

T.C.
KIRKLARELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

FOTOVOLTAİK SİSTEMLERDE MAKİNE ÖĞRENMESİ İLE
ENERJİ ÜRETİM TAHMİNİ

YÜKSEK LİSANS TEZİ

Onur AKTAŞ

HAZİRAN- 2024

T.C.
KIRKLARELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

**FOTOVOLTAİK SİSTEMLERDE MAKİNE ÖĞRENMESİ İLE ENERJİ
ÜRETİM TAHMİNİ**

YÜKSEK LİSANS TEZİ

Onur AKTAŞ

Tez Danışmanı: Dr. Öğr. Üyesi Murat Olcay ÖZCAN

Haziran – 2024

“FOTOVOLTAİK SİSTEMLERDE MAKİNE ÖĞRENMESİ İLE ENERJİ ÜRETİM TAHMİNİ” adlı tez çalışması adlı tez çalışması **Onur AKTAŞ** tarafından hazırlanmış olup aşağıdaki jüri tarafından **OY BİRLİĞİ** ile Kırklareli Üniversitesi Fen Bilimleri Enstitüsü Elektrik Elektronik Mühendisliği Anabilim Dalı’nda Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Danışmanı:

Dr. Öğr. Üyesi Murat Olcay ÖZCAN
Kırklareli Üniversitesi

.....

Jüri Üyeleri:

Doç. Dr. Fatih AYDIN
Balıkesir Üniversitesi

.....

Dr. Öğr. Üyesi Barış KÜÇÜKAYDIN
Kırklareli Üniversitesi

.....

Tez Savunma Tarihi: 12/06/2024

.....
Doç. Dr. H. Hale KARAYER
Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Kırklareli Üniversitesi Fen Bilimleri Enstitüsü Tez ve Proje Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum bilgileri, verileri ve dokümanları, değişik sonuç verebilecek şekilde araştırma araç gereçleri kullanmadan, işlem veya kayıt sonuçlarını değiştirmeden akademik ve etik kurallar çerçevesinde elde ettiğimi, bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, bu tezde sunduğum çalışmanın özgün olduğunu bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Onur AKTAŞ
12/06/2024

ÖZET

FOTOVOLTAİK SİSTEMLERDE MAKİNE ÖĞRENMESİ İLE ENERJİ ÜRETİM TAHMİNİ

Onur AKTAŞ

Yüksek Lisans Tezi

Kırklareli Üniversitesi

Fen Bilimleri Enstitüsü

Danışman: Dr. Öğr. Üyesi Murat Olcay ÖZCAN

Haziran 2024, 83 sayfa

Güneş ve rüzgâr gibi yenilenebilir kaynaklara ilgi ve yapılan yatırımlar her geçen gün artsa da bu tesislerin dağıtım şebekelerine sürekli enerji sağlayamama gibi bir dezavantajı bulunmaktadır. Bunun yanında tüketim kısmında da yaşanabilecek belirsizlikler, enerji arz-talep dengesizliğine sebep olabilmektedir. Bu durum tahmin çalışmalarını ön plana çıkarmaktadır. Bu tez çalışmasında iki farklı veri seti oluşturulmuş, makine öğrenimi yöntemleri ile enerji üretim tahmini yapılmıştır. İlk veri seti için model bir güneş enerji üretim tesisi kurularak, nesnelerin interneti teknolojisi (NİT) ile tesisin atmosferik verileri ve anlık akım değerleri on bir ay boyunca toplanmıştır. İkinci veri seti dünya genelinde hava tahminleri ve atmosferik veriler sunan bir web sitesinden elde edilmiştir. Elde edilen verilerden beş farklı veri kümesi oluşturularak eğitime sokulmuş, Doğrusal Regresyon, Rastgele Orman (RF), XGBoost ve LSTM makine öğrenme algoritmaları kullanılarak enerji üretim tahmini yapılmıştır. Tahmin performansını değerlendirmek için R^2 , RMSE ve MAE performans metrikleri kullanılmıştır. En iyi tahmin sonucu NİT veri seti ile yapılan eğitimde elde edilmiş, R^2 metrik değerlendirmesinden RF algoritması 0.9298, XGBoost algoritması 0.9303 sonuçlarını almıştır. İkinci veri seti ile yapılan eğitimde ise R^2 metrik değerlendirmesinden RF Algoritması 0,8869, XGBoost algoritması ise 0,8904 sonuçlarını almıştır. Bu sonuçlar, makine öğrenimi algoritmalarının enerji üretim tahmininde yüksek doğrulukla çalışabildiğini ve hem yerel atmosferik veriler hem de anlık akım verileri ile küresel hava tahmin verileri kullanılarak güvenilir tahminler elde edilebileceğini göstermektedir.

Anahtar Kelimeler: Makine Öğrenmesi, Nesnelerin İnterneti, Enerji Tahmini

ABSTRACT

ENERGY PRODUCTION PREDICTION WITH MACHINE LEARNING IN PHOTOVOLTAIC SYSTEMS

Onur AKTAŞ

MSc Thesis

Kirklareli University

Graduate School of Natural and Applied Sciences

Supervisor: Asst. Prof. Murat Olcay ÖZCAN

June 2024, 83 pages

Although interest and investments in renewable resources such as solar and wind are increasing day by day, these facilities have the disadvantage of not being able to provide continuous energy to distribution networks. In addition, uncertainties on the consumption side can cause energy supply-demand imbalances. This situation brings forecasting studies to the fore. In this thesis study, energy production prediction was made with machine learning using two different data sets. For the first data set, a sample solar power generation plant was established and atmospheric data and instantaneous current values of the plant were collected online for eleven months using the Internet of Things Technology. The second dataset was obtained from a website that provides weather forecasts and atmospheric data worldwide. Five different datasets were created from the obtained data and trained, and energy production prediction was made using Linear Regression, Random Forest, XGBoost and LSTM machine learning algorithms. R^2 , RMSE and MAE performance metrics were used to evaluate the prediction performance. The best prediction result was obtained in the training with NIT dataset, RFR algorithm obtained 0.9298 and XGBoost algorithm obtained 0.9303 results from R^2 metric evaluation. In the training with the second dataset, RF algorithm obtained 0.8869 and XGBoost algorithm obtained 0.8904 results from the R^2 metric evaluation. These results show that machine learning algorithms can work with high accuracy in energy production forecasting and reliable forecasts can be obtained using both local atmospheric data, instantaneous current data and global weather forecast data.

Anahtar Kelimeler: Machine Learning, Internet of Things, Energy Forecasting

TEŞEKKÜR

Bu çalışmanın hazırlanması sürecinde, bilgi, birikim ve tecrübeleriyle, yönlendirici ve yapıcı fikirleriyle bana yol gösteren, cesaretlendiren ve tez süresince yardımlarını esirgemeyen değerli tez danışmanım Dr. Öğr. Üyesi Murat Olcay Özcan'a en içten teşekkürlerimi sunarım.

Ayrıca, tez çalışmam boyunca büyük bir sabır ve özveri ile yanımda olan, desteğini sürekli hissettiğim hayat yoldaşım, sevgili eşim Büşra Aktaş'a ve hayatıma anlam katan biricik kızım Asya AKTAŞ' a sonsuz sevgilerimi sunarım.

İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER.....	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	xi
RESİMLERİN LİSTESİ.....	xii
SİMGELER VE KISALTMALAR.....	xiii
1. GİRİŞ.....	1
2. KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI	3
2.1. Literatür Taraması	3
2.1.1. NİT ve Akıllı Şebekeler	3
2.1.2. Makine öğrenmesi ile enerji tahmini alanındaki çalışmalar	6
2.2. Türkiye’ nin Elektrik Enerjisi Üretimi Potansiyeli.....	8
2.2.1 Enerji üretiminde Türkiye’ nin diğer ülkeler arasındaki konumu	8
2.2.2. Türkiye’ nin güneş enerjisi potansiyeli.....	10
2.3. Nesnelerin İnterneti Teknolojisi (NİT).....	11
2.3.1. NİT gelişimi, gelecek beklentisi ve kullanım alanları	13
2.3.2. NİT çalışma prensibi ve teknolojik altyapısı	14
2.4. Makine Öğrenmesi	15
2.4.1. Makine öğrenimi türleri	17
2.4.2. Makine öğrenmesi türlerinde kullanılan algoritmalar.....	20
2.4.3. Çalışmada kullanılan makine öğrenimi algoritmaları.....	21
3. MATERYAL VE DENEYSEL YÖNTEM.....	29
3.1. Materyal.....	29
3.2. DeneySEL Yöntem	31
3.2.1. Enerji üretimi	31
3.2.2. NİT ve oluşturulan sisteme uyarlanması	35
3.2.3. Veri setlerinin tanıtılması ve ön işleme	42
3.2.4. Makine öğrenmesi.....	49

4. BULGULAR VE TARTIŞMA	53
4.1. Hava Durumu Sunucusu Veri Seti ile Makine Öğrenimi	53
4.1.1. Tüm veriler ile makine öğrenimi	53
4.1.2. Atmosferik verileri kullanarak makine öğrenimi.....	57
4.1.3. Akım verileri ile makine öğrenimi.....	60
4.2. NİT ile Veri Seti Makine Öğrenimi.....	63
4.2.1 Tüm veri seti ile makine öğrenimi	63
4.2.2. NİT atmosferik verileri ile makine öğrenimi	67
5. SONUÇ VE ÖNERİLER	71
5.1. Sonuçlar	71
5.2. Öneriler.....	74
KAYNAKLAR.....	77
ÖZGEÇMİŞ	83

ÇİZELGELERİN LİSTESİ

Çizelge 2. 1.Yıllık üretimi en yüksek ülkeler ve alternatif kaynakların payı.....	8
Çizelge 2.2.Türkiye’ nin yıllık toplam enerji üretimindeki kaynakların dağılımı.....	9
Çizelge 2.3. Makine öğrenmesi türleri	17
Çizelge 2.4. Makine öğrenimi çeşitlerinin karşılaştırılması	19
Çizelge 2.5.Makine öğreniminde kullanılan algoritmalar	20
Çizelge 3.1.Kullanılan malzemeler ve teknik özellikleri	29
Çizelge 3.2.Verit setine ait otuz dakikalık dilim ve ortalamalar	44
Çizelge 3.3. Otuz dakikalık anlık değerler	45
Çizelge 3.4. RFR hiper parametreleri.....	50
Çizelge 3.5. XGB hiper parametreleri	51
Çizelge 3.6. LSTM hiper parametreleri.....	51
Çizelge 3.7. GridSearch CV bileşenleri	52
Çizelge 3.8. Girdilere ait sembol çizelgesi.....	52
Çizelge 4.1. Tüm veri seti ile elde edilen performans bulguları.....	54
Çizelge 4.2. RFR algoritması hiper parametre değerleri	54
Çizelge 4.3. XGB algoritması hiper parametre değerleri	55
Çizelge 4.4. LSTM algoritması hiper parametre değerleri.....	55
Çizelge 4.5. Hiper parametreler ile elde edilen performans bulguları.....	56
Çizelge 4.6. Girdilerin algoritmalara göre önem sıralaması.....	56
Çizelge 4.7. Atmosferik veriler ile elde edilen bulgular.....	57
Çizelge 4.8. RFR algoritması hiper parametre değerleri	58
Çizelge 4. 9. XGBoost algoritması hiper parametre değerleri	58
Çizelge 4.10. LSTM algoritması hiper parametre değerleri.....	59
Çizelge 4.11. Hiper parametreler ile elde edilen performans bulguları.....	59
Çizelge 4.12. Eğitime katılan verilerin performansa etkileri.....	60
Çizelge 4.13. Akım verileri ile elde edilen bulgular.....	60
Çizelge 4.14. RFR algoritması hiper parametre değerleri	61
Çizelge 4.15.XGB algoritması hiper parametre değerleri	61
Çizelge 4.16. LSTM algoritması hiper parametre değerleri.....	62
Çizelge 4.17.Hiper parametreler ile elde edilen performans bulguları.....	62

Çizelge 4.18. Eğitime katılan verilerin performansa etkileri.....	63
Çizelge 4.19. NİT verileri ile elde edilen bulgular	64
Çizelge 4. 20. RFR algoritması hiper parametre değerleri	64
Çizelge 4.21. XGB algoritması hiper parametre değerleri	65
Çizelge 4.22. LSTM algoritması hiper parametre değerleri.....	65
Çizelge 4.23. Hiper parametre değerleri ile elde edilen bulgular.....	65
Çizelge 4.24.Eğitime katılan verilerin performansa etkileri.....	66
Çizelge 4.25. Atmosferik veriler ile elde edilen bulgular.....	67
Çizelge 4.26. RFR algoritması hiper parametre değerleri	68
Çizelge 4.27. XGBoost algoritması hiper parametre değerleri	68
Çizelge 4.28. LSTM algoritması hiper parametre değerleri.....	69
Çizelge 4.29. Hiper parametre değerleri ile elde edilen bulgular.....	69
Çizelge 4.30. Eğitime katılan verilerin performansa etkileri.....	69
Çizelge 5.1. Atmosferik verilere ait değerlendirme sonuçları.....	71
Çizelge 5.2. Akım verileri ait değerlendirme sonuçları.....	73

ŞEKİLLERİN LİSTESİ

Şekil 2.1. Türkiye’ nin güneş enerjisi potansiyeli atlası.....	10
Şekil 2.2. Global Radyasyon değerleri ve güneşlenme süreleri.....	10
Şekil 2.3. Tekirdağ – Hayrabolu güneş enerjisi potansiyel haritası	11
Şekil 2.4. 2019 – 2030 yılları arası sektörlere göre internete bağlı cihaz sayısı	13
Şekil 2.5. NİT katmanları	14
Şekil 2.6. Yapay zekâ alt dalları	16
Şekil 2.7. Makine öğrenimi süreci.....	17
Şekil 2.8. Denetimli öğrenme akış diyagramı	18
Şekil 2.9. Denetimsiz öğrenme akış diyagramı	18
Şekil 2.10. Yarı denetimli öğrenme akış diyagramı	19
Şekil 2.11. RFR algoritması	22
Şekil 2.12. XGBoost algoritması.....	25
Şekil 3.1. PV güneş panellerinin seri bağlantısı	33
Şekil 3.2. PV güneş panellerinin paralel bağlantısı.....	33
Şekil 3.3. Akülerin seri ve paralel bağlantısı.....	34
Şekil 3.4. Kontrol panosu tek hat şeması	35
Şekil 3.5. Kurulan sistemin temel prensip şeması.....	36
Şekil 3.6. Sisteme ait bağlantı şeması	38
Şekil 3.7. Akış şeması	39
Şekil 3.8. Tüm sisteme ait blok şeması	43
Şekil 3.9. Isı haritası 1	45
Şekil 3.10. Sıcaklık, akım ve UV değerlerine ait grafik.....	48
Şekil 3.11. Rüzgâr ve akım değerlerine ait grafik	48
Şekil 3.12. Sıcaklık, nem ve akım değerlerine ait grafik.....	49
Şekil 3.13. Isı haritası 2	49
Şekil 4.1. Tüm veriler ile elde edilmiş tahmin verileri ve test verileri	57
Şekil 4.3. Akım verileri ile elde edilmiş tahmin ve test verileri grafiği	63
Şekil 4.4. Tüm veri seti ile edilmiş tahmin ve test verileri grafiği	67
Şekil 4.5 Atmosfer verileri ile elde edilmiş tahminler ile test verileri grafiği.....	70

RESİMLERİN LİSTESİ

Resim 3.1: PV güneş panellerinin montajlı görüntüsü	34
Resim 3.2. Aküler ve kontrol panosu	35
Resim 3.3: ThingSpeak platformu ara yüzü	40
Resim 3.4: Arduino IDE kodunda ThingSpeak eşleştirilmesi.....	41
Resim 3.5. ThingSpeak web görüntüsü	42
Resim 3.6. Mobil uygulama görüntüsü	42
Resim 3.7.NİT veri seti ilk 5 satırı	46
Resim 3.8. Accuweather veri setine ait ilk beş satır.....	47

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklamalar

A	Amper
KWh	Kilo Watt Saat
Ah	Amper Saat
°C	Santigrat Derece
V	Volt
%	Yüzde
W	Watt
UV	Ultraviyole
KWh/m²-gün	1m ² alanın günlük enerjisi
η	Öğrenme oranı
σ	Sigmoid fonksiyonu
Σ	Sigma, Dizi toplamı

Kısaltmalar

Açıklamalar

API	Application Programming Interface
DVM	Destek Vektör Makinesi
GEPA	Güneş Enerji Potansiyel Atlası
kNN	En Yakın Komşu
LR	Linear Regression
LSTM	Long – Short Term Memory
MAE	Mean Absolute Error
MLP	Çok Katmanlı Algılayıcılar

NİT	Nesnelerin İnterneti Teknolojisi
Preakim	Önceki Akım Değeri
Pre2akim	İki Önceki Akım Değeri
PV Panel	Photovoltaic Panel
RFR	Random Forest Regression
R²	R squared
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RF	Random Forest
TEİAŞ	Türkiye Elektrik İletim Anonim Şirketi
WEP	Wired Equivalent Privacy
WPA	Wi-Fi Protected Access
WPA2	Wi-Fi Protected Access
XGBoost	Extreme Gradient Boosting
YSA	Yapay Sinir Ağı

1. GİRİŞ

İklim değışiklikleri ve küresel ısınma karşı karşıya olunan sorunların en büyüğü olarak görülmektedir. Bugüne kadar elektrik enerjisinin üretilip kullanıma sunulması açısından en temel kaynaklar petrol, gaz, kömür gibi fosil tabandan üretilen ve sonlu kaynaklar olmuştur. Günümüzde dünya çapında tüketilen enerjinin %70' i fosil kaynaklardan elde edilmektedir [1]. Bu kaynakların tükenebilir özellikte olması ve çevreye zararları özellikle son elli yıllık süreçte büyük bir atılımın önünü açmıştır. Sera gazı emisyonları ve küresel ısınma durumlarındaki yaptırımların kısıtlayıcı hale gelmesiyle alternatif enerji kaynakları kullanımı artmaktadır [2]. Daha az maliyetle daha uzun süreli kullanım, uzun vadede bakım onarım işlemleriyle tükenmeyecek birer santral haline gelmeleri ve çevreye zararlarının daha az olması alternatif kaynakların önemli avantajları olarak sıralanmaktadır. Aynı zamanda, bu tip kaynakların sürekli enerji sağlayamaması ve üretilen enerjinin depolanamaması büyük bir sorun haline gelmektedir.

Öte yandan, tüketilen enerji miktarının saatlik, günlük hatta aylık dönemlerde farklılık göstermesi ve alternatif kaynakların hava şartlarına bağlı olarak enerji üretiminde farklılık göstermesi arz talep dengesizliği yaratmaktadır. Bu durumda alternatif kaynakların üretebileceği potansiyeli önceden tahmin etmek birçok problemin önüne geçilmesini sağlayacaktır.

Bu çalışmada temel amaç, PV panellerden elde edilecek elektrik enerjisini makine öğrenimi algoritmalarıyla tahmin etmektir. Bu kapsamda çalışma dört aşama belirlenerek gerçekleştirilmiştir.

1. İlk aşamada Tekirdağ ili Hayrabolu ilçesinde 165 W güce sahip iki adet PV panel kullanılarak model bir güneş enerjisi üretim tesisi kurulmuştur. Panellerden elde edilen elektrik enerjisi 200Ah saatlik iki adet jel akü ile depo edilmiş bir okul bahçesinin çevre aydınlatmasında kullanılmıştır.
2. İkinci aşamada nesnelerin interneti kapsamında panellerin bulunduğu ortamın sıcaklık, nem gibi atmosferik değerlerini ve panellerin ürettiği elektrik akımını ölçebilmek amacıyla Arduino mikroişlemci ve uygun duyargalar kullanılarak veriler toplanmış ve on bir ay boyunca ThingSpeak API adresine gönderilerek burada depo edilmiştir.

3. Üçüncü aşamada <https://www.accuweather.com> adresinden Hayrabolu ilçesine ait bir önceki veri seti ile aynı zaman aralığına ait hava durumu verileri on bir ay boyunca toplanmıştır.
4. Son aşamada ise elde edilen iki veri seti kullanılarak, makine öğrenimi algoritmalarıyla tahmin işlemi yapılmıştır.

Bu kapsamda, tezin ikinci bölümü literatür tarama ve kuramsal temellerin açıklanması olarak planlanmıştır. Burada makine öğrenimi ve nesnelerin interneti alanlarında yapılmış çalışmalar detaylı olarak incelenmiş ve devamında güneş enerjisinde Türkiye'nin sahip olduğu potansiyel ve bu konuda gelişmiş ülkelerin arasında bulunduğu konum anlatılmıştır. Ardından NİT' in gelişimi, beklentileri ve teknolojik alt yapısı açıklanmıştır. Son olarak yapay zekâ – makine öğrenimi kavramları açıklanarak, makine öğrenimi türleri ve türlerin özelliklerine göre kullanılabilecek algoritmalar sunulmuştur.

Üçüncü bölümde enerji üretim tesisinin kurulum aşamaları ve NİT kapsamında oluşturulan sistem ile entegre edilmesi aşamalarında uygulanan metotlar sisteme ait görseller ile anlatılmıştır. Devamında makine öğrenimi uygulamalarında kullanılan iki farklı veri setinin elde ediliş biçimleri, veri setlerinin tanıtılması ve ön işleme süreçleri açıklanmıştır.

Dördüncü bölümde, veri setleri kullanılarak uygulanan makine öğrenimi performansları paylaşılmıştır. Bu adımda iki veri seti ile toplamda beş farklı veri kümesi oluşturularak makine öğrenimi modelleri eğitilmiştir. RFR, XGBoost, LSTM ve LR algoritmaları ile yapılan uygulamaların performansları üç farklı metrik ile ölçülerek tablolar halinde paylaşılmıştır. R^2 , RMSE ve MAE metrikleri kullanılarak yapılan ilk uygulamanın ardından performansların yükseltilmesi amacıyla hiper parametre optimizasyonu ile eğitimler tekrarlanmış ve tüm sonuçlar ayrı ayrı çizelgeler ile sunulmuştur.

Son bölüm olan beşinci bölümde ise bir önceki adımda elde edilen bulgular üzerinden, algoritmaların performansları karşılaştırılarak, kullanılan özelliklerin performansa etkileri değerlendirilmiştir.

2. KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI

2.1. Literatür Taraması

Bu bölüm iki başlık altında sunulmuştur. İlk bölümde NİT'in akıllı şebekelere uygulanması ile ilgili yapılmış çalışmalar düzenlenmiştir. Diğer bölümde ise makine öğrenmesi teknikleri kullanılarak özellikle rüzgâr ve güneş gibi alternatif enerji kaynakları ile oluşturulmuş şebekelerin enerji üretim ve tüketim tahmini alanında yapılmış çalışmalar incelenmiştir.

2.1.1. NİT ve Akıllı Şebekeler

Singh vd. (2016), çalışmalarında, İrlanda'daki 6000 evden alınan akıllı sayaç verilerini kullanarak dağıtım trafolarındaki yükü tahmin etmeyi amaçlamaktadır. Çalışmada, IoT teknolojisi kullanılarak ayrıntılı tüketim verileri toplanmış ve bu verilerle yük tahmini yapılmıştır. Sonuçlar, IoT tabanlı veri toplama ve makine öğrenimi modellerinin, enerji dağıtımı ve yönetiminde daha doğru yük tahminleri yapılmasına olanak tanıdığını göstermektedir.

Kaur ve Kalra (2016), bu çalışmalarında, akıllı şebekelerin enerji verimliliğini artırmada ve enerji dağıtımında daha güvenilir bir sistem oluşturmadaki potansiyelini vurgulamaktadır. Ayrıca, IoT teknolojilerinin entegrasyonu ile enerji tüketimi ve üretiminin gerçek zamanlı izlenmesi ve yönetilmesi sağlanarak, şebekelerin daha esnek ve dayanıklı hale getirilebileceği belirtilmektedir.

Kleineidam, Krasser ve Reischböck (2016), çalışmalarında Wunsiedel' deki Akıllı Enerji Bölgesi'nde akıllı şebeke, akıllı sayaç ve akıllı ev çözümlerinin test edilmesini konu almaktadır. Bu çalışma, hücresel yaklaşım kullanarak enerji yönetimini optimize etmeyi ve sürdürülebilir enerji kullanımını teşvik etmeyi amaçlamaktadır. Ayrıca, gerçek dünyadaki uygulamaların sonuçları, enerji verimliliği ve güvenilirliğin artırılmasında önemli ilerlemeler kaydedildiğini göstermektedir.

Burritt ve Christ (2016), çalışmalarında elektrik enerjisinin veri takibi ve veri işleme işlemleriyle üretimin daha kaliteli hale geleceğini ve makine arızalarının azaldığını

belirtmiştir. Endüstri 4.0 ile arasındaki ilişkiyi inceleyerek, bu teknolojilerin dönüştürücü gücünü vurgulamaktadır.

Parmaksız vd. (2016), farklı açılarda yerleştirilmiş panellerden alınan anlık güç verileri eğim açısının güç üretimine olan etkisini gözleyebilmek için nesnelerin interneti tabanlı bir sistem oluşturmuştur.

Balamurugan ve Saravanakamalam (2017), çalışmalarında nesnelerin interneti ile tüketilen enerjinin gerçek zamanlı kontrol edilmesini sağlayarak ve sorun oluşması durumunda uzaktan müdahale edilebilecek bir tasarım oluşturmuşlardır.

Stusek vd. (2017), çalışmalarında nesnelerin interneti ile veri toplama işlemleri için kablolu veri aktarımı karşılaştırması yapabilmek için bir sistem tasarlamışlardır.

Abate vd. (2019), bu çalışmalarında NİT uygulamaları için düşük maliyetli bir akıllı güç ölçer geliştirmişlerdir. Cihaz, enerji tüketimini doğru bir şekilde izleyip raporlamak için tasarlanmış olup, hem donanım hem de yazılım bileşenlerini içermektedir. Ayrıca, güç ölçerin etkinliği ve doğruluğu test edilerek, IoT ortamlarında kullanılabilirliği gösterilmiştir.

Mudaliar ve Sivakumar (2020), çalışmalarında Raspberry Pi kullanarak IoT tabanlı gerçek zamanlı bir enerji izleme sistemi geliştirilmiştir. Bu sistem, enerji tüketimini anlık olarak izleyip kaydedebilmekte ve kullanıcıya enerji tüketim trendleri hakkında bilgi sağlamaktadır. Çalışmada, enerji verimliliğini artırmak ve enerji tasarrufunu teşvik etmek amacıyla sistemin nasıl kullanılabileceği detaylı bir şekilde açıklanmıştır. Ayrıca, sistemin kurulumu ve işleyişi hakkında teknik detaylar verilmiş ve performansı test edilmiştir. Sonuç olarak, Raspberry Pi tabanlı bu enerji izleme sistemi, enerji yönetiminde etkili ve uygun maliyetli bir çözüm olarak önerilmiştir.

Karaman, Taşkın ve Tokay (2021), çalışmalarında LabVIEW TM programı kullanılarak otomatik sayaç okuma sistemi geliştirilmiştir. Bu sistem, gerçek zamanlı enerji izleme ve tüketici farkındalığını artırmayı hedeflemektedir. Otomatik sayaç okuma sistemi, enerji tüketimini anlık olarak kaydedip analiz ederek kullanıcıların enerji kullanım alışkanlıklarını anlamalarına yardımcı olmaktadır. Çalışmada, sistemin donanım ve yazılım bileşenleri detaylı bir şekilde açıklanmış ve performansı test edilmiştir. Sonuç olarak, bu çözüm, enerji verimliliğini artırmak ve tüketicilerin enerji tasarrufu yapmalarına yardımcı olmak amacıyla etkili bir araç olarak sunulmuştur.

Kır ve Korkmaz (2021), çalışmalarında, güneş paneli parametrelerinin online ölçümü ve uzaktan izlenmesi üzerine bir sistem geliştirilmiştir. Bu sistem, güneş panellerinin verimliliğini artırmak ve performansını optimize etmek amacıyla tasarlanmıştır. Çalışmada, güneş paneli parametrelerinin gerçek zamanlı olarak izlenmesi ve analiz edilmesi için gerekli donanım ve yazılım bileşenleri detaylandırılmıştır. Sistem, güneş panellerinin çalışma koşullarını sürekli olarak izleyerek, olası arızaların ve verim kayıplarının önceden tespit edilmesine olanak tanımaktadır. Sonuç olarak, bu çalışma, güneş enerjisi sistemlerinin yönetimi ve bakımında önemli bir katkı sağlamaktadır.

