T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKA ALGORİTMALARI İLE ELEKTRİK TALEP TAHMİNİ: TEKİRDAĞ, ÇERKEZKÖY ÖRNEĞİ

ATABERK GÜÇLÜ

YÜKSEK LİSANS TEZİ

MEKATRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışmanı: Prof. Dr. Gürkan TUNA

EDİRNE-2025

Ataberk Güçlü'nün hazırladığı "**YAPAY ZEKA ALGORİTMALARI İLE ELEKTRİK TALEP TAHMİNİ: TEKİRDAĞ, ÇERKEZKÖY ÖRNEĞİ**" başlıklı bu tez, tarafımızca okunmuş, kapsam ve niteliği açısından **Mekatronik Mühendisliği** Anabilim Dalında bir **Yüksek Lisans tezi** olarak kabul edilmiştir.

Jüri Üyeleri: İmza

Prof. Dr. Gürkan TUNA ……

Prof. Dr. İlhan UMUT ……...

Dr. Öğr. Üyesi Özcan ÇETİNKAYA ……...

Tez Savunma Tarihi: 10/01/2025

Bu tezin Yüksek Lisans/Doktora tezi olarak gerekli şartları sağladığını onaylarım.

Prof. Dr. Gürkan TUNA ……

Tez Danışmanı

Trakya Üniversitesi Fen Bilimleri Enstitüsü onayı

……

Doç. Dr. Filiz UMAROĞULLARI

Fen Bilimleri Enstitüsü Müdürü

T.Ü.FEN BİLİMLERİ ENSTİTÜSÜ

MEKATRONİK MÜHENDİSLİĞİ YÜKSEK LİSANS PROGRAMI

DOĞRULUK BEYANI

Trakya Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında, tüm verilerin bilimsel ve akademik kurallar çerçevesinde elde edildiğini, kullanılan verilerde tahrifat yapılmadığını, tezin akademik ve etik kurallara uygun olarak yazıldığını, kullanılan tüm literatür bilgilerinin bilimsel normlara uygun bir şekilde kaynak gösterilerek ilgili tezde yer aldığını ve bu tezin tamamı ya da herhangi bir bölümünün daha önceden Trakya Üniversitesi ya da farklı bir üniversitede tez çalışması olarak sunulmadığını beyan ederim.

10/01/2025

Ataberk GÜÇLÜ

İmza

Yüksek Lisans Tezi

YAPAY ZEKA ALGORİTMALARI İLE ELEKTRİK TALEP TAHMİNİ: TEKİRDAĞ, ÇERKEZKÖY ÖRNEĞİ

T.Ü. Fen Bilimleri Enstitüsü

Mekatronik Mühendisliği Anabilim Dalı

ÖZET

Elektrik, günümüzde endüstride ve birçok alanda kullanılmaktadır. Endüstride elektrik, iş gücüne, üretim sahasının genişliğine ve mevsimsel şartlara bağlı olarak soğutma, ısıtma, aydınlatma, elektronik ve mekanik işlemlerde kullanılmaktadır. Bu tez çalışmasında, Tekirdağ, Çerkezköy bölgesinin vaka analizi, en küçük kareler yöntemi, k-en yakın komşu yöntemi, karar ağacı, destek vektör makinesi ve yapay sinir ağları algoritmaları kullanarak kısa dönem talep tahmini gerçekleştirilmiştir.

Bu tez çalışmasında Tekirdağ, Çerkezköy bölgesinin, 1 Ocak 2021- 31 Aralık 2023 dönemine ait saatlik tüketim verileri Trakya Elektrik Perakende Satış A.Ş'den temin edilmiştir. Aynı döneme ait saatlik meteorolojik veriler meteoblue veri tabanından sağlanmıştır. Tahmin çıktılarının grafikler üzerinde anlaşılabilir ve yorumlanabilir olması adına veri seti küçültülmüştür. En küçük kareler ve k-en yakın komşu yöntemi algoritmaları için 1 Mayıs 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 72 saatlik süreçteki değerler veri seti olarak kullanılmıştır. Karar ağacı, destek vektör makinesi ve yapay sinir ağları algoritmaları için 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler veri seti olarak kullanılmıştır. 240 saatlik veri setini %70'i eğitim %30'u test veri seti olarak kullanılmıştır.

MATLAB üzerinde elde edilen sonuçlar, ortalama mutlak yüzde hatası, korelasyon katsayısı ve kök ortalama kare hatası hata metriklerine dayalı olarak analiz

edilmiştir. Bu işlemler sonunda, çalışmada kullanılan yöntemler başarımlarına göre sıralanmış ve en az hata oranına sahip yöntem belirlenmiştir.

Yıl : 2025

Sayfa Sayısı : 161

Anahtar Kelimeler : Talep tahmini, karar ağacı, yapay sinir ağları, destek vektör makinesi, k-en yakın komşular yöntemi, en küçük kareler yöntemi

Master Thesis

ELECTRICITY DEMAND FORECASTING WITH ARTIFICIAL INTELLIGENCE ALGORITHMS: THE CASE OF TEKİRDAĞ, ÇERKEZKÖY

Trakya University Institute of Natural Sciences

Mechatronics Engineering

ABSTRACT

Electricity is widely used in industry and various other fields today. In industrial applications, electricity is utilized for cooling, heating, lighting, electronic, and mechanical operations depending on workforce requirements, the scale of the production site, and seasonal conditions. In this thesis study, a case analysis of the Tekirdağ, Çerkezköy region was conducted to perform short-term demand forecasting using the least squares method, k-nearest neighbors, decision tree, support vector machine, and artificial neural networks algorithms.

In this thesis, hourly consumption data for the Tekirdağ, Çerkezköy region covering the period from January 1, 2021, to December 31, 2023, were obtained. Hourly meteorological data for the same period were provided from the Meteoblue database. To ensure that the forecast outputs are interpretable and comprehensible in graphical form, the dataset was reduced. For the least squares and k-nearest neighbors algorithms, a 72-hour dataset covering the period from May 1, 2021, 00:00 to May 4, 2021, 00:00 was used. For the decision tree, support vector machine, and artificial neural networks algorithms, a 240-hour dataset covering the period from April 24, 2021, 00:00 to May 4, 2021, 00:00 was utilized. The 240-hour dataset was divided into 70% training and 30% testing sets. The results obtained in MATLAB were analyzed based on error metrics, including mean absolute percentage error, correlation coefficient, and root mean square

error. Following these processes, the methods used in the study were ranked based on their performance, and the method with the lowest error rate was identified.

Year : 2025

Number of Pages : 161

Keywords : Load forecasting, least squares method, k-nearest neighbors method, decision tree, support vector machine, artificial neural networks

İÇİNDEKİLER

ÖZET iv
ABSTRACT vi
ŞEKİLLER DİZİNİ x
ÇİZELGELER DİZİNİ xiv
BÖLÜM 1 1
GİRİŞ 1
BÖLÜM 2 4
LİTERATÜR TARAMASI 4
2.1. Giriş 4
BÖLÜM 3 25
TALEP TAHMİNİ 25
3.1. Giriş 25
3.2. Talep Tahmini Çeşitleri 25
3.2.1. Çok Kısa Dönem Talep Tahmini (ÇKDTT) 26
3.2.2. Kısa Dönem Talep Tahmini (KDTT) 26
3.2.3. Orta Dönem Talep Tahmini (ODTT) 26
3.2.4. Uzun Dönem Talep Tahmini (UDTT) 26
3.3. Talep Tahminini Etkileyen Faktörler 26
3.3.1. Meteorolojik Faktörler 27
3.3.2. Zamansal veya Takvimsel Faktörler 27
3.3.3. Tesadüfi Faktörler 27
3.3.4. Tüketici Kaynaklı Faktörler 27
BÖLÜM 4 28
MATERYAL ve METOTLAR 28
4.1. Giriş 28

4.2. Öğrenme Şekilleri 29
4.2.1. Denetimli öğrenme 29
4.2.2. Denetimsiz öğrenme 29
4.2.3. Pekiştirmeli öğrenme 29
4.3. Çalışmada Kullanılan Tahmin Yöntemleri 29
4.3.1. En Küçük Kareler Yöntemi 29
4.3.2. K-En Yakın Komşu 30
4.3.3. Karar Ağaçları 31
4.3.4. Destek Vektör Makinesi 32
4.3.5. Yapay Sinir Ağları 33
BÖLÜM 5 36
TALEP TAHMİNİ UYGULAMASI ÇERKEZKÖY BÖLGESİ VAKA ANALİZİ 36
5.1. Tahmin ve Analiz 36
5.1.1. En Küçük Kareler Yöntemi ile Talep Tahmini Uygulaması 38
5.1.2. K-En Yakın Komşu ile Talep Tahmini Uygulaması 58
5.1.3. Karar Ağacı ile Talep Tahmini Uygulaması 70
5.1.4. Destek Vektör Makinesi ile Talep Tahmini Uygulaması 93
5.1.5. Yapay Sinir Ağları ile Talep Tahmini Uygulaması 104
5.2. Hata Metrikleri 120
5.2.1. Mean Absolute Percentage Error (MAPE) 120
5.2.2. Root Mean Squared Error (RMSE) 120
5.2.3. Korelasyon Katsayısı (R): 121
5.3. Tahmin Sonuçlarının Değerlendirilmesi 121
BÖLÜM 6 131
SONUÇ 131
6.1. Öneriler 133
KAYNAKÇA 134
ÖZGEÇMİŞ 147

ŞEKİLLER DİZİNİ

Şekil 4.1. Yapay sinir ağları hücresi 34
Şekil 5.1. En küçük kareler yöntemi 1. Senaryo tahmin grafiği 38
Şekil 5.2. En küçük kareler yöntemi 2. Senaryo tahmin grafiği 40
Şekil 5.3. En küçük kareler yöntemi 3. Senaryo tahmin grafiği 41
Şekil 5.4. En küçük kareler yöntemi 4. Senaryo tahmin grafiği 42
Şekil 5.5. En küçük kareler yöntemi 5. Senaryo tahmin grafiği 43
Şekil 5.6. En küçük kareler yöntemi 6. Senaryo tahmin grafiği 44
Şekil 5.7. En küçük kareler yöntemi 7. Senaryo tahmin grafiği 46
Şekil 5.8. En küçük kareler yöntemi 8. Senaryo tahmin grafiği 47
Şekil 5.9. En küçük kareler yöntemi 9. Senaryo tahmin grafiği 48
Şekil 5.10. En küçük kareler yöntemi 10. Senaryo tahmin grafiği 49
Şekil 5.11. En küçük kareler yöntemi 11. Senaryo tahmin grafiği 51
Şekil 5.12. En küçük kareler yöntemi 12. Senaryo tahmin grafiği 52
Şekil 5.13. En küçük kareler yöntemi 13. Senaryo tahmin grafiği 54
Şekil 5.14. En küçük kareler yöntemi 14. Senaryo tahmin grafiği 55
Şekil 5.15. En küçük kareler yöntemi 15. Senaryo tahmin grafiği 57
Şekil 5.16. K-en yakın komşu 1. Senaryo tahmin grafiği 59
Şekil 5.17. K-en yakın komşu 2. Senaryo tahmin grafiği 60
Şekil 5.18. K-en yakın komşu 3.Senaryo tahmin grafiği 61
Şekil 5.19. K-en yakın komşu 4. Senaryo tahmin grafiği 61
Şekil 5.20. K-en yakın komşu 5. Senaryo tahmin grafiği 62
Şekil 5.21. K-en yakın komşu 6. Senaryo tahmin grafiği 63
Şekil 5.22. K-en yakın komşu 7. Senaryo tahmin grafiği 64
Şekil 5.23. K-en yakın komşu 8. Senaryo tahmin grafiği 65
Şekil 5.24. K-en yakın komşu 9. Senaryo tahmin grafiği 66
Şekil 5.25. K-en yakın komşu 10. Senaryo tahmin grafiği 66

Şekil 5.26. K-en yakın komşu 11. Senaryo tahmin grafiği 67
Şekil 5.27. K-en yakın komşu 12. Senaryo tahmin grafiği 68
Şekil 5.28. K-en yakın komşu 13. Senaryo tahmin grafiği 68
Şekil 5.29. K-en yakın komşu 14. Senaryo tahmin grafiği 69
Şekil 5.30. K-en yakın komşu 15. Senaryo tahmin grafiği 70
Şekil 5.31. Karar ağacı 1. Senaryo özellik önemi grafiği 71
Şekil 5.32. Karar ağacı 1. Senaryo tahmin grafiği 71
Şekil 5.33. Karar ağacı 2. Senaryo özellik önemi grafiği 72
Şekil 5.34. Karar ağacı 2. Senaryo tahmin grafiği 73
Şekil 5.35. Karar ağacı 3. Senaryo özellik önemi grafiği 74
Şekil 5.36. Karar ağacı 3. Senaryo tahmin grafiği 74
Şekil 5.37. Karar ağacı 4. Senaryo özellik önemi grafiği 75
Şekil 5.38. Karar ağacı 4. Senaryo tahmin grafiği 76
Şekil 5.39. Karar ağacı 5. Senaryo özellik önemi grafiği 77
Şekil 5.40. Karar ağacı 5. Senaryo tahmin grafiği 77
Şekil 5.41. Karar ağacı 6. Senaryo özellik önemi grafiği 78
Şekil 5.42. Karar ağacı 6. Senaryo tahmin grafiği 79
Şekil 5.43. Karar ağacı 7. Senaryo özellik önemi grafiği 80
Şekil 5.44. Karar ağacı 7. Senaryo tahmin grafiği 80
Şekil 5.45. Karar ağacı 8. Senaryo özellik önemi grafiği 81
Şekil 5.46. Karar ağacı 8. Senaryo tahmin grafiği 82
Şekil 5.47. Karar ağacı 9. Senaryo özellik önemi grafiği 83
Şekil 5.48. Karar ağacı 9. Senaryo tahmin grafiği 83
Şekil 5.49. Karar ağacı 10. Senaryo özellik önemi grafiği 84
Şekil 5.50. Karar ağacı 10. Senaryo tahmin grafiği 85
Şekil 5.51. Karar ağacı 11. Senaryo özellik önemi grafiği 86
Şekil 5.52. Karar ağacı 11. Senaryo tahmin grafiği 86
Şekil 5.53. Karar ağacı 12. Senaryo özellik önemi grafiği 87
Şekil 5.54. Karar ağacı 12. Senaryo tahmin grafiği 88
Şekil 5.55. Karar ağacı 13. Senaryo özellik önemi grafiği 89
Şekil 5.56. Karar ağacı 13. Senaryo tahmin grafiği 89
Şekil 5.57. Karar ağacı 14. Senaryo özellik önemi grafiği 90

Şekil 5.58. Karar ağacı 14. Senaryo tahmin grafiği 91
Şekil 5.59. Karar ağacı 15. Senaryo özellik önemi grafiği 92
Şekil 5.60. Karar ağacı 15. Senaryo tahmin grafiği 92
Şekil 5.61. Destek vektör makinesi 1. Senaryo tahmin grafiği 93
Şekil 5.62. Destek vektör makinesi 2. Senaryo tahmin grafiği 94
Şekil 5.63. Destek vektör makinesi 3. Senaryo tahmin grafiği 95
Şekil 5.64. Destek vektör makinesi 4. Senaryo tahmin grafiği 95
Şekil 5.65. Destek vektör makinesi 5. Senaryo tahmin grafiği 96
Şekil 5.66. Destek vektör makinesi 6. Senaryo tahmin grafiği 97
Şekil 5.67. Destek vektör makinesi 7. Senaryo tahmin grafiği 98
Şekil 5.68. Destek vektör makinesi 8. Senaryo tahmin grafiği 99
Şekil 5.69. Destek vektör makinesi 9. Senaryo tahmin grafiği 100
Şekil 5.70. Destek vektör makinesi 10. Senaryo tahmin grafiği 100
Şekil 5.71. Destek vektör makinesi 11. Senaryo tahmin grafiği 101
Şekil 5.72. Destek vektör makinesi 12. Senaryo tahmin grafiği 102
Şekil 5.73. Destek vektör makinesi 13. Senaryo tahmin grafiği 102
Şekil 5.74. Destek vektör makinesi 14. Senaryo tahmin grafiği 103
Şekil 5.75. Destek vektör makinesi 15. Senaryo tahmin grafiği 104
Şekil 5.76. Yapay sinir ağları 1. Senaryo regresyon grafiği 105
Şekil 5.77. Yapay sinir ağları 2. Senaryo regresyon grafiği 106
Şekil 5.78. Yapay sinir ağları 3. Senaryo regresyon grafiği 107
Şekil 5.79. Yapay sinir ağları 4. Senaryo regresyon grafiği 108
Şekil 5.80. Yapay sinir ağları 5. Senaryo regresyon grafiği 109
Şekil 5.81. Yapay sinir ağları 6. Senaryo regresyon grafiği 110
Şekil 5.82. Yapay sinir ağları 7. Senaryo regresyon grafiği 111
Şekil 5.83. Yapay sinir ağları 8. Senaryo regresyon grafiği 112
Şekil 5.84. Yapay sinir ağları 9. Senaryo regresyon grafiği 113
Şekil 5.85. Yapay sinir ağları 10. Senaryo regresyon grafiği 114
Şekil 5.86. Yapay sinir ağları 11. Senaryo regresyon grafiği 115
Şekil 5.87. Yapay sinir ağları 12. Senaryo regresyon grafiği 116
Şekil 5.88. Yapay sinir ağları 13. Senaryo regresyon grafiği 117
Şekil 5.89. Yapay sinir ağları 14. Senaryo regresyon grafiği 118

Şekil 5.90 Yapay sinir ağları 15. Senaryo regresyon grafiği 119

ÇİZELGELER DİZİNİ

Çizelge 2.1. Literatürdeki Çalışmalarda Elde Edilen Hata Değerleri 23
Çizelge 5.1. Senaryolar Özelinde Kullanılan Bağımsız Değişkenler 37
Çizelge 5.2. 1. Senaryo Hata Değerleri 121
Çizelge 5.3. 2. Senaryo Hata Değerleri 122
Çizelge 5.4. 3. Senaryo Hata Değerleri 122
Çizelge 5.5. 4. Senaryo Hata Değerleri 123
Çizelge 5.6. 5. Senaryo Hata Değerleri 124
Çizelge 5.7. 6. Senaryo Hata Değerleri 124
Çizelge 5.8. 7. Senaryo Hata Değerleri 125
Çizelge 5.9. 8. Senaryo Hata Değerleri 125
Çizelge 5.10. 9. Senaryo Hata Değerleri 126
Çizelge 5.11. 10. Senaryo Hata Değerleri 126
Çizelge 5.12. 11. Senaryo Hata Değerleri 127
Çizelge 5.13. 12. Senaryo Hata Değerleri 127
Çizelge 5.14. 13. Senaryo Hata Değerleri 128
Çizelge 5.15. 14. Senaryo Hata Değerleri 128
Çizelge 5.16. 15. Senaryo Hata Değerleri 129

BÖLÜM 1

GİRİŞ

İnsanoğlu, varoluşunun başından beri kendini geliştirmek ve hayat şartlarını daha iyi bir duruma getirmek için çabalamaktadır. Bu süreçte, sanayi devriminin gerçekleşmesiyle birlikte katlanarak artan endüstriyel büyüme birçok alanda önemli değişikliklere yol açmıştır. Bununla birlikte enerji, hayat standartlarını yükseltmek ve sanayileşmeyi sürdürmek için günümüz dünyasının vazgeçilmez unsurlarından biri haline gelmiştir. İletilebilir, dönüştürülebilir, depolanabilir, temiz ve verimli bir enerji türü olan elektrik enerjisi, teknolojinin hızlı gelişimiyle birlikte büyük önem kazanmıştır. Üretim, iletişim, sağlık, aydınlatma ve ulaşım gibi birçok alanda vazgeçilmez bir kaynaktır. Endüstride elektrik, iş gücüne, üretim sahasının genişliğine ve mevsimsel şartlara bağlı olarak soğutma, ısıtma, aydınlatma, elektronik ve mekanik işlemlerde kullanılmaktadır.

Gün geçtikçe artış gösteren elektrik talebinin etkili bir şekilde karşılanma gereksinimi, sürdürebilirlik ve enerji verimliliği konularını ön plana çıkartmaktadır. Talep tahmini, elektrik enerjisinin üretimini, dağıtımını ve kullanımını en uygun şartlarda yönetmekte büyük rol oynar ve endüstriyel tesislerde kritik öneme sahiptir. Bu tahminler fabrikaların veya dağıtım istasyonlarının enerji ihtiyaçlarını belirlemelerine ve enerji planları oluşturmalarına yardımcı olur. Enerji üretiminin yanı sıra dağıtımının doğru zaman ve miktarda gerçekleştirilmesi de önemli bir husustur. Endüstride talep tahmini gelecekteki ürün, hizmet veya enerji ihtiyaçlarını öngörerek üretim ve tedarik süreçlerinin daha verimli bir şekilde yönetilmesine katkı sağlar. Bu nedenle elektrik yükü talep tahmini, endüstriyel araştırma ve geliştirmede önemli bir konu haline gelmiştir.

Bu tez çalışmasında Tekirdağ, Çerkezköy ilçesinin vaka analizi ve talep tahmini gerçekleştirilmiştir. Bölgesel hava durumu şartları ve tüketim karakteristiği incelenerek veri setleri düzenlenmiştir ve uygun tahmin modelleri oluşturulmuştur. Bu çalışmada Tekirdağ, Çerkezköy bölgesinin talep tahmini vaka analizi ile birlikte En Küçük Kareler Yöntemi (LS), K-En Yakın Komşu (KNN), Karar Ağacı (DT), Destek Vektör Makinesi (SVM) ve Yapay Sinir Ağları (ANN) algoritmalarının performansları karşılaştırmalı olarak değerlendirilmiştir.

Çalışma kapsamında Tekirdağ, Çerkezköy bölgesinin, 1 Ocak 2021- 31 Aralık 2023 dönemine ait saatlik tüketim verileri Trakya Elektrik Perakende Satış A.Ş (TREPAŞ)'tan temin edilmiştir. Çerkezköy ilçesinin aynı döneme ait sıcaklık, bağıl nem ve rüzgâr hızı hava durumu verileri meteoblue veri tabanından günlük ve saatlik olarak temin edilmiştir. Tahmin çıktılarının grafikler üzerinde anlaşılabilir ve yorumlanabilir olması adına veri seti küçültülmüştür. LS ve KNN algoritmaları için 1 Mayıs 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 72 saatlik süreçteki değerler veri seti olarak kullanılmıştır. DT, SVM ve ANN algoritmaları için 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler veri seti olarak kullanılmıştır. 240 saatlik veri setinin %70'i eğitim %30'u test veri seti olarak kullanılmıştır. Parametre ayarları, algoritmaların veri seti üzerindeki davranışları gözlemlenerek belirlenmiştir. Analizi yapılan vaka özelinde en iyi hata değerlerini sağlayan parametre değerleri kullanılmıştır.

Algoritmalar özelinde hangi bağımsız değişkenlerin daha etkili olacağını gözlemlemek amacıyla 15 farklı tahmin senaryosu oluşturulmuştur. Senaryolarda bağımsız değişken olarak günün saati, haftanın günü, hafta sonu olma durumu, resmi tatil olma durumu, sıcaklık, nem, rüzgâr, 1 hafta öncesinin aynı saatteki tüketim verisi, 1 gün öncesinin aynı saatteki tüketim verisi ve 1 saat öncesinin tüketim verisi kullanılmıştır. Senaryolarda bağımlı değişken olarak saatlik tüketim verileri kullanılmıştır. Tahmin performansını arttırmak amacıyla eksik veya hatalı verilerin olduğu dönemler tespit edilerek veri setinden çıkarılmıştır. Algoritmaların her senaryo özelinde performans değerlendirmesi gerçekleştirilmiştir.

Tez çalışması toplamda altı bölümden oluşmaktadır. Birinci bölüm, çalışmanın amacını ve kapsamını aktarmaktadır. İkinci bölüm literatürde gerçekleştirilen talep tahmini uygulamalarından oluşmaktadır. Üçüncü bölümde, talep tahmini, türleri ve

tahmin sonucunu etkileyen faktörler hakkında bilgi verilmiştir. Dördüncü bölümde, çalışmada kullanılan yapay zekâ ve makine öğrenmesi algoritmalarının veriyi işleme yöntemlerine yer verilmiştir. Beşinci bölümde, tahmin süreci için gerçekleştirilen veri seti düzenlemesinden, algoritmaların parametre ayarlarından ve çalışmada kullanılan hata metriklerinden bahsedilmektedir. Bunlarla birlikte beşinci bölümde tahmin grafiklerinin yorumlanması ve hata çizelgelerinin değerlendirilmesine yer verilmiştir. Altıncı bölüm tez çalışmasının sonucunu ve gelecek çalışmalar için geliştirilebilecek yönlerini içermektedir.

BÖLÜM 2

LİTERATÜR TARAMASI

2.1. Giriş

Talep tahmini için kullanılan her bir yöntemin belirli durumlarda avantajları ve dezavantajları bulunmaktadır. Örneğin, ANN istatistiki öğrenme teorisine göre, tecrübe riskini azaltmayı sağlayan ampirik risk azaltma ilkesine dayanmaktadır ve niceleyici analizden yoksundur (Berner, Grohs, & Jentzen, 2018). SVM, ampirik risk azaltma ilkesinin yerine yapısal risk azaltma ilkesini kullanmaktadır (Campbell, Campbell, Reynolds, Singer, & Torres-Carrasquillo, 2006). Bu sayede aşırı uyum (overfitting) engellenerek yüksek genelleştirme performansı yakalanmaktadır (Atalay, 2013).

Talep tahmini, literatürde farklı yaklaşımlar ile gerçekleştirilmiştir. Şener (2005)'te aktif ve reaktif enerji tüketimleri, çekilen maksimum ve minimum güçler ve saatleri, güç katsayıları ve nüfus gibi bağımsız değişkenler ile senaryolar oluşturmuştur. Ankara ilinde bir bölgenin, regresyon analizi ile talep tahminini gerçekleştirmiştir. Talep tahmini için LS, ANN, Otoregresif Hareketli Ortalama (ARIMA), Çoklu Doğrusal Regresyon (MLR) ve eğilim analizi kullanılmıştır. En iyi tahmin performansı %1.46 MAPE değeri ile LS algoritmasına aittir (Şener, 2005). Esener (2012)'de bağımsız değişken olarak sıcaklık veya başka bir hava verisi kullanmadan, tahmin gününün öncesindeki 3 günün, bir hafta öncesindeki ve iki hafta öncesindeki günlerin saatlik yük verilerini kullanmıştır. Çıkış olarak tahmin gününün bir hafta öncesi ve iki hafta öncesindeki günlerin saatlik yük verisini kullanılarak, tüketim verisi üzerinden bir eğitim senaryosu oluşturmuştur ve Türkiye geneli 24 saatlik talep tahmini gerçekleştirmiştir.

ANN, Dalgacık Dönüşümü (WT), Dalgacık Dönüşümü ve Radyal Taban Fonksiyonlu Sinir Ağları (RTFA), Görgül Kip Ayrışımı (EMD) ve Radyal Taban Fonksiyonu (RBF) gibi yöntemler kullanılarak kabul edilebilir hata oranları içerisinde 24 saatlik talep tahmini yapılmıştır. WT, zaman serilerini dalgacık adı verilen fonksiyonlar yardımıyla ayrıştırmak için kullanılan bir yöntemdir (Yelgeç vd., 2022). EMD, orijinal zaman serisi verilerinin, kısa süreli karakteristik zaman ölçeğine dayalı olarak sonlu salınım modlarına ayrıştırarak analiz gerçekleştiren bir modeldir (Gün, 2023). En düşük Ortalama Mutlak Yüzde Hatası (MAPE) değerini WT ve RTFA vermiştir (Esener, 2012). Tutu, 2017'de günlük kısa dönem talep tahmini gerçekleştirmiştir. Eğitim verisi için 2012-2015 yılları arası Türkiye günlük elektrik tüketim verileri, test verisi olarak ise 2016 yılı Ocak, Şubat ve Mart ayı değerleri kullanılmıştır. Tahmin yöntemi olarak LS kullanılmıştır. Giriş parametresi olarak sıcaklık, nem ve bir önceki güne ait elektrik değerlerinin kullanılmıştır. Ocak, Şubat ve Mart ayı için sırasıyla %0.08, %0.04 ve %0.1 MAPE değerleri sağlanmıştır (Tutu vd., 2017).

Burger ve Moura (2015), 32 binanın tüketim verilerini kullanarak Sırt Regresyonu (RR), SVM, DT, KNN algoritmaları ile talep tahmini gerçekleştirmişlerdir. Binalardaki elektrik tüketim değeri dönemsel olarak değişiklik gösterdiği ve tahmini gerçekleştirecek algoritmaların tutarlılığını etkilediği için kapılı birleşim (gated ensemble) yöntemi kullanılmıştır. Kapılı birleşim yöntemi eğitim, doğrulama, kapılama ve test aşamalarından oluşmaktadır. Kapılama adımında hangi modelin hangi girdilerle en iyi performansı vereceği belirlenmektedir. Girdiler sonucu her bir modelin çıktıları ağırlıklandırılmaktadır. Ağırlık katsayısı büyük olan girdinin belirleyiciliği yüksektir (Gensler & Sick, 2018). Kapılama adımının sonunda hata değeri en az olan model seçilmektedir. Eğitim veri seti, Ocak 2012 ile Temmuz 2012 arasındaki saatlik tüketim değerlerinden oluşmaktadır. İlgili bina için saatlik Tüketim Değeri(D), Seyrek İkili Vektör(T) olarak temsil edilen Zaman Kayıtları, Saatlik Tüketim Verileri ve Zaman Kayıtları(TD), Saatlik Tüketim Verileri, Zaman Kayıtları ve Temsil Edilen Zamana Bağlı Hava Şartları(TE) olmak üzere 4 veri seti oluşturulmuştur. Doğrulama adımında her bir bina için talep tahmini gerçekleştirilmiştir ve Temmuz 2013- Ocak 2014 arası değerlerle karşılaştırılmıştır. Kapılama adımında, Çapraz Doğrulama (CV) yöntemi ile en düşük Kök Ortalama Kare Hatası (RMSE) değerini üreten algoritma ve Tekrarlı Eğitimli Doğrusal Regresyon (ITLR) yöntemi kullanılarak en iyi test sonucunu verecek algoritma

tahmini gerçekleştirmiştir. Kapılama adımında en iyi sonucu veren algoritma kullanılarak test aşaması uygulanmıştır. En iyi tahmin sonucunu %7.5 MAPE değeriyle KNN algoritması gerçekleştirmiştir (Burger & Moura, 2015).

Fan ve ark. (2019), zaman serisi metodu, Geri Yayılım Sinir Ağı (BPNN), Gri Model (GM) ve SVM gibi yaygın kullanılan talep tahmini algoritmalarının dezavantajlarını göz önünde bulundurarak yeni bir yaklaşım önerisinde bulunmuşlardır. BPNN, hata geri yayılımı ile bağlantı ağırlıklarını ve eşik değerlerini en uygun hale getiren bir sinir ağı modelidir (Bashir, Haoyong, Tahir, & Liqiang, 2022). GM, veri setinin bilgi seviyesini renklerle ifade eden, küçük ve eksik veri setleri ile gerçekçi desenleri yakalayabilme kabiliyetine sahip algoritmadır (Ho, 2010). Fan ve ark. (2019), bu çalışmasında, elektrik tüketim değerlerini etkileyen değişkenler ile tüketim değerleri arasında matematiksel model oluşturmak, elektrik yükünün durumunu gerçek zamanlı yansıtabilmek ve daha isabetli tahminler gerçekleştirebilmek için Ağırlıklı K-En Yakın Komşu (W-KNN) algoritması önerisinde bulunmuşlardır. Eğitim veri seti olarak Avusturalya Ulusal Elektrik Pazarı'nın 1 Ocak 2007 saat 08:00'den 1 Ocak 2008 saat 00:00'a kadar olan elektrik tüketim verileri kullanılmıştır. Bu veriler 8 saatlik dönemlere bölünerek ortalama tüketim değerleri alınmıştır ve kısa dönem talep tahmini gerçekleştirilmiştir. 3. ay ve 4. hafta olmak üzere 2 tahmin senaryosu gerçekleştirilmiştir. Önerilen algoritmanın üstünlüğünü göstermek adına tahmin sonuçları Otokorelasyonlu Hareketli Uygulama (ARMA), BPNN ve KNN algoritmaları ile karşılaştırılmıştır. 2 tahmin senaryosu için en düşük MAPE değerini veren algoritma %3.87 oranı ile W-KNN olmuştur (Fan, Guo, Zheng, & Hong, 2019). ARMA, otoregresif ve hareketli ortalama algoritmaların bir birleşimidir. Bir değişkenin tahmin edilmek istenen değerini, geçmiş gözlemler ve rastgele hataların doğrusal bir fonksiyonu varsayarak tahmin gerçekleştiren algoritmadır (Ervural, Beyca, & Zaim, 2016).

Talep tahmini sadece kısa vade için değil, orta ve uzun vade için de gerçekleştirilebilir. Lin ve ark. (2016), mevsimsel zaman serisi verilerini etkili bir şekilde değerlendirebilmek için Evrimsel Mevsimsel Ayrıştırma En Küçük Kareler Destek Vektör Regresyonu(ESDLS-SVR) metodunu geliştirmişlerdir. Tayvan Güç Şirketi (Taipower)'ne ait 16 güneş enerjisi istasyonunun aylık enerji çıktılarını göz önünde bulundurarak orta vadeli yük tahmini gerçekleştirmişlerdir. Bu çalışma için Ocak 2010'dan Nisan 2014'e kadar olan aylık enerji çıktıları eğitim, doğrulama ve test olmak

üzere 3 veri setine ayrılmıştır. Ocak 2010- Aralık 2011 arası veriler eğitim seti, Ocak 2012- Aralık 2012 arası veriler doğrulama seti ve Ocak 2013- Nisan 2014 arası veriler test seti olarak kullanılmıştır. Mevsimsel ayrıştırma işlemi ve uygun parametrelerin uygulanması için Genetik Algoritma (GA) kullanılarak Bütünleşik ARIMA, Mevsimsel Bütünleşik Otoregresif Hareketli Ortalama (SARIMA), Genel Regresyon Sinir Ağı (GRNN), En Küçük Kareler Destek Vektör Regresyonu (LS-SVR) ile RBF metotlarıyla tahmin gerçekleştirilmiştir. ARIMA, durağan olmayan zaman serilerini fark alma yöntemiyle durağan hale getiren ve tahmin edilmek istenen değeri geçmiş gözlemler ve rastgele hataların doğrusal bir fonksiyonu varsayarak tahmin gerçekleştiren algoritmadır (Namini vd. 2018). SARIMA, ARIMA modelinin mevsimsel parametreleri dâhil edebildiğimiz, geliştirilmiş halidir (Farsi vd., 2021). GRNN, doğrusal olmayan regresyon analizi gerçekleştiren bir sinir ağı modelidir (Qiao, Wang, & Zhu, 2020). ESDLS-SVR ile lineer çekirdek fonksiyonu yöntemi 0.1618 RMSE ve %7.8434 MAPE değerleri ile diğer yöntemlere göre daha iyi sonuç vermiştir (Lin & Pai, 2016).

Talep tahmini çalışmalarında tek kaynaktan elde edilen veriler kullanılabileceği gibi farklı kaynaklardan elde edilen veri setleri de kullanılabilir. Atalay (2013)'te veri seti olarak 1996-1998 yılları European Network on Intelligent Technologies for Smart Adaptive Systems (EUNITE) ve 2013 yılının ilk 5 ayını kullanmıştır. Veri setlerinin analizi SVM, ANN ve Rastlantısal Ormanlar (RF) metotlarıyla gerçekleştirilmiştir. Elektrik tüketim değerleri, ekonomik faktörler ve mevsimsel parametrelerin varyasyonlarıyla 13 farklı senaryo oluşturulmuştur. Çalışma sonucunda meteorolojik verinin azaltılması ile hata oranının düşürülebildiğine ve diferansiyel evrim metoduyla yapılan SVM'nin, bu problem için en uygun model olduğu sonucuna varılmıştır. İki veri seti için de en düşük hata oranını veren yöntem %1.34 MAPE değeri ile destek vektör makinesi olmuştur (Atalay, 2013).

Chia ve ark. (2015), güneş enerjisi sistemlerinde teslim edilen gücün güvenilirliğini attırmak ve pil ömrünü iyileştirmek amacıyla SVM kullanarak yük öngörü önerisinde bulunmuşlardır. Ani tepe güç talebinin sisteme olan etkilerini azaltmak, süperkapasitör-batarya birleşik enerji depolama bileşimi ve yük arasındaki akışı kontrol etmek hedeflenmiştir. Yük nitelik verileri deney düzeneğinden MATLAB ortamında 10 Hz örnekleme oranıyla toplanmıştır. SVM RBF çekirdek için ceza parametresi C = 2 ve RBF gama parametresi g = 0.25 olarak belirlenmiştir. En iyi performansa ulaşmak amaçlı

CV ve Izgara Arama (GS) uygulanmıştır. %4.9574 MAPE değeri elde edilmiştir (Chia, Lee, Shafiabady, & Isa, 2015).

(Demren, 2011)'de İstanbul Avrupa yakasının 3 yıllık elektrik tüketim değerleri, ekonomik faktörleri ve mevsimsel şartları veri seti olarak kullanarak Destek Vektör Regresyonu (SVR) ve ANN ile orta vadeli talep tahmini gerçekleştirmiştir. Elektrik tüketim değerleri, ekonomik faktörler ve mevsimsel parametreler üzerinden 17 farklı senaryo oluşturularak algoritmaların eğitimi gerçekleştirilmiştir. Orta vadeli talep tahmininde SVR modelinin %2.35 MAPE değeri ile ANN modeline göre doğruluk oranının daha yüksek olduğu sonucuna varılmıştır (Demren, 2011).

Kaytez (2020), ARIMA-LS-SVM birleşimi model ile Türkiye'nin uzun dönem talep tahminini gerçekleştirmiştir. Veri seti oluşturulurken Türkiye İstatistik Kurumu'ndan 1970-2017 yılları arasındaki nüfus verileri temin edilmiştir. Aynı döneme ait elektrik ithalat, ihracat, kurulu kapasite ve brüt elektrik üretimi verileri Türkiye Elektrik Ticaret ve Taahhüt A.Ş. tarafından sağlanmıştır. Net elektrik tüketimi verileri ve abone sayısı verileri ise Türkiye Elektrik Dağıtım A.Ş. yıllık raporlarından alınmıştır. Verilerin %85'i eğitim, %11'i test ve %4'ü doğrulama için kullanılmıştır. Kaytez (2020), ARIMA algoritması ile elektrik tüketim karakteristiğindeki gizli doğrusal desenleri, ARIMA-LS-SVM algoritması ile doğrusal olmayan desenleri yakalamıştır. ARIMA-LS-SVM modeli, LS ile SVM'de kullanılan karmaşık 2. dereceden programlama problemini doğrusal bir denklem sistemine indirgeyen ve SVM çekirdek fonksiyonları ile veri uzayındaki doğrusal olmayan desenleri yakalama yeteneğine sahip birleşik bir modeldir (Kang, Tan, Yuan, Lei, & Yuan, 2017). MAED, enerji üretim sistemlerinde ekonomik yük dağılımını birden fazla bölgeyi kapsayacak şekilde daha verimli hale getirmeyi amaçlayan bir yöntemdir (Naderipour vd., 2021). MAED algoritması ile kıyaslanan ARIMA-LS-SVM birleşik modeli, %1.1002 MAPE değeri ile gerçek değerlere daha yakın tahmin gerçekleştirmiştir. MAED algoritmasının MAPE değeri %11.81'dir (Kaytez, 2020).

Ahmad ve ark. (2019), ANN algoritmasının tahmin doğrulunu arttırmak ve yakınsama süresini düşürmek amacıyla DALF modelini önermişlerdir. Önerilen model Bi-level ve MI+ANN modelleriyle karşılaştırılmıştır. Bi-level, birinci seviye karar vericinin kararlarını, ikinci seviye karar vericinin tepkileri ile en uygun hale getiren iki

seviyeli bir yöntemidir (Mao, Zeng, Leng, Zhai, & Keane, 2009). DALF modeli ön işlem, tahmin ve uygun hale getirme birimlerinden oluşmaktadır. Ön işlem biriminde tahmin sonuçlarını iyileştirmek ve eğitim sürecini hızlandırmak için gereksiz ve bağlantısı olmayan veriler ayıklanır. Verilerin ayıklanması Karşılıklı Bilgi (Mutual Information) (MI) tekniği ile gerçekleştirilir. MI, 2 değişkenin içerdiği ilişkili bilgileri belirleyerek bir oran çıkartır. MI değeri yüksek olan özellikler güçlü ilişkiyi, düşük olanlar ise zayıf ilişkiyi belirtmektedir (Bozic, Stojanović, Stajić, & Floranović, 2013). Tahmin birimi, ertesi günün 1 saatlik yükünü kademeli olarak tahmin eden 24 tahminciden oluşmaktadır. Her bir tahminci sigmoid aktivasyon fonksiyonunu kullanan ve Çok Değişkenli Otoregresif Algoritma (MARA) ile eğitilen bir ANN'dir. Eğitim veri setinin bir kısmı doğrulama amaçlı kullanılmıştır ve doğrulama hatası ile ANN'lerin ağırlıkları ayarlanmıştır. Optimizasyon biriminde mEDE algoritması kullanılarak seçim sürecindeki rastgelelik etkisini azaltarak tahmin hatasını en aza indirgemek hedeflenmiştir. Ahmad ve ark. (2019), PJM elektrik piyasasından alınan gerçek zamanlı verilerle DAYTOWN(Ohio, ABD) ve EKPC(Kentcky, ABD) şebekelerinde MATLAB ortamında saatlik talep tahmini gerçekleştirmiştir. Tahmin sonuçları Bi-level ve MI+ANN modelleriyle karşılaştırılan DALF, %1.24 ortalama ile en düşük MAPE değerlerini elde etmiştir (Ahmad, Javaid, Mateen, Awais, & Khan, 2019).

Nalcaci ve ark. (2018), Çok Değişkenli Uyumlaştırılmış Regresyon Sınırlandırması (MARS), ANN ve LR modellerini kullanarak Türkiye'nin uzun dönem talep tahminini gerçekleştirmişlerdir. MARS, bağımlı değişken ile bağımsız değişkenlerin parçalı doğrusal ve kübik regresyonunu kullanarak ilişki analizini gerçekleştiren bir modeldir (Al-Musaylh, Deo, Adamowski, & Li, 2018). Eğitim ve test verileri tatiller, sıcaklık, göreceli nem, rüzgâr hızı ve hafta sonları göz önünde bulundurularak incelenip, düzenlenmiştir. 2011-2015 dönemine ait saatlik elektrik tüketim değerleri Türkiye Elektrik İletim A.Ş.'den temin edilmiştir. Aynı döneme ait hava durumu verileri Türkiye Devlet Meteoroloji İşleri Genel Müdürlüğü'nden alınmıştır. MATLAB ve Salford MARS ortamlarında oluşturulan MARS, ANN ve LR modellerinin eğitimi için 2011-2013 verileri, test için 2013-2015 verileri kullanılmıştır. Tahmin sonuçlarının hata oranlarını karşılaştıran Nalcaci ve ark. (2018), MARS modelinin 0.865 R^2, 0.0012 Ortalama Mutlak Hata (MAE), 0.017 RMSE ve %0.040 MAPE değeri ile en

tutarlı tahmin sonucunu verdiğini ortaya çıkartmışlardır (Nalcaci, Özmen, & Weber, 2019).

Shirzadi ve ark. (2021) akıllı şebekeler için daha verimli bir talep tahmini hedeflemiştir. SVR, RF, Uzun Kısa Süreli Bellek(LSTM) ve Doğrusal Olmayan Otoregresif Dışsal Sinir Ağı (NARX) modellerini kullanan Shirzadi ve ark. (2021), Kanada'nın Bruce County bölgesinin orta dönem yük talep tahminini gerçekleştirmişlerdir. RF karar ağaçlarının birleştirilmesi ile elde edilen toplu öğrenme yöntemidir (Rigatti vd., 2016). LSTM zamansal verilerdeki uzun süreli bağımlılıkları öğrenmek için özel olarak tasarlanmış RNN türüdür (Bouktif, Fiaz, Ouni, & Serhani, 2018). 2010-2019 dönemine ait saatlik elektrik tüketim verileri, sıcaklık ve rüzgâr hızından oluşan saatlik meteorolojik veriler Kanada Hükümeti'nin açık kaynaklı çevresel ve doğal kaynaklar veri tabanından sağlanmıştır. Her model için 9 yıllık veri eğitim seti, 1 yıllık veri test seti olarak kullanılmıştır. Veri setindeki eksikler bir önceki ve bir sonraki tüketim değerlerinin ortalaması alınarak düzenlenmiştir. RF hiper parametreleri 150 ağaç boyutu ve 30 maksimum derinlik olarak belirlenmiştir. SVR için C değeri 0 alınmıştır ve nokta çekirdek fonksiyonu kullanılmıştır. LSTM modeli için Adam uygun hale getirme yöntemi gerçekleştirilmiştir. Adam uygun hale getirme yöntemi, birinci dereceden gradyanlara dayanarak öğrenme ağırlıklarını optimize eden bir yöntemdir (Chandriah & Naraganahalli, 2021). NARX modelin her adımda kaç veriyi işleyeceğini belirleyen grup boyutu parametresi 32, döngü sayısı 10 olarak belirlenmiştir ve aşırı uyumu önlemek amacıyla her döngüde gizli katmandaki nöronların %20'si rastgele devre dışı bırakılmıştır. Gizli katman sayısı 10 ve gecikme değeri 168 olarak belirlenmiştir. Tahmin sonuçları Belirleme Katsayısı (R^2), RMSE ve MAPE kullanılarak değerlendirilmiştir. Eğitim süresi olarak diğer algoritmalara göre yavaş kalsa da en iyi tahmin sonucunu 0.96 R^2, 5.81 MW RMSE ve %4.2 MAPE değeriyle NARX modeli vermiştir (Shirzadi, Nizami, Khazen, & Nik-Bakht, 2021).

Elektrik Güç Sistemlerinin (EPS) yönetimi ve geliştirilmesini kolaylaştırmak adına Tacikistan Cumhuriyeti'nin Gorno-Badakhshan Özerk Bölgesi'ndeki izole bir EPS'in orta dönem talep tahmini gerçekleştiren Matrenin ve ark. (2022), makine öğrenmesi metotlarının uygulanması için en önemli özellikleri belirlemeyi amaçlamıştır. Bu çalışmada LR, SVR, Karar Ağacı Regresyonu (DTR), Rastgele Orman (RF), Karar Ağaçları Gradyan Arttırma(XGBoost), Karar Ağaçları Adaptif Arttırma(AdaBoost), LR

üzerinde AdaBoost yöntemleri kullanılmıştır. XGBoost, kayıp fonksiyonunu minimize etmek için ağaç ekleyen ve her yeni modelin önceki modellerin tahmin hatalarını öğrenen, karar ağacı destekli bir yöntemdir (Nielsen, 2016). AdaBoost, her adımda yanlış sınıflandırılmış verilerin ağırlıklarını artırarak hatayı minimize etmeyi amaçlayan toplu öğrenme yöntemidir (Ratsch, Onoda, & Müller, 2001). Veri seti Kasım 2015'ten Mart 2019'a kadar olan dönemin günlük maksimum yük değerlerinden oluşmaktadır. Zaman serisinden belirli özeliklerin seçilmesi ilgileşim katsayılarının analizi ile gerçekleştirilmiştir. Veri setinin %90'ı eğitim, %10'u test seti olarak kullanılmıştır. Modellerin hiper parametrelerini ayarlamak için CV uygulanmıştır. Her model 8 farklı alt veri seti ile eğitilmiş ve doğrulama veri setlerinde kontrol edilmiştir. En iyi tahmin sonucunu %1.30 MAPE değeriyle LR üzerinde AdaBoost modeli gerçekleştirmiştir (Matrenin vd., 2022).

