



Yapay Zekâ ve Makine Öğrenimi Araçları ile Elektrik Tüketimi Tahmini

Yazılım Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Umut Yıldız

ORCID 0000-0001-9431-8797

Tez Danışmanı: Doç. Dr. Sıla Övgü Korkut Uysal

Haziran 2023

İzmir Kâtip Çelebi Üniversitesi Fen Bilimleri Enstitüsü öğrencisi **Umut Yıldız** tarafından hazırlanan **Yapay Zeka ve Makine Öğrenimi Araçları ile Elektrik Tüketimi Tahmini** başlıklı bu çalışma tarafımızca okunmuş olup, yapılan savunma sınavı sonucunda kapsam ve nitelik açısından başarılı bulunarak jürimiz tarafından **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

ONAYLAYANLAR:

ONAYLAYANLAR:

Tez Danışmanı:

Doç. Dr. Sıla Övgü Korkut Uysal
İzmir Kâtip Çelebi Üniversitesi

Jüri Üyeleri:

Doç. Dr. Aytuğ Onan
İzmir Kâtip Çelebi Üniversitesi

Dr. Öğr. Üyesi Emin Borandağ
Celal Bayar Üniversitesi

Savunma Tarihi: 16.06.2023

Yazarlık Beyanı

Ben, **Umut Yıldız**, başlığı **Yapay Zekâ ve Makine Öğrenimi Araçları ile Elektrik Tüketimi Tahmini** olan bu tezimin ve tezin içinde sunulan bilgilerin şahsıma ait olduğunu beyan ederim. Ayrıca:

- Bu çalışmanın bütünü veya esası bu üniversitede Yüksek Lisans derecesi elde etmek üzere çalıştığım süre içinde gerçekleştirilmiştir.
- Daha önce bu tezin herhangi bir kısmı başka bir derece veya yeterlik almak üzere bu üniversiteye veya başka bir kuruma sunulduysa bu açık biçimde ifade edilmiştir.
- Başkalarının yayımlanmış çalışmalarına başvurduğum durumlarda bu çalışmalara açık biçimde atıfta bulundum.
- Başkalarının çalışmalarından alıntıladığımda kaynağı her zaman belirttim. Tezin bu alıntılar dışında kalan kısmı tümüyle benim kendi çalışmamdır.
- Kayda değer yardım aldığım bütün kaynaklara teşekkür ettim.
- Tezde başkalarıyla birlikte gerçekleştirilen çalışmalar varsa onların katkısını ve kendi yaptıklarımı tam olarak açıkladım.

Tarih:

16.06.2023

Electricity Consumption Forecasting via Artificial Intelligence and Machine Learning Tools

Abstract

This study focuses on the critical issue of projecting electricity usage for diverse businesses so that they may establish strategies and take the required measures. The dissertation's specific goal is to estimate the electricity consumption of machines on a production line using data obtained from sensors and analyzers, which include thermocouples as well as devices that record electricity usage. The Prophet technique, noted for its ability to capture the core properties of time-series data in predicting models, was used to do this. Furthermore, the thesis investigates the usefulness of the Prophet approach in predicting power usage by taking into account both the utilization and exclusion of temperature data. The findings, which are shown in tables and figures for both univariate and multivariate scenarios, demonstrate the Prophet method's efficacy. A comparison of the RMSE, MAE, and SMAPE scores also shows that removing temperature data leads to better forecasts than having temperature data.

In addition to the Prophet technique, this thesis recommends the use of LSTM and ARIMA models for forecasting power usage. LSTM models are well-known for their ability to recognize important patterns in time-series data, making them an invaluable tool in this context. LSTM models can provide reliable estimates for power usage by exploiting their capacity to capture long-term relationships. ARIMA models, which combine autoregressive, moving average, and differencing components, have a solid foundation in time-series analysis and can efficiently capture trends and seasonality patterns. The inclusion of LSTM and ARIMA models in future studies would provide additional insights into their performance and aid in determining the best technique for forecasting power usage.

Keywords: Electricity consumption, Prophet method, LSTM, ARIMA, Forecasting

Yapay Zekâ ve Makine Öğrenimi Araçları ile Elektrik Tüketimi Tahmini

Öz

Bu çalışma, stratejiler oluşturabilmeleri ve gerekli önlemleri alabilmeleri için çeşitli işletmeler için elektrik kullanımının tahmin edilmesi gibi kritik bir konuya odaklanmaktadır. Çalışmanın özel amacı, elektrik kullanımını kaydeden cihazların yanı sıra termokupl içeren sensörler ve analizörlerden elde edilen verileri kullanarak bir üretim hattındaki makinelerin elektrik tüketimini tahmin etmektir. Bunu yapmak için, tahmin modellerinde zaman serisi verilerinin temel özelliklerini yakalama becerisiyle dikkat çeken Prophet tekniği kullanıldı. Ayrıca, çalışma, sıcaklık verilerinin hem kullanımını hem de hariç tutulmasını dikkate alarak güç kullanımını tahmin etmede Prophet yaklaşımının yararlılığını araştırmaktadır. Hem tek değişkenli hem de çok değişkenli senaryolar için tablo ve şekillerde gösterilen bulgular, Prophet yönteminin etkinliğini göstermektedir. RMSE, MAE ve SMAPE puanlarının karşılaştırılması, sıcaklık verilerinin kaldırılmasının, sıcaklık verilerine sahip olmaktan daha iyi tahminlere yol açtığını da gösterir.

Prophet tekniğine ek olarak, bu çalışma güç kullanımını tahmin etmek için LSTM ve ARIMA modellerinin kullanılmasını önermektedir. LSTM modelleri, zaman serisi verilerindeki önemli kalıpları tanıma yetenekleriyle tanınır ve bu da onları bu bağlamda paha biçilmez bir araç haline getirir. LSTM modelleri, uzun vadeli ilişkileri yakalama kapasitelerinden yararlanarak güç kullanımı için güvenilir tahminler sağlayabilir. Otoregresif, hareketli ortalama ve fark bileşenlerini birleştiren ARIMA modelleri, zaman serisi analizinde sağlam bir temele sahiptir ve trendleri ve mevsimsellik kalıplarını verimli bir şekilde yakalayabilir. LSTM ve ARIMA modellerinin gelecekteki çalışmalara dâhil edilmesi, performansları hakkında ek bilgiler sağlayacak ve güç kullanımını tahmin etmek için en iyi tekniğin belirlenmesine yardımcı olacaktır.

Anahtar Kelimeler: Elektrik Tüketim, Prophet metodu, LSTM, ARIMA, Tahminleme

Teşekkür

Yorucu ancak bir o kadar da yeni şeyler öğrendiğim bu serüvenin sonuna gelmiş bulunuyorum. Öncelikle, saygıdeğer tez danışmanım Sıla Övgü Korkut Uysal'a içtenlikle teşekkürlerimi sunmak istiyorum. Bilgeliliğiniz, uzmanlığınız ve akademik gelişimime olan katkılarınız büyük bir motivasyon kaynağı oldu. Tez sürecindeki rehberliğiniz, yapıcı geri bildirimleriniz ve kararlı desteğiniz benim için paha biçilemez değerinde oldu. Mentorluğunuz ve kişisel gelişimim üzerindeki derin etkiniz için sonsuz minnettarım.

Aileme, eğitim hayatım boyunca yaptıkları fedakarlıklar ve sağladıkları destek için minnettarım. Başarılı bir şekilde ilerleyebilmemde sizlerin inancı ve yardımları büyük önem taşıyor. Sevgili kız arkadaşım Zeynep Özge Yalçın, tüm zorlu anlarda bana güç veren ve beni en iyi şekilde anlayan kişi olduğun için sana minnettarım. Tez sürecini bu denli başarılı bir şekilde tamamlayabilmemde senin desteğin olmadan mümkün olmazdı.

Klemsan Elektrik Elektronik A.Ş.'ye özel olarak sayın müdürüm Serkan Kaba aracılığıyla tez sürecindeki katkılarından dolayı teşekkürlerimi sunuyorum.

İçindekiler

Yazarlık Beyanı	ii
Öz	iii
Abstract	iv
Teşekkür	vi
Şekiller Listesi	x
Tablolar Listesi	xi
Kısaltmalar Listesi	xii
Semboller Listesi	xiii
1 Giriş	1
1.1 Literatür Özeti	2
1.2 Tezin Ana Hatları.....	7
2 Materyal ve Metotlar.....	9
2.1 Veri Toplama.....	9
2.2 Metotlar	12
2.2.1 Uzun Kısa Süreli Bellek (LSTM).....	13
2.2.2 Otoregresif Entegre Hareketli Ortalama (ARIMA).....	16
2.2.3 Prophet Metot	18
3 Analiz Sonuçları ve Tartışmalar	22
3.1 Veri Analizi	22
3.2 Zaman Serisi Analizi	24
3.3 Modellerin Karşılaştırması	28
4 Sonuç.....	36
Kaynaklar	38
Özgeçmiş	43

Şekiller Listesi

Şekil 2.1	Kullanılan Termokuplların bir örneği	10
Şekil 2.2	Termokuplların makineye yerleştirilmesini gösteren resim.....	11
Şekil 2.3	Veri toplamanın akış şeması	12
Şekil 2.4	t zamanında standart bir LSTM modelinin şematik gösterimi (t anından $t+1$ anıda kadar olan bir adımlık sürecin gösterimidir).....	14
Şekil 3.1	İşlenmemiş veri kümesinin tamamlayıcı istatistikleri	23
Şekil 3.2	Ön işleme yapılmış kümesinin tamamlayıcı istatistikleri	23
Şekil 3.3	Veri kümesindeki özelliklerin zamana göre çizgisel grafikleri.....	25
Şekil 3.4	Veri kümesinin trend, periyodiklik, gürültü analizini gösteren şekil.....	25
Şekil 3.5	Veri kümesinin korelasyonunu gösteren şekil	27
Şekil 3.6	Çok değişkenli ARIMA modeli sonucunun şekli	28
Şekil 3.7	Tek değişkenli ARIMA modeli sonucunun şekli.....	30
Şekil 3.8	Çok değişkenli LSTM model sonucunun şekli	31
Şekil 3.9	Tek değişkenli LSTM sonucunun şekli.....	33
Şekil 3.10	Çok değişkenli Prophet model sonucunun şekli (nokta-gerçek değer, çizgi-tahmin değerleri)	34
Şekil 3.11	Tek değişkenli Prophet model sonucunun şekli (nokta-gerçek değer, çizgi-tahmin değerleri)	36

Tablolar Listesi

Tablo 3.1	Çok değişkenli ARIMA modelinin saatlik SMAPE ve MAE hata oranları karşılaştırılması	29
Tablo 3.2	Tek değişkenli ARIMA modelinin saatlik SMAPE ve MAE hata oranları karşılaştırılması	30
Tablo 3.3	Çok değişkenli LSTM modelinin saatlik SMAPE ve MAE hata oranları karşılaştırılması	31
Tablo 3.4	Tek değişkenli LSTM modelinin saatlik SMAPE ve MAE hata oranları karşılaştırılması	32
Tablo 3.5	Çok değişkenli Prophet modelinin saatlik SMAPE ve MAE hata oranları karşılaştırması.....	33
Tablo 3.6	Çok değişkenli Prophet modelinin saatlik SMAPE ve MAE hata oranları karşılaştırması.....	33

Algoritmalar Listesi

Algoritma 3.1 ARIMA'daki optimal değerlerin belirlenmesi için sözde kod 17



Kısaltmalar Listesi

ARIMA	Otoregresif entegre hareketli ortalama (Autoregressive Integrated Moving Average)
LSTM	Uzun Kısa Süreli Bellek (Long Short-Term Memory)
WMA	Ağırlıklı hareketli ortalama (Weighted Moving Average)
SMA	Basit hareketli ortalama (Simple Moving Average)
HL	Holt lineer trend (Holt linear trend)
SES	Basit üstel yumuşatma (Simple Exponential Smoothing)
HW	Holt-Winters
CMA	Merkezli hareketli ortalama (Centered Moving Average)
SARIMA	Sezonsal Otoregresif Entegre Hareketli Ortalama (Seasonal Autoregressive Integrated Moving Average)
RNN	Tekrarlayan sinir ağı (Recurrent Neural Network)
RMLP	Tekrarlayan çok katmanlı algılayıcı ağı (Recurrent Multilayer Perceptron Network)
ARMA	Otoregresif hareketli ortalamalar modeli (Autoregressive–moving-average)
TBATS	Trigonometrik Box–Cox dönüşümü, ARMA hataları, Trend ve Mevsimsel bileşenler (Trigonometric Box–Cox transform, ARMA errors, Trend, and Seasonal components)
MLP	Çok katmanlı algılayıcı (Multilayer Perceptron)
RMLP	çok katmanlı bir algılayıcı ağının (Recurrent Multilayer Perceptron Network)
NNAR	sinir ağı otoregresyon (Neural Network Autoregression)

BPNN	Geri yayılım sinir ağı (Backpropagation Neural Network)
GBR	Gradyan artırma regresyonu (Gradient Boosting Regression)
RBF	Radyal tabanlı fonksiyon (Radial Basis Function)
SVR	Destek vektör regresyonu (Support Vector Regression)
RF	Rasgele orman regresyonu Random Forest regression
KNN	K-En Yakın Komşu Regresyonu (K-Nearest Neighbor regression)
GRU	Geçitli tekrarlayan birim (Gated Recurrent Unit)
ANN	Yapay sinir ağları (Artificial Neural Networks)
CNN	Evrişimli sinir ağları (Convolutional Neural Networks)
ANFIS	Uyarlanabilir ağ tabanlı bulanık çıkarım sistemi (Adaptive Network-based Fuzzy Inference System)
DOPH	Doğrudan optimum paralel hibrit (Direct Optimum Parallel Hybrid)
EMD	Ampirik Mod Ayırıklaştırma (Empirical Mode Decomposition)
ACKF	Uyarlanabilir Cubature Kalman Filtreleme (Adaptive Cubature Kalman Filtering)
FTS	Fuzzy zaman serisi (Fuzzy Time Series)
SWT	Durağan dalgacık dönüşümü (Stationary Wavelet Transform)
SMAPE	Simetrik ortalama mutlak yüzde hatası (Symmetric Mean Absolute Percentage Error)
MAPE	Ortalama mutlak yüzde hatası Mean Absolute Percentage Error
MAE	Ortalama mutlak hata (Mean Absolute Error)
MSE	Ortalama hataların karesi (Mean Square Error)
RMSE	Kök Ortalama Kare Hatası (Root Mean Square Error)

Semboller Listesi

∇	Fark operatörü
B	Geri fark operatörü
ϵ, ϵ_t	İndirgenemez hata terimi
C_{t-1}	Bir önceki zamana (t-1) ait hücre bilgisi
H_{t-1}	Bir önceki zamana (t-1) ait geçerli gizli katman
x_t	t. anda sisteme giren yeni bilgi
$\omega_i, \omega_o, \omega_f, \omega_c$	t anındaki x_t 'nin ağırlık katsayısı
f_t	LSTM metodu Unut Kapısı (Forget Gate)'nin çıkış fonksiyonu
U_c, U_i, U_o, U_f	H_{t-1} 'ın ağırlık katsayısı
b_f, b_c, b_i, b_o	İlgili kapı fonksiyonlarının yanlılık terimidir.
i_t	LSTM metodu Giriş Kapısı (Input Gate)'nin çıkış fonksiyonu
$\hat{C}_t$	LSTM metodunda
γ, p, d, q	ARIMA modelinde kullanılan sabitler
$y_t^{(d)}$	ARIMA modelinde d times differenced series
$y(t)$	Prophet modelinde t anında tahmini değer
$s(t)$	Prophet modelinde tekrar eden paternleri takip eden sezonsal bileşen
$g(t)$	Prophet modelinde Trend bileşeni
$h(t)$	Prophet modelinde, tatil bileşeni (istisnai olayların model katılması amacıyla)
k_{adj}, m_{adj}	Ayarlanmış artış oranı, ayarlanmış telafi parametresi
$X(t)$	Fourier serisi terimleri
$\beta \gamma \delta$	Katsayı vektörleri

Bölüm 1

Giriş

Elektrik tüketimi, imalat ve perakende satıştan konaklama ve sağlık hizmetlerine kadar pek çok işletme için esastır. Elektrik depolama ve üretim maliyetlerinin pratik olmaması, talep tahminin ve elektrik kullanımının doğru yapılmasını birçok endüstri için çok önemli kılar. Firmalar bunu yaparak enerji kullanımlarını kontrol edebilir, giderleri azaltabilir ve sürdürülebilirliği teşvik edebilir.

