



**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
YENİLENEBİLİR ENERJİ SİSTEMLER ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**DALGACIK DÖNÜŞÜMÜ VE SHAP DESTEKLİ ÖZELLİK SEÇİMİ
İLE CNN-LSTM TABANLI GÜNEŞ IŞINIMI TAHMİNİ**

Songül KAYIK

**Ağustos-2025
BATMAN**

T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
YENİLENEBİLİR ENERJİ SİSTEMLER ANA BİLİM DALI

YÜKSEK LİSANS TEZİ

DALGACIK DÖNÜŞÜMÜ VE SHAP DESTEKLİ ÖZELLİK SEÇİMİ
İLE CNN-LSTM TABANLI GÜNEŞ IŞINIMI TAHMİNİ

Songül KAYIK

Danışman
Dr. Öğr. Üyesi Selahattin Barış ÇELEBİ

Diğer Jüri Üyeleri

Doç. Dr. Ömer Ali KARAMAN

Dr. Öğr. Üyesi Ahmet YILMAZ

Dr. Öğr. Üyesi Selahattin Barış ÇELEBİ

Ağustos-2025
BATMAN

TEZ KABUL VE ONAYI

Songül KAYIK tarafından hazırlanan “DALGACIK DÖNÜŞÜMÜ VE SHAP DESTEKLİ ÖZELLİK SEÇİMİ İLE CNN-LSTM TABANLI GÜNEŞ İŞİNİMİ TAHMİNİ” adlı tez çalışması 12/08/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Yenilenebilir Enerji Sistemleri Ana Bilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Doç. Dr. Ömer Ali KARAMAN

Danışman

Dr. Öğr. Üyesi Selahattin Barış ÇELEBİ

Üye

Dr. Öğr. Üyesi Ahmet YILMAZ

İmza

.....

.....

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYANI

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sorumluluğu kabullendiğimi bildiririm.

ETHICAL DECLARATION

I declare that all the information in this thesis has been obtained within the framework of ethical behavior and academic rules, and that the source of any statements and information that do not belong to me in this study prepared in accordance with the thesis writing rules has been fully cited, and I declare that I accept all kinds of legal responsibility in case of any contrary situation.

Songül KAYIK
Tarih: 10/08/2025

ÖZET

YÜKSEK LİSANS TEZİ

DALGACIK DÖNÜŞÜMÜ VE SHAP DESTEKLİ ÖZELLİK SEÇİMİ İLE CNN-LSTM TABANLI GÜNEŞ İŞINIMI TAHMİNİ

Songül KAYIK

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü

Yenilenebilir Enerji Sistemleri Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Selahattin Barış ÇELEBİ

2025, 68 Sayfa

Bu tez çalışması, güneş ışıınımı tahmininde yüksek doğruluk ve açıklanabilirlik sağlamak amacıyla ayrık dalgacık tabanlı hibrit bir Evrişimsel Sinir Ağı (CNN)- Uzun Kısa Süreli Bellek Ağı) modeli ile SHAP destekli analiz ve özellik seçimini bir araya getiren özgün bir çerçeve önermektedir. Çalışmada, Suudi Arabistan'ın Riyad şehrine ait 15 dakikalık aralıklarla toplanan üç yıllık meteorolojik veriler kullanıldı. Manuel özellik mühendisliği ile veri seti zenginleştirildi, ardından Ayrık Dalgacık Dönüşümü (DWT) uygulanarak gürültü azaltımı ve çok ölçekli bileşen çıkarımı gerçekleştirildi. CNN katmanları uzamsal desenleri, LSTM katmanları ise zamansal bağımlılıkları modelledi. Optuna tabanlı Bayesçi optimizasyon ile hiperparametreler belirlendi. Shapley Katkı Açıklamaları (SHAP) analizi ile ortalama katkısı en yüksek %20'lik dilimdeki değişkenler seçildi ve model yeniden eğitilerek ablasyon çalışmaları ile bileşen katkıları doğrulandı. 5 katlı zaman serisi bölme çapraz doğrulama yöntemi ile yapılan değerlendirmelerde modelin, her bir fold sonucunda düşük Kök Ortalama Kare Hatası (RMSE) ve yüksek Belirleme Katsayısı (R^2) değerleri elde edebildiği görüldü. Özellikle zenith açısı ve bulut opaklığı gibi değişkenlerin tahminde baskın rol oynadığı belirlendi. Ayrıca modelin sağlamlığı belirsizlik tahmini ile değerlendirildi. SHAP yardımıyla seçilen özellikler Temel Bileşenler Analizi (PCA) ile de doğrulandı. Elde edilen sonuçlar, önerilen yöntemin güneş enerjisi potansiyelinin özellikle kısa vadeli değerlendirilmesinde, şebeke entegrasyonu ve enerji yönetiminde güvenilir bir araç olduğunu göstermektedir. Önerilen yaklaşım benzer iklim koşullarına sahip bölgelerde uygulanabilme potansiyeli taşımakta olup, esnek bir tahmin sistemi sunarak literatürdeki açıklanabilir Derin Öğrenme (DL) tabanlı güneş ışıınımı tahmini çalışmalarına önemli bir katkı sağlama potansiyeline sahiptir.

Anahtar Kelimeler: Güneş ışıınımı tahmini, ayrık dalgacık dönüşümü, hibrit CNN-LSTM, SHAP, özellik seçimi

ABSTRACT

MASTER THESIS

SOLAR IRRADIANCE FORECASTING WITH CNN-LSTM BASED ON WAVELET TRANSFORMATION AND SHAP-SUPPORTED FEATURE SELECTION

Songül KAYIK

Batman University Graduate Education Institute

Department of Renewable Energy Systems

Advisor: Asst. Prof. Dr. Selahattin Barış ÇELEBİ

2025, 68 Pages

This thesis proposes a novel framework that combines a wavelet-based hybrid CNN-LSTM model with SHAP-supported feature selection and analysis to achieve high accuracy and interpretability in solar irradiance forecasting. The study used three years of meteorological data from Riyadh, Saudi Arabia, collected at 15-minute intervals. Manual feature engineering was applied to enrich the dataset. Then, Discrete Wavelet Transform (DWT, db4, level 2) was used for noise reduction and multi-scale component extraction. CNN layers modeled spatial patterns, while LSTM layers captured temporal dependencies. Hyperparameters were optimized using Optuna-based Bayesian optimization. SHAP analysis was employed to select the top 20% of variables with the highest average contribution, and the model was retrained. Ablation studies were conducted to validate component contributions. The evaluation, performed using 5-fold TimeSeriesSplit cross-validation, showed low RMSE and high R^2 scores. Variables such as zenith angle and cloud opacity played a dominant role in predictions. Monte Carlo Dropout was applied for uncertainty estimation, and PCA-based dimensionality reduction was used to improve generalization capability. The results show that the proposed method is a reliable tool for assessing solar energy potential, grid integration, and energy management. This approach offers a flexible prediction system applicable to regions with similar climate conditions and provides a valuable contribution to explainable deep learning-based solar irradiance forecasting studies in the literature.

Keywords: Solar irradiance forecasting, wavelet transform, hybrid CNN-LSTM, SHAP, feature selection

ÖN SÖZ

Bu tez çalışması, güneş enerjisinin sürdürülebilir enerji üretimindeki önemini vurgulamak ve güneş ışıını tahmininde DL tabanlı yöntemlerin etkinliğini ortaya koymak amacıyla hazırlanmıştır. Çalışmanın her aşamasında, yalnızca akademik bir gerekliliği yerine getirmek değil, aynı zamanda yenilenebilir enerji alanına katkı sunabilecek bilimsel bir yaklaşım geliştirmek hedeflenmiştir.

Tez sürecinde bana yol gösteren, bilgi ve tecrübeleriyle her zaman destek olan değerli danışmanım Selahattin Barış Çelebi hocama en içten teşekkürlerimi sunarım. Ayrıca yüksek lisans eğitimim boyunca bilgi birikimlerinden faydalandığım tüm hocalarıma minnettarım.

Bu süreçte sabır, anlayış ve moral desteğiyle yanımda olan aileme ve dostlarıma da şükran borçluyum. Onların desteği, bu çalışmayı tamamlamamda en büyük motivasyon kaynağım olmuştur.

Son olarak, çalışmamda kullandığım verilerin sağlanmasında katkı sunan ilgili kurum ve kuruluşlara da teşekkür ederim.

Bu çalışmanın, yenilenebilir enerji alanında yapılacak yeni araştırmalara ışık tutması dileğiyle...

Songül KAYIK
BATMAN-2025

İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
TABLolar LİSTESİ	ix
ŞEKİLLER LİSTESİ	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
1.1. Tezin Amacı	4
1.2. Literatür Taraması	6
2. GÜNEŞ ENERJİSİNİN TEMELLERİ VE UYGULAMALARI.....	12
2.1. Güneş Enerjisi	12
2.2. Güneş Radyasyon Çeşitleri	12
2.3. Direkt (Doğrudan) Güneş Radyasyonu.....	13
2.4. Dağınık (Difüz) Güneş Radyasyonu.....	13
2.5. Yansıyan Güneş Radyasyonu	14
2.6. Global (Küresel) Güneş Radyasyonu	14
2.7. Yoğunlaştırıcı Termal Sistemler.....	15
2.8. Fotovoltatik Güneş Teknolojileri	16
2.9. Güneş Paneli (Fotovoltaik Modül) Çeşitleri	16
2.10. Gelişen PV Teknolojileri.....	17
2.11. Yükselen Teknolojiler.....	17
2.12. Şebeke Bağlantılı Sistemler	17
2.13. Şebeke Bağlantısız Sistemler	18
2.14. Dünyada Güneş Enerjisi Kullanımı	19
2.15. Riyad'ın Güneş Enerjisi Potansiyeli	20
3. YÖNTEM.....	22
3.1. Yapay Zeka	23
3.2. Makine Öğrenmesi.....	23
3.2.1. Geleneksel makine öğrenmesi yöntemleri	24
3.2.2. Yapay sinir ağları	24
3.3. Derin Öğrenme	26
3.3.1. Evrişimli sinir ağları (CNN)	26
3.3.2. Tekrarlayan sinir ağları (RNN)	28

3.3.3. Uzun kısa süreli bellek ağları (LSTM)	29
3.4. Veri Kümesi	30
3.5. Özellik Çıkarımı	32
3.5.1. Manuel özellik mühendisliği.....	32
3.5.2. Dalgacık dönüşümü ile çok ölçekli özellik çıkarımı	33
3.6. Optuna ile Bayesçi Optimizasyon ve TPE Algoritması.....	34
3.7. Model Değerlendirme	36
3.8. Çapraz Doğrulama ve Veri Bölme	38
4. BULGULAR VE TARTIŞMA	40
4.1. Model Performansı Değerlendirilmesi.....	40
4.2. Zaman Serisi Üzerinde Tahmin Başarımı	41
4.3. Model Belirsizlik Analizi	44
4.4. Öznitelik Seçimi ve Açıklanabilirlik	45
4.5. Ablation Çalışması ve Baseline Karşılaştırmaları	47
4.6. Hiperparametre Optimizasyonu Sonuçları.....	48
4.7. Güncel Literatürle Karşılaştırma	49
4.8. Sınırlamalar ve Gelecekteki Yönelimler.....	51
5. SONUÇLAR VE ÖNERİLER	53
KAYNAKLAR	54
TEŞEKKÜR	67

TABLÖLAR LİSTESİ

Tablo 3.1. Güneş ışınlamı tahmininde kullanılan atmosferik, çevresel ve meteorolojik veri özellikleri	31
Tablo 3.2: Optuna Bayesian Optimizasyonu ile belirlenen hiperparametreler, arama uzayı ve en iyi değerler.....	35
Tablo 4.1. DWT-CNN-LSTM tabanlı modelin 5 katlı çapraz doğrulama sonuçlarına göre test setindeki performans metrikleri	40
Tablo 4.2. Ablasyon ve azaltma analizleri ile temel model karşılaştırmaları (katmanlar arası ortalama).....	48
Tablo 4.3. Güneş ışıması tahmininde ML modellerinin karşılaştırmalı analizi	50
Tablo 4.4. Boru hattı sürecindeki ana adımların süresi	52

ŞEKİLLER LİSTESİ

Şekil 1.1. GHI, DNI ve DHI arasındaki geometrik ilişki ve toplam yatay ışınımın hesaplanması	2
Şekil 2.1. Parabolik oluk tipi yoğunlaştırıcı güneş enerjisi sistemi	15
Şekil 2.2. Türkiye’de Farklı PV Panel Tiplerine Göre Alan-Bazlı Yıllık Elektrik Üretimi (kWh/m ²).....	16
Şekil 2.3. Şebeke bağlantılı (on grid) sistemler	18
Şekil 2.4. Şebeke bağlantısız (off grid) sistemler	19
Şekil 2.5. Suudi Arabistan’ın yıllık ortalama güneş enerjisi potansiyeli dağılımı	21
Şekil 3.1. Önerilen DWT-CNN-LSTM güneş ışınımı tahmin modeli	22
Şekil 3.2. Bir nöronun girdi-çıkışı ilişkisi.....	25
Şekil 3.3 Tekrarlayan sinir ağında girdi, gizli ve çıkış birimleri (uyarlandı)	28
Şekil 3.4. Bir LSTM hücresinin iç yapısı ve çalışma prensibi	29
Şekil 3.5. 5-Katlı çapraz doğrulama süreci	39
Şekil 4.1. Farklı fold’larda modelin RMSE performansı	41
Şekil 4.2. Farklı fold’larda modelin R ² performansı.....	41
Şekil 4.3. Fold 5 için zaman serisi test tahmin sonuçları (İlk 20 Gün)	42
Şekil 4.4. Eğitim ve doğrulama kayıp eğrileri (a), gerçek ve tahmin değerlerinin dağılımı (b), artıkların frekans dağılımı (c)	42
Şekil 4.5. Saatlik RMSE analiziyle modelin gün içi tahmin hatası	43
Şekil 4.6. MC Dropout ile elde edilen $\pm 1\sigma$ belirsizlik bandı ve tahmin güvenilirliği	44
Şekil 4.7. Fold bazında MC Dropout ile elde edilen test belirsizliği değerlerinin karşılaştırılması	45
Şekil 4.8. SHAP tabanlı analizle belirlenen en etkili özneliliklerin görselleştirilmesi (En etkili yüksek %80’lik dilim)	46
Şekil 4.9. Bileşen sayısına göre kümülatif açıklanan varyans.....	47
Şekil 4.10. Fold 5 için Optuna optimizasyon geçmişi.	49

SİMGELER VE KISALTMALAR

Simgeler

R^2 : Coefficient of Determination (Belirleme Katsayısı)

Kısaltmalar

ARIMA : AutoRegressive Integrated Moving Average (Otomatik Regresif Entegre Hareketli Ortalama)
cA : Approximation Coefficients (Yaklaşım Katsayıları, wavelet bileşeni)
cD1/cD2 : Detail Coefficients (Ayrıntı Katsayıları, wavelet bileşenleri)
CNN : Convolutional Neural Network (Evrişimsel Sinir Ağı)
DHI : Diffuse Horizontal Irradiance (Dağınık Yatay Işınım)
DNI : Direct Normal Irradiance (Doğrudan Normal Işınım)
DWT : Discrete Wavelet Transform (Ayrık Dalgacık Dönüşümü)
DWT : Discrete Wavelet Transform (Ayrık Dalgacık Dönüşümü)
LSTM : Long Short-Term Memory (Uzun Kısa Süreli Bellek)
MAE : Mean Absolute Error (Ortalama Mutlak Hata)
MCD : Monte Carlo Dropout (Monte Carlo Dropout, belirsizlik tahmini yöntemi)
MSE : Mean Squared Error (Ortalama Kare Hata)
NN : Sinir Ağı
Optuna : Hiperparametre optimizasyon kütüphanesi
PCA : Principal Component Analysis (Temel Bileşen Analizi)
PCA : Principal Component Analysis (Temel Bileşen Analizi)
PV : Photovoltaic (Fotovoltaik)
RMSE : Root Mean Square Error (Kök Ortalama Kare Hata)
RMSE : Root Mean Square Error (Kök Ortalama Kare Hata)
SHAP : SHapley Additive exPlanations
SHAP : SHapley Additive exPlanations (SHapley Katkı Açıklamaları)

1. GİRİŞ

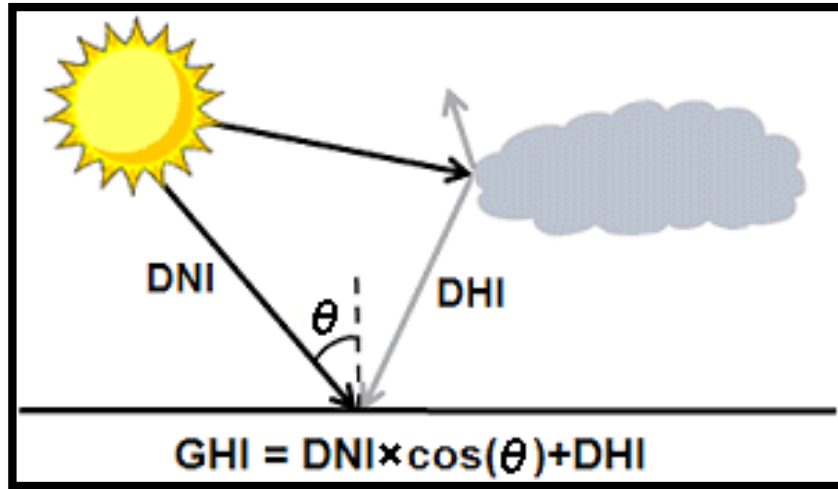
Enerji, dünya genelinde ekonomik ve sosyal kalkınmanın temel unsurlarından biri olarak kabul edilmektedir. Sanayileşme ve kentleşme süreçlerinde itici güç görevi gören enerji, aynı zamanda teknolojik gelişmelerin vazgeçilmez bir bileşenidir (Duffy vd., 2015). Ancak hızla artan küresel nüfus, enerji ihtiyacını önemli ölçüde derinleştirmekte ve endüstrileşme ile teknolojik ilerlemelerin yarattığı talep artışını tetiklemektedir. Bu durum, fosil yakıtlara olan bağımlılığı artırmakta ve çevreye yönelik olumsuz etkileri beraberinde getirmektedir (Abas vd., 2015). Fosil kaynakların kullanımı, küresel ısınma, iklim değişikliği ve hava kirliliği gibi ciddi çevresel sorunlara yol açmaktadır. Bu da temiz ve sürdürülebilir enerji kaynaklarına yönelik küresel ilgiyi artırmaktadır (W. Bach, 1981; Korkmaz vd., 2024; O. Yaman, 2021). Bu bağlamda, yenilenebilir enerji üretimi son yıllarda hızlı bir artış göstermiş olup, uluslararası kuruluşlar tarafından çeşitli politikalarla teşvik edilmektedir. Örneğin, Birleşmiş Milletler ve Avrupa Birliği gibi kurumlar, yenilenebilir enerji entegrasyonunu desteklemek amacıyla karbon vergisi gibi düzenleyici kurallar geliştirmiştir (Z. Ahmed vd., 2022; Elaouzy & El Fadar, 2023; Tsangas vd., 2022).

Mevcut enerji kaynaklarının verimli kullanımı ve tek bir kaynağa bağımlılığın azaltılması, ülkeler için stratejik bir avantaj sağlamaktadır. Küresel enerji talebinin günden güne artması, yenilenebilir enerji kaynaklarına olan ihtiyacı da aynı oranda yükseltmektedir. Bu kaynaklar arasında rüzgâr, güneş, hidroelektrik, biyoyakıt ve yeşil hidrojen enerji teknolojileri öne çıkmakta olup, çevresel sürdürülebilirlik açısından giderek daha fazla önem kazanmaktadır (Aadya Sharma, 2024). Bunlar arasında özellikle güneş enerjisi, kurulum ve kullanım kolaylığı, düşük işletme maliyeti, ekonomik verimlilik ve yaygın erişilebilirlik gibi avantajlarıyla dikkat çekmektedir. Güneş enerjisi, hem termal (ısı) hem de fotovoltaik (elektrik) üretim amacıyla kullanılabilen ve gelecekteki sanayi uygulamalarında (örneğin, ısıtma sistemleri veya güç üretimi) umut vaat eden bir potansiyele sahiptir (Elkelawy vd., 2024).

Yenilenebilir enerji kaynakları içerisinde güneş enerjisi, sınırsız enerji potansiyeline sahip olup çevre dostudur. Bu potansiyelin temelinde, yeryüzüne ulaşan güneş radyasyonu yer almaktadır. Ancak Dünya, bol miktarda güneş radyasyonu

kaynağına sahip olsa da, bu enerjinin kesintili (intermittent) doğası (bulutluluk, mevsimsel değişiklikler ve gece-gündüz döngüsü) doğası, üretimdeki belirsizlikleri artırmakta ve yaygın kullanımını sınırlamaktadır (Pommeret & Schubert, 2022). Bu kesintililik, güneş enerjisinin elektrik şebekesine güvenilir entegrasyonunu zorlaştırmaktadır ayrıca doğru ve sürekli tahmin yöntemlerine olan ihtiyacı doğurmaktadır (Tandon vd., 2025; Wang vd., 2025). Bu nedenlerle literatürde, güneş enerjisi tahminine yönelik ilgi artarak devam etmektedir. Özellikle enerji arz-talep dengesinin sağlanması, şebeke istikrarının korunması ve kaynak optimizasyonu için geleceğe yönelik tahminler kritik rol oynamaktadır (Travieso-González vd., 2024). Bu bağlamda, YZ tabanlı gelişmiş modellerin (klasik MÖ, DL ve hibrit yaklaşımlar) entegrasyonu, güneş enerji sistemlerinin sürdürülebilirliğini artırabilecek bir potansiyel sunmaktadır.

Güneş ışınımı, birim alan başına düşen güç (irradiance, W/m²) olarak tanımlanmakta olup, güneş enerjisinin temel ölçüsüdür. Bu kavram, Küresel Yatay Işınım (GHI), Doğrudan Normal Işınım (DNI) ve dağınık Yatay Işınım (DHI) gibi alt bileşenleri içermektedir (Şekil 1.1). GHI, bir bölgeye düşen toplam güneş enerjisi miktarını gösterir. Buna karşın DNI, özellikle yoğunlaştırılmış güneş enerjisi sistemlerinin performansı için hayati öneme sahiptir (Karaman vd., 2021).



Şekil 1.1. GHI, DNI ve DHI arasındaki geometrik ilişki ve toplam yatay ışınımın hesaplanması (Boutahir vd., 2022)

Güneş ışınımının doğru tahmin edilmesi, şebeke operatörlerinin üretim kalitesini ve kontrolünü iyileştirmesini sağlamaktadır. Ayrıca bu sonuçlar, değişken elektrik üretimine uygun şekilde planlama, zamanlama ve güç düzenlemesini kolaylaştırır (Wan

vd., 2015). Tahminler, enerji üretim planlaması, yönetim ve sistem verimliliği açısından vazgeçilmez olup, güneş panellerinin kapasite belirlenmesi ile arz-talep dengesinde rol oynamaktadır (Krishnan vd., 2023). Tüm bu nedenlerle, güneş ışıınının modellenmesi ve tahmin edilmesi, geleceğin enerji sistemleri için stratejik bir çalışma alanı olarak öne çıkmaktadır.

Güneş ışıınını tahmini, güneş enerjisinin şebekeye etkili entegrasyonu açısından büyük faydalar sunabilir. Literatürde, bu tahmin için çeşitli yöntemler geliştirilmiştir, ancak atmosferik koşullar (bulutluluk, nem, vb), coğrafi konum ve panel açısı gibi faktörler nedeniyle tahminler zorlaşmakta ve karmaşıklaşmaktadır (Geetha vd., 2022; Winster Praveenraj vd., 2024). Bu bağlamda, meteorolojik veriler (hava sıcaklığı, rüzgar hızı vb) ile güneş ışıınını arasındaki ilişkiler analizi edilmelidir (Çelebi & Aslan, 2024). Çünkü veri kümelerinin zaman ve mekânsal değişkenliği tahmin doğruluğunu etkilemektedir (Yunus Khan vd., 2020). Güneş ışıınını tahmini için geleneksel istatistiksel yöntemler (ARIMA modelleri gibi), uzun yıllar kullanılmış olsa da, günümüzde ML tabanlı zaman serisi analizleri istatistiksel yöntemlere bir alternatif olarak öne çıkmaktadır (Gürel vd., 2023). ML, yapay zekânın bir alt kümesi olarak son yıllarda yoğun ilgi gören bir çalışma alanı olup, verilerdeki karmaşık özellikleri yüksek doğrulukla temsil edebilen algoritmalar sunmaktadır. Bu algoritmalar, hava kirliliği izleme, hisse senedi fiyat tahmini ve enerji talebi öngörüsü gibi sıralı zaman serisi problemlerinde yaygın olarak kullanılmaktadır (D. M. Ahmed vd., 2022). Geçmişte zaman serisi sorunlarının çözümü için lineer istatistiksel yöntemler tercih edilirken, günümüzde veri hacmi ve karmaşıklığının artması nedeniyle bu yöntemler yetersiz kalmaktadır. ML tabanlı yaklaşımlar, bu alanda önemli ilerlemeler sağlamış olup, geleneksel yöntemlerin manuel özellik belirleme zorluklarını aşmaktadır (Arseven & Çınar, 2025). Bu nedenle literatürde, klasik ML tabanlı yaklaşımlar (Karar ağaçları, Naive Bayes vb) sıkça önerilmektedir. Ancak bu yöntemler karmaşık veri paternlerini yakalamada hala sınırlı kalmaktadır. Bu soruna çözüm olarak, araştırmacılar YSA ya da YSA türevleri gibi yöntemlerin kullanıldığı tahminlere yönelmiştir. Literatür, YSA tabanlı modellerin enerji üretimi, tüketimi ve rüzgâr hızı gibi zaman serisi tahmin problemlerinde başarılı olduğunu göstermektedir. Ayrıca dalgacık dönüşümü gibi sinyal işleme teknikleriyle entegre edildiğinde tahmin doğruluğu artmaktadır (Gasparin vd., 2022).

Sinir ağı tabanlı tekniklerin gelişimi, uzun süreli veriler arasındaki bağlantı paternlerini daha etkin yakalamayı sağlayabilmektedir. Bu da zaman serisi tahminlerinde

yüksek doğruluk elde edilmesine olanak tanımaktadır. Zira YSA'lar, verilerdeki doğrusal olmayan ilişkileri öğrenme yeteneğine sahiptir (Travieso-González vd., 2024). Bu modellerin büyük veri kümeleri işleyebilmesi, yüksek doğruluklu tahminlere ulaşmayı mümkün kılmaktadır (K. Wang vd., 2019). Ancak, YSA'lar ve YSA türevli yöntemler birçok avantajına rağmen, klasik ML tabanlı yöntemlerine kıyasla bazı zorluklar içermektedir. Bunların başında, optimal hiperparametrelerin belirlenmesi gelmektedir. Hiperparametreler, model eğitimi öncesi kullanıcı tarafından ayarlanan ve öğrenme sürecinde sabit kalabilen ya da güncellenebilen parametrelerdir (öğrenme oranı, katman sayısı veya dropout oranı vb). Hiperparametreler tahmin performansını büyük oranda etkiler. Bu ayarların YSA ya da türevlerindeki komplike yapısı, suboptimal seçimlerde performansı düşürebilmektedir (Ali vd., 2023; Çalışkan, 2023). Bu nedenle güneş ışıını tahmini gibi sinir ağı tabanlı problemlerde hiperparametre optimizasyonu ve model mimarisi tasarımı kritik öneme sahiptir.