Madrid ve ark. (2021), güç santrallerindeki talep belirsizliğini ve buna bağlı olarak enerji üretim maliyetini minimize etmek amacıyla kısa dönem talep tahmini gerçekleştirmişlerdir. Bu çalışmada MLR, KNN, SVM, RF, XGBoost makine öğrenmesi metotları kullanılmıştır. Panama Devleti'nin Ocak 2015-Haziran 2020 dönemine ait saatlik tüketim değerleri veri seti olarak kullanılmıştır. Meteorolojik şartlar EarthData uydu verilerinden alınmıştır. Panama Eğitim Bakanlığı'ndan ilgili dönemdeki tatiller ve okul dönemi tarih bilgileri sağlanmıştır. Tahmin sonuçlarını iyileştirmek için haftalık yük gecikmeleri, haftalık hareketli ortalama ve haftalık standart sapma gibi özellikler hesaplanmıştır. 35 özelliğin analizi sonucunda 13 özelliğin tahmin sonucunu önemli derecede etkilediği, geçmiş yük verileri, günün saati, tatil ve hafta sonu olma durumunun en kritik özellikler olduğu gözlemlenmiştir. Veri setindeki bir ayın ilk 3 haftası eğitim, son haftası ise test seti olarak kullanılmıştır. Modellerin hiper parametreleri kayma penceresi yaklaşımı ve zaman tabanlı CV kullanılarak belirlenmiştir. Tahmin sonuçları incelendiğinde en düşük hata oranlarına sahip modelin XGBoost olduğu gözlemlenmiştir (Madrid vd., 2021).

Lahouar ve ark. (2015), RF algoritmasıyla 24 saatlik yük tahmini gerçekleştirerek kısa dönem talep tahmin modeli önerisinde bulunmuştur. Önerilen model ANN ve SVM ile karşılaştırılmıştır. 1 Ocak 2009 – 31 Ağustos 2014 dönemine ait yarım saatlik elektrik tüketim verileri Tunus Elektrik ve Gaz Şirketi'nden temin edilmiştir. Model girdileri otokorelasyon grafiği ve if-then kuralları kullanılarak seçilmiştir. Eksik veya hatalı veriler

önceki değerlerle değiştirilmiştir ve veriler 0-1 ölçeğine indirgenmiştir. RF parametreleri, ağaç sayısı varsayılan olarak 500 ve her düğümde bölünecek değişken sayısı 4 olarak belirlenmiştir. Test seti 1 haftadır ve sonrasında model, çevrimiçi öğrenme süreci uygulayarak bir gün önceki yük değerlerini de eğitim veri setine eklemektedir. Tahmin sonuçları incelendiğinde en iyi sonucu veren modelin ortalama %2.3 MAPE değeri ile RF olduğu gözlemlenmiştir (Lahouar vd., 2015).

Rüzgâr enerjisinin belirsizliği ve kontrol edilemezliği nedeniyle etkili bir rüzgâr gücü tahmini, rüzgâr çiftliklerinde yüksek öneme sahiptir. Li ve ark. (2018), veri madenciliği algoritmaları ve geliştirilmiş SVM'ye dayalı DM-WT-CS-SVM birleşik yöntemi önerisinde bulunmuştur. Rüzgâr gücü tahmini Kuzeybatı Çin'de bulunan bir rüzgâr çiftliği üzerinde uygulanmıştır ve 12 Ocak 2016 – 14 Mart 2016 dönemi saatlik veri toplanmıştır. Girdi değişkenleri olarak rüzgâr hızı ve yönü kullanılmıştır. Toplamda 1512 örnek analiz edilmiştir. Sıfır ya da negatif değerler gibi hatalı veriler rüzgâr hızı ve rüzgâr gücü arasındaki ilişkinin Fourier eğri uydurma yöntemi ile tekrar hesaplanması ile düzenlenmiştir. Eksik veriler bir önceki ve bir sonraki günün ortalama değeri ile düzenlenmiştir. Veriler, WT yöntemi ile sadeleştirilip ölçeklendirilmiştir. SVM'nin çekirdek fonksiyonu ve ceza faktörü, Cuckoo Search(CS) algoritması ile en uygun hale edilmiştir ve sırasıyla 45.87 ve 0.15 olarak belirlenmiştir. Önerilen yöntem Çift Sınırlı Sınıfa Özgü Destek Vektör Makinesi (DM-CS-SVM), Çift Sınırlı Ağırlıklı İkiz Destek Vektör Makinesi (DM-WT-SVM), Ağırlıklı İkiz Sınıfa Özgü Destek Vektör Makinesi (WT-CS-SVM) ve SVM modelleriyle karşılaştırılmıştır. DM-CS-SVM modeli normal SVM'den farklı olarak sınırını her bir sınıf özelinde belirler. DM-WT-SVM modeli hiper düzlemleri belirli bir ağırlıklandırma ile oluşturur. WT-CS-SVM iki düzlem kullanan ve her sınıfa özgü ayrı ağırlıklandırma gerçekleştiren bir modeldir. Önerilen model 0.16 MAE, 0.05 MSE ve %6.8 MAPE değerleri ile en iyi tahmin sonucunu vermiştir (Li vd., 2018).

Yaslan ve ark. (2017) uyarlanabilir, doğrusal olmayan ve durağan olmayan zaman-frekans veri analiz algoritması olan EMD ile SVR algoritmasını kullanarak birleşik bir model önerisinde bulunmaktadır. Doğuşkan Mod Fonksiyonu (IMF) değerleri veri setindeki her bir bileşenin, doğrusal ve durağan olmayan zaman serilerinin sıfır ortalamalı ve yarı periyodik sinyallere ayrılması sonucu sağlanır. Bu çalışmada EMD yöntemi, eğitim veri seti olarak kullanılan Polonya, Bosna-Hersek ve İngiltere'nin

elektrik tüketim değerlerinden, ortalama IMF değerlerini çıkararak gürültü temizleme amacıyla kullanılmıştır. IMF değerleri veri setindeki her bir bileşenin, doğrusal ve durağan olmayan zaman serileri sıfır ortalamalı ve yarı periyodik sinyallere ayrılması sonucu sağlanır. Gürültü temizleme işlemi sonucu yeni bir veri seti oluşturulmuştur. Gürültüden arındırılmış veri setinin, orijinal veri setinin eğilimini koruduğu gözlenmiştir. Polonya, Bosna-Hersek ve İngiltere elektrik tüketim veri setlerinin sırasıyla %13, %20, %15' i test verisi olarak kullanılmıştır. SVR algoritmasının parametreleri eğitim veri setinde GS kullanarak belirlenmiştir. Sonuçlar RBF kullanılarak, LIBSVM kütüphanesi ile elde edilmiştir. EMD tabanlı gürültüsüz SVR, Polonya, Bosna-Hersek ve İngiltere veri setleri için sırasıyla %2.82, %1.69 ve %3.21 MAPE değeri ile en iyi performansı göstermiştir. (Yaslan vd., 2017).

Takeda ve ark. (2016), Birleşik Kalman Filtresi(EnKF) ile daralma/çoklu regresyon tekniklerini birleştirerek yeni bir yaklaşım önerisinde bulunmuştur. Tokyo bölgesine ait Mart 2016-Şubat 2017 döneminin saatlik elektrik tüketim değerleri veri seti olarak kullanılmıştır. Veriler Tokyo Elektrik Güç Şirketi'nden sağlanmıştır. Meteorolojik veriler Japonya Meteoroloji Ajansı'nın web sitesinden temin edilmiştir. Algoritmaların eğitimi için sıcaklık, nem, rüzgâr hızı ve güneş radyasyonu meteorolojik parametreleri kullanılmıştır. Bu bağımsız değişkenlerin yanı sıra tatil ve haftanın günü gibi takvimsel bilgiler de göz önünde bulundurulmuştur. Veri setinin %80'i eğitim, %20'si ise test seti olarak kullanılmıştır. Önerilen modelde öncelikle, Durum-Uzay Modelleri (SSM) ile tüketim değerlerinin yük modeli oluşturulmuştur. Sonrasında EnKF ile tahmin gerçekleştirilmiştir ve daralma/çoklu regresyon modeli ile tahmin doğruluğu arttırılmıştır. EnKF doğrusal olmayan sistemlerin modellenmesini ve sıcaklık tepki oranı gibi ayrıntılı analitik bilgiler sağlamaktadır. Daralma/çoklu regresyon (SMR) ile aşırı uyum durumunun oluşması engellenmiştir. SMR, bağımsız değişken sayısıyla birlikte artan model katsayılarındaki kararsızlığı azaltmak ve aşırı öğrenmeyi engellemek için kullanılan bir yöntemdir (Takeda vd., 2016). Önerilen model standart EnKF, ARIMA, SARIMA, SVR ve MLR modelleri ile karşılaştırılmıştır. MLR, bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi modellemek için kullanılan bir regresyon yöntemidir (Yildiz vd., 2017). %3.5 MAPE değeri ile en iyi sonucu veren model, önerilen model olmuştur (Takeda vd., 2016).

Ahmadi ve ark. (2020), ağaç tabanlı öğrenme algoritmalarını kullanarak rüzgâr gücü tahmini gerçekleştirmiştir. Çalışmada DT, RF, XGBoost ve Bağlama Tekniği (Bootstrap) yöntemleri kullanılmıştır. Bagging yöntemi, eğitim veri setinden rastgele veri alt kümeleri oluşturmayı içerir. Her model rastgele belirlenmiş alt veri kümeleri ile eğitilir, her modelin sonuçlarının ortalaması alınarak çıktı elde edilir (Deng vd., 2022). Rüzgâr hızı verileri İran'nın Ghadamgah ve Kahf bölgelerindeki rüzgâr çiftliklerinden temin edilmiştir. Veri seti 18 aylık bir dönemden oluşmaktadır ve veri girişinin tahmin doğruluğu üzerindeki etkisini incelemek için 4 farklı senaryo oluşturulmuştur. İlk senaryoda 10 dakika zaman aralığına sahip 40 metre yükseklikte ölçülmüş Ghadamgah veri seti, rüzgâr hızı ortalama ve standart sapma değerleriyle 1 saatlik, 12 saatlik ve 24 saatlik zaman aralıklarına tamamlanmıştır. Algoritmaların eğitimi için veri setinin %70'i kullanılmıştır. Performans katsayıları, uzun zaman aralıklarının veri setinin temsil gücünü azalttığından daha düşük performans gösterdiğini ortaya çıkarmıştır. En iyi sonucu veren algoritma XGBoost olmuştur. İkinci senaryoda, ekstrapolasyon işleminin tahmin sonuçlarına etkisi incelenmiştir.

Ekstrapolasyon, bilinen veri noktalarının ayrık kümesi dışında yeni veri noktaları oluşturma işlemidir (Karaca vd., 2016). Rüzgâr hızı ölçümü için kullanılan anemometreler genellikle 10m'lik yükseklikten veri toplar. Rüzgâr tribünlerinin yüksekliğindeki rüzgâr hızı, 10m'deki ve 30m'deki rüzgâr hızının ekstrapole edilmesi sonucu sağlanmıştır. 10 dakikalık örnek aralığı ile ekstrapole değerlere sahip olan yeni Ghadamgah veri seti oluşturulmuştur. XGBoost en iyi performansı gösteren algoritma olmuştur. Üçüncü senaryo algoritmaların farklı yüksekliklerdeki tahmin tutarlılığını araştırmıştır. 10 ve 30 metre yükseklikteki rüzgâr hızları göz önünde bulundurulmuştur. En iyi sonucu 6.34 MAE değeri ile XGBoost vermiştir. Son senaryo, algoritmaların yeni bir veri seti ile nasıl sonuç vereceğini gözlemlemek için oluşturulmuştur. Kahf rüzgâr çiftliğinden, saatlik örnek aralıkları ile 40 metredeki rüzgâr hızı verileri alınmıştır. Veri seti önceden eğitilmiş algoritmalara test seti olarak verilmiştir. En iyi tahmin sonucumu XGBoost algoritması vermiştir (Ahmadi vd., 2020).

Cerne ve ark. (2018), Takagi-Sugeno (TS) bulanık mantık modelini kullanarak kısa dönem yük tahmini gerçekleştirmiştir. 2010-2012 arasındaki döneme ait saatlik yük verileri Slovenya'nın güney batısındaki bir elektrik dağıtım şirketinden alınmıştır. Saatlik hava durumu verileri bölgedeki meteoroloji istasyonundan sağlanmıştır. Modelin eğitimi

için sıcaklık, güneş radyasyonu, yağış durumu, rüzgâr hızı, saat, haftanın günü, ayın günü, ay, yıl ve tatil günleri gibi değişkenler kullanılmıştır. Yağış durumunun tahmin sonuçlarına bir etkisinin olmadığı gözlemlenmiştir. TS modeli problem uzayının daha küçük birimlere bölmektedir. Her problem birimini ayrı ayrı çözümleyerek sonuçları birleştirir ve tahmin doğruluğunu arttırmaktadır. rGK algoritması ayrılan birimleri tanımlı kümelere böler, her küme bir ilk örnek ve bulanık kovaryans matrisi ile tanımlanır. rWLMS algoritması, ayrıştırılan kümelerin parametre uyarlamasını yaparak tahmin doğruluğunu arttırır. Bu çalışmada, rGK ve rWLMS algoritmalarını birlikte uygun bir şekilde kullanılması için birleşim üyelik fonksiyonu önerilmiştir. Birleşim üyelik fonksiyonu, mahalanobis mesafesi ve birimlerdeki hata oranlarını birleştirerek dışlayıcıların etkisini azaltan bir üyelik derecesi oluşturur. Modelin eğitimi için 2010-2011 verileri kullanılmıştır. Test verisi için 2012 verileri kullanılmıştır. Önerilen yöntemin MAPE değeri %3.68 olarak hesaplanmıştır (Cerne vd., 2018).

Yu ve ark. (2016), seyrek kodlama yöntemini uygulayarak kısa ve orta dönem talep tahmini gerçekleştirmiştir. Akıllı şebekeler sayesinde Tennessee'nin Chattanooga şehrindeki 5000 hanenin Eylül 2011-Ağustos 2013 dönemine ait 15 dakikalık aralıklarla elektrik tüketim değerleri toplanmıştır. Yük sinyalleri, modelin tüketim desenine odaklanabilmesi için öklidyen normu 1 olacak şekilde normalize edilmiştir. Veriler saatlik tüketim verilerine dönüştürülmüştür. Aynı döneme ait saatlik sıcaklık verileri Ulusal Okyanus ve Atmosfer İdaresi'nden temin edilmiştir. Veriler modellerin daha etkili bir şekilde tahmin yapabilmesi adına 168 saatlik parçalara bölünmüştür. Eylül 2011-Ağustos 2012 dönemi verileri eğitim seti olarak kullanılmıştır. Eylül 2012-Ağustos 2013 dönemi verileri ise test seti olarak kullanılmıştır. Yu ve ark. (2016), bir sonraki günün elektrik talep tahmini gerçekleştirirken geçmiş 7 günün verilerini kullanmıştır. Bir sonraki haftanın elektrik talep tahminini gerçekleştirmek için geçmiş 28 günün verileri kullanılmıştır. Seyrek kodlama yöntemi RR, ARIMA ve Holt-Winters yöntemi ile uygulanmıştır. Holt-Winters yöntemi üçlü üstel düzeltme ile zaman serisindeki eğilimleri ve mevsimsel faktörleri modelleyerek tek değişkenli veri setlerini tahmin etmek için kullanılır (Sudheer vd., 2014). En iyi tahmin sonucunu veren model %24 MAPE değeriyle seyrek kodlama ile RR modeli olmuştur. Bireysel hane seviyesindeki yük tüketiminde daha çok dalgalanma ve rastgelelik olduğundan MAPE değerleri yüksek çıkmıştır (Yu vd., 2016).

Zor ve ark. (2017), ANN, SVM, Uyarsamalı Sinirsel-Bulanık Çıkarım Sistemi(ANFIS) yapay zekâ algoritmalarını kullanarak kısa dönem talep tahmini gerçekleştirmiştir. ANFIS modeli bulanıklaştırma, if-then kuralları, çıkarsama ve durulaştırma adımlarını içeren, mantık ve ANN birleşimi bir modeldir (Çevik vd., 2014). Tüketim verileri, New England ve ABD bölgelerinden saatlik, Avustralya bölgesinden yarım saatlik aralıklarla sağlanmıştır. Veri setinin %70 ile eğitim, %30'u ile test seti oluşturulmuştur. ANN'in gizli katman sayısı 1 olarak belirlenmiştir. Aktivasyon fonksiyonu olarak sigmoid ve RBF kullanılmıştır. SVM için lineer, tanh, polinomal ve gauss çekirdek fonksiyonları kullanılmıştır. ANFIS birleşik modeli giriş katmanı, kural katmanı, normalizasyon katmanı, sonuç katmanı ve çıkış katmanı olmak üzere 5 katmandan oluşmaktadır. Izgara Bölütleme (GP), bulanık c-ortalama kümeleme ve çıkarımsal kümeleme ile girdi parametreleri bulanık kümelere ayrılmıştır. ANFIS birleşik modeli LS ve Gradyan Düşürme (GD) yöntemi kullanılarak eğitilmiştir. GD hatayı düşürmek için bir maliyet fonksiyonunun türevini alarak ağırlıkların ne kadar değiştirilmesi gerektiğini belirleyerek döngüsel olarak güncelleyen yöntemdir (Jihad vd., 2018). Saatlik tahmin sonuçları karşılaştırılan ANN ve SVM'nin MAPE değerleri sırasıyla %1.950 ve %1.790'dır. 30 dakikalık tahmin gerçekleştirilen ANFIS birleşik modeli %2.876 MAPE değerine sahiptir. En iyi tahmin sonucunu veren model SVM olmuştur (Zor vd., 2017).

Saber ve Alam (2017), MLR algoritması kullanarak kısa dönem talep tahmini gerçekleştirmiştir. Farklı bölgelerdeki akıllı ve yarı akıllı sayaçlardan 3 yıllık elektrik tüketim verisi toplanmıştır. İlgili döneme ait sıcaklık ve nem hava durumu verileri, tüketim verilerinin toplandığı bölgelerdeki hava durumu servislerinden temin edilmiştir. Çok çekirdekli paralel işleme yöntemi ile regresyon katsayıları belirlenmiştir ve büyük veri setinin matris haline getirilmesi hızlandırılmıştır. Haftalık dönem talep tahmini gerçekleştirilmiştir ve MLR algoritmasının MAPE değeri %3.99 olarak hesaplanmıştır (Saber & Alam, 2017).

Valgaev ve ark. (2016), KNN modeli ile kısa dönem talep tahmini gerçekleştirmiştir. Tüketim verileri İrlanda Enerji Düzenleme Komisyonu tarafından akıllı sayaçlar ile sağlanmıştır. Veri seti, 6000'den fazla kullanıcının bilgisini içermektedir ve Temmuz 2009-Aralık 2010 arasındaki dönemi kapsamaktadır. Tüketim verileri 30 dakikalık aralıklarla alınmıştır. KNN modelinin benzerlik ölçümü için permütasyonlu Lp

normu kullanılmıştır. Modelin komşu sayısının belirlenmesi, maksimum permütasyon mesafesi ve kullanılan Lp normunun belirlenmesi gibi adımları içeren parametre belirleme işlemleri CV yöntemi ile gerçekleştirilmiştir. 10-100 rastgele seçilmiş tekil konut ve ticari tüketicilerin örnekleri ile 4 benzetim oluşturulmuştur. 5-50-100 rastgele seçilmiş konut ve ticari tüketicilerin toplam yükü ile 6 benzetim oluşturulmuştur. Düşük voltajlı veya hiç tüketim olmayan dönemlerde MAPE değerinin sonsuza gitmesi gibi hataların oluşmasından dolayı doğruluk ölçümü her tüketici örneği için Bireysel Yük Profili (ILP) ve KNN modellerinin 365 günlük hatalarının medyanı hesaplanarak ve Lp normu kullanılarak ayarlanmıştır. ILP, tarihsel verileri inceleyerek tahmin edilecek döneme ait değişkenlere en çok benzeyen verileri çıktı olarak veren bir modeldir. KNN modeli benzetimlerin hepsinde ILP tekniğinden daha iyi sonuç vermiştir (Valgaev, Kupzog, & Schmeck, 2016).

Kong ve ark. (2017), LSTM, BPNN, KNN, Aşırı Öğrenme Makinesi(ELM) ve Birleşik Tahmin Çerçevesi(IS-HF) yaklaşımlarını kullanarak kısa dönem talep tahmini gerçekleştirmiştir. ELM yöntemi, bağlantı ağırlıklarını rastgele bir şekilde atayarak tek adımlık bir hesaplama ile en uygun çıkış ağırlıklarını elde eden yeni bir Tek Gizli Katmanlı İleri Beslemeli Yapay Sinir Ağı (SLFN)'dır (Li, Qian, & He, 2020). IS-HF, WT ve Bayesian Sinir Ağı (BNN) bileşiminden oluşan bir tahmin yöntemidir (Kong vd., 2017). Tüketim verileri, Avustralya'nın Yeni Güney Galler bölgesinden Smart Grid Smart City(SGSC) projesi kapsamında 69 farklı tüketiciye aittir. Veri seti 30 dakikalık aralıklarla toplanmıştır ve 1 Haziran 2013 ile 31 Ağustos 2013 tarihleri arasındaki dönemi içermektedir. Veriler eğitim, doğrulama ve test setleri olarak düzenlenmiştir. 1 Haziran 2013-5 Ağustos 2013 dönemi eğitim, 6 Ağustos 2013-22 Ağustos 2013 dönemi doğrulama ve 23 Ağustos-31 Ağustos dönemi test veri seti için kullanılmıştır. ELM modelinin gizli nöron sayısı en uygun hale getirilmiştir. LSTM ve BPNN modelleri için 2 gizli katman belirlenmiştir ve her katmanda 20 gizli düğüm olacak şekilde oluşturulmuştur. KNN algoritması için komşu sayısı 20 olarak belirlenmiştir. Tüm modeller için döngü sayısı 150 olarak belirlenmiştir. Tahmin sonuçları incelendiğinde bireysel konutların talep tahmini için en iyi performansı %44.06 MAPE değeri ile LSTM modeli vermiştir (Kong vd., 2017).

Khan ve ark. (2019), Geliştirilmiş-KNN (E-KNN) modeli önerisinde bulunarak kısa dönem yük ve fiyat tahmini gerçekleştirmiştir. Ocak 2016-Şubat 2016 arası saatlik

tüketim verileri ve fiyatlandırmaları New York Independent System Operator (NYISO) kuruluşundan temin edilmiştir. Hava şartlarının tüketim üzerindeki etkisi dikkate alınmamıştır. Veri seti, bazı yöntemlerin kayan nokta değeri ile işlem yapamamasından dolayı aynı ölçeğe indirgenmiştir. Özyinelemeli Özellik Eleme (RFE) ve DTR yöntemleriyle yük ve fiyat için önemli olan özellikler belirlenmiştir. RFE yöntemi bağımsız değişkenlerin belirleyiciliğini, bağımlı değişkenle olan ilişkisini dikkate alarak önem sıralaması yapar (Abedinia, Amjady, & Zareipour, 2016). Model, her adımda düşük önem seviyesine sahip değişken kaldırılarak kullanıma en uygun hale getirilmiştir. Tekil Değer Ayrıştırması (SVD) yöntemiyle yük ve fiyat ile ilgisi az olan özellikler çıkarılarak özellik boyutları azaltılmıştır. SVD, orijinal veri kümesindeki değişkenliğin çoğunu yakalayan ve birbiriyle ilişkisiz değişkenlerden oluşan bir veri seti kümesine indirgemesini sağlayan yöntemdir (Taylor, 2012). Veri setinin %70'i eğitim ve %30'u test veri seti olarak kullanılmıştır. Denemeler yapılarak uygun K değeri ve mesafe fonksiyonu belirlenmiştir. Önerilen tahmin modeli SVR ve normal KNN modelleriyle karşılaştırılmıştır. SVR, KNN, E-KNN modellerinin fiyatlandırma tahmin sonuçlarının MAPE değeri sırasıyla %68.81, %10.50, %2.37 ve talep tahmin sonuçlarının MAPE değeri sırasıyla %68.72, %16.48, %4.52 olarak hesaplanmıştır. Önerilen E-KNN modeli en iyi tahmin sonuçlarını vermiştir (Khan, Khan, Noshad, Javaid, & Javaid, 2019).

Dehalwar ve ark. (2016), ANN ve BRT yöntemlerini kullanılarak kısa dönem talep tahmini gerçekleştirmiştir. 2006-2010 dönemine ait Sydney/NSW bölgesinin saatlik tüketim verileri Avustralya Enerji Piyasası İşletmecisi (AEMO) tarafından sağlanmıştır. İlgili bölgeye ve döneme ait hava durumu verileri Avustralya Meteoroloji Bürosu (BOM) tarafından sağlanmıştır. ANN modeli için gizli katmanındaki nöron sayısı 20 olarak belirlenmiştir. Aktivasyon fonksiyonu olarak hiperbolik tanjant sigmoid fonksiyonu kullanılmaktadır. ANN modeli için veri seti, %70 eğitim, %15 doğrulama ve %15 test veri seti olmak üzere 3 gruba ayrılmıştır. BRT modeli için ağaç sayısı 20 ve minimum yaprak boyutu 40 olarak belirlenmiştir. Sıcaklık, nem, önceki günün 24 saatlik ortalama yükü, önceki günün 24 saatlik gecikmeli yükü, önceki 168 saatlik ortalama yük ve haftanın günü bilgilerine dayanarak bir sonraki günün yük talebini tahmin eden Dehalwar ve ark. (2016), ANN ve BRT modelleri için sırasıyla %1.90 ve %1.54 MAPE değerlerini elde etmiştir. BRT modeli daha iyi bir performans sergilemiş olsa da ANN modeline göre daha uzun sürede sonuca ulaşmıştır (Dehalwar, Kalam, Kolhe, & Zayegh, 2016).

Moon ve ark. (2018), iki aşamalı öngörü analitiği kullanarak kısa dönem talep tahminini gerçekleştirmiştir. Tüketim verileri Kore, Seul'daki özel bir üniversitenin 32 ana bina ve 16 yurt binasından temin edilmiştir. 2012-2016 arasındaki döneme ait yük tüketimleri 15 dakikalık, saatlik ve günlük olarak kaydedilmiştir. Hava durumu verileri Kore Meteoroloji Dairesi'nden sağlanmıştır ve günlük ortalama sıcaklık olan 12.5 santigrat dereceye sabitlenmiştir. Veri setinin %75'i eğitim seti, %25'i test seti olarak kullanılmıştır. İlk aşamada Basit Hareketli Ortalama (SMA) ile tüketim verilerinin deseni oluşturulmuştur. SMA, zaman serisi değerlerini, mevcut ve önceki değerlerin ortalaması olarak ifade eden bir yöntemdir (Haripuddin, Yahya, Zulhajji, & Yantahin, 2022). Daha sonrasında tüketim desenlerini dikkate alarak RF algoritması ile tahmin gerçekleştirilmiştir. Önerilen modelin tahmin doğruluğu, Zaman Serisi Çapraz Doğrulama (TSCV) yöntemiyle farklı zaman aralıklarında kontrol edilmiştir ve güvenilirliği arttırılmıştır. Önerilen model ANN ve SVR algoritmalarıyla karşılaştırılmıştır. Akademik binalar için %1.84 ve yurt binaları için %1.31 MAPE değerleriyle en iyi performansı önerilen model sağlamıştır (Moon, Kim, Son, & Hwang, 2018).

Rafiei ve ark. (2018), kısa dönem talep ve fiyat tahmini gerçekleştirmek amacıyla yeni bir model önerisinde bulunmuştur. 2014 yılına ait tüketim verileri Ontario ve Avustralya elektrik piyasalarından saatlik olarak temin edilmiştir. Veri setinin %70'i eğitim, %15'i doğrulama ve %15'i test seti olarak kullanılmıştır. Veri seti, Daubechies dalgacık fonksiyonu ile yaklaşık ve detay serilere ayrılmıştır. Geleneksel yöntemlere göre hız avantajı olduğundan, işlenmiş verilerle İyileştirilmiş Dalgacık Sinir Ağları (IWNN) modelini eğitmek için Genelleştirilmiş Ekstrem Öğrenme Makinesi (GELM) yöntemi kullanılmıştır. GELM yöntemi, gizli katmanın sapma değerlerini ve giriş katsayılarını rastgele belirleyerek çıkış matrisini genelleştirilmiş ters dönüşüm ile sağlar. Tahmin aralıklarını ve modelin belirsizliklerini görebilmek için Bootstrap yöntemi kullanılmıştır. Tahmin aralıklarının güvenilirliği Ortalama Kapsama Hatası (ACE) ile değerlendirilmiştir. Önerilen yöntem, Geri Yayılım (BP) + Bootstrap, Dolphin Algoritması + Bootstrap, RELM + Bootstrap ve Quantile Regresyon(QR) ile karşılaştırılmıştır. BP yöntemi ile çıkış katmanındaki sonucun hata değeri hesaplanarak geri yayılım gerçekleştirilir ve hatayı en aza indirgemek amacıyla nöron ağırlıkları tekrardan ayarlanır (Reddy vd., 2014). Dolphin algoritması, yunusların sürü mantığından

esinlenerek geliştirilmiş, başlangıç arama, çağrı, alıcı ve avlanma aşamalarından oluşan bir uygun hale getirme aracıdır (Ahmed, Allu, & Jirjees, 2022). RELM, geri besleme bağlantıları giriş katmanına bağlanan bağlam nöronları aracılığıyla gecikmeli değerleri kullanan ve ağa tekrardan girdi olarak sağlayan yöntemdir (Ertugrul vd., 2015). QR, açıklayıcı değişkenlerin değerlerine bağlı olarak koşullu bir dağılımın yüzdeliklerini tahmin etmek için kullanılan istatiksel bir yöntemdir (He, Liu, Li, Wang, & Lu, 2017). En iyi performansı sağlayan model ortalama %1 MAPE değeriyle önerilen model olmuştur (Rafiei, Niknam, Aghaei, Shafie-Khah, & Catalao, 2018).

Uydur ve ark. (2023), Trakya Üniversitesinin kısa dönem talep tahminini gerçekleştirmiştir. Tüketim verileri Trakya Üniversitesi'nin TREPAŞ-OSOSOS sisteminden ve meteoblue veri tabanından temin edilmiştir. 2022 yılına ait olan veriler saatlik ölçülmüştür. 2022 yılına ait günlük ve saatlik sıcaklık, nem verileri Meteoroloji Genel Müdürlüğü'nden sağlanmıştır. 2022 yılının ilk 8 ayı eğitim ve son 4 ayı test veri seti olarak kullanılmıştır. 5 farklı tahmin modeli oluşturulmuştur. Tahmin ve analiz için DT, LS ve ANN yöntemleri kullanılmıştır. DT yöntemi Python programı ile oluşturulmuştur. Ağaç sayısı 500 olarak belirlenmiştir. LS yöntemi MATLAB programıyla oluşturulmuştur. Polinom derecesi 6 olarak belirlenmiştir. ANN, MATLAP programı kullanılarak Levenberg-Marquardt fonksiyonu ile eğitilmiştir. Gizli katmandaki nöron sayıları 3 ve 12 olarak belirlenmiştir. 5 tahmin modelinde de en iyi performansı sergileyen yöntem ANN olmuştur. Model 1, 2, 3, 4 ve 5 için ANN algoritmasının MAPE değerleri sırasıyla %4.96, %5.64, %4.88, %4.21 ve %4.08'dir. (Uydur, Akin, Dursun, & Keskin, 2023).

Dietrich ve ark. (2020) bu çalışmasında, endüstriyel enerji esnekliğini arttırmayı amaçlayarak çok kısa vadeli talep tahmini gerçekleştirmiştir. Tüketim verileri Darmstadt Teknik Üniversitesi ETA araştırma fabrikasındaki 2 adet makineden toplanmıştır. PLC ve enerji sayaçları ile saniyelik olarak elektrik yükü, kontrol sinyalleri, arıza sinyalleri, yardımcı ünitelerin durumu, üretim komutları ve sıcaklık, basınç, hız gibi fiziksel özellikler okunmuştur. Eksik veriler medyan işlemi ile tamamlanmıştır ve StandartScaler ile ayrık değerler işlenip sayısal özellikler standartlaştırılmıştır. StandartScaler, veri setindeki özeliklerin ortalamasını 0, standart sapmasını 1 olacak şekilde dönüştüren bir ölçeklendirme yöntemidir. 2 makine için 2 farklı işlem yaptırılarak gözlem aralığı genişletilmiştir. Analiz ve talep tahmini KNN, DT, RF ve ANN yöntemleriyle

gerçekleştirilmiştir. KNN için K değeri 5 olarak belirlenmiştir. DT için maksimum derinlik 10 olarak belirlenmiştir. RF için maksimum derinlik 10 ve ağaç sayısı 100 olarak belirlenmiştir. ANN için gizli katman sayısı 1 olarak belirlenmiştir ve ReLU aktivasyon fonksiyonu kullanılmıştır. Gizli katman düğüm sayısı 90 olarak belirlenmiştir. En verimli tahmin ufkunun 100 saniye olduğu gözlemlenmiştir. 4 tahmin modeli arasında en iyi performansı sergileyen algoritmanın 0.64 R^2 değeri ile ANN olduğu sonucuna varılmıştır (Dietrich, Walther, Weigold, & Abele, 2020).

Rafi ve ark. (2021), yeni bir birleşik yöntem önerisinde bulunarak kısa dönem talep tahmini gerçekleştirmiştir. Ocak 2014-Aralık 2019 dönemine ait olan 30 dakikalık tüketim verileri Bangledeş Güç Sistemi'nden temin edilmiştir. Eksik veriler önceki tüketim karakteristiğine uyumlu bir şekilde tamamlanmıştır. Veri setinin Ocak 2014-Aralık 2018'e ait dönemi eğitim, Ocak 2019-Aralık 2019'a ait dönemi test veri seti için kullanılmıştır. Her örnek için 7 zaman adımı ve toplam günlük güç tüketimi 1 özellik sayılarak çok adımlı zaman serileri oluşturulmuştur. Önerilen birleşik yöntem, Konvolüsyonel Sinir Ağı (CNN) ve LSTM yöntemlerini içermektedir. CNN modeli, çok kanallı ve küp verilerdeki doğrusal olmayan davranışları algılama yeteneğine sahip bir derin öğrenme modelidir (Imani, 2021). CNN algoritması, giriş serilerini okuyarak özellik ayrımı yapar ve bu özellikleri güçlendirir. Elde edilen sonuçlar LSTM algoritması için giriş verisi olarak kullanılır. Birleşik model, Python programı kullanılarak Keras kütüphanesi ile oluşturulmuştur. CNN algoritması için bir boyutlu konvolüsyon ve kernel boyutu 3 olan 64 adet filtre kullanılmıştır. Aktivasyon fonksiyonu olarak ReLU seçilmiştir. LSTM için gizli katman nöron sayısı 200 olarak belirlenmiştir. Döngü sayısı 20 ve yığın boyutu 16 olarak belirlenmiştir. Önerilen birleşik model normal LSTM, Radyal Temelli Fonksiyon Ağı (RBFN) ve XGBoost algoritmalarıyla karşılaştırılmıştır. En iyi tahmin sonucunu veren modelin %3.07 MAPE değeri ile CNN-LSTM birleşik modeli olduğu gözlemlenmiştir (Rafi, Masood, Deeba, & Hossain, 2021).

Tarmanini ve ark. (2023), ARIMA ve ANN modellerini karşılaştırmayı amaçlayarak kısa dönem talep tahmini gerçekleştirmiştir. Tüketim verileri 2009-2010 dönemi arasında 709 bireysel haneden saatlik olarak sağlanmıştır. Veri seti toplamda 12.865 saatlik veri içermektedir. Veriler İrlanda Sosyal Bilim Veri Arşivi'nden temin edilmiştir. Verilerin %90'ı eğitim, %10'u test veri seti oluşturmak için kullanılmıştır. ANN'in gizli katman sayısı 1 olarak belirlenmiştir ve 250 döngü uygulanmıştır. ARIMA

modelinin MAPE değeri %2.61 iken ANN modeli %1.80 MAPE değeri ile daha iyi performans sergilemiştir (Tarmanini, Sarma, Gezegin, & Ozgonenel, 2023).

Ibrahim ve ark. (2022), akıllı şebekelerde makina öğrenmesi ve derin öğrenme yöntemleri ile kısa dönem talep tahmini gerçekleştirmeyi amaçlamıştır. Veriler, Panama Ulusal Dağıtım Merkezi ve Panama Kanal Otoritesi tarafından sağlanmıştır. Bu veriler, 2016-2019 dönemine ait saatlik yük tüketim verileri, sıcaklık ve bağıl nem verilerinden oluşmaktadır. Veriler, tüm giriş parametrelerinin aynı aralıkta olması için 0-1 arasında ölçeklendirilmiştir. 1 Ocak 2016-25 Ocak 2019 dönemi verileri eğitim, 6 Ocak 2019-25 Ocak 2019 dönemi verileri test veri seti oluşturmak için kullanılmıştır. Sıcaklık, bağıl nem, ay, haftanın günü, günün saati, tatiller, önceki gün aynı saatteki talep ve önceki 24 saatlik ortalama talep gibi özellikler modellerde kullanılmak üzere belirlenmiştir. Analiz ve talep tahmini için SVR, XGBoost, AdaBoost, RF, LightGBM, Derin Öğrenme Regresyon (DLR) modeli, Bi-LSTM ve Kapılı Tekrarlayan Birim (GRU) modelleri kullanılmıştır. LightGBM modeli, Microsoft Research tarafından 2017 yılında önerilmiş ve karar ağaçlarına dayanan bir gradyan arttırma çerçevesidir (Wang vd., 2020). GRU modeli LSTM ağlarına benzer şekilde tekrar eden bir sinir ağı çeşididir (Bacanin vd., 2023). DVR için RBF fonksiyonu kullanılmıştır ve C değeri 0.1 olarak belirlenmiştir. XGBoost için maksimum derinlik 3 ve ağaç sayısı 100 olarak belirlenmiştir. AdaBoost için ağaç sayısı 100 olarak belirlenmiştir. RF için ağaç sayısı 100 olarak belirlenmiştir. LightGBM için ağaç sayısı 100 ve yaprak sayısı 31 olarak belirlenmiştir. DLR için katman sayısı 3 ve gizli katmandaki düğüm sayısı 95 olarak belirlenmiştir. Aktivasyon fonksiyonu ReLU ve döngü sayısı 500 olarak belirlenmiştir. Bi-LSTM modeli için 3 gizli katman ve 70 düğüm oluşturulmuştur. Aktivasyon fonksiyonu ReLU ve döngü sayısı 500 olarak belirlenmiştir. GRU modeli için katman sayısı 3 ve gizli düğüm sayısı 80 olarak belirlenmiştir. Aktivasyon fonksiyonu ReLU ve döngü sayısı 500'dür. Karşılaştırılan modeller arasında en iyi tahmin sonucunu %2.9 MAPE değeri ile DLR gerçekleştirmiştir (Ibrahim, Rabelo, Gutierrez-Franco, & Clavijo-Buritica, 2022). Literatürdeki çalışmalarda elde edilen hata değerleri Çizelge 2.1'deki gibidir. Çizelge 2.1'de çalışma özelindeki en iyi hata oranı belirtilmektedir.

Çizelge 2.1. Literatürdeki Çalışmalarda Elde Edilen Hata Değerleri

Makale/ Tez	**Yöntem**	**MAPE Değeri**	**RMSE Değeri**	**R^2 Değeri**
(Tutu, 2017)	LS	%0.1	-	-
(Ahmad vd., 2019)	DALF, Bi-Level, MI-ANN	%1.24	-	-
(Takeda vd., 2016)	EnKF-SMR, EnKF, ARIMA, SARIMA, DVR, MLR	%3.5	-	-
(Ahmadi vd., 2020)	XGBoost, DT, RF, Bagging	%0.98	-	-
(Cerne vd., 2018)	rGK-rWLMS-TSBM	%3.68	-	-
(Yu vd., 2016)	SCRR, RR, ARIMA, Holt-Winters	%24	-	-
(Zor vd., 2017)	ANN, SVM, ANFIS	%1.790	-	-
(Saber vd., 2017)	MLR	%3.99		-
(Valgaev vd., 2016)	KNN		-	-
(Kong vd., 2017)	LSTM, BPNN, KNN, ELM, IS-HF	%44.06	-	-
(Khan vd., 2019)	E-KNN, DVR, KNN	%4.52	-	-
(Dehalwar vd., 2016)	BRT, ANN	%1.54	-	-

(Moon vd., 2018)	SMA-RF, ANN, DVR	%1.31	-	-
(Rafiei vd., 2018)	IWNN-GELM-ACE, BP-Bootstrapping, Dolphin algoritması-Bootstrapping, RELM-Bootstrapping, Quantile regresyon	%1	-	-
(Uydur vd., 2023)	ANN, LS, DT	%4.08	-	-
(Dietrich vd., 2020)	ANN, KNN, DT, RO	-	-	0.64
(Rafi vd., 2021)	CNN-LSTM, RBFN, XGBoost	%3.07	-	-
(Tarmanini vd., 2023)	ANN, ARIMA	%1.80	-	-
(Ibrahim vd., 2022)	DLR, SVM, XGBoost, AdaBoost, RO, LightGBM, Bi-LSTM, GRU	%2.9	-	-

BÖLÜM 3

TALEP TAHMİNİ

3.1. Giriş

Talep tahmini, öngörüsü yapılmak istenen zaman aralığına göre farklı parametreleri de analize dâhil ederek, geçmiş ve güncel yük göz önünde bulundurularak beklenen elektrik talebinin toplu düzeylerde tahmin edilmesini ifade eder (Sadeghi Keyno, Ghaderi, Azade, & Razmi, 2009). İşlemlerle, planlamayla ve yatırımlarla başa çıkmada talep tahmininin sağladığı faydalar piyasa teklifleri, kontrol edilebilir yük planlaması, üretim, yakıt alımları, yük geçişleri ve altyapı bakım programı dâhil olmak üzere birçok konuda kullanımının etkisini göstermiştir (Azeem, Ismail, Jameel, & Harindran, 2021).

3.2. Talep Tahmini Çeşitleri

Talep tahmini süreci, tahmin süresine göre 4 farklı kategoride gruplanabilir (Wen, Zhou, Yang, & Li, 2018). Bunlar:

- Çok Kısa Dönem Talep Tahmini (ÇKDTT)
- Kısa Dönem Talep Tahmini (KDTT)
- Orta Dönem Talep Tahmini (ODTT)
- Uzun Dönem Talep Tahmini (UDTT)

3.2.1. Çok Kısa Dönem Talep Tahmini (ÇKDTT)

Birkaç dakika ile 1 saat aralığında tahminlerdir. Daha çok gerçek zamanlı kontroller ve ani talep değişikliklerine hızlı düzenlemeler gerçekleştirmek ve operasyonel karar destek sistemleri için kullanır (Kuster, Rezgui, & Mourshed, 2017).

3.2.2. Kısa Dönem Talep Tahmini (KDTT)

1 saat ile 1 hafta aralığındaki tahminlerdir. Daha çok elektrik üretimi, iletiminin plan düzenlenmesinde ve günlük uygulamaların kontrolünde tercih edilir (Kuster vd., 2017).,

3.2.3. Orta Dönem Talep Tahmini (ODTT)

1 hafta ile 1 yıl aralığındaki tahminlerdir. Hizmet sağlayıcılarının kar değerlendirmesi, enerji ticareti, bakım planlaması gibi konular için tercih edilir (Kuster vd., 2017).

3.2.4. Uzun Dönem Talep Tahmini (UDTT)

1 yıl ile 20 yıl aralığındaki tahminlerdir. Stratejik planlama ve güç arzının tahmini için tercih edilir. Uzun dönem talep tahmini enerji politikalarının oluşturulmasında ve yatırım planlamalarında büyük öneme sahiptir (Kuster vd., 2017).

3.3. Talep Tahminini Etkileyen Faktörler

Talep tahmini, sürecin hassasiyetini etkileyen birçok faktöre bağlıdır. Araştırmacıların, analiz edilen sistemden doğru bir sonuç elde etmesi için bağımsız faktörleri dikkatli bir şekilde seçmesi gerekir (Mamun vd., 2020). Talep tahmini için dikkate alınması gereken faktörler şu şekilde sıralanabilir (Chapagain, Kittipiyakul, & Kulthanavit, 2020):

- Meteorolojik Faktörler
- Zamansal veya Takvimsel Faktörler
- Tesadüfi Faktörler
- Tüketici Kaynaklı Faktörler

3.3.1. Meteorolojik Faktörler

Sıcaklık, yük tüketimini etkileyen en önemli meteorolojik faktördür. Hava sıcaklığı, elektrikle çalışan soğutma ve ısıtma sistemlerindeki talep dalgalanmalarıyla doğrudan bağlantılıdır. Bunun yanı sıra nem de sıcaklığın yarattığı etkiyi güçlendiren bir faktördür (Bouktif vd., 2018). Rüzgâr hızı, güneş radyasyonu ve güneşlenme süreleri gibi meteorolojik faktörlere, temiz enerji kaynaklarının kullanıldığı bölgelerde özellikle dikkat edilmelidir (Wang, Lu, Wang, & Zhang, 2022).

3.3.2. Zamansal veya Takvimsel Faktörler

Zamansal ve takvimsel faktörlerin yük tüketimi üzerinde önemli bir etkisi vardır. Günün saati, haftanın günü, bulunulan mevsime bağlı olarak yaz saati uygulaması ve güneşlenme süresi soğutma, ısıtma ve aydınlatma sistemlerinin yük tüketimini direkt olarak etkilemektedir. Bununla birlikte tatiller ve bayramlar tüketim karakteristiği üzerinde büyük öneme sahiptir (Bouktif vd., 2018).

3.3.3. Tesadüfi Faktörler

Tesadüfi faktörler öngörülemeyen durumlardan oluşmaktadır. Fırtınalar veya soğuk hava dalgaları gibi ani ve aşırı hava durumu değişiklikleri yük talebinde kısa süreli dalgalanmalara neden olabilir. Bunun yanı sıra endüstriyel bölgelerdeki ya da elektrik şebekelerindeki arıza ve bakım çalışmaları belirli bölgelerde talebin düşmesine veya artmasına neden olabilir (Bouktif vd., 2018).

3.3.4. Tüketici Kaynaklı Faktörler

Konut, ticari ve endüstriyel tüketim alışkanlıkları gibi tüketici kaynaklı faktörler elektrik talebi üstünde doğrudan etkiye sahiptir. Hanelerdeki elektrik talebini kişinin veya kişilerin yaşam tarzı, enerji tasarrufu bilinci ve hanedeki elektrikli cihazların kullanım sıklığı oluşturmaktadır. Bununla birlikte iş yerleri ve fabrikaların mesai saatleri, vardiya sistemleri, üretim döngüleri ve kullanılan teknoloji harcanan yükü etkiler (Bouktif vd., 2018).