Kalemci (2021), Karadeniz Teknik Üniversitesi'nde tamamladığı yüksek lisans tezinde, akıllı şebeke uygulamalarına yönelik NİT tabanlı bir uzaktan izleme sistemi geliştirilmiştir. Bu sistem, enerji dağıtım şebekelerinin etkinliğini artırmak ve anlık izleme sağlamak amacıyla tasarlanmıştır. Çalışmada, NİT' in akıllı şebeke sistemlerine entegrasyonu detaylı bir şekilde ele alınmış ve sistemin bileşenleri tanıtılmıştır. Ayrıca, geliştirilen sistemin performansı test edilerek, enerji yönetiminde sağladığı avantajlar ortaya konulmuştur. Sonuç olarak, bu tez çalışması, akıllı şebeke teknolojilerinin gelişimine katkıda bulunarak, enerji yönetiminde verimliliği artırmayı hedeflemektedir.

Akkaya (2022), Necmettin Erbakan Üniversitesi'nde tamamladığı yüksek lisans tezinde, gerçek zamanlı enerji izleme için NİT tabanlı bir akıllı sayaç tasarımı yapılmıştır. Bu akıllı sayaç, enerji tüketimini anlık olarak izleyip raporlayarak, kullanıcıların enerji verimliliğini artırmayı hedeflemektedir. Çalışmada, akıllı sayacın donanım ve yazılım bileşenleri detaylandırılmış ve entegrasyonu sağlanmıştır. Ayrıca, geliştirilen sayacın performansı test edilerek, doğruluk ve güvenilirlik açısından değerlendirilmiştir. Sonuç olarak, bu çalışma, enerji yönetimi ve tasarrufunda önemli bir araç olarak NİT tabanlı akıllı sayaçların potansiyelini göstermektedir.

Vural (2022), çalışmada fotovoltaik panellerde güç üretiminin NİT tabanlı gerçek zamanlı uzaktan izlenmesi ve verilerin saklanması ele alınmıştır. Çalışma, güneş panellerinin verimliliğini artırmak ve performansını optimize etmek amacıyla tasarlanmış bir sistem geliştirmiştir. Bu sistem, fotovoltaik panellerin güç üretim verilerini anlık olarak izleyip kaydederek, kullanıcıların enerji üretim süreçlerini daha etkin bir şekilde yönetmelerine olanak tanımaktadır. Çalışmada, izleme ve veri saklama işlemleri için gerekli olan donanım ve yazılım bileşenleri ayrıntılı olarak açıklanmıştır. Sonuç olarak,

bu çalışma, güneş enerjisi sistemlerinin yönetimi ve izlenmesinde önemli bir katkı sunarak, enerji üretim süreçlerinin daha verimli hale getirilmesini sağlamaktadır.

Taşkesen, Üren ve Üren (2023) NİT ve akıllı şebekeler arasındaki ilişkiyi ele alan bir çalışmada, NİT cihazlarının akıllı şebekeler içindeki rolünü incelemektedir. Araştırma, NİT cihazlarının enerji verimliliği, güvenlik ve yönetim açısından akıllı şebekeler için sağladığı potansiyeli vurgulamakta ve bu teknolojilerin enerji sektöründeki uygulama alanlarını tartışmaktadır. Çalışma, IoT tabanlı çözümlerin şebeke operasyonlarını optimize etmedeki önemini vurgulayarak, bu alandaki gelecek araştırmalar için yönlendirme sağlamayı amaçlamaktadır.

2.1.2. Makine öğrenmesi ile enerji tahmini alanındaki çalışmalar

Genel hatlarıyla makine öğrenmesi verilerin bulunduğu her alanda kullanılabilmektedir. Bir şirketin depo, muhasebe gibi tüm alanlarında, taşımacılık, eğitim, sağlık, ticaret ve enerji gibi tüm sektörlerde, sensörler aracılığıyla toplanacak uygun verilerle makine öğrenmesi uygulanabilmektedir. Aşağıda alternatif enerji kaynaklarından özellikle rüzgâr ve güneş enerjisi için makine öğrenmesi ile yapılmış tahmin çalışmalarına yer verilmiştir.

Sharma vd. (2011), hava tahminleri kullanarak güneş enerjisi üretimini tahmin etmek için makine öğrenimi yöntemlerini ele alan bir çalışmayı sunmaktadır. Araştırma, meteorolojik verileri ve güneş enerjisi üretim verilerini kullanarak makine öğrenimi modelleri geliştirilmesini ve bu modellerin tahmin gücünü değerlendirmeyi amaçlamaktadır. Çalışma, güneş enerjisi sistemlerinin verimliliğini artırmak için ileriye dönük tahmin yeteneklerinin önemini vurgulayarak, enerji sektöründe uygulamalar için potansiyel sağlamaktadır.

Shi vd. (2012), fotovoltaik sistemlerin güç üretimini hava sınıflandırması ve destek vektör makineleri kullanarak tahmin etmeye yönelik bir çalışma sunmaktadır. Araştırma, meteorolojik verilerin kullanımının güç üretimi tahminlerinin doğruluğunu artırmada önemli olduğunu göstererek, destek vektör makinelerinin karmaşık veri ilişkilerini ele alma yeteneğini vurgulamaktadır. Çalışma, fotovoltaik sistemlerin enerji üretimindeki tahminlerin kesinliğinin artırılmasının, enerji verimliliğinin ve enerji şebekelerine entegrasyonunun optimize edilmesinde kritik bir rol oynayabileceğini ortaya koymaktadır.

Hossain vd. (2012), yapılan çalışmada 10 kat çaprazlama kullanılarak güneş panellerinden 6 saat sonraki elde edilecek güneş enerjisinin tahmin edilmesi gerçekleştirilmiştir. Çalışmada en iyi korelasyon katsayısını MLP gösterdiği, DVM ve LMS nin de yakın değerler sergilediği belirtilmiştir.

Zeng ve Qiao (2013), kısa vadeli güneş enerjisi tahminlerini destek vektör makinesi kullanarak gerçekleştirdikleri bir çalışmayı Renewable Energy dergisinde sunmaktadır. Araştırma, güneş enerjisi üretim tahminlerinde destek vektör makinelerinin etkinliğini değerlendirerek, bu makine öğrenimi yönteminin karmaşık veri yapılarını başarılı bir şekilde işleyebilme yeteneğini vurgulamaktadır. Çalışma, güneş enerjisi sistemlerinin verimliliğini artırmak ve enerji yönetiminde daha kesin tahminler sağlamak amacıyla kısa vadeli tahmin tekniklerinin önemini vurgulamaktadır.

Yang vd. (2014), yaptıkları çalışmada güneş enerji tahmini için hibrit bir sistem sunulmuştur. Güneşlenme, yağışlılık ve sıcaklık verileriyle yapılan güç tahmin modelinde, veri eğitimi için destek vektör regresyon, bir sonraki gün enerji potansiyeli tahmini için ise bulanık mantık yapısı kullanılmıştır. Oluşturulan hibrit sistemin destek vektör makinesi ve standart sinir ağlarına göre daha başarılı olduğunu belirlemişlerdir.

Zhang ve Wang (2015), çalışmalarında güneş enerjisi tahmini k-En yakın komşu ve Çekirdek yoğunluk tahmini algoritmaları kullanılmıştır. kNN algoritması veri setinde benzer özelliklere sahip günleri ayırıştırırken, KDE ile olasılık tahmini yapılmıştır. Elde edilen sonuçların “Global Energy Forecasting Competition 2014” den elde edilen veriler ile teyit edildiği belirtilmiştir.

Yousif ve Kazem (2016), çalışmalarında güneş ışıınım değeri ve ortam sıcaklığına bağlı olarak güneş panellerinden alınabilecek güç tahmin edilmeye çalışılmıştır. Tahmin işlemleri için destek vektör ve makinesi ve çok katmanlı algılayıcılar (MLP) kullanılmış ve DVM ile 0.9844, MLP ile 0.9701 R^2 değerleri elde edilerek başarılı çalışma gerçekleştirildiği belirtilmiştir.

Wang vd. (2017), çalışmada güneş panellerinden bir gün sonra elde edilecek enerji ve güç miktarını tahmin edebilecek bir model sunulmuştur. Destek vektör makinesi ve k-en yakın komşu (kNN) algoritmaları kullanılarak oluşturulan modelde, küçük ölçekli verilerde destek vektör makinesinin, büyük ölçekli verilerde ise k-en yakın komşu algoritmasının daha yüksek performans gösterdiği belirtilmiştir.

Tang vd. (2017), yaptıkları çalışmada Lasso regresyon algoritması kullanılarak güneş panellerinden üretilecek enerjinin tahmini için bir model oluşturulmuştur. Uygulanan modelin küçük verilerde diğer yöntemlere göre daha yüksek performans sergilediği belirtilmiştir.

2.2. Türkiye' nin Elektrik Enerjisi Üretimi Potansiyeli

2.2.1 Enerji üretiminde Türkiye' nin diğer ülkeler arasındaki konumu

Yıllık radyasyon toplamı 2610 saat civarında olan ülkemizde güneş santrallerine yapılacak yatırımlar, çevre kaygısı ve uluslararası taahhütlerin yanında ekonomik dengeler sebebiyle de zorunlu görülmektedir (Melikoğlu, 2013). 2021 yılı verilerine göre Türkiye yıllık enerji üretimi sıralamasında 317 milyar kWh ile dünya ülkeleri arasında 40. Sırada bulunurken, 8152 milyar KWh ile Çin ilk sırada yer almaktadır.

Çizelge 2.1 incelendiğinde en yüksek üretim kapasitesine sahip ülkelerde alternatif enerji kaynaklarının toplam üretimdeki payı (rüzgâr ve güneş enerjisi üretimi temel alınarak hesaplanmıştır) %30 seviyesine yaklaşırken ülkemizde bu oranın %14 seviyesinde olduğu görülmektedir (<https://www.eia.gov>, 28.12.2023)

Çizelge 2.1. Yıllık üretimi en yüksek ülkeler ve alternatif kaynakların payı

Ülke	Yıllık Üretim (Milyar kWh)	Fosil + Nükleer (Milyar kWh)	Güneş Enerjisi (Milyar kWh)	Rüzgâr Enerjisi (Milyar kWh)	Alt. Kaynaklar Payı (%)
Çin	8152	5800	341	611	11,678
ABD	4165	3282	164	380	13,061
Hindistan	1702	1362	67	78	8,519
Rusya	1110	885	2,3	3,9	0,558
Brezilya	663	160	17	71	13,273
Kanada	626	198	5,1	35	6,405
Almanya	557	326	50	114	29,443
Türkiye	317	201	14	31	14,195

Türkiye Elektrik İletim A.Ş. tarafından sunulan verilere göre hazırlanmış olan çizelge 2.2, 2.3 ve 2.4’ te Türkiye’ nin yıllık enerji üretiminde kullandığı kaynakların oranları görülmektedir. Buna göre son 3 yılda yıllık toplam üretimde fosil yakıtların oranında önemli miktarda azalma görülürken, özellikle güneş enerjisi üretim oranlarında artış görülmektedir. Devam eden süreçte bölgesel, teknik, ekonomik ve çevresel kısıtlamalar ve avantajlar göz önüne alınarak uzun vadeli bir planlama ile kurulacak yeni tesislerin teknolojisi, kapasitesi ve devreye katılma durumları belirlenmeli ve yatırımlara yön verilmedir(<https://www.teias.gov.tr>, 30.12.2023).

Çizelge 2.2. Türkiye’ nin yıllık toplam enerji üretimindeki kaynakları dağılımı (2021)

Kaynak	Pay (milyar kWh)	(%)
Fosil Yakıt	211,25	63,9
Hidroelk.	55,54	16,8
Jeotermal	10,83	3,3
Güneş	12,96	3,9
Rüzgâr	31,13	9,4
Diğer	9,04	2,7
Toplam	330,77	100

Çizelge 2.3. Türkiye’ nin yıllık toplam enerji üretimindeki kaynakları dağılımı (2022)

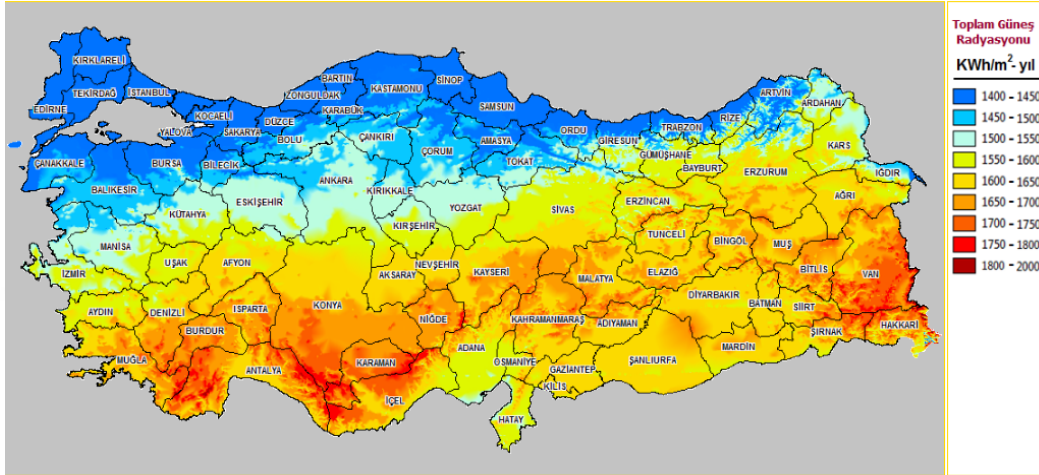
Kaynak	Pay (milyar kWh)	(%)
Fosil Yakıt	185,58	56,9
Hidroelk.	66,75	20,5
Jeotermal	11,07	3,4
Güneş	15,11	4,6
Rüzgâr	34,92	10,7
Diğer	12,54	3,8
Toplam	326	100

Çizelge 2.4. Türkiye’ nin yıllık toplam enerji üretimindeki kaynakları dağılımı (2023)

Kaynak	Pay (milyar kWh)	(%)
Fosil Yakıt	169,29	57,8
Hidroelk.	56,28	19,2
Jeotermal	10	3,4
Güneş	17,06	5,8
Rüzgâr	30,76	10,5
Diğer	9,57	3,3
Toplam	292,97	100

2.2.2. Türkiye' nin güneş enerjisi potansiyeli

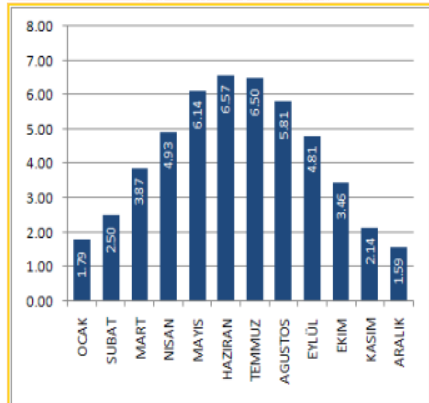
Güneş enerjisi potansiyeli atlası incelendiğinde (şekil 2.1), ülkemizin coğrafi konumu sebebiyle güney kısımlardan yukarılara gidildikçe güneşlenme ve radyasyon değerlerinin düştüğü gözlemlenmektedir. Buna bağlı olarak da hazırlanan haritada kuzey kesimlerde renklendirme sıcak tonlardan soğuk tonlara dönük olarak düzenlenmiştir.



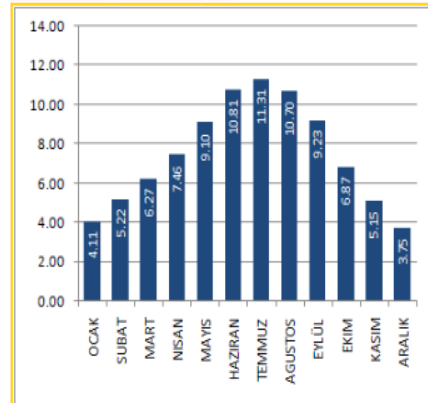
Şekil 2.1. Türkiye' nin güneş enerjisi potansiyeli atlası

Türkiye' nin yıllık güneşlenme süresi 2741 saat ve toplam ışıınım değeri 1527,46 kWh/m² olarak ölçülmüştür. 2023 yılı kasım ayı sonu itibariyle 10950 güneş santrali bulunmaktadır (<https://enerji.gov.tr/>, 28.12.2023). Bu sayısal verilere paralel olarak yıllık güneş enerji potansiyelinin 6150 milyar KWh ve yıllık ekonomik potansiyelinin 305 TWh olduğu belirtilmektedir.

TÜRKİYE Global Radyasyon Değerleri (KWh/m²-gün)



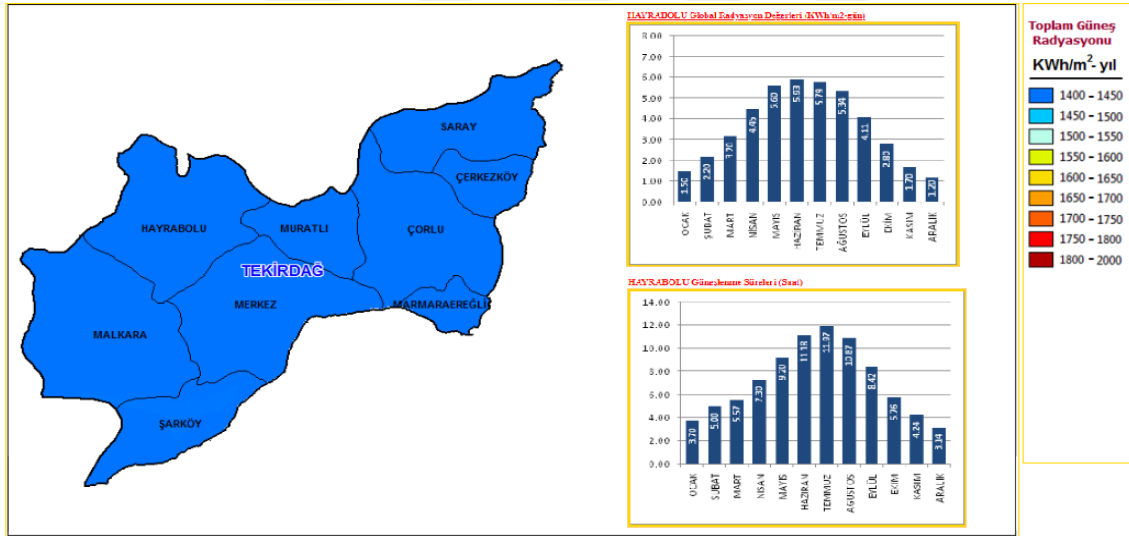
TÜRKİYE Güneşlenme Süreleri (Saat)



Şekil 2.2 Global radyasyon değerleri ve güneşlenme süreleri

GEPA' üzerinden alınan veriler ile hazırlanmış şekil 2.2' de Türkiye' de mayıs, haziran, temmuz ve ağustos aylarında en yüksek güneşlenme ve buna paralel olarak en yüksek radyasyon değerleri tespit edilmiştir. Buna göre en düşük radyasyon değeri aralık ayında gerçekleşirken, en yüksek radyasyon değeri temmuz ayı olarak belirtilmiştir. Güneşlenme verileri de aynı paralel de belirlenmiştir (Varınca ve ark., 2006). Grafik verileri günlük olarak verilmiştir. Günlük 3,75 saat ve aylık olarak 116,25 saat ile aralık ayı en az güneşlenmenin olduğu ay ve aylık 350,6 saatlik güneşlenme ile temmuz ayı en fazla güneşlenmenin olduğu ay olarak belirlenmiştir.

Bölgesel olarak bakıldığında ise güneş enerjisinden faydalanma açısından yıllık olarak en fazla güneşlenmenin gerçekleştiği bölge 1460 saat ile Güneydoğu Anadolu Bölgesi ve 1120 saat ile en az güneşlenmeye sahip bölge Karadeniz Bölgesi' dir. Marmara Bölgesi ise 1168 saat yıllık ortalama ile 6. sırada bulunmaktadır. Akdeniz ve Güneydoğu Anadolu Bölgeleri başta olmak üzere, Karadeniz bölgesi hariç tüm bölgelere yatırım yapılabilir (Onat, 2018).



Şekil 2.3. Tekirdağ – Hayrabolu güneş enerjisi potansiyel atlası

2.3. Nesnelerin İnterneti Teknolojisi (NİT)

Nesnelerin interneti kavramı gündelik hayatta kullanılan herhangi bir nesneye internete bağlanabilme yeteneğinin kazandırılmasını ifade etmektedir. Bir nesnenin farklı bir cihazla ya da bir insan ile zaman ve mekân fark etmeksizin iletişim kurabilmesi sürekli gelişen teknolojinin en son yeniliklerinden biri olarak görülmektedir. Daha geniş bir ifadeyle, her biri farklı şekilde adreslenmiş nesnelerin belirli protokoller ile dünya

apında yaygın bir aę ile birbirleriyle iletiřim kurabildikleri yeni bir dnya olarak tanımlanmaktadır. NİT’ in hızlı bir řekilde geliřmesiyle birlikte mekanik ve elektronik cihazların yanı sıra sıradan basit eřyalar bile eřitli sensr ve yazılımlarla birlikte internete baęlanabilir cihazlara dnřmřlerdir (Gndz ve ark., 2017)

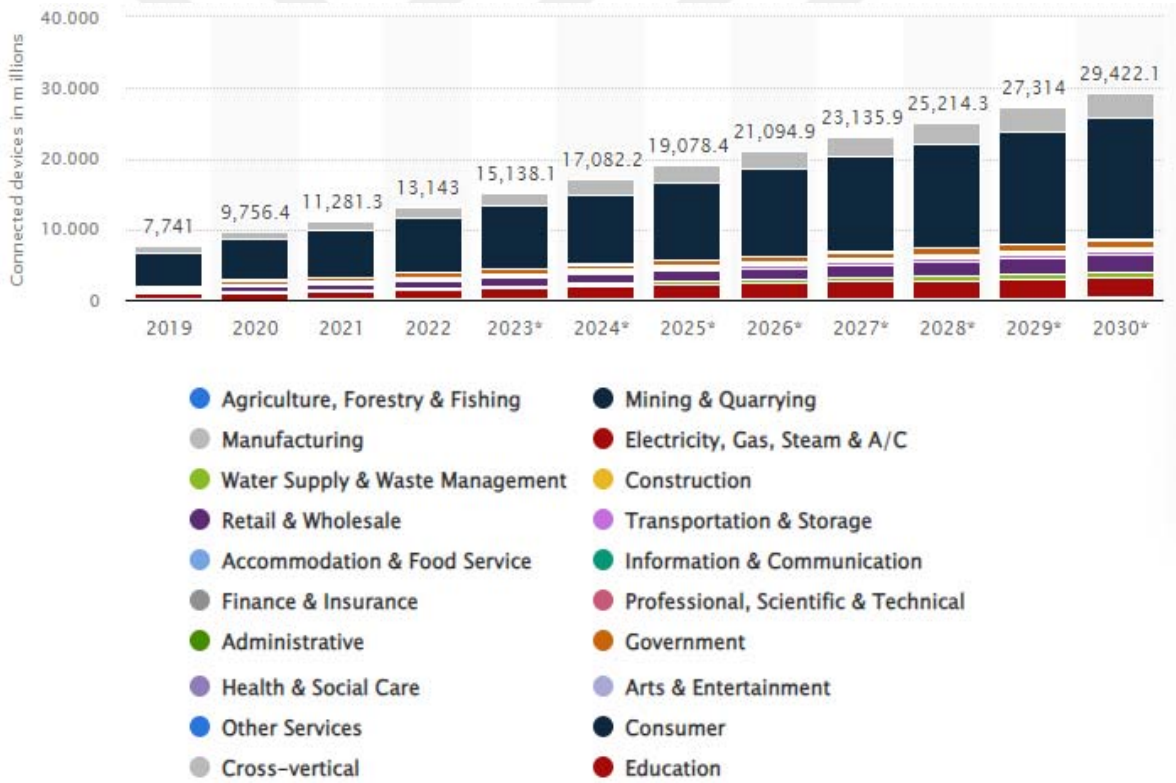
Tarımdan tařımacılıęa, spordan řehir dzenlemelerine kadar hayatın birok alanında yer almaya bařlayan NİT enerji sektrnde de grevler stlenmeye bařlamıřtır. Enerji iletim ve daęıtımında kullanılmaya bařlamasıyla akıllı řebekeler kavramı gndeme gelmiřtir (Kosmatos ve ark., 2011). Daęıtım ve iletim hatlarında kullanılan transformatrlerin periyodik bakımlar dıřında para mr tahminleri ve bakım sreleri dzenlenerek oluřacak arızaların nne geilebilmektedir. te yandan srekli artan dnya nfusu enerji tketimini de arttırmaktadır. Geleneksel řekilde fosil kaynaklardan elde edilen enerji, sera gazı emisyonlarını etkilemekte ve dnya lkeleri fosil kaynaklardan yenilenebilir enerji kaynaklarına geiř iřlemlerine hızlı řekilde devam etmektedir. Rzgr, gneř gibi yenilenebilir kaynaklar ile kurulacak santrallerde, enerji retim- tketim tahminleri, enerji optimizasyonu ve yeni yatırımlar iin doęru karar srelerinde NİT ‘ten faydalanılmaktadır.

eřitli iletiřim protokolleri ile oluřturulan ve adreslenebilen nesnelerin interneti vasıtasıyla haberleřebilmesine imkn tanıyan IOT “İnsandan- İnsana”, “İnsandan- Nesneye”, “Nesneden- Nesneye” iletiřim modlarına sahiptir. İnsan tarafından veri giriřine ihtiya duymayan, duyargalarla donatılmıř cihazlar, evreden topladıkları verilerle iletiřim kurup bir hareket iin karar verebilmektedir. Bu srecin saęlıklı bir řekilde devam ettirilebilmesi iin IOT tabanlı alıřan cihazların her birinin bir kimlięi ve iletiřim kurabilmek iin algılama yeteneklerine ihtiya vardır (Bassi ve ark., 2020).

Genel anlamda bir NİT yapısı 3 temel bileřenden oluřmaktadır. Bunlar; cihazlar, iletiřim protokolleri ve elde edilen verilerin saklandığı bulut yapıları olarak sıralanabilir. Ekosistemin en nemli bileřeni akıllı cihazlar sahip olduęu duyargalar yardımıyla evreden anlık veri toplayabilmekte ve sahip oldukları iletiřim nitesi ile dięer cihazlarla iletiřim kurabilmektedir. Bu ekosistemde ok nemli bir yeri olan son bileřen de cihazın verileri bulut ortamına aktarılmasını ve cihazların birbirleriyle haberleřmesini saęlayan, web servisi olarak da tanımlanabilecek olan iletiřim protokolleridir. Verilerin aktarımı ve cihaz haberleřmeleri aısından seilecek haberleřme protokol nem tařımaktadır (Abdel ve ark., 2019).

2.3.1. NİT gelişimi, gelecek beklentisi ve kullanım alanları

Dünya tarihi açısından kısa bir zaman dilimi sayılabilecek kadar yakın bir geçmişe sahip olan NİT, yaklaşık yirmi beş yıl önce bir fikir olarak ortaya atılmıştır. 2009 yılında bir üniversitede kahve makinesinin görüntülerinin internette paylaşılması, ilk NİT örneği olarak gösterilmektedir (Andy ve ark., 2017). Devam eden süreçte hızlı bir yükseliş gösteren bu teknoloji, günümüzde yapay zekâ ve makine öğrenimi kullanılarak veri analitiği yapılmasıyla, milyarlarca nesnenin internete bağlı olduğu bir konuma ulaşmıştır. 2030 yılına gelindiğinde internete bağlı NİT tabanlı nesne sayısının otuz milyar adete ulaşacağı tahmin edilmektedir. NİT araştırmacıları ve uygulayıcıları tarafından takip edilen bir pazar araştırma şirketi olan Statista tarafından yayınlanan 2019-2030 yılları arasındaki internete bağlı cihazların sayısını gösteren raporun özeti Şekil 2.4’ te görülmektedir (Lopez de Armentia ve ark., 2012).



