




# Makine Öğrenmesi Yöntemleriyle Anormal İçme Suyu Tüketimlerinin Tespit Edilmesi ve Tahmin Modellerinin Geliştirilmesi\*

İsmail Güney<sup>1\*</sup> , İhsan Hakan Selvi<sup>2</sup> 

<sup>1</sup> Sakarya Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, Sakarya, Türkiye

<sup>2</sup> Sakarya Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, Sakarya, Türkiye

ismailguney@gmail.com, ihselvi@sakarya.edu.tr

## Öz

Bu çalışmada, içme suyu gibi önemli bir ihtiyacın hane halkı tarafından tüketiminde belirli bir düzen olabileceği gibi, farklı etkenlere bağlı olarak düzensiz tüketimin de olabileceği öngörülmektedir. Artan nüfus, sınırlı içme suyu kaynakları, gelişen alt yapı ve teknoloji, içme ve kullanma suyuna olan talebi artırmıştır. Artan talebi karşılamak için alternatif su kaynağı arayışları yanında mevcut suların israf edilmemesinin ve daha verimli kullanılmasının da etkili olacağı öngörülmektedir. Yapay zekanın (AI) alt dalı olan makine öğrenmesi (ML) yöntemleriyle geçmiş dönemlerdeki içme suyu tüketimleri analiz edilmiş, olağan ve olağan dışı tüketim davranış modelleri çıkarılmıştır. İçme suyu mesken abonelerinin anormal tüketimlerinin tespiti ve bilgilendirilmeleri durumunda, hane içi tüketimlerin normal tüketim aralığında kalmasının sağlanacağı öngörülmektedir. Çalışmada Kayseri ili genelinde 2006 – 2022 (ilk 6 ay) tarihleri arasında sayaç endeks okuması 160 dönemden fazla olan 8.224 adet mesken abonesine ait sayaç, abone ve tüketim verileri dikkate alınmıştır. Veriler konumsal abone temelinde birleştirilmiş, 41 öznitelikli veri kümesi elde edilmiş, veri ön işlemleri sonucunda 24 öznitelikli bir veriseti oluşturulmuştur. Çalışmada 6 farklı öznitelik seçim yöntemi kullanılarak alt verisetleri elde edilmiştir. Bütün verisetler 7 farklı anomali analiz yöntemi kullanılarak anormal ve normal içme suyu tüketimleri tespit edilmiştir. Anomali analizleri sonucunda hesaplanan aykırılık puanları kullanılarak bütün tüketim değerleri 4 farklı tüketim sınıfı ile etiketlenmiş, veriseti gözetimli hale getirilmiş, 7 farklı ML sınıflandırma algoritması ile tüketim sınıfı tahmin modelleri geliştirilmiştir. Çalışma sonucunda anormal içme suyu tüketimlerinin ML yöntemleri ile tespit edilebileceği, tüketim sınıflarının tahmin edilebileceği ispatlanmış, suyun israf edilmeden daha verimli kullanımıyla ilgili gerekli politikaların oluşturulabileceği ve bunun için önlemler alınabileceği ortaya konmuştur.

**Anahtar kelimeler:** İçmesuyu, abone, tüketim analizi, makine öğrenmesi, anomali analizi, anomali tespiti, aykırı değer, etkili durum, etkili durumlar, etkili gözlemler.

## Detecting Abnormal Drinking Water Consumptions And Developing Forecast Models By Machine Learning Methods

### Abstract

In this study, it is predicted that there may be a certain order in the consumption of an important need such as drinking water by the household, as well as irregular consumption depending on different factors. Increasing population, limited drinking water resources, developing infrastructure and technology have increased the demand for drinking and utility water. There is a search for alternative water sources to meet this demand, but it is foreseen that these demands can be met by not wasting existing water and using it more efficiently. By using machine learning (ML) methods, which is a sub-branch of artificial intelligence (AI), drinking water consumption data in the past periods were analyzed, and ordinary and unusual consumption behavior models were extracted. It is envisaged that by detecting abnormal consumptions that may occur in drinking water consumption and informing the subscribers about this issue, it will be ensured that the consumption in the household remains within the normal consumption range. Although the amount of data collected, recorded and processed in today's IT world has increased significantly, it is known that the exact analysis is difficult in terms of time and cost. In this study, subscriber, meter, consumption, bill and payment data of 8,224 residential subscribers, whose water meter index reading is more than 160 periods throughout the province of Kayseri, between 2006 and 2022 (first 6 months) were taken into account. The data are combined on a spatial subscriber basis and a 41-features dataset is obtained. The dataset was transformed into a dataset

\* Bu çalışma İsmail GÜNEY'in yüksek lisans tezinden üretilmiştir.

Sorumlu yazar.

E-posta adresi: ismailguney@gmail.com

Alındı : 30 Aralık 2022

Revizyon : 21 Temmuz 2023

Kabul : 1 Ağustos 2023

with 24 features as a result of data preprocessing. In the study, 6 sub-datasets were obtained by using information gain (IG), gain ratio (GR), symmetric uncertainty coefficient (SU), pearson correlation coefficient (r), f-score and random forest (RF) feature selection methods. The 7th sub-dataset was obtained from the intersections of the selected features in the sub-datasets. In all datasets, abnormal and normal drinking water consumptions were determined by using 7 different ML anomaly analysis methods: tukey outlier labeling (TOL), forest of isolation (IF), z-score, copula-based outlier detection (COPOD), median absolute deviation (MAD), local outlier factor (LOF), and elliptical envelope (EE). At the beginning of the study were unsupervised drinking water consumption data at the end of the study, labeled as 4 different classes and the dataset was made supervised. Using the finally obtained supervised dataset, decision trees (DT), gaussian naive bayes (NB), k-nearest neighbors (KNN), logistic regression (LJR), multilayer perceptron neural network (MLP-NN), RF and gradient boosting (GB) have been developed consumption class estimation models with 7 different ML methods. As a result of the study, it has been proven that abnormal drinking water consumption can be detected by ML methods, and it has been revealed that necessary policies can be created for more efficient use of water without wasting water and measures can be taken for this.

**Keywords:** Drinking water consumption analysis, machine learning, anomaly analysis, anomaly detection, outlier, effective case, effective cases, effective observations.

## 1. Giriş (Introduction)

Su, canlıların beslenme, temizlenme, ulaşım ve taşıma gibi temel nedenlerle sürekli ihtiyaç duyduğu bir araç olmakla birlikte özelde insanoğlunun da üzerinde her zaman hesap yapageldiği önemli bir değer olmuştur. Milattan önceki çağlardan günümüze kadar insanlık suyun temini, yönetilmesi, paylaşılması, taşınması ve kullanılması hususlarına önem vermiş, yer yer bu önem Tablo 1’de gösterilen çatışmaların yaşanmasına neden olmuştur. Çatışmalar suyun bir silah olarak veya savaşların tetikleyici nedeni olarak görülmesiyle veya su kaynaklı kazaların oluşmasıyla gerçekleşmiş böylece insanlığın tarih içindeki mücadelesinde önemli izler bırakan su, halen tarihe, coğrafyaya şekil vermeye devam edegelmiştir (Water Conflict Chronology [Internet], 2022).

**Tablo 1.** Tarihte su ile ilgili yapılan çatışma nedeni ve sayıları. (The reasons and numbers of conflicts related to water in history.)

Dönem Aralığı (Yıllar)	Çatışma Sayısı	Çatışmada	Dönem Aralığı (Yıllar)	Çatışma Sayısı
Milattan Önce	26	Silah, Kaza	10	6
0 - 1000	2	Silah, Tetikleyici,	1	2
1001 - 1900	25	Silah, Kaza, Tetikleyici	13	9
1901 - 2000	209	Silah, Kaza, Tetikleyici	101	14
2001 - 2022	1.036	Silah, Kaza, Tetikleyici	109	16
Toplam:	1.298	Silah, Kaza, Tetikleyici	178	16

Çatışmaların büyük çoğunluğunun son 20 yılda meydana geldiği, ayrıca 20. yüzyılda nüfus artışının dört kat, su talebinin dokuz kat artması dikkate alındığında su güvenliği konusunun günümüz dünyasında hissedilecek kadar hızla büyüyen sosyal, politik ve ekonomik sorunları tetiklediği, giderek yayılan bir çevresel krize dönüştüğü artık kabul edilen bir gerçektir.

Yapılan araştırmalarda 2030 yılına kadar küresel tatlı su talebinin mevcut arzı %40’ın üzerinde aşacağı ve etkilenecek insan sayısının 4 milyara yaklaşacağı, artan rekabet koşulları ve beraberinde su üzerinde oluşturduğu stresin, dünya genelinde gıda, enerji, üretim ve insan güvenliği açısından önemli bir risk ve etki oluşturacağı öngörülmektedir (Cini, Mung ve Waughray, 2014).

Kentsel yaşamın giderek yaygınlaştığı günümüzde en önemli tüketim kaynaklarından biri olan su, potansiyel olarak yüzye, yeraltında, havada farklı formlarda bulunmakla birlikte formlar arası değişim ve geçiş döngüsüne sahip olarak hayatın içinde yer almaktadır. Ülkemizde ve dünyada 20 yılı aşkın bir süre zarfında nesnelerin interneti (IOT) tabanlı cihazların, kablosuz haberleşmenin ve AI’nın gelişimiyle akıllı şehir konseptleri hızla yaygınlaşmış ve giderek alanını genişletmiş, hatta oldukça kompleks sorunları da çözebilecek etkinliğe kavuşmuştur. Akıllı şehirler kavramı beraberinde akıllı su şehirleri kavramına zemin oluşturmuş böylece Uluslararası Su Birliği (IWA) tarafından 4 aşamalı Akıllı Su Şehri İlkeleri Uygulama Planı oluşturulmuştur. Çalışma konusu ve sonuçlarının bu plan içerisinde tüketilen su ve enerji miktarının azaltılması, mevcut suyun en verimli şekilde kullanılması, tüketim davranış analizlerinin sunduğu geri dönütler ile abonelerin bilinçlendirilmesi, tüketimin bu sayede disipline edilmesi ve nihai olarak da tespit edilen aykırı tüketimlerin dikkate alındığı kayıp ve kaçak denetimleri ile su yönetiminin daha bilimsel yapılabilmesi hususlarında katkı sağlayacağı düşünülmektedir (International Water Association, 2017).

