

**T.C.
BALIKESİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**



BURSA İLİ ELEKTRİK GÜCÜ TALEP TAHMİN ANALİZİ

YÜKSEK LİSANS TEZİ

MUSTAFA EREN KOÇBEY

BALIKESİR, HAZİRAN - 2019

**T.C.
BALIKESİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**



BURSA İLİ ELEKTRİK GÜCÜ TALEP TAHMİN ANALİZİ

YÜKSEK LİSANS TEZİ

MUSTAFA EREN KOÇBEY

Jüri Üyeleri: Dr. Öğr. Üyesi M. Kubilay EKER (Tez Danışmanı)

Dr. Öğr. Üyesi Özkan AKIN

Dr. Öğr. Üyesi Mustafa Ahmet Beyazıt OCAKTAN

BALIKESİR, HAZİRAN - 2019

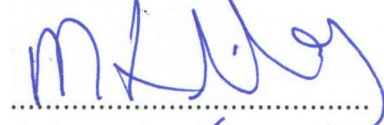
KABUL VE ONAY SAYFASI

Mustafa Eren KOÇBEY tarafından hazırlanan “BURSA İLİ ELEKTRİK GÜCÜ TALEP TAHMİN ANALİZİ” adlı tez çalışmasının savunma sınavı 19.06.2019 tarihinde yapılmış olup aşağıda verilen jüri tarafından oy birliği ile Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Danışman
Dr. Öğr. Üyesi Mehmet Kubilay EKER



Üye
Dr. Öğr. Üyesi Özkan AKIN



Üye
Dr. Öğr. Üyesi M. Ahmet Beyazıt
OCAKTAN



Jüri üyeleri tarafından kabul edilmiş olan bu tez Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulunca onanmıştır.

Fen Bilimleri Enstitüsü Müdürü

Prof. Dr. Necati ÖZDEMİR

.....

ÖZET

BURSA İLİ ELEKTRİK GÜCÜ TALEP TAHMİN ANALİZİ
YÜKSEK LİSANS TEZİ
MUSTAFA EREN KOÇBEY
BALIKESİR ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
(TEZ DANIŞMANI: DR. ÖĞR. ÜYESİ MEHMET KUBİLAY EKER)
BALIKESİR, HAZİRAN - 2019

Sinüzoidal sinyal yapısında akım taşınan enterkonnekte sistemin kararlı olarak çalışabilmesi için şebeke frekansının kabul edilebilir limitler içinde kalması gerekmektedir. Şebeke frekansı değişimi, tüketicilerin güç taleplerine ve üretim santrallerinin sisteme aktardığı güç miktarına bağlıdır. Üretim ve tüketim güç dengesinin sürekli olarak sağlanabilmesi için tüketicilerin anlık güç taleplerinin önceden tahmin edilmesi önemli bir problemdir.

Talep tahmini, elektrik enerjisi üretim miktarının belirlenebilmesinin yanında, uygun elektrik enerjisi fiyatlandırması açısından da önemlidir. Üretim kaynakları açısından son yıllarda yenilenebilir enerji kaynaklarına yönelimin artması nedeniyle üretim miktarlarında da değişkenlik olabileceği dikkate alındığında, talep edilecek anlık güç miktarının yüksek doğrulukta analiz edilmesi çok daha önemli hale gelmiştir. Enterkonnekte sistemin yatırım planlamaları açısından uzun dönemli talep tahminlerinin yapılması da önemlidir.

Elektrik gücü tüketimi, zaman, meteorolojik şartlar, nüfus ve ekonomik faktörlere göre değişkenlik göstermektedir.

Bu çalışmada, 2014-2019 yılları arası dönem için, Bursa iline ait TEİAŞ'tan temin edilen saatlik güç tüketim verilerinin yanında, Bursa Meteoroloji İl Müdürlüğü'nden temin edilen saatlik sıcaklık, nem ve basınç verileri kullanılarak elektrik gücü talep tahmin analizi yapılmıştır.

Çalışma kapsamında, elektrik gücü talep tahmini uygulaması için regresyon, yapay sinir ağları, regresyon ağacı, destek vektör regresyon, gaussian proses regresyon yöntemleri kullanılmıştır. Bu yöntemler için veri doğruluğu ve analiz hızları açısından yapılan karşılaştırmada, gaussian proses regresyon yönteminin daha doğruluklu, yapay sinir ağları ve regresyon ağacı yöntemlerinin ise daha hızlı sonuç ürettiği gözlenmiştir.

ANAHTAR KELİMELELER: Elektrik yük talebi, yük talep tahmini, elektrik yük talep kestirimi, elektrik dağıtımı.

ABSTRACT

**ELECTRIC LOAD DEMAND FORECASTING FOR BURSA PROVINCE
MSC THESIS
MUSTAFA EREN KOCBEY
BALIKESIR UNIVERSITY INSTITUTE OF SCIENCE
ELECTRICAL AND ELECTRONICS ENGINEERING
(SUPERVISOR: ASSIST. PROF. DR. MEHMET KUBILAY EKER)
BALIKESİR, JUNE 2019**

In order to operate the interconnected system which carries current in the sinusoidal form, the network frequency must remain within acceptable limits. The frequency change in grid depends on the load demand of the consumers and amount of power transmitted by plants to the system. Forecasting of the instantaneous load demand in order to maintain the balance of generation and consumption power is an important issue.

Load forecasting is not only important in determining the amount of electricity generation, but it is also important in terms of appropriate electricity pricing. Because of the increasing trend towards renewable energy sources in recent years, it has become much more important to analyze the amount of instantaneous power demand with high accuracy. It is also important to make long-term load forecasts in terms of investment planning of the interconnected system.

Electrical power consumption also varies according to time, meteorological conditions, population and economic factors. In this study, electrical load demand forecast analysis was performed using hourly power consumption data obtained from TEIAS of Bursa province besides hourly temperature, humidity and pressure data obtained from Bursa Meteorology Provincial Directorate for the period between 2014-2019 years.

Regression, artificial neural networks, regression tree, support vector regression, gaussian process regression methods were used for the estimation of electric load demand in this study. In the comparison of these methods, it was observed that the gaussian process regression method produced more accuracy results, artificial neural networks and regression tree methods yield faster results.

KEYWORDS: Electric load demand, load demand forecast, electric load demand estimation, electric power distribution.

İÇİNDEKİLER

Sayfa

ÖZET	i
ABSTRACT	ii
İÇİNDEKİLER	iii
ŞEKİL LİSTESİ	v
TABLO LİSTESİ	vi
KISALTIMA LİSTESİ	vii
SEMBOL LİSTESİ	ix
ÖNSÖZ	x
1. GİRİŞ	1
1.1 Tahmin ve Önemi	2
1.2 Elektrikte Talep Tahmini	3
1.3 Türkiye’de Yük Talep Tahmini	4
1.4 Enterkonnekte Sistemde Arz Talep İlişkisi	5
1.5 Yük Talebini Etkileyen Faktörler	8
1.5.1 Meteorolojik Etmenler	8
1.5.1.1 Hava Sıcaklığı	8
1.5.1.2 Bağıl Nem	9
1.5.2 Ekonomik Faktörler	9
1.5.3 Sosyal Faktörler	10
1.5.4 Beklenmeyen Faktörler	10
1.6 Elektrikte Talep Tahmini Konusunda Literatür Taraması	11
2. TALEP TAHMİNİNDE İSTATİSTİKSEL METODLAR	16
2.1 Regresyon Analizi	16
2.1.1 Doğrusal Regresyon Analizi	16
2.1.2 Eğrisel Regresyon Analizi	18
2.2 Yapay Sinir Ağları	19
2.3 Destek Vektör Makineleri	21
2.3.1 Destek Vektör Regresyon	24
2.4 Gauss Süreç Regresyonu	26
2.5 Karar Ağacı	28
2.5.1 Regresyon Ağacı	32
2.6 Hata Fonksiyonları	34
2.6.1 Ortalama Hata	34
2.6.2 Ortalama Yüzde Hata	35
2.6.3 Ortalama Mutlak Hata	35
2.6.4 Ortalama Hata Karesi	35
2.6.5 Kare Ortalamalarının Karekökü	36
2.6.6 Ortalama Mutlak Yüzde Hata	36
2.6.7 Simetrik Ortalama Mutlak Yüzde Hata	36
3. BURSA İLİNE AİT ELEKTRİK YÜK TALEP TAHMİNİNİN	37
YAPILMASI	37
3.1 UEDAŞ Tarafından Yapılan Yük Talep Tahmini Sonuçları	37
3.2 Bursa İli’nin Geçmiş Yıllara Göre Elektrik Tüketimi	40
3.3 Bursa İli Elektrik Yük Talep Tahmini Uygulaması	42
3.3.1 Verilerin Ön İşlemden Geçirilmesi	42

3.3.2	Regresyon ile Yük Talep Tahmini.....	46
3.3.3	Yapay Sinir Ağları Yöntemi ile Yük Talep Tahmini	49
3.3.4	Destek Vektör Regresyon Yöntemi ile Yük Talep Tahmini	51
3.3.5	Regresyon Karar Ağacı Yöntemi ile Yük Talep Tahmini	53
3.3.6	Gauss Süreç Regresyonu Uygulaması	56
4.	SONUÇ VE ÖNERİLER.....	58
5.	KAYNAKLAR.....	62
6.	EKLER.....	71
	EK A: Bursa ili 2 günlük tüketim verileri	71
	EK B: Bursa ili 1 aylık sıcaklık, basınç, nem verisi örneği	72
	EK C: Microsoft ® Excel ® den MATLAB ® e Veri Aktarım Kodu.....	75
	EK D: Tahmin Parametrelerinin Oluşturulmasını Sağlayan	
	MATLAB ® Kodu.....	82
	EK E: Lineer Regresyon MATLAB ® Kodu	88
	EK F: ANN Eğitim Performansı Grafiği	91
	EK G: ANN Eğitim Durumu	92
	EK H: Regresyon Ağacı Kodu.....	93
	EK I: Bagged Regresyon Ağacı MATLAB ® Kodu	95
	EK J: SVR MATLAB ® Kodu.....	97
	EK K: GPR MATLAB ® Kodu.....	99

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1: Doğrusal regresyon doğrusu.	18
Şekil 2.2: Tek katmanlı yapay sinir ağı.	20
Şekil 2.3: Bir sinir ağının akış diyagramı.	20
Şekil 2.4: Doğrusal olarak ayrılan destek vektörü.	21
Şekil 2.5: Doğrusal olarak ayrılmayan destek vektörü.	23
Şekil 2.6: Destek vektör regresyon örneği.	25
Şekil 2.7: GP prior SE kernel (a) ve GP posterior (b) karşılaştırması.	27
Şekil 2.8: Basit karar ağacı I.	28
Şekil 2.9: Basit karar ağacı II.	29
Şekil 2.10: Örnek regresyon karar ağacı.	32
Şekil 2.11: Örnek kümeleme yöntemi ile ayırıştırma.	33
Şekil 3.1: UEDAŞ 2017 yılı aylık demant grafiği.	39
Şekil 3.2: 2014-2019 Bursa saatlik yük tüketimi.	44
Şekil 3.3: İki günlük talep için tahmin parametrelerinin etkisi (31 Aralık 2018-1 Ocak 2019).	47
Şekil 3.4: 21 Nisan 2018 Doğrusal regresyon tahmin sonucu karşılaştırma. ...	48
Şekil 3.5: 2018-2019 Doğrusal regresyon tahmin sonucu.	49
Şekil 3.6: ANN modeli 49	49
Şekil 3.7: ANN tahmin grafiği 50	50
Şekil 3.8: ANN modeli uyumu (2019 ilk hafta). 51	51
Şekil 3.9: Gaussian Kernel SVR modeli tahmin sonuçları. 52	52
Şekil 3.10: Gaussian Kernel SVR modeli tahmin sonuçları (12 Parametre). ... 53	53
Şekil 3.11: Parametre optimizasyonu ile SVR 1 Ocak 2019 tahmin sonucu karşılaştırma (19 Parametre). 53	53
Şekil 3.12: Regresyon karar ağacı ile oluşturulan modelin tahmin sonuçları. . 54	54
Şekil 3.13: Regresyon karar ağacı 21-28 Şubat 2018 tahmin sonuçları. 55	55
Şekil 3.14: Bagged regresyon DT ile oluşturulan modelin tahmin sonuçları... 55	55
Şekil 3.15: Bagged regresyon ağacı Ağustos 2018 15 günlük tahmin sonucu. 56	56
Şekil 3.16: GPR 2018-2019 tahmin sonucu. 57	57
Şekil 3.17: GPR 21-28 Şubat 2018 tahmin sonucu. 57	57
Şekil 4.1: Yüke bağlı 1 haftalık tahmin karşılaştırması. 58	58
Şekil 4.2: Yüke bağlı 1 Ocak 2019 günlük tahmin karşılaştırması. 59	59
Şekil 4.3: Yük ve meteorolojik veriye bağlı 1 haftalık tahmin karşılaştırması. 59	59
Şekil 4.4: Yük ve meteorolojik veriye bağlı bir günlük tahmin karşılaştırması .. (1 Ocak 2019). 60	60

TABLO LİSTESİ

Sayfa

Tablo 2.1: Karar ağacı hava tahmini.....	29
Tablo 3.1: 2017 Tüketici türüne göre dağılım (MWh-%).....	37
Tablo 3.2: UEDAŞ 2007-2016 tüketim miktarı.....	38
Tablo 3.3: Brüt elektrik tüketimi tahmini (GWh).....	38
Tablo 3.4: UEDAŞ 2017 demant tablosu.....	39
Tablo 3.5: Bursa ili tüketim verileri (a), (b).....	40
Tablo 3.6: Bursa ili meteorolojik verileri.....	41
Tablo 3.7: Türkiye zaman dilimi değişim tarihleri.....	43
Tablo 3.8: LR ile parametre sayısının etkisinin gözlenmesi.....	46
Tablo 3.9: Yıllara göre regresyon analizi MAPE kıyaslama (%).	47
Tablo 3.10: Parametrelere göre regresyon analizi MAPE kıyaslama (%).	48
Tablo 3.11: ANN MSE Hata karşılaştırma.....	50
Tablo 3.12: Karar ağacı hata fonksiyonu karşılaştırma.....	56
Tablo 4.1: MAPE hata fonksiyonu karşılaştırma (%).	58
Tablo 4.2: Algoritma eğitim süresi karşılaştırma.....	60

KISALTMA LİSTESİ

ANN	: Yapay Sinir Ağları (Artificial Neural Network)
BPNN	: Geri yayılım sinir ağları (Backpropagation Neural Network)
BRT	: Torbalama Regresyon Karar Ağacı (Bagged Regression Tree)
CART	: Sınıflandırma ve Regresyon Ağaçları (Classification and Regression Trees)
CSV	: Virgülle Ayrılmış Değişken (Comma Separated Value)
DMD	: Dinamik Mod Dekompozisyon (Dynamic Mode Decomposition)
EMO	: Elektrik Mühendisleri Odası
EPIAŞ	: Enerji Piyasaları İşletme A. Ş.
EPK	: Elektrik Piyasası Kanunu
EPDK	: Enerji Piyasası Denetleme Kurulu
FS	: Özellik Seçimi (Feature Selection)
GA	: Genetik Algoritma (Genetic Algorithm)
GHS	: Evrensel Uyum Arama Algoritması (Global Harmony Search Algorithm)
GP	: Gaus Proses (Gaussian Process)
GPR	: Gaus Proses Regresyon (Gaussian Process Regression)
HSA	: Uyum Arama Algoritması (Harmony Search Algorithm)
IEA	: Uluslararası Enerji Ajansı (International Energy Agency)
LDA	: Doğrusal Diskriminant Analizi (Linear Discriminant Analysis)
libSVM	: Destek Vektör Makineleri için Kütüphane (Library for Support Vector Machine)
LMS	: En Küçük Ortalama Kareler (Least Means Square)
LR	: Doğrusal Regresyon (Linear Regression)
LS	: En Küçük Kareler (Least Squares)
MAE	: Ortalama Mutlak Hata (Mean Absolute Error)
MAPE	: Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error)
ME	: Ortalama Hata (Mean Error)
MÖ	: Milattan Önce (B.C)
MPE	: Ortalama Yüzde Hata (Mean Percentage Error)
MSE	: Ortalama Kare Hata (Mean Squared Error)
NASA	: Ulusal Havacılık ve Uzay Dairesi (The National Aeronautics and Space Administration)
OECD	: Ekonomik İşbirliği ve Kalkınma Örgütü (Organization for Economic Co-operation and Development)
OSB	: Organize Sanayi Bölgesi
PSO	: Parçacık Sürü Optimizasyonu (Particle Swarm Optimization)
RBFN	: Radyal Temelli Fonksiyon Ağı (Radial Basis Function Network)
RMSE	: Kare Ortalamaların Kare Kökü Hata (Root Mean Square Error)
sMAPE	: Simetrik Ortalama Mutlak Yüzde Hata (Symmetric Mean Absolute Percentage Error)
SE	: Üssel Kare (Squared Exponential)
SS	: Karelerin toplamı (Sum of Square)
STD	: Standart deviasyon (Standard Deviation)
SVM	: Destek Vektör Makinesi (Support Vector Machine)
SVO	: Sıralı Minimal Optimizasyon (Sequential Minimal Optimization Algorithm)
SVR	: Destek Vektör Regresyon (Support Vector Regression)

- QDA** : Kuadratik Diskriminant Analizi (Quadratic Discriminant Analysis)
TEİAŞ : Türkiye Elektrik İletim Anonim Şirketi
UCTE : Elektrik İletim Koordinasyon Birliđi (Union for the Coordination of Transmission of Electricity)
UEDAŞ : Uludađ Elektrik Dađıtım Anonim Şirketi
UTC : Uluslararası zaman (Coordinated Universal Time)
XLS : Excel Spreadsheet

SEMBOL LİSTESİ

β	: Regresyon doğrusu eğimi
n	: Regresyonda hata terimi
w	: Yapay sinir ağında ağırlık
C	: SVM’de Cortes ve Vapnik’e göre serbest değişken
ξ_i	: SVM’de gevşek değişken (Slack variable)
ε	: SVM’de Sınır düzlemin hiper düzleme olan uzaklığı
ω	: SVM’de ağırlık vektörü
v	: Öz vektör (Schölkopf)
ρ	: SVR’de bir katsayı
l	: SVR’de bir katsayı
α	: SVR’de modelde eğitim gözlemindeki bir katsayı (gevşek değişken)
b	: SVR’de modelde eğitim gözlemindeki bir katsayı (gevşek değişken)
a	: Eğitim gözlemindeki bir katsayı (gevşek değişken)
Q	: Rastgele yürüyüş dizisi (Random Walk Sequence)
L	: Doğrusal olarak ayrışan destek vektör fonksiyonu (lagrangian fonksiyonu)
Φ	: Çok boyutlu uzay transferi fonksiyonu
K	: GPR’da Kovaryans matrisi
\mathbb{E}	: Öklid uzayı (Euclidean space)

ÖNSÖZ

Bu çalışma kapsamında kestirimci yaklaşım ile mevcut yük verileri ve çevresel parametreler göz önünde bulundurularak Bursa iline ait yük talep tahmin analizi gerçekleştirilmiştir. Analizdeki temel amaç mevcutta yaygın olarak kullanılan veri madenciliği yöntemleri incelenerek elektrik tüketimini etkileyen en uygun parametrelerin belirlenmesi ve en etkin algoritmanın belirlenmesidir.

Üç bölüme ayrabilecek bu tezin birinci kısmında yük talebi tahminin önemi ve elektrik tüketimini etkileyen parametreler, ikinci bölümde tahmin uygulamalarında yaygın olarak kullanılan yöntemlerin teorik olarak incelenmesi, üçüncü bölümde ise Bursa iline ait elektrik yük talebi tahmini uygulamalarının yapılması ele alınmaktadır.

Bu çalışmayı hazırlarken desteğini her an yanımda hissettiğim canım eşim Dr. Fulya KOÇBEY'e, bana hayatımın her anında destek olan aileme, bana birçok alanda özgün bakış açısı kazandıran matematik öğretmenim Şevket ONGUN'a ayrıca yüksek öğrenim hayatımda ve çalışmalarımnda benden yardımlarını esirgemeyen danışmanım Dr. Öğr. Üyesi M. Kubilay EKER'e teşekkür ederim.

Balıkesir, Haziran 2019

Mustafa Eren KOÇBEY

1. GİRİŞ

Ağırlıklı olarak senkron generatörler üzerinden sinüzoidal bir sinyal yapısına sahip olarak üretilen elektrik enerjisinin bu haliyle depolanması mümkün değildir. Senkron generatörlerin paralel çalışmasında kararlılığın devam ettirilebilmesi için de şebeke frekansının kabul edilebilir limitler dışına çıkılmadan enterkonnekte sistemin çalışması sağlanmalıdır. Elektrik Şebeke Yönetmeliği'nde belirtildiği gibi şebeke frekansının 49,8 Hz ile 50,2 Hz aralığında TEİAŞ tarafından kontrol edilmesi gerekmektedir [1]. Şebekenin frekansını belirleyen senkron generatörün frekans değişimi, çekilen güçteki değişimle orantılı olarak değişmektedir. Şebeke frekansının anma değerinde kalabilmesi için tek yol, anlık olarak çekilen elektrik gücü miktarınca gücün üretici generatörler tarafından enterkonnekte sisteme aktarılmasıdır. Kısacası, anlık olarak ne kadar elektrik gücü tüketimi varsa o kadar gücün santraller tarafından şebekeye aktarımının sağlanması enterkonnekte sistemin sürekliliği açısından önemlidir. Elektrik Piyasası Talep Tahminleri Yönetmeliği'nde belirtildiği üzere uzun süreli tüketici talep tahmini yapma sorumluluğu TEİAŞ'ın yanında OSB ve dağıtım şirketleri için de mevcuttur [1]. TEİAŞ bu tahminleri değerlendirerek 10 yıllık talep tahminleri raporunu hazırlamaktadır. TEİAŞ ayrıca Yük Tevzi Bilgi Sistemi [2] üzerinden saatlik çekilen puant değerlerini de yayınlamaktadır.

Elektrik tüketiminde talep tahmini, uygun fiyatlandırma açısından da önem arz etmektedir. İhtiyaç duyulan güç miktarı saatlik olarak TEİAŞ Piyasa Yönetim Sistemi [3] üzerinden ilan edilmekte, elektrik enerjisi piyasa pazarının %30'unu oluşturan Gün Öncesi Piyasası üzerinden uygun teklifler dikkate alınarak alımlar yapılmaktadır [4].

Elektrik talebi zamana, çevresel ve ekonomik faktörlere göre değişmektedir. Şehirselleşme bazlı değerlendirildiğinde bir şehrin nüfus durumu, mevsimsel şartlar, sanayi üretimindeki değişiklikler, ülke bazlı değerlendirildiğinde ise ülkenin ekonomik durumu, geçmiş yıllara göre büyüme miktarı, nüfus değişimi gibi etmenler elektrik talebini değiştirmektedir.

Özellikle yenilenebilir enerji kaynaklarına yönelimin artması ile talep edilecek güç miktarının daha doğru olarak analiz edilmesi önem kazanmıştır. Yenilenebilir enerji kaynaklarından yapılan elektrik üretiminin diğer santrallerden farklı olarak kontrollü bir şekilde değiştirilebilen bir büyüklük olmadığı unutulmamalıdır.

Bu çalışma kapsamında kestirimci yaklaşım ile mevcut veriler ve çevresel parametreler göz önünde bulundurularak Bursa iline ait yük talep tahmin analizi gerçekleştirilmiştir. Çalışma kapsamında 2014-2019 yılları arasındaki Bursa/TÜRKİYE'ye ait saatlik veriler kullanılmış olup, çalışma günleri ve resmî tatiller de analize dahil edilmiştir.

Çalışma üç bölümden oluşmaktadır. Birinci bölümde; konu ile ilgili temel kavramlar ve veri madenciliği tanımları genel olarak özetlenmiştir.

İkinci bölümde; veri madenciliği yöntemlerinin kestirimci yaklaşım alanında ve talep tahmininde kullanılan metotlardan en yaygın kullanılanları teorik olarak anlatılmıştır.

Üçüncü bölümde; Bursa ili elektrik tüketiminin gelecekteki miktarının tahminini yapmak üzere modeller tasarlanmış ve analizleri yapılmıştır.

Sonuç bölümünde, çalışmadan elde edilen tespitler doğrultusunda değerlendirmeler ve yorumlar yapılmıştır.

1.1 Tahmin ve Önemi

Toplum ihtiyaçlarının kesintisiz karşılanması ve şartlara daha iyi adapte olunmasını sağlamak amacı ile kestirimci yaklaşım istatistiğinin incelediği en eski alanlardan birisi olmuştur. Kestirimci yaklaşımın tarihteki bilinen bir örneği M.Ö. 650 yılında Babilliler tarafından yapılan kısa zamanlı hava tahminidir. Söz konusu hava tahmininde bulutların hareketleri ve ışık halkalarının oluşumu baz alınarak analiz yapılmıştır [5].

Günümüzde tahmin yöntemleri nicel ya da nitel verilerin hesaplanmasında sıklıkla kullanılmaktadır. Ekonomik planlamadan sanayi uygulamalarına kadar taleplerin karşılanması amacı ile istatistiksel analizlerin yapılması zorunlu hal almıştır. Enerji, meteoroloji, tarım, sanayi, ithalat, ihracat, nüfus artışı, enflasyon gibi

hayatımızı etkileyen birçok alanda kestirimci yaklaşım metotları ile tahmin hesaplamaları yapılmaktadır.

1.2 Elektrikte Talep Tahmini

Dünya genelinde her geçen gün enerji talebi artmaktadır. Toplumun en temel ihtiyaçları arasında yer alan elektrik enerji ihtiyaçlarının karşılanabilmesi amacı ile gerçekleşecek taleplerin belirlenmesi ekonomik anlamda önem arz etmektedir. Elektrik piyasalarında güç kesintilerinin yaşanmaması için ileriye dönük tahminlerin yapılmasının yanında, yatırım önceliklendirme ve kaynak bulunması açısından da talep tahmini önemlidir.

IEA'dan alınan veriler doğrultusunda OECD üyesi ülkelerdeki toplam elektrik enerji tüketimi 2017 yılında %3,1 artarak 25570 TWh'e ulaşmıştır [6]. Dünyada, elektrik enerjisi ihtiyacının giderilmesi büyük oranda fosil ve nükleer kaynaklı santraller kullanılarak yapılmaktadır. Bu elektrik enerjisi tüketiminin 15597 TWh'lik kısmı, kömür ve doğalgaz santrallerinden [6] elde edilmektedir. Fosil yakıtların çevreye verdiği zarar ve nükleer santrallerin sürdürülebilirliği göz önünde bulundurulduğunda dünyada yenilenebilir enerjiye yönelim her geçen gün artmaktadır. Yenilenebilir enerji santrallerinin elektrik enerjisi üretimine katkısı ise tüketim değişimine bağlı olarak kontrol edilebilir bir büyüklük değildir.

Yük talebi analizlerinde yapılan geçmiş dönemlerdeki hesaplamalarda sadece önceki dönemlere ait talep miktarları dikkate alınırken [7]; sıcaklık, nem, büyüme, milli gelir, ihracat, hammadde birim fiyatı gibi veriler de analize eklenerek tahminin tutarlılığı artırılmaya çalışılmaktadır [8]. Söz konusu analizlerin hatalı olması ya da yapılamaması durumunda arz talep dengesi bozulmakta ve elektrik kesintisi ya da bölgesel şebeke çökmelerine sebep olmaktadır. Benzer analizler yenilenebilir enerji üretim kaynakları (rüzgâr ve güneş) açısından da gerçekleştirilerek, etkin ve ekonomik elektrik üretimi planlanmaktadır [9,10].

Yapılan talep tahminleri doğrultusunda, ilgili birimler tarafından ileriye dönük altyapı ve santral yatırımlarının planlanması belirlenebilmektedir. Özellikle gelişmekte olan ve yeterli enerji kaynağına sahip olmayan ülkelerde elektrik enerjisi üretim kaynakları ülke dışından temin edilmekte ve dışa bağımlılık artmaktadır. İleriye

dönük ihtiyaç olması durumunda beklenen enerji miktarı doğrultusunda ithalat anlaşmaları yapılmaktadır.

Yenilenebilir enerji kaynaklarından olan rüzgâr ve güneş enerji santralleri her geçen gün, artan miktarlarda elektrik enerjisi üretim sektöründe yerini almaktadır [11]. 2018 yılında Türkiye'deki elektrik üretiminin, %37,3'ü kömürden, %29,8'i doğal gazdan, %19,8'i hidrolik enerjiden, %6,6'sı rüzgârdan, %2,6'sı güneşten, %2,5'i jeotermal enerjiden ve %1,4'ü diğer kaynaklardan elde edilmiştir [12]. Elektrik talep miktarına karşılık oluşturulacak arzın karşılanmasında yenilenebilir enerji kaynaklarına öncelik verildiği dikkate alındığında yüksek doğrulukta talep tahmininin önemli olduğu açıktır.

Ayrıca yük talep analizi neticesinde belirlenen miktar, söz konusu gün ve saatteki rüzgâr ve güneşlenmeye bağlı elektrik üretim kapasitesi tahminleri ile birlikte değerlendirilmesi mümkün olmakta ve buna bağlı olarak kaynak tüketim türü için önceliklendirme planı yapılması mümkün olmaktadır.

Elektrik talep tahmini analizi yapılırken geçmiş yıllardaki eşdeğer veriler ve enerji tüketimini etkileyen parametreler veri madenciliği esaslarına göre kullanılmaktadır. Analizin detaylandırılması elektrik tüketimini etkileyen parametrelerin de eklenerek değerlendirilmesi esasına dayanmaktadır.

1.3 Türkiye'de Yük Talep Tahmini

Ülkemizde Enerji ve Tabii Kaynaklar Bakanlığı, özellikle gelecekteki yatırım planlarının yapılması, kapasite iyileştirilmesinin yapılması, elektrik arz güvenliğinin sağlanması gibi sebeplerden dolayı 14/3/2013 tarihli ve 6446 sayılı Elektrik Piyasası Kanunu doğrultusunda EPDK ve TEİAŞ denetiminde elektrik talep tahmini yapılmasını zorunlu kılmaktadır. Elektrik Piyasası Kanunu'nda belirtilen yük talep tahminleri; 7/5/2016 tarihli 29705 Sayılı Resmî Gazete'de yayınlanan Elektrik Piyasası Talep Tahminleri Yönetmeliği doğrultusunda gerçekleştirilmektedir. Söz konusu yönetmeliğe göre, elektrik enerjisi talep tahmini yapmakla yükümlü olan kurum ve kuruluşlar TEİAŞ, OSB'ler, dağıtım ve görevli tedarik şirketlerdir [1].

Ayrıca, 6446 sayılı Elektrik Piyasası Kanunu'nun 9. Madde 4. Bendine göre Dağıtım Şirketlerinin dağıtım lisansında belirlenen bölgelerde talep tahminlerinin

hazırlanması ve TEİAŞ'a bildirilmesi zorunluluğu belirtilmektedir. EPDK bu talep tahminlerini gerekli ise revize etmekle ve onaylamakla yükümlüdür. Onaylanan tahminler TEİAŞ tarafından yayınlanmaktadır [13].

Mevcut arz potansiyelini, potansiyel arz imkânlarını, yakıt kaynaklarını, iletim ve dağıtım sisteminin yapısı, gelişme planlarını, ithalat veya ihracat imkânlarını ve kaynak çeşitliliği politikalarını dikkate alarak enerji politikalarının belirlenmesini sağlayan Uzun Dönem Elektrik Enerjisi Üretim Gelişim Planının hazırlanması TEİAŞ'ın sorumluluğundadır (EPK Madde 20/1(a)). TEİAŞ bu raporu, organize sanayi bölgeleri, dağıtım ve görevli tedarik şirketlerinden aldığı talep tahminlerini değerlendirerek bakanlığa sunmaktadır. Bakanlık tarafından oluşturulacak olan arz güvenliği raporu, TEİAŞ'ın oluşturacağı kısa ve orta dönem arz-talep dengesi raporunun da içinde yer aldığı uzun dönem elektrik enerjisi üretim gelişim planı doğrultusunda hazırlamaktadır.