BÖLÜM 4

MATERYAL ve METOTLAR

4.1. Giriş

Bu tez çalışmasında Tekirdağ, Çerkezköy ilçesinin vaka analizi ve talep tahmini gerçekleştirilmiştir. Bölgesel hava durumu şartları ve tüketim karakteristiği incelenerek veri setleri düzenlenmiştir ve uygun tahmin modelleri oluşturulmuştur. Bu çalışmada Tekirdağ, Çerkezköy ilçesinin talep tahmini vaka analizi ile birlikte LS, KNN, DT, SVM ve ANN algoritmalarının performansları karşılaştırmalı olarak değerlendirilmiştir.

Çalışma kapsamında Tekirdağ, Çerkezköy bölgesinin, 1 Ocak 2021- 31 Aralık 2023 dönemine ait saatlik tüketim verileri TREPAŞ'tan temin edilmiştir. Çerkezköy ilçesinin aynı döneme ait sıcaklık, bağıl nem ve rüzgâr hızı hava durumu verileri, Meteoroloji Genel Müdürlüğü Bilgi Sistemi (MEVBİS)'nden temin edilmiştir. MEVBİS'ten sağlanan veriler incelendiğinde 1 Ocak 2021- 31 Aralık 2023 dönemine ait hava durumu verilerinin aylık ve saatlik ortalama biçiminde paylaşıldığı gözlemlenmiştir. Bu veri seti 3 yıllık dönemde gün içerisindeki herhangi bir saatin ayın her günü aynı ortalama hava durumu değerine sahip olmasından dolayı küçük bir veri uzayı oluşturmaktadır. Hava durumu verilerinin aynı değere sahip olması algoritmaların eğitiminde hava şartlarının belirleyici etkisini azaltarak performans kaybına neden olmuştur. MEVBİS'ten sağlanan verilerin kullanımının beklenen sonucu vermemesi üzerine ilgili döneme ait Çerkezköy ilçesi hava durumu verileri meteoblue veri tabanından günlük ve saatlik olarak temin edilmiştir.

4.2.Öğrenme Şekilleri

4.2.1. Denetimli öğrenme

Denetimli öğrenme yönteminde ilişkilendirilmiş verinin geri yayılım algoritması sayesinde istenilen çıktı ile elde edilen çıktı arasındaki hatalar hesaplanır. Bu hata değerleri ile ağırlıklandırma ayarlaması gerçekleştirilir. Çıktı düğümünde istenilen sinyaller dış bir öğretmen tarafından verildiği için bu şekilde adlandırılır (Kumar & Singh, 2012).

4.2.2. Denetimsiz öğrenme

Denetimsiz öğrenme yönteminde herhangi bir dış öğretmen bulunmamaktadır. Veriler girdi desenine göre kümelenir ve sınıflandırılır. Bu tür öğrenme yöntemi kendi kendine organizasyon olarak da adlandırılır (Kumar & Singh, 2012).

4.2.3. Pekiştirmeli öğrenme

Pekiştirmeli öğrenme yöntemi denetimli ve denetimsiz öğrenmenin birleşimidir. Durumlardan eylemlere nasıl bir haritalama gerçekleştireceğini öğrenmeye çalışır. Bu öğrenme yönteminde doğru çıktılar için ödül, yanlış çıktılar için ceza sinyalleri vardır (Kumar & Singh, 2012).

4.3. Çalışmada Kullanılan Tahmin Yöntemleri

4.3.1. En Küçük Kareler Yöntemi

LS, Lagrange polinomu, bölünmüş farklar yöntemi ve ara tahmin gibi sayısal analiz çeşitlerinden bir tanesidir. LS, birbirine bağlı iki değer arasındaki ilişkinin gerçeğe yakın bir denklemini oluşturmak için kullanılan regresyon yöntemidir (Ayyub & McCuen, 2015). Yöntemin amacı, modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkların karelerinin toplamını en aza indirgemektir. Toplam hatayı minimize eden parametreleri bulmak için türev alınır. Türevin sıfıra eşit olduğu noktalar, en küçük kareler hesaplaması için gerekli parametreleri belirtir (Boloix-Tortosa, Murillo-Fuentes, & Tsaftaris, 2019). LS, bağımlı değişken (Y) ile bağımsız değişken (X) arasındaki lineer ilişki denklem (4.1)'deki gibi modellenmektedir (Boloix-Tortosa vd., 2019).

$$Y = X\beta + \epsilon \quad (4.1) \qquad (4.1)$$

Parametre vektörü beta(β) Denklem (4.2)'deki gibidir.

$$\beta = (X^T X)^{-1} X^T Y \qquad (4.2)$$

Burada beta(β) parametre vektörü, epsilon(ϵ) hata terimidir. LS, yapılmak istenen analiz şartlarına veya veri tipine göre doğrusal, polinomal, düzlemsel, splin, genelleştirilmiş doğrusal ve bayes gibi analiz modellerine sahiptir (Boloix-Tortosa vd., 2019). Doğrusal ilişkiye sahip veri setinde kolay ve hızlı uygulanabilirliği ile tercih edilmektedir. Aykırı değerlere sahip doğrusal ilişkiye sahip olmayan veri setlerinde iyi performansa sahip değildir.

4.3.2. K-En Yakın Komşu

KNN yöntemi, parametrik olmayan bir sınıflandırma algoritmasıdır ve denetimli öğrenme yöntemidir. Veriyi, bulunduğu bir bölgedeki en yakın veya komşu eğitim örneklerine dayalı olarak sınıflandırır. Öklid, Manhattan, Chebyshev, Hamming, Minkowski gibi farklı mesafe hesaplama yöntemleri vardır (Nayak, Bhat, Reddy, & Rao, 2022). Sürekli verilerin en yakın komşularını hesaplamak için öklid mesafe ölçümü tercih edilir (Hu, Huang, Ke, & Tsai, 2016). KNN işaretlenmiş veri seti kullanarak, yeni ve işaretlenmemiş veriler için tahminde bulunur. K değeri işaretlenmemiş veriyi sınıflandırmada büyük rol oynar (Taunk, De, Verma, & Swetapadma, 2019). K değeri ilişkilendirilecek komşu sayısını temsil etmektedir ve bu değer genellikle vaka özelinde deneysel bir şekilde belirlenir. Küçük k değeri modelin aşırı uyum, büyük k değeri ise modelin yetersiz uyum sergilemesine sebep olur (Yildiz vd., 2017). Öklid mesafe ölçüm yöntemi denklem (4.3)'teki gibidir (Cunningham & Delany, 2021).

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \qquad (4.3)$$

KNN yöntemi, basit ve kolay anlaşılabilir bir dildir. Bununla birlikte esnek bir yapıya sahip olduğu için farklı analiz türlerine kolayca uygulanabilir. Her bir tahmin için tüm veri seti ile mesafe hesaplama işleminin gerçekleşmesi, büyük veri setleri ile çalışırken tahmin sürecini yavaşlatabilir. Tüm komşular eşit öneme sahiptir. Bu durum gürültülü veri setlerinde modelin tahmin performansını düşürebilir (Yildiz vd., 2017).

4.3.3. Karar Ağaçları

Karar ağaçları, kök, düğümler ve yapraklar olmak üzere 3 ana kısma ayrılır. Bir dizi testi mantıksal olarak birleştiren sıralı modellerdir (Zhao, 2007). Karar ağaçları nominal, sayısal ve alfabetik verileri işleyebilir. Her test sayısal bir özelliği, bir eşik değeriyle veya yazılı bir özelliği bir dizi olası değerle karşılaştırır. Bir veri noktası bir bölgeye düştüğünde, o bölgedeki en sık rastlanan sınıfa ait olur (Kotsiantis, 2011). 4 tane karar ağacı modeli vardır. Bunlar:

- ID3
- C4.5
- Sınıflandırma ve Regresyon Ağaçları (CART)
- Ki-Kare Otomatik Etkileşim Belirleme (CHAID)

Kök hücrenin veri setini olabildiğince çok açıklaması istenir. Düğümlerin saflığını ölçmek ve kök hücreyi belirlemek için bazı yöntemler kullanılmaktadır. Bu yöntemler:

- Bilgi kazancı yöntemi
- Bilgi kazancı oranı
- Ki-kare yöntemi
- Dağılım azaltma yöntemi
- Gini indeksi

Kök hücreyi belirlemek için genellikle gini indeksi tercih edilir. Gini indeksi, bir düğümdeki örneklerin farklı sınıflara ait olma olasılıklarının toplamını alarak düğümün saflığını belirler. En düşük gini indeksine sahip düğüm kök hücre olarak seçilir ve ağaç

oluşturulur (Moon, Kim, Kim, & Hwang, 2018). Gini indeksi denklem (4.4)' teki gibidir (Quinlan, 1986).

$$Gini\ indeksi = 1 - \sum_{i=1}^{m} p_i^2 \tag{4.4}$$

Burada p_i, bir öğenin i'inci sınıfa ait olma olsılığıdır. M, toplam sınıf sayısını belirtmektedir.

4.3.4. Destek Vektör Makinesi

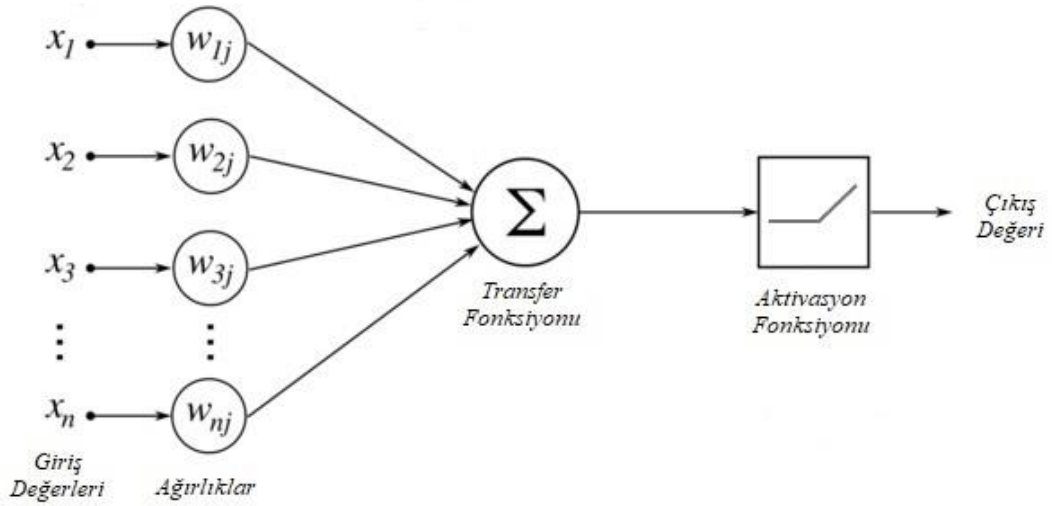
SVM istatiksel sınıflandırma ve regresyon analizinde geniş bir şekilde kullanılan, denetimli öğrenme prensibiyle çalışan makine öğrenmesi yöntemidir. Yapısal risk azaltma prensibini kullanarak hata değerlerini düşürmektedir. Bu sayede ANN'in aksine yerel uygun değer sorunlarının oluşmasını engellemektedir (Guo, Che, Shahidehpour, & Wan, 2021). SVM, eğitim verisindeki özellik uzayında negatif ve pozitif örnek aralıklarının en büyük olduğu en iyi hiper düzlemi belirler (Zhou vd., 2021). Verimli sonuçlar alabilmek için veri dağılımına uygun sınır ayarı yapılmalıdır. Sınır, sert ve esnek sınır olmak üzere iki durumdan oluşur (Ayhan & Erdoğmuş, 2014). Sert sınır, veri değerlerini doğrusal bir şekilde böler ve aykırı değerlere karşı duyarlıdır. Sınır bölgesinde veri değerleri olması durumuna esnek sınır denir (Wang vd., 2021). C hiper parametresi hataları görmezden gelme seviyesini denetler. Yüksek C değeri, eğitim veri setindeki hataları en aza indirgemek için dar bir sınır oluşturur. Bu durum aşırı uyuma yol açar. Düşük bir C değeri ise daha geniş bir sınır oluşturur. Bu durum modelin genelleme yeteneğini arttırır, model yeni verilerde daha iyi performans gösterir (Guo & Chou, 2020). 2 boyutta yeterli seviyede açıklanamayan verilere çekirdek fonksiyonları uygulanır (Sarma, Durlofsky, & Aziz, 2008). Çekirdek fonksiyonlar şu şekildedir:

- Doğrusal çekirdek fonksiyonu
- Polinomal çekirdek fonksiyonu
- Radyal tabanlı çekirdek fonksiyonu (RBF)
- Sigmoid çekirdek fonksiyonu

Sık kullanılan 2 fonksiyon vardır (Ayhan & Erdoğmuş, 2014). Bunlar RBF ve polinomal çekirdek fonksiyonlarıdır. Polinomal çekirdek fonksiyonu ile boyut arttırarak veriler daha ayrıntılı şekilde anlamlandırabilir (Wang, Xing, & Schaid, 2015). RBF ise örnek uzaydaki belirli iki noktanın benzerliğini veya yakınlığını karşılaştırarak sınıflandırma yapar (Yang vd., 2020). Bu sayede tahmin modellerindeki zayıf sınıflandırma ve aşırı uyum gibi sorunlara etkili bir yaklaşım gerçekleştirir. Benzerlik veya yakınlık karşılaştırması yapılan alanın genişliği gamma hiper parametresi ile kontrol edilir.

4.3.5. Yapay Sinir Ağları

ANN yaklaşımı, Warren McCulloh ve Walter Pitts tarafından 1990 yılında zaman serilerine alternatif bir yöntem olarak geliştirilmiştir. Özellikle tahmin ve sınıflandırma yapmak için farklı alanlarda başarıyla uygulanmıştır. Geliştirilmesinden kısa bir süre sonra talep tahmini için kullanılan popüler bir araç haline gelmiştir (Hammad, Jereb, Rosi, & Dragan, 2020). ANN modelleri, insan beyninin bilgi işleme ve organizasyon ilkeleri örnek alınarak oluşturulmuştur. ANN çok sayıda bağlantıya sahip olan fazla sayıda işlemciden oluşan devasa paralel bilgi işlem sistemleridir. Giriş katmanı, gizli katman ve çıkış katmanından oluşmaktadır. Giriş katmanı dış dünya ile doğrudan etkileşim içerisindedir. Gizli katmanda belirlenen aktivasyon fonksiyonları ile hesaplamalar gerçekleştirilir. Çıkış katmanında elde edilen sonuçlar alınır (Kumar & Singh, 2012). ANN, belirli değerlere sahip girdiler ve çıktı değeri ilişkilendirilmiş veri setlerini inceleyerek eğitilmektedir (Jain, Mao, & Mohiuddin, 1996). Ağırlık ilişkilendirmelerinin döngüsel olarak ayarlandığı bir eğitim sürecine sahiptir. Girdi ve çıktı değeri ilişkilendirilmiş veri seti, olasılık ağırlıklı ilişkiler oluşturmaya yaramaktadır. ANN ağırlık ilişkilendirmelerini, öğrenme kurallarına bağlı kalarak ve hedef çıktı ile işlenen çıktı arasındaki hata değerini kullanarak ayarlamaktadır (Montesinos López, Montesinos López, & Crossa, 2022). ANN hücresi Şekil 4.1'deki gibidir.



Şekil 4.1. Yapay sinir ağları hücresi

Elektrik talep tahmini için yaygın olarak kullanılan ANN modelleri şunlardır (Ahmed, Alam, & Hassan, 2023):

- İleri Beslemeli Sinir Ağı (FFNN)
- NARX sinir ağı
- BPNN
- RBF sinir ağı
- Rastgele Sinir Ağları (RNN)
- Geri Döngülü Sinir Ağları (RNNs)
- Kendini Organize Eden Rekabetçi Sinir Ağları (SOCNN)

Aktivasyon fonksiyonları, toplam, çarpım, maksimum, minimum, çoğunluk ve birikmiş toplam gibi işlemleri kullanarak hücrelere gelen verileri işler ve üretilecek çıkış değerini belirler (Salah Alaloul & Hannan Qureshi, 2020). Aktivasyon fonksiyonu belirlenirken doğrusal olmayan ve türevi hesaplanabilir bir fonksiyon seçilmelidir (Mohammad-Zaheri, Chen, Ghaffari, & Willison, 2009). ANN'lerde uygunluğu

nedeniyle genellikle sigmoid fonksiyonu kullanılmaktadır (Kılıçarslan, Adem, & Çelik, 2021). Yaygın olarak kullanılan ANN aktivasyon fonksiyonları şunlardır:

- Doğrusal aktivasyon fonksiyonu
- Sigmoid aktivasyon fonksiyonu
- Hiperbolik tanjant aktivasyon fonksiyonu
- ReLU aktivasyon fonksiyonu
- Sızıntılı ReLU aktivasyon fonksiyonu
- Swish aktivasyon fonksiyonu

BÖLÜM 5

TALEP TAHMİNİ UYGULAMASI ÇERKEZKÖY BÖLGESİ VAKA ANALİZİ

5.1. Tahmin ve Analiz

Bu çalışmada algoritmalar özelinde hangi bağımsız değişkenlerin daha etkili olacağını gözlemlemek amacıyla 15 farklı tahmin senaryosu oluşturulmuştur. Senaryolarda bağımsız değişken olarak saat, haftanın günü, hafta sonu olma durumu, resmi tatil olma durumu, sıcaklık, nem, rüzgâr, 1 hafta öncesinin aynı saatteki tüketim verisi, 1 gün öncesinin aynı saatteki tüketim verisi ve 1 saat öncesinin tüketim verisi kullanılmıştır. Saat değişkeni 1-24 arası değer almaktadır ve günün saatini belirtmektedir. Haftanın günü değişkeni hangi gün olduğunu belirtmek amacıyla 1-7 arası değer almaktadır. Hafta sonu olma durumunu belirten değişken 0 veya 1 değerleri almaktadır. Hafta sonu olan günler 1 ile belirtilmiştir. Resmi tatil olup olmama durumunu belirten değişken de 0 veya 1 değeri almaktadır. Resmi tatil olma durumu 1 ile belirtilmektedir. Senaryolarda bağımlı değişken olarak saatlik tüketim verileri kullanılmıştır. Tahmin performansını arttırmak amacıyla eksik veya hatalı verilerin olduğu dönemler tespit edilerek veri setinden çıkarılmıştır. Senaryolar özelinde kullanılan bağımsız değişkenler Çizelge 5.1'de belirtildiği gibidir.

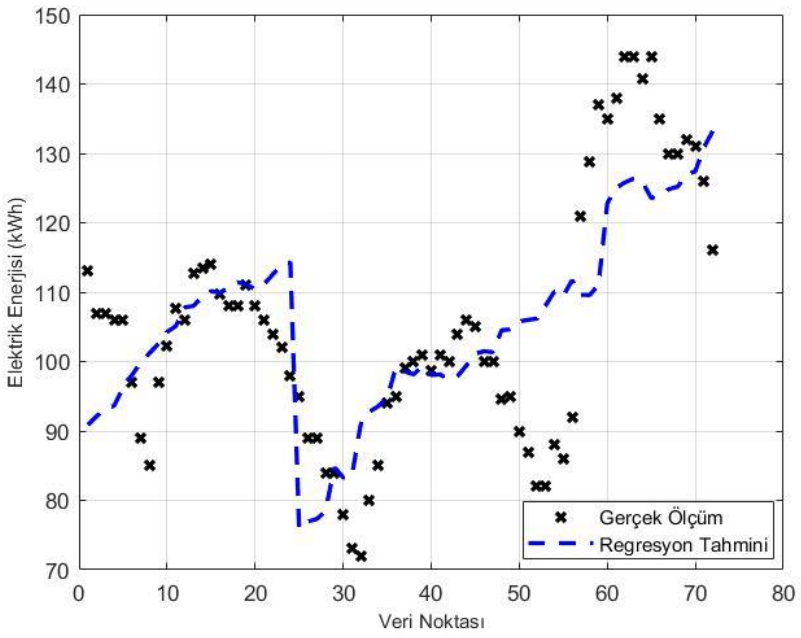
Çizelge 5.1. Senaryolar Özelinde Kullanılan Bağımsız Değişkenler

Senaryo/Bağımsız değişkenler	Sıcaklık	Nem	Rüzgar hızı	Günün saati	Haftanın günü	Hafta sonu olma durumu	Resmi tatil olma durumu	1 hafta önceinin aynı saatteki tüketim değeri	1 gün önceinin aynı saatteki tüketim değeri	1 saat önceinin aynı saatteki tüketim değeri
1. Senaryo	✓	✓	✓	✓	✓	✓	✓			
2. Senaryo	✓	✓	✓		✓	✓	✓			
3. Senaryo	✓	✓	✓			✓	✓			
4. Senaryo	✓	✓	✓		✓		✓			
5. Senaryo	✓	✓	✓				✓			
6. Senaryo	✓	✓	✓	✓						
7. Senaryo	✓	✓	✓							
8. Senaryo	✓	✓		✓	✓	✓	✓			
9. Senaryo	✓			✓	✓	✓	✓			
10. Senaryo		✓		✓	✓	✓	✓			
11. Senaryo			✓	✓	✓	✓	✓			
12. Senaryo	✓	✓	✓	✓	✓	✓	✓	✓		

13. Senaryo	✓	✓	✓	✓	✓	✓	✓		✓	
14. Senaryo	✓	✓	✓	✓	✓	✓	✓			✓
15. Senaryo	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

5.1.1. En Küçük Kareler Yöntemi ile Talep Tahmini Uygulaması

Uydurulan eğrinin ve tüketim verilerinin grafik üzerinde anlaşılabilir olması için veri seti olarak 1 Mayıs 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 72 saatlik süreçteki değerler kullanılmıştır. X ekseni eğri uydurma işleminde kullanılan veri sayısını, y ekseni tüketilen elektrik enerjisini belirtmektedir. 1. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.1'deki gibidir.



Şekil 5.1. En küçük kareler yöntemi 1. Senaryo tahmin grafiği

1. Senaryonun sonuçları incelendiğinde algoritmanın genel davranışları yakalama performansının düşük olduğu gözlemlenmektedir. Düşüş ve yükseliş eğrilerini anlamlı

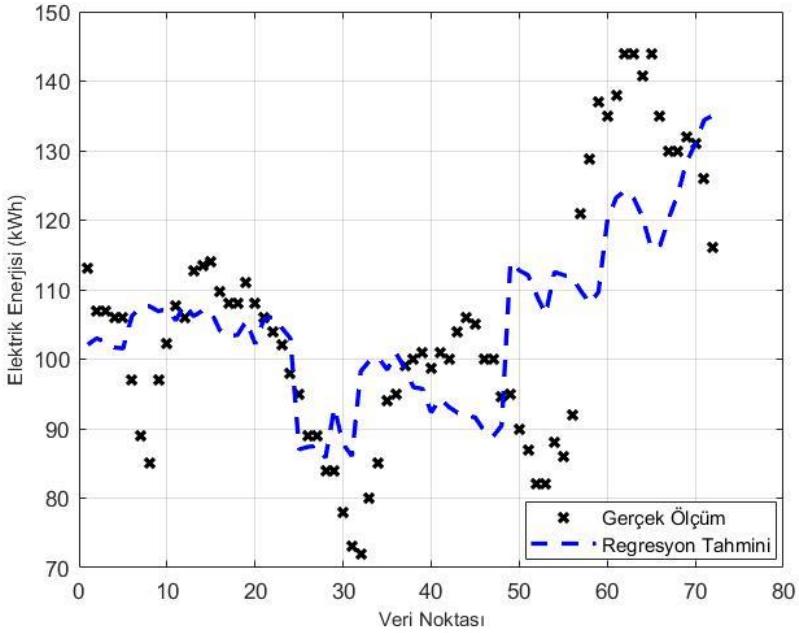
şekilde takip etse de tüketim değerlerinin dip ve zirve yaptığı noktalarda sapmalar mevcuttur. 1. Senaryo için uydurulan eğrinin formülü denklem (5.1)'de belirtilmektedir.

$$Tüketim = b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * sıcaklık + b_6 * nem + b_7 * rüzgar \quad (5.1)$$

Denklem (5.1)'de b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6, b_7 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.1)'in sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 59.5473
- b_1: 1.1186
- b_2: -4.3008
- b_3 : 0
- b_4: 7.667
- b_5: 0.99296
- b_6: 0.37108
- b_7: 0.55019

2. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.2'deki gibidir.



Şekil 5.2. En küçük kareler yöntemi 2. Senaryo tahmin grafiği

Günün saati bağımsız değişkeninin çıkarılması sonucu algoritma zamansal dalgalanmaları yakalamakta zayıf kalmıştır. Tüketim değerinin dip ve zirve yaptığı noktalarda sapmaların arttığı gözlemlenmektedir. 2. Senaryo için uydurulan eğrinin formülü denklem (5.2)'de belirtilmektedir.

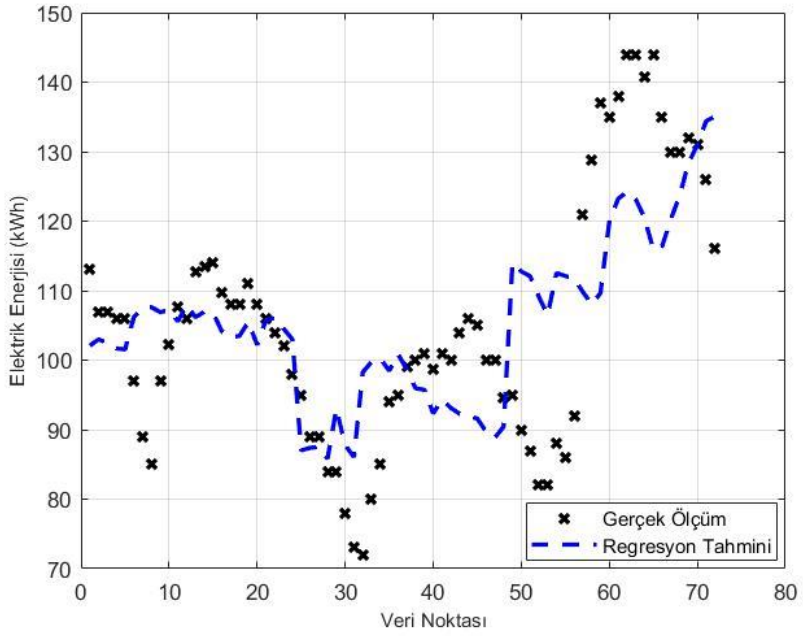
$$\begin{aligned} T\ddot{u}ketim = {} & b_0 + b_1 * haftanın\ g\ddot{u}n\ddot{u} + b_2 * hafta\ sonu\ durumu + b_3 \\ & * resmi\ tatil\ durumu + b_4 * sıcaklık + b_5 * nem + b_6 \\ & * r\ddot{u}zgar \end{aligned} \tag{5.2}$$

Denklem (5.2)'de b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.2)'nin sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 130.8407
- b_1: -4.0134
- b_2: 0

- b_3 : 11.98
- b_4: -0.56885
- b_5: -0.17967
- b_6: 0.78638

3. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.3'teki gibidir.



Şekil 5.3. En küçük kareler yöntemi 3. Senaryo tahmin grafiği

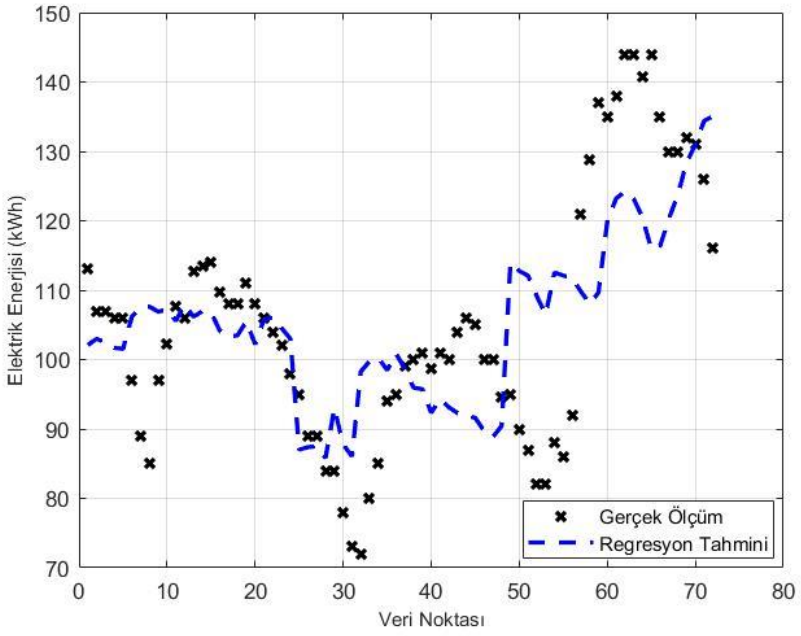
3. Senaryo için uydurulan eğrinin formülü denklem (5.3)'te belirtilmektedir.

$$Tüketim = b_0 + b_1 * hafta\ sonu\ durumu + b_2 * resmi\ tatil\ durumu + b_3 * sıcaklık + b_4 * nem + b_5 * rüzgar \quad (5.3)$$

Denklem (5.3)'te b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.3)'ün sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 126.8273
- b_1: -24.0804
- b_2: 15.9934
- b_3 : -0.56885
- b_4: -0.17967
- b_5: 0.78638

4. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.4'teki gibidir.



Şekil 5.4. En küçük kareler yöntemi 4. Senaryo tahmin grafiği

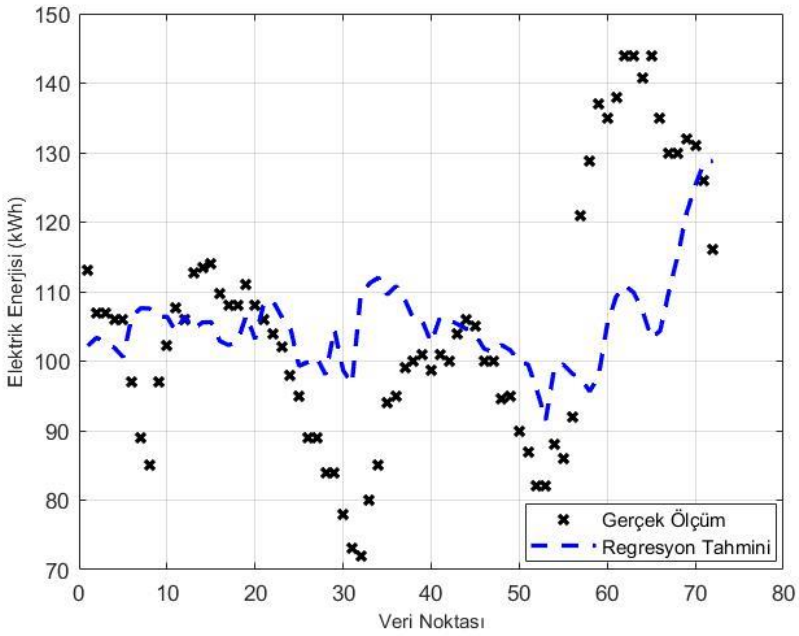
4. Senaryo için uydurulan eğrinin formülü denklem (5.4)'te belirtilmektedir.

$$Tüketim = b_0 + b_1 * haftanın\ günü + b_2 * resmi\ tatil\ durumu + b_3 * sıcaklık + b_4 * nem + b_5 * rüzgar \quad (5.4)$$

Denklem (5.4)'te b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.4)'ün sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 130.8407
- b_1: -4.0134
- b_2: 11.98
- b_3 : -0.56885
- b_4: -0.17967
- b_5: 0.78638

5. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.5'teki gibidir.



Şekil 5.5. En küçük kareler yöntemi 5. Senaryo tahmin grafiği

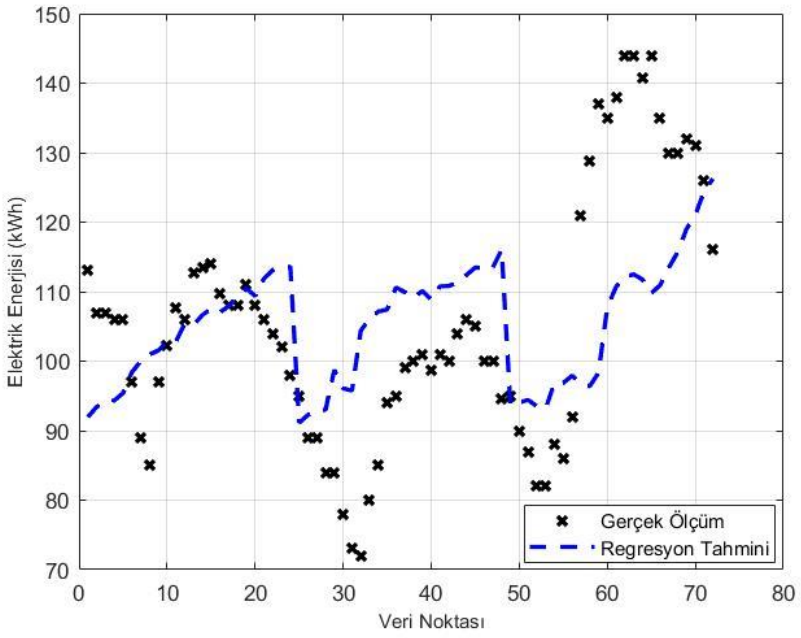
5. Senaryo için uydurulan eğrinin formülü denklem (5.5)'te belirtilmektedir.

$$Tüketim = b_0 + b_1 * resmi\ tatil\ durumu + b_2 * sıcaklık + b_3 * nem + b_4 * rüzgar \quad (5.5)$$

Denklem (5.5)'te b_0 sabit terim ve b_1, b_2, b_3, b_4 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.5)'in sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 150.3876
- b_1: 5.8986
- b_2: -1.4713
- b_3 : -0.47951
- b_4: 0.74623

Algoritmanın 3. Senaryo, 4. Senaryo ve 5. Senaryodaki davranışları incelendiğinde yakın sonuçlar verdiği gözlemlenmektedir. Veri uzayının küçük olduğu tahminlerde haftanın günü, hafta sonu olma durumu ve resmi tatil olma durumu bağımsız değişkenlerinin zamansal etkileri istenilen şekilde göstermediği gözlemlenmiştir. 6. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.6'daki gibidir.



Şekil 5.6. En küçük kareler yöntemi 6. Senaryo tahmin grafiği

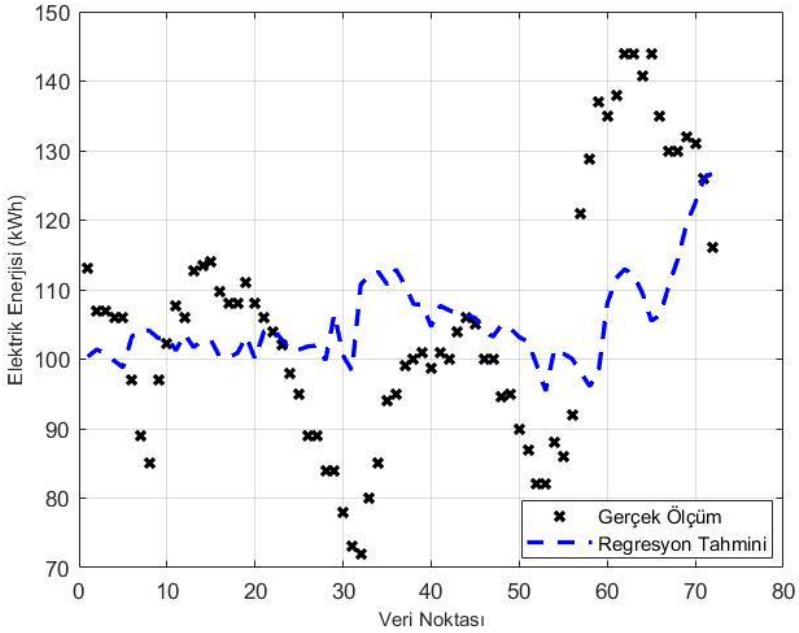
Günün saati bağımsız değişkeni dâhil edildiğinde algoritmanın genel davranışları ve zamansal dalgalanmaları algılaması artmıştır ancak dip ve zirve noktalardaki sapmalar devam etmektedir. 6. Senaryo için uydurulan eğrinin formülü denklem (5.6)'da belirtilmektedir.

$$Tüketim = b_0 + b_1 * saat + b_2 * sıcaklık + b_3 * nem + b_4 * rüzgar \quad (5.6)$$

Denklem (5.6)'da b_0 sabit terim ve b_1, b_2, b_3, b_4 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.6)'nın sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 82.8463
- b_1: 0.96604
- b_2: -0.010658
- b_3 : 0.050673
- b_4: 0.55581

7. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.7'deki gibidir.



Şekil 5.7. En küçük kareler yöntemi 7. Senaryo tahmin grafiği

7. Senaryo tahmin grafiği incelendiğinde saat bağımsız değişkeninin çıkarılmasıyla tahmin sonuçlarının ortalama bir çizgiye sahip olduğu gözlemlenmektedir. 6. Senaryo ve 7. Senaryo grafiği incelendikten sonra günün saati bağımsız değişkeninin, algoritmanın performansı üzerindeki olumlu etkisi kesinleşmiştir. Az sayıda bağımsız değişkenin olması algoritmanın tahmin performansını düşürmektedir. 7. Senaryo için uydurulan eğrinin formülü denklem (5.7)'deki gibidir.

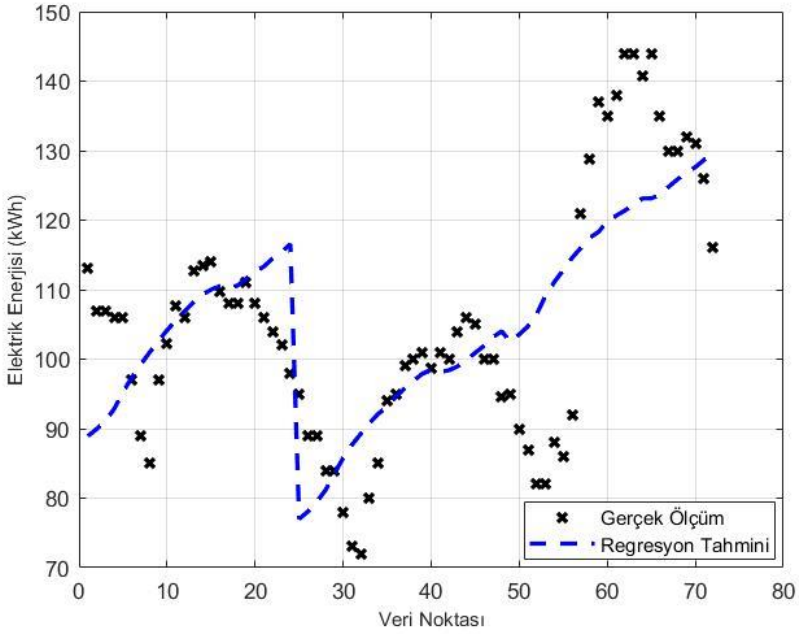
$$Tüketim = b_0 + b_1 * sıcaklık + b_2 * nem + b_3 * rüzgar \quad (5.7)$$

Denklem (5.7)'de b_0 sabit terim ve b_1 , b_2 , b_3 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.7)'nin sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 126.8293
- b_1: -0.92134

- b_2: -0.23968
- b_3 : 0.79889

8. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.8'deki gibidir.



Şekil 5.8. En küçük kareler yöntemi 8. Senaryo tahmin grafiği

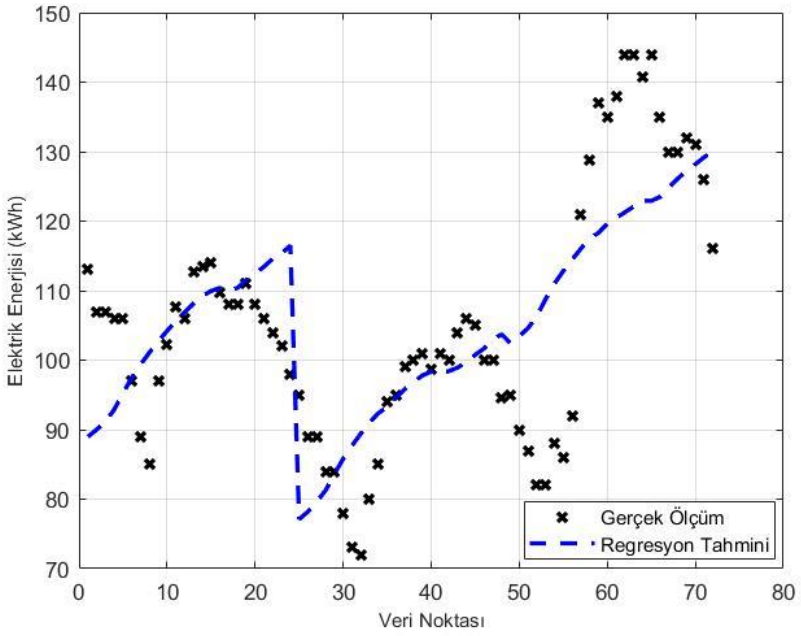
Tüm zamansal bağımsız değişkenler uygulamaya dâhil edildiğinde algoritmanın performansı artmaktadır. 8. Senaryo için uydurulan eğrinin formülü denklem (5.8)'de belirtilmektedir.

$$Tüketim = b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * sıcaklık + b_6 * nem \quad (5.8)$$

Denklem (5.8)'de b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.8)'in sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 96.3781
- b_1: 1.1837
- b_2: -4.2998
- b_3 : 0
- b_4: 7.9476
- b_5: 0.44136
- b_6: 0.034821

9. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.9'daki gibidir.



Şekil 5.9. En küçük kareler yöntemi 9. Senaryo tahmin grafiği

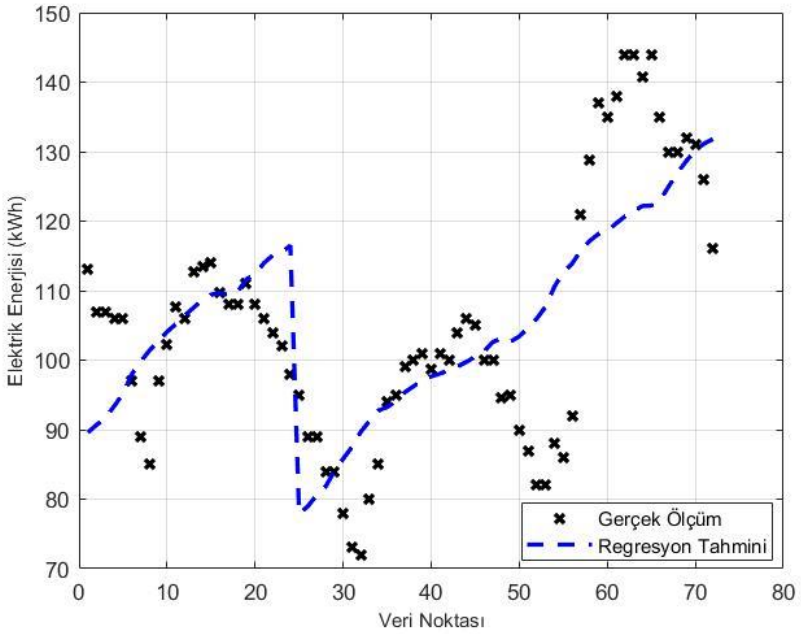
9. Senaryo için uydurulan eğrinin formülü denklem (5.9)'daki gibidir.

$$Tüketim = b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * sıcaklık \quad (5.9)$$

Denklem (5.9)'da b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.9)'un sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 100.1978
- b_1: 1.1682
- b_2: -4.2859
- b_3 : 0
- b_4: 8.1718
- b_5: 0.35388

10. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.10'daki gibidir.



Şekil 5.10. En küçük kareler yöntemi 10. Senaryo tahmin grafiği

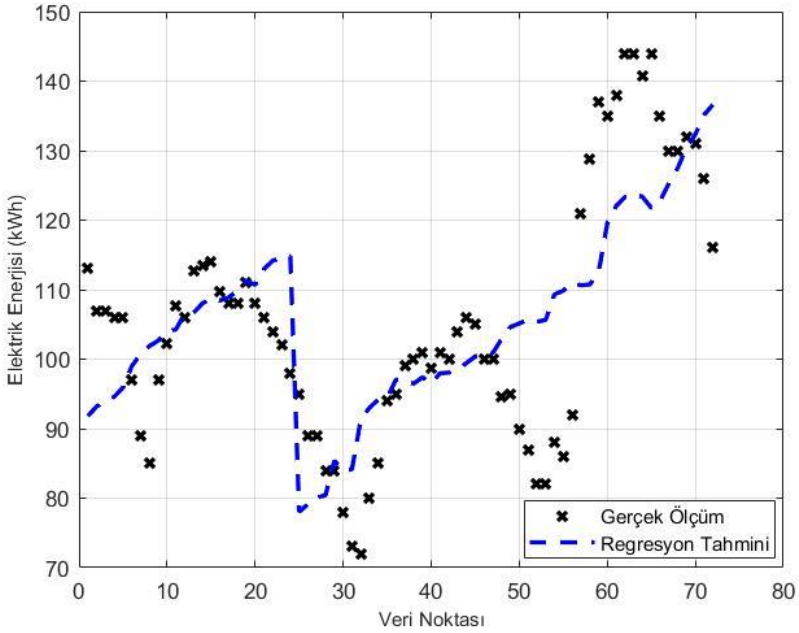
10. Senaryo için uydurulan eğrinin formülü denklem (5.10)'da belirtilmektedir.

$$Tüketim = b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * nem \quad (5.10)$$

Denklem (5.10)'da b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.10)'un sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 114.382
- b_1: 1.1133
- b_2: -4.2275
- b_3 : 0
- b_4: 8.8856
- b_5: -0.11791

11. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.11'deki gibidir.



Şekil 5.11. En küçük kareler yöntemi 11. Senaryo tahmin grafiği

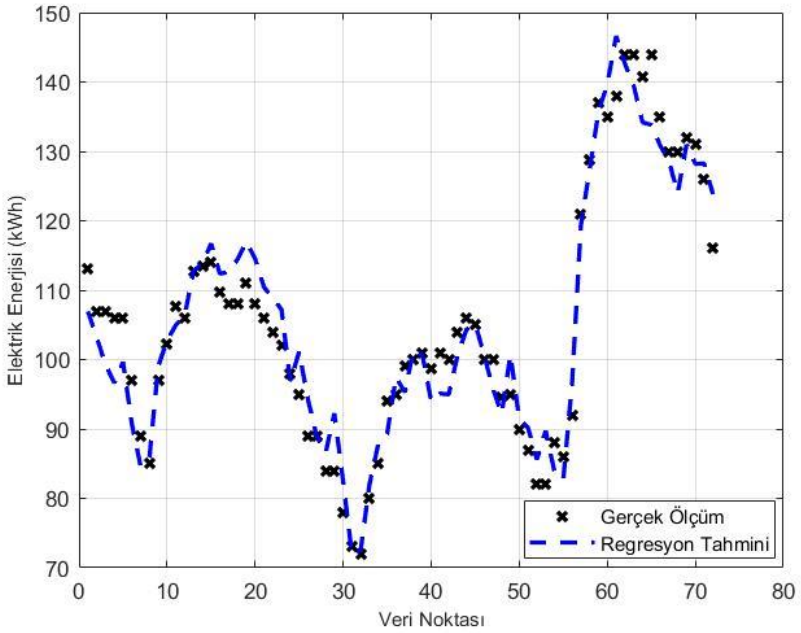
9. Senaryo, 10. Senaryo ve 11. Senaryo çıktıları incelendiğinde meteorolojik bağımsız değişkenlerin belirleyici etkisinin birbirine çok yakın olduğu gözlemlenmiştir. 11. Senaryo çıktıları incelendiğinde rüzgâr bağımsız değişkeninin diğer meteorolojik değişkenlere göre daha iyi belirleyiciliğe sahip olduğu gözlemlenmektedir. Genel tüketim karakteristiği başarılı bir şekilde yakalanmıştır fakat dip ve zirve noktalarda büyük sapmalar mevcuttur. 11. Senaryo için uydurulan eğrinin formülü denklem (5.11)'de belirtilmektedir.

$$Tüketim = b_0 + b_1 * saat + b_2 * haftanın günü + b_3 * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * rüzgar \quad (5.11)$$

Denklem (5.11)'de b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.11)'in sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 103.0888
- b_1: 0.99987
- b_2: -4.1591
- b_3 : 0
- b_4: 9.4981
- b_5: 0.41746

12. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.12'deki gibidir.