1.1. Tezin Amacı

ML son yıllarda popüler bir şekilde çalışılan bir yapay zekâ alt kümesi olup, yüksek seviyelerde etkili özellik temsil edebilen algoritmalara sahiptir. ML tabanlı yöntemler hastalık tanısı, hava kirliliği, hisse senedi ve enerji fiyat aralığı tahmini gibi bir çok sıralı zaman serisi problemlerinde yoğun şekilde kullanılmaktadır (D. M. Ahmed vd., 2022; I. Yaman vd., 2025). ML tabanlı yöntemlerinin kullanımı bu alanlarda çok daha başarılı sonuçlar elde edilmesini sağlamıştır (Chodakowska vd., 2024). Klasik zaman serisi veri analizleri, geçmiş ve şimdiki durum bilgileri için manuel olarak etkin özellikleri belirler. Bu ise zorlu bir süreç olup, istenilen performansın yakalanamamasına neden olabilmektedir. Bu problemlerin çözümü için de sinir ağı tabanlı yöntemler tercih edilmektedir (Ismail Fawaz vd., 2019). Sinir ağı tabanlı yöntemlerin sıralı veriler arasında ilişkileri yüksek seviyede algılayabilmesi sayesinde geleneksel yöntemlerin kullanım alanları zaman içinde azalmaktadır. Özellikle LSTM, Yinelemeli Sinir Ağı (RNN) ve CNN gibi modern sinir ağı tabanlı teknikler, tahmin işlemlerinde daha etkili yöntemler olarak öne çıkmaktadır (T. Zhu vd., 2021). Ancak literatürde, bu tür yeni nesil ML tabanlı yöntemlere odaklanan sınırlı sayıda güneş ışıını tahmin çalışması bulunmaktadır. Bununla birlikte bu yöntemlerin birlikte kullanıldığı hibrit yaklaşımlar şu an popüler bir çalışma alanıdır. Bu tez çalışmasında, güneş ışıını tahmininde kullanılan yeni nesil ML yöntemlerinden yararlanarak, hibrit bir mimari ile kullanılmıştır. Yöntemimiz

literatürdeki bu boşluğu gidermeyi amaçlamaktadır. Bunun için yüksek tahmin doğruluğuna sahip özgün bir framework geliştirilmiştir. Meteorolojik parametrelerin kullanıldığı hibrit modellerin tahmin performansı özellik önem analizi ile birlikte kullanıldığında performansta iyileşme sağlanmıştır (Durap, 2025; Matrenin vd., 2024). Bu çalışmalar, yeni modeller için referans oluşturmakta ve gelecek çalışmalara yön vermektedir.

Tez çalışmasında, DL tabanlı zaman serisi analiz yöntemleriyle güneş ışıınımı tahmini için yüksek doğruluklu bir model geliştirilmiştir. CNN ve LSTM algoritmaları hibrit bir mimaride birleştirilerek, verilerin hem uzamsal (CNN ile özellik çıkarma) hem de zamansal (LSTM ile bağımlılık yakalama) özelliklerinden eşzamanlı olarak yararlanılması sağlanmıştır. Modelin tahmin performansını yükseltmek ve güvenilir sonuçlar elde etmek amacıyla, manuel özellik mühendisliği adımları (Görünür sıcaklık ve DNI/GHI oranı hesaplamaları) uygulanmıştır. Ayrıca DWT (db4 dalgacık ile level=2 ayrıştırma) ile veri zenginleştirilerek hibrit bir yöntem tercih edilmiştir. Modelin yorumlanabilirliğini artırmak için SHAP analizi gerçekleştirilmiştir. Ardından ortalama SHAP değerlerinin önem değeri %80 ve üzerindeki özellikler (%20'lik en etkili özellikler) seçilmiştir. Bu seçilen özelliklerle model yeniden eğitilmiş ve ablasyon çalışmalarıyla (LSTM, dropout katmanları vb) bileşenlerin katkısı doğrulanmıştır. En optimum hiperparametrelerin (dalgacık değeri, filtreler vb) tespiti için Optuna tabanlı Bayesçi optimizasyon yöntemi (20 deneme) ile aranmıştır. Modelin zaman serisi doğasına uygun olarak, değerlendirmeler 5 farklı fold üzerinden zaman serisi bölme yöntemi çapraz doğrulaması ile yapılmıştır. Bu sayede başarı hem güvenilir hem de genelleştirilebilir hale getirilmiştir. Ayrıca, Monte Carlo Dropout (MC Dropout) ile belirsizlik analizi (std_pred) ve PCA ile boyut indirgeme gerçekleştirilerek SHAP tabanlı özellik seçimi güvenilirliği kontrol edilmiştir. Veri seti olarak, Suudi Arabistan'ın Riyad şehrine ait meteorolojik veriler bir API aracılığıyla (Solcast'dan 3 yıllık veriler) alınmıştır. Ardından veri kümesi özellik mühendisliği ve DWT ile genişletilerek kısa vadeli güneş ışıınımı tahminleri yapılmıştır. Geliştirilen optimal model, çeşitli istatistiksel metriklerle (RMSE, R^2 , ortalama belirsizlik) değerlendirilmiştir. Sonuç olarak geliştirilen model sayesinde bir bölgenin güneş enerjisi potansiyeli analiz edilmiş, enerji yatırımlarının ve yönetiminin daha verimli planlanmasına katkı sağlaması amaçlanmıştır. Ayrıca, önerilen yöntem benzer iklim özelliklerine sahip bölgelerde kullanılabilecek esnek ve güvenilir bir çerçeve olarak literatüre kazandırılmıştır. Özet olarak, çalışmada güneş ışıınımı tahmini amacıyla Optuna ile optimize edilmiş, SHAP ile özellik seçimi

yapılan ve ablasyon ile mimari yapısı doğrulanmış yeni bir hibrit DWT-CNN-LSTM modeli önerilmektedir. Çalışmanın temel hedefleri şunlardır:

- (i) Dalgacık temelli sinyal ayrıştırma (frekans bileşenlerini ortaya çıkararak) ve manuel özellik mühendisliği ile giriş verisinin niteliğini artırmak,
- (ii) Hem uzamsal hem de zamansal desenleri yakalayabilen bir CNN-LSTM hibrit mimarisi tasarlamak,
- (iii) Otomatik hiperparametre optimizasyonu kullanarak tahmin performansını en üst düzeye çıkarmak,
- (iv) SHAP analizi ile güneş ışıınımı için en önemli özellikleri tespit etmek ve belirsizlik metrikleriyle model güvenilirliğini doğrulamak.

Bu tez çalışması, gelişmiş sinyal işleme tekniklerini DL yöntemleri ve hiperparametre optimizasyonu bir araya getirmektedir. Böylece yenilenebilir enerji alanında güvenilir, yüksek doğruluklu ve açıklanabilir bir tahmin aracı geliştirmek amaçlanmıştır.

1.2. Literatür Taraması

Güneş ışıınının doğru tahmini, özellikle GHI açısından, yenilenebilir enerji çalışmalarının temel taşlarından biri haline gelmiştir. Bu durum, şebeke kararlılığı, verimli enerji dağıtımı ve PV sistemlerin entegrasyonu gibi ihtiyaçlardan kaynaklanmaktadır (Zhang vd., 2025). Geleneksel istatistiksel ve fiziksel modeller bulut opaklığı, atmosferik değişkenlik gibi faktörlerin etkilediği güneş verilerinin doğrusal olmayan ve durağan olmayan dinamiklerini yakalamada yetersiz kalmaktadır (Papachristopoulou vd., 2024). Örneğin, ARIMA tabanlı yöntemler istikrarlı ve ideal koşullar altında makul doğruluk sağlayabilmektedirler ve özellikle yaz aylarında güneş ışıınım verileri için saatlik tahminlerde yüksek başarılı sonuçlar elde edebilmektedirler. Ancak bu yöntem değişken hava koşullarında ve ani değişiklikleri yönetmede sınırlılık göstermektedir. Kış değişkenliği gibi durumlarda tahmin R^2 değerleri 0.90'larında altında kalabilmektedir (Chodakowska vd., 2024). Sonuç olarak gelecek değerlerin mevcut değerleri yansıttığını varsayan istatistiksel modeller basit temel hatlar olarak hizmet ederler. Ancak karmaşık zamansal bağımlılıkları modelleyemediği için yüksek hatalar üretebilir (Hammond vd., 2024).

ML alanındaki ilerlemeler, bu zorlukları daha iyi ele alan veri odaklı tekniklere odaklanmayı sağlamıştır (Chu vd., 2024). Destek Vektör Makineleri (SVM) ve Rastgele

Ormanlar (RF) gibi yöntemler kısa vadeli güneş ışıyım tahminlerinde umut verici sonuçlar göstermiştir. Ayrıca SVM varyantlı algoritmalar hava sıcaklığı ve rüzgar hızı gibi meteorolojik girdileri dahil ederek istatistiksel ve fiziksel modellere kıyasla RMSE sonuçlarında iyileştirmeler ve azalmalar sağlayabildiği görülmüştür (Aouidad & Bouhelal, 2024; Çelebi & Aslan, 2024). Bölgesel bir çalışmada, Kocaeli ilinde yatay düzleme gelen aylık ortalama günlük global güneş ışıyımını tahmin etmek için literatürdeki 30 regresyon modeli değerlendirilmiştir. 1973-2006 yılları arasında Meteoroloji Genel Müdürlüğü verileriyle yapılan analizde, yaz aylarında (Haziran-Ağustos) tahmin hatalarının yüksek olduğu tespit edilmiştir. Modeller, lineer, ikinci ve üçüncü dereceden polinom denklemleri içerecek şekilde tasarlanmıştır. Performans için MABE, MBE, MAPE, MPE ve RMSE metrikleri kullanılmıştır. Daha doğru sonuçlar için bölgeye özgü yeni regresyon modelleri önerilmiştir (Arslanoğlu, 2016). Ancak bu modeller sıralı verilerle başa çıkmada zorlandığı anlaşılmıştır ve bu durum sinir ağıları tabanlı mimarilerin benimsenmesini teşvik etmiştir (Malalha vd., 2024).

Yapılan bir çalışmada, PV modüllerinin maksimum güç üretimini tahmin etmek için YSA tabanlı bir yaklaşım önerilmiştir. Modelde giriş olarak güneş ışıyımı, sıcaklık ve rüzgâr hızı verileri kullanılmıştır. Çıkış olarak ise çevresel koşullar altında elde edilecek maksimum güç öngörülmüştür. Yöntemin performansı, günlük, aylık ve yıllık gerçek verilerle değerlendirilmiş olup, geleneksel çoklu doğrusal regresyon modellerine kıyasla daha yüksek doğruluk sağladığı görülmüştür (Hiyama & Kitabayashi, 1997). Diğer bir çalışmada, monokristalin PV modüllerinin güç üretimini tahmin etmek için YSA tabanlı bir model geliştirilmiştir. Çevresel veriler ve bir yıllık üretim ölçümleri kullanılarak, herhangi bir güne ait güç değeri %95 doğrulukla öngörülmüştür. Geliştirilen model, polinomal regresyon, çoklu doğrusal regresyon, analitik ve tek diyot modellerinden daha üstün performans göstermiş olup, zaman serisi analizi ile sinir ağı tabanlı tahminlerin etkinliğini vurgulamıştır (Mellit & Pavan, 2010). Bölgesel diğer bir çalışmada, Türkiye'nin 125 farklı bölgesinde fotovoltaiik santrallerin elektrik gücü üretimini tahmin etmek için ML tabanlı yöntemler test edilmiştir. YSA, Çoklu Doğrusal Regresyon (ÇDR) ve k-En Yakın Komşu Regresyonu (k-EYKR) kullanılmış, yedi bağımsız değişkenle güç tahmini yapılmıştır (Uğuz vd., 2019).

Klasik zaman serisi analizleri, geçmiş ve şimdiki veriler arasındaki ilişkileri yakalamada sınırlı kalırken, sinir ağı tabanlı yöntemler bu problemleri etkili bir şekilde çözmektedir (S.-C. Lim vd., 2022). Sinir ağı tabanlı modeller, sıralı verilerdeki geçici bağımlılıkları yüksek seviyede algılayabildiği için geleneksel yöntemlerin kullanım

alanları giderek azalmaktadır. DL tabanlı modellerin zaman serisi analizindeki başarıları, birçok araştırmacı bu alana yöneltmiştir (Gasparin vd., 2022). Bu modeller, enerji üretimi-tüketimi ve rüzgâr hızı tahmininde etkili olup, yüksek doğruluk sağlamaktadır (Ifaei vd., 2023).

Yapılan DL tabanlı bir çalışmada LSTM yöntemi ile günlük ısınım tahmini yapılmıştır. Önerilen yöntem Karar Ağacı Regresyonu, RF, Gradyan Güçlendirme ve K-En Yakın Komşu yöntemleri ile kıyaslanmıştır. Yöntemi test etmek için Temmuz 1983-Aralık 208 tarihleri arasında Çorum ilinin doğrusal olmayan ve karmaşık güneş ısınım değerleri kullanmıştır. Sonuçlar bir RNN modeli uzantısı olan LSTM tabanlı zaman serisi analizinin, diğer yöntemlerden daha iyi sonuç elde edilebildiğini göstermiştir. Veri seti olarak Amerika'nın Ulusal Havacılık ve Uzay Dairesi (NASA) tarafından sunulan Prediction of Worldwide Energy Resource (POWER) hizmetinden faydalanılmıştır. Veri setinde toplam 12969 adet veriden yararlanılmıştır (Kara, 2019). LSTM ağları zaman serisi bağımlılıklarını modellemede iyi performans gösterebilmektedirler ve tarihsel ısınım veri setleri üzerinde eğitildiğinde tahmin işlemlerinde $R^2 > 0.95$ değerlerine ulaşabilmektedirler (Li vd., 2023). Bununla birlikte, LSTM'ler çok değişkenli girdilerde uzamsal paternleri göz ardı edebilir ve bu da çoklu özellik senaryolarında suboptimal performansla sonuçlanabilir (L. Zhu vd., 2025).

Tezimizde gerçek zamanlı meteorolojik veriler (Riyad lokasyonu) kullanılmış olup, gerçek hayat uygulamalarını da desteklemektedir. Özellikle LSTM, geçmiş bağımlılıkları hatırlayan bir Yinelemeli Sinir Ağı (RNN) varyantı, RNN ve CNN, konvolüsyonel katmanlarla özellik çıkarma, gibi teknikler güneş ısınım tahmininde daha etkili sonuçlar vermektedir (Mishra & Palanisamy, 2019). Hibrit DL modelleri, güneş ısınım tahmininde hem uzamsal hem de zamansal özellikleri yakalamada baskın bir paradigma olarak ortaya çıkmaktadır (Hong vd., 2020). CNN-LSTM hibritlerinde, CNN katmanlarını yerel patern çıkarma ve LSTM ise sıralı modelleme için kullanılabilir. Bu da daha yüksek doğruluk elde edilmesine yardımcı olabilir (Çevik Bektaş & Altaş, 2024). Bir çalışmada Cezayir lokasyonuna ait verilerden özellik çıkarma için CNN modeli, zaman serisi analiz için ise LSTM kullanılmıştır. Model günlük ısınım tahminlerde normalize RMSE 0.0833 ve R^2 0.99 değerlerine ulaşmıştır. Ancak çalışmada hiperparametre ayarlaması ve açıklanabilirlik eksik kalmıştır (Ladjal vd., 2025). Benzer şekilde, Ekvador lokasyonu için 1–12 saat aralığı için GHI tahminleri yapılmıştır. Çalışmada Gürültü ile Tamamlayıcı Topluluk Ampirik Mod Ayırıştırması (CEEMDAN)-CNN-ATT-LSTM hibrit modeli geliştirmiştir. Mevsimler arası RMSE değerlerini 135–

149 W/m² olarak rapor edilmiştir. Dikkat mekanizmaları DNI gibi ana değişkenlere odaklanmayı iyileştirmiş, ancak yöntem sistematik optimizasyon veya belirsizlik nicelleştirmesi olmadan yapılmıştır (Camacho vd., 2025). Bu çalışmalar hibrit yöntemler durağan olmayan verileri yönetmede avantajlı olduklarını göstermiştir (Saf LSTM ve ya da CNN'lere kıyasla RMSE iyileştirmesi görüşmüştür). Ancak tüm bu çalışmalar manuel parametre seçimi ve kara kutu karar verme süreçleri açısından sınırlılıkları vardır (Zhou vd., 2024).

Hiperparametre optimizasyonu, model ayarlarını otomatikleştirerek uygun parametrelerin seçimini kolaylaştıran bir yöntemdir. Bayesçi yaklaşımlar kullanan Optuna gibi kütüphaneler, geniş arama alanlarında verimli çözümler sunmaktadır (Akiba vd., 2019). Literatürde metaheuristik algoritmalar da yaygın biçimde kullanılmıştır. Örneğin, özellik faktörizasyonuna dayalı bir CNN-LSTM modeli, metaheuristik optimizasyon ile geliştirilmiş ve Tokyo'da saatlik güneş ışıınımı tahmini için RMSE 0.183 elde etmiştir (D. Wang vd., 2025). Benzer biçimde, günlük PV tahminlerinde CNN-SLSTM modelinin hiperparametreleri Slime Mould Optimization (SMO) yöntemi ile ayarlanmış ve $R^2 = 0.97$ değeri elde edilmiştir. Bununla birlikte, bu çalışma sınırlı sayıda örnek (~896 veri) üzerinde yapılmış ve herhangi bir açıklanabilirlik tekniği kullanılmadığı için yorumlanabilirlik açısından zayıf kalmıştır (Mariappan vd., 2025). Hiperparametre optimizasyonuna ek olarak, sinyal ayrıştırma yöntemleri de model performansını artırmak için kullanılmaktadır. Örneğin, DWT ve CEEMDAN gürültüyü azaltmak ve çok ölçekli özellikleri yakalamak amacıyla entegre edilmiştir. Özellikle DWT, zaman serilerini yaklaşık (düşük frekanslı eğilimler) ve detay (yüksek frekanslı dalgalanmalar) bileşenlerine ayırarak modelin sağlamlığını artırdı (Jin vd., 2025). Benzer şekilde, Türkiye ve İsviçre'de yapılan bir çalışmada DWT-BiLSTM yaklaşımı uygulanmış, sinyaller dalgacık seviyesi 6'ya kadar ayrıştırılmış ve RMSE = 53.20 W/m² ile MAE = 27.67 W/m² değerleri elde edilmiştir. Ancak çalışmada SHAP gibi açıklanabilir yapay zekâ teknikleri uygulanmamış, bu durum modelin yorumlanabilirliğini sınırlamıştır (Çevik Bektaş & Altaş, 2024). Bu bağlamda, XAI yorumlamalı DL tabanlı teknikler güneş ışıınımı tahminde literatüre önemli katkı sunacaktır. Sonuç olarak literatürde, DL tabanlı yöntemlere odaklanan güneş ışıınımı tahmin çalışmaları sınırlıdır. Özellikle bu yöntemlerin birleşimin oluşan hibrit modellerin kullanımı oldukça popüler bir ve yenilikçi teknikler olarak öne çıkmaktadır. Bu nedenle hibrit yaklaşımlar (örneğin CNN-LSTM entegrasyonu) ile yapılacak çalışmalar literatüre katkı sunacaklardır.

XAI yöntemleri, SHAP LIME vb, enerji tahmininde model şeffaflığını artırmak amacıyla giderek daha fazla kullanılmaktadır (Van Zyl vd., 2024). Örneğin bir çalışmada SHAP tabanlı özellik seçimi ve Gauss gürültü artırma tekniklerini entegre eden hibrit bir DL modeli geliştirmiştir. Bu yaklaşım, CNN, GRU, LSTM ve CNN-GRU gibi temel modeller üzerinde önemli performans artışları sağlamış ve aynı zamanda eğitim süresini kısaltmıştır (Thipwangmek vd., 2024). Bununla birlikte, söz konusu modelin genellenebilirliği ve güvenilirliği yeterince doğrulanmamıştır. Literatürde eksik kalan bu yön, Monte Carlo Dropout kullanılarak belirsizlik nicelleştirmesi ile tamamlanabilir. Bu yöntem, kalibre edilmiş güven aralıkları sunarak model güvenilirliğini artırmaktadır (Gal & Ghahramani, 2016).

Yukarıda da değerlendirmelere göre güncel literatürdeki ilerlemelere rağmen hala boşluklar devam etmektedir. Genel olarak bakıldığında çoğu hibrit mimari parametrelerinin optimizasyonunu (Optuna yoluyla dalga tipi/seviyesi adaptasyonu gibi), uyarlanabilir özellik seçimi için XAI kullanımı ve model doğruluk analizini -belirsizliğini ihmal etmektedir. Çalışmamız, DWT yardımıyla özellik zenginleştirmekte ve Bayesçi bir yaklaşımla DWT-CNN-LSTM mimari optimizasyonu yapmakta, en etkili %80'lik yüzdelik seçimi için SHAP kullanmakta ve MC Dropout ile model sağlamlığını test etmektedir. Böylece uçtan uca bir boru hattı önererek literatürdeki boşlukları gidermektedir. Ayrıca yüksek doğruluk oranı ile öne çıkmaktadır. Modelimiz 5 fold çapraz doğrulamada ortalama RMSE 6.62 W/m² ve R² 0.98 elde etti.

Literatür genel olarak değerlendirildiğinde, DL tabanlı modeller güncel bir çalışma alanı olup birçok araştırmacı bu alana yönelmiştir (Aslan & Çelebi, 2022; Gasparin vd., 2022). Böylelikle DL tabanlı modeller, enerji üretimi-tüketimi ve rüzgâr hızı tahmini gibi çeşitli zaman serisi alanlarına uygulanmıştır (Ifaei vd., 2023). Bu yöntemlerin zaman serileri analizindeki başarıları göz önüne alındığında, enerji tahmini için yüksek doğrulukta sonuçlar elde edilmesi beklenir. Bu bilgiler doğrultusunda DL tabanlı tekniklerle güneş ışıınımı tahmini literatüre önemli bir katkı sağlayacaktır. Ayrıca tez çalışmamızda gerçek veriler kullanılmıştır. Zamansal veri analizi ile güneş ışıınının gerçek veriler ile tahmin edilmesi gerçek zamanlı uygulamaların kullanımını büyük ölçüde destekleyecektir (Z. Wang vd., 2025) Sonuç olarak, yapılan literatür taraması DL tabanlı hibrit modellerin güneş ışıınımı tahmininde başarılı sonuçlar verebildiğini ve bu modellerin özellikle meteorolojik parametrelerin entegrasyonu ve özellik önem analizleriyle daha da güçlendirilebileceğini göstermektedir. Bu çalışmalar, geliştirilecek yeni modeller için önemli bir referans oluşturmakta ve gelecekte yapılacak çalışmalara

yol göstermektedir. Yapılan tez çalışmasında, gerçek zamanlı meteoroloji verileri kullanılarak kısa vadeli ısıtım tahmini yapılması için etkili bir hibrit yöntem önerilmektedir. Bununla birlikte çalışmada literatürdeki ham meteoroloji verilerden başka manuel özellik mühendisliği ile yeni özellikler üretildi. Ardından DWT ile veri kümesi genişletildi. DWT zaman serisi verisindeki gizli desenleri ve farklı frekans bileşenlerini ortaya çıkarmak için kullanıldı. Bu sayede, modelin öğrenebileceği daha anlamlı ve ayrıntılı özellikler elde edildi.

Tezin ilerleyen bölümlerinde, güneş enerjisi ile ilgili teorik bilgiler 2. bölümde sunulmaktadır. Kullanılan yöntemlerin ve veri setlerinin ayrıntılı açıklaması 3. bölümde verilmektedir. 4. bölümde ise yapılan analizlerin sonuçları ile elde edilen bulgular kapsamlı bir şekilde değerlendirilmektedir. Son bölümde ise, tezde elde edilen bulguların genel bir değerlendirmesi, sınırlılıklar ve gelecek çalışma önerileri yer almaktadır.

2. GÜNEŞ ENERJİSİNİN TEMELLERİ VE UYGULAMALARI

2.1. Güneş Enerjisi

Güneş, Dünya'nın da içinde yer aldığı güneş sisteminin merkezinde konumlanmakta olup, sistemin kütesinin yaklaşık %99.8'ini oluşturmaktadır (E. Hossain, 2023). Bu devasa kütle, esas olarak hidrojen ve helyum gazlarından meydana gelmekte olup, güneşin çekirdeğinde gerçekleşen nükleer füzyon reaksiyonları sayesinde muazzam miktarda enerji üretmektedir. Özellikle, hidrojen atomlarının helyuma dönüşümü sırasında açığa çıkan bu enerji, güneş ışıınımı olarak adlandırılmakta ve uzay boşluğunda yaklaşık 1366 W/m^2 yoğunluğa sahip bir solar sabit (solar constant) ile yayılmaktadır (Hirshberg, 1973). Bu ışıınım, Dünya atmosferine ulaşmadan önce çeşitli süreçlerden geçse de yeryüzüne varan miktarı bile insanlığın küresel enerji tüketimini kat kat aşacak potansiyele sahiptir. Nükleer enerji kaynakları dışında, fosil yakıtlar, rüzgar, hidrolik ve biyokütle gibi tüm yenilenebilir enerji formlarının doğrudan veya dolaylı kaynağı güneştir (Zaidi, 2018). Güneşin yüzey tabakası olan fotosferin sıcaklığı yaklaşık 5800 K olup, bu sıcaklıkta madde kısmen iyonize plazma halinde bulunmakta ve ışıınının spektrumunu belirlemektedir (İbrahim, 2015). Güneş çekirdeğinden yayılan bu ışıınlar, yüksek enerji potansiyeli taşıyarak Dünya'ya ulaşmakta; ancak, atmosferik absorpsiyon, saçılma ve yansıma nedeniyle enerjinin yaklaşık %60-70'i uzayda veya atmosferde kaybedilmektedir (NASA Earth Observatory, 2019). Bu bağlamda, güneş ışıınlarından elde edilen enerjinin etkin kullanımı, sürdürülebilir enerji stratejileri açısından kritik öneme sahiptir. 1970'lerden itibaren dünya genelindeki teknolojik ilerlemeler güneş enerjisi üretim maliyetlerini önemli ölçüde düşürmüştür. Örneğin, 2023 itibarıyla güneş fotovoltaik Sistemlerinin Seviyelendirilmiş Enerji Maliyeti (LCOE) %89 oranında azalmıştır (Özgür, 2021). Bu gelişmeler sayesinde, güneş enerjisi sınırsız bir kaynak potansiyeli olarak daha verimli bir şekilde değerlendirilmektedir.