Dağıtım ve tedarik şirketlerinin talep tahmini yapma yükümlülüğü

İlgili yönetmelikte talep tahmini ile ilgili belirli şartlar verilmiştir. Özellikle talep tahmini yapacak olan dağıtım şirketleri gelecek on, tedarik şirketleri de gelecek beş yıllık dönem için yıllık bazda analiz yapmaları gerekmektedir (EPK Madde 6). Söz konusu talep tahmininde esas veri seti; ekonomik, sosyal, demografik, iklimsel, çevresel değişkenler ve açıklanan değişkenin gerçekleşmeleri kullanılarak oluşturulmaktadır (EPK Madde 7/1).

EPK Madde 20/1(b)'ye göre görevli tedarik şirketleri, her yıl aralık ayı sonuna kadar gelecek beş yıl için, tahmin ettikleri elektrik enerjisi puant güç taleplerini, ihtiyaç duydukları elektrik enerjisi miktarını, bu miktarın temini için yaptıkları sözleşmeleri ve ilave enerji veya kapasite ihtiyaçlarını EPDK'ya bildirmek zorundadırlar. Oluşturulan bu hiyerarşi ile EPDK'nın sistemdeki arz ve talepleri dengeli bir şekilde yönetmesi sağlanabilmektedir.

1.4 Enterkonnekte Sistemde Arz Talep İlişkisi

Elektrik üretiminde enterkonnekte sisteme bağlı olan santraller birbirleri ile senkron olarak eşdeğer gerilim ve frekans aralığında çalışmakta ve mevcuttaki tüketim ile dengeli bir üretim gerçekleşmesi gerekmektedir.

Toplam talebin karşılanamaması öncelikle bölgesel kesintilere sebep olurken, talepten daha fazla üretim gerçekleşmesi, ulusal ve/veya uluslararası şebekelerde aşırı yüklenmeye, şebeke çökmesine sebep olabilmektedir.

Özellikle ani taleplerin karşılanamaması durumunda enerji açığı meydana gelebilmekte ve frekans düşmesi yaşanabilmektedir. Bu durumda enterkonnekte sisteme bağlı olan bütün santraller riske girebilmekte ve şebeke çökmesi durumu ile karşılaşılabilir. Bunun yanı sıra bir enterkonnekte sistemde üretimin fazla olması enerji sarfiyatını artırırken sistemdeki güç kalitesini de olumsuz etkilemektedir.

Ayrıca dağıtım sistemlerinde hatların kapasitelerine bağlı olarak doğru yönetilememeleri hatta arızalara yol açabilmekte ve buna bağlı olarak enterkonnekte sistemin zarar görmesi durumu ile karşılaşılabilir [14]. Özellikle sistemin çökmesi durumunda birbirine bağlı diğer enterkonnekte sistemler etkilenmekte, çöken sistemi yeniden devreye alma süresi saatler sürebilmektedir. Bu durum da kurum ve kuruluşlara büyük mali kayıplara ve cezai yaptırımlara sebebiyet vermektedir.

Taleplerin arızaya bağlı olarak karşılanamaması sonucu 28 Eylül 2003 yılında Avrupa'da aşırı yüklenmeye bağlı elektrik kesintileri yaşanmıştır. İsviçre'de yer alan 380 kV'luk Mettlen-Lavorgo iletim hattının İtalya tarafından 300 MW fazla yük talebi doğrultusunda %86 seviyesinde yüklenmesine bağlı olarak meydana gelen ısıdan dolayı sarkması, rüzgârın etkisi ile bölgede bulunan ağaçlar üzerinden faz-toprak arızası meydana getirmiştir. Tek fazın kapatılması girişimlerinin yetersiz olması sebebi ile koruma ekipmanları hattı devre dışı bırakmıştır. Olay esnasında hattın manuel olarak devreye alınması yüksek yüke bağlı olarak faz açısının fazla olması sebebi ile koruma elemanları tarafından engellenmiştir. Matten - Lavorgo iletim hattının devre dışı kalması ile buradan akan yükün paralel iletim hatları tarafından taşınması diğer iletim hatlarının yükünü artırmıştır. İsviçre'de yer alan 380 kV'luk Sils-Soazza iletim hattı nominal kapasitesinin %110 seviyesine ulaşmış ve söz konusu aşırı yüklenmenin giderilememesi sebebi ile iletim hattı sarkarak faz-toprak arızası meydana getirmiştir. Koruma ekipmanları hattı açarak devre dışı bırakmıştır. İsviçre'deki iki büyük iletim hattının devre dışı kalması ile talep edilen yük diğer hatlardan karşılanmaya çalışılmış ve aşırı yüklenmeye bağlı olarak İsviçre'nin kuzey bölgesinin ulusal şebekeden kopmasına sebep olmuştur [14].

İsviçre-İtalya arasındaki iletim hattının kopması ile birlikte talep edilen yük İtalya-Fransa, İtalya-Slovenya, İtalya-Avusturya enterkonnekte sisteminden karşılanmaya çalışılmıştır. Aşırı yüklenmeye bağlı olarak şebekede meydana gelen dalgalanmalar sebebi ile iletim hatları devre dışı kalmıştır. İtalya-Fransa hattının devre dışı kalması ile İtalya UCTE'den tamamen ayrılmıştır. İtalya ulusal şebekesindeki büyük üretim açığı beraberinde frekansın düşmesine sebep olmuştur. Frekans düşüşü ile birçok santral devre dışı kalmış ve akabinde İtalya'nın enerji sistemi tamamen çökmüştür [14].

Söz konusu örnek olayda bir hattın kapasitesinin üzerinde yüklenmesi ile hatta öngörülemez bir arıza oluşturmuş ve hattın devre dışı kalmasına sebebiyet vermiştir. Hattın devre dışı kalması ile arzda oluşacak yük fazlalığı alternatif hatlar ile karşılanmaya çalışılsa da hatların kapasitesinin üzerinde yüklenmesine sebep olmuş ve diğer hatların devre dışı kalmasına sebebiyet vermiştir. İtalya içerisindeki arzın yetersiz olması frekans düşüşüne sebep olmuş ve bu durum sistemin çökmesine sebebiyet vermiştir. Bu noktada UCTE ile bağlı diğer şebekelerde de dengesizlikler oluşmuş ve gerilim dalgalanmaları meydana gelmiştir. Söz konusu örnekte olduğu gibi öngörülen ve tasarlanan sistemin ileriye dönük olarak tahmin metotları ile takip edilmesi, arızaların engellenmesi ve kötü durum senaryolarının geliştirilmesi ile bu duruma benzer sorunların meydana gelmesi önceden öngörülerek engellenebilecektir.

Avrupa'da 2004 yılında yaşanan başka bir olayda, Almanya'da rüzgâr ile elektrik üretiminin beklenilenden daha yüksek olmasına bağlı olarak Hollanda iletim hatlarında aşırı yüklenmeye sebep olmuş ve enterkonnekte sistemi tehlikeye sokmuştur [15].

Talep dengesinin sağlanması ile ilgili farklı bir olay 28 Ekim 2017 tarihinde yaşanmıştır. Söz konusu tarihte Almanya'da rüzgâr enerjisinden maksimum anlık 39190 MWh elektrik üretimi gerçekleşmiştir. Söz konusu üretim miktarı beklenenin üzerinde olduğu için talep dengesinin sağlanması amacı ile fiyat tarifesinde indirim yapılarak tüketim artırılması sağlanmıştır [16]. Gerçekleştirilen rüzgâr ve güneş tahminleri neticesinde üretim miktarının önceden belirlenerek planlama yapılmış olup arz-talep dengesizliği yaşanmaması adına elektrik birim fiyatında indirim sağlanarak denge sağlanması çözümüne gidilmiştir.

Ulusal ve/veya uluslararası enterkonnekte sistemlerin sağlıklı bir şekilde çalışması gereğince talep edilecek elektrik enerjisi miktarının, üretimin öncesinden tahmin edilmesi önemli bir yer teşkil etmektedir. Bu sayede mevcuttaki sistemin güvenliği sağlanırken kesintisiz enerji aktarımı sağlanabilmekte ve kurum ya da kuruluşların ileriye dönük yatırım planlarını yapmasına olanak tanımaktadır.

1.5 Yük Talebini Etkileyen Faktörler

Bir bölgenin yük talebini hava sıcaklığı, nemi, basıncı gibi meteorolojik etmenler; sanayi ve ticaretin durumu, nüfus artışı ve sosyal gelişmişlik düzeyi gibi parametreler doğrudan etkilemektedir. Elektrik tüketimi, meteorolojik, zamansal ve ekonomik faaliyetlerden etkilenen durağan olmayan ve rastlantısal bir zaman serisi olarak düşünülebilir [17].

Türkiye'deki 2000-2017 yılları arasındaki puant talebine [18] bakıldığında sanayinin ve nüfusun artışına bağlı olarak tüketimde bir artış durumu görülmektedir. Ayrıca yıllar içindeki artış oranları birbirine yakın olup 2017 yılı göz önünde bulundurulduğunda en yüksek tüketimin temmuz ayı olduğu, saatlik bazda bakıldığında ise yıl içi toplam kullanımın en yüksek olduğu saatin de 11:00 olduğu gözlemlenmektedir.

1.5.1 Meteorolojik Etmenler

Meteorolojik faktörler yük karakteristiğinin oluşmasında önemli parametre olarak rol almaktadır. Birçok tesis ısıtma, soğutma ya da tarımsal sıcaklık dengesi gibi hava durumuna bağlı ekipman kullanmaktadır [19].

Temel yük, hava değişkenlerinin yükselmesi ya da düşmesi ile etkilenmektedir. Yük ile hava değişkenleri arasındaki ilişki doğrusal olmadığı için talep tahmininde önemli bir parametredir [20].

1.5.1.1 Hava Sıcaklığı

Özellikle uzun süreli yük talebi analizlerinde sıcaklık tepe değerlerindeki talep tahmininde hava sıcaklığı en önemli parametrelerden birisidir [8]. Hava sıcaklığı

özellikle ısıtma-soğutma için elektrik kullanılan ve tarımsal sulama yapılan bölgelerde önemli tüketim artışına sebep olmaktadır. Ayrıca anlık sıcaklığın yanı sıra birbirini takip eden günlerdeki sıcaklık durumları da sistemde farklılıklara yol açmaktadır [17]. Örneğin hava sıcaklığının yüksek olduğu bölgelerde bireysel ya da ticari olarak ortam soğutma doğrultusunda tüketim miktarı artarken hava sıcaklığının çok düşük olduğu dönemlerde tam aksine ısınmaya bağlı elektrik tüketimi artmaktadır. Bu noktada talep tahmini doğruluğunu sıcaklık birinci dereceden etkilemektedir.

1.5.1.2 Bağıl Nem

Bağıl nemin elektrik tüketimine en büyük etkisi; insan vücudu düşünüldüğünde fazla nemin insan vücudunu rahatsız etmesine bağlı olarak özellikle yaz aylarında kullanımı artan iklimlendirme sistemleri vasıtası ile olmaktadır.

Hava sıcaklığının 20 °C'nin üzerinde olduğu yaz aylarında, yüksek nem ile yüksek sıcaklık kombinasyonu kişilerin kullanmış olduğu klimaların daha fazla soğutma yükü talebine sebep olmaktadır. Ayrıca nemin ve sıcaklığın yüksek olması, soğutma yükünün yüksek olmasına ve bu durumda söz konusu soğutma yükünün karşılanabilmesi için daha fazla elektrik ihtiyacına sebep olmaktadır. Bunun yanı sıra düşük seviyedeki bağıl nem ihtiyaç olan soğutma yükünün düşük olmasını sağlamaktadır [21].

1.5.2 Ekonomik Faktörler

Ekonomik düzen elektrik talebini etkileyen önemli faktörlerden bir tanesidir [19]. Elektrik tüketiminin büyük çoğunluğu ticari ve tarımsal alanda kullanılmaktadır. Özellikle endüstrinin gelişmesi tüketimi birinci dereceden etkileyen faktörler arasında yer almaktadır. Bu sebeple bir ülkenin genel olarak ekonomik durumundaki iyileşme, örneğin ihracatın artması, sanayinin birçok alanında üretimin arttığını göstermekte, üretimin artması da benzer şekilde enerji tüketimini artırmaktadır.

Özellikle tarımın geliştiği bölgelerde, mevsimsel olarak dönemler ele alındığında, elektrik tüketiminde tarımın etkisi de ortaya koyulabilir. Ayrıca yaz veya

kış turizmi, bölgelere bağlı olarak elektrik tüketimini etkileyen önemli ekonomik parametredir.

İthalat ve ihracat miktarları, ülke ekonomisini ve sanayisini birinci dereceden etkilediği için elektrik tüketim miktarını belirleyen parametreler arasında yer almaktadır.

1.5.3 Sosyal Faktörler

Bir ülke ya da bölge içerisinde bireylerin gelir düzeylerinin artışı ve buna bağlı olarak iç ticaretin gelişmesi de elektrik tüketimini artıran parametreler arasındadır. Bu durum hizmet sektörünün de gelişmesine ve ticaret alanında elektrik tüketiminin artmasına sebep olmaktadır.

Ayrıca kişi başına düşen gelirin artması, toplumun refah düzeyinin arttığını göstermektedir. Buna bağlı olarak bireylerin sosyal aktiviteleri, teknolojiye olan erişimlerini artırmakta ve bu durum da enerji tüketimini artıran parametreleri oluşturmaktadır.

Elektrikli araçların kullanımının yaygınlaşması, petrol tüketimini azaltmakta iken elektrik tüketiminde artışa sebep olmaktadır. Bu sebeple talep tahminlerinde, özellikle elektrikli araçların kullanımının yüksek olduğu ülkelerde şarj istasyonları elektrik tüketimi analiz sonucunu etkileyecek parametreler arasında yer almaktadır.

1.5.4 Beklenmeyen Faktörler

Elektrik tüketiminde tüketiminin önceden tahmin edilmesini güçleştiren değişken parametreler bulunmaktadır. Büyük yük tüketiminde bulunan demir-çelik fabrikaları, rüzgâr tünelleri¹ veya parçacık hızlandırma deneyleri gibi çalışma ve sektörlerin taleplerini öngörmek oldukça güçtür [19].

Ayrıca uluslararası ticarete ülkeler arası ilişkiler de elektrik tüketiminde öngörülemeyen faktörleri oluşturabilir. İki ülke arasındaki sıcak ilişkilerin zedelenmesi ticareti olumsuz etkileyeceği için sanayi üretimi de azaltacaktır. Bu durum üretimde öngörülemeyen bir düşüş olarak tahmin analizini etkileyecektir.

¹ Hava akışlarına ilişkin kuramsal ve sayısal bilgilerin elde edilmesini sağlayan, hava akımının nesnelere veya ortamlara etkisinin incelendiği deney ortamı

Uluslararası ticaret savaşları veya ambargolar ve akabinde alınan ani ticaret kararları neticesinde, üretim kapasiteleri aniden düşürülebilmektedir. Örneğin demir çelik sektöründe ya da turizm faaliyetlerinde bu durumun yaşanması şebekede tahmin edilemeyen talep düşüşüne sebebiyet verecektir.

1.6 Elektrikte Talep Tahmini Konusunda Literatür Taraması

Yük talebi tahmininde genellikle regresyon ve makine öğrenme yöntemleri kullanılmaktadır [22]. Özellikle teknolojinin ilerlemesi ve yapay zekâ çalışmalarının artması ile makine öğrenme yöntemleri talep tahminlerinde etkin sonuçlar vermektedir [8].

Dünya’da elektrik yük talepleri ile ilgili birçok tahmin çalışması yapılmaktadır [23]. Bursa bölgesinde en büyük dağıtım firması olan UEDAŞ, TEİAŞ’a gerçekleştirmiş olduğu yük talebi analizlerini EPK doğrultusunda sunmaktadır.

Chen ve arkadaşları [24], yük tahmini konusunda en küçük kareler destek vektör makineleri ile fuzzy zaman serileri ve harmoni arama algoritması kullanmıştır. Bu çalışmada destek vektör makinelerinin tek başına giriş değişkenlerinin etkinliğini belirlemenin yetersiz olması, tahmin sonuçlarının zayıf olduğu ve yavaş çalıştığı eleştirilmiştir. Yapılan analiz sonucunda oluşturulan hibrit algoritma ile MAPE hata fonksiyonu %3,709 olarak tespit edilmiş ve karşılaştırılan GHSA, HSA, PSO ve GA algoritmalarına göre daha iyi çalıştığı belirlenmiştir.

Kaynar ve arkadaşları [25], destek vektör regresyon algoritması ile kaotik yöntemlerin bir arada kullanıldığı hibrit algoritmanın Türkiye’de yük talep tahmininde %5’den daha küçük hata ile çalıştığını tespit etmiş ve klasik regresyon yöntemlerine ve yaygın olarak kullanılan yapay sinir ağları ile oluşturulan algoritmalara alternatif olabileceğini belirtmiştir.

Türkay ve Demren [26], libSVM algoritması ile İstanbul Avrupa Yakasındaki 2006-2009 yılları arasındaki günlük yük verileri kullanılarak talep tahmini uygulaması yapmıştır. Söz konusu uygulamada, günlük ortalama sıcaklık verileri ve tatil günleri de analize dahil edilmiştir. Ellatar ve arkadaşları [27], destek vektör regresyon algoritmasının hibrit modeli LWSVR ile yaptıkları analizde MAPE değerini %3,5’den daha küçük tespit etmiştir. Hussain ve arkadaşları [28], SVM kernel yaklaşımının

özellikle kısa zamanlı yük tahmininde LDA ve QDA algoritmalarına göre çok daha iyi çalıştığını belirtmektedir.

Nazarko ve Zalewski [29], yapmış oldukları fuzzy regresyon ile pik yük tahmini ile ilgili bir çalışmada, fuzzy regresyon ile standart regresyon kullanarak gerçekleştirilen analiz karşılaştırmaları yapmış ve yük talebinde doğrusal fuzzy regresyonun standart doğrusal regresyona göre daha iyi çalıştığı tespitinde bulunmuştur.

Mostafa ve Nagasaka [8], SVM yaklaşımı ve ANN yöntemlerini kullanarak yaptıkları yük talep tahmini uygulamasında, SVM için MAPE değerini %7 bulurken ANN için ise %13,6 bulmuştur. Söz konusu analizde, çalışma günleri değerlendirilerek tahmin yapılmış olup talebi etkileyen sıcaklık, rüzgâr hızı, nem, aydınlanma süresi gibi meteorolojik ve çevresel etmenler ele alınmamış ve hata yüksek bulunmuştur. Çalışmada SVM algoritmalarının ANN'e göre daha iyi çalıştığı bulgusuna varılmıştır.

Sarhani ve El Afia [30] tarafından hibrit çalışma olarak tespit edilen bir diğer çalışmada SVR ile PSA algoritmaları kullanılmış ve özellik seçiminin etkisi gözlenmiştir. Özellik seçimi uygulaması yapıldığında MAPE değeri %5,55 olarak tespit edilmiştir. Söz konusu çalışmada, farklı bir veri üzerinde benzer algoritma kullanılarak ikinci bir analiz yapılmıştır. Bu analiz kapsamında, hava koşullarına ait veriler ve haftanın günleri/resmî tatiller ile ilgili değişkenler de kullanılmıştır. Analiz sonucunda özellik seçimi uygulandığında MAPE değeri %2,9 olarak tespit edilmiştir. Söz konusu analiz doğrultusunda SVR uygulamalarının hibrit kullanımlarının daha iyi sonuç verdiği belirlenmiş ve FS uygulamasının tahmin performansını iyi yönde etkilediği belirtilmiştir.

Matijaš [31] tarafından gerçekleştirilen bir çalışmada, Robust LS-SVM algoritmasının SVR yöntemlerine göre daha iyi çalıştığı öne sürülmüştür. Analiz dört farklı veri setine uygulanmış ve gerekli karşılaştırmalar yapılmıştır. Analiz kapsamında, kullanılan veri ile SVR algoritmasının kabul edilebilecek düzeyde çalıştığı gözlenmiştir. Çalışmada LS-SVM ve Robust LS-SVM sonuçlarının diğer yöntemlere göre çok daha iyi sonuç verdiği ifade edilmiştir.

Božić ve Stojanović [32] tarafından SVM modeli ile yapılan bir çalışmada, takvimsel veri, sıcaklık ve rüzgâr parametreleri de kullanılarak 1 yıllık yük talep verisi

kapsamında orta zamanlı talep tahmini uygulaması dahilinde MAPE değeri %4,25 olarak bulunmuştur. Analiz mevsimsel olarak bölünerek tekrarlanmış ve detaylı olarak incelenmiştir. Analiz kapsamında özellikle yeni yıl ve Christmas döneminde hatanın fazla olduğu gözlenmiş ve bu tatil günleri yıl içerisinde çok az sayıda olduğu için gürültü olarak değerlendirilmiştir.

Setiawan ve arkadaşları [33], SVR kullanılarak gerçekleştirmiş oldukları çok kısa zamanlı elektriksel yük talebi tahmini uygulamasında ise 5 dakikalık yük talep tahmini yapmıştır. Analizde LR, LMS, BPNN ve SVR algoritmaları bir arada kullanılarak karşılaştırma yapılmıştır. Analiz kapsamında birden fazla hata hesaplama yöntemi kullanılmış olup 4 farklı durum ele alınmıştır. Analiz sonucu değerlendirildiğinde SVR yönteminin bütün durumlarda en iyi sonuç verdiği tespit edilmiştir. SVR yönteminin dezavantajı olarak eğitim süresinin LR ve LMS algoritmalarına göre daha uzun olması belirtilmiştir.

Akataş [9], yapay sinir ağları kullanılarak gerçekleştirmiş olduğu uzun süreli elektrik talep tahmini çalışmasında, RBFN yöntemi ile Japonya'daki 9 tesise ait 1975-2000 yıllarına ait veriler kullanmıştır. Söz konusu analizde elektrik tüketimini etkileyen faktör verileri parametre olarak kullanılarak tahmin sonuçları karşılaştırılmıştır. Söz konusu çalışma neticesinde eğitim serisinde MAPE değeri %1,15 bulunurken tahmin sonuçlarında ise %3,465 olarak belirlenmiştir.

Ghanbari ve arkadaşları [34], yapay sinir ağları algoritması ile gerçekleştirdiği çalışmalarında doğrusal regresyon, logaritmik doğrusal regresyon ve ANN sonuçlarını karşılaştırmıştır. Yapılan çalışma neticesinde ANN için MAPE değeri %2,08 tespit edilirken, LR için %27,06, Log-LR için %12,60 bulunmuştur. ANN algoritmasının doğrusal regresyon algoritması ve logaritmik doğrusal regresyon algoritmalarına göre daha iyi çalıştığı ortaya koyulmuştur.

Omidi ve arkadaşları [35], Nu-SVM metodu ile kısa dönem yük tahmini uygulaması gerçekleştirmiştir. Yapılan çalışmada, veri setine nu-SVM ve ANN algoritmaları uygulanmıştır. Sonuçlar mevsimsel olarak incelenmiş olup dört mevsimde de nu-SVR algoritması ANN algoritmasından daha iyi çalıştığı ortaya koyulmuştur.

Sarhani ve El Afia [36], özellik seçiminin (FS) ve parametre optimizasyonunun elektrik yük talep tahmininde SVR ile kullanılması üzerine yaptıkları çalışmada özellikle SVR-PSO hibrit algoritmasını kullanmıştır. Standart SVR sonucunda MAPE değeri %0,1366 bulunurken FS etkisi yapılan tahminde görülememiştir. Geliştirilen SVR-PSO algoritmasında ise MAPE değeri %0,1250 bulunmuştur.

Khan ve arkadaşları [37], ANN ve BRT kullanılarak gerçekleştirdikleri bir yük talebi analiz çalışmasında ANN için MAPE değerini %3,72 bulurken BRT algoritması için %3,33 olarak tespit etmiştir. Ayrıca günlük maksimum değer tahmininde, MAPE hata fonksiyonu değeri %2,19 ile BRT algoritmasında daha iyi sonuç vermiştir. Söz konusu çalışmada geçmiş yük taleplerine ilişkin veri seti ile birlikte hava durumu verisi kullanılmıştır.

Baghel ve arkadaşları [38], python kodu ile ve libSVM paketi kullanılarak geliştirdikleri bir algorithmada SVR yöntemini kullanarak kısa zamanlı yük talep analizi gerçekleştirilmiştir. Analiz kapsamında verinin %75'i eğitim verisi olarak ayrılmış, geriye kalan %25'lik kısmı ile regresyon fonksiyonu ile hata fonksiyonu oluşturulmuştur. Oluşturulan bu fonksiyon ile MAPE değeri düşürülerek daha kesin sonuçlar elde edilmiştir.

Mohan ve arkadaşları [39], kısa zamanlı yük talebi analizinde, haftalık ve aylık veriler kullanılarak farklı ülkelere ait çalışma yapmış ve özellikle günlük veriler ile çalışılırken DMD metodu kullanılmasının daha iyi tahmin yapılmasını sağlayacağı bulgusuna ulaşmıştır. Uygulamada birçok yöntem kıyaslanmış, öne sürülen metodun SVR ve hibrit modellerinden daha iyi çalıştığı gözlemlenmiştir.

Leith ve arkadaşları [40], İrlanda'nın yıllık yük verisini kullanılarak, Gaussian Process Prior yöntemi ile yük talebi tahmin çalışması yapmıştır. Modelin yaşayan bir yöntem olmasının alternatiflerine göre avantajı olarak belirtilmiştir. Özellikle hiper parametrelerin belirlenmesinde optimizasyon problemleri ve modelin çalışması sırasında her bir adımda matrisin tersinin alınması işleminin olması dezavantaj olarak belirtilmiştir.

Blum ve Riedmiller [41], çalışmasında özellikle haftalık ve günlük güneş enerjisinden elektrik üretimi, rüzgâr enerjisinden elektrik üretimi ve elektrik tüketim verilerini kullanarak GPR yöntemi ile modelleme yapmıştır. Analizde uygun

kovaryans fonksiyonu ile esnek tahmin modeli ortaya koyulmuş olup farklı tipteki müşteriler için yüksek kesinlikle tahmin yapıldığı belirtilmiştir.

Hachine ve arkadaşları [42], GPR metodu kullanılarak kısa dönemli tahmin yaptıkları bir uygulamada gerçek değerlere %95,5 yakınlıkla sonuç bulunmuştur. Bu uygulamada saatlik veriler ile tahmin modeli oluşturulmuş ve gaussian process prior modeli kullanılmıştır [42].

Bu tezin amacı ise, yaygın kullanılan algoritmaların çalışma performansını değerlendirmek, veri setine çalışma günleri ve ortalama tüketim parametrelerinin etkisini dikkate alarak, tutarlı ve pratik bir algoritma ile yük talebi analizi gerçekleştirmektedir.

Bu çalışma kapsamında veri madenciliği yöntemlerinden doğrusal regresyon, destek vektör makineleri regresyon (Support Vector Machine Regression), Yapay Sinir Ağları (Artificial Neural Network), Gaussian Process Regresyon ve Regresyon Karar Ağacı (Regression Decision Tree) modelleri kullanılarak, sonuçlar birbiri ile karşılaştırılmıştır ve performanslı çalışan modeller değerlendirilmiştir.

2. TALEP TAHMİNİNDE İSTATİSTİKSEL METODLAR

2.1 Regresyon Analizi

Regresyon analizi, aralarında sebep-sonuç ilişkisi bulunan iki veya daha fazla değişken arasındaki ilişkiyi, o konu ile ilgili tahminler (prediction) ya da kestirimler (estimation) yapabilmek amacı ile oluşturulan bir istatistiksel modeldir.

Regresyon analizi, bir kriter değişkeni ile bir veya daha fazla sayıda tahmin değişkenleri arasındaki ilgiyi sayısal hale dönüştürmede kullanılan istatistiksel analiz yöntemidir. Regresyon analizinde amaç, her tahmin değişkeninin kriter değişkenindeki toplam değişmeye olan katkısının saptanması ve dolayısıyla tahmin değişkenlerinin Y' doğrusal kombinasyon değerlerinden hareketle kriter değerlerinin tahmin edilmesidir [43].

Diğer bir deyiş ile bir örnek üzerindeki X_1, X_2, X_3, \dots serbest ve Y bağımlı değişken ölçülerine dayanarak Y ile X_1, X_2, X_3, \dots değişkenleri arasındaki fonksiyonel ilişkiyi kestirme işlemine regresyon analizi adı verilmektedir [44] ve denklem (2. 1) şeklinde gösterilmektedir.

$$Y = f(X_1, X_2, X_3, \dots) \quad (2. 1)$$

Bu modellemenin amacı koşullar değiştiğinde bağımlı değişken olan $E(Y_i)$ nasıl değiştiğini tanımlamaktır [45].

2.1.1 Doğrusal Regresyon Analizi

Basit regresyon analizi değişkenler arasında bulunduğu kabul edilen gerçek doğrusal ilişkiyi vermektedir [44].

Doğrusal regresyonda $E(Y_i)$ ile X_i arasındaki fonksiyonel ilişki (2.2) denklemini ile bulunmaktadır.

$$E(Y_i) = \beta_0 + \beta_1 X_i \quad (2.2)$$

Bu modelde $X = 0$ iken $E(Y_i) = \beta_0$ olmaktadır β_1 ise doğrunun eğimi olup X deki birim değişim ile $E(Y_i)$ arasındaki değişim oranını tanımlamaktadır [45].

Bir modelde Y_i kendi ana kütle ortalamasından $E(Y_i)$ sapması matematiksel modele bir hata terimi n eklenerek açıklanmaktadır. Bu noktada doğrusal regresyon denklemi (2.3) şeklinde olmaktadır [45].

$$E(Y_i) = \sum_{i=1}^n [\beta_0 + \beta_1 X_i] + n \quad (2.3)$$

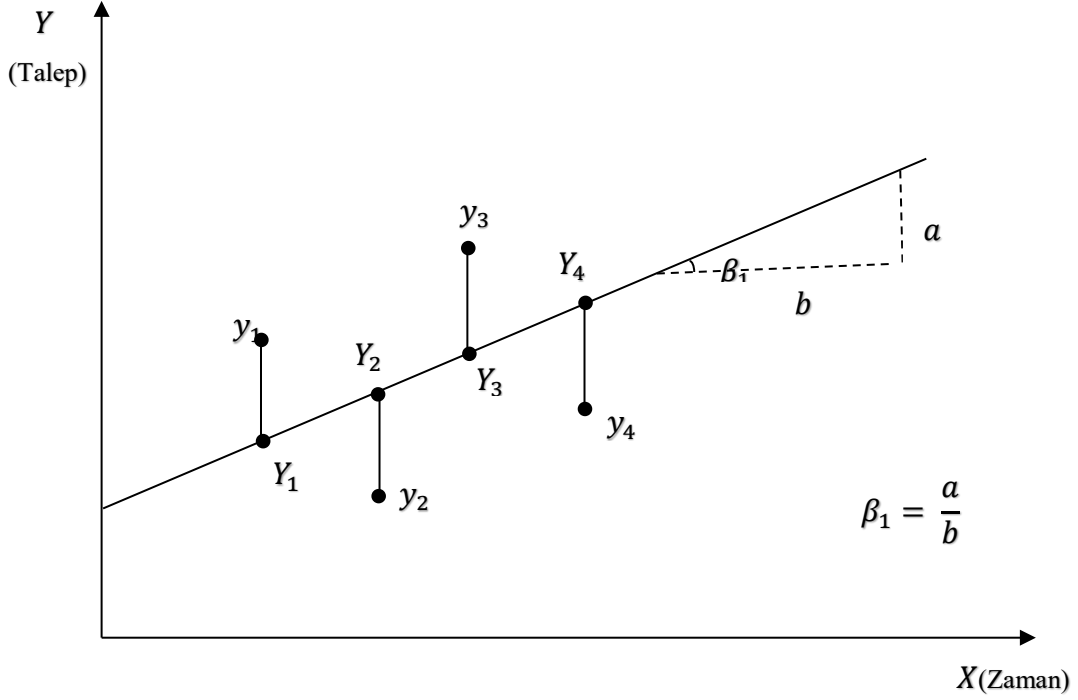
Bir doğrusal regresyon denkleminde yer alan β değerleri en küçük kareler yöntemi ile bulunmaktadır.

