



## Prediction of heat transfer value in the automotive industry with an approach based on internet of things and machine learning

Makbule Nalkıran<sup>1\*</sup>, Serkan Altuntaş<sup>2</sup>

<sup>1</sup>Tofaş Turkish Automotive Factory I.C., 16000, Osmangazi, Bursa, Türkiye

<sup>2</sup>Department of Industrial Engineering, Faculty of Machine, Yıldız Technical University, 06680, Besiktas, Istanbul, Türkiye

### Highlights:

- Predicting the heat sent to buildings by machine learning.
- Integration of internet of things and machine learning technologies
- A real application for the automotive industry

### Keywords:

- Heat Prediction
- Machine Learning
- Internet of Things

### Article Info:

Research Article

Received: 25.01.2024

Accepted: 04.08.2024

### DOI:

10.17341/gazimmfd.1406869

### Acknowledgement:

This work was prepared within the scope of 2244 TÜBİTAK Industry-Doctorate Project. Number: 119C064.

### Correspondence:

Author: Makbule Nalkıran  
e-mail: makbule.nalkiran@tofas.com.tr  
phone: +90 553 649 3759

### Graphical/Tabular Abstract

this study, an IoT and machine learning based approach is proposed, which includes six steps to estimate the heat generated in buildings. The steps of the proposed approach are shown in Figure A.

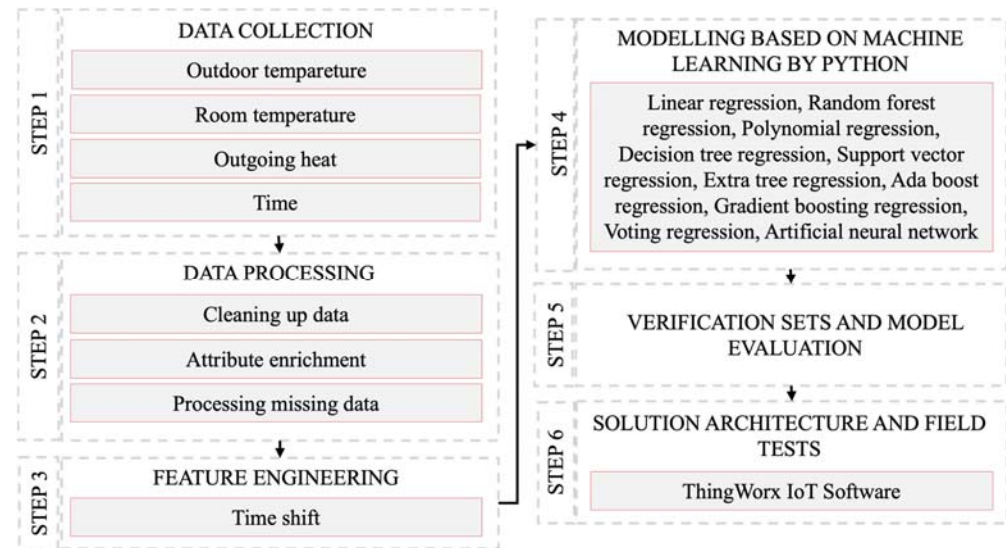


Figure A. Flow chart of the study

### Purpose:

The aim of this study is to predict the heat that will be sent to the buildings from the heating center in enterprises or facilities with many buildings and a single heating center.

### Theory and Methods:

Machine learning-based regression models were developed with an expanded data set by generating new variables from existing temperature data to predict the heat required for the selected pilot plant of a factory in the automotive industry.

### Results:

Among the nine different machine learning algorithms evaluated, the Linear Regression algorithm with the highest prediction accuracy was selected.

### Conclusion:

Temperature regulation was made with the developed model. Costs have been reduced thanks to the effects of many negative factors such as heat losses in the heating process of the facility, changes in outdoor conditions, overheating or cooling of the environment, loss of effect of the sent heat after a while, and the prevention of heat losses.



## Otomotiv sektöründe nesnelerin interneti ve makine öğrenmesine dayalı bir yaklaşımla ısı transfer değerinin tahmini

Makbule Nalkıran<sup>1\*</sup>, Serkan Altuntaş<sup>2</sup>

<sup>1</sup>Tofaş Türk Otomobil Fabrikası A.Ş., 16000, Osmangazi, Bursa, Türkiye

<sup>2</sup>Yıldız Teknik Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, 06680, Beşiktaş, İstanbul, Türkiye

### ÖNEÇIKANLAR

- Makine öğrenmesi ile binalara gönderilen ısınmın önceden tahmin edilmesi
- Nesnelerin interneti ve makine öğrenimi teknolojilerinin entegrasyonu
- Otomotiv sektörü için gerçek bir uygulama

### Makale Bilgileri

Araştırma Makalesi

Geliş: 25.01.2024

Kabul: 04.08.2024

DOI:

10.17341/gazimmfd.1406869

### Anahtar Kelimeleri:

Isı tahmini,  
makine öğrenmesi,  
nesnelerin interneti

### ÖZ

Enerji tüketiminin hâkim olduğu gelişen dünyada yapay zekâ çözümleriyle enerji tahmini, verimliliği ve maliyet etkinliği önem kazanmıştır. İşletmelerde harcanan enerji bu tüketimin önemli bir kısmıdır. Nesnelerin interneti (IoT) ve makine öğrenmesiyle bir otomobil fabrikası için ısı transfer değerleri tahminlenerek bu çalışmanın önemi ortaya koyulmuştur. Çalışma, IoT ve makine öğrenimi modellerinin bir kombinasyonuyla gerçek bir endüstriyel sisteme uygulanmıştır. Bu sistem üzerinde çalışan ve iki farklı endüstri 4.0 konusunun bütünlük kullanılması, çalışmanın özgün yanıdır. Fabrikada seçilen pilot tesis ısınmın tahmini için mevcut sıcaklık verilerinden yeni değişkenler üreterek, genişletilmiş veri seti ile makine öğrenmesi tabanlı regresyon modelleri geliştirilmiştir. Geliştirilen modelle sıcaklık regülasyonu sağlanarak, tesisin ısı kayıpları, dış ortam koşullarındaki değişiklikler, ortamın aşırı ısınması veya soğuması, ısınmın transfer sırasındaki kayıpları gibi olumsuz faktörler önlenerek maliyetler azaltılmıştır. Lineer, Rastgele Orman, Polinom, Karar Ağacı, Destek Vektör, Ekstra Ağaç, Adaboost, Gradient Boosting, Oylama Regresyon modelleri ve Yapay Sinir Ağı algoritmaları kullanılarak pilot tesise gönderilecek ısı tahminlenmiştir. Bu algoritmalarından Doğrusal Regresyon modeli en yüksek tahminleme değerine sahiptir. Son olarak çalışma, işletmenin canlı takip sistemi olan SCADA (Supervisory Control and Data Acquisition)'ya entegre edilerek gerçek zamanlı test edilmiştir. Çalışmanın sonuçlarında, önerilen yaklaşımla ısıtma sisteminden kaynaklanan aşırı ısınma maliyetlerinin, pilot tesiste yıllık % 90 azaldığı ve çalışan memnuniyetinin arttığı gözlemlenmiştir.

## Prediction of heat transfer value in the automotive industry with an approach based on internet of things and machine learning

### HIGHLIGHTS

- Predicting the heat sent to buildings by machine learning
- Integration of internet of things and machine learning technologies
- A real application for the automotive industry

### Article Info

Research Article

Received: 25.01.2024

Accepted: 04.08.2024

DOI:

10.17341/gazimmfd.1406869

### Keywords:

Heat prediction,  
machine learning,  
internet of things

### ABSTRACT

In the developing world where energy consumption is dominant, energy estimation, efficiency and cost effectiveness have gained importance with artificial intelligence solutions. The energy consumption in businesses is a significant part of this consumption. The importance of this study was revealed by estimating heat transfer values for an automobile factory with the Internet of Things (IoT) and machine learning. The study was applied to a real industrial system with a combination of IoT and machine learning models. The integrated use of two different Industry 4.0 topics working on this system is the original aspect of the study. Machine learning-based regression models were developed with an expanded data set by generating new variables from the existing temperature data for the prediction of the selected pilot plant temperature in the factory. Temperature regulation was provided with the developed model, and costs were reduced by preventing negative factors such as heat losses of the plant, changes in external environmental conditions, overheating or cooling of the environment, and losses of heat during transfer. The heat to be sent to the pilot plant was predicted using Linear, Random Forest, Polynomial, Decision Tree, Support Vector, Extra Tree, Adaboost, Gradient Boosting, Voting Regression models and Artificial Neural Network algorithms. Of these algorithms, the Linear Regression model has the highest prediction value. Finally, the study was integrated into the business's live monitoring system, SCADA (Supervisory Control and Data Acquisition), and tested in real time. In the results of the study, it was observed that with the proposed approach, overheating costs caused by the heating system decreased by 90% yearly in the pilot plant and employee satisfaction increased.

## 1. Giriş (Introduction)

Günümüzde, insanlığın tüketim eğilimindeki yaygın büyüme, en büyük etkisini üretim sektörlerinde göstermektedir. İşletmelerin bu artışa cevap vermesi ve pazarda rekabet edebilmeleri için firmaların Endüstri 4.0 ile imalat sektörlerindeki teknolojik gelişmeleri takip etmeleri ve ilgili teknolojileri kendi bünyelerinde kullanabilmeleri zorunlu hale gelmiştir. Bu teknolojiler aynı zamanda birçok üretim alanında enerjiyi daha verimli kullanmak ve maliyet tasarrufu sağlamak için kullanılabilir olacak yenilikçi araçlardır. Endüstrilerdeki tüm üretim tesisleri, enerji tüketim faaliyetlerini en verimli ve minimum maliyetli şekilde planlamaya çalışmaktadır. Bunu başarmada rol oynayan birçok faktör olabileceğinden, bu hedefe her zaman ulaşamamaktadır. Bu nedenle işletmeler, üretim sistemlerinde gelecekte daha kesin ve daha iyi hesaplanmış eylemlerde bulunmak için enerji tüketimlerini doğru bir şekilde tahmin etmek durumundadır.

Endüstri 4.0'ın temel teknolojilerinden olan nesnelerin interneti ve makine öğrenimi, çeşitli üretim alanlarında yaygın olarak kullanılmaktadır. Nesnelerin interneti teknolojileri ile üretim alanlarındaki veriler sensörler ile birleştirilmekte, bu sayede canlı olarak izlenebilmekte ve kontrol mekanizmaları kurulabilmektedir. Bu veri toplama ve izleme kolaylığı sayesinde, verilerden çıkarım yapma ve değer üretme çabaları son zamanlarda hem endüstride hem de akademiye son derece yaygın hale gelmiştir. Makine öğreniminin mahareti ve yukarıda belirtilen tahmine dayalı çalışmalar sayesinde işletmeler, gelecekteki eylemlerini daha kolay planlayarak başarılı olabilmektedirler.