2.2. Güneş Radyasyon Çeşitleri

Güneş ışıınımı, Güneş'ten yayılan elektromanyetik enerji olarak tanımlanmakta olup, Dünya atmosferine ulaşan en büyük doğal enerji kaynağıdır. Bu enerji, geniş bir

dalga boyu spektrumunda dağılmakta ve 0.1 nm ile 1000 μm arasında değişen aralıkta yoğunlaşmaktadır. Özellikle ultraviyole (UV: 0.1-0.4 μm), görünür ışık (0.4-0.7 μm) ve kızılötesi (IR: 0.7-1000 μm) bantlarında belirgindir (Moan, 2001). Dünya yüzeyine ulaşan güneş ışımasını, üç ana bileşenden oluşmakta olup, bu bileşenler enerji sistemlerinin tasarımı ve performans analizinde temel rol oynamaktadır. Bunlar sırasıyla Doğrudan (direkt), dağınık (difüz) ve yansıyan radyasyon (Moan, 2001) olarak karşımıza çıkmaktadır (Kilic, 2016).

2.3. Direkt (Doğrudan) Güneş Radyasyonu

Doğrudan güneş radyasyonu, güneşten herhangi bir engel, yansıma veya saçılma olmaksızın Dünya yüzeyine ulaşan ışımasını ifade etmekte olup, atmosferik etkilerden en az etkilenen bileşendir. Bu radyasyon, atmosfer veya diğer nesneler tarafından soğurulmayan, yansıtılmayan ve dağılmayan güneş ışığının doğrudan yeryüzüne varan kısmıdır; bu nedenle, güneş enerjisinin en yoğun ve konsantre formunu oluşturmaktadır olup, konsantre güneş enerjisi (CSP) sistemlerinde kritik rol oynamaktadır (Wald, 2009). Doğrudan radyasyon, yeryüzüne önemli miktarda enerji taşıyarak doğrudan ısıtma, aydınlatma ve fotovoltaik dönüşüm süreçlerini desteklemekte; ancak, bulut örtüsü, aerosol yoğunluğu ve zenith açısı gibi faktörler nedeniyle günlük ve mevsimsel değişkenlik göstermektedir (Thekaekara, 1976). Bu değişkenliklerin anlaşılması, güneş enerjisi üretiminin verimliliğini artırmak, iklim modellerini geliştirmek ve radyasyon ölçüm stratejilerini optimize etmek açısından büyük önem arz etmektedir.

Doğrudan güneş radyasyonu, pirheliometre adı verilen özel bir cihazla ölçülmekte olup, bu cihaz güneşi sürekli olarak doğru bir açıyla izleyerek hassas ölçümler yapmaktadır (Lee vd., 2009). Pirheliometre, yatay düzleme gelen doğrudan radyasyon şiddetini ve metrekaare başına düşen enerji miktarını W/m^2 cinsinden belirlemekte; içindeki sensör, alınan radyasyona göre voltaj üreterek bu veriyi kaydetmektedir (Vignola vd., 2019). Bu ölçümler, güneş enerjisi tahmin modellerinde temel veri kaynağı olarak kullanılmakta olup, atmosferik etkilerin (optik derinlik vb) modellenmesini sağlamaktadır.

2.4. Dağınık (Difüz) Güneş Radyasyonu

Doğrudan güneş radyasyonunun aksine, dağınık güneş radyasyonu, atmosferde moleküller, aerosoller ve bulutlar tarafından saçılarak farklı yönlerde dağılan ve yeryüzüne

ulaşan enerji bileşenini ifade etmektedir. Bu süreç, Rayleigh ve Mie saçılmaları gibi fiziksel mekanizmalarla gerçekleşmekte olup, radyasyonun yoğunluğunu azaltarak daha geniş bir alana yayılmasına neden olmaktadır (LIU & JORDAN, 1993). Atmosferik faktörler, örneğin bulut örtüsü, hava kirliliği, yağmur ve kar, dağınık radyasyonun oluşumunda temel rol oynamakta; bu faktörler, güneş ışınlarının doğrudan yolundan sapmasına yol açarak, bulutlu günlerde toplam radyasyonun büyük kısmını dağınık bileşenin oluşturmaya sebep olmaktadır. Dağınık radyasyon, doğrudan radyasyona kıyasla daha düşük yoğunlukta olup, yumuşak gölgeler ve düşük kontrastlı eşit aydınlatma sağlamakta; bu özellik, güneş enerjisi tahmin modellerinde (örneğin, bulut opaklığı gibi değişkenlerle modellenen sistemlerde) kritik öneme sahiptir. Dağınık güneş radyasyonu, gölge bantlı piranometre cihazı ile ölçülmekte olup, bu cihaz doğrudan radyasyonu engelleyerek sadece dağınık bileşeni kaydetmektedir (Pearcy, 2000). Çalışma kapsamında, dağınık radyasyonun modellenmesi, veri işleme adımlarında dalgacık dönüşümü gibi tekniklerle gürültü azaltma için kullanılmakta olup, bu, atmosferik saçılmanın tahmin doğruluğunu artırmaktadır.

2.5. Yansıyan Güneş Radyasyonu

Yansıyan ışınım ise, yüzeyden geri yansıyan radyasyon miktarını belirtmekte olup, albedo (yansıma katsayısı) ile doğrudan ilişkilidir. Albedo değeri kar, su veya açık renkli zeminlerde yüksek olduğundan (örneğin, kar için 0.8-0.9), yansıyan radyasyon artmaktadır (Gates, 1980).

2.6. Global (Küresel) Güneş Radyasyonu

Küresel radyasyon, toplam güneş radyasyonu olarak da adlandırılmakta olup, yeryüzüne ulaşan doğrudan ve dağınık radyasyon bileşenlerinin toplamını temsil etmektedir. Bu kavram, güneş enerjisi sistemlerinin tasarımı ve performans analizinde temel bir parametre olup, yatay bir yüzeye gelen toplam enerji miktarını ifade etmektedir. Matematiksel olarak, yatay küresel radyasyon GHI denklem (2.1) ile hesaplanmaktadır (Qiu & Li, 2019):

$$GHI = DNI \times \cos(\theta) + DHI \quad (2.1)$$

Burada DNI doğrudan normal radyasyon, θ güneş geometrisini (zenith açısı), ve DHI dağınık yatay radyasyon değerlerini temsil etmektedir (Korachagaon & Bapat,

2012). Bu denklem, zenith açısı dikkate alarak doğrudan bileşenin yatay projeksiyonunu hesaplamaktadır. Dikey yüzeyler için ise ek eğim faktörleri (tilt açısı gibi) dahil edilmektedir. Küresel radyasyonun ölçü birimi, metrekare başına watt (W/m^2) olarak ifade edilmekte olup, bu değer meteorolojik verilere, güneş geometrisine ve atmosferik parametrelere (aerosol optik derinliği, su buharı vb) dayalı karmaşık matematiksel modellerle tahmin edilmektedir (Badescu, 2008). Gerçek küresel radyasyon hesaplamaları, güncel yöntemler içermekte olup bu modeller uzay tabanlı solar sabitten başlayarak atmosferik kayıpları simüle etmektedir (Awad vd., 2018). Çalışma kapsamında GHI, solar enerji tahmininde temel girdi olarak işlenmekte olup, veri ön işleme adımlarında standartlaştırma ve sekans oluşturma ile model mimarisine entegre edilmektedir.

2.7. Yoğunlaştırıcı Termal Sistemler

Yoğunlaştırıcı Termal Sistemler (CSP), güneş ışınlarını aynalar veya mercekler yardımıyla tek bir odak noktası veya hattına toplayarak yüksek sıcaklıklara ulaşmayı sağlayan teknolojilerdir. Bu sistemler, doğrudan radyasyonun yoğunlaştırılmasına dayalı olup, elde edilen ısı enerjisini buhar türbinleri, Stirling motorları veya süperkritik CO_2 çevrimleri gibi güç dönüştürme mekanizmalarıyla elektriğe çevirmektedir (Carlqvist, 2009). CSP'nin temel avantajı, termal depolama kapasitesidir. Örneğin, erimiş tuz veya faz-değişim malzemeleri kullanılarak ısı depolanabilmekte ve bu sayede gece veya bulutlu dönemlerde sürekli üretim sağlanabilmektedir. Sistem türleri arasında parabolik trough (hattı odak), solar tower (nokta odak) ve Fresnel reflektörler öne çıkmakta olup, bunlar yüksek verimlilik (%20-30) ve ölçeklenebilirlik sunmaktadır (Şekil 2.1).



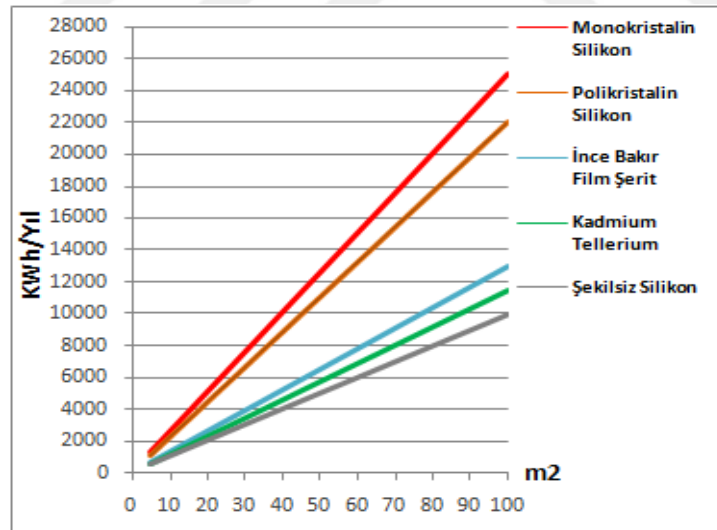
Şekil 2.1. Parabolik oluk tipi yoğunlaştırıcı güneş enerjisi sistemi (Montanes vd., 2017)

2.8. Fotovoltatik Güneş Teknolojileri

PV güneş hücreleri, güneş ışığının elektrik enerjisine dönüştürülmesinde temel rol oynamakta olup, mevcut güneş enerjisinden maksimum elektrik üretimi sağlamayı amaçlamaktadır. Bu teknolojinin temel fiziki, yarı iletken malzemelerdeki foton absorpsiyonu ve yük taşıyıcı ayrımı üzerine kuruludur; ancak, Shockley-Queisser sınırı gibi kısıtlamalar verimliliği sınırlamakta ve malzeme sentezi zorluklarını artırmaktadır (Polman vd., 2016).

2.9. Güneş Paneli (Fotovoltaik Modül) Çeşitleri

PV paneller, güneş enerjisinin elektrik dönüşümünde kritik bir rol oynamakta olup, hücre malzemesi ve üretim süreçlerindeki gelişmeler modül verimliliğini artırırken maliyetleri düşürmektedir. Şekil 2.2, Türkiye’de farklı PV panel tiplerine göre alan-bazlı yıllık elektrik üretimini (kWh/m^2) göstermekte olup, monokristalin tipler ortalama 1800 kWh/m^2 , polikristalin 1600 kWh/m^2 civarında üretim sağlamaktadır (Enerji ve Tabii Kaynaklar Bakanlığı, Türkiye, 2023).



Şekil 2.2. Türkiye’de Farklı PV Panel Tiplerine Göre Alan-Bazlı Yıllık Elektrik Üretimi (kWh/m^2) (Enerji ve Tabii Kaynaklar Bakanlığı, Türkiye, 2023)

Bu karşılaştırma, panel tipi seçiminde alan başına yıllık üretimin doğrudan etkili olduğunu vurgulamakta olup, özellikle sınırlı çatı alanı veya yüksek verim gerektiren uygulamalarda monokristalin silikon teknolojisi ön plana çıkmaktadır. Ancak, maliyet,

iklim dayanıklılığı ve yatırım geri dönüş süresi gibi ek faktörlerin de dikkate alınması gerekmektedir (Mitrašinović, 2021).

2.10. Gelişen PV Teknolojileri

Son yıllarda PV pazarında teknolojik gelişmeler hız kazanmış olup, özellikle n-tip Monokristal TOPCon (Tunnel Oxide Passivated Contact) paneller, yüksek verimlilikleri (%25+) ve uzun vadeli performans avantajları sayesinde sektörde ön plana çıkmaktadır. TOPCon modüller 2025 yılı sevkiyatlarının yaklaşık %70'ini oluşturarak p-tip PERC teknolojisinin önüne geçmiştir (Tiismus vd., 2025). Bu yükselişin başlıca sebepleri şunlardır: Yüksek modül verimi ve bifacial (çift yüzlü) kazanç, seviyelendirilmiş elektrik maliyetini (LCOE) önemli ölçüde düşürmekte; üretim altyapısının dönüşümü, PERC hatlarını düşük maliyetle TOPCon'a uyarlamakta; ışık kaynaklı bozunmaya (LID) direnç, n-tip wafer'ların dayanıklılığını artırmakta; büyük üreticilerin (LONGi, Jinko, Trina) GW ölçekli geçişi sektörü yönlendirmektedir (European Commission. Joint Research Centre., 2024).

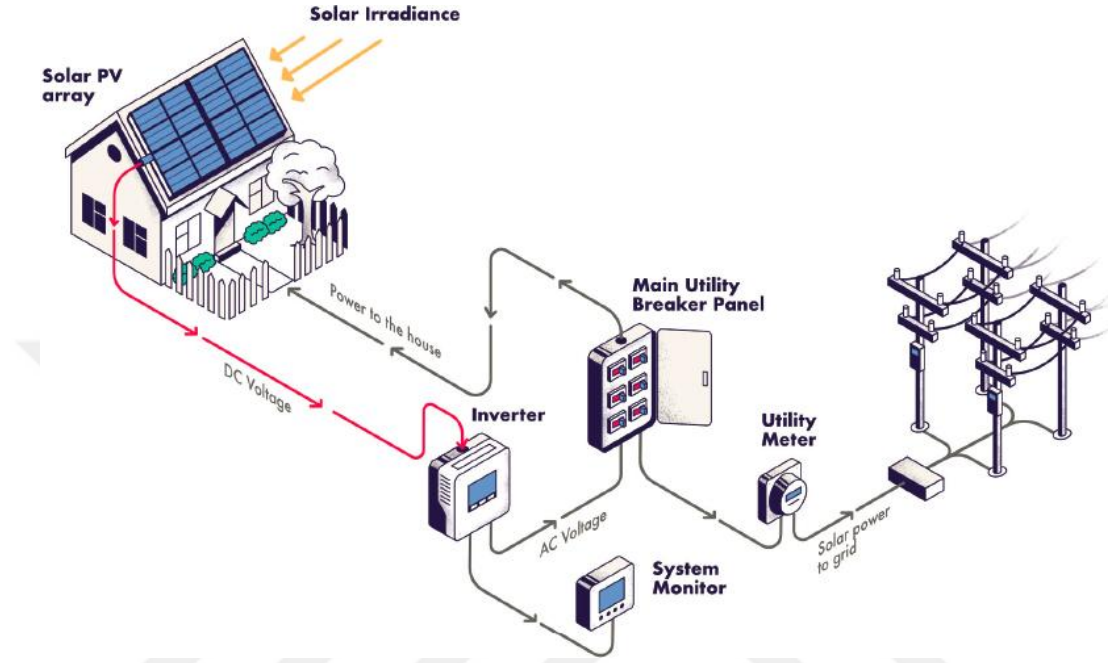
2.11. Yükselen Teknolojiler

Fotovoltaik alanda sadece TOPCon değil, farklı ileri teknolojiler de geliştirilmekte olup, Araştırma ve Geliştirme (Ar-Ge) çalışmaları yeni hücre tasarımlarının ticari pazara girişini hızlandırmaktadır. Öne çıkan yükselen teknolojiler şu şekilde özetlenebilir: Silikon-Perovskit Tandem hücreler, laboratuvar ortamında %34.85'in üzerinde verimlilik değerlerine ulaşmış olup, bu teknolojinin 2026-2027 yıllarında GW ölçekli üretim hatlarına geçmesi beklenmektedir (European Commission. Joint Research Centre., 2020). Tam Perovskit, Organik ve III-V hücreler, özellikle uzay uygulamalarında ve Ar-Ge çalışmalarında öne çıkmakta; bakır metallizasyon ise gümüş tüketimini azaltmak amacıyla hücre üretiminde güçlenmekte olup, maliyet düşüşü sağlamaktadır (Murarka & Hymes, 1995).

2.12. Şebeke Bağlantılı Sistemler

Şebeke bağlantılı (on-grid) sistemler, PV dizilerinin ürettiği Doğru Akımı (DC), şebeke ile senkron çalışabilen bir inverter aracılığıyla Alternatif Akıma (AC) çevirerek doğrudan dağıtım şebekesine besleyen, depolama zorunluluğu olmayan yenilenebilir

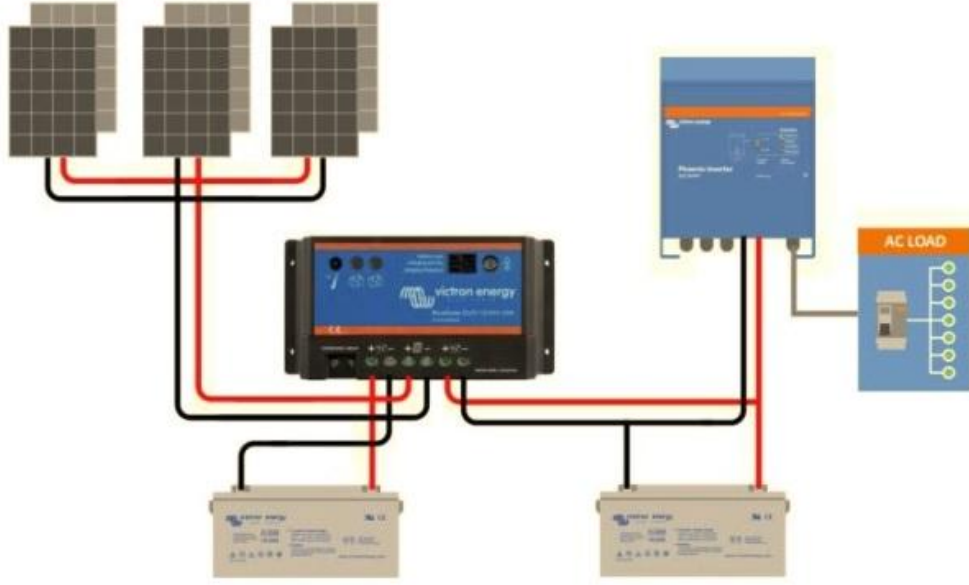
enerji kurulumlarıdır. Bu sistemlerde kullanıcı, üretim fazlasını şebekeye satar veya mahsuplaştırır. Üretimin yetersiz kaldığı anlarda enerjiyi şebekeden çeker (Şahin vd., 2022). Avantajları arasında düşük maliyet, kolay entegrasyon ve yüksek verimlilik yer almaktadır. Dezavantajları ise şebeke bağımlılığı ve kesinti riskidir.



Şekil 2.3. Şebeke bağlantılı (on grid) sistemler (National Academies of Sciences, Engineering, and Medicine, 2023)

2.13. Şebeke Bağlantısız Sistemler

Şebeke bağlantısız (off-grid) fotovoltaik sistemler, genel dağıtım şebekesine erişimi olmayan veya enerji bağımsızlığı hedefleyen yapılara kesintisiz elektrik sağlamak için PV modüllerini batarya depolaması ve akü yönetimiyle birleştirir (Şekil 2.4). Enerji, bir şarj regülatörü üzerinden akü bankasına aktarılır. Inverter AC yükleri besler. Sistemin güvenilirliği, “özerklik günleri (battery autonomy)” ve tasarım derin deşarj oranlarıyla belirlenir (Zenk, 2018). Avantajları enerji özerkliği ve uzak bölgeler için uygunluk; dezavantajları yüksek batarya maliyeti ve bakım gereksinimidir (Zenk, 2018).



Şekil 2.4. Şebeke bağlantısız (off grid) sistemler (Victron Energy BV, 2023)

2.14. Dünyada Güneş Enerjisi Kullanımı

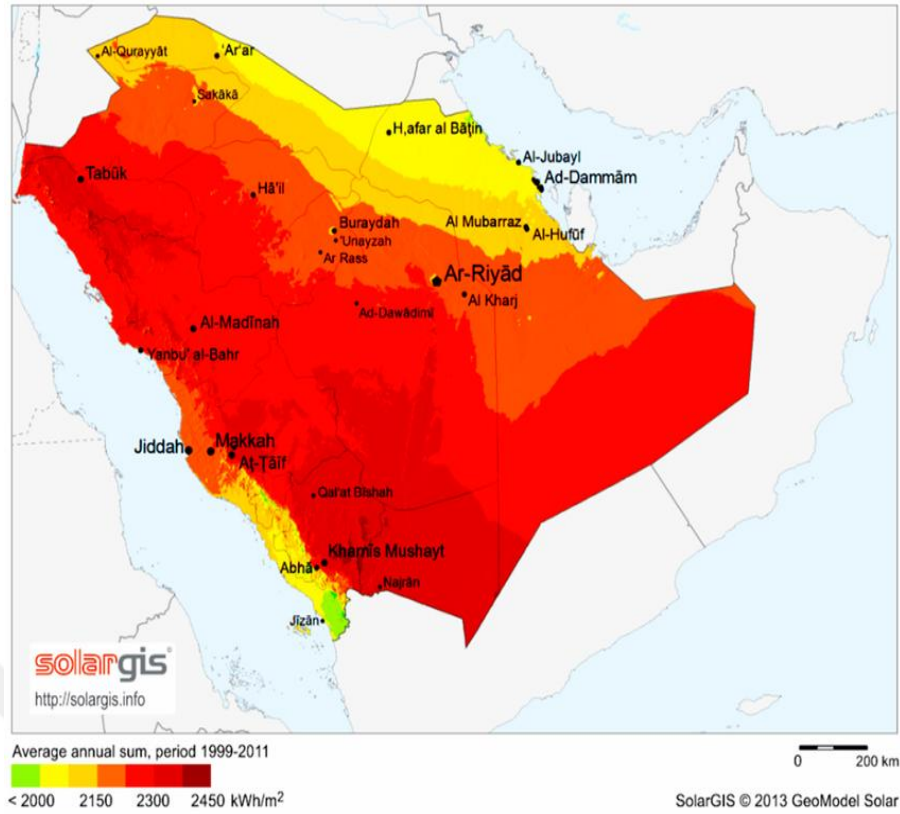
Dünya genelinde enerji tüketiminin hızla artması, yenilenebilir enerji kaynaklarına olan ihtiyacı aynı oranda yükseltmekte olup, bu kaynaklar arasında güneş, rüzgar, hidro, biyoyakıt ve hidrojen teknolojileri ön plana çıkmaktadır. Mevcut enerji kaynaklarının verimli kullanımı ve çeşitlendirilmesi, ülkeler için stratejik bir öncelik haline gelmiş olup, güneş enerjisi bu bağlamda hem ısı hem de elektrik üretimi amacıyla umut vaat eden bir seçenek olarak değerlendirilmektedir. Enerji, sanayileşme ve kentleşme süreçlerinde kritik bir faktör olup, ekonomik ve teknolojik gelişmelerde hayati rol oynamaktadır. Ayrıca, nüfus artışı enerji krizini yoğunlaştırmaktadır (Demirbas, 2009).

Güneş, devasa miktarda enerji yayarak tüm güneş sistemini beslemekte olup, Dünya enerji ihtiyacını karşılayabilecek kapasiteye sahiptir. Ancak bu enerjinin yalnızca bir kısmı yeryüzüne ulaşmaktadır. Güneşten gelen radyasyon, atmosferde absorpsiyon, saçılma ve yansıma süreçlerinden geçmektedir Doğrudan yeryüzüne ulaşan kısım tutulmaktadır. (Sharma & Puri, 2023; Wikipedia, 2025). Yeryüzüne ulaşan ortalama radyasyon, coğrafi konuma göre değişmekte olup, ekvator bölgeleri en yüksek değerlere sahiptir (Sharma & Puri, 2023). “Güneş kuşağı” olarak adlandırılan 35° Kuzey ve Güney enlemleri arası bölgeler, yıl boyunca 2000-3500 saat güneşlenme süresine ve 3.5-7 kWh/m²/gün potansiyele sahip olup, kurak alanlarda 2000-2500 kWh/m²/yıl’a ulaşmaktadır (Energy Sector Management Assistance Program (ESMAP), World Bank,

2019; Shabaniverki, 2017). Dünya üzerindeki global radyasyon dağılımı coğrafi konuma göre farklılık göstermekte olup, ekvatorda yaklaşık 2200 kWh/m²/yıl, Akdeniz’de 1700 kWh/m²/yıl, Orta Avrupa’da 1000 kWh/m²/yıl ve Kuzey Avrupa’da 800 kWh/m²/yıl değerleri gözlenmektedir. 2023 sonu itibarıyla küresel PV kurulu gücü 1.6 TW’yi aşmış olup, 2024’te 600 GW ek kapasite ile 2.2 TW’ye ulaşması beklenmektedir (Masson vd., 2024; Sandro, 2025). Güneş enerjisi üretimi, gelişmiş ülkelerde (Çin, ABD, Japonya, Almanya vb) daha yüksek olup, PV kapasitesinde bu ülkeler başı çekmektedir. Bununla birlikte çevresel sorunlar ve talep artışı, güneş enerjisini cazip kılmaktadır (Soto vd., 2022). Önümüzdeki 20 yılda güneş enerjisi küresel piyasada %30-35 kapasite artışı sağlayacak olup, enerji yatırımlarının önemli kısmını oluşturması beklenmektedir. Kurulum maliyetleri teknolojik ilerlemelerle %80’den daha fazla düşmesine rağmen, üretim potansiyeli düşük ülkelerde (Çin, ABD) üretim yüksekken, Afrika gibi yüksek üretim potansiyelli bölgelerde altyapı eksikliğinden kaynaklı olarak üretim düşük kalmaktadır (M. S. Hossain vd., 2024). Bu durum ise günümüzde hala devam eden bir problem olarak sürmektedir.

2.15. Riyad’ın Güneş Enerjisi Potansiyeli

Riyad (24.713552 N, 46.675296 E koordinatları), Suudi Arabistan’ın başkenti olup çöl iklimine sahiptir. Yüksek güneş radyasyonu potansiyeliyle yenilenebilir enerji dönüşümünde kritik rol oynayan bir şehir olarak karşımıza çıkmaktadır (Şekil 2.5). Yıllık GHI yaklaşık 2200 kWh/m² düzeyinde olup, bu değer ekvator bölgeleriyle yarışır niteliktedir (Almasoud & Gandayh, 2015). Atmosferik faktörler (bulut opaklığı, aerosol vb) değişkenlik gösterse de, Riyad’da ortalama 3000-3200 saat güneşlenme süresiyle yüksek verimlilik sağlanabilecek bölgeler arasındadır (Energy Sector Management Assistance Program (ESMAP), World Bank, 2019).