En küçük kareler yöntemine göre, bir zaman serisine en iyi uyan başka bir deyişle bir değerler serisini en iyi ifade eden doğru veya eğri, denklemin uygulanması ile bulunacak teorik değerler arasındaki farkların karelerinin toplamını (sapmaların kareleri toplamını) minimum yapan doğru veya eğridir [46].

Denklem (2.3) de verilen doğru denklemindeki hata n en küçük kareler yöntemi ile bulunmaktadır. ε değerini (2.4) eşitliği vermekte ve (2.5) şeklinde ifade edilmektedir.

$$n = \sum (Y' - Y)^2 \quad (2.4)$$

$$SS(n) = \sum_{i=1}^k [E(Y_i) - (\beta_0 + \beta_1) X_i]^2 \quad (2.5)$$



Şekil 2.1: Doğrusal regresyon doğrusu.

Şekil 2.1’de verilen regresyon doğrusuna ait grafikte talebin zamana bağlı tahmini yapılmıştır. Şekil 2.1’de görüldüğü gibi, bağımlı değişken değerleri ile gerçek değerler arasındaki farkın karelerinin toplamını minimum yapacak en uygun doğru belirlenmeye çalışılmaktadır.

2.1.2 Eğrisel Regresyon Analizi

Eğrisel regresyonda bağımlı değişken olan $E(Y_i)$ ile serbest değişken olan X_i arasındaki ilişki eğri biçiminde görülürse bu polinom fonksiyonu (2.6) eşitliği ile ifade edilmektedir.

$$E(Y_i) = \sum_{i=1}^k [\beta_0 + \beta_i X_i^i] + n \quad (2.6)$$

2.2 Yapay Sinir Ağları

Yapay sinir ağları, insan beyninin özelliklerinden olan öğrenme yolu ile yeni bilgiler türetebilme ve keşfedebilme gibi yetenekleri herhangi bir yardım almadan kendiliğinden gerçekleştirmek amacı ile biyolojik sinir ağlarından esinlenilerek geliştirilen bilgisayar sistemleridir [47].

Yapay sinir ağları ile ilgili çalışmalar, 1938 yılında Rashevsky tarafından başlatılmış olup güçlü bilgisayarların olmaması nedeniyle 1960 yılında kesintiye uğramıştır. 1980'li yıllarda bilgisayar teknolojisinin ilerlemesi ile yapay sinir ağları konusunda yeniden çalışmalar başlamış olup yeni algoritmalar ortaya koyulmuştur [48].

Yapay sinir ağları bütün modelleri için geçerli olan genel karakteristik özelliklere sahiptir. Bu özellikler yapay sinir ağlarının tercih edilmesinde en büyük etmenlerdir. Doğrusal olmayan yapıları modelleyebilmesi, farklı problemler için uygulanabilirliği, öğrenme ve genelleme yapma yeteneği, hata toleransına sahip olması en önemli özellikleri arasındadır. Ayrıca paralel dağılmış yapısı sayesinde hızlı ve etkin bir şekilde çalışabilmesi tercih sebebidir [49].

Yapay sinir ağları genel olarak üç ana katmandan oluşmakta olup bu katmanlar;

- Giriş katmanı
- Gizli katmanlar
- Çıkış katmanı

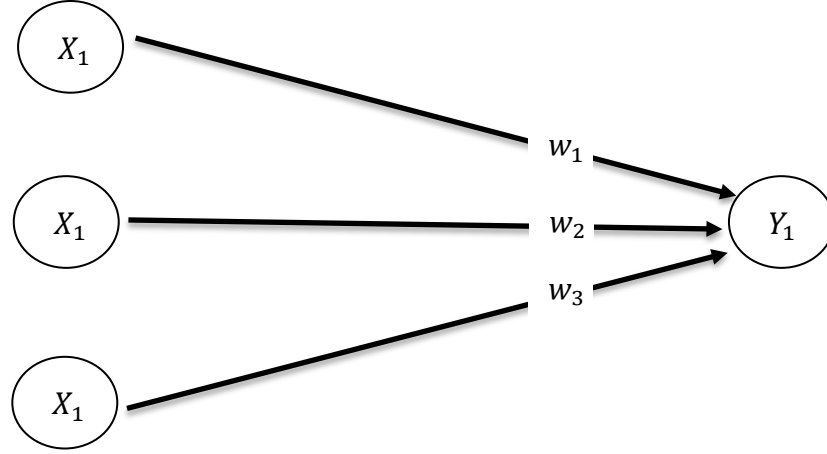
olarak sınıflandırılabilir.

Yapay sinir ağları, görüntü tanıma, veri madenciliği, sınıflandırma, tahmin ve proses modelleme gibi birçok alanda kullanılmaktadırlar [34].

Basit bir yapay sinir ağının (yapay nöronun) fonksiyonu (2.7) denkleminde görüldüğü gibi ifade edilmektedir.

$$y_{in} = w_1X_1 + w_2X_2 + w_3X_3 \quad (2.7)$$

Bu fonksiyonun diyagramı Şekil 2.2'de gösterilmektedir [50].

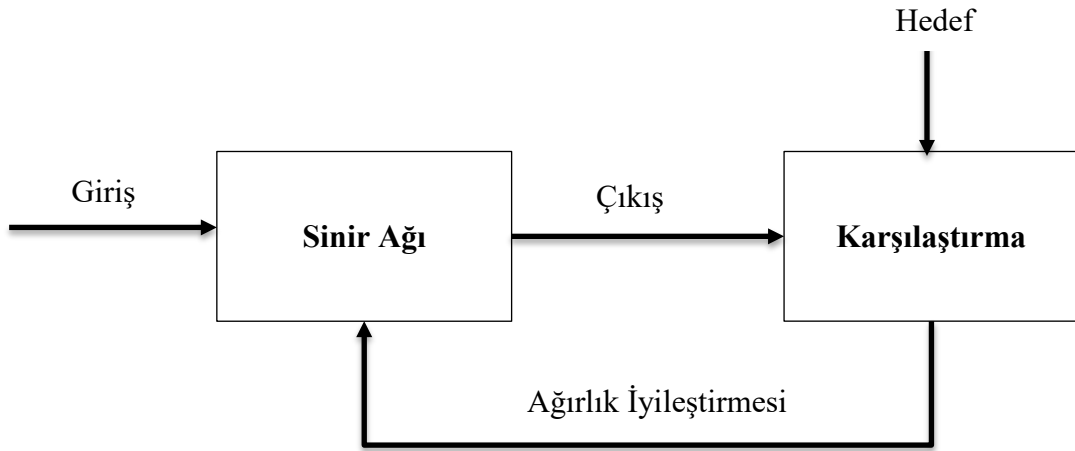


Şekil 2.2: Tek katmanlı yapay sinir ağı.

NET girişi denklem (2.8)'deki gibi olurken lojistik sigmoid fonksiyonu (2.9) eşitliğindeki gibi olmaktadır [50].

$$y = f(y_{in}) \quad (2.8)$$

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (2.9)$$



Şekil 2.3: Bir sinir ağının akış diyagramı.

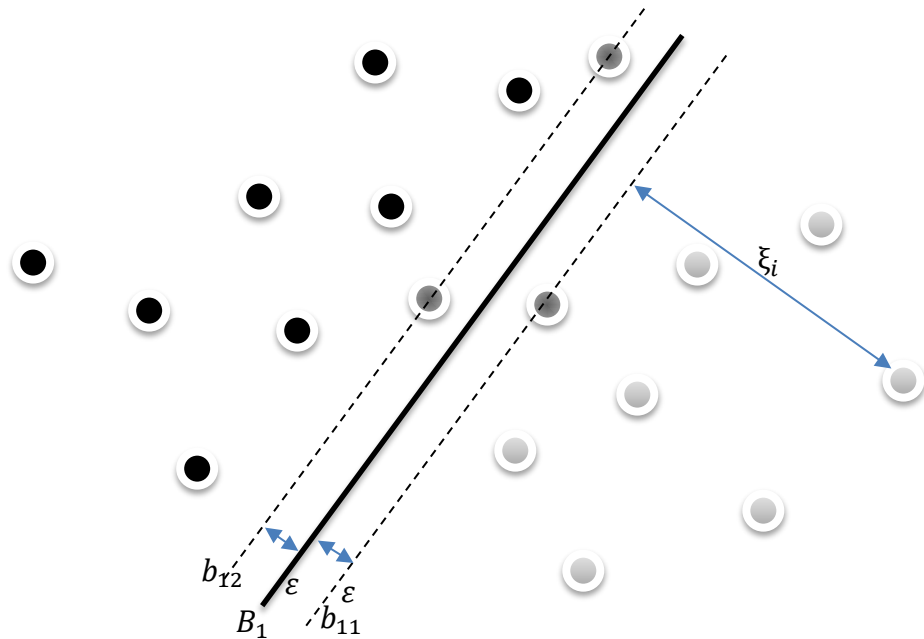
2.3 Destek Vektör Makineleri

Destek vektör makineleri, orijinal adı ile support vector machine (SVM) 1995 yılında Vapnik [51] tarafından ortaya koyulmuştur. SVM özellikle makine öğrenmesi alanında önemli yapay zekâ uygulama metodudur. Özellikle regresyon kestirimi ve sinyal işlemede kullanıma uygun bir yöntemdir [35].

Destek vektör makinelerinin en önemli avantajı, sınıflandırma problemini kareli optimizasyon problemine dönüştürerek çözüm getirebilmesidir. Böylece problemin çözümü sırasında gerçekleşen öğrenme işlem adımları azalmakta ve diğer algoritmalara göre daha hızlı çözümlenebilmektedir [52].

Destek vektör makineleri genellikle doğrusal olmayan regresyon problemlerinin çözümünde başarılı sonuç veren bir metot olarak kullanılmaktadır.

Şekil 2.4’de destek vektör makineleri için iki sınıflı problem örneği verilmiştir. Söz konusu örnekte destek vektörleri doğrusaldır. Destek vektörleri, ayırma hiper düzlemine en yakın olan her iki sınıfa da ait örnekler olarak ifade edilmektedir. Şekil 2.4’de görülen b_{11} ve b_{12} doğruları sınır düzlemleri oluşturmakta ve sınır düzlemlerin arasında bulunan ve her iki sınır düzleme eşit uzaklıkta bulunan B_1 düzlemi hiper düzlem olarak ifade edilmektedir [52].



Şekil 2.4: Doğrusal olarak ayrılan destek vektörü.

İki sınır düzlem (b_{11} ve b_{12}) arasında kalan bölgeye marjin adı verilmektedir. Marjin; karar sınırının bir veri noktası ile kesişmeden artırılacağı genişlik olarak da tanımlanabilmektedir. Şekil 2.4'deki gibi doğrusal olarak ayrılan bir problemde doğru denklemleri, denklem (2.10), (2.11) ve (2.12)'deki gibi olur iken eğitim örneklerini en iyi şekilde ayıran hiper düzlem en küçük; denklem (2.13)'deki gibi [53], en büyük; denklem (2.14)'deki gibi olmaktadır.

$$B_1 = \bar{\omega} \cdot \bar{x} + b = 0 \quad (2.10)$$

$$b_{11} = \bar{\omega} \cdot \bar{x} + b = +1 \quad (2.11)$$

$$b_{12} = \bar{\omega} \cdot \bar{x} + b = -1 \quad (2.12)$$

$$L(\omega) = \frac{\|\bar{\omega}\|^2}{2}$$

$$f(\bar{X}_i) = \begin{cases} 1, & \text{eğer } \bar{\omega} \cdot \bar{x} + b \geq 1 \\ -1, & \text{eğer } \bar{\omega} \cdot \bar{x} + b \leq -1 \end{cases} \quad (2.13)$$

$$Margin = \frac{2}{\|\bar{\omega}\|^2}$$

$$f(\bar{X}_i) = \begin{cases} 1, & \text{eğer } \bar{\omega} \cdot \bar{x} + b \geq 1 \\ -1, & \text{eğer } \bar{\omega} \cdot \bar{x} + b \leq -1 \end{cases} \quad (2.14)$$

Şekil 2.5'de doğrusal olarak ayrılmayan destek vektör makinesi için örnek verilmiştir. Bu durumda Cortes ve Vapnik'in teorisine göre bir C serbest değişkeni kullanılmaktadır. Denklem ve sınır değerler (2.15) şeklinde olmaktadır [54].

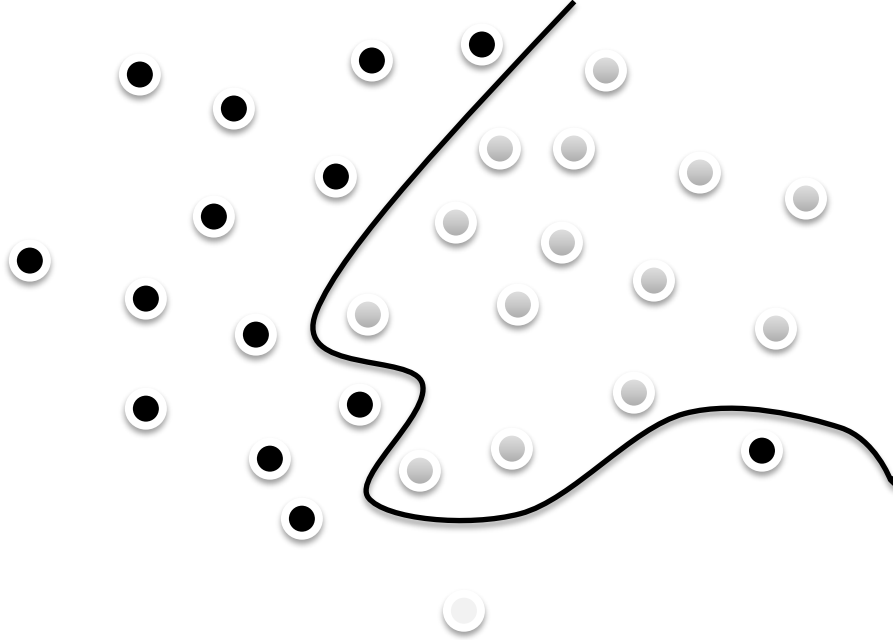
$$min_{\omega, b, \xi} = \frac{1}{2} \omega^T \omega + C \left(\sum_{i=1}^l \xi_i^k \right)$$

$$y_i(\omega^T \Phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, i = 1, \dots, l, \quad (2.15)$$

Optimal ω , (2.16) eşitliği kullanılarak, karar fonksiyonu ise (2.17) denklemi ile ifade edilmektedir [54].

$$\omega = \sum_{i=1}^l y_i \alpha_i \Phi(x_i) \quad (2.16)$$

$$\text{sgn}(\omega^T \Phi(x) + b) = \text{sgn}\left(\sum_{i=1}^l y_i \alpha_i K(x_i, x) + b\right) \quad (2.17)$$



Şekil 2.5: Doğrusal olarak ayrılmayan destek vektörü.

Denklem (2.15)'de belirtilen C sabiti aşırı uyumu kontrol etmeyi sağlamaktadır. Bulunan C sabiti küçük ise marjin geniş, büyük ise marjin dardır. $C = \infty$ ise marjin tüm kısıtları zorlamaktadır.

C-Destek Vektör Sınıflandırmasına alternatif olarak 2000 yılında Schölkopf tarafından $\nu \in (0,1]$ olan yeni parametre geliştirilmiştir [54].

$$\begin{aligned}
\min_{\omega, b, \xi} &= \frac{1}{2} \omega^T \omega + v\rho + \frac{1}{l} \left(\sum_{i=1}^l \xi_i \right) \\
& y_i(\omega^T \Phi(x_i) + b) \geq \rho - \xi_i, \\
& \xi_i \geq 0, i = 1, \dots, l, \quad \rho \geq 0
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
\min_{\alpha} &= \frac{1}{2} \alpha^T Q \alpha \\
& 0 \leq \alpha_i \leq 1/l, \quad i = 1, \dots, l, \\
& e^T \alpha \geq v, \quad y^T \alpha = 0,
\end{aligned} \tag{2.19}$$

$$\begin{aligned}
Q_{ij} &= y_i y_j K(x_i, x_j) \\
v &\leq \frac{2 \min(\#y_i = +1, \#y_i = -1)}{l} \leq 1,
\end{aligned} \tag{2.20}$$

Denklem (2.20)'de verilen bağıntıya göre v , $(0,1]$ aralığından küçüktür [54]. Karar fonksiyonu yine denklem (2.17)'deki gibi olmaktadır.

$$\begin{aligned}
\min_{\bar{\alpha}} &= \frac{1}{2} \alpha^{-T} Q \bar{\alpha} \\
& 0 \leq \bar{\alpha}_i \leq 1/l, \quad i = 1, \dots, l, \\
& e^T \bar{\alpha} \geq vl, \quad y^T \bar{\alpha} = 0,
\end{aligned} \tag{2.21}$$

2.3.1 Destek Vektör Regresyon

SVR algoritmasındaki en temel amaç bilinen girdiler ile bilinmeyen çıktıların tahmin edilmesidir [32]. SVM 1998 yılında Vapnik tarafından ortaya koyulmuş olup, binary sınıflandırma problemlerini, konveks optimizasyon problemlerine dönüştürerek çözümleneyen yöntemdir. Hiper düzlem üzerinde mümkün olan çok sayıda noktayı doğru bir şekilde sınıflandırarak çalışmaktadır [55].

SVR problem formülasyonunda bir boyutlu örneği Şekil 2.6'da verilmiştir. Sürekli fonksiyonu denklem (2.22)'deki gibi yazılabilmektedir. Çok boyutlu veri

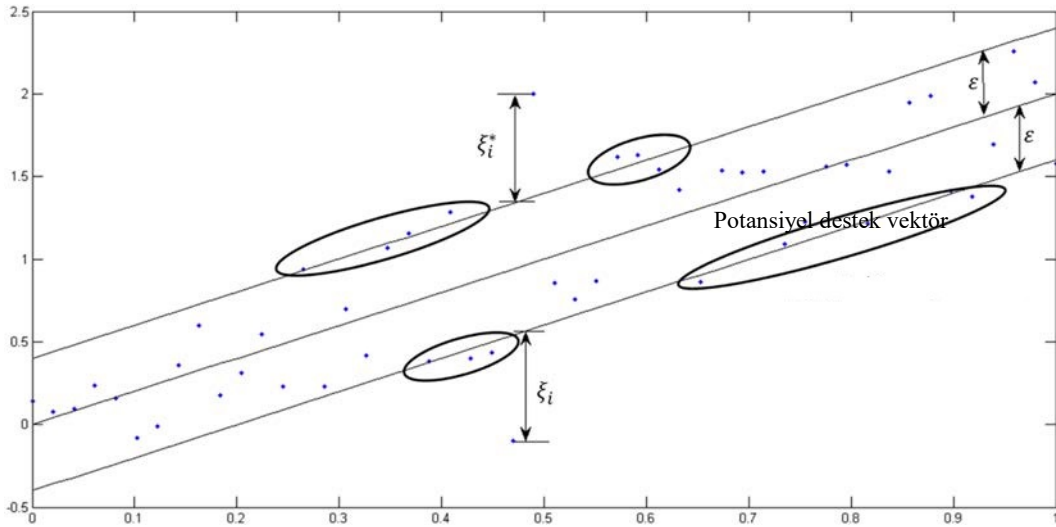
matematiksel notasyonda çok deęişkenli regresyon modeli denklem (2.23) şeklinde ifade edilmektedir [55].

$$y = f(x) = \langle \omega, x \rangle + b = \sum_{j=1}^M \omega_j x_j + b$$

$$y, b \in R, x, \omega \in R^M \quad (2.22)$$

$$f(x) = \begin{bmatrix} \omega \\ b \end{bmatrix}^T \begin{bmatrix} x \\ 1 \end{bmatrix} = \omega^T x + b$$

$$x, \omega \in R^{M+1} \quad (2.23)$$



Şekil 2.6: Destek vektör regresyon örneęi.

Ağırlıkların büyüklüğünün yassılığı göz önünde bulundurulduğunda fonksiyon (2.24) denkleminde gösterilmektedir [55].

$$f(x, \omega) = \sum_{i=1}^M \omega_i x^i$$

$$x \in R, \omega \in R^M \quad (2.24)$$

2.4 Gauss Süreç Regresyonu

Bir Gaussian süreç modeli, bir Gaussian dağıtımında bağlantı sınırlı numaralardan rastgele değişkenlerin toplanmasıdır. GPR doğrusal olmayan regresyon problemlerinin çözümünde kullanıma uygun, parametrik olmayan bir modeldir [56].

GP, Gauss dağılımında prior fonksiyonlarının, posterior fonksiyonlarına dönüştürülmesi temeline dayanmaktadır [57]. Tahmin, test vektörlerinin sürekli sayıdaki eğitim etiketlerini kullanarak gerçekleştirmek ve bu noktada tahmin değerleri sistematik ya da rassal varyasyonlar ile olabilmektedir [56].

GP fonksiyonlar üzerine olasılık dağılımını tanımlar ve denklem (2.25)'da verildiği gibi ifade edilebilmektedir [58].

$$f(x) \sim \mathcal{GP}(m(x), K(x, x')) \quad (2.25)$$

$m(x)$ ortalama, $K(x, x')$ ise kovaryans fonksiyonu olup denklem (2.26) ve (2.27)'deki gibi ifade edilmektedir.

$$m(x) = \mathbb{E}[f(x)] \quad (2.26)$$

$$K(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))^T] \quad (2.27)$$

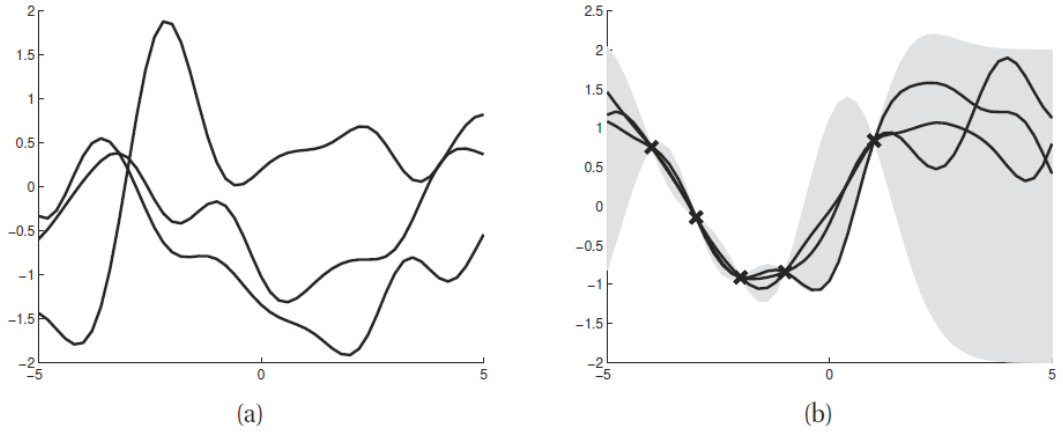
Kovaryans fonksiyonunda genellikle tercih edilen SE, denklem (2.28)'da verilmiştir.

$$K_{SE}(x, x') = \theta_f^2 \exp\left(-\frac{1}{\theta_l^2} \|x - x'\|^2\right) \quad (2.28)$$

Kovaryans fonksiyonunda belirtilen θ_f ve θ_l sırası ile x-ölçekleme (Genlik) ve y-ölçekleme (uzunluk ölçeği) [59] olarak tanımlanmaktadır.

Kovaryans ve ortalama, deęerlendirilecek girdi bileşenlerin X bütün sınırlı kombinasyonlarını deęerlendirilmesini saęlamaktadır [60]. Kovaryans matrisi denklem (2.29)'da belirtildięi gibi olmaktadır.

$$\begin{aligned}
 \mathbf{K} &:= \mathcal{K}((x_1, \dots, x_n), (x_1, \dots, x_n)) \\
 &= \begin{bmatrix} \mathcal{K}(x_1, x_1) & \mathcal{K}(x_1, x_2) & \dots & \mathcal{K}(x_1, x_n) \\ \mathcal{K}(x_2, x_1) & \mathcal{K}(x_2, x_2) & \dots & \mathcal{K}(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{K}(x_n, x_1) & \mathcal{K}(x_n, x_2) & \dots & \mathcal{K}(x_n, x_n) \end{bmatrix} \quad (2.29)
 \end{aligned}$$



Şekil 2.7: GP prior SE kernel (a) ve GP posterior (b) karşılaştırması.

Şekil 2.7'de GP prior SE kernel ile kullanımı (a) ve 5 gürültüsüz gözlem sonrası GP posterior (b) sonuçları verilmiştir. (b) bölümündeki taranmış alan denklem (2.30) ile ifade edilebilmektedir [57].

$$\mathbb{E}[f(x)] \pm 2std(f(x)) \quad (2.30)$$

Kernel parametreleri (hyper-parameters) denklem (2.31) ve (2.32)'de verilen eşitlik ile tahmin edilebilmektedir [57].

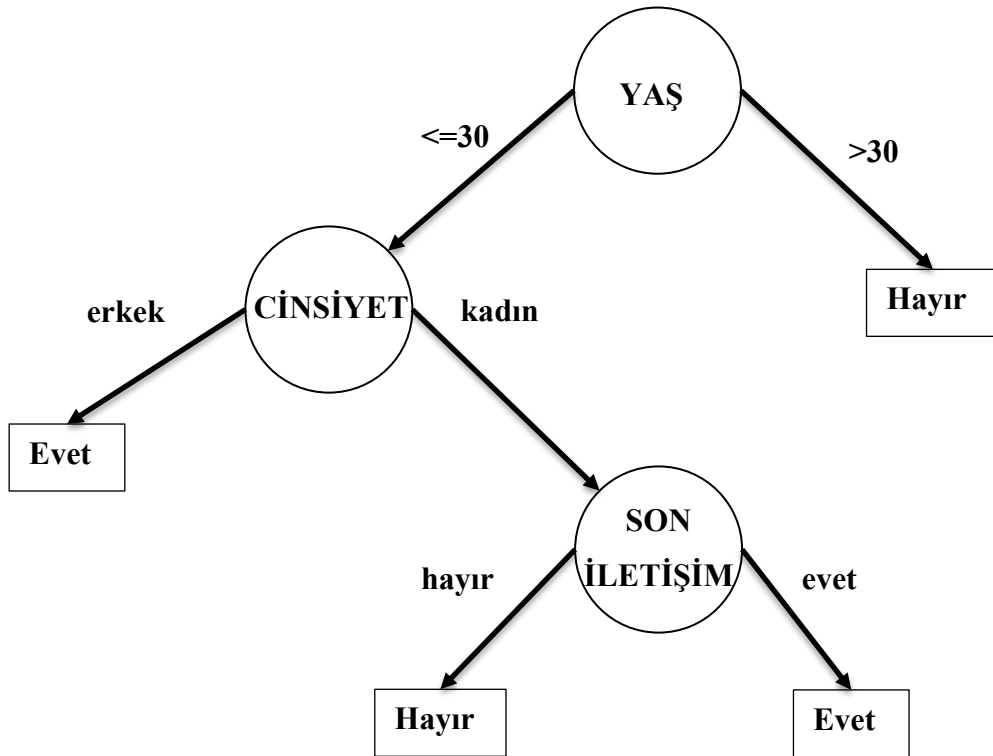
$$\frac{\partial}{\partial \theta_j} \log p(y|X) = \frac{1}{2} y^T K_y^{-1} \frac{\partial K_y}{\partial \theta_j} K_y^{-1} y - \frac{1}{2} \text{tr}(K_y^{-1} \frac{\partial K_y}{\partial \theta_j}) \quad (2.31)$$

$$\frac{\partial}{\partial \theta_j} \log p(y|X) = \frac{1}{2} \text{tr} \left((\alpha \alpha^T - K_y^{-1}) \frac{\partial K_y}{\partial \theta_j} \right) \quad (2.32)$$

2.5 Karar Ağacı

Karar ağaçları kategorik ya da sürekli (nümerik) verilerin sınıflandırılmasında kullanılan bir veri madenciliği sınıflandırma yöntemidir. Veri madenciliği modelinde ortaya koyulan model ağaç diyagramı yapısında olup, bu diyagramda veri sınıflandırma için kuralları görsel olarak temsil edilmektedir [61].

Karar ağacının hem nominal hem de nümerik veri ile çalışabilmesi en büyük avantajı arasında yer almaktadır. Örnek olarak potansiyel müşterilerin analizinin yapılmasında kullanılan bir diyagram Şekil 2.8’de verilmiştir. Bu yapıda “Eğer müşteri yaşı 30’dan küçük ya da eşit ise ve müşterinin cinsiyeti erkek ise müşteri postaya geri dönüş yapacaktır.” şeklinde analiz yapılabilmektedir [62].



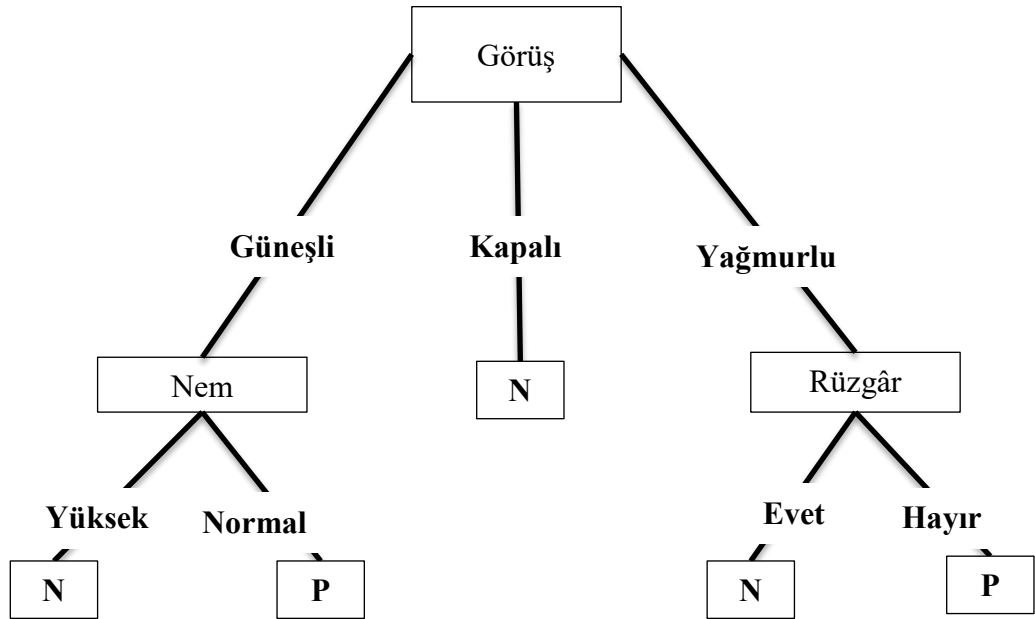
Şekil 2.8: Basit karar ağacı I.

Küçük bir eğitim seti ile hava tahmini örneğinde örnek veri seti Tablo 2.1'deki gibidir [63].

Tablo 2.1: Karar ağacı hava tahmini.

Değişkenler					
No.	Görüş	Sıcaklık	Nem	Rüzgâr	Sınıf
1	Güneşli	Sıcak	Yüksek	Hayır	N
2	Güneşli	Sıcak	Yüksek	Evet	N
3	Bulutlu	Sıcak	Yüksek	Hayır	P
4	Yağmurlu	Ilık	Yüksek	Hayır	P
5	Yağmurlu	Serin	Normal	Hayır	P
6	Yağmurlu	Serin	Normal	Evet	N
7	Bulutlu	Serin	Normal	Evet	P
8	Güneşli	Ilık	Yüksek	Hayır	N
9	Güneşli	Serin	Normal	Hayır	P
10	Yağmurlu	Ilık	Normal	Hayır	P
11	Güneşli	Ilık	Normal	Evet	P
12	Bulutlu	Ilık	Yüksek	Evet	P
13	Bulutlu	Sıcak	Normal	Hayır	P
14	Yağmurlu+	Ilık	Yüksek	Evet	N

Bu veri setine göre karar ağacı Şekil 2.9'daki gibi olmaktadır.



Şekil 2.9: Basit karar ağacı II.