Su dağıtımının Yönetimsel Kontrol ve Veri Toplama (SCADA) sistemleri ile yönetilmeye başlanması, verilerin toplanması ve kayıt altına alınması ile ciddi bir veri havuzu oluşmuş, beraberinde verilerle anlamlı sonuçlar elde edilebilecek analizlerin önü açılmış daha kompleks çözümler sunan AI modelleri geliştirilmiştir. Mevcut durumda su sektöründe AI ile geliştirilen uygulamalar su alt yapısının fiziki durumunun tespiti ve bakımı, su talebi ve tüketiminin tahmini, su

rezervuarlarının ve barajların sağlık ve çevresel etkilerinin izlenmesi, su kalitesinin izlenmesi ve su ile ilgili felaketlerin öngörülmesi ve izlenmesi olarak 5 başlıkta raporlanmıştır. Öyle ki AI destekli yeniliklerin 2030 yılına kadar küresel ekonomiye 200 milyar dolar katkı sağlayacağı tahmin edilmektedir (Yıldız ve Özgüler, 2020).

Küresel su döngüsü dikkate alındığında suyun gelecek yıllarda çoğu ülkeleri etkileyeceği, su ile ilgili yapılan menfi çalışmaların etkisinin yeterli olamayacağı, ülkelerin birlikte hareket etmesinin özellikle de gelişmiş ülke tecrübelerinin ve imkanlarının gelişmekte olan ve gelişmemiş ülkelere aktarılmasının gerekliliği bir zorunluluk halini almıştır. Bundan dolayıdır ki Birleşmiş Milletler (BM) bünyesinde bağlı bütün üye ülkelerin 2030 yılına kadar uygulayabilecekleri Sürdürülebilir Kalkınma Hedefleri (BM SKH) belirlenmiş, bu hedeflerden özellikle su ile ilgili olanlar için geliştirilen AI çalışmalarıyla hedeflenen en önemli gaye 2030 yılına kadar herkesin güvenli ve erişilebilir içme suyuna kavuşmasını sağlamak için, altyapıya yatırım yapmak, sıhhi tesisleri inşa etmek ve her düzeyde hijyeni teşvik etmek zorunluluğu olmuştur. Ayrıca hedefler arasında su kıtlığını hafifletmek için ormanlar, dağlar, sulak alanlar ve nehirler gibi suyla bağlantılı eko-sistemleri korumak ve eski haline getirmek zorunluluğu yanında gelişmekte olan ülkelerde su verimliliğini teşvik etmek ve arıtma teknolojilerini desteklemek için uluslararası iş birliğine de vurgu yapılmıştır (Türkiye Cumhuriyeti Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2020).

Günümüzde musluğumuzu çok fazla hesap yapmadan açabiliyor olsak da bu durumun süreklilik arz etmeyeceği, su kaynakları kullanımının doyuma ulaşacağı, nüfus ve suya olan ihtiyacın artmaya devam edeceği konunun uzmanları tarafından da öngörülmektedir. Gelişen teknoloji ile birlikte suyun daha profesyonel anlamda değerlendirilmesine ve gelecekteki su ihtiyacına yönelik tahmin senaryolarının üretilmesine yönelik çok ciddi imkanlar elde edilmiş olup, bu senaryolara göre 2040 yılına gelindiğinde yerel su kaynaklarının yeniden kullanılacağı, geri dönüşümün zorunluluk halini alacağı, yağmur suyu yönetiminin yaygınlaşacağı, yeterince kullanılmayan su kaynaklarına erişimlerin sağlanacağı, doğrudan suyu arıtma ve arıtım yollarının yeşil alanlara veya parklara dönüştürüleceği, kentsel su hizmetlerinin gizli hizmetler olmaktan çıkarılıp daha görünür hale getirileceği ve teknolojinin bireysel haneler üzerinde büyük bir etkisi olacağı, bir mobil uygulama aracılığıyla gün boyunca su kullanımlarının gerçek zamanlı takip edilmesinin sağlanacağı, bu vesileyle tüketim davranış değişikliklerinin, hangi amaçlarla ne kadar su kullanıldığının ve hangi saatlerde kullanıldığının, şebekedeki su kayıp ve kaçaklarının tespitinin, su kullanımlarının faturaları nasıl etkileyeceğinin

görülebileceği bir kentsel su modeli bizleri beklemektedir (Arup ve Sydney Water, 2015).

Hane içinde su tüketimleri incelendiğinde iç mekandaki tüketimlerin Tablo 2’de gösterildiği oranlarda olduğu, hane tüketimlerinin hane büyüklüğü ile orantılı olduğu ancak büyüklüğe oranla tüketimin daha az arttığı ve bir tasarruf yapılacaksa nerelerde yapılabileceği öngörülebilecektir (Mayer, DeOreo, Opitz, Kiefer, Davis, Dziegielewski, ve Diğerleri, 1999).

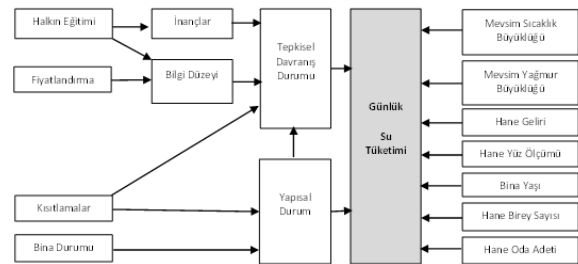
**Tablo 2.** Suyun hane içinde nihai kullanım yerleri ve oranları. (Uses and rates of water in the household.)

Son Kullanım	İç Mekân Su Kullanım Yüzdeleri
Tuvaletler	% 26,7
Çamaşır Makineleri	% 21,7
Duşlar	% 16,8
Musluklar	% 15,7
Sızıntı	% 13,7
Diğer Lokal	% 2,2
Banyolar	% 1,7
Bulaşık Makineleri	% 1,4
<b>Toplam *</b>	<b>% 100,0</b>

\*ondalık tek haneye indirgenmişinden %99,9

Ayrıca bilimsel çalışmalarda hane halkı dışındaki tüketim oranları her ne kadar kültüre ve coğrafyaya göre değişiklikler gösterse de nüfus ve yüz ölçüm ölçekleri dikkate alınarak yaklaşık olarak hesaplanabilmektedir (Mayer, Dziegielewski, Kiefer, Lantz, Opitz ve Porter, 2000).

Hanelerin günlük su tüketimlerini etkileyen Şekil 1’de gösterildiği gibi birçok faktör belirlenmiş ve günlük su tüketim değerlendirmesi yapılırken bunlar göz önünde bulundurulmuş, bunların haricinde de birçok etmenin var olduğu dikkate alınarak bu etmenler arasındaki ilişkilerin tüketimi tetiklediği gerçeğinden hareketle bir konut suyu koruma modeli uyarlanmıştır (Billings ve Jones, 2008).



**Şekil 1.** Konut suyu koruma modeli (Residential water protection model.)

Bu çalışmanın amaçlarından biri olan su tasarrufuyla ilgili Tablo 3’te gösterildiği gibi 5 ana başlık altında bir dizi önerilerde bulunulmuş, su yöneticileri ve nihai tüketicilerin bu önerilerden birçoğunu yaptıkları, ilgili planlamaların giderek yaygınlaştığı sonraki dönemlerde

daha fazlasının hayatın içinde yer alacağı görülmektedir (Campbell, Johnson ve Larson, 2004).

**Tablo 3.** Su tasarrufu ile ilgili öneriler. (Tips on saving water.)

Yönetmelikler	Kamuya Yönelik Bilgilendirmeler	Mühendislik	Fiyatla İlgili Stratejiler	Dolaylı Mühendislik
<ul style="list-style-type: none"> <li>Az su kullanan armatürleri belirten tesisat kodları</li> <li>Düşük su kullanımı gerektiren peyzaj düzenlemeleri</li> <li>kullanımı gerektiren peyzaj düzenlemeleri</li> <li>Su israfı yasağı</li> <li>Belirli dış mekân amaçları için ham veya işlemez su gereksinimleri</li> <li>Yeniden satışta güçlendirme gereksinimleri</li> </ul>	<ul style="list-style-type: none"> <li>Genel tanıtım: reklam panoları, postalar, su faturası ekleri</li> <li>Herkese açık forumlar</li> <li>Gazete, radyo, TV, internet kullanımı</li> <li>İlköğretim ve ortaöğretimlerde eğitim programları</li> <li>Verimli sulama sistemlerini teşvik etmek için çalıştaylar</li> <li>Su bazlı çevre düzenlemesini teşvik eden atölyeler (xeriscape)</li> <li>Günlük hava durumu raporlarına peyzaj sulama ihtiyaçlarının eklenmesi</li> </ul>	<ul style="list-style-type: none"> <li>Yeni nesil tuvaletler için ücretsiz dağıtım veya sübvansiyonlar</li> <li>Düşük akışlı duş başlıkları ve musluk havalandırıcılarının ücretsiz dağıtımı</li> <li>Az su kullanan giysiler veya bulaşık makineleri için sübvansiyonlar</li> <li>Peyzaj ve yağmurlama sistemi dönüşümü için sübvansiyonlar</li> <li>Yağmur suyunun toplanması ve peyzajda kullanımı için sübvansiyonlar</li> <li>Çevre düzenlemesinde gri suyun yeniden kullanımı için sübvansiyonlar</li> <li>Geri kazanılmış su dağıtım sisteminin genişletilmesi</li> <li>Sızıntı tespit programları</li> </ul>	<ul style="list-style-type: none"> <li>Emtia ücretleri olan sayaçlar</li> <li>Su dağıtımlarının evrensel ölçümü</li> <li>Suyun marjinal fiyatının faturalarda belirgin şekilde gösterilmesi</li> <li>Su fiyatındaki artış</li> <li>En yoğun talep dönemleri için daha yüksek su fiyatları</li> <li>Su için artan kademe oranları</li> <li>Su bütçeleri</li> <li>Ekonomik açıdan çekici hale getirmek için geri kazanılan suyun sübvansiyonu</li> </ul>	<ul style="list-style-type: none"> <li>Kişisel etkileşim içeren herhangi bir mühendislik çözümü</li> <li>Yüksek su kullanımı olan müşterilere yönelik su kullanımı denetimleri</li> </ul>

Çalışma ile birlikte gerek sular idarelerinin gerekse abonelerin kendilerine sorabilecekleri hızla artan kent nüfusu ve su ihtiyacı nasıl karşılanabilir, giderek artan bir şekilde hissedilen su kıtlığı ve çevresel bozulma ile karşı karşıya kalınan bir dünyada adil su hizmetleri sağlanabilir mi, içme ve kullanma su tüketimini etkileyen faktörler nelerdir, su tüketiminde bir düzen bir davranış kalıbı var mıdır, varsa bu bir mevsimsellik taşıyor mudur, tüketim düzenini bozan dönemler var mıdır, tüketim düzenini bozan aykırı tüketimler ile sayaç okuma hatası, kayıp kaçak tüketim, sayaç ölçüm arızası, tüketimde israf vb. tespitler yapmak mümkün müdür, tüketim davranışı AI alt dalı olan ML algoritmaları ile modellenilebilir mi, tüketim davranışı disipline edilebilir mi, tüketim üzerinden su tasarrufu ile kaynak arayışına ciddi bir katkı sağlanabilir mi sorularına cevaplar sunulmaya çalışılmıştır.