Türkiye'de enerji tüketiminde, içerisinde fabrika binalarının tükettikleri enerji miktarı ile birlikte en yüksek payı % 33,8 oranında sanayi sektörü, ardından % 31'lik bir pay ile mesken ve hizmetler sektörü almıştır [1]. Buna ek olarak sanayi sektörü kapsamında değerlendirilen fabrika binalarında iç harcamaların en büyük kalemlerinden birisi enerji giderleridir ve bu giderlerin arasında ısıtma sistemleri önemli bir yer tutar [2]. Dolayısıyla fabrika binalarında enerji tüketimini etkileyen en önemli faktörlerden biri ısıtma sistemleridir. Sağlıklı endüstriyel çalışma koşullarının sağlanması yasal olarak (6331 İSG Kanunu İşyeri Bina ve Eklentilerinde Alınacak Sağlık ve Güvenlik Önlemlerine İlişkin Yönetmelik) zorunludur ve fabrika binalarındaki ısı seviyeleri her zaman bu koşullara uygun olmayabilir. Isıtma sistemleri ile ilgili bu tür sorun ve rahatsızlıklara yol açan değişkenler çok çeşitlidir. Bu değişkenleri belirleyip değişkenlerle ilgili geçmiş veriler incelenerek, ısı seviyeleri kontrol edilebilir ve gelecek ısı davranışları tahmin edilebilir. Bu aşamada söz konusu verilerden değer üretme (yani veri toplama ve yorumlama teknikleri) devreye girmektedir. Doğru veriler, uygun zamanda ve uygun şekilde toplandığında, uygun algoritmalar kullanıldığında daha hassas bir sıcaklık ve gönderilecek ısı tahmini yapmak mümkün olmaktadır. Bu tür tahminleme çalışmaları (başka bir deyişle, makine öğrenimi (ML)), bilgisayar bilimi ve istatistik disiplinlerini birleştirerek bilgisayarların programlama olmadan belirli bir görevi nasıl yapacaklarını öğrenmelerini sağlamaktadırlar [3]. Bu çalışmada, otomotiv sektöründe faaliyet gösteren lider bir firmanın farklı üretim tesislerinde (boya, montaj, gövde, süspansiyon ve pres) ihtiyaç duyulan ısıyı nesnelerin interneti ve makine öğrenimi yöntemleri kullanılarak önceden tahmin edilmesi ile mevcut durumda binalardaki aşırı ısınmanın önüne geçilmesi, maliyet kayıplarının ve çalışanların aşırı ısıya maruz kalma durumlarının önlenmesi amaçlanmıştır. Bu amaç doğrultusunda gerçekleştirilen çalışma sonuçları etraflıca tartışılmıştır ve devreye alma faaliyetleri öncesinde işletmenin seçilen bir pilot tesisinde (gövde) test çalışmaları yapılmıştır. Bu çalışmanın itici gücü iki yönlüdür. İlk olarak, kullanılan makine öğrenimi regresyon modellerinin tahmin

performansı; fabrika tesisleri ihtiyacı doğrultusunda belirlenen hedef sıcaklık, IoT sensörleri sayesinde elde edilen oda sıcaklıkları verileri ve web servisleri ile elde edilen dış ortam sıcaklıkları temelinde belirlenmiştir. İkinci olarak, bir otomotiv firmasının seçilen pilot tesisinden nesnelerin interneti teknolojisi kullanılarak 1 ay boyunca birer dakikalık aralıklarla toplanan verilerle makine öğrenimi teknikleri uygulanmıştır. Son olarak ise makine öğrenimi metodlarını fabrikanın endüstriyel otomasyon sistemine entegre ederek seçilen pilot tesise gönderilecek ısı değeri tahmin edilmiştir.

Çalışmanın takip eden bölümleri şu şekilde özetlenebilir; İkinci bölümde, ısı tahmini ve makine öğrenmesi çalışmaları ile ilgili literatür özetlenmiştir. Üçüncü bölümde; makine öğrenimi, regresyon modelleri ve regresyon modellerinin performans metrikleri ile çalışmanın iş akışlarının açıklayıcı anlatımları materyal ve yöntem kapsamında yer almaktadır. Dördüncü bölümde; çalışmada makine öğrenimi ve nesnelerin interneti teknolojilerinden faydalanılan gerçek bir uygulama gerçekleştirilmiştir. Son bölümde ise araştırmanın veri analizi ve değerlendirmesini kapsayan sonuç ve tartışma bölümüne yer verilmiştir.

## 2. Literatür Taraması (Literature Review)

Literatürde makine öğrenmesi ile tahmine dayalı çalışmalar son yıllarda yaygın bir tartışma konusu olmuştur. Isı tahminleme ile ilgili çok sayıda makine öğrenimi modeli de bu çalışmalara konu olmuş ve uygulama çalışmaları yapılmıştır. Bu bölümde özellikle son beş yıldaki çalışmalar kronolojik olarak değerlendirilmiştir.

Büyük bir konut binası ile ilgili veri setlerini kullanarak yıllık bina enerji tüketimini tahmin etmek isteyen Olu-Ajayı vd. [4] Yapay Sinir Ağı (ANN), Gradient Boost (GB), Derin Sinir Ağı (DNN), Rastgele Orman (RF), Yığınlama ve K En Yakın Komşu (KNN) gibi çeşitli makine öğrenimi tekniklerini kullanmışlardır ve çalışmaları ile aynı zamanda veri setlerinden küme oluşturma yönteminin model performansı üzerindeki etkisini de araştırmışlardır. Pavelcik ve Cuba [5] makine öğrenimi algoritmalarını kullanarak fren disklerinin sıcaklık tahminini gerçekleştirmiştir. Bhowmik vd. [6] ise makine öğrenme yöntemlerini uygulayarak oda sıcaklığında sabit basınçta polimerlerin özgül ısıyı tahmin etmeye yönelik bir çalışma yürütmüşlerdir ve burada karar ağaçları regresyon algoritmalarına ve temel bileşen analiz yöntemlerine başvurmuşlardır. Bir diğer çalışmada Liu vd. [7] yapay sinir ağları algoritması ve makine öğrenimine dayalı parçacık sürüsü optimizasyonu kullanarak polimerik nanokompozitlerin termal iletkenliğini tahmin etmek için hibrit bir makine öğrenimi yöntemi sunmuşlardır. Swartz vd. [8] mikro yapılarla değiştirilmiş belirli bir yüzey için kritik ısı akışlarını, analiz etmek ve daha iyi tahmin etmek için rastgele orman regresyon yöntemini esas alan yeni bir analiz algoritması türeterek, kritik ısı akışını tahmin eden en önemli parametrelerin verimli bir şekilde keşfedilmesini kolaylaştırmışlardır.

Liu vd. [9] motor performansını ve emisyonları optimize etmek için harcanan ısı enerjisini t yapay sinir ağı, rastgele orman, destek vektörü regresyonu ve gradyan artırma regresyon ağaçları olmak üzere dört farklı makine öğrenimi algoritmasını tahmin etmişlerdir. He ve Lee [10] sayrek dağılmış deneysel veri noktalarına ve kritik ısı akışına dayalı verilerle v-Destek Vektör Makinesi algoritması üzerinde çalışmışlardır.

Sajjadi vd. [11] çalışmalarında bölgesel ısıtma sistemine bağlı tüketicinin ısı yükünü, Aşırı Makine Öğrenmesi (ELM) algoritmasını kısa vadeli ve çok adımlı bir tahmin modeli olarak tasarlayarak 1 ila 24 saat arasında değişen süreler için dokuz farklı ELM tahmin modeli elde etmişlerdir ve sonuçları genetik programlama (GP) ve yapay sinir

ağı (YSA) modelleri ile karşılaştırarak elde ettikleri deneysel sonuçlar ile ELM yaklaşımının tahmin doğruluğunda ve genelleme potansiyelinde iyileşme sağlanabileceğini göstermişlerdir.

Literatürde ısı tahminleme ile ilgili benzer birçok çalışma bulunmaktadır. Özellikle son yıllarda yapılmış çalışmalar Tablo 1'de sunulmaktadır.

Literatürde ısı tahminleme dışında pek çok sayıda makine öğrenimi temelli benzer tahmin çalışmaları sunulmaktadır. Lu vd. [24] konut günlük elektrik tüketim verilerini kullanarak gelecekteki elektrik yükünü tahmin etmek için bir Seyrek Grup Kement-Kuantile Regresyon Derin Sinir Ağı (SGLQRDNN) modeli önermişlerdir. Yüce ve Kabak [25] bir üretim tesisine ait detay üretim alanlarında 4 farklı prosese ait toplam 8 iş merkezine üretim süresi tahminleme için makine öğrenmesi algoritmalarından yapay sinir ağı, destek vektör regresyonu ve gradyan artırma algoritmaları uygulanmış ve her iş merkezi için en iyi sonucu veren algoritma belirleyerek elde edilen sonuçlara göre destek vektör regresyonu ve gradyan artırma algoritmalarının üretim süresi tahminlemede başarılı sonuçlar verdiği görülmüştür.

Fan vd. [26] büyük bir su kaynağı ağının geçmiş bakım verilerini kullanarak su borusu arızalarını tahmin etmek için Light GBM, Neural Network, Logistic Regresyon, K-En Yakın Komşular ve Destek Vektörü Sınıflandırması dahil olmak üzere beş farklı makine öğrenimi modellerini karşılaştırarak elde ettiği Light GBM'nin en iyi performansı bulmuştur. Chen vd. [27] iki rüzgâr türbinindeki arızaların teşhisi için SCADA verileri ile transfer öğrenme algoritmaları Inception V3 ve TradaBoost'u kullanan bir çerçeve sunmuş ve bu çerçeveyi doğrulamışlardır. Xu vd. [28] çalışmalarında tuzlu su varlığında metan hidrat oluşum sıcaklığının tahmini için Çoklu Doğrusal Regresyon, k-En Yakın Komşu, Destek Vektörü Regresyonu, Rastgele Orman ve Gradyan Artan Regresyon olmak üzere beş makine öğrenme algoritması uygulamışlar ve bu makine öğrenimi algoritmalarını karşılaştırdıktan sonra bu beş makine öğrenimi algoritması arasında Gradient Boosting'in en iyi tahmini verdiğini gözlemlemişlerdir. Shapi vd. [29] Microsoft Azure bulut tabanlı ML platformunda enerji tüketim tahmini için Destek Vektör Makinesi, Yapay Sinir Ağı ve k-En Yakın Komşu olmak üzere üç farklı metodoloji önererek tahmine dayalı bir model geliştirmişlerdir. Reşat [30] enerji yönetim sistemlerinde önemli bir yere sahip olan kısa vadeli enerji tahmin süreçleri için ARIMA ve yapay sinir ağlarını kullanarak geliştirdiği melez tahmin modeli sayesinde, sadece ARIMA modeli ile yapılan tahmin verilerine göre %39,9'luk bir iyileşme sağlamıştır. Zhu vd. [31] günlük enerji tüketimi profillerini ölçümü amacıyla akıllı ölçüm verilerinin ve tarihsel verilerin kullanıldığı istatistiklerin tahmin modellerini oluşturmuş ve anormal

durumları tespit etmek için tahmin modellerinin kalıntılarını analiz eden istatistiksel kontrol çizelgelerini kullanmıştır. Başka bir çalışmada, Wang vd. [32], EVI (Economized Steam Injection) ile R134a ısı pompasının performansını tahmin etmek için SVR (Support Vector Regression) ve Adaboost.R2 algoritmalarını kullanarak bir yöntem önermişlerdir. Köse vd. [33] Yenilenebilir bir enerji olan rüzgâr enerjisinin potansiyel hesabı ve rüzgâr karakterizasyonunda Weibull olasılık dağılım fonksiyonu (Wodf) ile klasik yöntemlerin hız dağılımlarındaki uyum zayıflığını gidermek için, dağılım parametrelerini tahmin etmek üzere Yusufçuk Algoritması (YA) önerilmişlerdir ve çalışmalarının sonucunda YA yönteminin determinasyon katsayısı (R<sup>2</sup>) ölçütüne göre en iyi performansı verdiğini gözlemlemişlerdir.

Yapılan literatür araştırması gereği, nesnelere interneti ve birden fazla makine öğrenmesi teknikleri birlikte kullanılarak anlık veriler ile anlık ya da belirlenen süreler içerisinde tahmin sonuçları üretilen çalışmalara az rastlanmıştır. Bilgilerimize göre bu iki farklı Endüstri 4.0 teknolojisini kullanarak doğrudan sanayi tesislerini ele alan ve bu tesislerdeki üretim binalarının ısınması ile ilgili bir tahminleme çalışmasına rastlanmamıştır. Makine öğrenimi algoritmaları ve nesnelere interneti teknolojisi ile gerçek zamanlı tahminler elde edip, tahmin sonuçlarının eş zamanlı olarak fabrika binalarında canlı bir şekilde kullanılması ve ısıtma sistemlerindeki iyileştirici sonuçların canlı bir sistemde görülmesi ise çalışmanın en güçlü yanlarından birisidir. Bu fark sayesinde bir nesnelere interneti platformu kullanılarak burada makine öğrenimi algoritmaları çalıştırılabilir ve gerçek zamanlı veri ile çalışılabilir. Yine aynı zamanda bu durum çalışma sonuçlarının da çalışanlar tarafından gerçek zamanlı olarak kullanılmasını sağlamıştır. Bu kapsamda da geliştirilen yöntemin, gerçek bir endüstriyel sistemde uygulanması ve doğrulanması ile literatüre katkıda bulunulduğu düşünülmüştür.