Makine öğrenimi algoritmaları, elektrik talebini tahmin etmek için mükemmel araçlardır. Bu bağlamda, zaman serisi verilerindeki karmaşık kalıpları yakalarken eş zamanlı olarak belirsizlik tahminleri oluşturma yetenekleri, onları işletmeler için cazip hale getirir. Enerji kullanımını tahmin etmek için bu araçları kullanmanın bazı sınırlamaları olsa da birçok firma için faydaları dezavantajlarına ağır basar. Ayrıca, elektrik tüketimini tahmin etmek için herhangi bir teknoloji geliştirmek veya kullanmak, endüstride hala ilgi çekicidir.

Bu tez, endüstrideki elektrik tüketimini tahmin etmeye yönelik bazı önemli bilgiler sunmayı amaçlamıştır. Anlaşılır olması açısından, klemensler, kablo tavaları ve enerji analizörleri üreten bir fabrikadan alınan veriler dikkate alınmıştır. Amaç olarak iki görev merkezlenmiştir: 1) elektrik tüketiminin tahmininde çeşitli makine öğrenimi ve yapay zekâ araçlarının kullanılması, 2) farklı kullanımla veya kullanılmadan sonuç tahmini üzerindeki etkilerin incelenmesi. Daha açık bir şekilde, bu tezde doğru elektrik tüketimi tahmini oluşturmak için endüstri örneğinde üç farklı metodoloji tartışılmıştır. Ayrıca, sıcaklığın elektrik tüketimini tahmin etme üzerindeki etkisinin araştırılmasına da odaklanılmıştır.

1.1 Literatür Özeti

Literatürde sadece endüstride değil, bireysel kullanım için de elektrik tüketimini tahmin etmek için meta-sezgisel, yapay zekâ, istatistiksel ve hibrit yaklaşımlar dahil olmak üzere çeşitli stratejiler uygulanmaktadır. Bu tezin ana metodolojileri Uzun Kısa Süreli Bellek (LSTM), Otoregresif Entegre Hareketli Ortalama (ARIMA) ve Facebook'un Prophet Modeli olsa da daha kapsamlı ve güncel bir inceleme sağlamak için çok önemlidir, son yıllarda ise çeşitli stratejiler olabildiğince üretilmeye çalışılmıştır.

Çeşitli stratejiler kullanılarak elektrik tüketiminin tahmini üzerine araştırmalar yapılmıştır. Güncel istatistiksel yaklaşımlar şu şekilde örneklendirilebilir: Lee ve diğ. [1] çalışmalarında, Malezya Tun Hussein Onn Üniversitesi'nin (UTHM) gelecekteki elektrik tüketimini tahmin etmek için, ağırlıklı hareketli ortalama (WMA), basit hareketli ortalama (SMA), Holt lineer trend (HL), basit üstel yumuşatma (SES), Holt-Winters (HW) ve merkezli hareketli ortalama (CMA) metotlarının içeren altı farklı metodun karşılaştırılmasına odaklanmışlardır. [1]'de araştırmacılar, ortalama mutlak hata (MAE), ortalama mutlak hata yüzdesi (MAPE), ortalama karesel hata (MSE) ve ortalama karesel hatanın karekökü (RMSE) gibi metrikleri dikkate alarak ve tahminlerin verilerle uyumlu olup olmadığını kontrol ederek HW'nin diğer beş yöntemden daha iyi performans gösterdiğini belirtmiştir. Durağan olmayan verileri durağan hale dönüştürme yeteneklerinden yararlanarak, Sezonlar Otoregresif Entegre Hareketli Ortalama (SARIMA) ve ARIMA yöntemleri ise Kaur & Ahuja [2] tarafından, çeşitli sağlık binalarının gelecekteki elektrik tüketimini tahmin etmek için değerlendirilmiştir. Yürütülen karşılaştırmalı analiz dikkate alındığında, Kaur ve Ahuja [2], SARIMA'nın elektrik tüketimini tahmin etmede ARIMA yönteminden daha iyi olduğu sonucuna varmışlardır. Literatürde, zaman serilerini tahmin etmek için istatistiksel yöntemleri ve makine öğrenimi modellerini karşılaştırmaya odaklanan bir dizi çalışma mevcuttur. Örneğin, Makridakis ve diğ. [3] LSTM, tekrarlayan sinir ağları (RNN) ve sekiz istatistiksel yaklaşım içeren on makine öğrenmesi modellerini hesaplama gereksinimleri ve model doğruluklarını açısından performanslarını karşılaştırmışlardır. Bunun amaçla, M3-yarışmalarında [4] kullanılan 3003 veri kümelerinin 1045 örneği üzerine çalışılmıştır. Araştırmacılar, makine öğrenimi yöntemlerinin doğruluğunun istatistiksel yöntemlerden daha iyi

performans gösterdiğini [3]'te ortaya koymuşlardır. Bununla birlikte, makine öğrenimi modellerinin hesaplama karmaşıklığının, istatistiksel yöntemlerden önemli ölçüde büyük olduğu sonucuna varılmıştır. [3]'te sunulan çalışmayı Kalimoldayev ve diğ. [5] tarafından yürütülen bir başka kapsamlı karşılaştırmalı çalışma takip eder. Araştırmacılar, panel verilerin elektrik tüketimini tahmin etmek için istatistiksel yöntemlerin ve makine öğrenimi yöntemlerinin doğruluk performanslarını tek değişkenli ve çok değişkenli modeller için tartışmışlardır. Çalışmada ayrıca faktörlerin önemini de belirlemiştir. Kalimoldayev ve diğ. [5], belirlenen dört girdi özellikleriyle, tekrarlayan çok katmanlı bir algılayıcı ağının (RMLP) modelinin en doğru tahmin olduğu vurgulanmıştır. Kaydedilen doğruluk sonuçlarına göre, bu modeli sırasıyla, Gradyan artırıcı ağaç modeli (girdi değişkenleri tarih, saat, haftanın günü, ay, çeyrek, yıl, haftanın günü, ayın günü ve yılın haftası gibi faktörleri kullanarak); parçalı lineer regresyon (hava sıcaklığı (segmentasyon için kullanılan), haftanın günü, günün saati gibi dış değişkenler kullanarak); parametrik olmayan çoklu sezonsal ayrıştırma modeli (günlük, haftalık ve yıllık mevsimsellik ile); ve ARIMA model (günlük, haftalık ve yıllık mevsimsel ayrıştırma modeli ve hava sıcaklığı, haftanın günü, günün saati gibi değişkenlerle) takip ettiği belirtilmiştir, [5]. Ayrıca, Leite Coelho da Silva ve diğ. [6], Brezilya endüstrisindeki elektrik tüketimini tahmin etmeyi amaçlamıştır. Bunu amaçla, araştırmacılar, sinir ağı otoregresyonunu (NNAR) ve çok katmanlı algılayıcıyı (MLP) içeren nöral verileri içeren çeşitli yapay zekâ yöntemlerinin yanı sıra SARIMA, HW, Trigonometric Box-Cox dönüşümü; ARMA hataları, Trend ve Mevsimsel bileşenler (TBATS), dinamik doğrusal modeli gibi çeşitli istatistiksel yaklaşımları değerlendirilmiştir. [6]'da yapılan analize göre MLP, MAPE ölçeğini dikkate alarak diğer yöntemlere göre en iyi sonuçları elde ettiği belirtilmiştir.

Mevcut araştırmalar, elektrik tüketimini tahmin etmek için makine öğrenimi modellerinin ve hibritleştirilmiş modellerin kullanımına gösterilen ilgiyi de vurgulamaktadır. Bu çalışmalara birkaç örnek olarak şunlar verilebilir: Wahid ve diğ. [7], konut dairelerinin enerji tüketimini tahmin etmek için iki farklı sınıflandırma metodu kullanmıştır: Çok Katmanlı Algılayıcılar ve Rastgele Orman. [7]'nin kaydedilen model doğruluklarına ve Kappa istatistik sonuçları Çok Katmanlı Algılayıcılar'ın Rastgele Orman modeline göre daha üstün olduğunu göstermiştir. Zhong ve diğ., [8]'de Çin'deki bir kıyı kasabasındaki bir konut binasının enerji

tüketimini tahmin etmek için yeni bir vektör alan tabanlı destek vektörü regresyon modeli önermiştir. [8] çalışmasında önerilen modelin temel fikri, girdiler ve çıktılar arasındaki yüksek doğrusal olmayan ilişkiyi, doğruluk, sağlamlık ve genelleme yeteneğini artırmak için optimal özellik uzayını belirleyerek doğrusal ilişki formuna dönüştürmektir. Kaydedilen doğruluk, yetenek ve sağlamlık değerlendirmelerine göre, önerilen yöntemin elde edilen sonuçlarının; geri yayılım sinir ağı (BPNN) modeli, çoklu doğrusal regresyon modelleri, gradyan artırıcı regresyon (GBR), Aşırı Derin Öğrenme ve destek vektör regresyonu (SVR) modeli (hem radyal baz fonksiyon (RBF) çekirdekli hem de sigmoid çekirdekli) gibi diğer yöntemlerin sonuçlarına göre daha iyi olduğu görülmüştür. Ayrıca, Peng ve diğ. [9] Rastgele Orman Regresyonu (RF), GBR, BPNN, Lineer Regresyon ve K-En Yakın Komşu Regresyonu (KNN) olan beş makine öğrenimi modeli aracılığıyla çeşitli özellikleri kullanarak limandaki gemilerin enerji tüketimini tahmin etmeye çalışmıştır. [9]'de MSE, MAE, R^2 ve açıklanan varyans puanı dâhil olmak üzere alınan metriklere göre GBR ve RF diğer yöntemlerden daha iyi performans göstermiştir. Fan ve diğ. [10] tarafından Ocak 2007 ve Şubat 2007 için Avustralya elektrik tüketimi verileri için “gray catastrophe” (GC (1,1))’lu SVR ile RF tarafından birleştiren hibrit bir model (SVR-GC-RF) önerilmiştir. Fan ve diğ., önerilen SVR-GC-R’nin sonuçlarını SVR, RF ve RF ile birleştirilmiş SVR (SVR-RF), ve ayrıca, literatürde ARIMA, LSTM ve SVR içeren hibridize edilmiş çeşitli çalışmaların sonuçları ile karşılaştırmışlardır. SVR-GC-RF’nin elde edilen MAPE, MAE ve kök ortalama kareler hatası (RMSE), hem Ocak 2007 hem de Şubat 2007 veri setlerinde en düşük hataları sağlamıştır.

Ribeiro ve diğ. [11] tarafından yapılan çalışmada, Brezilya’daki bir termoplastik reçine üretim tesisinin enerji tüketiminin tahminine ilişkin makine öğrenimi ve derin öğrenme yöntemlerinin bir karşılaştırması sunulmuştur. Bu karşılaştırmaya kapsamında, Ribeiro ve diğ. [11], ARIMA, RNN, LSTM, geçitli tekrarlayan birim (GRU), SVR ve RF gibi farklı yöntemler dikkate alınmıştır. Hem grid araması hem de Diebold-Mariano testi sonuçları ışığında, [11]’de GRU’nun diğer tüm modelleri geride bıraktığı sonucuna varılmıştır. Bu çalışmalara ek olarak, derin öğrenme modelleri ve yapay zekâ modelleri aracılığıyla enerji tüketimini tahmin etmek için de literatürde çeşitli girişimlerde bulunulmuştur. Güncel araştırmaları özetlemeye geçmeden önce, elektrik tüketim tahmininde makine öğrenimi ve derin öğrenme modellerine daha derinlemesine bir inceleme yapmak isteyenler ilgili okuyucular [12-14] çalışmalarına

ve bu çalışmalardaki referanslara başvurulabilir. Ayrıca, Seyedzadeh ve diğ. [15] tarafından çeşitli makine öğrenimi metodolojilerinin tahmin ve geliştirme üzerine yapılan kapsamlı derlemeye de göz atılabilir. Belirtilen derlemeleri takiben, son yıllarda yapılan güncel çalışmalardan bazıları şu şekilde örneklendirilebilir: Koprinska ve diğ. [16] hem güneş enerjisi tahmini hem de elektrik yükü tahmini için Evrişimli Sinir Ağlarını (CNN) üzerine çalışmıştır. Yöntem, Çok Katmanlı Algılayıcı Sinir Ağları (MLP NN) ve LSTM metotları ile dört veri seti üzerinde bir kalıcılık temeli test edilerek karşılaştırılmıştır. RMSE ve MAE skorları dikkate alındığında, CNN ve MLP NN'nin birbirleri ile yakın sonuçlar verdiği ve hesaplama maliyeti açısından bu yöntemlerin tercih edilebileceği sonucuna varılmıştır. [17]'de Ramos ve diğ. farklı veri setleri ve günün farklı zaman dilimleri için çeşitli öğrenme parametrelerini dikkate alarak, endüstrideki elektrik tüketimini tahmin etmek için ANN ve SVM metotlarını kullanmış ve metotların performans özellikleri karşılaştırmıştır. SMAPE, MAPE, RMSE ve MAE değerlerini hesaba katarak, model performanslarının günün periyodlarına, tüketim kalıpları için farklı ayarlara, tahmin algoritmasının farklı özelliklerine göre değişkenlik gösterdiği gözlemlenmiştir. Böylece, Ramos ve diğ. [17]'de iki metodun da algoritmanın özelliklerine veya günün periyod seçimlerine göre birbirlerine göre tercih edilebilir olduğu sonucuna varmışlardır. Chahkoutahi ve diğerleri tarafından MLP NN, Uyarlanmış network-tabanlı bulanık çıkarım sistemi (ANFIS) ve SARIMA'ya dayanan hibritleştirilmiş bir model, doğrudan optimum paralel hibrit (DOPH) modeli, elektrik yükü tahmini için [18]'de önerilmiştir. Elektrik talebi tahmin modellerinin doğruluğunu artırmak için Ampirik Mod Ayırıklaştırma (EMD) ve LSTM ağlarına dayanan başka bir hibrit yöntem, Bedi ve Toshniwal tarafından [19]'da önerilmiştir. RMSE ve mutlak yüzde hata skorları, [19] çalışmasında önerilen hibrit modelin performansının LSTM, RNN ve RNN ile hibritleştirilmiş EMD'den daha iyi olduğunu göstermiştir. Yazarlar bir sonraki çalışmaları olan [20]'de ise derin öğrenme yapılarını uzun vadeli tarihsel bağımlılıklarla kullanarak yeni bir çerçeve önermiştir. [20]'de önerilen model, Hindistan'ın Chandigarh Birlik Toprakları'nın elektrik talebini tahmin etmeye yönelik olarak ANN, RNN ve SVR ile karşılaştırılmıştır. Bu çalışmada, [20], önerilen modelin RMSE ve MAPE metriklerinin diğer metotlara kıyasla minimum hata skorlarını kaydettiği belirtilmiştir. Jain ve diğ. [21], Uttar Pradesh'in gelecekteki elektrik talebini tahmin etmek için K-ortalamalar kümeleme ve bulanık zaman serileri (FTS) temelli KmFuzz adlı yeni bir yöntem geliştirmiştir. [21]'de KmFuzz