Şekil 2.9’da verilen karar ağacı uygulamasında karar ağacı eğitim veri setindeki veriler doğrultusunda her objenin doğru sınıf ile eşleştiği görülmektedir. Bu yapıda karar ağacı anlamlı sınıflandırma ilişkisini kurmalıdır [63].

Karar ağacı yönteminde en sık kullanılan algoritmalar şunlardır:

- Hunt algoritması
- CART
- ID3, C4.5, C5.0
- SLIQ, SPRINT

Bu algoritmalar ile veri setinin daha seçici olarak sınıflandırılmasını sağlayan analizler yapılması mümkün olmaktadır.

Karar ağaçlarında düğümlerin safsızlığı GINI indeksi, Entropi hesabı ya da sınıflandırma hatası kullanılarak belirlenebilmektedir.

Bir karar ağacı algoritmasındaki t düğümü için GINI indeksi (2.33) eşitliğinde görüldüğü gibi tanımlanmaktadır.

$$GINI(t) = 1 - \sum_j [p(j|t)]^2 \quad (2.33)$$

$p(j|t)$: t düğümündeki j sınıfının göreceli frekansı

CART, SLIQ ve SPRINT algoritmalarında eğer bir p düğümü k bölmeye (child) ayrılıyor ise bölmenin safsızlığı denklem (2.34)’de olduğu gibi ifade edilmektedir.

$$GINI_{split} = \sum_{i=1}^k \frac{n_i}{n} GINI(i) \quad (2.34)$$

n_i = i. çocuktaki nesne sayısı

n = p düğümündeki nesne sayısı

Bir karar ağacı algoritmasında rastgeleliği, belirsizliği ve beklenmeyen durumun ortaya çıkma olasılığının hesabı entropi ile gerçekleştirilmektedir. T düğümündeki entropi (2.35) eşitliğindeki gibi gösterilmektedir.

$$Entropy(t) = - \sum_j p(j|t) \log(p(j|t)) \quad (2.35)$$

$p(j|t)$: t düğümündeki j sınıfının göreceli frekansı

Özellikle ID3 ve C4.5 algoritmalarında kullanılan bilgi kazancı ise denklem (2.36)'da gösterilen eşitlik ile bulunmaktadır.

$$GAIN_{split} = Entropy(t) - \left(\sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right) \quad (2.36)$$

n_i = i. bölmesindeki nesne sayısı

Karar ağacı algoritmasının avantajları ise;

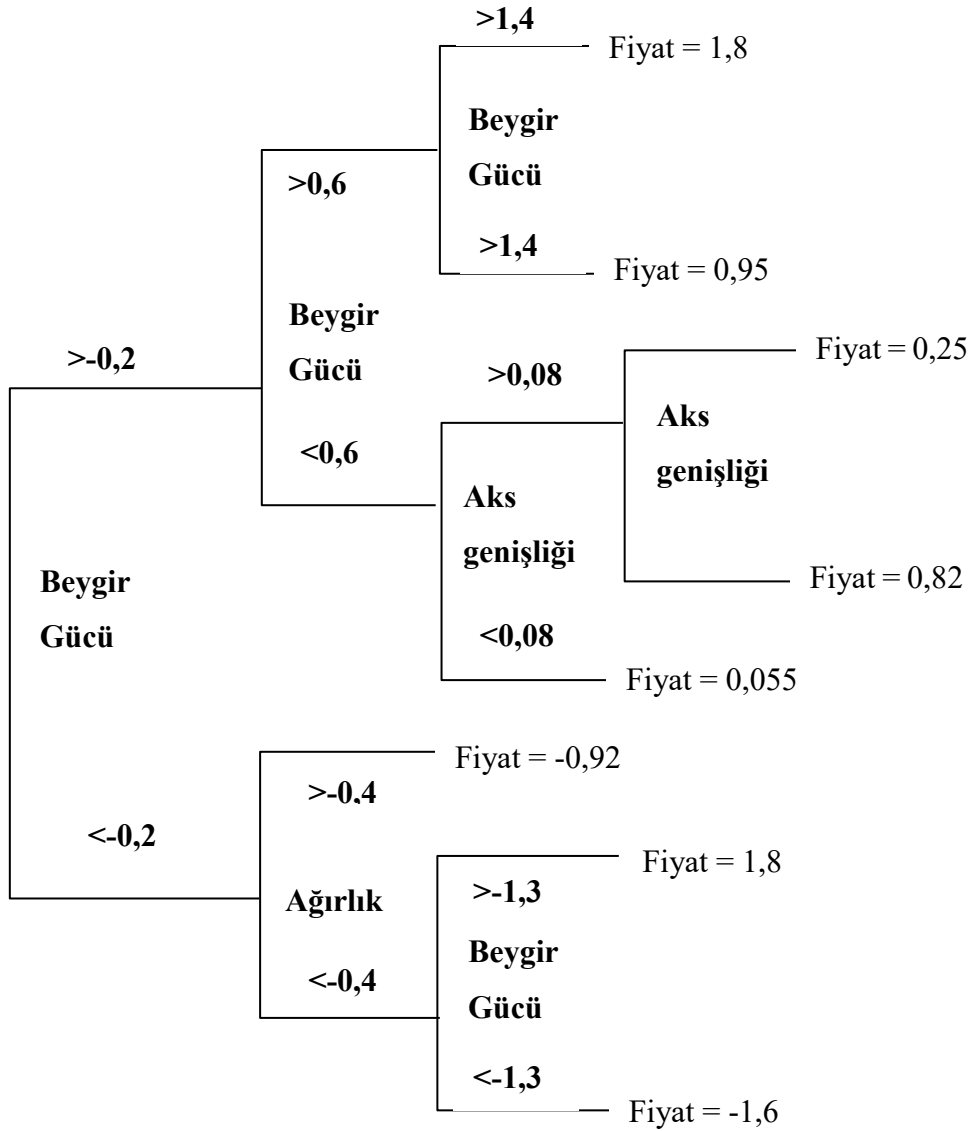
- Anlaşılması ve yorumlanması kolaydır.
- Karar ağacını görselleştirme imkânı tanır.
- Veri ön işlemeye ihtiyaç azdır.
- Ağacı oluşturma ve kullanma maliyeti düşüktür: Tahmin maliyeti öğrenme kümesi büyüklüğünün logaritmasıyla orantılıdır [64].
- Sayısal ve kategorik verinin bir arada işlenebilmesine olanak tanır.
- İstatistiksel testler ile uygulanabilmektedir.
- Gömülü öznitelik seçimi imkânı vardır.
- Gürültü ve ilgisiz özniteliklerden etkilenmemektedir.
- Karar ağacı algoritmasının dezavantajları ise;
- Aşırı öğrenmeye bağlı olarak karmaşık karar ağaçlarının oluşması durumu ile karşılaşılabilir. Sonuçlar her zaman anlamlı olmayabilir [64].

- Verideki küçük bir deęişiklik için dahi yeni ve farklı bir karar ağacı oluşur [64].
- Baskın sınıfa baęlı olarak yanlış karar ağacı oluşturulabilmesidir.

2.5.1 Regresyon Ağacı

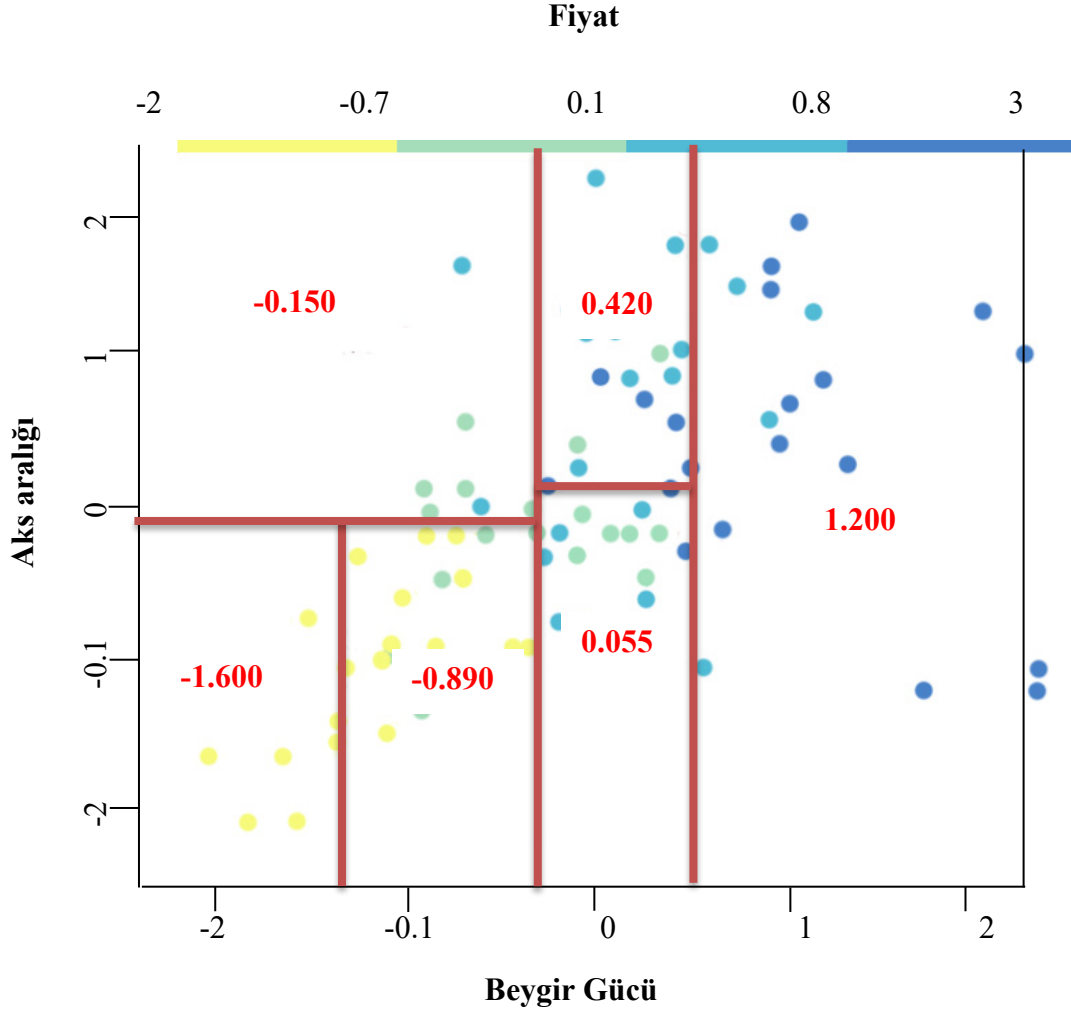
CART, Breiman Friedman Olsen ve Stone tarafından geliştirilmiş olup özellikle sınıflandırma ve regresyon ağaçlarının oluşturulmasında kullanılmaktadır [65].

Şekil 2.10'de güç, dingil aralığı ve aęırlığa baęlı olarak araç satışının analizinin yapıldığı regresyon ağaç diyagramı verilmiştir [66].



Şekil 2.10: Örnek regresyon karar ağacı.

Şekil 2.10’da 2003 model arabalar için fiyat tahmini çalışmasının karar ağacı diyagramı verilmiştir. Şekil 2.11’de söz konusu regresyon ağacını oluşturan verilerin kümeleme yöntemi ile ayrışımı verilmiştir.



Şekil 2.11: Örnek kümeleme yöntemi ile ayrıştırma.

Beygir gücü 0.6’dan büyük olduğunda aks aralığı önemini yitirmektedir. Söz konusu ağaçtaki hata denklem (2.37) şeklinde, $m_c = \frac{1}{n_c} \sum_{i \in C} y_i$ için kümeleme yapıldığında denklem (2.38) şeklinde yazılabilmektedir.

$$S = \sum_{c \in \text{leaves}(T)} \sum_{i \in C} (y_i - m_c)^2 \quad (2.37)$$

$$S = \sum_{c \in \text{leaves}(T)} n_c V_c \quad (2.38)$$

Regresyon karar ağaçlarında aşırı öğrenme ve model performansını artırmaya yönelik geliştirilmiş olan Bagging ve boosting adı ile bilinen iki yaklaşım kullanılmaktadır [67]. Bagging yaklaşımı veri seti içerisinde yerinde bırakılarak yerine koyulacak şekilde yeniden örneklem alınarak yeni ağaçlar oluşturulması esasına dayanmaktadır. Boosting yaklaşımında ise artık optimizasyonuna dayalı olarak performans artışı sağlanmaktadır.

Temel regresyon ağacı algoritması yazılırken aşağıdaki işlem adımlarına dikkat edilmektedir;

- Bütün noktaları içeren tekil düğüm ile başla;
- Eğer düğümdeki bütün noktalar aynı bağımsız değişkenler için aynı değerlere sahip ise dur. Aksi durumda bütün binary kesitler hatayı olabildiğince azaltmaktadır.
- Her bir düğüm için ilk aşamaya geri dönülür [66].

Veri madenciliği dahilinde öncelikli veri toplama, veriyi test ve eğitim veri seti olmak üzere bölme, veriyi doğrulama ve ileriye dönük tahmin edilmesi şeklinde tamamlanmaktadır [65].

2.6 Hata Fonksiyonları

Tahmin modellerinde en yaygın olarak kullanılan hata fonksiyonları ME, MPE, MAE, RMAE, MAPE, sMAPE fonksiyonlarıdır [68].

2.6.1 Ortalama Hata

Ortalama hata fonksiyonu bir makine öğrenmesi modelinin öngördüğü tahmin değerleri ile gerçek değerlerin arasındaki ortalama hatadır. Bir ölçümde tahmin değeri ile gerçek değer arasındaki fark olarak da tanımlanabilmektedir. Fonksiyon, denklem (2.39)'de verildiği gibidir.

$$ME = \frac{1}{n} \sum_{j=1}^n e_j \quad (2.39)$$

2.6.2 Ortalama Yüzde Hata

Ortalama yüzde hata, modelin tahminlediği değerler ile gerçek değerler arasındaki farkın ortalama yüzdesidir. MPE değerleri hesaplanırken mutlak değer kullanılmadığı için pozitif ve negatif değerler birbirini dengeleyebilmektedir. Fonksiyon, denklem (2.40)'da verildiği gibidir.

$$MPE = \frac{100}{n} \sum_j \frac{e_j}{A_j} \quad (2.40)$$

2.6.3 Ortalama Mutlak Hata

Ortalama mutlak hata iki sürekli değişken arasındaki farkın ölçüsüdür. MAE, her gerçek değer ile veriye en iyi uyan çizgi arasındaki ortalama dikey veya yatay mesafe olarak da tanımlanabilmektedir. Fonksiyon, denklem (2.41)'de verildiği gibidir.

$$MAE = \frac{1}{n} \sum_{j=1}^n |e_j| \quad (2.41)$$

2.6.4 Ortalama Hata Karesi

Ortalama hata karesi, bir regresyon eğrisinin bir dizi noktaya ne kadar yakın olduğunun belirlenmesinde kullanılmaktadır. Bu yöntem ile modelin tahmin etme performansı ölçülebilmektedir. Fonksiyon, denklem (2.42)'de verildiği gibidir.

$$MSE = \frac{1}{n} \sum_{j=1}^n e_j^2 \quad (2.42)$$

2.6.5 Kare Ortalamalarının Karekökü

Kare ortalamalarının karekökü fonksiyonu tahmin hatalarının standart sapması olarak tanımlanabilmektedir. Düşük değere sahip tahminleyiciler daha iyi performans göstermektedir. MSE'nin kare kökü olarak da tanımlanabilmektedir. Fonksiyon, denklem (2.43)'de verildiği gibidir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n e_j^2} \quad (2.43)$$

2.6.6 Ortalama Mutlak Yüzde Hata

Ortalama mutlak yüzde hata, regresyon ve zaman seri modellerinde tahminlerin doğruluğunu ölçmek için kullanılmaktadır. Fonksiyon, denklem (2.44)'de verildiği gibidir.

$$MAPE = \frac{100}{n} \sum_j \frac{|e_j|}{|A_j|} \quad (2.44)$$

2.6.7 Simetrik Ortalama Mutlak Yüzde Hata

Simetrik ortalama mutlak yüzde hata, sıfır ya da sıfıra yakın hata oranlarında MAPE'ye alternatifi olarak kullanılmaktadır. Fonksiyon, denklem (2.45)'de verildiği gibidir.

$$sMAPE = \frac{100}{n} \sum_j \frac{2 * |e_j|}{|A_j| + |P_j|} \quad (2.45)$$

3. BURSA İLİNE AİT ELEKTRİK YÜK TALEP TAHMİNİNİN YAPILMASI

TEİAŞ 2. Bölge Müdürlüğü yapılanması Bursa lokasyonunda konumlanmış olup Bursa, Balıkesir, Yalova ve Çanakkale illerini kapsamaktadır.

EPDK'dan alınan veriler doğrultusunda 2017 yılında Türkiye'nin toplam elektrik enerjisi tüketimi 225713,5 GWh olup bu tüketimin %5,21 kadarı olan 11764,5 GWh'i Bursa ilinde 18 trafo merkezi üzerinden gerçekleştirilmiştir [18]. Bursa ilinde gerçekleştirilen bu tüketiminin 7224,7 GWh'i dağıtım gerilimi seviyesinden olurken, 4539,8 GWh'i iletim gerilimi seviyesindedir.

2017 yılında Bursa ilinde faturalanan tüketiminin tüketici türüne göre dağılımı Tablo 3.1'de verilmiştir [18].

Tablo 3.1: 2017 Tüketici türüne göre dağılım (MWh-%).

Aydınlatma	Mesken	Sanayi	Tarımsal Sulama	Ticarethane	Genel Toplam
130061,90	1931360,11	7320343,93	84245,68	2298507,23	11764518,85
1,11	16,42	62,22	0,71	19,54	100

2017 yılı sonu itibari ile Bursa ilinde toplamda yer altı ve havai toplam hat uzunluğu 18,556 km olup toplam kapasitesi 4438 MVA olan 10689 adet trafo bulunmaktadır [18].

3.1 UEDAŞ Tarafından Yapılan Yük Talep Tahmini Sonuçları

TEİAŞ tarafından 2018 yılında yayınlanan 10 yıllık yük talep tahmini raporunda UEDAŞ görev bölgesine ait 2007-2016 dönemindeki tüketim ve artış Tablo 3.2'de verilmiştir [69].

Tablo 3.2: UEDAŞ 2007-2016 tüketim miktarı.

YIL	(GWh)	ARTIŞ (%)
2007	8689	-
2008	9331	7,4
2009	8961	-4,0
2010	9651	7,7
2011	10330	7,0
2012	10397	0,6
2013	10381	-0,1
2014	10648	2,6
2015	11170	4,9
2016	11708	4,8

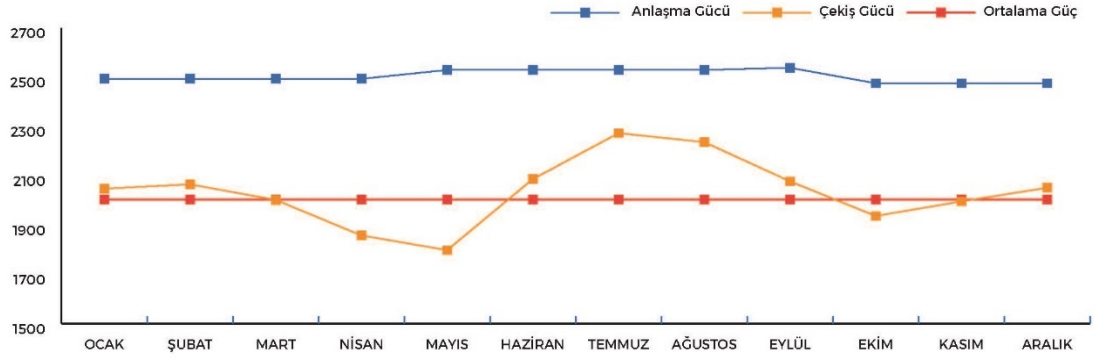
Tablo 3.2’de görüldüğü gibi 2007-2016 döneminde, Uludağ EDAŞ bölgesinin son 10 yıl elektrik enerjisi tüketimi artış ortalaması yıllık %3,4 olarak gerçekleşmiştir. 2018-2027 döneminde düşük, baz (referans) ve yüksek brüt elektrik talep tahminleri Tablo 3.3’de verilmiştir [69].

Tablo 3.3: Brüt elektrik tüketimi tahmini (GWh).

Yıl	Düşük	Artış (%)	Baz	Artış (%)	Yüksek	Artış (%)
2018	12550	-	13073	-	13596	-
2019	13180	5,0	13730	5,0	14279	5,0
2020	13696	3,9	14266	3,9	14837	3,9
2021	14219	3,8	14811	3,8	15404	3,8
2022	14746	3,7	15360	3,7	15975	3,7
2023	15275	3,6	15911	3,6	16548	3,6
2024	15765	3,2	16422	3,2	17079	3,2
2025	16253	3,1	16930	3,1	17608	3,1
2026	16738	3,0	17436	3,0	18133	3,0
2027	17222	2,9	17940	2,9	18658	2,9

TEİAŞ ve EPDK'nın talebi doğrultusunda dağıtım şirketleri düşük, baz (referans) ve yüksek tüketim senaryolarına göre Tablo 3.3'deki gibi tahmin analizi yapmaktadır.

Şekil 3.1'de ve Tablo 3.4'de UEDAŞ bölgesindeki 2017 yılına ait aylık demant bilgileri verilmiştir [70].



Şekil 3.1: UEDAŞ 2017 yılı aylık demant grafiği.

Tablo 3.4: UEDAŞ 2017 demant tablosu.

Aylar	Anlaşma Gücü (MW)	Çekiş Gücü (MW)	Ortalama Güç (MW)
Ocak	2493,00	2047,51	2033,46
Şubat	2493,00	2064,98	2033,46
Mart	2493,00	2001,66	2033,46
Nisan	2493,00	1857,59	2033,46
Mayıs	2539,00	1797,88	2033,46
Haziran	2529,00	2086,26	2033,46
Temmuz	2539,00	2272,73	2033,46
Ağustos	2529,00	2236,20	2033,46
Eylül	2538,00	2077,20	2033,46
Ekim	2475,00	1936,43	2033,46
Kasım	2475,00	1995,67	2033,46
Aralık	2475,00	2051,09	2033,46

3.2 Bursa İli'nin Geçmiş Yıllara Göre Elektrik Tüketimi

Bu çalışma kapsamında TEİAŞ Yük Tevzi Başkanlığından 4982 sayılı Bilgi Edinme Hakkı Kanunu doğrultusunda talep edilen, Bursa iline ait 1 Ocak 2014 – 22 Şubat 2019 tarihleri arasında saatlik puant tüketim verileri kullanılarak talep tahmini yapılmıştır. Verilerin 12 saatlik örneği Tablo 3.5'deki gibidir. İki günlük örnek veri EK A'da verilmiştir.

Alınan veriler 01-01-2014 tarihinden saat 01:00 dan başlamakta ve 22-02-2019 tarihi saat 16:00'ya kadar saatlik MW cinsinden tüketim verileridir.

Veriler ham veri olup string olarak CSV* formatında edinilmiştir.

Tablo 3.5: Bursa ili tüketim verileri (a), (b).

TARİH	SAAT	MW
2014-01-01	01:00	757,60
2014-01-01	02:00	629,10
2014-01-01	03:00	582,20
2014-01-01	04:00	533,70
2014-01-01	05:00	536,90
2014-01-01	06:00	504,02
2014-01-01	07:00	542,10
2014-01-01	08:00	602,90
2014-01-01	09:00	644,70
2014-01-01	10:00	677,70
2014-01-01	11:00	816,60
2014-01-01	12:00	905,70

TARİH	SAAT	MW
2014-01-01	13:00	951,10
2014-01-01	14:00	1056,6
2014-01-01	15:00	1099,3
2014-01-01	16:00	872,99
2014-01-01	17:00	1047,4
2014-01-01	18:00	1022,2
2014-01-01	19:00	1218,4
2014-01-01	20:00	1148,6
2014-01-01	21:00	1065,6
2014-01-01	22:00	1039,8
2014-01-01	23:00	1067,9
2014-01-01	24:00	1078,5

Çalışmada elektrik talebine etki eden çevresel faktörlerden sıcaklık, nem ve basınç verileri de tahmin parametresi olarak kullanılmaktadır. Söz konusu meteorolojik veriler 4982 sayılı Bilgi Edinme Hakkı Kanunu doğrultusunda Bursa

Meteoroloji Müdürlüğü'nden talep edilmiştir. Söz konusu veriler 1 Ocak 2014 – 28 Şubat 2019 tarihleri arasındadır.

Veri setinde yer alan sıcaklık birimi °C, Nispi nem %, aktüel basınç hPa biriminde olup yük verilerinin bulunduğu tablo ile birleştirilerek tek bir tablo elde edilmiştir. Bir aylık örnek veriler EK B'de verilmiştir. Örnek veriler Tablo 3.6'da verilmiştir.

Tablo 3.6: Bursa ili meteorolojik verileri.

TARİH	SAAT	SICAKLIK	NEM	BASINÇ
2014-01-01	01:00	7,1	77	1012,1
2014-01-01	02:00	7,1	77	1012
2014-01-01	03:00	7	80	1012,5
2014-01-01	04:00	7,2	77	1012,1
2014-01-01	05:00	7,1	77	1011,8
2014-01-01	06:00	7,2	75	1011,8
2014-01-01	07:00	6,8	78	1011,9
2014-01-01	08:00	6,8	78	1012
2014-01-01	09:00	6,8	78	1012,1
2014-01-01	10:00	7,1	77	1012,5
2014-01-01	11:00	7,7	71	1012,8
2014-01-01	12:00	8	69	1012,5
2014-01-01	13:00	8,2	68	1011,9
2014-01-01	14:00	8,6	67	1011,3
2014-01-01	15:00	9,7	64	1011,1
2014-01-01	16:00	9,6	64	1010,9
2014-01-01	17:00	7,2	77	1011,5
2014-01-01	18:00	5,9	82	1011,7
2014-01-01	19:00	4,7	87	1011,8
2014-01-01	20:00	3,8	89	1011,7
2014-01-01	21:00	2,9	91	1011,7
2014-01-01	22:00	2,2	93	1011,8
2014-01-01	23:00	1,9	94	1011,8
2014-01-02	24:00	1,6	95	1011,3

3.3 Bursa İli Elektrik Yük Talep Tahmini Uygulaması

3.3.1 Verilerin Ön İşlemden Geçirilmesi

TEİAŞ Yük Tevzi Başkanlığı'ndan ve Bursa Meteoroloji İl Müdürlüğü'nden alınan veriler ham veri olup direkt olarak işlenmeye müsait değildir. Bu sebeple birtakım ön işlemlerden¹ geçirilmiştir. Uygulama kapsamında veri dizini anlamlı bir hale getirilmiş olup, aykırı, eksik ve/veya tanımsız (NULL) veriler analizi yanılmaması amacı ile gözden geçirilmiştir. Yük tevziden alınan verilerde saat 24:00'da tarih değiştirilmemiş olduğu tespit edilmiştir. Söz konusu tarihteki veriler saat 01:00 itibari ile başlamaktadır. Bütün veri setinin bu akışta olduğu tespit edilmiştir. Her günün saat 24:00 (12 am) itibari ile başladığı göz önünde bulundurularak verilerde düzeltme yapılmıştır. 2014-01-01 tarihi için 24:00'daki ölçüm bilinmediği için söz konusu tarihteki eksik veriye en yakın 4 saatlik verinin aritmetik ortalaması alınarak bir değer varsayımı gerçekleştirilmiştir.

Veri seti içerisinde tespit edilen -zaman dilimi değişiminden kaynaklı olarak oluşan- tanımsız (NULL) veriler için söz konusu tarih ve saatteki eksik veriye en yakın 4 saatlik tüketimin aritmetik ortalaması alınarak bir değer kabulü yapılmıştır. Söz konusu verilerin veri setinden çıkartılması saatlik bakımdan analizde sapmalara sebebiyet vermemesi için ortalamaya bağlı bir kabul daha uygun görülmüştür.

Verilere ek olarak Ocak 2014 – Şubat 2019 tarihleri arasındaki resmi tarihler belirlenmiş olup *Holidays* adında bir veri seti oluşturulmuştur. Bu sayede analiz kapsamında resmî tatilin etkisi de incelenmiştir.

Meteorolojik veriler aylık tablo halinde olup işleme uygun halde değildir. Bu sebeple Excel² ® üzerinde macro³ oluşturularak yük verilerinde olduğu gibi saatlik sütun formatına getirilmiştir.

Meteoroloji Genel Müdürlüğü ölçüm verilerini UTC zaman formatına göre tutmaktadır. Bu sebeple alınan verilerin yük talep verileri ile uyumlu hale getirilmesi

¹ Söz konusu uygulamalar worksheet olarak Microsoft ® Office ® Excel ® ortamında gerçekleştirilmiş olup işlemin devamına MATLAB ® 2018a üzerinde devam edilmiştir.

² Microsoft ® Office ® Excel ®

³ Microsoft ® Office ® üzerinde sıralı işlemlerin gerçekleştirilmesi için komut dizilimi

gerekmektedir. UTC'ye göre olan meteorolojik veriler, Türkiye'nin söz konusu tarihteki zaman dilimine (Tablo 3.7) göre revize edilmiştir.

Tablo 3.7: Türkiye zaman dilimi deęişim tarihleri.

Tarih	Saat	Zaman Dilimi
31 Mart 2014	03:00	UTC+3
26 Ekim 2014	04:00	UTC+2
29 Mart 2015	03:00	UTC+3
8 Kasım 2015	04:00	UTC+2
27 Mart 2016	03:00	UTC+3

Meteorolojik verilerde de günün ilk verisi saat 1:00'da ölçülmüş görünmekte ve günün son verisi 24:00 olarak girilmiştir. Bu noktada yük verilerinde olduğu gibi uygun bir şekilde düzeltme gerçekleştirilmiştir.

Veri setinde yer alan meteorolojik verilerde yer alan eksik veriler ortalama deęer öngörüsünde bulunularak tamamlanmıştır. Ortalama deęerler belirlenirken eksik veriye en yakın 4 saatlik verinin aritmetik ortalaması alınmıştır. 5 saatten daha fazla olarak birbirini takip eden eksikliklerde, aynı tarih içerisinde eksik veri grubuna en yakın 12 saatlik aritmetik ortalama alınarak varsayımda bulunulmuştur. 24 saatlik¹ eksikliklerde bir gün öncesinin verileri kullanılarak varsayımda bulunulmuştur.

Ayrıca yük verileri ve meteorolojik verilerden oluşan veri setinde yeni bir sütun oluşturularak varsayımların yapıldığı satırları temsil eden 0-1 deęerlerinden oluşan sütun matrisi de oluşturulmuştur. Bu matriste yer alan 0 deęerleri varsayımın yapıldığını, 1 deęerleri ise orijinal verilerin olduğunu göstermektedir. Bu matris vasıtası ile eksik verilerin bulunduğu satır bilgisi alınabilmekte ve gerekli görülen işlemlerde eksik verilerin veri setinden çıkartılması sağlanabilmektedir.

¹ 29 Şubat 2016 tarihinde meteorolojik ölçüm sonucu bulunmamaktadır.