### 3. Malzeme ve Yöntem (Material and Method)

#### 3.1. Makine Öğrenmesi (Machine Learning)

Makine öğrenimi, özellikle denetimli, yarı denetimli ve denetimsiz olmak üzere üç gruba ayrılabilir. Denetimli öğrenme, bir dizi girdi değişkenine dayalı olarak çıktı değerini tahmin etmektir ve denetimsiz öğrenme ise bir çıktı değeri olmaksızın bir dizi girdi değişkeni arasındaki ilişkilerin ve örüntülerin tanımlanmasıdır [34]. Yarı denetimli öğrenme ise öğrenme sistemine her girdi seti için üretilmesi gereken çıktı setini sunmaz. Bunun yerine sistem, kendisine gösterilen girdilere yanıt olarak çıktı üretir ve üretilen çıktının doğru ya da yanlış olduğuna dair bir sinyal üretilir. Sistem üretilen sinyali dikkate alarak öğrenme işlemini gerçekleştirir [35]. Her şeyden önce, sistemi iyi kavramak için tam bir kavrayış geliştirmek zorunludur. Sistemin kapsamlı bir şekilde anlaşılması, makine öğrenimine iyi bir yaklaşıma

**Tablo 1.** Makine öğrenimi kullanılarak yapılan ısı tahminleme ile ilgili literatürün özeti  
(Summary of the literature on heat forecasting using machine learning)

Yazar [Yıl]	Yöntem
Zhang vd. [12]	Tahminleme, Kümeleme, Optimizasyon, Talep Yönetimi
Sundar vd. [13]	Tahminleme, Derin Öğrenme
Dai vd. [14]	Aşırı Makine Öğrenimi (ELM), Tek Katmanlı İleri Beslemeli Sinir Ağı (SLFN)
Lin vd. [15]	Hibrit Uzay-Zamansal Dikkat Uzun Kısa Süreli Bellek (STALSTM)
Song vd. [16]	Tabanlı Isı Yükü Tahmin Modeli, LTSM Modeli
Yuan vd. [17]	Evrışimli Sinir Ağı (CNN), Geçici Evrışimli Sinir Ağı (TCN), Tekrarlayan Sinir Ağı (RNN)
Hughes vd. [18]	Sekizinci Kesirli Yargı Temelli Ortalama Yöntemi, Destek Vektör Makinesi (SVM)
Vu vd. [19]	Makine Öğrenimi Regresyon Modelleri (Destek Vektörü, Rastgele Orman, Yapay Sinir Ağları)
Zhu vd. [20]	Makine Öğrenimi Algoritmaları
Park vd. [21]	Makine Öğrenimi, Yapay Sinir Ağları
Kim vd. [22]	Çoklu Regresyon Modelleri, Makine Öğrenimi Algoritmaları
Wu vd. [23]	Makine Öğrenimi Algoritmaları
	Makine Öğrenimi Algoritmaları, Yapay Sinir Ağları

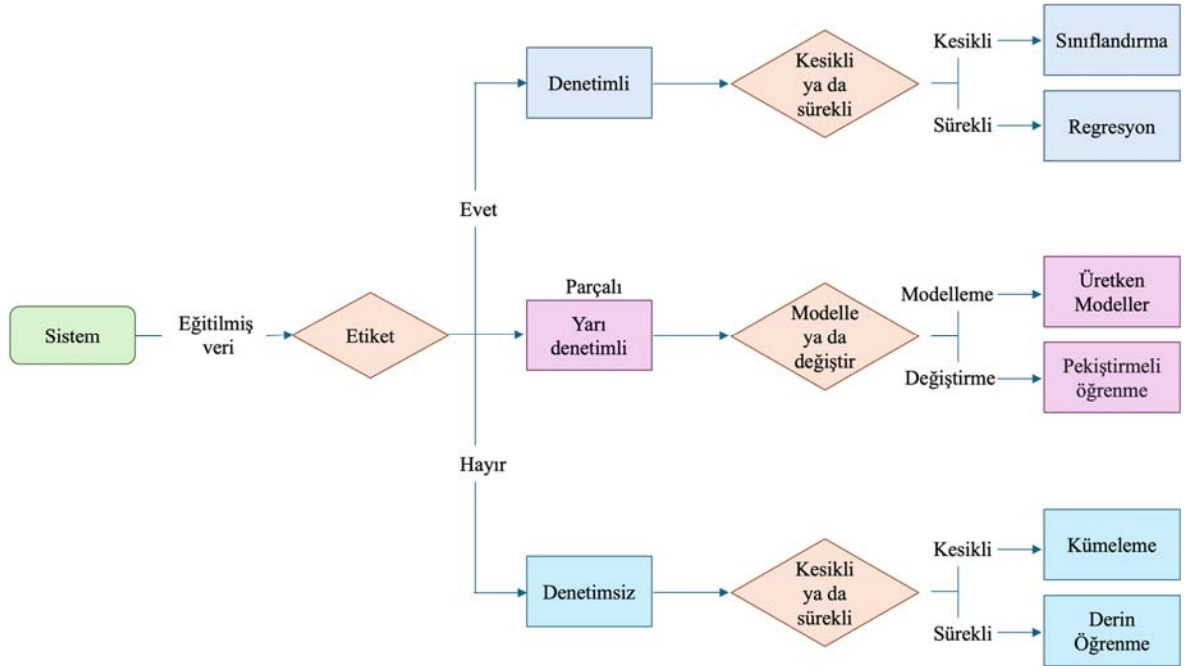
sahip olmak için belirleyicidir. Sistemi tanıdıktan sonra, makine öğrenmesi için kullanılacak veriler incelenir ve verinin türüne bağlı olarak ayrıklık mı yoksa süreklilik mi soruları ele alınır. Veri analizi ve yorumlaması için sınıflandırma veya regresyon algoritmaları seçilebilir. Çalışılma yapılacak girdi verilerine dayalı makine öğrenimi türleri Şekil 1’de gösterilmiştir.

Makine öğrenmesi tabanlı ısı tahmini sistemi için kullanılacak algoritmaların seçimi önemlidir. Burada kullanılacak veri setinin özelliklerine göre algoritmalar değişkenlik göstermektedir. Aynı zamanda bir veri setine birden fazla algoritma da uygulanabilmektedir. Her bir veriye uyum sağlayan tek bir makine öğrenmesi algoritması da yoktur. Her bir algoritmanın her bir veri seti ve problem karşısında birbirlerine karşı üstün ve zayıf yönleri vardır [37]. Bu yüzden çalışma yaparken birden fazla makine öğrenimi algoritması seçilerek problemlere uygulanması, çıktılarına göre karşılaştırma yaparak doğru algoritmayı seçmede çok etkilidir.

Bu çalışmada kullanılan veri setinde bağımsız değişkenlere bağlı bir bağımlı değişkenin tahmin edilmesi istenildiğinden dolayı aşağıda açıklanan makine öğrenmesi regresyon modelleri kullanılmıştır.

- Doğrusal Regresyon: Bağımsız değişkenlere dayalı bir hedef tahmin değeri modellemeyi sağlar [38]. Genellikle bu bağımsız değişkenler ile tahmin edilen değerler arasında bir ilişki kurmak için kullanılır.
- Polinom Regresyon: Veriler arasındaki doğrusal olmayan ilişkileri açıklar. Bu doğrusal olmayan ilişkiler polinomsal olarak modellenir. Benzer bir model, birden fazla değişkeni olan ve verilerin doğrusal olmadığı durumlarda kullanılan doğrusal çoklu regresyon modelidir. Verileri bir polinom özelliğine sığdırmaya çalışır. Yeni model, lineer regresyonda düz bir çizgi yerine polinom çizgileri kullanır.
- Karar Ağacı Regresyonu: Bir karar ağacı regresyon algoritması, nitelikleri ifade eden düğümlerden ve bu düğümlerin bağlantılarını ifade eden dallardan oluşan bir ağaç olarak ifade edilir. Burada her bir düğüm sınıflandırılacak gruplardaki öznitelikleri, her bir dal ise düğümün alabileceği değerleri temsil etmektedir.

- Rastgele Orman Regresyonu: Rastgele orman regresyon algoritması, geniş bir regresyon ağacı kümesini birleştiren bir topluluk öğrenme algoritmasıdır [39]. Bir regresyon ağacı, hiyerarşik olarak düzenlenen ve ağacın kökünden yaprağına sırayla uygulanan bir dizi koşulu veya kısıtlamayı temsil eder [40].
- Ekstra Ağaç Regresyonu: Rastgele orman algoritmasının bir uzantısı olarak geliştirilmiştir ve bir veri kümesine uyma olasılığı daha düşüktür [41]. Ekstra ağaç (ET), rastgele ormanla aynı prensibi kullanır ve her temel tahminciyi eğitmek için özelliklerin rastgele bir alt kümesini kullanırken düğümü bölme için karşılık gelen değere sahip en iyi özelliği rastgele seçer [42].
- Gradient Boosting Regresyon: Gradient boosting (GB) yöntemi, regresyon ve sınıflandırmada sıklıkla kullanılan bir yöntemdir. Bu yöntemde her bir tahmin için tahmin değeri için kullanılan parametre ağırlıklarını değiştirmek yerine önceki tahminlerdeki hata ile karşılaştırılarak sonuca ulaşılmaya çalışılmaktadır [43].
- Voting Regresyon: Çok sayıda modelden oluşan bir grup üzerinde eğitim veren ve çıktı olarak seçilen sınıfın en yüksek olasılığına dayalı olarak bir çıktıyı (sınıfı) tahmin eden bir makine öğrenimi modelidir [44].
- Adaboost Regresyon: Temel tahmin algoritmasının performansını artıran bir büyüme algoritmasıdır. Aynı veri kümesi üzerinde aynı anda regresyon modelinin birkaç kopyasını eğitir ve örneklerin ağırlıkları mevcut tahminin hatasına göre ayarlanır [45].
- Destek Vektör Makinesi: Bir Destek Vektör Makinesi (SVM), regresyon veya sınıflandırma gibi analizler için verilerin kodunu çözebildiğinden denetimli bir öğrenme modeli olarak kabul edilebilir [46]. Yüksek boyutlu bir uzayda birçok hiper düzlem oluşturabilir. Örneğin, regresyon ve sınıflandırma veya aykırı değer açıklama görevleri için kullanılacak G boyutunun alanı; burada G, özelliklerin sayısını temsil eder [47].
- Yapay Sinir Ağları: Yapay sinir ağları, insan beyninin öğrenme stili temel alınarak geliştirilmiş olup, öğrenme sürecinin gerçekleştirildiği yapı insan vücudundaki sinir ağlarına benzediği için bu şekilde isimlendirilmiştir. Belirli karmaşık problemlerin çözümü için, sinir ağlarında olduğu gibi, bilgi akışını ve öğrenmeyi sağlayan düğümler ve bu düğümler arasında bağlantılar vardır. Yapay sinir ağlarında öğrenme süreci örneklerle



Şekil 1. Girdi verilerine dayalı makine öğrenimi türleri (Types of machine learning based on input data) [36].