modeli, RMSE ve MAPE skorlarına göre XGboost, ANN, ARIMA, LSTM ve çeşitli FTS yöntemlerine kıyasla daha iyi sonuçlar verdiği belirtilmiştir. Guo ve diğ. endüstriyel güç tüketicilerinin maksimum güç talebini tahmin etmeyi [22]'de araştırmıştır. Bunun için müşterilerin maksimum güç taleplerinin karakteristik özelliklerine dayanarak Facebook Prophet yöntemi ve uyarlanmış Cubature Kalman filtreleme (ACKF) yönteminin kombinasyonuna dayanan geliştirilmiş bir maksimum güç talep tahmin algoritması önerilmiştir. [22] çalışmasında ortaya koyulan deneysel sonuçlar, önerilen modelin iyi performans sergilediğini göstermiştir. Yukarıda sunulan çalışmalar ek olarak, daha yakın tarihli çalışmalar is şu şekilde örneklendirilebilir: Amantegui ve diğ. [23], bir endüstriyel mutfağın elektrik tüketimini tahmin etmede Prophet yöntemi, RF ve LSTM'nin kullanımını değerlendiren karşılaştırmalı bir çalışma yürütmüştür. Bu bağlamda, Portekiz'deki bir restorandan elde edilen bir aylık veri seti üzerinde uygulanan sonuçlara göre, [23]'de Prophet yöntemi ve RF iyi bir seçenek olduğu belirtilirken, Prophet modeli kolay kullanımı nedeniyle RF'e tercih edilebilir olduğu da vurgulanmıştır. Ayrıca Amantegui ve diğ. [23]'deki çalışmalarında LSTM'nin daha büyük veri setleri için tercih edilebilir bir yöntem olabileceğini de belirtmişlerdir. Zhou ve diğ. [24] tarafından yürütülen çalışmada LSTM modelleri ve pekiştirmeli öğrenme ajanları kullanarak gelecek günün elektrik tüketimini ve pik elektrik talebini tahmin etmek için yeni bir teknik önerilmiştir. Bashir ve diğ. [25] tarafından ortaya konulan bir başka çalışmada ise kısa vadeli elektrik yükünü tahmin etmek için Prophet ve LSTM modellerine dayalı bir hibrit strateji önerilmiştir. [25]'de sunulan çalışmada, önerilen metodun, literatürde çok iyi bilinen LSTM, ARIMA gibi bağımsız modellere kıyasla daha iyi sonuçlar verdiği belirtilmiştir. Ayrıca, Yan ve diğ. [26] kesikli durağan dalgacık dönüşümü (SWT) ile birleştirilen karmaşık çok kanallı bir ikili gömeli LSTM çerçevesi (MC-BiNLSTM) kullanarak, hassas ve verimli bir enerji tüketimi tahmini üretmek amacıyla bir metod önerilmiştir. Gerçek bir IoT enerji tüketimi veri seti üzerinde RMSE, MAPE, MAE ve R^2 metrikleri dikkate alınarak yürütülen kapsamlı bir karşılaştırmalı analiz sonuçları, [26], önerilen yöntemin tüm karşılaştırılan yöntemlerden üstün olduğunu göstermiştir. Bu çalışmalara ek olarak, Chaturvedi ve diğ. [27], Hindistan'ın Merkezi Enerji Otoritesi (CEA) algoritmasını mevcut trend-tabanlı modellemesini SARIMA, LSTM RNN ve Facebook Prophet ile ele almıştır. RMSE ve MAPE skorları göz önünde bulundurularak yapılan karşılaştırmalar doğrultusunda, Facebook Prophet modeli ve CEA modelinin

SARIMA ve LSTM RNN'i geride bıraktığı sonucuna varılmıştır. Son olarak, uzun ve orta vadeli elektrik tüketimini tahmin etmek için, çıkarmalı kümeleme (SC), ızgara bölme (GP) ve regresyon denklemi senaryoları (RES) ile kohort zekasına dayalı uyarlanabilir sinirsel bulanık çıkarım sistemlerini (CI-ANFIS) kullanan yeni bir çerçeve Tutun ve diğ. Tarafından [28]'de önerilmiştir. Türkiye'deki bir durum çalışmasında gerçek veri setine uygulanan aylık ve yıllık net elektrik tüketimini tahmin etmek için önerilen çerçevenin [28], aynı veri kümesi üzerinde yapılan literatürdeki mevcut yöntemleri geride bıraktığı belirtilmiştir.

1.2 Tezin Anahatları

Enerji tüketimini tahmin etmek için kapsamlı araştırmalar yapılmış olmasına rağmen, sonuçlar iyi bir tahmin gerçekleştirmenin tek bir yöntemi olmadığını göstermektedir. Yukarıda belirtilen tüm çalışmalar, enerji tüketiminin hala ilgi çekici konu olmaya devam ettiğine dair öngörünün kanıtıdır. Dolayısıyla, enerji tüketimini tahmin etme konusunda daha fazla araştırma ve tartışma yapılmasına hâlâ ihtiyaç vardır. Bu tezin iki ana amacı vardır: biri sektördeki bir şirketin gelecekteki elektrik tüketimi için tek değişkenli ve çok değişkenli seçenekleri dikkate alarak bir tahminde bulunmak, diğeri de kullanılan yöntemlerin, yani ARIMA yöntemi, LSTM ve Prophet modeli sonuçlarını karşılaştırmaktır. Bu yöntemlerin popülaritesinin yanı sıra, seçilen modeller farklı stratejilere karşılık gelmektedir.

Bu tez dört temalı bölümden oluşmaktadır: Giriş bölümü, güncel bir literatür taramasına ışık tutmakta, ardından Bölüm 2'de araştırmanın temel ve teorik boyutları ortaya konmaktadır. Veri toplama ve analizine ek olarak, Bölüm 2'de sunulan dikkate alınan yöntemlerin metodolojileri. Bölüm 2'nin genel hatları şu şekilde özetlenebilir: veri toplamak için montaj hattındaki bir endüstriyel makine örnek olarak alınmıştır. Yüksek elektrik tüketimine yol açan nedensel bir faktör, makinedeki sıcaklıkla ilişkilendirilebilir. Bu nedenle, makinenin olası önemli yerlerine dört sensör yerleştirilmiştir. Ayrıca, ortam sıcaklığı da kaydedilmiştir. Toplanan veriler, verilerin doğasına uygun olarak temizlenmiş ve manipüle edilmiştir. Ayrıca Bölüm 2'de tahmin yapma metodolojileri (ARIMA, LSTM ve Prophet metodu) ve bunların matematiksel arka planı da açıklanmaya çalışılmıştır. Bu bölümü, araştırmanın

bulgularını sunmak için Bölüm 3 izlemektedir. Betimleyici veri analizinin sonuçları sunulduktan sonra, önceden işlenmiş veriler, tek değişkenli ve çok değişkenli zaman serileri olmak üzere iki araştırma kolu dikkate alınarak incelenmiştir. Bunu gerçekleştirmek için öncelikle zaman etiketleri kullanılarak ARIMA, LSTM ve Prophet yöntemi kullanılarak makinenin elektrik tüketimi tahmin edilmiştir. Ardından sıcaklığın elektrik tüketiminde önemli bir faktör olabileceği düşünülerek ARIMA, LSTM ve Prophet yöntemi ile çok değişkenli bir analiz yapılmıştır. Bölüm 3 boyunca hem tek değişkenli hem de çok değişkenli durumlar için karşılaştırmalı bir çalışma yapılmıştır. Karşılaştırma, birkaç metrik tarafından desteklenmiştir. Bölüm 4'te tez; sonuçları tartışarak, yöntemlerin kullanımında karşılaşılan artıları ve eksileri vurgulayarak ve gelecekteki çalışmaların olası yollarına işaret ederek sonuçlandırılmıştır.

Bölüm 2

Materyal ve Metotlar

2.1 Veri Toplama

Veri toplama, araştırma hipotezlerini yanıtlamak, bir model için doğru içgörüler oluşturmak ve olası sonuçları elde etmek için çeşitli kaynaklardan veri toplama ve analiz etme prosedürüdür. Araştırmanın herhangi bir hipotezi, toplanan verilere dayanarak değerlendirilebilir. Bu nedenle veri toplama, araştırma alanından bağımsız olarak veri analizinin en önemli ve en değerli kısımlarından biridir. Veri toplamak için standart onaylanmış teknikler kullanılsa da, gereksinimlere bağlı olarak veri toplama yaklaşımı alandan alana değişebilir. Bu tezde enjeksiyon hattı makinelerinden, sensörler aracılığıyla veriler toplanmıştır. Termokupllar 1000 Santigrat dereceye kadar sıcaklıklara dayanabilir ve sıcaklık ölçebilir. Makinenin belirli noktalarına termokupllar yerleştirilmiş ve bu noktalardan belirli periyotlarda sıcaklık verileri okunmuştur. Aynı zamanda 15 dakikalık elektrik tüketim verileri de aynı veri tabanında saklanmaktadır. Bu verilerin saklanması için MS-SQL ve MongoDB veritabanları kullanılmıştır.

MSSQL sunucusu, verileri depolamak için kullanılan bir veritabanı yönetim sistemidir. Kullanılan tablolar arasında ilişki sağlar. Bu nedenle daha karmaşık verilerin birbirleri ile ilişki kurarak basit bir şekilde saklanmasını sağlar.

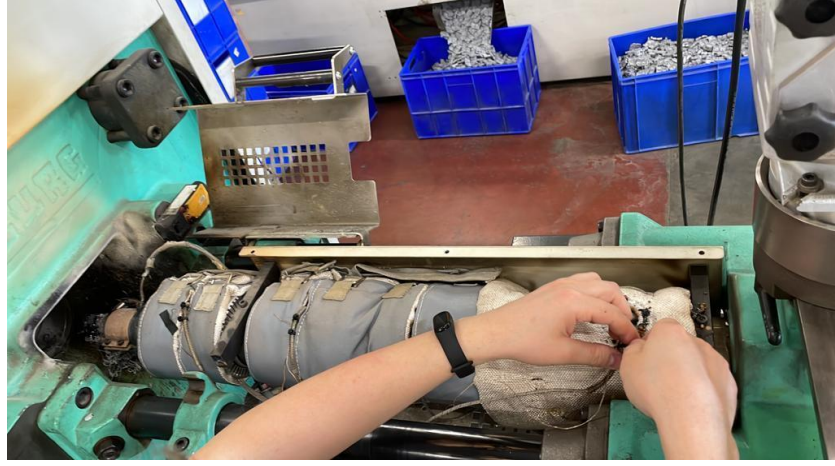
MongoDB, açık kaynaklı bir NoSQL veritabanı uygulamasıdır. MongoDB'de her kayıt bir belge olarak ifade edilir. Ve bu belgeler JSON formatında saklanır. Bu tezin amacı, üretim hattındaki değişen sıcaklıklara göre makinelerin elektrik tüketimini analiz etmektir. Bu nedenle üretim hattındaki makinelerin tüketimlerini analiz edecek bir sistem kurulması öngörülmektedir. Bu nedenle mantolama yapılan bir makinenin sıcaklık değerlerinin hesaplanması amaçlanmaktadır. Makinenin baskı bölümünün

ön, orta ve üst taraflarına yüksek sıcaklıklarda ölçüm yapan termokupllar monte edilmiştir. Bu termokupllar 1000 dereceye kadar ölçüm yapmak için kullanılır. Bunların dışında ortam sıcaklığını ölçmek için de sensör kullanılmıştır. Kullanılan termokupllar Şekil 2.1'de gösterilmiştir.



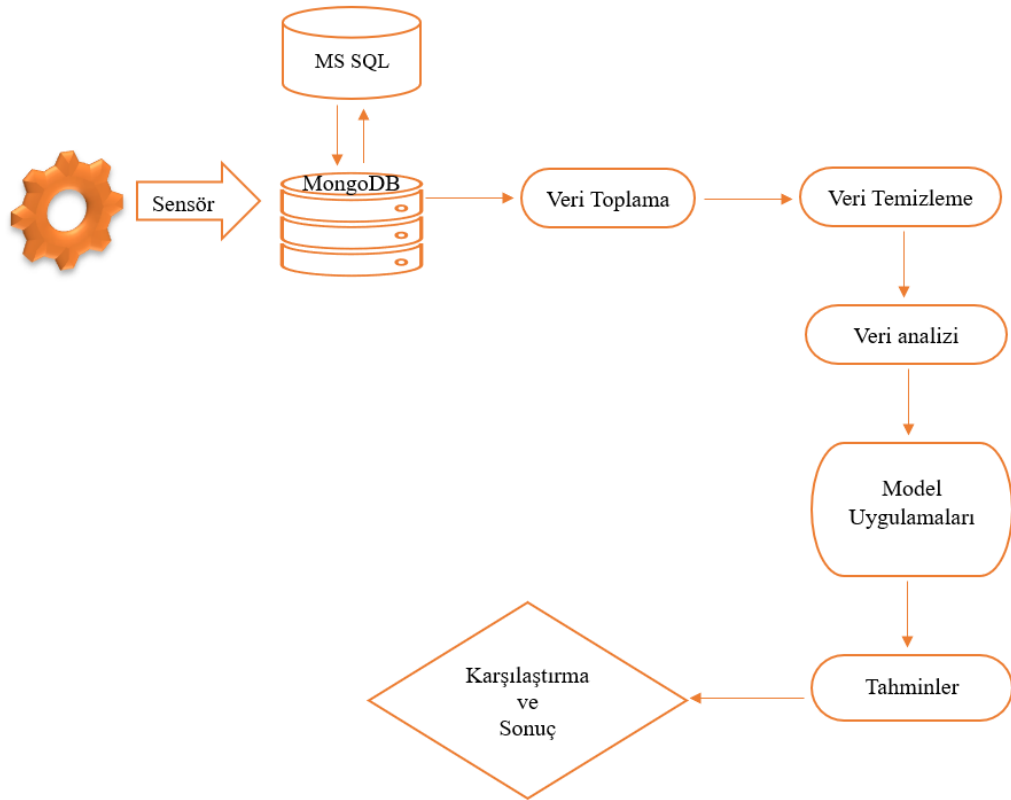
Şekil 2.1: Kullanılan Termokuplların bir örneği

Şekil 2.1'in yanı sıra Şekil 2.2'de, termokuplların makine üzerinde belirtilen yerlere nasıl yerleştirildiğini gösterilmiştir.



Şekil 2.2: Termokuplların makineye yerleştirilmesini gösteren resim

Termokuplları makineye yerleştirdikten sonra, termokupllardan toplanan veriler Wifi To RS-485 (WTOR) aracılığıyla iletilerek sistemin hizmetlerini paylaşmasına olanak sağlandı. WTOR, kablosuz iletişimi sağlayan bir ara elemandır. Kullandığı servisleri kendi üzerinde tanımlı Modbus tablosu üzerinden haberleşerek karşılar. Sistemin servisleri her 15 dakikada bir termokupllar ile haberleşerek tüketim verilerini veri tabanına yazmaktadır. Bir veritabanı olarak, cihaz türleri, özellikler ve yapılandırma ayarları gibi bilgiler MS-SQL sunucusunda saklanır. Ayrıca makinelerin yaptığı tüketim bilgileri de MongoDB'de tutulmaktadır. Burada, MongoDB verileri aynı anda dakika ve saniye cinsinden kaydetmediği için veri setinde MongoDB'den alınan sıcaklık verilerinin ortalama değerleri kullanılmaktadır. Aynı şekilde tüketim verilerinin toplam değeri kullanıldığını belirtmekte fayda vardır.



Şekil 2.3: Tez kapsamında yapılan işlemlerin akış şeması

Gösterilen şekil, veri toplama sürecinden başlayarak Bölüm 2 ve Bölüm 3'te yürütülen işlemler için basit akış şemasını sunar. Şekil 2.3'te sunulan veri toplama sürecinde, MS-SQL'den alınan belirli sorgularla meşgul olan MongoDB'de depolanan veriler, erişilebilirliği ve anlaşılabilirliği nedeniyle Jupyter dosyasına aktarılarak geri kalan işler; veri analizi, model uygulamaları, tahminler, tamamen Jupyter Notebook arayüzü vasıtasıyla yürütülmüştür. Yürütülen tüm analizler, Bölüm 3'te detaylı bir şekilde anlatılmıştır.