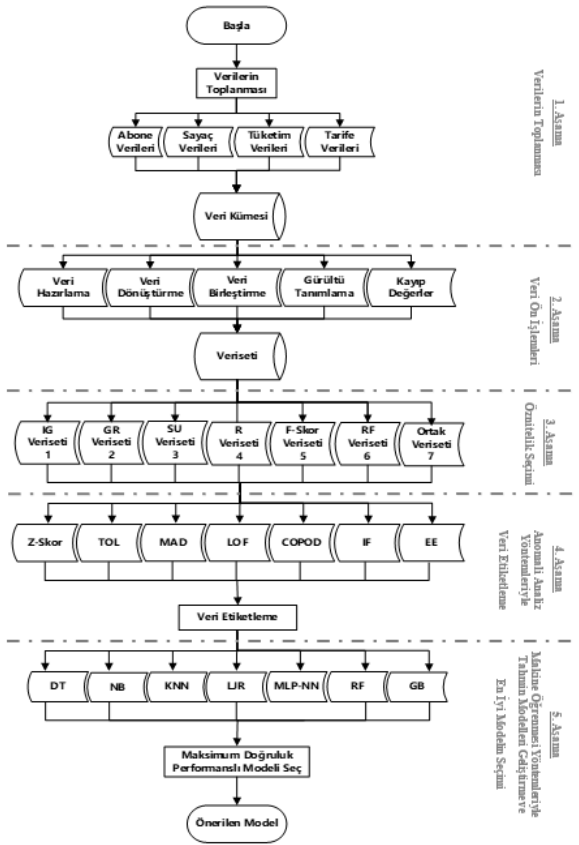
Çalışmanın amacı içmesuyu tüketim davranışlarını tespit etmek, alternatif su kaynağı arayışının aksine varolan kaynakları daha verimli kullanmaya katkı sağlamak, tüketimi etkileyen faktörlere göre aykırı tüketim davranış sınırları belirlemek ve bunları takip etmek, abonelerin israf etmeden su kullanımına dolaylı olarak da ekonomisine katkıda bulunmak, kurumun çok yüksek taleplerde yaşayacağı zorlukların üstesinden gelmesine katkıda bulunmak, kayıp kaçak kullanım potansiyeli olan aboneleri tespit etmek ve takip edilmesini sağlamaktır.

Bu çalışma kapsamında Kayseri ili içme ve kullanma su abonelerinin hane içi su tüketimleri dikkate alınmış, dünyada ülkemizde su tasarrufuna, bu konularda toplumsal bilincin artırılmasına ve su kayıp kaçak denetimlerine yönelik yapılan ML çalışmalarına katkıda bulunacak bir model geliştirilmiş, bununla birlikte

sadece bireysel davranışları etkilemeye yönelik geri dönütler ile bir tasarrufun sağlanması, beraberinde su yöneticilerinin yatırım planlamalarına ve yönetim yaklaşımlarına farklı bir bakış açısı sunacak analizlerin yapılması ve bulguların paylaşılması hedeflenmiştir.

## 2. Yöntem (Method)

Çalışma Şekil 2’de gösterildiği gibi beş aşamada gerçekleştirilmiş olup ilk aşamada ihtiyaç duyulacağı öngörülen ham veriler temin edilerek bir veri kümesi elde edilmiş akabinde veri kümesi veri ön işlemlerinden geçirilerek analiz için uygun bir zaman serisi veriseti elde edilmiştir. Üçüncü aşamada öznitelik seçim yöntemleriyle alt verisetler türetilmiş ardından gözetimsiz ML ve istatistikî teknikler kullanılarak anormal tüketim tespitleri ve tüketim sınıf etiketlemesi yapılmıştır. Son aşamada ise etiketlenmiş gözetimli hale getirilmiş veriseti ile ML teknikleri kullanılarak tüketim sınıfı tahmin modelleri geliştirilmiş ve model performansları karşılaştırılarak en iyi model seçimi yapılmıştır.



Şekil 2. Çalışma kapsamında önerilen model şeması. (Model scheme proposed in the study.)

## 2.1. Çalışma alanı ve verilerin toplanması (Workspace and data collection)

Çalışma kapsamında Kayseri il genelinde okuması yapılan toplam 668.823 aboneden 610.821 adet konut abonesine ait 1980 – 2022 (ilk 6 ay dahil) tarihsel dönem aralığında abone, sayaç okuma ve tarife verilerinden oluşan bir veriseti kullanılmıştır. Çalışmada Tablo 4'te gösterildiği gibi çok fazla sayıda abone analizinin yapılmasının çalışmaya ayrıca bir katkı sunmayacağı ve çok fazla zaman ve işleme neden olacağı için okuma dönem adet sayısı 160 ve üzeri olan 8.224 adet konut abonesinin verileri incelenmiştir.

Tablo 4. KASKİ okuma dönem sayılarına göre abone sayısı tablosu (Number of subscribers according to KASKİ reading period numbers.)

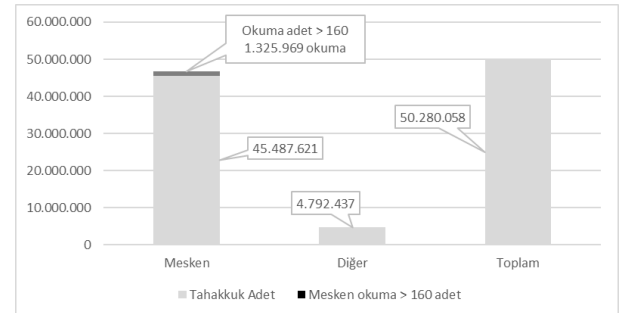
Dönem Adet	Abone Sayısı	Mesken Sayısı
120'den az	454.574	409.292
120 ve üzeri	214.249	201.529
140 ve üzeri	140.491	133.914
160 ve üzeri	8.286	8.224

Literatürdeki çalışmaların çoğu, abonelerin tamamı üzerinden kümülatif talepleri dikkate almış, elde edilen modeller değerlendirilirken bütün abonelerin ortak davranışı varsayımıyla hareket edilmiş ve dolayısıyla il geneli hesaplanan hane başına düşen ortalama TÜFE ve

nüfus artış/azalış oranları da dikkate alınmıştır. Bu çalışmada ise her bir aboneye özel aykırı tüketim tespit modeli oluşturulduğundan ekonomik göstergeler ve nüfusa ilişkin istatistik verilerin her haneye özel, dikkate alınması gerekmektedir. Ancak aboneye özel istatistik değerlerin temini ve sürekli güncel tutulması birçok kurumda tutulan verilerin (maaş, harcama, hanede yaşayanların doğum, ölüm, evlenme, evcil hayvan, bitki vb.) online entegrasyonu ile veya su yönetimi idareleri tarafından ABYS'de haneye özel veri alanları tanımlanıp bilgilerin sahadan personel marifetiyle toplanması ile mümkün olacaktır. Söz konusu bu durumda abonelerin rızası, sular idarelerinin ihtiyaç öncelikleri ve bazı yasal düzenlemeler dikkate alındığında mümkün olmadığı ancak uzun vadeli teknik ve teknolojik gelişmeler ile birlikte (e-Devlet vb.) elde edilebileceği öngörülmüş, haneye özel ekonomik göstergeler ve nüfusa ilişkin istatistik veriler girdi değerlerine dahil edilmemiştir.

Ayrıca sayaç okuması yapılan bölgelerin iklimsel verileri okumacı veya sabit meteoroloji istasyonları marifetiyle tespit edilse dahi hane içi iklim şartlarının da haneden haneye farklılıklar taşıyacağı gerçeğinden hareketle iklim verileri girdi değeri olarak dahil edilmemiştir.

Analiz için temin edilen 2022 yılı 6. Ay sonu itibarıyla toplam 50.280.058 adet okuma verisinden, en az 160 dönem sayaç okuma gören mesken abonelere ait Şekil 2.6'da da gösterildiği gibi 1.325.969 adet tüketim verisi değerlendirilmiştir.



Şekil 3. Mesken ve diğer abone türlerine göre tahakkuk sayıları grafiği. (Graph of accrual numbers by residence and other subscriber types.)

Çalışma ile birlikte hanelere özel bazı istatistik veriler kullanılamamış olsa da çalışma sonunda elde edilen modele KASKİ veya aboneler tarafından tüketimi etkilediği düşünülen özel alanların da sonradan eklenebileceği modüler bir yapı sunulmuştur.

## 2.2. Veri ön işlemleri ve verisetinin hazırlanması (Data preprocessing and dataset preparation)

### Veri Hazırlama

Çalışma konusu analiz için ihtiyaç duyulan sayaç okuma, abone ve tarife verileri KASKİ'den csv formatında temin edilmiştir.

Çalışmada içme suyu abonelerinin tüketim miktarı çıktı değişkeni olarak belirlenmiş, hedef entropi  $H(Y)$

tüketim değerleri kullanılarak hesaplanmıştır. Tüketimi etkileyen özniteliklerin temel istatistiksel değerler Tablo 5'te gösterildiği gibi hesaplanmıştır.

**Tablo 5.** Veriseti öznitelik seçiminde kullanılan temel istatistiksel değerler tablosu (Table of base statistic values used in dataset feature selection.)

