özel durumlar da mevcuttur. Bu özel durumda gece saatlerinde sızıntı yokmuş gibi görünse de gerçekte olabilir. Bunun için gün içindeki aktif tüketim zaman aralıkları analiz edilmektedir.

Elektronik vana oranının ayarlanması için IoT (Nesnelerin İnterneti) tabanlı bir gömülü yazılım uygulaması yapılmıştır. Donanım olarak da, mevcut Iot uygulamalarında yaygın olarak kullanılan ESP8266 NodeMCU geliştirme kartının kullanılması uygun görülmüştür. Bu kart bir IoT modülüdür. Iot modülünün internet veri tabanından aldığı veriler ile oransal vananın kontrolünü sağlaması için gerekli olan elektronik kart tasarlanmıştır. Bu kart sinyal voltaj uyumu için tasarlanmıştır.

Arayüz yazılımı tarafından analizler yapıldığında Firebase veritabanına aktarılan vana oranları ve kullanıcı onay durumu tasarlanan elektronik karttaki gömülü yazılım sayesinde gerçek zamanlı kontrol edilmektedir. Kullanıcı onayı Android uygulama tarafından alındıktan sonra, nesnelerin interneti özelliği olan elektronik tasarım sayesinde oransal vananın da oranı değişmektedir. Böylece su hattından geçen su sınırlandırılmış olur. Analiz kısmında akıllı su sayacının kendi arayüzündeki veriler ile analiz sonuçlarının karşılaştırılması yapılmıştır. Yapılan çalışma sonucunda ortaya çıkan otomasyon sistemi kullanıcıya daha detaylı bilgiler vererek kullanıcı gözlemini zenginleştirmiştir. Kullanımı, sızıntı olaylarının meydana geldikten hemen sonra ve müşteri ciddi şekilde etkilenmeden önce düzeltilmesi olasılığı ile sızıntı yönetimine daha hızlı çözümlü bir yaklaşımı teşvik eder. Sistem, abonelerin anormal tüketimlerini tespit etmeyi ve düzeltmeyi mümkün kılar. Bu sayede sistem 'sızıntı ve aşırı tüketim seviyesini' azaltmayı sağlar.

6. SONUÇLAR

Hükümetler, Yapay Zeka ve Makine Öğrenimi'nden yararlanarak müşterilerin ihtiyaçlarını tahmin edebilir, önleyici eylemleri otomatikleştirebilir ve hizmetlerini ve ürünlerini nicel ve nitel ölçülere göre uyarlayabilir. Ayrıca, şebekeler ve akıllı sayaçlar kaynakları optimize ettiğinden ve uzaktan izleme çözümleri su şebekesinin verimliliğini artırdığından IoT'nin fırsatları sınırsızdır. Ayrıca analitik, her başarılı IoT uygulamasının temel bir bileşeni olarak kabul edilir. Bu nedenle, IoT teknolojisi gerçek zamanlı analizler sağlayabilir ve ulusal refahı iyileştiren veriye dayalı kararları güçlendirebilir.

Yapılan çalışmada su tüketimini temsil edebilecek bir haritayı kapsayacak şekilde akıllı su sayaçları kuruldu. Gerçek zamanlı izleme ve kritik yönetim durumları için gerekli işlemler kritik öneme sahiptir. Bu çalışmada kullanılan yöntem olarak, su tüketimi verilerini toplamak için kullanılan ulusal ölçekteki IoT sisteminde performansı artırmak ve maliyeti azaltmak için akıllı su sayaçlarından yararlanıldı. Önerilen sistem, farklı bölgeler için alt modellerle tüm ülkeyi kapsayacak şekilde genişletilebilir.

Yapılan çalışma, güvenlik, maliyet ve ölçeklenebilirlik dikkate alınarak entegre hizmetlerle ulusal ölçekte ölçeklendirilebilen bir su talebi yönetim sistemine dahil edilecek su tüketimi için bir test ortamı sunmaktadır.

Ortaya çıkan sistemin ana avantajı, tüketim sınıfına dayalı bir tüketim modeli elde etmek için kullanılmak üzere farklı abonelerin su tüketimini analiz etmektir. Bu sayede, uyarlanabilir bir makine öğrenimi modeli ile su talebi için periyodik analiz mümkün oldu. Su talebi tahminine dayanarak, nihai su tüketiminin azaltılması hedefi için hem su hizmetleri yönetimi hem de tüketici davranış yönetimi için bir dizi senaryo birleştirilebilir. Ayrıca bu araştırma, yoğun zamanlarda su talebinin azaltılması, daha iyi su talebi dağılımı ve daha düşük su tüketimi için nicel önlemler sunması gereken bir yönetim sistemi sunmaktadır. Ayrıca, planlı şehir gelişimi ve su şebekesi altyapısına ve performansına dayatılan genişlemenin etkisini ölçmek mümkündür.