2.2 Yöntemler

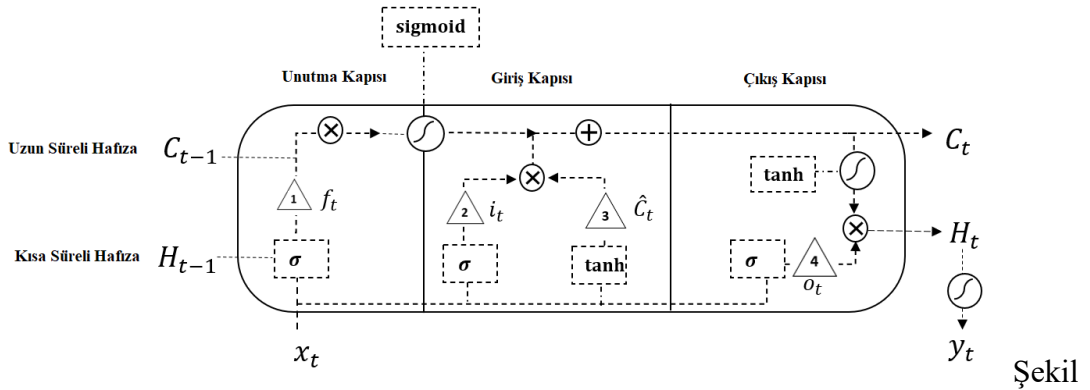
Tez kapsamında, makinedeki değişken sıcaklıklara göre elektrik tüketimini tahmin etmek için yapay zekâ tabanlı modeller aracılığıyla bir regresyon modeli önerisi amaçlanmıştır. Böylelikle, sıcaklığın elektrik tüketimini nasıl etkilediğine dair içgörüler oluşturması planlanmaktadır. Bu bölümde uygulanacak modellerin tanıtımına odaklanılmıştır. Bu amaçla, Uzun Kısa Süreli Bellek (LSTM), Otoregresif

Entegre Hareketli Ortalama (ARIMA) modeli ve Facebook Prophet modeli seçilmiştir.

2.2.1 Uzun Kısa Süreli Bellek (LSTM)

Uzun Kısa Süreli Bellek (LSTM), uzun süreli bağımlılık dizilerini tahmin etmek için mükemmel performansa sahip özel bir tekrarlayan sinir ağı türüdür. Bu gelişmiş sıralı ağ, LSTM, bilginin sürekliliğini sağlar. Ayrıca, RNN kaybolan gradyan sorununu çözme yeteneği, LSTM'yi daha kullanışlı hale getirir. Ayrıca, LSTM, geleneksel ileri beslemeli sinir ağlarının aksine geri besleme bağlantılarına sahiptir. Anlık verilere ek olarak veri dizilerini de yönetebilir [29].

Yöntem ilk olarak 1997 yılında Jürgen Schmidhuber ve Sepp Hochreiter [30] tarafından sunuldu. Sabit Hata Karuseli (CEC) birimleri, patlayan ve kaybolan gradyan problemlerini çözmek için LSTM'de kullanılmıştır. LSTM bloğunun ilk tasarımı, hücrelerin yanı sıra giriş ve çıkış bağlantı noktalarını da içerdiği [31]'de belirtilmiştir.



2.4: t zamanında standart bir LSTM modelinin şematik gösterimi (t anından $t+1$ anında kadar olan bir adımlık sürecin gösterimidir)¹

¹ Yöntem üç ana kapıdan oluşur: Unut Kapısı, yeni bilgilerin unutulup unutulmayacağına karar verilmesi anlamına gelir. Giriş Kapısı, girdi tarafından çıkarılan yeni bilginin önemini değerlendirir. Ayrıca bu kapı hücre durumu bilgisini de günceller. Çıkış Kapısı, bir sonraki gizli katmanda hangi bilginin taşınacağına karar verilmesidir. Burada $\oplus$ noktasal toplamayı belirtmek için kullanılır ve $\otimes$ noktasal çarpmayı ifade eder. Üstelik, $\odot$ operatörü de aktivasyon fonksiyonunu ifade eder.

Anlaşılr olması adına, Şekil 2.4, tek değışkenli bir LSTM yönteminin tek seferlik bir adımını şematik olarak göstermektedir. Kesin süreç şu şekilde tarif edilebilir: Temelde LSTM'nin mimarisi, sırasıyla Kapıyı Unutma, Giriş Kapısı ve Çıkış Kapısı olmak üzere üç ana bileşen içermektedir. Önceki durumdan mevcut duruma iletilen iki bileşen vardır. En önemli bileşen hangisinin doğrusal bir kendi kendine döngü olduğudur [32-33]. Yine de ağırlığı bir unutma kapısı tarafından kontrol edilir. Unutma kapısının ana odak noktası; ağ, yeni verilerle karşılaştığında hangi bilgilerin unutulacağına karar vermektir. Unutma kapısının çıktı işlevi şu şekilde tanımlanır:

$$f_t = \sigma(x_t \omega_f + H_{t-1} U_f + b_f)$$

Burada x_t o anki yeni bilgiyi, ω_f girdiye ait ağırlığı, H_{t-1} önceki zaman damgasına ait geçerli gizli katmanı ve U_f yinelenen ağırlığı gösterir. Hücre durumuna göre birleşen yeni bilginin bir biçiminin C_{t-1} ile birleşen giriş kapısına girip girmeyeceğine karar vermek. Son karar için bu adımdaki aktivasyon fonksiyonu aşağıdaki gibi tanımlanan sigmoid fonksiyonudur:

$$\begin{cases} C_{t-1} f_t = 0, & \text{herşeyi unut} \\ C_{t-1} f_t = 1, & \text{hiçbirşeyi unutma} \end{cases}$$

Unutma kapısını, hücre durumunu beslemek için hangi bilgilerin dikkate alındığını belirlemekten sorumlu olan giriş kapısı takip etmektedir. Hücre durumunun LSTM'nin hafızasını temsil ettiğini belirtmekte fayda vardır. Sigmoid aktivasyon fonksiyonu kullanılarak giriş kapısı fonksiyonu aşağıdaki gibi hesaplanabilir:

$$i_t = \sigma(x_t \omega_i + H_{t-1} U_i + b_i)$$

Burada x_t o anki yeni bilgiyi, ω_i girdiye ait ağırlığı, H_{t-1} önceki zaman damgasına ait mevcut gizli katmanı ve U_i yinelenen ağırlığı gösterir. Daha önceki bir adımda gizli katmanın doğrusal bir kombinasyonu ve yeni bilgi, aday değeri oluşturmak için tanh aktivasyon fonksiyonu altında değerlendirilir, öyle ki:

$$\hat{C}_t = \tanh(x_t \omega_c + H_{t-1} U_c + b_c)$$

$\hat{C}_t$, -1 ve 1 aralığında deęişen deęerler alabilir. Burada, $\hat{C}_t = -1$ ise bilgilerin çıkarılacağı, $C_t = 1$ olduğunda ise ekleneceęi anlamına gelir. Güncellenen hücre durumu aşığıdaki şekilde hesaplanabilir.

$$C_t = C_{t-1} f_t + i_t \hat{C}_t$$

Son kapı olan çıkış kapısı, hem sigmoid hem de $\tanh$ aktivasyon fonksiyonları kullanılarak sonlandırılır. Sigmoid aktivasyon işlevi, aşığıdakiler tarafından tanımlanan bilgilerin kaldırılıp kaldırılmayacağına karar verir:

$$o_t = \sigma(x_t \omega_o + H_{t-1} U_o + b_o)$$

Burada ω_o ve U_o sırasıyla x_t ve H_{t-1} in ağırlıklarıdır. Ayrıca b_o yanlılık terimidir. Öte yandan, bir gizli katman bir sonraki gizli durum için taşınan bilgileri hesaplar. Geçerli gizli katman daha sonraki bir zaman damgasına aittir ve noktasal çarpma yardımıyla aşığıdaki gibi hesaplanabilir:

$$H_t = o_t \tanh(C_t).$$

Son olarak, yeni hücre durumu ve yeni gizli durum, bir sonraki adımın iki girdi bileşeni olarak saklanır. LSTM konusu hakkında daha fazla bilgi için ilgili okuyucu [34]'de sunulan çalışmayı inceleyebilir.

2.2.2 Otomatik Gerileyen Entegre Hareketli Ortalama (ARIMA)

Otomatik Regresif Entegre Hareketli Ortalama (ARIMA) modeli, 1970 yılında Box ve Jenkins [35] tarafından geliştirildi. ARIMA modeli, iyi bilinen iki yöntemden oluşan ARMA modelinin geliştirilmiş bir versiyonu olarak kabul edilebilir: Otomatik Regresyon modeli (AR) ve Hareketli Ortalama modeli (MA). Hem ARMA modeli hem de ARIMA modeli stokastik modellerdir. Ancak ARMA modelinin temel eksikliği durağan veri gereksiniminin ARIMA modeli ile ortadan kaldırılmış olmasıdır. Bu bağlamda, ARIMA modelinin temel bileşenleri şu şekilde tanımlanabilir, bkz. [36-37]:

- AR(p): p mertebesinde Otomatik Gerileyen, verileri sığdırmak için kendi plugged gözlemlerini kullanır. Zaman serisinin mevcut değeri ile geçmiş değerleri arasındaki ilişkiyi modeller.
- I(d): d sırasının entegrasyonu. Bu operatör, bir metodoloji kullanarak geçmişin d gözlemini şimdiki zamandan çıkararak verileri durağan hale getirmeyi ifade eder [38-39]. Verilerdeki herhangi bir eğilimi veya mevsimselliği ortadan kaldırmak için d parametresi, durağan verilerin durağan olmayan (ham) verilere bağımlılığını simüle eden entegre bileşenin sırasını belirtir.
- MA(q): Hareketli ortalama bileşeninin sırası veya q parametresi, zaman serisinin mevcut değerinin, daha önceki beklenen değerlerden kalanlara veya hatalara nasıl bağlı olduğunu açıklar.

Matematiksel bir bakış açısından, model aşağıdaki gibi kompakt bir şekilde ifade edilebilir:

$$y_t^{(d)} = \gamma + \sum_{i=1}^p \phi_i y_{t-i}^{(d)} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}$$

Burada ϵ_t sıfır ortalama ve sonlu varyans σ_ϵ^2 ile gözlenemeyen bir rastgele beyaz gürültü olarak kabul edilir. Yani $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. Ayrıca γ sabit $\phi_i, i \in [1, p]$ ve $\theta_j, j \in [1, q]$ model parametreleridir. Model parametrelerindeki değişikliğin, değişen zaman serisi modellerine yol açtığını belirtmekte fayda var. Ancak hata teriminin varyansı, zaman serisinin ölçeğini değiştirebilir [40]. Ayrıca, $y_t^{(d)}$, d kat farkı olan serileri ifade eder. Daha kesin olarak, B geri kaydırma operatörünü temsil etsin, öyle ki

$$By_t = y_{t-1}$$

Fark operatörü $\nabla, \nabla y_t = y_t - y_{t-1}$ ile tanımlanır. Geri kaydırma operatörü yardımıyla ∇ fark operatörü $\nabla y_t = (1 - B)y_t$ olarak ifade edilebilir. Öyleyse,

$$y_t^{(d)} = \nabla^d y_t = (1 - B)^d y_t$$

ARIMA modeli hakkında daha fazla ayrıntı için okuyucuyu [40] numaralı referansa yönlendiriyoruz.

Optimum değerlerin belirlenmesi, oluşturulan model için güvenilir tahmin elde etmek için en önemli görevlerden biridir. Ayrıntılı zaman serileri analizine dayanan ve maksimum değerlerinin belirlenmesinden sonra, optimumu keşfetmenin otomatikleştirilmesi Algoritma 1'de sunulmuştur.

Fonksiyon: arma_order_select_ic:

Başla

'Sonuçlar' ve 'best_orders' adlı boş bir sözlük oluşturun

Girdiler:

- A. y - analiz edilecek zaman serisi verileri;
- B. max_ar - dikkate alınacak maksimum otoregresif terim sayısı
- C. max_ma - dikkate alınacak maksimum hareketli ortalama terim sayısı
- D. ic - model seçimi için kullanılacak bilgi kriterleri listesi; ARIMA modelinin fit() yöntemine iletilecek ek bağımsız değişkenler sözlüğü
- E. trend - modele sabit terim dahil edilip edilmeyeceğini belirtir

Yap: 0'dan max_ar'a kadar her p değeri için

0'dan max_ma'ya kadar her q değeri için a

- A. (p,0,q) ve trend=trend sırasına sahip bir ARIMA modelini y zaman serisi verilerine uydurmaya çalışın.
- B. Uydurma başarılı olursa, ic'deki her bir bilgi kriteri için sırayı ve karşılık gelen bilgi kriteri değeri(ler)ini sonuç sözlüğüne ekleyin.
- C. Uydurma başarısız olursa bir sonraki yinelemeye devam edin.

Yap: ic'deki her bilgi kriteri için aşağıdakileri yapın:

- A. Sonuçlar sözlüğünde bilgi kriterinin minimum değerini veren sırayı bulun.
- B. Geçerli bilgi kriteri için en iyi sırayı best_orders sözlüğüne ekleyin.

Bitir: best_orders sözlüğünü döndür.

Algoritma 1: ARIMA'daki optimal değerlerin belirlenmesi için sözde kod

2.2.3 Prophet Modeli

Yüksek performanslı tahminler oluşturmak zorlu bir problemdir çünkü hem alan hakkında derin bir bilgi sahibi olmayı hem de parametrelerini ayarlamayı ve teknikleri seçmeyi gerektirir. Facebook'tan yazarlar bu zorluklarla karşılaştılar ve bir çare olarak Prophet modelini önerdiler [41]. Modelin arkasındaki ana felsefe, ölçekte yorumlanabilir parametrelerle modüler yapılandırılabilir bir regresyon modeli oluşturmaktır. Prophet modelini önererek yazarlar, gerçek zamanlı seri veri setlerinin aykırı değerleri içermesi ve değişen mevsimsel etkileri gibi doğal özelliklerini dikkate alan bir model oluşturmayı amaçlamışlardır. Buna ek olarak, yazarlar [41] zaman serisi problemlerini bir eğri uydurma alıştırması olarak gördüklerini ve bunun Prophet modeline esneklik, düzenlilik aralığını kaldırma gibi birçok avantajı im ettiğini vurgulamışlardır [41]. ARIMA modelinden farklı olarak ölçümler, uyarlanabilirlik, yorumlanabilirlik. Yukarıda belirtilen tüm özellikler ve avantajlar, hızı ve doğruluğu ile birleşerek, Prophet modelini birçok uygulamada diğer yöntemlerden üstün kılabilir.

Prophet modeli, [42]'dekine benzer bir toplama modeli kullanarak, trend, mevsimsellik ve tatiller gibi üç ana bileşeni dikkate alarak zaman serisi verilerini ayrıştırır.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

burada $y(t)$, t süresi için tahmin edilen değeri ifade eder ve $g(t)$, doğrusal veya doğrusal olmayan fonksiyonlarla modellenen trend (periyodik olmayan) bileşeni temsil eder. Ayrıca, $s(t)$ yinelenen kalıpları yakalamak için mevsimsellik bileşenidir ve $h(t)$, istisnai olayların zaman serileri üzerindeki etkisini modellemek için tatil bileşenidir. Tatil parametresinin programlarının düzenli aralıklara sahip olması gerekmediğini burada hatırlatmak önemlidir. Ayrıca ϵ_t , verilerdeki diğer herhangi bir değişkenlik kaynağını temsil eden indirgenemez artık bileşeni ifade eder.

