



T.C.
EGE ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü



FOTOVOLTAİK TEKNOLOJİLERİN ÜRETİM OPTİMİZASYONUNDA MAKİNE ÖĞRENMESİNİN KULLANILMASI

Doktora Tezi

Burak KAHRAMAN

Güneş Enerjisi Anabilim Dalı

İzmir
2024

EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

(DOKTORA TEZİ)

FOTOVOLTAİK TEKNOLOJİLERİN ÜRETİM
OPTİMİZASYONUNDA MAKİNE ÖĞRENMESİNİN
KULLANILMASI

Burak KAHRAMAN

Tez Danışmanı: Doç. Dr. Ahmet YILANCI

Güneş Enerjisi Anabilim Dalı

İzmir

2024



EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ETİK KURALLARA UYGUNLUK BEYANI

EÜ Lisansüstü Eğitim ve Öğretim Yönetmeliğinin ilgili hükümleri uyarınca Yüksek Lisans Tezi / Doktora Tezi olarak sunduğum “**Fotovoltaik Teknolojilerin Üretim Optimizasyonunda Makine Öğrenmesinin Kullanılması**” başlıklı bu tezin kendi çalışmam olduğunu, sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını, bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversitede başka bir tez çalışması içinde sunmadığımı, bu tezin planlanmasından yazımına kadar bütün safhalarda bilimsel etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

06/09/2024

İmzası

Burak KAHRAMAN



ÖZET**FOTOVOLTAİK TEKNOLOJİLERİN ÜRETİM
OPTİMİZASYONUNDA MAKİNE ÖĞRENMESİNİN
KULLANILMASI**

KAHRAMAN, Burak

Doktora Tezi, Güneş Enerjisi Anabilim Dalı

Tez Danışmanı: Doç. Dr. Ahmet YILANCI

Eylül 2024

Tez kapsamında fotovoltaik teknolojilerin üretim optimizasyonunda makine öğrenmesinin kullanılması gerçekleştirilmiştir. Makine öğrenmesi (ML), bilgisayarların bilinen verilerden belirli bir görevi nasıl gerçekleştireceklerini öğrenmeleri amacıyla son dönemde oldukça farklı alanda kullanılan bir disiplindir. Bu farklı alanlardan biri olan fotovoltaik teknolojilerin tasarımının ve fabrikasyon süreçlerinin optimizasyonu konuları da çalışılan konular arasına hızla dâhil olmuştur.

Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” adlı projenin hedeflerinden biri, yeni nesil güneş hücrelerinin endüstriyel boyutlara taşınması aşamasında yöntem geliştirmek ve optimizasyonunu sağlamaktır. Bu tezde, Güneş Enerjisi Enstitüsü bünyesinde yapılan çalışmalardan elde edilen veriler kullanılarak oluşturulmuş olan veri seti üzerinde, farklı makine öğrenmesi algoritmaları kullanılarak yeni nesil fotovoltaik teknolojilerden olan perovskit güneş hücresi yapısının optimizasyonu gerçekleştirilmiştir. Altı farklı makine öğrenmesi algoritmasının kullanıldığı model eğitim süreci sonucunda en yüksek başarımlı değerini XGBoost modeli göstermiştir. Geliştirilmiş olan model, kök ortalama kare hatası: 0.6737 ve Pearson korelasyon katsayısı: 0.8952 performans sonuçlarına sahiptir. XGBoost modeli üzerinde, slot-die (yarık kaplama) ve air knife (hava bıçağı) cihazlarındaki belirlenmiş olan farklı parametre değerleri

iin g dnřm verimi (PCE) deęeri tahmin edilmiř ve karřılařtırma yapılabilmesi adına elde edilmiř olan parametreler kullanılarak perovskit gneř hcrei retimi gerekleřtirilmiřtir. Kullanılan parametreler sonucunda retimi gerekleřtirilmiř olan perovskit gneř hcresinin performans lmlerinde, dz lm sonucunda %8,52 ve ters lm sonucundaysa %8,71 PCE deęeri elde edilmiřtir. Elde edilen sonular, deneysel alıřmalar srecinde retilen en yksek PCE deęerine sahip perovskit gneř hcrei iin sırasıyla %11,37'lik ve %4,94'lk bir artıřın saęlandıęını gstermektedir.

Anahtar szckler: Makine ęrenmesi, Perovskit Gneř Hcreleri, Optimizasyon



ABSTRACT**UTILIZING MACHINE LEARNING IN PRODUCTION
OPTIMIZATION OF PHOTOVOLTAIC TECHNOLOGIES**

KAHRAMAN, Burak

PhD in Solar Energy Department

Supervisor: Assoc. Prof. Dr. Ahmet YILANCI

September 2024

In this thesis, the use of machine learning in the production optimization of photovoltaic technologies has been realized. Machine learning (ML) is a discipline that has recently been used in many different fields in order for computers to learn how to perform a specific task from known data. Optimization of the design and fabrication processes of photovoltaic technologies, which is one of these different fields, has rapidly become one of the topics studied.

One of the objectives of the project titled "Development of the Manufacturing and Characterization Infrastructure of New Generation Photovoltaic Technologies", which is being carried out within the Solar Energy Institute of Ege University, is to develop and optimize the methodology for the transfer of new generation solar cells to industrial dimensions. Optimization of the perovskite solar cell structure, which is one of the new generation photovoltaic technologies, was performed using different machine learning algorithms on the data set created using data obtained from the studies carried out within the Solar Energy Institute. As a result of the model training process using six different machine learning algorithms, the XGBoost model showed the highest performance value. The developed model has the performance results of RMSE: 0.6737 and Pearson correlation coefficient: 0.8952. Using the XGBoost model, the power conversion efficiency (PCE) value was predicted for different parameter values determined in slot-die and air-knife devices, and perovskite solar cell (PSC) production was carried out using the obtained parameters for comparison. In the performance measurements of the PSC device produced using the parameters, a PCE value of 8.52% was obtained in the

forward measurement and 8.71% in the reverse measurement. The obtained results show an increase of 11.37% and 4.94%, respectively, for the PSC with the highest PCE value produced during the experimental studies.

Keywords: Machine Learning, Perovskite Solar Cells, Optimization



ÖNSÖZ

Yeni nesil fotovoltaik teknolojilerden olan perovskit güneş hücreleri; yüksek absorpsiyon katsayıları, ayarlanabilir bant aralığına sahip olmaları, farklı yüzeylere uygulanabilmeleri, düşük maliyetleri ve kolay üretilebilmeleri gibi avantajlarından ötürü son yıllarda fotovoltaik alanında devrim niteliğinde bir teknoloji olarak öne çıkmaktadır. Yapılan çalışmalarda büyük gelişmeler kaydedilmiş olsa da hala çözülmesi gereken kararlılık, ölçeklenebilirlik ve güvenilirlik problemleri, en iyi malzeme kombinasyonları ve üretim süreçleri, temel standartlar ve metrikler üzerindeki çalışmalar devam etmektedir. Ayrıca ticari uygulanabilirliğe ulaşmak için perovskit güneş hücresi üretim süreçlerinin optimize edilmesi gerekmektedir. Geleneksel yöntemlerle üretim sürecinin optimizasyonu, zaman alıcı ve maliyetli olabilmektedir. Bu nedenle, üretim süreçlerinde yenilikçi yaklaşımlar büyük önem taşımaktadır.

Makine öğrenmesi, büyük veri setlerini analiz ederek karmaşık sistemlerin davranışlarını modelleyebilme ve öngörülerde bulunabilme kapasitesiyle, çeşitli endüstriyel süreçlerin optimizasyonunda etkili bir araç olarak kullanılmaktadır. Bu tez çalışmasında, perovskit güneş hücresi üretim prosesinin optimizasyonunda makine öğrenmesi kullanılmıştır.

Çalışmanın amacı, perovskit güneş hücrelerinin üretim verimliliğini artırmak ve üretim maliyetlerini düşürmektir. Bu bağlamda, farklı makine öğrenmesi algoritmaları ve optimizasyon yöntemleri kullanılarak en uygun üretim koşulları belirlenmiş ve deneysel verilerle doğrulanmıştır. Elde edilen sonuçlar, perovskit güneş hücrelerinin daha verimli ve ekonomik bir şekilde üretilmesine katkı sağlamayı hedeflemektedir. Ayrıca yapılan tez çalışması, büyük alan perovskit güneş hücrelerinin üretim optimizasyonunda makine öğrenmesinin potansiyelini ortaya koymaktadır.



İÇİNDEKİLER

Sayfa

ÖZET	vii
ABSTRACT	ix
ÖNSÖZ	xi
İÇİNDEKİLER	xiii
ŞEKİLLER DİZİNİ	xv
TABLolar DİZİNİ	xvii
SİMGELER VE KISALTMALAR	xix
1. GİRİŞ	1
2. LİTERATÜR ÖZETİ	3
2.1. Literatür Özeti	3
2.2. Tezin Önemi	24
2.3. Tezin Amacı	25
2.4. Tez Kapsamı	26
3. MAKİNE ÖĞRENMESİ	27
3.1. Yapay Zekânın Tarihçesi	27
3.2. Makine Öğrenmesi	30
3.2.1. Denetimli Öğrenme	32
3.2.2. Denetimsiz Öğrenme	34
3.2.3. Pekiştirmeli Öğrenme	35
4. FOTOVOLTAİK TEKNOLOJİLER	36
4.1. Güneş Hücresi Teknolojileri	36
4.1.1. Perovskit Güneş Hücreleri	36
5. GEREÇ VE YÖNTEM	39
6. YAPILAN DİĞER ÇALIŞMALAR	51
6.1. Bant Boşluğu Değerinin Tahmin Edilmesi	51
6.2. Perovskite Veri Tabanı Projesi	54
6.3. Empedans Spektroskopisi Sonuçlarının Tahmin Edilmesi	55
7. BULGULAR VE TARTIŞMA	58
8. SONUÇ VE ÖNERİLER	73
KAYNAKLAR DİZİNİ	77
TEŞEKKÜR	85
ÖZGEÇMİŞ	86



ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
3.1 Yapay zekâ, makine öğrenmesi ve derin öğrenme	27
3.2 Klasik programların şematik olarak gösterimi.....	31
3.3 Makine öğrenmesinin şematik olarak gösterimi.....	31
3.4 Makine öğrenmesi kategorileri	32
4.1 Araştırma ölçeğindeki Perovskit güneş hücrelerinin güç dönüşüm verimleri	37
4.2 Yaygın olarak kullanılan perovskit güneş hücresi yapıları.....	37
5.1 Kullanılan makine öğrenmesi akış şeması.....	41
6.1 Veri setinin ilk 5 satırı	51
6.2 Parametrelerin ortalama, standart sapma, minimum, Q1, medyan, Q3 ve maksimum değerleri.....	52
6.3 Korelasyon matrisi.....	52
6.4 PCE ve BandGap parametreleri arasındaki ilişkinin görselleştirilmesi.....	53
6.5 XGBoost modelinin hata ve tahmin edilen değer – gerçek değer uyum grafikleri.....	55
6.6 Farklı ön gerilim değerleri için deneysel EIS grafiği	57
6.7 Farklı ön gerilim değerleri için tahmin edilen EIS grafiği	57
7.1 Slot-die ve air-knife cihazları	58
7.2 Slot-die cihazının şematik olarak gösterimi	59
7.3 PSC aygıt mimarisi	59
7.4 Karakterizasyon çalışması sonuçları a) Absorpsiyon grafiği, b) Emisyon grafiği, c) XRD grafiği ve d) AFM görüntüsü	60
7.5 Slot-die cihazıyla kaplanmış film ve tamamlanmış aygıt görüntüsü.....	61
7.6 En yüksek verim değerine sahip olan aygıtın performans ölçümleriyle I-V grafikleri	62

7.7 Oluşturulmuş olan veri setinin bir bölümünün görüntüsü	62
7.8 Belirlenmiş olan parametreler	63
7.9 ML modellerini eğitmek için seçilmiş olan parametreler	64
7.10 PCE_rev parametresinde yer alan değerlerin dağılım grafiği	64
7.11 Korelasyon matrisi	65
7.12 LR modelinin tahmin edilen değer – gerçek değer uyum grafikleri	66
7.13 KNN modelinin tahmin edilen değer – gerçek değer uyum grafikleri	67
7.14 SVR modelinin tahmin edilen değer – gerçek değer uyum grafikleri	67
7.15 RF modelinin tahmin edilen değer – gerçek değer uyum grafikleri	68
7.16 XGBoost modelinin tahmin edilen değer – gerçek değer uyum grafikleri ...	69
7.17 CatBoost modelinin tahmin edilen değer – gerçek değer uyum grafikleri ...	70
7.18 Tahmin sonuçlarına göre üretimi gerçekleştirilen PSC	71
7.19 Tahmin sonuçlarına göre üretilmiş olan aygıtın performans ölçümleriyle I-V grafikleri.....	72

TABLOLAR DİZİNİ

<u>Tablo</u>	<u>Sayfa</u>
5.1 ML modellerinin eğitimi için belirlenmiş olan parametreler	42
6.1 Beş farklı ML modelinin performans ölçüm sonuçlarının karşılaştırılması...	53
6.2 ML modellerinin performans ölçüm sonuçlarının karşılaştırılması	55
6.3 Z' ve -Z" değerli için elde edilen RMSE, ortalama ve hata yüzdesi sonuçları.....	56
7.1 LR modelinin performans ölçüm sonuçları	66
7.2 KNN modelinin performans ölçüm sonuçları.....	66
7.3 SVR modelinin performans ölçüm sonuçları	67
7.4 RF modelinin performans ölçüm sonuçları	68
7.5 XGBoost modelinin performans ölçüm sonuçları	69
7.6 CatBoost modelinin performans ölçüm sonuçları	70
7.7 ML modeli kullanılarak elde edilen en yüksek verim değerinin tahmin sonuçları	71



SİMGELER VE KISALTMALAR

<u>Simgeler</u>	<u>Açıklama</u>
V_{oc}	Açık Devre Gerilimi
bar	Basınç
rpm	Dakikadaki Devir Sayısı
FF	Dolum Faktörü
eV	Elektrovolt
V	Gerilim
I_{sc}	Kısa Devre Akımı
J_{sc}	Kısa Devre Akım Yoğunluğu
μl	Mikrolitre
μm	Mikrometre
μs	Mikrosaniye
mm	Milimetre
ms	Milisaniye
nm	Nanometre
$\Delta \tau$	Operasyonel Kararlılık Test Süresi

OpF	Optimizasyon Faktörü
R	Pearson Korelasyon Katsayısı
E_g	Yasak Enerji Aralığı
InF	Yenilik Faktörü



<u>Kısaltmalar</u>	<u>Açıklama</u>
AFM	Atomik Kuvvet Mikroskobu
Ag	Gümüş
ANN	Yapay Sinir Ağları
AVT	Ortalama Görünür Bölge Geçirgenliği
Bi	Bizmut
BO	Bayesian Optimizasyonu
Br	Brom
CatBoost	Category Gradient Boosting
Cl	Klor
Cs	Sezyum
Cu	Bakır
DC	Diskriminant Sınıflandırıcı
DFT	Yoğunluk Fonksiyonel Teorisi
DL	Derin Öğrenme
DT	Karar Ağaç
EIS	Elektrokimyasal Empedans Spektroskopisi
EL	Çoklu Öğrenme
ETL	Elektron Taşıyıcı Katman

EQE	Dış Kuantum Verimliliği
FA	Formamidyum
FTO	Flor Katkılı Kalay Oksit
FV	Fotovoltaik
GBDT	Gradyan Destekli Karar Ağaçları
GBR	Gradyan Destekli Regresyon
GBRT	Gradyan Destekli Regresyon Ağaçları
GPR	Gaussian Proses Regresyon
HTL	Boşluk Taşıyıcı Katman
HOMO	En Yüksek Dolu Moleküler Orbital
I	İyodür
K	Potasyum
KNN	K-En Yakın Komşu
KRR	Kernel Ridge Regresyonu
LightGBM	Light Gradyan Destekli Makine
LR	Doğrusal Regresyon
LUE	Işık Kullanım Verimliliği
LUMO	En Düşük Boş Moleküler Orbital
MA	Metilamonyum

MAE	Ortalama Mutlak Hata
ML	Makine Öğrenmesi (Machine Learning)
MLP	Çok Katmanlı Algılayıcı
MPPT	Maksimum Güç Noktası İzleme
MSE	Ortalama Kare Hatası
Na	Sodyum
NBC	Naive Bayes Sınıflandırma
OPV	Organik Fotovoltaik
Pb	Kurşun
PCE	Güç Dönüşüm Verimi
PSC	Perovskit Güneş Hücresi
Rb	Rubidyum
RF	Rastgele Orman
RMSE	Kök Ortalama Kare Hatası
RSPP	Hızlı Sprey Plazma İşleme
Sb	Antimon
SHAP	SHapley Additive exPlanations
Sn	Kalay
SOM	Kendi Kendini Organize Eden Harita

SVR	Destek Vektör Regresyon
TDOE	Taguchi Deney Tasarımı
XGBoost	eXtreme Gradient Boosting
XRD	X-Işını Kırılımı





1. GİRİŞ

Makine öğrenmesi, genellikle büyük veri setleri üzerinde çalışarak ve veri üzerinde algoritmalar ve matematiksel modeller kullanılarak veriden bilgi çıkarma ve model oluşturma süreçlerine dayanmaktadır. Özellikle son dönemde bilgisayar donanımlarının gelişmesi, kullanılan algoritmalarındaki gelişmeler ve kullanılabilir veri sayısındaki artış nedeniyle, birçok alanda kullanılmaya başlanmıştır. ML, büyük veri setlerini analiz ederek karmaşık sistemlerin davranışlarını modelleyebilme ve öngörülerde bulunabilme kapasitesiyle, çeşitli üretim süreçlerinin optimizasyonunda da etkili bir araç olarak kullanılmaktadır.

Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan, Kalkınma Bakanlığı destekli 2016K12-2841 numaralı “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” adlı projenin hedeflerinden biri olan ve kurulumu tamamlanan Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarında, yeni nesil fotovoltaik teknolojilerden olan perovskit güneş hücresi üretim çalışmaları devam etmektedir.

Perovskit güneş hücreleri, yüksek verimlilikleri ve düşük üretim maliyetleri nedeniyle son yıllarda büyük ilgi görmektedir. Ancak, bu teknolojilerin laboratuvar ölçeğinden ticari üretim ölçeğine taşınması, yani hücre ölçeğinden modül ölçeğine boyut artırımı, önemli bir zorluk teşkil etmektedir. Bu süreçte, ölçek büyütme işlemine uygun üretim proseslerinin geliştirilmesi ve optimizasyonu kritik bir rol oynamaktadır. Üretim süreçlerinin optimizasyonu hem maliyetlerin düşürülmesi hem de ürün kalitesinin artırılması açısından büyük önem taşımaktadır. ML uygulamaları, üretim süreçlerinin daha hızlı ve etkili bir şekilde optimize edilmesine olanak tanımaktadır. Bu sayede, perovskit güneş hücrelerinin ticari üretimi için gerekli olan süreçler daha kısa sürede geliştirilebilecek ve bu teknolojilerin yaygınlaşması sağlanabilecektir. Ayrıca, bu çalışmalar, Türkiye'nin yenilenebilir enerji alanındaki rekabet gücünü artıracak ve sürdürülebilir enerji çözümlerine katkıda bulunacaktır.