Verilerin MATLAB ® 2018a'ya aktarılması

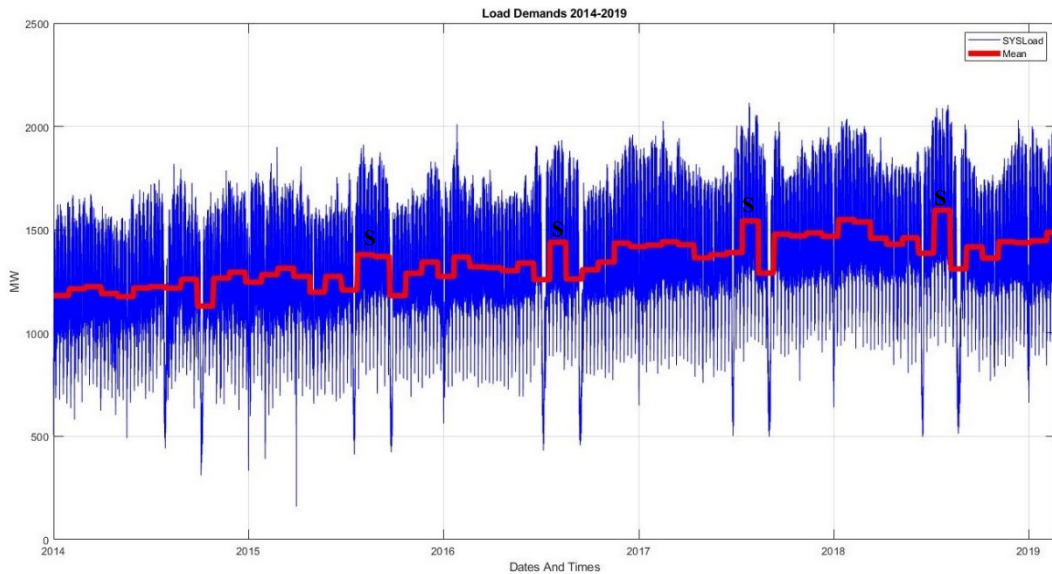
Analiz kapsamında veri setinin 45000 satırın üzerinde olması dolayısıyla analize MATLAB ® üzerinden devam edilmesi uygun görülmüş olup hazırlanan verilerin Database Explorer uygulaması üzerinden SQL kullanılarak MATLAB ®'e aktarılması sağlanabileceği gibi oluşturulan bir script (.m* file) üzerinden de gerçekleştirilebilmektedir.

Aktarımda *xlsread* komutu (bkz. EK C) kullanılmış olup;

```
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) = {'';
```

komut dizilimi ile boş olmayan, nümerik olan ve numara olmayan ifadelerin matris formunda aktarılması sağlanmıştır. Aktarılan tarih, saat ve yük verileri data adında bir veri tabanı oluşturularak matris formunda saklanmıştır. Veri seti 21-02-2019, 23:00 tarihine ilişkin satıra kadar aktarılmış olup geriye kalan 16 veri analiz dışında tutulmuştur.

Oluşturulan matrisler *table2array* komutu ile diziye dönüştürülmüştür. Ayrıca aktarım sırasında saat formatında hatalı aktarımın tespit edildiği ifadeler her bir satır koduna istinaden düzeltilerek aykırı veriler ortadan kaldırılmıştır. Oluşturulan *xlsLoad* dosyası kullanılarak kaynak *.xls** dosyası dizini belirtilmek koşulu ile verilerin matris olarak saklandığı bir veri yapısı (*struct*) oluşturulmuştur.



Şekil 3.2: 2014-2019 Bursa saatlik yük tüketimi.

Beş yıllık veri dahilinde 30 günlük ortalama tüketim değerleri göz önünde bulundurulduğunda elektrik talebinde mevsimselliğin etkisi net bir şekilde gözlemlenmektedir. Şekil 3.2’de “S” olarak işaretlenen bölgeler yaklaşık olarak Temmuz-Ağustos aylarına tekabül etmektedir.

Analiz kapsamında öncelikli olarak birtakım parametreler oluşturulmuştur. Bu kapsamda tahmin sırasında kullanılacak belirleyici parametreler mevcut veri seti kullanılarak aşağıda belirtildiği gibi belirlenmiştir. Parametreler belirlenirken parametre sayısı doğrusal regresyon uygulamaları ve MAPE hata fonksiyonu karşılaştırmaları ile belirlenmiştir.

- Bir önceki gün aynı saatteki yük
- Bir önceki hafta aynı saatteki yük
- Bir önceki dört hafta aynı saatteki yük
- Bir önceki sekiz hafta aynı saatteki yük
- Bir önceki 24 saatlik ortalama yük
- Bir önceki 12 saatlik ortalama yük
- Bir önceki 6 saatlik ortalama yük
- Resmî tatillerdeki 24 saatlik yük
- Hafta içlerindeki 24 saatlik yük
- Hafta sonları 24 saatlik yük
- Çalışılmayan günlerdeki¹ 24 saatlik yük

parametreleri kullanılarak saatlik toplamda 12 farklı özellik (*feature*) elde edilmiştir. Bunun yanı sıra meteoroloji verilerden sıcaklık, basınç ve nem bilgileri kullanılırken ayrıca

- Bir önceki gün aynı saatteki sıcaklık
- Bir önceki hafta aynı saatteki sıcaklık
- Bir önceki dört hafta aynı saatteki sıcaklık
- Bir önceki sekiz hafta aynı saatteki sıcaklık

Parametrelerinin de kullanılması ile toplamda 19 özellik (*feature*) elde edilmiştir.

¹ Çalışma günleri Pazartesi-Cuma alınmıştır.

3.3.2 Regresyon ile Yük Talep Tahmini

Talep tahmininde öncelikli olarak doğrusal regresyon analizi yapılmış olup MAPE hata fonksiyonu değerleri incelenmiştir. Yapılan analizde hata oranının düşürülmesi için ön değerlendirme yapılmıştır. Hata oranında iyileştirmenin yapılması için ek parametre değerlendirilmesi de bu aşamada yapılmıştır.

Tahmin sırasında meteorolojik veriler kullanılmadan da regresyon işlemi gerçekleştirilmiş olup sonuçlar karşılaştırılmıştır.

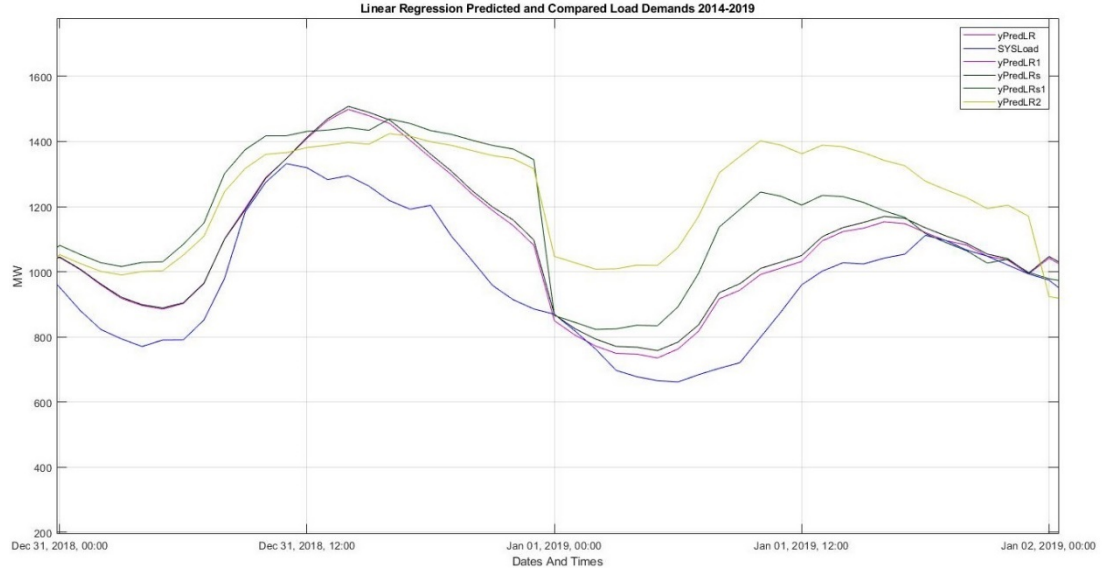
Tablo 3.8: LR ile parametre sayısının etkisinin gözlenmesi.

Parametre	Variable	MAPE (%)
Çalışma günleri ve ortalama yük alınmadan	yPredLRs2	7,8191
Çalışma günlerinin ve tatil bilgisinin eklenmesi ile	yPredLRs1	7,4196
Geçmiş ortalama yük verisinin eklenmesi ile	yPredLRs	4,6758
Meteorolojik verilerin eklenmesi ile	yPredLR1	4,6461
Meteorolojik ek parametrelerin eklenmesi ile	yPredLR	4,6438

yPredLRs1 ve yPredLRs2 tahminlerini oluşturan parametrelerde tatil bilgileri ve bir gün öncesinin ortalama yük bilgileri kullanılmıştır. yPredLRs tahmininde ise optimize edilen parametreler kullanılmıştır. 12 farklı öznelik ile birlikte gerçekleştirilen doğrusal regresyon sonucunda MAPE hata fonksiyonu %4,6758 bulunmuştur. Meteorolojik verilerin eklenmesi ile birlikte oluşturulan ek parametrelerin etkisinin çok büyük olmadığı gözlenmiştir (yPredLR ve yPredLR1). Buna bağlı olarak oluşturulan 19 parametre ile birlikte yapılan regresyon analizinde (yPredLR) MAPE hata fonksiyonu %4,6438 olarak bulunmuştur.

Gerçek değere yakınlığı ve hata fonksiyonunun %5'in altında olması sebebi ile bir önceki 12 saatlik ortalama, geçmiş 4 hafta ve geçmiş 8 haftalık ortalama tüketim verilerinin, çalışma ve tatil günleri olup olmamasının ve meteorolojik verinin tahmin sonucunu anlamlı şekilde etkilediği gözlenmektedir.

Doğrusal regresyon uygulaması sırasında üç farklı parametre grubu ile yapılan çalışma kıyaslanarak parametrelerin tahmin sonuçlarında etkisini gösteren grafik Şekil 3.3’de verilmiştir.



Şekil 3.3: İki günlük talep için tahmin parametrelerinin etkisi (31 Aralık 2018-1 Ocak 2019).

Ayrıca çalışmada belirlenen yüklere bağlı 12 parametre ile birlikte 4 yıllık veri ile 5 yıllık verinin regresyon değerleri arasında karşılaştırma yapılmış olup MAPE arasındaki ilişki incelenmiştir. Karşılaştırma Tablo 3.9’daki gibidir.

Tablo 3.9: Yıllara göre regresyon analizi MAPE kıyaslama (%).

2014-2018 Yılları arası	2014-2019 Yılları arası
4,7911	4,6758

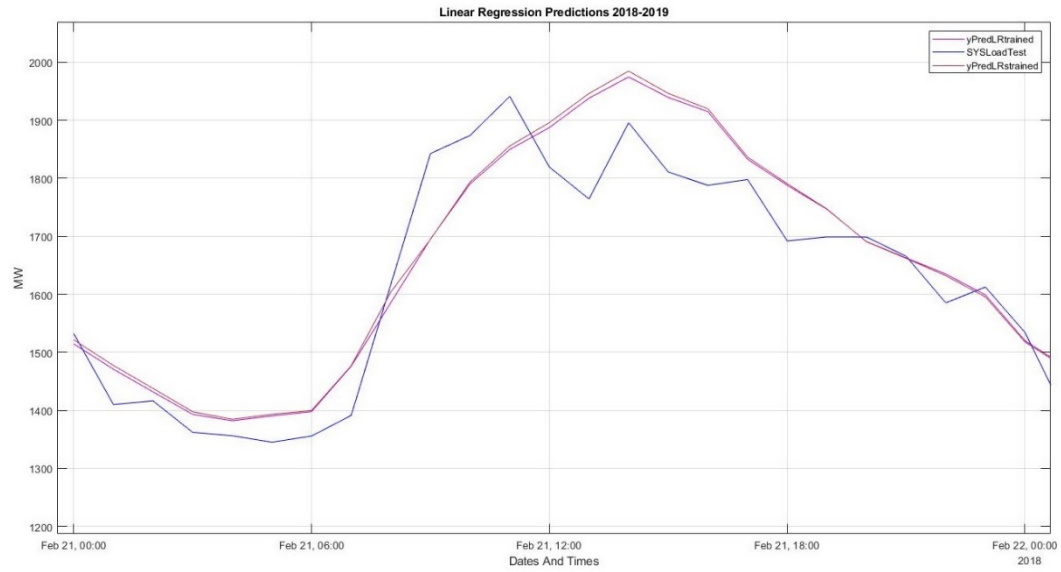
Yapılan değerlendirmeler neticesinde en uygun veri seti oluşturulmuştur. Eğitim veri seti için 20-02-2018 tarihinden önceki veriler seçilmiş olup geriye kalan veriler test verisi olarak ayrılmıştır. Eğitim veri setinde yer alan geçmiş tarihlere ait veri bulundurmayan satırlar (Nan içeren) ön işlem adımı olarak veri setinden

çıkartılmıştır. Analiz, kullanıma uygun eğitim için 34945 adet veri, test için ise 8760 adet veri ile gerçekleştirilmiştir.

Eğitim sonucu hata fonksiyonu MAPE hata fonksiyonu %4,7812 bulunmuştur. Regresyonda kullanılan girdi değişken sayısı hata fonksiyonu sonucunu doğrudan etkilemiştir. Sadece yük verilerinden oluşan eğitim ve test veri seti ile oluşturulan model ile meteorolojik verinin de yer aldığı 19 parametre kullanılarak oluşturulan modelin hata fonksiyonlarının karşılaştırılması Tablo 3.10'da, bir günlük tahmin sonuçlarının karşılaştırmasını gösteren grafik Şekil 3.4'de verilmiştir.

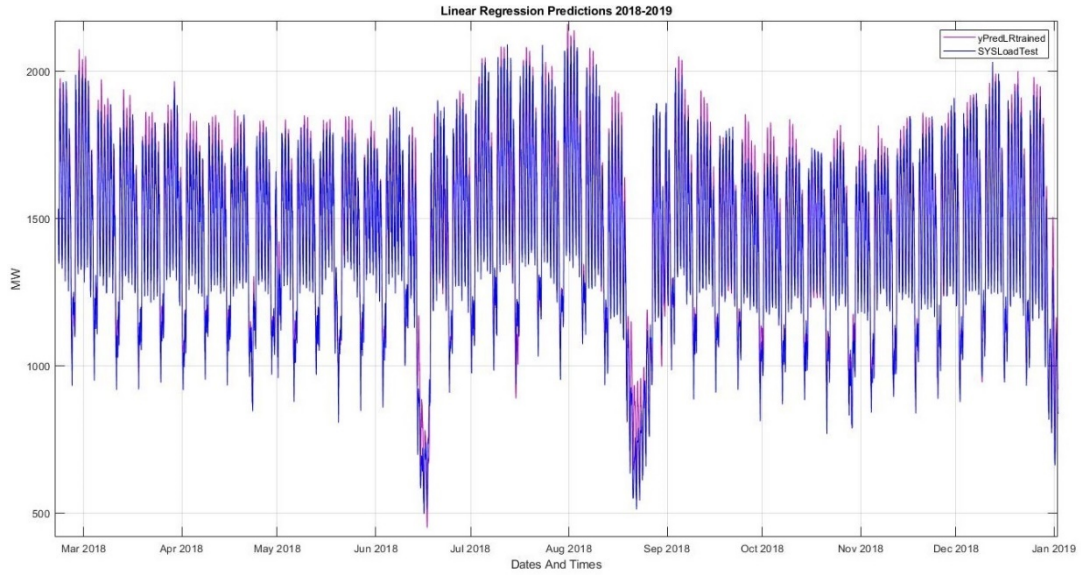
Tablo 3.10: Parametrelere göre regresyon analizi MAPE kıyaslama (%).

Yük verileri ile	Yük ve meteorolojik veriler ile
4,8065	4,7812



Şekil 3.4: 21 Nisan 2018 Doğrusal regresyon tahmin sonucu karşılaştırma.

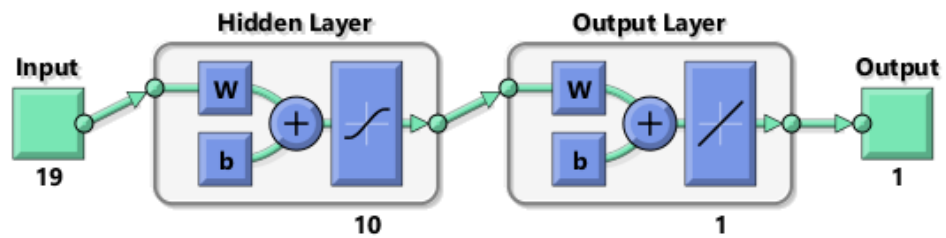
2018-2019 yılları arasındaki eğitilmiş doğrusal regresyon tahmin sonuçlarını gösteren grafik Şekil 3.5'de verilmiştir.



Şekil 3.5: 2018-2019 Doğrusal regresyon tahmin sonucu.

3.3.3 Yapay Sinir Ağları Yöntemi ile Yük Talep Tahmini

Veri seti üzerinde hazırlanan ek parametreler ile birlikte Neural Network Fitting Toolbox¹'i kullanılarak yapay sinir ağları uygulaması yapılmıştır. Bu noktada girdi verileri 45048x19 boyutunda bir matris, çıktı verileri 45048x1 boyutunda bir matris olarak belirlenmiştir. Eğitim veri seti olarak, veri setinin tamamından rastgele %70'i ayrılmış olup toplam örnek sayısı 31534 olarak belirlenmiştir. Validasyon için veri setinin rastgele %15'i, test verisi için de kalan %15'i 6757 adet örnek ayrılmıştır. Toolbox üzerinde gizli nöron sayısı, yapılan testler doğrultusunda 10 olarak belirlenmiştir. Oluşturulan model Şekil 3.6'da verilmiştir.



Şekil 3.6: ANN modeli.

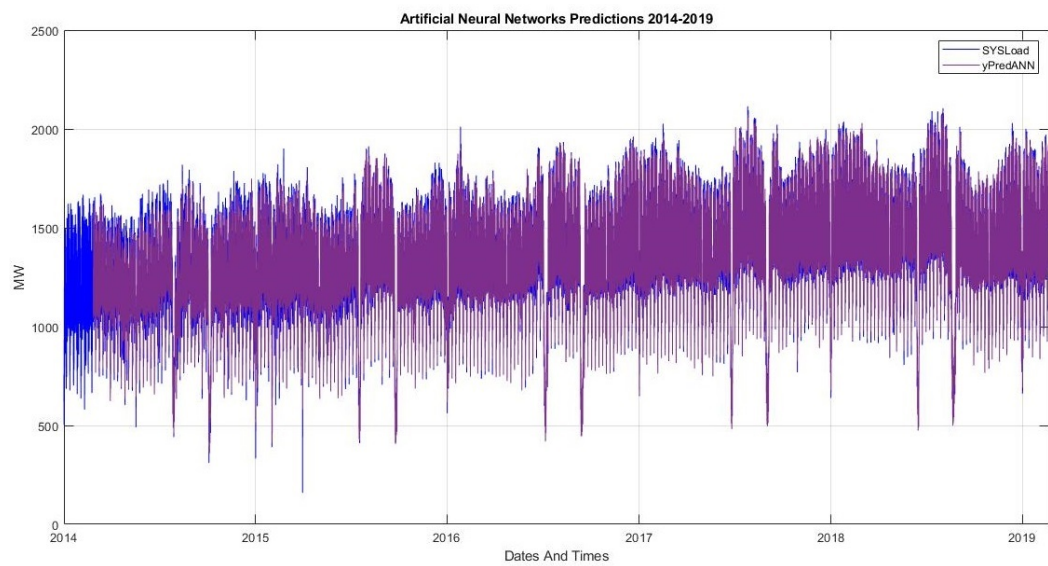
¹ nfool

Eğitim Levenberg-Marquardt modeline göre yapılmış olup meteorolojik veri kullanılmadan (12 parametre ile) ve meteorolojik veri kullanılarak (19 parametre ile) tekrarlanmıştır. Modellerin MSE hata sonuçları Tablo 3.11'deki gibidir.

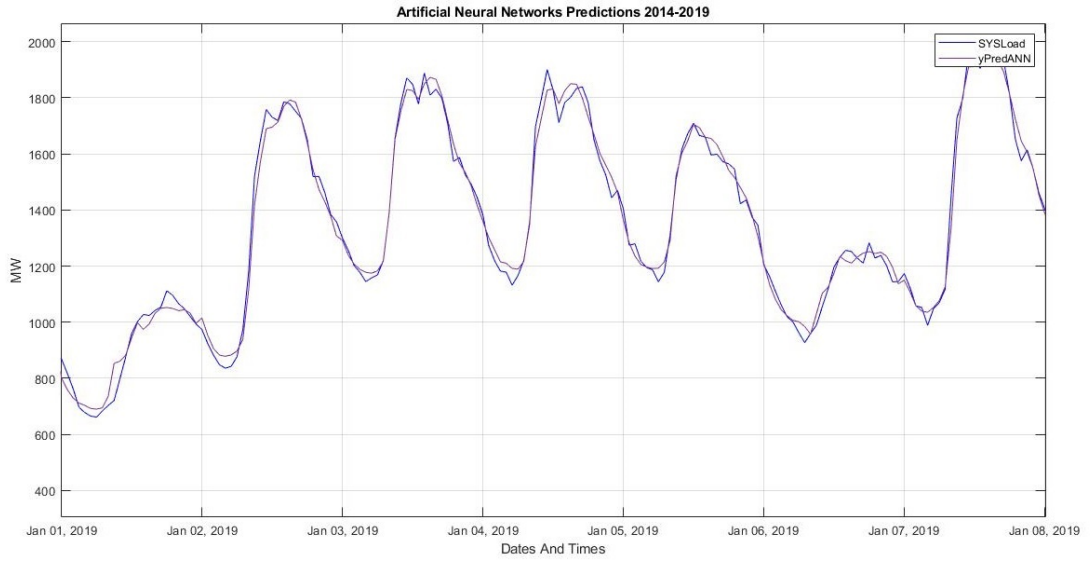
Tablo 3.11: ANN MSE Hata karşılaştırma.

12 Parametre ile	Örnek sayısı	MSE	Regülerizasyon
Eğitim Veri Seti	31534	2186.54416e-0	9.86946e-1
Validasyon	6757	2213.00566e-0	9.87086e-1
Test	6757	2208.36203e-0	9.86603e-1
19 Parametre ile	Örnek sayısı	MSE	Regülerizasyon
Eğitim Veri Seti	31534	1956.19447e-0	9.88333-1
Validasyon	6757	1923.13097e-0	9.88633-1
Test	6757	1964.46859e-0	9.88225-1

Yalnızca yük verileri kullanılarak geliştirilen parametreler ile oluşturulan yapay sinir ağları modelinin MAPE hata fonksiyonu %2,7281 olarak bulunmuştur. Meteorolojik veriler kullanılarak geliştirilen ek parametreler ile oluşturulan modelde MAPE hata fonksiyonu %2,5821 olarak bulunmuştur. Bu modele göre oluşturulan analiz sonucunda meteorolojik ek parametreler ile geliştirilmiş modelin tahmin grafiği Şekil 3.7'de, 1 günlük uyumu ise Şekil 3.8'deki gibidir.



Şekil 3.7: ANN tahmin grafiği.



Şekil 3.8: ANN modeli uyumu (2019 ilk hafta).

3.3.4 Destek Vektör Regresyon Yöntemi ile Yük Talep Tahmini

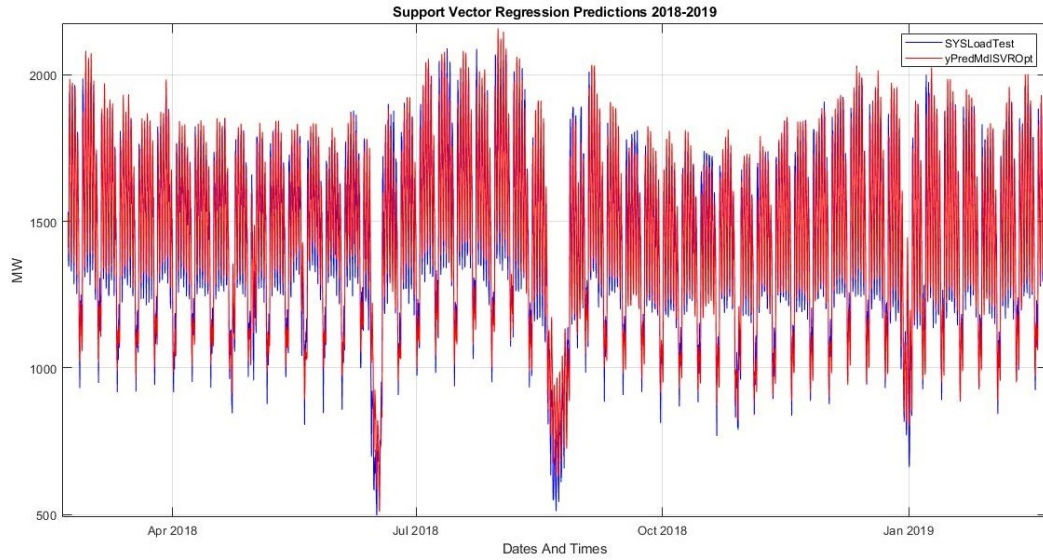
36808x19 boyutundaki veri seti üzerinde MATLAB ® ile destek vektör regresyon ile model oluşturulmasına olanak tanıyan *fitrsvm* fonksiyonu kullanılmıştır. Söz konusu fonksiyon modelin eğitilmesini ya da üzerinde cross-valide edilmesini sağlamaktadır.

Veri seti üzerinde doğrusal regresyon ile anlamlı bir sonuç bulunmadığı için model Gaussian kernel fonksiyonu kullanılarak yeniden kurgulanmıştır. Öncelikli olarak verilerin belirli bir tarihten önceki zamana ait olan kısmı eğitim verisi için, geriye kalan kısmı ise test verisi olarak ayrılmıştır.

Eğitim veri seti için 20-02-2018 tarihinden önceki veriler seçilmiş olup geriye kalan veriler test verisi olarak ayrılmıştır. Eğitim veri setinde yer alan geçmiş tarihlere ait veri bulundurmayan satırlar (Nan içeren) ön işlem adımı olarak veri setinden çıkartılmıştır. Analizde gerçekleştirilen eğitim; 34581 adet, test; 8741 adet veri ile gerçekleştirilmiştir.

Kernel fonksiyonu *Gaussian* olarak belirlenmiş ve verinin standardizasyonu ile birlikte otomatik kernel scale yapılarak model yeniden oluşturulmuştur. SVO

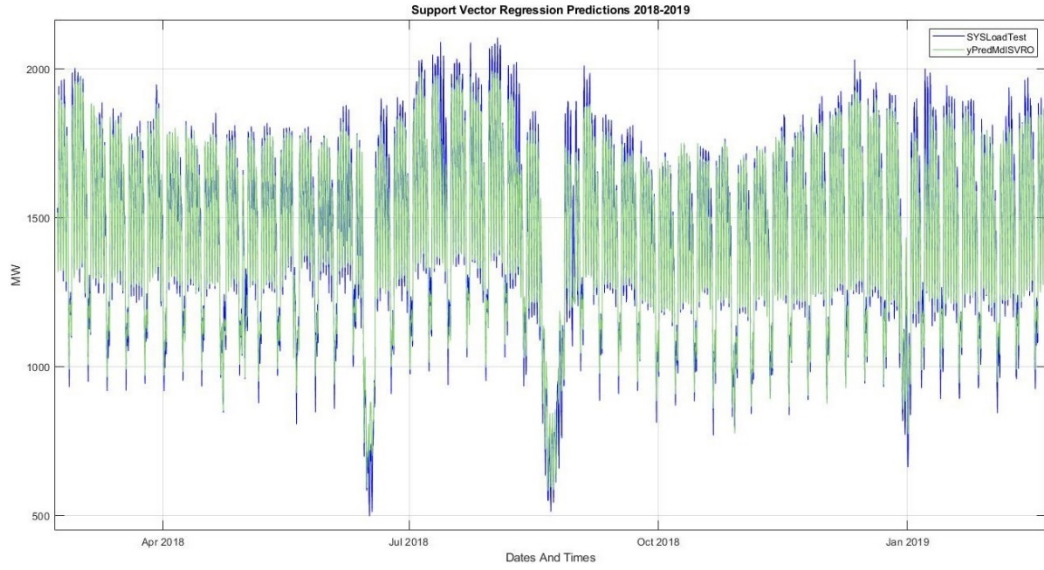
algoritması ile oluşturulan modelin en uygun gaussian kernel scale 116,9698 olarak tespit edilmiştir. *fitrsvm* fonksiyonu içerisindeki hiper parametre optimizasyonu ile model için en uygun parametrelerin belirlenmesi doğrultusunda model tekrardan oluşturulmuştur. MAPE hata fonksiyonu %4,4179 olarak bulunmuştur.



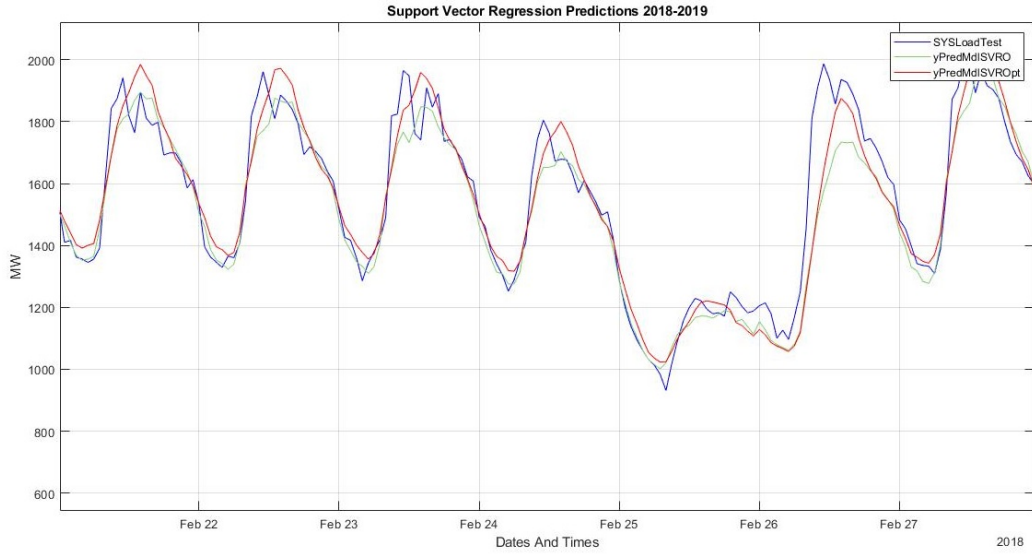
Şekil 3.9: Gaussian Kernel SVR modeli tahmin sonuçları.

Meteorolojik veriler kullanılmadan 12 parametre ile oluşturulan modelde; kernel fonksiyonu *Gaussian* olarak belirlenmiş ve verinin standardizasyonu ile birlikte otomatik kernel scale yapılarak model yeniden oluşturulmuştur.

SVO algoritması ile oluşturulan modelin en uygun gaussian kernel scale 2,4310 olarak tespit edilmiştir. *fitrsvm* fonksiyonu içerisindeki hiper parametre optimizasyonu ile model için en uygun parametrelerin belirlenmesi doğrultusunda model tekrardan oluşturulmuştur. MAPE hata fonksiyonu %3,0308 olarak bulunmuştur. Elde edilen modelin tahmin sonuçlarını gösteren grafik Şekil 3.10'da verilmiştir.



Şekil 3.10: Gaussian Kernel SVR modeli tahmin sonuçları (12 Parametre).



Şekil 3.11: Parametre optimizasyonu ile SVR 1 Ocak 2019 tahmin sonucu karşılaştırma (19 Parametre).

3.3.5 Regresyon Karar Ağacı Yöntemi ile Yük Talep Tahmini

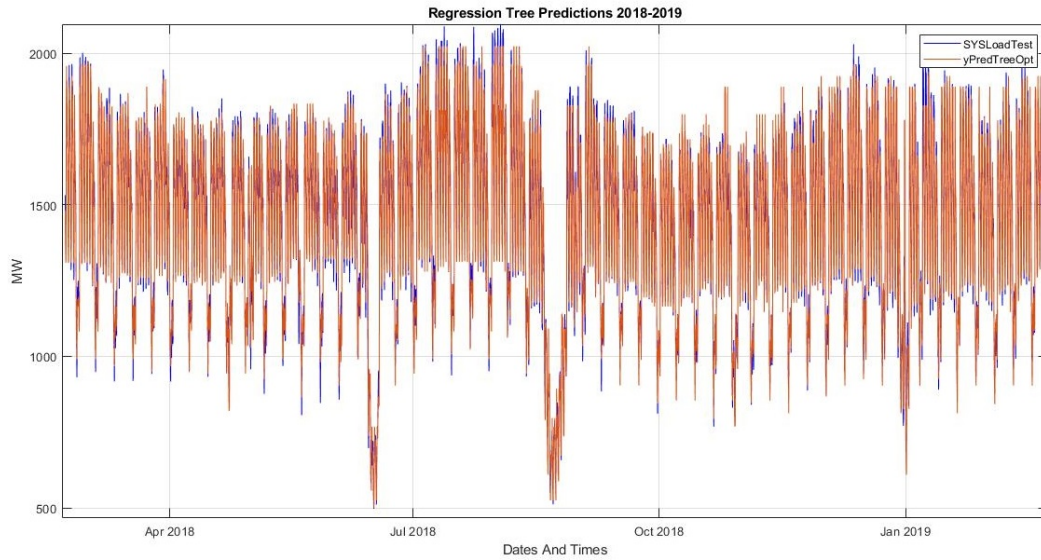
Veri seti üzerinde MATLAB ® ile regresyon karar ağacı uygulaması, *fitrtree* fonksiyonu kullanılarak gerçekleştirilmiştir.