Şekil 2.4. 2019-2030 yılları arası sektörlere göre internete bağlı cihaz sayısı

Tabloya göre tarım, ormancılık, balıkçılık sektöründe kullanılan internete bağlı nesne sayısı 2019 yılında 14,3 milyon iken, 2024 yılında 45,9 milyon, 2030 yılında 194,7 milyon olacağı tahmin edilmektedir. Perakende ve toptan satış sektöründe 2019 yılında 136,3 milyon iken, 2024 yılında 1.282,9 milyon, 2030 yılına gelindiğinde ise 2.518,5

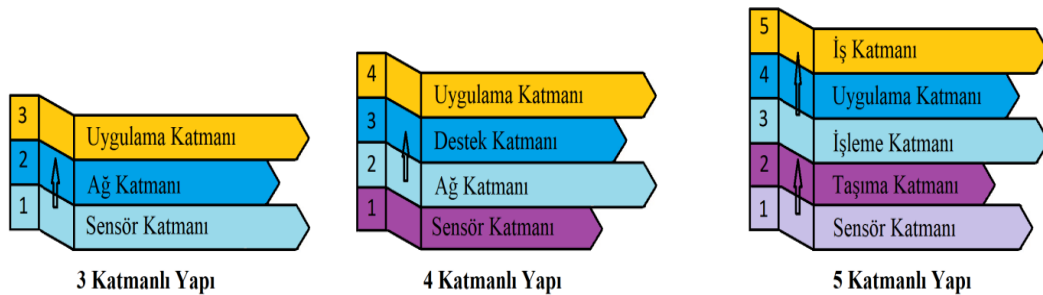
milyon olacağı öngörülmektedir (www.statista.com/, 12.01.2024). NİT sektör bazı değerlendirilip artış oran tahminleri incelendiğinde ise en fazla artış beklenen sektörlerin perakende-toptan satış, tüketim, taşımacılık ve atık yönetimi gibi sektörler olduğu görülmektedir.

2.3.2. NİT çalışma prensibi ve teknolojik altyapısı

NİT, adreslenebilme özelliğine sahip nesnelerin belirli iletişim protokolleri kullanarak internet aracılığıyla iletişim kurabilmesine olanak sağlamaktadır. İnsandan insana, nesneden nesneye ve nesneden insana iletişim modları içermektedir (Abdel Basset ve ark., 2019). Sensörlerle donatılmış cihazların çevreden aldıkları verilerle iletişim kurabilme sürecinin istenilen biçimde devam edebilmesi için her nesnenin bir kimliğe ve algılama kabiliyetine sahip olması istenmektedir.

Genel olarak bir NİT tabanlı bir cihaz nesne, veri, insan ve süreç olarak dört bileşen içermektedir. Nesne bileşeni internete bağlı cihazları tarif etmektedir. İnsan bileşeni ise etkin internet bağlantısını sağlayan bileşen ve veri bileşeni ise insan ve nesne tarafından üretilen bileşendir. Süreç bileşeni ise bu etkileşim içerisinde erişim bileşenidir (Gündüz ve ark., 2018).

NİT için kesin olarak kabul edilmiş bir mimari katman yapısı henüz söz konusu değildir (Nasiri ve ark., 2019). İlk NİT uygulamalarında üç katmanlı mimari yapı kullanılmıştır ve süreçte dört katmanlı yapı önerilmiştir. Dört katlı mimarinin veri saklama ve güvenlik özelliklerine katkı sağlamak için beş katmanlı mimari yapı önerilmiştir. Şekil 2.5’ te NİT mimari yapıları görülmektedir.



Şekil 2.5. NİT Katmanları

Sensör Katmanı: NİT cihazının çalışma ortamına göre sahip olduğu sıcaklık, nem, titreşim, solunum, nabız sayısı çevresel özelliklerin toplanabilme yeteneğini ifade etmektedir.

Ağ Katmanı: Uygulama katmanı ile sensör katmanı arasında iletişimci katman olan ağ katmanında veriler analog ve dijital sinyallere çevrilerek işlenmeden aktarılmaktadır. RFID, LoRa, NF, M2M gibi protokoller burada yer almaktadır.

Uygulama Katmanı: NİT cihazının kendisinden istenen özellikleri kullanıcıya sunduğu katmandır.

Destek Katmanı: Güvenlik gerekçesiyle eklenmiş, üç katmanlı yapıda işlenmeden gönderilen verileri bir takım ön işlemden geçiren ve depolayabilen katmandır.

Taşıma Katmanı: 5 katmanlı mimari bir yapıda işleme ve sensör katmanı arasında bulunan iletişimci katmandır.

İşleme Katmanı: Taşıma katmanından iletilen verilerde gereksiz yığınları temizleme işlemleri yapan katmandır (Rao ve ark., 2018).

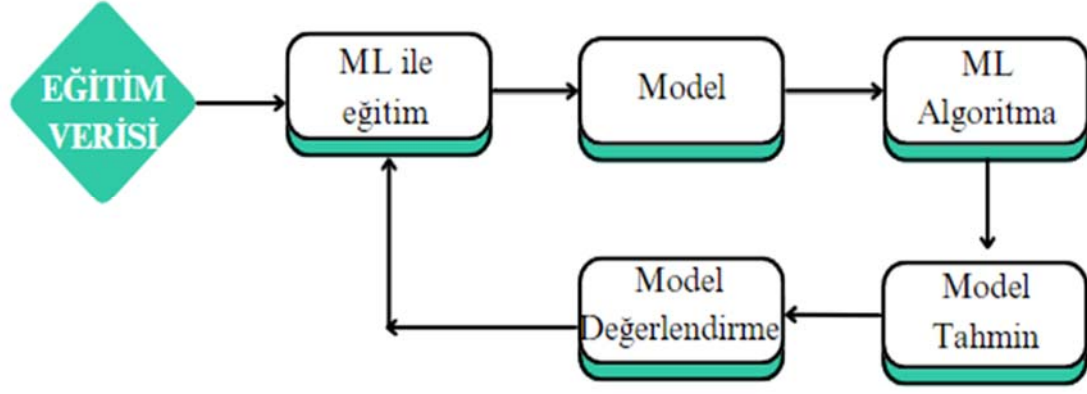
2.4. Makine Öğrenmesi

Günümüzde bilgisayar bilimlerinin en ilgi gören alanlarından biri yapay zekâ olarak kabul edilmektedir. Yapay zekâ, temelde insanların bilişsel öğrenme yöntemlerini taklit ederek problem çözme yetilerini simüle edebilme kapasitesine sahiptir (Copeland ve ark., 2007). Görüntü işleme, sinyal analizi gibi çeşitli çalışmaları içermesine rağmen, yapay zekâ aynı zamanda öğrenme kapsamında olmayan farklı yaklaşımları da barındırabilir (Sarkar ve ark., 2018). Bu yaklaşımlar, semantik analiz, doğal dil işleme gibi alanlarda yapay zekâ tekniklerinin uygulandığı ancak öğrenme sürecinden farklı şekilde çalışan metodolojileri ifade eder. Derin öğrenme ise makine öğreniminin bir türü olsa da makine öğrenmesinden ayrılan kısmı büyük ve karmaşık verilerle çalışabilmesi ve ses, video gibi farklı veri setlerini rahatça işleyebilmesidir. Şekil 2.6' de yapay zekâ, makine öğrenmesi ve derin öğrenme ilişkisi görülmektedir.



Şekil 2.6. Yapay Zekâ Alt Dalları

Makine öğrenmesi, insanların öğrenme yetilerini bilgisayarlarla taklit ederek, ham verileri analiz edip yeni bilgiler ve beceriler elde etmeyi amaçlayan bir araştırma ve çalışma alanıdır. Bu alanda yapılan çalışmalar, bilgisayar sistemlerinin öğrenme süreçlerini optimize etmeye ve karmaşık veri kümelerini işlemeye yönelik stratejiler geliştirmeyi hedefler (Wang ve ark., 2009). Yapay zeka sistemleri, insan öğrenme süreçlerinden farklı olarak daha hızlı öğrenebilir ve çabuk sonuçlar üretebilir, ancak genellikle belirli bir konuda eğitilmişlerdir (Özekici, 2021). Temel olarak makine öğrenmesi bilgisayar kodlamaları olmaksızın, büyük veri setleri ile bilgisayarlara girdi sağlayıp bunun sonucunda birtakım işlemler gerçekleştirmesini sağlayan yapay zekâ araştırmalarının bir alt dalıdır. Burada makine öğrenmesinin yüksek performans gösterebilmesi için veri setinin genişliği, veri içerisindeki örüntü ve desenlerin tutarlı olması gerekmektedir. Kompleks veri setleriyle çalışmalarda kolaylık sağlayan makine öğrenmesi, veri seti üzerinden öğrenmeler gerçekleştirip genellemelere varmaktadır (Zilyas ve ark., 2023). Genel olarak bir makine öğrenimi süreci, veri setinin elde edilmesi, gerekli veri ön işleme adımlarından sonra toplam verinin %70-%80 civarında kısmını eğitim verisi olarak ayırıp makine öğrenimi algoritmalarını kullanarak bir model oluşturulması devamında ise verinin artı kalan kısmı ile modelin test edilmesinden oluşmaktadır (Yurtsever, 2024). Bu sürece ait görsel anlatım şekil 2.7’ de gösterilmiştir.



Şekil 2.7 Makine öğrenimi süreci

2.4.1. Makine öğrenimi türleri

Makine öğrenmesi uygulamalarında problemin amacına ve istenen çıktıya göre kullanılması gereken öğrenme türü değişiklik göstermektedir. Uygulama amacına uygun yapılacak bir öğrenme türü seçimi doğru orantılı olarak modelin başarısını da artmaktadır. Çizelge 2.3’ te makine öğrenimi görevlerinin çeşitleri görülmektedir.

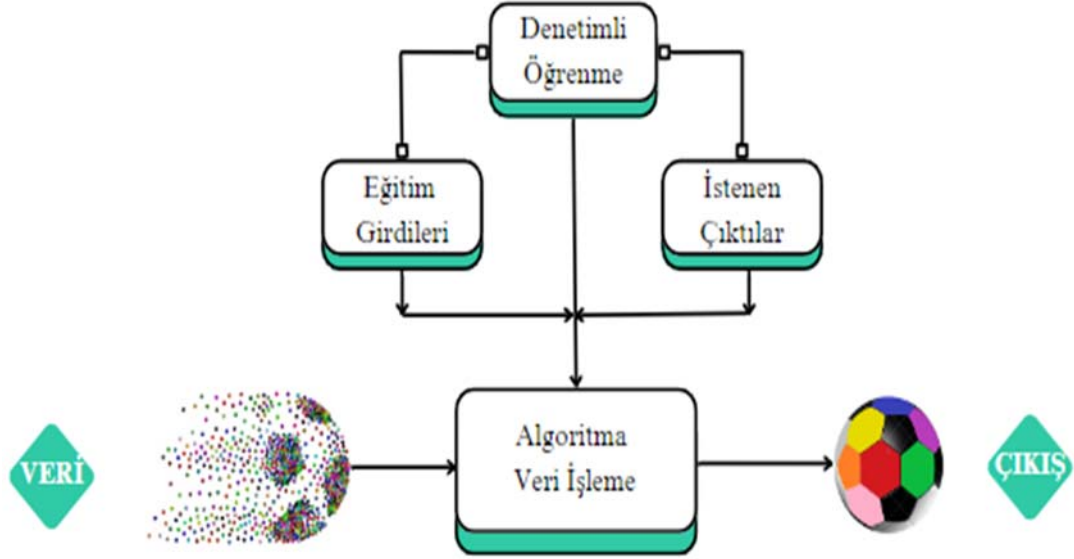
Çizelge 2.3. Makine öğrenimi türleri

Pekiştirmeli Öğrenme	Denetimli Öğrenme	Derin Öğrenme	Denetimsiz Öğrenme	Yarı Denetimli Öğrenme
Sezgisel Yöntemler	Sınıflandırma	CNN	Kümeleme	
Monte Carlo Yöntemleri	Regresyon	RNN	Boyut Azaltma	
Dinamik Programlama	Doğrusal Olmayan Reg Doğrusal Reg.			

Denetimli Öğrenme

Denetimli öğrenme genel bir tanım olarak, girilen veri ile sonuç arasındaki ilişkiyi öğrenen ve yeni veriler geldiğinde sonucu tahmin edebilecek bir makine öğrenimi türüdür (Zhang ve ark., 2021). Denetimli öğrenmenin denetimsiz öğrenmeden ayrılan en temel farkı anlamlı çıktı sağlamak için insanlar tarafından müdahaleye ihtiyaç duymasındır. Genellikle özellikleri verilen bir girdinin şartlı olasılık tahminini yapmaktadır ve birçok makine öğreniminin temeli sayılmaktadır (Liu ve ark., 2011). Sınıflandırma ve regresyon problemlerinde kullanılabilen algoritmalar ise Lineer Regresyon, Destek Vektör

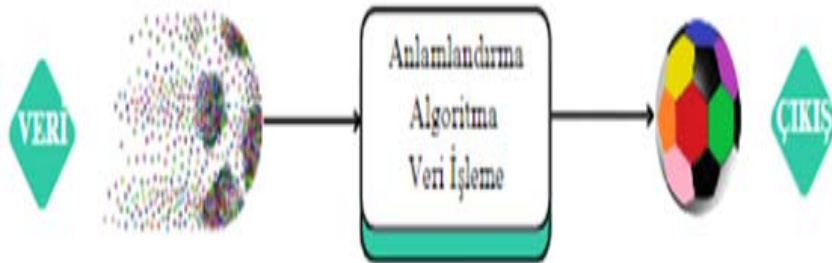
Makinesi (SVM), Karar Ağacı, En Yakın Komşu (K-Nearest Neighbors) ve RFR (Random Forest Regression) olarak sıralanabilir (Lennartsson ve ark., 2021). Şekil 2.8’ de denetimli öğrenmeye ait iş akış diyagramı görülmektedir.



Şekil 2.8. Denetimli öğrenme akış diyagramı

Denetimsiz Öğrenme

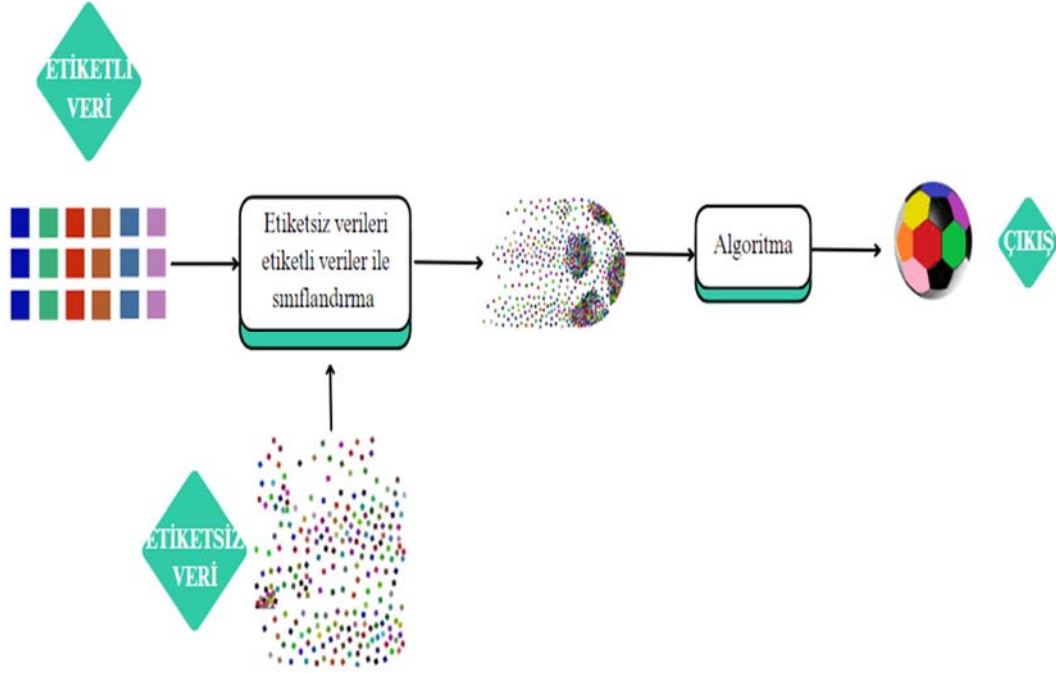
Denetimsiz öğrenme, uygulayıcı müdahalesi olmadan, etiketlenmemiş ve sınıflandırılmamış verilerle çalışabilen bir öğrenme türüdür. Temel bileşen analizi, nöral ağlar, Markov karar süreci gibi algoritmalar denetimsiz öğrenme için kullanılabilir (Chollet, 2021). Denetimsiz öğrenme genellikle boyut azaltma ve kümeleme olmak üzere iki temel yaklaşımdan yararlanır. Şekil 2.9'da denetimsiz öğrenmeye ait akış diyagramı gösterilmektedir.



Şekil 2.9. Denetimsiz öğrenme akış diyagramı

Yarı Denetimli Öğrenme

Denetimsiz öğrenme ve denetimli öğrenme arasında yer alan, veri seti içerisinde hem etiketli hem etiketsiz veri bulunduran, etiketsiz verilerin daha fazla bulunduğu veri setlerinde kullanılan öğrenme türüdür (Kızılkaya ve Oğuzlar, 2018). Şekil 2.10’ da yarı gözetimli öğrenmeye ait akış şeması görülmektedir.



Şekil 2.10. Yarı denetimli öğrenme akış diyagramı

Çizelge 2.4. Makine öğrenimi çeşitlerinin karşılaştırılması

	Denetimli Öğrenme	Denetimsiz Öğrenme	Yarı Denetimli Öğrenme
Uygulayıcı Müdahalesi	Gerek duymaz	Gerekli	Kısmen Gerekli
Veri Etiketi	Gerek duymaz	Gerekli	Kısmen Gerekli
Veri Seti	Tüm veri girdi	Test – Eğitim	Gerekli

Pekiştirmeli öğrenme

Pekiştirmeli öğrenme ismi literatürde takviyeli öğrenme olarak da geçmektedir. Takviyeli öğrenme, kendi kendine kararlar alabilen, bulunduğu ortamı algılayıp karar senaryoları oluşturan bir öğrenme türüdür (Tan ve Karaköse, 2022). Robotlar ve oyunlarda sıkça

kullanılmaktadır. Oluşturulan her problem için bir senaryo hazırlanır ve bu senaryo sayesinde hangi eylemlerin yapılması gerektiği, hangilerinin yapılmaması gerektiği dönütlerle sağlanarak öğrenme gerçekleşir (İmren ve ark., 2012).

Derin öğrenme

Derin öğrenme, makine öğrenmelerinde kullanılan verilerden daha büyük ve karmaşık verileri kullanarak anlamlı örüntüler elde etmede kullanılan bir yapay zekâ uygulamasıdır. Derin öğrenme sinir ağlarının büyüklüğüyle orantılı katmanlardan meydana gelmektedir ve bir katmandan çıkan veri bir sonraki katmana girdi olarak verilmektedir (Deng ve Yu, 2014). Nesne takibi ve görüntü işleme uygulamaları başta olmak üzere endüstri, sağlık, askeri ve akademik alanlarda sıklıkla kullanılmaktadır.

2.4.2. Makine öğrenmesi türlerinde kullanılan algoritmalar

Problemin amacına göre seçilecek olan makine öğrenimi modeli değişmektedir. Çizelge 2. 5’ te makine öğrenme modellerinde sıklıkla kullanılan algoritmalar gösterilmiştir. Bu çalışmada denetimli öğrenme başlığında, LR, LSTM, RFR ve XGBoost algoritmaları kullanılmıştır.

Çizelge 2.5. Makine öğreniminde kullanılan algoritmalar

Denetimli Öğrenme		Denetimsiz Öğrenme		Pekiştirmeli Öğrenme
Sınıflandırma	Regresyon	Kümeleme	Boyut azaltma	
SVM	LR	K-means	PCA	Markov karar
Ayırıcı analiz	SVM	K-medoids		
Navie Bayes	GPR	Fussy-C means	SVD	Q-öğrenme
En yakın komşu	XGBoost	Hiyerarşik kümeleme		
Karar ağaçları	Karar ağaçları	Gauss karışım	YSA	YSA
Rastgele orman	Rastgele orman	Gizli Markov		
YSA	YSA	YSA		

2.4.3. Çalışmada kullanılan makine öğrenimi algoritmaları

LR

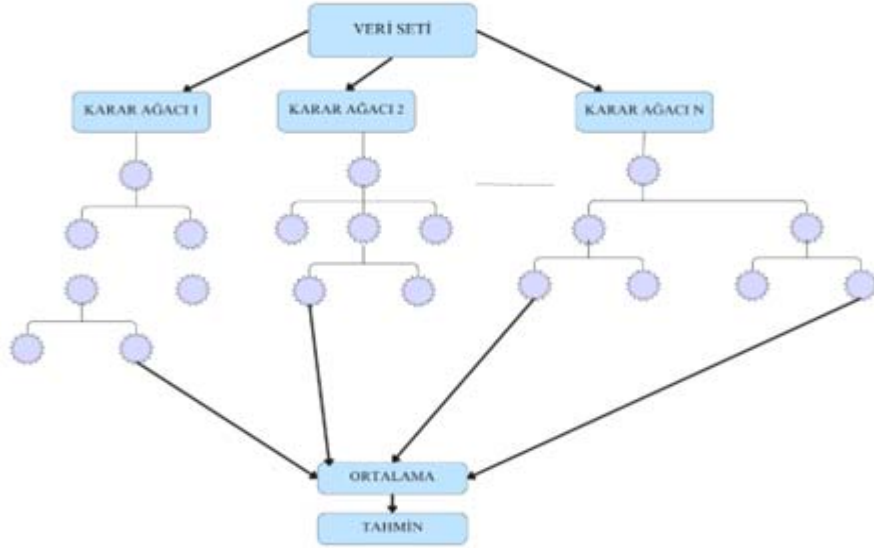
Regresyon analizinde temel amaç değişkenlerden birinin diğer değişken üzerinde olan ilişkisini matematiksel ifadeler ile ortaya çıkarmaktır. Eğer, y bağımlı değişken ve x bağımsız değişken olmak üzere, y değişkeni ve x değişkeni arasında bir ilişki açıklanabilirse, matematiksel anlamda y , x üzerinde bir regresyon olarak değerlendirilebilir [59]. Temel olarak LR bilinmeyen değerleri en iyi şekilde tahmin etmek üzere kurulmuş, birden fazla bilinen değeri içinde barındıran bir denklemin katsayılarını, aralarında doğrusal bir ilişki bulunduğunu kabul ederek tahmin etme işlemidir. Denklem 2.1’ de LR’ nin matematiksel ifadesi yer almaktadır ve Y_i bağımlı değişkeni, X_i bağımsız değişkeni, β bilinmeyenleri ve ϵ ise hatayı ifade etmektedir (Kılıç, 2013).

$$Y_i = f(x) = f(X_i, \beta) + \epsilon \quad (2.1)$$

RFR Regresyonu

Makine öğrenmesi alanında yaygın olarak kullanılan güçlü bir sınıflandırma ve regresyon yöntemi olan RFR algoritması birçok karar ağacının birleşiminden oluşur ve bu ağaçların kararlarının çokluğuna göre tahmin sonucu sunmaktadır. RFR algoritmasında her karar ağacı veri seti içerisinde rastgele seçilen veriler ile eğitilmektedir. Bu da aşırı öğrenmeyi engelleyerek genelleme kabiliyetini arttırmaktadır (Akar ve Güngör, 2012)

Büyük veri setlerinde üzerinde etkili bir şekilde çalışabilen RFR algoritması siber saldırıların tespit edilmesinde ve birçok mühendislik problemlerinde kullanılmaktadır (Bilen ve Özer, 2022). Aynı zamanda özniteliklerin seçimi ve önemlerinin sıralanması gibi işlevleri de yerine getirmektedir. Bu şekilde performansa katkı sağlayan ve gereksiz öznitelikleri belirleyerek, performansı artırılabilir (Çelik ve Özdemir, 2023). Şekil 2.11’ de RFR algoritmasına ait çalışma prensibi görülmektedir.



Şekil 2.11. RFR algoritması

Karar Ağaçlarının Oluşturulması

RFR algoritması örneklemeler kullanılarak oluşturulan çokça ağaç yapısından oluşmaktadır. Bu ağaçların her biri veri setinden rastgele atanmış N sayıda veri ile eğitilmektedir.

Karar Ağaçlarının Eğitilmesi

Oluşturulan ağaçlardan her biri N sayıda veri ile eğitilirken her düğümdeki en anlamlı bölünme Ortalama Kare Hata (Mean Squared Error- MSE) Gini Endeksi ile hesaplanabilmektedir.

*Ortalama Kare Hata (Regresyon)

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \quad (2.2)$$

Ortalama Kare Hata regresyon modellerinin performansını değerlendirmek amacıyla kullanılmaktadır. Düşük hesaplanan bir MSE değeri, modelin doğru tahminler yaptığını ifade etmektedir. Formülde 2.2' de y_i gerçek değerleri, N veri örneği sayısını ve $\hat{y}$ ise tahmin edilen değerleri ifade etmektedir

*Gini endeksi (sınıflandırma)

$$Gini = 1 - \sum_{i=1}^C p_i^2 \quad (2.3)$$

Gini endeksi' nin amacı veri setinde var olan düzensizlikleri ve çeşitlilikleri belirleyip düğümler arasındaki bölünmeleri en iyi hale getirmektir. Formülde 2.3' te C veri setindeki toplam sınıf sayısını, D değerlendirilen veri kümesini ve P_i ise alınan örneklerin toplam örnek sayısına oranını ifade etmektedir. Gini İndeksi 0 ile 0,5 arasındaki değerler alıp, 0 hesaplandığında tamamen homojenlikten söz edilebilmektedir (Chen ve ark., 2019).

Topluluk Tahmini

*Sınıflandırma

$$y = mode(y_1, y_2, \dots, y_k) \quad (2.4)$$

Topluluk Tahmini, genellikle sınıflandırma problemlerinde kullanılan bir yöntemdir. Bu yöntemde, $y_1, y_2, \dots, y_k$ gibi farklı sınıflandırıcıların tahminleri (y) topluluk içinde bir araya getirilerek, genellikle mod (yani en sık görülen sınıf) kullanılarak nihai tahmin oluşturulur. Bu şekilde, topluluk tahmini yöntemi, farklı sınıflandırıcıların tahminlerini birleştirerek daha güvenilir sonuçlar elde etmeyi amaçlar. Her bir ağacın tahmini alınarak en çok oyu alan sınıf seçilmektedir. Formül 2.4' te Y_i her bir ağacın tahminini ve k ağaç sayısını ifade etmektedir.

*Regresyon

$$y = \frac{1}{k} \sum_{i=1}^k y_i \quad (2.5)$$

Tüm ağaçların tahminlerinin ortalaması alınmaktadır. Formül 2.5' te y_i her bir ağacın tahminini ve k ağaç sayısını ifade etmektedir (Chen, 2019).

Öznitelik Önem Sırası

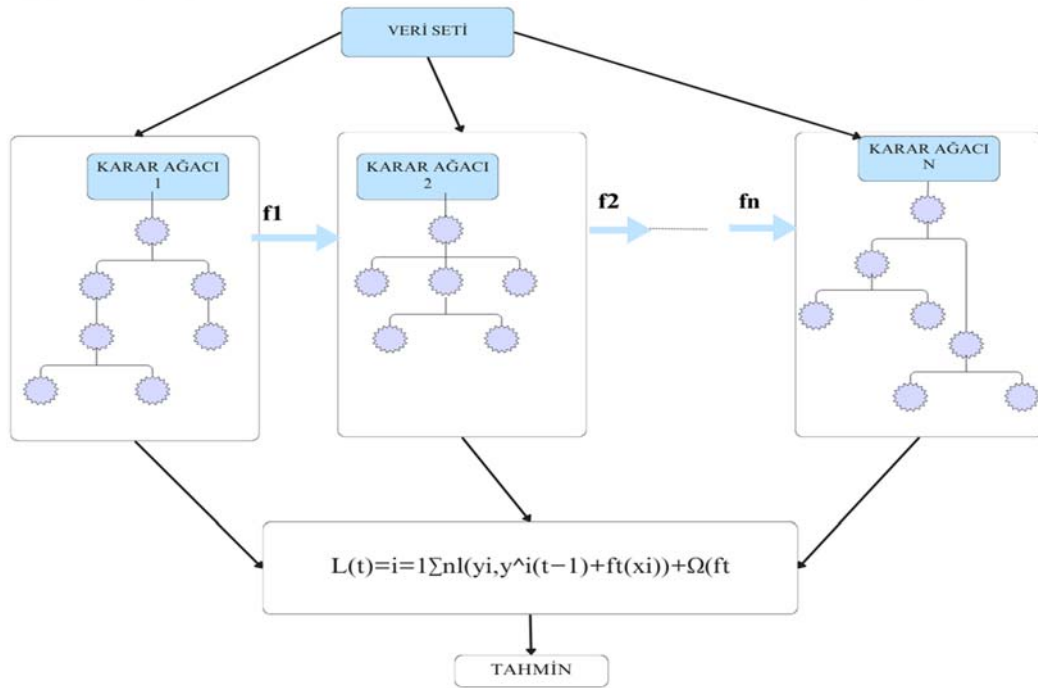
Gini endeksi ve Ortalama Kare Hata sonuçlarına göre karar ağaçlarının performansına etki eden özniteliklerin önem derecelerinin belirlenmesi adıımıdır.