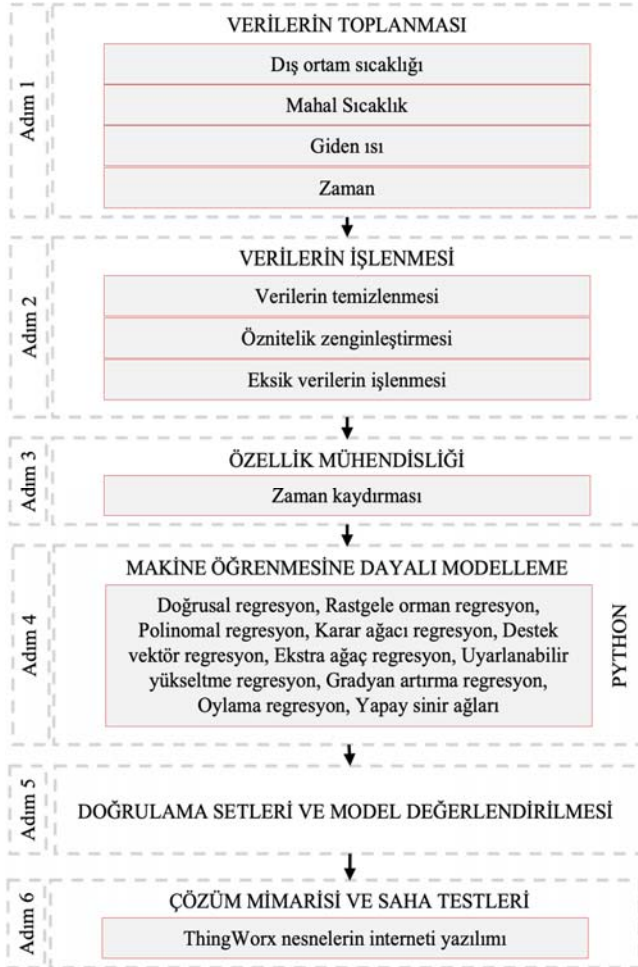


gerçekleştirilmektedir. Yapay sinir ağlarında ağırlıklara verilen değerler ile öğrenme işlemi gerçekleştirilir. Öğrenilen bilgiler tüm ağa yayılmış ağırlıklarda saklıdır [48].

### 3.2. Çözüm Metodolojisi (Solution Methodology)

Bu çalışmada, binalarda üretilen ısıyı tahmin etmek için altı adımı kapsayan nesnelerin interneti ve makine öğrenimine dayalı bir yaklaşım önerilmiştir. Önerilen yaklaşımın adımları Şekil 2'de sunulmuştur.

Burada ilk olarak adım 1'de çalışmanın yapıldığı fabrika binasına ait dış ortam sıcaklığı, mahal sıcaklık, giden ısı ve bu verilerin hangi zamanlarda tutulduğuna dair veriler toplanmıştır. Adım 2'de verilerin işlenmesi kapsamında sırasıyla veri temizliği, mevcut veriler kullanılarak öz nitelik zenginleştirilmesi ve veri setindeki eksik verilerin temizlenmesi çalışmaları yapılmıştır. Adım 3'te veri setinde zaman kaydırması işlemleri yapılmıştır. Adım 4'te python programlama dili kullanılarak makine öğrenmesine dayalı 10 farklı model üzerinde çalışılmıştır. Adım 5'te ise Adım 4'te kullanılan modellerin doğrulanması için doğrulama veri setleri oluşturulmuş ve modellerin değerlendirmeleri yapılmıştır. Son olarak ise Adım 6'da Thingworx IoT yazılımı kullanılarak çalışmanın ve uygulamanın çözüm mimarisi ortaya koyularak çalışmanın gerçek ve çalışan bir sistemde saha testleri yapılmıştır.



Şekil 2. Makine öğrenmesine dayalı çözüm metodolojisi (Solution methodology based on machine learning)

### 4. Deneysel Metot (Experimental Method)

İş sağlığı ve ergonomi açısından çalışanlara uygun çalışma koşullarının sağlanması için oda sıcaklığının belirli bir seviyede tutulması zorunludur. Bina ortamının aşırı ısınması veya soğuması ciddi rahatsızlığa neden olabilir veya ek maliyetlere yol açabilir. Çalışmanın yapıldığı fabrikadaki gibi enerji merkezlerine sahip tesislerde enerji merkezinden diğer binalara gönderilen ve bu binaların ısınma sürecindeki ısı kaybı, dış ortam koşullarındaki değişiklikler ve üretilen ısının etkisi gibi belirsizlikler durumu daha da kötüleştirir ve hedef sıcaklığa ulaşmayı veya mevcut ısıyı korumayı zorlaştırır.

Çalışmanın yapıldığı işletmede ısı, merkezi bir sistemle sağlanmaktadır. Şekil 3'te çalışmanın yapılacağı tesisin merkezi ısıtma sistemi düzeni görsel olarak ifade edilmiştir. Bu merkezi sistem ile kazanlarda su ısıtılmakta ve Şekil 4'deki gibi bir akıştan sonra elde edilen ısı işletmenin enerji merkezinden belirli uzaklıklardaki binalara gönderilmektedir. Isı merkezi ile binalar arasındaki mesafe nedeniyle sıcak su transferinde ısı kaybı artmakta ve ayrıca merkezden transfer edilen sıcak suyun etkisi yaklaşık 30 dakika sonra ortaya çıkmaktadır. Bu durumlar dikkate alınarak 30 dakika sonra hedef sıcaklığa ulaşmak için şu an gönderilmesi gereken ısı miktarının tahmin edilmesi gerekmektedir.

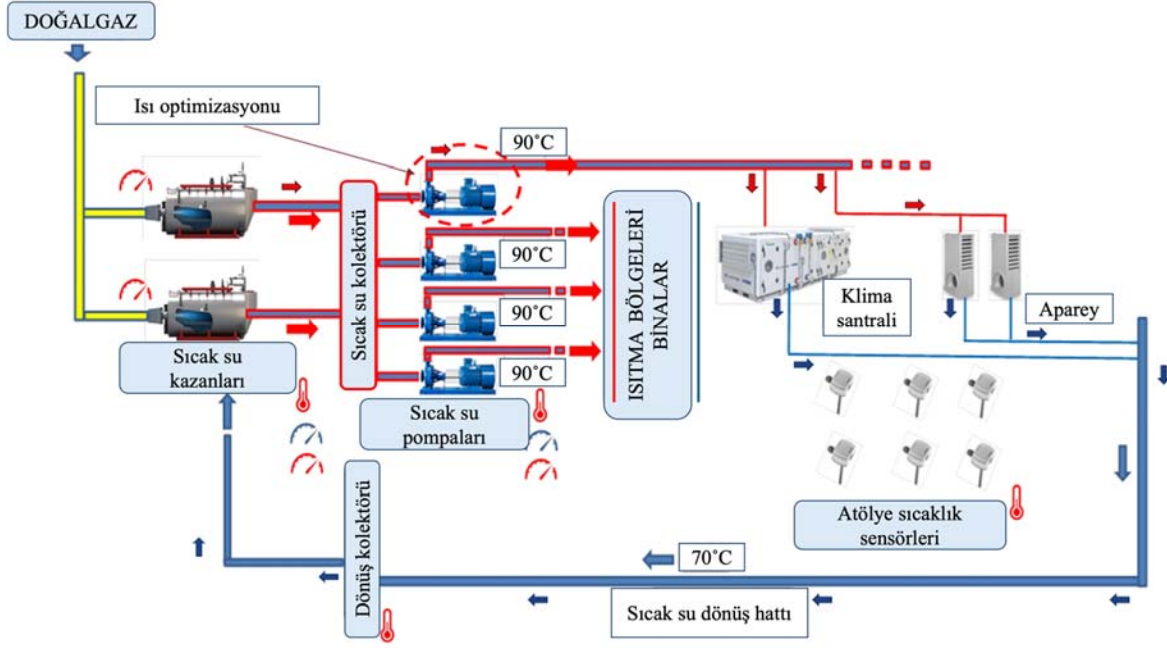
Bu bölümde yukarıda belirtilen problemlerin çözümü için işletme binasında ortamın hedef sıcaklıkta tutulmasını sağlayacak ısı tahminlemesi çalışması yapılmıştır. Çalışma kapsamında dış hava sıcaklığı, oda sıcaklığı değişimleri, geçmişte farklı dönemlere ait hava durumu tahminleri girdi değişkenleri olarak kullanılmıştır. Isı tahmini algoritması, makine öğrenmesi algoritmalarının toplanan veriler ve bu değişkenlerden üretilen yeni verilerle eğitilmesiyle geliştirilmiştir. Bu algoritma ile fabrikanın önemli hedeflerinden biri olan enerji maliyetlerinin iyileştirilmesi için aşırı ısınmadan kaynaklanan kayıpların azaltılması amaçlanmaktadır. Şekil 4'te tesis oda sıcaklığına ve ayarlanmış sıcaklığa göre mevcut problem ve ideal durum bir görsel ile ifade edilmiştir. Çalışmanın gerçekleştirildiği tesiste belirlenen hedef sıcaklık 18° C'dir. Bu kapsamda tesislerde kurulu olan sensörlerden ve ısıtma otomasyon sisteminden birer dakikalık aralıklarla yaklaşık 1 aylık veri toplanmıştır. Bu veriler tesis mahal sıcaklığı, dış ortam sıcaklığı ve gönderilen ısı değerleridir. Bu bağlamda, işletmenin endüstriyel alanda endüstriyel ekipmanlardan gerçek zamanlı veri toplamaya imkân sağlayan ve nesnelerin interneti (IoT) platformu olan Thingworx IoT platformunun analitik araçları kullanılarak, gerçek zamanlı olarak ısıtma kazanlarının göndereceği ısı miktarı tahminlenmektedir.

Tahmin sonrası elde edilen ısı değeri ısı merkezinin SCADA sistemine yayınlanır. Bu sayede hedef sıcaklık değerine ulaşmak için sisteme optimuma yakın bir ısı değeri gönderilerek aşırı ısınma olarak tanımlanan enerji kayıpları azaltılır. Şekil 5, tesisteki genel durumu ve aşırı ısınmanın oluştuğu durumu özetlemektedir.

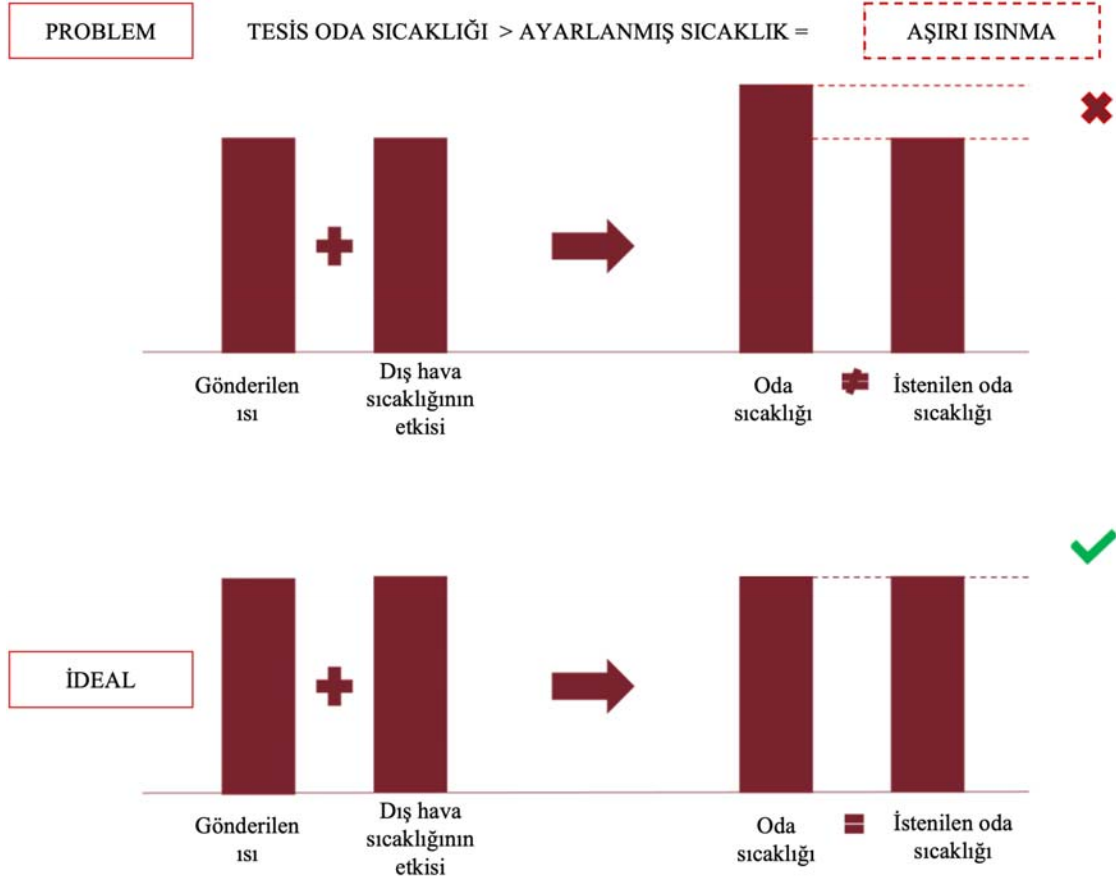
#### 4.1. Veri Toplama (1. Adım) (Data Collection (Step 1))