Literatürde yapılan çalışmalar, yeni nesil fotovoltaik teknolojiler için en uygun tasarım ve üretim yapısının oluşturulması aşamasında makine öğrenmesi

uygulamalarının yaygın olarak kullanıldığını göstermektedir. ML, büyük veri setlerini analiz ederek üretim süreçlerinin daha verimli hale getirilmesine olanak tanır. Özellikle, üretim parametrelerinin optimizasyonu, hata tespiti ve kalite kontrol süreçlerinde ML algoritmaları etkili bir şekilde kullanılmaktadır. Bu algoritmalar, üretim sürecinde toplanan verileri analiz ederek, en uygun üretim koşullarını belirleyebilmekte ve böylece üretim verimliliğini artırabilmektedir. Bununla birlikte, literatürdeki çalışmaların büyük bir kısmının halen hücre ölçeğinde olduğu görülmektedir. Bu durum, modül ölçeğinde yapılacak çalışmaların önemini artırmaktadır. Modül ölçeğinde yapılacak araştırmalar, perovskit güneş hücrelerinin ticari uygulamalar için uygun hale getirilmesi açısından kritik öneme sahiptir. Bu bağlamda, Ege Üniversitesi Güneş Enerjisi Enstitüsü'nde yürütülen çalışmalarla hem ulusal hem de uluslararası düzeyde önemli katkılar sağlanması hedeflenmektedir.

Yapılan tez çalışmasında, Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan araştırmaların sonuçlarının dâhil edilmesiyle oluşturulan veriler kullanılarak, en yüksek verimin elde edileceği büyük ölçek üretim yapısının optimizasyonunda ML kullanılması amaçlanmaktadır. Kullanılacak olan ML algoritmaları olarak Doğrusal Regresyon, K-En Yakın Komşu, Destek Vektör Makinesi, Rastgele Orman, XGBoost ve CatBoost seçilmiştir. ML modellerinin değerlendirilmesi için kullanılacak metrikler; kök ortalama kare hatası ve Pearson korelasyon katsayısı olarak belirlenmiştir. Bu optimizasyon sürecinde, geliştirilen modellerin en yüksek başarımlarına ulaşması hedeflenmektedir. Sonuç olarak tez çalışmasında elde edilen çıktıların, Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı'nda yapılacak deneysel çalışmalarda elde edilecek güneş hücreleri aracılığıyla doğrulanması gerçekleştirilecektir.

2. LİTERATÜR ÖZETİ

2.1. Literatür Özeti

Makine öğrenmesi, bilgisayarların bilinen verilerden belirli bir görevi nasıl gerçekleştireceklerini öğrenmeleri amacıyla kullanılan bir disiplindir (Yin et al., 2020). Bilgisayar donanımlarının gelişmesi ve kullanılabilir veri sayısındaki artış nedeniyle ML, özellikle son dönemde, birçok alanda kullanılmaya başlanmıştır (Sharifzadeh et al., 2019). Bu alanlardan biri olan fotovoltaik teknolojilerin üretim aşamalarında da yapılan çalışmalar giderek artmaktadır. Bu bölümde çalışma alanıyla ilgili olarak literatürde yapılan araştırmalar incelenmekte ve ana hatlarıyla paylaşılmaktadır.

Perovskit güneş hücreleri (PSC'ler), ilk üretilmeye başladıkları 2009 yılında %3,8 verim ile çalışmaktayken, günümüzde verimleri %26,7 seviyesine ulaşmış durumdadır (NREL, 2024). PSC'ler kolay üretilebilmeleri, farklı yüzeylere uygulanabilmeleri, yüksek absorpsiyon katsayıları gibi avantajlarından ötürü son yıllarda yeni nesil fotovoltaik teknolojileri alanında büyük potansiyel göstermektedir. ABX_3 kimyasal formülüyle gösterilen perovskit yapısında A ve B katyonları, X ise anyonu temsil etmektedir. Perovskit güneş hücresi; metal kontak, seçici yük taşıma tabakaları olan ETL/HTL, perovskit, flor katkılı kalay oksit tabaka (FTO) ve cam gibi farklı katmanların bir araya getirilmesiyle oluşturulmaktadır (Fu et al., 2016).

Perovskit katmanının farklı kimyasallar kullanılarak oluşturulması yanında, her bir katman da farklı malzemeler ve teknikler kullanılarak elde edilebilmektedir. Bu katmanlarda yapılan her bir değişiklik güç dönüşüm verimi veya kararlılık gibi parametrelere olumlu/olumsuz etki etmektedir. PSC'ler alanında yapılan çalışmalarda büyük gelişmeler kaydedilmiş olsa da hala çözülmesi gereken kararlılık, ölçeklenebilirlik ve güvenilirlik problemleri, en iyi malzeme kombinasyonları ve üretim süreçleri, temel standartlar ve metrikler üzerindeki çalışmalar devam etmektedir (Park and Zhu, 2020).

Yeni nesil fotovoltaik teknolojilerin geliştirilmesi aşamasında kullanılmak üzere, enerji hasadı uygulamaları için %87'si 1.2eV-2.4eV bant boşluğu

değerlerine sahip olan 75 farklı perovskit yapısı, ML kullanılarak 2 aylık bir dönem içinde üretilip karakterize edilmektedir (Sun et al., 2019). Deneysel XRD verilerine dayalı bileşikler 0D, 2D ve 3D yapılarına sınıflandırmak için yapay sinir ağları kullanılıp insan analizinden 10 kat daha hızlı bir şekilde ve %90 başarı oranına ulaşılmaktadır. Bu çalışmada ABX_3 , $A_3B_2X_9$ ve ABX_4 şeklinde ($A = MA, FA, Cs, Rb, K$ veya Na | $B = Pb, Sn, Ag, Cu, Na, Bi$ veya Sb | $X = Cl, Br$ veya I) 17 bileşik, 58 alaşım toplam 75 halojenür perovskit sentezi ve karakterizasyonu gerçekleştirilip tamamı kurşunsuz olan 4 bileşik ilk kez ince film formunda rapor edilmektedir. Doğrusal olmayan bant boşluğu davranışı sergileyen yapının keşfi çok eklemli güneş hücreleri için kurşunsuz inorganik perovskit elde etme noktasında yeni bir yol açmaktadır.

Fotovoltaik sistemler için kararlı ve toksik olmayan hibrit organik-inorganik perovskit yapısının ML kullanılarak elde edilmesi amacıyla yapılan bir diğer çalışmada, 132 adet kararlı ve toksik olmayan perovskit yapısına ulaşılmaktadır (Wu and Wang, 2019). Çalışmanın ilk adımında veri kümesinde yer alan 230.808 adet yapı, 77.748'e düşürülmekte, sonrasında ise belirlenen parametreler doğrultusunda (tolerans ve oktahedral faktörü) 38.086 adet veriye farklı ML yöntemleri uygulanmaktadır. Araştırmada Gradyan Destekli Regresyon (GBR), Destek Vektör Regresyon (SVR) ve Kernel Ridge Regresyonu (KRR) algoritmaları kullanılmaktadır. Geliştirilmiş olan modellerin performanslarını ölçmek için R^2 , Ortalama Mutlak Hata (MAE) ve Ortalama Kare Hatası (MSE) metrikleri kullanılıp en iyi sonucun SVR algoritması kullanılarak geliştirilen modele ait olduğu belirtilmektedir.

Perovskit güneş hücresi alanında ML kullanılarak yapılan uygulamalara odaklanılmış olan çalışmada; hücre mimarileri, kullanılan materyaller ve hücre üretim prosedürleri karşılaştırılmaktadır. Verimliliği iyileştirmeye odaklanan çalışmalar, son zamanlarda bu teknolojinin ticarileştirilmesi için fazlasıyla önem arz eden kararlılık konusunda yoğunlaşmaktadır (Yılmaz ve Yıldırım, 2020). Ayrıca malzemelerin histerezis, tekrarlanabilirliği ve toksisitesinin yanında sıcaklık, nem ve oksijen gibi çevresel stres faktörleri de diğer çalışma alanlarını oluşturmaktadır. Daha fazla kararlı ve daha az toksik alternatiflerin keşfi noktasında ML ile ilişkili çalışmalar özellikle son iki yılda ciddi bir artış göstermektedir.

Çalışmanın referans olarak aldığı bir diğer çalışmaya göre fotovoltaik güneş hücrelerinin fabrikasyonu aşamalarında kullanılan ML uygulamaları dört adımda toplanmaktadır (Li et al., 2019). Bu adımları aşağıdaki şekilde sıralamak mümkündür.

- Malzeme özelliklerinin tahmini,
- Cihaz yapısının optimizasyonu,
- Üretim süreçlerinin optimizasyonu ve
- Elde edilen verilerin yeniden yapılandırılması

Yeni nesil güneş hücrelerinin fabrikasyonu sürecinde kullanılacak ML uygulamalarında da aynı yolun izlenebileceği belirtilmektedir.

Güneş hücrelerinde kullanılacak olan kurşunsuz perovskit yapısının ML kullanılarak tanımlanmasını hedefleyen çalışmada 540 bileşik için, bir bileşiğin kimyasal bileşenlerinin temel fazlarına kıyasla kararlılığını gösteren ΔH_f ve optimum değeri 1.1eV ile 1.8eV arasında tanımlanan E_g değerlerinden oluşan bir veri seti oluşturulmaktadır (Im et al., 2019). Mevcut veriler ve Gradyan Destekli Regresyon Ağaçları (GBRT) algoritması kullanılarak elde edilen model, farklı yapıların ΔH_f ve E_g değerlerini tahmin edebilmek için eğitilmektedir. Modelin amacı, her düğümün giriş özelliklerinden birini temel alan bir bölünmeyi ve her yaprağın hedef değişkenin tahminini temsil ettiği bir ağaç oluşturmaktır. Hazırlanmış olan veri seti 32 farklı parametreden oluşmaktadır. Geliştirilen modelde Kök Ortalama Kare Hatası (RMSE) metriği kullanılmaktadır ve test verileri kullanılarak yapılan tahminlerde ΔH_f değeri 0.021eV/atom ve E_g değeri 0.223eV yakınlıkta tahmin edilmektedir.

Yüksek performanslı PSC'lerin geliştirilmesi sırasında ML kullanılarak öngörü ve stratejilerin belirlenmesi amacıyla 2000 makalenin incelendiği çalışmada malzeme bileşimini optimize etmek, tasarım stratejileri geliştirmek ve PSC'lerin performansını tahmin etmek için ML kullanımı amaçlanmaktadır (Li et al., 2019). İncelenmiş olan makalelerden elde edilen verilerle 333 farklı veriden oluşan bir veri seti oluşturulmaktadır. Bu veriler; A yapısını (MA, FA, Cs), B yapısını (Pb ve Sn)

ve C yapısını oluşturan (Br, Cl, I) malzemelerin oranlarıyla bant boşluğu, ΔH , ΔL , PCE (güç dönüşüm verimliliği), V_{oc} (açık devre gerilimi), J_{sc} (kısa devre akımı) ve FF (dolum faktörü) parametrelerinden oluşmaktadır. Doğrusal Regresyon (LR), K-En Yakın Komşu (KNN), SVR, Rastgele Orman (RF) ve Yapay Sinir Ağları (ANN) algoritmaları kullanılarak beş farklı model geliştirilmektedir. Bu modellerde, tahmin edilen değerler ile etiketlenmiş gerçek değerler arasındaki ilişkiyi değerlendirmek için RMSE ve Pearson korelasyon katsayısı (r) metrikleri kullanılmaktadır. Model tahminlerinde PSC'lerdeki en yüksek güç dönüşüm veriminin 1.2eV bant boşluğu civarında elde edilebileceği tahmin edilmektedir ve bu sonuca güneş hücrelerindeki teorik öngörülerle uyumlu olacak şekilde ulaşılmaktadır. Ayrıca daha yüksek güç dönüşüm verimi için daha düşük değerlere sahip ΔH ve ΔL ve onlarla ilişkili olacak şekilde daha düşük bant boşluk değerine sahip yapıların kullanılması gerektiği ortaya konmaktadır.

PSC'ler için bant boşluğu değerinin doğru bir şekilde hesaplanması büyük öneme sahiptir. Yoğunluk Fonksiyonel Teorisi (Density Functional Theory) malzeme bant boşluk değerlerinin hesaplanması için kullanılabilse de bu yöntem zengin malzeme hesaplama bilgisi gerektirmektedir ve sonuçların gerçek deneysel sonuçlarla tutarlı olup olmadığı konusunda genellikle bazı soru işaretleri mevcuttur. Yapılan çalışmada perovskit bant boşluklarını tahmin etmek için ML kullanılmaktadır (Liu et al., 2021). Yakın dönemdeki 1254 makaleden, 227 deneysel bant boşluğu verisini toplanarak, LR, KNN, SVR, RF, Çok Katmanlı Algılayıcı (MLP) ve XGBoost algoritmaları kullanılarak 6 farklı ML modeli oluşturulmaktadır. Modellerin performans ölçümlerinde RMSE ve Pearson korelasyon katsayısı metrikleri kullanılmaktadır. Sonuç olarak 0,55'e kadar düşen RMSE değeri ve 0,99'a varan Pearson korelasyon katsayısı değeri elde edilmektedir. Buna ek olarak, perovskit yapısını (ABX_3) oluşturan her bir kimyasal bileşimin bant boşluğu üzerindeki etkisi, SHAP (SHapley Additive exPlanations) değeri kullanılarak verilmektedir. SHAP değeri; öznelik parametrelerinin önemini ve nihai sonuç üzerindeki pozitif ve negatif korelasyon etkilerini yansıtabilmektedir. Çalışmada SHAP aracılığıyla, B katmanındaki metal ve X katmanındaki halojen iyonu içeriklerinin ABX_3 tipi perovskit bant boşluğu üzerinde önemli bir etkiye sahip olduğu bulunmaktadır. Bu sonuçlarla, PSC'ler için bant

boşluk değerlerini hızlı ve doğru bir şekilde tahmin etmekte ML kullanımının büyük potansiyele sahip olduğu vurgulanmaktadır.

Yapay zekanın bir dalı olan ML, son zamanlarda çeşitli özelliklerin tahmini ve rasyonel malzeme tasarımı için optoelektronik cihazlarda güçlü bir teknoloji olarak kullanılmaktadır. Metal halojenür perovskitler, üstün foto fiziksel özellikleri ve güneş hücresi uygulamalarındaki hızlı gelişimi nedeniyle odak noktasında yer almakta ve son dönemde üretim sürecini optimize etmek ve işleme maliyetini azaltmak için ML tercih edilmektedir. Yapılan çalışmada perovskit güneş hücrelerinin tasarımında farklı ML uygulamaları kapsamlı bir şekilde incelenmektedir (Parikh et al., 2021). PSC'lerin; tasarımında, termodinamik kararlılık, bant boşluğu, yapısal deformasyon ve yapıyı tahmin etmede, malzeme bileşiminin optimizasyonunda, rekombinasyon ve degradasyonun anlaşılması alanlarında yapılan ML uygulamalarına yer verilmektedir. Ayrıca gelecekteki olası araştırma alanları ile ML zorlukları incelenmektedir. Sonuç olarak ML, malzeme bileşimini optimize etmek ve gelecekte perovskit teknolojisi alanındaki tasarım stratejilerini tahmin etmek için vazgeçilmez bir yol haritası haline geleceği yorumu yapılmaktadır.

Yapılan bir diğer çalışma geliştirmekte olan fotovoltaik teknolojilerin cihaz performansları alanında inceleme niteliği taşımaktadır (Almora et al., 2021). Güç dönüşüm verimi, dış kuantum verimliliği (EQE), deney şartlarındaki PCE değerinin ideal PCE değerine oranı, ortalama görünür bölge geçirgenliği (AVT), ışık kullanım verimliliği (LUE), operasyonel kararlılık test süresi ($\Delta\tau$), operasyonel kararlılık testi için etkili toplam degradasyon oranı ($DR_{\Delta\tau}$), yenilik faktörü (Inf) ve optimizasyon faktörü (OpF) parametreleri belirlenmektedir. En yüksek verim veren güneş hücreleri, esnek yüzey fotovoltaik güneş hücreleri, transparan ve yarı transparan güneş hücreleri ve geliştirmekte olan güneş hücrelerinin operasyonel kararlılığı şeklinde dört gruba ayrılmış olan geliştirmekte olan fotovoltaik teknolojiler, yukarıdaki parametrelere göre belirlenmekte ve çalışma sonucunda yayınlanmaktadır.

PSC'lerde küçük alanlı laboratuvar ölçekli cihazların güç dönüşüm verimi son on yılda hızlı bir gelişme kaydetmiştir. Bununla birlikte, başarılı bir

ticarileştirme için düşük maliyetli, ölçeklenebilir ve yüksek verimli üretim tekniklerinin geliştirilmesi gerekmektedir. Perovskit güneş hücreleri için ölçeklenebilir bir üretim tekniği geliştirmek, teknolojinin ticarileşme yolundaki kritik araştırma alanlarından biridir. Deneysel çalışmalarda spin kaplama tekniği kullanılarak %25 verimli perovskit güneş hücreleri üretilmesine rağmen, bu işleme yöntemi bir üretim hattına ölçeklenebilir değildir. Yapılan çalışmada, perovskit güneş hücreleri süreç optimizasyonu için ML rehberliğinde bir sıralı öğrenme çerçevesi sunulmaktadır (Liu et al., 2022a). Geliştirilen yöntem, açık hava perovskit cihaz üretimi için hızlı sprey plazma işleme (RSPP) tekniğine uygulanmaktadır. Çalışmada hedef değişken olarak PCE belirlenmektedir. Genel bir çerçeve için, ilk olarak modelsiz bir örnekleme yöntemiyle süreç koşullarının deneysel planlaması yapılmaktadır. Ardından, PSC'ler RSPP yöntemiyle üretilmekte ve PCE, standart test koşulları altında bir güneş simülatörüyle ölçülmektedir. Süreç parametreleri; sıcaklık (°C), hız (cm/s), sprey akışı (mL/min), plazma yüksekliği (cm), plazma gaz akışı (L/min) ve plazma görev döngüsü (%) ve cihaz PCE'lerinin deneysel verileriyle, süreç-verimlilik ilişkisini öğrenmek için ML kullanılarak regresyon modeli eğitilmekte ve regresyon modeli daha sonra örneklenmemiş bölgeler için PCE'yi tahmin etmekte kullanılmaktadır. Son olarak, tahmin sonuçları kısıtlama bilgileriyle birlikte değerlendirilip yeni bir deneysel çalışma gerçekleştirilmektedir. Sınırlı bir deneysel bütçe ile 100 işlem koşulunu tarayarak, RSPP ile üretilen bir cihazdan elde edilen en iyi sonuçta %18,5'lik bir verim artışı elde edilmektedir. RSPP tekniği ile PSC'lerin optimize edilmesi, diğer geleneksel araştırma odaklı deney tasarımı yöntemlerine kıyasla daha hızlı bir optimizasyon sağlamaktadır. Çalışma RSPP için yapılmış olmasına rağmen, ML perovskit güneş hücrelerinin üretim teknolojilerinin geliştirilmesi için geniş çapta kullanılabilmektedir.