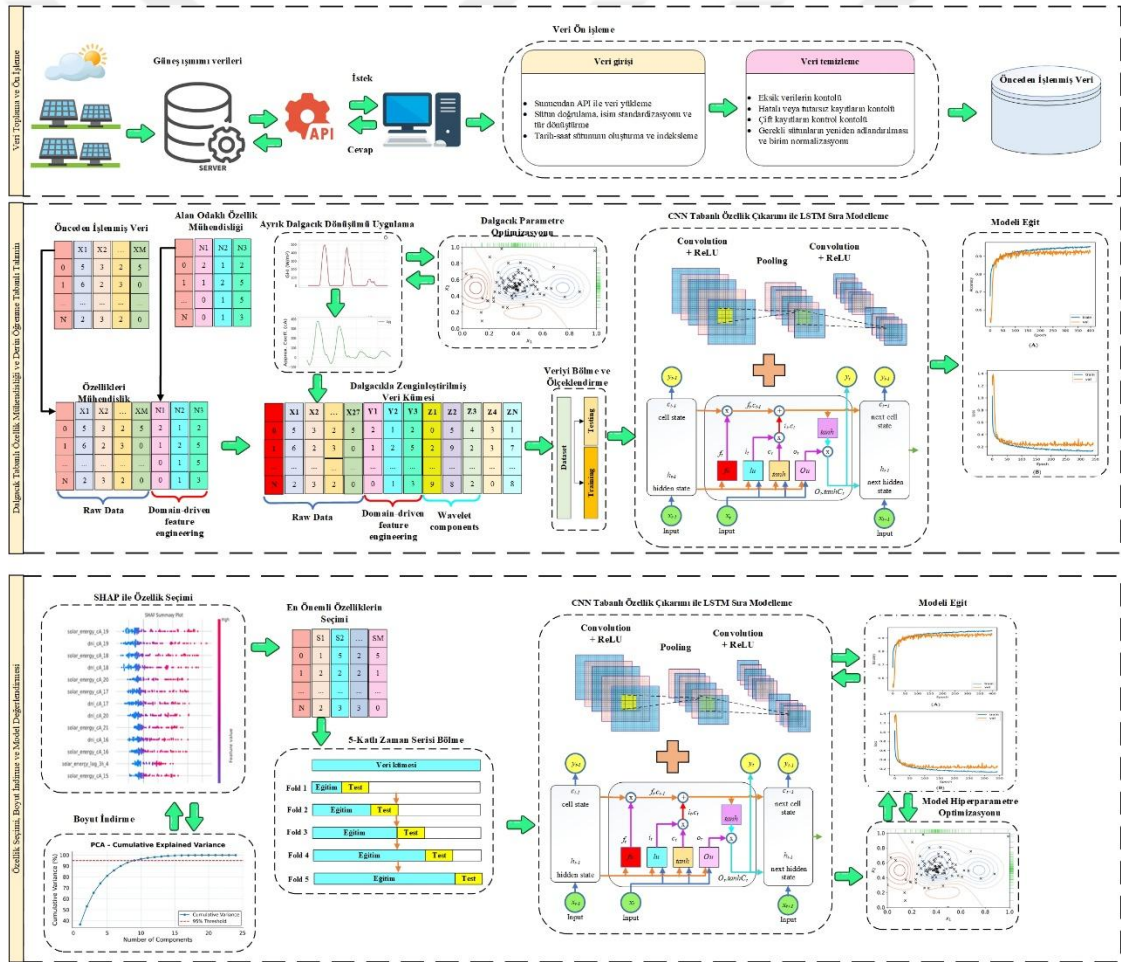


Şekil 2.5. Suudi Arabistan'ın yıllık ortalama güneş enerjisi potansiyeli dağılımı (Riyad bölgesi) (Solargis, t.y)

Çalışmada Riyad'ın seçilme gerekçesi, coğrafi avantajlarından dolayıdır. Yüksek DNI ve düşük nem oranı, CNN-LSTM modelimiz için ideal bir test bölgesi görevi görmektedir. Ayrıca Riyad'ın Suudi Arabistan'ın enerji merkezinde olması, tezdeki amaçlarımızdan birisi olan gerçek dünya uygulamalarını temsil etmektedir. Güncel çalışmalarda kararımızı desteklemektedir. Geniş kapsamlı bir çalışma 43 farklı veri kaynaktan Riyad radyasyon iklimini analiz etmiştir. Sonuçlar Riyad'ın yüksek GHI potansiyelini doğrulamıştır (Almasoud & Gandayh, 2015). Ayrıca Riyad dünyada yüksek solar enerji potansiyelli bölgeler arasında da gösterilmektedir (Energy Sector Management Assistance Program (ESMAP), World Bank, 2025).

3. YÖNTEM

Önerilen güneş ışınımı tahmini çerçevesi sıralı aşamalardan oluşmaktadır: Veri ön işleme, çok ölçekli özellik çıkarımı için DWT, hibrit CNN -LSTM modellemesi, Optuna ile hiperparametre optimizasyonu, SHAP ile özellik seçimi ve performans değerlendirmesidir. Bu entegre yaklaşım, zaman serisi verilerindeki gürültüyü azaltmayı, mekânsal ve zamansal desenleri yakalamayı ve modelin yorumlanabilirliğini sağlamayı hedefleyerek tahmin doğruluğunu artırmayı amaçlamaktadır (Şekil 3.1).



Şekil 3.1. Önerilen DWT-CNN-LSTM güneş ışınımı tahmin modeli

Bu çalışmadaki tüm deneyler, 64 çekirdekli iki adet Intel Xeon Platinum 8480+ CPU ve 128 GB RAM'e sahip, Linux tabanlı yüksek performanslı bir hesaplama sunucusunda gerçekleştirilmiştir. Yazılım ortamı olarak Python 3.11.9 ve TensorFlow 2.12 kullanılmıştır.

3.1. Yapay Zeka

YZ, bilgisayar sistemlerinin insan benzeri zeka sergilemesini sağlayan teknolojiler ve yöntemler bütünü olarak tanımlanabilir. YZ, öğrenme, problem çözme, karar verme ve algılama gibi bilişsel işlevleri taklit ederek, karmaşık görevleri otonom veya yarı otonom bir şekilde gerçekleştirebilen sistemlerin geliştirilmesini hedefler (Ozturk, 2024). Bu alanda yapılan çalışmalar sağlık, eğitim, finans, ulaşım endüstri gibi birçok sektörde dönüştürücü etkiler yaratmış ve modern toplumun teknolojik altyapısında merkezi bir rol üstlenmektedir (Aslım & Çalışkan, 2025; Kumar vd., 2024; Sunar vd., 2021).

YZ'nin temel bileşenleri arasında öğrenme, çıkarım yapma ve uyarlanabilirlik yer alır. ML, YZ'nin en önemli alt dallarından biri olarak, verilerden öğrenen algoritmaların geliştirilmesine odaklanır (Aslan & Çelebi, 2024). Denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme gibi yaklaşımlar, farklı problem türlerine çözüm sunar. Örneğin, denetimli öğrenme, etiketlenmiş verilerle modellerin eğitilmesini sağlarken; pekiştirmeli öğrenme, ödül temelli öğrenme süreçleriyle otonom sistemlerin geliştirilmesinde etkilidir (Sutton vd., 1998).

3.2. Makine Öğrenmesi

Klasik ML algoritmaları genelde el ile belirlenen özellik seçimine ve daha az karmaşık model mimarilerine dayanır. Zaman serisi verilerinde ise kısa ve uzun vadeli karmaşık ilişkiler, gürültü ve dalgalanmalar önemli bir zorluktur (Gamboa, 2017; Kayık & Çelebi, 2025). Özellikle meteorolojik verilerde çok boyutluluk, doğrusal olmayan ilişkiler ve ani değişimler sıktır. DL gibi ML algoritmaları, bu tür karmaşık ilişkileri insan müdahalesi olmadan otomatik olarak öğrenebilir. CNN ve LSTM gibi DL modelleri, özellikle çok sayıda giriş değişkeni içeren çok değişkenli zaman serisi verilerinde manuel özellik mühendisliği gerektirmeksizin otomatik olarak geçici ve non-lineer ilişkileri çıkarabilir. Hibrit CNN-LSTM yapılar, klasik yöntemlerle açıklanamayan yüksek düzeyde karmaşıklığı etkili şekilde modelleyerek daha yüksek tahmin doğruluğu sağlayabilir ve çok değişkenli zaman serilerinde hata performansının belirgin şekilde iyileştirebilir (Çelebi & Emiroğlu, 2023b; Widiputra vd., 2021).

DWT gibi ileri analizler özellik mühendisliğiyle entegre edilebilir. Son yıllarda yapılan birçok bilimsel çalışmada, meteoroloji verileri ile kısa vadeli enerji tahmini için hibrit DL modellerinin geleneksel yöntemlere göre daha iyi sonuçlar verdiği göstermiştir

(Ladjal vd., 2025; Widiputra vd., 2021). Bu yüzden, tezde DL tabanlı hibrit bir yaklaşım tercih edilmiştir. Bunun için çalışmamızda DL tabanlı bir pipeline tasarlanmıştır. Kullanılan yöntem hem klasik zaman serisi tekniklerinden hem de geleneksel ML modellerinden daha esnek ve güçlü bir yaklaşım sunmaktadır.

3.2.1. Geleneksel makine öğrenmesi yöntemleri

ML yöntemleri özellikle zaman serisi tahminleri gibi alanlarda büyük ilerlemeler göstermiştir. Geleneksel yöntemler, örneğin ARIMA, SARIMA, regresyon ve sezgisel teknikler, hedef değişkenin istatistiksel karakteristiklerine dayalı modellere dayanır. Bu yaklaşımlar, özellikle düşük veri hacmi, sınırlı değişkenli veri yapısı ve düşük hesaplama gereksinimi gibi durumlarda tercih edilmektedir (Kontopoulou vd., 2023). Ancak, çok boyutlu ve non-lineer ilişkiler içeren gerçek dünya veri kümelerinde bu yöntemlerin performansı düşebilir. Bu nedenle bu yöntemler örneğin bir finansal zaman serisine ait sadece son gözlemlere dayanarak kısa vadeli temel tahminler için tercih edilebilir (Tsay, 2005). Ancak güneş enerjisi tahmini gibi hava koşullarına bağımlı dışsal değişkenler eklenince model yapısı karmaşıklaşır. Bu nedenle, geleneksel yöntemlerin güçlü olduğu alanlar ile sınırlılıklarını analiz ederek, hibrit modellenen tekniklerin daha gerçekçi çözümler olarak düşünülebilir (B. Lim & Zohren, 2020).

3.2.2. Yapay sinir ağları

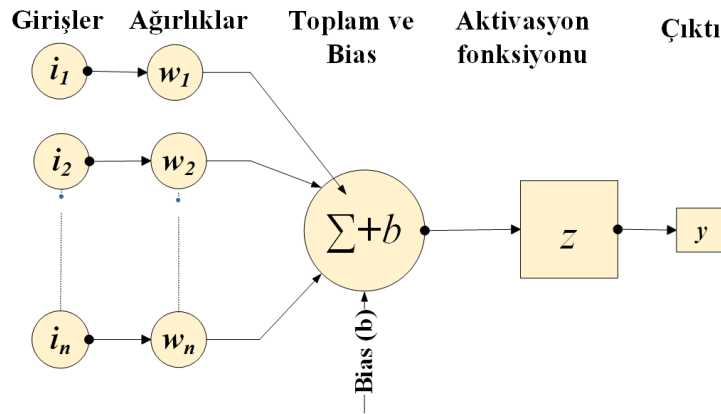
YSA, biyolojik sinir sistemlerinden esinlenilerek geliştirilmiş, çok katmanlı yapılar olarak tanımlanır. Yapılar, nöron adı verilen düğümlerin ağırlıklı bağlantılarla birbirine bağlandığı katmanlardan oluşur. Bu yapı sayesinde, eğitim verisi üzerinden öğrenme yaparak çok sayıda girdi arasındaki ilişkileri modelleyebilirler. YSA'lar temel olarak aşağıdaki bileşenlerden oluşur (S.-C. Wang, 2003):

1-Giriş katmanı, YSA'nın temel yapı taşlarından biridir ve ağın çevreyle ilk temas noktası olarak düşünülebilir. Bir sinir ağı modeli, gerçek dünyadan veya çeşitli veri kaynaklarından aldığı ham bilgiyi önce giriş katmanında toplar. Bu katmandaki her bir nöron, bir veri örneğinde yer alan bir özelliğe (feature) karşılık gelir.

2-Gizli katmanlar, YSA'nın asıl öğrenmeyi ve karmaşık örüntüleri keşfetmeyi gerçekleştirdiği katmanlardır. Giriş katmanından gelen ham veriler, gizli katmanlarda birçok farklı dönüşüme uğrar ve işlenir. Her bir gizli katmandaki nöron (düğüm), önceki

katmandan aldığı ağırlıklı girişleri toplar ve bu toplama, bir aktivasyon fonksiyonu uygular. Bu sayede doğrusal olmayan ilişkiler ve karmaşık veri desenleri öğrenilebilir. Gizli katmanlardaki nöronların kullandığı aktivasyon fonksiyonları sinir ağına esneklik kazandırır. Bu katmanda kullanılan en yaygın kullanılan aktivasyon fonksiyonlarından bazıları sigmoid ve ReLU'dur.

3-Çıkış katmanı, YSA'da modelin öğrenme sürecinin son aşamasıdır ve tüm katmanlardan geçen bilginin özetlenip bir tahmin veya karar olarak dışarıya aktarıldığı bölümdür. Ağdaki önceki katmanlardan gelen sinyaller, çıkış katmanında birleştirilerek son işleme tabi tutulur. Buradaki her bir nöron, modelin üreteceği çıktı ile doğrudan ilişkilidir. Çıkış katmanının yapısı, çözülmek istenen probleme göre değişir. Eğer bir sınıflandırma problemi söz konusuysa, çıkış katmanında genellikle “softmax” veya “sigmoid” gibi aktivasyon fonksiyonları kullanılır. Softmax, birden fazla sınıfın bulunduğu çoklu sınıflandırma problemlerinde, her sınıfa ait olasılığı hesaplar. Sigmoid ise, genellikle iki sınıflı (binary) problemler için uygundur ve sonucu 0 ile 1 arasında bir değere sıkıştırır. Regresyon problemlerinde ise çıkış katmanında çoğunlukla aktivasyon fonksiyonu kullanılmaz ya da “linear” aktivasyon seçilir. böylece model gerçek değer aralığında sayısal tahminler üretebilir (Çelebi & EmiRoğlu, 2023; S.-C. Wang, 2003). Şekil 3.2, YSA'ya ait bir nöronun temel matematiksel modelini göstermektedir. Şekil 3.2'de görüleceği üzere soldan sağa bir bilgi akışı olur.



Şekil 3.2. Bir nöronun girdi-çıkı ilişkisi (Birecikli vd., 2020)

Şekil 3.2'deki girişler ($i_1, i_2, \dots, i_n$) modele dışarıdan gelen ham verileri temsil ederler. Her bir giriş, genellikle bir veri özelliğini ifade eder. Ağırlıklar ($w_1, w_2, \dots, w_n$), olarak temsil edilir. Her giriş, bir ağırlık ile çarpılır. Bu ağırlıklar, sinir ağının öğrenme sürecinde güncellenen ve her girişin önemini belirleyen parametrelerdir. Girişler ile

ağırlıkların çarpımı toplandıktan sonra, modelin esnekliğini artırmak için bir “bias (b)” eklenir. Ardından aktivasyon fonksiyonu uygulandıktan sonra, elde edilen değer “Çıktı” olarak adlandırılır ve genellikle “y” ile gösterilir. Bu çıktı, ya bir sonraki katmana iletilir ya da ağının nihai çıktısı olur (S.-C. Wang, 2003).

3.3. Derin Öğrenme

DL çok katmanlı bir YSA’nın gelişmiş bir biçimi olarak ML’nin en önemli alanlarından biri haline gelmiştir. DL, özellikle karmaşık ve yüksek boyutlu verilerde, veri içindeki örüntüleri ve ilişkileri otomatik olarak öğrenme kapasitesiyle öne çıkar. Geleneksel YSA’larda genellikle tek gizli katman bulunurken, DL modellerinde çok daha fazla sayıda katman (bazen onlarca veya yüzlerce) ardışık olarak kullanılabilir. Bu çok katmanlı yapı, verinin daha yüksek düzeyde ve daha soyut özelliklerini çıkarabilmeyi mümkün kılar (LeCun vd., 2015).

DL, özellikle büyük veri setlerinin ve güçlü hesaplama kaynaklarının erişilebilir hale gelmesiyle birlikte, ses ve görüntü işleme, doğal dil işleme, tıbbi tanı, otomatik sürüş gibi birçok alanda çığır açıcı sonuçlar vermiştir (Duru vd., 2021; Yılmaz & Poli, 2022). Çok katmanlı ağlarda her katman verinin farklı bir temsilini öğrenir. Alt katmanlar genellikle temel ve basit özellikleri (örneğin kenarlar, renkler), daha üst katmanlar ise daha karmaşık örüntüleri (örneğin nesne parçaları, yüzler) modelleyebilir (Aslan, 2025). Bu özelliği sayesinde DL modelleri, manuel özellik mühendisliği ihtiyacını da büyük ölçüde ortadan kaldırır (Schmidhuber, 2015).

YSA’lar, DL’nin temel taşıdır. Ancak klasik YSA’lar tek katman ile öğrenme sınırlı kalabilirken, DL’de ise katman sayısı arttıkça modelin öğrenme gücü ve temsil kapasitesi de artar. Fakat bu durum beraberinde aşırı uyum, hesaplama maliyeti ve uzun eğitim süreleri gibi zorluklar da getirebilir. Bu yüzden dropout, batch normalization, erken durdurma gibi düzenleme teknikleri yaygın şekilde kullanılır. Sonuç olarak DL, çok katmanlı ve parametrelili YSA mimarileriyle modern ML’nin en etkili araçlarından biri olmuş, karmaşık verilerde üstün performans sağlamıştır (F. Bach, 2015).

3.3.1. Evrişimli sinir ağları (CNN)

CNN mimarisi, giriş verisinden uzamsal ve/veya zamansal örüntüleri otomatik olarak öğrenmek üzere geliştirilmiş katmanlı yapılardan oluşur. Özellikle bir boyutlu

konvolüsyon katmanları, zaman serisi veya ardışık veri analizinde öne çıkar (Yılmaz, 2025). Her konvolüsyon katmanında, giriş üzerinde kayan, öğrenilebilir çekirdekler (w^k) kullanılır. Bu çekirdekler, model eğitimi sırasında optimize edilerek veri içindeki karakteristik desenleri yakalar. Bir CNN’de, çekirdek (kernel) k için konum (i, j) ‘deki aktivasyon olmak üzere ağ parametresi Denklem (3.1)-(3.2) ile hesaplanır (O’Shea & Nash, 2015):

$$z_{(i,j)}^k = \sum_{m,n} x_{(i+m,j+n)} \cdot w_{(m,n)}^k + b^k \quad (3.1)$$

$$a_{(i,j)}^k = f(z_{(i,j)}^k) \quad (3.2)$$

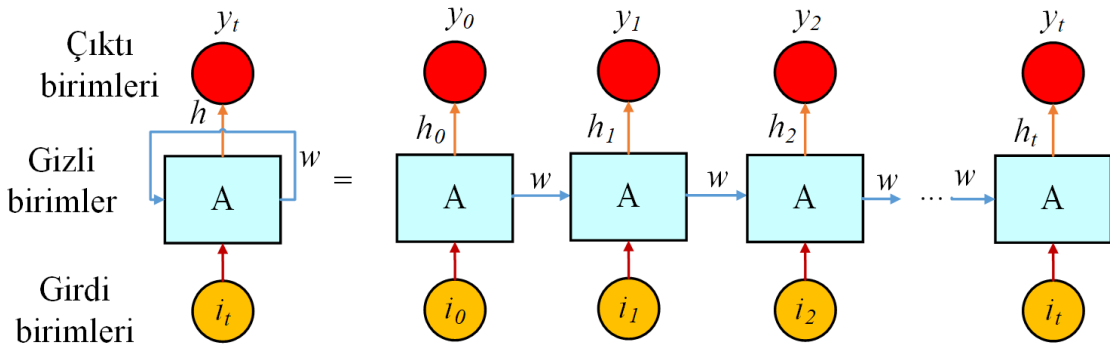
Burada $x_{(i+m,j+n)}$, giriş matrisinden ilgili pencerenin değerlerini, $w_{(m,n)}^k$ çekirdeğin ağırlıklarını, b^k ise bias terimini gösterir. f aktivasyon fonksiyonudur ve genellikle $\text{ReLU}(f(z) = \max(0, z))$ tercih edilir. Bu yapı sayesinde CNN katmanları, giriş verisindeki yerel (lokal) desenleri otomatik olarak çıkarır ve bu desenleri bir sonraki katmanlara aktarır (Çelebi & Emiroğlu, 2023a). Tipik olarak, konvolüsyon katmanlarının ardından max-pooling gibi alt örnekleme katmanları gelir. Bu işlemler uzamsal veya zamansal boyutun küçültülmesini ve daha soyut özelliklerin oluşturulmasını sağlar. Sonraki aşamada, flatten işlemi ile çok boyutlu çıktılar tek boyuta indirilir ve tam bağlantılı (fully connected) katmanlara giriş olarak verilir (Aslan, 2025). Tez çalışmasında, CNN mimarisinin bu özelliklerinden faydalanarak özellikle dalgacıkla ayrıştırılmış zaman serisi verisinde yerel ve geçici örüntülerin otomatik şekilde öğrenilmesi hedeflenmiştir. Dalgacık dönüşümü ile elde edilen detay ve yaklaşık bileşenler, giriş olarak CNN’e sunulduğunda ağ, hem frekans bazlı hem de zamansal değişimleri modelleyebilir. Bu sayede, klasik yöntemlerin kaçırabileceği kısa süreli dalgalanma veya ani değişimlerin model tarafından otomatik olarak yakalanması mümkün olur. Sonuç olarak, CNN tabanlı mimari hem geleneksel özellik mühendisliği ihtiyacını azaltır hem de yüksek boyutlu verilerde güçlü performans gösterir. Tez kapsamında geliştirilen modelde, bu yapının avantajlarından yararlanılmaya çalışmıştır.

3.3.2. Tekrarlayan sinir ağırları (RNN)

Tekrarlayan sinir ağırları (RNN), ardışık ve zamana bağlı veri yapıları ile çalışmak için özel olarak tasarlanmış YSA türevleridir. RNN'ler, klasik ileri beslemeli (feedforward) ağların aksine, ağ içinde geri bağlantılar (feedback) içerir. Böylece ağ, önceki adımlardan gelen bilgiyi hafızasında tutarak, güncel girdileri geçmişteki bağlama göre değerlendirebilir. RNN'in temel farkı, bir "gizli durum" (hidden state) üzerinden, ağın geçmişte işlediği bilgileri yeni gelen veri ile birleştirmesidir. Matematiksel olarak, her adımda ağın durumu Denklem 3.3'deki tanımladığı gibi güncellenir (Hochreiter & Schmidhuber, 1997):

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b) \quad (3.3)$$

Burada x_t güncel giriş, h_{t-1} bir önceki gizli durum, W_{xh} ve W_{hh} ağırlık matrisleri, b bias ve f aktivasyon fonksiyonudur (Şekil 3.3).

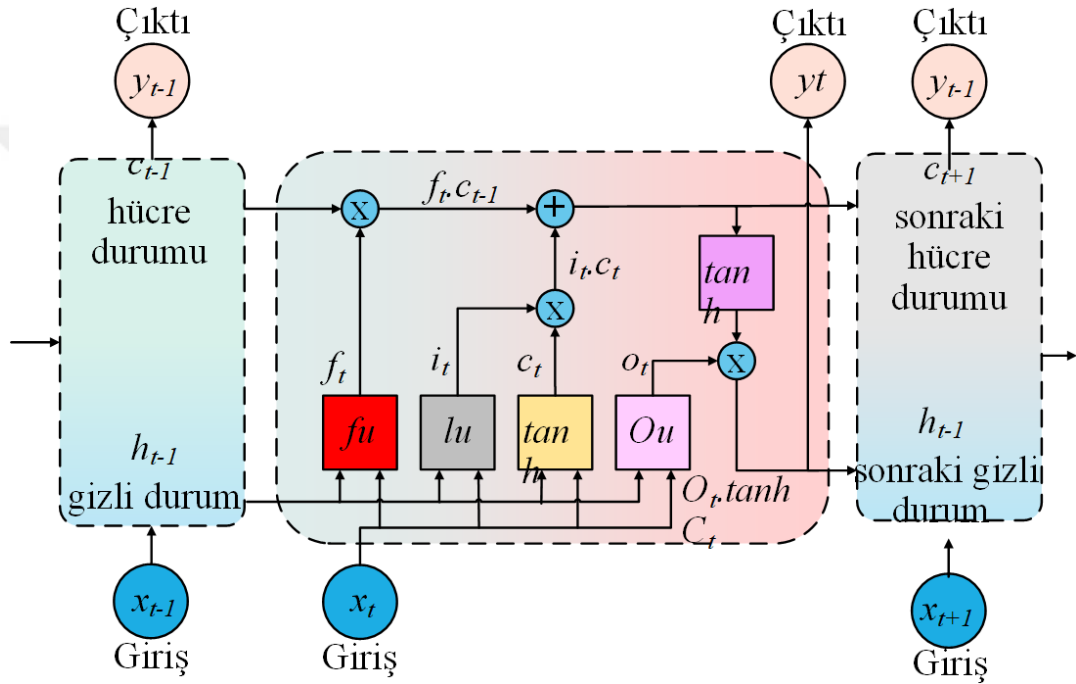


Şekil 3.3 Tekrarlayan sinir ağında girdi, gizli ve çıkış birimleri (uyarlandı (Çelebi & Fidan, 2024))

RNN'ler dil modelleme, zaman serisi analizi, ses işleme ve makine çevirisi gibi ardışık bağımlılıkların önemli olduğu birçok alanda yaygın olarak kullanılmaktadır. RNN yapısı, verinin sırasını ve geçmişini dikkate aldığı için özellikle zamansal bağımlılıklar içeren problemlerde avantaj sağlar. Ancak RNN'ler uzun ardışık veri dizilerinde "gradyan sönmesi (vanishing gradient)" ve "gradyan patlaması (exploding gradient)" gibi öğrenme sorunlarına yatkındır. Bu nedenle pratikte genellikle Gated Recurrent Unit (GRU), LSTM gibi gelişmiş varyantları tercih edilir (Graves vd., 2013).