Çalışmanın gerçekleştirildiği tesislerden alınan veriler IoT sensörleri yardımıyla ThingWorx IoT yazılımı üzerinden toplanmıştır. Toplanan veriler, makine öğrenimi modellerini eğitmek için dışa aktarılmıştır. Çalışma toplamda 47350 satırlık bir veri seti ile gerçekleştirilmiştir. Dotzauer [49] mevcut bir merkezi ısıtma sisteminin çalışmasını optimize etmek için ısı enerjisi talebinin tahmin edilmesini gerektirdiğini vurgulamış ve enerji yükünü etkileyen en önemli parametrelerin dış ortam sıcaklığı ve kullanım özellikleri olduğunu belirtmiştir. Bu kapsamda, mevcut birimin verileri IoT sensörleri tarafından birer dakikalık frekanslarla toplanmaktadır. Bu veriler:

- Oda (mahal) sıcaklığı (°C)
- Dış ortam sıcaklığı (°C)
- Gönderilen ısı (kcal)



Şekil 3. Çalışmanın yapılacağı tesisin ısıtma düzeni (Heating scheme of the facility where the work will be done)



Şekil 4. Mevcut ve ideal durum (Current and ideal situation)

#### 4.2. Veri işleme (2. Adım) (Data processing (Step 2))

Bu adımda öncelikle toplanan veriler Python programlama dili kullanılarak analiz edilmiş, ardından eksik veriler tespit edilerek temizlenmiş ve son olarak veriler birleştirilmiştir. Mevcut veriler model geliştirme ortamına aktarıldıktan sonra hedef sıcaklığa göre özellik çıkarımı yapılmıştır. Tahmin performansını artırmak için, mevcut verilerden hedef sıcaklık ile mevcut oda sıcaklığı arasındaki fark ve oda-dış sıcaklık farkı olmak üzere iki yeni değişken türetilmiştir. Son olarak, eksik veriler işlenmiş ve veri ön işlemesi tamamlanmıştır.

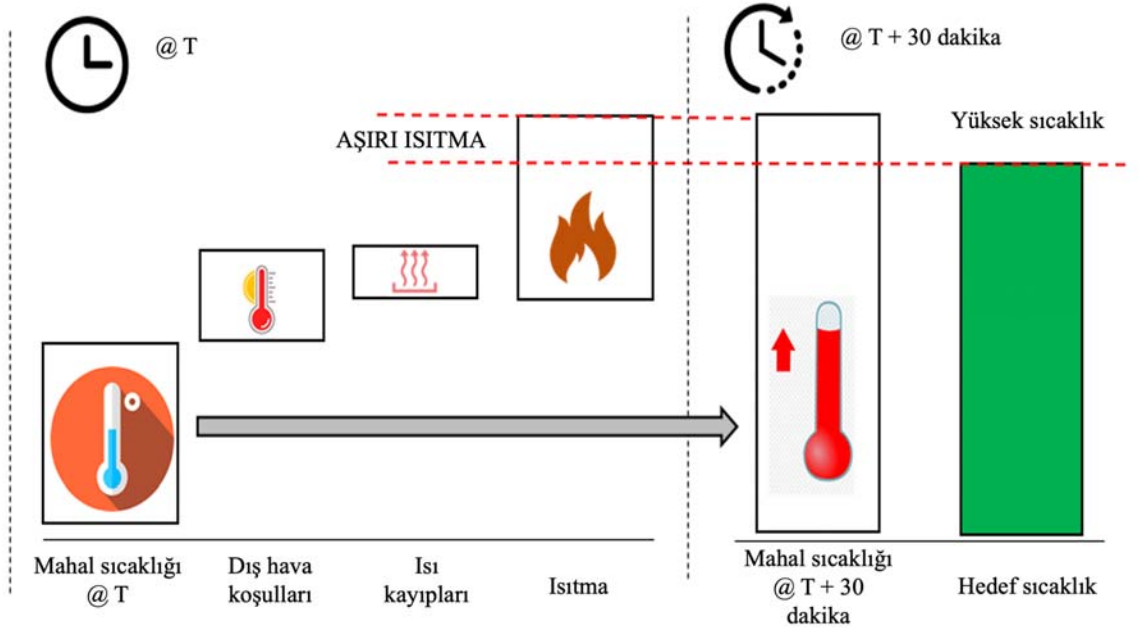
#### 4.3. Özellik Mühendisliği (3. Adım) (Feature Engineering (Step 3))

Bu adımda veri seti zamana göre büyükten küçüğe sıralanmış ve geçmiş verilerin de etkisini görebilmek için veri setindeki belirlenmiş ve sıralanmış parametrelere değerlerinin satır bazında aşağı ya da yukarı hücrelere iletilmesi ile belirlenen zamana göre kaydırma işlemleri yapılmıştır. Oda sıcaklığı, dış ortam sıcaklığı ve giden ısı parametrelerinde 15 ve 30 dakikalık zaman kaymalarıyla yeni veriler elde edilmiştir. Türetilen yeni değişkenler ve zaman kaymaları Şekil 6'da görselleştirilmiştir.

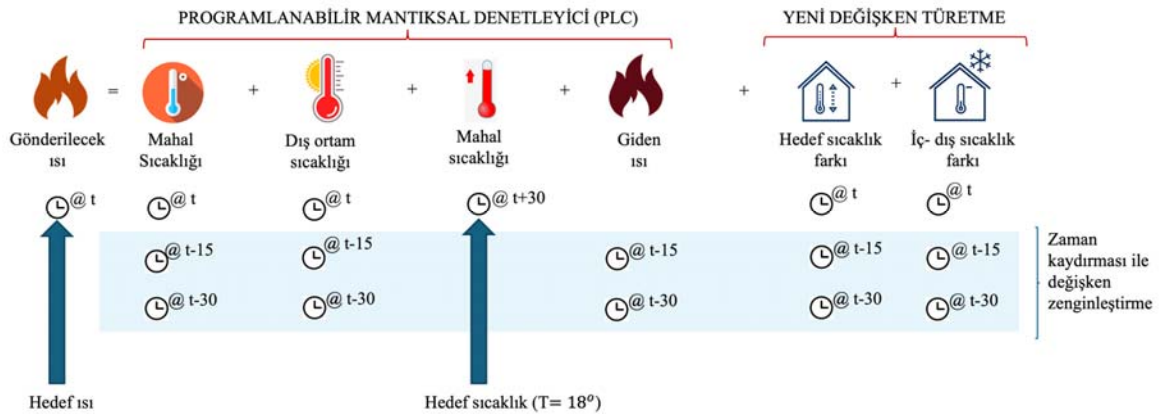
Zaman kaydırma işlemlerinden sonra kayıp verileri kaydırılmış ve boş değerleri içeren satırlar veri setinden çıkarılmıştır. Son olarak performans metrikleri olan. MAE (Ortalama Mutlak Hata), MSE (Ortalama Hata Karesi), RMSE (Kök Ortalama Kare Hatası), MAPE (Ortalama Mutlak Yüzde Hatası) ve  $R^2$  değerleri hesaplanmıştır.

#### 4.4. Makine Öğrenimine Dayalı Modelleme (4. Adım) (Modeling Based on Machine Learning (Step 4))

Modelleme aşamasında veri seti eğitim, test ve doğrulama setleri olarak üç gruba ayrılmıştır. Verilerin ilk %80'i sırasıyla %60 ve %20 oranlarında eğitim ve test seti olarak rasgele bölünmüştür. Veri setinin geri kalanı, model değerlendirmesi için kullanılan doğrulama seti için kullanılmıştır. Lineer Regresyon, Random Forest Regresyon, Polinom Regresyon, Karar Ağacı Regresyon, SVR, Ekstra Ağaç Regresyon, Ada Boost Regresyon, Gradient Boosting Regresyon, Oylama Regresyon ve YSA olmak üzere 10 farklı regresyon modeli eğitilmiştir. Girdi değişkenlerinin tahmin modelleri üzerindeki etkisini görmek ve bu modelleri eğitmek için dört farklı girdi değişken grubu kullanılmıştır. Deneylerin girdileri ve girdi verilerinin hangi deneylerde kullanıldığı Tablo 2'de gösterilmektedir.



Şekil 5. Tesisteki genel durum (General condition of the facility)



Şekil 6. Yeni değişken oluşturma (Creating a new variable)



Eğitim ve test setlerini içeren verilerin ilk kısmı, tahmin performansını artırmak için çeşitlendirilmiş örnekler sağlamak üzere rastgele bölünmüştür. Modellerin performansları yanıtıcı olabilir çünkü karakteristik problem zaman serisindedir ve test eğitim seti ile aynı değer altında yürütülmüştür. Bu nedenle, makine öğrenmesi modellerinin performansını analiz etmek için doğrulama seti de kullanılmıştır. Bu adımda, belirtilen makine öğrenimi modelleri eğitilmiştir ve ön performansı değerlendirmek için test setleri kullanılmıştır. 1. Deneye göre performans metriklerinin sonuçları Tablo 3, 2. Deneye göre performans metriklerinin sonuçları Tablo 4, 3. Deneye göre performans metriklerinin sonuçları Tablo 5 ve 4. Deneye göre performans metriklerinin sonuçları Tablo 6'da gösterilmiştir.

#### 4.5. Doğrulama Setleri ve Model Değerlendirmesi (Adım 5) (Validation Sets and Model Evaluation (Step 5))

Zaman serileri dışındaki çoğu problemde veri seti belli bir orana göre rastgele eğitim ve test gruplarına ayrılır [50]. Ancak bu, zaman serileri

için yapılamaz çünkü veri noktaları yalnızca gerçek sıralarında anlamlıdır. Zaman serisi analizinde eğitim verisinden hemen sonra test verisi başlamalı ve karıştırma yapılmamalıdır [51]. Test seti olarak ayrılan gözlem değerleri, zaman açısından eğitim seti olarak ayrılan gözlem değerlerinden sonra gelmelidir [52].

Bu durumda problem, zaman serisi şeklinde ilerleyen verilerden tahmine dayalıdır ve Adım 2'de belirtildiği gibi verilerde zaman kayması ile türetilen yeni değişkenler elde edilmiştir. Bu nedenle, doğrulama verilerini zaman serisi formuyla kullanmak daha uygun bulunmuştur. Adım 4'te açıklandığı gibi, toplam veri setinin yaklaşık %20' sini oluşturan en yakın zaman aralığı, doğrulama seti olarak bölünmüştür ve Adım 4'te açıklanan model testinden sonra tahmin ve gerçek sonuçları karşılaştırmak için kullanılmıştır.