Verilerin belirli bir tarihten önceki zamana ait olan kısmı eğitim verisi için ayrılmıştır. Geriye kalan kısmı ise test verisi olarak ayrılmıştır.

Eğitim veri seti için 20-02-2018 tarihinden önceki veriler seçilmiş olup geriye kalan veriler test verisi olarak ayrılmıştır. Eğitim veri setinde yer alan geçmiş tarihlere ait veri bulundurmayan satırlar (Nan içeren) ön işlem adımı olarak veri setinden çıkartılmıştır. Analizde kullanıma uygun eğitim 34851 adet, test 8741 adet veri ile gerçekleştirilmiştir.

Fonksiyon girdi değişkenleri doğrultusunda regresyon ağacı oluşturmaktadır. Fonksiyonda kategorik tahmin değişken sayısı 4^1 olarak belirtilmiştir. Veri seti için daha anlamlı sonuç elde edilebilmesi için parametre optimizasyonu gerçekleştirilmiştir. Optimizasyon için *fitree* fonksiyonu içerisinde yer alan Hiper parametrelerin otomatik olarak gerçekleştirilmesi (*Bayesian Optimizastion*) sağlanmıştır.

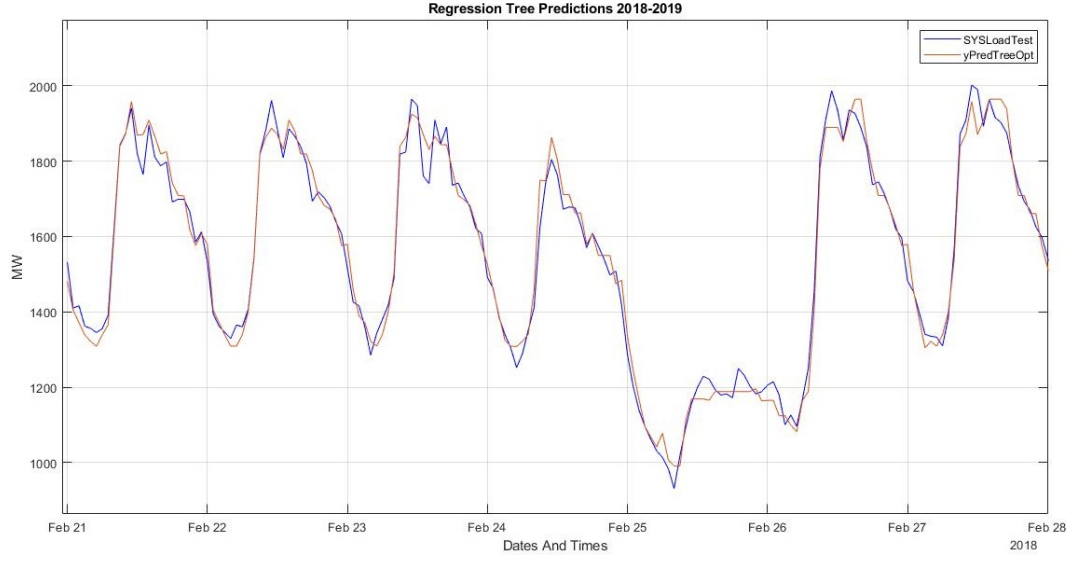
Modelin MAPE hata fonksiyonu %2,8991 olarak bulunmuştur. Şekil 3.12'de oluşturulan modelin tahmin sonuçlarını gösteren grafik verilmiştir.



Şekil 3.12: Regresyon karar ağacı ile oluşturulan modelin tahmin sonuçları.

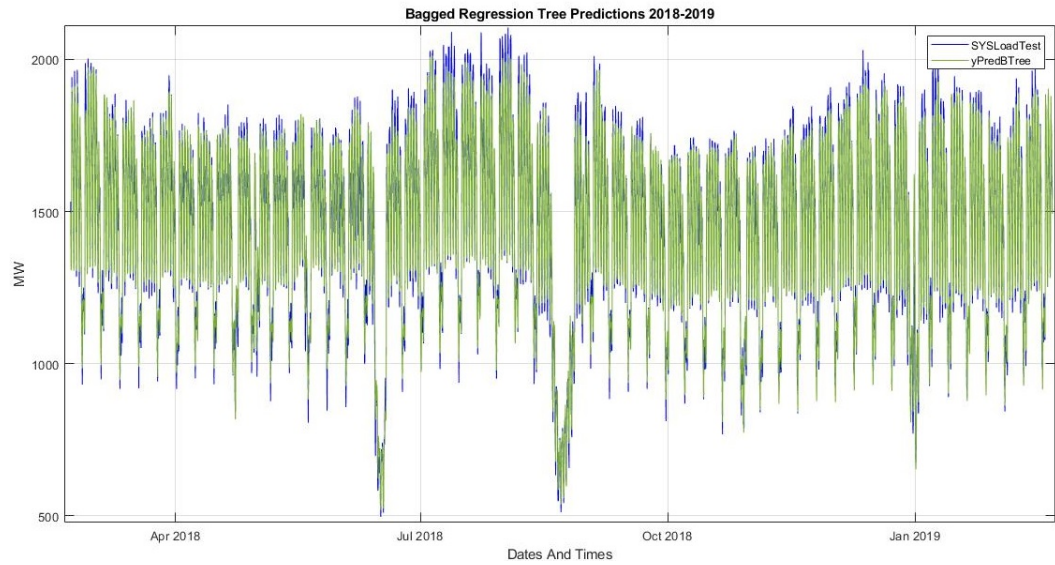
¹ Resmi tatil, hafta içi, hafta sonu, çalışılmayan gün

Eđitim veri setindeki test verilerinin ilk haftası olan 21-28 Şubat 2018 tarihindeki tahmini gösteren grafik Şekil 3.13’de verilmiştir.



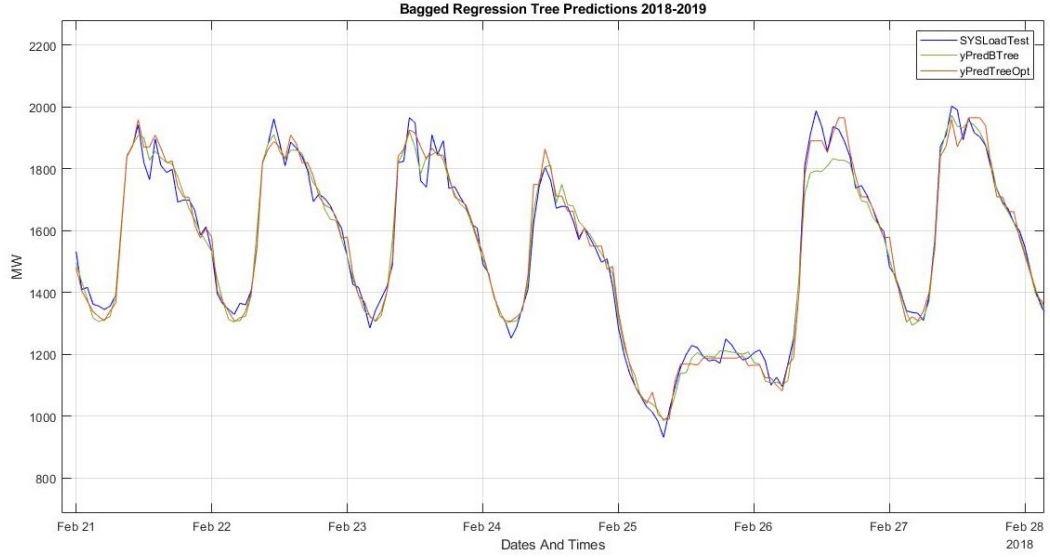
Şekil 3.13: Regresyon karar ağacı 21-28 Şubat 2018 tahmin sonuçları.

Ön işleminden geçirilmiş eğitim veri seti üzerinde aşırı öğrenmeye karşı bagged regresyon karar ağacı algoritması kullanılmıştır. Uygulama için MATLAB ® kütüphanesindeki *TreeBagger* fonksiyonu kullanılmış olup ağaç sayısı 20, *minleaf* parametre optimizasyonu ise 17 seçilmiştir. Oluşturulan modelin doğruluđu Şekil 3.14’de verilmiştir. MAPE hata fonksiyonu %2,4298 olarak bulunmuştur.



Şekil 3.14: Bagged regresyon DT ile oluşturulan modelin tahmin sonuçları.

Regresyon karar ağacı ile ilgili iki farklı model karşılaştırıldığında bagged ağaç uygulamasının daha verimli çalıştığı gözlemlenmiştir. Şekil 3.15’de görüldüğü gibi gerçek veriye Bagged regresyon karar ağacı daha yakın davranış sergilemektedir.



Şekil 3.15: Bagged regresyon ağacı Ağustos 2018 15 günlük tahmin sonucu.

Meteorolojik verilerin kullanılmadığı ve yük verileri ile geçmiş verilere dayanan 12 parametre kullanılarak yapılan uygulamada regresyon karar ağacı modelinin MAPE hata fonksiyonu %2,8583 olarak bulunmuştur. Söz konusu parametreler ile oluşturulan Bagged regresyon ağacı modelinde ise MAPE hata fonksiyonu %2,6363 olarak bulunmuştur. Karşılaştırma sonuçları Tablo 3.12’deki gibidir.

Tablo 3.12: Karar ağacı hata fonksiyonu karşılaştırma.

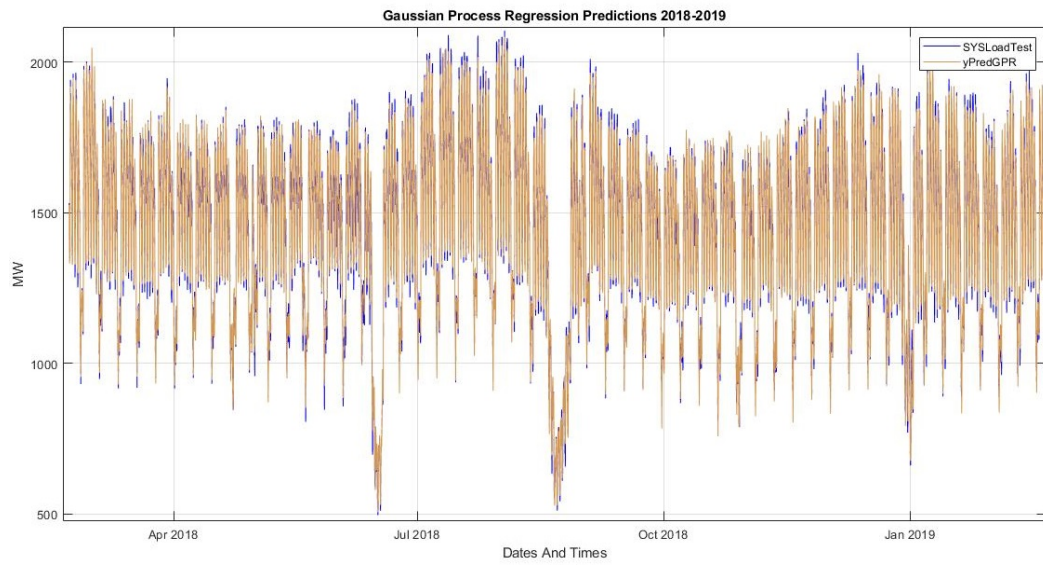
	<u>Yük verileri ile</u>	<u>Yük ve meteorolojik veriler ile</u>
<u>Bagged Tree</u>	2,6363	2,4298
<u>Reg Tree</u>	2,8583	2,8991

3.3.6 Gauss Süreç Regresyonu Uygulaması

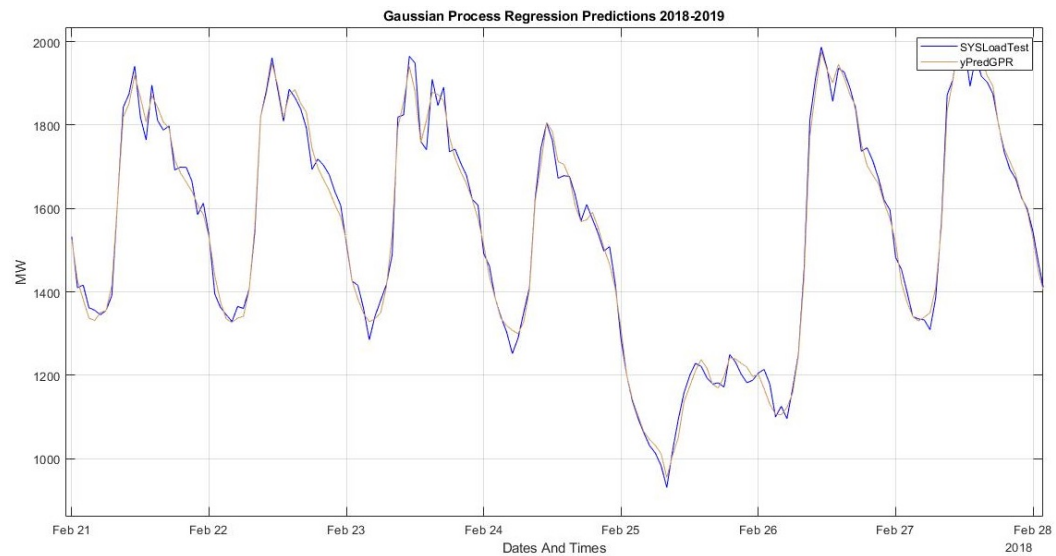
GPR uygulamasında diğer yöntemlerde izlenen benzer yol izlenmiş olup çalışmada, eğitim veri seti olarak 20-02-2018 tarihinden önceki veriler seçilmiş olup

geriye kalan veriler test verisi olarak ayrılmıştır. Eğitim veri setinde yer alan geçmiş tarihlere ait veri bulundurmeyen satırlar (Nan içeren) ön işlem adımı olarak veri setinden çıkartılmıştır. Analizde kullanıma uygun eğitim 34581 adet, test 8741 adet veri ile gerçekleştirilmiştir.

GRP uygulaması MATLAB ® kütüphanesindeki *fitrgp* fonksiyonu ile gerçekleştirilmiştir. Oluşturulan modelde MAPE hata fonksiyonu %1,8729 olarak bulunmuştur. Tahmin sonucu Şekil 3.16'da, 1 haftalık karşılaştırmayı gösteren grafik Şekil 3.17'de verilmiştir.



Şekil 3.16: GPR 2018-2019 tahmin sonucu.



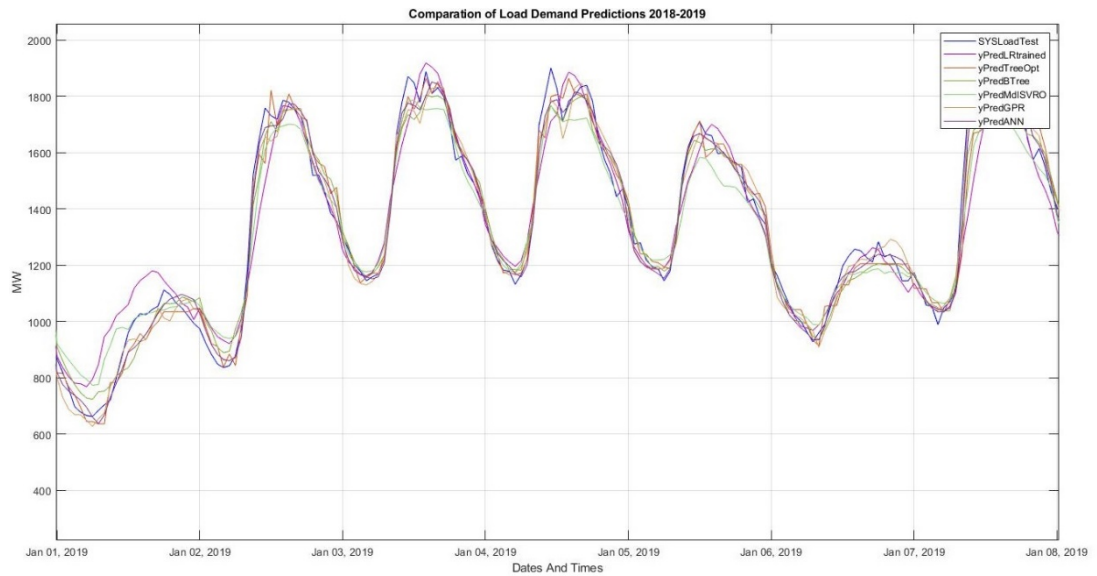
Şekil 3.17: GPR 21-28 Şubat 2018 tahmin sonucu.

4. SONUÇ VE ÖNERİLER

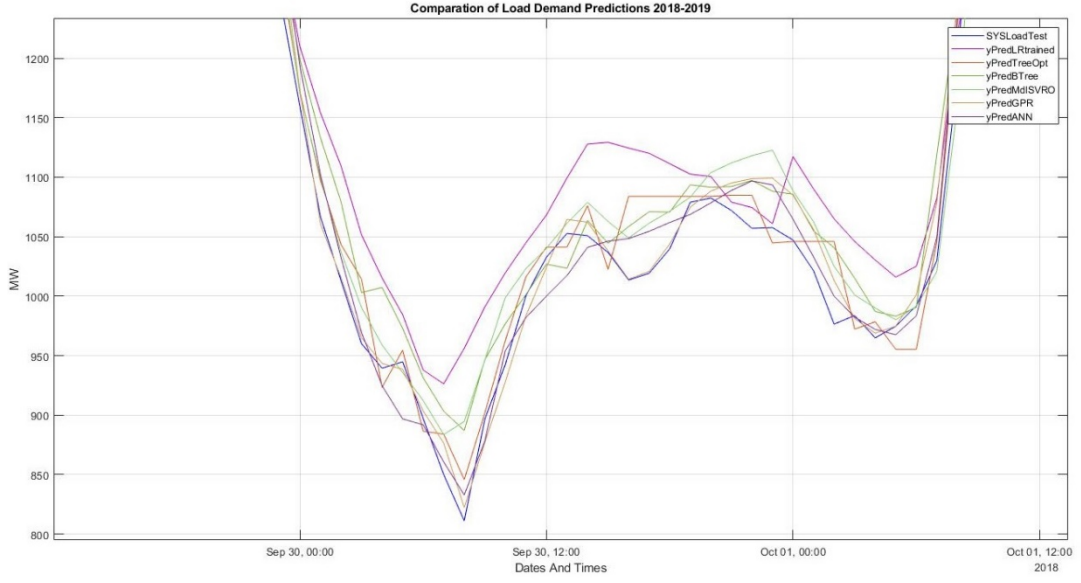
Çalışma kapsamında gerçekleştirilen elektrik talep tahmini uygulaması sonucunda regresyon, yapay sinir ağları, regresyon ağacı, destek vektör regresyon, gaussian proses regresyon yöntemleri ile kullanılarak oluşturulan modellerin hata fonksiyonu karşılaştırmaları Tablo 4.1’de ve özellikle 2019 yılının ilk haftasındaki saatlik yük tahminlerini gösteren grafik Şekil 4.1’de bir günlük tahmin sonucunu gösteren grafik Şekil 4.2’de verilmiştir.

Tablo 4.1: MAPE hata fonksiyonu karşılaştırma (%).

	Regresyon		Karar Ağacı		ANN	SVR	GPR
	LR	Trained LR	Reg. D. Tree	Bagged D. Tree			
12 Par.	4,6758	4,8065	2,8583	2,6363	2,7281	3,0308	2,0921
19 Par.	4,6438	4,7812	2,8991	2,4298	2,5821	4,4149	1,8729

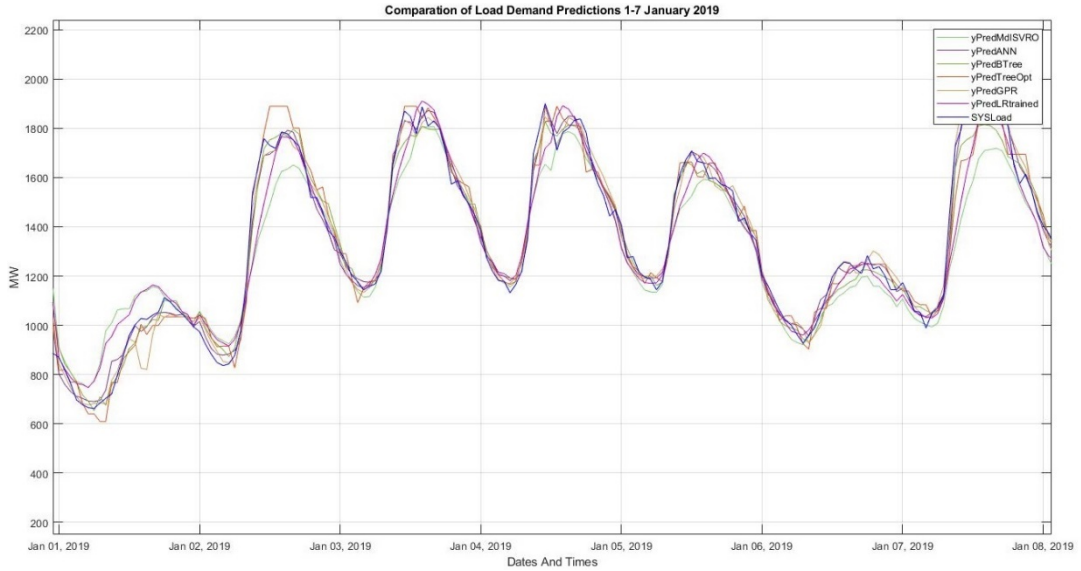


Şekil 4.1: Yüke bağlı 1 haftalık tahmin karşılaştırması.



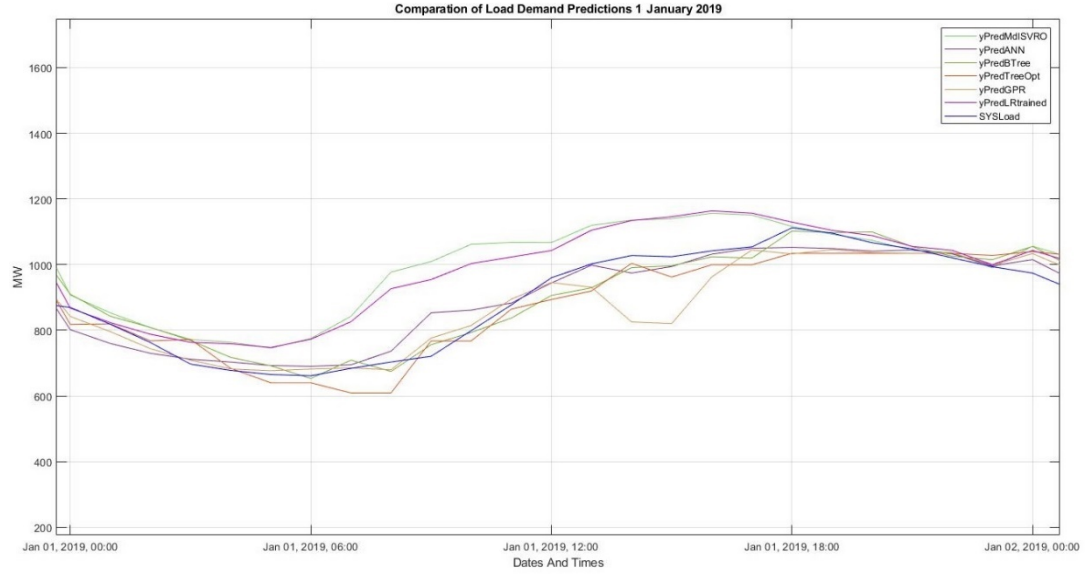
Şekil 4.2: Yüke bağlı 1 Ocak 2019 günlük tahmin karşılaştırması.

Yük ve meteorolojik veriye bağlı olarak gerçekleştirilen tahmin uygulamasında algoritmalar arası 1 haftalık karşılaştırma şekil 4.3’de verilmiştir.



Şekil 4.3: Yük ve meteorolojik veriye bağlı 1 haftalık tahmin karşılaştırması.

Yük ve meteorolojik veri ile yapılan uygulamada 1 Ocak 2019 tarihindeki tahmini gösteren grafik şekil 4.4'deki gibidir. Şekilde de görüldüğü gibi gerçek veriye en yakın GPR, ANN ve Bagged Karar Ağacı algoritmaları çalışmaktadır.



Şekil 4.4: Yük ve meteorolojik veriye bağlı bir günlük tahmin karşılaştırması (1 Ocak 2019).

Tablo 4.2: Algoritma eğitim süresi karşılaştırma.

	Regresyon		Karar Ağacı		ANN	SVR	GPR
	LR	Trained LR	Reg. D. Tree	Bagged D. Tree			
Eğitim Hızı	Çok Hızlı	Çok Hızlı	Yavaş	Çok Hızlı	Çok Hızlı	Çok Yavaş	Hızlı

Tablo 4.2'de oluşturulan algoritmaların eğitim sürelerinin karşılaştırmaları verilmiştir. Özellikle hata fonksiyonları da göz önünde bulundurulduğunda ANN ve GPR algoritmalarının elektrik gücü talep tahmini uygulamalarında kullanımının uygun olabileceği bulgusuna ulaşılmıştır.

Meteorolojik verilerin tahmin sonucunu özellikle regresyon karar ağacı modeli dışında daha iyi etkilediği gözlemlenmiştir.

Analiz sonuçları gerçek değerler ile karşılaştırıldığında LR ani güç değişimlerine tam uyum sağlayamadığı gözlenmiştir. Regresyon karar ağacı çözümlerinde hata fonksiyonu düşük olmasına karşın, belirli saat aralıklarında güç talebini sabit olarak değerlendirmiştir. Bu sebeple algoritma zamana bağlı olarak yanıltıcı olabilmektedir. Ayrıca karar ağacı algoritmalarının ezberci yönü de tahmin güvenilirliğini olumsuz etkilemektedir.

Bagged regresyon ağacı genel anlamda olumlu performans göstermesine karşın bazı günlerdeki tahminlerinde %10 mertebelerinde hata yaptığı tespit edilmiştir. Genel uyum açısından kabul edilebilir bir yöntem olarak kullanılması uygundur.

Tahminde ANN ve GPR algoritmalarının genel anlamda doğru bir şekilde çalıştığı gözlenmiştir. Büyük ve tekrarlayan verilerde ANN algoritmasının ezberci yaklaşımda bulunabilmesi durumu algoritmanın dezavantajları arasında yer almaktadır. Modelleme hızı açısından bagged regresyon ağacı ve ANN algoritmaları hızlı ve pratik çözüm sunabilmektedir.

Sonuç olarak geçmiş tüketim yük verilerinin yanında elektrik tüketimini önemli ölçüde etkileyen sıcaklık, basınç ve nem verileri ile kullanılarak oluşturulan GPR yöntemi ile oluşturulan modelin kullanılarak gerçekleştirilen yük talebi tahmininde diğer metotlara göre doğru ve gerçeğe yakın tahminlerin elde edileceği bulgusuna ulaşılmıştır. Hızlı modelleme ihtiyacının olması durumunda ise ANN ve bagged regresyon ağacı modellemeleri ile kabul edilebilir tahminler gerçekleştirilebilmektedir.

İleriki çalışmalarda, ekonomik ve sosyal faktörlerin de parametre olarak kullanılması ile elektrik yük talep tahmini gerçekleştirilebilir. Özellikle kişi başına düşen milli gelir miktarı, incelenen bölgedeki ekonomik durum ve sosyal etmenler, ihracat oranları, ulusal ve uluslararası siyasi ilişkiler incelenerek analiz kapsamı detaylandırılabilir.

5. KAYNAKLAR

- [1] EPDK, “Elektrik Piyasası Talep Tahminleri Yönetmeliği”, *Resmî Gazete*, 29705, (2016).
- [2] TEİAŞ, “Yük Tevzi Bilgi Sistemi [online]”, (22 Şubat 2019), <https://ytbs.teias.gov.tr>.
- [3] TEİAŞ, “Piyasa Yönetim Sistemi [online]”, (22 Şubat 2019), <https://tpys.teias.gov.tr>.
- [4] EPIAŞ, “2017 Yılı Faaliyet Raporu”, İstanbul, (2018).
- [5] NASA, “NASA Earth Observatory [online]”, (17 Kasım 2018), <http://earthobservatory.nasa.gov/Features/WxForecasting/wx2.php>, (2002).
- [6] IEA, “Global energy and CO2 status report,” *International Energy Agency*, France, (2018).
- [7] Vinagre E., Pinto T., Ramos S., Vale Z. and Corchado J. M., “Electrical Energy Consumption Forecast Using Support Vector Machines”, *27th International Workshop on Database and Expert Systems Applications*, Porto, 171-175 (2016).
- [8] Mostafa Al M. and K. Nagasaka K., “Artificial Neural Networks Applied to Long-term Electricity Demand Forecasting,” *Proceedings of the Fourth International Conference on Hybrid Intelligent Systems*, Kitakyūshū, (2004).
- [9] Akataş, N., “Kısa süreli rüzgar tahmini için WRF model performansının analizi ve rüzgar gücü uygulamaları”, Yüksek Lisans Tezi, *İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü*, Meteoroloji Mühendisliği Anabilim Dalı, Atmosfer Bilimleri Bilim Dalı, İstanbul, (2015).

- [10] Ener Rüßen S., “Küresel güneş ışınımı tahmin etmek için uydu görüntülerini ve güneşlenme süresini birleştirme,” Doktora Tezi, Orta Doğu Teknik Üniversitesi Fen Bilimleri Enstitüsü, Fizik Anabilim Dalı, Ankara, (2013).
- [11] Elektrik Mühendisleri Odası (EMO), “Enerji İstatistikleri”, TMMOB, Ankara, (2019).
- [12] T.C. Enerji ve Tabii Kaynaklar Bakanlığı, “Elektrik [online]”, (24 Nisan 2019), <https://www.enerji.gov.tr/tr-TR/Sayfalar/Elektrik>, (2019)
- [13] “Elektrik Piyasası Kanunu (6446)”, *Resmî Gazete*, 53 (28603), (2013).
- [14] Bulut A. ve Sohtaoglu N. H., “Ulusal Elektrik Şebekelerinin Avrupa Kıtası Odaklı Entegrasyonu”, *Elektrik-Elektronik-Bilgisayar Mühendisliği 12. Ulusal Kongresi ve Fuarı*, Eskişehir, (2007).
- [15] Arthur N., “Almanya’yı elektrik kesintilerinden güneş enerjisi kurtardı [online]”, (04 Kasım 2018), <https://yesilekonomi.com/almanyayi-elektrik-kesintilerinden-gunes-enerjisi-kurtardi>, (2012).
- [16] Starn J., “Germany set to pay customers for electricity usage as renewable energy generation creates huge power surplus [online]”, (04 Kasım 2018), <https://www.independent.co.uk/environment/germany-grids-paying-electricity-customers-renewable-energy-power-surplus-wind-solar-generation-a8022576.html>, (2017).
- [17] Atalay M., “Makine Öğrenme Yöntemleri Yardımıyla Tüketim İstatistiklerine Göre Talep Tahmini”, Yüksek Lisans Tezi, *T. C. İstanbul Kültür Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, (2013).
- [18] EPDK, “2017 Yılı Elektrik Piyasası Gelişim Raporu,” Elektrik Piyasaları Düzenleme Kurulu, Ankara, (2018).
- [19] Gross G. and Galiana F. D., “Short-term load forecasting” *Proceedings of the IEEE*, 75(12), 1558-1573, (1987).