Şekil 5.12. En küçük kareler yöntemi 12. Senaryo tahmin grafiği

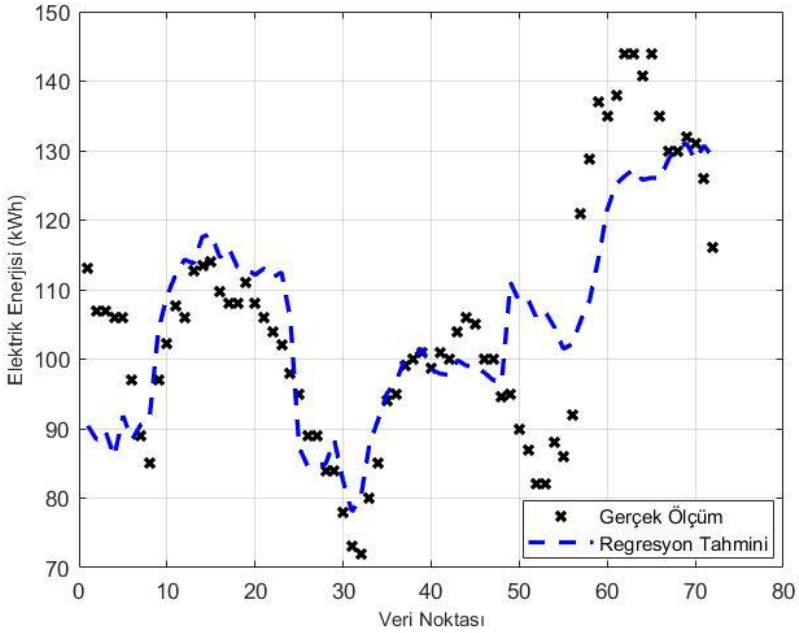
12. Senaryo çıktıları incelendiğinde tüm zamansal ve meteorolojik değişkenlere 1 hafta öncesinin aynı saatteki tüketim değerinin eklenmesi algoritmanın çalışma performansını yükseltmiştir. Ancak hala dip ve zirve noktalarda küçük sapmalar gözlemlenmektedir. 12. Senaryo için uydurulan eğrinin formülü denklem (5.12)'deki gibidir.

$$
\begin{aligned}
T\ddot{u}ketim = \; & b_0 + b_1 * saat + b_2 * haftanın\ g\ddot{u}n\ddot{u} + b_3 \\
& * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 \\
& * sıcaklık + b_6 * nem + b_7 * r\ddot{u}zgar + b_8 \\
& * 1\ hafta\ \ddot{o}ncesinin\ aynı\ saatteki\ t\ddot{u}ketim\ verisi
\end{aligned}
\tag{5.12}
$$

Denklem (5.12)'de b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.12)'nin sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: -45.1594
- b_1: -0.1675
- b_2: -0.25708
- b_3 : 0
- b_4: -1.7644
- b_5: 0.80113
- b_6: 0.46949
- b_7: 0.58778
- b_8: 0.95209

13. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.13'teki gibidir.



Şekil 5.13. En küçük kareler yöntemi 13. Senaryo tahmin grafiği

13. Senaryo çıktıları incelendiğinde zamansal ve meteorolojik değişkenlere 1 gün öncesinin aynı saatteki tüketim değerinin eklenmesiyle algoritma performansı düşmüştür. 13. Senaryo için uydurulan eğrinin formülü denklem (5.13)'te belirtilmektedir.

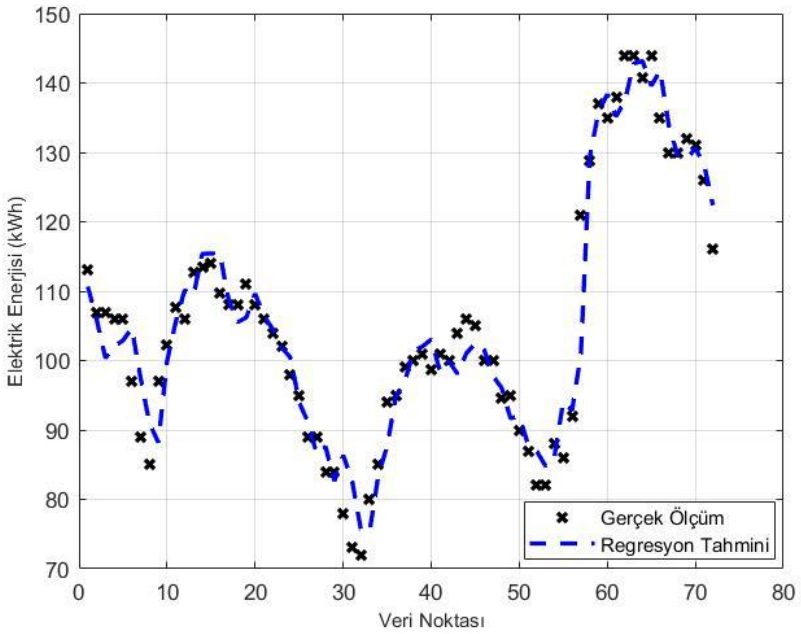
$$\begin{aligned} Tüketim = {} & b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 \\ & * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 \\ & * sıcaklık + b_6 * nem + b_7 * rüzgar + b_8 \\ & * 1\ gün\ öncesinin\ aynı\ saatteki\ tüketim\ verisi \end{aligned} \quad (5.13)$$

Denklem (5.13)'te b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.13)'ün sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 38.3825
- b_1: 0.70473
- b_2: -5.3858

- b_3 : 0
- b_4: -6.0264
- b_5: 0.45519
- b_6: 0.1837
- b_7: 0.30478
- b_8: 0.56426

14. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.14'teki gibidir.



Şekil 5.14. En küçük kareler yöntemi 14. Senaryo tahmin grafiği

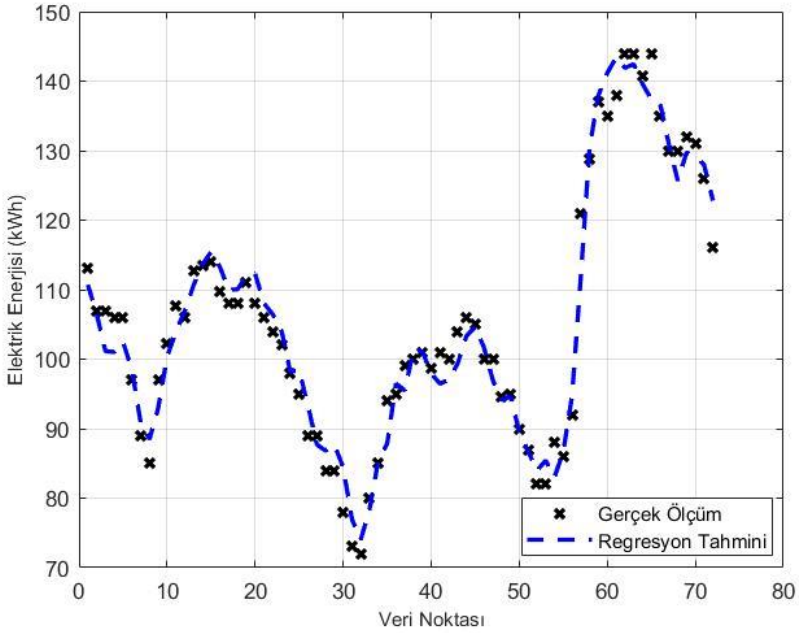
14. Senaryo çıktıları incelendiğinde 1 saat öncesinin tüketim değeri bağımsız değişkenlere dâhil edildiğinde tüketim karakteristiği oldukça başarılı bir şekilde yakalanmıştır. Dip ve zirve değerlerde küçük sapmalar mevcuttur. 14. Senaryo için uydurulan eğrinin formülü denklem (5.14)'te belirtilmektedir.

$$
\begin{aligned}
T\ddot{u}ketim = \; & b_0 + b_1 * saat + b_2 * haftanın\ g\ddot{u}n\ddot{u} + b_3 \\
& * hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 \\
& * sıcaklık + b_6 * nem + b_7 * r\ddot{u}zgar + b_8 \\
& * 1\ saat\ \ddot{o}ncesinin\ t\ddot{u}ketim\ verisi
\end{aligned}
\tag{5.14}
$$

Denklem (5.14)'te b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.14)'ün sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: 14.1739
- b_1: 0.097956
- b_2: -0.54865
- b_3 : 0
- b_4: 0.13609
- b_5: 0.30173
- b_6: -0.15948
- b_7: -0.30008
- b_8: 0.93562

15. Senaryo için gerçekleştirilen LS talep tahmini grafiği Şekil 5.15'teki gibidir.



Şekil 5.15. En küçük kareler yöntemi 15. Senaryo tahmin grafiği

15. Senaryo, tüm bağımsız değişkenlerin kullanılmasıyla birlikte genel davranışı yakalamada en başarılı performansı sergilemiştir. Dip ve zirve değerlerdeki uyum oldukça yüksektir. 15. Senaryo için uydurulan eğrinin formülü denklem (5.15)'te belirtilmektedir.

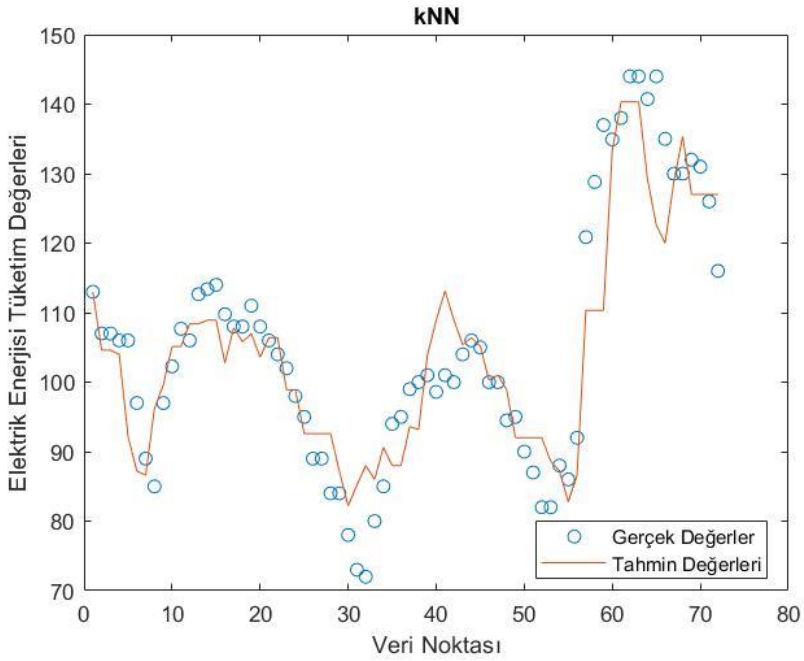
$$\begin{aligned} &Tüketim = b_0 + b_1 * saat + b_2 * haftanın\ günü + b_3 * \\ &hafta\ sonu\ durumu + b_4 * resmi\ tatil\ durumu + b_5 * sıcaklık + b_6 * \\ &nem + b_7 * rüzgar + b_8 * \\ &1\ hafta\ öncesinin\ aynı\ saatteki\ tüketim\ verisi + b_9 * \\ &1\ gün\ öncesinin\ aynı\ saatteki\ tüketim\ verisi + b_{10} * \\ &1\ saat\ öncesinin\ tüketim\ verisi \end{aligned} \tag{5.15}$$

Denklem (5.15)'te b_0 sabit terim ve b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8 senaryoda kullanılan bağımsız değişkenlerin regresyon katsayılarıdır. Denklem (5.15)'in sabit terimi ve regresyon katsayıları şu şekildedir:

- b_0: -24.0085
- b_1: -0.10931
- b_2: 0.25743
- b_3 : 0
- b_4: 0.5547
- b_5: 0.62866
- b_6: 0.20324
- b_7: 0.19688
- b_8: 0.59021
- b_9: -0.10139
- b_{10}: 0.46294

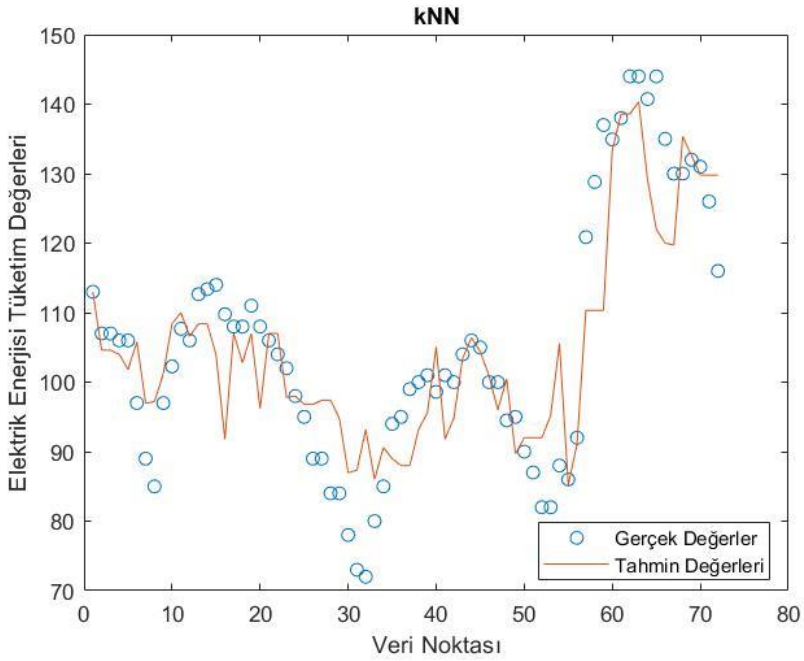
5.1.2. K-En Yakın Komşu ile Talep Tahmini Uygulaması

Tahmin değerleri ve tüketim verilerinin grafik üzerinde anlaşılabilir olması için veri seti olarak 1 Mayıs 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 72 saatlik süreçteki değerler kullanılmıştır. Yapılan denemeler sonucu k değeri 5 olarak belirlenmiştir. Ağırlıkandırma türü mesafe olarak belirlenmiştir ve en yakın komşuları hesaplamak için Öklid mesafe ölçümü kullanılmıştır. X ekseni karşılaştırma yapmak üzere referans alınan tüketim değerlerinin örnek sayını, y ekseni tüketilen elektrik enerjisini belirtmektedir. 1. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.16'daki gibidir.



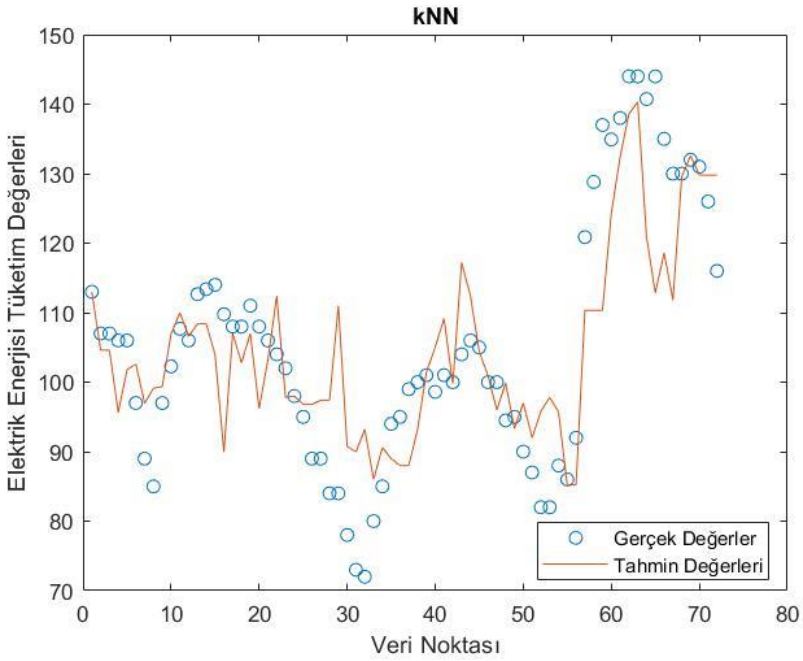
Şekil 5.16. K-en yakın komşu 1. Senaryo tahmin grafiği

KNN algoritmasının 1. Senaryo için grafiği incelendiğinde tüketim değerlerinin genel davranışlarını başarılı bir şekilde yakaladığı gözlemlenmektedir. Tüketim değerlerinin ani düşüşlerinde ve zirve değerlerinde küçük sapmalar mevcuttur. 2. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.17'deki gibidir.



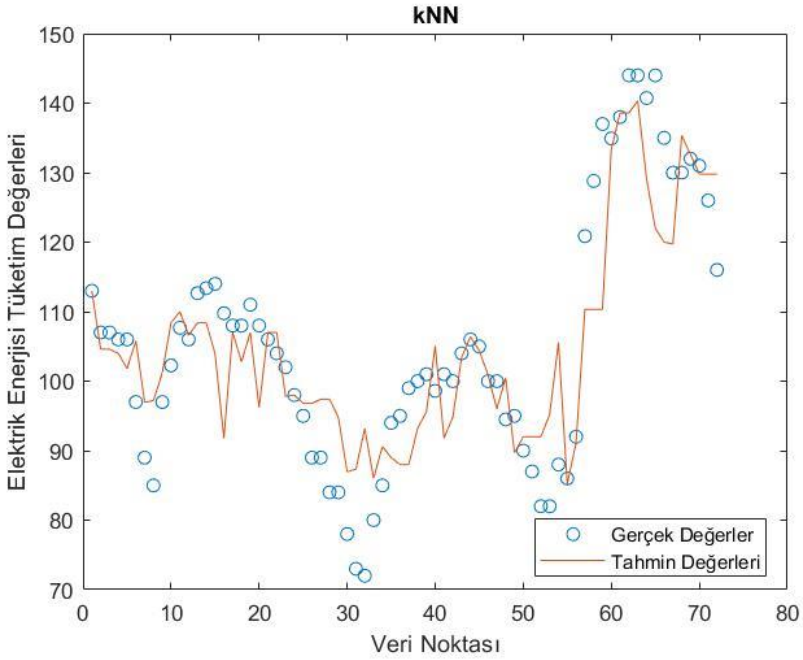
Şekil 5.17. K-en yakın komşu 2. Senaryo tahmin grafiği

Bağımsız değişkenler arasından günün saati çıkarıldığı zaman algoritmanın tüketim karakteristiğini yakalamadaki performansının düştüğü gözlemlenmektedir. Ani düşüş ve zirve değerlerdeki tutarsızlık artmıştır. 3. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.18'deki gibidir.



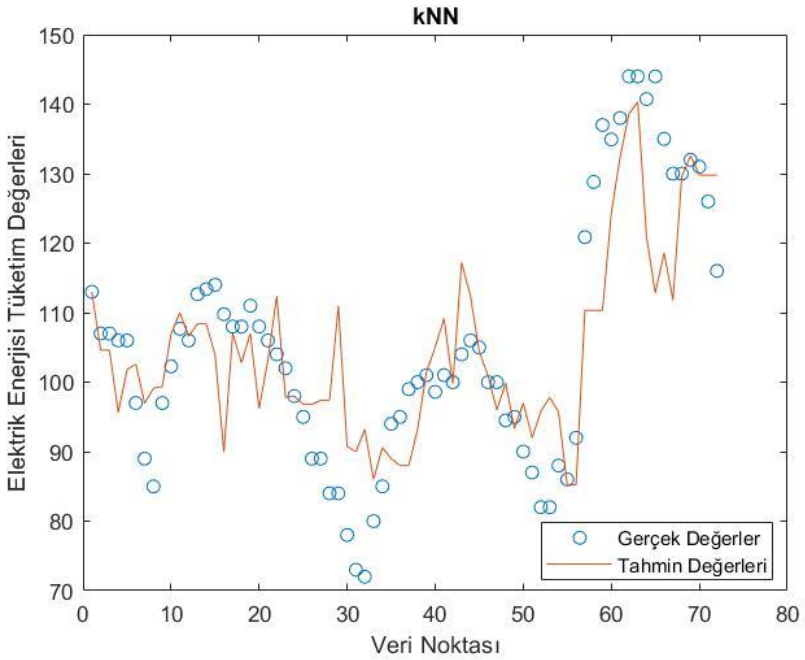
Şekil 5.18. K-en yakın komşu 3.Senaryo tahmin grafiği

4. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.19'daki gibidir.



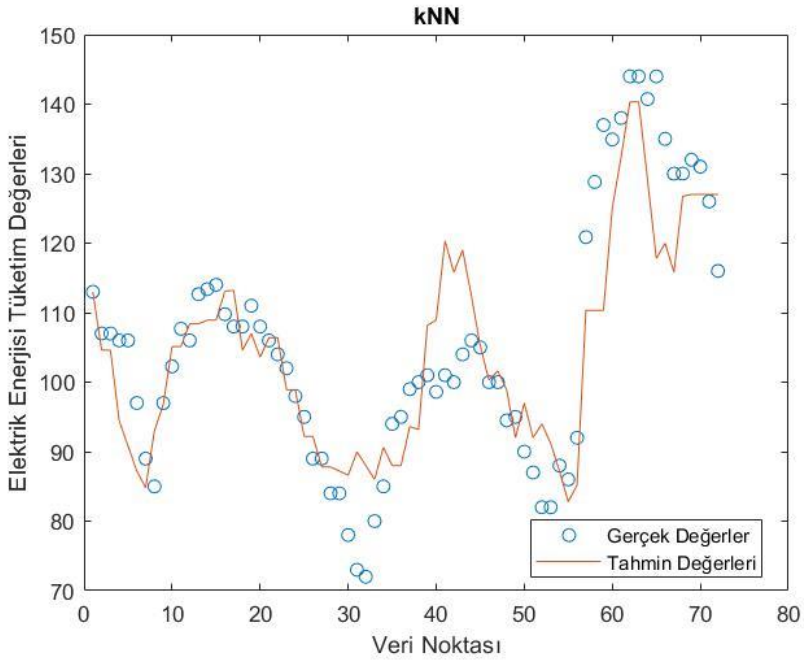
Şekil 5.19. K-en yakın komşu 4. Senaryo tahmin grafiği

Algoritmanın 3. Senaryo, 4. Senaryo ve 5. Senaryo özelindeki çıktıları incelendiğinde haftanın günü, hafta sonu olma durumu ve resmi tatil olma durumu bağımsız değişkenlerinin tek başına önemli derecede bir etkisi olmadığı gözlemlenmektedir. Eğri tüketim karakteristiğini takip ederken dalgalanmalara uğramıştır. Ani düşüşler ve zirve değerlerdeki sapmalar giderilememiştir. 5. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.20'deki gibidir.



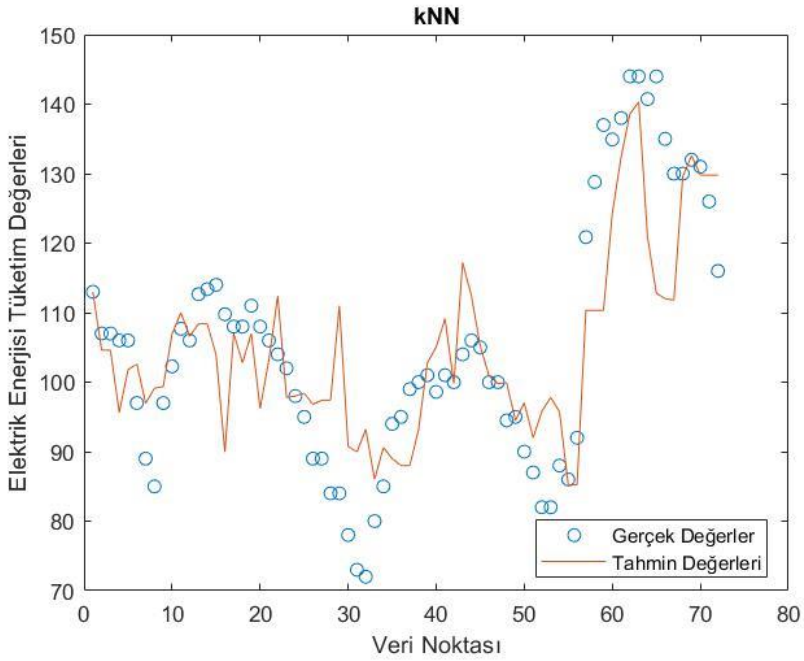
Şekil 5.20. K-en yakın komşu 5. Senaryo tahmin grafiği

6. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.21'deki gibidir.



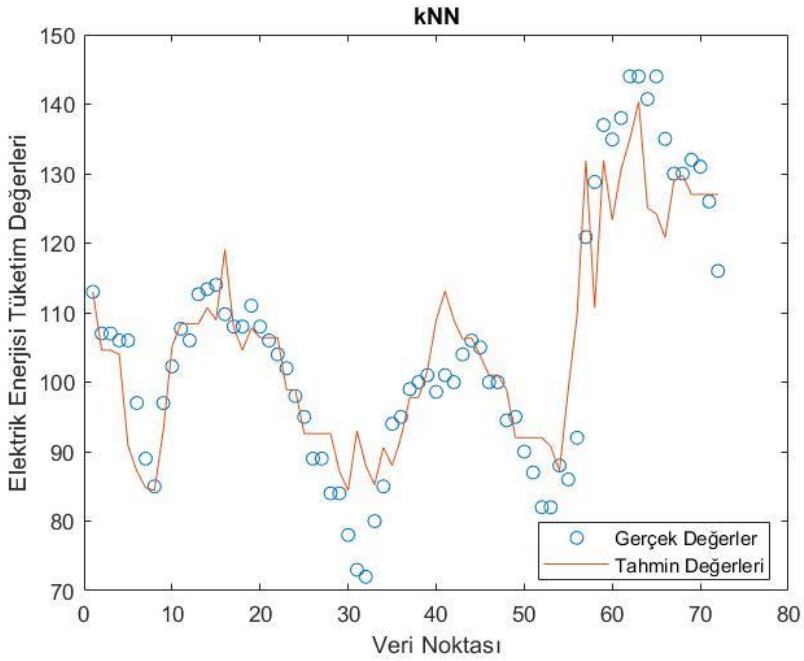
Şekil 5.21. K-en yakın komşu 6. Senaryo tahmin grafiği

6. Senaryonun grafiği incelendiğinde günün saati bağımsız değişkeninin, zamansal döngü ifade eden değişkenler arasında en yüksek belirleyiciliğe sahip olduğu anlaşılmaktadır. Günün saati değişkenin tahmin senaryosuna dâhil edilmesiyle birlikte tüketim değerlerinin genel davranışı başarılı bir şekilde yakalanmıştır. Bağımsız değişken sayısının az olması modelin performansını kısıtlamıştır. 7. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.22'deki gibidir.



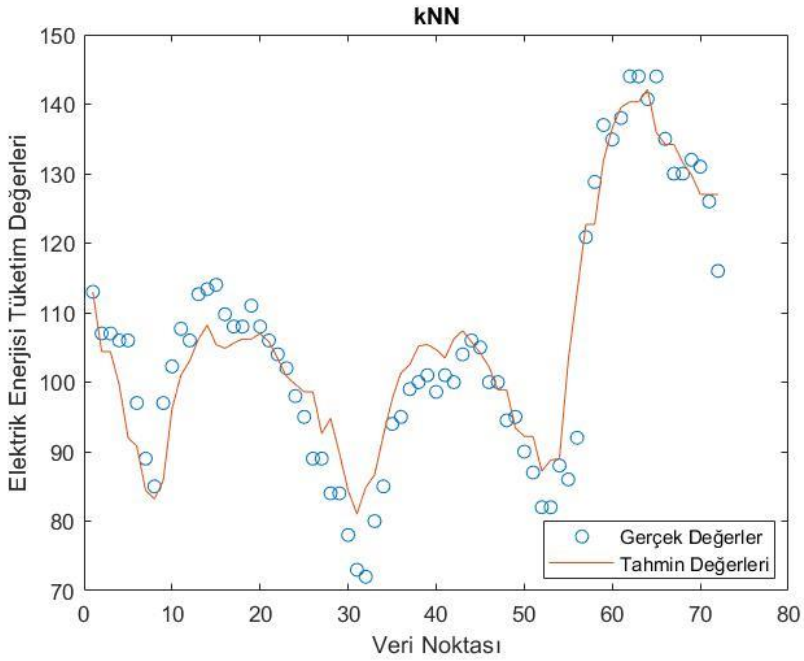
Şekil 5.22. K-en yakın komşu 7. Senaryo tahmin grafiği

Algoritmanın 7. Senaryoya olan tepkileri incelendiğinde hiçbir zamansal bağımsız değişkenin olmaması nedeniyle tutarlı bir grafik ve tahmin sonucu sağlanamamıştır. 6. Senaryonun grafiği ile karşılaştırıldığında tüketimin genel davranışlarında dalgalanmalar, ani tüketim değeri düşüşleri ve zirve noktalardaki sapmalar tekrar ortaya çıkmıştır. 8. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.23'teki gibidir.



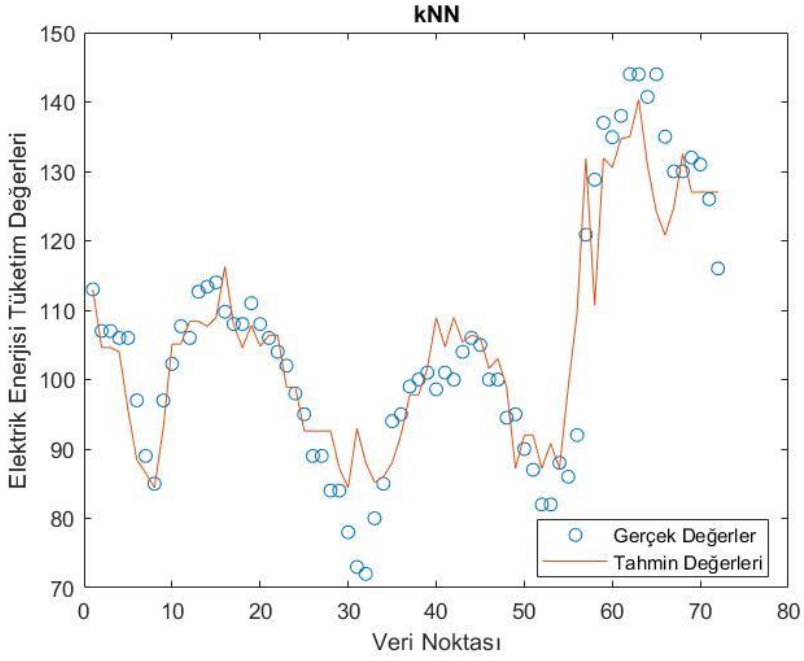
Şekil 5.23. K-en yakın komşu 8. Senaryo tahmin grafiği

8. Senaryonun çıktıları 1. Senaryonun çıktıları ile karşılaştırıldığında rüzgâr bağımsız değişkeninin algoritma üzerindeki belirleyiciliğinin düşük olduğu gözlemlenmektedir. 9. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.24'teki gibidir.



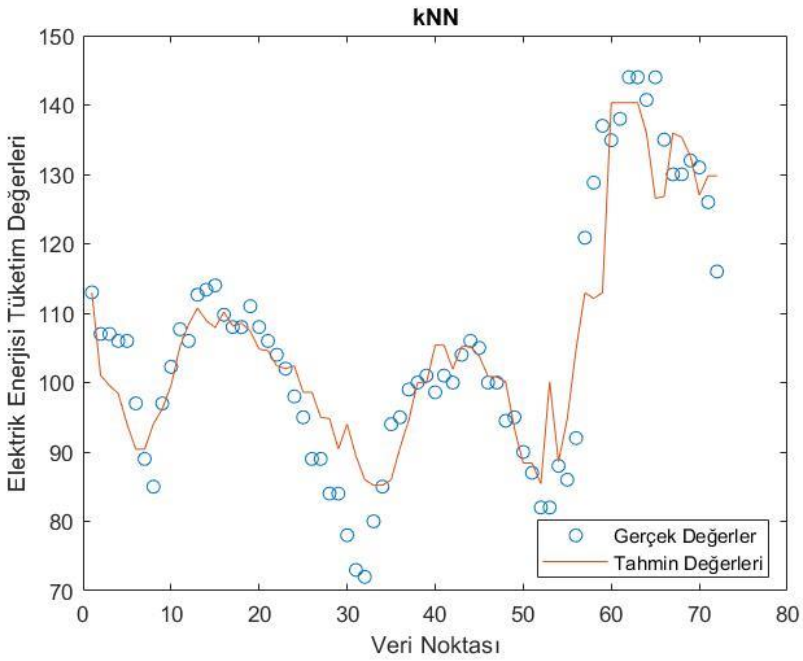
Şekil 5.24. K-en yakın komşu 9. Senaryo tahmin grafiği

10. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.25'teki gibidir.



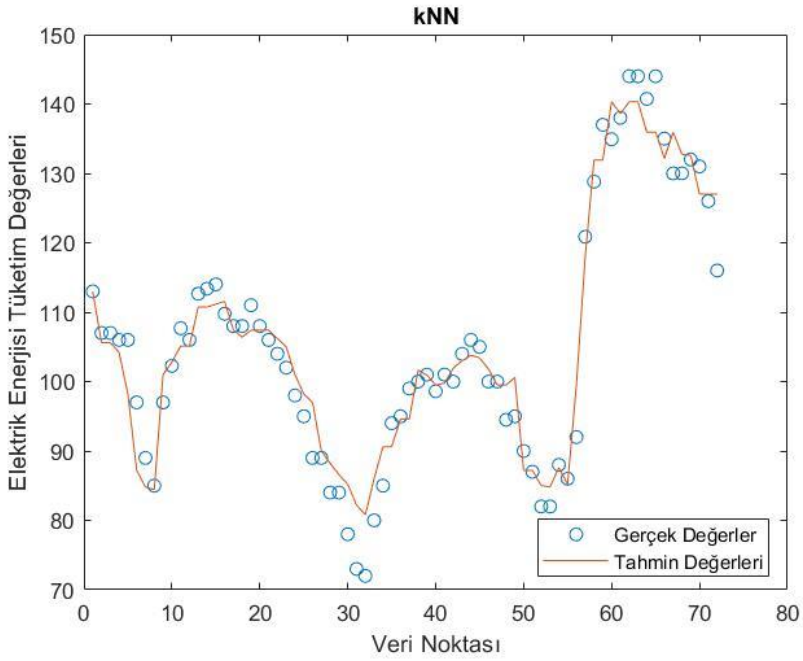
Şekil 5.25. K-en yakın komşu 10. Senaryo tahmin grafiği

11. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.26'daki gibidir.



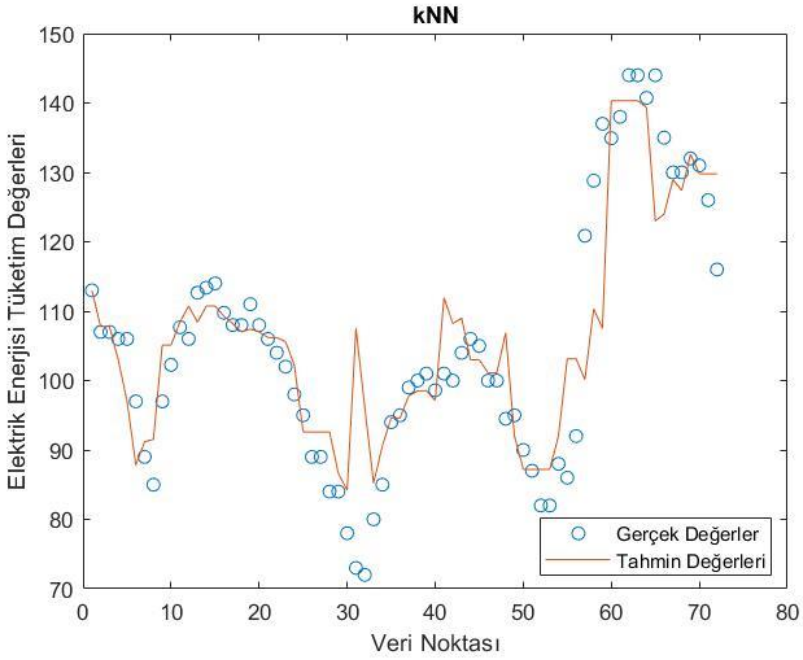
Şekil 5.26. K-en yakın komşu 11. Senaryo tahmin grafiği

9. Senaryo, 10. Senaryo ve 11. Senaryo grafikleri incelendiğinde tüm zamansal bağımsız değişkenlerin senaryoya dâhil edilmesiyle tüketim karakteristiğinin yakalanması yüksek oranda artmıştır. Ani düşüşlerde ve zirve noktalarda küçük sapmalar mevcuttur. Meteorolojik bağımsız değişkenlerin algoritma üzerindeki belirleyiciliği incelendiğinde sıcaklık değişkeninin önemi ortaya çıkmıştır. 12. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.27'deki gibidir.



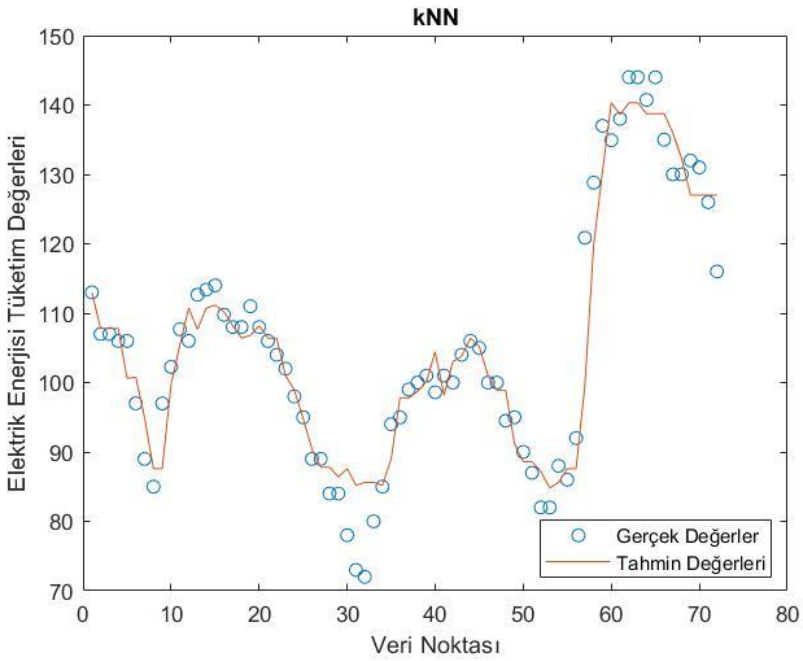
Şekil 5.27. K-en yakın komşu 12. Senaryo tahmin grafiği

13. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.28'deki gibidir.



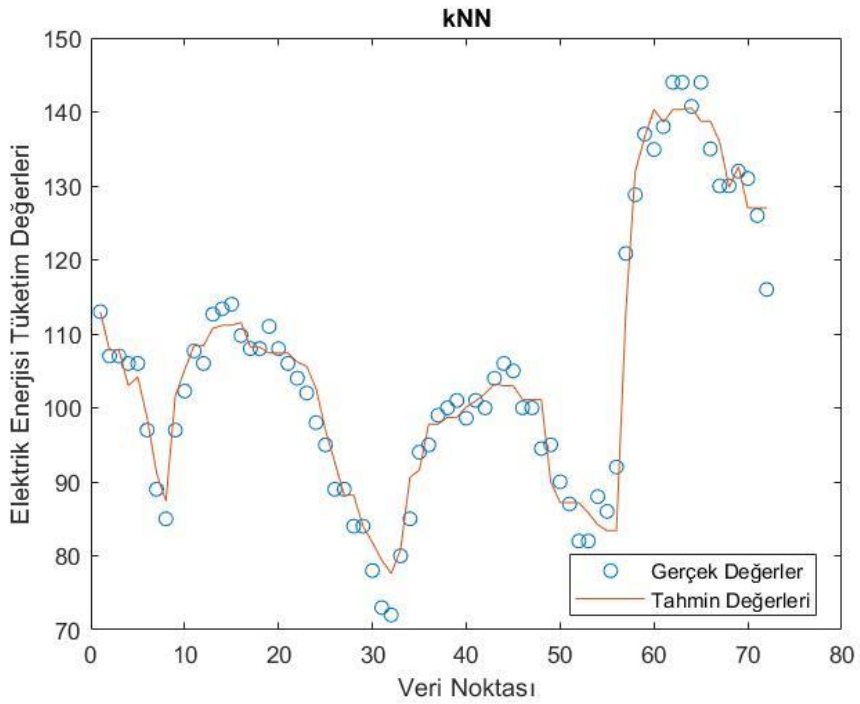
Şekil 5.28. K-en yakın komşu 13. Senaryo tahmin grafiği

14. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.29'daki gibidir.



Şekil 5.29. K-en yakın komşu 14. Senaryo tahmin grafiği

12. Senaryo, 13. Senaryo, 14. Senaryolara meteorolojik ve zaman temelli bağımsız değişkenlere ek olarak sırasıyla 1 hafta öncesinin aynı saatteki, 1 gün öncesinin aynı saatteki ve 1 saat öncesinin tüketim değeri dâhil edilmiştir. Buradaki amaç geçmiş tüketim değerlerinin belirleyicilik karşılaştırmasını yapmaktır. Bahsi geçen senaryoların sonuçları incelendiğinde bağımsız değişken olarak sonradan eklenmiş geçmiş tüketim değerlerin arasında en belirleyici olanı 1 saat öncesinin tüketim değerleridir. 12. Senaryo ve 13. Senaryoya yakın bir çıktı elde edilse de 14. Senaryonun grafiği incelendiğinde tüketim değerlerindeki dalgalanmalar oldukça başarılı bir şekilde yakalanmıştır. Dip ve zirve tüketim değerlerindeki tutarsızlığın büyük bir oranda azaldığı gözlemlenmektedir. 15. Senaryo için gerçekleştirilen KNN talep tahmini grafiği Şekil 5.30'daki gibidir.

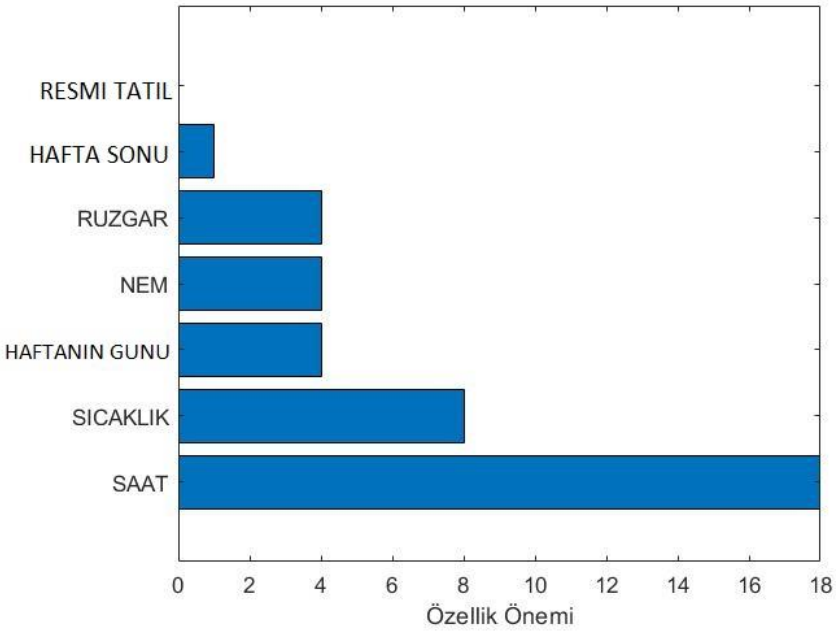


Şekil 5.30. K-en yakın komşu 15. Senaryo tahmin grafiği

15. Senaryoda tüm zamansal, meteorolojik ve ek bağımsız değişkenlerin kullanılması beklendiği şekilde sonuç vermemiştir. 14. Senaryo ile karşılaştırıldığında zirve noktalardaki tutarsızlığın arttığı gözlemlenmektedir. Model daha karmaşık bir hale gelerek hata değerlerini yükseltmiştir.

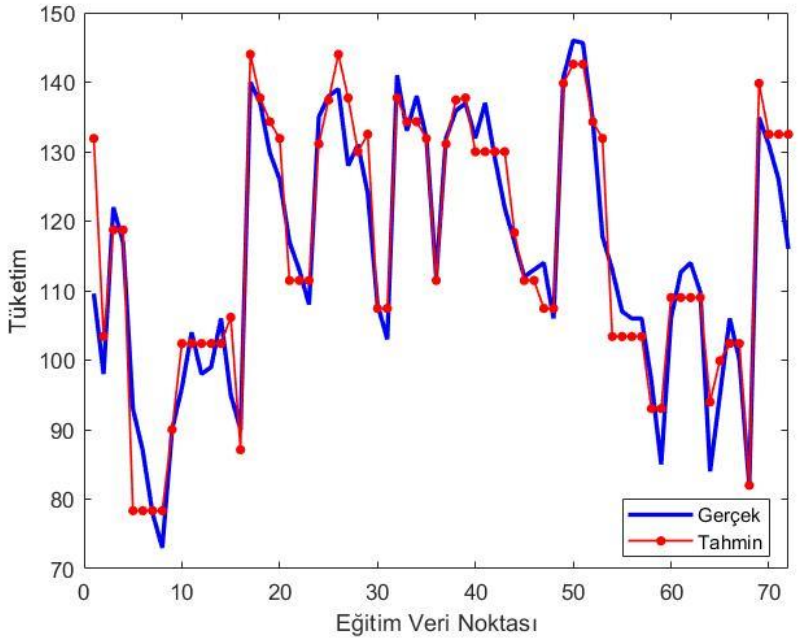
5.1.3. Karar Ağacı ile Talep Tahmini Uygulaması

Tahmin değerleri ve tüketim verilerinin grafik üzerinde anlaşılabilir olması için veri seti olarak 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler kullanılmıştır. Veri setinin %70'i eğitim, %30'u test veri seti olarak kullanılmıştır. DT'nin yaprak sayısı senaryo ve bağımsız değişkenlerinin önem katsayısına göre algoritma tarafından otomatik olarak belirlenmiştir. Yaprak sayısı dışındaki parametreler varsayılan olarak kullanılmıştır. 1. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.31'deki gibidir.



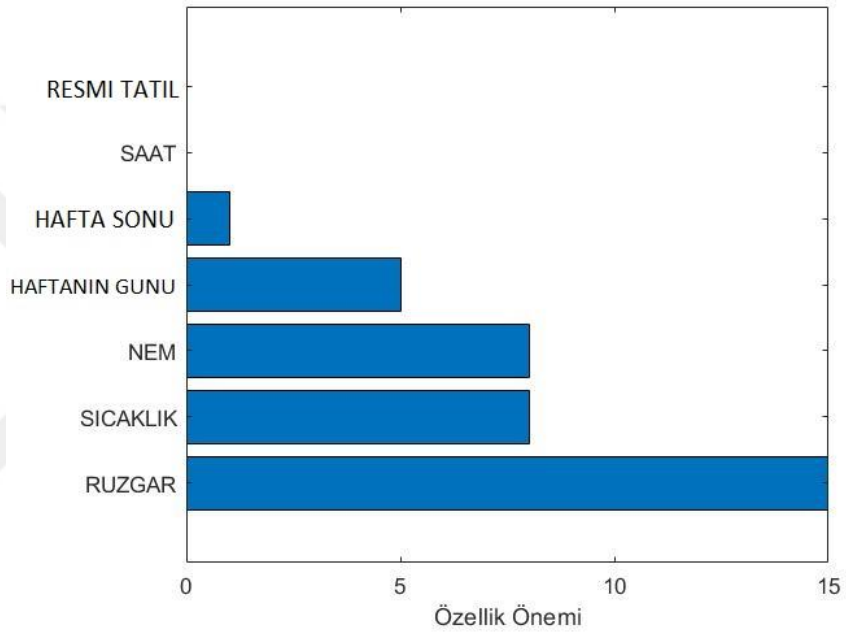
Şekil 5.31. Karar ağacı 1. Senaryo özellik önemi grafiği

1. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.32'deki gibidir.