Şu anda sistem, her hane için eğitilecek yeni bir LSTM ağı gerektirmektedir. Fakat bu yaklaşım gerçek dünyada teknik kapasiteden dolayı mümkün değildir. Bu durum ayrıca ele alınmalıdır. Bir sonraki adım, bu sınırlamanın üstesinden gelmek için haneleri su tüketim modellerine göre gruplandırmak için kümeleme tekniklerini keşfetmektir. Ek olarak, daha fazla iyileştirme için ve her bir anormallik türünü tahmin etmek için ayrı modeller ve çok etiketli sınıflandırma eğitimi yapılması planlanmaktadır. Böylece abone sınıfının manuel olarak girilmesine gerek kalmayıp daha başarılı sonuçlar elde edilebilir. Model çıktıları daha güvenli hale geldiğinde, tavsiye sisteminin kapsamlı bir şekilde geliştirilmesi sağlanacaktır.

İkinci aşama olarak, daha geniş bir bölgede 300 den fazla akıllı sayaçtan veri toplanarak analiz yapılabilecek sistem planlanmaktadır. Bu yüksek veri, ağ etkinliğinin her zamankinden daha ayrıntılı bir şekilde yorumlanmasına ve müşterilere daha uygun yanıt verilmesine olanak tanıyacaktır. Bir sonraki hedef, gerekli altyapıların sağlanarak su hizmetleri şirketinin makine öğrenmesini de kapsayan bir analiz sistemini kullanmasını teşvik etmektir.

7. KAYNAKLAR

- Abadi, M., et al. «TensorFlow: a system for large-scale machine learning.» *OSDI'16: Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation*. Savannah: USENIX, 2016. 265-283.
- Abadi, M.L., et al. «Predictive classification of water consumption time series using non-homogeneous markov models.» *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. Tokyo: IEEE, 2017.
- Aggarwal, S., S Chauhan, ve R.J. Prakash. «An automated system to monitor the usage of water in apartments using IOT and artificial neural network.» *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)*. Kannur: IEEE, 2019.
- Anzanello, M.J., ve F.S. Fogliatto. «Learning curve models and applications: Literature review and research directions.» *International Journal of Industrial Ergonomics*, 2011: 573-583.
- Chaplin, M.F. «Water: its importance to life.» *Biochemistry and Molecular Biology Education*, 2001: 54-59.
- Choi, K., J. Yi, C. Park, ve S. Yoon. «Deep learning for anomaly detection in time-series data: review, analysis, and guidelines.» *IEEE Access*, 2021: 120043-120065.
- Chollet, F. «Keras: The python deep learning library.» *Astrophysics Source Code Library*. 2018.
- Hochreiter, S., ve J. Schmidhuber. «Long short-term memory.» *Neural Computation*, 1997: 1735-1780.
- Ibrahim, T., Y. Omar, ve F.A. Maghraby. «Water demand forecasting using machine learning and time series algorithms.» *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)*. Pune: IEEE, 2020.
- Li, D., D. Chen, B. Jin, L. Shi, J. Goh, ve S. Ng. «MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks.» *ICANN 2019 - 28th International Conference on Artificial Neural Networks*. Münih, 2019. 703-716.
- Malinowski, M.R.B. , ve R.J. Povinelli. «Using smart meters to learn water customer behavior.» *IEEE*, 2020: 1-13.
- Merta, J., ve J. Fikejz. «Utilization of machine learning to detect sudden water leakage for smart water meter.» *International Conference on Radioelektronika*. Pardubice: IEEE, 2019.

- Mounce, S.R., ve J.B. Boxall. «Implementation of an on-line artificial intelligence district meter area flow meter data analysis system for abnormality detection: A case study.» *ResearchGate* 10, no. 3 (2010): 437-444.
- Nasser, A.A., M.Z. Rashad, ve S.E. Hussein. «A two-layer water demand prediction system in urban areas based on micro-services and LSTM neural networks.» *IEEE*, 2020: 147647 - 147661.
- Rahim, M.S., K.A. Nguyen, R.A. Stewart, D. Giurco, ve M. Blumenstein. «Predicting household water consumption events: towards a personalised recommender system to encourage water-conscious behaviour.» *International Joint Conference on Neural Networks (IJCNN)*. Budapeşte: IEEE, 2019.
- Terzi, F., ve M. Ocakçı. «Kentlerin geleceği: Akıllı kentler.» *İTÜ Vakıf Dergisi*, 2017.
- Yıldız, T.Z., ve T. Aytekin. «Short term water demand forecasting using regional data.» *2019 27th Signal Processing and Communications Applications Conference (SIU)*. Sivas: IEEE, 2019.

ÖZGEÇMİŞ

MERT CELAL NİŞANCI

ÖĞRENİM BİLGİLERİ

Yüksek Lisans 2018- 2022	Akdeniz Üniversitesi Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Bölümü, Antalya
Lisans 2014-2017	Akdeniz Üniversitesi Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Antalya

MESLEKİ VE İDARİ GÖREVLER

AR-GE ve Proje Mühendisi 2020-Devam Ediyor	Wagner Kablo Mühendislik ve AR-GE, Antalya
---	---