3.3.3. Uzun kısa süreli bellek ağları (LSTM)

LSTM ağları, klasik RNN'nin zaman içindeki uzun vadeli bağımlılıkları öğrenmede yaşadığı “gradyan sönmesi” sorununu aşmak için geliştirilmiş, özel kapı mekanizmalarına sahip RNN varyantlarıdır. LSTM hücreleri, özellikle uzun aralıklı verilerde geçmişten gelen bilgiyi koruyabilmek ve unutulması gereken detayları silmek için iç yapısında unutma (forget), giriş (input) ve çıkış (output) kapısı olmak üzere üç ana kapı kullanır (Hochreiter & Schmidhuber, 1997).



Şekil 3.4. Bir LSTM hücresinin iç yapısı ve çalışma prensibi (Çelebi & Karaman, 2023)

LSTM her bir her zaman adımında kapılarını ve durumlarını Denklem (3.4)-(3.9) ile tanımlandığı gibi günceller:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3.4)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (3.5)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (3.6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (3.7)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (3.8)$$

$$h_t = o_t \odot \tanh(c_t) \quad (3.9)$$

Burada x_t giriş vektörüdür (tezimizde CNN çıktısı), h_{t-1} ve c_{t-1} önceki gizli durum ve hücre durumudur ve σ sigmoid aktivasyon fonksiyonudur. Gerekirse birden fazla LSTM katmanı üst üste yerleştirilebilir. Her zaman adımında, unutmakapısı önceki hücre bilgisinin ne kadarının korunacağını, giriş kapısı ve aday hücre yeni bilgi güncellemesini ve çıkış kapısı ise yeni gizli durumu belirler. Bu sayede, LSTM hücresi geçmişteki önemli bilgileri gerektiği kadar uzun süre saklayabilir.

Tez çalışmasında, zaman serisi verilerinde LSTM katmanları kullanılarak hem kısa süreli hem de uzun vadeli zamansal bağımlılıkların otomatik olarak öğrenilmesi hedeflenmiştir. LSTM katmanlarına giriş olarak genellikle ön işlenmiş özellikler veya CNN katmanlarının çıktıları verilebilir. Böylece, verideki karmaşık geçici ve lokal örüntüler LSTM aracılığıyla zaman boyutunda işlenir. Ayrıca, LSTM mimarisinde birden fazla katman ardışık olarak kullanılabilir. Çok katmanlı LSTM yapıları, daha derin ve soyut zaman serisi temsili modelleyerek, öğrenme kapasitesini artırabilir. Son LSTM katmanının çıktısı ise genellikle tam bağlantılı (dense) bir katmandan geçirilerek nihai tahmin (örneğin ısınım değeri) elde edilir.

3.4. Veri Kümesi

Tez çalışmasında, güneş enerjisi sistemlerinin tasarımı ve performans analizi için temel bir bileşen olan meteorolojik ve güneş ısınımı verilerinden oluşan kapsamlı bir veri kümesi olan Solcast kullanılmıştır. Bu veri kümesi, bir API aracılığıyla temin edilmiş olup, sistemlerin enerji üretimini etkileyen çok çeşitli atmosferik ve çevresel parametreleri içermektedir. Çalışmada Riyad lokasyonu, Suudi Arabistan (24.713552° K, 46.675296° D), için 15 dakikalık meteorolojik değişkenlerden oluşan bir veri kümesi kullanılmaktadır. Veri kümesi, 1 Nisan 2022'den 1 Nisan 2025'e kadar olan dönemi kapsamaktadır. 105.312 örnek ve 28 özellikten (zaman bilgisi dahil) oluşmaktadır. Ham veriler, Solcast platformunun açık erişimli API'si aracılığıyla 30 günlük bloklar halinde elde edilmiştir (Solcast, 2025).

Veri özellikleri hava sıcaklığı (air_temp), yüzeyin yansıtma kapasitesini ifade eden albedo, güneşin yatay düzlemdeki açısız konumunu gösteren güneş azimutu (azimuth), açık gökyüzü koşullarında ısınımın farklı bileşenlerini temsil eden doğrudan dağınık ısınım (clearsky_dhi), doğrudan normal ısınım (clearsky_dni), küresel yatay ısınım (clearsky_ghi) ve eğik yüzeye gelen ısınım (clearsky_gti) bulunmaktadır. Ayrıca, atmosferik koşulları detaylandıran bulut opaklığı (cloud_opacity), çiy noktası sıcaklığı

(dewpoint_temp), gerçek dünya koşullarında ölçülen doğrudan dağınık ışınım (dhi), doğrudan normal ışınım (dni), küresel yatay ışınım (ghi) ve eğik yüzeye gelen ışınım (gti) gibi parametreler veri kümesinde yer almaktadır. Çevresel faktörler arasında ise yağış miktarı (precipitable_water), yağış oranı (precipitation_rate), bağıl nem (relative_humidity), yüzey basıncı (surface_pressure), kar birikimini ve etkilerini ölçen kar derinliği (snow_depth), kar su eşdeğeri (snow_water_equivalent), çatıdaki kar kirlenmesi (snow_soiling_rooftop) ve zemindeki kar kirlenmesi (snow_soiling_ground) parametreleri bulunmaktadır. Rüzgar dinamiklerini değerlendirmek için rüzgar yönü (wind_direction_10m ve wind_direction_100m) ve rüzgar hızı (wind_speed_10m ve wind_speed_100m) 10 metre ve 100 metre yüksekliklerde ölçülmüş; güneşin dikey konumunu belirleyen güneş zenith açısı (zenith) ile veri toplama zaman aralıklarını ifade eden veri periyodunun sonu (period_end) ve periyot (period) parametrelerinden oluşmaktadır (Tablo 3.1).

Tablo 3.1. Güneş ışınımı tahmininde kullanılan atmosferik, çevresel ve meteorolojik veri özellikleri (Solcast, 2025)

Kategori	Parametre Adı	Açıklama
Meteorolojik	air_temp	Hava sıcaklığı
	albedo	Yüzeyin yansıtma kapasitesi
	azimuth	Güneşin yatay düzlemdeki açısal konumu
	clearsky_dhi	Açık gökyüzü koşullarında doğrudan dağınık ışınım
	clearsky_dni	Açık gökyüzü koşullarında doğrudan normal ışınım
	clearsky_ghi	Açık gökyüzü koşullarında küresel yatay ışınım
	clearsky_gti	Açık gökyüzü koşullarında eğik yüzeye gelen ışınım
	cloud_opacity	Bulut opaklığı
	dewpoint_temp	Çiy noktası sıcaklığı
	dhi	Gerçek koşullarda ölçülen doğrudan dağınık ışınım
	dni	Gerçek koşullarda ölçülen doğrudan normal ışınım
	ghi	Gerçek koşullarda ölçülen küresel yatay ışınım
	gti	Gerçek koşullarda ölçülen eğik yüzeye gelen ışınım
	precipitable_water	Yağışa dönüşebilir atmosferik su miktarı
Çevresel	precipitation_rate	Yağış oranı
	relative_humidity	Bağıl nem
	surface_pressure	Yüzey basıncı
	snow_depth	Kar derinliği
	snow_water_equivalent	Kar su eşdeğeri
	snow_soiling_rooftop	Çatıdaki kar kirliliği
	snow_soiling_ground	Zemindeki kar kirliliği
Rüzgar Dinamikleri	wind_direction_10m	10 m yükseklikte rüzgar yönü
	wind_direction_100m	100 m yükseklikte rüzgar yönü
	wind_speed_10m	10 m yükseklikte rüzgar hızı
	wind_speed_100m	100 m yükseklikte rüzgar hızı
Güneş Konumu ve Zaman	zenith	Güneşin zenith (dikey) açısı
	period	Veri toplama süresini ifade eden zaman aralığı başlangıcı
	period_end	Veri toplama süresini ifade eden zaman aralığı sonu

Çalışma ön işlemede, tutarlı zaman indekslemesi sağlamak için zaman hizalamayı içermektedir. Bunu eksik veya anormal değerlerin lineer interpolasyon ve z-skor tabanlı aykırı değer tespiti (eşik ± 3 standart sapma) ile ele alınması izlemiştir. Zaman bağımlılıklarını korumak için, veri kümesi kronolojik olarak %80 eğitim (doğrulama dahil) ve %20 test olarak bölünmüştür. Optuna ile hiperparametre optimizasyonu için, 5 fold zaman serisi bölme kullanılarak, zaman-serisi farkındalığına sahip iç çapraz doğrulama uygulanmıştır. Böylece bölümler arasında veri sızıntısı önlenmek amaçlanmıştır.

3.5. Özellik Çıkarımı

Tez çalışmasında özellik çıkarımı iki farklı yöntemle gerçekleştirilmiştir. İlk olarak, manuel özellik mühendisliği yapılmıştır. Yani mevcut ham verilerden anlamlı yeni özellikler elde edilmiştir. Bu tür yeni özellikler, fiziksel anlamı olan ve literatürde sıkça kullanılan yöntemlerdir (Owens vd., 2025). İkinci olarak, DWT uygulanarak her bir sayısal özellikten otomatik olarak yeni alt özellikler türetilmiştir. Bu dönüşüm ile hem Solcast kaynaklı verilerden hem de manuel özellik mühendisliği ile üretilen verilerden düşük frekanslı (trend/cA) ve yüksek frekanslı (detay/cD1) bileşenleri çıkarılmıştır. Böylece, modelin hem kısa süreli dalgalanmalara hem de uzun vadeli trendlere duyarlı hale gelmesi amaçlanmıştır. Db4 dalgacık seçimi, dalga ayrıştırma için etkili olduğu için tercih edilmiştir. Seviye 2, yeterli frekans ayrıştırması sağladığı için kullanılmıştır.

3.5.1. Manuel özellik mühendisliği

Tez çalışmasında manuel özellik mühendisliği ile üç yeni özellik elde edilmiştir. Oluşturulan yeni özellikler, literatürde fotovoltaik güç ve güneş enerjisi tahmini alanında sıkça kullanılan ve fiziksel olarak anlamlı olan oranlardan seçilmiştir. “apparent_temp” (hissedilen sıcaklık), rüzgarın soğutucu etkisini hesaba kattığı için gerçek enerji üretimini daha iyi yansıtır (Lateef vd., 2021). “dni_ghi_ratio” ve “dhi_ghi_ratio” oranları ise, atmosferdeki bulutluluk durumu, ışık dağılımı ve güneşlenme koşullarını özetleyen göstergelerdir. Bu nedenle, tezde kullanılan manuel özellikler fiziksel anlamı güçlü ve önceki çalışmalarda performans artışı sağladığı kanıtlanmış göstergelerden oluşmaktadır apparent_temp, dni_ghi_ratio ve dhi_ghi_ratio Denklem (3.10)-(3.12) kullanılarak hesaplanmıştır (Liu vd., 2021):

$$\text{apparent_temp} = T_{\text{air}} - 0.7 \times V_{\text{wind}} \quad (3.10)$$

$$\text{dni_ghi_ratio} = \frac{DNI}{GHI + \epsilon} \quad (3.11)$$

$$\text{dhi_ghi_ratio} = \frac{DHI}{GHI + \epsilon} \quad (3.12)$$

Burada T_{air} hava sıcaklığı, V_{wind} rüzgâr hızı, DNI doğrudan gelen güneş ışıınımı, ϵ ise sıfıra bölmeyi engellemek için küçük bir değerdir (tezde 1×10^{-6} olarak alındı). Özellik mühendisliğimizde amacımız modelinin yalnızca ham sensör verilerinden değil, aynı zamanda bu tür anlamlı ve açıklayıcı türev bilgilerini kullanılarak model başarısını artırmayı amaçlamaktadır.

3.5.2. Dalgacık dönüşümü ile çok ölçekli özellik çıkarımı

Çalışmada, çok değişkenli zaman serisi verilerindeki kısa vadeli dalgalanmaları ve uzun vadeli eğilimleri yakalamak amacıyla, ön işleme adımı olarak DWT uygulanmıştır. DWT, sinyalleri frekans bileşenlerine ayırarak gürültüyü etkili bir biçimde azaltması ve ilgili örüntüleri ön plana çıkarması nedeniyle, güneş ışıınımı tahmininde DL modellerinin dayanıklılığını ve doğruluğunu önemli ölçüde artırır. Örneğin, 15 dakikalık güneş ışıınımı verilerinde, DWT yüksek frekanslı bileşenlerle ani bulut geçişlerini, düşük frekanslı bileşenlerle ise günlük eğilimleri ayrıştırarak modelin non-stationarity ile başa çıkmasını sağlar. Son dönem çalışmalar, bu tekniğin hibrit modellerde kullanımının RMSE değerlerinde iyileşme sağladığını göstermektedir (Çevik Bektaş & Altaş, 2024). Benzer biçimde, DWT'yi CNN-LSTM mimarileriyle bütünleştiren hibrit yaklaşımlar, değişken hava koşullarını yönetmede üstün başarı sergilemiştir. Bu entegrasyon, dalgacık bileşenlerinin CNN katmanlarında yerel özellik çıkarımı için kullanılmasını sağlar ve LSTM'nin zamansal bağımlılıkları modellemesini güçlendirir (Camacho vd., 2025).

Optuna optimizasyonuna dayalı olarak (Bölüm 3.6'da ayrıntılı biçimde ele alınmıştır), Daubechies-4 (db4) ana dalgacık ve 2 seviyeli ayrıştırma en iyi parametreler olarak seçilmiştir. Db4 dalgacığı, yüksek vanishing moments db4 için 4 moment) özelliği sayesinde düzgün sinyalleri (güneş ışıınımı gibi) etkili biçimde temsil eder ve sınır etkilerini minimize eder (Shensa, 1992). Bu işlem, uzun vadeli eğilimleri temsil eden bir düşük frekanslı yaklaşım bileşeni (cA) ile kısa vadeli varyasyonları ve gürültüyü yansıtan bir yüksek frekanslı detay bileşeni (cD1) üretmiştir. cD2 hariç tutulmuştur çünkü seviye 2'de yeterli ayrıştırma sağlanmıştır. Böylece ayrıştırılmış katsayılar orijinal zaman

çözünürlüğüne uyarlanmış (yeniden örneklenmiş) ve CNN-LSTM modeline giriş verisi olarak birleştirilmiştir. DWT, matematiksel olarak denklem (3.13) -(3.14) ile tanımlanır:

$$cA_j[k] = \sum_n x[n] \cdot g[2k - n] \quad (3.13)$$

$$cD_j[k] = \sum_n x[n] \cdot h[2k - n] \quad (3.14)$$

Burada, g düşük geçiren filtreyi, h yüksek geçiren filtreyi, $x[n]$ giriş sinyalini, cA_j seviye j 'deki yaklaşım katsayısını ve cD_j detay katsayısını temsil etmektedir. Dalgacık parametreleri için arama alanı, Optuna'nın TPE algoritması ile optimize edilmiştir. Bu algoritma, Bayesçi optimizasyon temelli olup, hiperparametre uzayını verimli tarayarak en iyi db4 ve seviye 2 kombinasyonunu belirlemiştir.

3.6. Optuna ile Bayesçi Optimizasyon ve TPE Algoritması

Geleneksel hiperparametre belirleme yöntemleri, manuel deneme-yanılma veya grid/random search gibi temel yaklaşımları içermektedir. Ancak, hiperparametre uzayının karmaşıklığı ve parametreler arasındaki etkileşimler, bu yöntemleri hem zaman alıcı hem de verimsiz hale getirebilir. Optuna, Bayesian optimizasyon temelli bir yaklaşım sunar ve TPE gibi algoritmalarla hiperparametre arama alanını etkin bir şekilde keşfeder. Ayrıca, erken durdurma (pruning) mekanizması ile düşük performanslı denemeleri filtreleyerek hesaplama kaynaklarını optimize eder (Akiba vd., 2019).

Çalışmada, hiperparametre optimizasyonu için Optuna'nun TPE algoritması tercih edilmiştir. TPE, Bayesyen optimizasyon çerçevesinde, model eğitiminin maliyetli olduğu durumlarda en iyi parametre kombinasyonunu az sayıda deneme ile bulmayı hedefler. TPE, hiperparametre uzayını olasılıksal yoğunluk tahminleri ile modelleyerek, önceki denemelerin başarımlarını (örneğin doğrulama kaybı) dikkate alır. Bu algoritma, hiperparametre kombinasyonu (θ) e ilgili performans metriği y (çalışmamızda RMSE) üzerinden bir hedef fonksiyonu optimize eder. TPE'de, en iyi sonuçları veren hiperparametrelerin seçilme olasılığını artırmak için iki yoğunluk dağılımı tanımlanır: $l(\theta)$ (en iyi y değerlerine karşılık gelen dağılım) ve $g(\theta)$ ise (daha kötü y değerlerine karşılık gelen dağılım). Buradan optimizasyon Denklem 3.15 ile ifade edilebilir (Hakhverdyan, 2024).

$$\theta^* = \arg \max_{\theta \in \mathcal{H}} \frac{l(\theta)}{g(\theta)} \quad (3.15)$$

Burada, $l(\theta)$ ve $g(\theta)$ Parzen estimator ile tahmin edilir ve y için bir eşik değeri belirlenerek arama alanı daraltılır.

Tablo 3.2’de DL modelinin mimarisinde Optuna optimizasyon algoritması ile ayarlanan tüm hiperparametreler ve arama aralıkları gösterilmektedir. Öğrenme oranı için sürekli ve logaritmik bir aralık tanımlanırken, diğer katmanlara ait parametreler ise literatüre uygun olarak belirli seçenekler üzerinden optimize edilmiştir. Dropout oranları sürekli bir aralıkta, filtre ve nöron sayıları ise kategorik olarak taranmıştır (Gal & Ghahramani, 2015; Tsai vd., 2024). Bu parametrelerin her biri, modelin tahmin performansını en üst düzeye çıkarmak amacıyla optimize edilmiştir.

Tablo 3.2: Optuna Bayesian Optimizasyonu ile belirlenen hiperparametreler, arama uzayı ve en iyi değerler

Kategori	Parametre Adı	Arama Uzayı (Tip)	En İyi Değer	İlgili Fold/Not
Dalgacık	dalgacık	Kategorik: ['db2', 'db4', 'db6', 'sym4', 'coif1']	'db4'	Tüm fold'lar; düşük gürültü için.
Dalgacık	dalgacık_level	Tamsayı: [1, 2, 3]	2	Fold 5'te en düşük RMSE (4.49).
Model	learning_rate	Logaritmik: 1e-5 ile 1e-3 arası	0.0001	Stabil eğitim; Fold 1-5 ortalama RMSE 6.62.
Model	filters1	Kategorik: [32, 64, 128]	64	Dengeli özellik çıkarma.
Model	filters2	Kategorik: [32, 64, 128]	64	Dengeli özellik çıkarma.
Model	kernel_size1	Kategorik: [3, 5]	3	Küçük kernel, yerel pattern'lar için.
Model	kernel_size2	Kategorik: [3, 5]	5	Orta kernel, geniş pattern'lar için.
Model	lstm_units1	Kategorik: [64, 128, 256]	128	Zaman bağımlılıkları için.
Model	lstm_units2	Kategorik: [64, 128, 256]	64	Hafıza verimliliği.
Model	dropout1	Kategorik: [0.1, 0.2, 0.3, 0.4, 0.5]	0.3	Düzenleme.
Model	dropout2	Kategorik: [0.1, 0.2, 0.3, 0.4, 0.5]	0.2	Düzenleme.
Model	dense_units	Kategorik: [32, 64, 128]	32	Basit çıktı katmanı.

Bu hiperparametreler, Optuna ile 20 trial sonrası belirlenmiştir. Optuna ile gerçekleştirilen bu optimizasyon, güneş enerjisi tahmininde hiperparametrelerin etkileşimini dikkate alarak, modelin genel performansını iyileştirdiği ablasyan çalışmalarından anlaşılmaktadır.

3.7. Model Değerlendirme

Bir ML modelinin başarısı, daha önce görmemiş olduğu test veri seti üzerinde değerlendirilir. Bunun için birçok istatistiksel metot vardır. Önerilen DWT–CNN–LSTM modelinin performansı, birçok istatistiksel yöntem ile değerlendirilmiştir.

Değerlendirme Metrikleri: Çalışmamızda tahmin doğruluğu, regresyon metrikleri kullanılarak nicel olarak değerlendirilmiştir. Gerçek değerler kümesi y_i , tahminleri $\hat{y}_i$ ve örnek sayısını n kabul edelim. Buradan Denklem (3.16)-(3.26)'da (Botchkarev, 2019; Chicco vd., 2021; Lewinson, 2025):

- **Kök Ortalama Kare Hata (RMSE):** Gerçek ve tahmin edilen değerler arasındaki kare farkların ortalamasının kareköküdür (Denklem 3.16):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.16)$$

- **Ortalama Mutlak Hata (MAE):** Gerçek ve tahmin edilen değerler arasındaki mutlak farkların ortalamasıdır (Denklem (3.17)).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.17)$$

- **Açıklanan Varyans Oranı (R^2):** Modelin toplam değişkenliğin ne kadarını açıkladığını gösterir (Denklem (3.18)):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.18)$$

- **Ortalama Mutlak Yüzde Hata (MAPE):** Gerçek değerlere göre yüzdesel hatadır (Denklem (3.19)):

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.19)$$

• **Medyan Mutlak Hata (MedAE):** Tahmin ve gerçek değerler arasındaki mutlak farkların medyanıdır (Denklem (3.20)):

$$\text{MedAE} = \text{median}(|y_i - \hat{y}_i|) \quad (3.20)$$

• **Açıklanan Varyans Skoru (EVS):** Modelin tahminlerinin gerçek değerlerin varyansını ne kadar iyi açıkladığını ve verideki değişkenliğin ne kadarını başarıyla yakaladığını gösteren bir performans ölçütüdür (Denklem (3.21)):

$$\text{EVS} = 1 - \frac{\text{Var}(y - \hat{y})}{\text{Var}(y)} \quad (3.21)$$

Burada Var, varyansı temsil eder.

• **Kök Ortalama Kare Hatanın Ortalama Değere Oranı (CVRMSE):** Modelin hata büyüklüğünü gerçek değerlerin ortalaması göre yüzde olarak ifade eden bir hata ölçütüdür (Denklem (3.22)):

$$\text{CVRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100 \quad (3.22)$$

• **Ortalama sapma hatası (MBE):** modelin tahminlerinin gerçek değerlere göre ortalama olarak ne kadar fazla ya da eksik olduğunu gösteren bir hata ölçütüdür (Denklem (3.23)):

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (3.23)$$

• **Simetrik Ortalama Mutlak Yüzde Hata (sMAPE):** tahmin ile gerçek değerler arasındaki farkı, her iki değer in ortalamasına oranlayarak yüzde olarak hesaplayan ve ölçeğe duyarlı olmayan bir hata ölçütüdür (Denklem (3.24)):

$$\text{sMAPE} = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \quad (3.24)$$

- **Ölçeklenmiş Ortalama Mutlak Hata (MASE):** Modelin ortalama mutlak hatasını, basit bir naif tahmin yönteminin ortalama mutlak hatasına bölerek ölçekten bağımsız şekilde karşılaştırma yapmaya olanak tanıyan bir hata ölçütüdür. Naive tahmin baz alınarak normalize edilir (Denklem (3.25)):

$$\text{MASE} = \frac{\text{MAE}}{\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|} \quad (3.25)$$

- **Test_Uncertainty:** Test veri seti üzerinde yapılan tahminlerin belirsizliğini nicel olarak ifade eden bir metriktir. Bu belirsizlik genellikle Monte Carlo Dropout (MC-Dropout) gibi stokastik çıkarım yöntemleri ile elde edilir. MC-Dropout kullanıldığında, aynı giriş verisi üzerinde model birden fazla tahmin üretir. Her bir veri noktası için bu tekrar tahminlerin standart sapması hesaplanır. Ardından, tüm veri noktaları için elde edilen standart sapmaların ortalaması alınarak ortalama tahmin belirsizliği elde edilir (Denklem (3.26)):

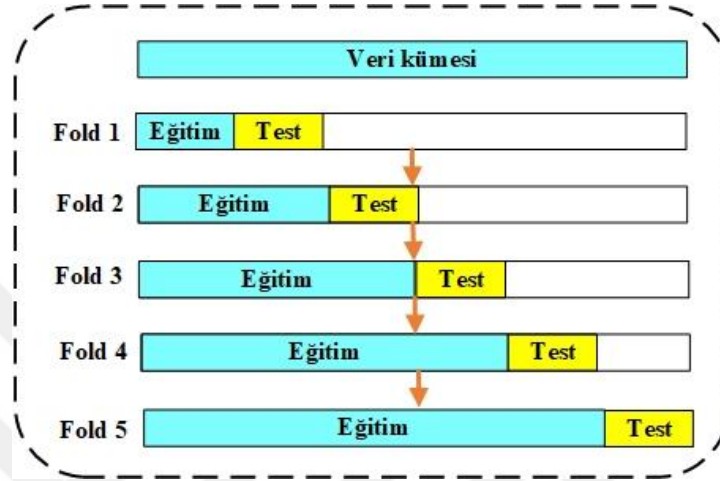
$$\text{Test Uncertainty} = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_i^{(t)} - \bar{y}_i)^2} \quad (3.26)$$

Burada $\hat{y}_i^{(t)}$, i -inci veri noktası için t -inci MC-Dropout tahminini; $\bar{y}_i$ ise aynı veri noktası için tüm tekrarların ortalamasını ifade eder.

3.8. Çapraz Doğrulama ve Veri Bölme

Bu çalışmada modelin genelleme yeteneğini değerlendirmek ve overfitting'i önlemek amacıyla 5 katlı çapraz doğrulama (5 fold CV) uygulanmıştır. Çapraz doğrulama, veri setini birden fazla alt küme olarak bölerek modelin farklı veri parçaları üzerindeki performansını test eder. Bu yöntem, özellikle güneş enerjisi gibi zaman serisi verilerinde modelin değişken hava koşullarına uyumunu ölçer. Literatürde, CV'nin tahmin modellerinde hata tahminini azalttığı ve güvenilirliği artırdığı vurgulanır (Kohavi & others, 1995). Bu yaklaşım, geleceğe dayalı veri sızıntısını önler ve geçmiş verilerle eğitim, gelecek verilerle test sağlar. Zaman serisi bölme ile veri seti 5 ($n_splits=5$) kata bölünmüştür. Her kat için eğitim seti önceki verilerden, doğrulama ve test setleri ise

sonraki verilerden oluşur (Bergstra & Bengio, 2012) Veri boyutu (Riyad 15 dakikalık güneş verisi) dikkate alınarak, $test_size=0.2$ ile ayrılmış, kalan kısım 5 kata bölünmüştür. Her fold'da, veriler Z-score normalization ile normalize edilmiş ve ardından zaman adımlı sekanslar ($time_steps=24$) oluşturulmuştur. CV sonuçları bir modelin genelleme yeteneğini, aşırı öğrenme riskini ve tahmin güvenilirliğini gösterir (Şekil 3.5).