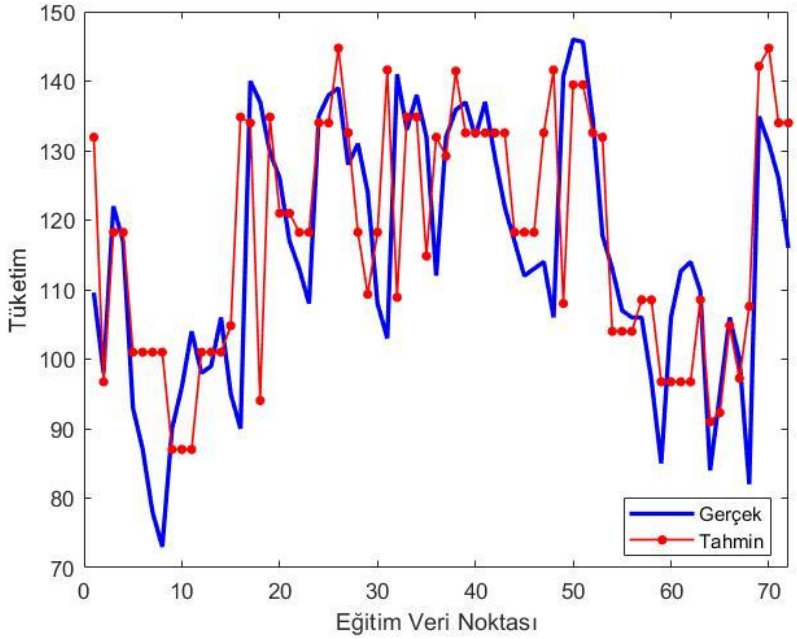
Şekil 5.32. Karar ağacı 1. Senaryo tahmin grafiği

Karar ağacı algoritmasının 1. Senaryo için tahmin grafiği incelendiğinde tahmin değerlerinin gerçek değerlerler ile uyumunun oldukça yüksek olduğu gözlemlenmektedir. Algoritma 1. Senaryo özelinde en uygun yaprak sayısını 40 olarak belirlemiştir. Tüketim değerlerinin genel davranışı başarılı bir şekilde yakalanmıştır fakat dip ve zirve noktalarda ufak sapmalar gözlemlenmektedir. 2. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.33'teki gibidir.



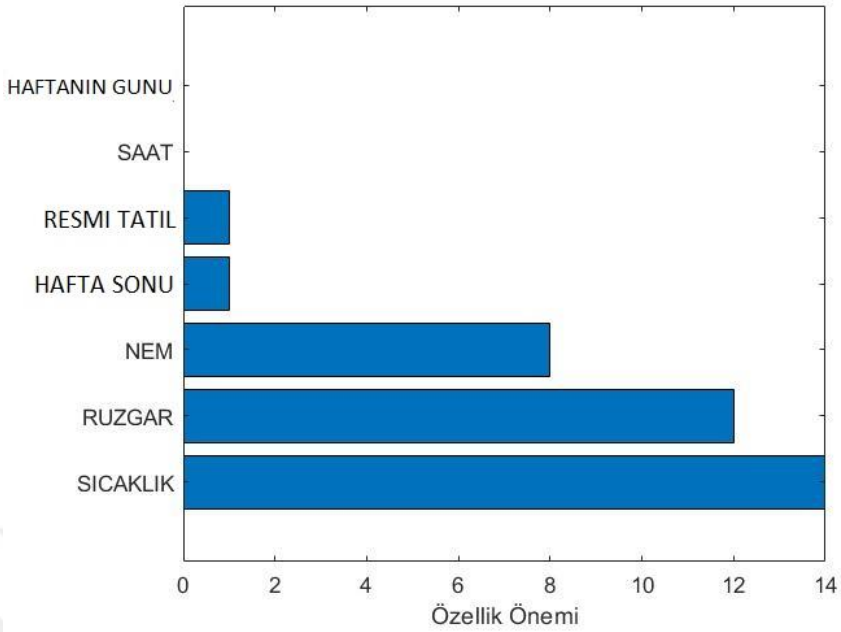
Şekil 5.33. Karar ağacı 2. Senaryo özellik önemi grafiği

2. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.34'teki gibidir.



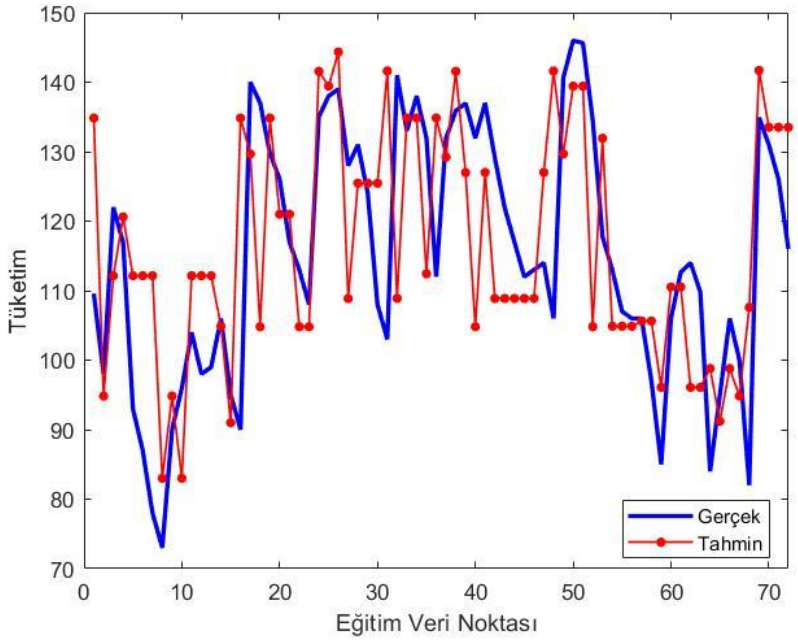
Şekil 5.34. Karar ağacı 2. Senaryo tahmin grafiği

Günün saati bağımsız değişkeni çıkarıldıktan sonra tahmin değerlerinin tutarsızlığının arttığı gözlemlenmektedir. Haftanın günü ve hafta sonu olma durumu gibi zamansal bağımsız değişkenlerin tahmin gücü sınırlı kalmıştır. Algoritma 2. Senaryo özelinde en uygun yaprak sayısını 38 olarak belirlemiştir. Tüketim değerinin dip, zirve yaptığı noktalarda ve dalgalanmalardaki sapmaların arttığı gözlemlenmektedir. 3. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.35'teki gibidir.



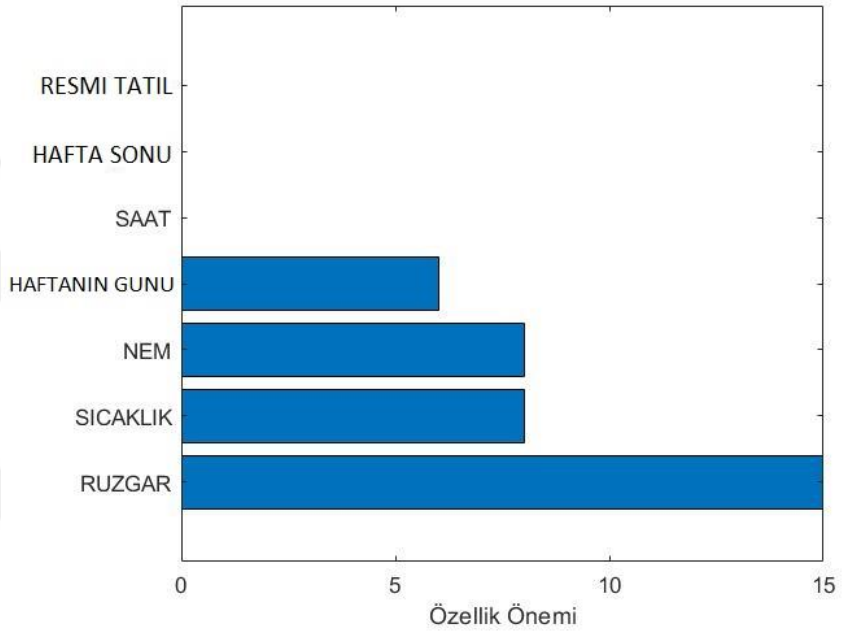
Şekil 5.35. Karar ağacı 3. Senaryo özellik önemi grafiği

3. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.36'daki gibidir.



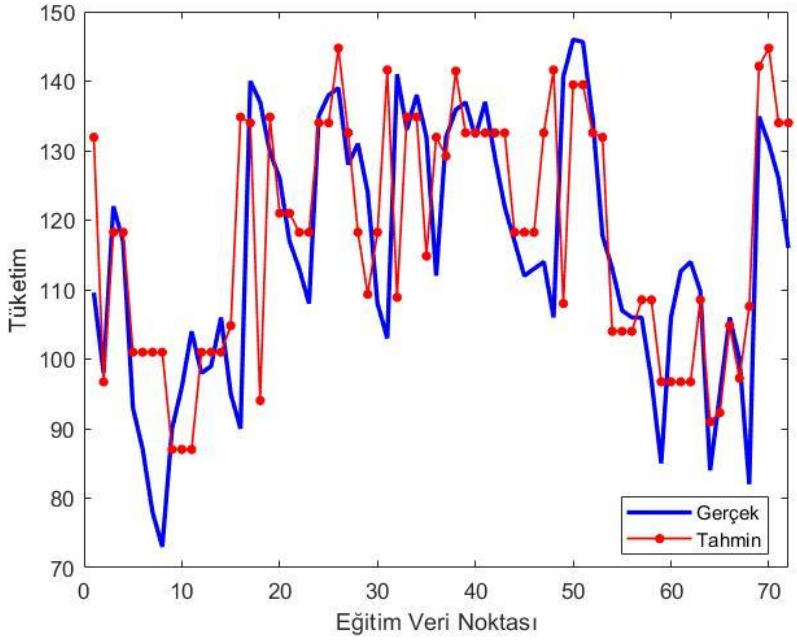
Şekil 5.36. Karar ağacı 3. Senaryo tahmin grafiği

Algoritmanın 3. Senaryo çıktıları incelendiğinde haftanın günü bağımsız değişkeninin çıkarılması ile tahmin değerlerindeki tutarsızlığın arttığı gözlemlenmektedir. Algoritma 3. Senaryo özelinde en uygun yaprak sayısını 37 olarak belirlemiştir. 4. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.37'deki gibidir.



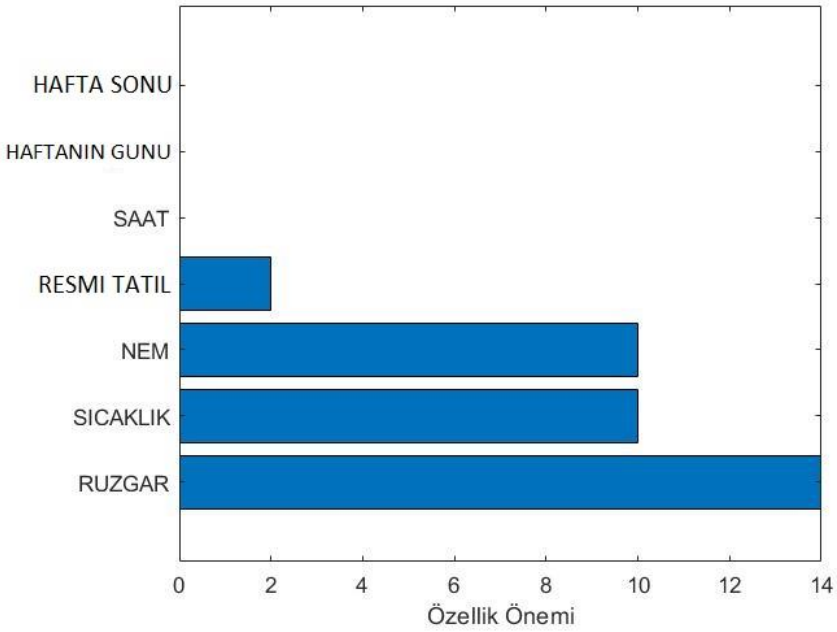
Şekil 5.37. Karar ağacı 4. Senaryo özellik önemi grafiği

4. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.38'deki gibidir.



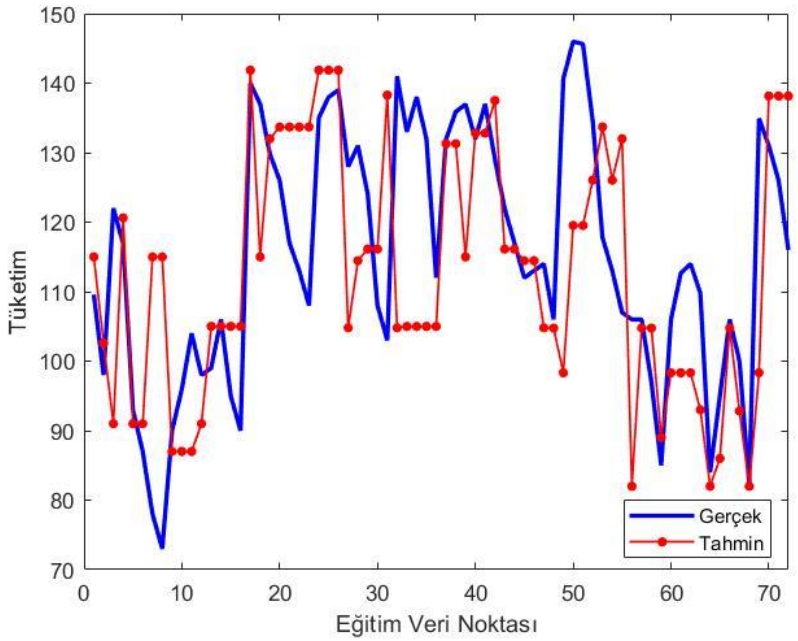
Şekil 5.38. Karar ağacı 4. Senaryo tahmin grafiği

Algoritmanın 4. Senaryo için çıktıları incelendiğinde dip ve zirve noktalarda benzer sapmaların devam ettiği gözlemlenmektedir. Haftanın günü bağımsız değişkeninin, hafta sonu olma durumu değişkenine göre belirleyiciliğinin yüksek olması genel eğilimi yakalanma performansı arttırmıştır. Algoritma 4. Senaryo özelinde en uygun yaprak sayısını 38 olarak belirlemiştir. 5. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.39'taki gibidir.



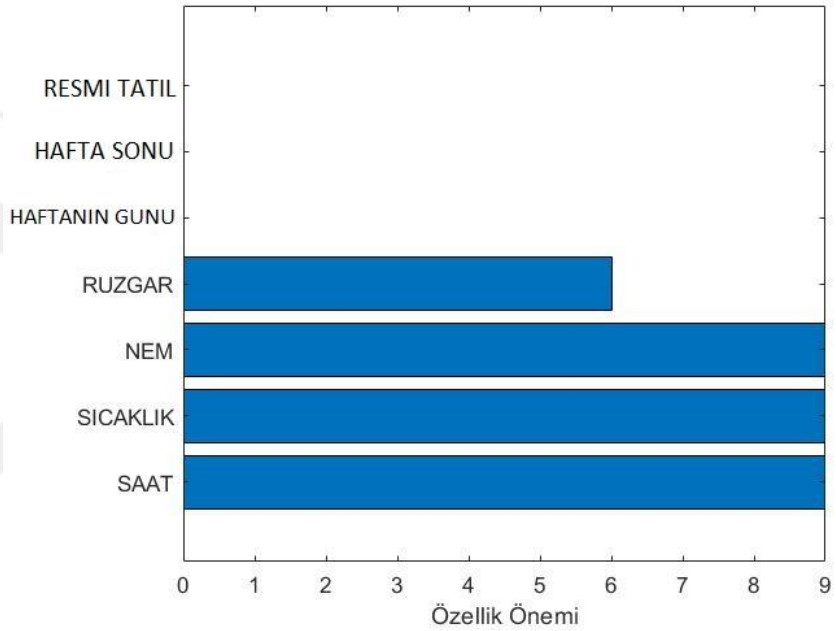
Şekil 5.39. Karar ağacı 5. Senaryo özellik önemi grafiği

5. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.40'taki gibidir.



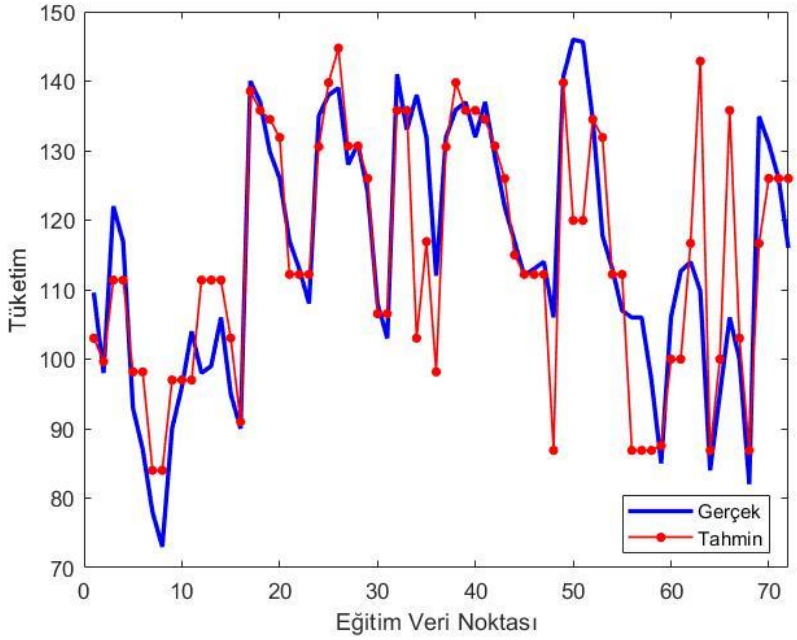
Şekil 5.40. Karar ağacı 5. Senaryo tahmin grafiği

Algoritmanın 5. Senaryo için çıktıları incelendiğinde tahmin verilerindeki sapmaların arttığı gözlemlenmektedir. Resmi tatil değişkeninin özellik önemi katsayısı 2 olsa da zamansal bağımsız değişken olarak tek başına yeterli belirleyiciliği sağlayamamıştır. Algoritma 5. Senaryo özelinde en uygun yaprak sayısını 37 olarak belirlemiştir. 6. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.41'deki gibidir.



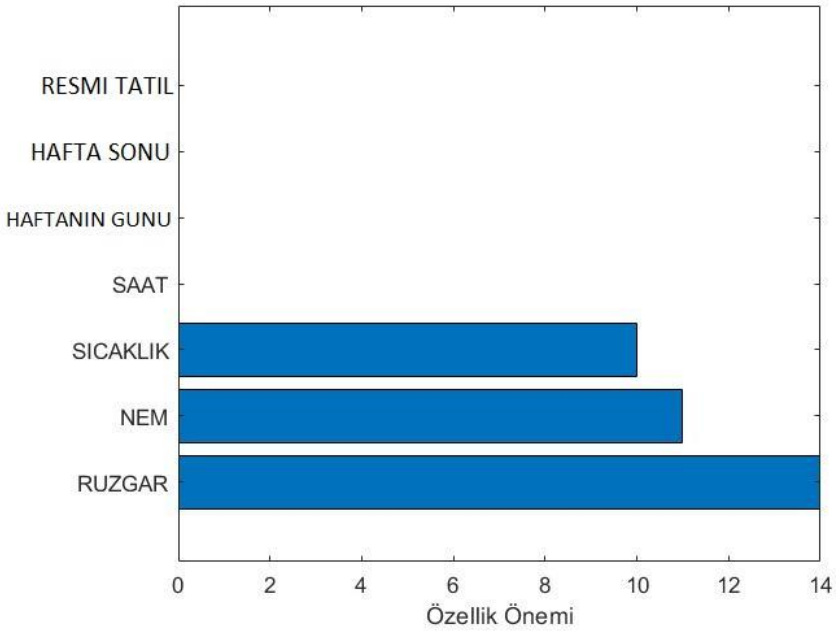
Şekil 5.41. Karar ağacı 6. Senaryo özellik önemi grafiği

6. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.42'deki gibidir.



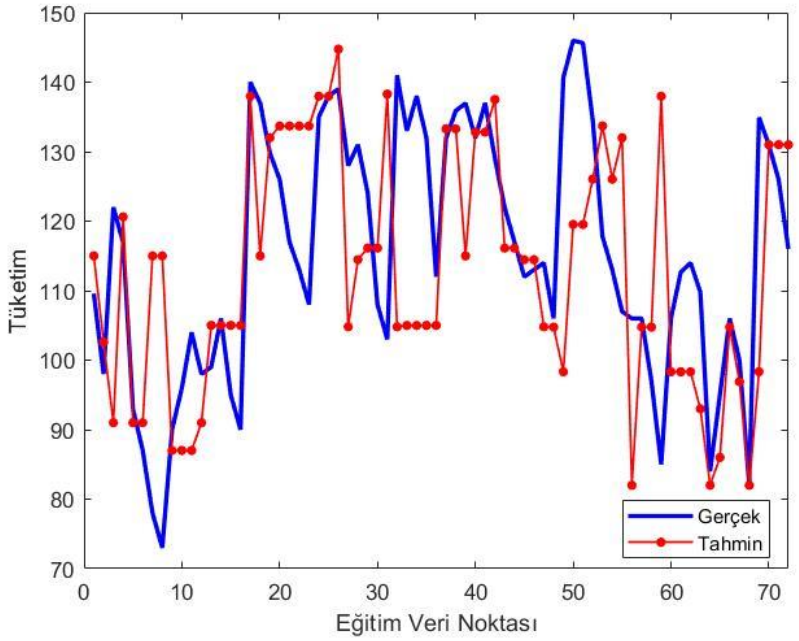
Şekil 5.42. Karar ağacı 6. Senaryo tahmin grafiği

6. Senaryonun tahmin grafiği incelendiğin günün saati bağımsız değişkeninin önemi net bir şekilde fark edilmektedir. Algoritma genel tüketim karakteristiğini oldukça başarılı bir şekilde yakalamıştır. Günün saati bağımsız değişkeninin dâhil edilmesiyle birlikte zamansal dalgalanmaların yakalanma performansı artmıştır fakat dip ve zirve noktalarda küçük sapmalar mevcuttur. Algoritma 6. Senaryo özelinde en uygun yaprak sayısını 34 olarak belirlemiştir. 7. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.43'teki gibidir.



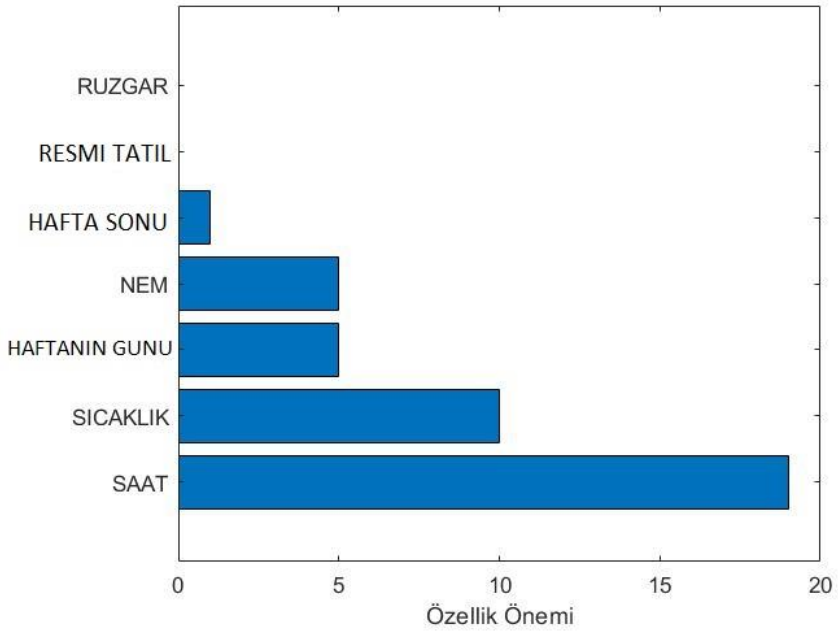
Şekil 5.43. Karar ağacı 7. Senaryo özellik önemi grafiği

7. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.44'teki gibidir.



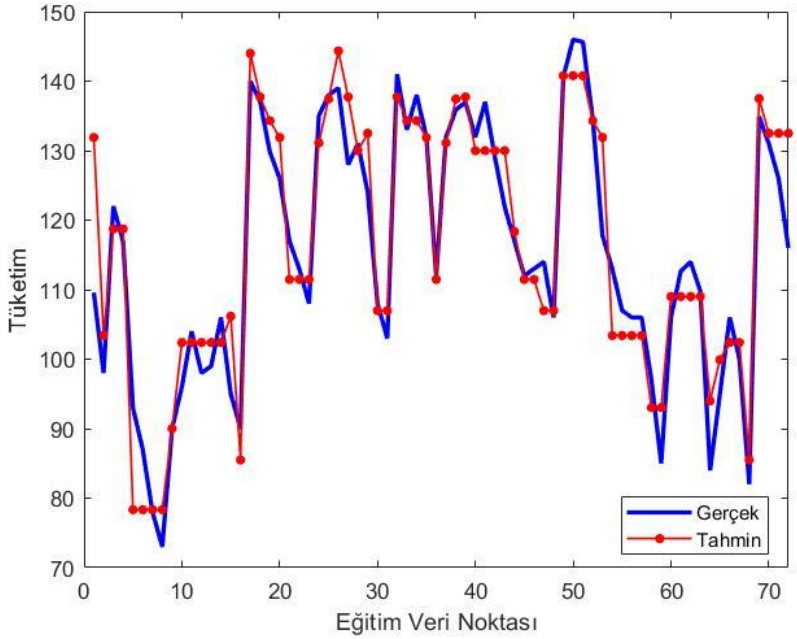
Şekil 5.44. Karar ağacı 7. Senaryo tahmin grafiği

7. Senaryonun tahmin grafiği incelendiğinde günün saati bağımsız değişkeninin dâhil edilmemesi algoritmanın performansını önemli ölçüde düşürdüğü gözlemlenmektedir. Özellikle zamansal dalgalanmaların olduğu bölgelerde sapma oranı artmıştır. Algoritma 7. Senaryo özelinde en uygun yaprak sayısını 36 olarak belirlemiştir. 8. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.45'teki gibidir.



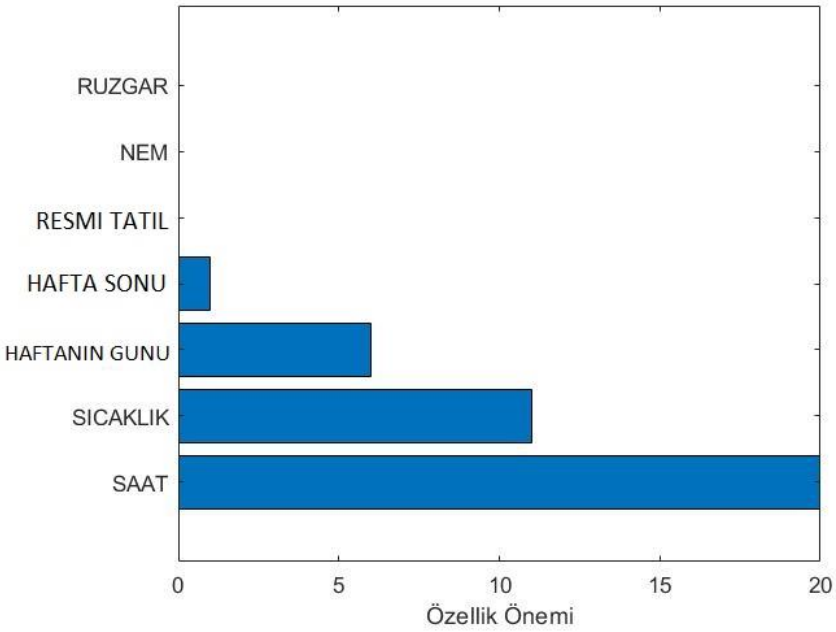
Şekil 5.45. Karar ağacı 8. Senaryo özellik önemi grafiği

8. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.46'daki gibidir.



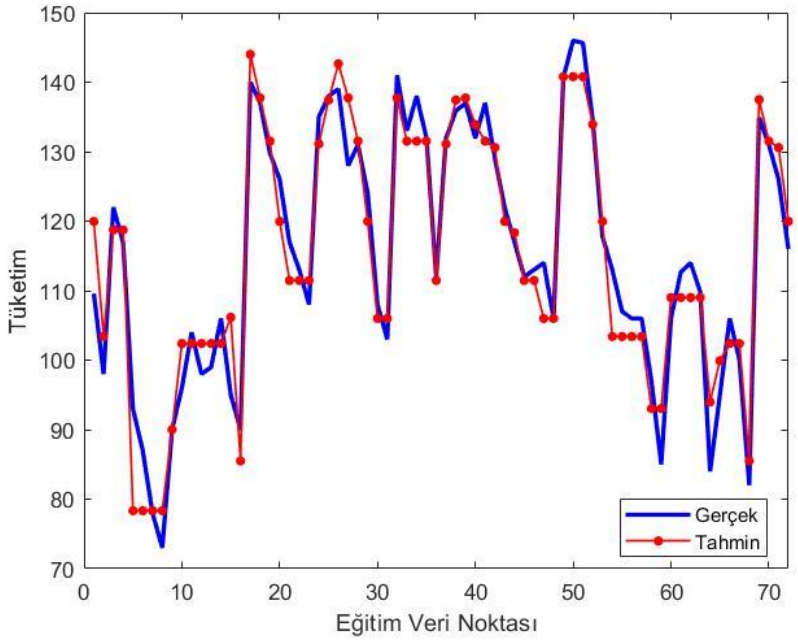
Şekil 5.46. Karar ağacı 8. Senaryo tahmin grafiği

Algoritmanın 8. Senaryo için tahmin grafiği incelendiğinde tüm zaman temelli bağımsız değişkenlerin dâhil edilmesinin olumlu etkisi açık bir şekilde gözlemlenmektedir. Tüketimin zirve yaptığı noktalarda küçük sapmalar olsa da tahmin edilen ve gerçek değerler arasındaki uyum oldukça yüksektir. Algoritma 8. Senaryo özelinde en uygun yaprak sayısını 41 olarak belirlemiştir. 9. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.47'deki gibidir.



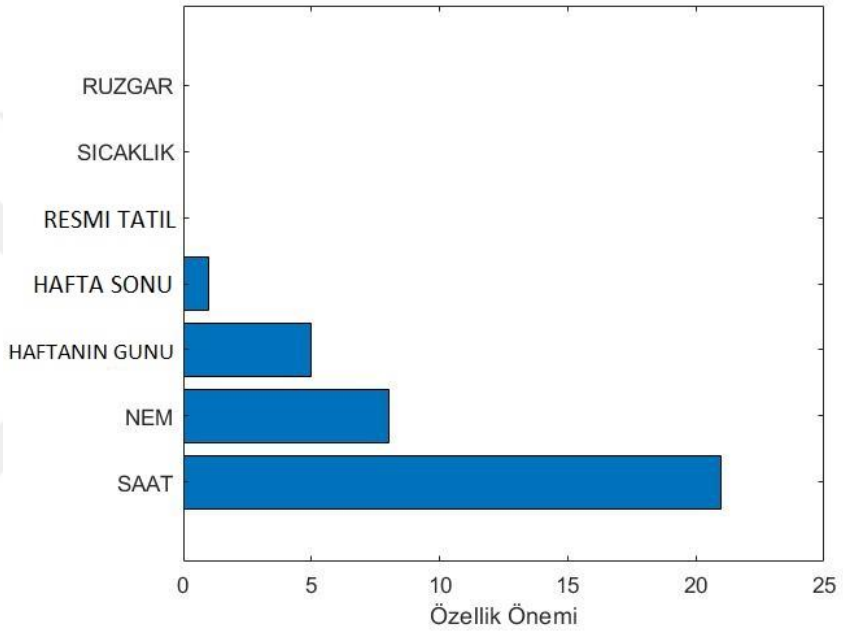
Şekil 5.47. Karar ağacı 9. Senaryo özellik önemi grafiği

9. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.48'deki gibidir.



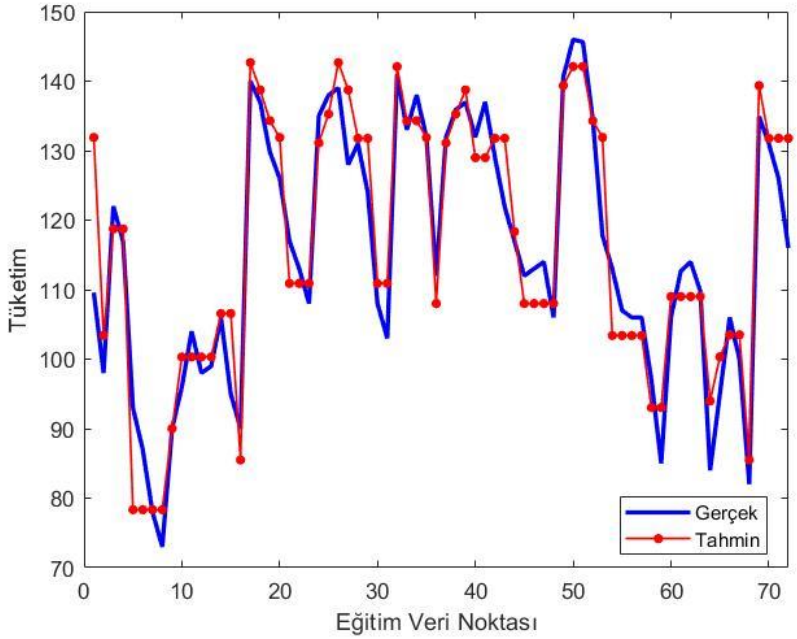
Şekil 5.48. Karar ağacı 9. Senaryo tahmin grafiği

9. Senaryo tahmin grafiği incelendiğinde belirleyiciliği yüksek değişkenler dâhil edildiği için nem bağımsız değişkeninin eksikliğinin olumsuz bir etkisi olmamıştır. 8. Senaryo ile karşılaştırıldığında algoritma genel tüketim karakteristiğini aynı şekilde yakalayabilmiştir. Algoritma 9. Senaryo özelinde en uygun yaprak sayısını 39 olarak belirlemiştir. 10. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.49'daki gibidir.



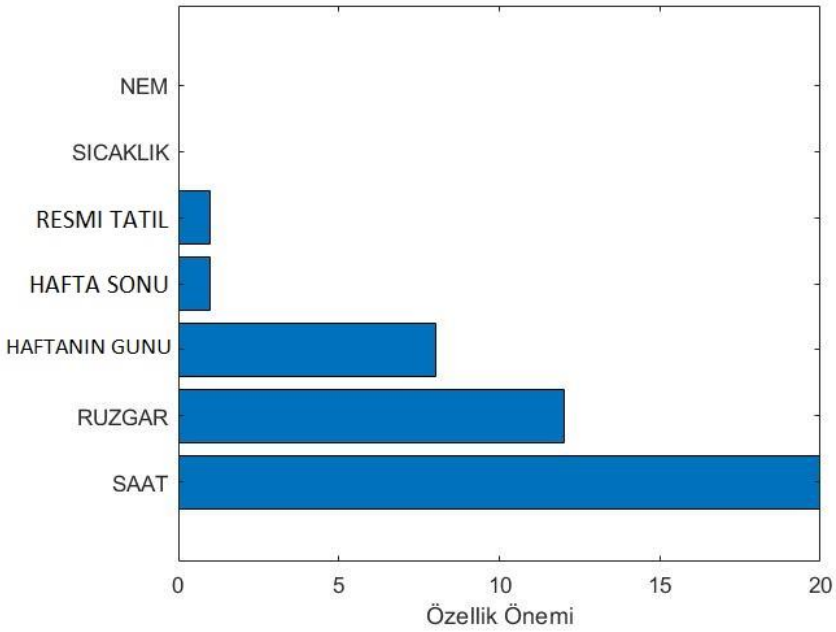
Şekil 5.49. Karar ağacı 10. Senaryo özellik önemi grafiği

10. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.50'deki gibidir.



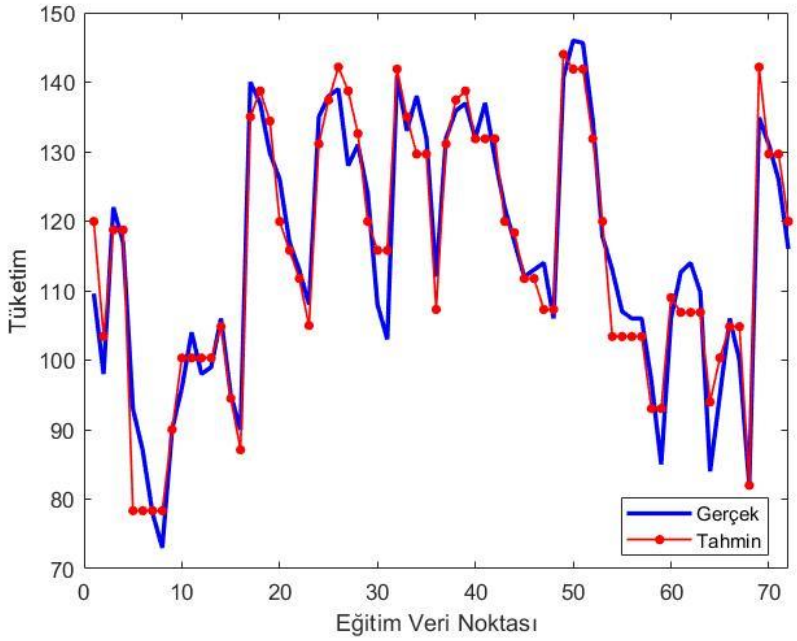
Şekil 5.50. Karar ağacı 10. Senaryo tahmin grafiği

Algoritmanın 10. Senaryo için çıktıları incelendiğinde sıcaklık ile değiştirilen nem bağımsız değişkeninin belirleyiciliğinin daha düşük olduğu gözlemlenmektedir. Gerçek ve tahmin değerleri arasındaki uyum oldukça iyi olsa da dip ve zirve değerlerde küçük sapmalar gözlemlenmektedir. Algoritma 10. Senaryo özelinde en uygun yaprak sayısını 36 olarak belirlemiştir. 11. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.51'deki gibidir.



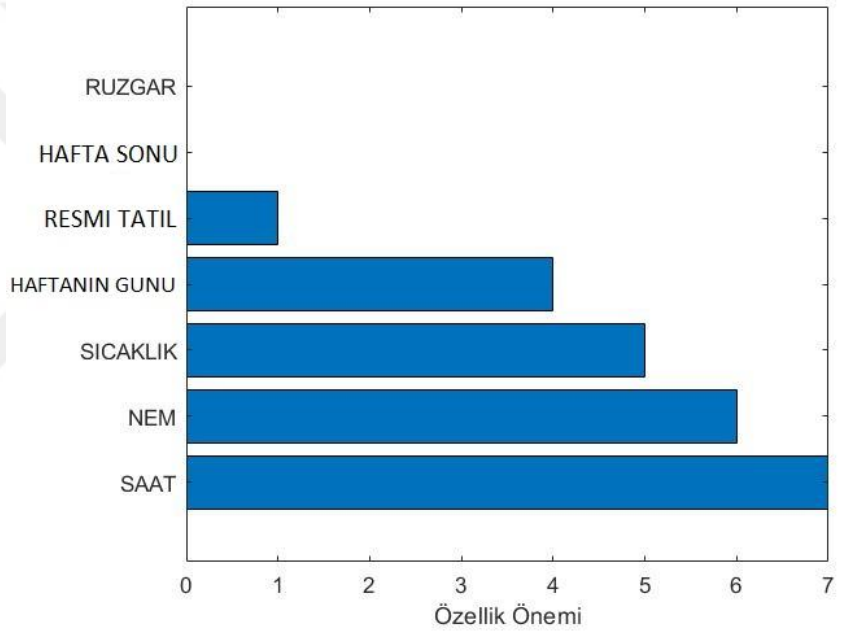
Şekil 5.51. Karar ağacı 11. Senaryo özellik önemi grafiği

11. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.52'deki gibidir.



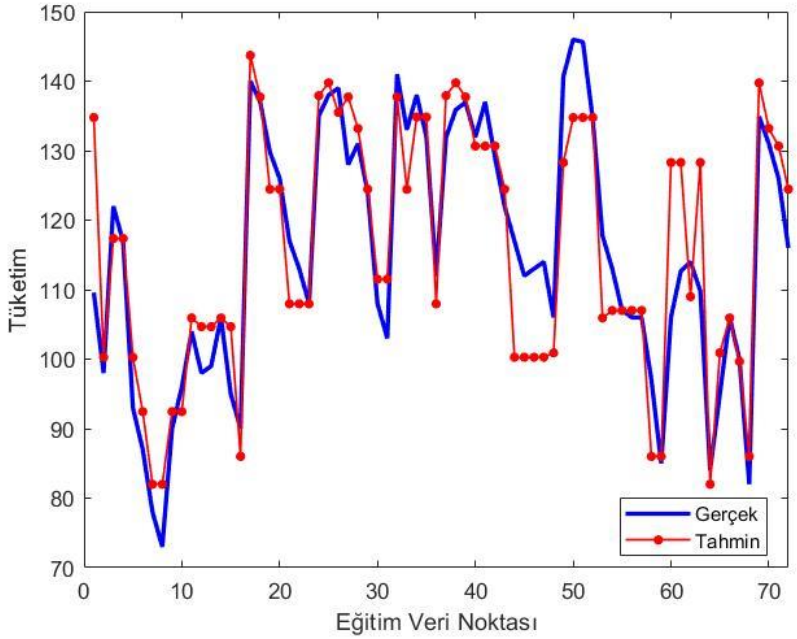
Şekil 5.52. Karar ağacı 11. Senaryo tahmin grafiği

11. Senaryonun tahmin grafiği 9. Senaryo ve 10. Senaryonun grafikleri ile karşılaştırıldığı zaman sıcaklık bağımsız değişkeninin belirleyiciliğinin daha iyi olduğu anlaşılmaktadır. Sıcaklık, nem ve rüzgâr hızı bağımsız değişkenlerinin belirleyiciliği yakın olduğu için tüketim karakteristiği aynı başarı ile yakalanmıştır. Algoritma 11. Senaryo özelinde en uygun yaprak sayısını 43 olarak belirlemiştir. 12. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.53'teki gibidir. Burada HAFTA, 1 hafta öncesinin aynı saatteki tüketim değeri bağımsız değişkenini ifade etmektedir.



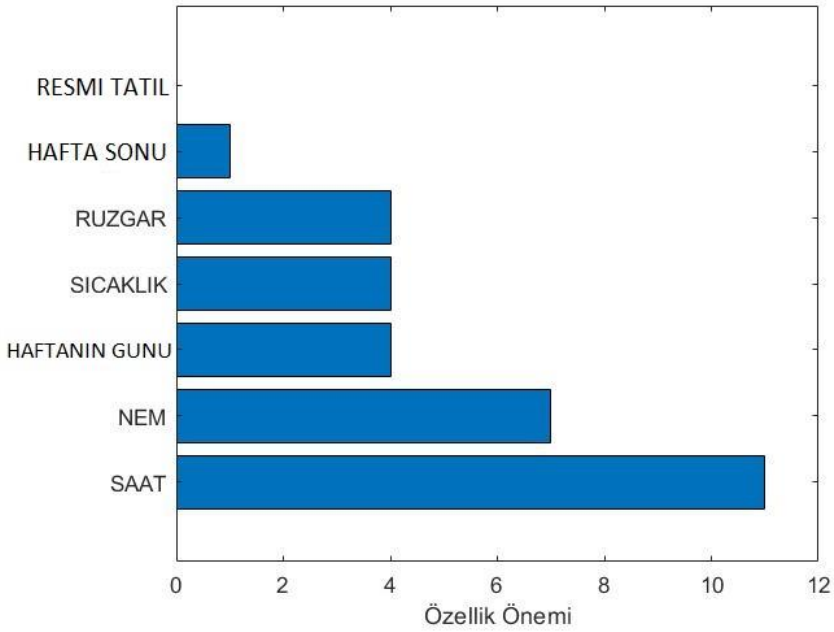
Şekil 5.53. Karar ağacı 12. Senaryo özellik önemi grafiği

12. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.54'teki gibidir.



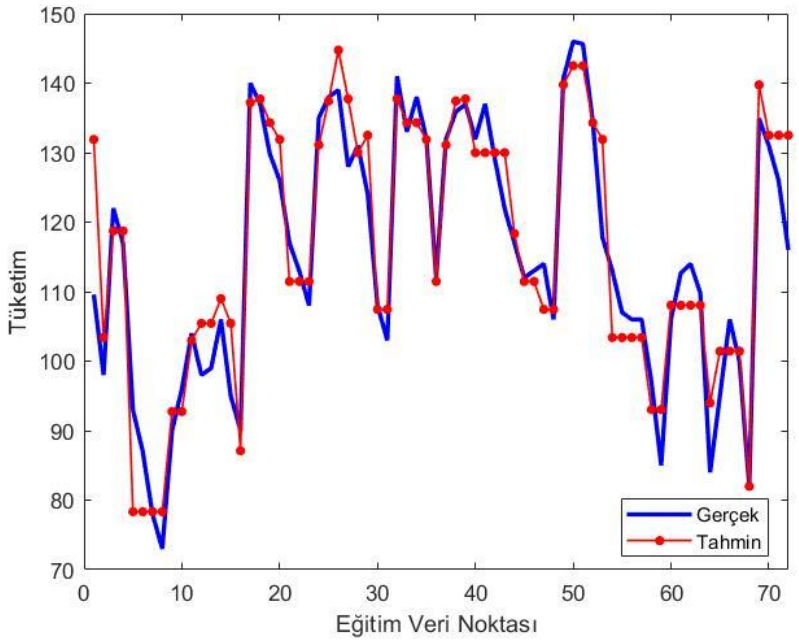
Şekil 5.54. Karar ağacı 12. Senaryo tahmin grafiği

12. Senaryo tahmin grafiği incelendiğinde 1 hafta öncesinin aynı saatteki tüketim değeri bağımsız değişkeni en yüksek öneme sahip olsa da 1. Senaryo tahmin grafiği ile karşılaştırıldığında genel uyumu algılamadaki sapma oranlarının arttığı gözlemlenmektedir. Algoritma 12. Senaryo özelinde en uygun yaprak sayısını 33 olarak belirlemiştir. 13. Senaryo için gerçekleştirilen DT talep tahmininde kullanılan bağımsız değişkenlerin önem sıralaması Şekil 5.55'teki gibidir. Burada GUN, 1 gün öncesinin aynı saatteki tüketim değeri bağımsız değişkenini ifade etmektedir.