- [20] Feng W., Keng Y. E., Qi L. Y., Jun L. and Shan Y. C., “Short-term Load Forecasting Based On Weather Information”, *1998 International Conference on Power System Technology*, Beijing, 572-575, (1998).
- [21] Contaxi E., Delkis C., Kavatza S. D. and Vournas C. D., “The Effect of Humidity in a Weather-Sensitive Peak Load Forecasting Model”, *Power Systems Conference and Exposition*, Atlanta, 1528-1534, (2006).
- [22] Türkay B. E., “Türkiye'nin Uzun Dönem Puant Yük Talebinin ve Enerji İhtiyacının Tahmin Edilmesi”, *Elektrik Mühendisliği*, 453, 31-33, (2015).
- [23] UCTE, “System Adequacy Forecast 2008-2020”, UCTE, Brussel, (2008).
- [24] Chen Y. H., Hong W., Shen W. and Huang N. N., "Electric Load Forecasting Based on a Least Squares", *Energies*, 9 (70), 70-83, (2016).
- [25] Kaynar O., Özekicioğlu H. ve Demirkoparan F., “Forecasting of Turkey’s Electricity Consumption with Support Vector Regression and Chaotic Particle Swarn Algorithm”, *Journal of Administrative Sciences*, 15 (29), 2011-224, (2017).
- [26] Türkay B. E. ve Demren D., “Electrical Load Forecasting Using Support Vector Machines”, *International Conference on Electrical and Electronics Engineering*, Nagpur, 49-53, (2011).
- [27] Elattar E.E, Goulermas J. and Wu Q. H., “Electric Load Forecasting Based on Locally Weighted Support Vector Regression”, *IEEE Transactions on Systems, Man, And Cybernetics—Part C: Applications And Reviews*, 40 (4), 438-447, (2010).
- [28] Hussain L., Nadeem M. S. and Shah S. A. A., “Short Term Load Forecasting System Based On Support Vector Kernel Methods”, *International Journal of Computer Science & Information Technology*, 6 (3), 93-102, (2014).
- [29] Nazarko J. and Zalewski W., “The Fuzzy Regression Approach to Peak Load Estimation”, *IEEE Transactions on Power Systems*, 14 (3), 809-814, (1999).

- [30] Sarhani M. and El Afia A., “Electric Load Forecasting Using Hybrid Machine Learning Approach Incorporating Feature Selection”, *Proceedings of the International Conference on Big Data Cloud and Applications*, Morocco, (2015).
- [31] Matijaš M., “Electric Load Forecasting Using Multivariate Meta-Learning”, Ph.D Thesis, *University of Zagreb*, Zagreb, (2013).
- [32] Božić M. and Stojanović M., “Application of SVM Methods for Mid-Term Load Forecasting”, *Serbian Journal Of Electrical Engineering*, 8 (1), 73-83, (2011).
- [33] Setiawan A., Koprinska I. and Agelidis V. G., “Very Short-Term Electricity Load Demand Forecasting Using Support Vector Regression”, *2009 International Joint Conference on Neural Networks*, Atlanta, GA, USA, (2009).
- [34] Ghanbari A., Naghavi A., Ghaderi S., and Sabaghian M., “Artificial Neural Networks and Regression Approaches Comparison for Forecasting Iran’s Annual Electricity Load”, *Powereng*, Lisbon, 675-679, (2009).
- [35] Omidi A., Barakati S. M. and Tavakoli S., “Application of nuSupport Vector Regression in Short-Term Load Forecasting”, *The 20th Iranian Electrical Power Distribution Conference*, Zahedan, 32-36, (2015).
- [36] Sarhani M. and El Afia A., “Feature Selection and Parameter Optimization of Support Vector Regression for Electric Load Forecasting”, *2nd International Conference on Electrical and Information Technologies*, Rabat, (2016).
- [37] Khan A. R., Razzaq S., Alquthami T., Moghal M. R., Amin A. and Mahmood A., “Day Ahead Load Forecasting for IESCO Using Artificial Neural Network and Bagged Regression Tree”, *1st International Conference on Power, Energy and Smart Grid*, Mirpur, (2018).
- [38] Baghel M., Ghosh A., Singh N. K. and Singh A. K., “Short –Term Electric Load Forecasting Using SVR Implementing LibSVM Package and Python Code”, *IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering*, Varanasi, 485-489, (2016).

- [39] Mohan N., Kumar S. S. and Soman K. P., “A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model”, *Applied Energy*, 232, 229-224, (2018).
- [40] Leith D. J., Heidl M. and Ringwood J. V., “Gaussian Process Prior Models for Electrical Load Forecasting”, *8th International Conference on Probabilistic Methods Applied to Power Systems*, Ames, (2004).
- [41] Blum M. and Riedmiller M., “Electricity Demand Forecasting using Gaussian Processes”, *Trading Agent Design and Analysis*, Freiburg, 10-13, (2013).
- [42] Hachino T., Takata H., Fukushima S. and Igarashi Y., “Short-Term Electric Load Forecasting Using Multiple Gaussian Process Models”, *World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering*, 8 (2), 447-452, (2014).
- [43] Gültekin F., “Regresyon Analizi [online]”, (17 Kasım 2018) <https://www.scribd.com/document/60162167/Regresyon-Analizi>.
- [44] Balcı H., Işıklı Esenler İ. ve Kurban M., “Regresyon Analizi Kullanılarak Kısa Dönem Yük Tahmini”, *ELECO '2012 Elektrik - Elektronik ve Bilgisayar Mühendisliği Sempozyumu*, Bursa, 796-801, (2012).
- [45] Şehirli K., “Regresyon Analizi [online]”, (24 Kasım 2018) <http://debis.deu.edu.tr/userweb//kemal.sehirli/dosyalar/regresyon1-2.pdf>, (2008).
- [46] Çiloğulları C., “Talep Tahmin Yöntemleri”, *Yıldız Teknik Üniversitesi Makine Fakültesi*, Endüstri Mühendisliği Bölümü, İstanbul, (2010).
- [47] Yakut E., Elmas B. ve Yavuz S., “Predicting Stock-Exchange Index Using Methods Of Neural Networks And Support Vector Machines”, *Suleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences*, 19(1), 139-157, (2014).

- [48] Hamzaelebi C., *Yapay Sinir Ađları: Tahmin Amalı Kullanımı Matlab ve Neurosolutions Uygulamalı*, Bursa: Ekin Kitabevi Yayınları, (2011).
- [49] Es H. A., “Yapay Sinir Ađları ile Türkiye'de Net Enerji Talep Tahmini”, Yüksek Lisans Tezi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü*, Endüstri Mühendisliği Anabilim Dalı, Ankara, (2013).
- [50] Fausett L., *Fundamentals of Neural Networks*, Upper Saddle River: Prentice Hall, (1993).
- [51] Vapnik V. N., *The Nature of Statistical Learning Theory*, Holmdel: Springer, (1995).
- [52] Ayhan S. ve Erdoğan Ş., “Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi”, *Eskişehir Osmangazi Üniversitesi İİBF Dergisi*, 9(1), 175-198, (2014).
- [53] J.-P. Vert, “Centre for Computational Biology CBIO at MINES ParisTech [online]”, (09 Mart 2019), <http://members.cbio.mines-paristech.fr/~jvert/talks/110401mines/mines.pdf>.
- [54] Chang C.-C. and Lin C.-J., “A Library for Support Vector Machines [online]”, (08 Aralık 2018), <http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf>, (2011).
- [55] Awad M. and Khanna R., *Efficient Learning Machines*, Apress, (2015).
- [56] McDuff D., “Gaussian Processes [online]”, (21 Şubat 2019). <http://courses.media.mit.edu/2010fall/mas622j/ProblemSets/slidesGP.pdf>, (2010).
- [57] Murphy K. P., *Machine Learning: A Probabilistic Perspective*, Cambridge, Massachusetts: The MIT Press, (2012).
- [58] Richardson R. R., Osborne M. A. and Howey D. A., “Gaussian process regression for forecasting battery state of health”, *Journal of Power Sources*, 357, 209-219, (2017).

- [59] E. Snelson, “Gatsby Computational Neuroscience Unit, UCL [online]”, (22 Şubat 2019), <http://mlg.eng.cam.ac.uk/tutorials/06/es.pdf>, (2006).
- [60] Skagestad R., “Electricity Demand Forecasting with Gaussian Process Regression”, M. Sc. Thesis, Norwegian University of Science and Technology, Trondheim, (2018).
- [61] Mori H., Sakatani Y., Fujino T. and Numa K., “Feature Extraction of One-Step-Ahead Daily Maximum Load with Regression Tree”, *Electrical Engineering in Japan*, 146 (2), 381-388, (2006).
- [62] Rokach L. and Maimon O., *Data Mining And Knowledge Discovery Handbook*, Tel-Aviv: Springer, (2005).
- [63] Quinlan J. R., “Induction of Decision Trees”, *Machine Learning*, Kluwer Academic Publishers, Boston, 81-106, (1986).
- [64] Gooroochurn K., “Decision Trees HL 8Only, Le Bocage International School IB Business & Management (F7) [online]”, (10 Temmuz 2019), <https://www.coursehero.com/file/24367578/53371221-DECISION-TREES-HLpdf>.
- [65] Tran V. T., Yang B.-S., Oh M.-S. and Tan A. C. C., “Machine condition prognosis based on regression trees and one-step-ahead prediction”, *Mechanical Systems and Signal Processing*, 22, 1179-1193, (2008).
- [66] Shalizi C., “Regression Trees”, Carnegie Mellon University, Pittsburgh, (2006).
- [67] Keskin M. V., “Ağaca Dayalı Yöntemlerde Bagging ve Boosting Arasında ne Fark var? [online]”, (27 Şubat 2019), <https://www.veribilimi.co/agaca-dayali-yontemlerde-bagging-ve-boosting-arasinda-ne-fark-var>, (2019).
- [68] “MSE, RMSE, MAE, MAPE ve Diğer Metrikler, [online]”, (12 Kasım 2018), <https://veribilimcisi.com/2017/07/14/mse-rmse-mae-mape-metrikleri-nedir>.
- [69] TEİAŞ, “10 Yıllık Talep Tahminleri Raporu”, Ankara, (2017).

[70] UEDAŞ, “2017 Faaliyet Raporu”, Bursa, (2018).

[71] Şahinler S., “En Küçük Kareler Yöntemi ile Doğrusal Regresyon Modeli Oluşturmanın Temel Prensipleri”, *MKÜ Ziraat Fakültesi Dergisi*, 5(1-2), 57-73, (2000).

EKLER

6. EKLER

EK A Bursa ili 2 günlük tüketim verileri

TARİH	SAAT	MW
2014-01-01	01:00	757.60
2014-01-01	02:00	629.10
2014-01-01	03:00	582.20
2014-01-01	04:00	533.70
2014-01-01	05:00	536.90
2014-01-01	06:00	504.02
2014-01-01	07:00	542.10
2014-01-01	08:00	602.90
2014-01-01	09:00	644.70
2014-01-01	10:00	677.70
2014-01-01	11:00	816.60
2014-01-01	12:00	905.70
2014-01-01	13:00	951.10
2014-01-01	14:00	1056.6
2014-01-01	15:00	1099.3
2014-01-01	16:00	872.99
2014-01-01	17:00	1047.4
2014-01-01	18:00	1022.2
2014-01-01	19:00	1218.4
2014-01-01	20:00	1148.6
2014-01-01	21:00	1065.6
2014-01-01	22:00	1039.8
2014-01-01	23:00	1067.9
2014-01-01	24:00	1078.5

TARİH	SAAT	MW
2014-01-02	01:00	775.30
2014-01-02	02:00	786.54
2014-01-02	03:00	819.77
2014-01-02	04:00	746.80
2014-01-02	05:00	846.90
2014-01-02	06:00	887.50
2014-01-02	07:00	851.70
2014-01-02	08:00	1067.8
2014-01-02	09:00	1248.5
2014-01-02	10:00	1303.4
2014-01-02	11:00	1350.6
2014-01-02	12:00	1276.9
2014-01-02	13:00	1248.3
2014-01-02	14:00	1248.5
2014-01-02	15:00	1276.2
2014-01-02	16:00	1228.9
2014-01-02	17:00	1489.1
2014-01-02	18:00	1381.4
2014-01-02	19:00	1262.0
2014-01-02	20:00	1186.9
2014-01-02	21:00	1214.9
2014-01-02	22:00	1137.7
2014-01-02	23:00	1142.8
2014-01-02	24:00	898.60

EK B Bursa ili 1 aylık sıcaklık, basınç, nem verisi örneği

Gün/Saat	Saatlik Sıcaklık (°C)																							
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	7.0	7.2	7.1	7.2	6.8	6.8	6.8	7.1	7.7	8.0	8.2	8.6	9.7	9.6	8.6	7.2	5.9	4.7	3.8	2.9	2.2	1.9	1.6	2.4
2	3.0	4.4	4.9	5.7	5.7	5.9	6.1	6.6	7.7	9.1	10.1	11.4	12.2	12.1	11.5	10.7	9.7	9.1	8.2	8.3	8.2	8.2	8.2	8.1
3	8.2	8.2	8.1	8.2	8.3	8.3	8.3	8.6	8.9	9.5	10.2	10.5	11.6	11.6	10.1	9.1	7.3	6.1	5.5	5.1	4.6	6.7	7.1	5.9
4	6.2	5.7	5.2	5.3	4.9	5.7	6.8	6.7	8.6	10.1	10.2	10.5	10.7	10.2	9.7	8.7	8.4	8.5	8.1	7.4	7.1	7.2	7.3	7.3
5	7.2	6.9	4.8	4.2	4.1	2.7	2.5	4.3	6.3	8.6	11.0	12.5	13.2	13.7	13.0	9.9	7.6	6.1	5.3	4.1	3.3	3.5	2.9	2.9
6	2.8	5.8	7.1	7.3	7.1	7.1	7.2	7.9	9.1	10.4	11.9	12.8	13.6	13.7	12.6	11.2	10.3	9.9	9.4	8.9	8.6	8.3	7.7	7.2
7	7.0	6.9	6.6	6.3	6.3	6.2	6.1	6.7	8.0	9.2	10.8	11.9	12.7	13.0	12.8	10.7	8.5	7.9	7.7	5.8	6.2	4.9	5.3	5.3
8	4.2	3.6	3.6	3.3	3.6	3.2	3.8	5.2	6.5	8.4	9.8	11.2	12.1	12.1	11.6	9.9	7.6	6.3	5.3	6.3	5.9	4.5	4.2	2.8
9	2.4	1.3	1.6	0.9	1.0	0.0	1.7	3.4	6.3	8.0	9.9	10.9	11.8	10.3	9.3	7.6	6.2	4.6	3.6	2.8	1.8	1.2	1.7	1.5
10	1.1	2.2	2.2	2.8	2.9	3.1	3.1	4.4	5.8	7.6	9.4	10.8	11.7	12.2	11.6	9.9	8.9	7.8	6.9	5.4	6.0	5.4	4.7	4.8
11	4.4	4.4	4.4	4.1	4.0	3.2	3.7	5.0	6.0	7.7	9.5	11.4	12.7	13.2	13.0	9.2	8.6	5.7	3.4	3.3	2.5	2.1	2.3	3.9
12	3.0	3.1	3.0	2.8	3.9	3.8	4.1	5.5	6.6	7.6	9.3	10.7	12.2	11.8	11.5	10.9	10.2	9.7	9.5	9.5	9.2	8.4	8.3	7.8
13	7.6	7.6	7.6	6.9	5.2	5.0	5.6	6.3	8.7	11.2	12.2	13.0	13.7	13.6	13.3	9.7	6.9	6.4	4.9	5.2	4.0	3.1	3.5	2.9
14	2.6	3.4	5.1	4.7	5.3	5.6	5.7	7.9	8.3	9.2	10.2	12.4	13.2	13.5	13.2	11.8	10.8	9.6	8.9	10.1	7.1	5.1	4.4	3.4
15	2.7	3.5	3.0	4.5	5.1	8.6	8.6	9.8	11.0	12.6	13.4	14.7	15.5	16.3	15.8	14.2	11.7	12.2	11.5	10.7	10.1	9.9	9.8	8.6
16	8.7	8.8	8.7	8.5	10.1	9.6	9.4	9.8	10.9	13.6	16.9	17.7	16.5	15.4	13.7	12.1	14.0	12.8	11.0	10.2	10.3	10.7	10.4	10.6
17	11.1	10.6	10.2	9.8	9.3	8.9	9.4	10.3	11.4	13.9	15.2	16.2	16.9	15.9	15.8	14.4	10.6	9.3	8.9	8.6	7.3	8.0	7.1	5.3
18	4.6	4.2	3.8	3.2	2.8	3.6	4.0	6.3	9.9	12.9	15.0	16.4	17.6	17.8	16.0	16.1	13.9	14.9	15.9	16.1	15.8	11.5	10.3	11.7
19	16.0	16.2	16.1	16.2	16.0	16.2	15.8	17.0	18.3	18.2	19.5	19.7	18.9	18.8	18.9	17.3	16.2	16.1	15.0	12.9	10.6	8.7	7.2	7.2
20	6.9	5.8	5.7	6.3	7.7	6.7	8.3	10.0	10.6	13.0	14.4	16.8	18.2	20.2	18.5	16.9	16.3	15.9	15.7	12.6	11.9	16.0	16.8	16.4
21	15.9	13.7	12.4	11.5	9.9	10.7	10.4	13.7	15.2	19.7	21.2	21.4	20.7	20.2	19.2	18.6	17.3	16.7	15.6	17.3	17.1	16.8	16.9	14.8
22	15.1	15.0	10.0	10.0	10.8	11.4	12.1	13.5	14.7	15.8	16.8	17.5	16.1	16.2	16.1	15.1	14.5	13.8	13.9	12.6	11.0	10.3	11.4	11.6
23	11.4	11.3	10.9	11.4	7.9	7.1	7.6	8.8	10.4	13.4	16.5	16.7	15.8	15.7	14.7	10.6	10.4	9.8	9.7	9.5	9.6	9.4	8.9	8.5
24	8.2	8.5	7.7	7.6	8.0	7.1	6.8	8.3	10.7	11.8	13.1	14.0	14.1	14.5	14.3	12.8	11.2	10.6	9.9	9.7	9.3	10.1	10.8	10.7
25	10.8	10.7	10.7	11.4	11.9	12.1	10.9	10.0	11.0	13.5	16.4	17.9	18.3	18.7	18.6	16.4	12.8	11.4	10.5	11.3	11.7	12.0	11.5	10.3
26	10.8	11.3	11.6	11.4	11.1	11.0	10.8	11.1	11.7	12.3	12.1	13.2	13.2	13.9	14.0	9.0	8.6	8.6	8.4	8.0	7.6	7.6	7.6	7.0
27	6.3	6.0	5.5	5.1	4.7	4.2	3.9	3.7	3.7	3.8	3.8	4.3	4.5	3.9	3.9	4.1	4.2	4.2	4.4	4.3	4.6	4.7	4.5	4.5
28	4.6	4.7	4.7	4.8	4.8	5.0	5.0	5.1	6.1	6.6	7.2	8.3	8.6	9.2	8.7	8.1	8.2	8.2	8.2	8.3	8.2	7.9	7.5	7.4
29	7.6	8.3	9.9	11.2	12.3	14.1	13.9	14.6	13.3	13.0	11.9	13.1	12.3	11.9	11.9	11.2	10.3	10.3	9.2	7.5	5.9	6.4	5.8	6.1
30	6.2	6.0	5.5	5.3	5.3	5.0	5.0	5.1	5.5	5.7	6.0	6.2	6.5	7.0	6.7	6.0	4.9	4.3	4.1	3.8	3.8	3.9	3.4	3.8
31	3.3	3.3	3.0	3.1	3.0	2.8	2.8	3.7	4.0	4.2	5.3	6.0	7.0	7.2	6.7	5.1	4.5	4.3	4.0	4.0	4.0	4.1	4.1	3.8

Gün/Saat	Saatlik Nispi Nem (%)																														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23							
1	80.0	77.0	77.0	75.0	78.0	78.0	77.0	77.0	71.0	68.0	88.0	67.0	64.0	64.0	69.0	77.0	82.0	87.0	89.0	91.0	93.0	94.0	95.0	95.0							
2	96.0	83.0	77.0	71.0	70.0	70.0	69.0	67.0	62.0	62.0	59.0	56.0	53.0	53.0	56.0	57.0	59.0	61.0	63.0	64.0	64.0	64.0	64.0	64.0							
3	63.0	65.0	65.0	65.0	64.0	67.0	67.0	68.0	65.0	65.0	63.0	62.0	61.0	60.0	76.0	80.0	84.0	90.0	90.0	94.0	94.0	92.0	92.0	92.0							
4	93.0	93.0	94.0	95.0	96.0	83.0	94.0	89.0	78.0	78.0	79.0	76.0	74.0	79.0	82.0	84.0	84.0	85.0	87.0	90.0	90.0	90.0	90.0	91.0							
5	89.0	88.0	91.0	94.0	95.0	97.0	96.0	83.0	76.0	76.0	63.0	56.0	52.0	51.0	54.0	71.0	81.0	86.0	88.0	91.0	93.0	93.0	94.0	95.0							
6	95.0	86.0	73.0	72.0	73.0	71.0	67.0	55.0	58.0	53.0	49.0	47.0	46.0	46.0	50.0	60.0	54.0	55.0	58.0	59.0	59.0	58.0	60.0	63.0							
7	84.0	64.0	65.0	66.0	67.0	67.0	65.0	61.0	57.0	53.0	49.0	46.0	46.0	46.0	46.0	52.0	59.0	65.0	57.0	64.0	63.0	72.0	65.0	66.0							
8	69.0	68.0	69.0	68.0	69.0	67.0	66.0	63.0	57.0	52.0	48.0	45.0	43.0	42.0	42.0	46.0	54.0	58.0	62.0	57.0	58.0	62.0	65.0	69.0							
9	72.0	79.0	76.0	80.0	83.0	82.0	79.0	72.0	64.0	57.0	50.0	48.0	45.0	61.0	71.0	74.0	81.0	85.0	91.0	93.0	94.0	94.0	83.0	79.0							
10	80.0	75.0	72.0	71.0	69.0	68.0	67.0	62.0	57.0	52.0	46.0	43.0	35.0	31.0	41.0	45.0	47.0	50.0	54.0	59.0	57.0	59.0	61.0	64.0							
11	62.0	65.0	64.0	65.0	68.0	68.0	64.0	60.0	54.0	54.0	51.0	47.0	45.0	48.0	51.0	68.0	61.0	66.0	76.0	80.0	81.0	86.0	86.0	77.0							
12	78.0	80.0	79.0	80.0	77.0	75.0	73.0	70.0	64.0	63.0	57.0	54.0	50.0	53.0	56.0	60.0	59.0	60.0	61.0	61.0	64.0	68.0	69.0	73.0							
13	74.0	73.0	73.0	76.0	83.0	86.0	84.0	87.0	72.0	66.0	67.0	65.0	62.0	63.0	63.0	72.0	87.0	84.0	91.0	90.0	93.0	95.0	96.0	95.0							
14	96.0	97.0	88.0	87.0	83.0	82.0	81.0	69.0	60.0	62.0	58.0	38.0	49.0	50.0	50.0	55.0	59.0	62.0	66.0	60.0	67.0	79.0	84.0	89.0							
15	91.0	88.0	89.0	89.0	82.0	64.0	62.0	60.0	56.0	50.0	48.0	46.0	44.0	43.0	44.0	49.0	55.0	60.0	60.0	63.0	61.0	65.0	63.0	67.0							
16	70.0	75.0	76.0	77.0	71.0	69.0	70.0	68.0	64.0	56.0	46.0	45.0	47.0	60.0	75.0	84.0	67.0	73.0	92.0	92.0	89.0	87.0	87.0	89.0							
17	85.0	86.0	88.0	92.0	94.0	95.0	96.0	93.0	90.0	75.0	68.0	62.0	61.0	61.0	61.0	65.0	80.0	89.0	87.0	92.0	89.0	91.0	92.0	95.0							
18	95.0	97.0	97.0	98.0	98.0	99.0	99.0	99.0	78.0	50.0	50.0	52.0	44.0	44.0	49.0	54.0	55.0	60.0	50.0	50.0	51.0	59.0	65.0	65.0							
19	49.0	49.0	50.0	50.0	52.0	52.0	53.0	53.0	49.0	50.0	47.0	47.0	46.0	46.0	47.0	51.0	54.0	53.0	55.0	59.0	69.0	76.0	83.0	87.0							
20	89.0	90.0	91.0	93.0	83.0	85.0	79.0	78.0	76.0	66.0	58.0	48.0	43.0	35.0	43.0	44.0	46.0	45.0	45.0	58.0	60.0	44.0	43.0	43.0							
21	45.0	48.0	54.0	55.0	64.0	63.0	60.0	53.0	52.0	31.0	29.0	29.0	28.0	30.0	33.0	33.0	37.0	42.0	45.0	37.0	38.0	38.0	41.0	55.0							
22	54.0	55.0	94.0	96.0	88.0	76.0	70.0	63.0	56.0	52.0	46.0	44.0	46.0	45.0	43.0	49.0	51.0	52.0	54.0	58.0	66.0	72.0	74.0	59.0							
23	63.0	64.0	63.0	57.0	80.0	80.0	81.0	77.0	72.0	61.0	44.0	41.0	47.0	49.0	53.0	84.0	89.0	94.0	96.0	96.0	96.0	93.0	94.0	93.0							
24	94.0	95.0	95.0	95.0	94.0	90.0	86.0	86.0	88.0	75.0	68.0	63.0	65.0	68.0	55.0	65.0	75.0	79.0	83.0	83.0	87.0	83.0	76.0	79.0							
25	68.0	66.0	67.0	60.0	58.0	57.0	65.0	87.0	78.0	64.0	60.0	37.0	38.0	40.0	40.0	47.0	63.0	62.0	66.0	72.0	69.0	63.0	63.0	71.0							
26	66.0	72.0	63.0	62.0	62.0	62.0	63.0	64.0	62.0	62.0	65.0	62.0	64.0	62.0	60.0	87.0	89.0	90.0	90.0	91.0	92.0	92.0	90.0	91.0							
27	94.0	95.0	95.0	95.0	96.0	97.0	97.0	97.0	97.0	97.0	97.0	91.0	89.0	92.0	92.0	92.0	92.0	92.0	89.0	93.0	93.0	89.0	93.0	94.0							
28	95.0	94.0	93.0	90.0	90.0	86.0	91.0	93.0	95.0	94.0	90.0	83.0	83.0	80.0	76.0	77.0	78.0	79.0	79.0	80.0	79.0	83.0	86.0	87.0							
29	84.0	80.0	67.0	69.0	60.0	49.0	50.0	46.0	51.0	59.0	69.0	65.0	66.0	67.0	66.0	70.0	74.0	77.0	82.0	87.0	90.0	90.0	90.0	90.0							
30	88.0	83.0	83.0	80.0	83.0	82.0	84.0	82.0	81.0	81.0	81.0	78.0	78.0	73.0	75.0	76.0	79.0	80.0	80.0	81.0	81.0	83.0	83.0	84.0							
31	84.0	83.0	83.0	82.0	81.0	80.0	77.0	76.0	74.0	74.0	71.0	69.0	66.0	63.0	65.0	72.0	74.0	74.0	75.0	76.0	77.0	77.0	77.0	75.0							

EK C Microsoft® Excel® den MATLAB® e Veri Aktarım Kodu

```
% Insert Datex

%% Import the data, extracting spreadsheet dates in Excel serial
date format
[~, ~, raw, dates] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-
2019_rev_with_meteodata_v2.xlsx', 'BURSA', 'A2:A45049', '', @convertSpre
adsheetExcelDates);
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x), raw)) =
{' '};
dates = dates(:,1);

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x), raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x), dates); % Find non-
numeric cells
dates(R) = {NaN}; % Replace non-numeric Excel dates with NaN

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.Date = datetime([dates(:,1)].',
'ConvertFrom', 'Excel');

% For code requiring serial dates (datenum) instead of datetime,
uncomment
% the following line(s) below to return the imported dates as
datenum(s).

%
bursa20142019revwithmeteodav2.Date=datenum(bursa20142019revwithmet
eodav2.Date);

%% Clear temporary variables
clearvars Date raw dates R;

%Write Datex
data.Datex=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Insert SYSLoadx

%% Import the data
```



```

[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx', 'BURSA', 'F2:F45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x), raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x), raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
SYSLoad = reshape([raw{:}], size(raw));

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.SYSLoad = SYSLoad(:,1);

%% Clear temporary variables
clearvars SYSLoad raw R;

%Write SYSLoadx
data.SYSLoadx=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Insert Dateh

%% Import the data, extracting spreadsheet dates in Excel serial
date format
[~, ~, raw, dates] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-
2019_rev_with_meteodata_v2.xlsx', 'BURSA', 'E2:E45049', '', @convertSpre
adsheetExcelDates);
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x), raw)) =
{' '};
dates = dates(:,1);

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x), raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x), dates); % Find non-
numeric cells
dates(R) = {NaN}; % Replace non-numeric Excel dates with NaN

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.Dateh = datetime([dates{:},1] .',
'ConvertFrom', 'Excel');

% For code requiring serial dates (datenum) instead of datetime,
uncomment

```

```

% the following line(s) below to return the imported dates as
datenum(s).

%
bursa20142019revwithmeteodav2.Dateh=datenum(bursa20142019revwithme
teodav2.Dateh);

%% Clear temporary variables
clearvars Dateh raw dates R

%Write Dateh
data.Dateh=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Inset Hourx (Numeric Hour data)

%% Import the data
[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','D2:D45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
Hour = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.Hour = Hour(:,1);

%% Clear temporary variables
clearvars Hour raw R;

%Write Hourx
data.Hourx=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Insert Holidays between 2014 and 2019

%% Import the data, extracting spreadsheet dates in Excel serial
date format
[~, ~, raw, dates] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-
2019_rev_with_meteodata_v2.xlsx','Holidays','A2:A83','',@convertSpre
adsheetExcelDates);

```

```

raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};
dates = dates(:,1);

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),dates); % Find non-
numeric cells
dates(R) = {NaN}; % Replace non-numeric Excel dates with NaN

%% Create table
bursa20142019revwithmeteodativ2S1 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodativ2S1.Holidays =
datetime([dates{: ,1}].', 'ConvertFrom', 'Excel');

% For code requiring serial dates (datenum) instead of datetime,
uncomment
% the following line(s) below to return the imported dates as
datenum(s).

%
bursa20142019revwithmeteodativ2S4.Holidays=datenum(bursa20142019revw
ithmeteodativ2S4.Holidays);

%% Clear temporary variables
clearvars Holidays raw dates R;

%Write Holidays
Holidays=bursa20142019revwithmeteodativ2S1;
clear bursa20142019revwithmeteodativ2S1;

%meteodata
%% Import the temperature data
[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','G2:G45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
Temp = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodativ2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodativ2.Temp = Temp(:,1);

%% Clear temporary variables

```

```

clearvars Temp raw R;

data.Temp=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Import Humidity data
[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','H2:H45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
Humidity = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.Humidity = Humidity(:,1);

%% Clear temporary variables
clearvars Humidity raw R;

data.Hum=bursa20142019revwithmeteodav2;
clear bursa20142019revwithmeteodav2;

%% Import pressure data
[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','I2:I45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
Pressure = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.Pressure = Pressure(:,1);

%% Clear temporary variables
clearvars Pressure raw R;

```

```

data.Pres=bursa20142019revwithmeteodtav2;
clear bursa20142019revwithmeteodtav2;

%dmis
%% Import the data
[~, ~, raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','J2:J45049');
raw(cellfun(@(x) isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
original = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodtav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodtav2.original = original(:,1);

%% Clear temporary variables
clearvars original raw R;

dmis=bursa20142019revwithmeteodtav2;

clear bursa20142019revwithmeteodtav2;

dmis=table2array(dmis);
indices = find(dmis(:,1)==0);
dmis(indices,:) = [NaN];

%% -creating struct and converting data to array

Holidays=table2array(Holidays);
data.SYSLoad=table2array(data.SYSLoadx);
data.Date=table2array(data.Dateh);
data.Temp=table2array(data.Temp);
data.Hum=table2array(data.Hum);
data.Pres=table2array(data.Pres);
%data.Hour=table2array(data.Hour);
data.Hourx=table2array(data.Hourx);
%data.Datex=table2array(data.Datex);

%% edit errored dates

data.Date(3) = '01-Jan-2014 02:00:00';