Perovskit güneş hücreleri alanında yapılan araştırmalar 2013 yılından bu yana artış göstermektedir ve son dönemde kurşunsuz perovskit malzemelerin keşfedilmesi çalışmaları da büyük önem kazanmaktadır. Yapılan çalışmada, kurşunsuz halojenür çift perovskit malzemelerin termodinamik kararlılığını ve bant boşluğunu yüksek hızda ve yüksek hassasiyette tespit etmek, seçilen özelliklerin önemini analiz etmek ve potansiyel kurşunsuz perovskit malzemeleri keşfetmek için ML kullanılmaktadır (Guo and Lin, 2021). Çalışma için 540 veri ve 32 parametreden oluşan veri tabanı üzerinde RF, Ridge Regresyon, SVR ve XGBoost

dahil olmak üzere dört farklı ML algoritması kullanılmaktadır. En yüksek tahmin performansını, termodinamik kararlılık için XGBoost (R^2 : 0.9935 ve MAE: 0.0126), bant boşluğu için RF (R^2 : 0.9410 ve MAE: 0.1492) sağlamaktadır. Ayrıca termodinamik kararlılık ve elektronegatifliğin en önemli özelliği (eleneg_x) arasında doğrusal ilişkinin, bant boşluğuyla; en yüksek işgal enerji seviyesi (hoe_b1) ve kübik faz arasındaysa doğrusal olmayan ilişkinin olduğu sonucuna ulaşılmaktadır. Bu çalışma, perovskit güneş hücresi malzemelerinin keşfini hızlandırmak için makine öğrenmesinin büyük bir potansiyeli olduğunu göstermektedir.

ML, son dönemde PSC'lerin araştırılmasını ve geliştirilmesini hızlandırmak için sıklıkla kullanılmaktadır. Yapılan çalışmada, düşük bant boşluğu değerlerine sahip beş keşfedilmemiş $(\text{FAPbI}_3)_x(\text{MAPbBr}_{2.8}\text{Cl}_{0.2})_{1-x}$ perovskit yapısını tahmin etmek için XGBoost ve RF algoritmalarıyla ML kullanılmaktadır (Yan et al., 2022). Kısa devre akım yoğunluğu (J_{sc}) ve açık devre gerilimi (V_{oc}) de beş farklı PSC yapısı için öngörülmektedir. Deneysel olarak, en yüksek güç dönüşüm verimi %22,5 ile düzlemsel $(\text{FAPbI}_3)_{0,95}(\text{MAPbBr}_{2.8}\text{Cl}_{0,2})_{0,05}$ PSC için elde edilmektedir. Bu yapıda J_{sc} 24,6mA/cm², V_{oc} 1,11V ve dolum faktörü %82,4 olarak ölçülmektedir. Ölçülen ve tahmin edilen bant boşlukları arasında %2'den daha az bağıl hata ile bir uyum olduğu sonucuna ulaşılmaktadır. Ölçülen J_{sc} ve V_{oc} değerleri ML tahminleri ile tutarlılık gösterirken, J_{sc} değeri ayrıca optik modelleme ve simülasyondan bağımsız olarak doğrulanmaktadır. Buna ek olarak düzlemsel hücrenin üst kısmında, yansıma önleyici PDMS (polydimethylsiloxane) nanokoni diziler kullanılarak akım yoğunluğunun 1,2mA/cm² artırılabilceği ve verimlilik değerinin %23,6'ya yükseltilebileceği bulunmaktadır. Kararlılık ölçümleri de mevcut $(\text{FAPbI}_3)_{0,95}(\text{MAPbBr}_{2.8}\text{Cl}_{0,2})_{0,05}$ PSC cihaz kararlılığının, kapsülleme olmadan 600 saat boyunca $(\text{FAPbI}_3)_{0,95}(\text{MAPbBr}_3)_{0,05}$ yapısından önemli ölçüde daha iyi olduğu sonucunu göstermektedir.

Çalışmada, 2D/3D PSC'ler üzerine 146 yayından 599 veri noktası içeren bir veri kümesi, ML kullanılarak analiz edilmektedir (Yılmaz et al., 2022). ML modelleri, XGBoost, RF ve ANN kullanılarak PCE değerini tahmin etmek için geliştirilmektedir. Ayrıca PCE değerinin tahminlerinde kullanılmak üzere, veri setindeki eksik değerleri bulmak için bant boşluğu tahmin modelleri de

geliştirilmektedir. Hem bant boşluğu hem de PCE tahminleri için geliştirilen modellerde oldukça başarılı sonuçlara ulaşılmaktadır. Bant boşluğu tahminleri için inorganik katmanın kalınlığı (n), anyon yarıçapı (R_x) ve 2D katyon (R_m) en önemli parametreler olarak bulunmaktadır. Geliştirilen XGBoost modelinin, bant boşluğunu tahmin etmek için 0,14 RMSE değeriyle en başarılı olduğu görülmektedir. PCE tahminlerinde n-i-p yapıdaki PSC'ler için n , R_m ve bant boşluğu, p-i-n yapıdakiler içinse n , bant boşluğu ve boşluk taşıyıcı katmanın (HTL) iletim bant enerjisi en önemli parametreler olarak bulunmaktadır. PCE tahminlerinde, hem n-i-p hem de p-i-n yapısındaki PSC'ler için sırasıyla 3,04 ve 3,71 RMSE değerleriyle en iyi tahminler, RF modelleri kullanılarak elde edilmektedir. Ayrıca çalışmada, katmanlı perovskit yapılara sahip hücrelerin daha kararlı olduğu ve en kararlı hücrelere yol açan 2D ve 3D katyonların sırasıyla butylammonium ve formamidinium-Cs karışık katyon olduğu bulunmaktadır.

Perovskit güneş hücresi alanında nitelikli verinin önemine vurgu yapılan makalede, PSC'ler üzerine hakemli makalelerden tüm anlamlı cihaz verilerinin toplandığı ve bu verilerden oluşturulan veri tabanının kullanıma sunulduğu kapsamlı bir araştırma yapılmaktadır (Jacobsson et al., 2021). Yaklaşık 15.000 makalenin incelendiği veri tabanında, her biri için 100'e yakın parametrenin olduğu 42.460 cihaz listelenmektedir. Parametreler; çalışmaların yapıldığı makalelerin verileri, hücre yapısının özellikleri (kullanılan mimari, boyut, esneklik, şeffaflık, modül özellikleri vb.), cihaz yapısı (altlık, ETL, perovskit, HTL, metal kontak vb.), kullanılan kimyasalların elde edilme yöntemleri ve performans parametreleri şeklinde gruplanmaktadır. Çalışmada; bulunabilir, erişilebilir, birlikte çalışabilir ve yeniden kullanılabilir verilerin, yapay zekâ ve ML kullanımına olanak tanıdığı ve böylelikle PSC alanındaki gelişmelerin hızlanabileceği paylaşılmaktadır.

Organik fotovoltaiklerin (OPV'ler) verimliliği, malzeme tasarımı ve cihaz mühendisliği alanındaki çalışmaların sonucunda yakın zamanda %17'yi aşmıştır ve OPV'lerin köklü silikon güneş hücrelerine rakip olabileceği düşünülmektedir. Bir OPV'nin PCE değeri genellikle Scharber modeliyle, en yüksek dolu moleküler orbital (HOMO) ve en düşük boş moleküler orbital (LUMO) enerjileri kullanılarak 0,65'lik sabit bir dolum faktörü (FF) ile hesaplanmaktadır. Diğer birçok moleküler özellik cihaz parametrelerini etkilediğinden, Scharber modelinin performansının

oldukça zayıf olduğu ortaya çıkmaktadır. Moleküler özellikler ve cihaz parametreleri arasındaki ilişkilerin anlaşılması, organik güneş hücresinin genel performansını iyileştirmenin yanında, farklı uygulama alanları için kullanılacak cihaz gereksinimlerinin belirlenmesi için de oldukça önemlidir. Yapılan çalışmada, önemli olan 13 moleküler özellik kullanılarak, üç önemli cihaz parametresinin (V_{oc} , J_{sc} ve FF) tahmin edilmesi için bir dizi ML modeli oluşturulmaktadır (Sahu and Ma, 2019). Modeller, 300 veri üzerinde GBRT ve RT makine öğrenmesi algoritmaları kullanılarak geliştirilip elde edilen modellerin başarımlar değerleri Pearson korelasyon katsayısı, RMSE ve MAPE metrikleri kullanılarak ölçülmektedir. Sonuç olarak bu modellerin, belirli bir fotovoltaiik uygulama için uygun organik malzemelerin tasarlanmasında hayati bir rol oynayabileceği yorumu yapılmaktadır.

Fulleren olmayan üçlü organik güneş hücrelerinde uygun enerji seviyesi hizalaması, yük üretimi/taşınmadaki eşzamanlı iyileşme ve gerilim kaybındaki azalma nedeniyle PCE değerini artırabilmektedir. Bu çalışmada PCE'yi tahmin etmek ve yüksek performanslı malzeme kombinasyonlarını oluşturabilmek için ML kullanılmaktadır (Hao et al., 2021). Veri kümesi; üçlü organik güneş hücrelerinin her bir katmanındaki HOMO ve LUMO seviyeleri, V_{oc} , J_{sc} , dolum faktörü ve PCE parametre değerlerinin yer aldığı 157 veriden oluşmaktadır. RF, XGBoost, KNN, Karar Ağaç (DT), SVM, Ridge Sınıflandırma ve MLP algoritmaları kullanılarak 7 farklı ML modeli oluşturulmaktadır. Enerji seviyesi parametrelerine dayalı olarak oluşturulan regresyon ve sınıflandırma modellerinden RF algoritmasının en iyi tahmin yeteneğini gösterdiği sonucuna ulaşılmaktadır. Ayrıca, deneysel ve öngörülen sonuçları karşılaştırmak için iki doğrulama deneyi tasarlanmıştır. Sonuç olarak, fulleren olmayan akseptörlerin LUMO seviyesinin açık devre gerilimini azaltabileceği, ancak kısa devre akım yoğunluğunu önemli ölçüde artırabileceği ve açık devre geriliminin fulleren olmayan üçlü organik güneş hücrelerinde üçüncü bileşenin LUMO seviyesiyle optimize edilebileceği ortaya konmaktadır.

Organik fotovoltaiikler ve kimyasal yapıları arasındaki temel ilişkilerin hızlı bir şekilde incelenmesi, karmaşık bağlantıları nedeniyle zorlu bir süreçtir. Çalışmada ML yaklaşımını Taguchi Deney Tasarımı (TDOE) ile birleştirerek, OPV'lerin yapı-özellik eşleştirmeleri ve cihaz performanslarının tahmini için yeni bir metodoloji tasarlanmaktadır (Abadi et al., 2021). Geliştirilen modeller, küçük

moleküllü OPV sistemlerinden oluşan 240 veri ve OPV'lerin seçilen on mikroskopik özelliğinden oluşan veri tabanı kullanılarak oluşturulmaktadır. ML modelinin oldukça dikkat çekici performansı (Pearson katsayısı: 0,79), OPV'lerin gizli fiziksel ilkelerini ortaya çıkarma yeteneğini göstermektedir. TDOE modeli, OPV'lerin tasarım sürecinde HOMO ve LUMO seviyelerinin, gelecek vaat eden OPV malzemelerinin geliştirilmesinde oldukça önemli roller oynadığını göstermektedir. Elde edilen sonuçlar, ML'nin deney tasarımı ile birleştirilmesinin, tasarım sürecini etkili ve verimli bir şekilde yönlendirmek için oldukça önemli olduğunu ortaya koymaktadır.

Yapılan bir diğer çalışmada, organik fotovoltaiik (OPV) güneş hücrelerinin kararlılığını artırmak için ML kullanılmaktadır (David et al., 2020). Cihaz özellikleri, performans ve kararlılık parametrelerinin yer aldığı 1850 veriden oluşan veri tabanı üzerinde, güneş hücresi kararlılığı ve PCE arasındaki en etkili faktörleri belirlemek için sıralı minimal optimizasyon regresyon (SMOreg) modeli kullanılmaktadır. Çalışma, cihaz mimarisindeki her bir ince film için kararlılık ve PCE'de iyileştirmelere yol açabilecek malzemelerin tanımlanmasına ve OPV'lerin bozulmasında farklı stres faktörlerinin rolünün vurgulanmasına olanak sağlamaktadır. ISOS-L protokolleri altında yapılan testlerde ışık spektrumu ve aktif katman malzemesi seçiminin kararlılığı önemli ölçüde etkilediği, ISOS-D protokolleri altında yapılan testlerde ise birincil özelliklerin malzeme ve kapsüllemeye bağlı olduğu bulunmaktadır. Sonuç olarak, en iyi kararlılık ve performansa sahip malzemenin belirlenmesinde ML kullanımının hızlı ve verimli bir yöntem olduğu sonucuna ulaşılmaktadır. Gelecekteki OPV teknolojilerini geliştirmek için ML modelleriyle çok değerli bilgiler elde edilebileceği, böylece zaman alan deney ve optimizasyon ihtiyacının önemli ölçüde azaltılabileceği paylaşılmaktadır.

Halojenür perovskitler son yıllarda akademik olarak çokça çalışılan bir alan olsa da endüstriyel üretim örnekleri oldukça sınırlıdır. Yapılan perspektif çalışmasında, halojenür perovskitlerin ticarileştirilmesi noktasında karşılaşılan zorluklar gözden geçirilip ML araçlarının bu süreçte nasıl yardımcı olabileceği araştırılmaktadır (Kumar et al., 2022). Çalışmada sektör ihtiyaçları:

- 1) İstikrarlı bir temel proses sürdürülmesi (verimlilik ve çevresel istikrar ile ölçüldüğü üzere) ve prosesin en son gelişmeleri içerecek şekilde sık sık güncellenmesi
- 2) Küçük alanlı cihazlara yakın performans gösteren geniş alanlı cihazlara ve/veya büyük ölçekli üretime ulaşılması
- 3) Düşük performansın temel nedenlerini giderme (ideal olarak hat içinde ve üretim prosesinin başlarında) ve proses veya cihaz mimarisi değişikliklerinden kaynaklanan performans iyileştirmelerinin tahmin edilmesi
- 4) Kurum içi veri bilimi yeteneklerinin geliştirilmesi olarak sıralanmaktadır.

Uygulamalı ML üzerine devam eden akademik araştırmalar, potansiyel olarak dönüştürücü araçlar sağlayarak perovskitlerin ticarileştirilmesine anlamlı bir şekilde yardımcı olabilmektedir. Bu "oyun değiştirici" teknikleri; gerçek dünya problemlerini çözebilen ters tasarım araçları, ML modellerinden fiziksel öngörüler elde etme, proses transferini tesisler, ekipmanlar ve operatörler arasında kolaylaştırma ve literatürün araştırma ve geliştirme hattına otomatik olarak entegre edilmesi şeklinde sıralamak mümkündür.

Çalışma sonucunda;

- 1) Kurumsal bilgi ve insan uzmanlığını harmanlayan aktif öğrenme algoritmalarının, temel üretim süreçlerini istikrarlı hale getirmeye ve hızla güncellemeye yardımcı olabileceği
- 2) ML tabanlı sınıflandırma araçlarıyla bilgisayar görüntüleme yöntemlerinin, büyük ve küçük alanlı cihazlar arasındaki performans farkını daraltmaya yardımcı olabileceği
- 3) Çıkarım yöntemlerinin, birden fazla veri akışını ve simülasyonu uzlaştırarak ve araştırma çabalarını en yüksek olasılıklı alanlara odaklayarak, kök neden analizini hızlandırmaya yardımcı olabileceği yorumları yapılmaktadır.

Metal halojenür PSC'lerin performansı, üretim süreçleri ve perovskit yapısındaki bileşenler dahil olmak üzere büyük ölçüde deneysel parametrelere bağlıdır. Bu faktörleri optimize etmek için çok sayıda çalışma yapılmaktadır (Liu et al., 2022b; Li et al., 2021). Bununla birlikte PSC cihaz performanslarını sadece

üretim parametrelerini kullanarak tahmin etmek hala fazlasıyla zordur. Yapılan çalışmada bu alana odaklanılıp hakemli yayınlardan elde edilen 1072 cihazın parametrelerini içeren bir veri tabanı üzerinde ML kullanılarak farklı modeller geliştirilmektedir (Lu et al., 2022). Veri tabanı, PSC üretim süreçlerinden elde edilen 21 farklı parametreden oluşmaktadır. Parametreler; ABX₃ yapısında kullanılan malzeme oranları, ETL elektron hareketliliği, katkı maddesi kullanımı, cihaz performans değerleri, anti-çözgen özellikleri gibi gruplardan oluşmaktadır. LR, ANN, RF, XGBoost ve Gradient Boosting Karar Ağacı (GBDT) algoritmalarından oluşan 5 farklı ML modeli geliştirilmektedir. Modellerin hiper parametre optimizasyonunda 10-fold çapraz doğrulama kullanılmakta ve başarımların ölçümü için RMSE ve Pearson korelasyon katsayısı metrikleri belirlenmektedir. Optimize edilen ML modelleri içinden XGBoost, üretim parametreleri kullanılarak PCE değerini 1,28 RMSE ve 0,768 Pearson korelasyon katsayısı ile yüksek bir şekilde doğru tahmin etmektedir. Ayrıca, cihaz performansını etkileyen faktörler SHAP değeri kullanılarak sıralanmaktadır. SHAP aracılığıyla, A katmanındaki katyonun yüksek verimli PSCs elde etmek için çok önemli olduğu sonucuna ulaşılmıştır. ML modeli ile tahmin edilen sonuçları doğrulamak için deneyler ve DFT hesaplamaları kullanılmaktadır. Çalışma, deneylerden önce üretim parametreleri kullanılarak cihaz performansını tahmin etmede ML'nin uygulanabilirliğini ortaya koymaktadır.

Yapılan bir diğer çalışma, ML yöntemleri kullanarak perovskit güneş hücrelerinin kararlılığını analiz etmeyi amaçlamaktadır (Mammeri et al., 2023). Farklı malzemeler, biriktirme yöntemleri ve depolama koşullarına sahip 1050 PSC'den oluşan bir veri kümesi üzerinde farklı ML algoritmaları kullanılmaktadır. Kullanılan veri kümesi, PSC kararlılıkları ile ilgili yayınlanmış makalelerden elde edilen deneysel verilerin taranmasıyla oluşturulmaktadır. ETL ve HTL katman bileşenleri ve kalınlıkları, arka kontak malzemeleri ve çevresel depolama koşulları gibi parametreler kullanılıp bu parametrelerin cihaz kararlılığı üzerindeki etkisinin belirlenmesi amaçlanmaktadır. Kategorik değişkenler, girdiler arasındaki sıralı ilişkiye duyarlıdır. Bu nedenle Ekstra Ağaçlar (ET) sınıflandırıcısı kullanılmaktadır. Bu yöntem, birçok karar ağacından gelen tahmin ve sınıflandırmayı birleştiren bir ML algoritmaları topluluğudur. Çalışma sonunda yüksek kararlılık için hem normal ve hem de ters çevrilmiş güneş hücreleri için iki

farklı perovskit yapısı belirlenmektedir. Perovskit normal hücreler için; ETL olarak $\text{TiO}_2/\text{m-TiO}_2$, aktif katman olarak (2D-3D) perovskit, HTL ve HTL ikinci katman olarak P3HT ve LiTFSi + TBP ve arka kontak malzeme olarak Karbon, öncül çözügen olarak DMF + DMSO ve anti-çözügen olarak Klorobenzen kullanılan yapıların cihaz kararlılığını artırdığı bulunmaktadır. Perovskit ters çevrilmiş hücreler için ise, ETL ve ETL ikinci katmanı olarak BCP ve PCBM, perovskit aktif katman olarak $\text{MAPbI}_{3-x}\text{Cl}_x$, HTL NiO_x ve DEA, arka kontak malzeme olarak Al, öncül çözügen olarak DMF ve anti-çözügen olarak Klorobenzen kullanılan yapıların cihaz kararlılığını artırdığı bulunmaktadır. Bulgular PSC katmanlarında hidrofobik malzemelerin kullanılmasının cihaz kararlılığını artırdığını göstermektedir. Sonuçta ayrıca, perovskit katmanlar olarak çoklu katyon ve 2D/3D kristal yapı kullanılmasının uzun vadeli kararlılık sağladığı paylaşılmaktadır.