Prophet modelinin arkasındaki matematik şu şekilde açıklanabilir: Model, iki parçalı trend modelini kullanır, öyle ki

$$g(t) = \begin{cases} \frac{C(t)}{1 + \exp(-k_{adj}(t - m_{adj}))}, & \text{lineer olmayan, parçalı – tanımlı} \\ & \text{lojistik büyüme} \\ k_{adj}t + m_{adj}, & \text{lineer, parçalı – tanımlı} \\ & \text{lineer büyüme} \end{cases}$$

Burada k_{adj} ve m_{adj} sırasıyla ayarlanmış oran ve ayarlanmış ofset parametreleridir, öyle ki $k_{adj} = k + \mathbf{a}(t)^T \boldsymbol{\delta}$ burada k sabiti büyüme oranıdır, hız ayarlama vektörü $\boldsymbol{\delta} = [\delta_1, \delta_2, \dots, \delta_S]$, burada S bir zamandaki değişim noktalarını $s_j, j = 1, 2, 3, \dots, S$ gösterir ve δ_j , deki orandaki değişikliktir. Ayrıca ayarlanan ofset parametresi $m_{adj} = m + \mathbf{a}(t)^T \boldsymbol{\gamma}$ ile tanımlanır, burada $\boldsymbol{\gamma}$, tarafından tanımlanan değişim noktasındaki doğru ayardır.

$$\gamma_j = \begin{cases} \left(s_j - m - \sum_{p < j} \gamma_p \right) \left(1 - \frac{k + \sum_{p < j} \delta_p}{k + \sum_{p \leq j} \delta_p} \right), & \text{lineer olmayan trend için} \\ -s_j \delta_j, & \text{lineer trend için} \end{cases}$$

γ_j nin esas olarak δ_j değerine bağlı olduğuna ve tümevarımsal olarak bulunabileceğine dikkat edin, öyle ki

$$j=0 \text{ için: } \gamma_0 = (s_j - m) \left(1 - \frac{k}{k + \delta_0} \right),$$

$$j=1 \text{ için: } \gamma_1 = (s_j - m - \gamma_0) \left(1 - \frac{k + \delta_0}{k + \delta_0 + \delta_1} \right),$$

$$j=2 \text{ için: } \gamma_1 = (s_j - m - \gamma_0 - \gamma_1) \left(1 - \frac{k + \delta_0 + \delta_1}{k + \delta_0 + \delta_1 + \delta_2} \right),$$

Ayrıca, $\mathbf{a}(t) = [a_1(t), a_2(t), \dots, a_S(t)]$ burada

$$a_j(t) = \begin{cases} 1, & \text{eğer } t \geq s_j \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Doğrusal model için $\boldsymbol{\delta}$ üzerine önceden bir seçim konulmuştur, öyle ki $\boldsymbol{\delta} \sim \text{Laplace}(0, \tau)$ değişiklik noktaları seçimini otomatikleştirir. Burada τ , modelin hızını değiştirme esnekliği için kontrol parametresidir. τ parametresi, belirsizliğin belirlenmesi nedeniyle tahminlerin yapılmasında çok önemli bir rol oynamaktadır. Bir Bayesci çalışma, hiyerarşik bir önsel ile sonsalını belirlemek için kullanılabilir;

aksi takdirde, maksimum olabilirlik tahmini kullanılabilir. Daha fazla ayrıntı için ilgili okuyucuyu [41]'de yapılan en temel çalışmaya yönlendiririz.

Prophet modelinin trend bileşenine ek olarak, Prophet modelinin mevsimsel bileşeni esas olarak Fourier serisine bağlıdır. Yani mevsimsel bileşen şu şekilde ifade edilebilir:

$$s(t) = X(t)\beta$$

Burada

$$X(t) = \begin{bmatrix} \cos\left(\frac{2\pi(1)t}{P}\right), \sin\left(\frac{2\pi(1)t}{P}\right), \cos\left(\frac{2\pi(2)t}{P}\right), \sin\left(\frac{2\pi(2)t}{P}\right), \dots, \\ \cos\left(\frac{2\pi(N-1)t}{P}\right), \sin\left(\frac{2\pi(N-1)t}{P}\right), \cos\left(\frac{2\pi(N)t}{P}\right), \sin\left(\frac{2\pi(N)t}{P}\right) \end{bmatrix}$$

Burada P normal periyodu temsil eder. Ayrıca β , $\beta = [a_1, b_1, a_2, b_2, \dots, a_N, b_N]$ olacak şekilde $1 \times 2N$ vektörüdür. Fourier katsayıları için ilk belirleme $\beta \sim \text{Normal}(0, \sigma_\beta^2)$ dur. Ayrıca (Taylor, S. J., & Letham, B., 2018)'de birçok problem için yıllık ve haftalık mevsimselliğin $N=3$ ve $N=10$ olduğu belirtilmiştir.

Ayrıca, tatil bileşeni şu şekilde modellenilebilir:

$$h(t) = W(t)\kappa$$

burada $W(t) = [1(t \in D_1), \dots, 1(t \in D_L)]$ ve $\kappa \sim \text{Normal}(0, \sigma_\kappa^2)$. D_j 'nin j . tatili için tatil tarihlerini temsil ettiğine dikkat edilmeli.

Prophet modeli, bu bileşenlere ek olarak, Uzmanlar tarafından Kapasiteler, Değişim Noktaları, Tatiller ve mevsimsellik ve Düzeltme parametrelerine müdahale edilmesine de olanak tanır. Bir analistin, modelin daha güvenilir tahmin yakalama yeteneğini geliştirebileceği gerçeğini göz önünde bulunduran yazarlar, modelin tahmin kalitesini otomatik olarak artırmak için çeşitli tahmin stratejilerini birleştiren alternatif bir döngü içinde analist modelleme yaklaşımı geliştirdiler. Tahmin sürecinin ayrıntılı analizleri ve açıklamaları için ilgili okuyucuyu bu konudaki ana referansımız olan [41]'e yönlendiriyoruz.

Özetle, Prophet modeli, çok çeşitli uygulamalar için kullanılabilen, zaman serileri tahmini için çok yönlü ve kullanıcı dostu bir araçtır. Otomatik değişim noktası tespiti, belirsizlik aralıkları ve eksik veya aşırı değerleri işleme yeteneği gibi özellikler sunarak onu büyük ölçekli tahmin için sağlam bir seçim haline getirir.



Bölüm 3

Analiz Sonuçları ve Tartışmalar

Bu bölüm, tez kapsamında yürütülen analizlerin sonuçlarını sunmak için düzenlenmiştir. Bu bağlamda, tezin amacına uygun olacak şekilde, öncelikle, veri analizi sürecinde yapılan işlemler ve verinin tanımlayıcı analizleri sunulmuştur. Daha sonra, hazırlanan veri kümesi kullanılarak Bölüm 2 de sunulan LSTM, ARIMA ve Prophet metotlarının karşılaştırmalı analizi verilmiştir. Daha kapsamlı karşılaştırma yapılabilmesi amacıyla hem tek değişkenli analiz hem de çok değişkenli analiz modelleri de incelenmiştir.

3.1 Veri Analizi

Veri analizi, verilerin toplanması ve işlenmesi ile ilişkili iyi bilinen bir süreçtir. Ayrıca, veri analizi, çıkarımlar yapmak ve veriler hakkında daha derin içgörüler elde etmek için verileri inceleme çalışmasıdır. Verilerin analizi, ana hipotezin daha derin bir şekilde anlaşılmasının yanı sıra yeni hipotezler üretme fırsatları da sunar. Ayrıca, yapılan analiz, çeşitli modeller arasından model seçimi için bir bakış açısı sunabilir. Bahsedilen tüm avantajlar, daha güvenilir bir model elde etmek için veri analizini kaçınılmaz bir kavram haline getirmektedir.

Bölüm 3 boyunca yürütülen analizler 05 Haziran 2022-26 Mayıs 2023 tarihleri arasında veritabanından sorgulama işlemleri ile toplanan veriler üzerinde uygulanmıştır. Bu bağlamda, sıcaklık verileri 5 dakikalık aralıklarla, tüketim verileri ise 15 dakikalık aralıklarla kayıt altına alınmıştır. Veritabanında saklanmasından kaynaklı veri kümelerinin kullanımını optimize etmek için, ortak bir grupta oluşturulması hedeflenmiştir. Bu amaçla, hem sıcaklık verileri hem de tüketim

verileri saatlik bazda gruplandırılmıştır. Bu gruplandırma işlemleri sıcaklık verisinde her saat için ortalama değer, tüketim verisinde ise her saat içindeki kümülatif tüketim miktarı olacak şekilde hesaplanmıştır. Bu yaklaşım, saatlik eğilimler ve kalıplarla ilgili temel bilgileri korurken, veri kümesinin boyutunu etkili bir şekilde azaltmıştır. Yani, veritabanındaki verinin ilk boyutu 69295×6 iken, gruplandırılmış veri kümesinin boyutu 6474×6 'dır.

Ele alınan gruplandırılmış veri kümesindeki boş satırların değerlendirilmesi, eksik veya hatalı girişlerin belirlenmesi ve uç değerlerin incelenmesi gibi prosedürleri kapsayan çeşitli analizler gerçekleştirilmiştir. Bu bağlamda, gruplandırılmış veri kümesinin temizleme işleminden öncesi ve sonrasının tamamlayıcı istatistikleri Şekil 3.1 ve Şekil 3.2'de sırasıyla okuyucuya sunulmuştur.

	on	orta	arka	ortam	kWsa
count	6474.000000	6472.000000	6474.000000	6474.000000	6408.000000
mean	266.265998	251.308197	248.113006	54.456530	12.969126
std	16.078473	1095.280409	18.384772	21.585865	2.482670
min	17.828400	-69926.285489	17.663300	-6.134336	0.000000
25%	265.274675	270.411215	247.112204	60.789416	11.693825
50%	268.535677	272.735377	251.602572	63.324522	13.932350
75%	269.713381	275.610350	254.464754	65.238361	14.726650
max	301.568125	303.986678	276.724333	77.787400	26.727900

Şekil 3.1: İşlenmemiş veri kümesinin tamamlayıcı istatistikleri

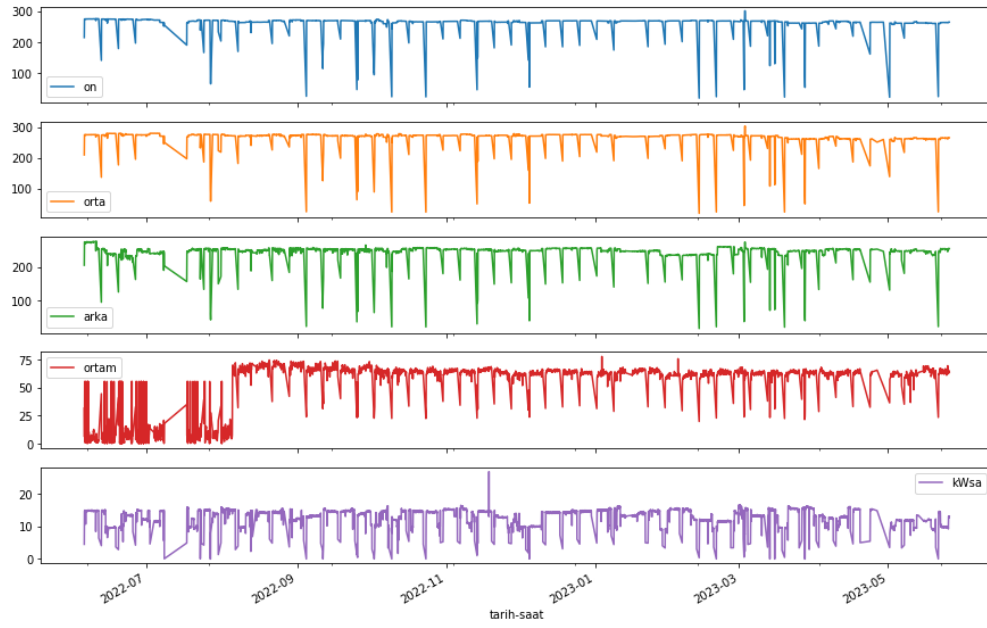
	on	orta	arka	ortam	kWsa
count	6474.000000	6474.000000	6474.000000	6474.000000	6474.000000
mean	266.265998	270.376741	248.113006	55.464145	12.969126
std	16.078473	16.472168	18.384772	20.242866	2.469980
min	17.828400	18.638300	17.663300	0.014317	0.000000
25%	265.274675	270.408696	247.112204	60.789416	11.711375
50%	268.535677	272.735085	251.602572	63.324522	13.881850
75%	269.713381	275.609748	254.464754	65.238361	14.719300
max	301.568125	303.986678	276.724333	77.787400	26.727900

Şekil 3.2: Ön işleme yapılmış kümesinin tamamlayıcı istatistikleri

Şekil 3.1 ve Şekil 3.2’den görüldüğü gibi, veri kümesi, zaman sütununa ek olarak ön, orta, arka, ortam sıcaklıkları ile elektrik tüketim verisi içermektedir. Burada, sıcaklık verileri Celsius ve tüketim verileri kilowatt saat cinsindendir. MongoDB’den gruplandırılarak aktarılan, Şekil 3.1, veriler dikkatle incelenmiş, sensörden kaynaklı elde edilen aykırı değerler, kontrolümüz dışındaki dış etkenler nedeniyle engellendiği dönemlerle karşılaşıldığında, ortaya çıkan boşluklar zamansal sürekliliği sağlamak için ortalama değerlerle doldurulmuştur, bkz Şekil 3.2. Böylelikle, eksiklikler ve düzensizlikler düzelterek veri kümesinin iyileştirilmesi ve, dolayısıyla, model eğitimi sürecinde daha yüksek doğruluklu sonuçlar üretilmesi amaçlanmıştır.

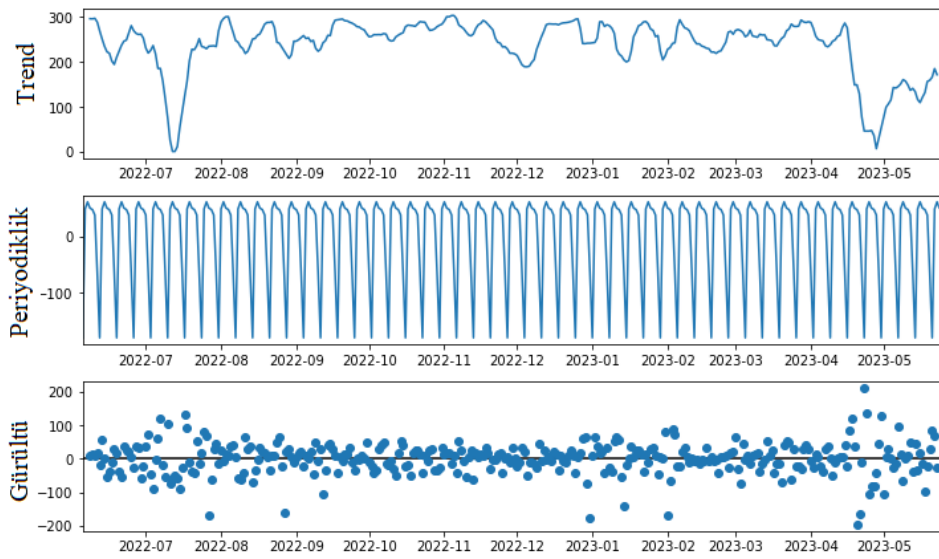
3.2 Zaman Serisi Analizleri

Bir zaman serisi, belirli bir zaman aralığında kaydedilen bir dizi gözlemden oluşan veri kümesidir, [43]. Finanstan tıp alanına, astronomiden meteorolojiye kadar farklı bir çok alanda çok önemli kullanım alanına sahip olan zaman serileri özellikle gelecek tahminleri konusunda önemli bir yere sahiptir. Tezin hedefi doğrultusunda ele alınan veri bir zaman serisi örneğidir. Bu amaçla, model çalışmalarına geçmeden önce veri kümesinin daha iyi anlaşılması ve gerekli parameter ve model seçimleri için yürütülen tanımlayıcı zaman serisi analizleri ve yorumları bu bölümde sunulmuştur. İlk olarak, veri kümesinde her bir özelliğin zamana göre olan grafikleri Şekil 3.3’de sunulmuştur.