Şekil 3.5. 5-Katlı çapraz doğrulama süreci

4. BULGULAR VE TARTIŞMA

4.1. Model Performansı Değerlendirilmesi

Tezde önerilen SHAP tabanlı özellik seçimi kullanan ve CNN-LSTM mimarisi ile oluşturulan modelin performansı, 5 fold zaman serisi çapraz doğrulama yöntemiyle değerlendirilmiştir. Her fold için test veri kümesine ait RMSE, MAE, R^2 ve belirsizlik metriği hesaplanmıştır. Sonuçlar hem normalize edilmiş hem de orijinal ölçekli değerler üzerinden ayrı ayrı hesaplanmıştır. Böylece model performansının daha anlaşılır ve yorumlanabilir olması sağlanmaya çalışılmıştır (Tablo 4.1).

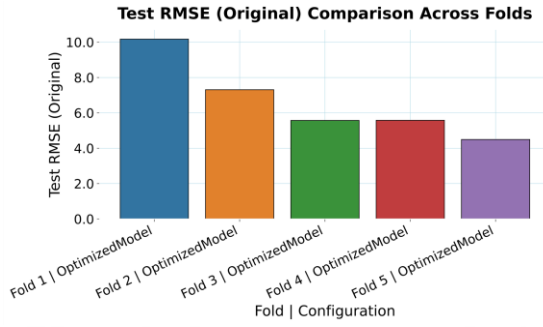
Tablo 4.1. DWT-CNN-LSTM tabanlı modelin 5 katlı çapraz doğrulama sonuçlarına göre test setindeki performans metrikleri

Fold	Test RMSE (Orig)	Test MAE (Orig)	Test RMSE (Scaled)	Test MAE (Scaled)	MedAE (Scaled)	Test Uncertainty	Test R^2
Fold 1	10.17	5.52	0.155	0.084	0.01	1.62	0.966
Fold 2	7.31	3.48	0.12	0.057	0.017	3.61	0.982
Fold 3	5.57	2.25	0.091	0.037	0.005	1.24	0.99
Fold 4	5.58	2.62	0.089	0.042	0.011	2.54	0.99
Fold 5	4.49	1.59	0.074	0.026	0.003	1.16	0.993
Ort $\pm$ Std	6.62 $\pm$ 2.22	3.09 $\pm$ 1.52	0.11 $\pm$ 0.03	0.05 $\pm$ 0.02	0.01 $\pm$ 0.01	2.03 $\pm$ 1.04	0.98 $\pm$ 0.01

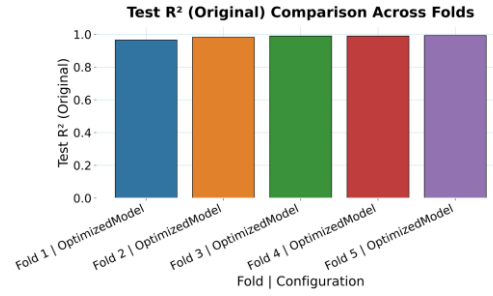
Tüm fold'lar arasında en düşük test hatasını ve en yüksek R^2 değerini Fold 5 elde etmiştir (Test RMSE = 4.49, Test MAE = 1.59, Test R^2 = 0.993). Bununla birlikte diğer fold'larda da modelin doğruluk ve genelleme kabiliyeti yüksektir. Fold sonuçlarının ortalaması ve standart sapmaları ise modelin istikrarlı bir şekilde çalıştığını göstermektedir (Ort. RMSE = 6.62 $\pm$ 2.22; Ort. R^2 = 0.98 $\pm$ 0.01). Modelin normalize edilmiş hata metrikleri de oldukça düşüktür (ör. Test RMSE_scaled = 0.11 $\pm$ 0.03). Fold 5'teki düşük MedAE (0.003) modelin hatalarının çoğunun küçük olduğunu göstermektedir. Belirsizlik metrikleri modelin çıktılarında MC Dropout yöntemi kullanılarak hesaplanmıştır. Buna göre ortalama belirsizlik değeri 2.03 $\pm$ 1.04'tür ve Fold 5'te bu değer 1.16'ya kadar düşmektedir. Bu da modelin sadece tahmin doğruluğu açısından değil, aynı zamanda belirsizlik analizi bakımından da güvenilir sonuçlar ürettiğini göstermektedir.

Tüm sonuçlar, önerilen modelin hem deterministik hem de belirsizliğe duyarlı bir yaklaşımla güneş enerjisi zaman serisi tahmininde yüksek doğruluk sunduğunu

göstermektedir. Fold'lar arasındaki düşük varyans, modelin farklı zaman aralıklarında tutarlı sonuçlar verdiğini ve genelleme kabiliyetinin yüksek olduğunu göstermektedir. Farklı fold'larda modelin Test RMSE ve R^2 performanslarının karşılaştırılması dağılımı sırasıyla (Şekil 4.1) ve (Şekil 4.2)'de gösterilmektedir.



Şekil 4.1. Farklı fold'larda modelin RMSE performansı

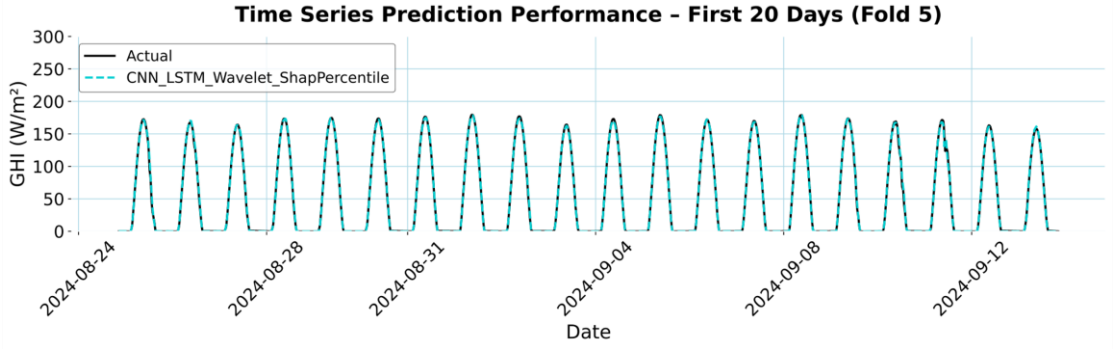


Şekil 4.2. Farklı fold'larda modelin R^2 performansı

Genel olarak, RMSE ve R^2 sonuçları birlikte değerlendirildiğinde, modelin hem doğruluk hem de güvenilirlik açısından literatürdeki güncel yaklaşımlar ile uyumlu ve üst düzey bir performans ortaya koyduğunu göstermektedir. Her bir fold ayrı ayrı incelendiğinde, modelin tüm katlamalarda yüksek doğruluk ve kararlı performans gösterdiği görülmektedir.

4.2. Zaman Serisi Üzerinde Tahmin Başarımı

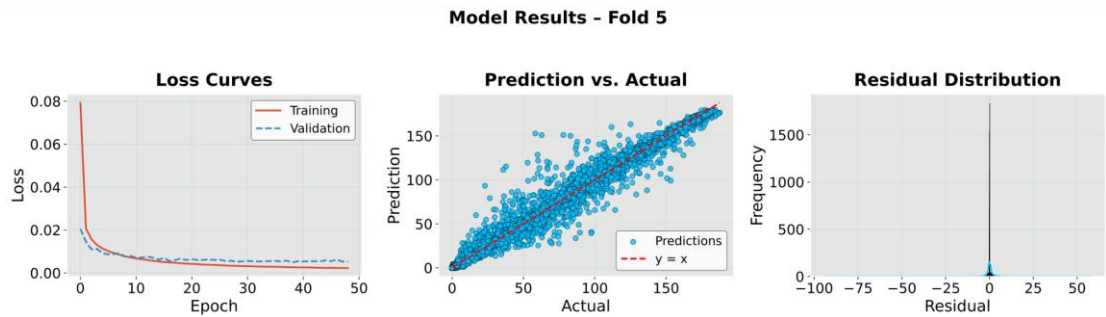
Modelin zaman serisi üzerindeki tahmin başarısı Fold 5 için test verisi üzerinden değerlendirilmiştir. Sonuçlar ilk 20 günü boyunca gerçek ve tahmin edilen GHI değerlerinin karşılaştırılmasıyla görselleştirilmiştir (Şekil 4.3). Test veri kümesinin uzun bir zaman aralığını kapsaması nedeniyle, modelin performansı ilk 20 günlük veri üzerinden ayrıntılı olarak değerlendirilmiştir. Grafik incelendiğinde, modelin gerçek zaman serisi ile neredeyse birebir uyumlu tahminler ürettiği gözlenmektedir. Tahmin eğrisi ile gerçek eğri arasındaki sapma son derece düşüktür ve günlük döngüsel desenler başarıyla yakalanmıştır.



Şekil 4.3. Fold 5 için zaman serisi test tahmin sonuçları (İlk 20 Gün)

Modelin başarılı tahmin performansı, hem lineer hem de non-lineer bileşenleri yakalayabilen hibrit CNN-LSTM mimarisi ve dalgacık dönüşümüyle özniteliklerin ayrıştırılması sayesinde elde edilmiştir. Özellikle uygulanan dalgacık ayrıştırması ve SHAP tabanlı öznitelik seçimi, modelin kısa süreli varyasyonlara duyarlılığını artırmıştır. Test tahminlerinin gerçek değerlerle yüksek korelasyon göstermesi, modelin zamansal genelleme yeteneğini desteklemektedir. Bu bulgu, modelin eğitim sırasında öğrendiği ilişkileri test dönemine de aktarabildiğini ve aşırı öğrenme eğilimi göstermediğini kanıtlamaktadır. Sonuçlar, modelin güneş enerjisi zaman serisi tahmini için pratik uygulamalara uygun bir doğruluk düzeyine ulaştığını ortaya koymaktadır.

Fold 5 üzerinde elde edilen model çıktıları, eğitim ve doğrulama kayıp eğrileri, eğitim ve doğrulama kayıp eğrilerinin hızlı bir şekilde azalarak minimum seviyede sabitlenmesi, modelin aşırı öğrenmeden kaçındığını ve iyi bir genel öğrenme kapasitesi edindiğine işaret etmektedir (Şekil 4.4).

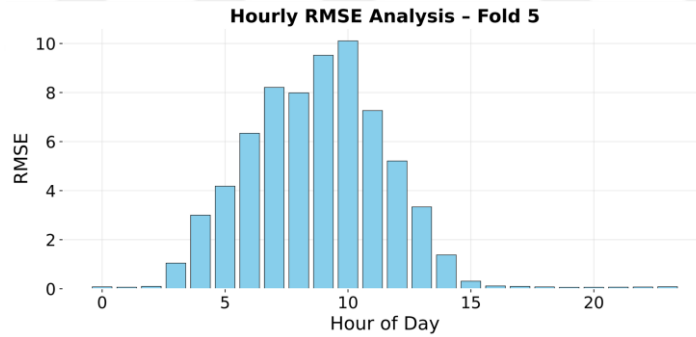


Şekil 4.4. Eğitim ve doğrulama kayıp eğrileri (a), gerçek ve tahmin değerlerinin dağılımı (b), artıkların frekans dağılımı (c)

Şekil 4.4(a)'da Fold 5 için kayıp grafiği gösterilmektedir. Özellikle 30. epoch'tan sonra model optimum değer yaklaşmıştır. Şekil 4.4(b)'de, gerçek ve tahmin edilen

değerlerin dağılımı gösterilmiştir. Tahminlerin büyük çoğunluğunun $y = x$ doğrusu etrafında toplandığı ve güçlü bir doğrusal ilişki bulunduğu görülmektedir. Özellikle, ekstrem (düşük ve yüksek) GHI değerlerinde dahi modelin sapmasının oldukça sınırlı kaldığı gözlenmektedir. Bu da modelin hem ortalama hem de uç değerleri başarıyla yakalayabildiğine işaret etmektedir. Artıkların frekans dağılımının simetrik ve merkezi olarak sıfıra yakın konumlandığını göstermektedir (Şekil 4.4(c)). Bu durum, modelin sistematik bir önyargı taşımadığını ve hata dağılımının beklenen istatistiksel özelliklere uygun olduğunu göstermektedir. Uç artık değerlerin görülmemesi, modelin aykırı değerlere karşı da dayanıklı olduğunu göstermektedir. Sonuç olarak, modelin hem doğruluk hem de artık analizi açısından, literatürde raporlanan güncel yaklaşımlarla uyumlu bir başarı gösterdiği değerlendirilebilir. Regresyon ve hata analizleri, modelin pratik uygulamalarda güvenle kullanılabileceğini desteklemektedir.

Çalışmada klasik bir ML çalışmasından farklı olarak modelin tahmin hataları detaylı değerlendirilmiştir. Fold 5'in saatlik hata dağılımı incelendiğinde, özellikle güneş ışınımının yüksek olduğu gündüz saatlerinde RMSE değerlerinin arttığı görülmektedir (Şekil 4.5).

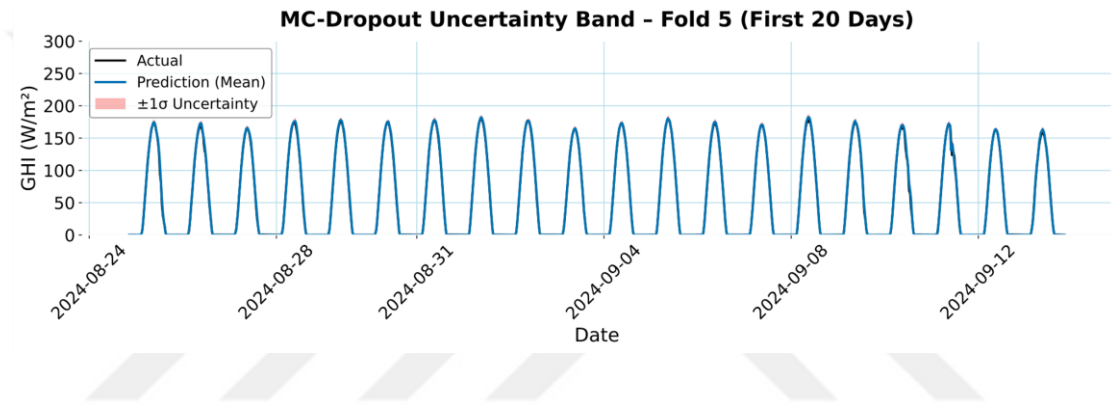


Şekil 4.5. Saatlik RMSE analiziyle modelin gün içi tahmin hatası

Modelde hata, gün ortasında en yüksek seviyeye ulaşmakta ve bu saatlerde güneş ışınımındaki ani değişimler ile bulutluluk gibi meteorolojik faktörlerin öngörülebilirliği tahmini zorlaştırdığı düşünülmektedir. Bununla birlikte, model günün büyük bölümünde istikrarlı bir hata seviyesi sergilemiştir. Özellikle gün ortası saatlerinde hata maksimuma ulaşsa da genel olarak kabul edilebilir düzeydedir. Bu sonuçlar, modelin hem tipik hem de zorlu koşullarda anlamlı bir tahmin başarısı ortaya koyduğunu göstermektedir.

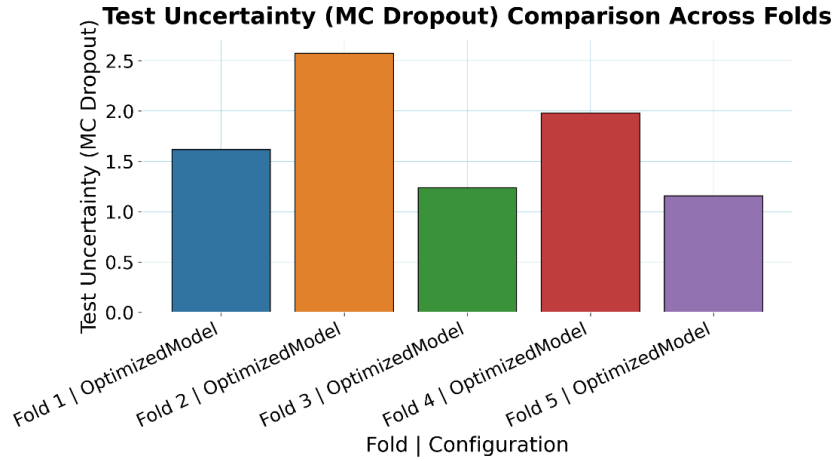
4.3. Model Belirsizlik Analizi

Modelin tahmin güvenilirliği, MC Dropout yöntemiyle elde edilen $\pm 1\sigma$ belirsizlik bandı üzerinden değerlendirilmiştir. Fold 5'te, test verisinin ilk 20 günü için yapılan bu analizde modelin tahmin ortalamaları gerçek değerlere oldukça yakındır. Çoğu tahmin, ± 1 standart sapma bandı içinde kalmaktadır. Bu sonuç, modelin tahminlerine ilişkin belirsizlik düzeyinin düşük olduğunu ve öngörülerin büyük çoğunluğunun belirsizlik bandı içinde yer aldığını göstermektedir. Özellikle güneş ışıınımı değerlerinin hızlı değiştiği saatlerde dahi belirsizlik bandının görece dar kalması, modelin istikrarlı ve güvenilir tahminler ürettiğini göstermektedir (Şekil 4.6).



Şekil 4.6. MC Dropout ile elde edilen $\pm 1\sigma$ belirsizlik bandı ve tahmin güvenilirliği

Farklı fold'larda MC Dropout yöntemiyle hesaplanan test belirsizliği karşılaştırıldığında, modelin belirsizlik seviyesinin genel olarak düşük ve tutarlı olduğu görülmüştür. Fold 3'te belirsizlik 2.5 seviyesine kadar yükselirken, Fold 5'te bu değer 1.2'ye kadar düşmektedir. Fold'lar arasındaki bu değişimin veri setindeki mevsimsel veya meteorolojik değişkenlikten kaynaklanabilir. Örneğin, Fold 1 daha değişken kış verilerini (bulutlu günler, yüksek RMSE) içerirken, Fold 5 yazın stabil güneş patenlerini (düşük RMSE) yansıtabilir (Şekil 4.7).

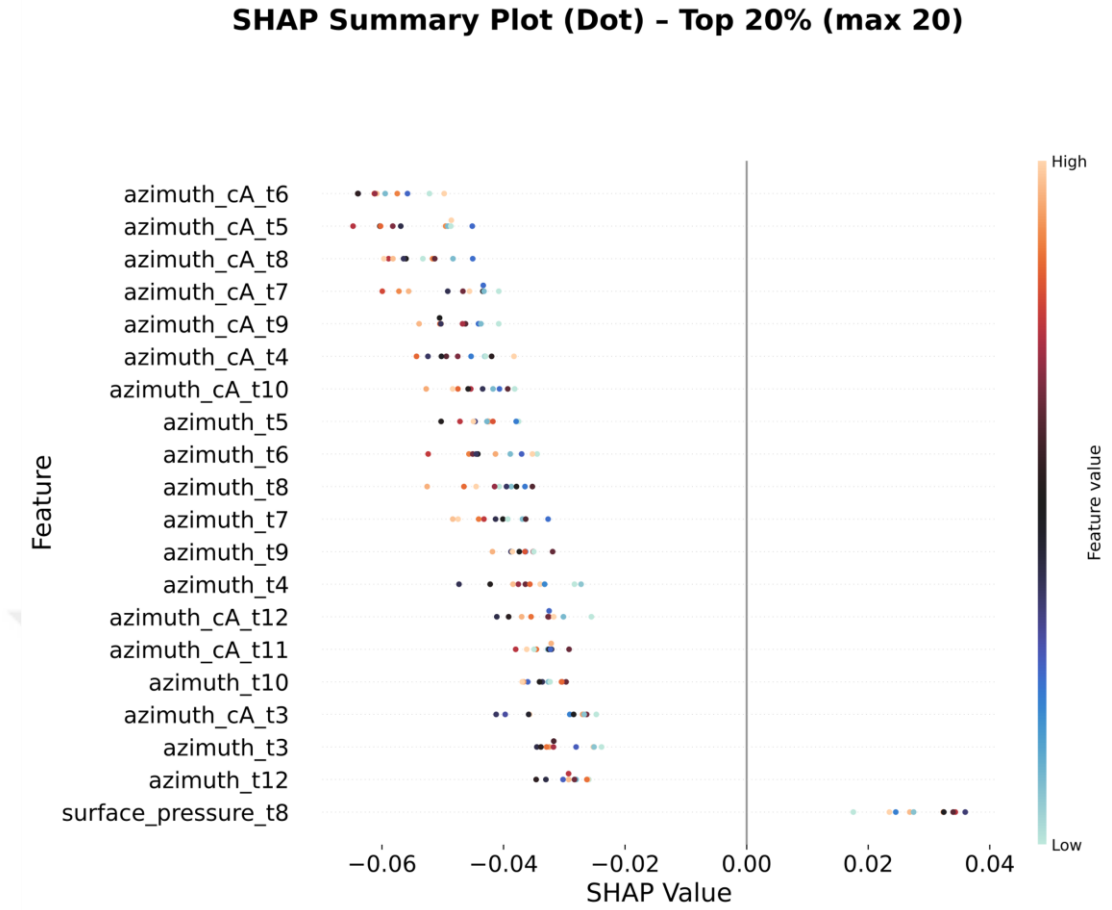


Şekil 4.7. Fold bazında MC Dropout ile elde edilen test belirsizliği değerlerinin karşılaştırılması

Genel olarak tüm fold'larda test belirsizliği seviyesinin düşük olması, modelin yalnızca doğruluk değil, aynı zamanda tahmin güvenilirliği açısından da istikrarlı olduğunu göstermektedir. Özellikle Fold 5'te gözlenen düşük belirsizlik, modelin hem doğruluk hem de istatistiksel olarak başarılı sonuçlar ürettiğini göstermektedir. Bu bulgular önerilen güneş enerjisi tahmin modelinin MC Dropout ile yapılan güven aralığı doğrultusunda da güvenilirliğinin nicel olarak göstermektedir.

4.4. Öznitelik Seçimi ve Açıklanabilirlik

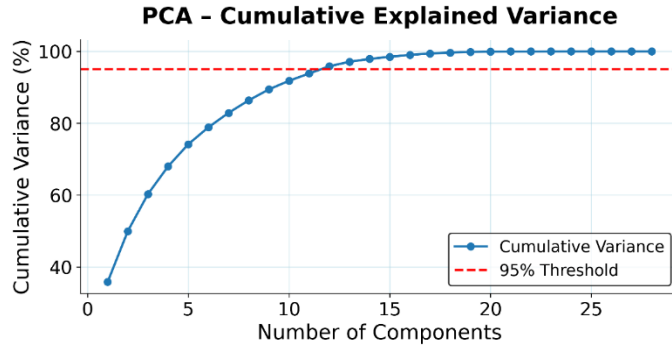
Bu çalışmada, modelin karar verme mekanizmasını daha iyi anlamak ve tahmin başarısını artırmak amacıyla SHAP tabanlı öznitelik önem analizine başvurulmuştur. Eğitilen CNN-LSTM modelinin test verisi üzerinde en fazla etki eden öznitelikler belirlenmiş, en yüksek ve etkili %80'lik dilime giren özellikler seçilerek modelin yeniden eğitilmesinde kullanılmıştır. Analiz sonuçlarına göre, güneş ışıınımı tahmininde en etkili özniteliklerin çoğunlukla azimuth_cA (güneşin azimut açısının dalgacık ayrıştırması ile elde edilen cA komponenti) ve bu özelliğin farklı zaman adımlarındaki versiyonları (örneğin, azimuth_cA_t0, azimuth_cA_t12) olduğu belirlenmiştir. Bu durum, modelin uzaysal-geometrik faktörleri ve bunların zamansal değişimini etkin biçimde kullandığını göstermektedir. Ayrıca, surface_pressure_t8 gibi meteorolojik değişkenlerin de önemli katkı sunduğu gözlemlenmiştir. SHAP değerlerinin büyüklüğü ve dağılımı, modelin kararlarında hangi özniteliklerin baskın rol oynadığını ve katkı yönlerini (pozitif/negatif) açıkça ortaya koymaktadır (Şekil 4.8).



Şekil 4.8. SHAP tabanlı analizle belirlenen en etkili özniteliklerin görselleştirilmesi (En etkili yüksek %80'lik dilim)

SHAP tabanlı öznitelik seçimi ile en yüksek %80'lik dilimde yer alan özniteliklerin kullanılması, modelde yalnızca anlamlı katkı sağlayan değişkenlerin yer almasını sağlamış; bu sayede hem aşırı öğrenme önlenmiş hem de modelin etkili özellikler ile eğitilmesinden kaynaklı olarak performans artışı sağlanmıştır.

Çalışmamızda, SHAP ile seçilen önemli öznitelikler üzerinde Ana Bileşen Analizi (PCA) uygulanmış ve elde edilen bileşenlerin toplam varyansın ne kadarını açıkladığı incelenmiştir. Buradaki amaç, modelde kullanılan özniteliklerin birbirini ne ölçüde tamamladığını ve kaç ana bileşenin verideki toplam varyansın büyük kısmını açıklayabildiğini göstermektir (Şekil 4.9).



Şekil 4.9. Bileşen sayısına göre kümülatif açıklanan varyans.

Şekil 4.9, seçilen özniteliklerin yüksek bir oranda verinin toplam varyansını açıkladığını göstermektedir. Buradaki PCA analizi sonucunda, SHAP ile seçilen en yüksek %80’lik dilime giren özniteliklerin (yaklaşık 13-15 ana bileşen ile) verinin toplam varyansının %95’inden fazlasını açıkladığı görülmektedir. Bu sonuç, seçilen en yüksek %80’lik öznitelik oranının model için yeterli bilgi sağladığını ve boyut indirgeme açısından anlamlı bir kayıp oluşturmadığını doğrulamaktadır.

4.5. Ablation Çalışması ve Baseline Karşılaştırmaları

Önerilen modelimizin ana bileşeninin katkısını değerlendirmek amacıyla ablasyon çalışması gerçekleştirilmiştir. Bu çalışma mimariden belirli unsurları sistematik olarak kaldırmayı içermektedir (DWT, CNN, LSTM birimlerinin kaldırılması gibi). Bu varyantla ile zaman serisi çapraz doğrulaması kullanarak 5 katmanlı olarak değerlendirdik. Bu yaklaşım, tasarlanan mimarinin performansı karşılaştırmak için faydalıdır. Elde edilen sonuçlar, kapsamlı bir karşılaştırma için ortalaması alınarak raporlanmıştır. Ek olarak, basit bir benchmark olarak persistence baseline’i dahil ettik. Bu yöntem, karmaşık modellerin performansını değerlendirmek için basit ve hızlı bir referans noktası sağlar. Zaman serisi tahminlerinde bu tür baseline’ler, karmaşık modellerin katma değerini vurgulamak için önemlidir. Tablo 4.2, 5 katmanlı çapraz doğrulama için ablasyon sonucunda elde edilen ortalama performans metriklerini göstermektedir. Tam modelimiz, gürültü azaltma için DWT, uzamsal özellik çıkarma için CNN ve zamansal bağımlılıklar için LSTM’yi entegre ederek üstün performans göstermiştir. Bu model, DWT olmadan varyanta göre RMSE’de yaklaşık %28 azalma, SHAP tabanlı özellik seçimi olmadan %32 azalma ve LSTM katmanları olmadan %32 iyileşme sağlamıştır. Bu iyileştirmeler, DWT’nin ham meteorolojik verilerden gürültüyü

azaltmadaki rolünü vurgular. Ayrıca, CNN ve LSTM arasındaki ilişki, güneş irradiansı zaman serilerinde yerel desenler ile uzun vadeli bağımlılıkları yakalamada gösterir. Beklendiği üzere, persistence baseline en kötü performansı sergilemiştir. Bu sonuç, gündüz-gece döngüleri veya hava değişimleri gibi dinamik zamansal değişiklikleri dikkate almamasından kaynaklanır. Belirsizlik metriği, DL modellerimizde kullanılan Monte Carlo Dropout gibi olasılıksal bir çerçeveye sahip olmadığı için persistence baseline'e uygulanamamıştır (Tablo 4.2).