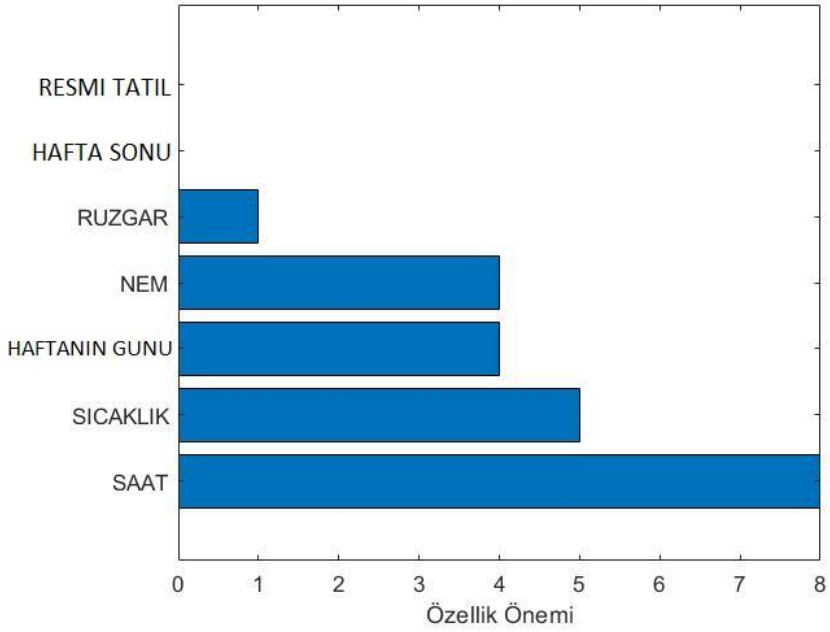
Şekil 5.55. Karar ağacı 13. Senaryo özellik önemi grafiği

13. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.56'daki gibidir.



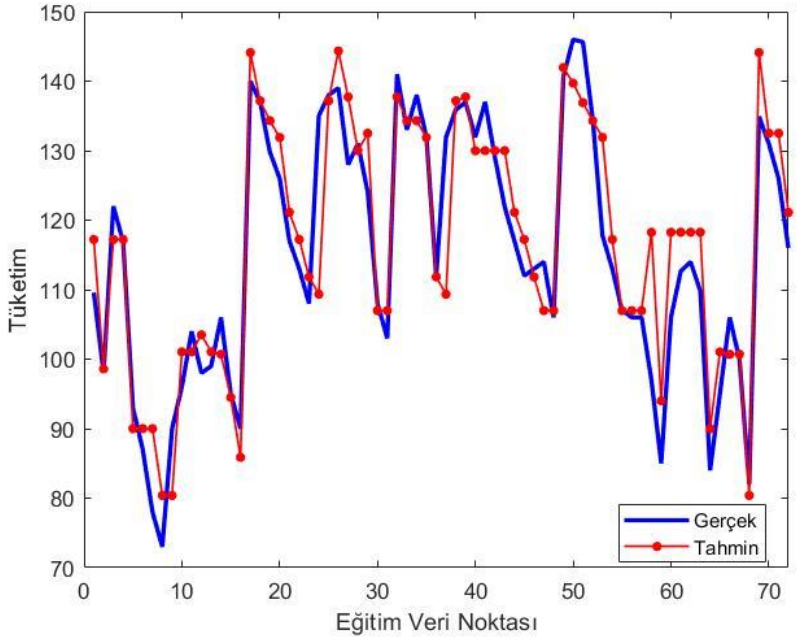
Şekil 5.56. Karar ağacı 13. Senaryo tahmin grafiği

13. Senaryo tahmin grafiği incelendiğinde 1 gün öncesinin aynı saatteki tüketim değeri bağımsız değişkeninin büyük bir etkisinin olmadığı gözlemlenmektedir. 1. Senaryo tahmin grafiği ile karşılaştırıldığında genel uyum ve sapma oranları benzer şekildedir. Algoritma 13. Senaryo özelinde en uygun yaprak sayısını 37 olarak belirlemiştir. 14. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.57'deki gibidir. Burada SAAT1, 1 saat önceki tüketim bağımsız değişkenini ifade etmektedir.



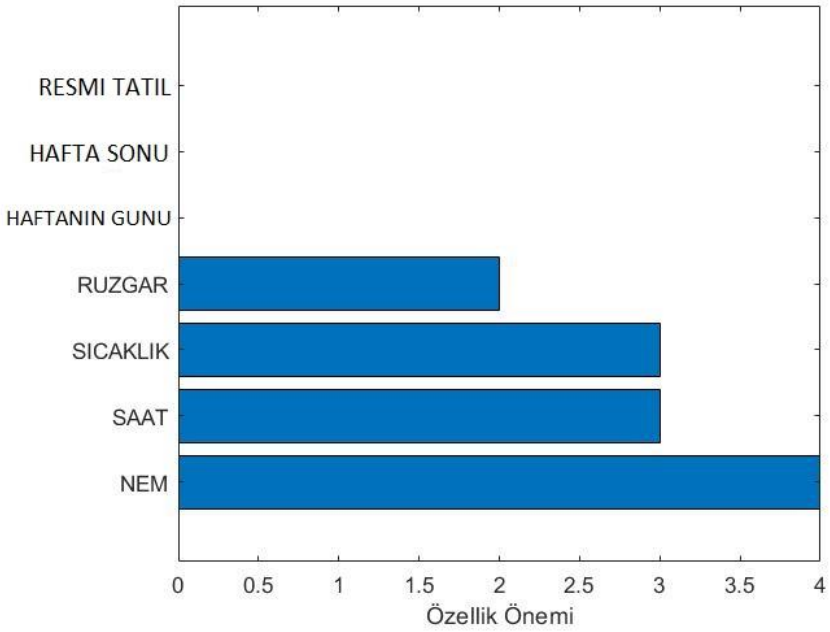
Şekil 5.57. Karar ağacı 14. Senaryo özellik önemi grafiği

14. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.58'deki gibidir.



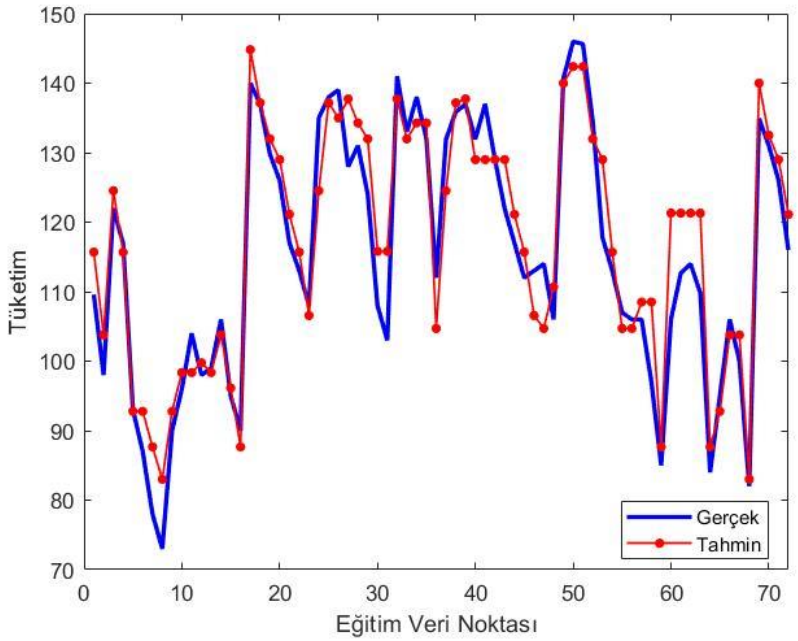
Şekil 5.58. Karar ağacı 14. Senaryo tahmin grafiği

Algoritmanın 14. Senaryo için olan tahmin grafiği incelendiğinde, 1 saat öncesinin tüketim değeri bağımsız değişkeninin en yüksek özellik önemi katsayısına sahip olmasına rağmen tahmin performansına ekstra bir etkisinin olmadığı gözlemlenmektedir. Algoritma 14. Senaryo özelinde en uygun yaprak sayısını 33 olarak belirlemiştir. 15. Senaryo için gerçekleştirilen DT talep tahmininde bağımsız değişkenlerin önem sıralaması Şekil 5.59'daki gibidir.



Şekil 5.59. Karar ağacı 15. Senaryo özellik önemi grafiği

15. Senaryo için gerçekleştirilen DT talep tahmini grafiği Şekil 5.60'taki gibidir.

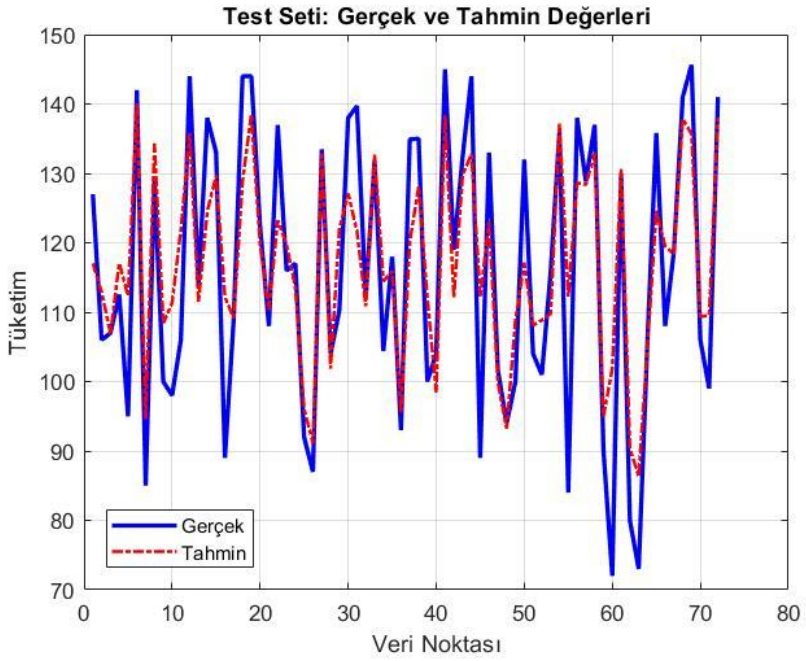


Şekil 5.60. Karar ağacı 15. Senaryo tahmin grafiği

15. Senaryonun tahmin grafiği incelendiğinde tüm ek bağımsız değişkenlerin dâhil edilmesiyle birlikte algoritmanın genel davranışı yakalama performansında artış gözlemlenmektedir. Algoritma 15. Senaryo özelinde en uygun yaprak sayısını 37 olarak belirlemiştir.

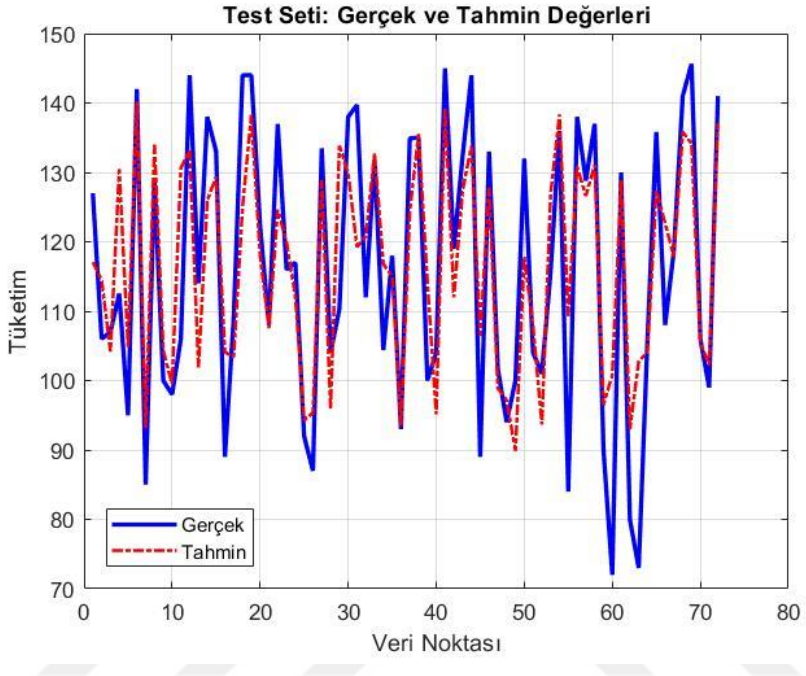
5.1.4. Destek Vektör Makinesi ile Talep Tahmini Uygulaması

Tahmin değerleri ve tüketim verilerinin grafik üzerinde anlaşılabilir olması için veri seti olarak 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler kullanılmıştır. Veri setinin %70'i eğitim, %30'u test veri seti olarak kullanılmıştır. SVM'nin kernel fonksiyonu Gaussian RBF olarak belirlenmiştir. C hiper parametresi 1.349 olarak belirlenmiştir. RBF kernelinin genişliğini belirleyen kernel ölçek değeri 0.66 olarak belirlenmiştir. Tahmin edilen değerlerin gerçek değere ne kadar yaklaşacağını belirleyen epsilon hiper parametresi 13.49 olarak belirlenmiştir. 1. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.61'deki gibidir.



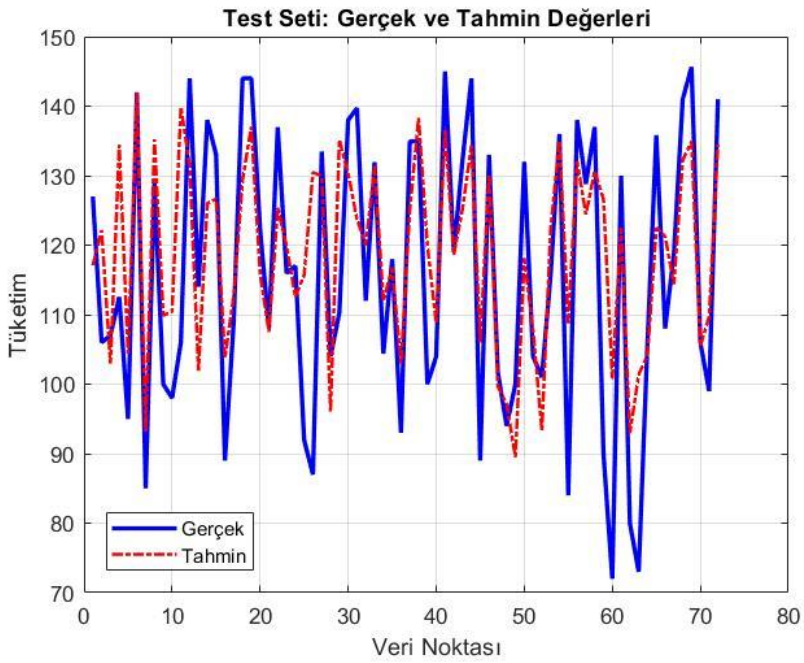
Şekil 5.61. Destek vektör makinesi 1. Senaryo tahmin grafiği

1. Senaryonun tahmin grafiği incelendiğinde tahmin eğrisinin gerçek değerleri iyi bir şekilde takip ettiği gözlemlenmektedir. Dip ve zirve noktalarda sapmalar mevcuttur. 2. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.62'deki gibidir.



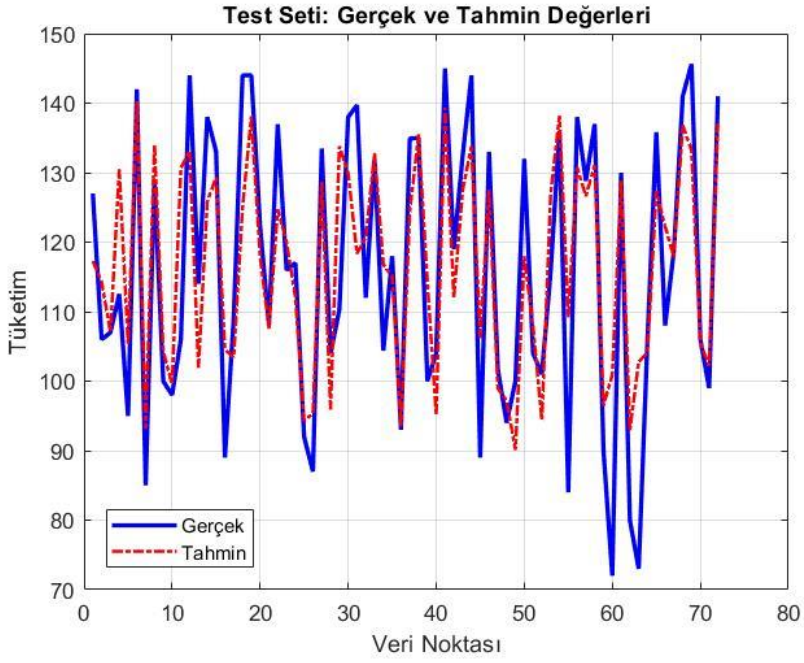
Şekil 5.62. Destek vektör makinesi 2. Senaryo tahmin grafiği

2. Senaryonun tahmin grafiği incelendiğinde günün saati bağımsız değişkeninin dâhil edilmemesiyle birlikte dip ve zirve değerlerdeki sapmaların arttığı gözlemlenmektedir. Günün saati bağımsız değişkeninin çıkarılması diğer algoritmalarda olduğu gibi büyük bir performans düşüşüne neden olmamıştır. 3. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.63'teki gibidir.



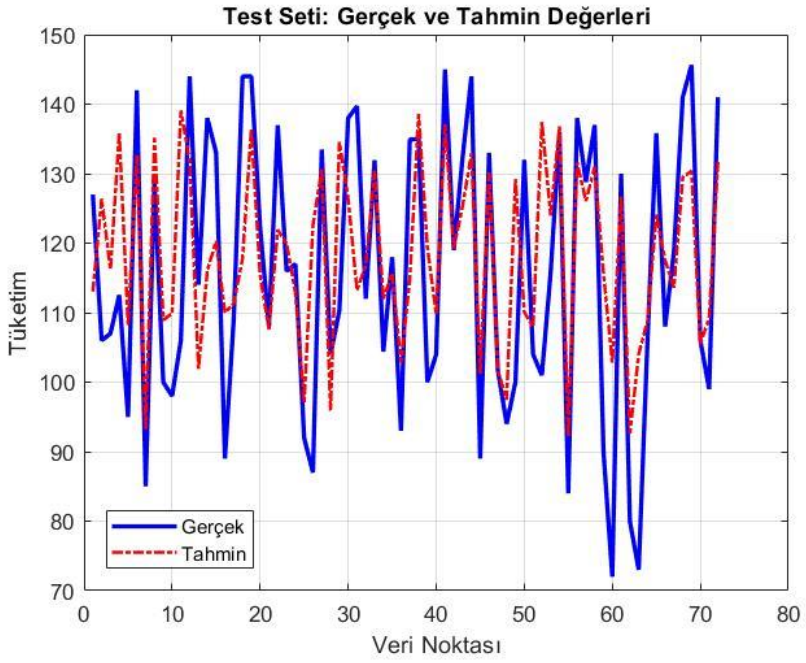
Şekil 5.63. Destek vektör makinesi 3. Senaryo tahmin grafiği

4. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.64'teki gibidir.



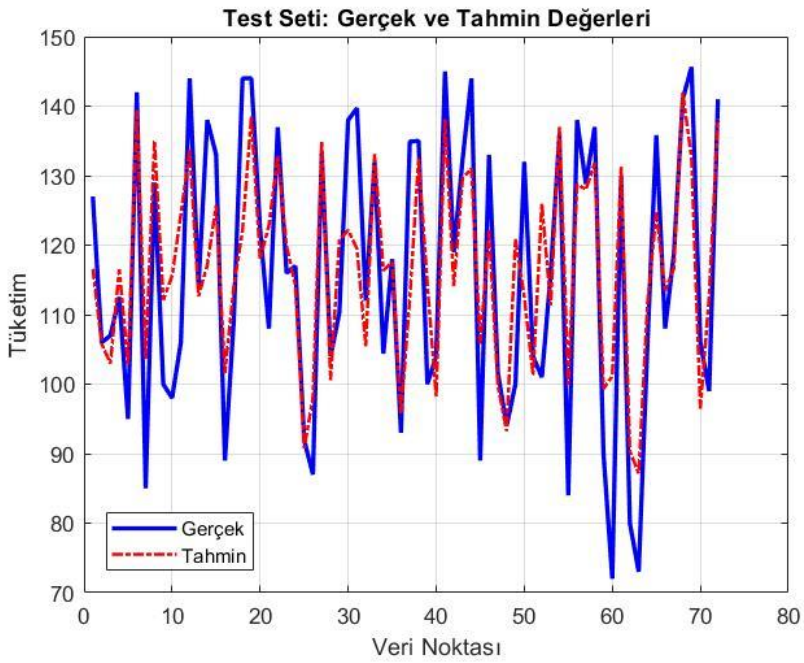
Şekil 5.64. Destek vektör makinesi 4. Senaryo tahmin grafiği

3. Senaryo ve 4. Senaryo çıktıları incelendiğinde haftanın günü bağımsız değişkeninin, hafta sonu olma durumu bağımsız değişkenine göre belirleyiciliğinin daha yüksek olduğu anlaşılmaktadır. Tahmin grafikleri karşılaştırıldığında 4. Senaryodaki gerçek ve tahmin değerleri arasındaki uyum daha iyidir, dip ve zirve noktalardaki tutarsızlık daha azdır. 5. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.65'teki gibidir.



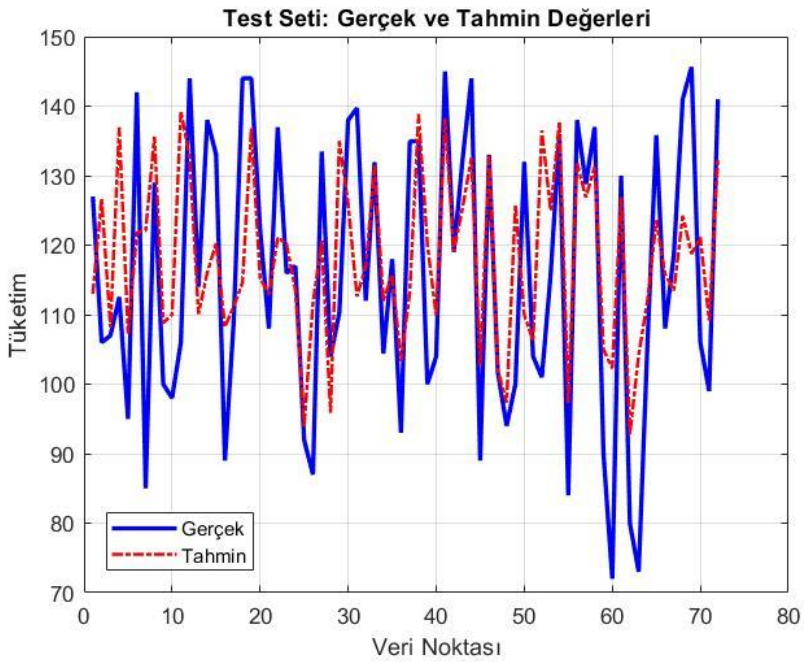
Şekil 5.65. Destek vektör makinesi 5. Senaryo tahmin grafiği

5. Senaryo grafiği incelendiğinde zaman temelli resmi tatil olma durumu bağımsız değişkeninin tek başına zamansal dalgalanmaları yakalama konusunda yeterli olmadığı gözlemlenmektedir. Tahmin eğrisi gerçek değerlerden büyük sapmalar göstermiştir. 6. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.66'daki gibidir.



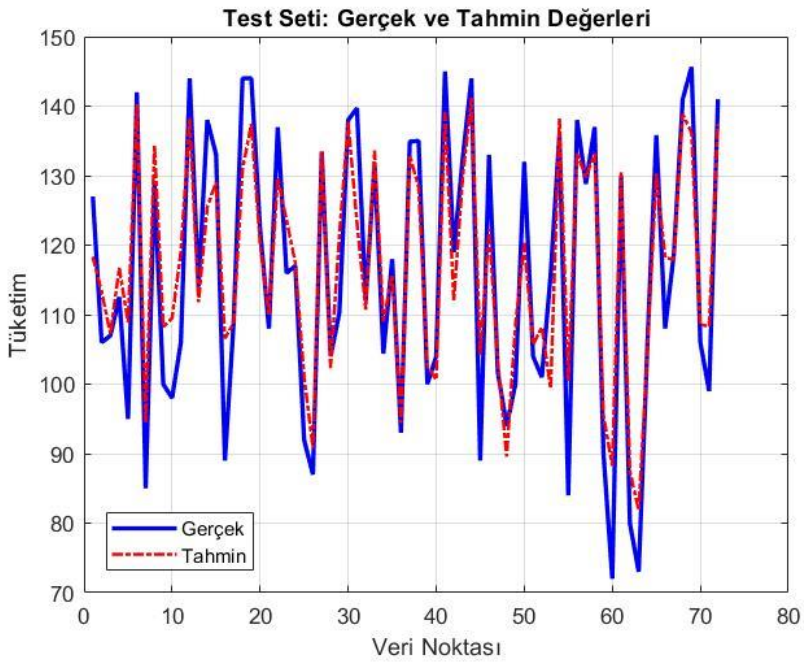
Şekil 5.66. Destek vektör makinesi 6. Senaryo tahmin grafiği

6. Senaryonun tahmin grafiği incelendiğinde günün saati bağımsız değişkeninin dâhil edilmesiyle birlikte algoritmanın genel tüketim davranışlarını yakalama performansının arttığı gözlemlenmektedir. Dip ve zirve değerlerde sapmalar mevcuttur. 7. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.67'deki gibidir.



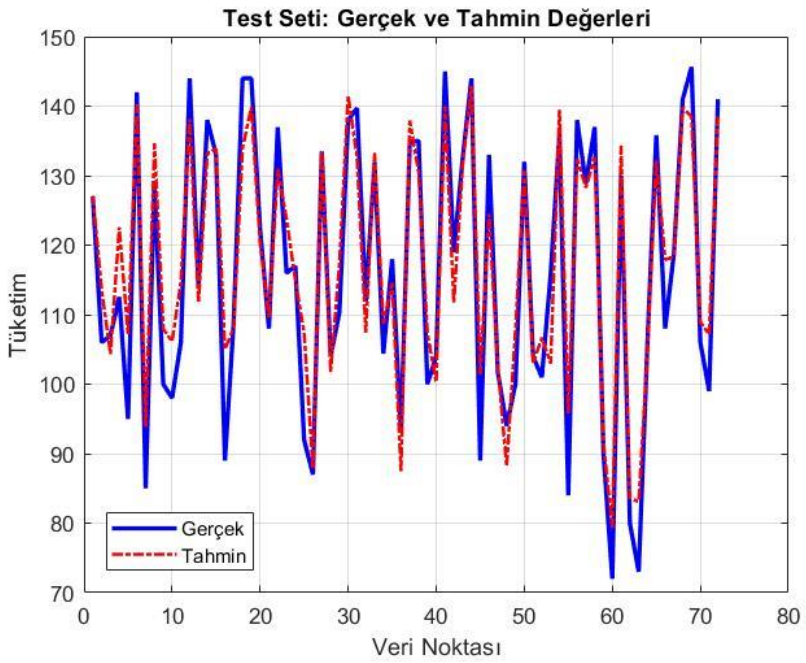
Şekil 5.67. Destek vektör makinesi 7. Senaryo tahmin grafiği

7. Senaryonun tahmin grafiği incelendiğinde günün saati bağımsız değişkeninin, zamansal dalgalanmaları yakalama konusundaki önemini tekrar görmekteyiz. 6. Senaryo ile karşılaştırıldığında 7. Senaryonun gerçek ve tahmin değerleri arasında ciddi sapmalar olduğu gözlemlenmektedir. 8. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.68'deki gibidir.



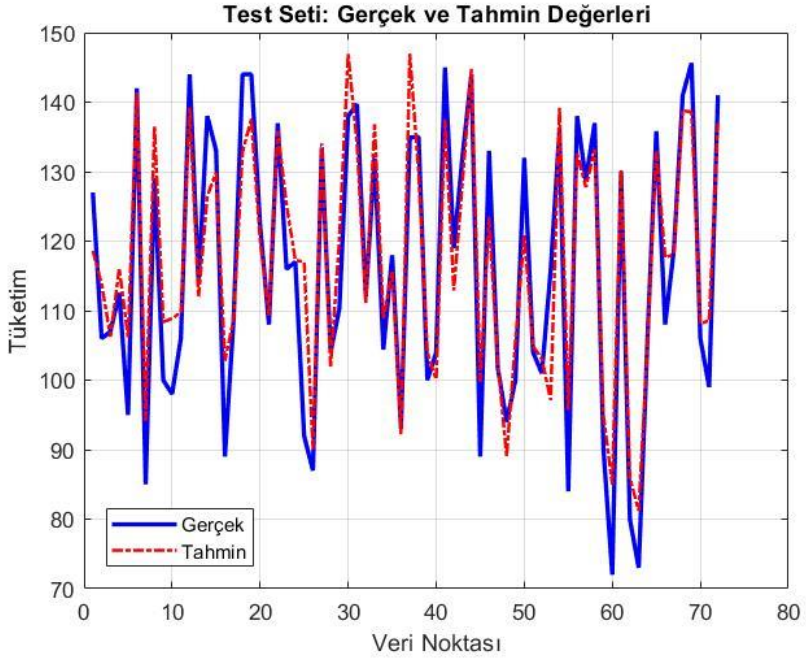
Şekil 5.68. Destek vektör makinesi 8. Senaryo tahmin grafiği

8. Senaryo tahmin grafiği incelendiğinde rüzgâr bağımsız değişkeninin dâhil edilmemesi algoritmanın performansını arttırmıştır. 1. Senaryonun tahmin grafiğiyle karşılaştırıldığında genel tüketim karakteristiği benzer bir şekilde yakalanmış olsa da 8. Senaryoda dip ve zirve değerlerdeki uyum daha yüksektir. 9. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.69'daki gibidir.



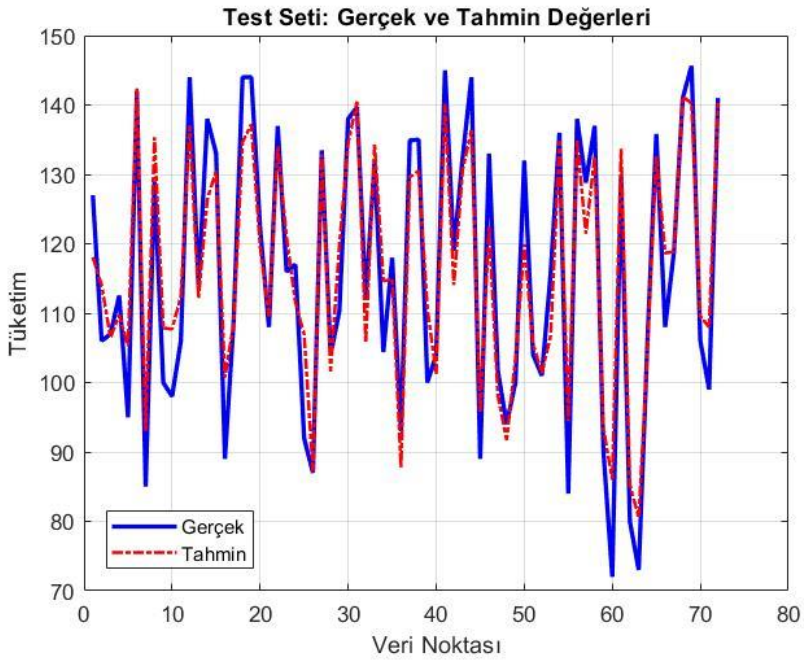
Şekil 5.69. Destek vektör makinesi 9. Senaryo tahmin grafiği

10. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.70'teki gibidir.



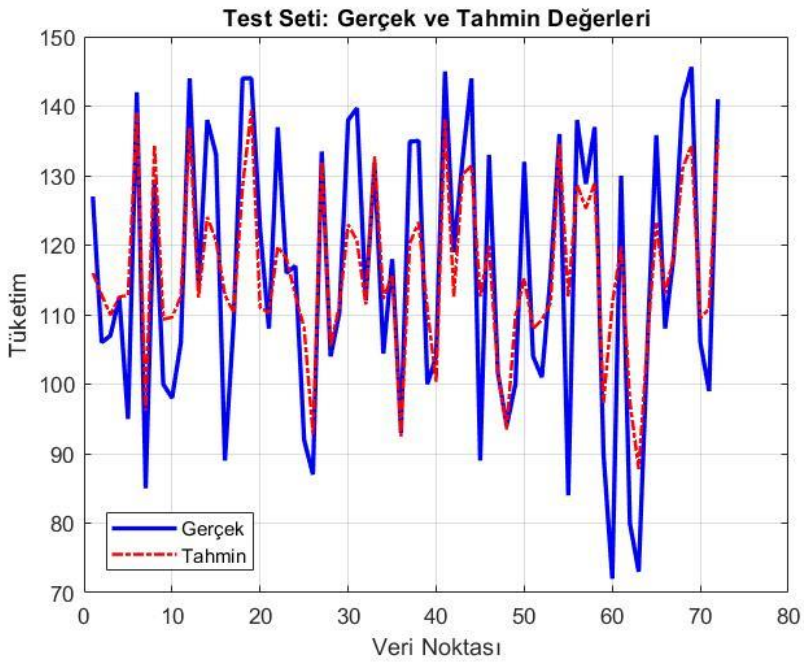
Şekil 5.70. Destek vektör makinesi 10. Senaryo tahmin grafiği

11. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.71'deki gibidir.



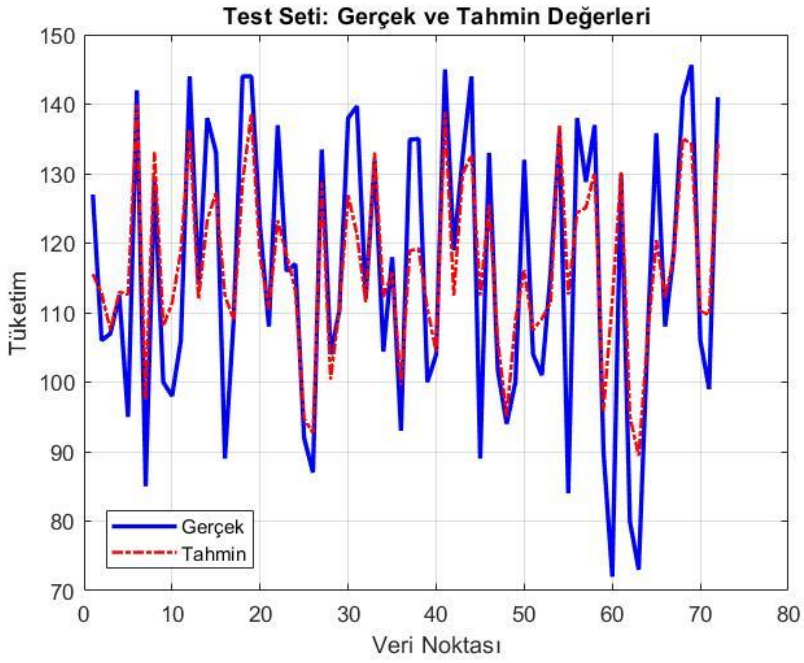
Şekil 5.71. Destek vektör makinesi 11. Senaryo tahmin grafiği

9. Senaryo, 10. Senaryo ve 11. Senaryo çıktıları incelendiğinde belirleyiciliği en yüksek meteorolojik bağımsız değişkenin sıcaklık olduğu gözlemlenmektedir. Tahmin ve gerçek değerler arasındaki en iyi uyumu 9. Senaryo sağlamıştır. Dip ve zirve noktalarda küçük sapmalar mevcuttur. 12. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.72'deki gibidir.



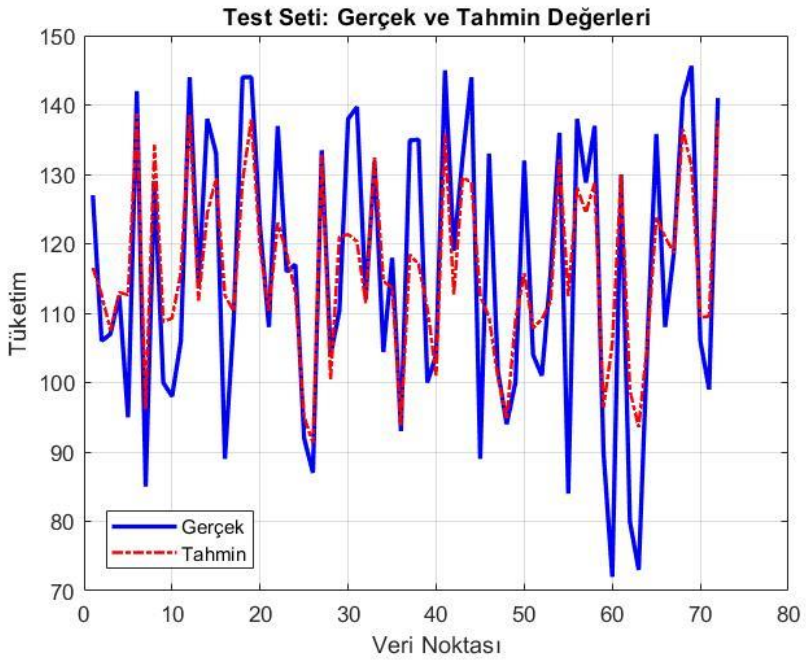
Şekil 5.72. Destek vektör makinesi 12. Senaryo tahmin grafiği

13. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.73'teki gibidir.



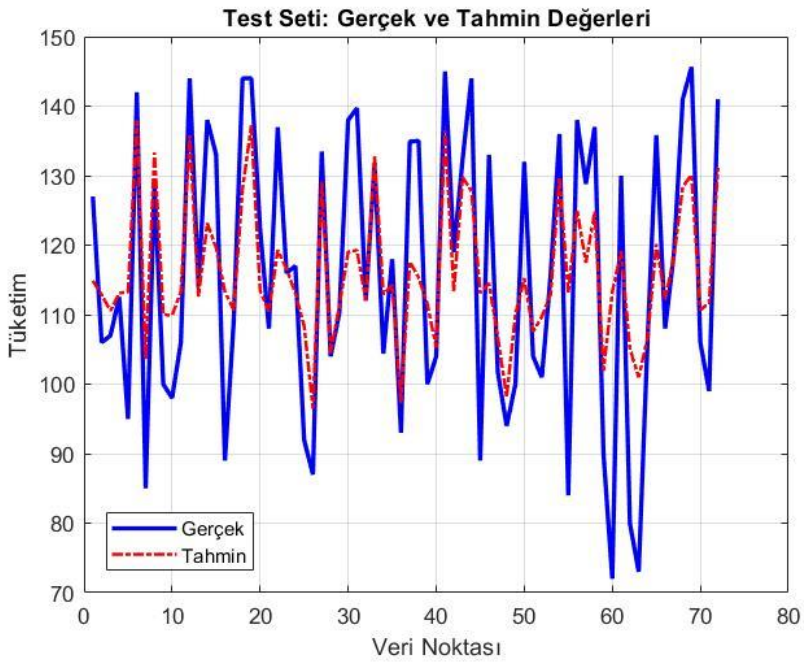
Şekil 5.73. Destek vektör makinesi 13. Senaryo tahmin grafiği

14. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.74'teki gibidir.



Şekil 5.74. Destek vektör makinesi 14. Senaryo tahmin grafiği

12. Senaryo, 13. Senaryo ve 14. Senaryo çıktıları incelendiğinde geçmiş tüketim verileri olan ek bağımsız değişkenlerin algoritmanın performansında fark yaratacak bir etkisinin olmadığı gözlemlenmektedir. Ek bağımsız değişkenlerin dâhil edilmediği senaryolarla karşılaştırıldığında uyum seviyesinde düşüş ve hata değerlerinde artış gözlemlenmektedir. 15. Senaryo için gerçekleştirilen SVM talep tahmini grafiği Şekil 5.75'teki gibidir.

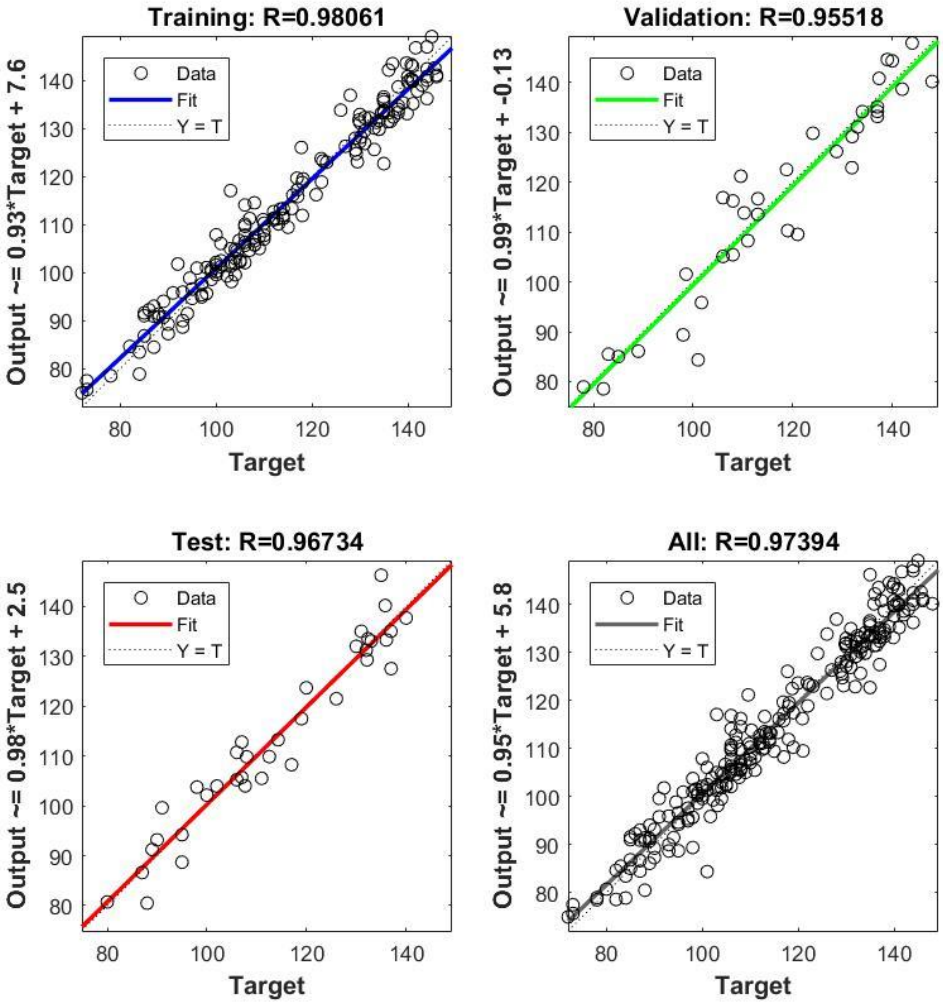


Şekil 5.75. Destek vektör makinesi 15. Senaryo tahmin grafiği

15. Senaryo çıktıları incelendiğinde tüm bağımsız değişkenlerin kullanımının yarattığı kargaşa net bir şekilde gözlemlenmektedir. Zamansal dalgalanmalar iyi bir şekilde takip edilmiştir fakat dip ve zirve noktalarda büyük sapmalar mevcuttur.

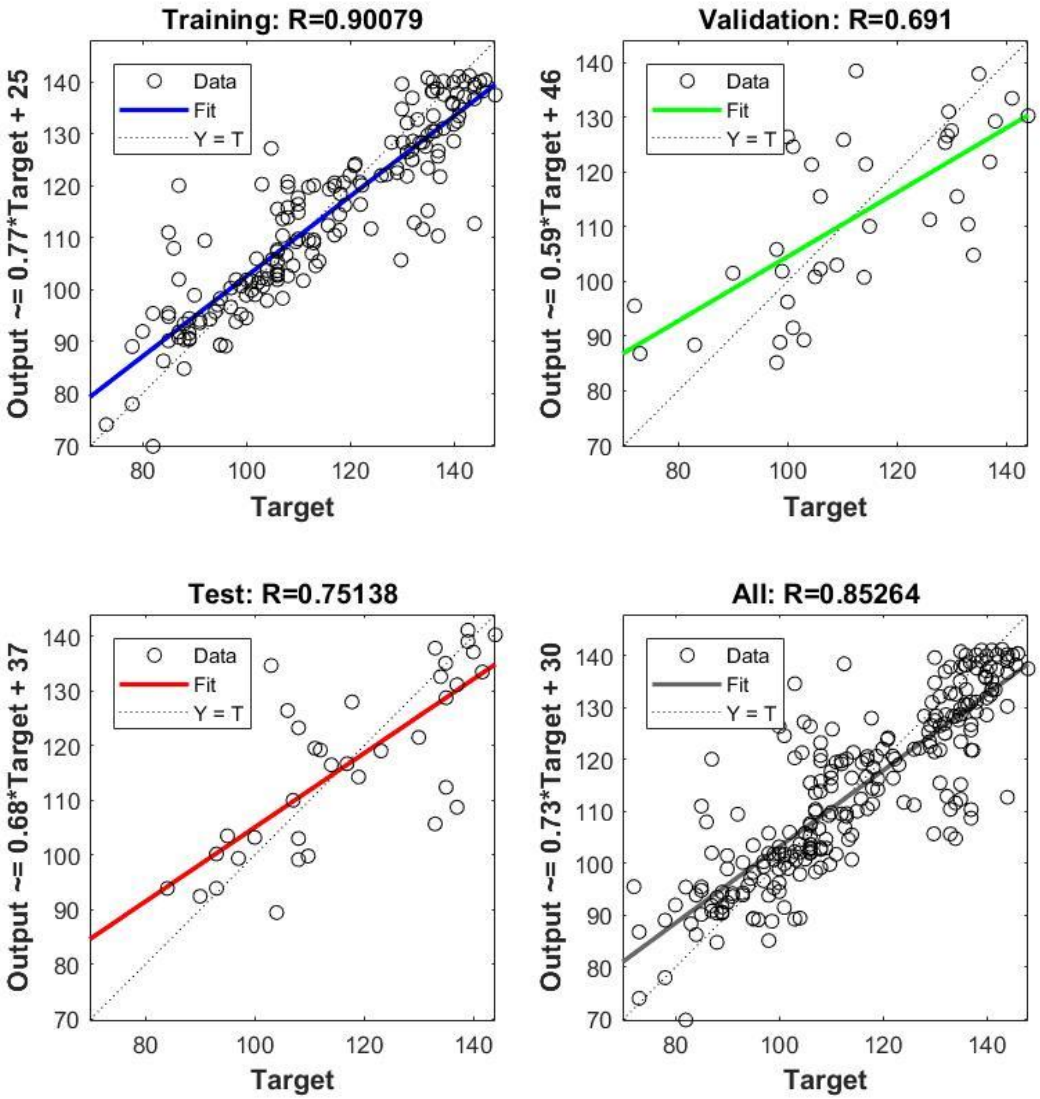
5.1.5. Yapay Sinir Ağları ile Talep Tahmini Uygulaması

Tahmin değerleri ve tüketim verilerinin grafik üzerinde anlaşılabilir olması için veri seti olarak 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler kullanılmıştır. Talep tahmini uygulaması için 2 gizli katmanlı yapay sinir ağı modeli oluşturulmuştur. Modelin 1. gizli katmanındaki nöron sayısı 13, 2. gizli katmanındaki nöron sayısı 18 olarak belirlenmiştir. Levenberg-Marquardt eğitim fonksiyonu kullanılmıştır. Modelin transfer fonksiyonu sigmoid fonksiyonudur. Talep tahmini uygulaması 10 döngü ile gerçekleştirilmiştir. Denemeler sonucu en iyi hata değerlerini veren eğitim parametreleri kullanılmıştır. Veri setinin %70'i eğitim, %15'i doğrulama ve %15'i test için kullanılmıştır. 1. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.76'daki gibidir.