Bir sonraki adımda değinilecek olan ve çözüm mimarisinde kullanılacak olan ThingWorx IoT Yazılımında bulunan modellerin kullanım kısıtlaması nedeniyle deneyler için üç farklı makine öğrenmesi modeli kullanılmıştır. Bu modeller ekstra ağaç regresyonu,

**Tablo 2.** Deneyler ve giriş değişkenleri (Experiments and input variables)

Girdi Değişkenleri	Deney 1 (Tablo 2)	Deney 2 (Tablo 3)	Deney 3 (Tablo 4)	Deney 4 (Tablo 5)
Oda sıcaklığı	X	X	X	X
Dış ortam sıcaklığı	X	X	X	X
Gönderilen ısı	X	X	X	X
Zaman periyodu	X	X		
Oda sıcaklığı (t+30)	X	X	X	X
Gönderilen ısı (t-15)		X	X	X
Gönderilen ısı (t-30)		X	X	X
Dış ortam sıcaklığı tahmini (t+60)				X

**Tablo 3.** 1. Deney (Experiment 1)

Metrik	Linear	Random Forest	Polynomial	Decision Tree	SVR	Extra Tree	Ada Boost	Gradient Boosting	Voting	ANN
MAE	716785,37	29961,89	475634,16	25932,51	374584,86	21373,68	703805,8	360986,22	188460,8	325948,59
MSE	8,99	1,32	4,77	4,05	4,04	8,17	9,15	3,29	9,46	3,24
RMSE	948616	114808,48	690727,03	636097,87	636097,87	90409,54	956537,7	574083,17	307496,96	2693157,99
MAPE	2,85	2,63	1,24	8,39	8,39	1,81	2,91	9,56	4,90	7,70
R <sup>2</sup>	0,76	0,99	0,87	0,99	0,89	0,99	0,76	0,91	0,98	0,91

**Tablo 4.** 2. Deney (Experiment 2)

Metrik	Linear	Random Forest	Polynomial	Decision Tree	SVR	Extra Tree	Ada Boost	Gradient Boosting	Voting	ANN
MAE	149670,19	34145,93	156978,92	33677,34	141581,17	22915,25	597744,17	112452,78	68644,19	133202,52
MSE	1,41	2,20	1,55	3,42	1,18	1,07	4,97	1,01	4,20	1,17
RMSE	376721,18	148379,44	394888,15	1850002,63	344834,47	103840,57	705069,89	317586,74	204992,57	2799160,92
MAPE	4,43	3,48	3,26	2,18	5,07	1,98	3,45	2,59	1,46	7,80
R <sup>2</sup>	0,96	0,99	0,96	0,99	0,97	0,99	0,88	0,97	0,99	0,97

**Tablo 5.** 3. Deney (Experiment 3)

Metrik	Linear	Random Forest	Polynomial	Decision Tree	SVR	Extra Tree	Ada Boost	Gradient Boosting	Voting	ANN
MAE	153447,68	34735,95	168653,84	32853,88	150168,82	21037,91	650676,03	116863,86	71771,54	138152,34
MSE	1,84	2,08	2,00	3,14	1,62	7,86	7,06	1,21	4,89	1,58
RMSE	429369,59	144374,70	447728,11	177387,73	402582,77	88672,93	840658,68	348208,21	221145,86	2778528,39
MAPE	3,75	2,12	2,78	1,01	4,65	8,66	4,02	2,14	1,19	7,71
R <sup>2</sup>	0,95	0,99	0,95	0,99	0,96	0,99	0,82	0,97	0,99	0,96

**Tablo 6.** 4. Deney (Experiment 4)

Metrik	Linear	Random Forest	Polynomial	Decision Tree	SVR	Extra Tree	Ada Boost	Gradient Boosting	Voting	ANN
MAE	149288,57	30497,81	142992,92	29465,69	141754,14	19477,18	387009,90	109624,16	65586,12	152782,05
MSE	1,49	1,73	1,35	2,44	1,29	7,45	3,56	1,01	3,88	1,30
RMSE	387010,31	131819,61	368701,46	156449,25	359225,88	86351,56	597165,63	318212,97	197027,48	2795600,49
MAPE	4,20	2,76	3,44	8,70	4,90	1,46	1,86	2,33	1,29	7,93
R <sup>2</sup>	0,96	0,99	0,97	0,99	0,97	0,99	0,91	0,97	0,99	0,97

lineer regresyon ve karar ağacı regresyonudur. 1. Deneye göre kullanılan makine öğrenimi algoritmalarının doğrulama sonuçları Tablo 7'de, 2. Deneye göre kullanılan makine öğrenimi algoritmalarının doğrulama sonuçları Tablo 8'de, 3. Deneye göre kullanılan makine öğrenimi algoritmalarının doğrulama sonuçları Tablo 9'da, 4. Deneye göre kullanılan makine öğrenimi algoritmalarının doğrulama sonuçları Tablo 10'da gösterilmiştir.

Dağıtım için bu makine öğrenme algoritmaları arasında en yüksek  $R^2$  değerine sahip doğrusal regresyon modeli seçilmiştir. Ayrıca deneyler arasında yüksek performans gösteren ve daha az özellik eklenerek daha az işlem gerektiren 2. deney seçilmiştir. (Tablo 8).

Doğrusal Regresyon algoritması ile elde edilen sonuçlar değerlendirildiğinde 1. deney dışındaki tahmin değerlerinin  $R^2$  değerinin 0.92 olduğu görülmektedir. Yani modelden elde edilen tahmin için kullanılacak eğitilmiş verilerin test verileriyle %92 oranında örtüştüğü ve bu oranda doğru tahmin yaptığı görülmektedir. Isı ( $kcal/h$ ), sıcaklık ( $^{\circ}C$ ) ve modele giren verilerin sırasına göre lineer regresyon model sonucunun ifade edildiği Şekil 7'de  $Y_{\text{pred}}$ 'in gösterdiği öngörü değerleri ile  $Y_{\text{test}}$ 'in ifade ettiği doğrulama değerlerinin büyük ölçüde aynı grafiğe sahip olduğu görülmektedir. Bu durum tahmin değeri ile doğrulama değerinin eşleşmesini görsel olarak göstermektedir.

#### 4.6. Çözüm Mimarisi ve Saha Testleri (Adım 6) (Solution Architecture and Field Tests (Step 6))

PTC (Parametric Technology Corporation) Thingworx, geliştirilen ve artık evrensel olan konseptte yaygın olarak kullanılan tanımlama, veri toplama, veri görselleştirme ve veri analitiğinde kullanılan yazılımların öncülerinden biri olarak pratikte kullanılan bir nesnelerin interneti (IoT) yazılımıdır. Analitik araçları için sınırlı makine öğrenimi modelleri içermesi gibi bazı kısıtlamalara sahip olsa da uçtan uca problem çözümü için güçlü bir araçtır. Çalışmanın yaptığı fabrikanın koşulları da göz önüne alındığında geliştirilen modeli canlı bir ortamda deneyimleyebilmek için önceki adımlarda uygulanan modelleme yaklaşımları, PTC ThingWorx IoT Foundation ve Analytical Module'ün canlı sunucusu tarafında gerçekleştirilmiştir.

Python'da zaman kayması ve değişken açılım çalışmaları (zaman kayması veya değişken farkları vb.) yapılmış ve önceden işlenmiş veri seti ThingWorx IoT ortamına aktarılmıştır. Canlı sunucuda ilgili değişken genişletme çalışması ThingWorx IoT altyapısı kullanılarak yapılmıştır. Bu çalışmalardan sonra makine öğrenimi modeli gerçek zamanlı olarak çalıştırılmış ve tahmin değeri karar destek sistemi olarak kullanılmıştır. Bu aşamada gerçekleştirilen çözüm mimarisi Şekil 8'de özetlenmiştir.

**Tablo 7.** 1. Deneyin Model Değerlendirmeleri (1. Experiment Model Evaluations)

Metrik	Test			Doğrulama		
	Linear	Decision Tree	Extra Tree	Linear	Decision Tree	Extra Tree
MAE	716785,37	25932,51	21373,68	754933,22	724599,26	653061,07
MSE	8,99	4,05	8,17	8,92	1,99	1,06
RMSE	948616	636097,87	90409,54	944798,79	1411749,80	1033506,08
MAPE	2,85	8,39	1,81	2,89	1,03	1,44
$R^2$	0,76	0,99	1,00	0,77	0,49	0,73

**Tablo 8.** 2. Deneyin Model Değerlendirmeleri (2. Experiment Model Evaluations)

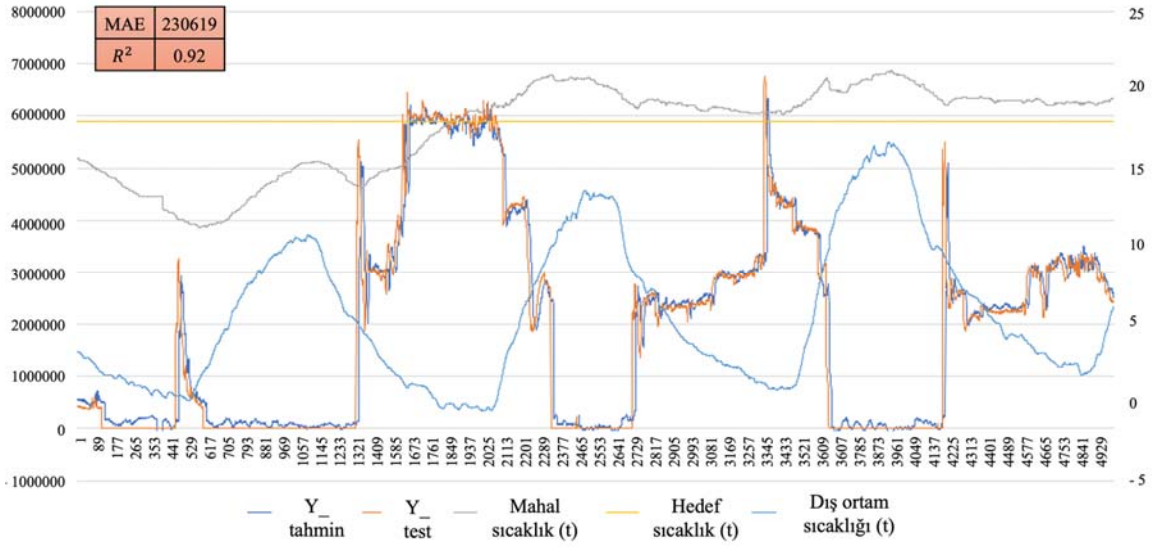
Metrik	Test	Doğrulama				
		Decision Tree	Extra Tree	Linear	Decision Tree	Extra Tree
MAE	149670,19	33677,34	22915,25	235135,60	425615,57	321419,30
MSE	1,41	3,42	1,07	3,07	8,71	4,43
RMSE	376721,18	1850002,63	103840,57	554665,24	933514,83	665957,13
MAPE	4,43	2,18	1,98	5,08	1,59	5,18
$R^2$	0,96	0,99	1,00	0,92	0,78	0,89

**Tablo 9.** 3. Deneyin Model Değerlendirmeleri (3. Experiment Model Evaluations)

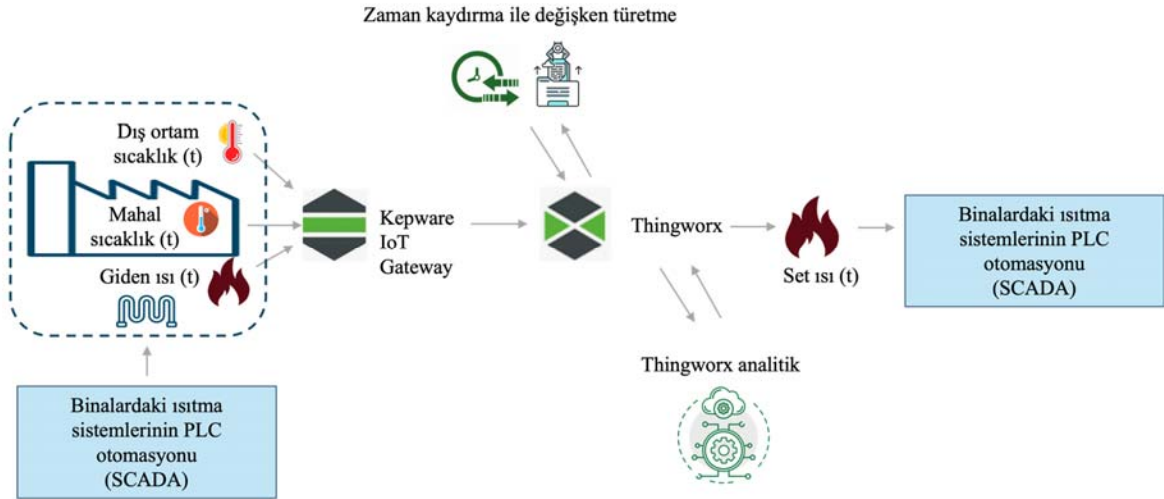
Metrik	Test			Doğrulama		
	Linear	Decision Tree	Extra Tree	Linear	Decision Tree	Extra Tree
MAE	153447,68	32853,88	21037,91	228781,33	380835,74	328691,02
MSE	1,84	3,14	7,86	3,08	7,41	4,60
RMSE	429369,59	177387,73	88672,93	555716,26	861287,48	678775,98
MAPE	3,75	1,01	8,66	4,63	3,52	4,68
$R^2$	0,95	0,99	1,00	0,92	0,81	0,88