XGBoost (Extreme Gradient Boosting) Algoritması

Bu makine öğrenme algoritması karar ağaçları temeline dayanan, yüksek boyutlu veri setleriyle çalışabilen bir algoritmadır. Temel amacı, ardışık karar ağaçları oluşturup bu ağaçları güçlendirerek başarılı tahmin performansları sergilemektedir (Duman, 2022).

Bu bakımdan finans piyasalarında, sağlık sektöründe ve mühendislik uygulamalarında sıklıkla kullanılmaktadır. Kalp hastalıkları üzerine yapılan bir çalışmada XGBoost algoritması uygulanan diğer modellerden daha yüksek performans sergilemiştir (Yang ve Guan, 2022). Güç transformatörlerinde hata teşhisi üzerine yapılan bir çalışmada ise öznelik seçiminde ve model veri kümelerinde başarılı sonuçlar verdiği belirtilmektedir (Wu ve ark., 2021).

XGBoost algoritması eksik değerlerin işlenmesi, modeli geliştirmeyen alanların kaldırılması gibi düzenleme yeteneklerine sahip olduğundan yüksek boyutlu verilerle çalışmaya uygun şartları sağlamaktadır (Chen ve Guestrin, 2016). Şekil 2.12’ de XGBoost algoritmasının çalışma prensibi görülmektedir.



Şekil 2.12. XGBoost algoritması

XGBoost, birçok alanda başarıyla uygulanmış, öznelik önem analizi yapabilen, eksik verilerle çalışabilen ve isabetli tahmin modelleri oluşturmak amacıyla gradyan güçlendirme kullanan bir makine öğrenme algoritmasıdır. Şekilde görülen amaç fonksiyonu ve fonksiyondaki bileşenler denklem 2.6’ da gösterilmiştir.

$$L(t) = \sum_{i=1}^n l(y_i, y_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (2.6)$$

n: veri集中的 toplam örnek sayısı

y^i : i-inci örneğin gerçek değeri

$y_i^{(t-1)}$: model tarafından tahmin edilen değeri

$f_t(x_i)$: t-inci iterasyonda öğrenilen karar ağacının i-inci örnek üzerindeki tahmini

l : gerçek değer ile modelin tahmini arasındaki farkı ifade etmektedir.

Modelin eğitim sürecinde en aza indirmeye çalıştığı toplam hata ölçüsünü ifade etmektedir. Kayıp fonksiyonu ve düzenleme fonksiyonunu içerisinde barındırır. Kayıp fonksiyonu, bir modelin eğitim sürecinde en aza indirmeye çalıştığı toplam hata ölçüsünü ifade eder. Bu fonksiyon genellikle modelin tahminlerinin gerçek değerlerden ne kadar uzak olduğunu ölçer. Modelin eğitimi sırasında amaç, kayıp fonksiyonunu minimize ederek modelin tahminlerini mümkün olduğunca doğru hale getirmektir.

Kayıp fonksiyonu aynı zamanda düzenleme terimini de içerebilir. Düzenleme terimi, modelin karmaşıklık düzeyini kontrol etmeye yardımcı olur. Modelin çok karmaşık olması, genelleme yapma yeteneğini azaltabilir ve aşırı uyuma yol açabilir. Dolayısıyla düzenleme terimi, modelin aşırı uyumu önlemek ve daha iyi genelleme yapmasını sağlamak için kullanılır.

Özetle, kayıp fonksiyonu modelin eğitim sürecinde performansını ölçen bir metrik olarak kullanılırken, düzenleme terimi ise modelin karmaşıklığını kontrol ederek genelleme yeteğini iyileştirmeye yardımcı olur (Balcı, 2022).

Yapay Sinir Ağları- LSTM (Long – Short Term Memory) Algoritması

Yapay sinir ağları, biyolojik sinir sistemlerinin işleyişinden esinlenerek geliştirilmiş, veri işleme ve modelleme için kullanılan hesaplama modelleridir. Bu ağlar, çeşitli katmanlarda organize edilmiş düğümler veya "nöronlar" ile yapılandırılmıştır. Her nöron, kendisine gelen girdileri işler, bir ağırlıklandırma işlemi uygular ve belirli bir aktivasyon fonksiyonuna göre çıktılar üretir.

YSA'nın temel bileşenleri giriş katmanı, gizli katmanlar ve çıkış katmanıdır. Giriş katmanı, ham verileri alır ve ağırlıklandırma işlemi için gerekli ilk işlemeyi yapar. Gizli katmanlar, ağırlıklandırma kapasitesini artıran çok sayıda nöronla oluşur ve bu katmanlar arasındaki bağlantılar sayesinde veri daha yüksek seviyede özellik çıkarımı yapılarak

işlenir. Çıkış katmanını ise son işlemeyi yaparak nihai sonucu üretir (Ertürk ve ark., 2023). Bir yapay sinir ağındaki nöronun matematiksel modeli denklem 2.7' de gösterilmektedir.

$$y = f(\sum_{i=1}^n w_i x_i + b) \quad (2.7)$$

Burada:

- y nöronun çıktısını,
- f aktivasyon fonksiyonunu (örneğin, sigmoid, tanh, ReLU),
- w_i ağırlıkları,
- x_i giriş değerlerini,
- b bias terimini temsil eder.

Geri yayılım (backpropagation) algoritması, ağı ağırlıklarını güncelleyerek hatayı minimize etmek için kullanılır. Ağırlık güncellemesi, türevler kullanılarak denklem 2.8' de görüldüğü şekilde yapılmaktadır.

$$w_i^{\text{new}} = w_i^{\text{old}} - \eta \frac{\partial E}{\partial w_i} \quad (2.8)$$

Burada:

- η öğrenme oranını,
- E hata fonksiyonunu,
- $\frac{\partial E}{\partial w_i}$ ise hata fonksiyonunun ağırlık w_i göre türevini temsil eder.

LSTM

LSTM, zaman serisi verileri ve dizisel bilgilerin analizinde üstün performans gösteren bir tür tekrarlayan sinir ağı (RNN) yapısıdır. RNN, özellikle sıralı ve zaman serisi verilerini işlemek için geliştirilmiş bir yapay sinir ağı türüdür. Standart yapay sinir ağlarından farklı olarak, RNN' ler geçmiş adımlardan gelen bilgileri hatırlayarak kullanabilme yeteneğine sahiptir. Bu özellik, RNN' leri dil modelleme, konuşma tanıma ve zaman serisi tahmini gibi görevlerde son derece etkili kılmaktadır.

RNN'ler, ağı iç yapısında geri besleme döngüleri içermekte; bu, ağı çıktılarının bir sonraki zaman adımında girdi olarak kullanılmasına olanak tanımaktadır. Bu geri besleme mekanizması, RNN'lerin geçmiş bilgiyi tutarak yeni bilgiyi bu geçmişle birlikte işleyebilmesini sağlar. Böylece, zaman bağımlılıklarını ve sıralı verilerdeki ilişkileri

modelleme kapasitesine sahip olurlar. Standart yapay sinir ağıları, her bir girdi örneğini bağımsız olarak işler ve çıktıları da bağımsız olarak üretir. Bu ağılar, zaman veya sıra bilgisini dikkate almaz ve katmanlar arasındaki bilgi akışı ileri yönde gerçekleşir. Bu tür ağılar, genellikle görüntü sınıflandırma gibi bağımsız ve sırasız veri setlerinde kullanılır.

LSTM ise, geleneksel RNN'lerin uzun süreli bağımlılıkları öğrenmede yaşadığı sorunları aşmak için geliştirilmiştir. Temel olarak, LSTM hücreleri, bilgi akışını daha iyi kontrol etmek için giriş, çıkış ve unutma kapıları gibi özel bileşenler kullanılır (Xin ve ark., 2023).

LSTM hücresinin matematiksel modeli şu bileşenlerden oluşur:

*Unutma kapısı (Forget gate)

Bu kapı, hücre durumunda hangi bilgilerin unutulacağını belirler. Formül 2.9' da Sigmoid aktivasyon fonksiyonu kullanarak, 0 ve 1 arasında bir değer üretir. Bu değer, hücre durumundaki bilginin ne kadarının korunacağını belirler.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.9)$$

*Giriş kapısı (Input gate)

Bu kapı, hücreye hangi yeni bilgilerin ekleneceğini belirler. Formül 2.10 ve 2.11' de Sigmoid aktivasyon fonksiyonu, hücre durumuna ne kadar yeni bilginin ekleneceğini kontrol ederken, tanh aktivasyon fonksiyonu, bu yeni bilginin ölçeğini belirler.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.10)$$

$$\check{c} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.11)$$

Burada:

- x_t Giriş vektörünü,
- h_t Hücre durumunu,
- C_t Hücre durumunu,
- f_t, i_t, o_t Sırasıyla unutma, giriş ve çıkış kapılarının aktivasyonlarını,
- W_f, W_i, W_c, W_o ağırlık matrislerini,
- b_f, b_i, b_c, b_o bias terimlerini,
- σ sigmoid fonksiyonunu,

- $\tanh$ aktivasyon fonksiyonunu temsil eder.

Bu kapılar sayesinde LSTM hücreleri hem kısa vadeli hem de uzun vadeli bağımlılıkları etkili bir şekilde öğrenebilir. Bu özellik, LSTM'yi zaman serisi tahmini, dil modelleme, konuşma tanıma ve video analizi sıralı veri problemlerinde oldukça etkili kılar.

LSTM ağlarının eğitimi, genellikle büyük veri setleri ve güçlü hesaplama kaynakları gerektirir. Ancak, doğru bir şekilde eğitildiklerinde, bu ağlar karmaşık sıralı veri ilişkilerini anlamada ve tahmin etmede üstün performans sergilemektedir (Shao ve ark., 2021).



3. MATERYAL VE DENEYSEL YÖNTEM

3.1. Materyal

Çizelge 3.1. Kullanılan malzemeler ve teknik özellikleri

Sıra	Malzeme	Malzeme Görseli	Teknik Bilgiler
1	Fotovoltaik Güneş Paneli		Polikristal Solar Modül Nominal Maksimum Güç: 165 Wp Maksimum Güçte Gerilim: 19,5 V Maksimum Güçte Akım: 8,47 A Açık Devre Gerilimi: 23,0 V Kısa Devre Akımı: 8,93 A Ebatlar: 1488x675x35 mm
2	Akü		Derin Döngülü Jel Akü Nominal Gerilim: 12 V Nominal Kapasite: 200 Ah Deşarj Süresi: 10 Hr * Nominal gerilim değeri, akü tamamen şarj ya da deşarj edildiğinde ölçülen gerilim değeridir. 200 Ah değeri, tam şarj ya da deşarjda akünün 12 V sabit gerilimde bir saat boyunca 200A akım verebileceğini göstermektedir. 10 Hr değeri ise akünün tam şarj durumunda 10 saat boyunca 20 A akım sağlayabileceğini ifade etmektedir.
3	Şarj Kontrol Cihazı		Çalışma Gerilimi: 12 / 24 V Maksimum Akım Kapasitesi: 30 A

Sıra	Malzeme	Malzeme Görseli	Teknik Bilgiler
4	Akım Sensörü		Model: ACS 712 Çalışma Gerilimi: 3/5.5 V Akım Ölçüm Aralığı: -30 / +30 A Çalışma Sıcaklığı: -40 / +85 °C
5	Sıcaklık ve Nem Sensörü		Model: DHT 11 Çalışma Gerilimi: 3.3 / 5 V Sıcaklık Ölçüm Aralığı: 0 °C / 50 °C Nem Ölçüm Aralığı: %20 / % 80 Sıcaklık Hassasiyeti: %2 Nem Hassasiyeti: %5
6	Işık Sensörü		Model: LDR Çalışma Gerilimi: 3 / 5 V Çalışma Sıcaklığı: - 30 °C / 50 °C
7	Wi-Fi Modülü		Model: ESP 8266 Çalışma Gerilimi 3.3 V Wi-Fi Standardı: IEEE 802.11 b/g/n Frekans Aralığı: 2.4 GHz Güvenlik Protokolleri: WEP, WPA/WPA2
8	Mikroişlemci		Model: Arduino Uno Mikro denetleyici: ATmega 328P Çalışma Gerilimi: 5V Giriş/Çıkış Pinleri: 14 adet (6PWM)
9	Kontrol Panosu		Polikarbon malzeme Sac tabanlı 60x40x20 cm
10	Solar İnverter		Giriş Gerilimi: 12 V Çıkış Gerilimi: 230 V Maksimum Güç: 700 W Kare Dalga Sinüs

Çalışmada, güneş enerjisi sistemleri ve NİT tabanlı alt sistemin entegre edilerek güneş panellerinden elde edilen verilerin anlık olarak toplanması ve analiz edilmesi amaçlanmıştır. Bu bölümde, bu tez kapsamında kurulan sisteme ait devre bağlantı, blok ve akış şemaları, kullanılan tüm malzemelerin teknik detayları ve projeyi gerçekleştirmede kullanılan üç ayrı bölümden oluşan yazılım kısmı anlatılmaktadır.

3.2. Deneyisel Yöntem

Bu bölüm dört farklı başlık altında anlatılmıştır. İlk bölümde enerji üretimi için kurulan deneyisel tesisin çalışması, malzemelerin bağlantı şemaları ve akış şemaları ile açıklanmıştır.

İkinci bölümde, kurulan deneyisel model ile ilişkilendirilen Nesnelerin İnterneti uygulaması, bu uygulamada seçilen donanımlar ve bulut bilişim platformu tanıtılmıştır.

Üçüncü bölümde makine öğrenme algoritmalarında kullanılacak iki adet veri setinin tanıtımı yapılmıştır. Verilerin elde edilme yöntemleri ve algoritmalara uygulanmadan önce ön işlemlerden geçirildiği detaylı olarak anlatılmıştır.

Son bölümde ise kullanılan makine öğrenimi algoritmaları ve metriklerin sonuçlarına ayrılmıştır. Girdi olarak seçilen özelliklerin tahmin başarılarına etkileri tablolar halinde sunulmuştur.

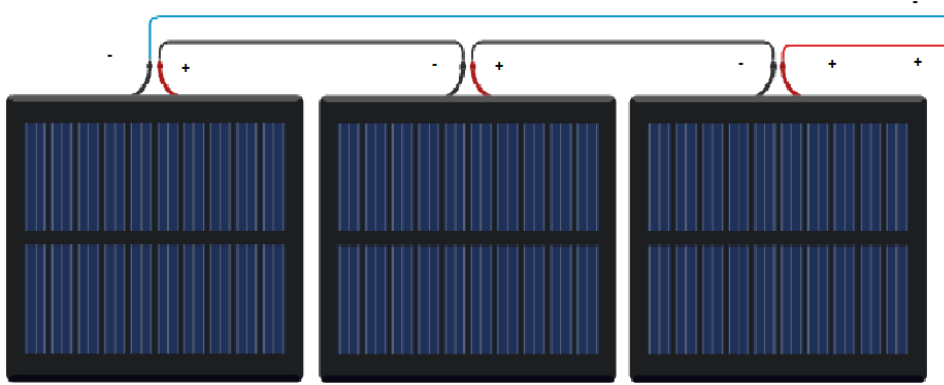
3.2.1. Enerji üretimi

PV güneş panelleri ve bağlantıları

PV güneş panelleri, güneş enerjisinden yararlanarak elektrik enerjisi üretme amacıyla tasarlanmış ekipmanlardır. Güneş panellerinin paralel bağlantısında çıkış geriliminin sabit kalması ve akım değerinin toplam panel sayısı çıkışı kadar artması beklenmektedir. Şekil 3.1 'de gösterilen seri bağlantı şeması sistem ihtiyacının doğuracağı gerilim seviyesine ulaşabilmek amacıyla yapılmaktadır. Bu bağlantı tipinde çıkış akımı sabit kalırken çıkış gerilimi ise panel sayısı kadar artabilmektedir. Şekil 3.2' de gösterilen paralel bağlantı şemasında ise çıkış yine sistem ihtiyacından belirlenecek akım değerine ulaşabilmek için yapılan bağlantı şeklindedir. Paralel bağlantıda kollar üzerinde gerilim aynı kalırken akım değeri panel sayısı kadar artabilmektedir.

$U_1= 12V, I_1=5A, P_1= 60W$ $U_2= 12V, I_2=5A, P_2= 60W$ $U_3= 12V, I_3=5A, P_3= 60W$

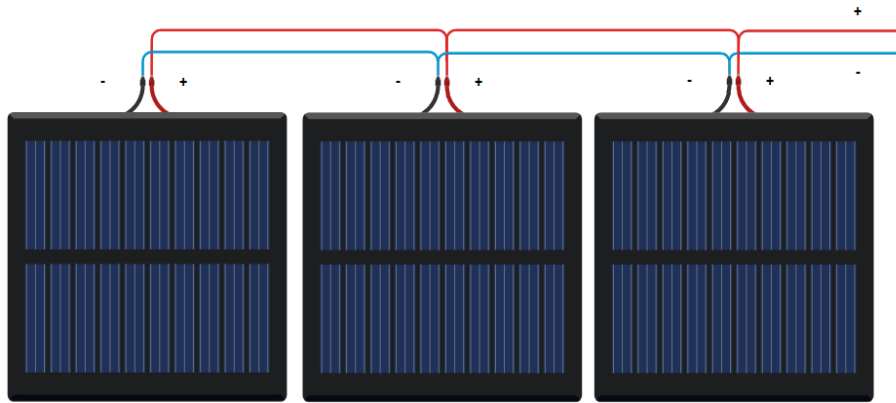
Çıkış gücü: 180 W, Çıkış Gerilimi:36 V, Çıkış Akımı: 5 A



Şekil 3.1. PV güneş panellerinin seri bağlantısı

$U_1= 12V, I_1=5A, P_1= 60W$ $U_2= 12V, I_2=5A, P_2= 60W$ $U_3= 12V, I_3=5A, P_3= 60W$

Çıkış gücü: 180 W, Çıkış Gerilimi:12 V, Çıkış Akımı: 15 A



Şekil 3.2. PV güneş panellerinin paralel bağlantısı

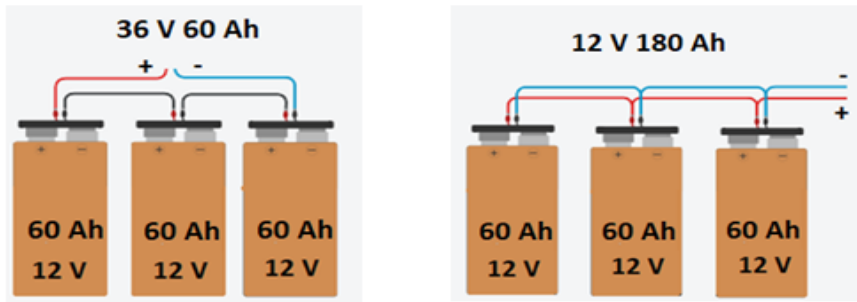
Bu çalışmada 19.5 V, 8.47 A ve 165 W etiket değerlerine sahip iki adet Polikristal güneş paneli kullanılmıştır. Panel çıkışları paralel kablolama yöntemiyle bağlanarak çıkışta maksimum 19.5V gerilim değeri, 16.94 A akım değeri ve 330 W güç değerine sahip bir sistem oluşturulmuştur. Resim 3.1’ de PV panellerin monte edildiği platform görülmektedir. Platform orta kısmında görülen 100 W değerine sahip PV panel farklı karakteristik özelliklere sahip olup elde edilecek toplam gücü azaltıcı yönde etki gösterdiğinden uygulamaya dahil edilmemiştir.



Resim 3.1. PV Güneş panellerinin montajlı görüntüsü

Akü grubu ve bağlantıları

Aküler temel olarak kimyasal enerjiyi elektrik enerjisine dönüştürmek amacıyla kullanılan elektro kimyasal donanımlardır. Üretim şekline göre farklı yerlerde kullanılmakta olup fiyat, enerji yoğunluğu, güç yoğunluğu ve şarj metodu bakımından farklılıklar göstermektedirler (Kayıklı ve Balıkçı, 2011). Aküler genel anlamda ıslak ve kuru tip aküler olarak iki grupta üretilmektedir. Islak tip bataryalarda elektrot görevi sıvı ile, kuru tip akülerde ise pasta veya jel kullanılmaktadır (megep.meb.gov.tr, 28.04.2024). Akülerin bağlantı çeşitlerinde de iki grupta bulunmaktadır. Şekilde 3.3’ te görülen seri bağlantıda akü kapasitesi sabit kalıp çıkış gerilimi artırılırken, paralel bağlantıda çıkış gerilimi sabit kalmakta ve akü kapasitesi artırılmaktadır.



Şekil 3.3. Akülerin seri ve paralel bağlantısı

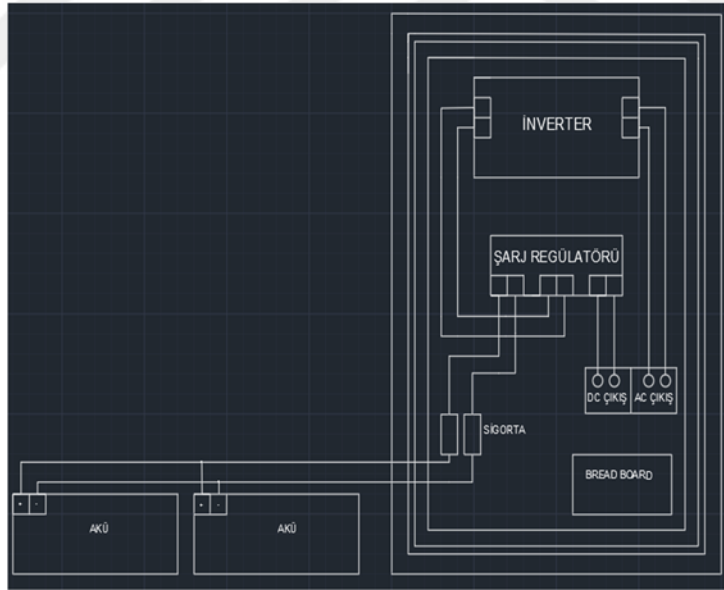
Kontrol tablosu

Yapılan çalışmada 12V çıkış gerilimine sahip 200 Ah’ lik iki adet jel akü kullanılmıştır (Resim 3.2). Bağlantı kısmında ise paralel kablolama tercih edilmiş olup 12 V gerilim değerine ve 400 Ah kapasite değerine sahip bir enerji depolama sistemi oluşturulmuştur.



Resim 3.2. Aküler ve kontrol panosu

PV paneller ile üretilen enerji 30 A 12 V değerlerine sahip şarj kontrol cihazı ile akülere düzenli bir şekilde depolanmıştır. Depolanmış bu enerji 700 W 230 V sinüsoidal çıkışa sahip inverter vasıtasıyla kullanıma hazır hale getirilmiştir.



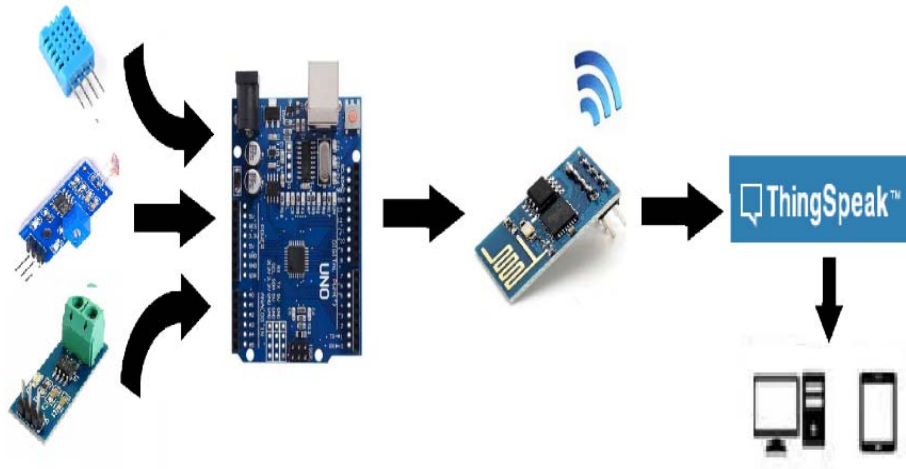
Şekil 3.4. Kontrol panosu tek hat şeması

Çalışmada enerji tüketimi ile ilgili bir araştırma yapılmamakla birlikte oluşturulan sistemden elde edilen enerji okul bahçe aydınlatmasında kullanılmıştır. Şekil 3.4'te CAD programında çizilmiş pano tek hat şeması görülmektedir.

3.2.2. NİT ve oluşturulan sisteme uyarlanması

NİT'in günlük hayatta kullanımı teknolojik gelişmelere paralel olarak artmaya devam etmektedir. Kullanım alanlarının genişlemesi ve farklı alanlarda pratik çözümler sunabilmesi bu teknolojiye ilginin artmasını sağlamaktadır (Swan, 2012). Bugün tarım ve hayvancılık, ulaşım ve taşımacılık, akıllı evler ve şehirler, akıllı şebekeler gibi fazlaca alanda NİT' in kullanıldığı görülmektedir.

Bu çalışmada NİT PV paneller üzerinde uygulanarak akıllı şebekeler alanında yer alabilecek bir tasarım oluşturulmuştur. Çalışmada PV panellerden anlık olarak alınan akım değerlerini ve panellerin bulunduğu ortamdaki ışık, nem ve sıcaklık verilerini uygun sensörler ile periyodik olarak mikroişlemciye göndermek, burada işlenen verilerin wi-fi ile açık kaynaklı internet uygulamasına gönderilmesi ve burada depolanması amaçlanmıştır. Aynı zamanda toplanıp depolanan bu veriler diğer araştırmacılar tarafından farklı projelerde kullanılabilmesi amacıyla kullanıma açık kanal olarak paylaşılmıştır. Sistemin prensip olarak çalışması şekil 3.5' te gösterilmiştir.



Şekil 3.5. Kurulan sistemin temel prensip şeması

Arduino UNO

Arduino UNO, basit kullanımı, esnek programlama seçenekleri ve geniş donanım uyumluluğuyla, çeşitli projelerde ve uygulamalarda kullanılan ve güçlü özellikleri ile mikro denetleyici tabanlı geliştirme platformudur. ATmega328P mikro denetleyiciye dayalı bir karttır ve çeşitli giriş/çıkış pinleri aracılığıyla sensörler, motorlar ve diğer harici bileşenlerle iletişim kurabilir.

ESP 8266 wi-fi modülü

ESP8266, IoT (Nesnelerin İnterneti) ve kablosuz bağlantı uygulamaları için tasarlanmış, yüksek entegrasyonlu bir Wi-Fi modülüdür. 2,4 GHZ frekans bandında iletişim sağlamak ve TCP/IP dahili haberleşme protokolü ile çalışmaktadır. Nesnelerin interneti uygulamaları, uzaktan kontrol sistemleri ve endüstriyel otomasyon sistemleri gibi alanlarda kullanılmaktadır.

DHT 11 ısı ve nem sensörü

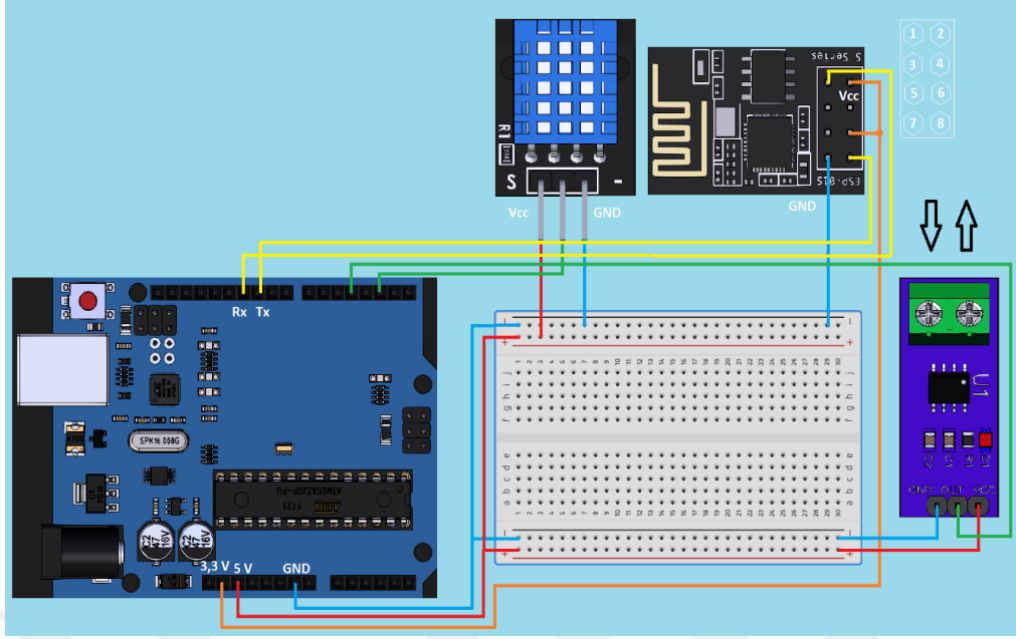
DHT 11, çevresel parametrelerin incelenmesi ve takip edilebilmesi için kullanılan, güvenli ve hassas ölçümler yapabilen bir sensör çeşididir. DHT 11 sensörü, NTC termistör ve nem duyarlı polimer kapasitif bir nem ölçer içermektedir. Hava koşullarını izleme, tarımsal uygulamalar ve ev otomasyonlarında kullanılabilmektedir.