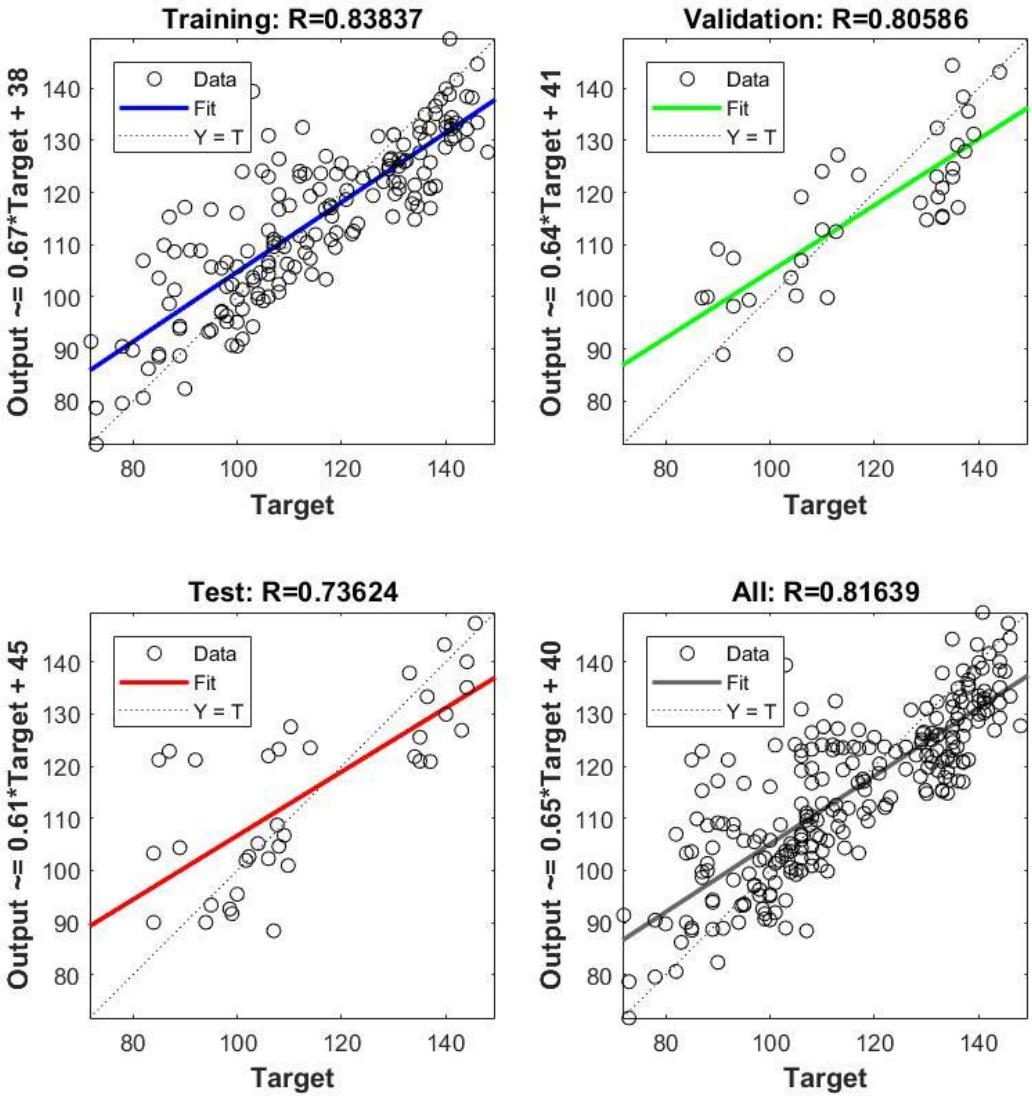
Şekil 5.76. Yapay sinir ağları 1. Senaryo regresyon grafiği

1. Senaryo tahmin sonuçları incelendiğinde gerçek ve tahmin değerleri arasında oldukça tatmin edici bir uyum olduğu gözlemlenmektedir. Algoritmanın genelleme kapasitesi yüksektir. 2. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.77'deki gibidir.



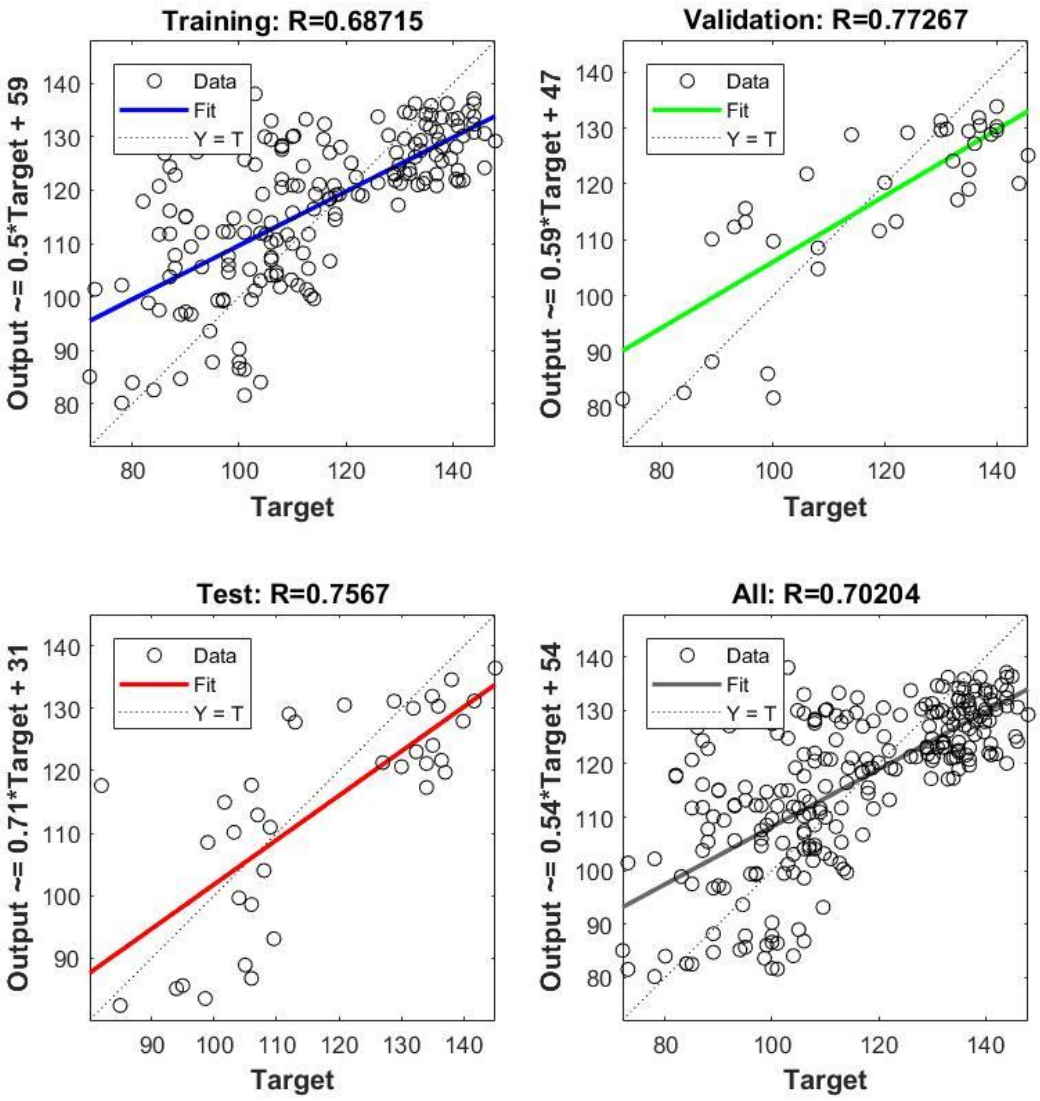
Şekil 5.77. Yapay sinir ağları 2. Senaryo regresyon grafiği

2. Senaryonun regresyon grafikleri incelendiğinde günün saati bağımsız değişkeni çıkarıldıktan sonra modelin performansında ciddi bir düşüş gözlemlenmektedir. Modelin genelleme yeteneği zayıflamıştır. 3. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.78'deki gibidir.



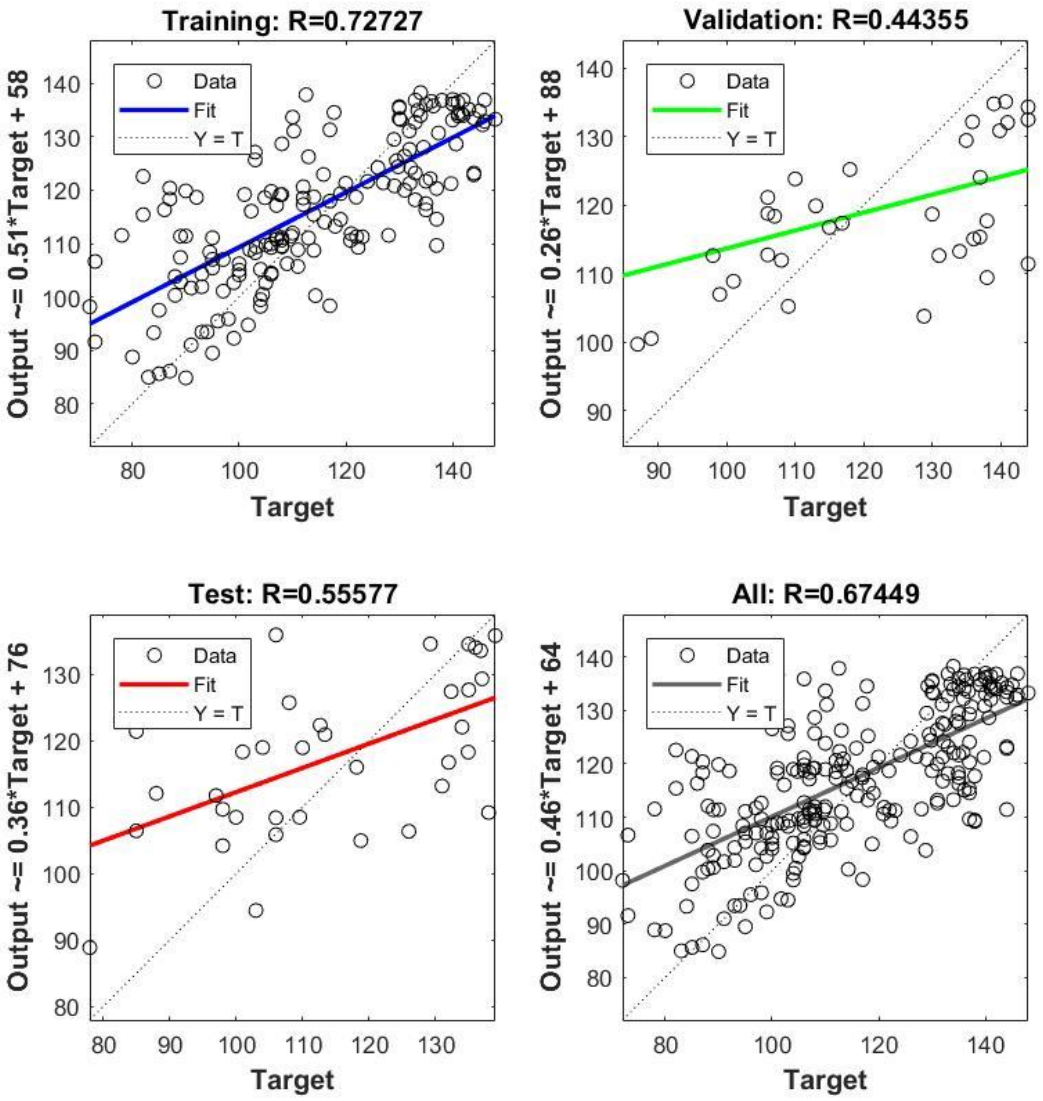
Şekil 5.78. Yapay sinir ağları 3. Senaryo regresyon grafiği

4. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.79'daki gibidir.



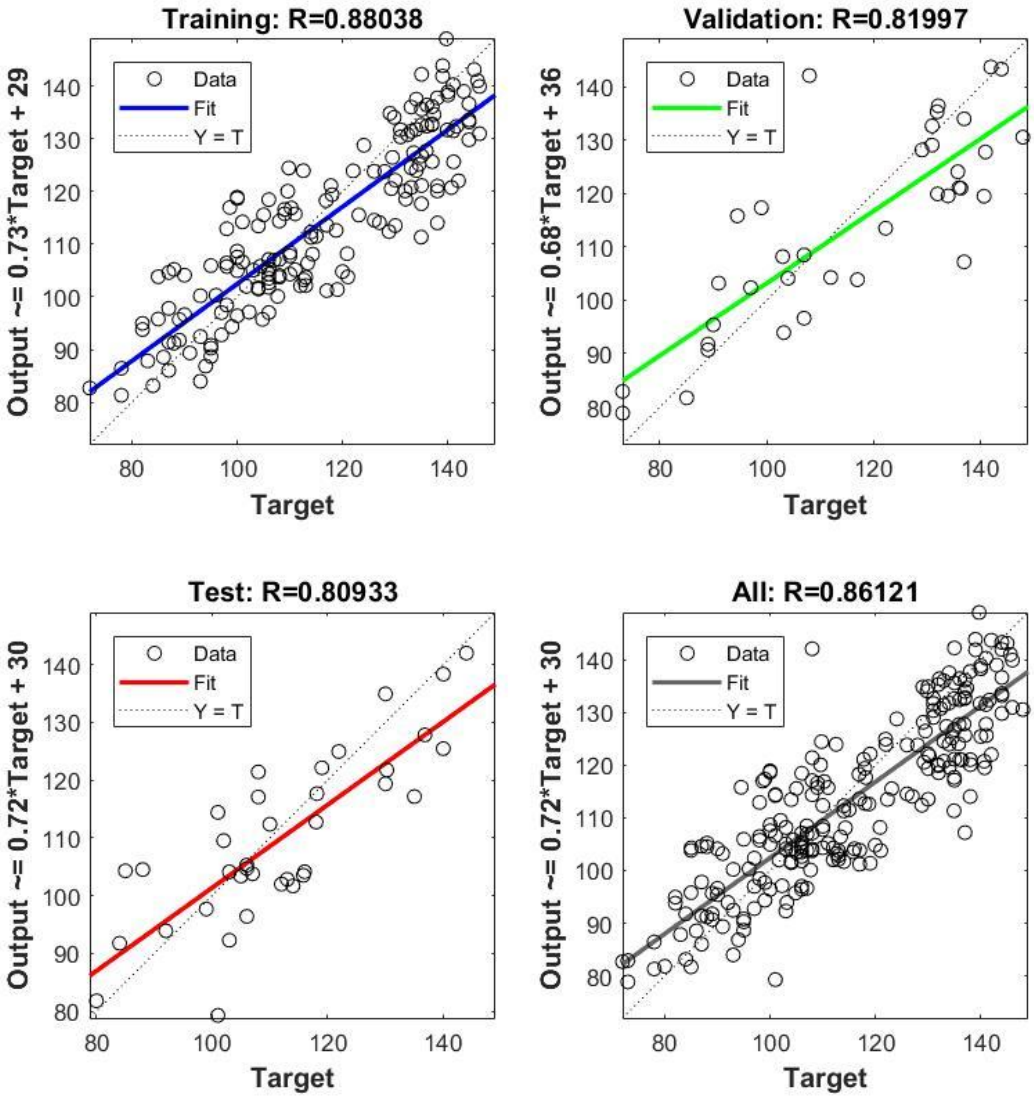
Şekil 5.79. Yapay sinir ağları 4. Senaryo regresyon grafiği

3. Senaryo ve 4. Senaryonun regresyon grafikleri incelendiğinde haftanın günü bağımsız değişkeninin belirleyici etkisinin daha yüksek olduğu sonucuna varılmaktadır. 5. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.80'deki gibidir.



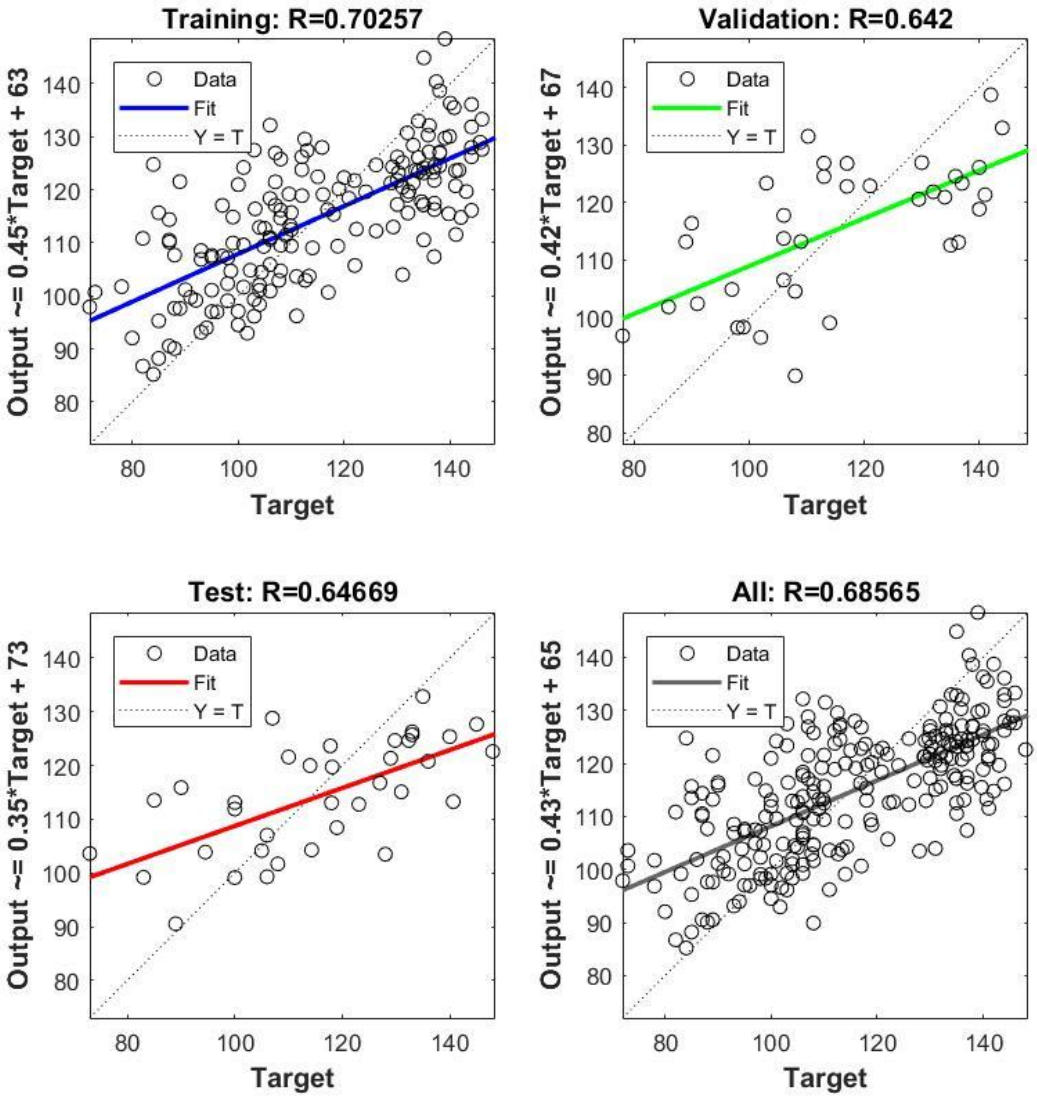
Şekil 5.80. Yapay sinir ağları 5. Senaryo regresyon grafiği

5. Senaryonun regresyon grafikleri incelendiğinde haftanın günü bağımsız değişkeninin de çıkarılması üzerine zamansal bağımsız değişken sayısının yetersizliği açık bir şekilde gözlemlenmektedir. Modelin genelleme performansı düşmüştür. 6. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.81'deki gibidir.



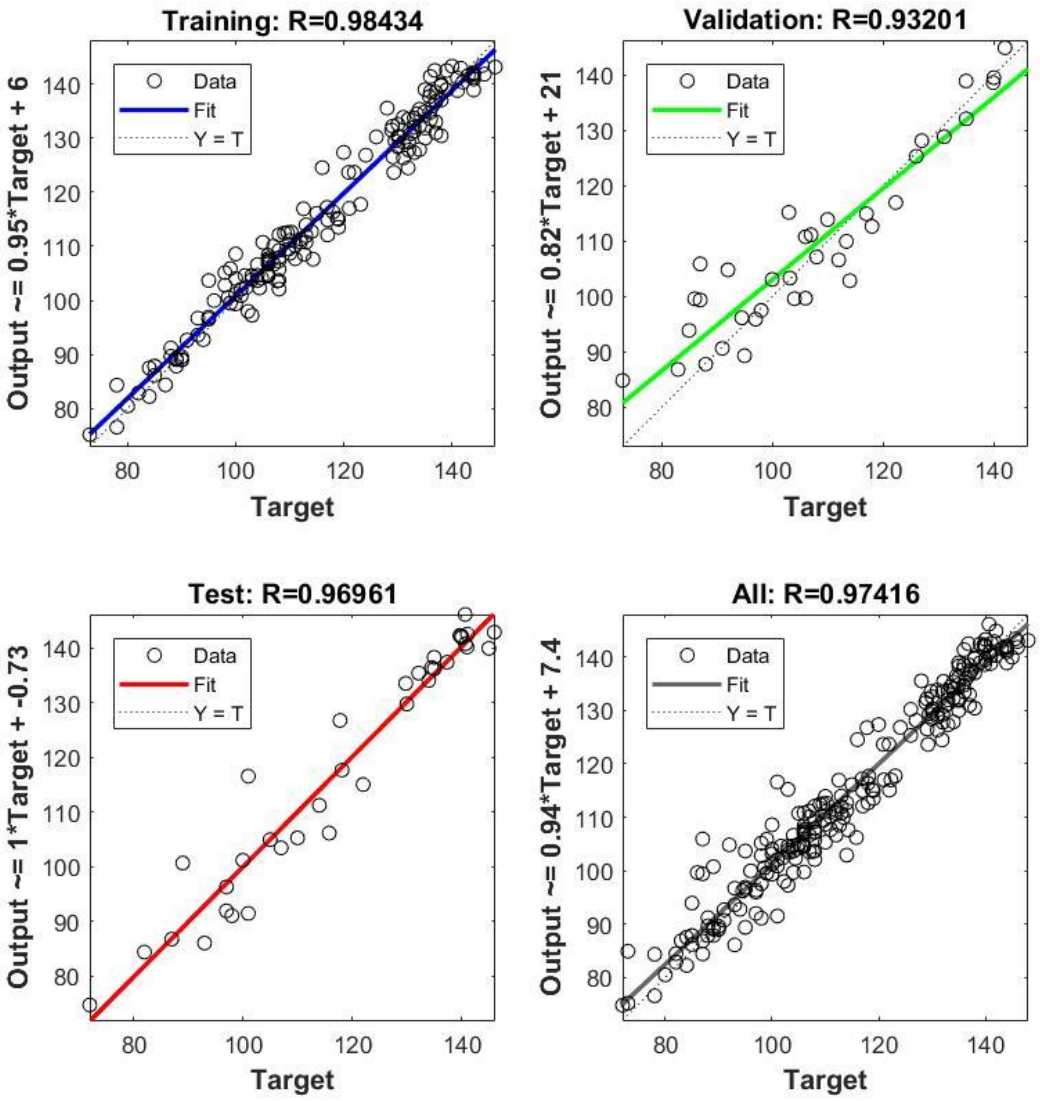
Şekil 5.81. Yapay sinir ağları 6. Senaryo regresyon grafiği

6. Senaryonun regresyon grafikleri incelendiğinde günün saati bağımsız değişkeninin etkisi açık bir şekilde gözlemlenmektedir. Modelin doğruluk oranı ve genelleme kapasitesi artmıştır. 7. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.82'deki gibidir.



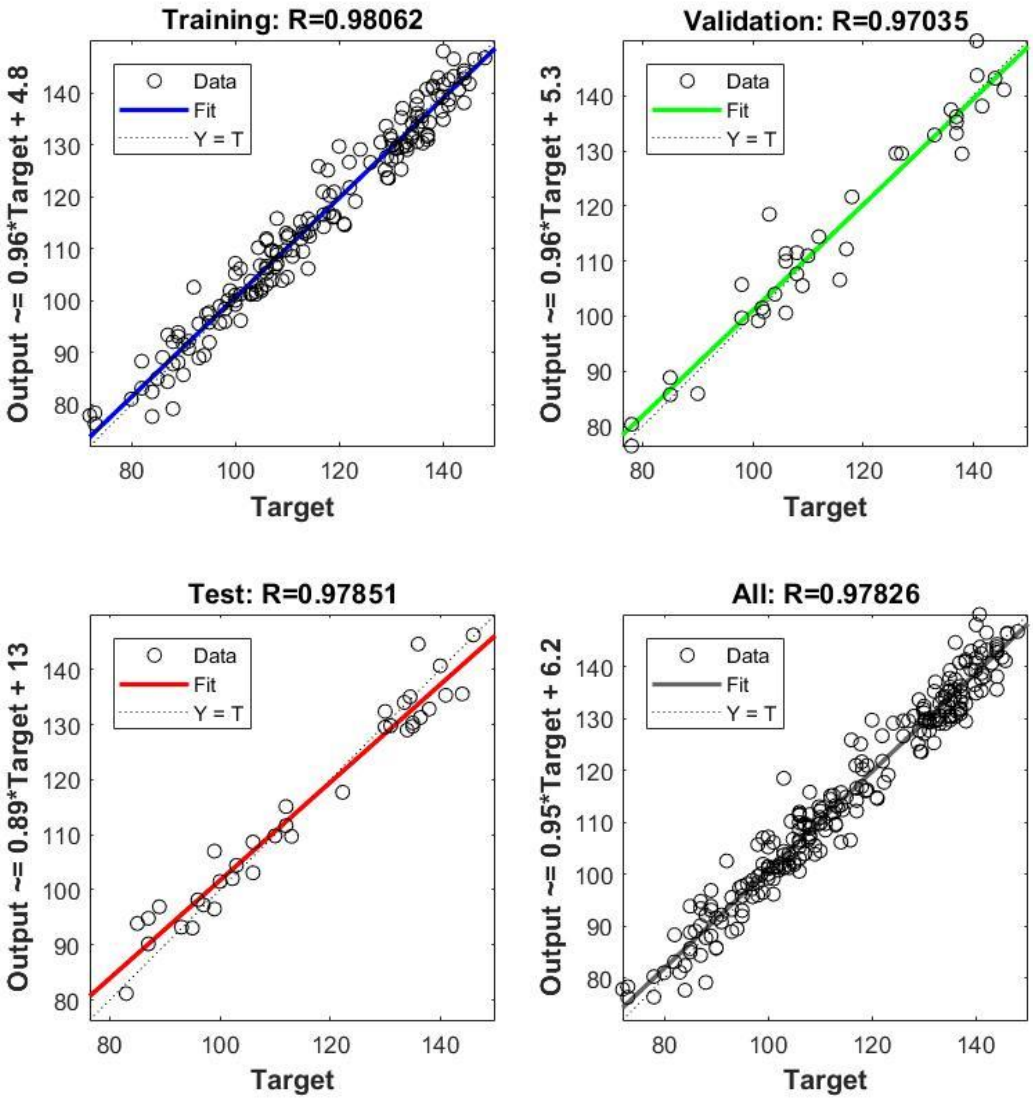
Şekil 5.82. Yapay sinir ağları 7. Senaryo regresyon grafiği

7. Senaryo, zamansal bağımsız değişkenlerin çıkarılmasının üzerine basit ve genelleme yeteneği düşük bir model haline gelmiştir. Doğruluk seviyesi ve genelleme yeteneğinin düştüğü gözlemlenmiştir. 8. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.83'teki gibidir.



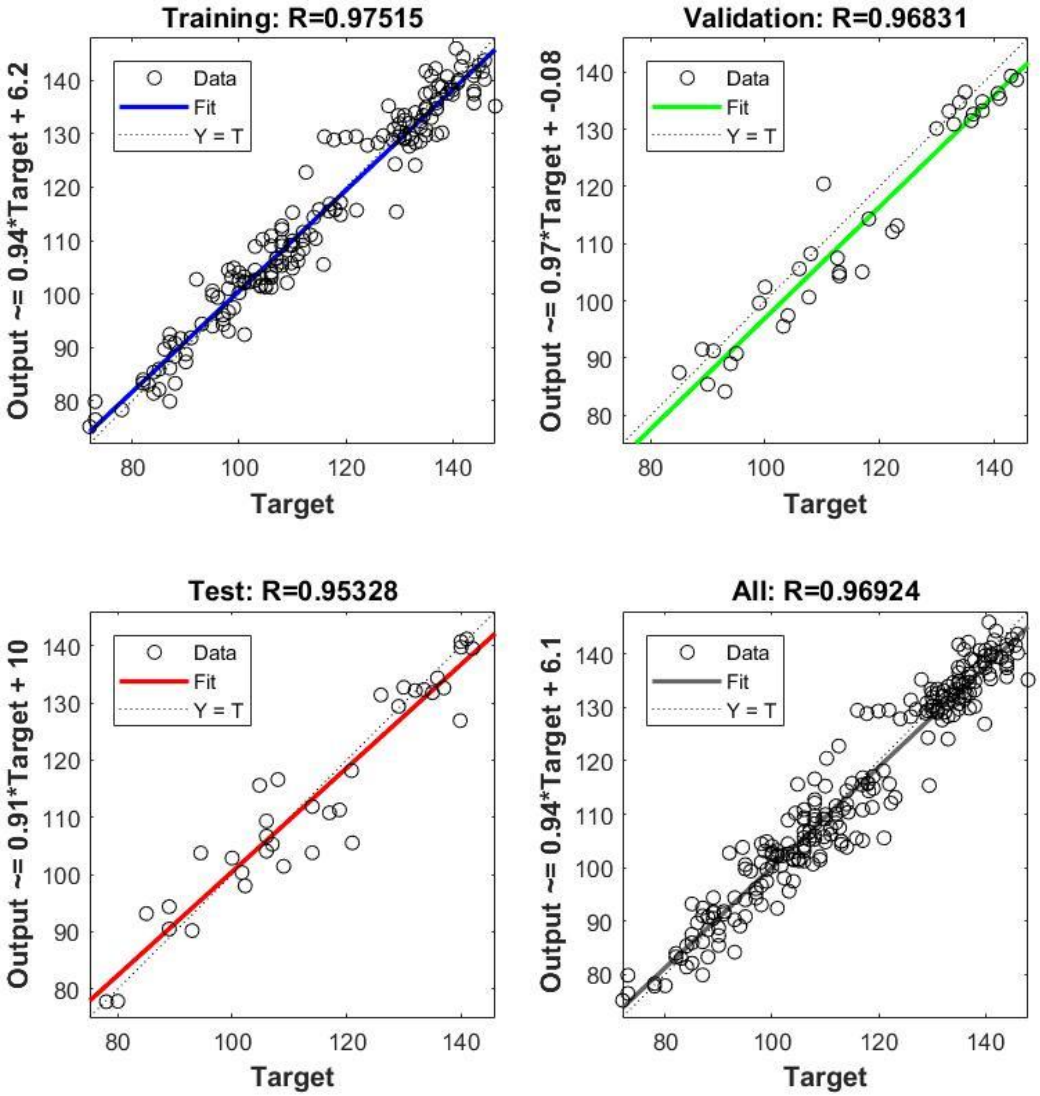
Şekil 5.83. Yapay sinir ağları 8. Senaryo regresyon grafiği

Algoritmanın 8. Senaryo için talep tahmini çıktıları incelendiğinde zamansal bağımsız değişkenlerin hepsinin kullanılmasının performansa yansıyan olumlu etkisi fark edilmektedir. 1. Senaryo sonuçları ile karşılaştırıldığında 8. Senaryonun, rüzgâr bağımsız değişkeninin çıkarılmasıyla daha basit bir model haline gelerek daha başarılı tahmin performansı sergilediği gözlemlenmiştir. 9. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.84'teki gibidir.



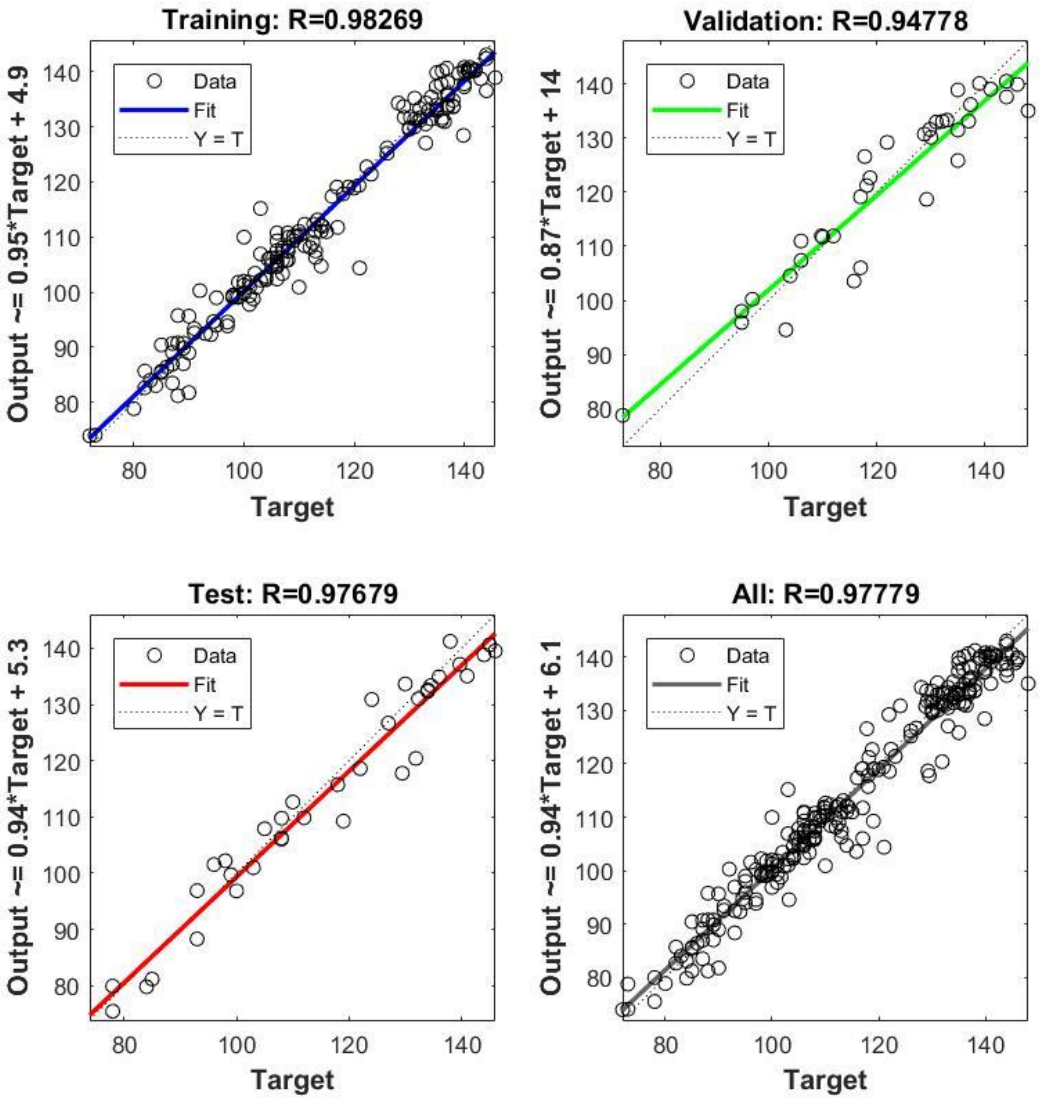
Şekil 5.84. Yapay sinir ağları 9. Senaryo regresyon grafiği

10. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.85'teki gibidir.



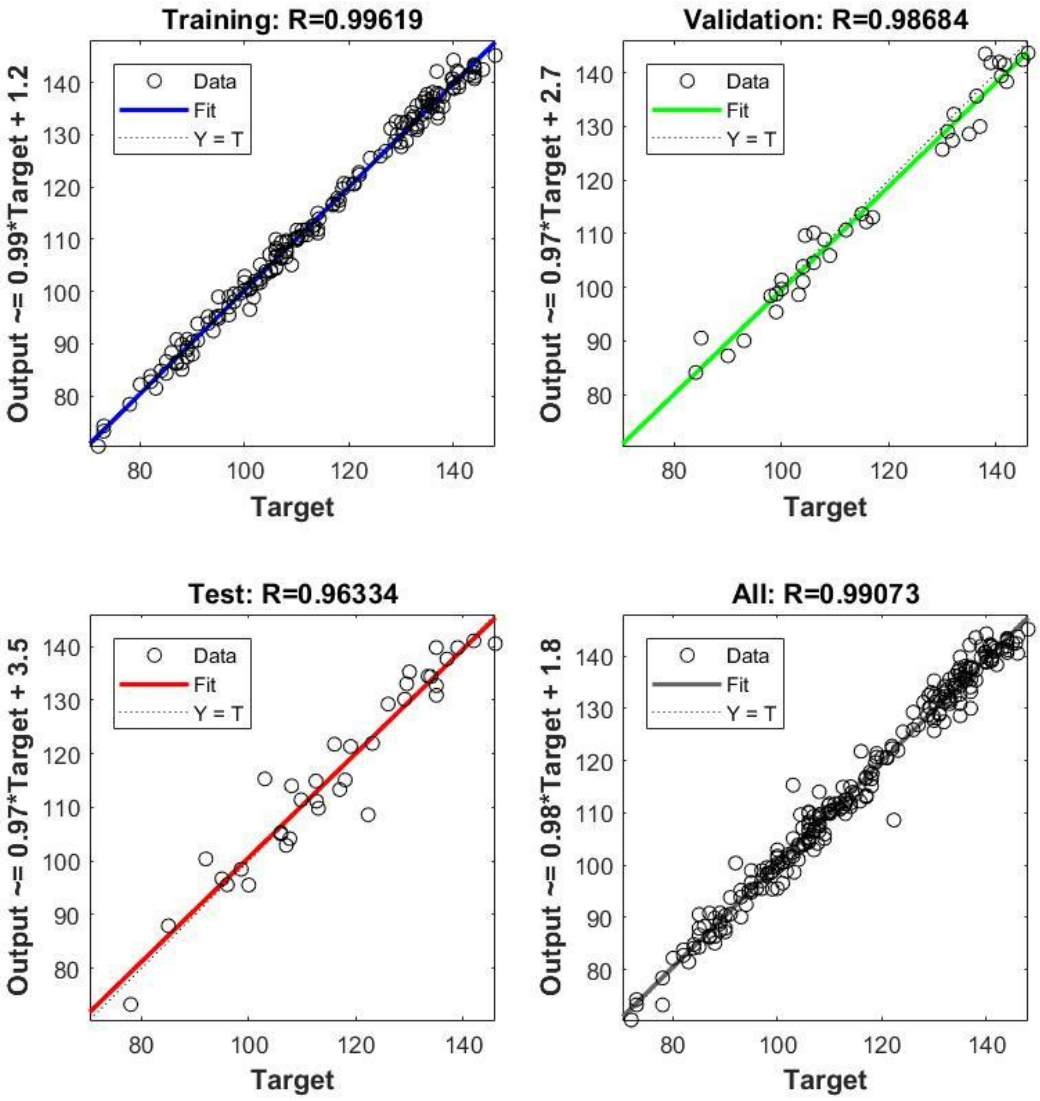
Şekil 5.85. Yapay sinir ağları 10. Senaryo regresyon grafiği

11. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.86'daki gibidir.



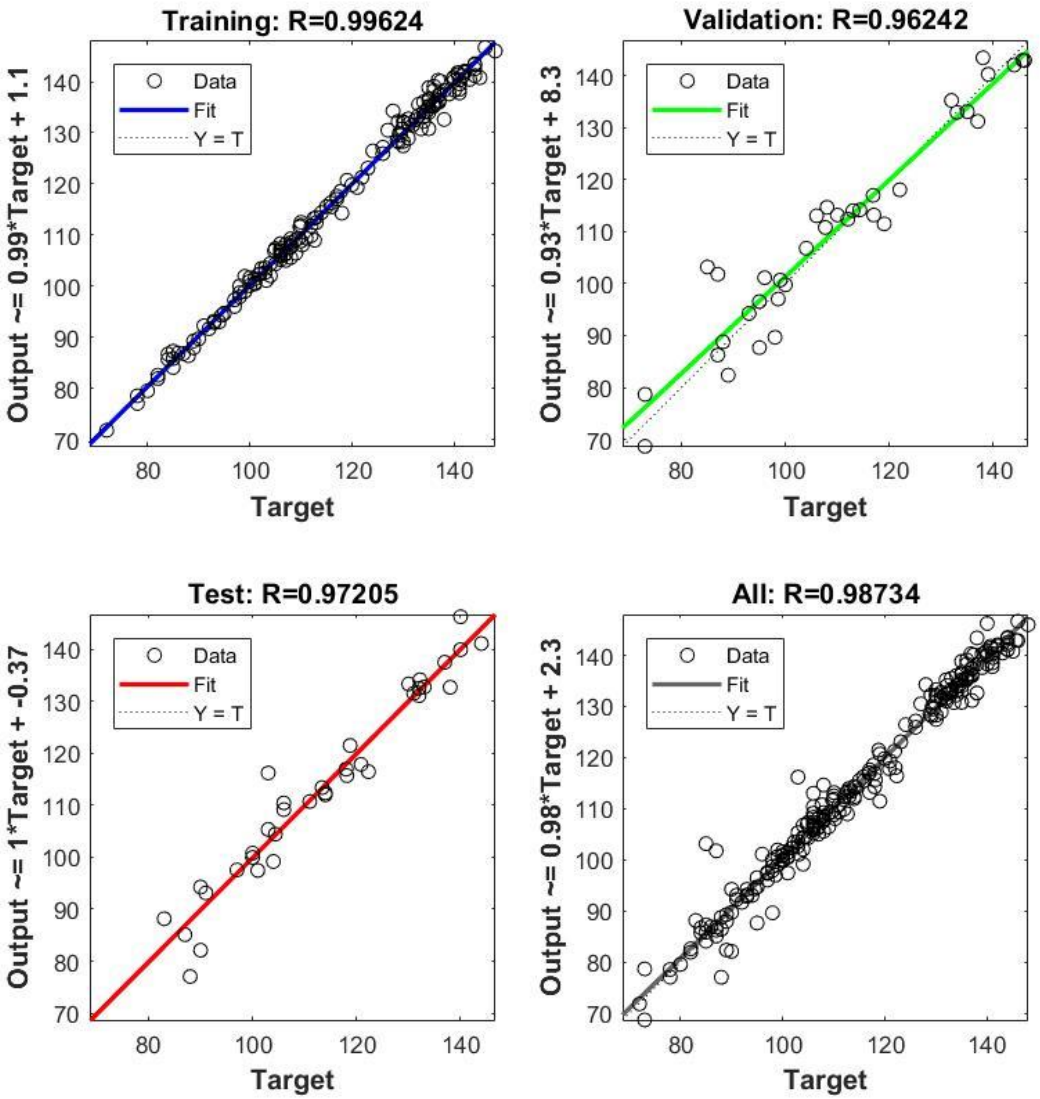
Şekil 5.86. Yapay sinir ağları 11. Senaryo regresyon grafiği

9. Senaryo, 10. Senaryo ve 11. Senaryo regresyon grafikleri incelendiğinde sıcaklık, nem ve rüzgâr bağımsız değişkenleri arasında belirleyiciliği en yüksek değişkenin sıcaklık olduğu sonucuna varılıyor. 3 senaryo arasında en iyi sonuç sıcaklık bağımsız değişkeninin kullanıldığı 9. Senaryoya aittir. 12. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.87'deki gibidir.



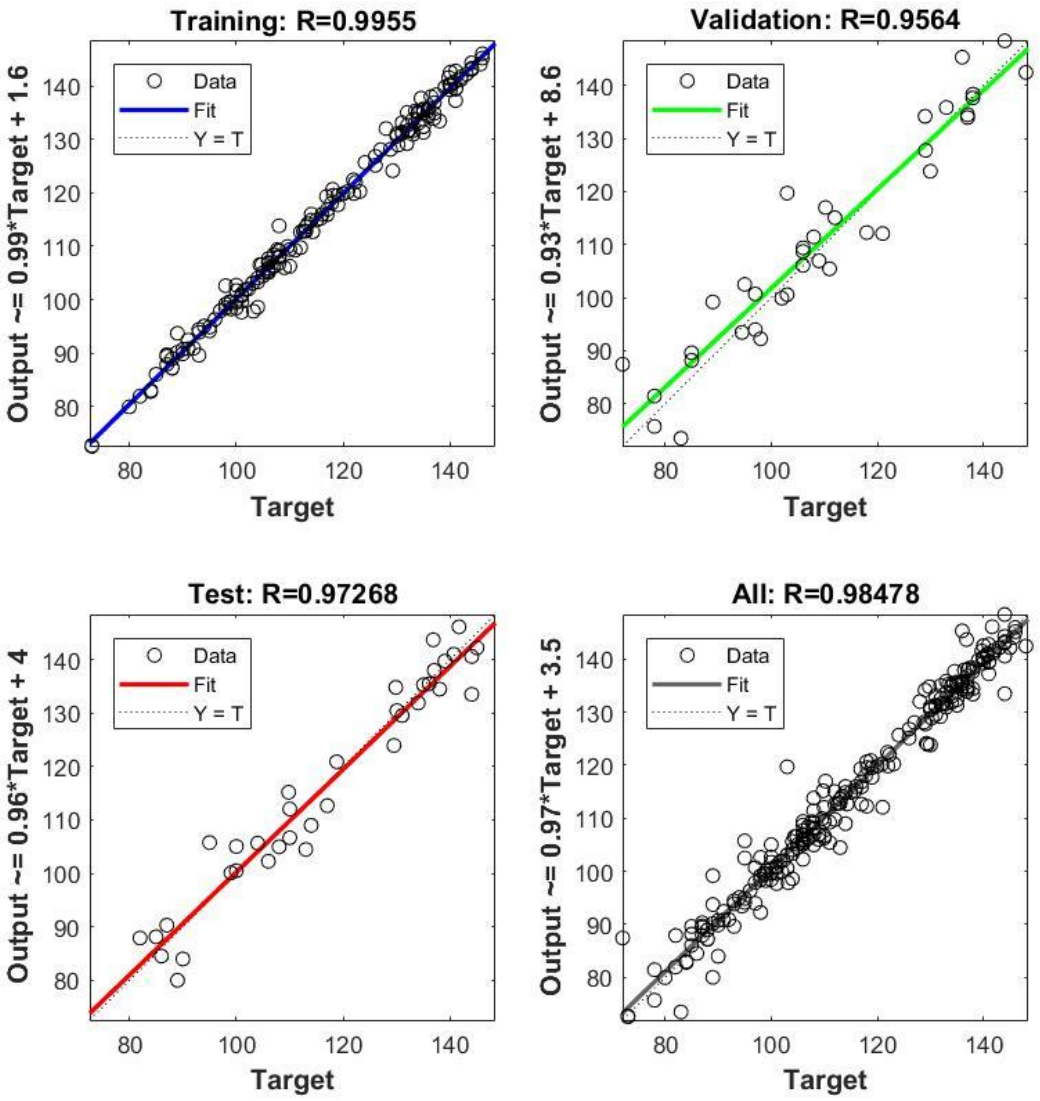
Şekil 5.87. Yapay sinir ağları 12. Senaryo regresyon grafiği

13. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.88'deki gibidir.