Öznitelikler	Min	Max	Ortalama - $\mu$ -	Ortanca Medyan	Standart Sapma - $\sigma$ -	Varyans	Hedef Entropi $H(Y)$	Girdi Entropi $H(X)$
gecen_gun	3	93	35,49	35	10,19	103,84	4,24	1,98
fatura	4,35	165,37	35,85	29,24	26,85	720,72	4,24	4,14
odeme	4,35	165,37	35,93	29,24	26,88	722,65	4,24	4,14
ocak	0	30	3,13	0	7,62	58,12	4,24	0,55
subat	0	27	2,90	0	6,51	42,43	4,24	0,64
mart	0	31	3,16	0	7,08	50,08	4,24	0,71
nisan	0	30	2,96	0	7,63	58,19	4,24	0,61
mayıs	0	31	3,06	0	7,99	63,77	4,24	0,52
haziran	0	26	2,91	0	6,42	41,23	4,24	0,78
temmuz	0	31	2,97	0	6,91	47,73	4,24	0,71
agustos	0	31	2,97	0	7,22	52,10	4,24	0,76
eylul	0	30	2,78	0	7,38	54,47	4,24	0,59
ekim	0	27	2,90	0	6,70	44,94	4,24	0,71
kasım	0	30	2,77	0	7,06	49,91	4,24	0,60
aralik	0	31	2,98	0	7,83	61,31	4,24	0,54
kademe	1	2	1,01	1	0,08	0,01	4,24	0,04
K1	0	1	0,69	1	0,46	0,21	4,24	0,17
K2	0	1	0,10	0	0,29	0,09	4,24	0,12
K3	0	1	0,21	0	0,41	0,17	4,24	0,25
S1	0	1	0,08	0	0,27	0,07	4,24	0,15
S2	0	1	0,59	1	0,49	0,24	4,24	0,17
S3	0	1	0,33	0	0,47	0,22	4,24	0,15

#### Veri Dönüştürme

Sayaç okuma dönemlerinde biriken endekslerin bir sonraki okuma döneminde kayıt altına alındığı, dolayısıyla bazı dönemsel tüketimlerin olduğundan fazla görülebileceği, bu durumun analiz sonuçlarını yanıltabileceği öngörülmüş ve öncelikle Denklem 1'de gösterildiği gibi tüketim miktarı özneliği oluşturulmuştur.

$$su\_toplam\_m3 = last\_index - first\_index \quad (1)$$

Sayaç okuma tarihi ile önceki okuma tarihi farkından da Denklem 2'de gösterildiği gibi tüketim süresi hesaplanmış yeni bir öznitelik olarak eklenmiştir.

$$gecen\_gun = reading\_date - pre\_reading\_date \quad (2)$$

Hesaplanan tüketim gün sayısının yılın hangi ayından kaç günü kapsadığı bilgisine ulaşmak için yılın her ayı için 12 yeni öznitelik oluşturulmuş, hangi ayda kaç günlük bir tüketim olduğu hesaplanarak ilgili aya işlenmiştir. Böylece tüketim yapılan günler ilgili ayın ağırlık çarpanları olarak kullanılmıştır.

Her bir aboneye ait sayaç endeks değerlerinden farklı sayaç kullanımları tespit edilmiş, tüketimlerinin kaçınıcı sayaç ile yapıldığı bilgisini gösteren yeni bir öznitelik ("sayac\_durum") Tablo 6'da gösterildiği gibi geliştirilmiştir.

**Tablo 6.** Verisetindeki endeks hareketleri üzerinden sayaç bilgisi dönüşüm tablosu (Counter conversion table over index movements in the dataset.)

Okuma Tarihi	İlk End.	Son End.	Sayac durum	S 1	S 2	S 3
22.02.2006	262	268	S1	1	0	0
---	---	---	---	-	-	-
14.11.2007	0	3	S2	0	1	0
---	---	---	---	-	-	-
6.06.2022	17	29	S3	0	0	1

Her abonenin tüketim gözlem noktalarındaki sözleşme numarası bilgisinden hareketle tüketimin farklı bir kullanıcı tarafından yapıldığı bilgisine ulaşılmış yeni bir öznitelik ("kullanici\_durum") Tablo 7'de gösterildiği gibi geliştirilmiştir.

**Tablo 7.** Verisetindeki sözleşme numarası üzerinden kullanıcı bilgisi dönüşüm tablosu (User conversion table over agreement number in dataset.)

Okuma Tarihi	Sözleşme No	Kullanıcı Durum	K 1	K 2	K 3
22.02.2006	12001815	K1	1	0	0
---	---	---	-	-	-
17.05.2018	357910	K2	0	1	0
---	---	---	-	-	-
6.06.2022	402841	K3	0	0	1

Verisetinde her bir okuma dönemine ait 5 kademeli bir faturalandırma bilgisi (bill\_lv1, bill\_lv2, bill\_lv3, bill\_lv4, bill\_lv5) mevcut olduğundan, bu bilgiler kullanılarak kademe durumunu gösteren bir öznitelik ("kademe") Tablo 8'de gösterildiği gibi geliştirilmiş kademe tutar bilgi öznitelikleri verisetinden çıkarılmıştır.

**Tablo 8.** Verisetindeki kademe tutar bilgi alanları ve dönüşüm tablosu (Price level amount information fields and conversion table in the dataset.)

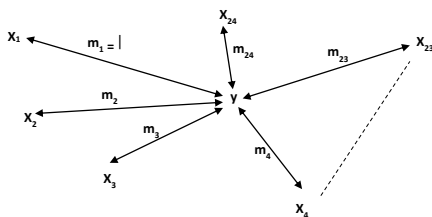
Okuma Tarihi	bill lv1	bill lv2	bill lv3	bill lv4	bill lv5	kademe
22.02.2006	6,67	0,00	0,00	0,00	0,00	1
---	-	-	-	-	-	---
31.03.2022	13,82	16,38	0,00	0,00	0,00	2

Ayrıca verisetini oluşturan öznitelik değerlerinin gün, ay, yıl, ₺, m3 gibi farklı birimlerden oluşması dikkate alınarak, grafik gösterim ve performans hesaplamalarının daha sağlıklı yapılabilmesi için her bir öznitelige ait değerler Denklem 3'te gösterildiği gibi normalize işlemi uygulanarak 0 ile 1 aralığına indirgenmiştir.

$$X_n = \frac{(X_g - X_{min})}{(X_{max} - X_{min})} \quad (3)$$

Burada  $X_n$  normalize edilmiş gözlem değeri,  $X_g$  gözlemin gerçek değeri ve  $X_{max}$ ,  $X_{min}$ 'de sırasıyla ilgili gözlemin en büyük ve en küçük değerleridir.

Ayrıca çalışmada mesafe ölçüm yöntemi kullanılarak ta analiz yapılmıştır. Mesafe ölçüm yöntemlerinden girdi öznitelik sayıları ikiden fazla olduğu için Şekil 4'te de gösterildiği gibi minkowski yöntemi kullanılmış, gözlem noktaları mesafe değerleri hesaplanmıştır.



**Şekil 4.** Bağımsız değişkenlerin bağımlı değişkene mesafelerinin minkowski ile ölçüm gösterimi. (Measurement representation of the distances of the independent variables to the dependent variable with the minkowski method.)

Bağımsız değişkenler olan  $X_i$  girdi değerlerinin, bağımlı değişken olan  $y$  değerine mesafeleri  $m_i$  olarak alınmış olup Denklem 4'te gösterildiği şekilde her bir gözlem noktasına yani sayaç okuma noktası özniteliklerine ait değerler tek bir mesafe değerine dönüştürülmüştür.

$$Minkowski = d(X, y) = [\sum_{i=1}^n (|X_i - y|^p)]^{\frac{1}{p}} \quad (4)$$

Burada  $p$  girdi öznitelik yani boyut sayısı,  $y$  çıktı değeri ve  $X_i$ 'de her bir girdi öznitelik değeridir. Ayrıca negatif korelasyona sahip özniteliklerin  $X$  değerleri  $-X$  olarak alınmış, her bir tüketim gözlem noktası mesafesi toplam mesafe olarak hesaplanmıştır.

#### Veri Birleştirme

Çalışmada MS Access veritabanı kullanılarak csv formatında temin edilen 15 öznitelikli abone, 21 öznitelikli tüketim ve 5 öznitelikli tarife verileri, tablolar arası ilişkisel yapı dikkate alınarak birleştirilmiş, her bir abone için bütün özniteliklerin yer aldığı 41 öznitelikli TS bir veriseti elde edilmiştir.

#### Gürültü Tanımlama

Çalışma, hanelerin tüketimlerine odaklandığı için öncelikle içme suyu abonelerinden sadece mesken abonelerine ait veriler seçilmiş, ayrıca en az 160 okuma dönemi okuma yapılan aboneler dikkate alınmış diğer abonelere ait veriler çıkarılmıştır. Örneklem kapsamında belirlenen abonelere ait 41 öznitelikli verisetindeki mükerrer öznitelikler çıkarılmış veriseti nihayetinde 33 öznitelige indirgenmiştir. Ayrıca bütün gözlem değerleri aynı olan öznitelikler de öznitelik seçim yöntemleri kullanılarak analize dahil edilmemiştir.

#### Kayıp Değerler

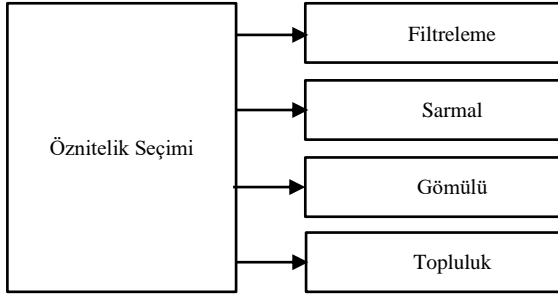
Geçmiş sistem kayıtlarında sayaç değişimlerine ait iş emirlerinin ve sayaç numara bilgilerinin sağlıklı olmaması nedeniyle, sayaç değişim bilgileri, okuma endeks hareketlerinden tespit edilerek elde edilmiştir.

Geçmiş dönemlere ait tarife bilgisine ulaşılmadığında ilgili tarihten bir önceki dönem birim fiyat bilgisi veya fatura tutarı, tüketim miktarına bölünerek elde edilen tutar bilgisi dikkate alınmıştır.

#### 2.3. Öznitelik seçimi ve alt verisetlerin türetilmesi (Feature selection and derivation of subdatasets)

Öznitelik seçimi Şekil 5'te gösterildiği gibi filtreleme, sarmal, gömülü ve topluluk öznitelik seçim

yöntemleri olmak üzere 4 başlıkta incelenebilir (Cai, Luo, Wang ve Yang, 2018).



Şekil 5. Öznitelik seçme yöntemleri şeması. (Feature selection methods schema.)

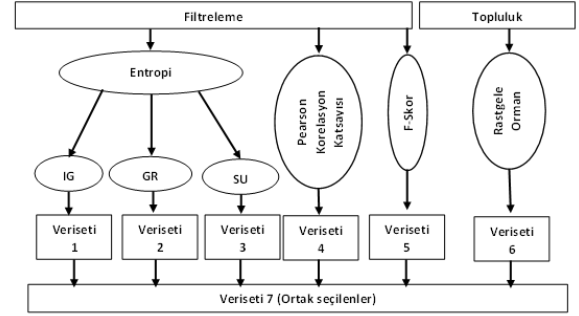
Filtreleme öznitelik seçim yaklaşımı ile veriseti istatistiksel olarak analiz edilir. Temel olarak verisetindeki bütün bilgiyi içeren benzersiz özelliklerin öncelikle puanlanmasına sonrasında sıralanmasına dayalı olarak değerlendirilmesini ve seçilmesini sağlayan bir yöntemdir (Cai, Luo, Wang ve Yang, 2018) (Zhang, Nie, Li ve Wei, 2019).