ACS 712 akım sensörü

Elektrik akımı iletken bir yoldan aktığında aynı zamanda bir manyetik alan üretir. Akım değerinin ölçülebilmesi için Hall sensörü kullanılarak, manyetik alan ile orantılı bir gerilim üretilir (Sâri ve ark., 2021). ACS 712 akım sensörü de Hall prensibiyle çalışarak, elektrik ve elektronik devrelerde akım ölçümlerini hassas ve güvenilir bir şekilde yapabilmektedir. ACS 712 ile 5A ve 25A akım sınırlamalarında ölçüm yapılabilmektedir.

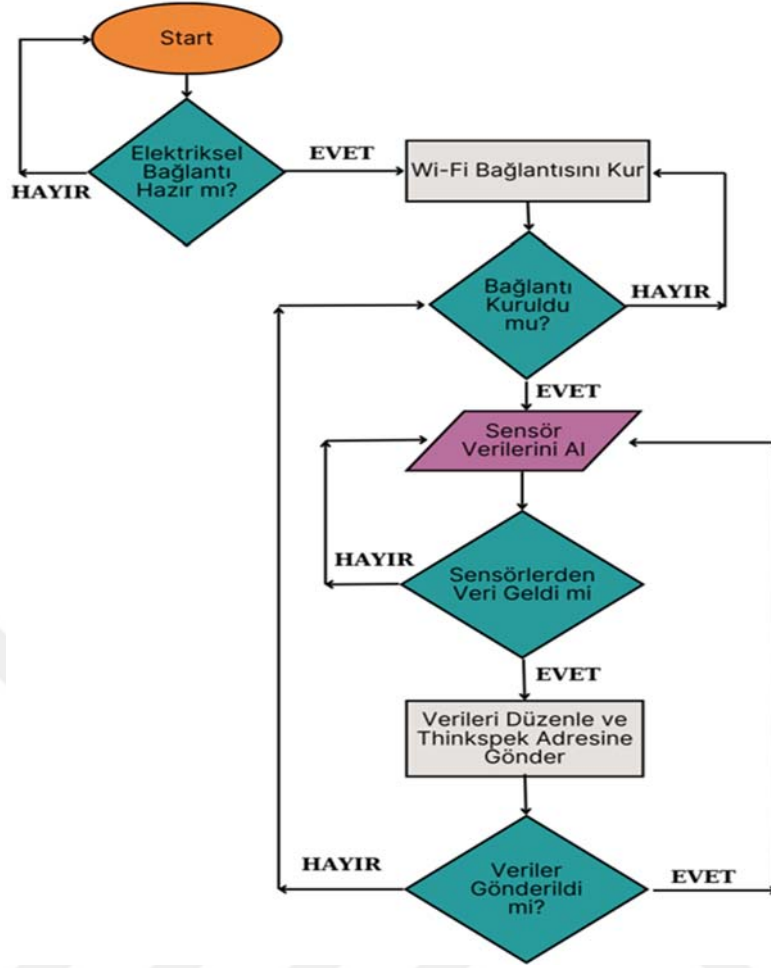
Sistem ait bağlantı şeması ve çalışma prensibi

Çalışmada kullanılan mikroişlemci, sensör ve wi-fi modülüne ait bağlantı şeması şekil 3.6' da görülmektedir. Burada DHT11 sıcaklık – nem sensörü ve ACS 712 akım sensörü enerji besleme girişleri 5V ancak ESP8266 modülünün besleme gerilimi 2 ve 6 numaralı pinler üzerinden 3,3 V olarak yapılmaktadır. Esp8266 modülünün veri alışverişinde kullanılması gereken 2 ve 6 numaralı pinleri devre şemasında sarı hatlar ile gösterilmiştir ve mikroişlemcinin 10-11 numaralı pin girişlerine irtibatlandırılmıştır.



Şekil 3.6. Sisteme ait bağlantı şeması

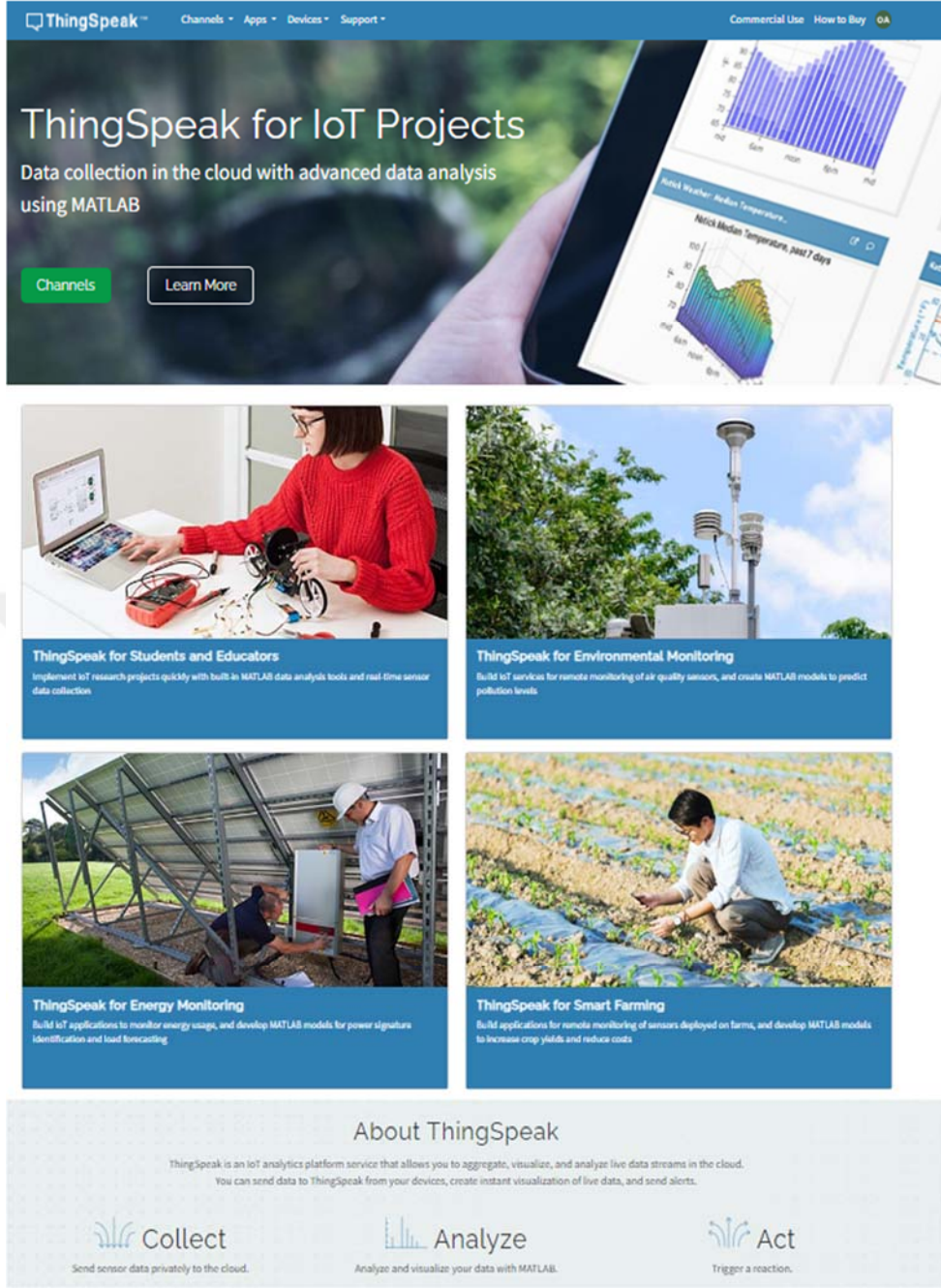
Burada Tx veri gönderme, Rx ise veri almada kullanılmaktadır. 3 ve 5 numaralı pinler dijital giriş ve çıkış olarak kullanılan pinler olmasına rağmen bu çalışmada kullanılmamıştır. Boşta kalan 4 numaralı pin ise wi-fi modülünü sıfırlama işleminde kullanılmaktadır. Devrede kullanılan diğer bir sensör olan ACS 712 akım sensörünün üç pin girişi bulunmaktadır. Bunlardan ikisi besleme girişleri ve son giriş ise mikroişlemcinin analog pin çıkışlarına bağlanmaktadır. Sensörün arka kısmında bulunan klemense ise akım değeri ölçülecek alıcının elektriksel bağlantısı yapılmaktadır.



Şekil 3.7 Akış şeması

ThingSpeak veri izleme ve depolama platformu

ThingSpeak, buluttaki veri akışlarını toplamaya, görselleştirmeye ve analiz etmeye olanak tanıyan bir NİT analiz platformudur (<https://thingspeak.com/>, 16.05.2024). İnternet vasıtasıyla Arduino, Raspberry gibi mikroişlemcilere bağlı sensörlerden algılanan çevresel verileri depolayıp, görsel olarak analiz etmeye imkân sunmaktadır. Resim 3.3 da ThingSpeak platformu ara yüzü görülmektedir.



Resim 3.3. ThingSpeak platformu ara yüzü

ThingSpeak platformuna veri göndermek ve bu verileri analiz edebilmek için ilgili web sitesine üyelik oluşturmak gerekmektedir. Ardından verilerin gönderileceği bir kanal oluşturulup, ilgili kanala ait ayarlamaları yaptıktan sonra, kanala ait API ve IP adresini sistem kod şemasına işlemek gerekmektedir. Resim 3.4'te anahtar ve IP adresinin kodlanması görülmektedir.

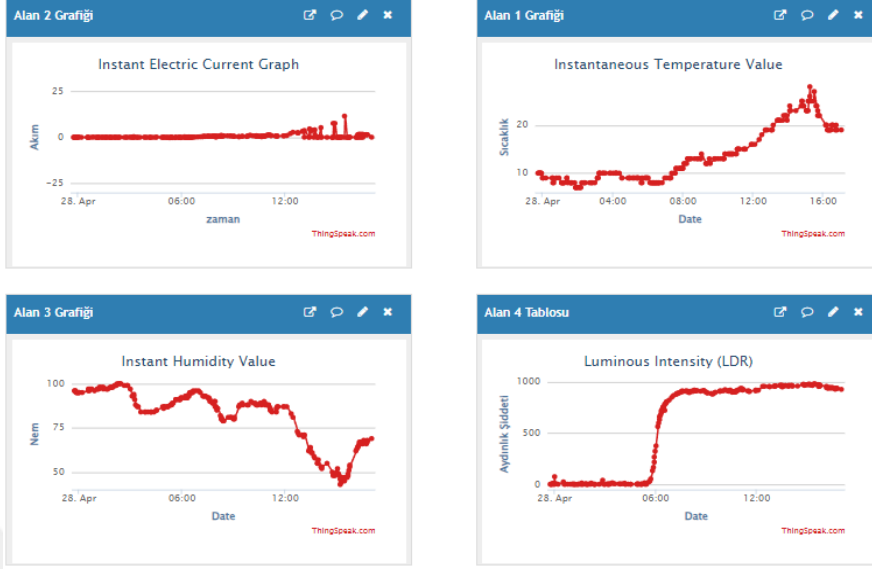


Resim 3.4. Arduino IDE kodunda ThingSpeak eşleştirilmesi

ThingSpeak platformu üzerinde ücretsiz üyelik ile 4 farklı kanal oluşturup her kanalda sekiz farklı veri girişi yapılabilen üç milyon veri saklanabilmektedir. Daha sonra saklanan bu veriler .csv uzantılı dosya şeklinde indirilebilmektedir. Resim 3.5'te uygulamaya ait bilgisayar ekranı görüntüsü ve resim 3.6' da mobil uygulama görüntüsü görülmektedir.

Kanal İstatistikleri

Oluşturuldu: yaklaşık bir yıl önce
Son giriş: 2 gün önce
Girişler: 103951



Resim 3.5. ThingSpeak web görüntüsü



Resim 3.6. Mobil uygulama görüntüsü

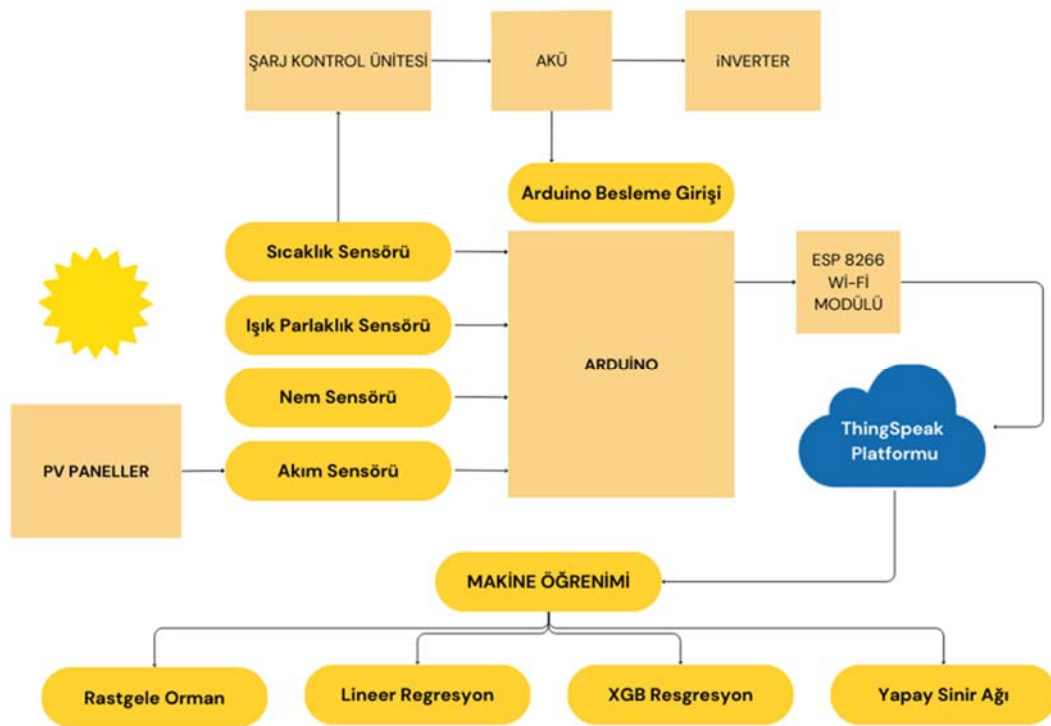
Çalışmada ThingSpeak platformuna ilk veri 18.04.2023 saat 13.30' da gönderilmiştir ve 18.03.2024 saat 11.30' da son veri alımıyla birlikte veri transferi durdurulmuştur. 11 ayda, 2 dakikalık periyodik aralıklarla akım, sıcaklık, nem ve ışık parlaklık değerleri toplanarak 93851 veriden oluşan bir veri seti oluşturulmuştur.

Veri toplama sürecinde, sensör arızaları, internet bağlantısının kesilmesi gibi sebeplerle veri akışında kısa süreli aksamalar ve iki dakikalık periyotlarda süre farklılıkları

yaşanmış, sistemin tekrar işler hale getirilmesi esnasında anlamsız ve aykırı değerler elde edilmiştir.

Elde edilen veri seti, makine öğrenme algoritmalarında kullanılmadan önce, Python programlama diliyle ön işleme sürecine alınmıştır. Veri işleme sürecinde yukarıda bahsedilen aykırı değerler temizlendikten sonra 2 dakikalık periyotlarla alınan verilerin ortalamaları otuz dakikalık periyotlar haline getirilerek 16650 veriden oluşan daha sade bir seti haline getirilmiştir.

Şekil 3.8’ de çalışma kapsamında kurulan tüm sistemin blok şeması görülmektedir.



Şekil 3.8. Tüm sisteme ait blok şeması

3.2.3. Veri setlerinin tanıtılması ve ön işleme

Tez kapsamında iki adet veri seti ile çalışılmıştır. Bu iki veri setinin içerdiği veriler ve elde edilmiş yöntemleri farklılık göstermektedir.

Nesnelerin interneti veri seti

İlk veri seti, PV paneller ile irtibatlandırılan NİT tasarımı ile elde edilmiştir. Bu veri setinde sıcaklık, nem, ışık parlaklığı ve akım verileri bulunmaktadır. Bu veri setinde ilk

anlamlı veri 04.05.2023 günü, 12.01 de elde edilmiştir. Takip eden 11 aylık süreçte iki dakikalık periyotlarda elde edilen 91949 veriden oluşan bir veri seti oluşturulmuştur.

Anlık değişimlerini kaçırmamak ve daha sağlıklı bir veri elde edebilmek amacıyla seçilen iki dakikalık periyotlar, Phyton programlama diliyle 30 dakikalık ortalama veriler haline dönüştürülmüştür. Çizelge 3.2’da veri setine ait otuz dakikalık bir kesit görülmektedir. Burada 13:00 – 13:30 arasında toplanan sıcaklık verilerinin ortalaması 39,58 °C iken 13:30 da alınan sıcaklık verisi (çizelge 3.3) 41 °C, yine aynı saat dilimine ait ortalama akım değeri 6,25 A iken 10:30 ‘da alınan akım verisi 10,63 A olarak görülmektedir.

Çizelge 3.2. Veri setine ait otuz dakikalık dilim ortalamalar

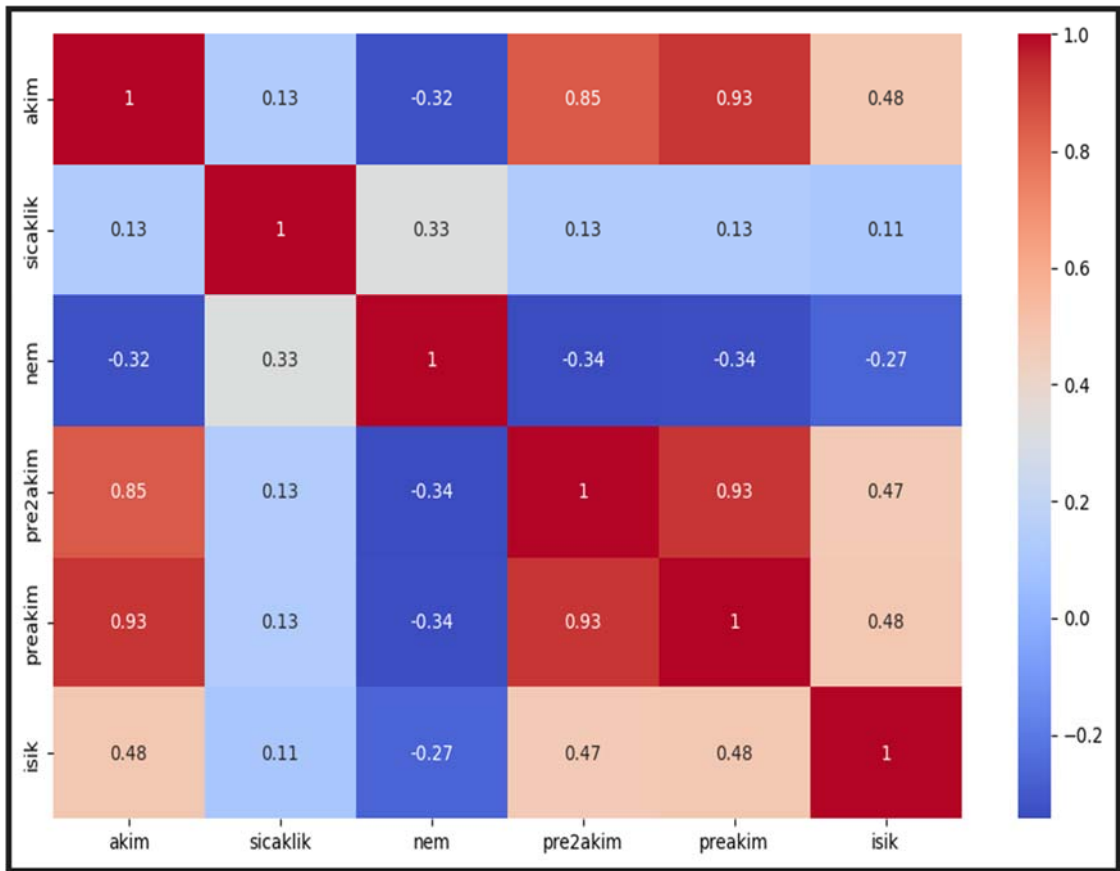
Tarih	Saat	Sıcaklık	Akım	Nem	Işık Şiddeti	Gündüz
2023-07-02	13:00:51	41	10,63	11	1022	1
2023-07-02	13:02:57	40	7,15	11	1020	1
2023-07-02	13:05:02	41	5,07	11	1015	1
2023-07-02	13:07:08	41	5,22	11	1022	1
2023-07-02	13:09:14	39	3,89	12	1014	1
2023-07-02	13:11:19	39	3,67	12	1014	1
2023-07-02	13:13:24	39	4,63	12	1015	1
2023-07-02	13:15:30	39	4,18	12	1014	1
2023-07-02	13:23:51	39	3,44	12	1013	1
2023-07-02	13:25:57	38	8,41	12	1016	1
2023-07-02	13:28:02	40	9,52	11	1017	1
13:30 Ortalama Değer		39,58	6,2583	11,583	1015,75	1

Çizelge 3.3. Otuz dakikalık anlık değerler

Tarih	Saat	Sıcaklık	Akım	Nem	Işık Şiddeti	Gündüz
2023-07-02	13:30:07	41	10,63	11	1022	1

Aynı zamanda çizelge incelendiğinde iki dakikalık periyotlar halinde veri toplama isteğiyle oluşturulan sistemikte aksamalar yaşandığı görülmektedir. Veri gönderme aralığı otuz dakika içerisinde 8 dakikaya kadar uzadığı görülmektedir. Bu durum veri setinin tamamında görülmektedir. Sebep olarak internet bağlantı hattında oluşan kopmalar ve sensörlerden kaynaklı geçici arızalar olduğu düşünülmektedir.

Sensörlerden alınan verilerin toplanmasındaki zaman farkları ve ortalama değerler ile anlık değerlerin farklı olması göz önüne alındığında, zaman dilimi ortalaması şeklinde elde edilen veri setinin makine öğrenmesinde daha anlamlı sonuçlar ortaya çıkaracağı düşünülmektedir. Bu durum sonuç bölümünde incelenecektir.



Şekil 3.9. Isı haritası 1

Şekil 3.9’da Python’daki Seaborn kütüphanesi ile oluşturulmuş ısı haritası görülmektedir. Isı haritaları, veri setlerindeki ilişkileri göstermek için bir renk skalası oluşturup bu renkleri matrislerle sunmaktadır. Seaborn kütüphanesi genellikle Pearson korelasyon katsayısı üzerinden ısı haritası oluşturur. Bu yöntem, değişkenler arasındaki lineer ilişkiyi ölçmek için kullanılan bir metriktir. Skalada kırmızı renk tonları yüksek pozitif ilişkiyi, açık kırmızıdan maviye doğru ilerleyen tonlar ise düşük korelasyonu ifade eder. Ayrıca,

korelasyon katsayıları genellikle -1 ile 1 arasında değer alır ve ilişkinin gücünü (büyüklüğünü) ile yönünü (pozitif veya negatif) gösterir.

Isı haritasında akım değeri ile pozitif yönde ilişki gösteren değerler ısı ve sıcaklık verileridir. Nem değeri negatif yönde bir korelasyon göstermektedir.

Bu işlemlerin dışında, tahmin performansının iyileştirilmesi adına, veri setine “önceki akım” ve “iki önceki akım” sütunları eklenmiştir. Bu işlemlerin sonunda resim 3.7’ de elde edilen veri setine ait 5 satır görselde görülmektedir.

```
#veri setindeki ilk 5 veri satırını
df.head()
```

✓ 0.0s

	data	date	saat	pre2akim	preakim	akim	isik	nem	sicaklik	gunduz
0	2023-06-20	14:00	3.665	4.150	5.060	255.0	21.0	21.0	1	
1	2023-06-20	14:30	4.150	5.060	2.272	1014.0	21.0	21.0	1	
2	2023-06-20	15:00	5.060	2.272	1.705	1008.0	22.0	21.0	1	
3	2023-06-20	15:30	2.272	1.705	0.740	1008.0	22.0	22.0	1	
4	2023-06-20	16:00	1.705	0.740	3.000	1014.0	22.0	22.0	1	

Resim 3.7. NİT veri seti ilk beş satırı

Hava durumu sunucusu veri seti

Bu veri seti, dünya genelinde hava durumu tahmin ve meteorolojik veriler sunan <https://www.accuweather.com/> web sitesinden elde edilmiştir. Otuz dakikalık periyotlarda sıcaklık, ultraviyole, rüzgâr, rüzgâr yönü, nem, basınç, bulutluluk, gece-gündüz olmak üzere sekiz farklı veri içermektedir. NİT ile elde edilen veri seti ile aynı zaman aralıklarına sahiptir ve buradan elde edilen akım değerleri, “akım”- “önceki akım”- “iki önceki akım” şeklinde eklenmiştir. Bu veri setinde "önceki akım" ve "iki önceki akım" terimleri, zaman serileri analizinde kullanılan tekniklerdir. "Önceki akım", mevcut zaman noktasından bir önceki zaman diliminde elde edilen akım değerini ifade ederken, “iki önceki akım” ise iki zaman dilimi öncesinde elde edilen akım değerini ifade eder.