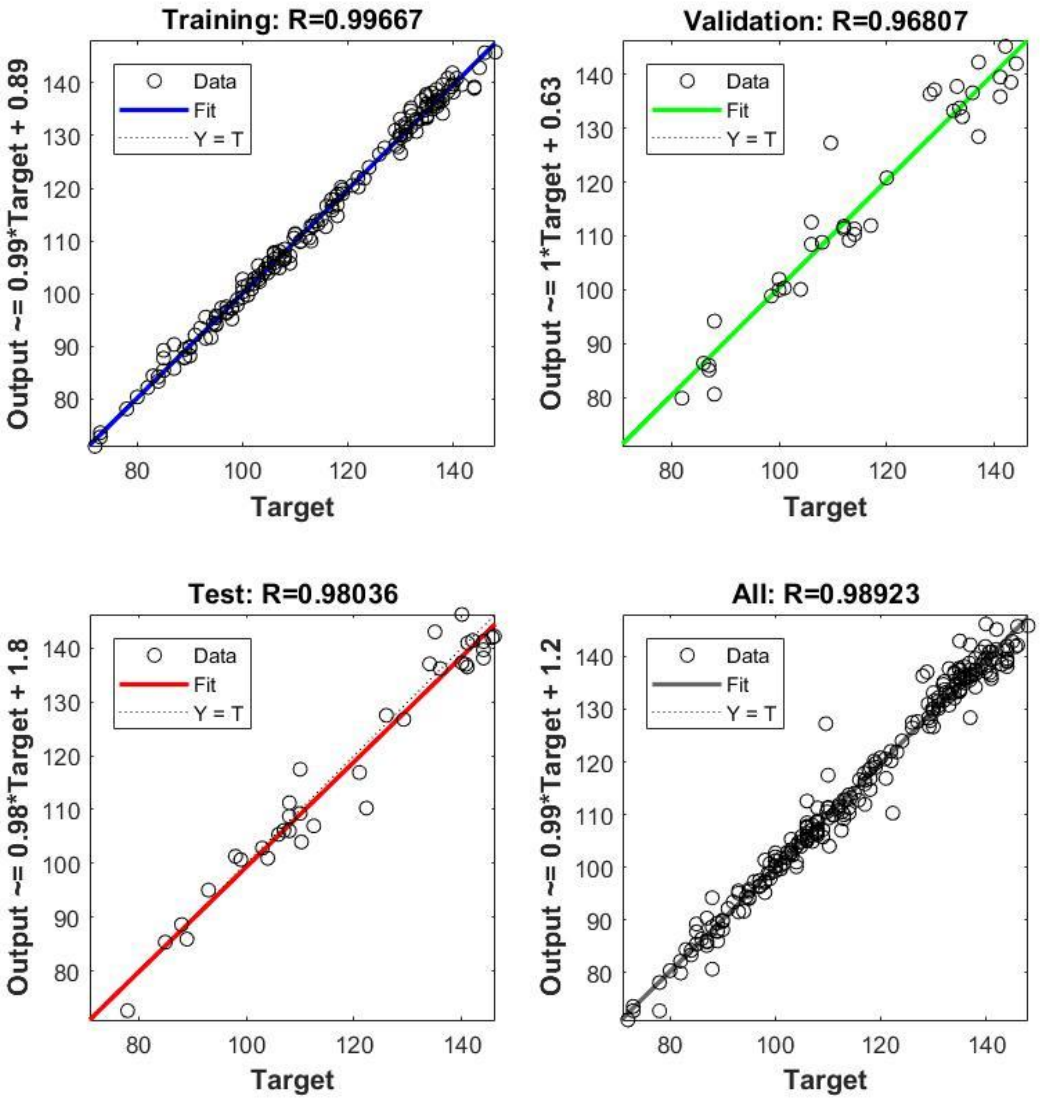
Şekil 5.88. Yapay sinir ağları 13. Senaryo regresyon grafiği

14. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.89'daki gibidir.



Şekil 5.89. Yapay sinir ağları 14. Senaryo regresyon grafiği

12. Senaryo, 13. Senaryo ve 14. Senaryo regresyon grafikleri incelendiğinde 1 hafta öncesinin aynı saatteki tüketim değeri bağımsız değişkeninin en yüksek belirleyiciliğe sahip olduğu gözlemlenmektedir. Algoritmanın 12. Senaryoda sergilediği uyum ve genelleme yeteneği en iyi performansa sahiptir. 15. Senaryo için gerçekleştirilen ANN talep tahmini eğitim, doğrulama, test ve genel regresyon grafikleri Şekil 5.90'daki gibidir.



Şekil 5.90 Yapay sinir ağları 15. Senaryo regresyon grafiği

15. Senaryonun regresyon grafiği incelendiğinde tüm bağımsız değişkenlerin kullanılmasıyla birlikte algoritmanın bağımlı ve bağımsız değişkenler arasındaki ilişkilendirme performansının arttığı gözlemlenmektedir. 12. Senaryo ile karşılaştırıldığında 15. Senaryo daha başarılı bir tahmin sonucuna sahiptir.

5.2. Hata Metrikleri

5.2.1. Mean Absolute Percentage Error (MAPE)

Türkçesi ortalama mutlak yüzde hatası olan MAPE, tahmin edilen değerler ile gerçek değerler arasındaki mutlak yüzdelik hatalarının ortalamasını alan hata metriğidir. Düşük MAPE değeri yüksek tahmin doğruluğu anlamına gelmektedir. MAPE formülü denklem (5.16)'da belirtilmektedir (Vivas, Allende-Cid, & Salas, 2020).

$$MAPE = \frac{1}{m}\sum_{k=1}^{m}\left|\frac{t_k - y_k}{t_k}\right| * 100 \tag{5.16}$$

Burada:

- t_k: Gerçek tüketim değeri,
- y_k: Algoritmanın tahmin ettiği değeri,
- m : Toplam örnek sayısını ifade eder.

5.2.2. Root Mean Squared Error (RMSE)

Türkçesi kök ortalama kare hatası olan RMSE, algoritmanın tahmin ettiği değerler ile gerçek değerler arasındaki sapmaların karelerinin ortalamasının karekökünü hesaplayan hata metriğidir. Düşük RMSE değeri yüksek tahmin doğruluğu anlamına gelmektedir. RMSE formülü denklem (5.17)'de belirtilmektedir (Hodson vd., 2022).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{5.17}$$

Burada:

- y_i: Gerçek tüketim değeri,

- $\widehat{y_i}$: Algoritmanın tahmin ettiği değer,
- n: Toplam örnek sayısını ifade eder.

5.2.3. Korelasyon Katsayısı (R):

R, iki değişken arasındaki doğrusal ilişki derecesini ölçen istatiksel bir ölçüttür. -1 ile 1 arasında değer almaktadır. 1 değeri mükemmel doğrusal ilişkiyi tanımlamaktadır. R formülü denklem (5.18)'deki gibidir (Espinoza, Suykens, Belmans, & De Moor, 2007).

$$\mathrm{R} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} \tag{5.18}$$

Burada:

- x_i, y_i: X ve y'nin i'inci değerlerini,
- $\bar{x}$, $\bar{y}$: X ve y'nin ortalamalarını ifade etmektedir.

5.3. Tahmin Sonuçlarının Değerlendirilmesi

Algoritmaların 1. Senaryo için hata değerleri Çizelge 5.2'deki gibidir.

Çizelge 5.2. 1. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.55173	%9.46	12.08
KNN	0.8302	%5.42	7.63
DT	0.95	%3.95	5.98
SVM	0.89	%7.30	10.17
ANN	0.96734	%2.98	4.33

Çizelge 5.2 incelendiğinde 1. Senaryo özelinde %2.98 MAPE değeri ile en düşük hata oranına sahip algoritmanın ANN olduğu gözlemlenmektedir. Bununla birlikte 0.96734 R değeri ile gerçek ve tahmin değerleri arasında en yüksek uyumu gösteren algoritma da ANN olmuştur. Algoritmaları 2. Senaryo için hata değerleri Çizelge 5.3'teki gibidir.

Çizelge 5.3. 2. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.42368	%10.95	13.70
KNN	0.7602	%6.77	8.98
DT	0.64	%9.87	14.94
SVM	0.84	%8.03	10.95
ANN	0.75138	%6.55	10.00

Çizelge 5.3 incelendiğinde 2. Senaryo özelinde %6.55 ile en düşük MAPE değerine sahip algoritma ANN olmuştur. RMSE ve R değerleri karşılaştırıldığında 8.98 ile en düşük RMSE değerine sahip, 0.7602 R değeri ile gerçek ve tahmin değerleri arasındaki ilişkiyi yakalama performansı en iyi olan algoritma KNN olmuştur. Algoritmaların 3. Senaryo için hata değerleri Çizelge 5.4'teki gibidir.

Çizelge 5.4. 3. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.42368	%10.95	13.70
KNN	0.6651	%7.84	10.57

DT	0.59	%10.95	15.70
SVM	0.74	%10.06	13.66
ANN	0.73624	%7.71	11.02

Çizelge 5.4 incelendiğinde 3. Senaryo özelinde 0.73624 R değeri ile en iyi tahmin performansını gösteren algoritma ANN'dir. Algoritmaların 4. Senaryo için hata değerleri Çizelge 5.5'teki gibidir.

Çizelge 5.5. 4. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.42368	%10.95	13.70
KNN	0.7602	%6.77	8.98
DT	0.64	%9.87	14.94
SVM	0.84	%7.99	10.97
ANN	0.7567	%10.18	13.70

Çizelge 5.5 incelendiğinde 4. Senaryo özelinde %6.77 MAPE ve 9.98 RMSE değeri ile en düşük hata oranlarına sahip algoritmanın KNN olduğu sonucuna varılmaktadır. Bununla birlikte 0.84 R değeri ile gerçek ve tahmin değerleri arasında en iyi ilişkiyi yakalayan algoritma SVM olmuştur. Algoritmaların 5. Senaryo için hata değerleri Çizelge 5.6'daki gibidir.

Çizelge 5.6. 5. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.13738	%12.33	16.76
KNN	0.6651	%7.84	10.57
DT	0.54	%11.45	17.37
SVM	0.68	%10.87	14.92
ANN	0.55577	%10.15	14.22

Çizelge 5.6 incelendiğinde 5. Senaryo özelinde 0.6651 R değeri ile en iyi performansı göstererek %7.84 MAPE ve 10.57 RMSE değerleri ile en düşük hata oranına sahip algoritmanın KNN olduğu gözlemlenmektedir. Algoritmaların 6. Senaryo için hata değerleri Çizelge 5.7'deki gibidir.

Çizelge 5.7. 6. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.22492	%0.22492	15.88
KNN	0.76	%6.58	8.98
DT	0.81	%6.53	10.82
SVM	0.84	%8.23	11.29
ANN	0.80933	%6.91	9.87

Çizelge 5.7 incelendiğinde en düşük MAPE değerine sahip algoritmanın %6.53 ile DT olduğu anlaşılmaktadır. En düşük RMSE değeri 8.98 ile KNN algoritmasına aittir. Bunların yanı sıra 0.84 R değeri ile daha tutarlı bir tahmin grafiği izleyerek en iyi

performansı gösteren algoritma SVM olmuştur. Algoritmaların 7. Senaryo için hata değerleri Çizelge 5.8'deki gibidir.

Çizelge 5.8. 7. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.11983	%13.10	16.93
KNN	0.6544	%7.87	10.74
DT	0.47	%11.89	18.26
SVM	0.63	%11.44	15.62
ANN	0.63647	%10.24	13.93

Çizelge 5.8 incelendiğinde 7. Senaryo özelinde %7.87 MAPE ve 10.74 RMSE değerleri ile en düşük hata oranlarına sahip algoritmanın KNN olduğu anlaşılmaktadır. Bununla birlik 0.6544 R değeri ile gerçek ve tahmin değerleri arasındaki en yüksek uyum da KNN algoritmasına aittir. Algoritmaların 8. Senaryo için hata değerleri Çizelge 5.9'daki gibidir.

Çizelge 5.9. 8. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.53436	%9.51	12.31
KNN	0.8279	%5.59	7.74
DT	0.95	%4.04	6.03
SVM	0.94	%5.70	7.78

ANN	0.96961	%2.87	4.32

Çizelge 5.9 incelendiğinde 8. Senaryo özelinde %2.87 MAPE, 4.32 RMSE ve 0.96961 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 9. Senaryo için hata değerleri Çizelge 5.10'daki gibidir.

Çizelge 5.10. 9. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.53427	%9.50	12.31
KNN	0.887	%4.93	6.20
DT	0.96	%3.42	4.75
SVM	0.96	%4.85	6.39
ANN	0.97851	%2.77	3.96

Çizelge 5.10 incelendiğinde 9. Senaryo özelinde %2.77 MAPE, 3.96 RMSE ve 0.97851 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 10. Senaryo için hata değerleri Çizelge 5.11'deki gibidir.

Çizelge 5.11. 10. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.53238	%9.49	12.34
KNN	0.8525	%5.25	7.16
DT	0.95	%4.19	6.11

SVM	0.94	%5.29	7.39
ANN	0.95328	%3.24	4.77

Çizelge 5.11 incelendiğinde 10. Senaryo özelinde 3.24 %MAPE, 4.77 RMSE ve 0.95328 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 11. Senaryo için hata değerleri Çizelge 5.12'deki gibidir.

Çizelge 5.12. 11. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.54318	%9.50	12.19
KNN	0.832	%5.41	7.50
DT	0.96	%3.54	4.98
SVM	0.95	%4.94	6.51
ANN	0.97679	%2.56	4.07

Çizelge 5.12 incelendiğinde 11. Senaryo özelinde %2.56 MAPE, 4.07 RMSE ve 0.97679 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 12. Senaryo için hata değerleri Çizelge 5.13'teki gibidir.

Çizelge 5.13. 12. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.93851	%3.59	4.47
KNN	0.9507	%3.20	4.11

DT	0.90	%5.15	7.82
SVM	0.86	%8.12	11.38
ANN	0.96334	%1.62	2.59

Çizelge 5.13 incelendiğinde 12. Senaryo özelinde %1.62 MAPE, 2.59 RMSE ve 0.96334 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 13. Senaryo için hata değerleri Çizelge 5.14'teki gibidir.

Çizelge 5.14. 13. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.62699	%8.29	11.02
KNN	0.7593	%5.74	8.95
DT	0.95	%4.05	6.05
SVM	0.87	%7.75	11.01
ANN	0.97205	%1.73	3.03

Çizelge 5.14 incelendiğinde 13. Senaryo özelinde %1.73 MAPE, 3.03 RMSE ve 0.97205 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 14. Senaryo için hata değerleri Çizelge 5.15'teki gibidir.

Çizelge 5.15. 14. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.93466	%3.41	4.61
KNN	0.9233	%3.52	5.04

DT	0.92	%4.57	7.12
SVM	0.85	%8.09	11.43
ANN	0.97268	%1.99	3.32

Çizelge 5.15 incelendiğinde 14. Senaryo özelinde %1.99 MAPE, 3.32 RMSE ve 0.97268 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır. Algoritmaların 15. Senaryo için hata değerleri Çizelge 5.16'daki gibidir.

Çizelge 5.16. 15. Senaryo Hata Değerleri

Yöntem	R	MAPE	RMSE
LS	0.96745	%2.56	3.26
KNN	0.9631	%2.73	3.52
DT	0.95	%4.08	5.66
SVM	0.82	%9.72	13.20
ANN	0.98036	%1.59	2.81

Çizelge 5.16 incelendiğinde 15. Senaryo özelinde %1.59 MAPE, 2.81 RMSE ve 0.98036 R değeri ile en iyi performansı gösteren algoritmanın ANN olduğu sonucuna varılmaktadır.

Tüm hata değerleri incelendiğinde günün saati dışındaki tüm zamansal ve meteorolojik bağımsız değişkenlerin kullanıldığı 2. Senaryoda en düşük MAPE değerine sahip algoritmanın ANN olmasına rağmen KNN, daha düşük RMSE ve daha yüksek R değeri sağlayarak daha iyi bir performans sergilemiştir. Bağımsız değişken sayısının az olduğu 4. Senaryo, 5. Senaryo, 6. Senaryo ve 7. Senaryoda KNN ve SVM algoritmalarının ön plana çıktığı gözlemlenmektedir. 4. Senaryo özelinde en düşük hata oranlarına sahip

olan algoritma KNN olmasına rağmen SVM, günün saati bağımsız değişkeninin eksikliğine daha fazla uyum sağlayarak daha tutarlı tahmin sonuçları gerçekleştirmiştir. Algoritmaların 5. Senaryo özelindeki performansları karşılaştırıldığında zaman temelli tek bağımsız değişkenin resmi tatil olma durumu olması basit bir model oluşturduğu için KNN algoritması ön plana çıkmıştır. 6. Senaryo sonuçları incelendiğinde KNN, modelin basitliğinden dolayı düşük MAPE ve RMSE hata değerlerine sahip olan algoritmadır fakat günün saati bağımsız değişkeninin eklenmesi SVM algoritmasının daha uyumlu bir grafik izlemesini sağlamıştır. 7. Senaryo sonuçları incelendiğinde sadece meteorolojik bağımsız değişkenlerin kullanılması, gerçek ve tahmin değerleri arasındaki ilişkiyi güçlendirmek için fazla bağımsız değişkene ihtiyaç duyan SVM ve ANN algoritmalarının performansını kısıtlayarak KNN algoritmasını ön plana çıkarmıştır. Bu senaryoların dışında bağımsız değişken sayısının fazla olduğu ve 1 saat öncesinin tüketim değeri gibi ek bağımsız değişkenlerin kullanıldığı 1. Senaryo, 3. Senaryo, 8. Senaryo, 9. Senaryo, 10. Senaryo, 11. Senaryo, 12. Senaryo, 13. Senaryo, 14. Senaryo ve 15. Senaryoda ANN algoritması ön plana çıkmıştır. En iyi tahmin performansını tüm zamansal, meteorolojik ve ek bağımsız değişkenlerin kullanıldığı 15. Senaryoda %1.59 MAPE, 2.81 RMSE ve 0.98036 R değeri ile ANN algoritması sağlamıştır.

BÖLÜM 6

SONUÇ

Bu tez çalışmasında LS, KNN, DT, SVM ve ANN yapay zekâ algoritmaları ile Tekirdağ, Çerkezköy bölgesinin kısa dönem talep tahmini ve algoritmaların performans değerlendirmesi gerçekleştirilmiştir. Algoritmalar MATLAB R2017b ortamında oluşturulmuştur. 1 Ocak 2021- 31 Aralık 2023 dönemine ait saatlik tüketim verileri TREPAŞ'tan temin edilmiştir. Aynı döneme ait saatlik meteoroloji verileri meteoblue veri tabanından temin edilmiştir. Oluşturulan veri seti incelenerek hatalı ve ölçüm yapılmamış değerler eğitim ve test sürecine dâhil edilmemiştir.

Tahmin çıktılarının grafikler üzerinde anlaşılabilir ve yorumlanabilir olması adına veri seti küçültülmüştür. LS ve KNN algoritmaları için veri seti olarak 1 Mayıs 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 72 saatlik süreçteki değerler kullanılmıştır. DT, SVM ve ANN algoritmaları için 24 Nisan 2021 00.00'dan 4 Mayıs 2021 00.00'a kadar olan 240 saatlik süreçteki değerler veri seti olarak kullanılmıştır. 240 saatlik veri setinin %70'i eğitim %30'u test veri seti olarak kullanılmıştır. Deneme yanılma sonucu en iyi performans gösteren hiper parametre değerleri kullanılmıştır. Bağımsız değişkenlerin algoritmalar üzerindeki etkisini değerlendirebilmek için 15 farklı senaryo oluşturulmuştur.

Senaryo sonuçları incelendiğinde zamansal döngü ifade eden bağımsız değişkenlerin tahmin sürecine dâhil edilmesi, algoritmaların genel tüketim karakteristiğini algılama performansını arttırarak sapma oranlarını azalttığı gözlemlenmiştir. Bağımsız değişken sayısının az olduğu senaryoların sonuçları

incelendiğinde, bu durumun algoritmaların tahmin performansını oldukça kısıtladığı gözlemlenmiştir. Bunların yanı sıra ek bağımsız değişkenlerin eklenmesi sonucu gerçek ve tahmin değerleri arasındaki uyumun arttığı, dip ve zirve değerlerdeki tutarsızlıkların oldukça azaldığı gözlemlenmiştir.

Bağımsız değişken sayısının az olduğu senaryolarda KNN ve SVM algoritmaları ön plana çıkmıştır. Haftanın günü, resmi tatil olma durumu ve tüm meteorolojik bağımsız değişkenlerin sürece dâhil edildiği 4. Senaryo özelinde 0.84 R değeri ile gerçek ve tahmin değerler arasında en iyi ilişkiyi yakalayan algoritma SVM olmuştur. Zamansal değişkenlerden sadece resmi tatil olma durumu ile tüm meteorolojik bağımsız değişkenlerin kullanıldığı 5. Senaryo özelinde 0.6651 R değeri ile en iyi performansı göstererek %7.84 MAPE ve 10.57 RMSE değerleri ile en düşük hata oranına sahip algoritma KNN olmuştur. Günün saati ve tüm meteorolojik verilerin kullanıldığı 6. Senaryoda DT ve KNN algoritmaları sırasıyla %6.53 MAPE ve 8.98 RMSE hata değerlerine sahip olmalarına rağmen, 0.84 R değeri ile gerçek ve tahmin değerleri arasındaki en yüksek uyumu gösteren algoritma SVM olmuştur. Sadece meteorolojik verilerin kullanıldığı 7. Senaryoda %7.87 MAPE, 10.74 RMSE ve 0.6544 R değeri ile en iyi tahmin performansını gerçekleştiren algoritma KNN olmuştur.

Bağımsız değişken sayının fazla olduğu ve ek bağımsız değişkenlerin dâhil edildiği 1. Senaryo, 8. Senaryo, 9. Senaryo, 10. Senaryo, 11. Senaryo, 12. Senaryo, 13. Senaryo, 14. Senaryo ve 15. Senaryoda ANN algoritmasının üstünlüğü dikkat çekmiştir. Tahmin sürecinin en başarılı sonuçları %1.59 MAPE, 2.81 RMSE ve 0.98036 R değeri ile 15. Senaryoda ANN algoritması ile elde edilmiştir. Çalışma bulguları veri seti özelliklerinin ve senaryo yapısının algoritma performansında önemli bir etkiye sahip olduğunu göstermektedir.

Tüm tahmin çıktıları değerlendirildiğinde bağımsız değişken sayısının artmasıyla birlikte ANN, yüksek genelleme ve karmaşık ilişkileri daha iyi modelleyebilme kabiliyeti ile ön plana çıkmıştır. Bu tez çalışmasının sonuçları doğrultusunda Tekirdağ, Çerkezköy bölgesinin yüksek doğruluk seviyesinde tüketim tahmini gerçekleştirilerek, altyapı bakım programı ve enerji planlaması öngörülerinin güvenilir veriler ile daha tutarlı bir şekilde gerçekleştirilebileceği ortaya konmuştur.

6.1. Öneriler

Veri seti özelliklerinin ve senaryo yapısının algoritmaların performansı üzerindeki etkisi göz önünde bulundurulduğunda, bağımsız değişkenlerin sayısı arttırılarak veya farklı senaryo varyasyonları denenerek algoritmaların öngörü yeteneği iyileştirilebilir. Bunun yanı sıra DT algoritması için gerçekleştiren özellik önem derecesinin belirlenmesi gibi süreçleri tüm algoritmalara uygulayarak daha özele indirgenmiş senaryo varyasyonları oluşturulabilir.

Birden fazla algoritmadan oluşan birleşik modeller kullanılarak tahmin tutarlılığı arttırılabilir. Böylelikle birleşik modelde kullanılacak olan algoritmaların güçlü yönleri birleştirilerek genel performans arttırılabilir (Zor vd., 2017).

Çalışmada kullanılan algoritmalar, enerji sistemlerine bütünleşik hale getirilerek gerçek zamanlı tahmin ve izleme sistemleri geliştirilebilir (Yun vd., 2008).

Algoritmaların çok yönlülüğü, rüzgâr (Ahmadi vd., 2020) veya güneş enerjisi santralleri (Fan vd., 2019) gibi farklı süreçlere sahip sektörlerin tüketim veri setleri ile çalışarak geliştirilebilir.

KAYNAKÇA

Abedinia, O., Amjady, N., & Zareipour, H. (2016). A New Feature Selection Technique for Load and Price Forecast of Electrical Power Systems. IEEE Transactions on Power Systems, 32(1), 62–74. https://doi.org/10.1109/TPWRS.2016.2556620

Aguilar Madrid, E., & Antonio, N. (2021). Short-term electricity load forecasting with machine learning. *Information*, *12*(2), 50. https://doi.org/10.3390/info12020050

Ahmad, A., Javaid, N., Mateen, A., Awais, M., & Khan, Z. A. (2019). Short-term load forecasting in smart grids: An intelligent modular approach. *Energies*, *12*(1), 164. https://doi.org/10.3390/en12010164

Ahmed, F. S., Allu, A. A., & Jirjees, M. A. (2022). Performance Enhancement of Speed Control for Induction Motor Using Dolphin Algorithm. Journal of Physics: Conference Series, 2312(1). https://doi.org/10.1088/1742-6596/2312/1/012026

Ahmed, S.F., Alam, M.S.B., Hassan, M. et al. Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev 56, 13521–13617 (2023). https://doi.org/10.1007/s10462-023-10466-8

Ahmadi, A., Nabipour, M., Mohammadi-Ivatloo, B., Amani, A. M., Rho, S., & Piran, M. J. (2020). Long-term wind power forecasting using tree-based learning algorithms. *IEEE Access*, *8*, 151511–151523. https://doi.org/10.1109/ACCESS.2020.3017442

Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F., & Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Advanced Engineering Informatics, 35(November 2017), 1–16. https://doi.org/10.1016/j.aei.2017.11.002

Atalay, M. (2013). *Makine öğrenme yöntemleri yardımıyla tüketim istatistiklerine göre talep tahmini* (Yüksek Lisans Tezi, İstanbul Kültür Üniversitesi, Fen Bilimleri Enstitüsü).

Ayhan, S., & Erdoğmuş, Ş. (2014). Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi. Eskişehir Osmangazi Üniversitesi İktisadi Ve İdari Bilimler Dergisi, 9(1), 175-201.

Ayyub, B., & McCuen, R. H. (2015). *Numerical analysis for engineers: methods and applications*. CRC Press.

Azeem, A., Ismail, I., Jameel, S. M., & Harindran, V. R. (2021). Electrical load forecasting models for different generation modalities: A review. *IEEE Access, 9*, 142239–142259. https://doi.org/10.1109/ACCESS.2021.3120731

Bacanin, N., Jovanovic, L., Zivkovic, M., Kandasamy, V., Antonijevic, M., Deveci, M., & Strumberger, I. (2023). Multivariate energy forecasting via metaheuristic tuned long-short term memory and gated recurrent unit neural networks. Information Sciences, 642(January), 119122. https://doi.org/10.1016/j.ins.2023.119122

Bashir, T., Haoyong, C., Tahir, M. F., & Liqiang, Z. (2022). Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. Energy Reports, 8, 1678–1686. https://doi.org/10.1016/j.egyr.2021.12.067

Berner, J., Grohs, P., & Jentzen, A. (2018). Analysis of the generalization error: Empirical risk minimization over deep artificial neural networks overcomes the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations. *SIAM J. Math. Data Sci., 2*, 631-657.

Boloix-Tortosa, R., Murillo-Fuentes, J. J., & Tsaftaris, S. A. (2019). The Generalized Complex Kernel Least-Mean-Square Algorithm. IEEE Transactions on Signal Processing, 67(20), 5213–5222. https://doi.org/10.1109/TSP.2019.2937289

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2018). Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies, 11*(7), 1636. https://doi.org/10.3390/en11071636

Bozic, M., Stojanović, M., Stajić, Z., & Floranović, N. (2013). Mutual information-based inputs selection for electric load time series forecasting. Entropy, 15(3), 926–942. https://doi.org/10.3390/e15030926

Burger, E. M., & Moura, S. J. (2015). Gated ensemble learning method for demand-side electricity load forecasting. *Energy and Buildings, 109*, 23–34. https://doi.org/10.1016/j.enbuild.2015.10.019

Campbell, W.M., Campbell, J.P., Reynolds, D.A., Singer, E., & Torres-Carrasquillo, P.A. (2006). Support vector machines for speaker and language recognition. *Comput. Speech Lang.*, *20*, 210-229.

Cerne, G., Dovžan, D., & Škrjanc, I. (2018). Short-term load forecasting by separating daily profiles and using a single fuzzy model across the entire domain. *IEEE Transactions on Industrial Electronics*. Advance online publication. https://doi.org/10.1109/TIE.2018.2795555

Chapagain, K., Kittipiyakul, S., & Kulthanavit, P. (2020). Short-Term Electricity Demand Forecasting: Impact Analysis of Temperature for Thailand. *Energies*, 13(10), 2498. https://doi.org/10.3390/en13102498

Chandriah, K. K., & Naraganahalli, R. V. (2021). RNN / LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting. Multimedia Tools and Applications, 80(17), 26145–26159. https://doi.org/10.1007/s11042-021-10913-0

Chia, Y. Y., Lee, L. H., Shafiabady, N., & Isa, D. (2015). A load predictive energy management system for supercapacitor-battery hybrid energy storage system in solar application using the Support Vector Machine. *Applied Energy*, *137*, 588–602. https://doi.org/10.1016/j.apenergy.2014.09.026

Cunningham, P., & Delany, S. J. (2021). k-Nearest Neighbour Classifiers - A Tutorial. *ACM Computing Surveys, 54*(6), Article 128. https://doi.org/10.1145/3459665

Çevik, H. H., & Çunkaş, M. (2015). Short-term load forecasting using fuzzy logic and ANFIS. Neural Computing and Applications, 26(6), 1355–1367. https://doi.org/10.1007/s00521-014-1809-4

Dehalwar, V., Kalam, A., Kolhe, M. L., & Zayegh, A. (2016). Electricity load forecasting for urban area using weather forecast information. *2016 IEEE International Conference on Power and Renewable Energy (ICPRE)*, 355–360. https://doi.org/10.1109/ICPRE.2016.7871231

Deng, X., Ye, A., Zhong, J., Xu, D., Yang, W., Song, Z., Zhang, Z., Guo, J., Wang, T., Tian, Y., Pan, H., Zhang, Z., Wang, H., Wu, C., Shao, J., & Chen, X. (2022). Bagging–XGBoost algorithm based extreme weather identification and short-term load forecasting model. Energy Reports, 8, 8661–8674. https://doi.org/10.1016/j.egyr.2022.06.072

Demren, D. (2011). *Electrical load demand forecasting application using support vector machines* (M.Sc. Thesis, Istanbul Technical University, Institute of Science and Technology).

Dietrich, B., Walther, J., Weigold, M., & Abele, E. (2020). Machine learning based very short term load forecasting of machine tools. *Applied Energy, 276*, 115440. https://doi.org/10.1016/j.apenergy.2020.115440

Ertugrul, Ö. F. (2015). Forecasting electricity load by a novel recurrent extreme learning machines approach. International Journal of Electrical Power and Energy Systems, 78, 429–435. https://doi.org/10.1016/j.ijepes.2015.12.006

Esener, İ. I. (2012). *Akıllı sistemler kullanılarak güç sistemlerinde yük tahmini analizi ve uygulaması* (Yüksek Lisans Tezi, Bilecik Şeyh Edebali Üniversitesi, Fen Bilimleri Enstitüsü).

Espinoza, M., Suykens, J. A. K., Belmans, R., & De Moor, B. (2007). Electric load forecasting using kernel-based modeling for nonlinear system identification. *IEEE Control Systems Magazine*, *27*(5), 43–57. https://doi.org/10.1109/MCS.2007.904656

Ervural, B. C., Beyca, O. F., & Zaim, S. (2016). Model Estimation of ARMA Using Genetic Algorithms: A Case Study of Forecasting Natural Gas Consumption. Procedia - Social and Behavioral Sciences, 235(October), 537–545. https://doi.org/10.1016/j.sbspro.2016.11.066

Fan, G. F., Guo, Y. H., Zheng, J. M., & Hong, W. C. (2019). Application of the weighted K-nearest neighbor algorithm for short-term load forecasting. *Energies*, *12*(5), 916. https://doi.org/10.3390/en12050916

Farsi, M., Hosahalli, D., Manjunatha, B. R., Gad, I., Atlam, E. S., Ahmed, A., Elmarhomy, G., Elmarhoumy, M., & Ghoneim, O. A. (2021). Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data. Alexandria Engineering Journal, 60(1), 1299–1316. https://doi.org/10.1016/j.aej.2020.10.052

Gensler, A., & Sick, B. (2018). A Multi-Scheme Ensemble Using Coopetitive Soft-Gating With Application to Power Forecasting for Renewable Energy Generation. http://arxiv.org/abs/1803.06344

Guo, C.-Y., & Chou, Y.-C. (2020). A novel machine learning strategy for model selections - Stepwise Support Vector Machine (StepSVM). *PLoS ONE*, *15*(8), e0238384. https://doi.org/10.1371/journal.pone.0238384

Guo, W., Che, L., Shahidehpour, M., & Wan, X. (2021). Machine-Learning based methods in short-term load forecasting. *The Electricity Journal*, *34*, 106884. https://doi.org/10.1016/j.tej.2020.106884

Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. (2020). Methods and models for electric load forecasting: A comprehensive review. *Logistics & Sustainable Transport*, *11*(1), 51–76. https://doi.org/10.2478/jlst-2020-0004

Haripuddin, H., Yahya, M., Zulhajji, Z., & Yantahin, M. (2022). Peak Load Forecasting Methods of Sulbagsel Electrical Systems. Indonesian Journal of Fundamental Sciences, 8(1), 18. https://doi.org/10.26858/ijfs.v8i1.33164

He, Y., Liu, R., Li, H., Wang, S., & Lu, X. (2017). Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Applied Energy, 185, 254–266. https://doi.org/10.1016/j.apenergy.2016.10.079

Ho, P. H. K. (2010). Forecasting Construction Manpower Demand by Gray Model. Journal of Construction Engineering and Management, 136(12), 1299–1305. https://doi.org/10.1061/(asce)co.1943-7862.0000238

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geoscientific Model Development, 15(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). The distance function effect on k-nearest neighbor classification for medical datasets. SpringerPlus, 5(1), 1304. https://doi.org/10.1186/s40064-016-2941-7

Ibrahim, B., Rabelo, L., Gutierrez-Franco, E., & Clavijo-Buritica, N. (2022). Machine learning for short-term load forecasting in smart grids. *Energies*, *15*(21), 8079. https://doi.org/10.3390/en15218079

Imani, M. (2021). Electrical load-temperature CNN for residential load forecasting. Energy, 227, 120480. https://doi.org/10.1016/j.energy.2021.120480

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *IEEE Computer*, *29*(3), 31–44. https://doi.org/10.1109/2.485891

Jihad, A. S., & Tahiri, M. (2018). Forecasting the heating and cooling load of residential buildings by using a learning algorithm “gradient descent”, Morocco. Case Studies in Thermal Engineering, 12(February), 85–93. https://doi.org/10.1016/j.csite.2018.03.006

Khan, S., Khan, Z. A., Noshad, Z., Javaid, S., & Javaid, N. (2019). Short Term Load and Price Forecasting using Tuned Parameters for K-Nearest Neighbors. ITT 2019 - Information Technology Trends: Emerging Technologies Blockchain and IoT, 89–93. https://doi.org/10.1109/ITT48889.2019.9075062

Karaca, C. (2016). Çoklu Regresyon MetoduylaElektrikTüketiTalebiniEtkileyenFaktörleriİncelenmesi. Selcuk University Journal of Engineering ,Science and Technology, 4(3), 182–182. https://doi.org/10.15317/scitech.2016320514

Kaytez, F. (2020). A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption. *Energy*, *197*, 117200. https://doi.org/10.1016/j.energy.2020.117200

Kılıçarslan, S., Adem, K., & Çelik, M. (2021). An overview of the activation functions used in deep learning algorithms. Journal of New Results in Science, 10(3), 75-88. https://doi.org/10.54187/jnrs.1011739

Kang, A., Tan, Q., Yuan, X., Lei, X., & Yuan, Y. (2017). Short-Term Wind Speed Prediction Using EEMD-LSSVM Model. Advances in Meteorology, 2017. https://doi.org/10.1155/2017/6856139

Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*. Advance online publication. https://doi.org/10.1109/TSG.2017.2753802

Kotsiantis, S. B. (2011). Decision trees: A recent overview. *Artificial Intelligence Review*. Advance online publication. https://doi.org/10.1007/s10462-011-9272-4

Kumar, P., & Singh, Y. (2012). An empirical study of software reliability prediction using machine learning techniques. *International Journal of System Assurance Engineering and Management*, *3*(3), 194–208. https://doi.org/10.1007/s13198-012-0123-8

Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: A critical systematic review. *Sustainable Cities and Society, 35*, 257–270. https://doi.org/10.1016/j.scs.2017.08.009

Lahouar, A., & Ben Hadj Slama, J. (2015). Day-ahead load forecast using random forest and expert input selection. *Energy Conversion and Management, 103*, 1040–1051. https://doi.org/10.1016/j.enconman.2015.07.041

Li, C., Lin, S., Xu, F., Liu, D., & Liu, J. (2018). Short-term wind power prediction based on data mining technology and improved support vector machine method: A case study in Northwest China. *Journal of Cleaner Production, 205*, 909–922. https://doi.org/10.1016/j.jclepro.2018.09.143

Li, T., Qian, Z., & He, T. (2020). Short-Term Load Forecasting with Improved CEEMDAN and GWO-Based Multiple Kernel ELM. Complexity, 2020. https://doi.org/10.1155/2020/1209547

Lin, K. P., & Pai, P. F. (2016). Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression. *Journal of Cleaner Production, 134*,456–462. https://doi.org/10.1016/j.jclepro.2015.08.099

Mamun, A. A., Sohel, M., Mohammad, N., Sunny, M. S. H., Dipta, D. R., & Hossain, E. (2020). A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access, 8*, 134911–134938. https://doi.org/10.1109/ACCESS.2020.3010702

Mao, H., Zeng, X. J., Leng, G., Zhai, Y. J., & Keane, J. A. (2009). Short-term and midterm load forecasting using a bilevel optimization model. IEEE Transactions on Power Systems, 24(2), 1080–1090. https://doi.org/10.1109/TPWRS.2009.2016609

Matrenin, P., Safaraliev, M., Dmitriev, S., Kokin, S., Ghulomzoda, A., & Mitrofanov, S. (2022). Medium-term load forecasting in isolated power systems based on ensemble machine learning models. *Energy Reports, 8*, 612–618. https://doi.org/10.1016/j.egyr.2021.11.175

Mohammad-Zaheri, M., Chen, L., Ghaffari, A., & Willison, J. (2009). A combination of linear and nonlinear activation functions in neural networks for modeling a de-superheater. *Simul. Model. Pract. Theory*, *17*, 398-407.

Montesinos López, O.A., Montesinos López, A., Crossa, J. (2022). Fundamentals of Artificial Neural Networks and Deep Learning. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer, Cham. https://doi.org/10.1007/978-3-030-89010-0_10

Moon, J., Kim, Y., Son, M., & Hwang, E. (2018). Hybrid short-term load forecasting scheme using random forest and multilayer perceptron. *Energies*, *11*(12), 3283. https://doi.org/10.3390/en11123283

Moon, J., Kim, K. H., Kim, Y., & Hwang, E. (2018). A short-term electric load forecasting scheme using 2-stage predictive analytics. *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 219–225. https://doi.org/10.1109/BigComp.2018.00040

Naderipour, A., Kalam, A., Abdul-Malek, Z., Davoudkhani, I. F., Mustafa, M. W. Bin, & Guerrero, J. M. (2021). An effective algorithm for maed problems with a new reliability model at the microgrid. Electronics (Switzerland), 10(3), 1–23. https://doi.org/10.3390/electronics10030257

Nalcaci, G., Özmen, A., & Weber, G. W. (2019). Long-term load forecasting: Models based on MARS, ANN, and LR methods. *Central European Journal of Operations Research, 27*(4), 1033–1049. https://doi.org/10.1007/s10100-018-0531-1

Nayak, S., Bhat, M., Reddy, S., & Rao, A. (2022). Study of distance metrics on k - nearest neighbor algorithm for star categorization. *Journal of Physics: Conference Series, 2161*.

Nielsen, D. (2016). *Tree Boosting with XGBoost* (Master's thesis, Norwegian University of Science and Technology, Department of Mathematical Sciences).

Qiao, L., Wang, Z., & Zhu, J. (2020). Application of improved GRNN model to predict interlamellar spacing and mechanical properties of hypereutectoid steel. Materials

Science and Engineering: A, 792(June), 139845. https://doi.org/10.1016/j.msea.2020.139845

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106. https://doi.org/10.1007/BF00116251

Rafi, S. H., Masood, N.-A., Deeba, S. R., & Hossain, E. (2021). A short-term load forecasting method using integrated CNN and LSTM network. *IEEE Access*, *9*, 32436–32445. https://doi.org/10.1109/ACCESS.2021.3060654

Rafiei, M., Niknam, T., Aghaei, J., Shafie-Khah, M., & Catalao, J. P. S. (2018). Probabilistic load forecasting using an improved wavelet neural network trained by generalized extreme learning machine. IEEE Transactions on Smart Grid, 9(6), 6961–6971. https://doi.org/10.1109/TSG.2018.2807845

Ratsch, G., Onoda, T., & Müller, K. R. (2001). Soft margins for AdaBoost. Machine Learning, 42(3), 287–320. https://doi.org/10.1023/A:1007618119488

Saber, A. Y., & Alam, A. K. M. R. (2017). Short term load forecasting using multiple linear regression for big data. 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings, 2018-January, 1–6. https://doi.org/10.1109/SSCI.2017.8285261

Sadeghi Keyno, H. R., Ghaderi, F., Azade, A., & Razmi, J. (2009). Forecasting electricity consumption by clustering data in order to decline the periodic variable's affects and simplification the pattern. *Energy Conversion and Management*, *50*(3), 829–836. https://doi.org/10.1016/j.enconman.2008.09.036

Salah Alaloul, W., & Hannan Qureshi, A. (2020). Data Processing Using Artificial Neural Networks. IntechOpen. https://doi.org/10.5772/intechopen.91935

Sarma, P., Durlofsky, L.J., & Aziz, K. (2008). Kernel Principal Component Analysis for Efficient, Differentiable Parameterization of Multipoint Geostatistics. *Mathematical Geosciences*, *40*, 3-32.

Shirzadi, N., Nizami, A., Khazen, M., & Nik-Bakht, M. (2021). Medium-term regional electricity load forecasting through machine learning and deep learning. *Designs*, *5*(2), 27. https://doi.org/10.3390/designs5020027

Sudheer, G., & Suseelatha, A. (2015). Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models. International Journal of Electrical Power and Energy Systems, 64, 340–346. https://doi.org/10.1016/j.ijepes.2014.07.043

Şener, F. (2005). *Yük tahmin yöntemleri ve Ankara Merkez Metropol alan için regresyon analizi yöntemi kullanılarak uygulanması* (Yüksek lisans tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü).

Takeda, H., Tamura, Y., & Sato, S. (2016). Using the ensemble Kalman filter for electricity load forecasting and analysis. *Energy, 104*, 184–198. https://doi.org/10.1016/j.energy.2016.03.070

Tarmanini, C., Sarma, N., Gezegin, C., & Ozgonenel, O. (2023). Short-term load forecasting based on ARIMA and ANN approaches. *Energy Reports, 9*, 550–557. https://doi.org/10.1016/j.egyr.2023.01.060

Taunk, K., De, S., Verma, S., & Swetapadma, A. (2019). A brief review of nearest neighbor algorithm for learning and classification. In *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1255–1260). IEEE. https://doi.org/10.1109/ICICCS.2019.9065747

Taylor, J. W. (2012). Short-term load forecasting with exponentially weighted methods. IEEE Transactions on Power Systems, 27(1), 458–464. https://doi.org/10.1109/TPWRS.2011.2161780

Tutu, B. E. (2017). *Türkiye için kısa vadeli elektrik enerjisi talep tahmini* (Yüksek Lisans Tezi, TOBB Ekonomi ve Teknoloji Üniversitesi, Sosyal Bilimler Enstitüsü).

Wang, Y., Chen, J., Chen, X., Zeng, X., Kong, Y., Sun, S., Guo, Y., & Liu, Y. (2021). Short-term load forecasting for industrial customers based on TCN-LightGBM.

IEEE Transactions on Power Systems, 36(3), 1984–1997. https://doi.org/10.1109/TPWRS.2020.3028133

Wang, J., Lu, S., Wang, S. H., & Zhang, Y. D. (2022). A review on extreme learning machine. Multimedia Tools and Applications, 81(29), 41611–41660. https://doi.org/10.1007/s11042-021-11007-7

Wang, X., Xing, E. P., & Schaid, D. J. (2015). Kernel methods for large-scale genomic data analysis. Briefings in bioinformatics, 16(2), 183-192. https://doi.org/10.1093/bib/bbu024

Wen, L., Zhou, K., Yang, S., & Li, L. (2018). Compression of smart meter big data: A survey. *Renewable and Sustainable Energy Reviews*.

Uydur, C. C., Akin, F., Dursun, B., & Keskin, R. (2023). Load forecasting for Trakya University. *Conference Paper*. Retrieved from https://www.researchgate.net/publication/374743757

Valgaev, O., Kupzog, F., & Schmeck, H. (2016). Low-voltage power demand forecasting using K-nearest neighbors approach. *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, 1019–1024. https://doi.org/10.1109/ISGT-Asia.2016.7796525

Vivas, E., Allende-Cid, H., & Salas, R. (2020). A systematic review of statistical and machine learning methods for electrical power forecasting with reported mape score. Entropy, 22(12), 1–24. https://doi.org/10.3390/e22121412

Yang, J., Wu, Z., Peng, K., Okolo, P. N., Zhang, W., Zhao, H., & Sun, J. (2020). Parameter selection of Gaussian kernel SVM based on local density of training set. *Inverse Problems in Science and Engineering*, *29*(4), 536-548. https://doi.org/10.1080/17415977.2020.1797716

Yaslan, Y., & Bican, B. (2017). Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting. *Measurement*, *103*, 52–61. https://doi.org/10.1016/j.measurement.2017.02.007

Yelgeç, M. A., & Bingöl, O. (2022). Ayrık dalgacık dönüşümü ve Xgboost ile rüzgâr gücü tahmini. Uluslararası Teknolojik Bilimler Dergisi, 14(2), 58–65. https://doi.org/10.55974/utbd.1132336

Yildiz, B., Bilbao, J. I., & Sproul, A. B. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews, 73*, 1104–1122. https://doi.org/10.1016/j.rser.2017.02.023

Yu, C. N., Mirowski, P., & Ho, T. K. (2017). A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids. IEEE Transactions on Smart Grid, 8(2), 738–748. https://doi.org/10.1109/TSG.2015.2513900s

Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L., & Yang, S. (2008). RBF Neural Network and ANFIS-Based Short-Term Load Forecasting Approach in Real-Time Price Environment. *IEEE Transactions on Power Systems, 23*(3), 853–860. https://doi.org/10.1109/TPWRS.2008.922249

Zhao, H. (2007). A multi-objective genetic programming approach to developing Pareto optimal decision trees. *Decis. Support Syst., 43*, 809-826.

Zor, K., Timur, O., & Teke, A. (2017). A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting. 2017 6th International Youth Conference on Energy, IYCE 2017. https://doi.org/10.1109/IYCE.2017.8003734

ÖZGEÇMİŞ

Lisans öğrenimini 2022 yılında Namık Kemal Üniversitesi elektronik ve haberleşme mühendisliği programında tamamladı. Şu anda Trakya Üniversitesi mekatronik mühendisliği programında öğrenimine devam etmektedir.