Sarmal öznitelik seçim yaklaşımında, veriseti denetimli ise sınıflandırıcı algoritmalar kullanılarak, denetimsiz ise kümeleme algoritmaları ile özellikler üzerinde arama işlemi yapılarak alt kümeler oluşturulmakta ve belli sınıflar sadece değerlendirmeye alınmaktadır. Çok büyük ölçekli verisetlerinde bu yaklaşımın kullanılması durumunda her bir alt kümenin değerlendirilmesi fazla miktarda hesaplama işlemleri gerektirdiği için zaman ve kapasite dikkate alınarak tercih edilmesi gerekmektedir (Cai, Luo, Wang ve Yang, 2018) (Zhang, Nie, Li ve Wei, 2019).

Gömülü öznitelik seçim yönteminde model geliştirme sürecinde harcanan zamanın azaltılması amaçlanmakta olup hem sınıflandırma hem de seçim işlemini eşzamanlı olarak model eğitimi aşamasında gerçekleştiren DT, SVM gibi algoritmalar kullanılmaktadır (Cai, Luo, Wang ve Yang, 2018) (Zhang, Nie, Li ve Wei, 2019).

Topluluk öznitelik seçim yönteminde farklı özellikler alt kümeleri elde edilerek oluşturulan verisetleri tarafından eğitilen temel sınıflandırma algoritmaları kullanılarak yapılmış özellik seçim topluluklarının rastgele alt uzay yöntemi (RSM) ve RF yöntemi gibi çoğunluk oylamasıyla daha doğru sınıflandırma ve seçim yapılmasına dayanmaktadır (Cai, Luo, Wang ve Yang, 2018).

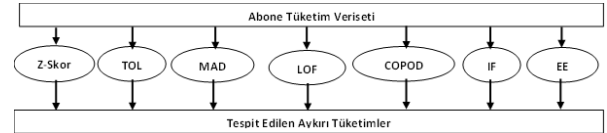
Bu çalışmada özellik seçimi yöntemlerinden Şekil 6'da gösterildiği gibi IG, GR, SU, R, F-Skor ve RF sınıflandırma yöntemleri kullanılmıştır.



Şekil 6. Öznitelik seçme yöntemleri şeması. (Feature selection methods schema.)

### 2.3. Anomali analiz yöntemleriyle veri etiketleme (Data labeling with anomaly analysis methods)

Çalışmada anormal içme suyu tüketim tespitleri, Şekil 7'de gösterildiği gibi 7 farklı anomali analiz yöntemi kullanılarak yapılmış, abonelerin içme suyu tüketimlerini sınıflandırmak ve etiketlemek için yeterli olmuştur. Daha spesifik sınıflandırma ve etiketleme için diğer ML sınıflandırma tekniklerinden de istifade edilebilir.



Şekil 7. Topluluk öğrenme yöntemi ile aykırı tüketim tespit şeması. (Outlier consumption detection scheme with ensemble learning method.)

Aykırı tüketim tespiti yapan her bir yöntemin çalışma şeması Şekil 2.16'da gösterilmiş olup, devamında bütün yöntemlerden elde edilen sonuçlar aykırılık puanlarını elde etmek için değerlendirilmiştir.

#### Veri Etiketleme

Çalışmada tüketim tespit yöntemlerinin sonuçlarına göre aykırılık puan hesabı yapılmış, Tablo 9'da gösterildiği şekilde 4 farklı tüketim sınıfı ve 4 farklı renk ile etiketlenmiştir.

Tablo 9. Aykırılık puanlarına göre sınıflandırma etiketleri tablosu. (Table of classification labels by outlier scores.)

Aykırılık Puanı	Tüketim Sınıfı
0-1	Normal - Yeşil
2-3	Dikkat - Sarı
4-5	Riskli - Turuncu
6-7	Aşırı - Kırmızı

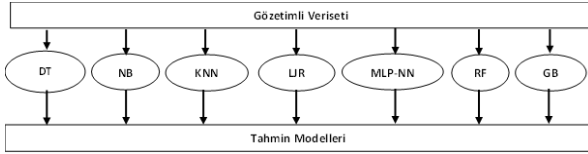
### 2.4. ML yöntemleriyle tüketim sınıfı tahmin modelleri geliştirme (Developing consumption class prediction models with ML methods.)

Çalışma sonunda tüketim sınıfı etiketleme ile gözetimli hale getirilmiş veriseti, Şekil 8'de gösterildiği



gibi DT, NB, KNN, LJR, MLP-NN, RF ve GB olmak üzere 7 farklı ML sınıflandırma algoritmaları kullanılarak içme suyu tüketim sınıfı tahmin modelleri geliştirilmiştir.

Tüketim verileri TS verilerden oluştuğu için ilk yıllar eğitim, son yıllar da kontrol verisi olmak üzere %90'a, %10 oranlarında bölünerek ve bölünmeden kullanılmıştır.



**Şekil 8.** İçmesuyu tüketim sınıfı tahmin modelleri şeması. (Diagram of drinking water consumption class prediction models.)

#### 2.4. Performans ölçümü ve en iyi modelin seçimi (Performance measurement and selection of the best model.)

Çalışmada bütün anomali tespit modellerinin öncelikle TP, TN, FP, FN değerlerinden oluşan hata matrisleri hesaplanmış, Doğruluk, Duyarlılık, Kesinlik, Hassasiyet, Özgüllük, MAE ve MSE performans metrikleri kullanılarak modeller karşılaştırılmıştır.

Her bir abone için geliştirilen anormal tüketim tespit modellerinden en yüksek doğruluk oranındaki model en iyi model olarak belirlenmiştir. Anormal tüketim tespit modellerinden doğruluk oranı aynı olanlar var ise anomali duyarlılık ve anomali kesinlik oranlarına göre yüksek oranda olan en iyi model olarak belirlenmiştir.

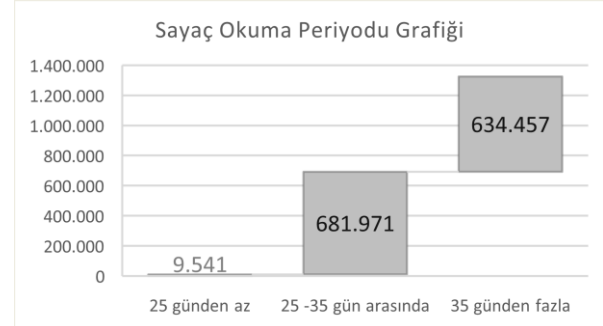
Tüketim sınıfı tahmin modelleri de Doğruluk ve determinasyon katsayısı ( $R^2$ ) performans metrikleriyle karşılaştırılmış, her bir içme suyu abonesi için en iyi tahmin modeli seçilmiştir.

### 3. Bulgular (Findings)

#### 3.1. Anomali analiz sonuçları (Anomaly analysis results.)

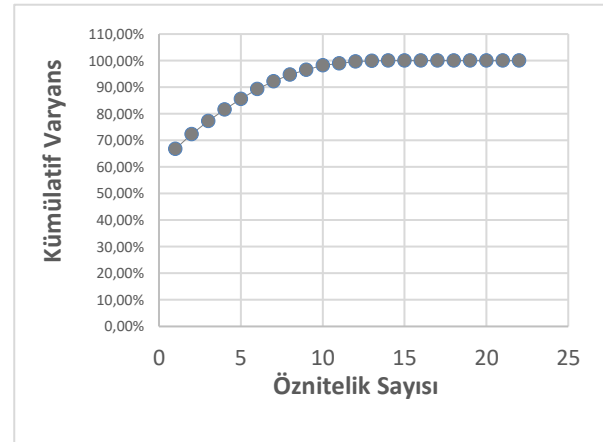
Veriseti incelendiğinde özetle 2006–2022 (ilk 6 ay) yılları arasında kış aylarında içmesuyu tüketimlerinin düştüğü, yaz aylarında tüketim değerlerinin diğer aylara göre daha yüksek olduğu, çalışma kapsamında analizi yapılan abonelere ait toplamda 17.908.783 m<sup>3</sup> tüketim yapıldığı, 1 okuma döneminde konut/mesken su tüketim miktarı en yüksek 818 m<sup>3</sup>, en düşük 1 m<sup>3</sup> ve ortalama 13 m<sup>3</sup> olduğu, yine 1 okuma döneminin en geç 819 gün, en erken aynı gün ve ortalama 36 gün olduğu tespit edilmiştir.

Abonelere ait Şekil 9'da gösterilen sayaç okuma periyodu adetleri incelendiğinde 1.325.969 adet okumanın %0,72'sinin 25 günden az bir zamanda, %47,85'inin 35 günden fazla bir zamanda ve %51,43'ünün ise 25 ile 35 gün aralığında yapıldığı tespit edilmiştir.



**Şekil 9.** Sayaç okuma dönem grafiği (Meter reading period chart.)

Temel veriseti 24 öznitelikten, öznitelik seçim yöntemleri ile elde edilen alt verisetleri ise en az 7 en fazla 12 adet öznitelikten oluşmuştur. Burada kümülatif varyans %90 ve üzerinde olarak Şekil 10'da gösterildiği gibi öznitelik adetleri belirlenmiş, bir sonraki öznitelik kümülatif varyansı %95'in üzerine çıkarıyorsa alt veriseti öznitelik sayısının bir fazlası olarak belirlenmiştir.



**Şekil 10.** Kümülatif varyans oranına göre öznitelik sayıları grafiği (Graph of feature counts by cumulative variance ratio.)

Her bir aboneye ait gerçek değerlerden oluşan ve veri ön işlemleri ile elde edilen 24 öznitelik tamınının olduğu 8.224 adet veriseti ve öznitelik seçim yöntemleri sonrasında elde edilen 57.568 alt veriseti olmak üzere toplamda 65.792 adet veriseti analizde kullanılmıştır. Ayrıca gerçek değerler minkowski mesafe ölçüm yöntemiyle birleştirilerek ve dönüştürülerek 65.792 adet veriseti daha elde edilmiş, toplamda 131.584 adet alt veriseti kullanılmıştır.

Alt verisetler 6 farklı öznitelik seçim algoritması ile belirlenmiş olup, öznitelik seçme algoritmalarının Tablo 10'da gösterildiği gibi aynı öznitelikleri seçme sayıları

ve oranları gösterilmiştir. Öznitelik seçme yöntemlerinden IG ve SU yöntemlerinin ortak olarak belirlenen özniteliklere en yakın oranlarda öznitelik seçimi yaptıkları anlaşılmıştır. Sırasıyla RF ve GR yöntemlerinin çok az da olsa aynı öznitelikleri seçtikleri,

korelasyon ve f skor yöntemleri kullanılarak yapılan öznitelik seçimlerinin diğer yöntemlere göre neredeyse hiç benzerlik göstermediği, birbirleriyle de çok az bir benzerlik taşıdıkları anlaşılmıştır.