FV güç sistemlerinde (şebekeye bağlı, şebekeden bağımsız ve hibrit), sistemin verimliliği, güvenilirliği ve güvenliği üzerinde olumsuz etkisi olan farklı türde hata ve arızalarla karşılaşmak mümkündür. FV dize üzerinde etkisini gösteren önemli hatalar, kalıcı ve geçici hatalar olarak iki grupta sınıflandırılabilir. Kalıcı hatalar arasında kablo bağlantıları ve kabarcıklar, nem, delaminasyon, salyangoz izleri, sararma, görünmez hücre çatlakları ve yanmış hücreler, çizikler ve yaşlanan paneller gibi FV panel arızaları yer almaktadır. Bu tür arızalar, arızalı FV paneller değiştirilerek çözülebilmektedir. Toz birikimi, kar ve yapay gölgelenme gibi geçici hatalar, etkilenen FV paneli değiştirmeden operatörler tarafından giderilebilmektedir. Bu arızalar, FV panel verimliliğinin bazen %80'e kadar düşmesine neden olabilmektedir (Mellit & Kalogirou, 2022a).

FV endüstrisinin hızla gelişmesiyle birlikte, arıza koruma, tespit ve teşhis, sistemlerin güvenli ve istikrarlı çalışmasını sürdürmek çok önemli bir konu haline gelmektedir. FV sistemlerin arıza ve hata tespitinde ML ve Derin Öğrenme (DL) son zamanlarda yaygın olarak kullanılmaktadır. Genel olarak, arıza tespit ve teşhis yöntemleri üç adımdan oluşmaktadır:

- 1) Arıza tespit adımı: Doğru bir modele dayalı olarak ölçülen ve hesaplanan gücün karşılaştırılmasıyla basit bir şekilde gerçekleştirilebilmektedir. Arızaları tespit etmek için ML ve DL kullanmak mümkündür.

- 2) Arıza sınıflandırma ve tanılama adımı: Arızanın nedenlerini sınıflandırmayı ve tanılamayı amaçlar. Bu adım uygun algoritmalar geliştirilerek gerçekleştirilebilmektedir. Algoritmanın karmaşıklığı esas olarak tanılacak arızaların niteliğine bağlıdır. Bazı durumlarda, arızalar aynı belirtileri gösterirken nedenleri farklıdır, bu durumda ML veya DL algoritmaları bu sorunu ele almak için en uygun yöntemlerdir.
- 3) Hata lokalizasyonu ve izolasyonu adımı: Özel ekipman ve uzman bilgisi gerektirdiğinden, özellikle büyük ölçekli FV sistemler için en önemli ve zorlu adımdır. Örneğin sıcak noktalar, salyangoz çatlağı vs. gibi kusurları tespit etmek ve tanılamak için DL ile FV panellerin/dizilerin termal veya elektrolüminesans görüntülerinin kullanılması gerekmektedir.

Yapılan çalışmada FV dizelerinin arıza teşhisi için farklı ML ve çoklu öğrenme (EL) yöntemleri değerlendirilmektedir (Mellit & Kalogirou, 2022b). Çoklu öğrenme, birden fazla model oluşturan ve daha sonra bunları iyileştirilmiş sonuçlar üretmek için birleştiren tekniklerdir. Odak noktası, FV dizeleri etkileyebilecek bazı karmaşık hataların, çoklu hataların ve daha önce değerlendirilmemiş benzer I-V eğrilerine sahip hataların tespiti ve sınıflandırılmasıdır. Çalışmada veri tabanındaki 575 örneğin I-V eğrisi için normalleştirilmiş parametreler kullanılmaktadır. Verilerin elde edildiği sistemde 5 farklı arıza durumu oluşturulmakta ve bu durumların tespitinde LR, Naive Bayes Sınıflandırma (NBC), KNN, SVM, DT, RF ve ANN gibi farklı ML algoritmaları kullanılmaktadır. Sonuç olarak, RF ve DT modelleriyle arızaların %99 gibi yüksek doğrulukla tespit edilebildiği, arızaların sınıflandırılmasında ise ileri EL yöntemlerinden olan stacking (yığın) kullanılarak %81,73 gibi kabul edilebilir bir doğruluk sonucuna ulaşılmaktadır.

Fotovoltaik güç sistemlerinde sıcak noktalar, bir hücrenin veya hücre grubunun önemli ölçüde ısındığı ve FV modüllerin performansında kayıplara hatta uzun vadede daha büyük problemlere yol açabilecek bir güvenilirlik sorunudur. Genellikle kısmi gölgeleme nedeniyle hücrelerin sıcaklıklarındaki eşit olmayan artışlar sıcak noktalara neden olmaktadır. Bunun yanında FV modüllerde meydana gelen çatlakların (“salyangoz izi” veya mikro çatlaklar) varlığı ile sıcak noktaların arasında bir korelasyon olduğu yapılan çalışmalarla doğrulanmaktadır. Sıcak noktalar nedeniyle enerji hasadında %15’e varan düşüşler meydana

gelebilmektedir. Bu nedenle yapılan çalışmada, sıcak noktaların erken evrede teşhis edilmesi için ML tabanlı bir araç geliştirilmiştir. En uygun ML modelinin belirlenmesi için DT, SVM, KNN ve Diskriminant Sınıflandırıcı (DC) olmak üzere dört farklı sınıflandırma algoritması kullanılmaktadır (Dhimish, 2021). Belirlenmiş olan bir FV güç sistemi üzerinde ölçümler yapıp elde edilen verilere göre sıcak noktalar üç farklı kategoriye ayrılmaktadır. Bununla birlikte farklı parametre verilerinden, $S1(P_{mpp})$, $S2(P_{mpp}$ ve I_{sc}) ve $S3(P_{mpp}$, I_{sc} ve V_{oc}) şeklinde 3 farklı veri seti oluşturulmaktadır. ML modellerinin sıcak noktaları tespit etmekte kullanılabilmesi için bu veri hazırlama işlemi fazlasıyla önemlidir. Çalışma sonucunda bulunan sonuçlar karşılaştırılıp Diskriminant Sınıflandırıcıların %98 ile en yüksek tespit doğruluğuna sahip olduğu, en düşük tespit doğruluğunun ise %84 ile DT algoritması için bulunduğu paylaşılmaktadır. Ayrıca kullanılan farklı ML algoritmalarının tamamında, en yüksek sonucun en fazla parametreden oluşan S3 veri seti kullanıldığında elde edildiği görülmektedir.

Yapılan derleme çalışmasında (Liu et al., 2023), PSC ve bileşen malzemeleri için ML kullanımında temel teknolojiler ve beklentiler özetlenmektedir. Yeni nesil fotovoltaik teknolojilerden olan PSC'ler, birçok araştırmacı tarafından büyük ilgi görmektedir. PSC'ler ticarileşme için en umut verici üçüncü nesil güneş hücreleridir ve güç dönüşüm verimliliği, 10 yıldan kısa süre içinde %25,7'ye yükselmektedir. PSC, birden fazla malzemeden oluşan karmaşık bir sistem olduğundan, özellikleri (elektriksel özellikler, kararlılık, toksisite vb.) birçok faktörden etkilenmektedir. Tüm etki faktörlerini; malzeme özellikleri (perovskit malzeme katmanının kimyasal bileşimi, tanecik boyutu, yüzey kusurları ve film kalitesi, taşıma malzemesi katmanının mobilitesi ve iletkenliği, elektrot malzemesinin özellikleri vb.), üretim yöntemleri (damla kaplama yöntemi, atomik katman biriktirme yöntemi, serigrafi yöntemi, çözgen ve anti-çözgen seçimi, tavlama sıcaklığı ve süresinin seçimi vb.), malzeme uyumluluğu (enerji seviyesi hizalaması ve taşıma katmanı ile perovskit katmanının kimyasal uyumluluğu, arka arayüz kontak vb.) ve aygıt özellikleri (hücre yapısı, hücre boyutu, kapsülleme yöntemi vb.) şeklinde dört grupta özetlemek mümkündür. Bu derlemede, PSC alanında yer alabilecek ML teknolojilerini tanımlanmakta ve ML teknolojilerinin PSC'ler için mevcut uygulamaları sınıflandırılmaktadır. Bunun için 67 makale ayrıntılı olarak incelenmekte ve kullanılan teknikler özet olarak paylaşılmaktadır.

ML kavramı 5 başlıktan oluşmaktadır. Bunlar; veri, öznitelik, algoritma, doğrulama ve yorumlamadır. PSC alanında veriye ulaşmak için son dönemde geliştirilmiş olan veri tabanları gösterilmektedir. Derlemede yararlanılmak için belirlenmiş olan çalışmalarda en yaygın olarak kullanılan ilk on ML algoritması sıralanmaktadır. Sonuçta PSC alanında yayınlanmış olan 67 makale, 10 boyutta yapı-söküme tabi tutulmuştur. Bunlar:

- Tarih (Makalenin yayınlanma zamanına göre)
- Veri ve kod durumu (Veri kaynağı, verilerin ilk durumundan bağımsız olarak yazarın toplama kaynağıdır. Veri koşulu, model eğitimi ve testi için kullanılan verilerdir ve kod koşulu, yazarların çalışan kodu paylaşıp paylaşmadığına dayanmaktadır.)
- Nesne (Etiket kaynağına göre perovskit güneş hücreleri veya perovskit malzemeler)
- Sorunlar (Algoritmanın çözdüğü problemin özellikleri)
- Girdi (Girdilerin sayısı ve türü de dahil olmak üzere optimal modelin girdi veri durumu)
- ML algoritmaları (Çalışmada kullanılan ML algoritmaları)
- Etiket (Model eğitimi için kullanılan etiket özellikleri)
- Performans (Optimal algoritma altında model performansı)
- Açıklama yöntemi (ML modelini ve özniteliklerini açıklama yöntemi)
- ML Amacı (ML modelinin uygulaması ve anlamı)

Son birkaç yıldır, PSC'lere yönelik araştırmalara yardımcı olacak ML kullanılan makale sayısı giderek artmaktadır. Yapılan derleme çalışmasında (Chen et al., 2023), araştırılan soruların ve anlatıların türüne, hangi tür verilerin yararlı olabileceğine ve ML modellerini eğitmek için ilgili özniteliklerin nasıl çıkarılacağına odaklanarak genel bir bakış sunulmaktadır. Yapılan çalışma üç konu başlığı üzerine yoğunlaşmaktadır. Bunlar:

1) Deneysel verilerin anlamlandırılması

Literatürden toplanan verilerin kullanılması

- Bütünsel cihaz performansının modellenmesi
- İçsel malzeme özelliklerinin modellenmesi

- Molekül özelliklerinin modellenmesi

Deneyssel olarak elde edilen verilerin kullanılması

- Sonuçların modellenmesi
- Öznitelik mühendisliği
- Görüntü ve zaman serisi verileri

2) Hızlandırılmış analiz ve otomasyon için ML

3) Yeni malzemeler için teori odaklı araştırma

- Kimyasal uzayın hangi bölümünün keşfedileceğini seçme
- Hangi özelliklerin tahmin edileceğine karar verme
- Eğitim verisi bulma ve oluşturma
- Öznitelikler ve öznitelik oluşturma şemaları
- Doğru modeli seçme
- Öngörüler

Organik, boyaya duyarlı ve perovskit güneş hücrelerinde yaygın olarak kullanılan çözelti ile işlenmiş ince filmler, daha kolay üretim süreçleri, daha düşük üretim maliyeti ve malzemelerin ağırlık başına potansiyel olarak daha yüksek gücü nedeniyle önemli bir alternatif olmaktadır. Yapılan çalışmada (Zhang et al., 2023) çözelti ile işlenmiş fonksiyonel ince filmlerin mühendisliği için SPINBOT adı verilen tam otomatik bir platformun tasarımı ve kullanımı sunulmaktadır. Bilgiye dayalı malzeme bilimi, manuel rutinlerin ve geleneksel deneme-yanılma yaklaşımlarının hala baskın olduğu bir alandır. Bunun tipik bir örneği, aygıt üretimi için fonksiyonel ince film proses optimizasyonudur. Özellikle yüksek performanslı aygıtlar için proses koşullarının belirlenmesi, hala büyük ölçüde tek bir laboratuvardaki uzmanların ampirik kararlarına dayanmakta ve bir defaya mahsus denemelerin ve bütünsel olmayan yaklaşımların kısıtlamalarının temelini oluşturmaktadır. Sonuç olarak, bu "optimize edilmiş" prosedürler tipik olarak tek bir laboratuvar için işe yaramakta, ancak laboratuvar içi tekrarlanabilirlik ve laboratuvarlar arası aktarılabilirlik zayıf kalmaktadır. Özellikle perovskit alanında bu durum, yayınlanmış çok sayıda proses koşulu ve aygıt performansındaki yüksek standart sapma ile ortaya konmaktadır. Çalışmada geliştirilmiş olan SPINBOT,

olağanüstü deneysel kontrol ile yüzlerce alt tabakanın denetimsiz işlenmesi yoluyla yüksek örnekleme değişkenliğine sahip deneyler gerçekleştirebilmektedir. Genellikle Gaussian Proses Regresyon (GPR) modelini kullanan Bayesian Optimizasyonu yönteminin, girdi ve hedef özellikler arasında görselleştirilmiş bir eşleme sağlayarak çok boyutlu değişken uzaylarına sahip problemleri optimize etmede etkili olduğu kanıtlanmaktadır. Bayesian optimizasyon algoritmasının sağladığı yinelemeli optimizasyon süreci sayesinde SPINBOT, karmaşık bir parametre alanını araştırarak üretilen ince filmlerin kalitesini ve tekrarlanabilirliğini sürekli olarak iyileştirmektedir. ML rehberliğinde SPINBOT platformu, filmlerin basit bir fotoluminesans karakterizasyonu yoluyla perovskit güneş hücrelerinin optimizasyon prosesinin hızlandırılmasını sağlamaktadır. Sonuç olarak, ortam atmosferinde güneş hücresi üretim prosesiyle, yüksek tekrarlanabilirlikle %21,6'lık PCE sağlayan optimum bir film elde edilmektedir. Ayrıca enkapsüle edilmemiş aygıtların, 60-65°C'de 1100 saat sürekli çalıştıktan sonra, ilk verimliliklerinin %90'ını koruduğu paylaşılmaktadır.

PSC'ler için son on yılda rekabetçi verimlilik değerlerine ulaşılmış olsa da kararlılık sorunları, teknolojinin ticarileşme noktasında ele alınması gereken kritik bir zorluk olmaya devam etmektedir. Yapılan çalışmada (Hartono et al., 2023) Yüksek Verimli Yaşlandırma Sistemiyle son 3 yıl içinde toplanmış olan Maksimum Güç Noktası İzleme (MPPT) veri setinin analizi yapılmaktadır. Farklı perovskit malzemeler (hem organik hem de inorganik), yük seçici katmanlar (küçük moleküller, polimer, inorganik), kontak katmanlar (gümüş, bakır ve altın) ve mimarilerden (hem n-i-p hem de p-i-n) oluşan aygıtlardan, kontrollü koşullar altında (sürekli aydınlatma, kontrollü sıcaklık ve atmosfer) elde edilen 2.245 MPPT yaşlanma eğrisi analiz edilmektedir. Yaşlandırma testi sırasında ulaşılan maksimum verimliliğe göre 150 saat sonra PCE kaybının analizi yapılmaktadır. Maksimum verimliliğin, test edilen güneş hücresinin maksimum potansiyelini veya kapasitesini yansıttığını düşünülmektedir. Buna ek olarak, PSC'ler farklı gruplara ayrılabilen çeşitli bozunma eğrisi şekilleri göstermektedir. Çalışmada, veri setinden bozunma eğrisi kümeleri elde etmek için kendi kendini organize eden harita (SOM) adı verilen denetimsiz bir ML yöntemi uygulanmakta ve yaşlanma eğrilerinin baskın şekli belirlenebilmektedir. Sonuçta belirli şekil kümelerinin oluşumu ile ulaşılan maksimum PCE arasında bir korelasyon gözlemlenmektedir. Yaşlanma

testi verilerini kategorize etmenin ve ana şekilleri tanımlamanın PSC'ler için uygun yaşam süresi ölçütlerini bulmaya yardımcı olacağı sonucu paylaşılmaktadır.

PSC'ler, gelişmekte olan en dinamik fotovoltaik teknolojidir ve dünya çapında binlerce araştırmacının ilgisini çekmektedir. Son zamanlarda, bu teknoloji için temel zorluklardan biri olan cihaz kararlılığı sorunları üzerinde de çalışmalar yapılmış ve bu da önemli miktarda veri toplanmasıyla sonuçlanmıştır. Bunun en iyi örneği, kararlılıkla ilgili ölçümleri de içeren Perovskite Veri Tabanı projesidir. Yapılan çalışmada (Graniero et al., 2023), bu veri tabanından cihaz kararlılığının rapor edildiği 1.800 PSC hakkındaki veriler kullanılarak, hücre kararlılığı için en önemli faktörleri belirlemek ve incelemek için RF algoritması kullanılmaktadır. Cihaz performansının oldukça standartlaştırılmış bir değeri olan PCE'nin aksine, yaşlanma deneylerinden elde edilen zaman serisini tek bir sayıya indirgeyecek genel kabul görmüş bir cihaz kararlılığı değeri yoktur. T80 en yaygın kararlılık ölçütlerinden biridir ve tek tip bozunma eğrileri gösteren güneş teknolojileri için yeterince iyi çalışmaktadır. T80, Perovskite Veri Tabanında raporlanmış ve bu çalışmada ML modellemesi için hedef değişken olarak kullanılmaktadır. Ayrıca, tek bir sayıya dayalı kararlılık ölçütlerinin aksine her bir hücre için tüm yaşlanma eğrisinin mevcut olduğu 1.000'den fazla güneş hücresi verisine sahip kurum içi veri tabanı incelenmektedir. Kurum içi veri seti, Helmholtz-Zentrum Berlin'deki "Kararlı Perovskit Güneş Hücreleri için Aktif Malzemeler ve Arayüzler" bölümünde toplanmıştır ve 2019-2022 yıllarında çeşitli tiplerde 1.000'den fazla PSC üzerinde gerçekleştirilen yaşlandırma deneylerinin zaman serisi verilerini içermektedir. RF modelinin tüm veri kümesi ve aynı zamanda diğer iki veri kümesi bölümü için belirlediği en önemli 20 öznitelik ve bu özniteliklerin önemi gösterilmektedir. RF modeli, PSC'lerin termal, nem veya foto-kararlılığını etkilediği bilinen birçok özelliği yakalamaktadır. Sıcaklık, bağıl nem gibi değişkenler tahmin edilebileceği gibi en etkili faktörler arasındadır. Perovskit yapısında yer alan, özellikle tek değerlikli katyon olarak MA veya FA organik türlerinin veya anyon olarak iyotun varlığı genellikle daha az dirençli perovskit güneş hücreleriyle sonuçlanmaktadır. Elektron veya boşluk taşıma katmanları için seçenek olarak araştırılan çoklu taşıma malzemeleri cihaz kararlılığını önemli ölçüde etkiler ve RF tahminleri de bu noktada uyum göstermektedir.