Şekil 3.3: Veri kümesindeki özelliklerin zamana göre çizgisel grafikleri

İkinci olarak, sunulan tezin nihai hedefi doğrultusunda, veri kümesinin daha kapsamlı incelemesinin yapılabilmesi, amacıyla zamansal elektrik tüketim verisinin temel bileşenleri, bir başka deyişle trend, periyodiklik ve gürültü terimleri, Şekil 3.4'te sunulmuştur.

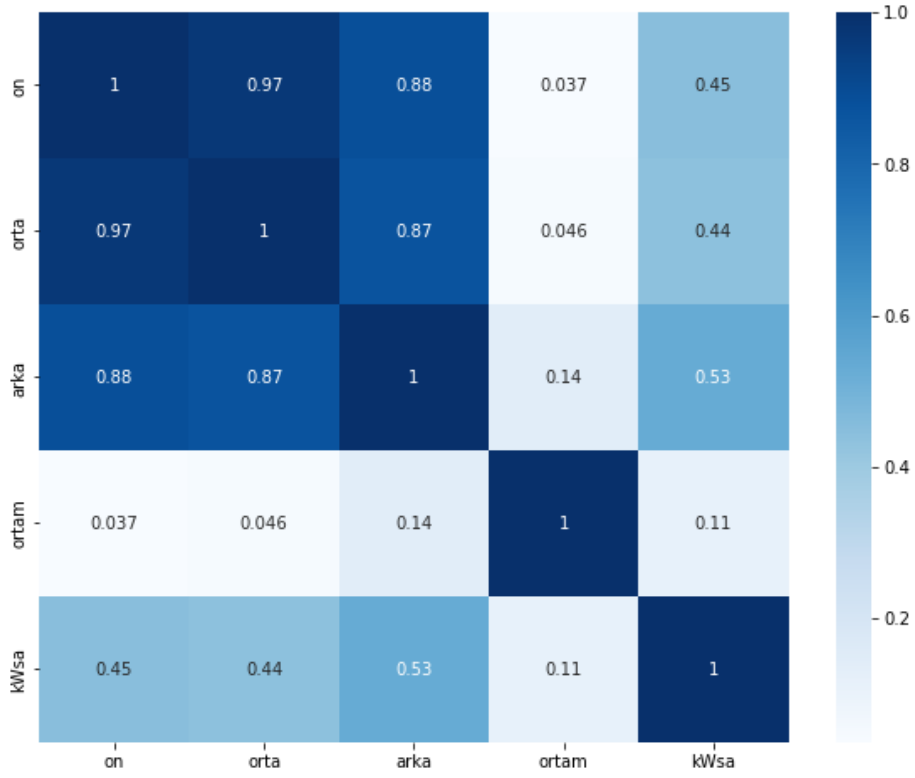


Şekil 3.4: Veri kümesinin trend, periyodiklik, gürültü analizini gösteren şekil

Şekil 3.4'te elektrik tüketim verisinin zaman içerisinde sabit bir eğilim (trend) sergilemediği söylenebilir. Bir başka deyişle, elektrik tüketim tahmininde lineer olmayan bir modelin, lineer modele nazaran daha güvenilir olabileceği söylenebilir. Ayrıca, verinin periyodikliğini incelediğimizde ise aylık ortalama dört tepe oluşturduğu gözlenmektedir. Burada, verinin haftalık bir periyodik davranış gösterdiği çıkarımında bulunabiliriz. Şekil 3.4'te, son olarak, sunulan gürültü terimi ise, verinin indirgenemez hata terimleri olarak düşünülebilir. Burada, sıfır çizgisini baz alarak genel anlamda simetrik bir yapı oluşturması, elektrik tüketim verisini etkileyen gürültü teriminin beyaz-gürültü (White Noise) olduğu söylenebilir. Bu da, zaman serisi analizlerinde, istenilen bir durumdur.

Bu analizlere ek olarak zaman serilerinde önemli bir yeri olan durağanlık kontrolü ise artırılmış Dickey-Fuller hipotez testi ile incelenmiştir. Analiz sonucunda, p -değeri 0.0926 olarak kaydedilmiştir. Elde edilen p -değeri, 0.05 eşik değerinden büyük olması sebebi ile sıfır-hipotez reddedilemez. Bu sonuç vasıtasıyla, özellikle ARIMA modelinde fark parametresinin etkili olabileceği çıkarımı yapılabilir.

Model uygulamalarına geçmeden önce, son olarak, elektrik tüketim verisinin en etkileşimli olduğu, başka bir deyişle elektrik tüketimini etkileyen önemli özelliklerin belirlenmesi amacıyla korelasyon incelemesi yapılmıştır. Veri kümesi tamamen sayısal değerlerden oluştuğundan, Pearson korelasyonu kullanılarak korelasyon katsayıları çıkarılmıştır. Yapılan analizin görseli Şekil 3.5'te sunulmuştur.



Şekil 3.5: Veri kümesinin korelasyonunu gösteren şekil

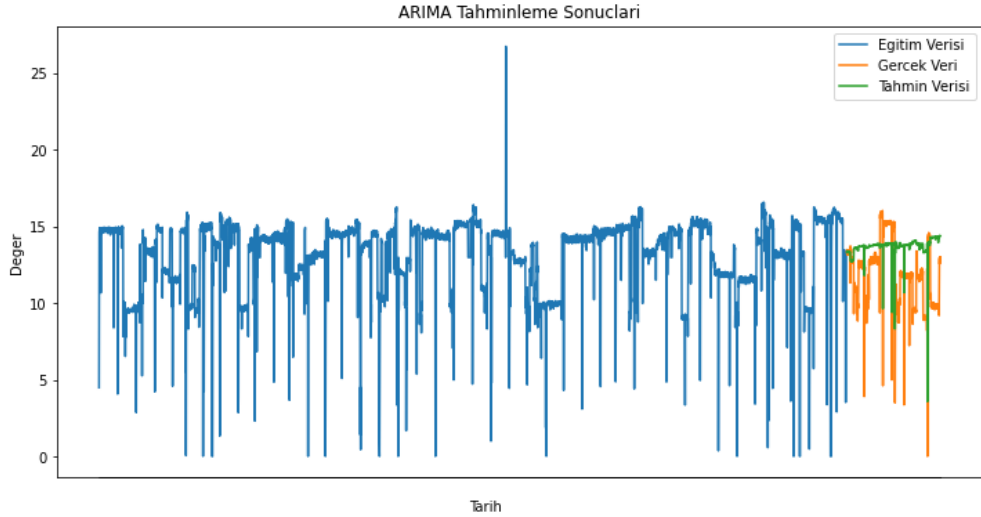
Şekil 3.5'ten açıkça görülmektedir ki elektrik tüketimi, ortam sıcaklığı ile doğrusal olarak ilişkisi azdır. Bununla birlikte, arka sıcaklıklarla ön ve orta sıcaklıklardan daha güçlü bir korelasyon ilişkisi vardır. Korelasyonun, giriş değişkenleriyle hedef değişkeni arasındaki doğrusal ilişkiyi temsil ettiğini belirtmek önemlidir. Şekil 3.5 ayrıca arka sıcaklıklar ile hem ön hem de orta sıcaklıklar arasında pozitif ve yüksek bir korelasyon ilişkisi olduğunu vurgulamaktadır. Bu durum aynı zamanda arka sıcaklıkların ön ve orta sıcaklıkları doğrusal olarak etkilediği anlamına gelir. Eş-doğrusallık (colinearity) dikkate alındığında, modellerimizin güvenilirliği artırmak amacıyla çok değişkenli analizlerde, zaman parametresine ek olarak, elektrik tüketimi ile en yüksek korelasyon katsayısı olan arka sıcaklık parametresi ele alınmıştır.

3.3 Model Karşılaştırmaları

Bu çalışma, Klemsan Elektrik Elektronik şirketinin bir klemens üretim makinesinden toplanan sıcaklık ve elektrik tüketimi verilerini kullanarak bir ARIMA, LSTM, Prophet modellerinin günlük zaman serisi verilerini tahmin etmedeki performansını araştırmak için yapılmıştır. 05 Haziran 2022 ila 26 Mayıs 2023 tarihleri arasındaki bir aylık dönemi içeren verileri üzerinde modellerin etkinliği, Denklem 1 ve Denklem 2’de sırasıyla belirtilen simetrik ortalama mutlak yüzde hata (SMAPE), ortalama mutlak hata (MAE) ölçütler kullanılarak değerlendirilmiştir. Bu bölümde çalışılan tüm modellerde eğitim-test kümesi, toplam kümenin %80-%20’si olarak seçilmiştir.

$$\text{SMAPE: } \frac{1}{N} \sum_{i=1}^N \frac{|y(t_i) - \hat{y}(t_i)|}{(|y(t_i)| + |\hat{y}(t_i)|)/2} \quad (1)$$

$$\text{MAE: } \frac{1}{N} \sum_{i=1}^N |y(t_i) - \hat{y}(t_i)| \quad (2)$$



Şekil 3.6: Çok değişkenli ARIMA modeli sonucunun şekli

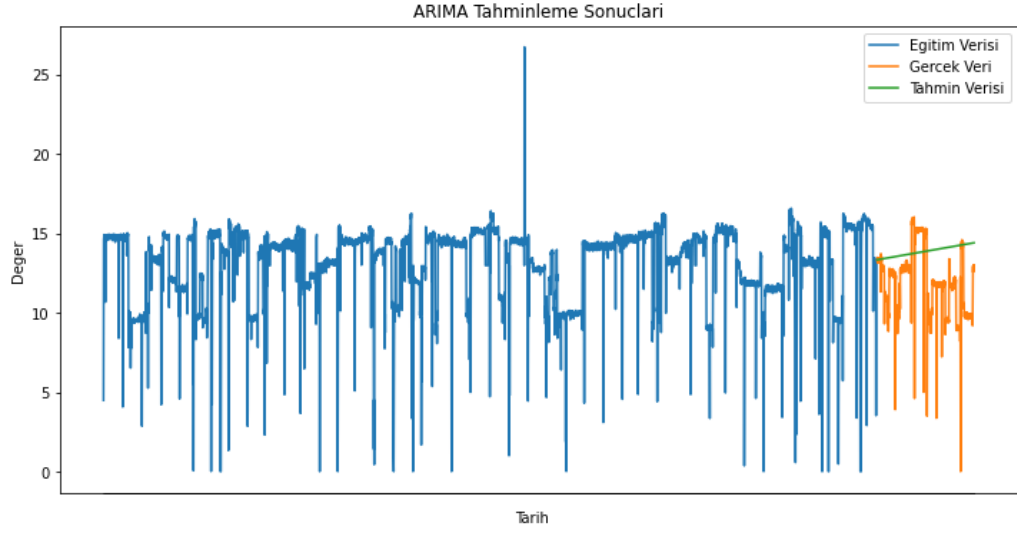
Saatlik elektrik tüketim tahminlerini gösteren Şekil 3.6, $p=1$, $d=2$, ve $q=1$ parametreleri ile çok değişkenli ARIMA modeli kullanılarak elde edilmiştir. Modelin doğruluk analizi Tablo 3.1’de SMAPE ve MAE metrikleri gözetilerek verilmiştir.

Verilen hatalar tahminlerin, gerçek değerlerle arasında daha iyi bir uyum gösterdiğini gösterir.

Tablo 3.1: Çok değişkenli ARIMA modelinin saatlik SMAPE ve MAE hata oranları karşılaştırılması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	1.1286	7.2468
04-04-2023 12:00	2.6352	17.7804
05-05-2023 12:00	4.4598	32.0651
12-05-2023 12:00	6.3895	51.7892
19-05-2023 12:00	7.1336	56.7431
25-05-2023 12:00	6.3741	49.2152

Tablo 3.1 ele alınan çok değişkenli ARIMA modelinin saatlik SMAPE ve MAE hatalarının temsilen sunulmuştur. Tüm verinin %20'lik kısmını oluşturan test veri kümesi 1295 satır içermesi sebebi ile Tablo 3.1 bu verilerin birkaç temsili yansıtmaktadır. Daha güvenilir bir sunuş amacıyla Mart ve Nisan aylarını temsilen birer örneklem ve tahminin daha az güvenilir olabileceğini düşünülen Mayıs ayını temsilen farklı 4 haftadan örneklem hatalar sunulmuştur. Tabloya göre MAE değerleri 1.1286 ile 7.1336 arasında değiştiği ve SMAPE değerlerinin ise 7.2468 ile 56.7431 gözlemlenmektedir. Bununla birlikte, tüm test veri kümesi boyunca en düşük MAE (0.0053) ve SMAPE (0.0331) değerleri 28-03-2023'te gözlemlenmiştir. Ayrıca, maksimum MAE (16.1386) ve SMAPE (200) ile 15-03-2023, 27-03-2023, ve 21-05-2023 tarihlerinde gözlemlenmiştir.



Şekil 3.7: Tek değişkenli ARIMA modeli sonucunun şekli

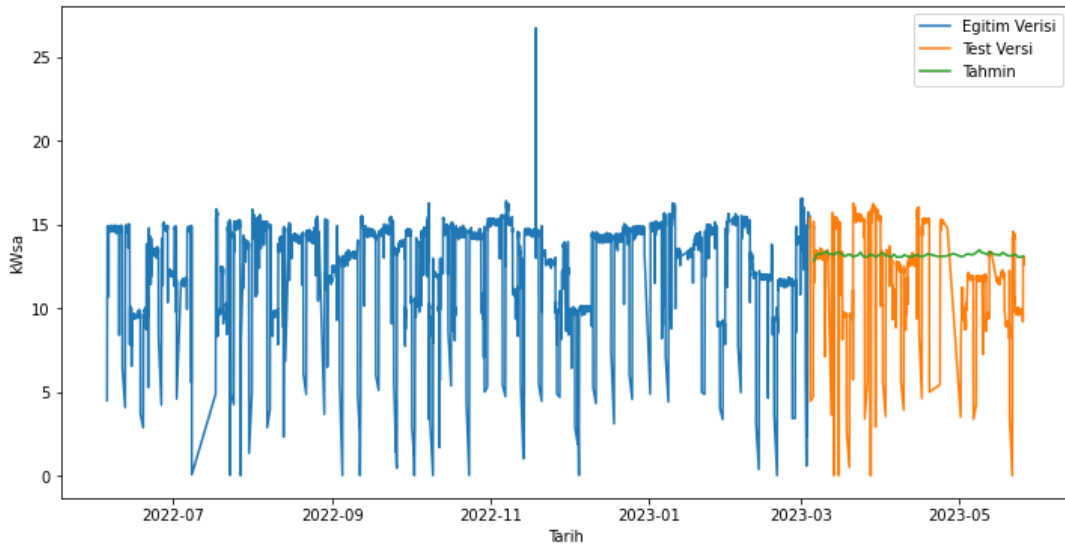
Şekil 3.7 tek değişkenli ARIMA modeli kullanarak, p , d ve q parametreleri, sırasıyla, 1,2 ve 1 seçilerek, saatlik elektrik tüketim tahminlerini göstermektedir. Modelin doğruluk analizi ise, Tablo 3.2’de farklı zaman örneklemelerinde SMAPE ve MAE değerleri verilerek sunulmuştur.

Tablo 3.2: Tek değişkenli ARIMA modelinin günlük günlük SMAPE ve MAE hata oranları karşılaştırılması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	1.1979	7.6745
04-04-2023 12:00	3.1095	21.2207
05-05-2023 12:00	4.5291	32.4822
12-05-2023 12:00	6.7082	52.1885
19-05-2023 12:00	7.2029	57.1367
25-05-2023 12:00	6.4434	49.6173

Tablo 3.2'ye göre MAE değerlerinin 1.1979 ile 7.2029 arasında değiştiği; SMAPE değerlerinin ise 7.6745 ile 57.1367 gözlemlenmektedir. Tüm test veri kümesi boyunca bakıldığında, en düşük MAE (0.0011) ve SMAPE (0.0073) değerleri 21-03-2023'de gözlemlenmiştir. En düşük skorlara ek olarak, maksimum MAE (16.2086) ve SMAPE (200) ile 15-03-2023, 27-03-2023, ve 21-05-2023 tarihlerinde gözlemlenmiştir. Her ne kadar ARIMA tek değişkenli analiz sonuçları, çok değişkenli analiz sonuçlarına benzer görünse de, Şekil 3.6 ve Şekil 3.7 de hesaba katıldığında ARIMA çok değişkenli modelin, tek değişkenli modele göre daha uyumlu sonuçlar verdiği gözlenmektedir.