Bu terimler, zaman serilerindeki geçmiş değerleri modelleme ve tahmin etme amacıyla sıklıkla kullanılır. Resim 3.8’ de veri setine ait ilk beş satır görülmektedir.

```
#veri setindeki ilk 5 veri satırı
df.head()
```

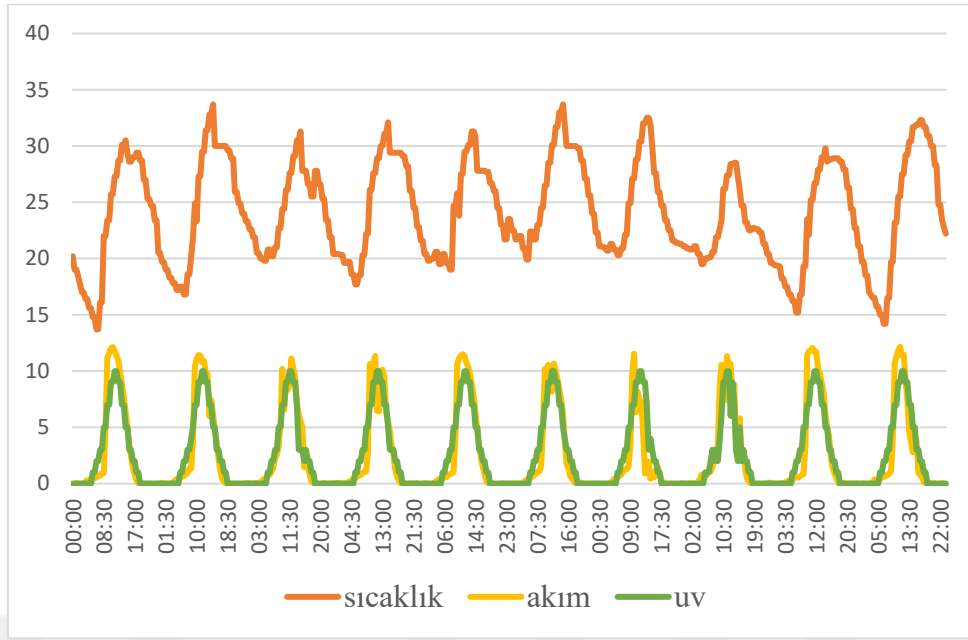
✓ 0.0s

	id	date	saat	sicaklik	uv	ruzgar	ruzgaryonu	nem	basinc	bulut	pre2akim	preakim	akim	gunduz
2	4206.0	2023-05-02	07:30	8.9	3.0	12.6	D	67.0	1022.0	2.0	0.412857	0.581905	0.685238	1.0
3	4207.0	2023-05-02	08:00	8.9	3.0	12.6	D	67.0	1022.0	2.0	0.581905	0.685238	0.818750	1.0
4	4208.0	2023-05-02	08:30	11.0	4.0	12.6	D	58.0	1022.0	2.0	0.685238	0.818750	3.210526	1.0
5	4209.0	2023-05-02	09:00	11.0	4.0	15.2	D	63.0	1022.0	0.0	0.818750	3.210526	12.540000	1.0
6	4210.0	2023-05-02	09:30	14.2	6.0	15.2	D	51.0	1022.0	0.0	3.210526	12.540000	13.136111	1.0

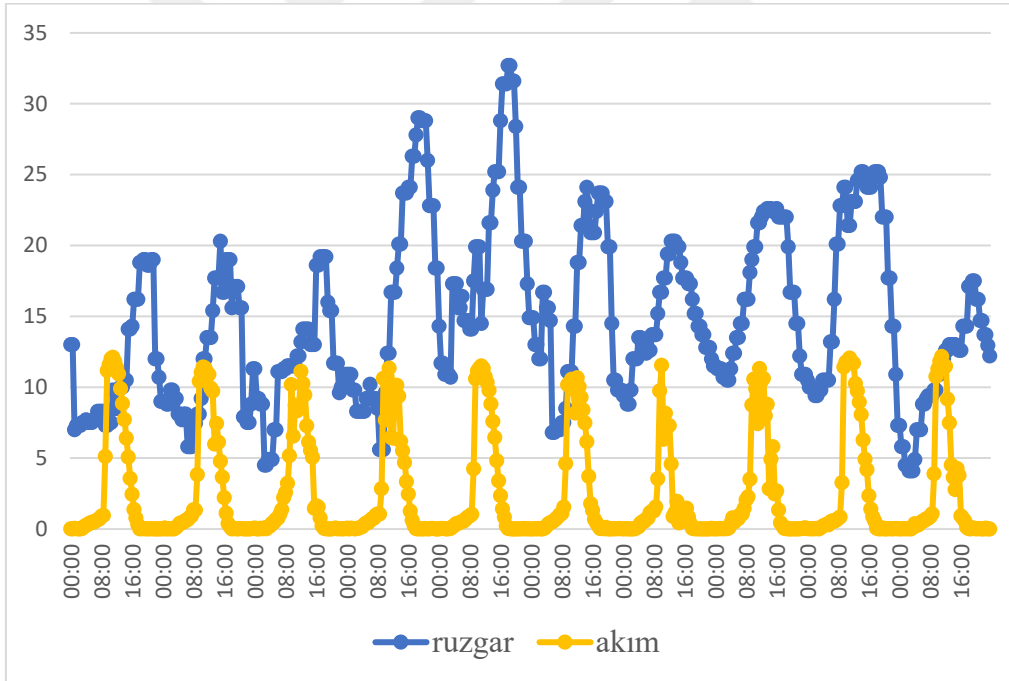
Resim 3.8. Accuweather veri setine ait ilk beş satır

Veri seti toplamda, otuz dakikalık periyotlarla alınmış 14427 veri satırından oluşmaktadır. Veri setinin bazı bölümlerinde, internet sunucusu ile ilgili olduğu düşünülen otuz dakikalık periyotlarında bozulmalar görülmüştür. Bu bozulmaların bulunduğu veri satırları tahmin işlemine katılmamıştır.

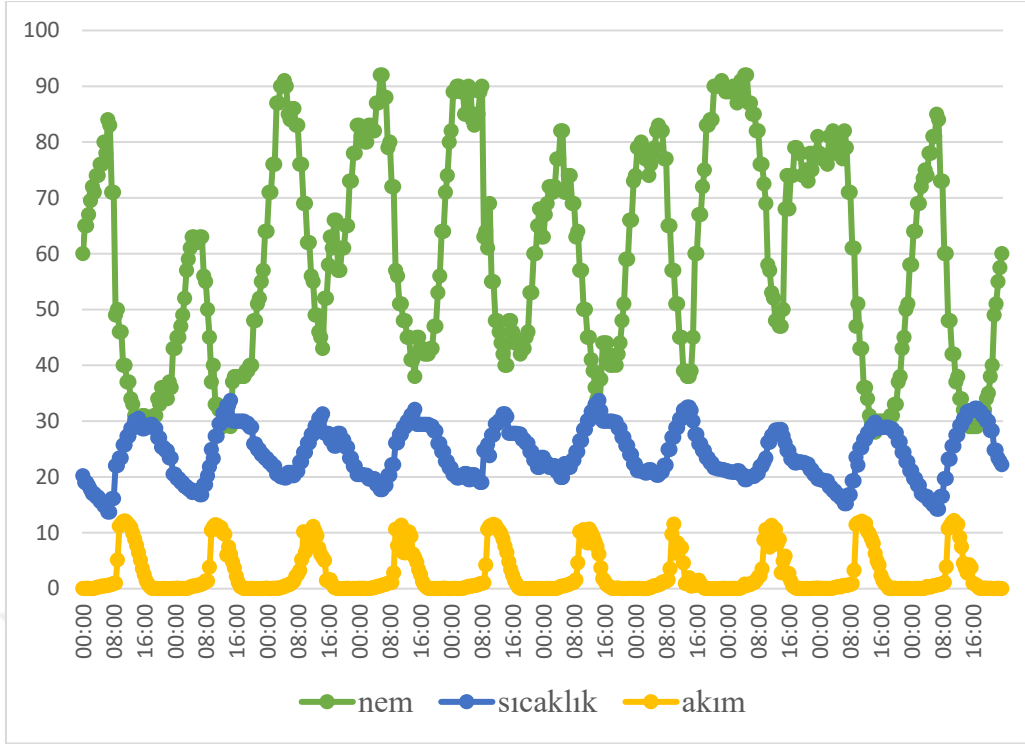
Şekil 3.10 – 3.11 ve 3.12’de veri setinde bulunan bazı verilerin akım değeri ve diğer veriler ile karşılaştırılması görülmektedir. Buna göre akım değeri ile en fazla tutarlılık gösteren veri değeri ultraviyole olarak görülmektedir. Akım değeriyle paralel olarak devam eden oransal olarak ultraviyolede az olsa da sıcaklık verisi olmaktadır. Rüzgâr değeri ile akım değerinin belirli bir paralelliği görülmesi de nem değerleri sıcaklık ile ters orantılı değiştiği görülmekte ve buradan çıkarımla nem ve akım değerleri arasında ters orantılı bir ilişki bulunduğu söylenebilmektedir.



Şekil 3.10. Sıcaklık, akım ve UV değerlerine ait grafik

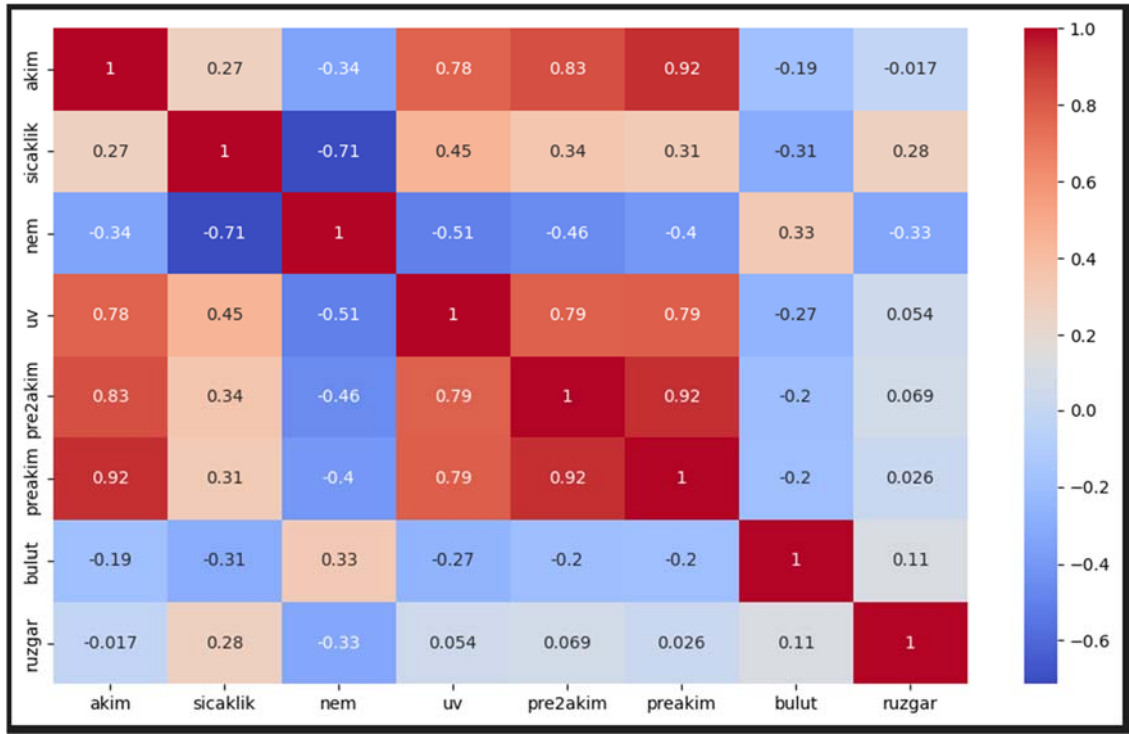


Şekil 3.11. Rüzgâr ve akım değerlerine ait grafik



Şekil 3.12. Sıcaklık, nem ve akım değerlerine ait grafik

Oluşturulan ısı haritası (şekil 3.13) yukarıdaki grafikleri destekler niteliktedir. Buna göre; akım verileriyle en yüksek ilişkiyi (önceki akım ve iki önceki akım değerleri hariç tutularak) ultraviyole değeri 0,78 değeri ile sağlamaktadır. UV değerinden sonra en yüksek ilişki ise 0,27 değeri ile sıcaklık olarak görülmektedir. Bulutluluk, nem ve rüzgâr verileriyle pozitif ilişki bulunmadığı, ayrıca en güçlü negatif ilişkinin-0,019 rüzgâr olduğu görülmektedir.



Şekil 3.13. Isı haritası 2

3.2.4. Makine öğrenmesi

Makine öğrenmesi, bilgisayar sistemlerinin belirlenmiş bir görevi yerine getirebilmek amacıyla veriler kullanarak kendi performansını iyileştirebilen bir yapay zekâ alt dalı olarak tanımlanmaktadır. Makine öğrenimi programlanmış talimatlarla değil, veriler üzerinden öğrenme yeteneğine dayanmaktadır.

Bu çalışmada iki farklı veri seti üzerinden makine öğrenimi uygulaması yapılmıştır. Tezin bir önceki bölümünde anlatıldığı gibi bu veri setlerinin içerdiği veriler aynı zaman dilimine ait ancak elde edilmiş biçimleri ve veri cinsleri farklılıklar göstermektedir. Bu bakımdan iki veri seti aynı makine öğrenimi algoritmalarına uygulanarak performansları paylaşılmıştır.

Hiper parametre optimizasyonu

Makine öğrenimi algoritmaları olarak Lineer Regresyon, RFR Regresyon, LSTM ve XGBoost algoritmaları seçilmiştir. İlk uygulamanın ardından makine öğrenimi model performanslarının artırılması ve en etkili kombinasyonu belirlemek amacıyla hiper parametre ayarları ile uygulama tekrarlanmıştır ve sonuçlar ayrı tablolar halinde sunulmuş ve karşılaştırılmıştır. Hiper parametre değerleri modelin genelleme kabiliyetini

arttırmak için kullanılmaktadır. En iyi hiper parametre değerleri farklı kombinasyonları rastgele deneme yöntemiyle ya da GridSearch CV tekniğinden yararlanarak bulunabilmektedir. Hiper parametre kombinasyonlarını sistematik şekilde deneyip en iyi değerleri bulmayı amaçlamaktadır (Bergsta ve Bengio, 2012). Kullanılan parametrelerin işlevleri çizelge 3.4, çizelge 3.5 ve çizelge 3.6’ da gösterilmiştir.

Çizelge 3.4. RFR hiper parametreleri

Hiper parametre	Görevi ve İçeriği
N_estimators	RFR modelinde kullanılacak ağaç sayısını belirlemektedir.
Max_depth	Bir ağacın maksimum derinliği bir dalın en fazla kaç seviyeye kadar dallanabileceğini göstermektedir.
Min_samples_split	Bir iç düğümün bölünmesi gereken en az örnek sayısıdır.
Min_samples_leaf	Bir yaprağın sahip olması gereken örnek sayısını belirler.
Bootstrap	Eğitim verisi üzerinden yapılacak örneklemeyi ifade etmektedir.

Çizelge 3.5. XGB hiper parametreleri

Hiper parametre	Görevi ve İçeriği
N_estimators	Oluşturulacak ağaç sayısını ifade etmektedir. Ağaç sayısının fazla olması performansı arttırmaktadır.
Learning rate	Oluşturulan her bir ağacın katkısını kontrol eden öğrenme oranıdır. Düşük öğrenme oranı istikrarlı bir öğrenme, yüksek bir öğrenme oranı daha hızlı öğrenme oluşturmaktadır.
Subsample	Eğitimin her bir aşamasında kullanılacak alt örnekleme sayısını oranını göstermektedir.
Min_child_weight	Bir iç düğümün bölünmesi için gereken en az ağırlık sayısını göstermektedir.
Colsample_bytree	Ağaç oluşturma aşamasında kullanılacak özelliklerin oranlarını belirlemekte kullanılmaktadır.
Gamma	Bir düğümün bölünmesinde düşüş sağlamak için gereken en az kaybı belirlemektedir.

Çizelge 3.6. LSTM hiper parametreleri

Hiper parametre	Görevi ve İçeriği
Units	LSTM katmanındaki nöron sayısını belirler.
Activation	Aktivasyon fonksiyonunu belirler. LSTM hücreleri içinde kullanılacak aktivasyon fonksiyonlarını ayarlar.
Batch_size	Eğitim sırasında kullanılacak minibatch boyutunu belirler. Bir minibatch, her adımda modele beslenen eğitim örneklerinin sayısını ifade etmektedir.
Learning_rate	Öğrenme oranını belirler. Modelin her adımda ağırlıklarını ne kadar değiştireceğini kontrol eder.
Epochs	Eğitim veri setinin model tarafından kaç kez işleneceğini belirler. Bir epoch, tüm eğitim veri setinin bir kez modelden geçirilmesi anlamına gelir.

Seçilen makine öğrenimi algoritmalarına uygulanacak hiper parametreler, modelin performansını doğrudan etkilemektedir. Hiper parametrelerin doğru bir şekilde belirlenmesi, modelin tahmin doğruluğunu ve genelleme yeteneğini artırmada kritik bir rol oynar. Hiper parametre değeri seçiminde, her bir parametrenin ne işe yaradığını ve model üzerindeki etkisini anlamak, en doğru seçimi belirlemede yardımcı olmaktadır. GridSearch CV yöntemi, bu parametrelerin optimize edilmesinde yaygın olarak kullanılan bir tekniktir ve bu bileşenlerin açıklamaları Çizelge 3.7'de sunulmuştur.

Çizelge 3.7. GridSearch CV bileşenleri

GridSearch CV	Açıklama
Model	GridSearch CV' in optimize edeceği model seçilmektedir
Param grid	Hiper parametre ızgarasındaki denenecek kombinasyonları içermektedir.
Cv	Çapraz doğrulama kat sayısı
Scoring	Modelin performansını belirlemek için seçilen metrik bilgisi

Farklı veri kümeleri için yapılacak makine öğrenimi uygulamaları için veri setinde bulunan özellikler çizelge 3.8' de semboller halinde sunulmuştur.

Çizelge 3.8. Girdilere ait sembol çizelgesi

Sütun Başlıkları	Semboller	Sütun Başlıkları	Semboller
Gündüz	S (sun)	Rüzgâr	W (wind)
Tarih	D (date)	Akım	I _n
Saat	T (time)	Preakim	I _{n-1}
Sıcaklık	H (heat)	Pre2Akım	I _{n-2}
Nem	Hu(humidity)	Ultraviyole	UV (ultraviyole)
Bulut	C (cloud)	Veri Sayısı	N(number)
Işık	L(light)		

4. BULGULAR VE TARTIŞMA

4.1. Hava Durumu Sunucusu Veri Seti ile Makine Öğrenimi

Bu veri seti, dünya genelinde güncel hava durumu verileri ve tahminleri paylaşan bir meteoroloji sayfasından elde edilmiştir. Yaklaşık on bir aylık bir zaman dilimine ait otuz dakikalık periyotlarla toplanmış sıcaklık, nem, ultraviyole, rüzgâr, rüzgâr yönü, basınç, gece-gündüz ve bulut verilerini içermektedir. Ayrıca tahmin performansını arttırabilmek amacıyla pv panellerden elde edilen akım değerleri, ‘önceki akım’ ve ‘iki önceki akım’ sütunları olacak şekilde veri setine eklenmiştir. Veri seti son haliyle 13 sütun ve 19618 satırdan oluşmaktadır.

Veri kümesi üç ayrı şekilde kullanılarak eğitim uygulanmış ve her eğitimden sonra hiper parametreler ile tekrar eğitim uygulanarak algoritmaların performansları tablolar halinde sunulmuştur. Ayrıca GridSearch CV tarafından seçilmiş en iyi hiper parametre değerleri ve veri türlerinin algoritmalarda önem sıraları ile birlikte sunulmuştur.

4.1.1. Tüm veriler ile makine öğrenimi

Bu uygulamada veri setinde bulunan tüm atmosferik veriler ve PV panellerden elde edilen akım değerleri birlikte kullanılmıştır.

Çizelge 4.1. Tüm veri seti ile elde edilen performans bulguları

GİRDİLER				
“D”- “T”- “H” – “Hu” – “C”- “W” – “UV” – “I _n ” – “I _{n-1} ” – “I _{n-2} ”				
	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.886	0.854	0.879	0.876
RMSE	1.018	1.152	1.051	1.061
MAE	0.440	0.560	0.454	0.510

Çizelge 4.1’de uygulanan ilk eğitimden sonra tahmin sonuçlarında RFR algoritmasının tüm metriklerde en başarılı sonucu verdiği, XGBoost algoritmasının ise ilk algoritmaya yakın performans sergilediği görülmektedir. Aynı zamanda XGBoost ve LSTM algoritmaları özellikle R^2 metrik değerlendirmesinde çok yakın sonuçlar göstermiştir.

İlk uygulanan eğitimin ardından çizelge 4.2, çizelge 4.3 ve çizelge 4.4’ te sırasıyla RFR, XGBoost ve LSTM algoritmaları için GridSearch tarafından seçilmiş en iyi hiper parametre değerleri ile tekrar eğitim uygulanmıştır. GridSearch tarafından seçilen en iyi değerler aşağıdaki tablolarda sunulmuştur. Ayrıca bunlara ait performans sonuçları da çizelge 4.5’ te gösterilmiştir.

Çizelge 4.2. RFR algoritmasına ait hiper parametre değerleri

RFR Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[250, 500]	250
Max_depth	[10, 50, 100]	100
Min_samples_split	[2, 4, 8]	2
Min_samples_leaf	[1, 2, 4]	4
Bootstrap	[True, False]	True

Çizelge 4.3. XGB algoritması hiper parametre değerleri

XGBoost Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[100, 500]	500
Learning rate	[0.01, 0.1, 0.2, 0.3]	0,01
Subsample	[0.5,0.8,1]	0,5
Min_child_weight	[5,10,20,40]	10
Colsample_bytree	[0.5,0.8,1]	0,8
Gamma	[0,0.1,1,10]	0

Çizelge 4.4. LSTM algoritması hiper parametre değerleri

LSTM		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
Units	[64, 128, 256]	128
Activation	[relu, tanh]	tanh
Batch_size	[16, 32, 64]	16
Learning_rate	[0.1, 0.01, 0.001]	0.001
Epochs	[50, 100, 200]	200

GridSearch tarafından seçilmiş en uygun Hiper parametreler ile tekrarlanan uygulamanın sonuçları çizelge 4.5’ te sunulmuştur.

Çizelge 4.5. Hiper parametreler ile elde edilen performans bulguları

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R ²	0.886	-	0.890	0.884
RMSE	1.016	-	1.000	1.028
MAE	0.433	-	0.441	0.437

Hiper parametre optimizasyonu sonrasında XGBoost algoritmasının RMSE ve R² metrikleriyle değerlendirmede diğer algoritmalara göre daha başarılı olduğu görülmektedir. RFR ve LSTM algoritmaları tüm metrik değerlendirmelerinde birbirlerine çok yakın performanslar sergilemişlerdir. RFR algoritması ise MAE metriği değerlendirmesinde diğer algoritmaların üzerinde yer almıştır.

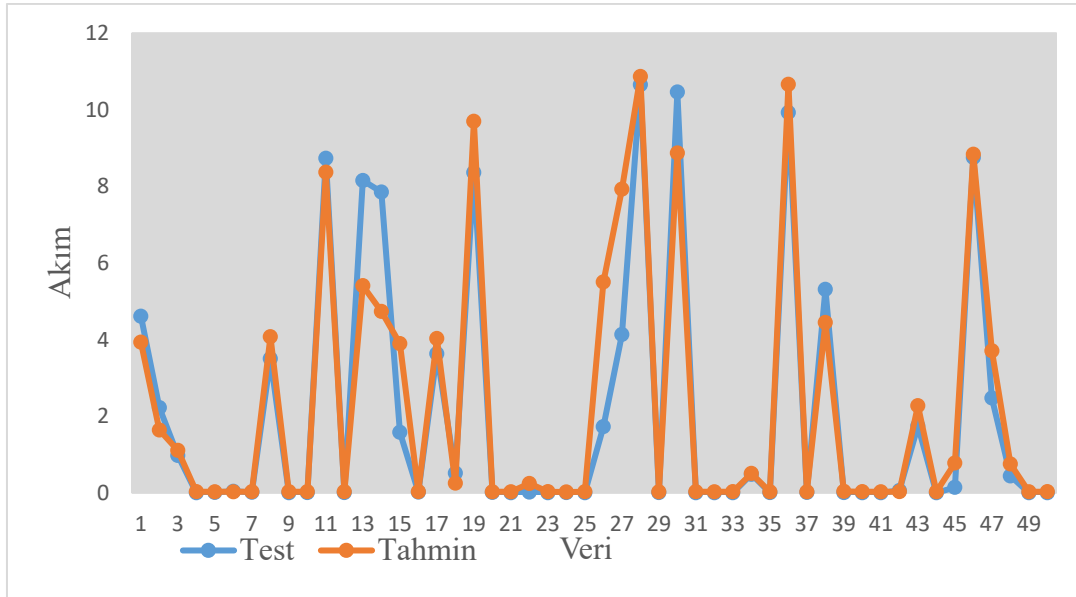
Veri seti içerisinde bulunan her veri türü, algoritma performansını pozitif ya da negatif yönde etkilemektedir. Hangi özelliğin performansı ne kadar etki ettiğinin bilinmesi performans geliştirmede işe yarar bir şekilde kullanılabilir. Buna göre, Çizelge 4.6’ da eğitime katılan verilerin tahmin performansını etkilemedeki güçleri oransal olarak sıralanmıştır. Bu model eğitiminde saat ve tarih verileri dışında sekiz adet özellik kullanılmıştır.

Çizelge 4.6. Özelliklerin algoritmalara göre önem sıralamaları

	Rastgele Orman	Lineer Regresyon	XGBoost
Veri Setindeki Girdilerin Önem Sıralaması	1. preakim: %87.87	1. preakim: %83.16	1. preakim: %80.71
	2. pre2akim: %3.72	2. uv: %19.40	2. uv: %3.88
	3. uv: %1.67	3. gunduz: %10.14	3. gunduz: %3.47
	4. sicaklik: %1.64	4. nem: %0.71	4. pre2akim: %2.98
	5. nem: %1.56	5. bulut: %-0.05	5. bulut: %2.24
	6. ruzgar: %1.36	6. basinc: %-0.21	6. nem: %1.93
	7. bulut: %1.08	7. sicaklik: %-0.23	7. ruzgar: %1.75
	8. basinc: %1.04	8. ruzgar: %-0.92	8. basinc: %1.70
	9. gunduz: %0.06	9. pre2akim: %-12.01	9. sicaklik: %1.34

Çizelge 4.6' ya göre tahmin işleminde tüm algoritmalara en yüksek verimi sağlayan veri “bir önceki akım” değeri olmuştur. Oransal olarak tamamında farklı olsa da tüm algoritmalarından %80 üzerinde bir skor almıştır. Tüm verilerin içerisinde ultraviyole değeri de lineer regresyon ve XGBoost algoritmalarında ikinci sırada yer almıştır.

Tahmin çalışması sonucunda elde edilen veriler .csv formatında alınmıştır. Elde edilen veri setine ait görsel karşılaştırma şekil 4.1' de görülmektedir. Grafik oluşturma sürecinde anlamlı bir görüntü elde edebilmek amacıyla veri setinin tamamı kullanılmamıştır.



Şekil 4.1. Tüm veriler ile elde edilmiş tahmin verileri ve test verileri

4.1.2. Atmosferik verileri kullanarak makine öğrenimi

Bu uygulamada PV panellerden alınan “önceki akım” ve “iki önceki akım” verileri tahmin işlemine katılmamıştır. <https://www.accuweather.com/> web sitesinden alınan on bir aylık atmosferik veriler kullanılarak uygulanan makine öğrenimine ait metrik değerlendirmeleri çizelge 4.7’ de sunulmuştur.

Çizelge 4.7. Atmosferik veriler ile elde edilen bulgular

GİRDİLER				
D” - “T” - “H” – “Hu” – “C”- “W” – “UV				
	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.793	0.623	0.772	0.690
RMSE	1.374	1.854	1.441	1.682
MAE	0.685	1.063	0.775	0.886

Yalnızca atmosferik veriler kullanılarak gerçekleştirilen dört makine öğrenimi algoritma sonuçlarında üç metrikte de en başarılı sonucu veren algoritma RFR Regresyonu olarak görülmektedir. RFR algoritmasını yine tüm metrik değerlendirmelerinde ikinci sırada bulunan XGBoost algoritması takip etmiştir. Tüm metrik değerlendirmelerinde en düşük performansı gösteren ise LR algoritması olmuştur.

Çizelge 4.8. RFR Algoritması Hiper Parametre Değerleri

RFR Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[250, 500]	500
Max_depth	[10, 50, 100]	50
Min_samples_split	[2, 4, 8]	2
Min_samples_leaf	[1, 2, 4]	1
Bootstrap	[True, False]	True

Çizelge 4. 9. XGBoost Algoritması Hiper Parametre Değerleri

XGBoost Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[100, 500]	500
Learning rate	[0.01, 0.1, 0.2, 0.3]	0,1
Subsample	[0.5,0.8,1]	0,5
Min_child_weight	[5,10,20,40]	10
Colsample_bytree	[0.5,0.8,1]	0,8
Gamma	[0,0.1,1,10]	0

Çizelge 4.10. LSTM Algoritması Hiper Parametre Değerleri

LSTM		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
Units	[64, 128, 256]	256
Activation	[relu, tanh]	tanh
Batch_size	[16, 32, 64]	32
Learning_rate	[0.1, 0.01, 0.001]	0.001
Epochs	[50, 100, 200]	200

İlk uygulanan eğitimin ardından çizelge 4.8 RFR, çizelge 4.9 XGBoost ve çizelge 4.10 LSTM algoritmaları için GridSearch tarafından seçilmiş en iyi hiper parametre değerleri ile tekrar eğitim uygulanmıştır. Bu uygulamanın sonuçları çizelge 4.11’ de sunulmuştur.

Çizelge 4.11. Hiper parametreler ile Elde Edilen Performans Bulguları

	Rastgele Orman	Lineer Regr.	XGBoost Regr.	LSTM
R²	0.795	-	0.783	0.712
RMSE	1.367	-	1.406	1.621
MAE	0.685	-	0.774	0.840

Yalnızca atmosfer verileri kullanılarak, hiper parametre değerleri ile yapılan makine öğrenimi uygulaması sonucunda algoritma tahmin başarılarında artışın sınırlı kaldığı görülmektedir. Örnek olarak, RFR algoritmasında R^2 metrik değerlendirmesi ilk uygulamada 0,79316 iken hiper parametre ayarlarından sonra yapılan uygulamada 0,79536 olarak görülmektedir.