**Tablo 10.** 4. Deneyin Model Değerlendirmeleri (4. Experiment Model Evaluations)

Metrik	Test			Doğrulama		
	Linear	Decision Tree	Extra Tree	Linear	Decision Tree	Extra Tree
MAE	149288,57	29465,69	19477,18	236708,13	422949,46	359538,68
MSE	1,49	2,44	7,45	3,04	8,13	4,93
RMSE	387010,31	156449,25	86351,56	551955,33	902086,22	702367,68
MAPE	4,20	8,70	1,46	5,12	1,74	6,73
$R^2$	0,96	0,99	1,00	0,92	0,79	0,87



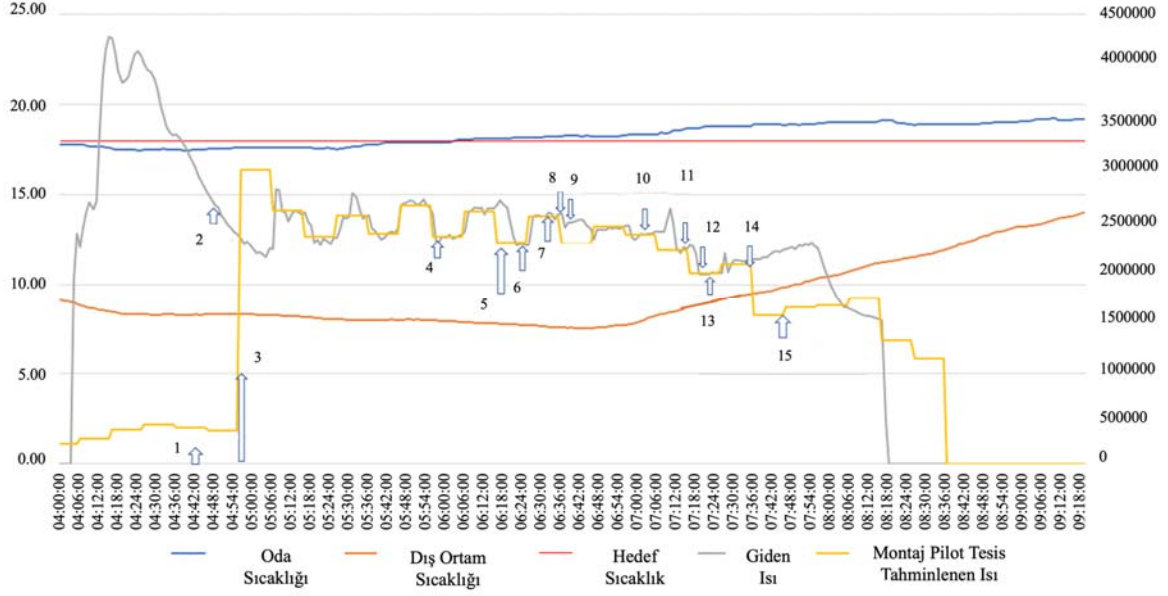
Şekil 7. Isı ( $kcal/h$ ), sıcaklık ( $^{\circ}C$ ) ve modele giren verilerin sırasına göre lineer regresyon model sonucu (Linear regression model result)



Şekil 8. Çözüm Mimarisi (Solution Architecture)

Çözüm mimarisi çalışmalarının ardından canlı saha testleri gerçekleştirilmiştir. Bu amaçla çalışmanın yapıldığı tesiste 5 saatlik canlı test çalışması yapılmıştır. Saha testi sonucunda elde edilen model sonucu sıcaklık ( $^{\circ}C$ ), ısı ( $kcal/h$ ) ve modele giren verilerin zamanına göre Şekil 9'da verilmiştir. Sonuç, makine öğrenimi modelinin, ısıtma sistemi manuel olarak devreye alındıktan sonra oda sıcaklığını istenilen seviyede tutabildiğini ve sağlayabildiğini göstermiştir. Ancak, özellikle başlangıç dönemi konusunda bazı iyileştirme noktalarının olduğu gözlemlenmiştir. Makine öğrenimi modeli tahmin değeri, hedef sıcaklığın üzerinde düşüktür. Farklı mevsimlerden yeni veri örnekleri ile modeli eğitilerek geliştirilebilir. Ayrıca gerçek zamanlı testte hedef sıcaklıktan bazı sapmalar gözlemlenmiştir. Çünkü ısıtma sisteminde tahmin değerini uygulamak için teknik kısıtlamalar vardır. Şekil 9'da gösterilen 1'den 15'e kadar olan sayılar ile, bölgesel bazda sapmalara neden olan müdahale kısıtlamaları ifade edilmektedir. Bu kısıtlamalar aşağıda verilmiştir.

1. Etiket güncellenmiştir.
2. Kalibrasyon yapılmıştır.
3. Pompa kademesi yükseltilmiştir.
4. Atölyedeki bazı klima santralleri devre dışı bırakılmıştır.
5. Pompaya dışarıdan müdahale edilmiştir (Pompa en düşük seviyede çalıştırıldı.).
6. Vana ayarı yapılmıştır.
7. Klima santrali devreye alınmıştır.
8. Sıcaklık düşürülmüştür.
9. Ana vanaya müdahale edilmiştir.
10. Vana % 80 kısılmıştır.
11. Vana % 90 kısılmıştır.
12. Vana % 95 kısılmıştır.
13. Vana % 99 kısılmıştır.
14. Vana % 95 kısılmıştır.
15. Atölye ısıtması ve fabrika geneli ısıtması devre dışı bırakılarak sıcak su kazanı kapatılmıştır.



Şekil 9. Sıcaklık (°C), Isı (kcal/h) ve modele giren verilerin zamanına göre saha testi sonrasında elde edilen doğrusal regresyon modeli sonucu (Linear regression model result obtained after field test)

## 5. Sonuçlar (Conclusions)

Bu çalışma, bir otomotiv firmasının mevcut tesisinde bir ay boyunca 1 dakikalık aralıklarla toplanan sıcaklık verilerini kullanarak belirli bir tesiste izin verilen ısı değerini pratik olarak tahmin etmek için 10 farklı makine öğrenme algoritmasının kullanımını detaylandırmaktadır. Başlangıçta, tahmin çalışmaları Python Skit-learn ve Keras kütüphaneleri kullanılarak toplanmıştır. Daha sonra ısıtma otomasyon sisteminin bütünleşmiş çalışabildiği PTC ThingWorx IoT yazılımının analitik modülü kullanılarak modeller canlı ortam ve saha testleri için devreye alınmıştır. Değerlendirmelerin ve testlerin sonuçları doğrusal regresyon modelinin en yüksek tahmin performansına sahip olduğunu göstermiştir. Burada en önemli performans ölçütlerinden biri olan R-kare değeri 0,92 çıkmıştır. Değerin bire yakın olması, çalışmada kullanılan verilerin seçilen modele uygunluğunu göstermiştir. Özellikle yapılan çalışmanın gerçek zamanlı olarak tahmin çıktısı üretebilmesi nedeniyle bu çalışmanın endüstriyel ortamlardaki benzer projelere örnek teşkil etmesi beklenmektedir.

Bu çalışma gerçek bir endüstriyel sistem yerinde gerçekleştirilmiş ve sadece teorik temellere dayanmamıştır. IoT yazılımı ve makine öğrenimi teknolojilerinin her ikisi de kullanılmıştır ve sonuçlar canlı bir ortamda test edilmiştir. Çalışma kapsamında geliştirilen model ile yarım saat öncesinde tahminlenen ısı Thingworx IoT platformu üzerinden fabrika binalarındaki SCADA ekranlarına yansıtılarak tesislere gönderilecek ısı bu değerlere göre ayarlanmıştır. Çalışmanın bir diğer sonucu olan maliyet kazancı kapsamında ise sadece çalışmanın yapıldığı pilot tesiste aşırı ısıtmadan kaynaklanan maliyetlerde yıllık % 90 azalma gözlemlenmiştir. Literatürde elde edilen sonuçları destekleyen, makine öğrenimi yöntemlerinin kullanılması ile fabrika binalarındaki ısı değerlerini araştıran çalışmalar açısından benzerlik gösteren çalışmalar mevcuttur. Bunlar arasında, IoT teknolojisinin kullanılması ve SCADA sistemi ile entegrasyonu açısından farklılık gösteren fakat benzer amaçlar doğrultusunda gerçekleştirilmiş bir çalışmada, Ra vd. [53] bir fabrika binasında çoklu derin sinir ağı (DNN) modellerini kullanarak model öngörülmesi kontrol (MPC) sonuçları sunarak soğutma ayar noktası sıcaklığını korurken yoğunlaşma ünitelerinin enerji tüketimini % 32

azaltmıştır. Bir diğer çalışmada, Ghofrani vd. [54] değişken hava hacmi (VAV) sisteminde doğrusal olmayan durumları yakalamak için, bina bölgesi seviyesinde çok değişkenli doğrusal regresyon ve YSA modellerini önererek, ön görülmesi kontrol ve iç sıcaklığın doğrusal olmasını sağlamışlardır.

Bu çalışmanın gerçek bir endüstriyel sisteme dayanması ve iki farklı Endüstri 4.0 konusunun bütünleşik bir şekilde kullanılması nedeniyle literatüre önemli bir katkı sağlaması beklenmektedir.

Bu çalışma, söz konusu işletmenin pilot tesisinde gerçekleştirilmiştir. Burada ağır basan varsayım, kullanılan yaklaşımların, yöntemlerin ve algoritmaların pilot tesiste veya diğer tesislerde iyileştirme için etkin olarak kullanılabilirdir. Bu varsayım hem çalışmanın yapıldığı fabrika ve endüstriyel sanayi kuruluşları hem de diğer binalar için maliyet avantajı sağlamaktadır. Bunun yanı sıra binalardaki ısıtma maliyetleri düşerken, fazla ısıya maruz kalan çalışanlar ve çevre içinde daha konforlu çalışma ortamı sağlanabilir olması da çalışmanın bir diğer avantajı arasındadır. Ayrıca gözlemlerle otomatik çevrimde ısı üretilerek yarı otomatik bir kullanım çerçevesi oluşturulabilir. Makine öğrenme algoritmalarının iyileştirilmesi ve ısı kontrol algoritmalarının tasarımı ile yarı otomatik kullanımın ötesinde tam otomatik bir kullanım çerçevesi tasarlanabilir. Çalışmada kullanılan IoT yazılımının, makine öğrenmesi modellemesi kısıtlamaları ortadan kaldırılarak daha kapsamlı makine öğrenmesi modelleri üzerinde çalıştırılabilir. Mevcut modellerin yanı sıra yeni makine öğrenmesi modelleri geliştirilip IoT yazılımlarına entegre edilebilir. Son olarak gelecekte optimizasyon çalışmaları ile yapılan çalışmalar birleştirilerek optimum ısı değerini tahmin edebilen hibrit bir model oluşturulabilir ve otomatize edilebilir.

## Teşekkür (Acknowledgement)

Bu çalışma 2244 TÜBİTAK Sanayi-Doktora Projesi kapsamında hazırlanmıştır. Yazar "Makbule Nalkıran", bu makale çalışmasını Yıldız Teknik Üniversitesi'nde doktora derecesi almak için yerine getirmesi gereken gerekliliklerden bir tanesi olarak gönderdiğini kabul etmektedir. Yazarlar, TÜBİTAK ve TOFAŞ Türkiye Otomobil Fabrikası A.Ş.'ye teşekkür eder.