```

```
data.Date(6) = '01-Jan-2014 05:00:00';  
data.Date(9) = '01-Jan-2014 08:00:00';  
data.Date(12) = '01-Jan-2014 11:00:00';  
data.Date(15) = '01-Jan-2014 14:00:00';  
data.Date(18) = '01-Jan-2014 17:00:00';  
data.Date(21) = '01-Jan-2014 20:00:00';  
data.Date(27) = '02-Jan-2014 02:00:00';  
data.Date(30) = '02-Jan-2014 05:00:00';  
data.Date(33) = '02-Jan-2014 08:00:00';  
data.Date(36) = '02-Jan-2014 11:00:00';  
data.Date(39) = '02-Jan-2014 14:00:00';  
data.Date(42) = '02-Jan-2014 17:00:00';  
data.Date(45) = '02-Jan-2014 20:00:00';  
data.Date(48) = '02-Jan-2014 23:00:00';
```

EK D Tahmin Parametrelerinin Oluşturulmasını Sağlayan MATLAB® Kodu

```
dates=data.Date;

% Numeric Representation of dates
data.Datex=yyyymmdd(dates);

%% Create Predictors-----
[m,n]=size(data.SYSLoad);
clear n;
% Short term forecasting inputs
% Lagged load inputs
%prevDaySameHourLoad = [NaN(24,1); data.SYSLoad(1:end-24)];
%prevWeekSameHourLoad = [NaN(168,1); data.SYSLoad(1:end-168)];%7*24
%prev4WeeksSameHourLoad = [NaN(672,1); data.SYSLoad(1:end-
672)];%168*4
%prev8WeeksSameHourLoad = [NaN(1344,1); data.SYSLoad(1:end-
1344)];%168*8
%prev24HrAveLoad = filter(ones(1,24)/24, 1, data.SYSLoad);
%prev24HrAveLoad = filter(ones(1,24)/24, 1, [NaN(24,1);
data.SYSLoad(1:end-24)]);

%% prevDaySameHourLoad

for k=1:23
    prevDaySameHourLoad(k)=NaN;
end
for k=24:m
    prevDaySameHourLoad(k)= data.SYSLoad(k-23);
end

prevDaySameHourLoad=transpose(prevDaySameHourLoad);

%% prevWeekSameHourLoad & prev4WeeksSameHourLoad &
prev8WeeksSameHourLoad

for k=1:167
    prevWeekSameHourLoad(k)=NaN;
end
for k=168:m
    prevWeekSameHourLoad(k)= data.SYSLoad(k-167);
end
clear k;
prevWeekSameHourLoad=transpose(prevWeekSameHourLoad);
%%%4 weeks
for k=1:671
```

```

        prev4WeeksSameHourLoad(k)=NaN;
end
for k=672:m
    prev4WeeksSameHourLoad(k)= data.SYSLoad(k-671);
end
clear k;
prev4WeeksSameHourLoad=transpose (prev4WeeksSameHourLoad);

%%8 weeks

for k=1:1343
    prev8WeeksSameHourLoad(k)=NaN;
end
for k=1344:m
    prev8WeeksSameHourLoad(k)= data.SYSLoad(k-1343);
end
clear k;
prev8WeeksSameHourLoad=transpose (prev8WeeksSameHourLoad);

%% prev24HrAveLoad test for other arguments 6 hour 3 hour load

for k=1:23
    prev24HrAveLoad(k)=NaN;%mean (data.SYSLoad(1:24));
end
for k=24:m
    prev24HrAveLoad(k)= mean (data.SYSLoad((k-23):k));
end
clear k;
prev24HrAveLoad=transpose (prev24HrAveLoad);

%% 12 Hour avarage Load

for k=1:11
    prev12HrAveLoad(k)= NaN;%mean (data.SYSLoad(1:12));
end
for k=12:m
    prev12HrAveLoad(k)= mean (data.SYSLoad((k-11):k));
end
clear k;
prev12HrAveLoad=transpose (prev12HrAveLoad);

%% 6 hours avarage Load

for k=1:5
    prev6HrAveLoad(k)=NaN;%mean (data.SYSLoad(1:6));
end
for k=6:m
    prev6HrAveLoad(k)= mean (data.SYSLoad((k-5):k));
end
clear k;
clear m;
prev6HrAveLoad=transpose (prev6HrAveLoad);

%%
[m,n]=size (data.Temp);
clear n;
%% prevDaySameHourTemp

```



```

for k=1:23
    prevDaySameHourTemp(k)=NaN;
end
for k=24:m
    prevDaySameHourTemp(k)= data.Temp(k-23);
end

prevDaySameHourTemp=transpose(prevDaySameHourTemp);

%% prevWeekSameHourTemp & prev4WeeksSameHourTemp &
prev8WeeksSameHourTemp

for k=1:167
    prevWeekSameHourTemp(k)=NaN;
end
for k=168:m
    prevWeekSameHourTemp(k)= data.Temp(k-167);
end
clear k;
prevWeekSameHourTemp=transpose(prevWeekSameHourTemp);
%%%4 weeks
for k=1:671
    prev4WeeksSameHourTemp(k)=NaN;
end
for k=672:m
    prev4WeeksSameHourTemp(k)= data.Temp(k-671);
end
clear k;
prev4WeeksSameHourTemp=transpose(prev4WeeksSameHourTemp);

%%%8 weeks

for k=1:1343
    prev8WeeksSameHourTemp(k)=NaN;
end
for k=1344:m
    prev8WeeksSameHourTemp(k)= data.Temp(k-1343);
end
clear k;
clear m;
prev8WeeksSameHourTemp=transpose(prev8WeeksSameHourTemp);

%% Holiday calculation

%%% Date tablosunun boyutunda yeni bir 0 Hol matrisi olu?turuluyor.
%%% Buradaki 0 isHoliday matrisindeki al??ma g?nlerini ifade etmi?
olacak

holix=datenum(Holidays);
[s1,s2] = size(holix);%tatil g?n say?s?n?n belirlenmesi s1
clear s2;
for k=1:s1
    Hol(k)=0;

```

```

end
Hol=transpose(Hol);
clear k;

%%Holiday içerisindeki nümerik veri ile veri seti içerisindekilerin
fark?
%%alınarak sradan tatil olan günlerin belirlenmesi sağlanıyor. Bu
sayede
%%her bir tatil gün nümerik olarak asıl veri setinde 1 olarak
%işaretlenmekte. Benzer çalınma hafta içi/haftasonu günler için de
%yapılabilir. Bu durumda aritmetik olarak tekrarlandığı için daha
kolay
%%bir şekilde uygulanabilir.

k=holix(1);%veri setindeki ilk tatil gün

for m=1:s1
    f=(holix(m)-k);%tatil günlerindeki her bir datenum değerinin
bunması ve ardından k datenumdan kaçta kaç (bir sonraki tatil datenum
farkını bul)
    f1=f*24;%Bir sonraki tatil gününün hol matrisi içerisindeki
satırını bul
    for t=1:24
        Hol(f1+t)=1;%bulunan satıra 24 saatlik 1 verisini yaz
    end
end

%define a new variable
%biten tatil günden sonra veri setinin sonuna kadar kalan değerlere
0 yaz

[s3,s4] = size(data.Date);%veri setindeki satır sayısı bul
clear s4;
for m=(f1+1+24):(s3)
    Hol(m)=0;
end

isHolidays=logical (Hol);%isholiday deikenine mantıksal olarak ata

%clear variables and assign new name
clear s1;
clear t;
clear m;
clear k;
clear s3;
clear f;
clear f1;
clear Hol;
clear holix;
data.Holidays=Holidays;
clear Holidays;
%% Date predictors
dayOfWeek = weekday(dates);
dayOfWeek = dayOfWeek-1;
dayOfWeek(dayOfWeek==0)=7;

```

```

%% Weekends
isWeekends = ismember(dayOfWeek,[6 7]);

%% Non-business days
isWorkingDay = ~isHolidays & ~ismember(dayOfWeek,[6 7]);

%missing datas

%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook:
C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast\V4\bursa
2014-2019_rev_with_meteodata_v2.xlsx
%   Worksheet: BURSA
%
% To extend the code for use with different selected data or a
different
% spreadsheet, generate a function instead of a script.

% Auto-generated by MATLAB on 2019/03/25 14:51:56

%% Import the data
[~,~,raw] =
xlsread('C:\Users\erenk\OneDrive\Documents\MATLAB\2019_load_forecast
\V4\bursa2014-2019_rev_with_meteodata_v2.xlsx','BURSA','J2:J45049');
raw(cellfun(@(x) ~isempty(x) && isnumeric(x) && isnan(x),raw)) =
{' '};

%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-
numeric cells
raw(R) = {NaN}; % Replace non-numeric cells

%% Create output variable
original = reshape([raw{:}],size(raw));

%% Create table
bursa20142019revwithmeteodav2 = table;

%% Allocate imported array to column variable names
bursa20142019revwithmeteodav2.original = original(:,1);

%% Clear temporary variables
clearvars original raw R;

dmis=bursa20142019revwithmeteodav2;

clear bursa20142019revwithmeteodav2;

dmis=table2array(dmis);
indices = find(dmis(:,1)==0);
dmis(indices,:) = [NaN];

```

```
%% Predict matrix and labels
```

```
X = [data.Datex data.Hourx dayOfWeek isWorkingDay isHolidays  
prev8WeeksSameHourLoad prev4WeeksSameHourLoad prevWeekSameHourLoad  
prevDaySameHourLoad prev24HrAveLoad prev12HrAveLoad prev6HrAveLoad  
data.Temp data.Hum data.Pres prev8WeeksSameHourTemp  
prev4WeeksSameHourTemp prevWeekSameHourTemp prevDaySameHourTemp  
dmis];  
XVariableNames = {'Date', 'Hour', 'Weekday', 'isWorkingDay'  
'Holidays', 'prev8WeeksSameHourLoad', 'prev4WeeksSameHourLoad',  
'PrevWeekSameHourLoad', 'prevDaySameHourLoad', 'prev24HrAveLoad',  
'prev12HrAveLoad', 'prev6HrAveLoad', 'Temperature', 'Humidity',  
'Pressure', 'prev8WeeksSameHourTemp', 'prev4WeeksSameHourTemp',  
'prevWeekSameHourTemp', 'prevDaySameHourTemp', 'dmis'};
```

```
Xs = [data.Datex data.Hourx dayOfWeek isWorkingDay isHolidays  
prev8WeeksSameHourLoad prev4WeeksSameHourLoad prevWeekSameHourLoad  
prevDaySameHourLoad prev24HrAveLoad prev12HrAveLoad prev6HrAveLoad];  
XsVariableNames = {'Date', 'Hour', 'Weekday', 'isWorkingDay'  
'Holidays', 'prev8WeeksSameHourLoad', 'prev4WeeksSameHourLoad',  
'PrevWeekSameHourLoad', 'prevDaySameHourLoad', 'prev24HrAveLoad',  
'prev12HrAveLoad', 'prev6HrAveLoad' };
```

```
X1 = [data.Datex data.Hourx dayOfWeek isWorkingDay isHolidays  
prev8WeeksSameHourLoad prev4WeeksSameHourLoad prevWeekSameHourLoad  
prevDaySameHourLoad prev24HrAveLoad prev12HrAveLoad prev6HrAveLoad  
data.Temp data.Hum data.Pres];  
X1VariableNames = {'Date', 'Hour', 'Weekday', 'isWorkingDay'  
'Holidays', 'prev8WeeksSameHourLoad', 'prev4WeeksSameHourLoad',  
'PrevWeekSameHourLoad', 'prevDaySameHourLoad', 'prev24HrAveLoad',  
'prev12HrAveLoad', 'prev6HrAveLoad', 'Temperature', 'Humidity',  
'Pressure'};
```

```
Xs1 = [data.Datex data.Hourx dayOfWeek isWorkingDay isHolidays  
prevWeekSameHourLoad prevDaySameHourLoad prev24HrAveLoad];  
Xs1VariableNames = {'Date', 'Hour', 'Weekday', 'isWorkingDay'  
'Holidays', 'PrevWeekSameHourLoad', 'prevDaySameHourLoad',  
'prev24HrAveLoad'};
```

```
Xs2 = [data.Datex data.Hourx dayOfWeek prevWeekSameHourLoad  
prevDaySameHourLoad prev24HrAveLoad];  
Xs2VariableNames = {'Date', 'Hour', 'Weekday',  
'PrevWeekSameHourLoad', 'prevDaySameHourLoad', 'prev24HrAveLoad'};
```

EK E Linear Regresyon MATLAB® Kodu

```
X=X(:,1:19);

y=data.SYSLoad;
[b,bint,r]=regress(y,X);

yPredLR = X *b;

plot(data.Date, yPredLR, 'DisplayName','yPredLR', 'YDataSource',
'yPredX', 'color',[0.75, 0, 0.75]);hold on;
plot(data.Date, y, 'DisplayName','SYSLoad', 'YDataSource',
'SYSLoad', 'color','b');
title('Linear Regression Predicted and Compared Load Demands 2014-
2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
grid on;

mapeLR = nanmean(abs(r./y*100)) %Mean absolute percentage error

%X1
y=data.SYSLoad;
[b,bint,r]=regress(y,X1);

yPredLR = X1 *b;

plot(data.Date, yPredLR1, 'DisplayName','yPredLR1', 'YDataSource',
'yPredX', 'color',[0.75, 0, 0.75]);hold on;
plot(data.Date, y, 'DisplayName','SYSLoad', 'YDataSource',
'SYSLoad', 'color','b');
title('Linear Regression Predicted and Compared Load Demands 2014-
2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
grid on;

mapeLR1 = nanmean(abs(r./y*100)) %Mean absolute percentage error

hold on;
[b,bint,r]=regress(y,Xs);

yPredLRs = Xs *b;
```

```

plot(data.Date, yPredLRs, 'DisplayName','yPredLRs', 'YDataSource',
'yPredXs', 'color','[0, 0.10, 0]');
%plot(y);

mapeLRs = nanmean(abs(r./y*100)) %Mean absolute percentage error
*7.4196 çal??ma günleri ve tatil bilgisi ile

[b,bint,r]=regress(y,Xs1);

yPredLRs1 = Xs1 *b;
plot(data.Date, yPredLRs1, 'DisplayName','yPredLRs1', 'YDataSource',
'yPredXs1', 'color','[0, 0.25, 0]');
hold on;
mapeLRs1 = nanmean(abs(r./y*100))

[b,bint,r]=regress(y,Xs2);

yPredLRs2 = Xs2 *b;

plot(data.Date,yPredLRs2, 'DisplayName','yPredLR2', 'YDataSource',
'yPredXs2', 'color', '[0.75, 0.75, 0]');

mapeLRs2 = nanmean(abs(r./y*100)) %Mean absolute percentage error *e
7.8191 çal??ma günleri ve tatil bilgisi ile, tatil bilgisi holiday'
den güncellendi. 4.9388 e dü?tü

%regTrainer

data.DateN=datenum(data.Date);
y=data.SYSLoad;
% Create training set
RegtrainInd = data.DateN < datenum('02-21-2018');
RegtrainX = X(RegtrainInd,:);
RegtrainY = y(RegtrainInd);

%eksik veriler ?kart?ld? Nan
R=rmmissing([RegtrainX RegtrainY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
RegtrainX = R(:,1:19);
RegtrainY = R(:,end);
clear R;

% Create test set and save for later
RegtestInd = data.DateN >= datenum('02-21-2018');
RegtestX = X(RegtestInd,:);
RegtestY = data.SYSLoad(RegtestInd);

R=rmmissing([RegtestX RegtestY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
RegtestX = R(:,1:19);
RegtestY = R(:,end);
clear R;

```

```

t = string(data.Date);
TestDate = [t dms];
RegtestDates = TestDate(RegtestInd,:);
clear TestDate;
clear t;
R=rmmising([RegtestDates]);%eksik verilerin tablodan ç?kart?lmas?
NAN veriler bu sayede analizden ç?kart?ld?.
RegtestDates = R(:,1);
RegtestDates = datetime(RegtestDates,'InputFormat','dd-MMM-yyyy
HH:mm:ss');
clear R;

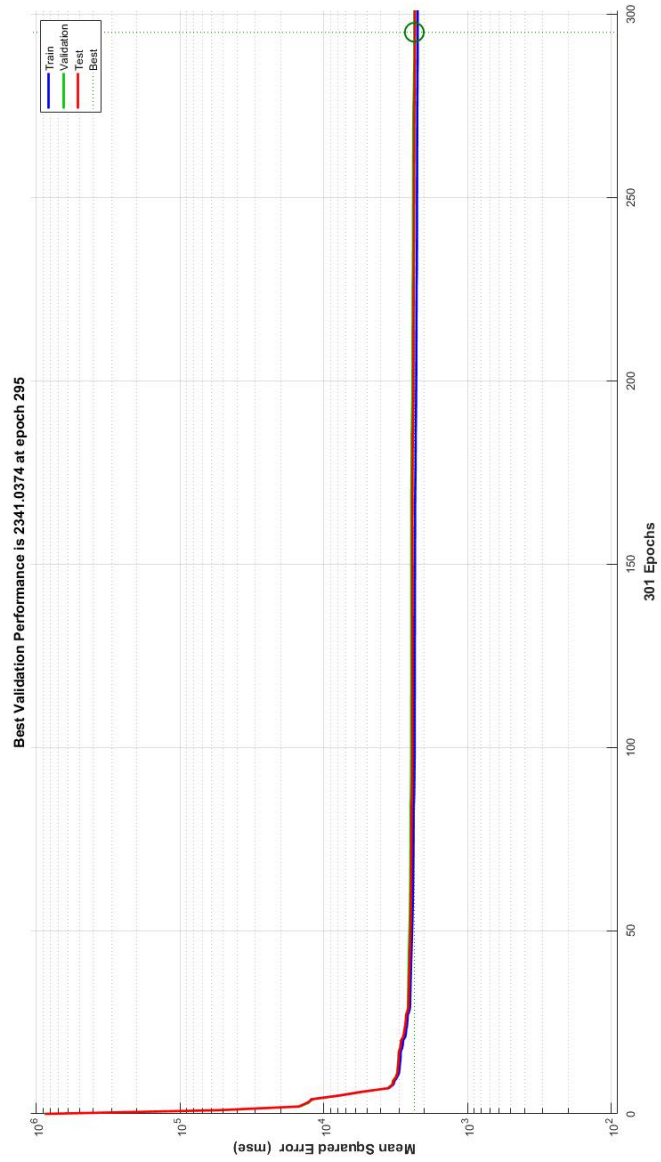
[b,bint,r]=regress(RegtrainY,RegtrainX);

yPredLRtrained = RegtestX *b;
mapeyPredLRtrained = nanmean(abs(r./RegtrainY*100))%4.8065    4.7812

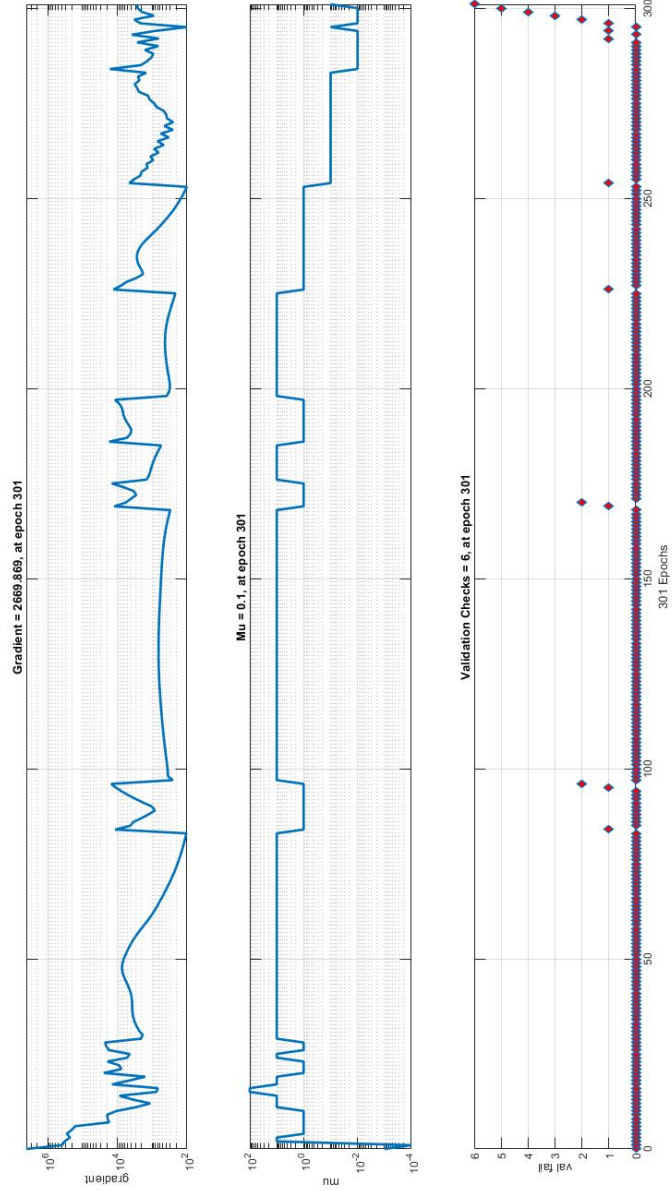
figure(1);
plot(RegtestDates, yPredLRtrained, 'DisplayName','yPredLRtrained',
'YDataSource','yPredLRtrained', 'color',[0.75, 0, 0.75]);hold on;
plot(RegtestDates, RTtestY,
'DisplayName','SYSLoadTest','YDataSource',
'SYSLoadTest','color','b');grid on;hold on;
title('Linear Regression Predictions 2018-2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
grid on;

```

EK F ANN Eğitim Performansı Grafiği



EKG ANN Eğitim Durumu



EK H Regresyon Ağacı Kodu

```
%train and test variables

data.DateN=datenum(data.Date);

% Create training set
RTtrainInd = data.DateN < datenum('02-21-2018');
RTtrainX = X(RTtrainInd,:);
RTtrainY = y(RTtrainInd);

%eksik veriler ç?kart?ld? Nan
R=rmmissing([RTtrainX RTtrainY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
RTtrainX = R(:,1:19);
RTtrainY = R(:,end);
clear R;

% Create test set and save for later
RTtestInd = data.DateN >= datenum('02-21-2018');
RTtestX = X(RTtestInd,:);
RTtestY = data.SYSLoad(RTtestInd);

R=rmmissing([RTtestX RTtestY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
RTtestX = R(:,1:19);
RTtestY = R(:,end);
clear R;

t = string(data.Date);
TestDate = [t dms];
RTtestDates = TestDate(RTtestInd,:);
clear TestDate;
clear t;
R=rmmissing([RTtestDates]);%eksik verilerin tablodan ç?kart?lmas?
NAN veriler bu sayede analizden ç?kart?ld?.
RTtestDates = R(:,1);
RTtestDates = datetime(RTtestDates, 'InputFormat', 'dd-MMM-yyyy
HH:mm:ss');
clear R;

%tree parameter optimization
MdlTreeOpt = fitrtree(RTtrainX, RTtrainY, 'categoricalpredictors', 4,
'OptimizeHyperparameters', 'auto', ...
'HyperparameterOptimizationOptions', struct('AcquisitionFunctionName'
, ...
'expected-improvement-plus'))
```

```
%• Resmi tatillerdeki 24 saatlik yük Hafta içlerindeki 24 saatlik
yük Hafta sonlar? 24 saatlik yük Çal???lmayan günlerdeki 24
saatlik yük
yPredTreeOpt=predict (MdlTreeOpt,RTtestX);
```

```
r=RTtestY-yPredTreeOpt;
mape = nanmean(abs(r./RTtestY*100))%%% 2.8583 mape 2.8901
```

```
figure(2)
plot(RTtestDates, RTtestY,
'DisplayName','SYSLoadTest','YDataSource',
'SYSLoadTest','color','b');grid on;hold on;
plot(RTtestDates, yPredTreeOpt,
'DisplayName','yPredTreeOpt','YDataSource',
'yPredTreeOpt','color',[0.8500, 0.3250, 0.0980]);grid on;hold on;
title('Regression Tree Predictions 2018-2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
```

```
clear RTtrainInd
clear RTtrainX
clear RTtrainY
clear RTtestInd
clear RTtestY
clear RTtestY
clear RTtestDates
```

EK I Bagged Regresyon Ağacı MATLAB ® Kodu

```
data.DateN=datenum(data.Date);

% Create training set
TBtrainInd = data.DateN < datenum('02-21-2018');
TBtrainX = X(TBtrainInd,:);
TBtrainY = data.SYSLoad(TBtrainInd);

R=rmmissing([TBtrainX TBtrainY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
TBtrainX = R(:,1:19);
TBtrainY = R(:,end);
clear R;

% Create test set and save for later
TBtestInd = data.DateN >= datenum('02-21-2018');
TBtestX = X(TBtestInd,:);
TBtestY = data.SYSLoad(TBtestInd);

R=rmmissing([TBtestX TBtestY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
TBtestX = R(:,1:19);
TBtestY = R(:,end);
clear R;

t = string(data.Date);
TestDate = [t dms];
TBtestDates = TestDate(TBtestInd,:);
clear TestDate;
clear t;
R=rmmissing([TBtestDates]);%eksik verilerin tablodan ç?kart?lmas?
NAN veriler bu sayede analizden ç?kart?ld?.
TBtestDates = R(:,1);
TBtestDates = datetime(TBtestDates, 'InputFormat', 'dd-MMM-yyyy
HH:mm:ss');
clear R;

MdlBTree = TreeBagger(20, TBtrainX, TBtrainY, 'method',
'regression', 'minleaf', 5)
%view(MdlBTree.Trees{1}, 'Mode','graph');
yPredBTree = predict(MdlBTree, TBtestX);

r=TBtestY-yPredBTree;
mapeBTree = nanmean(abs(r./TBtestY*100))%mean leaf 5 2.5767 meant
leaf 17=2.6309      18  2.5792  5 2.4298  Minimum number of
```

observations per tree leaf. Default is 1 for classification and 5 for regression.

```
%simpleTree = Prune(yPredBTree.Trees{1}, 500);
%simpleTree = Prune(simpleTree, simpleTree.prunelist(1)-10);

figure(3)
plot(TBtestDates, TBtestY,
'DisplayName', 'SYSLoadTest', 'YDataSource',
'SYSLoadTest', 'color', 'b');grid on;hold on;
plot(TBtestDates, yPredBTree,
'DisplayName', 'yPredBTree', 'YDataSource',
'yPredBTree', 'color', '[0.4660, 0.6740, 0.1880]');grid on;hold on;
title('Bagged Regression Tree Predictions 2018-2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
```

EK J SVR MATLAB ® Kodu

```
%SVM Regressionv2

data.DateN=datenum(data.Date);
y=data.SYSLoad;
% Create training set
SVRtrainInd = data.DateN < datenum('02-21-2018');
SVRtrainX = X(SVRtrainInd,:);
SVRtrainY = y(SVRtrainInd);

R=rmmissing([SVRtrainX SVRtrainY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
SVRtrainX = R(:,1:19);
SVRtrainY = R(:,end);
clear R;
% Create test set and save for later
SVRtestInd = data.DateN >= datenum('02-21-2018');
SVRtestX = X(SVRtestInd,:);
SVRtestY = y(SVRtestInd);

R=rmmissing([SVRtestX SVRtestY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
SVRtestX = R(:,1:19);
SVRtestY = R(:,end);
SVRtestDates = data.Date(SVRtestInd);

t = string(data.Date);
TestDate = [t dms];
SVRtestDates = TestDate(SVRtestInd,:);
clear TestDate;
clear t;
R=rmmissing([SVRtestDates]);%eksik verilerin tablodan ç?kart?lmas?
NAN veriler bu sayede analizden ç?kart?ld?.
SVRtestDates = R(:,1);
SVRtestDates = datetime(SVRtestDates,'InputFormat','dd-MMM-yyyy
HH:mm:ss');
clear R;

MdlSVRO = fitrsvm(SVRtrainX, SVRtrainY, 'KernelFunction','gaussian',
'Standardize',true,
'CategoricalPredictors',4,'BoxConstraint',444.1662,'KernelScale',116
.9698);
yPredMdlSVRO = predict(MdlSVRO, SVRtestX);

r=SVRtestY-yPredMdlSVRO;
mapeSVRO = nanmean(abs(r./SVRtestY*100))%3.0308-4.4241

figure(4);
```

```

plot(SVRtestDates, SVRtestY,
'DisplayName', 'SYSLoadTest', 'YDataSource',
'SYSLoadTest', 'color', 'b');grid on;hold on;
plot(SVRtestDates, yPredMdlSVRO,
'DisplayName', 'yPredMdlSVRO', 'YDataSource',
'yPredMdlSVRO', 'color', '[0.5, 0.8, 0.4250]');grid on;hold on;
title('Support Vector Regression Predictions 2018-2019');
xlabel('Dates And Times');
ylabel('MW');
legend;

%optimize

rng default
MdlSVROpt = fitrsvm(SVRtrainX,
SVRtrainY, 'OptimizeHyperparameters', 'auto', ...

'HyperparameterOptimizationOptions', struct('AcquisitionFunctionName'
, ...
'expected-improvement-plus'))

yPredMdlSVROpt = predict(MdlSVROpt, SVRtestX);

r=SVRtestY-yPredMdlSVROpt;
mapeSVROpt = nanmean(abs(r./SVRtestY*100))%%4.5394===4.4179

```

EK K GPR MATLAB ® Kodu

```
data.DateN=datenum(data.Date);

% Create training set
GPRtrainInd = data.DateN < datenum('02-21-2018');
GPRtrainX = X(GPRtrainInd,:);
GPRtrainY = data.SYSLoad(GPRtrainInd);

R=rmmissing([GPRtrainX GPRtrainY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
GPRtrainX = R(:,1:19);
GPRtrainY = R(:,end);
clear R;

% Create test set and save for later
GPRtestInd = data.DateN >= datenum('02-21-2018');
GPRtestX = X(GPRtestInd,:);
GPRtestY = data.SYSLoad(GPRtestInd);

R=rmmissing([GPRtestX GPRtestY]);%eksik verilerin tablodan
ç?kart?lmas? NAN veriler bu sayede analizden ç?kart?ld?.
GPRtestX = R(:,1:19);
GPRtestY = R(:,end);
clear R;

t = string(data.Date);
TestDate = [t dms];
GPRtestDates = TestDate(GPRtestInd,:);
clear TestDate;
clear t;
R=rmmissing([GPRtestDates]);%eksik verilerin tablodan ç?kart?lmas?
NAN veriler bu sayede analizden ç?kart?ld?.
GPRtestDates = R(:,1);
GPRtestDates = datetime(GPRtestDates, 'InputFormat', 'dd-MMM-yyyy
HH:mm:ss');
clear R;

MdlGPR =
fitrgp(GPRtrainX,GPRtrainY,'KernelFunction','ardsquaredexponential',
'FitMethod','sr','PredictMethod','fic','Standardize',1)

yPredGPR = predict(MdlGPR, GPRtestX);

L = loss(MdlGPR,GPRtestX,GPRtestY)%% 1.3888e+03 1.0828e+03
```



```
r=GPRtestY-yPredGPR;
mapeGPR = nanmean(abs(r./GPRtestY*100))%%2.0921 2.0836 1.8729
```

```
figure(7);
plot(GPRtestDates, GPRtestY,
'DisplayName','SYSLoadTest','YDataSource',
'SYSLoadTest','color','b');grid on;hold on;
plot(GPRtestDates, yPredGPR, 'DisplayName','yPredGPR','YDataSource',
'yPredGPR','color',[0.8290, 0.5940, 0.3250]);grid on;hold on;
title('Gaussian Process Regression Predictions 2018-2019');
xlabel('Dates And Times');
ylabel('MW');
legend;
```

```
%Optimize parameter
```

```
MdlGPROpt =
fitrgp(GPRtrainX,GPRtrainY,'KernelFunction','squaredexponential',...
'OptimizeHyperparameters','auto','HyperparameterOptimizationOptions'
,...
struct('AcquisitionFunctionName','expected-improvement-plus'))
```