Elektrokimyasal empedans spektroskopisi (EIS), metal halojenür PSC'lerin karakterizasyonu ve analizi için çok yönlü bir teknik olarak ortaya çıkmıştır. PSC'lerde iyon göçü ve taşıyıcı birikimi hakkındaki önemli bilgiler, EIS spektrumunun düşük frekanslı rejiminden çıkarılabilmektedir. Bununla birlikte, düşük frekanslarda uzun ölçüm süresi ve uzun süre ışığa ve ön gerilime maruz kalma nedeniyle malzeme bozulması gözlenmektedir. Yapılan çalışmada (Parikh et al., 2023) halojenür perovskit tek kristallerinin düşük frekanslı tepkisini tahmin etmek için ML modeli geliştirilmektedir. Çalışmada ilk olarak yüksek kaliteli MAPbBr₃ tek kristallerinin sentezi yapıp ardından farklı ışıyım (88, 250, 382, 490 ve 575 W/m²) ve ön gerilim (0, 0.2, 0.4, 0.6, 0.8 ve 1V) değerlerinde 8741 veri noktasından oluşan veri kümesini hazırlamak için EIS spektrumları kaydedilmektedir. LR, DT, GPR, GBR ve RF algoritmaları kullanılarak elde edilen veri kümesi üzerinde farklı ML modelleri oluşturulmaktadır. Geliştirilen ML modellerinden RF, test seti için 0,981 R² skoru ve 0,0196 RMSE değeri ile düşük frekanslı EIS yanıtının gerçek ve hayali kısımlarını tahmin edebilmektedir. Gerçek deneysel verilerden, daha yüksek uygulanan ön gerilimde negatif kapasitansın hâkim olduğu gözlemlenebilmektedir. Geliştirilen model, düşük frekansta (50Hz-300mHz) gerçek ve hayali kısımları yakından tahmin edebilmektedir. Böylece EIS için ML modellerinin kullanılmasında yeni bir yol açmaktadır.

Geçtiğimiz on yıl içinde, halojenür PSC'ler geleneksel ince film fotovoltaiik teknolojilerini geride bırakarak en yüksek güç dönüşüm verimliliğinde %25,9'luk değere ulaşmıştır. Tekno-ekonomik analizler, ticari uygulanabilirliğe ulaşmak için PSC'lerin üretiminin daha geniş bir alana ve daha yüksek verime ölçeklendirilmesi gerektiğini göstermektedir. Fotonik kütleme, yüksek verimli PSC üretimini mümkün kılabilen bir ince film işleme tekniğidir. Fotonik kütleme veya yoğun darbeleri ışık tavlama sırasında, geniş bant spektrumlu (200-1500 nm) yoğun foton darbeleri (20µs-100ms) sağlamak için bir flaş ksenon lambası kullanılmaktadır. İnce filmler, yüksek yoğunluklu ancak düşük enerjili ışık darbelerinden gelen enerjiyi absorbe ederek tamamen kristalleşmektedir. Fotonik kütlemede, perovskit kristalizasyonu 20ms gibi çok kısa sürelerde gerçekleşmekte, tipik olarak 10 dakika veya daha uzun süre gerektiren termal tavlama ile kıyasla yüksek verimli aygıtlar elde edilebilmektedir. Bu nedenle, bu teknik PSC cihazlarının üretim ölçeğinin büyütülmesi için uygundur. Fotonik kütleminin işleme sonucunu etkileyebilecek

birçok değişkeni vardır ve bu da optimizasyonu zorlaştırmaktadır. Yapılan çalışmada (Xu et al., 2023), ITO kaplı Willow Glass üzerinde fotonik olarak kürlenmiş MAPbI_3 PSC'lerin güç dönüşüm verimliliğini optimize etmek için bir ML çerçevesi olan Bayesian Optimizasyonu kullanılmaktadır. Fotonik kütleme, PSC'lerin üretimi için umut vaat etse de darbe voltajı, darbe uzunluğu, darbe sayısı ve tekrarlama hızı gibi birçok işlem değişkenini içermektedir. Buna ek olarak, perovskit prekürsör formülasyonu da PSC'lerin nihai güç dönüşüm verimi üzerinde önemli bir etkiye sahip olabilmektedir. Çok sayıda girdi değişkeni PCE değerinin maksimize edilmesini zorlaştırmaktadır. Araştırmacıların bilgilerinin rehberliğinde deneme yanılmaya dayanan, geleneksel tek seferde tek değişkenli yaklaşım zaman almakta ve genellikle yeni bir sürecin tam olarak anlaşılmasına yol açmaktadır. Buna karşılık, sofistike deney tasarımı çok boyutlu süreç optimizasyonu için büyük deneysel kaynaklar, yani büyük bir veri seti gerektirmektedir. Bu nedenle, akıllı ve verimli bir optimizasyon stratejisine ihtiyaç duyulmaktadır. Sıralı bir ML algoritması olan BO, bir amaç fonksiyonu üzerindeki önceki ve sonraki dağılımlara odaklanan küresel bir optimizasyon stratejisidir. BO, çok yönlülüğü ve küçük veri kümeleri için uygunluğu nedeniyle, birçok yüksek boyutlu, pahalı optimizasyon probleminde kullanılmaktadır. Yapılan çalışmada, esnek yüzey üzerindeki MAPbI_3 güneş hücrelerinin PCE değerini en üst düzeye çıkarmak için BO kullanılmış ve sınırlı deneysel 48 koşulla, %11,42'lik bir PCE değeri elde edilmiştir. Ayrıca PCE değerinin %10 ve üzerinde olduğu 14 yeni koşul öngörülmüştür. BO'nun süreç optimizasyonunda güçlü bir araç olduğu ve diğer PSC üretim çalışmalarına uyarlanabileceği sonucu paylaşılmaktadır.

ML fotovoltaiik araştırmalarda giderek daha fazla dikkat çekmekte ve yakın gelecekte rekor düzeyde yüksek PCE değerine ulaşmada hayati bir araç olacağı öngörülmektedir. ML'nin önemli ölçüde faydalı olduğu alanlardan biri de güneş hücresi üretiminde optimum parametre kombinasyonunu bulmak için gereken deney sayısının azaltılmasıdır. Sıralı ML yöntemleri, büyük boyutlu parametre uzaylarına sahip problemlerin hızlı optimizasyonuna olanak tanımaktadır. BO sıralı öğrenmenin pratik yaklaşımlarından biridir. Yeni örnekleme konumlarını verimli bir şekilde oluşturmak ve optimum değerleri hızlı bir şekilde belirlemek için mevcut verileri, Bayesian çıkarımını kullanarak birleştirmektedir. BO algoritmalarının grid arama gibi diğer yöntemlere göre başlıca avantajlarından bazıları, küçük veri

kümeleriyle iyi çalışması, amaç fonksiyonu üzerinde herhangi bir kısıtlama olmaması (yani, bir türevin varlığına gerek olmaması) ve çok sayıda girdi parametresine (>100 parametre) sahip problemlerde verimli optimizasyon sağlamak için uygulanabilmesidir. BO geniş parametre uzayına sahip problemlerde optimum parametreleri hızlı bir şekilde belirlemek için yollar sağlamaktadır. Yapılan çalışmada (Tsoi and Yerci, 2024), önceki bilgileri kullanan BO algoritmalarının, rekor düzeyde yüksek performans sunma potansiyeli nedeniyle hızla daha fazla ilgi gören bir teknoloji olan tandem güneş hücrelerinde kullanımı gösterilmektedir. Çalışmanın üzerinde kurulmuş olduğu parametre uzayı 9 farklı parametreden oluşmaktadır. Her bir parametreye karşılık gelen artış değerleriyle birlikte ($11 \times 11 \times 10 \times 4 \times 4 \times 7 \times 7 \times 7$) toplam 46.483.360 farklı kombinasyon oluşmaktadır. Çalışmada BO kullanılarak, değerlendirilmesi ≈ 88 yıl süren tüm olası parametre kombinasyonları uzayında, optimum PCE değerinin 20 dakikada elde edilebileceği gösterilmektedir. Ayrıca, önceki bilgileri kullanan yöntemlerin, 1. iterasyonda $\approx 8,9$ artış sağlayarak ve 20 farklı denemede $38,5$ 'lik bir PCE değerine ulaşmak için 5 kat daha az zaman gerektirdiği gösterilmektedir. Bunların yanında çalışmadan elde edilen sonuçların, optimum parametrelerin belirlenmesinde çok sayıda deney yapma ihtiyacını ortadan kaldırarak tandem güneş hücrelerinin geliştirilmesini hızlandırmaya yardımcı olacağı sonucu paylaşılmaktadır.

2.2. Tezin Önemi

Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” adlı projenin hedeflerinden biri olan ve kurulumu tamamlanan “Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı (Class 1000 Temiz Oda)” hâlihazırda kullanılmaya başlanmıştır. Bu laboratuvarda hammaddeden son ürüne kadar yeni nesil güneş hücresi üretimi yapılabilmektedir. “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” projesinin öncelikli amacı güneş hücrelerinin endüstriyel boyutlara taşınması aşamasında yöntem geliştirmek ve optimizasyonu sağlamaktır. En uygun yapının tespit edilip uygulanması noktasında yüksek miktarda kimyasal sarf malzeme kullanılmaktadır. Kullanılan malzemelerin maliyeti sebebiyle hedeflenen

sonucun olabildiğince az deneme sonucunda elde edilmesi kritik önem taşımaktadır. Ayrıca hedeflenen sonuca mümkün olan en kısa zamanda ulaşmak konusu da büyük önem arz etmektedir.

Yapılan tez çalışmasıyla, elde edilen veriler kullanılarak yeni nesil güneş hücresi yapısının en yüksek verim elde edilecek şekilde en uygun büyük ölçek yapısının tasarım ve üretim sürecinin belirlenmesi amaçlanmaktadır. Literatürdeki çalışmaların büyük bir kısmının halen hücre ölçeğinde olduğu görülmektedir. Bu durum, modül ölçeğinde yapılacak çalışmaların önemini artırmaktadır. Modül ölçeğinde yapılacak araştırmalar, perovskit güneş hücrelerinin ticari uygulamalar için uygun hale getirilmesi açısından kritik öneme sahiptir. ML kullanılarak geliştirilmesi planlanan modelin en yüksek başarımlı değere ulaşması için elde edilecek çıktıların Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı'nda deneysel çalışmalar vasıtasıyla doğrulanması da yapılacaktır.

2.3. Tezin Amacı

Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan, Kalkınma Bakanlığı destekli 2016K12-2841 numaralı “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” adlı projenin hedeflerinden biri olan ve kurulumu tamamlanan “Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı (Class 1000 Temiz Oda)” hâlihazırda kullanılmaya başlanmıştır. Bu projedeki nihai hedef, organik yarıiletken temelli fotovoltaik sistemlerin hücreden modül aşamasına kadar üretim süreçlerinin geliştirilmesi ve ilgili standartlara göre testlerinin yapılmasına imkân verecek altyapının geliştirilmesidir. Bu süreçteki en kritik nokta, fotovoltaik sistemlerde hücre ölçeğinden modül ölçeğine boyut artırımının yapılması; bu ölçek büyütme işlemine uygun üretim prosesinin geliştirilmesi ve optimizasyonudur. Literatürde yapılan çalışmalarda görüleceği üzere yeni nesil fotovoltaik teknolojiler için en uygun tasarım ve üretim yapısının oluşturulması aşamasında ML son dönemde yaygın olarak kullanılmaktadır. Literatürün gösterdiği bir diğer nokta ise yapılan çalışmaların çok büyük bir bölümünün halen hücre ölçeğinde olmasıdır. Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülmekte olan araştırmaların sonuçlarının da dâhil edilmesiyle oluşturulacak veriler kullanılarak en yüksek

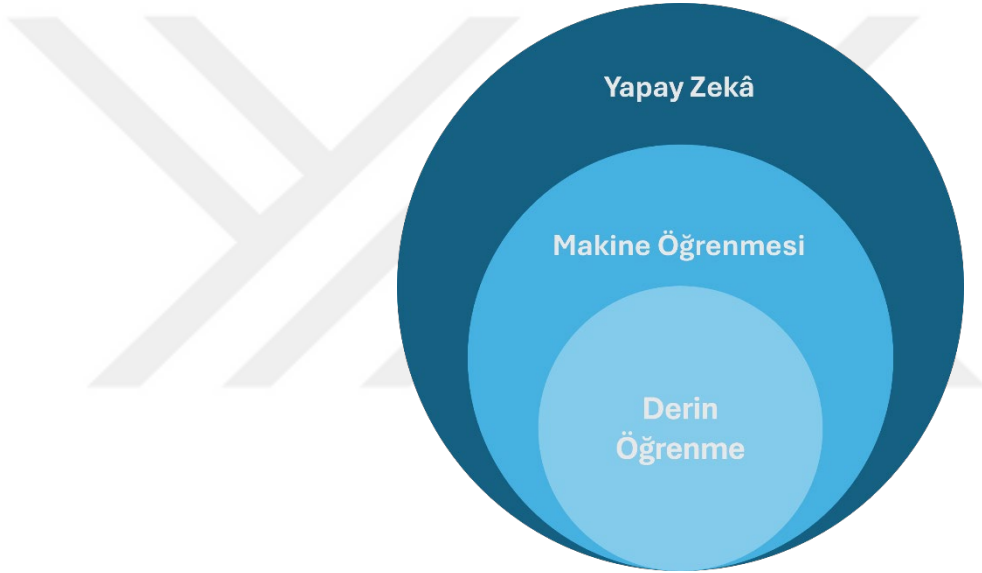
verimin elde edileceği büyük ölçek üretim yapısının optimizasyonunda ML kullanılması amaçlanmaktadır. Kullanılacak olan ML yöntemleri olarak LR, KNN, SVR, RF, XGBoost ve CatBoost seçilmiştir. Seçilmiş olan ML modellerinin değerlendirilmesi için kullanılacak metrikler; RMSE ve Pearson korelasyon katsayısı olarak belirlenmiştir. Bu optimizasyon sürecinde, geliştirilmesi planlanan modellerin en yüksek başarımlarına ulaşması hedeflenmektedir. Tez çalışmasında elde edilecek olan çıktılar, “Yeni Nesil Fotovoltaik Teknolojilerin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi” adlı Kalkınma Bakanlığı destekli projede girdi olarak kullanılacak olup, Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı’nda deneysel çalışmalar vasıtasıyla elde edilecek güneş hücreleri aracılığıyla doğrulanması amaçlanmaktadır.

2.4. Tez Kapsamı

Tez çalışması sekiz bölümden oluşmaktadır. Giriş bölümünün ardından, konu ile ilgili yapılan daha önceki çalışmalara, “Literatür Özeti” başlığında değinilmektedir. Ardından “Makine Öğrenmesi” başlığında, yapay zekânın tarihçesinden, makine öğrenmesinden ve makine öğrenmesi kategorilerinden üçüncü bölümde bahsedilmektedir. Dördüncü bölümde “Fotovoltaik Teknolojilerin” tanımları gerçekleştirilip, ayrıldığı alt bölümlerden ve bu bölümlerde yer alan fotovoltaik güneş hücrelerinden kısaca bahsedilip, perovskit güneş hücreleri hakkında temel bilgiler paylaşılmaktadır. Beşinci bölümde, tez çalışmasında kullanılan gereç ve yöntemlere değinilmektedir. Altıncı bölümde tez çalışmasına hazırlık olarak da değerlendirilebilecek olan “Yapılan Diğer Çalışmalar” bölümü yer almaktadır. Yedinci bölüm olan “Bulgular ve Tartışma” kısmında Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde üretimi gerçekleştirilmiş olan perovskit güneş hücresi parametreleri ve elde edilen veriler kullanılarak geliştirilmiş olan makine öğrenmesi modellerinin performans sonuçları yer almaktadır. Sekizinci ve son bölüm olan “Sonuç ve Öneriler” kısmında ise yapılan çalışmalarla ilgili genel sonuçlar verilmektedir. Ayrıca daha sonraki çalışmalar için öneriler bu bölümde yer almaktadır.

3. MAKİNE ÖĞRENMESİ

Yapay zekâ, 1950’li yıllarda yeni gelişmeye başlamış ve bilgisayar bilimcilerin bugün de cevabını aradığı “bilgisayarlar düşünebilir mi?” sorusuna cevap ararlarken ortaya çıkmış bir kavramdır. Kısaca yapay zekâ “insanlar tarafından yerine getirilebilen düşünsel faaliyetlerin otonom hale getirilmesi” olarak tanımlanabilmektedir. Yapay zekâ; makine öğrenmesini, derin öğrenmeyi ve daha fazlasını içine alan geniş bir alandır. Yapay zekâ ve alt kümelerinde yer alan makine öğrenmesi ve derin öğrenme kavramlarının küme olarak ifadesi Şekil 3.1’de gösterilmektedir (Chollet, 2017).



Şekil 3.1 Yapay zekâ, makine öğrenmesi ve derin öğrenme (Chollet, 2017)

3.1. Yapay Zekânın Tarihçesi

Başlangıç:

- 1943: Günümüzde yapay zekâ olarak geçen ilk çalışma Warren McCulloch ve Walter Pitts tarafından yapılmış ve ilk yapay nöron modelini önermişlerdir.
- 1950: Harvard'da iki lisans öğrencisi olan Marvin Minsky ve Dean Edmonds, 1950 yılında ilk sinir ağı bilgisayarını inşa etmişlerdir. SNARC adı verilen bu bilgisayar, 40 nörondan oluşan bir ağı simüle etmek için

3000 vakum tüpü ve bir B-24 bombardıman uçağından alınan ihtiyaç fazlası otomatik pilot mekanizması kullanılmaktaydı.

- 1950: Alan Turing yapay zekâ konusunda dönüm noktası çalışmalardan biri olan “Bilgisayar Makineleri ve Zekâ” başlıklı makalesinde, "Makineler düşünebilir mi?" sorusunu sorarak Turing Testini tanıtmıştır. Bu test, bir makinenin insan gibi davranabilme yeteneğini ölçmeyi amaçlamaktadır. Bu makalede ayrıca makine öğrenmesi, genetik algoritmalar ve pekiştirmeli öğrenmeyi tanıtmıştır.