Tez kapsamında ARIMA modeline ek olarak ele alınan bir diğer metot ise LSTM'dir. Benzer bir analiz yürütülerek LSTM modeli aynı veri kümesinde %80 eğitim ve %20 test oranları ile çalışılmıştır. LSTM modeli için de hem çok değişkenli analiz hem de tek değişkenli analiz dikkatle çalışılmıştır. Şekil 3.8 ve Tablo 3.3'te LSTM modelinin çok değişkenli analiz sonuçları sunulurken, Şekil 3.9 ve Tablo 3.4 ise LSTM modelinin tek değişkenli analiz sonuçlarını vurgulamaktadır.



Şekil 3.8: Çok değişkenli LSTM model sonucunun şekli

Şekil 3.8'de LSTM modelinin çok değişkenli modelinin görselini sunulmuştur. Şekil 3.8'deki görsel dikkate alındığında, modelin tahminlerinin gerçek verilerle çok

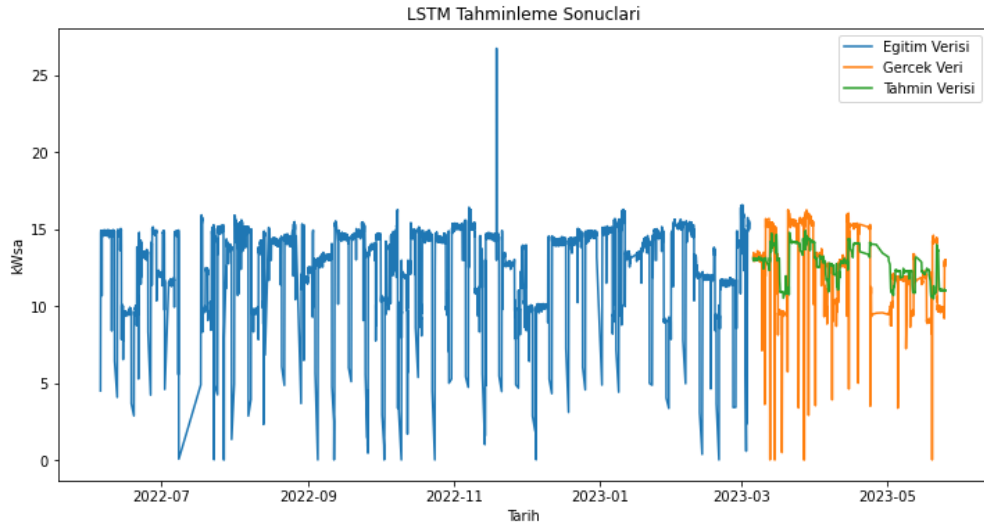
uyumlu olmadı gözlenmektedir. Bu amaçla, Tablo 3.3'te LSTM modelinin örnek MAE ve SMAPE skorları listelenmiştir.

Tablo 3.3: Çok değişkenli LSTM modelinin günlük SMAPE ve MAE hata oranları karşılaştırılması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	0.1870	1.4199
04-04-2023 12:00	0.8927	7.2606
05-05-2023 12:00	2.9391	29.0813
12-05-2023 12:00	2.1242	17.3667
19-05-2023 12:00	3.1230	31.2315
25-05-2023 12:00	2.0439	16.4755

Tablo 3.3'te de görüldüğü üzere, çok değişkenli LSTM modelinin MAE değerleri yaklaşık olarak 0.1870 ile 3.1230 arasında değişmektedir. Ek olarak, SMAPE değerleri ise 1.4199 ila 31.2315 aralığında değişkenlik gösterdiği gözlemlenmektedir. Tüm test veri kümesi üzerinde yapılan incelemelerde, en düşük MAE değeri 0.0011 olarak ve SMAPE değeri 0.0073 olarak 08-03-2023 tarihinde kaydedilmiştir. Yine tüm test veri kümesinde modelin en yüksek MAE skoru 5.3440 olarak ve en yüksek SMAPE skoru da 66.99 olarak 27-03-2023 tarihinde gözlemlenmiştir.

LSTM modelinin ele alınan veri kümesinde elektrik tüketim verisini tahminlemede çok değişkenli modelinin mi tek değişkenli modelinin mi daha iyi sonuç verdiğini anlamak için, LSTM modeli sadece zaman değişkeni dikkate alınarak da tek değişkenli model olarak çalışılmıştır. Çalışılan model sonuçları Şekil 3.9 ve Tablo 3.4'te okuyucuya sunulmuştur.



Şekil 3.9: Tek değişkenli LSTM sonucunun şekli

Her ne kadar iyi bir uyum yakalamamış görünse de, Şekil 3.9'un tahmin verilerinin, Şekil 3.8'e nazaran, test verilerinin yapısına daha uygun bir sonuç döndürdüğü çıkarımında bulunulabilir. Elbette, modelin güvenilirliğini test etmek amacıyla MAE ve SMAPE saatlik hata skorlarının örnekleri Tablo 3.4'te listelenmiştir.

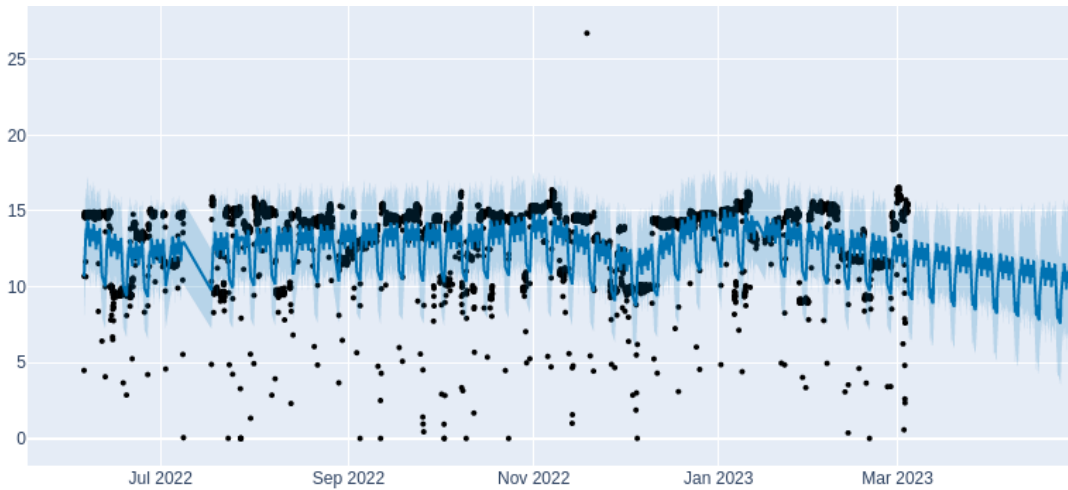
Tablo 3.4: Tek değişkenli LSTM modelinin günlük günlük SMAPE ve MAE hata oranları karşılaştırılması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	0.1687	1.2821
04-04-2023 12:00	0.8791	7.1858
05-05-2023 12:00	1.7617	20.2905
12-05-2023 12:00	1.4713	12.9703
19-05-2023 12:00	1.7647	19.4305
25-05-2023 12:00	1.7235	15.0232

Tablo 3.4'te verilen değerler baz alındığında MAE değerlerinin 0.1687 ila 1.7647 arasında değiştiği, buna ek olarak, SMAPE değerlerinin ise 1.2821 ile 20.2905 arasında değişkenlik gösterdiği görülebilir. Elbette, tablo değerleri tüm test verilerini

yansıtmamaktadır. Bu amaçla, tüm test verilerinde yapılan analiz sonucunda karşılaşılan en düşük MAE ve SMAPE skorları sırasıyla 0.1246 ve 0.9450'dir. Her iki metrik için de elde edilen en düşük skorlar 07-03-2023 tarihinde gözlemlenmiştir. Ayrıca, karşılaşılan en yüksek MAE ve SMAPE skorları ise sırasıyla MAE 5.3742 ve 66.527'dir. En yüksek skorlar ise, her iki ölçekte de, 27-03-2023 tarihinde kaydedilmiştir. SMAPE ve MAE değerleri ve Şekil 3.8 ile Şekil 3.9 birlikte ele alındığında, tek düzeyli LSTM modelinin, çok düzeyli LSTM modeline göre daha uyumlu sonuçlar ürettiği sonucuna ulaşılabilir.

Son olarak, Prophet metodu aynı veri kümesi ele alınarak elektrik tüketim tahminlemesi yapmak için uygulanmıştır. Tutarlılığı korumak amacıyla, tüm veri kümesi %80-%20 şeklinde eğitim ve test kümesine bölünmüştür. Ayrıca, trend bileşeni için doğrusal büyüme oranı seçilmiştir. Eğitilen model, veri setinin yapısına uygun olarak saatlik sıklıkta ve test veri kümesinin boyutuna uygun olacak şekilde tahminlerde bulunulmuştur. Buradan itibaren sunulan tablo ve şekiller, Prophet modelinin sonuçlarını göstermektedir. Daha açık bir şekilde, Şekil 3.10 ve Tablo 3.5 çok düzeyli Prophet modelinin sonuçlarını yansıtırken, Şekil 3.11 ve Tablo 3.6 tek düzeyli durum için Prophet modelinin sonuçlarını sunmaktadır.



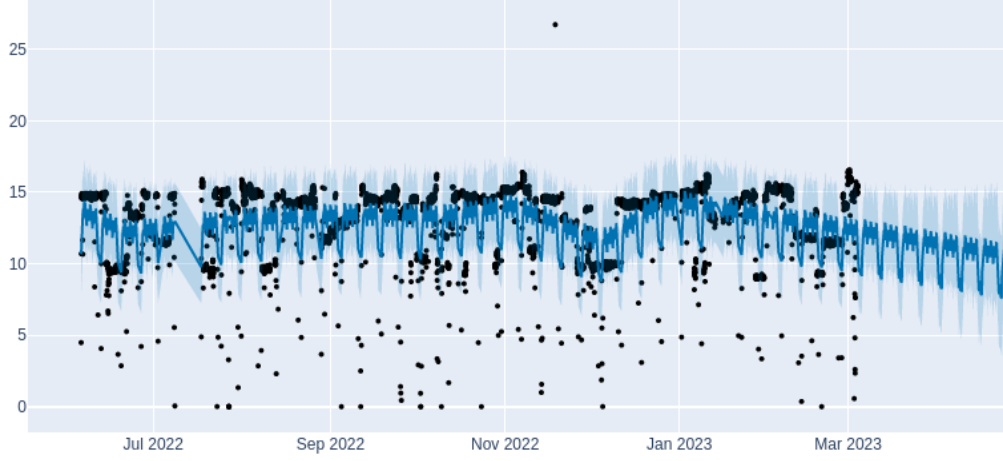
Şekil 3.10: Çok değişkenli Prophet model sonucunun şekli (nokta-gerçek değer, çizgi-tahmin değerleri)

Şekil 3.10’da tasvir edilen şekilde noktalar gerçek değerleri gösterirken mavi çizgi grafiği tahminleri göstermektedir. Modelin doğruluk analizi, diğer modellerde yaptığımız gibi, MAE ve SMAPE ölçekleri gözetilerek Tablo 3.5’te örneklendirilmiştir.

Tablo 3.5: Çok değişkenli Prophet modelinin günlük SMAPE ve MAE hata oranlarının karşılaştırması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	1.6572	11.7338
04-04-2023 12:00	1.3008	10.4082
05-05-2023 12:00	0.1908	1.6153
12-05-2023 12:00	2.3948	22.3371
19-05-2023 12:00	0.6457	5.5931
25-05-2023 12:00	0.7391	5.9580

ARIMA ve LSTM modellerindeki analize benzer bir analiz yürütülerek, Tablo 3.5’te Prophet modelinin test kümesi boyunca birkaç örneği verilmiştir. Tablo 3.5’e göre, MAE 0.1908 ile 2.3948 arasında değişkenlik gösterirken SMAPE değerleri 1.6153 ile 22.3371 arasında değişkenlik göstermektedir. Tüm veri kümesi dikkate alındığında elde edilen en düşük MAE ve SMAPE skorları, sırasıyla, 0.0101 ve 0.07828 olarak 12-04-2023’de gözlemlenmiştir. Ayrıca, tüm analiz kapsamında karşılaşılan en yüksek MAE ve SMAPE değerleri ise, sırasıyla, 10.7951 ve 199.9925 olarak 28-03-2023 tarihinde gözlemlenmiştir. Çok değişkenli Prophet modeline ek olarak, modelin farklı varyasyonlarını görebilmek adına, tek değişkenli Prophet modeli de çalışılmıştır. Bu bağlamda, çalışılan modelin görseli Şekil 3.11’de sunulmuştur.



Şekil 3.11: Tek değişkenli Prophet sonucunun şekli (nokta-gerçek değer, çizgi-tahmin değerleri)

Şekil 3.11, tek değişkenli Prophet model sonucunun bir tasviridir. Burada, noktalar gerçek değerleri gösterirken çizgi grafiği tahmin değerlerini yansıtmaktadır. Modelin daha detaylı analizinin tamamlanması amacıyla, model doğruluğu MAE ve SMAPE metrikleri yardımıyla Tablo 3.6’da örneklendirilmiş ve tablo yorumlamasında detaylandırılmıştır.

Tablo 3.6: Tek değişkenli LSTM modelinin günlük hata oranlarının karşılaştırması

Tarih- Zaman	Hata Oranları	
	MAE	SMAPE
04-03-2023 12:00	2.5185	18.3927
04-04-2023 12:00	0.6929	5.3289
05-05-2023 12:00	0.0061	0.5167
12-05-2023 12:00	2.3429	21.9063
19-05-2023 12:00	0.4420	3.7954
25-05-2023 12:00	1.0492	8.5647

Tablo 3.6’tek verilere göre MAE deęerleri 0.0061 ile 2.5185 arasında deęiřtięi ve SMAPE deęerlerinin ise 0.5167 ile 21.9063 arasında deęiřtięi gözlemlenmektedir. Bunula birlikte, tüm veri kümesi boyunca en düşük MAE skoru 0.0007 ve en düşük SMAPE skoru ise 0.0055 olarak 28-09-2022’te kaydedilmiřtir. En düşük deęerlere ek olarak, en yüksek MAE skoru 10.7951 ve en yüksek SMAPE skoru ise 199.9925 olarak 28-03-2023 tarihinde gözlemlenmiřtir. řekil 3.10-3.12 ve Tablo 3.5-3.6 birlikte ele alındığında çok deęiřkenli ve tek deęiřkenli benzer sonuçlar verdięi gözlemlenmiřtir. Bununla birlikte Tablo 3.5 ve Tablo 3.6 deęerleri dikkate alındığında, göreceli olarak çok deęiřkenli Prophet modelinin tek deęiřkenli modele göre tercih edilebilir olduęu söylenebilir.



Bölüm 4

Sonuç ve Gelecek Çalışmalar

Özet olarak, doğru tahminler enerji kullanımını etkili bir şekilde yönetme, maliyetleri azaltma ve sürdürülebilirliği teşvik etme açısından önemli bir rol oynamaktadır. Elektrik tüketiminin tahmin edilmesi, çeşitli sektörlerde büyük bir öneme sahiptir. Zaman serisi verilerinde karmaşık desenleri tanımlayabilme ve belirsizlik tahminleri sunabilme yetenekleri nedeniyle makine öğrenme, derin öğrenme ve yapa zekâ algoritmaları elektrik kullanımını tahmin etmek için güçlü araçlar olarak ortaya çıkmıştır. Bu çalışma, farklı stratejilere dayanan üç farklı modelin endüstriyel sektördeki makinaların elektrik tüketimini tahmin etmedeki etkinliğini değerlendirmeyi amaçlamıştır.