**Tablo 10.** Öznitelik seçim yöntemleriyle aynı özniteliklerin seçilme sayıları ve oranları (The number and rate of selection of the same features by feature selection methods)

Öznitelik Seçim Yöntemleri				Bilgi Kazancı	Kazanç Oranı	Simetrik Belirsizlik Katsayısı	Korelasyon	F Skor	Rastgele Orman	Ortak
				IG	GR	SU	R	F	RF	
				V1	V2	V3	V4	V5	V6	V7
Bilgi Kazancı	IG	V1	Adet	495	7.066	84	0	1.018	3.669	
			Oran	%6	%86	%1	%0	%12	%45	
Kazanç Oranı	GR	V2	Adet		587	114	0	340	345	
			Oran		%7	%1	%0	%4	%4	
Simetrik Belirsizlik Katsayısı	SU	V3	Adet			85	0	972	3.641	
			Oran			%1	%0	%12	%44	
Korelasyon	R	V4	Adet				11	51	44	
			Oran				%0	%1	%1	
F Skor	F	V5	Adet					0	0	
			Oran					%0	%0	
Rastgele Orman	RF	V6	Adet						707	
			Oran						%9	

Çalışmada elde edilen alt verisetlerinde Tablo 11’de de özniteliklerin seçilme sayısı ve oranları gösterilmiştir.

**Tablo 11.** Alt verisetlerinde öznitelik seçim yöntemlerinin seçim adetleri ve oranları tablosu (Table of selection numbers and rates of feature selection methods in subdatasets.)

Öznitelik	Alt Verisetlerinde	Seçilme Oranı
Haziran	33.406	%23,47
Temmuz	32.408	%22,77
Ekim	30.908	%21,72
Subat	30.665	%21,55
Eylül	30.451	%21,40
Mart	29.798	%20,94
Kasım	28.814	%20,25
Aralık	26.991	%18,96
Mayıs	26.172	%18,39
Ocak	24.735	%17,38
Agustos	23.464	%16,49
Nisan	22.197	%15,60
S1	14.732	%10,35
Odeme	13.248	%9,31
Kademe	10.852	%7,62
K1	10.150	%7,13
S2	9.856	%6,93
S3	1.879	%1,32
K2	1.428	%1,00
K3	148	%0,10
S4	108	%0,08
K4	16	%0,01
S5	7	%0,00
K5	2	%0,00

İkinci aşamada hem gerçek hem de minkowski mesafe değerleriyle elde edilen alt verisetler, çalışma kapsamında belirlenen 7 farklı anomali analiz algoritması ile analiz edilmiş, anormal tüketim tespitleri, anomali puanlamaları ve içmesuyu tüketim sınıflandırmaları yapılmıştır.

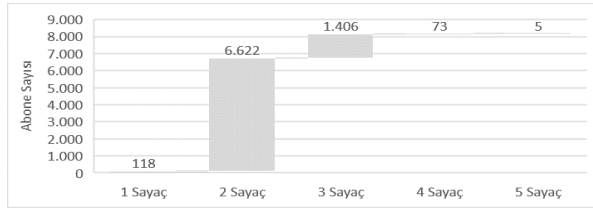
Toplam 1.325.969 adet içme suyu tüketim gerçek değerleri kullanılarak yapılan analizde, 1.196.181 adet normal tüketim, 129.788 adet de anormal tüketim tespit edilmiş olup, bütün tüketimlerde %9,79 oranında bir anomali durum olduğu anlaşılmıştır. Anomali tespit yöntemlerindeki eşik değeri ile bu oran yukarı veya aşağı yönlü değiştirilebilir. Anormal tüketimler de kendi içinde “Dikkat” etiketli tüketim sayısı 62.239 (anormal tüketimlere göre %47,95, bütün tüketimlere göre %4,69), “Riskli” etiketli tüketim sayısı 30.418 (anormal tüketimlere göre %23,44, bütün tüketimlere göre %2,29), “Aşırı” etiketli tüketim sayısı 37.131 (anormal tüketimlere göre %28,61 bütün tüketimlere göre %2,80) adet olarak tespit edilmiştir.

Toplam 1.325.969 adet içme suyu tüketiminin minkowski mesafe değerlerine dönüştürülerek yapılan analizinde, 1.152.927 adet normal tüketim, 173.042 adet de anormal tüketim olarak tespit edilmiş olup %13,05 oranında bir anomali durum söz konusu olduğu anlaşılmıştır. Ayrıca anormal tüketimlerde kendi içinde “Dikkat” etiketli tüketim sayısı 93.177 (anormal tüketimlere göre %53,85 bütün tüketimlere göre %7,03), “Riskli” etiketli tüketim sayısı 50.856 (anormal tüketimlere göre %29,39, bütün tüketimlere göre %3,84), “Aşırı” etiketli tüketim sayısı 29.009 (anormal

tüketimlere göre %16,76, bütün tüketimlere göre %2,19) adet olarak tespit edilmiştir.

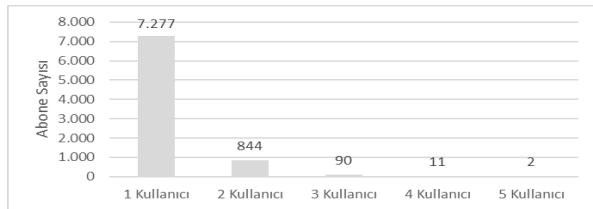
Çalışmada gerçek değerler ve minkowski mesafe değerleri ile yapılan anormal tüketim tespitlerinde, aynı anomali tespit yöntemleri ve eşik değerleri kullanılmasına rağmen farklı sonuçlar elde edilmiştir. Minkowski değerleriyle toplamda %3,26'lık bir oranda daha fazla anormal tüketim tespiti olmasına rağmen, "Aşırı" tüketim sınıf tespitinde %0,61 oranında bir düşüşün olması minkowski değerleri kullanılarak elde edilen tespitlerin gerçek değerlerle tespit edilenlerden daha iyi sonuçlar verdiği şeklinde yorumlanmıştır.

Ayrıca abonelerin en fazla 167 en az 160 dönem sayaç okuması yapılmış olup Şekil 11'de gösterildiği gibi %80,52'sinin tüketim süresince 2. sayacı kullandıkları tespit edilmiştir.



**Şekil 11.** Abonelerin tüketim süresince kullandıkları sayaç durum grafiği (Counter status graph used by subscribers during consumption)

Ayrıca içme suyu abonelerinin Şekil 12'de de gösterildiği gibi %88,48'inin tüketimlerinin aynı kullanıcı tarafından yapıldığı tespit edilmiştir.



**Şekil 12.** Aboneliklere tüketim süresince sözleşme yapan kullanıcı durum grafiği (User status graph contracting subscriptions during consumption)

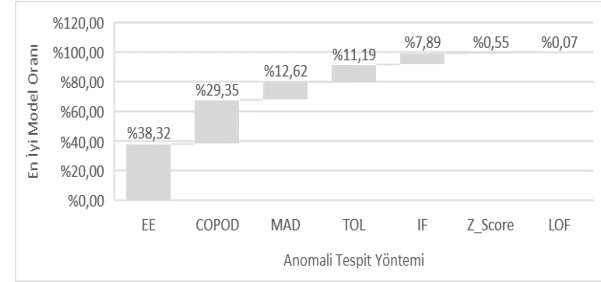
Abone tüketimlerinin tüketim süresince Tablo 12'de gösterildiği gibi yalnızca %0,24'ünün kademeye girdiği, %0,55'inin de ödenmediği tespit edilmiştir.

**Tablo 12.** Sayaç, kullanıcı, kademe ve ödeme durumları özet tablosu. (Counter, user, level and payment status summary table.)

	Kademesiz	Kademeli	Ödenen	Ödenmeyen	Toplam
Abone Sayısı	6.833	1.391	8.224	3.189	8.224
Okuma Sayısı	1.322.781	3.188	1.318.707	7.262	1.325.969

Çalışmada her bir abone için toplamda 112 adet anormal tüketim tespit modeli geliştirilmiş; doğruluk, duyarlılık, kesinlik, hassasiyet, özgüllük, MAE ve MSE

performans metrikleriyle modeller karşılaştırılmıştır. Toplamda 921.088 adet model geliştirilmiş, bunlardan 16.077 adeti en iyi model olarak maksimum doğruluk performansı göstermiştir. Modellerin performans doğruluk oranları en düşük %43, en yüksek %100 ve ortalama %85 olarak tespit edilmiş olup Şekil 13'te gösterildiği gibi EE yöntemiyle elde edilen modellerin çoğunlukla en iyi modeller olduğu, Z Skor ve LOF yöntemleri ile elde edilen modellerin çok düşük oranlarda maksimum doğruluk sağladığı anlaşılmıştır.



**Şekil 13.** Maksimum doğruluk sağlayan anomali analiz yöntemlerinin oranları grafiği (Graph of rates of anomaly analysis methods that provide maximum accuracy)

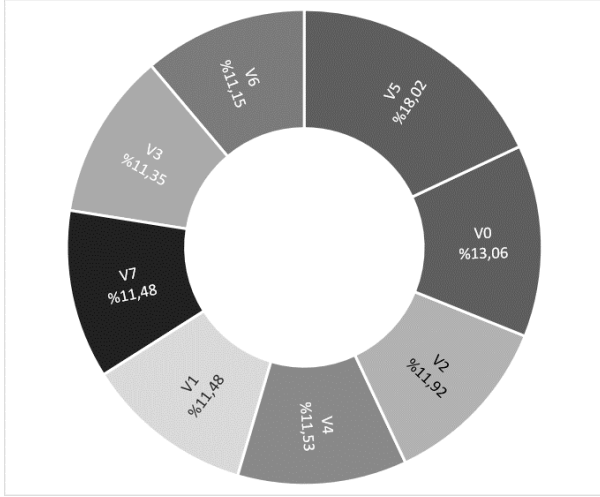
Her bir abone için geliştirilen anormal tüketim tespit modellerinden doğruluk performansı metriği %90 ve üzeri olanlar Tablo 13'te de gösterildiği gibi sıralanır ve doğruluk oranı en yüksek olanı veya kullanılan anomali analiz yöntemine göre veya kullanılan alt verisine göre idare tarafından belirlenebilir. Her bir abone için ayrı olmak üzere belirlenen en iyi anormal tüketim tespit modeli ile bir sonraki okuma döneminde anormal tüketim analizinin yapılması daha az işlem ile ve daha kısa sürede mümkün olacaktır.