Yapay Zekânın Doğuşu:

- 1956: John McCarthy, Marvin Minsky, Nathaniel Rochester ve Claude Shannon tarafından düzenlenen Dartmouth çalıştayı, bir araştırma alanı olarak yapay zekanın başlangıç noktasını temsil etmektedir. Duyurusunda yer alan "yapay zekâ" terimi McCarthy tarafından ilk kez burada kullanılmaktadır. Princeton'dan Trenchard More, IBM'den Arthur Samuel, MIT'den Ray Solomonoff ve Oliver Selfridge, Carnegie Tech'den Allen Newell ve Herbert Simon iki ay süren çalıştayda yer alan diğer katılımcılardır. Çalıştay, tüm önemli isimlerin birbiriyle tanışması ve sonraki 20 yıl boyunca alana bu kişiler, onların öğrencileri ve meslektaşlarının hâkim olacak olması sebebiyle de önem arz etmektedir.

Gelişmeler:

- 1958: McCarthy, MIT AI Laboratuvarında, gelecek 30 yıl boyunca hâkim yapay zekâ programlama dili olacak olan LISP dilini tanımlamıştır.
- 1966: Joseph Weizenbaum tarafından MIT'de geliştirilen erken dönem bir doğal dil işleme bilgisayar programı olan ELIZA geliştirilmiştir.
- 1969: Stanford Araştırma Enstitüsü'nde problem çözme yeteneklerine sahip, algılama, planlama ve yürütmeyi birlikte yerine getirebilen ilk robot olan Shakey geliştirilmiş ve daha sonra yapılacak olan pek çok araştırmayı etkilemiştir.

Uzman Sistemler:

- 1969: Stanford'da Ed Feigenbaum, Bruce Buchanan ve Joshua Lederberg tarafından bir kütle spektrometre cihazından sağlanan bilgilerden, kimyasal analiz ve moleküler yapı çıkarımı yapabilen DENTRAL programı geliştirilmiştir.
- 1972: Feigenbaum, Buchanan ve Dr. Edward Shortliffe tarafından kan enfeksiyonlarını teşhis etmek için MYCIN geliştirilmiştir. MYCIN yaklaşık 450 kuralla, kıdemsiz doktorlardan önemli ölçüde daha iyi ve neredeyse bazı uzmanlar kadar iyi performans göstermiştir.
- 1977: Yale'de dilbilim kökenli yapay zekâ araştırmacısı Roger Schank ve öğrencileri doğal dili anlama görevini üstlenen bir dizi program geliştirmiştir (Schank ve Abelson, 1977; Wilensky, 1978; Schank ve Riesbeck, 1981; Dyer, 1983).

Yapay Zekâ Endüstrisinin Doğuşu:

- 1981: Japonya, akıllı bilgisayarlar üretmeye yönelik 10 yıllık bir plan olan "Beşinci Nesil" projesini duyurmuştur. Buna karşılık olarak Amerika Birleşik Devletleri'nde, ulusal rekabet gücünü sağlamak için tasarlanmış bir araştırma konsorsiyumu olan Mikroelektronik ve Bilgisayar Teknolojisi Şirketi (MCC) kurulmuştur. Her iki gelişme de çip tasarımı ve insan arayüzü araştırmalarını içermesi açısından yapay zekâ konusunda oldukça önem arz etmektedir.
- 1982: Ticarileşmiş ilk başarılı uzman sistem olan R1, Digital Equipment Corporation'da kullanılmaya başlamıştır. Yeni bilgisayar sistemleri için siparişlerin yapılandırılmasına yardımcı olan program, 1986 yılına gelindiğinde şirkete yılda yaklaşık 40 milyon dolar tasarruf sağlamaktaydı (McDermott, 1982).

Yapay Sinir Ağları:

- 1986: Yapay sinir ağlarının eğitiminde kullanılan geri yayımlı öğrenme algoritması (back-propagation) tanıtılmıştır. Algoritma, bilgisayar

bilimleri ve psikolojideki pek çok öğrenme problemine uygulanmış ve sonuçların Parallel Distributed Processing adlı derlemede paylaşılması büyük heyecan uyandırmıştır (Rumelhart ve McClelland, 1986).

- 1980'lerde sinir ağları üzerine yapılan çalışmaların çoğu, neler yapılabileceğini ve sinir ağlarının geleneksel tekniklerden nasıl farklı olduğunu öğrenmek amacıyla yapılmıştır. Sinir ağlarının gelişmiş metodoloji ve teorik çerçeveler kullanılarak istatistik, örüntü tanıma ve makine öğrenmesi alanlarında uygulanabileceği bir anlayışa ulaşılmıştır. Ayrıca mevcut gelişmeler veri madenciliği olarak adlandırılan yeni bir endüstrinin doğmasına neden olmuştur.

Derin Öğrenme ve Büyük Veri:

- 1997: IBM'in satranç oynayan bilgisayarı Deep Blue, dünya satranç şampiyonu olan Garry Kasparov'u yenmeyi başarmıştır.
- 2005: Yarışmacılardan otonom taşıt üretmelerinin beklendiği DARPA Grand Challenge yarışmasını, Stanford Üniversitesi'nin robotik aracı Stanley kazanmıştır.
- 2006: Geoffrey Hinton tarafından "derin öğrenme" terimi tanıtılmıştır ve algoritmaların daha derin sinir ağları ile eğitilmesine olanak sağlanmıştır.
- 2011: IBM'in Watson adlı bilgisayarı, doğal dil işleme alanında büyük bir gelişim göstererek Jeopardy! Yarışmasının şampiyonlarını yenmeyi başarmıştır.
- 2016: DeepMind şirketinin geliştirmiş olduğu AlphaGo yazılımı, dünyanın en iyi Go oyuncularından biri olan Lee Sedol'u yenmeyi başarmıştır.
- 2020: OpenAI firması tarafından doğal dil işleme ve üretiminde büyük ilerleme sağlayan bir dil modeli olan GPT-3 geliştirilmiştir.

3.2. Makine Öğrenmesi

Makine öğrenmesi, bilgisayar sistemlerinin belirli görevleri, programlanmadan otomatik olarak öğrenip geliştirmesini sağlayan yapay zekâ alt dalıdır. Klasik programlamada sembolik yaklaşımlar kullanılarak geliştirici girdi

olarak kuralları ve verileri ortaya koymakta, veriler bu kurallardan geçirilerek bir cevap elde edilmektedir (Moroney, 2020). Klasik programlamada yer alan öğeler şematik olarak Şekil 3.2’de gösterilmektedir.



Şekil 3.2 Klasik programlamanın şematik olarak gösterimi (Moroney, 2020)

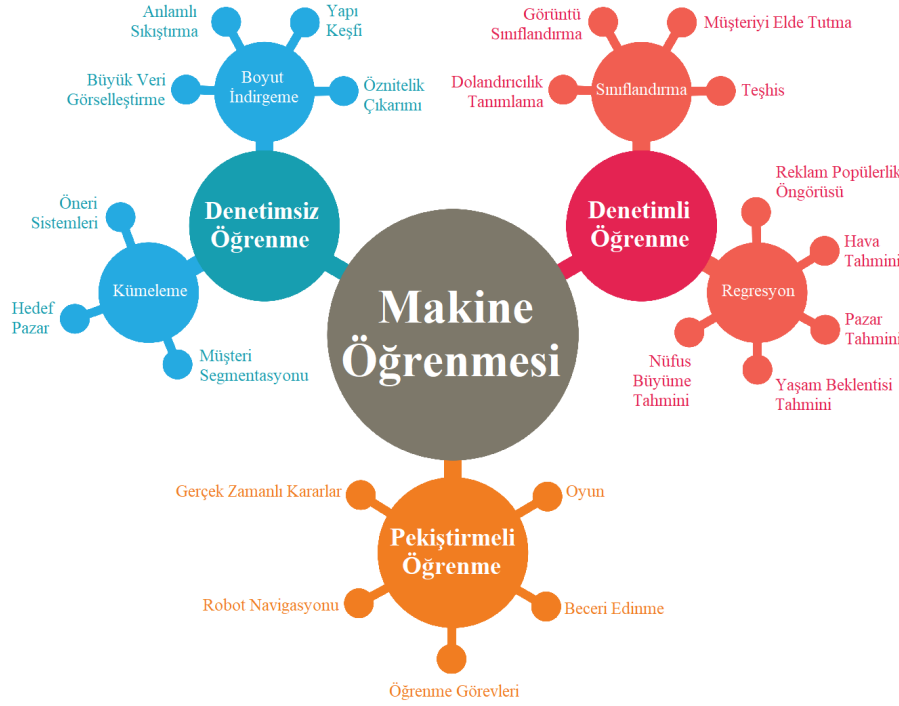
Makine öğrenmesindeyse klasik programlamada yer alan öğelerin yerleri değişmektedir. Kullanıcı tarafından girdi olarak belirlenen veri ve cevaplar kullanılarak yeni verilere özgün cevaplar üretebilecek kurallar öğrenilmektedir. Makine öğrenmesinin şematik olarak gösterimi Şekil 3.3’te yapılmaktadır.



Şekil 3.3 Makine öğrenmesinin şematik olarak gösterimi (Moroney, 2020)

Makine öğrenmesi sistemleri elle programlanmak yerine eğitilmektedir ve bu eğitim için ihtiyaç duyulan şeyler; veri giriş noktaları, beklenen çıktılar ve algoritmanın başarımının ölçülmesidir. Elde edilen sonucun değerlendirilebilmesi için beklenen çıktı ile gerçekleşen arasındaki farkın ölçülmesi gerekmektedir. Ölçüm, algoritmanın kendi başarımını en iyileştirmesi için geri bildirim olarak değerlendirilmektedir ve bu işlem “öğrenme” olarak adlandırılmaktadır (Chollet, 2017).

Makine öğrenmesi, genellikle büyük veri setleri üzerinde çalışarak ve veri üzerinde algoritmalar ve matematiksel modeller kullanılarak veriden bilgi çıkarma ve model oluşturma süreçlerine dayanmaktadır. Makine öğrenmesi; denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme olarak üç ana kategoriye ayrılmaktadır. Bu kategorilerin yer aldığı görsel Şekil 3.4’te gösterilmektedir.



Şekil 3.4 Makine öğrenmesi kategorileri

3.2.1. Denetimli Öğrenme

Denetimli öğrenme en yaygın kullanılan ML kategorisini oluşturmaktadır (Müller and Guido, 2017). Bir veri seti üzerindeki girdilerden bilinen çıktılara/hedeflere olan eşleştirmeyi öğrenmeyi kapsamaktadır. Kullanıcı algoritmaya girdi ve istenen çıktı çiftlerini sağlar ve algoritma bir girdi verildiğinde istenen çıktıyı üretmenin bir yolunu bulur. Girdi ve çıktılarından oluşan bir veri kümesi oluşturmak genellikle zahmetli bir süreç olsa da denetimli öğrenme algoritmaları iyi anlaşılır ve performanslarının ölçülmesi kolaydır. Sınıflandırma ve regresyon olarak adlandırılan iki ana denetimli makine öğrenmesi problemi türü vardır. Sınıflandırmada amaç, önceden tanımlanmış bir olasılıklar listesinden bir seçim olan sınıf etiketini tahmin etmektir. Sınıflandırma problemleri, tam olarak iki sınıf arasında ayırım yapmanın özel durumu olan ikili sınıflandırma ve ikiden fazla sınıf arasında sınıflandırma olan çok sınıflı sınıflandırma olarak ikiye ayrılmaktadır. E-postaları spam ya da spam değil olarak sınıflandırmak ikili sınıflandırma problemine bir örnektir. Regresyon problemlerinde amaç, sürekli bir

sayıyı veya matematiksel ifadeyle gerçek sayıyı tahmin etmeye çalışmaktır. Bir kişinin yıllık gelirini; eğitiminden, yaşından ve yaşadığı yerden tahmin etmek bir regresyon problemi örneğidir. Bir başka regresyon problemi örneği; önceki verim, hava durumu ve çiftlikte çalışan işçi sayısı gibi nitelikler göz önüne alındığında bir mısır çiftliğinin verimini tahmin etmektir.

Yapılan tez çalışmasında deneysel olarak üretimi gerçekleştirilen perovskit güneş hücrelerinin güç dönüşüm verimi hedef çıktı olarak kullanılarak üretim prosesindeki farklı parametre değerleri için aygıt verimleri tahmin edilmeye çalışılmaktadır. Girdi ve çıktı verilerinin yer aldığı veri seti kullanıldığı için yapılan çalışma denetimli öğrenme kategorisi altında yer almaktadır. En yaygın kullanılan denetimli öğrenme algoritmaları aşağıda sıralanmaktadır (Müller and Guido, 2017).

- En yakın komşular: Küçük veri kümeleri için, temel olarak iyi, açıklaması kolay.
- Doğrusal modeller: Denenecek ilk algoritma olarak, çok büyük veri kümeleri için iyi, çok yüksek boyutlu veriler için iyi.
- Naive Bayes: Sadece sınıflandırma için. Doğrusal modellerden bile daha hızlıdır, çok büyük veri kümeleri ve yüksek boyutlu veriler için iyidir. Genellikle doğrusal modellere göre performansları daha kötüdür.
- Karar ağaçları: Çok hızlıdır, verilerin ölçeklendirilmesine gerek yoktur, görselleştirilebilir ve kolayca açıklanabilir.
- Rastgele ormanlar: Neredeyse her zaman tek bir karar ağacından daha iyi performans gösterir, çok sağlam ve güçlüdür. Verilerin ölçeklendirilmesine gerek yoktur. Çok yüksek boyutlu seyrek veriler için iyi değildir.
- Gradyan destekli karar ağaçları: Genellikle rastgele ormanlardan biraz daha doğrudur. Rastgele ormanlara göre eğitilmesi daha yavaş ancak tahmin etmesi daha hızlıdır ve bellekte daha küçüktür. Rastgele ormanlara göre daha fazla parametre ayarı gerekir.

- Destek vektör makineleri: Benzer anlama sahip özelliklerden oluşan orta büyüklükteki veri kümeleri için güçlüdür. Verilerin ölçeklendirilmesini gerektirir, parametrelere duyarlıdır.
- Sinir ağları: Özellikle büyük veri kümeleri için çok karmaşık modeller oluşturabilir. Verilerin ölçeklendirilmesine ve parametrelerin seçimine duyarlıdır. Büyük modellerin eğitilmesi için uzun zaman gerekir.

3.2.2. Denetimsiz Öğrenme

Denetimsiz öğrenme, ML yöntemlerinden biridir ve bu yaklaşımda modeller etiketlenmemiş veri kullanılarak geliştirilmektedir (VanderPlas, 2017). Veriyi görselleştirmek, verilerdeki gürültüyü azaltmak veya eldeki veriler arasındaki korelasyonu anlamak bu kategorinin temel amacını oluşturmaktadır. Denetimsiz öğrenme, girdilerden herhangi bir bilinen çıktı olmaksızın veri setindeki gizli kalıpların ve yapıların anlaşılmasına odaklanmaktadır. Bu yaklaşım, verilerin doğal yapısını keşfetmek ve anlamak için kullanılmaktadır.

Denetimsiz öğrenme teknikleri arasında kümeleme ve boyut indirgeme önemli bir rol oynamaktadır. Kümeleme, veri noktalarını benzerliklerine göre gruplandırarak veri setindeki doğal kümeleri veya grupları ortaya çıkarmayı hedeflemektedir. Bu alanda yaygın olarak kullanılan algoritmalar arasında K-Ortalama, Hiyerarşik Kümeleme ve DBSCAN bulunmaktadır. Boyut indirgeme ise, yüksek boyutlu veri setlerini daha düşük boyutlu bir alana indirerek verilerin daha kolay anlaşılmasını ve görselleştirilmesini sağlamaktadır. Aynı zamanda, bu teknik veri setindeki gürültüyü azaltarak daha anlamlı özelliklerin belirlenmesine yardımcı olmaktadır. Boyut indirgeme için kullanılan algoritmalar arasında Temel Bileşen Analizi (PCA) ve t-Dağıtımli Stokastik Komşu Yerleştirme (t-SNE) gibi yöntemler bulunmaktadır.

Denetimsiz öğrenme, özellikle etiketlenmiş verinin az olduğu veya etiketlemenin maliyetli olduğu durumlarda oldukça faydalıdır. Ayrıca, veri setinin yapısını anlamak ve daha iyi öğrenme modelleri geliştirmek için veri keşfi ve ön işleme aşamalarında da sıkça kullanılmaktadır.

3.2.3. Pekiştirmeli Öğrenme

Pekiştirmeli öğrenme, bir ajanın belirli bir ortamda etkileşimde bulunarak ödül veya ceza alması üzerine kurulu bir ML türüdür (Russell and Norvig, 2010). Bu öğrenme sürecinde ajan, çevresiyle sürekli etkileşim halindedir ve her eylemi sonucunda bir ödül veya ceza almaktadır. Bu geri bildirim mekanizması, ajanın hangi eylemlerin daha iyi sonuçlar doğurduğunu öğrenmesine yardımcı olmaktadır.

Ajan, ödülleri maksimize etmek için stratejiler geliştirmekte ve bu stratejiler, zamanla daha iyi kararlar almasını sağlamaktadır. Bu süreç, deneme-yanılma yoluyla öğrenmeyi içermekte ve ajan, uzun vadede en yüksek toplam ödülü elde etmeyi hedeflemektedir. Pekiştirmeli öğrenme, ajanın karar verme sürecini optimize etmesine yardımcı olan Markov Karar Süreçleri (MDP) gibi matematiksel çerçevelerle modellenmektedir.

Pekiştirmeli öğrenmenin, çeşitli alanlarda geniş uygulamaları bulunmaktadır. Örneğin, robotik kontrol sistemlerinde, robotların çevreleriyle etkileşim kurarak görevleri öğrenmesi sağlanmaktadır. Oyun stratejilerinde, ajanlar, oyun ortamında en iyi hamleleri öğrenerek rakiplerini yenmeye çalışmaktadır. Otonom araçlar, çevrelerini algılayarak ve trafik kurallarına uyarak güvenli bir şekilde seyahat etmeyi öğrenmektedir. Ayrıca pekiştirmeli öğrenme; finansal portföy yönetimi, enerji sistemleri optimizasyonu ve sağlık hizmetleri gibi birçok farklı alanda da uygulanmaktadır.

4. FOTOVOLTAİK TEKNOLOJİLER

Fotovoltaik etki, güneş hücrelerinin çalışma prensibinin temelini oluşturmaktadır ve Fransız fizikçi Alexandre Edmond Becquerel tarafından 1839 yılında keşfedilmiştir. Becquerel yaptığı deneyde, ışığın asidik bir elektrolit çözeltisi içindeki platin elektrotlar üzerine düşmesiyle oluşan potansiyel farkı gözlemlemiştir. Becquerel'in gözleminin ardından, İngiliz mühendis Willoughby Smith 1873 yılında, fotovoltaik hücrelerin gelişiminde önemli bir adım olan selenyumun ışığa maruz kaldığında elektriksel iletkenliğinin arttığını keşfetmiştir. Bu gelişimin ardından Amerikalı mucit Charles Fritts 1883 yılında, selenyumun ince bir altın tabakası ile kaplanmasıyla %1'lik güç dönüşüm verimliliğine sahip olan ilk fotovoltaik hücreyi geliştirmiştir (Breeze, 2016).

4.1. Güneş Hücresi Teknolojileri

Fotovoltaik güneş hücresi teknolojileri, kullanılan malzemelere ve üretim yöntemlerine göre farklı kategorilere ayrılmaktadır (Fraas, 2015).