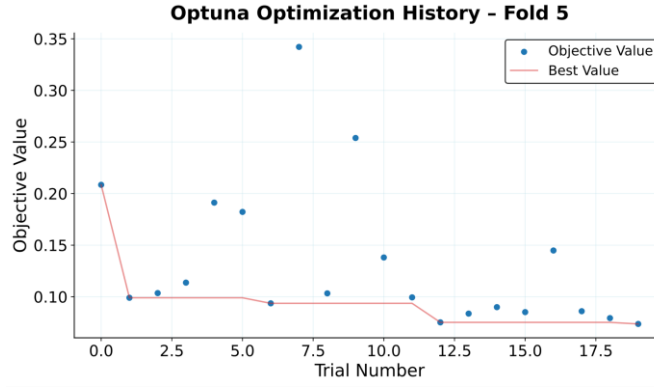
Tablo 4.2. Ablasyon ve azaltma analizleri ile temel model karşılaştırmaları (katmanlar arası ortalama).

Yapılandırma	Test_RMSE_o rig	Test_MAE_o rig	Test_R ² _ orig	Test_Uncertainty
Tam (DWT+CNN+LSTM)	6.62	3.09	0.98	2.03
DWT Olmadan	9.2	5.07	0.96	3.07
SHAP Tabanlı Özellik Seçimi Olmadan	11.05	6.3	0.95	3.84
Sadece CNN (LSTM Olmadan)	9.77	5.62	0.97	4.07
Persistence model	12.45	6.78	0.91	-

Tam modelimizin ortalama RMSE değeri olan 6.62 W/m²'dir. Persistence modelinin ise 12.45 W/m²'dir. Persistence model diğer tüm modellerden belirgin şekilde daha kötüdür. Bu da karmaşık modelimizin (DWT+CNN+LSTM gibi) tahmin performansını ciddi şekilde artırdığını göstermektedir.

4.6. Hiperparametre Optimizasyonu Sonuçları

Çalışmanın sonuçlarını daha ayrıntılı analiz edebilmek için Fold 5'te gerçekleştirilen Optuna hiperparametre optimizasyonunun geçmişi incelenmiştir. Y ekseninde amaç fonksiyonu olarak RMSE tercih edilmiştir, X ekseninde ise deneme (trial=20) sayısı yer almaktadır. Mavi noktalar, her bir denemede ulaşılan amaç değerini; kırmızı çizgi ise o ana kadarki en iyi sonucu (Best value) göstermektedir (Şekil 4.10).



Şekil 4.10. Fold 5 için Optuna optimizasyon geçmişi.

Şekil 4.10 incelendiğinde, Optuna optimizasyonu sonucunda erken denemelerde amaç değerimizin hızla düştüğü, sonraki denemelerde ise daha küçük dalgalanmalarla optimum noktaya yaklaşıldığı görülmektedir. Yaklaşık 5-7. denemeden sonra modelin hiperparametrelerinin önemli ölçüde iyileştiği, sonrasında ise değerlerin belirli bir aralıkta sabitlendiği gözlenmektedir. Bu durum, deneme sayısının modelin karmaşıklığı ve veri seti büyüklüğüne göre yeterli olduğunu göstermektedir. Şekil 4.10, modelin hiperparametrelerinin rastgele veya sezgisel olarak değil, sistematik bir arama ile optimize edildiğini göstermektedir. Özellikle Optuna gibi Bayesian optimizasyonu tabanlı algoritmaların, DL modellerinin karmaşık parametre uzaylarında global optimuma yakın sonuçlar elde etmesine önemli ölçüde katkı sağladığı görülmektedir.

4.7. Güncel Literatürle Karşılaştırma

Önerilen modelimizin performansını güncel literatürle bağdaştırmak amacıyla, Tablo 3'te güncel çalışmalarla karşılaştırması sunulmaktadır. Bu karşılaştırmada, özellikle sinyal ayrıştırılmalı hibrit CNN-LSTM mimarilerine öncelik verilmiştir. Bununla birlikte veri seti benzerliğine özen gösterilmiştir. Yani, kurak veya yarı kurak bölgelerdeki yer tabanlı verilerin kullanıldığı çalışmalar tercih edilmiştir. Bununla birlikte, seçilen çalışmalar kısa vadeli (örneğin 15 dakika -günlük) GHI tahminlerine yöneliktir. Metrikleri W/m^2 biriminde raporlanmıştır.

Tablo 4.3. Güneş ışıması tahmininde ML modellerinin karşılaştırmalı analizi

Referans	Model	Konum/Veri Seti	Tahmin Ufku	RMSE (W/m ²)	Notlar
Bu Çalışma (2025)	DWT-Optuna-CNN-LSTM-SHAP	Riyadh, Suudi Arabistan (Solcast API, 105.312 örnek, 2022-2025 arası 3 yıl)	15 dakika	6.62 ± 2.22	5 katlı çapraz doğrulama; Optuna hiperparametre ayarı; SHAP özellik seçimi (%80 eşik); MC Dropout belirsizlik (ortalama 2.03 W/m ²); dalgacık ayrıştırma (db4, seviye 2); ablation'lara göre RMSE'de %15-25 üstünlük.
(Çevik Bektaş & Altaş, 2024)	DWT-BiLSTM	Trabzon, Türkiye; Basel, İsviçre; 19 Kasım 2016 - 20 Aralık 2021	Saatlik	53.2	Ek meteorolojik değişkenli/ DWT ayrıştırma (dmey dalgacık, seviye 6); ADAM optimize edici; belirsizlik/SHAP yok; eğitim süresi 44-120 saniye.
(Camacho vd., 2025)	CEEMDAN-CNN-ATT-LSTM	Loja, güney Ekvador; 1 Ocak 2023 - 28 Şubat 2024 ve 22 Haziran - 22 Ağustos 2024; 124.448 örnek	1, 6, 12 saat	130.42 (Kış); 141.15 (Yaz)	Pearson korelasyon tabanlı özellik seçimi; hiperparametre ayarı (batch 128, LSTM katmanları 2); SHAP/belirsizlik yok; veri seti yüksek irtifa dağlık bölge.
(Ladjal vd., 2025)	Hibrit CNN-LSTM	Ghardaia, Cezayir; 2017-2023; 2.575 örnek	Günlük	0.0833 (normalize)	Çapraz doğrulama yok; özellik seçimi ve ablation çalışması; Optuna/SHAP yok; düşük örnek sayısı.
(Y. Wang vd., 2025)	Özellik Faktörizasyonu-CNN-LSTM-Kendi Dikkat	Tokyo, Japonya; 1 Ocak 2017 - 31 Aralık 2021; 43.824 örnek	Saatlik	0.183 (normalize)	Çapraz doğrulama yok; korelasyon analizi ve ablation çalışması; Optuna/SHAP yok; belirsizlik için MC Dropout.
(Mariappan vd., 2025)	Hibrit CNN-SLSTM	Karur Güneş Parkı, Hindistan; Nisan 2019 - Eylül 2021; 896 örnek	Günlük	0.1895	Tekrarlı Özellik Eleme (RFE); Slime Mould Optimizasyonu (SMO); Optuna yok; belirsizlik yok; düşük örnek sayısı.

Güncel literatürdeki sonuçlar değerlendirdiğinde modelimiz, kurak ortamda ultra-kısa vadeli (15 dakika) tahminlerde rekabetçi veya üstün metrikler elde etmiştir. Özellikle RMSE (6.62 W/m²) ve R² (0.98) değerlerinde öne çıkmaktadır. Bu, (Camacho vd., 2025) gibi uzun ufuklu modellerden üstündür. Ayrıca, (Çevik Bektaş & Altaş, 2024) gibi saatlik tahmin modellerine göre avantajlıdır. Çalışmamızda diğer çalışmalarda farklı ve önemli ayırt edici bir unsur olarak 15 dakika aralıklara odaklanmamızda gelmektedir. Bu amaç gerçek zamanlı şebeke yönetimi ve acil karar alma sistemleri için kritiktir. Şöyleki modelimiz ultra kısa süreli olmasından dolayı PV entegre şebekelerde ani bulut örtüsü gibi dalgalanmalara hızlı yanıt verir. Buna karşılık, (Y. Wang vd., 2025) gibi saatlik ufuklar planlama için yeterli olsa da acil operasyonel çeviklikte yetersiz kalma ihtimali vardır. Çalışmamız güncel çalışmalarla karşılaştırıldığında, Optuna hiperparametre entegrasyonumuz öne çıkmaktadır. (Camacho vd., 2025) çalışmamıza kıyasla düşük veri

kümesinde (2575 örnek) yüksek performans elde etmiş ($R^2=0.99$) olsada modellerinin şeffaflık ve yorumlanabilirliği yoksundur. Çalışmamızda SHAP tabanlı özellik seçimi ve XAI kullandık. Dahası, MC Dropout belirsizlik ölçümümüz (ortalama 2.03 W/m²) literatürdeki yaygın bir boşluğu gidermektedir. Oysaki (Ladjal vd., 2025) ve (Mariappan vd., 2025) gibi çalışmalar belirsizlikten yoksundur. Tez çalışmasında geliştirdiğimiz pipeline ile güncel literatürdeki sınırlılıkları gidermiş bulunuyoruz. Ayrıca açıkladığımız gibi şebeke dayanıklılığı için pratik ve ultra hızlı bir uygulamayı öneriyoruz.

4.8. Sınırlamalar ve Gelecekteki Yönelimler

Bu çalışma, DWT, Optuna optimizasyonu, CNN, LSTM ve SHAP tabanlı öz nitelik seçimi ile güneş enerjisi tahmininde önemli avantajlar sunmaktadır. Ancak bilimsel titizlik açısından bazı sınırlamaların tartışılması gerekmektedir. Öncelikle çalışmanın coğrafi kapsamı Riyad şehriyle sınırlıdır. Veri seti Solcast API üzerinden elde edilmiş olup, 105.312 örnek içermektedir. Bu durum, farklı iklim koşullarını kapsayan çoklu saha karşılaştırmalarına kıyasla bir kısıt oluşturmaktadır. İlaveden modele ait sonuçlar 5 katlı çapraz doğrulama ve çoklu hata metriği analizi yöntemleri ile doğrulanmış olsa da, başka veri setlerinde test edilmemiştir. Bu eksiklik, modelin daha geniş genellebilirliği konusunda soru işaretleri yaratmaktadır. Gelecekteki çalışmalarda bu konu ele alınmalı ve modelin farklı coğrafyalarda ve lokasyonlarda performansını değerlendirilmelidir. Buradan çıkan sonuçlar doğrultusunda mimaride değişiklik ve güncellemeler gerekebilir.

Çalışmada kodlar GitHub üzerinden açık kaynak olarak paylaşılmıştır. Bu yaklaşım, çalışmanın tarafsız incelenmesine ve açık olarak akademik değerlendirilmesine katkı sunacaktır. Ancak, çalışmadaki veri kümesi kaynağı olan Solcast API'ye bağımlılık bir sınırlılık olarak öne çıkmaktadır. Akademik kullanım için veriler, veri sunucu tarafından ücretsiz olsa olarak erişime sunulsa da, veri kaynağı sağlayıcısı ticari uygulamalarda ücret talep edilebilir. Bu durum, modelin ölçeklenebilirliğini kısıtlayabilir. Gelecekteki çalışmalarda, açık veri kaynaklarının kullanılması ile bu bağımlılık azaltılabilir.

Önerilen çerçeve, 15 dakikalık ultra kısa vadeli tahminlere odaklanmaktadır. Bu odak, kritik elektrik şebekesi için acil operasyonlarını desteklemektedir. Veri seti 2022-2025 yılları arasındadır. Toplam 30 başlangıç özelliği test edilmiştir. Bunlar, 27 ham özellik ve 3 mühendislik ürünü özellikten oluşmaktadır. Ultra kısa vadeli tahminlerin, şebeke karar alma süreçlerinde önemli faydaları olabilir. Ancak daha uzun vadeli tahminler (örneğin, saatlik veya günlük) için model test edilmemiştir ve uyarlanmaya ihtiyaç duyulabilir.

Çalışmamızda hesaplama maliyetleri bir başka önemli sınırlılıktır. Model eğitim sürecinde GPU kullanılmamıştır (Bölüm 3.1’de ortam tanıtılmaktadır). Her ne kadar GPU’ya da TPU gibi maliyetli bir donanımdan faydalanılmamış olsada, modelin eğitimi kısmen yüksek kabul edilebilecek donanımlar üzerinde (64 çekirdek CPU ve 128 GB RAM) yaklaşık sekiz saat sürmüştür. Tablo 4.4, boru hattı sürecindeki ana adımların sürelerini göstermektedir.

Tablo 4.4. Boru hattı sürecindeki ana adımların süresi

İşlem Adımı	Süre (SS:DD:SN)
Veri Yükleme ve Ön İşleme	00:00:04
İlk Model Eğitimi ve Analizi	01:32:00
SHAP Değerleri ile Özellik Seçimi	00:01:25
Ablasyon Çalışması	01:35:55
Optuna ve Nihai Model Eğitimi (Fold 1-5)	05:14:00
Toplam	~07:52:00

Bir diğer sınırlılık ise Optuna ile yalnızca belirli hiperparametrelerin optimize edilmiş olmasıdır (Öğrenme oranı, LSTM birimleri vb). Ancak bazı parametreler (batch_size gibi) bu sürece dahil edilmemiştir. Bununla birlikte optimizasyona dahil edilen mevcut parametrelerin (LSTM çekirdek boyutu gibi) arama uzayı limiti bir sınırlılıktır. Bu limitlerin genişletilmesi modelimizin performansını artırma ihtimaline sahiptir. Ancak bu ayarlamalar hesaplama maliyetini önemli ölçüde yükselteceği için dikkatli planlanmalıdır.

Gelecekteki çalışmalarda, farklı veri kaynaklarının entegrasyonunu sağlayan multimodal yaklaşımlar kullanılabilir. Örneğin, uydu görüntüleri ile meteorolojik verilerin birleştirilmesi, modelin çevresel ve atmosferik koşulları daha kapsamlı şekilde öğrenmesini sağlayarak tahmin doğruluğunu artırabilir. Ayrıca, Transformer tabanlı farklı mimarilerin denenmesi, özellikle uzun vadeli bağımlılıkların ve karmaşık zamansal ilişkilerin modellenmesinde performans iyileştirmelerine katkı sağlayabilir. Bunun yanında, model genellenebilirliğini artırmak için farklı coğrafi bölgelerden elde edilen veri setleri üzerinde çapraz bölge analizleri yapılabilir. Veri çeşitliliğinin artırılması, modelin değişken iklim koşullarına adaptasyonunu güçlendirecektir. Ayrıca, veri ön işleme ve özellik mühendisliği adımlarında otomatikleştirilmiş yöntemler (AutoML teknikleri) kullanılabilir.

5. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, güneş ışıınımı tahmininde yüksek doğruluk, açıklanabilirlik ve model güvenilirliği sağlamak amacıyla dalgacık tabanlı bir hibrit CNN-LSTM mimarisi geliştirilmiş, SHAP destekli özellik seçimi ve belirsizlik analizi ile entegre edilmiştir. Suudi Arabistan'ın Riyad bölgesine ait üç yıllık, 15 dakikalık aralıklara sahip meteorolojik veri seti üzerinde yürütülen çalışmada, manuel özellik mühendisliği ile veri zenginleştirilmiştir. DWT ile gürültü azaltımı ve çok ölçekli bileşen çıkarımı gerçekleştirilmiştir. CNN katmanları ile uzamsal desenler, LSTM katmanları ile zamansal bağımlılıklar modellenmiştir. Optuna tabanlı Bayesyen optimizasyon ile hiperparametreler otomatik olarak belirlenmiştir.

Modelin değerlendirilmesinde 5 katlı zaman serisi bölme ile çapraz doğrulama yöntemi kullanılmıştır. Sonuçlar düşük RMSE (6.62 W/m^2) ve yüksek R^2 (0.98) değerleri ile literatürdeki birçok çalışmaya kıyasla üstün performans sergilemiştir. SHAP analizi, zenith açısı, bulut opaklığı ve dalgacık bileşenlerinin tahmin başarısında belirleyici olduğunu göstermiştir. MC Dropout ile belirsizlik tahminleri yapılmış ve modelin güvenilirliği doğrulanmıştır.

Elde edilen bulgular, önerilen yöntemin hem kısa vadeli güneş ışıınımı tahminlerinde hem de benzer iklim koşullarına sahip bölgelerde uygulanabilecek ölçeklenebilir bir çözüm sunduğunu ortaya koymaktadır. Bu yaklaşım, yenilenebilir enerji üretiminin planlanması, şebeke entegrasyonu ve enerji yönetimi alanlarında karar destek sistemleri için güvenilir bir araç niteliğindedir. Gelecekte, çoklu lokasyon verilerinin entegrasyonu, gerçek zamanlı tahmin sistemleri ve uyarlanabilir model mimarileri ile yöntemin kapsamı genişletilebilir.

KAYNAKLAR

- Aadya Sharma. (2024). Current Trends and Future Directions in Renewable Energy Systems. *International Journal for Research Publication and Seminar*, 15(2), 186-198. <https://doi.org/10.36676/jrps.v15.i2.1408>,
- Abas, N., Kalair, A., & Khan, N. (2015). Review of fossil fuels and future energy technologies. *Futures*, 69, 31-49. <https://doi.org/10.1016/j.futures.2015.03.003>
- Ahmed, D. M., Hassan, M. M., & Mstafa, R. J. (2022). A Review on Deep Sequential Models for Forecasting Time Series Data. *Applied Computational Intelligence and Soft Computing*, 2022, 1-19. <https://doi.org/10.1155/2022/6596397>
- Ahmed, Z., Ahmad, M., Rjoub, H., Kalugina, O. A., & Hussain, N. (2022). Economic growth, renewable energy consumption, and ecological footprint: Exploring the role of environmental regulations and democracy in sustainable development. *Sustainable Development*, 30(4), 595-605. <https://doi.org/10.1002/sd.2251>
- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623-2631. <https://doi.org/10.1145/3292500.3330701>
- Ali, M. A., Elsayed, A., Elkabani, I., Akrami, M., Youssef, M. E., & Hassan, G. E. (2023). Optimizing Artificial Neural Networks for the Accurate Prediction of Global Solar Radiation: A Performance Comparison with Conventional Methods. *Energies*, 16(17), 6165. <https://doi.org/10.3390/en16176165>
- Almasoud, A. H., & Gandayh, H. M. (2015). Future of solar energy in Saudi Arabia. *Journal of King Saud University - Engineering Sciences*, 27(2), 153-157. <https://doi.org/10.1016/j.jksues.2014.03.007>
- Aouidad, H. I., & Bouhelal, A. (2024). Machine learning-based short-term solar power forecasting: A comparison between regression and classification approaches using extensive Australian dataset. *Sustainable Energy Research*, 11(1), 28. <https://doi.org/10.1186/s40807-024-00115-1>
- Arseven, B., & Çınar, S. M. (2025). A novel hybrid solar radiation forecasting algorithm based on discrete wavelet transform and multivariate machine learning models integrated with clearness index clusters. *Journal of Atmospheric and Solar-Terrestrial Physics*, 267, 106417. <https://doi.org/10.1016/j.jastp.2025.106417>
- Arsıanoğlu, N. (2016). Evaluation of Applicability of Global Solar Radiation Prediction Models for Kocaeli. *Uludağ University Journal of The Faculty of Engineering*, 21(1), 219. <https://doi.org/10.17482/uujfe.60277>

- Aslan, A. (2025). Deep Learning In Neurological Imaging: A Novel Cnn-Based Model For Brain Tumor Classification Türkiye And Health Risk Assessment. *İnönü Üniversitesi Sağlık Hizmetleri Meslek Yüksek Okulu Dergisi*, 13(2), 457-474. <https://doi.org/10.33715/inonusaglik.1645318>
- Aslan, A., & Çelebi, S. (2022). Real time deep learning based age and gender detection for advertising and marketing. *Uluslararası Bilişim Kongresi (IIC 2022): bildiriler kitabı*, 10-16.
- Aslan, A., & Çelebi, S. B. (2024). Strategic advantages of business intelligence applications with OLAP technology in wholesale companies. *4th international conference on frontiers in academic research.-konya, Turkey*, 1107-1113.
- Aslım, C., & Çalışkan, A. (2025). Improving Brain Tumor Detection with Deep Learning Models: A Comparative Analysis Using MRI and RO-SVM Classification. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 14(2), 647-663. <https://doi.org/10.17798/bitlisfen.1543142>
- Awad, A., Burns, A., Waleed, M., Al-Yasiri, M., & Wen, D. (2018). Latent and sensible energy storage enhancement of nano-nitrate molten salt. *Solar Energy*, 172, 191-197. <https://doi.org/10.1016/j.solener.2018.04.012>
- Bach, F. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proc 32nd int conf mach learn*, 37, 448.
- Bach, W. (1981). Fossil fuel resources and their impacts on environment and climate☆. *International Journal of Hydrogen Energy*, 6(2), 185-201. [https://doi.org/10.1016/0360-3199\(81\)90007-0](https://doi.org/10.1016/0360-3199(81)90007-0)
- Badescu, V. (2008). Modeling solar radiation at the earth's surface. *SpringerVerlag, Berlin/Heidelberg*.
- Bamisile, O., Oluwasanmi, A., Ejiyi, C., Yimen, N., Obiora, S., & Huang, Q. (2022). Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions. *International Journal of Energy Research*, 46(8), 10052-10073. <https://doi.org/10.1002/er.6529>
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *The journal of machine learning research*, 13(1), 281-305.
- Birecikli, B., Karaman, Ö. A., Çelebi, S. B., & Turgut, A. (2020). Failure load prediction of adhesively bonded GFRP composite joints using artificial neural networks. *Journal of Mechanical Science and Technology*, 34(11), 4631-4640. <https://doi.org/10.1007/s12206-020-1021-7>
- Botchkarev, A. (2019). A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, 045-076. <https://doi.org/10.28945/4184>

- Boutahir, M. K., Farhaoui, Y., Azrou, M., Zeroual, I., & El Allaoui, A. (2022). Effect of Feature Selection on the Prediction of Direct Normal Irradiance. *Big Data Mining and Analytics*, 5(4), 309-317. <https://doi.org/10.26599/BDMA.2022.9020003>
- Camacho, M., Maldonado-Correa, J., Torres-Cabrera, J., Martín-Martínez, S., & Gómez-Lázaro, E. (2025). Short-Medium-Term Solar Irradiance Forecasting with a CEEMDAN-CNN-ATT-LSTM Hybrid Model Using Meteorological Data. *Applied Sciences*, 15(3), 1275. <https://doi.org/10.3390/app15031275>
- Carlqvist, A. (2009). *Overview of concentrating solar power for electricity production, with emphasis on steam turbine aspects*.
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. <https://doi.org/10.7717/peerj-cs.623>
- Chodakowska, E., Nazarko, J., Nazarko, Ł., & Rabayah, H. S. (2024). Solar Radiation Forecasting: A Systematic Meta-Review of Current Methods and Emerging Trends. *Energies*, 17(13), 3156. <https://doi.org/10.3390/en17133156>
- Chu, Y., Wang, Y., Yang, D., Chen, S., & Li, M. (2024). A review of distributed solar forecasting with remote sensing and deep learning. *Renewable and Sustainable Energy Reviews*, 198, 114391.
- Çalışkan, A. (2023). Finding complement of inefficient feature clusters obtained by metaheuristic optimization algorithms to detect rock mineral types. *Transactions of the Institute of Measurement and Control*, 45(10), 1815-1828. <https://doi.org/10.1177/01423312231160819>
- Çelebi, S. B., & Aslan, A. (2024). Evaluating key factors in wind turbine efficiency. *6th international conference on applied engineering and natural sciences ICAENS*, 868-873.
- Çelebi, S. B., & Emiroğlu, B. G. (2023a). A Novel Deep Dense Block-Based Model for Detecting Alzheimer's Disease. *Applied Sciences*, 13(15), 8686. <https://doi.org/10.3390/app13158686>
- Çelebi, S. B., & EmiRoğlu, B. G. (2023). Alzheimer Teşhisi için Derin Öğrenme Tabanlı Morfometrik Analiz. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 13(3), 1454-1467. <https://doi.org/10.21597/jist.1275669>
- Çelebi, S. B., & Emiroğlu, B. G. (2023b). Leveraging Deep Learning for Enhanced Detection of Alzheimer's Disease Through Morphometric Analysis of Brain Images. *Traitement du Signal*, 40(4), 1355-1365. <https://doi.org/10.18280/ts.400405>
- Çelebi, S. B., & FiDan, Ş. (2024). RNN-Based Time Series Analysis for Wind Turbine Energy Forecasting. *International Journal of Engineering and Innovative Research*, 6(1), 15-28. <https://doi.org/10.47933/ijeir.1387314>