**Tablo 13.** Her bir aboneye ait doğruluk performansı metriği maksimum olan en iyi model tablosu (Top model table with maximum accuracy performance metric for each subscriber)

No	Abone No	En iyi Model	Doğruluk
1	12000541	c_V5_COPOD	0,993
2	12000950	g_V1_EE	0,993
3	10003387	c_V4_COPOD	0,993
4	10004150	c_V2_IF	0,993
5	4002847	c_V7_COPOD	0,993
...	...	...	...
8220	12001129	g_V3_IF	0,975
8221	4011157	c_V6_Z_score	0,962
8222	6009693	g_V1_EE	0,950
8223	216960	c_V2_TOL	0,900
8224	217139	c_V4_TOL	0,900

Çalışmada elde edilen bütün tüketim sınıfı tahmin modellerinden, doğruluk performansı metriği %90 ve üzeri olanlar dikkate alındığında, Şekil 12'de de gösterildiği üzere örneklem olarak alınan 8.224

abonenin 7.150 tanesi (%86,94) alt verisetlerinden (V1, V2, V3, V4, V5, V6, V7), diğer 1.074 adet abonenin (%13,06) ise temel verisetinden (V0) elde edilebilmektedir. Her bir aboneye ait geliştirilen maksimum doğruluk performanslı model en fazla oranda f skor öznelik seçim yöntemi ile elde edilen V5 alt veriseti kullanılarak elde edilmiştir. Diğer alt verisetleri ile de birbirine yakın oranlarda başarılı sonuçlar elde edildiği anlaşılmıştır.



**Şekil 12.** Doğruluk performansı %90 ve üzeri olan tahmin modellerinde kullanılan veriseti oranları grafiği (Graph of dataset ratios used in prediction models with an accuracy performance of 90% or more)

Burada işlem sayısı ve süresi azaltılmak istendiğinde her bir abone için alt verisetlerinden en az öznelikli olanı tercih edilebilir.

Model geliştirme sürecinde işlemci hızı 1.1 GHz, RAM kapasitesi 4GB olan bir PC kullanılarak her bir abone için model ön işlemleri ortalama 10 sn, anomali analizleri ortalama 8 sn ve performans ölçüm ve karşılaştırma işlemleri ortalama 3 sn olmak üzere toplamda yaklaşık 96 saat sürmüştür. Çalışma kapsamında analizi yapılan 8.224 abonenin belirlenmiş en iyi modeli kullanılarak anomali tespiti yapılmasında bu süre yaklaşık olarak 51 dk sürmektedir. Böylece önerilen yöntem ile 4 günlük bir analiz ortalama 1 saatten az bir zamanda yapılabilmektedir. Genellikle bütün sular idarelerinde mevcut olan profesyonel donanım altyapısının kullanılması ile çalışma kapsamında önerilen analiz işlem sürelerinin daha da kısaltılması mümkün olacaktır.

Çalışmanın temel amacı olan abonelerin duyarlı hale getirilmesi ile tüketim davranışlarında oluşan anormal tüketimlerinin azaltılması, tüketimlerin normal seviyelere çekilmesi durumunda, tespit edilen ‘Asiri’ sınıftaki tüketimlerde %20 tasarruf ile 139.427 m<sup>3</sup>, ‘Riskli’ sınıftaki tüketimlerden %15 tasarruf ile 154.058 m<sup>3</sup> ve ‘Dikkat’ sınıftaki tüketimlerden de %10 tasarruf ile 169.402 m<sup>3</sup> olmak üzere toplamda

sistemdeki %2,58’lik 462.887 m<sup>3</sup> suyun daha verimli kullanımının sağlanacağı öngörülmektedir.

Abone faturalarında ise ‘Asiri’ sınıftaki tüketimlerde %20 tasarruf ile 284.304,37 ₺, ‘Riskli’ sınıftaki tüketimlerden %15 tasarruf ile 288.181,09 ₺ ve ‘Dikkat’ sınıftaki tüketimlerden de %10 tasarruf ile 349.438,92 ₺ olmak üzere içme suyu fatura toplamlarında %5,15’lik 921.924,38 ₺ tasarrufun sağlanacağı öngörülmektedir.

### 3.2. ML analiz sonuçları (ML analysis results.)

Çalışma sonunda tüketim sınıf etiketlemesi ile gözetimli hale getirilmiş veriseti, DT, NB, KNN, LJR, MLP-NN, RF ve GB olmak üzere 7 farklı ML sınıflandırma algoritmaları kullanılarak içme suyu tüketim sınıfı tahmin modelleri geliştirilmiştir. Tüketim verileri TS verilerden oluştuğu için ilk yıllar eğitim, son yıllar da kontrol verisi olmak üzere %90’a, %10 oranlarında bölünerek ve bölünmeden kullanılmıştır. Geliştirilen modeller Tablo 14’te gösterildiği gibi ACC ve R<sup>2</sup> performans metrikleriyle karşılaştırılmış, her bir içme suyu abonesi için en iyi model seçilmiştir.

**Tablo 14.** İçme suyu tüketim sınıf tahmin modelleri performans tablosu (Performance table of drinking water consumption class prediction models)

ML Model	Veriseti Bölünmeden				%90 Eğitim - %10 Test			
	Minkowski Değerler		Gerçek Değerler		Minkowski Değerler		Gerçek Değerler	
	ACC	R <sup>2</sup>	ACC	R <sup>2</sup>	ACC	R <sup>2</sup>	ACC	R <sup>2</sup>
1 DT	1,00	1,00	0,99	0,73	1,00	1,00	1,00	1,00
2 Gaussian NB	0,94	-0,36*	0,69	-2,88*	1,00	1,00	0,69	0,00
3 KNN	1,00	1,00	0,99	0,73	1,00	1,00	0,94	0,00
4 LJR	0,87	0,00	0,86	-0,27*	1,00	1,00	1,00	1,00
5 MLP-NN	0,94	-0,52*	0,89	-0,97*	1,00	1,00	1,00	1,00
6 RF	1,00	1,00	0,99	0,73	1,00	1,00	1,00	1,00
7 GB	1,00	1,00	0,99	0,73	1,00	1,00	1,00	1,00

\*R<sup>2</sup> değerinin negatif olması modelin çok daha kötü olduğunu göstermektedir.

Minkowski mesafe değerleriyle elde edilen modeller, gerçek değerler kullanılarak geliştirilen modellerden daha iyi performans göstermiştir. Modeller, kurumun isteği doğrultusunda veya belli periyotlarda otomatik olarak yeni veriler de dikkate alınarak güncellenebilir ve böylece daha iyi performans sonuçları elde edilebilir.

## 4. Sonuçlar (Results)

Çalışma sonucunda abonelerin sayaç okuma dönemlerine ait abone, sayaç, sayaç okuma, tarife, tahakkuk ve tahsilat verilerinden tüketim davranış

modelleri geliştirilmiş ve tüketim sınıflandırma etiketlemesi yapılmıştır. Her bir aboneye özel olarak geliştirilen tahmin modeli ile her okuma döneminde sayaç okuma esnasında gerçekleşen tüketimin sınıfı belirlenmesi sağlanmıştır.

İçmesuyu abone tüketimlerinden anormal tüketimlerin tespit edilmesi ve tüketim sınıflarının belirlenmesi ile; tüketim alışkanlıklarının su israfına neden olacak düzeyde değişmiş olabileceği, iç tesisatta bir sızıntı olabileceği, kullanılan içme suyu sayacının ölçüm hassasiyetini kaybetmiş olabileceği, ölçüm dışı tüketim veya kaçak kullanımın söz konusu olabileceği, sayaç okuma memurunun endeksi yanlış girmiş veya sayacı yanlış okumuş olabileceği, sayaç endeks okuma döneminin ay aşırı geç veya daha kısa sürelerde yapılmış olabileceği, bahçe, peyzaj veya tarım amaçlı sulama yapılıyor olabileceği, hanenin boş veya tatile çıkmış olabileceği, hane halkı sayısının doğum, ölüm vb. nedenlerle artmış veya azalmış olabileceği, iklim şartlarının klima kullanımı vb. nedenler ile değişmiş olabileceği değerlendirilmelerinin daha sağlıklı yapılabileceği öngörülmektedir.

Ayrıca çalışmada önerilen yöntemle sular idaresinin her okuma döneminde, anormal tüketim analiz raporu ve anormal içmesuyu tüketimi yapan abone listesi oluşturabileceği, sahada yapılan sayaç arıza, fiziki kayıp kaçak gibi içme suyu denetimlerinde bu listelerden istifade edebileceği öngörülmektedir. Böylece denetimlerde genelleme yapılmayarak daha özenli ve itinalı davranılacağı, anomali tespitinin kontrolünün ve geri dönüşlerinin de sağlanmış olacağı öngörülmektedir. Çalışma ile analizi yapılan abonelerin, yaklaşık 16 yıllık tüketim verilerinden, "Aşırı" sınıflı 27.463 adet tüketim dikkate alındığında yaklaşık olarak aylık 137 adet kontrolün gerektiği, günlük 25 adet kontrol yapan bir ekip ile 1 haftaya yakın bir zamanda kontrol edilmesinin mümkün olduğu öngörülmektedir. Bütün abonelerin çalışmada önerilen yöntemle ait analizlerinin yapılmasıyla daha bütüncül bir kontrol listesi elde edilebilir, ekip sayısı ve anormal tüketim sayısı dikkate alınarak bir planlama yapılabilir. Ayrıca istenirse ekip sayısına ve ekiplerin kontrollerini yapabilecekleri anomali sayısına göre de anomali tespit yöntemlerinin eşik değerleri yukarı veya aşağı yönlü ayarlanarak planlama yapılabilir.

Abonelerin sayaç okumalarının yapılamaması veya okuma dönem sürelerinin ay aşırı günlerde veya daha kısa sürelerde yapılmasının önüne geçilerek, sayaç okumalarının her bir abone için kayıp dönem olmadan her ayın aynı gününde yapılmasının tüketim tahmin model performansını arttıracığı, ayrıca mevsimsel şartların dolayısıyla meteorolojik verilerin hane içindeki tüketimlere etkisi gerçeğinin tek başına yeterli olmayacağı, hanelerde klima gibi hane içindeki iklim şartlarını değiştiren sistemlerin varlığına ait bilgilerin de dikkate alınmasının faydalı olacağı öngörülmektedir.

Ayrıca önerilen yöntem kullanılarak, sayaçların içeride olması ve kullanıcıların evde olmaması durumlarında ya da sayaç arıza dönemlerinde yapılan kıyas tüketimlerin, her aboneye özel olarak tahmin edilmesi veya belirlenmesi daha adilane, daha spesifik ve daha bilimsel bir yaklaşım sağlayacaktır.

Ayrıca çalışmada abone devri, kapama ve açma işlemlerinden dolayı tüketim davranış modellerinin yenilenmesinin uygun olacağı, çalışmada tüketim gün sayısının, kullanıcı değişiminin, sayaç değişiminin, ödeme durumunun dikkate alındığı ancak başkaca parametrelerin de etkili olabileceği gerektiğinde önerilen modele bunlarında eklenebileceği önerilmektedir.