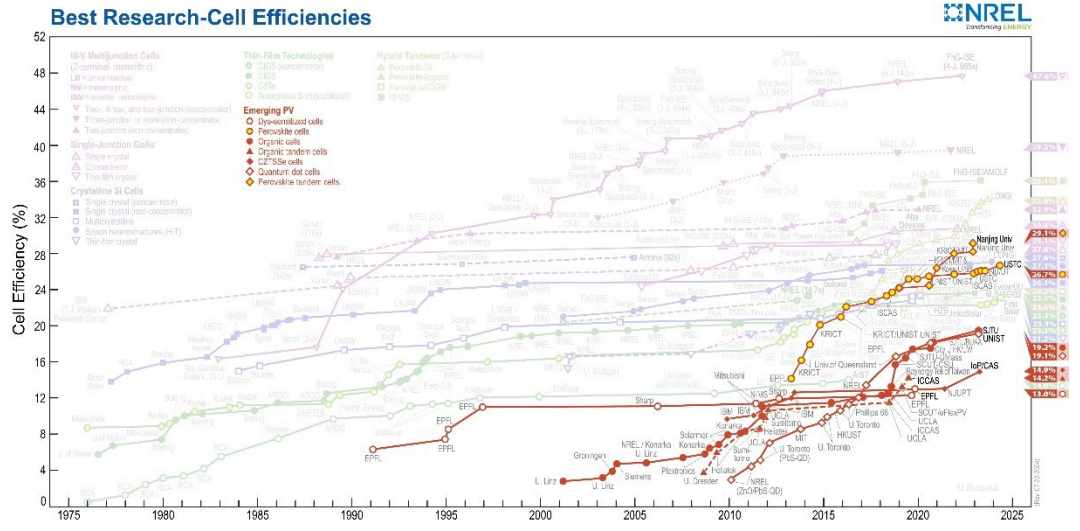
1.Nesil (Kristal silisyum güneş hücreleri): Tek kristal (monokristal) ve çok kristal (polikristal) silisyum güneş hücreleri

2.Nesil (İnce film güneş hücreleri): Amorf silisyum, kadmiyum tellür (CdTe), bakır indiyum galyum selenit (CIGS) ve bakır çinko kalay sülfür (CZTS)

3.Nesil (Yeni nesil fotovoltaik teknolojiler): Organik, polimer, DSSC (boya duyarlı güneş hücreleri), perovskit güneş hücreleri, kuantum nokta güneş hücreleri ve tandem (çok eklemlili) güneş hücreleri

4.1.1. Perovskit Güneş Hücreleri

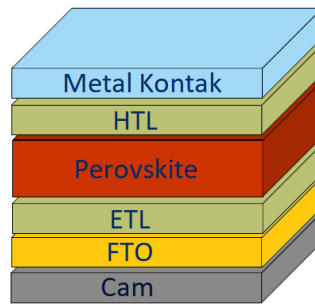
PSC'ler son on beş yılda laboratuvar ölçeğindeki verimlilikleri %3,8 seviyesinden %26,7 seviyesine yükselmiş ve bu süre zarfında akademik olarak fazlaca araştırmaya konu olmuş yeni nesil fotovoltaik teknolojilerdendir. Bahsedilen süre içindeki gelişiminin yer aldığı görsel Şekil 4.1'de gösterilmektedir (NREL, 2024).



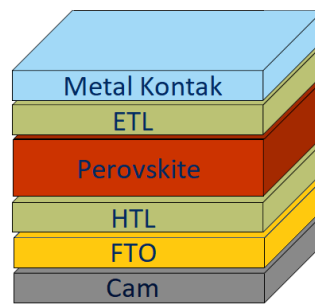
Şekil 4.1 Araştırma ölçeğindeki Perovskit güneş hücresi güç dönüşüm verimleri (NREL, 2024)

PSC'ler; yüksek absorpsiyon katsayıları, ayarlanabilir bant aralığına sahip olmaları, farklı yüzeylere uygulanabilmeleri, düşük maliyetleri ve kolay üretilibilmeleri gibi avantajlarından ötürü son yıllarda yeni nesil fotovoltaik teknolojileri alanında büyük potansiyel göstermektedir. ABX_3 kimyasal formülüyle gösterilen perovskit yapısında A ve B katyonları, X ise anyonu temsil etmektedir. Perovskit güneş hücresi; metal kontak, seçici yük taşıma tabakaları olan ETL/HTL, perovskit, flor katkılı kalay oksit tabaka (FTO) ve cam gibi farklı katmanların bir araya getirilmesiyle oluşturulmaktadır. Yaygın olarak kullanılan aygıt mimarileri Şekil 4.2'de gösterilmektedir (Fu et al., 2016).

a Düzlemsel n-i-p yapısı



b Düzlemsel p-i-n yapısı



Şekil 4.2 Yaygın olarak kullanılan perovskit güneş hücresi yapıları (Fu et al., 2016)

Perovskit katmanının farklı kimyasallar kullanılarak oluşturulması yanında, her bir katman da farklı malzemeler ve teknikler kullanılarak elde edilebilmektedir. Katmanlardaki her bir değişikliğin, güç dönüşüm verimi veya kararlılık gibi parametrelere olumlu/olumsuz etkileri olmaktadır. Yapılan çalışmalarda büyük gelişmeler kaydedilmiş olsa da hala çözülmesi gereken kararlılık, ölçeklenebilirlik ve güvenilirlik problemleri, en iyi malzeme kombinasyonları ve üretim süreçleri, temel standartlar ve metrikler üzerindeki çalışmalar devam etmektedir. Ayrıca ticari uygulanabilirliğe ulaşmak için PSC'lerin üretiminin daha geniş bir alana ve daha yüksek verime ölçeklendirilmesi gerektiğini göstermektedir.

PSC'lerde ışığa duyarlı aktif katman olarak kullanılan perovskit katmanının kaplama işlemi; dönü, slot-die, doctor blade, ink-jet, sprej kaplama gibi farklı teknikler kullanılarak gerçekleştirilebilmektedir (Park and Zhu, 2020). Bu tekniklerden en yaygın olarak kullanılanı dönü kaplama yöntemidir ancak bu teknik küçük alan PSC'lerin üretimi için uygun olsa da büyük alan için uygun değildir. Üretim sürecindeki kritik noktalardan biri de hücre ölçeğinden modül ölçeğine boyut artırımının yapılması; bu ölçek büyütme işlemine uygun üretim prosesinin geliştirilmesi ve optimizasyonudur. Yapılan tez çalışmasında, Güneş Enerjisi Enstitüsü bünyesindeki Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı'nda üretimi gerçekleştirilen $5 \times 5 \text{ cm}^2$ boyutundaki PSC'lerin perovskit katmanında slot-die kaplama tekniği kullanılmıştır.

5. GEREÇ VE YÖNTEM

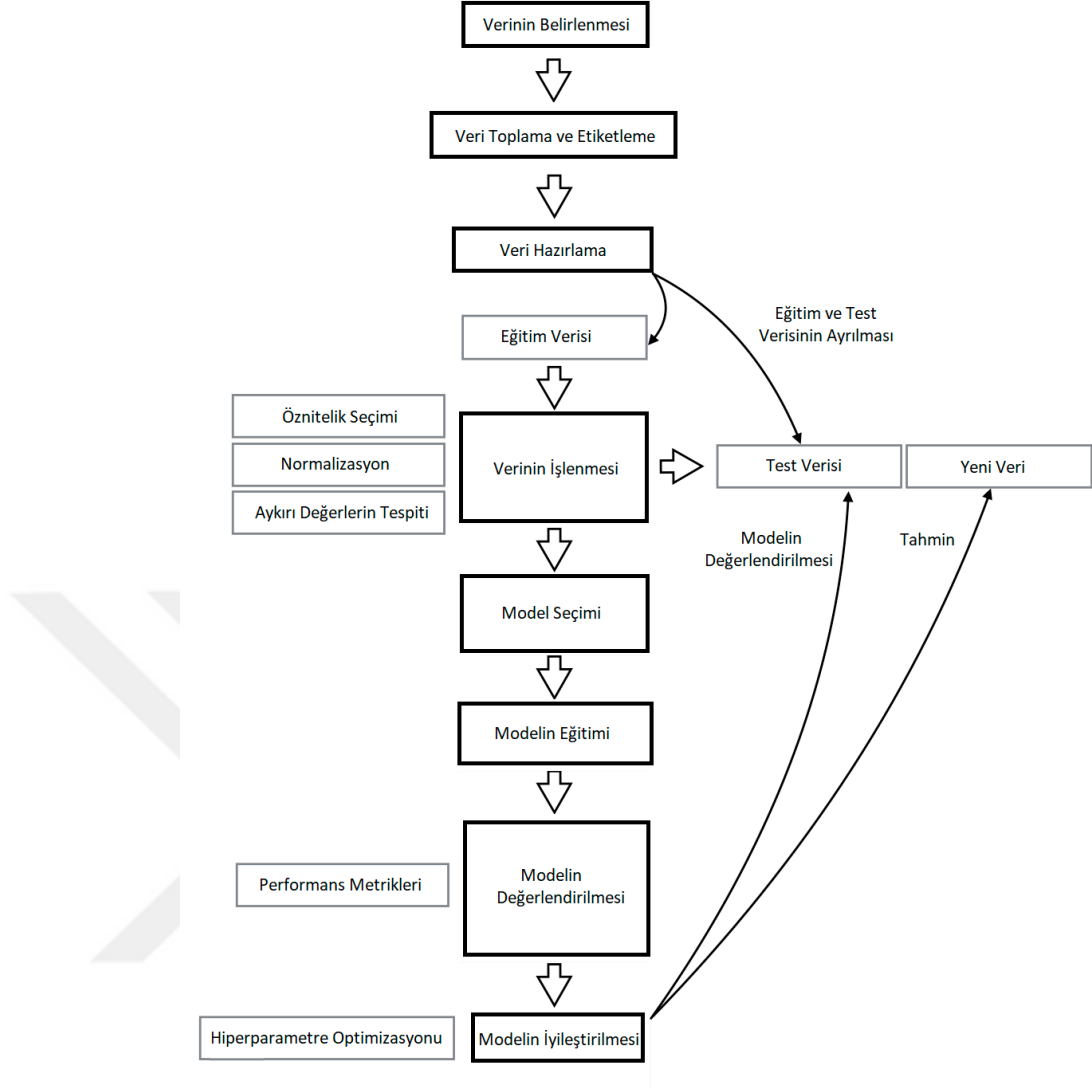
Makine öğrenmesi uygulamalarında genel olarak izlenen adımlar ve bu adımların kısa açıklamaları aşağıdaki gibi sıralamak mümkündür:

- Problemin Tanımı: İhtiyaçların net olarak belirlenmesi ve çözülmesi planlanan problemin tanımlanmasıdır.
- Performans Metriklerinin Belirlenmesi: Oluşturulacak ML modelinden nasıl tahminlerin elde edileceği ve bu modelin performansının nasıl ölçüleceğinin belirlenmesidir.
- Kullanılacak Verinin Belirlenmesi: Belirlenen problemin çözümü noktasında kullanılabilecek verinin seçilmesi aşamasıdır.
- Veri Toplama ve Etiketleme: Amaçlanan hedefe ulaşmak için mümkün olduğunca çok sayıda verinin toplanması ve bu verilerin uygun bir şekilde etiketlenmesi gerekmektedir. Herhangi bir ML uygulamasında fazlasıyla önem arz eden bu bölüm, yapılması planlanan tez çalışması için de aynı şekilde kıymetli olup çalışmanın büyük bir bölümünü oluşturmaktadır.
- Veri Hazırlama: Elde edilen verilerin hazırlanması; hedeflenen çözüme ulaşılması için kullanılacak olan veri setinin gereksiz ayrıntılardan arındırılması ve işlenebilecek şekilde düzenlenmesidir. Bu adımda verilerin normalize edilmesi ve görselleştirme de kullanılmaktadır.
- Model Seçimi: ML paradigmalarını genel olarak 3 başlıkta toplamak mümkündür. Bunlar: Denetimli öğrenme, Denetimsiz öğrenme, Pekiştirmeli öğrenmedir. Mevcut veri kümesi üzerinde çalışmaya uygun olan ve problemin çözümü için ihtiyaç duyulan modelin belirlenmesi adımıdır. Tek bir model yerine en uygun çözüm için birden fazla model üzerinde çalışıp en iyi modelin belirlenmesi yaygın olarak uygulanmaktadır.
- Modelin Eğitimi: Bu adım en fazla kaynak tüketen adımdır, çok sayıda veri üzerinde çalışılması durumunda yüksek hesaplama gücüne ihtiyaç

duyulabilmektedir. ML uygulamalarında amaçlanan genelleştirme yeteneği yüksek eğitilmiş modeller ortaya çıkarmaktır. Modelin eğitimi sonrasında öğrenilmiş parametrelerden oluşan eğitilmiş ML modeli elde edilmiş olur.

- Modelin Değerlendirilmesi: Bu adımda eğitilmiş ML modeli daha önce görmemiş olduğu veriler üzerinde test edilir. Modelin başarımlar oranının belirlenmesi için farklı performans metrikleri kullanılabilmektedir.
- Modelin İyileştirilmesi: Model performansını değerlendirdikten sonra parametrelerin optimize edilmesi için farklı ayarlamaların yapılması ve modelin başarımlar değerinin en iyi sonuca ulaşmasının sağlanmasıdır.
- Kullanıma Sunma: Modelin iyi bir şekilde çalıştığından emin olduktan sonra kullanıma sunulmasıdır.

Yapılan tez çalışmasında kullanılan ML adımlarının yer aldığı akış şeması Şekil 5.1’de gösterilmektedir.



Şekil 5.1 Kullanılan makine öğrenmesi akış şeması

Kullanılan veri seti, perovskit güneş hücresi üretim sürecindeki parametrelerle birlikte aygıt performans parametreleri yer alacak şekilde belirlenmiştir. Toplamda 36 farklı parametre değerine sahip olan 829 farklı veriden oluşan bir veri seti elde edilmiştir. Bu adımın sonrasında, ilk olarak üzerinde çalışılacak olan parametrelerin seçimi yapılarak diğer parametreler veri setinden çıkarılmıştır. Ardından daha yüksek verim değerlerine sahip yapılar üzerinde çalışmak için PCE parametresi %3'ten büyük verilerden oluşacak şekilde belirlenmiş ve daha küçük değerlerin yer aldığı verilerin veri setinden çıkarılması işlemi yapılmıştır. Sonuçta model eğitimi için üzerinde çalışılacak veri seti 10 parametre ve 662 farklı veriden oluşacak olan nihai haline ulaşmıştır. Tablo 5.1'de belirlenmiş olan parametreler ve tanımları yer almaktadır.

Tablo 5.1 ML modellerinin eğitimi için belirlenmiş olan parametreler

	Parametre	Birim	Açıklama
Slot-Die Parametreleri	Kaplama Boşluğu	μm	Başlık ve kaplama yüzeyi arasındaki yükseklik
	Kaplama Hızı	mm/s	X eksenindeki doğrusal hız
	Dağıtım Oranı	$\mu\text{l/s}$	Kaplama sıvısının akış miktarı
Air-Knife Parametreleri	Kaplama Boşluğu	μm	Başlık ve kaplama yüzeyi arasındaki yükseklik
	Kaplama Hızı	mm/s	X eksenindeki doğrusal hız
	Hava Basıncı	bar	Azot gazının basıncı
HTL Katmanı	Spiro-OMeTAD	rpm	Dönü kaplama hızı
I-V Ölçüm Parametreleri	Voc_fw	mV	Düz ölçümde elde edilen açık devre gerilimi
	FF_fw	-	Düz ölçümde elde edilen dolum faktörü
	PCE_rev	%	Ters ölçüm sonucunda elde edilen PCE değeri

Yapılan tez çalışmasında kullanılacak olan veri ML algoritmaları olarak LR, KNN, SVR, RF, XGBoost ve CatBoost seçilmiştir. Modellerin seçimi sonrasında çeşitli kütüphaneler (numpy, pandas, scikit-learn, tensorflow vb.) kullanılarak verinin hazırlanması ve modelin geliştirilmesi sağlanmaktadır. Tez çalışmasında ML uygulamalarında çok büyük oranda kullanılan Python programlama dili kullanılmaktadır. Seçilmiş olan ML modellerinin değerlendirilmesi için kullanılacak metrikler; kök ortalama kare hatası ve Pearson korelasyon katsayısı (r) olarak belirlenmiştir.

Doğrusal regresyon, veri noktaları arasındaki ilişkiyi modellemek için kullanılan temel bir ML algoritmasıdır. Bağımlı bir değişken (hedef veya yanıt değişkeni) ile bir veya daha fazla bağımsız değişken (öznitelik) arasındaki ilişkiyi modellemektedir. Doğrusal regresyonda, gözlemlenen verilere bir fonksiyon uydurarak değişkenler arasında doğrusal bir ilişki olduğunu göstermek ve henüz gözlemlenmemiş yeni veriler üzerinde tahminlerde bulunmak amaçlanmaktadır. (Nield, 2022).

Tek bir bağımsız değişkenin yer aldığı basit bir doğrusal regresyon modeli matematiksel olarak Denklem 5.1'deki gibi ifade edilmektedir.

$$y = \beta_0 + \beta_1 x + \epsilon \quad (5.1)$$

Burada; y bağımlı değişken, x bağımsız değişken, β_0 doğrunun y eksenini kestiği nokta, β_1 doğrunun eğimi ve ϵ doğrusal ilişkiyle açıklanamayan değişkenliğin ifadesi olan hata terimini göstermektedir. Birden fazla bağımsız değişkenin olduğu çoklu doğrusal regresyon denklemiyse Denklem 5.2'de gösterilmektedir.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon \quad (5.2)$$

Basit ve anlaşılır bir model olmasının yanında büyük veri setleriyle çalışırken bile hızlı ve kolay bir şekilde hesaplanabilmesi gibi avantajları nedeniyle doğrusal regresyon, veri analizi için temel bir araçtır ve genellikle daha karmaşık modellere geçmeden önce değişkenler arasındaki ilişkileri keşfetmenin ilk adımı olarak ekonomi, finans, pazarlama, sağlık gibi çok çeşitli alanda kullanılmaktadır.

K-En yakın komşu algoritması, denetimli öğrenme yöntemlerinden biri olup ML alanında hem sınıflandırma hem de regresyon problemlerinin çözümünde kullanılmaktadır. K-En yakın komşu regresyon algoritması, bağımlı değişken değerini, bağımsız değişkenlerin en yakın komşularının ortalamasına göre hesaplayıp elde edilen sonucu yaklaşık olarak tahmin eden bir yöntemdir. KNN algoritması, basitliği ve uygulama kolaylığı nedeniyle çok yönlü ve yaygın olarak kullanılan bir makine öğrenmesi algoritmasıdır. Algoritmanın kaç komşuya göre hesaplama yapacağı kullanıcı tarafından belirlenen K değerine göre

gerçekleşmektedir ve Öklid mesafesi gibi bir mesafe metriğine dayalı olarak belirli bir veri noktasına en yakın K komşuyu bularak çalışmaktadır. Veri noktaları arasındaki mesafenin ölçülmesinde en yaygın metrikler olarak Öklid, Manhattan ve Minkowski mesafe ölçümleri kullanılmaktadır (Müller and Guido, 2017).