Bu bağlamda, tez kapsamında, literatürde sıklıkla tercih edilen ARIMA, LSTM ve Prophet modelleri ele alınmıştır. ARIMA modelleri zaman serisi analizinde sağlam bir temele sahiptir ve tahmin için yaygın olarak kullanılır. Bu modeller, verinin entegrasyonunu (I) dikkate alırken oto regresif (AR) ve hareketli ortalama (MA) bileşenlerini yakalar. ARIMA modelleri, elektrik tüketimi verilerindeki trendleri ve mevsimsellik desenlerini tanımlayarak güvenilir tahminler sağlayabildiği bilinmektedir. LSTM modelleri, zaman serisi verilerinde uzun vadeli bağımlılıkları yakalama yetenekleriyle bilinir. Ayrıca, tekrarlayan sinir ağı mimarisi sayesinde LSTM modelleri ardışık verileri etkili bir şekilde işleyebilir ve doğru tahminler sağlayabilir. Bu avantajları sebebiyle elektrik tüketimi tahmininde potansiyel olarak değerli olduğu düşünülmektedir. Sunulan bu modellere ek olarak, daha güncel bir metot olan Prophet metodunun elektrik tüketimini tahmin etmek için geleneksel

yöntemlere göre etkili bir alternatif olabileceği düşünülmektedir. Model, kullanıcı dostu ara yüzü ve basitleştirilmiş geliştirme süreci gibi önemli avantajlar sunar.

Sunulan tez üç ana hattan oluşmaktadır. Bunlardan ilki verilerin toplanması aşamasıdır. Bu amaçla, üretim hattındaki makinelerden sensörler vasıtasıyla veri toplanmıştır. Veri toplama işlemi sırasında, çok değişkenli ve tek değişkenli analizlerin karşılaştırılmasını sağlamak amacıyla, iki farklı alt küme dikkate alınmıştır: biri tarih ve elektrik tüketimi, diğeri ise tarih, elektrik tüketimi ve sıcaklık bilgilerini içermektedir. Sensörler aracılığıyla veri tabanına aktarılan verilerin, veri yapılarının ve zaman ölçeklerinin tutarlılığı ve yürütülecek analizlerin güvenilirliğini artırabilmek amacıyla bir takım ön işleme sürecine tabi tutulmuştur. Ayrıca, detaylı tanımlayıcı istatistik ve zaman serisi analizleri yürütülmüş ve yorumlanmıştır. Temel prensibi veriyi derinlemesine anlamak üzerine olan bu analizlerden sonra ele alınan çok değişkenli analizler için kullanılacak özellikler, korelasyon ilişkisi vasıtasıyla tartışılmıştır. Böylelikle, yapılan çalışmalar daha bilimsel bir çerçeveye oturtulmuştur. Son aşamada ise, ARIMA, LSTM ve Prophet metotları, işlenmiş veri kümesinde tek değişkenli ve çok değişkenli senaryolar üzerinde çalışılmıştır.

Elde edilen sonuçları, her bir metodun farklı senaryoları için (tek değişkenli ve çok değişkenli) tartışılmıştır. Daha açık şekilde ifade etmek gerekirse, tek değişkenli model analizleri, girdi fonksiyonu olarak sadece zaman faktörü; çıktı olarak ise yalnızca elektrik tüketim verisini kullanarak modellerin çalıştırıldığı durumu ifade etmektedir. Çok değişkenli model analizleri ise, girdi olarak hem zaman hem de arka sıcaklık verilerini kabul eder; çıktı fonksiyonu olarak yine elektrik tüketim verisi ele alınmıştır. Yapılan analizler şekiller ve tablolar ile desteklenmiştir. Model illüstrasyonlarının yanı sıra model doğrulukları mutlak hata ve simetrik mutlak yüzdelik hata ölçekleri ile test edilmiştir.

Elde edilen şekiller ve tablolar birlikte dikkate alındığında, senaryolar kapsamında çok değişkenli ARIMA modelinin, tek değişkenli modele göre daha uyumlu sonuçlar verdiği gözlemlenmiştir. Bununla birlikte, tek değişkenli LSTM modelinin, çok değişkenli modele göre daha uyumlu sonuçlar ürettiği, dolayısıyla daha tercih edilebilir sonuçlar verdiği gözlemlenmiştir. Son olarak, Prophet modeli hem tek düzeyli hem de çok düzeyli model uygulamalarında benzer hatalar üretmiş olmakla birlikte, verilen tablo değerleri ve görseller dikkate alındığında çok değişkenli

Prophet modelinin tek deęişkenli modele göreceli olarak tercih edilebilir olduęu söylenebilir.

Bu tez kapsamında yapılan detaylı karşılaştırmalı analizler sonucunda, LSTM modelinin MAE ve SMAPE ölçeklerinde dięer tüm modellere nazaran daha iyi sonuç verdięi görölmüştür. LSTM modelini, Prophet modeli takip etmiştir. Çalışılan modeller arasında en az tercih edilebilir model ARIMA model olmuştur.

Bu çalışmadan elde edilen bilgiler, LSTM metodunun endüstriyel ortamda kullanımının etkisini maksimize etmeye büyük katkı sağlayabileceęine yönelik çıkarımlar sunmaktadır. Çalışılan modelleri, sadece bir örnek makinadan alınan sonuçlara uygulandığını vurgulamakta fayda vardır. Ayrıca, veri kümesinin çeşitlilięinin artırılması (başka makinadan alınacak veriler ile) ya da eklenebilecek özelliklerin artırılması modellerin sonuçlarında farklılıklar oluşturabileceęi de vurgulanmalıdır.

Bu karşılaştırmalı analiz, her modelin güçlü yönlerini ve zayıflıklarını ortaya koyarak, belirli uygulamalar için en uygun yaklaşımı belirlemeye yardımcı olacaktır. Gelecekteki araştırma çalışmaları, Prophet metot, LSTM ve ARIMA modellerinin endüstriyel ortamlarda elektrik tüketimini tahmin etmedeki performansını karşılaştırmaya odaklanmalıdır. Ayrıca, birden fazla modelin güçlü yönlerini birleştiren genel uyum yöntemlerinin araştırılması daha doğru ve güvenilir tahminlere ulaşmayı sağlayabilir.

Kaynaklar

- [1] Lee, Y.W., Tay, K.G., & Choy, Y.Y. (2018). Forecasting Electricity Consumption Using Time Series Model. International Journal of Engineering & Technology
- [2] Kaur, H., & Ahuja, S. (2019). SARIMA Modelling for Forecasting the Electricity Consumption of a Health Care Building. International Journal of Innovative Technology and Exploring Engineering.
- [3] Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLoS ONE 13(3): e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- [4] Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. International journal of forecasting, 16(4), 451-476.
- [5] Kalimoldayev, M., Drozdenko, A., Koplyk, I., Marinich, T., Abdildayeva, A. & Zhukabayeva, T. (2020). Analysis of modern approaches for the prediction of electric energy consumption. Open Engineering, 10(1), 350-361. <https://doi.org/10.1515/eng-2020-0028>
- [6] Leite Coelho da Silva, F., da Costa, K., Canas Rodrigues, P., Salas, R., & López-Gonzales, J. L. (2022). Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector. Energies, 15(2), 588.
- [7] Wahid, F., Ghazali, R., Shah, A. S., & Fayaz, M. (2017). Prediction of energy consumption in the buildings using multi-layer perceptron and random forest. IJAST, 101, 13-22.
- [8] Zhong, H., Wang, J., Jia, H., Mu, Y., & Lv, S. (2019). Vector field-based support vector regression for building energy consumption prediction. Applied Energy, 242, 403-414.

- [9] Peng, Y., Liu, H., Li, X., Huang, J., & Wang, W. (2020). Machine learning method for energy consumption prediction of ships in port considering green ports. *Journal of Cleaner Production*, 264, 121564.
- [10] Fan, G., Yu, M., Dong, S., Yeh, Y., & Hong, W. (2021). Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling. *Utilities Policy*.
- [11] Ribeiro, A.M., Carmo, P.R., Silva, I.R., Sadok, D.F., Lynn, T., & Endo, P.T. (2020). Short-Term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: A Comparison of Machine Learning and Deep Learning Models. *Algorithms*, 13, 274.
- [12] González-Briones, A., Hernandez, G., Corchado, J. M., Omatu, S., & Mohamad, M. S.: Machine learning models for electricity consumption forecasting: a review. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-6) (2019). IEEE.
- [13] Saurabh, V., Vaibhav, K., Mrunmayee, G., Atharva, C., & Sandip, H.: A SURVEY: ELECTRICITY DEMAND PREDICTION USING STATISTICAL, MACHINE LEARNING AND DEEP LEARNING METHODS. *International Research Journal of Modernization in Engineering Technology and Science* 05(02): 493–502 (2023). <https://www.doi.org/10.56726/IRJMETS33460>
- [14] Ozcanli, AK, Yaprakdal, F, Baysal, M.: Deep learning methods and applications for electrical power systems: A comprehensive review. *Int J Energy Res.* 44 7136– 7157 (2020). <https://doi.org/10.1002/er.5331>
- [15] Seyedzadeh, S., Rahimian, F., Glesk, I. et al.: Machine learning for estimation of build-ing energy consumption and performance: a review. *Vis. in Eng.* 6, 5 (2018). <https://doi.org/10.1186/s40327-018-0064-7>
- [16] Koprinska, I., Wu D., and Wang, Z.: Convolutional Neural Networks for Energy Time Series Forecasting, 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018, pp. 1-8, <https://doi.org/10.1109/IJCNN.2018.8489399>.

- [17] Ramos, D., Faria, P., Vale, Z. and Correia, R.: Short Time Electricity Consumption Forecast in an Industry Facility. In IEEE Transactions on Industry Applications, vol. 58, no. 1, pp. 123-130, (2022). doi:10.1109/TIA.2021.3123103.
- [18] Chahkoutahi, F., Khashei, M.: A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. Energy, 140, 988–1004 (2017). <https://doi.org/10.1016/j.energy.2017.09.009>
- [19] Bedi, J. and Toshniwal, D.: Empirical Mode Decomposition Based Deep Learning for Electricity Demand Forecasting. IEEE Access, 6, 49144–49156, (2018). doi: 10.1109/ACCESS.2018.2867681.
- [20] Bedi, J., Toshniwal, D.: Deep learning framework to forecast electricity demand. Applied energy, 238, 1312–1326 (2019).
- [21] Jain, R., Jain, N., Gupta, Y. et al. A Modified Fuzzy Logic Relation-Based Approach for Electricity Consumption Forecasting in India. Int. J. Fuzzy Syst. 22, 461–475 (2020). <https://doi.org/10.1007/s40815-019-00704-z>
- [22] Guo, C., Ge, Q., Jiang, H., Yao, G., and Hua, Q.: Maximum power demand prediction using fbprophet with adaptive Kalman filtering. IEEE Access, 8, 19236-19247. (2020). doi: 10.1109/ACCESS.2020.2968101
- [23] Amantegui J, Morais H, Pereira L. Benchmark of Electricity Consumption Forecasting Methodologies Applied to Industrial Kitchens. Buildings. 2022; 12(12):2231. <https://doi.org/10.3390/buildings12122231>
- [24] Zhou, X., Lin, W., Kumar, R., Cui, P., and Ma, Z.: A data-driven strategy using long short term memory models and reinforcement learning to predict building electricity consumption. Applied Energy 306 :18078 (2022). <https://doi.org/10.1016/j.apenergy.2021.118078>
- [25] Bashir, T., Haoyong, C., Tahir, M. F., and Liqiang, Z.: Short-term electricity load fore-casting using hybrid prophet-LSTM model optimized by BPNN. Energy Reports, 8, 1678-1686 (2022). <https://doi.org/10.1016/j.egyr.2021.12.067>

- [26] Yan, K., Zhou, X., & Chen, J. (2022). Collaborative deep learning framework on IoT data with bidirectional NLSTM neural networks for energy consumption forecasting. *Journal of Parallel and Distributed Computing*, 163, 248-255.
- [27] Chaturvedi, S., Rajasekar, E., Natarajan, S., and McCullen, N.: A comparative assessment of SARIMA, LSTM RNN, and Fb Prophet models to forecast total and peak monthly energy demand for India. *Energy Policy*, 168, 113097. (2022). <https://doi.org/10.1016/j.enpol.2022.113097>
- [28] Tutun, S., Tosyali, A., Sangrody, H., Khasawneh, M., Johnson, M., Albizri, A., & Harfouche, A. (2023). Artificial intelligence in energy industry: forecasting electricity consumption through cohort intelligence & adaptive neural fuzzy inference system. *Journal of Business Analytics*, 6(1), 59-76. DOI: 10.1080/2573234X.2022.2046514
- [29] Prabhakaran S. ARIMA Model – Complete Guide to Time Series Forecasting in Python [İnternet]. *Machine Learning Plus 2022* [erişim tarihi 10.10.2022]. <https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/#freepythonprogramming>
- [30] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [31] Uzun Kısa süreli bellek [İnternet]. Wikipedia. Wikimedia Foundation; 2022 [erişim tarihi 10.10.2022]. https://tr.wikipedia.org/wiki/Uzun_k%C4%B1sa_s%C3%BCreli_bellek
- [32] Goodfellow, I., Bengio, Y., Courville, A. (Ed.) *Deep Learning*, MIT Press, 2016 <http://www.deeplearningbook.org>
- [33] Saxena S Learn About Long Short-Term Memory (LSTM) Algorithms [İnternet]. *Analytics Vidhya*. 2021 [erişim tarihi 10.10.2022]. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/#h2_1
- [34] Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

- [35] Box G., Jenkins G., Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, 1970
- [36] Autoregressive Integrated moving average (ARIMA) [İnternet]. Corporate Finance Institute. 2022 [erişim tarihi 10.10.2022]. <https://corporatefinanceinstitute.com/resources/data-science/autoregressive-integrated-moving-average-arima/>
- [37] Hayes A. Autoregressive Integrated moving average (ARIMA) [İnternet]. Investopedia; 2022 [erişim tarihi 10.10.2022]. <https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp>
- [38] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). IEEE.
- [39] Zhang, M. Time series: Autoregressive models ar, ma, arma, arima. University Lecture; 2018 Lecture 11, 12-13. <https://people.cs.pitt.edu/~milos/courses/cs3750/lectures/class16.pdf>
- [40] Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 07 January 2023.
- [41] Taylor, S. J., Letham, B.: Forecasting at scale. The American Statistician 72(1):37-45 (2018)
- [42] Harvey, A., and Peters, S. (1990), Estimation Procedures for Structural Time Series Models. Journal of Forecasting, 9, 89–108
- [43] Brockwell, P. J., & Davis, R. A. (Eds.). (2002). Introduction to time series and forecasting. New York, NY: Springer New York.

Özgeçmiş

Eğitim:

2022–2023	İzmir Kâtip Çelebi Üniversitesi, Yazılım Müh. Tezli Yüksek Lisans Bölümü
2020–2022	İzmir Kâtip Çelebi Üniversitesi, Yazılım Müh. Tezsiz Yüksek Lisans Bölümü
2014-2018	Doğuş Üniversitesi Bilgisayar Mühendisliği Lisans Bölümü
2011-2014	Dokuz Eylül Üniversitesi Bilgisayar Programcılığı Önlisans Programı

İş Deneyimi:

12/21 -	Klemsan Elektrik Elektronik A.Ş
04/19 - 12/21	Asay Group
01/19 - 04/19	Mini Bilişim

Yayınlar:

1. Yıldız, U., Korkut S. O. (2023). Electricity Consumption Forecasting using the Prophet Model in Industry: A Case Study. In Proceedings of the 2nd International Conference on Computing, IoT and Data Analytics