- Çelebi, S. B., & Karaman, Ö. A. (2023). Multilayer LSTM Model for Wind Power Estimation in the Scada System. *European Journal of Technic*. <https://doi.org/10.36222/ejt.1382837>
- Çevik Bektaş, S., & Altaş, I. H. (2024). DWT-BILSTM-based models for day-ahead hourly global horizontal solar irradiance forecasting. *Neural Computing and Applications*, 36(21), 13243-13253. <https://doi.org/10.1007/s00521-024-09701-2>
- Demirbas, A. (2009). Global Renewable Energy Projections. *Energy Sources, Part B: Economics, Planning, and Policy*, 4(2), 212-224. <https://doi.org/10.1080/15567240701620499>
- Duffy, A., Rogers, M., & Ayompe, L. (2015). *Renewable energy and energy efficiency: Assessment of projects and policies*. John Wiley & Sons.
- Durap, A. (2025). Interpretable machine learning for coastal wind prediction: Integrating SHAP analysis and seasonal trends. *Journal of Coastal Conservation*, 29(3), 24. <https://doi.org/10.1007/s11852-025-01108-y>
- Duru, I., Sunar, A. S., White, S., & Diri, B. (2021). Deep Learning for Discussion-Based Cross-Domain Performance Prediction of MOOC Learners Grouped by Language on FutureLearn. *Arabian Journal for Science and Engineering*, 46(4), 3613-3629. <https://doi.org/10.1007/s13369-020-05117-x>
- Elaouzy, Y., & El Fadar, A. (2023). Role of carbon tax in the widespread integration of sustainable strategies in buildings. *Building and Environment*, 234, 110182. <https://doi.org/10.1016/j.buildenv.2023.110182>
- Elkelawy, M., M. El-Ashmawy, W., & Ahmed, S. M. (2024). State of the Art in Concentrated Solar Power: Latest Technological Advancements and Innovations in Efficiency and Energy Storage. *Pharos Engineering Science Journal*, 1(1), 17-28. <https://doi.org/10.21608/pesj.2025.344983.1007>
- Energy Sector Management Assistance Program (ESMAP), World Bank. (2019). *Global Solar Atlas 2.0, interactive solar resource map application*. <https://globalsolaratlas.info/map>
- Energy Sector Management Assistance Program (ESMAP), World Bank. (2025). *Saudi arabia – global solar atlas*. <https://globalsolaratlas.info/download/saudi-arabia>
- Enerji ve Tabii Kaynaklar Bakanlığı, Türkiye. (2023). *Güneş enerji potansiyel atlası (GEPA)*. <https://gepa.enerji.gov.tr/>
- European Commission. Joint Research Centre. (2020). *Photovoltaics: Technology development report 2020*. Publications Office. <https://data.europa.eu/doi/10.2760/827685>
- European Commission. Joint Research Centre. (2024). *Clean Energy Technology Observatory, Photovoltaics in the European Union: Status report on technology*

- development, trends, value chains and markets : 2024*. Publications Office. <https://data.europa.eu/doi/10.2760/1812909>
- Gal, Y., & Ghahramani, Z. (2015). *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning* (Versiyon 6). arXiv. <https://doi.org/10.48550/ARXIV.1506.02142>
- Gal, Y., & Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. *international conference on machine learning*, 1050-1059.
- Gamboa, J. C. B. (2017). *Deep Learning for Time-Series Analysis* (Versiyon 1). arXiv. <https://doi.org/10.48550/ARXIV.1701.01887>
- Gasparin, A., Lukovic, S., & Alippi, C. (2022). Deep learning for time series forecasting: The electric load case. *CAAI Transactions on Intelligence Technology*, 7(1), 1-25. <https://doi.org/10.1049/cit2.12060>
- Gates, D. M. (1980). Solar Radiation. İçinde D. E. Reichle (Ed), *Biophysical Ecology* (ss. 96-147). Springer New York. https://doi.org/10.1007/978-1-4612-6024-0_6
- Geetha, A., Santhakumar, J., Sundaram, K. M., Usha, S., Thentral, T. M. T., Boopathi, C. S., Ramya, R., & Sathyamurthy, R. (2022). Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms. *Energy Reports*, 8, 664-671. <https://doi.org/10.1016/j.egyr.2021.11.190>
- Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645-6649. <https://doi.org/10.1109/ICASSP.2013.6638947>
- Gürel, A. E., Ağbulut, Ü., Bakır, H., Ergün, A., & Yıldız, G. (2023). A state of art review on estimation of solar radiation with various models. *Heliyon*, 9(2).
- Hakhverdyan, A. (2024). *Tree-structured parzen estimator (TPE)*. <https://anna-ssi.github.io/blog/tpe/>
- Hammond, J. E., Orozco, R. A. L., Baldea, M., & Korgel, B. A. (2024). *Short-Term Solar Irradiance Forecasting Under Data Transmission Constraints*. <https://doi.org/10.48550/ARXIV.2403.12873>
- Hirshberg, J. (1973). Helium abundance of the Sun. *Reviews of Geophysics*, 11(1), 115-131. <https://doi.org/10.1029/RG011i001p00115>
- Hiyama, T., & Kitabayashi, K. (1997). Neural network based estimation of maximum power generation from PV module using environmental information. *IEEE Transactions on Energy Conversion*, 12(3), 241-247. <https://doi.org/10.1109/60.629709>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

- Hong, Y.-Y., Martinez, J. J. F., & Fajardo, A. C. (2020). Day-Ahead Solar Irradiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory. *IEEE Access*, 8, 18741-18753. <https://doi.org/10.1109/ACCESS.2020.2967900>
- Hossain, E. (2023). The Sun and the Earth. İçinde E. Hossain, *The Sun, Energy, and Climate Change* (ss. 1-67). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-22196-5_1
- Hossain, M. S., Wadi Al-Fatlawi, A., Kumar, L., Fang, Y. R., & Assad, M. E. H. (2024). Solar PV high-penetration scenario: An overview of the global PV power status and future growth. *Energy Systems*. <https://doi.org/10.1007/s12667-024-00692-6>
- Ibrahim, H. A. (2015). Determination of the temperature balancing in the atmosphere and surface of the earth by using physics techniques. *Journal of Physical Science and Innovation*, 7(2).
- Ifaei, P., Nazari-Heris, M., Tayerani Charmchi, A. S., Asadi, S., & Yoo, C. (2023). Sustainable energies and machine learning: An organized review of recent applications and challenges. *Energy*, 266, 126432. <https://doi.org/10.1016/j.energy.2022.126432>
- Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.-A. (2019). Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery*, 33(4), 917-963. <https://doi.org/10.1007/s10618-019-00619-1>
- Ji, Y., Song, J., & Shen, P. (2022). A review of studies and modelling of solar radiation on human thermal comfort in outdoor environment. *Building and Environment*, 214, 108891. <https://doi.org/10.1016/j.buildenv.2022.108891>
- Jin, Y., Shen, W., & Wu, C. Q. (2025). Power Prediction Based on Signal Decomposition and Differentiated Processing with Multi-Level Features. *Electronics*, 14(10), 2036. <https://doi.org/10.3390/electronics14102036>
- Kara, A. (2019). Uzun-Kısa Süreli Bellek Ağı Kullanarak Global Güneş Işınımı Zaman Serileri Tahmini. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 7(4), 882-892. <https://doi.org/10.29109/gujsc.571831>
- Karaman, Ö. A., Tanyıldızı Ağır, T., & Arsel, İ. (2021). Estimation of solar radiation using modern methods. *Alexandria Engineering Journal*, 60(2), 2447-2455. <https://doi.org/10.1016/j.aej.2020.12.048>
- Kayık, S., & Çelebi, S. B. (2025). Optimizing solar irradiance forecasting in batman, turkey using artificial neural networks with hyperparameter tuning. *Proceedings of the 3rd international mardin scientific research and innovation congress*, 377-390.
- Kilic, F. C. (2016). Evaluation of solar energy, its present status, potentials, policies and the latest developments in turkey. *International Journal of Ecosystems & Ecology Sciences*, 6(3).

- Kohavi, R. & others. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*, 14, Article 2.
- Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. *Future Internet*, 15(8), 255. <https://doi.org/10.3390/fi15080255>
- Korachagaon, I., & Bapat, V. N. (2012). General formula for the estimation of global solar radiation on earth's surface around the globe. *Renewable Energy*, 41, 394-400. <https://doi.org/10.1016/j.renene.2011.11.002>
- Korkmaz, M., Budakçı, M., & Kılınç, İ. (2024). Assessment of surface roughness in milling of wood with different material temperature and cutting parameters. *BioResources*, 19(4), 9343-9357. <https://doi.org/10.15376/biores.19.4.9343-9357>
- Krishnan, N., Kumar, K. R., & Inda, C. S. (2023). How solar radiation forecasting impacts the utilization of solar energy: A critical review. *Journal of Cleaner Production*, 388, 135860. <https://doi.org/10.1016/j.jclepro.2023.135860>
- Kumar, S., Verma, A. K., & Mirza, A. (2024). *Digital transformation, artificial intelligence and society*. Springer.
- Ladjal, B., Nadour, M., Bechouat, M., Hadroug, N., Sedraoui, M., Rabehi, A., Guermoui, M., & Agajie, T. F. (2025). Hybrid deep learning CNN-LSTM model for forecasting direct normal irradiance: A study on solar potential in Ghardaia, Algeria. *Scientific Reports*, 15(1), 15404. <https://doi.org/10.1038/s41598-025-94239-z>
- Lateef, K. F., Paul, J., & Jahan, Z. (2021). Weather Parameters and Heat Index Prediction Applying Deep Neural Networks. *2021 IEEE Region 10 Symposium (TENSYP)*, 1-6. <https://doi.org/10.1109/TENSYP52854.2021.9550852>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Lee, C.-Y., Chou, P.-C., Chiang, C.-M., & Lin, C.-F. (2009). Sun Tracking Systems: A Review. *Sensors*, 9(5), 3875-3890. <https://doi.org/10.3390/s90503875>
- Lewinson, E. (2025). *A comprehensive overview of regression evaluation metrics*. <https://developer.nvidia.com/blog/a-comprehensive-overview-of-regression-evaluation-metrics>
- LIU, B. H., & JORDAN, R. (1993). The interrelationship and characteristic distribution of direct, diffuse and total solar Radiation. *SPIE milestone series*, 54, 97-115.
- Li, Y., Wang, J., Dai, X., Wang, L., Yeh, C.-C. M., Zheng, Y., Zhang, W., & Ma, K.-L. (2023). How Does Attention Work in Vision Transformers? A Visual Analytics Attempt. *IEEE Transactions on Visualization and Computer Graphics*, 29(6), 2888-2900. <https://doi.org/10.1109/TVCG.2023.3261935>

- Lim, B., & Zohren, S. (2020). *Time Series Forecasting With Deep Learning: A Survey*. <https://doi.org/10.48550/ARXIV.2004.13408>
- Lim, S.-C., Huh, J.-H., Hong, S.-H., Park, C.-Y., & Kim, J.-C. (2022). Solar Power Forecasting Using CNN-LSTM Hybrid Model. *Energies*, 15(21), 8233. <https://doi.org/10.3390/en15218233>
- Liu, W., Liu, Y., Zhou, X., Xie, Y., Han, Y., Yoo, S., & Sengupta, M. (2021). Use of physics to improve solar forecast: Physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI. *Solar Energy*, 215, 252-265. <https://doi.org/10.1016/j.solener.2020.12.045>
- Malalha, S. A. K., Burhanuddin, M., & Yunos, N. B. M. (2024). Unveiling the tapestry of machine learning: A comparative analysis of support vector machines, random forests, and neural networks in diverse applications. *Tuijin Jishu/Journal of Propulsion Technology*, 45(3), 2024.
- Mariappan, Y., Ramasamy, K., & Velusamy, D. (2025). An optimized deep learning based hybrid model for prediction of daily average global solar irradiance using CNN SLSTM architecture. *Scientific Reports*, 15(1), 10761. <https://doi.org/10.1038/s41598-025-95118-3>
- Masson, G., Bosch, E., Van Rechem, A., & de l'Epine, M. (2024). *Snapshot 2024 – snapshot of global PV markets*. IEA Photovoltaic Power Systems Programme (IEA-PVPS). <https://doi.org/10.69766/VHRF4040>
- Matrenin, P. V., Gamaley, V. V., Khalyasmaa, A. I., & Stepanova, A. I. (2024). Solar Irradiance Forecasting with Natural Language Processing of Cloud Observations and Interpretation of Results with Modified Shapley Additive Explanations. *Algorithms*, 17(4), 150. <https://doi.org/10.3390/a17040150>
- Mellit, A., & Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy*, 84(5), 807-821. <https://doi.org/10.1016/j.solener.2010.02.006>
- Mishra, S., & Palanisamy, P. (2019). *An Integrated Multi-Time-Scale Modeling for Solar Irradiance Forecasting Using Deep Learning* (Versiyon 3). arXiv. <https://doi.org/10.48550/ARXIV.1905.02616>
- Mitrašinović, A. M. (2021). Photovoltaics advancements for transition from renewable to clean energy. *Energy*, 237, 121510. <https://doi.org/10.1016/j.energy.2021.121510>
- Moan, J. (2001). Visible light and UV radiation. *Radiation at Home, Outdoors and in the Workplace*, 2001, 69-85.
- Montanes, R., Windahl, J., Pålsson, J., & Thern, M. (2017). Dynamic modeling of parabolic trough solar thermal power plant with thermal storage using Modelica. *Heat Transfer Engineering*.

- Murarka, S. P., & Hymes, S. W. (1995). Copper metallization for ULSL and beyond. *Critical Reviews in Solid State and Materials Sciences*, 20(2), 87-124. <https://doi.org/10.1080/10408439508243732>
- NASA Earth Observatory. (2019). *Earth's energy budget*. <https://earthobservatory.nasa.gov/features/EnergyBalance/page4.php>
- National Academies of Sciences, Engineering, and Medicine. (2023). Appendixes. İçinde *The role of net metering in the evolving electricity system*. The National Academies Press. <https://doi.org/10.17226/26704>
- O'Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks* (Versiyon 2). arXiv. <https://doi.org/10.48550/ARXIV.1511.08458>
- Owens, C. B., Mathew, N., Olaveson, T. W., Tavenner, J. P., Kober, E. M., Tucker, G. J., Hart, G. L. W., & Homer, E. R. (2025). Feature engineering descriptors, transforms, and machine learning for grain boundaries and variable-sized atom clusters. *Npj Computational Materials*, 11(1), 21. <https://doi.org/10.1038/s41524-024-01509-x>
- Ozturk, O. (2024). The Impact of AI on International Trade: Opportunities and Challenges. *Economies*, 12(11), 298. <https://doi.org/10.3390/economies12110298>
- Özgür, M. D. (2021). *Ekonomik büyüme ve yenilenebilir enerji arasındaki ilişki: Türkiye örneği*.
- Papachristopoulou, K., Fountoulakis, I., Bais, A. F., Psiloglou, B. E., Papadimitriou, N., Raptis, I.-P., Kazantzidis, A., Kontoes, C., Hatzaki, M., & Kazadzis, S. (2024). Effects of clouds and aerosols on downwelling surface solar irradiance nowcasting and short-term forecasting. *Atmospheric Measurement Techniques*, 17(7), 1851-1877. <https://doi.org/10.5194/amt-17-1851-2024>
- Pearcy, R. W. (2000). Radiation and light measurements. İçinde R. W. Pearcy, J. R. Ehleringer, H. A. Mooney, & P. W. Rundel (Ed), *Plant Physiological Ecology* (ss. 97-116). Springer Netherlands. https://doi.org/10.1007/978-94-010-9013-1_6
- Polman, A., Knight, M., Garnett, E. C., Ehrler, B., & Sinke, W. C. (2016). Photovoltaic materials: Present efficiencies and future challenges. *Science*, 352(6283), aad4424. <https://doi.org/10.1126/science.aad4424>
- Pommeret, A., & Schubert, K. (2022). Optimal energy transition with variable and intermittent renewable electricity generation. *Journal of Economic Dynamics and Control*, 134, 104273. <https://doi.org/10.1016/j.jedc.2021.104273>
- Qiu, Z., & Li, P. (2019). Solar Energy Resource and Its Global Distribution. İçinde X. Zhao & X. Ma (Ed), *Advanced Energy Efficiency Technologies for Solar Heating, Cooling and Power Generation* (ss. 1-30). Springer International Publishing. https://doi.org/10.1007/978-3-030-17283-1_1

- Sandro, T. (2025). *Optimalizace domácího PV systému pro různé klimatické podmínky* [Master]. České vysoké učení technické v Praze. Vypočetní a informační centrum.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Shabaniverki, M. (2017). Design and analyze of 20 MW photovoltaic solar powerplant in Iran. *Journal of Solar Energy Research*, 2(3), 65-69.
- Sharma, N., & Puri, V. (2023). Solar Energy Fundamental Methodologies and its Economics: A Review. *IETE Journal of Research*, 69(1), 378-403. <https://doi.org/10.1080/03772063.2020.1822762>
- Shensa, M. J. (1992). The discrete wavelet transform: Wedding the a trous and Mallat algorithms. *IEEE Transactions on Signal Processing*, 40(10), 2464-2482. <https://doi.org/10.1109/78.157290>
- Solargis. (t.y). *Solargis: Online access to high-resolution solar and meteorological data, PV simulation tools, maps and services*. <https://solargis.info/>
- Solcast. (2025). *Global solar irradiance data and PV system power output data*. <https://solcast.com/>
- Soto, E. A., Hernandez-Guzman, A., Vizcarrondo-Ortega, A., McNealey, A., & Bosman, L. B. (2022). Solar Energy Implementation for Health-Care Facilities in Developing and Underdeveloped Countries: Overview, Opportunities, and Challenges. *Energies*, 15(22), 8602. <https://doi.org/10.3390/en15228602>
- Sunar, A. S., Leon-Urrutia, M., White, S., & Uhomoihi, J. (2021). Guest editorial: Distance learning, MOOCs and globalisation of engineering education. *Computer Applications in Engineering Education*, 29(4), 645-647. <https://doi.org/10.1002/cae.22392>
- Sutton, R. S., Barto, A. G., & others. (1998). *Reinforcement learning: An introduction* (No. 1; C. 1). MIT press Cambridge.
- Şahin, Z. R., Dinçer, F., & Yılmaz, A. S. (2022). 4 kişilik bir ailenin elektrik enerjisi ihtiyacı için şebeke bağlantılı güneş enerjisi santrali tasarımı ve simülasyonu. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 25(Özel Sayı), 46-56.
- Tandon, A., Awasthi, A., Pattnayak, K. C., Tandon, A., Choudhury, T., & Kotecha, K. (2025). Machine learning-driven solar irradiance prediction: Advancing renewable energy in Rajasthan. *Discover Applied Sciences*, 7(2), 107. <https://doi.org/10.1007/s42452-025-06490-8>
- Thekaekara, M. P. (1976). Solar radiation measurement: Techniques and instrumentation. *Solar Energy*, 18(4), 309-325. [https://doi.org/10.1016/0038-092X\(76\)90058-X](https://doi.org/10.1016/0038-092X(76)90058-X)

- Thipwangmek, N., Suetrong, N., Taparugssanagorn, A., Tangparitkul, S., & Promsuk, N. (2024). Enhancing Short-Term Solar Photovoltaic Power Forecasting Using a Hybrid Deep Learning Approach. *IEEE Access*, 12, 108928-108941. <https://doi.org/10.1109/ACCESS.2024.3440035>
- Tiismus, H., Maask, V., Astapov, V., Korõtko, T., & Rosin, A. (2025). State-of-the-Art Review of Emerging Trends in Renewable Energy Generation Technologies. *IEEE Access*, 13, 10820-10843. <https://doi.org/10.1109/ACCESS.2025.3528640>
- Travieso-González, C. M., Cabrera-Quintero, F., Piñán-Roescher, A., & Celada-Bernal, S. (2024). A Review and Evaluation of the State of Art in Image-Based Solar Energy Forecasting: The Methodology and Technology Used. *Applied Sciences*, 14(13), 5605. <https://doi.org/10.3390/app14135605>
- Tsai, M.-J., Lin, C.-H., Lai, J.-P., & Pai, P.-F. (2024). Using Deep Learning, Optuna, and Digital Images to Identify Necrotizing Fasciitis. *Electronics*, 13(22), 4421. <https://doi.org/10.3390/electronics13224421>
- Tsangas, M., Zorpas, A. A., & Jeguirim, M. (2022). Sustainable renewable energy policies and regulations, recent advances, and challenges. İçinde *Renewable Energy Production and Distribution* (ss. 449-465). Elsevier. <https://doi.org/10.1016/B978-0-323-91892-3.00009-1>
- Tsay, R. S. (2005). *Analysis of financial time series*. John Wiley & sons.
- Uğuz, S., Oral, O., & Çağlayan, N. (2019). Estimation of Energy to be Obtained from PV Power Plants Using Machine Learning Methods. *Uluslararası Muhendislik Arastirma ve Gelistirme Dergisi*, 769-779. <https://doi.org/10.29137/umagd.514933>
- Van Zyl, C., Ye, X., & Naidoo, R. (2024). Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP. *Applied Energy*, 353, 122079. <https://doi.org/10.1016/j.apenergy.2023.122079>
- Victron Energy BV. (2023). *BlueSolar PWM DUO-LCD&USB schematic diagram*. https://www.aceflex.de/wp-content/uploads/2023/03/BlueSolarDuo_schema.jpg
- Vignola, F., Michalsky, J., & Stoffel, T. (2019). *Solar and infrared radiation measurements*. CRC press.
- Wald, L. (2009). Solar radiation energy (fundamentals). *Encyclopedia of Life Support System (EOLSS)*, Eolss Publishers, Oxford, 44-99.
- Wan, C., Zhao, J., Song, Y., Xu, Z., Lin, J., & Hu, Z. (2015). Photovoltaic and solar power forecasting for smart grid energy management. *CSEE Journal of Power and Energy Systems*, 1(4), 38-46. <https://doi.org/10.17775/CSEEJPES.2015.00046>
- Wang, D., Xu, M., Guangming, Z., Luo, F., Gao, J., & Chen, Y. (2025). Enhancing wind power forecasting accuracy through LSTM with adaptive wind speed calibration

- (C-LSTM). *Scientific Reports*, 15(1), 5352. <https://doi.org/10.1038/s41598-025-89398-y>
- Wang, K., Qi, X., & Liu, H. (2019). Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy*, 189, 116225. <https://doi.org/10.1016/j.energy.2019.116225>
- Wang, S.-C. (2003). Artificial Neural Network. İçinde S.-C. Wang, *Interdisciplinary Computing in Java Programming* (ss. 81-100). Springer US. https://doi.org/10.1007/978-1-4615-0377-4_5
- Wang, Y., Li, Y., Zheng, Y., & Gao, and W. (2025). Solar Radiation Forecasting with Hybrid Deep Learning Framework Integrating Feature Factorization. *Human-centric Computing and Information Sciences*, 15(0), 1-19. <https://doi.org/10.22967/HCIS.2025.15.009>
- Wang, Z., Bu, S., Wen, J., & Huang, C. (2025). A comprehensive review on uncertainty modeling methods in modern power systems. *International Journal of Electrical Power & Energy Systems*, 166, 110534. <https://doi.org/10.1016/j.ijepes.2025.110534>
- Widiputra, H., Mailangkay, A., & Gautama, E. (2021). Multivariate CNN-LSTM Model for Multiple Parallel Financial Time-Series Prediction. *Complexity*, 2021(1), 9903518. <https://doi.org/10.1155/2021/9903518>
- Wikipedia. (2025). *Solar constant*. https://en.wikipedia.org/wiki/Solar_constant
- Winster Praveenraj, D. D., A, M., Pastariya, R., Sharma, D., Abootharmahmoodshakir, K., & Dhablia, A. (2024). Machine Learning Integration for Enhanced Solar Power Generation Forecasting. *E3S Web of Conferences*, 540, 04007. <https://doi.org/10.1051/e3sconf/202454004007>
- Yaman, I., Kilic, I., Yaman, O., Poyraz, F., Erdem Kaya, E., Ozgur Baris, V., & Ciris, S. (2025). A state-of-the-art new method for diagnosing atrial septal defects with origami technique augmented dataset and a column-based statistical feature extractor. *Computers in Biology and Medicine*, 196, 110967. <https://doi.org/10.1016/j.compbiomed.2025.110967>
- Yaman, O. (2021). An automated faults classification method based on binary pattern and neighborhood component analysis using induction motor. *Measurement*, 168, 108323. <https://doi.org/10.1016/j.measurement.2020.108323>
- Yilmaz, A. (2025). Bilateral Cross Hashing Image Retrieval Based on Principal Component Analysis. *Arabian Journal for Science and Engineering*, 50(15), 12495-12512. <https://doi.org/10.1007/s13369-025-10135-8>
- Yilmaz, A., & Poli, R. (2022). Successfully and efficiently training deep multi-layer perceptrons with logistic activation function simply requires initializing the weights with an appropriate negative mean. *Neural Networks*, 153, 87-103. <https://doi.org/10.1016/j.neunet.2022.05.030>

- Yunus Khan, T. M., Soudagar, M. Elahi. M., Kanchan, M., Afzal, A., Banapurmath, Nagaraj. R., Akram, N., Mane, S. D., & Shahapurkar, K. (2020). Optimum location and influence of tilt angle on performance of solar PV panels. *Journal of Thermal Analysis and Calorimetry*, 141(1), 511-532. <https://doi.org/10.1007/s10973-019-09089-5>
- Zaidi, B. (2018). Introductory chapter: Introduction to photovoltaic. *Solar Panels and Photovoltaic Materials*, 1.
- Zenk, H. (2018). *Low cost provides of the energy needs of plateau houses by using photovoltaic systems*.
- Zhang, L., Wang, J., Qian, Y., & Li, Z. (2025). Photovoltaic power uncertainty quantification system based on comprehensive model screening and multi-stage optimization tasks. *Applied Energy*, 381, 125061. <https://doi.org/10.1016/j.apenergy.2024.125061>
- Zhou, N., Shang, B., Xu, M., Peng, L., & Feng, G. (2024). Enhancing photovoltaic power prediction using a CNN-LSTM-attention hybrid model with Bayesian hyperparameter optimization. *Global Energy Interconnection*, 7(5), 667-681. <https://doi.org/10.1016/j.gloi.2024.10.005>
- Zhu, L., Huang, X., Zhang, Z., Li, C., & Tai, Y. (2025). A novel U-LSTM-AFT model for hourly solar irradiance forecasting. *Renewable Energy*, 238, 121955. <https://doi.org/10.1016/j.renene.2024.121955>
- Zhu, T., Guo, Y., Li, Z., & Wang, C. (2021). Solar Radiation Prediction Based on Convolution Neural Network and Long Short-Term Memory. *Energies*, 14(24), 8498. <https://doi.org/10.3390/en14248498>

TEŞEKKÜR

Veri Erişimi: Yazarlar, bu çalışmada kullanılan küresel güneş ışıınımı ve güneş enerji sistemi güç çıktısı verilerini sağlayan Solcast'a teşekkür eder. Veri seti, makul talepler doğrultusunda ve kamuya açık bir kurumunda akademik çalışmacı veya öğrenci olarak kayıt yaptırılması şartıyla, ticari olmayan çalışmalarda kullanım için Solcast tarafından ücretsiz sunulmaktadır. Uygun araştırmacılar, Solcast API Toolkit aracılığıyla ücretsiz erişim hakkı elde etmektedir. Daha fazla bilgi ve veri erişimi için lütfen <https://solcast.com/> adresini ziyaret ediniz.

Hesaplama Altyapısı: Bu araştırmada yer alan tüm/kısmi nümerik hesaplamalar TÜBİTAK ULAKBİM, Yüksek Başarım ve Grid Hesaplama Merkezi'nde (TRUBA kaynaklarında) gerçekleştirilmiştir.

Kod Erişimi: Tez kapsamında önerilen mimarinin, veri kümesi sağlayıcısının güvenlik gerekçeleriyle API kısmı hariç tam uygulaması aşağıdaki adreste kamuya açıktır.

<https://github.com/sbariscelebi/solar-irradiance-prediction>