Anomali analizleri sonucunda yapılan tespitler bir öngörü vermektedir ve gerçekliği ancak saha çalışmaları sonucunda gelen geri dönütler ile anlaşılacaktır. Dolayısıyla hangi anomali tespit yöntemi kullanılırsa kullanılsın, abone tüketimlerinin sınıflarının öngörülmesinde tam bir başarı sağlanması arzu edilen bir sonuçtur, fakat tam bir başarı sağlanmazsa dahi önemli olan başarı oranını artıracak adımlar atmak, etkili olduğu anlaşılan başkaca veriler varsa bunları temin etmek ve geri dönütlerle gelen hataları düzelterek ilerlemek olmalıdır.

Ayrıca sular idarelerinin çoğunda yapılan şebeke basınç denetimlerinden elde edilen kayıp kaçak lokasyonları ile bu çalışmadan elde edilen anormal tüketim yapan abonelere ait lokasyonlar eşleştirilerek kayıp kaçak denetimlerinde daha spesifik çözümler elde edilebilir.

Ayrıca abonelere özel tahmin modelleri geliştirildiğinde, bölgelere ayrılmış içmesuyu şebekesinde de ihtiyaç duyulacak aylık, yıllık içme suyu miktarları daha gerçekçi olarak tahmin edilebilecektir. Böylece arz yönlü değil talep yönlü bir bakış açısıyla su yönetimi sağlanmış olacaktır. Bu bakış açısıyla özellikle suya ulaşımın çok zor olduğu bölgelerde ve kuraklık dönemlerinde daha etkili analizler yapılmasına yardımcı olacak ve sonuçlar alınmasını sağlayacaktır. Bu çalışmadaki taleplerin disipline edilmesi hedefi bu amaca hizmet etmektedir.

Çalışmada her bir abone için modellerin geliştirilmesi ve en iyi modelin doğruluk performans metriği, işlem adeti ve işlem süreleri dikkate alınarak seçilmesi anlık ihtiyaç duyulacak bir durum olmayıp sayaç okuma işleminin dönemsel yapılıyor olması avantajı ile ilgili ay içinde analizin yapılmasının yeterli olacağı, dolayısıyla modellerin hız performanslarından çok, veri kalitesi ve model verimliliğinin öncelenebileceği ifade edilebilir. Sahada sayaç okuması esnasında anlık sınıflandırma yapılması daha önce geliştirilmiş modellerin kullanımıyla sağlanacaktır.

Tatil günleri, pandemi gibi etkenlerin sebep olduğu olağanüstü yaşam koşulları tüketimi dönemsel etkileyip davranış modellerini yanıltabileceği öngörülmektedir. Bu husus başka bir çalışmanın konusu olarak da değerlendirilebilir.

Çalışma sonucunda içmesuyu abonelerine ait anormal tüketim dönem sayısının toplam tüketim dönem sayısındaki oranının yaklaşık %2 olduğu düşünülürse verisetinin dengesiz bir yapıda olduğundan bahsedilebilir. Çalışmanın devamı ve tamamlayıcısı olarak tüketim sınıfı tahmin modellerinin geliştirilmesi düşünüldüğünde, elde edilen gözetimli verisetlerinin KNN algoritması temelli SMOTE gibi veri çoğaltma algoritmaları veya normal tüketimlerin anormal tüketim değerleri ile değiştirilmesi yöntemleri ile dengeli hale getirilmesinin tahmin modellerinin performanslarını arttıracığı öngörülmektedir. Çalışmada içme suyu tüketim sınıfı tahmin model performansları yeterli olduğu için buna gerek duyulmamıştır. İhtiyaç duyulması halinde normal tüketim değerleri yerine kullanılacak anormal tüketim değerleri, yaz ayları tüketiminin kış aylarında veya kış ayları tüketimlerinin yaz aylarında yapılmış gibi gösterilmesiyle de elde edilebilir.

Çalışma sonunda elde edilen gözetimli veriseti, tahmin modelleri geliştirilmeden önce tekrardan öznelik seçim işlemlerine tabi tutulabileceği, yeterli sayıda faktörün seçilebileceği, analiz hızının, verimin ve performans metrik değerlerinin arttırılabileceği öngörülmektedir.

İçme suyu abone sayaçlarının, ölçüm hassasiyetlerinden dolayı en geç 10 yılda bir ya kalibrasyon ayarlarının yapılması ya da değişiminin yapılması gerekmekte olup, çalışma ile yaklaşık 16 yıldır tek sayaç kullanan 118 abonenin sayaç değişimlerinin yapılmasının, benzer olarak sayaç değişim kayıtlarına ulaşılabilen diğer bütün abonelerin endeks kontrolü ile 10 yıllık sayaç denetimlerinin yapılmasının daha doğru tüketim ölçüm değerlerini sağlayacağı öngörülmektedir.

Abonelere sayaç okuma esnasında verilen su bildirim makbuzlarıyla birlikte tüketim alışkanlığının normal veya anormal seyirde olduğu, anormal ise hangi sınıf bir tüketim yaptıkları bilgisinin verilerek su tüketimi duyarlılığının artırılacağı öngörülmektedir. Ayrıca hane içi su tüketimlerinin genelde hangi amaçlarla yapıldığı, nerelerde yapıldığı, ne kadar yapılması gerektiği ve nerelerde bir tasarruf sağlanabileceği bilgilendirmelerinin veya önerilerinin kurum ve aboneler arasında var olan bütün iletişim kanalları aracılığıyla mesajlar, broşürler kısa filmler vb. yöntemler kullanılarak yapılmasının da duyarlılığa katkı sağlayacağı öngörülmektedir.

Tespit edilen anormal tüketimlerde vatandaşın bilgilendirilmesi ve duyarlı hale getirilmesi ile tüketimlerin normal tüketim sınıflarına çekildiği düşünüldüğünde öncelikle bu durumun abone sözleşme sahibinin ekonomisine katkı sağlamanın yanında su idarelerinin su talebini karşılamak için alternatif kaynakların sisteme getirilmesi yatırımları ve arayışlarına destekleyici bir çözüm olacağı düşünülmektedir.

Bu çalışmada sadece konutlardaki içmesuyu tüketimi dikkate alınmış ve modeller geliştirilmiştir. Bununla birlikte önerilen yöntem toplu tüketimlerin söz konusu olduğu diğer bütün sektörlerde de kullanılabilir bir altlığa sahiptir. Aynı yöntem ile elektrik, doğalgaz, internet, mobil hatlar, kredi kartları gibi tüketimlerin bir abonelik sistemi üzerinden dönemsel yapıldığı ve kayıt altına alındığı alanlarda her bir abone için tüketim davranış modeli çıkarılabilir, anormal tüketimler tespit edilip belli sınırlarda tüketimlerin yapılması veya tüketimin disipline edilmesi sağlanabilir, tüketici kitlesinde geri bildirimler yapılarak konuyla ilgili duyarlılıklar geliştirilebilir. Ayrıca bu sektörlerde hizmet veren kurumlar yapay AI bu sistemi artı bir hizmet olarak kendi mobil uygulamalarında abonelerine sunarak daha fazla müşteri memnuniyeti sağlamış olacaklardır.

## **Teşekkür (Acknowledgment)**

Abone verileri konusunda kurum desteği ve katkısını esirgemeyen Kayseri Su Kanalizasyon İdaresi'ne ve Genel Müdür Doç. Dr. Özgür ÖZDEMİR'e ayrıca su yönetimi ile ilgili birikimlerinden istifade ettiğim her konuda bilgi ve desteğini esirgemeyen Prof. Dr. Mahmut FIRAT'a teşekkür ederim.

## **Kaynaklar (References)**

- Arup and Sydney Water, 2015. The Future of Urban Water: Scenarios for Urban Water Utilities in 2040 [Internet]. Arup, Available from: <https://www.arup.com/perspectives/publications/research/section/the-future-of-urban-water>
- Billings RB., Jones CV., 2008. Forecasting Urban Water Demand (Second Edition) [Internet]. Second. Kitap- American Water Works Association, 1–367 p, Available from: [www.awwa.org](http://www.awwa.org)
- Cai J, Luo J, Wang S, Yang S., 2018. Feature Selection in Machine Learning: A New Perspective. *Neurocomputing*. 300, 70–9.
- Campbell HE, Johnson RM., Larson EH., 2004. Prices, devices, people, or rules: The relative effectiveness of policy instruments in water conservation. *Review of Policy Research*, 21(5), 637–62.
- Cini J, Mung A, Waughray D., 2014. Global Agenda Council on Water Security 2012-2014 [Internet]. World Economic Forum. 2014, Available from: <http://www.weforum.org/content/global-agenda-council-water-security-2012-2014>

- International Water Association. 2017, The IWA principles for water wise cities - for urban stakeholders to develop a shared vision and act towards sustainable urban water in resilient and liveable cities. Urban Stakeholders to Develop a Shared Vision and Act towards Sustainable Urban Water in Resilient and Liveable Cities [Internet], 2nd, 1–6, Available from: [https://iwa-network.org/wp-content/uploads/2016/10/IWA\\_Brochure\\_Water\\_Wise\\_Communities\\_SCREEN-1.pdf](https://iwa-network.org/wp-content/uploads/2016/10/IWA_Brochure_Water_Wise_Communities_SCREEN-1.pdf)
- Mayer P, Dziegielewski B, Kiefer JC, Lantz GL, Opitz EM, Porter G. Et al., 2000. Commercial and Institutional End Uses of Water. Power, 2014 p.
- Mayer PW, DeOreo WB, Opitz EM, Kiefer JC, Davis WY, Dziegielewski B, et al., 1999. Residential End Uses of Water Executive Summary. AWWA Research Foundation and American Water Works Association, 1–345 p.
- Türkiye Cumhuriyeti Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2020. Sürdürülebilir Kalkınma Amaçları ve Göstergeleri. In: MÜDÜRLÜĞÜ YHG, BAŞKANLIĞI BVBYD, editors. Strateji ve Bütçe Başkanlığı Yayınları [Internet]. Haziran 20. Ankara: STRATEJİ VE BÜTÇE BAŞKANLIĞI, p. 1–42, Available from: <http://www.surdurulebilirkalkinma.gov.tr/wp-content/uploads/2021/02/SKA-ve-Gostergeleri-Kapak-Birlestirilmis.pdf>
- Water Conflict, 2022. Water Conflict Chronology [Internet]. Vol. 1, Pacific Institute 2022, p. 1–52, Available from: <https://www.worldwater.org/conflict/list/>
- Yıldız D, Özgüler H., 2020. Yapay Zekâ ve Su Yönetimi. Vol. 30, Rapor. Ankara.
- Zhang R, Nie F, Li X, Wei X., 2019. Feature Selection with Multi-view Data: A Survey. Information Fusion [Internet], 50(March 2019), 158–67, Available from: <https://doi.org/10.1016/j.inffus.2018.11.019>