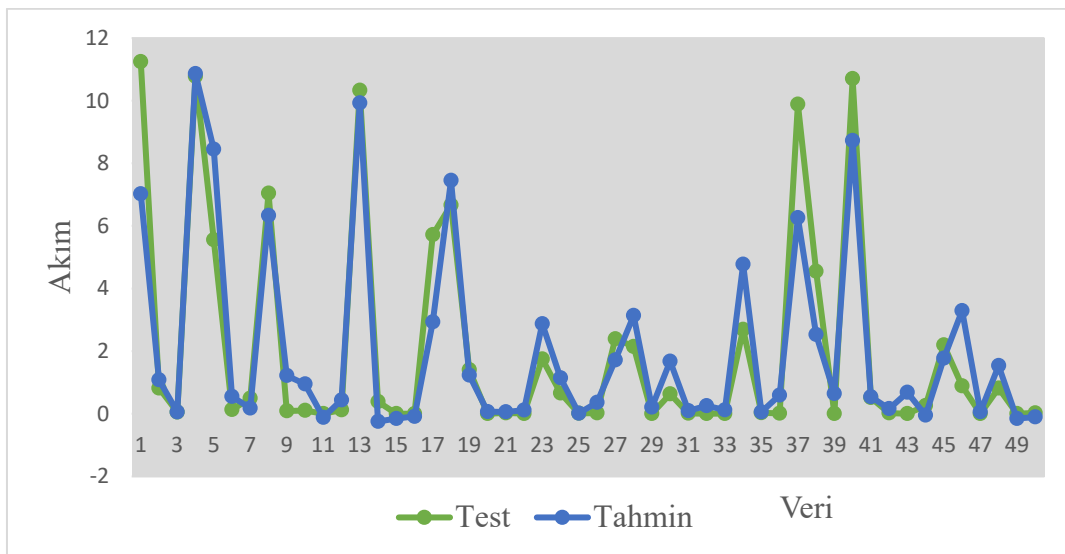
Çizelge 4.12’ de eğitime katılan verilerin tahmin performansına etkileri oransal olarak sıralanmıştır.

Çizelge 4.12. Eğitime Katılan Verilerin Performansa Etkileri

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon
Veri Setindeki Girdilerin Önem Sıralaması	1. uv: %63.12	1. uv: %69.09	1. uv: %71.77
	2. sicaklik: %10.59	2. gündüz: %33.09	2. bulut: %5.95
	3. ruzgar: %7.56	3. nem: %0.29	3. basinc: %5.38
	4. nem: %7.14	4. basinc: %0.24	4. sicaklik: %5.20
	5. basinc: %6.12	5. bulut: %0.01	5. ruzgar: %4.36
	6. bulut: %5.18	6. sicaklik: %-1.32	6. nem: %4.23
	7. gündüz: %0.30	7. ruzgar: %-1.40	7. gündüz: %3.12

Çizelge 4.12’ da yalnızca atmosfer verileri ile makine öğrenimi uygulamasında panellerden üretilen enerjinin ultraviyole değeri ile yakından ilgili olduğu görülmektedir.

Şekil 4.2’ de elde edilen tahmin verileri ile test verilerinin karşılaştırılması gösterilmiştir.



Şekil 4.2. Atmosferik Veriler ile Elde Edilmiş Tahmin Verileri ve Test Verileri

4.1.3. Akım verileri ile makine öğrenimi

Bu uygulamada atmosferik veriler kullanılmadan yalnızca önceki akım ve iki önceki akım değerleri üzerinden tahmin işlemleri yapılmıştır. Bu adımda asıl amaçlanan nokta akım verilerine ait performans değerleri ile atmosfer verilerinin performansını karşılaştırmaktır.

Çizelge 4.13. Akım Verileri ile Elde Edilen Bulgular

GİRDİLER				
“I _{n-1} ” – “I _{n-2} ”				
	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.844	0.840	0.842	0.876
RMSE	1.192	1.207	1.200	1.045
MAE	0.527	0.559	0.512	0.512

Atmosferik verilerin bulunmadığı, yalnızca akım değerleri veri seti ile yapılan uygulamada (çizelge 4.13), kullanılan tüm algoritmaların birbirine yakın performans sergilediği, yakın performans sonuçlarına rağmen LTSM algoritmasının tüm metrik değerlendirmelerinde en başarılı sonuçları verdiği görülmektedir. Aynı zamanda MAE metrik değerlendirmesinde de XGBoost ve LSTM algoritmaları eş performans sergilemişlerdir.

Çizelge 4.14. RFR Algoritması Hiper Parametre Değerleri

RFR Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[250, 500]	500
Max_depth	[10, 50, 100]	10
Min_samples_split	[2, 4, 8]	8
Min_samples_leaf	[1, 2, 4]	4
Bootstrap	[True, False]	True

Çizelge 4.15.XGB algoritması hiper parametre değerleri

XGBoost Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[100, 500]	500
Learning rate	[0.01, 0.1, 0.2, 0.3]	0,01
Subsample	[0.5,0.8,1]	0,5
Min_child_weight	[5,10,20,40]	10
Colsample_bytree	[0.5,0.8,1]	1
Gamma	[0,0.1,1,10]	10

Çizelge 4.16. LSTM algoritması hiper parametre değerleri

LSTM		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
Units	[64, 128, 256]	256
Activation	[relu, tanh]	tanh
Batch_size	[16, 32, 64]	64
Learning_rate	[0.1, 0.01, 0.001]	0.01
Epochs	[50, 100, 200]	200

İlk uygulanan eğitimin ardından, farklı algoritmalar üzerinde GridSearch yöntemi ile en iyi hiper parametre değerleri belirlenmiştir. Bu kapsamda, Rastgele Orman algoritması için seçilen en iyi hiper parametre değerleri ile yapılan eğitim sonuçları Çizelge 4.14'te, XGBoost algoritması için belirlenen en iyi hiper parametrelerle gerçekleştirilen eğitim sonuçları Çizelge 4.15'te ve LSTM algoritması için seçilmiş en iyi hiper parametrelerle yapılan eğitim sonuçları ise Çizelge 4.16'da sunulmuştur. Bu algoritmaların, GridSearch tarafından seçilen en iyi hiper parametre değerleri ile yeniden eğitilmesi sonucu elde edilen performans değerlendirmeleri ve sonuçlar ise Çizelge 4.17'de ayrıntılı olarak gösterilmiştir. Bu şekilde, farklı algoritmaların ve hiper parametrelerin model performansları karşılaştırılmıştır.

Çizelge 4.17.Hiper parametreler ile elde edilen performans bulguları

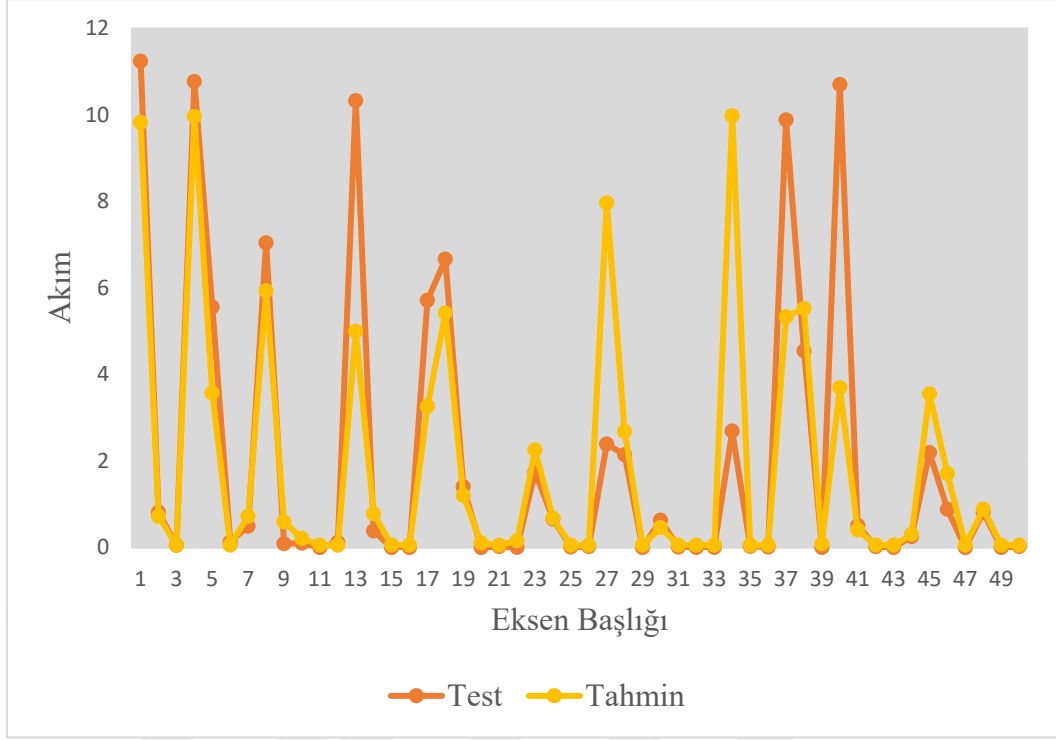
	Rastgele Orman	Lineer Reg.	XGBoost Regr.	LSTM
R²	0.858	-	0.863	0.859
RMSE	1.135	-	1.115	1.134
MAE	0.489	-	0.492	0.586

Yalnızca akım verileri kullanılarak, hiper parametre değerleri ile yapılan makine öğrenimi uygulaması sonucunda ilk uygulamanın aksine XGBoost algoritması R² ve RMSE metriklerinde daha başarılı sonuçlar elde etmiştir. RFR Algoritması ise MAE metriğinde daha iyi bir performans göstermiştir. Çizelge 4.18’ da veri setinde bulunan verilerin tahmin performansına etkileri oransal olarak görülmektedir.

Çizelge 4.18. Eğitime Katılan Verilerin Performansa Etkileri

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon
Veri Setindeki Girdilerin Önem Sıralaması			
1. preakim: %91.80	1. preakim: %111.87	1. preakim: %95.04	
2. pre2akim: %8.20	2. pre2akim: %-11.87	2. pre2akim: %4.96	

Çizelge 4.18’e bakıldığında, tahmin işleminde bir önceki akım değerinin tahmin performansına büyük ölçüde olumlu bir etki yaptığı açıkça görülmektedir. Bu durum, modelin bir önceki akım değerini dikkate alarak daha doğru tahminler yapabildiğini göstermektedir. Öte yandan, iki önceki akım değerinin ise özellikle LR (Doğrusal Regresyon) algoritmasında negatif bir etki yarattığı tespit edilmiştir. Bu durum, iki önceki akım değerinin modelin tahmin performansını olumsuz yönde etkilediğini işaret etmektedir. Şekil 4.3’te ise elde edilen tahmin verileri ile test verilerinin karşılaştırılması detaylı bir şekilde gösterilmiştir. Bu grafik sayesinde, modelin tahmin yeteneği ve gerçek verilerle ne kadar uyumlu olduğu görsel olarak değerlendirilebilmektedir. Bu tür karşılaştırmalar, modelin doğruluğunu ve güvenilirliğini değerlendirmek için önemli bir araç olarak görülmektedir.



Şekil 4.3. Akım verileri ile elde edilmiş tahmin ve test verileri grafiği

4.2. NİT ile Veri Seti Makine Öğrenimi

Bu çalışma kapsamında kullanılan veri seti, kurulmuş olan güneş enerjisi üretim santralinden toplanan verilerden oluşturulmuştur. Enerji santraline uygun sensörlerin ve NİT tabanlı sistemin eklenmesiyle, santralden sıcaklık, nem, ışık şiddeti ve akım değerleri gibi çeşitli veriler toplanmıştır. Bu veriler, 11 aylık bir süre boyunca düzenli olarak kaydedilmiş ve ThingSpeak platformunda depolanmıştır. Verilerin toplanma sıklığı her iki dakikada bir olacak şekilde ayarlanmış ve bu ham veriler daha sonra ön işlemden geçirilmiştir. Bu ön işleme sürecinde, iki dakikalık aralıklarla toplanan veriler, otuz dakikalık aralıklarda ortalama değerlere sahip olacak şekilde yeniden düzenlenmiştir. Sonuç olarak, elde edilen veri seti 15310 satır ve 7 sütundan oluşmaktadır. Bu geniş veri tabanı, analiz ve modelleme çalışmalarında kullanılmak üzere hazırlanmış ve sunulmuştur. Bu veri setinin kapsamlı yapısı, güneş enerjisi üretim santralinin performansının detaylı bir şekilde incelenmesine olanak sağlamaktadır.

4.2.1 Tüm veri seti ile makine öğrenimi

Bu uygulamada NİT kullanılarak elde edilen veri seti ile makine öğrenimi modeli inşa edilmiştir. Veri seti içerisinde hiçbir veri türü çıkarılmadan, tüm veriler kullanılarak

uygulanmıştır. Sıcaklık, nem, ışık parlaklığı verilerinin yanında önceki akım ve iki önceki akım değerleri yer almaktadır.

Kullanılan algoritmaların ve değerlendirme metriklerinin performans sonuçları, Çizelge 4.19'da ayrıntılı bir şekilde sunulmuştur. Her bir algoritmanın tahmin yeteneği, belirlenen metrikler ile değerlendirilmiş ve bu metrikler arasında doğruluk, hata oranı, ortalama mutlak hata, ortalama karesel hata ve kök ortalama karesel hata gibi ölçütler yer almıştır.

Çizelge 4.19. Nesnelerin interneti verileri ile elde edilen bulgular

GİRDİLER				
“I _{n-1} ” – “I _{n-2} ” – “H” – “Hu” – “L”				
	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.927	0.871	0.918	0.889
RMSE	0.815	1.081	0.863	0.991
MAE	0.336	0.527	0.356	0.466

İlk makine öğrenimi uygulama sonucunda RFR algoritmasının tüm metriklerde en iyi performans gösterdiği görülmektedir. XGBoost algoritmasının sunduğu performans ise özellikle R² ve MAE metriklerinde RFR algoritmasına ait performanslara çok yaklaşılmaktadır. Tüm sonuçlara bakarak, performans sıralamalarında üçüncü sırada LSTM algoritmasının son sırada ise LR algoritmasının yer aldığı görülmektedir.

Çizelge 4. 20. RFR algoritması hiper parametre değerleri

RFR Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[250, 500]	250
Max_depth	[10, 50, 100]	100
Min_samples_split	[2, 4, 8]	8
Min_samples_leaf	[1, 2, 4]	1
Bootstrap	[True, False]	True

Çizelge 4.21. XGB algoritması hiper parametre değerleri

XGBoost Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[100, 500]	500
Learning rate	[0.01, 0.1, 0.2, 0.3]	0,01
Subsample	[0.5,0.8,1]	0,5
Min_child_weight	[5,10,20,40]	10
Colsample_bytree	[0.5,0.8,1]	1
Gamma	[0,0.1,1,10]	0

Çizelge 4.22. LSTM algoritması hiper parametre değerleri

LSTM		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
Units	[64, 128, 256]	256
Activation	[relu, tanh]	tanh
Batch_size	[16, 32, 64]	32
Learning_rate	[0.1, 0.01, 0.001]	0.001
Epochs	[50, 100, 200]	200

İlk uygulanan eğitim sürecinin ardından, Rastgele Orman (RFR), XGBoost ve LSTM algoritmaları için GridSearch yöntemi ile en iyi hiper parametre değerleri belirlenmiştir. Bu belirleme işlemi sonucunda, her bir algoritma için optimal hiper parametreler seçilerek, modeller tekrar eğitilmiştir. Bu kapsamda, Rastgele Orman algoritması için seçilen en iyi hiper parametre değerleri ile gerçekleştirilen yeniden eğitim sonuçları Çizelge 4.20'de sunulmuştur. Aynı şekilde, XGBoost algoritması için belirlenen en iyi hiper parametrelerle yapılan eğitim sonuçları Çizelge 4.21'de, LSTM algoritması için seçilmiş en iyi hiper parametrelerle gerçekleştirilen eğitim sonuçları ise Çizelge 4.22'de detaylı bir şekilde gösterilmiştir. Bu algoritmaların, GridSearch tarafından seçilen en uygun hiper parametre değerleri ile yeniden eğitilmesi sonucu elde edilen performans

değerlendirmeleri ve sonuçlar ise Çizelge 4.23'te sunulmuştur. Bu şekilde, farklı algoritmaların ve hiper parametrelerin model performansları karşılaştırılarak, en etkili tahmin modelinin belirlenmesi amaçlanmıştır. Bu analizler, modelleme sürecinde hangi algoritmanın daha yüksek performans gösterdiğini ve belirli hiper parametrelerin etkisini anlamak için kritik bir rol oynamaktadır.

Çizelge 4.23. Hiper parametre değerleri ile elde edilen bulgular

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.929	-	0.930	0.889
RMSE	0.809	-	0.797	0.986
MAE	0.336	-	0.342	0.444

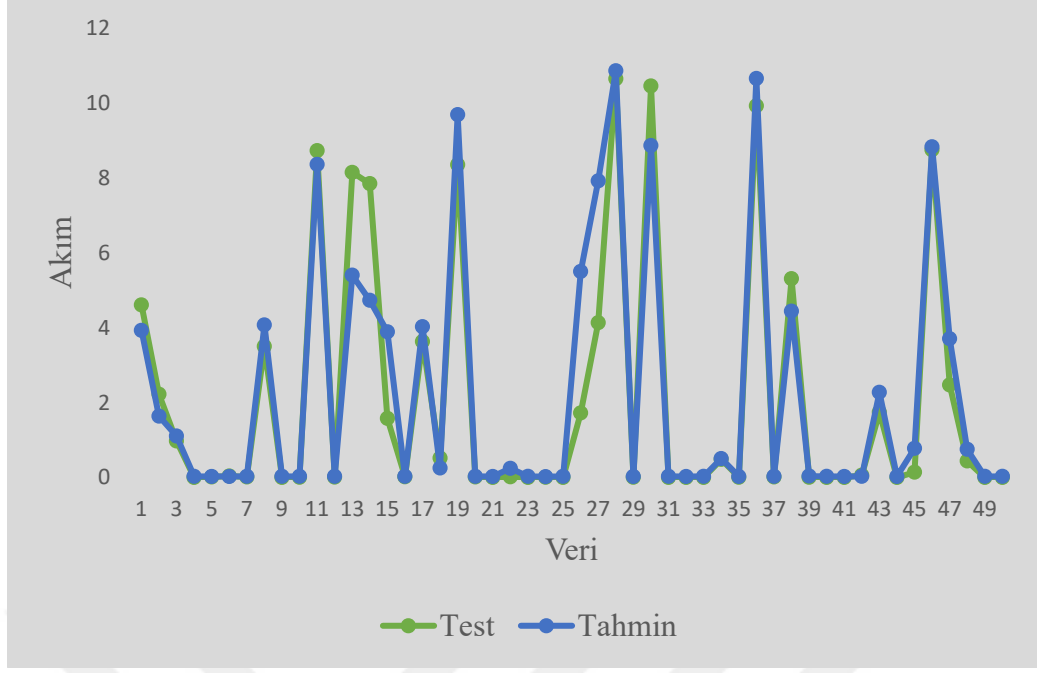
Çizelge 4.24' e göre hiper parametre değerleri ile yapılan uygulamada XGBoost uygulamasının performansı önemli ölçüde artarak R² ve RMSE metrik değerlendirmelerinde en iyi performansı göstermiştir. MAE metrik değerlendirmesinde ise RFR algoritması öne çıkmaktadır.

Çizelge 4.24.Eğitime katılan verilerin performansa etkileri

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon
Veri Setindeki Girdilerin Önem Sıralaması	1. preakim: %88.04	1. preakim: %109.61	preakim: %83.79
	2. isik: %5.47	2. gunduz: %6.90	isik: %8.36
	3. pre2akim: %3.17	3. sicaklik: %1.14	pre2akim: %2.35
	4. sicaklik: %1.65	4. isik: %0.03	sicaklik: %1.94
	5. nem: %1.62	5. nem: %-0.08	nem: %1.81
	6. gunduz: %0.04	6. pre2akim: %-17.60	gunduz: %1.75

Çizelge 4.24' de veri setinde bulunan verilerin tahmin performansına etkileri oransal olarak görülmektedir. Görüldüğü üzere, önceki akım verisi tahmin işleminde algoritmalar tarafından en çok faydalı bulunan veri olarak görülmektedir.

Elde edilen tahmin verileri ve gerçek değerler kullanılarak çizilen grafik şekil 4.4' te görülmektedir.



Şekil 4.4. Tüm veri seti ile edilmiş tahmin ve test verileri grafiği

4.2.2. NİT atmosferik verileri ile makine öğrenimi

Bu uygulamada nesnelerin interneti teknoloji ile toplanmış verilerden akım değerleri çıkarılarak makine öğrenimi uygulaması yapılmıştır. Kullanılan veri setinde sıcaklık, nem ve ışık parlaklığı özellikleri yer almaktadır. Elde edilen bulgular çizelge 4.21’ de gösterilmiştir.

Çizelge 4.25. Atmosferik veriler ile elde edilen bulgular

GİRDİLER				
“H” – “Hu” – “L”				
	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R²	0.764	0.453	0.753	0.652
RMSE	1.489	2.269	1.523	1.754
MAE	0.693	1.704	0.718	0.953

Atmosferik veriler ile yapılan ilk uygulamada RFR Algoritması tüm metrik değerlendirmelerinde en iyi performansı göstermiştir. RFR algoritmasına en yakın

değerleri ise XGBoost algoritması gösterirken, bu sıralamayı LSTM algoritması ve lineer regresyon algoritması olacak şekilde sıralanmıştır.

İlk uygulanan eğitimin ardından çizelge 4.26 RFR, çizelge 4.27 XGBoost ve çizelge 4.28 LSTM algoritmaları için GridSearch tarafından seçilmiş en iyi hiper parametre değerleri ile tekrar eğitim uygulanmıştır. Bu uygulamanın ardında hiper parametre optimizasyonu ile tekrarlanan eğitime ait bulgular çizelge 4.29’ da ve elde edilen performans sonuçlarına etki eden özelliklerin katkı oranları çizelge 4.30’ da görülmektedir.

Çizelge 4.26. RFR algoritması hiper parametre değerleri

RFR Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[250, 500]	250
Max_depth	[10, 50, 100]	10
Min_samples_split	[2, 4, 8]	8
Min_samples_leaf	[1, 2, 4]	1
Bootstrap	[True, False]	True

Çizelge 4.27. XGBoost algoritması hiper parametre değerleri

XGBoost Regresyonu		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
n_estimators	[100, 500]	500
Learning rate	[0.01, 0.1, 0.2, 0.3]	0,01
Subsample	[0.5,0.8,1]	0,5
Min_child_weight	[5,10,20,40]	5
Colsample_bytree	[0.5,0.8,1]	1
Gamma	[0,0.1,1,10]	0

Çizelge 4.28. LSTM algoritması hiper parametre değerleri

LSTM		
Hiper Parametreler	Girilen Değerler	GridSearch ile Seçilen Değerler
Units	[64, 128, 256]	256
Activation	[relu, tanh]	tanh
Batch_size	[16, 32, 64]	64
Learning_rate	[0.1, 0.01, 0.001]	0.001
Epochs	[50, 100, 200]	200

Çizelge 4.29. Hiper parametre değerleri ile elde edilen bulgular

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon	LSTM
R ²	0.783	-	0.788	0.786
RMSE	1.427	-	1.410	1.376
MAE	0.678	-	0.693	0.658

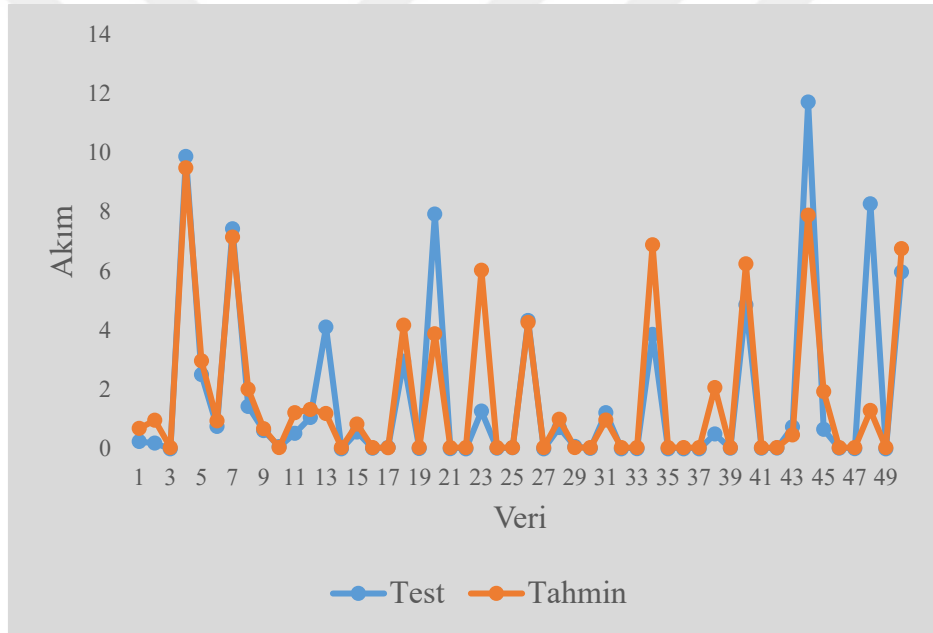
Çizelge 4.29'da hiper parametre değerlerinin optimizasyonu sonucunda elde edilen performans sonuçları incelenmiştir. Bu inceleme sonucunda, R2 metriği açısından XGBoost algoritmasının performansının arttığı açıkça görülmektedir. XGBoost, bu metrikle değerlendirildiğinde daha yüksek bir açıklayıcılık düzeyi sunarak tahminlerdeki doğruluğunu artırmıştır. Diğer yandan, kök ortalama karesel hata ve ortalama mutlak hata metriklerine göre LSTM algoritması daha üstün performans sergilemiştir. LSTM, bu metrikler doğrultusunda daha düşük hata oranlarına ulaşarak tahmin doğruluğunda öne çıkmaktadır. Çizelge 4.30'da ise, veri setinde bulunan çeşitli değişkenlerin tahmin performansına olan etkileri oransal olarak gösterilmiştir. Bu çizelgede, her bir değişkenin modelin tahmin yeteneği üzerindeki etkisi detaylı bir şekilde analiz edilerek sunulmuştur. Bu analizler, hangi değişkenlerin tahmin performansını daha fazla etkilediğini anlamak ve modelin optimizasyonunda hangi verilere daha fazla önem verilmesi gerektiğini belirlemek açısından büyük önem taşımaktadır.

Çizelge 4.30. Eğitime katılan verilerin performansa etkileri

	Rastgele Orman	Lineer Regresyon	XGBoost Regresyon
Veri Setindeki Girdilerin Önem Sıralaması	1. isik: %79.42 2. nem: %10.51 3. sıcaklık: %10.07	1. sıcaklık: %232.66 2. isik: %4.76 3. nem: %-137.42	1.isik: %87.74 2.nem: %6.88 3.sıcaklık: %5.38

NİT ile elde edilen dış ortam verileri ile yapılan makine öğrenimi uygulamasında ışıık şiddeti değeriinin en fazla yararlanılan veri olduğu görülmektedir.

Elde edilen tahmin verileri ve test verileri ile çizilen grafik şekil 4.5’ te görülmektedir.



Şekil 4.5 Atmosfer verileri ile elde edilmiş tahminler ile test verileri grafiği

5. SONUÇ VE ÖNERİLER

5.1. Sonuçlar

Bu çalışmanın temel amacı, fotovoltaik panellerden üretilebilecek enerji miktarını tahmin etmek olarak belirlenmiştir. Bu hedef doğrultusunda, en güvenilir verilere ve isabetli tahmin sonuçlarına ulaşmak maksadıyla verilerin doğrudan tarafımızdan toplanması gerektiği kanaatine varılmıştır. Bu bağlamda güneş enerjisinden elektrik enerjisi üreten ve bu enerjiyi depolayabilen bir sistem modeli oluşturulmuştur. Bu sistemin bulunduğu ortama ait atmosferik verilerin ve anlık akım değerlerinin toplanıp analiz edilebilmesi için Nesnelerin İnterneti Teknolojisi tabanlı modüler bir alt sistem tasarlanmış ve enerji üretim tesisiyle entegre edilmiştir. Ayrıca daha kapsamlı atmosferik verilere ulaşabilmek için, hava durumu tahminleri ve meteorolojik veriler sunan bir web sitesinden yararlanılmıştır. Elde edilen iki veri seti, XGBoost, RFR, LR ve LSTM algoritmaları ile analiz edilmiş ve RMSE, MAE ve R^2 metrikleriyle değerlendirilerek algoritmaların performansları karşılaştırılmış ve sonuçlar paylaşılmıştır.

İlk veri seti üç farklı veri kümesine, diğer veri seti ise 2 farklı veri kümesine sahip olacak şekilde kullanılmıştır. Bu beş farklı veri kümesi, dört makine öğrenimi algoritması ile eğitime tabi tutularak üç farklı metrik ile değerlendirilip alınan sonuçlar tablolar halinde sunulmuştur. Tüm senaryolardan elde edilen bulgulara dayanarak;

Çizelge 5.1’ de görüldüğü üzere, yalnızca meteorolojik veriler kullanılarak yapılan enerji üretimi tahmin çalışmalarında, makine öğrenimi algoritmalarının tahmin performansları düşük kalmaktadır. Hava durumu verilerinin birçok değişken içermesi ve bunların da anlık olarak fazlaca değişkenlik göstermesi buna neden olarak görülmektedir.