**Kaynaklar (References)**

1. Türkiye Cumhuriyeti Çevre Şehircilik ve İklim Değişikliği Bakanlığı. Sektörlere Göre Nihai Enerji Tüketimi. <https://cevreselgostergeler.csb.gov.tr/sektorlere-gore-nihai-enerji-tuketimi-i-85804#>. Yayın tarihi 2023. Erişim Tarihi Mayıs 24, 2024.
2. Yılmazoğlu, M. Z., & Durmaz, A., Fabrikalarda Enerji Ekonomisi için Absorpsiyonlu Sistemlerin ve Parabolik Oluk Tip Güneş Kollektörlerinin Kullanılması, 9. International Combustion Symposium, Kırıkkale, 393-401, 2006.
3. Langroudi, P. P., & Weidlich, I., Applicable Predictive Maintenance Diagnosis Methods in Service-Life Prediction of District Heating Pipes, *Environmental & Climate Technologies*, 24 (2), 2020.
4. Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S., Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques, *Journal of Building Engineering*, 45, 103406, 2022.
5. Pavelčík, V., & Kuba, E., Application of basic machine learning algorithms in railway brake disc temperature prediction, *Transportation Research Procedia*, 55, 715-722, 2021.
6. Bhowmik, R., Sihni, S., Pachter, R., & Vernon, J. P., Prediction of the specific heat of polymers from experimental data and machine learning methods, *Polymer*, 220, 123558, 2021.
7. Liu, B., Vu-Bac, N., & Rabczuk, T., A stochastic multiscale method for the prediction of the thermal conductivity of Polymer nanocomposites through hybrid machine learning algorithms, *Composite Structures*, 273, 114269, 2021.
8. Swartz, B., Wu, L., Zhou, Q., & Hao, Q., Machine learning predictions of critical heat fluxes for pillar-modified surfaces, *International Journal of Heat and Mass Transfer*, 180, 121744, 2021.
9. Liu, J., Huang, Q., Ulishney, C., & Dumitrescu, C. E., Machine learning assisted prediction of exhaust gas temperature of a heavy-duty natural gas spark ignition engine, *Applied Energy*, 300, 117413, 2021.
10. He, M., & Lee, Y., Application of machine learning for prediction of critical heat flux: Support vector machine for data-driven CHF look-up table construction based on sparingly distributed training data points, *Nuclear Engineering and Design*, 338, 189-198, 2018.
11. Sajjadi, S., Shamsheerband, S., Alizamir, M., Yee, L., Mansor, Z., Manaf, A. A., & Mostafaeipour, A., Extreme learning machine for prediction of heat load in district heating systems, *Energy and Buildings*, 122, 222-227, 2016.
12. Zhang, Q., Tian, Z., Ma, Z., Li, G., Lu, Y., & Niu, J., Development of the heating load prediction model for the residential building of district heating based on model calibration, *Energy*, 205, 117949, 2020.
13. Sundar, S., Rajagopal, M. C., Zhao, H., Kuntumalla, G., Meng, Y., Chang, H. C., & Salapaka, S., Fouling modeling and prediction approach for heat exchangers using deep learning, *International Journal of Heat and Mass Transfer*, 159, 120112, 2020.
14. Dai, Z., Chen, Z., Selmi, A., Jermittiparsert, K., Denić, N. M., & Nešić, Z., Machine learning prediction of higher heating value of biomass, *Biomass Conversion and Biorefinery*, 1-9, 2021.
15. Lin, T., Pan, Y., Xue, G., Song, J., & Qi, C., A Novel Hybrid Spatial-Temporal Attention-LSTM Model for Heat Load Prediction, *IEEE Access*, 8, 159182-159195, 2020.
16. Song, J., Xue, G., Pan, X., Ma, Y., & Li, H., Hourly heat load prediction model based on temporal convolutional neural network, *IEEE Access*, 8, 16726-16741, 2020.
17. Yuan, J., Zhou, Z., Huang, K., Han, Z., Wang, C., & Lu, S., Analysis and evaluation of the operation data for achieving an on-demand heating consumption prediction model of district heating substation, *Energy*, 214, 118872, 2021.
18. Hughes, M. T., Fronk, B. M., & Garimella, S., Universal condensation heat transfer and pressure drop model and the role of machine learning techniques to improve predictive capabilities, *International Journal of Heat and Mass Transfer*, 179, 121712, 2021.
19. Vu, A. T., Gulati, S., Vogel, P. A., Grunwald, T., & Bergs, T., Machine learning-based predictive modeling of contact heat transfer, *International Journal of Heat and Mass Transfer*, 174, 121300, 2021.
20. Zhu, G., Wen, T., & Zhang, D., Machine learning based approach for the prediction of flow boiling/condensation heat transfer performance in mini channels with serrated fins, *International Journal of Heat and Mass Transfer*, 166, 120783, 2021.
21. Park, S., Bailey, J. P., Pasman, H. J., Wang, Q., & El-Halwagi, M. M., Fast, easy-to-use, machine learning-developed models of prediction of flash point, heat of combustion, and lower and upper flammability limits for inherently safer design, *Computers & Chemical Engineering*, 155, 107524, 2021.
22. Kim, H., Moon, J., Hong, D., Cha, E., & Yun, B., Prediction of critical heat flux for narrow rectangular channels in a steady state condition using machine learning, *Nuclear Engineering and Technology*, 53 (6), 1796-1809, 2021.
23. Wu, W., Wang, J., Huang, Y., Zhao, H., & Wang, X., A novel way to determine transient heat flux based on GBDT machine learning algorithm, *International Journal of Heat and Mass Transfer*, 179, 121746, 2021.
24. Lu, S., Xu, Q., Jiang, C., Liu, Y., & Kusiak, A., Probabilistic load forecasting with a non-crossing sparse-group Lasso-quantile regression deep neural network, *Energy*, 122955, 2021.
25. Tuğçe, Y. Ü. C. E., & Kabak, M., Makine Öğrenmesi Algoritmaları ile Detay Üretim Alanları İçin İş Merkezi Kırılımında Üretim Süresi Tahminleme. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 37 (1), 47-60, 2021.
26. Fan, X., Wang, X., Zhang, X., & Yu, P. A. X. B., Machine learning based water pipe failure prediction: The effects of engineering, geology, climate and socio-economic factors, *Reliability Engineering & System Safety*, 219, 108185, 2022.
27. Chen, W., Qiu, Y., Feng, Y., Li, Y., & Kusiak, A., Diagnosis of wind turbine faults with transfer learning algorithms, *Renewable Energy*, 163, 2053-2067, 2021.
28. Xu, H., Jiao, Z., Zhang, Z., Huffman, M., & Wang, Q., Prediction of methane hydrate formation conditions in saltwater using machine learning algorithms, *Computers & Chemical Engineering*, 151, 107358, 2021.
29. Shapi, M. K. M., Ramli, N. A., & Awalın, L. J., Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Developments in the Built Environment*, 5, 100037, 2021.
30. Reşat H.G., Design and development of hybrid forecasting model using artificial neural networks and ARIMA methods for sustainable energy management systems: A case study in tobacco industry, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35 (3), 1129-1140, 2020.
31. Zhu, J., Shen, Y., Song, Z., Zhou, D., Zhang, Z., & Kusiak, A., Data-driven building load profiling and energy management. *Sustainable Cities and Society*, 49, 101587, 2019.
32. Wang, Y., Li, W., Zhang, Z., Shi, J., & Chen, J., Performance evaluation and prediction for electric vehicle heat pump using machine learning method, *Applied Thermal Engineering*, 159, 113901, 2019.
33. Bayram, K. Ö. S. E., Aygün, H., & Semih, P. A. K., Parameter estimation of the wind speed distribution model by dragonfly algorithm, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 38 (3), 1747-1756, 2023.
34. Friedman, J., Hastie, T., & Tibshirani, R., *The Elements of Statistical Learning*, New York: Springer Series in Statistics, 1 (10), 2021.
35. Öztemel, E., *Yapay Sinir Ağları*, İstanbul, Papatya Yayıncılık, 2003.
36. Hinton G., et al. *Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups*, *IEEE Signal Process. Mag.*, 29 (6):82-97, 2012.
37. Wolpert, D. H. ve W. G. Macready, No free lunch theorems for optimization, *IEEE transactions on evolutionary computation* 1 (1), 67-82, 1997.
38. Gültepe, Y., Makine Öğrenmesi Algoritmaları ile Hava Kirliliği Tahmini Üzerine Karşılaştırmalı Bir Değerlendirme, *Avrupa Bilim ve Teknoloji Dergisi*, (16), 8-15, 2019.
39. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J., *Classification and regression trees*. Routledge, 2017.
40. Rodriguez-Galiano, V., Mendes, M. P., Garcia-Soldado, M. J., Chica-Olmo, M., & Ribeiro, L., Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain), *Science of the Total Environment*, 476, 189-206, 2014.
41. Geurts, P., Ernst, D., & Wehenkel, L., Extremely randomized trees, *Machine learning*, 63, 3-42, 2006.
42. John, V., Liu, Z., Guo, C., Mita, S., & Kidono, K., Real-time lane estimation using deep features and extra trees regression, In *Image and Video Technology: 7th Pacific-Rim Symposium, PSIVT 2015*, Auckland, New Zealand, November 25-27, 2015, Revised Selected Papers 7 (pp. 721-733), Springer International Publishing, 2016.



43. Nitin D., Babita S., Chalak H.D., Gradient boosting-based regression modelling for estimating the time period of the irregular precast concrete structural system with cross bracing, *Journal of King Saud University Engineering Sciences*, 1018-3639, 2021.
44. Geeksforgeeks. Makine öğrenimi | Sklearn kullanarak Oylama Sınıflandırıcısı. <https://www.geeksforgeeks.org/ml-voting-classifier-using-sklearn/>. Güncelleme tarihi Kasım 25, 2019. Erişim Tarihi: Ekim 21, 2023.
45. Basu, V., Prediction of Stellar Age with the Help of Extra-Trees Regressor in Machine Learning, In Proceedings of the International Conference on Innovative Computing & Communications (ICICC), 2020.
46. Suykens, J. A., & Vandewalle, J., Chaos control using least-squares support vector machines, *International journal of circuit theory and applications*, 27 (6), 605-615, 1999.
47. El-Said, E. M., Abd Elaziz, M., & Elsheikh, A. H., Machine learning algorithms for improving the prediction of air injection effect on the thermohydraulic performance of shell and tube heat exchanger, *Applied Thermal Engineering*, 185, 116471, 2021.
48. Cinsdiki, M., *Neural Network Solutions for ATM Routing & Multicasting Problems*, 1997.
49. Dotzauer, E., Simple model for prediction of loads in district-heating systems, *Applied Energy*, 73 (3-4), 277-284, 2002
50. Han, J., Kamber, M., ve Pei, J., *Data mining concepts and techniques third edition*, The Morgan Kaufmann Series in Data Management Systems, 5 (4), 83-124, 2011.
51. Kamber, E., Körpüz, S., Melih, C. A. N., Aydoğmuş, H. Y., & Gümüş, M., Yapay Sinir Ağlarına Dayalı Kısa Dönemli Elektrik Yüğü Tahmini, *Endüstri Mühendisliği*, 32 (2), 364-379, 2021.
52. Bergmeir, C. ve J. M. Benítez, On the use of cross-validation for time series predictor evaluation, *Information Sciences*, 2012.
53. Ra, S. J., Kim, J. H., & Park, C. S., Real-time model predictive cooling control for an HVAC system in a factory building, *Energy and Buildings*, 285, 112860, 2023.
54. Ghofrani, A., Nazemi, S. D., & Jafari, M. A., Prediction of building indoor temperature response in variable air volume systems, *Journal of Building Performance Simulation*, 13 (1), 34-47, 2020.