Meteorolojik veriler içerisinde basınç, rüzgâr ve nem gibi faktörler tahmin performansına en az katkıyı sunan hatta negatif etki yaratabilen veriler olarak ortaya çıkmaktadır. Bununla beraber ultraviyole verisi, en güçlü pozitif etkiyi sunan veridir ve ardından sıcaklık ve bulutluluk verileri en fazla katkıyı sağlayan veriler olarak görülmektedir

Çizelge 5.1. Atmosferik verilere ait değerlendirme sonuçları

Veri Seti	Metrikler	RFR	LR	XGB	LSTM
Hava Durumu Sunucusu Veri Seti	R ²	0.79536	0.623846	0.78360	0.712
	RMSE	1.36760	1.854166	1.40632	1.621
	MAE	0.68525	1.063199	0.77432	0.840
NİT Veri Seti	R ²	0.78370	0.453795	0.78895	0.786
	RMSE	1.42789	2.26910	1.41045	1.376
	MAE	0.67875	1.70431	0.69399	0.658

İki farklı veri setine ait atmosferik verilerle gerçekleştirilen eğitimlerde, tahmin performanslarının birbirlerine oldukça yakın oranlarla sonuçlandığı gözlemlenmiştir. Bu durum, her iki veri setinin de atmosferik verilerle yapılan eğitimlerde benzer tahmin doğruluğu sağladığını göstermektedir. Ancak, atmosferik verilerin katılmadığı ve yalnızca fotovoltaik panellerden toplanan akım verileri ile gerçekleştirilen eğitimlerin sonuçları, atmosferik verilerle yapılan eğitimlerden daha başarılı tahmin performansı sunduğunu ortaya koymuştur. Çizelge 5.2'de görüldüğü üzere, akım verilerine dayalı eğitimlerin performansı, atmosferik verilerin eklenmesiyle yapılan eğitimlerden üstün olmuştur.

Daha detaylı incelendiğinde, atmosferik verilerin ve önceki akım verilerinin aynı veri seti içerisinde girdi olarak sunulmasının, tahmin performanslarını tüm algoritmalarda belirgin bir şekilde artırdığı tespit edilmiştir. Bu durum, özellikle Çizelge 5.2'de açıkça görülmektedir. Bu analizler, atmosferik veriler ve geçmiş akım verilerinin birlikte kullanıldığında, modelin tahmin gücünü önemli ölçüde güçlendirdiğini göstermektedir.

Sadece atmosferik veriler kullanılarak yapılan eğitimlerde, iki farklı veri seti için birbirine oldukça yakın performans değerleri elde edilmiştir. Ancak, akım verileri eklenerek yapılan eğitimlerde, Nesnelerin İnterneti (NİT) tabanlı sistemden elde edilen veri seti daha yüksek performans sergilemiştir. Bu, NİT tabanlı veri setinin, tahmin modellerinde daha güvenilir ve doğru sonuçlar sağladığını göstermektedir.

NİT tabanlı veri seti ile yapılan eğitimlerde, R² metrik değerlendirmesi açısından Rastgele Orman (RFR) algoritması 0.929, XGBoost algoritması ise 0.930 sonuçlarını

almıştır. Bu yüksek performans değerleri, NİT tabanlı veri setinin, özellikle bu iki algoritma için oldukça etkili olduğunu ve modelin doğruluk seviyesini önemli ölçüde artırdığını ortaya koymaktadır. Bu bulgular, gelecekteki çalışmalar için NİT tabanlı veri toplama sistemlerinin önemini ve kullanımının model performansını nasıl artırabileceğini vurgulamaktadır.

Çizelge 5.2. Akım verilerine ait değerlendirme sonuçları

Veri Seti	Metrikler	RFR	LR	XGB	LSTM
Akım Veri Seti	R ²	0.858	0.840	0.863	0.859
	RMSE	1.135	1.207	1.115	0.133
	MAE	0.489	0.559	0.492	0.586
Hava Durumu Sunucusu Veri Seti + Akım Verileri	R ²	0.886	0.854	0.890	0.884
	RMSE	1.016	1.152	1.000	1.028
	MAE	0.433	0.560	0.441	0.437
NİT Veri Seti + Akım Verileri	R ²	0.929	0.871	0.930	0.889
	RMSE	0.809	1.081	0.797	0.986
	MAE	0.336	0.527	0.342	0.444

Hava durumu sunucusundan alınan veri setinde ise; R² metrik değerlendirmesinden RFR algoritması 0.886, XGBoost Algoritması ise 0.890 sonuçlarını almıştır.

Bu durum tesisin bulunduğu ortamdan alınan verilerin, tahmin işlemlerinde, diğer veri setine göre daha gerçekçi ve işe yarar olduğunu göstermektedir. Elde edilen veri setleri ile yapılan analizler sonucunda, en iyi performans bu veri kümesinde elde edilmiştir ve en iyi sonuç XGBoost yöntemi ile alınmıştır. Tüm veri kümeleri üzerinde elde edilen performans metrikleri değerlendirildiğinde, XGBoost' un en başarılı sonuçları verdiği görülmektedir.

Ancak hava durumu sunucusundan elde edilen veri seti için herhangi ek maliyet gerekmezken, NİT veri seti için sistem kurulumu maliyeti, internet bağlantı ücreti ve internet bağlantı problemleri, sensör arızaları gibi istenmeyen durumlar söz konusu olmaktadır.

5.2. Öneriler

Bu çalışmada, makine öğrenimi algoritmalarının tahmin performansını değerlendirmek amacıyla çeşitli analizler gerçekleştirilmiştir. Kullanılan algoritmalar arasında Rastgele Orman, Doğrusal Regresyon, XGBoost ve Uzun Kısa Süreli Bellek (LSTM) bulunmaktadır. Her bir algoritmanın farklı özellikleri ve güçlü yönleri vardır, bu da onları çeşitli veri setleri ve problem türleri için uygun hale getirmektedir.

Rastgele Orman algoritması, birçok karar ağacının birleşiminden oluşan güçlü bir ensemble öğrenme yöntemidir. Bu algoritma, veri setindeki karmaşık ilişkileri yakalayabilme yeteneği ve aşırı uyum sorununu azaltma kapasitesi sayesinde geniş bir kullanım alanına sahiptir. Doğrusal Regresyon ise, özellikle bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiyi modellemek için kullanılan basit ve hızlı bir algoritmadır.

XGBoost, Gradient Boosting algoritmasının optimize edilmiş bir versiyonudur ve özellikle yüksek doğruluk gerektiren rekabetçi veri bilim projelerinde yaygın olarak kullanılır. Bu algoritma, hataları iteratif olarak düzelterek model performansını sürekli iyileştirir. Ayrıca, XGBoost 'un yüksek hesaplama hızları ve paralel işleme yetenekleri, büyük veri setleri üzerinde çalışmayı kolaylaştırmaktadır.

Uzun Kısa Süreli Bellek (LSTM) algoritması ise, özellikle zaman serisi verileri ve sekans tabanlı problemler için etkili olan bir tür RNN (Recurrent Neural Network) modelidir. LSTM, zaman içindeki uzun süreli bağımlılıkları öğrenme yeteneği sayesinde, geçmiş verilerden gelecekteki verileri tahmin etmede başarı göstermektedir.

Veri setinin temel özellikleri ve boyutu, hangi algoritmanın kullanılacağı konusunda önemli bir rol oynamaktadır. Farklı algoritmaların kullanılması, modelin çeşitli veri türlerine ve özelliklerine uyum sağlamasına yardımcı olabilir. Bu çalışmada kullanılan geniş yelpazedeki makine öğrenimi algoritmaları, model performansını iyileştirmek için çeşitli stratejilerin ve yaklaşımların denenmesine olanak tanımıştır. Sonuç olarak, farklı algoritmaların uygulanması ve karşılaştırılması, hangi modelin belirli bir veri seti ve problem türü için en iyi performansı sunduğunu belirlemek açısından kritik bir önem taşımaktadır. Bu yaklaşım, gelecekteki modelleme çalışmalarında daha etkili ve güvenilir tahminler yapmayı mümkün kılabilmektedir.

Süreç boyunca yapılan analizler, veri setindeki bazı özelliklerin model performansını hiç etkilemediğini, hatta bazen negatif bir etki yarattığını ortaya koymuştur. Özellikle, bazı bağımsız değişkenlerin hedef değişkenle olan ilişkisinin zayıf olduğu veya veri setinde gürültü yarattığı belirlenmiştir. Bu durum, modelin genel doğruluğunu ve genelleme kapasitesini düşürebilmektedir. Bu nedenle, özellik seçimi ve boyut indirgeme tekniklerinin uygulanması, model performansının optimize edilmesi için kritik öneme sahiptir. Gelecekteki çalışmalarda, özellik mühendisliği ve önem derecesi analizlerinin dikkatle yapılması, daha etkin ve verimli modellerin geliştirilmesine katkı sağlayacaktır.

Çalışmada, atmosfer verileri ile güneş panellerinden toplanan akım verilerini birleştirerek yeni bir veri seti oluşturulmuştur. Bağımlı değişken olarak akım değeri seçilmiştir. Modelin tahmin performansını artırmak amacıyla, veri setine önceki dönemlere ait akım değerleri de eklenmiştir. Spesifik olarak, bir önceki akım değeri ve iki önceki akım değeri özellikler arasına dahil edilmiştir. Bu yaklaşımın nedeni, akım değerlerinin zaman içinde süreklilik ve bağımlılık gösterebilmesidir. Geçmiş akım değerleri, mevcut akım değerleri üzerinde belirleyici bir etkiye sahip olabilmekte ve bu tür zaman serisi bileşenlerinin modele eklenmesi, gelecekteki akım değerlerinin daha doğru tahmin edilmesine olanak tanımaktadır. Böylece, modelin geçmiş verilerden öğrenerek daha yüksek doğrulukta tahminler yapması sağlanmıştır.

Gelecekteki çalışmalarda, modelin tahmin doğruluğunu daha da artırmak amacıyla, veri setine üç önceki, dört önceki ve n önceki akım değerlerinin de eklenmesi tavsiye edilmektedir. Bu eklemeler, modelin daha geniş bir zaman dilimindeki akım değişikliklerini yakalamasına ve daha kapsamlı bir öğrenme süreci geliştirmesine katkıda bulunabilir. Bu şekilde, modelin genel performansı ve genelleme kapasitesi iyileştirilebilir.

Çalışmada, veri ön işleme aşamasında gece saatlerinde elde edilen akım verilerinde bazı düzensizlikler tespit edilmiştir. Normal koşullarda, gece saatlerinde güneş panellerinden alınan akım değerlerinin 0 Amper olması beklenirken, bazen negatif yönde çok küçük akım değerleri gözlemlenmiştir. Bu tür negatif akım değerleri, ölçüm hataları veya sistem gürültüsünden kaynaklanabilir. Bu durumu düzeltmek amacıyla, negatif akım değerleri 0 Amper olarak kabul edilerek veriler işlenmiştir. Alternatif olarak, gece saatlerinde alınan tüm akım değerleri de analiz dışı bırakılarak yalnızca gündüz saatlerine ait veriler ile çalışılabilir. Ancak, bu çalışmada tercih edilen yöntem, negatif akım değerlerinin 0 olarak

düzeltilmesi yoluyla veri setinin bütünlüğünü korumak olmuştur. Bu yaklaşım, modelin doğruluğunu artırmak ve olası hatalı ölçümlerin model performansını olumsuz etkilemesini önlemek için uygulanmıştır.

Çalışmada elde edilen veri setinde, güneş panellerinden toplanan akım verileri ile atmosfer verileri arasında ölçek farklılıkları bulunmaktadır. Bu farklılıklar, modelin genelleme yeteneğini ve performansını olumsuz etkileyebilmektedir. Bu sebeple enerji üretim tahminlerinin doğruluğunu artırmak amacıyla çeşitli ön işleme adımlarından geçirilmiştir. Gelecekteki çalışmalarda, model performansının daha da iyileştirilmesi ve daha güvenli tahminler elde etmek için verilerin normalize edilmesi önerilmektedir. Aynı zamanda diğer bir ön işleme adımı olarak Fourier dönüşümü de kullanılabilir. Fourier dönüşümü, zaman serisi verilerini frekans domenine dönüştürerek verideki periyodik bileşenlerin ve frekansların belirlenmesine olanak tanımaktadır. Bu sayede, veri setindeki bileşenler daha iyi analiz edilebilir ve ayrıştırılabilir. Elde edilen frekans bileşenleri, modelin tahmin performansını artırmak amacıyla ek özellikler olarak veri setine dahil edilebilir. Ayrıca, Fourier dönüşümü ile gürültünün belirlenmesi ve filtrelenmesi, verilerin daha temiz ve güvenilir hale getirilmesine katkıda bulunmaktadır. Bu nedenle, Fourier dönüşümünün veri ön işleme sürecine dahil edilmesi, enerji üretim tahmin modellerinin genel doğruluğunu ve genelleme kapasitesini iyileştirmek için önemli bir adım olarak değerlendirilmelidir.

Gelecekteki enerji üretim tahmin çalışmalarında kullanılacak veri setlerinin seçimi konusunda, sonuçlar bölümünde tartışılan hususların da dikkate alınması, olası problemlere karşı önceden hazırlıklı olmayı sağlayacaktır.

KAYNAKLAR

- Abate, F., Carratù, M., Liguori, C. ve Paciello, V.** (2019). A Low Cost Smart Power Meter for IoT. *Measurement*, 136, 59–66.
- Abdel-Basset, M., Manogaran, G., Mohamed, M., & Rushdy, E.** (2019). Internet of things in smart education environment: Supportive framework in the decision- making process. *Concurrency and Computation: Practice and Experience*, 31(10), e4515.
- Akar, Ö., & Güngör, O.** (2012). Rastgele Orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması. *Jeodezi ve Jeoinformasyon Dergisi* (106), 139-146. <https://doi.org/10.9733/jgg.241212.1t>
- Akkaya, O. S.** (2022). Gerçek Zamanlı Enerji İzleme İçin Nesnelerin İnterneti Tabanlı Akıllı Sayaç Tasarımı, (Yüksek Lisans Tezi), Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü Elektrik ve Elektronik Anabilim Dalı, Konya.
- Andy, Syaiful, Budi Rahardjo, and Bagus H.** (2017). "Attack scenarios and security analysis of MQTT communication protocol in IoT system." *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. IEEE, 2017.
- Balamurugan, S. and Saravanakamalam, D.** (2017). Energy Monitoring and Management Using Internet of Things, *International Conference on Power and Embedded Drive Control*, 208-212., Chennai, India
- Balcı, F.** (2022). A Hybrid Attention-based LSTM-XGBoost Model for Detection of ECG-based Atrial Fibrillation. *Gazi University Journal of Science Part A: Engineering and Innovation*, 9(3), 199-210. <https://doi.org/10.54287/gujisa.1128006>
- Bassi, A. ve Horn, G.** (2008). Internet of Things in 2020: A Roadmap for the Future. European Commission: Information Society and Media, 22, 97-114
- Bergstra, J., & Bengio, Y.** (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- Bilen, A., & Özer, A. B.** (2022). Siber Saldırıları için Rastgele Orman Algoritması Kullanılarak Öznitelik Seçimi. *Fırat Üniversitesi Fen Bilimleri Dergisi*, 34(1), 31-37.
- Burritt, R. and Christ, K.** (2016). Industry 4.0 and environmental accounting: a new revolution?, *Asian Journal of Sustainability and Social Responsibility* 1, 23–38, <https://ajssr.springeropen.com/articles/10.1186/s41180-016-0007>
- Chen, R. C.** (2019). Using Deep Learning to Predict User Rating on Imbalance Classification Data. *IAENG International Journal of Computer Science*, 46(1).

- Chen, R. C., Dewi, C., Huang, S. W., & Caraka, R. E.** (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, 7(1), 52.
- Chen, T., & Guestrin, C.** (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Chollet, F.** (2021). *Deep learning with Python*. Simon and Schuster.
- Copeland, B. J., & Proudfoot, D.** (2007). Artificial intelligence: History, foundations, and philosophical issues. In *Philosophy of psychology and cognitive science* (pp. 429-482). North-Holland.
- Çelik, S., & Özdemir, D.** (2023). Rastgele orman regresyon algoritması ile bitcoin fiyat tahmini. *Journal of Scientific Reports-B*, (008), 55-64.
- Deng, L., & Yu, D.** (2014). Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7(3-4), 197-387.
- Duman, E.** (2022). Implementation of XGBoost Method for Healthcare Fraud Detection. *Scientific Journal of Mehmet Akif Ersoy University*, 5(2), 69-75.
- Ertürk, S., Kara, H., Akkuş, C., Genç, G.** (2023). Türkiye’de Farklı İklim Kuşakları İçin Yapay Sinir Ağları Kullanılarak Güneş Işınımının Tahmini. *Gazi University Journal of Science Part C: Design and Technology*, 11(4), 885-892. <https://doi.org/10.29109/gujsc.1331788>
- Gündüz, M. Z., & Resul, D. A. Ş.** (2017). “Nesnelerin İnterneti: Gelişimi, Bileşenleri ve Uygulama Alanları”, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 24(2), 327-335, (2017).
- Hossain, M. R., Oo, A. M. T., and Ali, A. B. M. S.** (2012). "Hybrid prediction method of solar power using different computational intelligence algorithms", 2012 22nd Australasian Universities Power Engineering Conference: “Green Smart Grid Systems”, AUPEC 2012, Bali, Indonesia, 26-26 September, 2012.
- İmren, Ş., Dede, M. C., & Özdemir, S.** (2012). Dört Kol Mekanizmasında Eklem Aşınmasından Tekrarlanan Pozisyon Hatalarını Minimize Etmek İçin Bir Takviyeli Öğrenme Algoritması.
- Kalemci, M.E.** (2021). Akıllı Şebeke Uygulamalarına Yönelik Nesnelerin İnterneti Tabanlı Uzaktan İzleme Sistemi, (Yüksek Lisans Tezi), Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü, Elektrik- Elektronik Mühendisliği Anabilim Dalı, Trabzon.
- Karaman, B., Taşkın, S. ve Tokay, M.** (2021). Gerçek Zamanlı Enerji İzleme ve Tüketici Farkındalığı için LabVIEW TM Programı ile Otomatik Sayaç Okuma. *International Journal of Advances in Engineering and Pure Sciences*, 33(2), <https://doi.org/10.7240/jeps>.
- Kartal, E., & Balaban, M. E.** (2018). Machine learning techniques in cardiac risk assessment. *Turkish Journal of Thoracic and Cardiovascular Surgery*, 26(3), 394.
- Kaur, M. and Kalra, S.** (2016). A, Review on IoT Based Smart Grid, *International Journal of Energy, Information and Communications*, Department of

Computer Science Guru Nanak Dev University, Regional Campus
Jalandhar, India.

- Kayıklı, T., & Balıkçı, A.** (2014). Elektrikli Araçlarda Kullanılan Lityum-Polimer Aküler İçin Bir Şarj Cihazı Tasarımı A Lithium-Polymer Battery Charger Design for Electric Vehicles.
- Kılıç, S.** (2013). Doğrusal regresyon analizi. *Journal of Mood Disorders*, 3(2), 90
- Kır, Y.** (2022) Güneş Paneli Parametrelerinin Online Ölçümü ve Uzaktan İzlenmesi. *Sürdürülebilir Mühendislik Uygulamaları ve Teknolojik Gelişmeler Dergisi*, 4(1), 46-52.
- Kızılkaya, Y. M., & Oğuzlar, A.** (2018). Bazı Denetimli Öğrenme Algoritmalarının R Programlama Dili İle Kıyaslanması. *Karadeniz Uluslararası Bilimsel Dergi*, 37(37), 90-98.
- Kleineidam, G., Krasser, M., & Reischböck, M.** (2016). The cellular approach: smart energy region Wunsiedel. Testbed for smart grid, smart metering and smart home solutions. *Electrical Engineering*, 98(4), 335-340.
- Kosmatos, E.A., Tselikas, N.D. and Boucouvalas, A.C.** (2011) Integrating RFIDs and Smart Objects into a Unified Internet of Things Architecture. *Advances in Internet of Things: Scientific Research*, 1, 5-12.
- Lennartsson, A., & Blomberg, M.** (2021). Fault Detection in Permanent Magnet Synchronous Motors using Machine Learning.
- Liu, B., & Liu, B.** (2011). Supervised learning. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, 63-132.
- López-de-Armentia, J., Casado-Mansilla, D., & López-de-Ipina, D.** (2012, July). Fighting against vampire appliances through eco-aware things. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing* (pp. 868-873). IEEE.
- Melikoğlu, M.** (2013). "Vision 2023: Feasibility analysis of Turkey's renewable energy projection", *Renewable Energy* 50, 570-575, 2013.
- Mudaliar, M. D. ve Sivakumar, N.** (2020). IoT Based Real Time Energy Monitoring System Using Raspberry Pi. *Internet of Things*, 12, 100292. <https://doi.org/10.1016/j.iot.2020.100292>
- Nasiri, H., Nasehi, S., & Goudarzi, M.** (2019). Evaluation of distributed stream processing frameworks for IoT applications in Smart Cities. *Journal of Big Data*, 6(1), 52.
- Onat, N.** (2018). Türkiye’de yenilenebilir kaynaklardan elektrik enerjisi üretimi: Mevcut durum ve gelecek beklentileri. *Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi*, 1(1), 8-15.
- Özbay, H., Parmaksız, H., Karafil, A., & Kesler, M.** (2016). Farklı Eğitim Açılarındaki Fotovoltaik Panellerin Elektriksel Ölçümlerinin Raspberry Pi ile İzlenmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 4(2), 711-718.
- Özkesici, M. Y., & Yılmaz, S.** (2021). Oral ve Maksillofasiyal Radyoloji’de Yapay Zekâ. *Sağlık Bilimleri Dergisi*, 30(3), 346-351.

- Rao, TA ve Haq, AB.** (2018). IoT katmanlarının karşılaştığı güvenlik sorunları ve koruyucu önlemleri. *Uluslararası Bilgisayar Uygulamaları Dergisi*, 179 (27), 31-35.
- Sarkar, D., Bali, R., & Ghosh, T.** (2018). *Hands-On Transfer Learning with Python: Implement advanced deep learning and neural network models using TensorFlow and Keras*. Packt Publishing Ltd.
- Sâri, M. B., Yuliantini, L., Prihtiadi, H., & Djamel, M.** (2021). Easy monitoring and data record system of electric current detected by ACS712 affordable non-destructive electrical current sensor. *Pillar of Physics*, 13(2).
- Sharma, N., Sharma, P., Irwin, D., & Shenoy, P.** (2011, October). Predicting solar generation from weather forecasts using machine learning. In *2011 IEEE international conference on smart grid communications (SmartGridComm)* (pp. 528-533). IEEE.
- Shao, Y., Zhao, Y., Yu, F., Zhu, H., & Fang, J.** (2021). The Traffic Flow Prediction Method Using the Incremental Learning-Based CNN-LTSM Model: The Solution of Mobile Application. *Mobile Information Systems*, 2021(1), 5579451.
- Shi, J., Lee, W. J., Liu, Y., Yang, Y., & Wang, P.** (2012). Forecasting power output of photovoltaic systems based on weather classification and support vector machines. *IEEE Transactions on Industry Applications*, 48(3), 1064–1069. <https://doi.org/10.1109/TIA.2012.2190816>
- Singh, M.J., Agarwal, P., and Padmanabh, K.** (2016). Load Forecasting at Distribution Transformer using IoT based Smart Meter Data from 6000 Irish Homes, Contemporary Computing and Informatics (IC3I) International Conference, 758- 763., <https://ieeexplore.ieee.org/document/7918062>.
- Stusek, M., Pokorny, J., Masek, P., Hajny, J., Hosek, J.** (2017). A Non-Invasive Electricity Measurement Within The Smart Grid Landscape: Arduino-Based Visualization Platform for IoT, 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, 423-429.
- Swan, M.** (2012). Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator networks*, 1(3), 217-253.
- Tan, Z., & Karaköse, M.** (2022). Dinamik ortamlarda derin takviyeli öğrenme tabanlı otonom yol planlama yaklaşımları için karşılaştırmalı analiz. *Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi*, 9(16), 248-262.
- Tang, N., Mao, S., Wang, Y., & Nelms, M.** (2018). LASSO-based single index model for solar power generation forecasting. 2017 IEEE Global Communications Conference, GLOBECOM 2017- Proceedings, 2018-Janua, 1–6. <https://doi.org/10.1109/GLOCOM.2017.8255070>
- Taşkesen, E., Üren, R., & Üren, S.** (2023). Nesnelerin İnterneti (IoT) ve Akıllı Şebekeler: IoT Cihazları Akıllı Şebekeler İçin Kullanımı. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(4), 43-49.
- Url-1** <<https://www.eia.gov/>>, erişim tarihi 28.12.2023.

- Url-2** <<https://www.teias.gov.tr/>>, erişim tarihi 30.12.2023.
- Url-3** <<https://enerji.gov.tr/>> erişim tarihi 28.12.2023
- Url-4** <<https://www.statista.com/>>, erişim tarihi 16.11.2023
- Url-5** <<https://megep.meb.gov.tr/>>, erişim tarihi 28.04.2024.
- Url-6** <<https://thingspeak.com/>>, Son erişim tarihi 01.05.2024.
- Varınca, K. B., & Gönüllü, M. T.** (2006). Türkiye’de güneş enerjisi potansiyeli ve bu potansiyelin kullanım derecesi, yöntemi ve yaygınlığı üzerine bir araştırma. *Ulusal güneş ve hidrojen enerjisi kongresi*, 21, 23.
- Vural, A.** (2022). Fotovoltaik Panellerde Güç Üretiminin Nesnelerin İnterneti Tabanlı Gerçek Zamanlı Uzaktan İzlenmesi ve Verilerin Saklanması, (Yüksek Lisans Tezi), Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü Yenilenebilir Enerji Sistemleri Anabilim Dalı, Afyonkarahisar
- Wang, F., Zhen, Z., Wang, B., & Mi, Z.** (2017). Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting. *Applied Sciences*, 8(1), 28. <https://doi.org/10.3390/app8010028>
- Wang, H., Ma, C., & Zhou, L.** (2009, December). A brief review of machine learning and its application. In *2009 international conference on information engineering and computer science* (pp. 1-4). IEEE.
- Wu, Z., Zhou, M., Lin, Z., Chen, X., & Huang, Y.** (2021). Improved genetic algorithm and xgboost classifier for power transformer fault diagnosis. *Frontiers in Energy Research*, 9, 745744.
- Xin, J., Zhou, C., Jiang, Y., Tang, Q., Yang, X., & Zhou, J.** (2023). A signal recovery method for bridge monitoring system using TVFEMD and encoder-decoder aided LSTM. *Measurement*, 112797.
- Yang, H. T., Huang, C. M., Huang, Y. C., and Pai, Y. S.** (2014). "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output", *IEEE Transactions on Sustainable Energy* 5(3), 917-926, 2014
- Yang, J., & Guan, J.** (2022). A Heart Disease Prediction Model Based on Feature Optimization and Smote-Xgboost Algorithm. *Information* 2022, 13, 475.
- Yousif, J., & A. Kazem, H.** (2016). Modeling of Daily Solar Energy System Prediction using Soft Computing Methods for Oman. *Research Journal of Applied Sciences, Engineering and Technology*, 13(3), 237–244. <https://doi.org/10.19026/rjaset.13.2936>
- Yurtsever, S. B.** (2023). Sosyal Medya İstihbaratının Makine Öğrenmesi Çerçevesinde İncelenmesi: Terörizm Çalışmaları. *Savunma ve Güvenlik Araştırmaları Dergisi*, 1(1), 97-119.
- Zeng, J., & Qiao, W.** (2013). Short-term solar power prediction using a support vector machine. *Renewable Energy*, 52, 118–127. <https://doi.org/10.1016/j.renene.2012.10.009>
- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J.** (2021). Dive into deep learning. *arXiv preprint arXiv:2106.11342*.

- Zhang, Y., & Wang, J.** (2015, July). GEFCom2014 probabilistic solar power forecasting based on k-nearest neighbor and kernel density estimator. In *2015 IEEE Power & Energy Society General Meeting* (pp. 1-5). IEEE.
- Zilyas, D., & Yılmaz, A.** (2023). Makine Öğrenmesi Yöntemleri ile Eğitim Başarısının Tahmini Modeli. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 14(3), 437-447