Öklid mesafesi, iki nokta arasındaki uzaklığı ölçmek için kullanılan metriklerden biridir ve Denklem 5.3'te gösterilmektedir.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5.3)$$

Manhattan mesafesi, iki nokta arasındaki uzaklığı, bu noktaların koordinatları arasındaki farkların mutlak değerlerinin toplamı olarak hesaplamaktadır. Matematiksel olarak Denklem 5.4'teki gibi ifade edilmektedir.

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (5.4)$$

Minkowski mesafesi hem Öklid hem de Manhattan mesafesinin genelleştirilmiş hali olarak düşünülebilecek bir metriktir (bkz. Denklem 5.5).

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (5.5)$$

Formülde $p=1$ olduğunda Manhattan, $p=2$ olduğundaysa Öklid mesafesinin formülü elde edilmektedir.

Destek Vektör Regresyonu, regresyon problemleri için kullanılan bir Destek Vektör Makinesi türüdür. Tahmin edilen ve gerçek değerler arasındaki hatayı en aza indiren geleneksel regresyon modellerinin aksine, SVR hatayı epsilon marjı olarak bilinen belirli bir eşiğe sığdırmaya çalışmaktadır (Geron, 2019). Algoritma, tüm eğitim verileri için gerçek hedef değerlerden en fazla epsilon sapması olan ve

aynı zamanda mümkün olduğunca düz olan bir fonksiyon bulmayı amaçlamaktadır. Bulunmaya çalışılan doğrusal fonksiyon Denklem 5.6'da gösterilmektedir.

$$f(x) = \langle w, x \rangle + b \quad (5.6)$$

Denklemden w ağırlık vektörünü, x girdi vektörünü ve b ise yanlılık (bias) terimini ifade etmektedir. Amaç, Denklem 5.7'de gösterilen maliyet fonksiyonunu minimize etmektir.

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (5.7)$$

Burada $\|w\|^2$ fonksiyonun düzlüğünü sağlayan düzenleme terimi, C fonksiyonun düzlüğü ile ϵ değerinden daha büyük sapmaların tolere edileceği miktar arasındaki dengeyi belirleyen düzenleme parametresi, ξ_i ile ξ_i^* ise ϵ marjından sapma derecesini belirleyen gevşek (slack) değişkenler olarak ifade edilmektedir. Denklem 5.8'de gösterilmekte olan kısıt koşulları, ξ_i ile ξ_i^* tarafından kontrol edilen hata miktarlarına izin verirken tahminlerin de ϵ marjı içinde kalmasını sağlamaktadır.

$$y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i$$

$$\langle w, x_i \rangle + b - y_i \leq \epsilon + \xi_i^* \quad (5.8)$$

$$\xi_i, \xi_i^* \geq 0$$

Rastgele Orman regresyonu, bir modelin tahmin performansını ve sağlamlığını artırmak için birden fazla karar ağacını birleştiren bir topluluk öğrenmesi yöntemidir. Sayısal değerleri tahmin etmek için çok yönlü bir ML algoritması olan RF, aşırı uyumu azaltmak ve model doğruluğunu artırmak için birden fazla karar ağacının tahminlerini birleştirmektedir (Alpaydın, 2020).

Regresyon ağaçlarında, her bir düğümdeki verileri, ortaya çıkan her bir alt küme içindeki hedef değişkenin varyansı en aza indirilecek şekilde bölmek

amaçlanmaktadır. Kullanılan yaygın bölme kriteri varyanstaki azalmadır ve Denklem 5.9'da gösterilmektedir.

$$Var(t) = \frac{1}{N_t} \sum_{i \in t} (y_i - \bar{y}_t)^2 \quad (5.9)$$

Bir t düğümünde; N_t düğümdeki örnek sayısını, y_i ifadesi i için hedef değeri, $\bar{y}_t$ ise t düğümündeki ortalama hedef değeri göstermektedir.

Algoritma tüm olası bölünmeleri değerlendirir ve varyansta en büyük azalmayı sağlayan bölünmeyi seçmektedir. Bölme işlemi Denklem 5.10'da yer almaktadır.

$$\Delta Var = Var(t) - \left(\frac{N_{t_L}}{N_t} Var(t_L) + \frac{N_{t_R}}{N_t} Var(t_R) \right) \quad (5.10)$$

Burada t_L ve t_R sırasıyla sol ve sağ alt düğümleri ifade etmektedir.

RF, her biri ayrı bir karar ağacını eğitmek için kullanılan orijinal veri setinin birden fazla alt kümesini oluşturmak için torbalama (bagging) yöntemini kullanmaktadır. Oluşturulan her alt küme, orijinal veri setinde sayıyla aynı olacaktır. Bu bazı örneklerin bir alt kümede birden çok kez görünebileceği, diğerlerinin ise hiç görünmeyebileceği anlamına gelmektedir. Karar ağacındaki her bölünmede rastgele bir öznitelik alt kümesi dikkate alınmaktadır. Bu rastgelelik ağaçlar arasındaki korelasyonu azaltmaya yardımcı olarak daha iyi genelleme sağlamaktadır.

Rastgele Orman regresyonunda, belirli bir girdi için nihai tahmin, tüm bireysel ağaçların tahminlerinin ortalaması alınarak elde edilmektedir (bkz. Denklem 5.11).

$$\hat{y}(x) = \frac{1}{M} \sum_{j=1}^M \hat{y}_j(x) \quad (5.11)$$

M tane ağacın olduğu bir ormanda; $\hat{y}_j(x)$ ifadesi, x değeri için j ağacının tahminini ve $\hat{y}(x)$ ifadesi de nihai tahmin değerini göstermektedir. Bu ortalama alma işlemi, aşırı uyum ve varyansın azaltılmasına yardımcı olarak tek bir karar ağacına kıyasla genel model performansını iyileştirmekte, daha doğru ve istikrarlı tahminler yapılmasını sağlamaktadır.

XGBoost algoritması bir gradyan destekli karar ağacı modelidir ve başarısının arkasındaki en önemli faktör, her senaryoda ölçeklenebilir olmasından kaynaklanmaktadır. Sistem, tek bir makinede mevcut popüler çözümlerden on kat daha hızlı çalışmakta, dağıtık veya bellek sınırlı ortamlarda milyarlarca örneğe ölçeklenebilmektedir. XGBoost'un ölçeklenebilirliği birkaç önemli sistem ve algoritmik optimizasyona dayanmaktadır. XGBoost'un sunmuş olduğu yenilikler arasında, seyrek verileri işlemek için yeni bir karar ağacı öğrenme algoritması ve örnek ağırlıklarının ele alınmasını sağlayan teorik olarak gerekçelendirilmiş bir ağırlıklı quantile taslağı bulunmaktadır (Chen and Guestrin, 2016).

XGBoost bir dizi karar ağacından oluşan bir modeldir ve matematiksel ifadesi Denklem 5.12'de gösterilmektedir.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (5.12)$$

Her bir ağaç f_k fonksiyonu ile temsil edilmekte ve modelin çıktısı bu ağaçların çıktılarının toplamlarından oluşmaktadır. Denklemde $\mathcal{F}$ regresyon ağaçlarının uzayını ifade etmektedir.

XGBoost, regüle edilmiş bir amaç fonksiyonunu minimize etmeye çalışmaktadır (bkz. Denklem 5.13).

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k), \quad \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (5.13)$$

Denklemde l , tahmin edilen değer olan $\hat{y}_i$ ile gerçek değer y_i arasındaki farkı ölçen kayıp fonksiyonudur. Modelin karmaşıklığının ifadesi $\Omega(f)$ ile gösterilmektedir.

Eklenmiş olan regülasyon terimi, aşırı uyumu önlemek için öğrenilen ağırlıkların yumuşatılmasına yardımcı olmaktadır.

Her bir iterasyonda, yeni bir ağaç eklenerek model güncellenmektedir. Bu süreç ikinci derece Taylor serisi açılımı kullanılarak Denklem 5.14'te ifade edilmektedir.

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^n \left[l(y_i, \hat{y}^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (5.14)$$

Burada g_i ve h_i sırasıyla kayıp fonksiyonun birinci ve ikinci türevleridir. Her bir ağaç, veri setini en iyi şekilde bölecek biçimde optimize edilmektedir. Bölme işlemi, her bir özneliğin tüm olası bölme noktaları üzerinde yapılan bir tarama ile gerçekleştirilmekte ve en iyi bölme noktası, kaybı en çok azaltan nokta olacak şekilde belirlenmektedir. XGBoost bölme noktalarını hızlı bir şekilde tespit ederek büyük veri setlerinde hesaplama verimliliğini artırmaktadır. XGBoost algoritmasını farklı kılan özelliklerinden bir başkasıysa veri setlerinde yer alan eksik değerler mevcutken de çalışabilmesidir.

CatBoost, Prokhorenkova et al., tarafından 2017 yılında geliştirilen bir gradyan destekli karar ağacı modelidir ve heterojen veri kümeleri için kullanılmaktadır. CatBoost diğer GBDT algoritmaları olan XGBoost ve LightGBM ile karşılaştırıldığında, klasik algoritmaya permütasyon odaklı bir alternatif olan sıralı boosting uygulaması ve kategorik özellik işleme gibi iki önemli farka sahiptir. Diğer GBDT algoritmalarında, her aşamada elde edilen öğrenme kendinden önceki öğrenme gradyanına göre hesaplanmakta bu da yanlış gradyan tahminlerine ve aşırı uygun (overfit) modellere neden olmaktadır. CatBoost, sıralı artırma ile diğer modellerde karşılaşılan yanlış gradyanın neden olduğu tahmin kaymasını önlemekte ve gradyan tahmin aralığını azaltmaktadır (Hancock ve Khoshgoftaar, 2020).

CatBoost algoritmalarında kategorik özellikler, one-hot veya binary kodlama gibi geleneksel yöntemlerden farklı olarak hedef istatistikleri adı verilen bir kodlama yöntemiyle ele alınmaktadır.

Hedef istatistikleri yönteminin matematiksel olarak ifadesi Denklem 5.15'te gösterilmektedir.

$$\hat{x}_k^i = \frac{\sum_{j=1}^n 1\{x_j^i = x_k^i\} \cdot y_j + a \cdot p}{\sum_{j=1}^n 1\{x_j^i = x_k^i\} + a} \quad (5.15)$$

Denklemde; i kategorik özniteliklerin seviyelerini gösterirken, n toplam örnek sayısını ifade etmektedir. Öte yandan; a ağırlık parametrelerini ve p eğitim alt kümesinden hesaplanan ortalama değeri temsil etmektedir. Kategorik öznitelik olan i 'nin örneklerini x_j^i ve x_k^i ifade etmektedir. Denklemde y_j , x_k^i 'nin hedefini temsil etmektedir. x ile y arasındaki ilişki x_k^i doğrudan y_j 'ye bağlı olduğundan, koşullu kaymaya neden olup eğitim ve test alt kümesi arasında farklılıklar gösterebilmektedir. Bunun önüne geçebilmek için sanal bir zaman yaratan ve Denklem 5.16'da gösterilmekte olan sıralı hedef istatistikleri yöntemi önerilmektedir.

$$\hat{x}_k^i = \frac{\sum_{x_j \in \mathcal{D}_k} 1\{x_j^i = x_k^i\} \cdot y_j + a \cdot p}{\sum_{x_j \in \mathcal{D}_k} 1\{x_j^i = x_k^i\} + a} \quad (5.16)$$

$\mathcal{D}_k$, x_1 'den x_k 'ya kadar olan alt kümeyi, p ise $\mathcal{D}_k$ alt kümesinin ortalama değerini ifade etmektedir. Bu nedenle oluşturulan örnek dizin sırası sayesinde, yalnızca mevcut dizin sırasından önceki örnekler dikkate alınmaktadır.

CatBoost algoritmasını XGBoost ve LightGBM'den ayıran bir diğer özellik ise asimetrik ağaçlar yerine, karar tabloları olarak da adlandırılan simetrik karar ağaçlarını kullanmasıdır. Simetrik ağaçlar, ağacın tüm seviyelerinde aynı bölme kriterinin kullanıldığı anlamına gelmektedir. Bu tür ağaçlar dengelidir, aşırı uyuma daha az eğilimlidir ve test zamanında yürütmeyi önemli ölçüde hızlandırmaya izin vermektedir (Prokhorenkova et al., 2017).

Kök Ortalama Kare Hatası, bir ML modelinin tahmin ettiği değerler ile gerçek değerleri arasındaki uzaklığın bulunmasında sıklıkla kullanılan, hatanın büyüklüğünü ölçen bir metriktir.

Veri setinde yer alan n tane değer için tahmin edilen değeri $\hat{y}$ ve gerçek değeri y olarak ifade edildiğinde RMSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad RMSE = \sqrt{MSE} \quad (5.17)$$

İki değişken arasındaki doğrusal ilişkinin derecesini belirleyen Pearson korelasyon katsayısı r ile gösterilmektedir. Pearson korelasyon katsayısı -1 ile +1 arasında değer almaktadır ve işareti ilişkinin yönünü belirlemektedir. Korelasyon katsayısının 0'a yaklaşması değişkenler arasında zayıf ilişkinin varlığını göstermektedir. Veri setinde yer alan n tane değer için tahmin edilen değeri y ve gerçek değeri x olarak aldığımızda r ifadesi:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (5.18)$$

6. YAPILAN DİĞER ÇALIŞMALAR

6.1. Bant Boşluğu Değerinin Tahmin Edilmesi

Li et al., tarafından 2019 yılında yayımlanan çalışmada, literatürde yer alan ve birçok farklı makaleden elde edilmiş olan bir veri seti kullanılmaktadır. Hazırlanmış olan veri seti 333 farklı veriden ve perovskite yapısını oluşturan ABX_3 formundaki A (MA, FA, Cs), B (Pb ve Sn) ve C yapısını oluşturan (Br, Cl, I) malzemelerin oranlarıyla bant boşluğu, ΔH , ΔL , PCE, V_{oc} , J_{sc} ve FF şeklinde 15 farklı parametreden oluşmaktadır. Veri setinin oluşturulması sonrasında, hedef değişken olarak belirlenmiş olan bant boşluğu değerinin tahmin edilebilmesi için LR, KNN, SVR, RF ve ANN algoritmaları kullanılarak beş farklı ML modeli geliştirilmektedir. Modellerin başarımlarının ölçülmesinde RMSE ve Pearson korelasyon katsayısı metrikleri kullanılarak elde edilen sonuçlar paylaşılmaktadır.

ML alanında yapılacak olan çalışmalarda nitelikli verinin ne kadar önemli olduğu bilindiğinden ve tez çalışmasının başlangıç döneminde üzerinde çalışmak için hazırlanmış olan bir veri setine sahip olunmadığından, bahsedilen çalışmadaki veri seti kullanılarak belirlenmiş olan beş farklı ML modelinin geliştirilmesi ve elde edilen sonuçların makalede yer alan sonuçlarla karşılaştırılması gerçekleştirilmiştir. Şekil 6.1’de oluşturulmuş olan veri setinin ilk 5 satırı gösterilmektedir.

	ID	MA	FA	Cs	Pb	Sn	Br	Cl	I	BandGap	ΔH	ΔL	PCE	V_{oc}	J_{sc}	FF
0	1	1.00	0.00	0.0	1.0	0.0	0.00	0.0	1.00	1.50	0.15	0.20	20.70	1.11	24.00	0.78
1	2	1.00	0.00	0.0	1.0	0.0	0.00	0.0	1.00	1.50	0.30	0.00	19.40	1.13	22.60	0.75
2	3	0.15	0.85	0.0	1.0	0.0	0.15	0.0	0.85	1.72	0.40	0.18	20.59	1.13	22.99	0.79
3	4	0.17	0.83	0.0	1.0	0.0	0.40	0.0	0.60	1.72	0.40	0.18	12.60	1.12	17.10	0.66
4	5	0.15	0.85	0.0	1.0	0.0	0.15	0.0	0.85	1.72	0.40	0.18	19.20	1.03	23.70	0.79

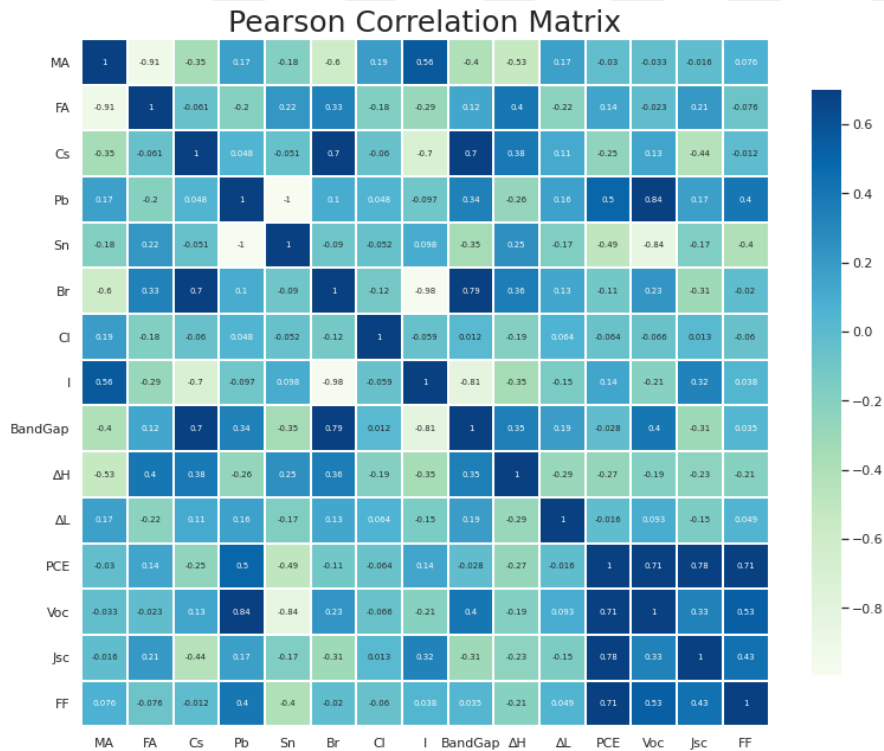
Şekil 6.1 Veri setinin ilk 5 satırı

Hazırlanmış olan veri hakkında daha fazla bilgi sahibi olabilmek adına her bir parametrenin ortalama, standart sapma, minimum, %25 (Q1: Birinci quantile), %50 (Q2: Medyan veya ikinci quantile), %75 (Q3: Üçüncü quantile) ve maksimum değerleri Şekil 6.2’de gösterilmektedir.

	ID	MA	FA	Cs	Pb	Sn	Br	Cl	I	BandGap	ΔH	ΔL	PCE	Voc	Jsc	FF
count	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000	333.000000
mean	167.000000	0.660746	0.296638	0.041730	0.966366	0.036036	0.064883	0.005796	0.930798	1.555961	0.282984	0.206505	15.688372	1.025812	20.661718	0.724970
std	96.273049	0.422600	0.395672	0.171672	0.172056	0.170864	0.132517	0.023459	0.131453	0.125436	0.188627	0.165985	4.009477	0.148837	3.636857	0.073066
min	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.170000	-0.100000	-0.640000	0.110000	0.150000	2.620000	0.278000
25%	84.000000	0.150000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.850000	1.500000	0.170000	0.120000	14.490000	1.000000	20.200000	0.700000
50%	167.000000	1.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	1.000000	1.500000	0.200000	0.200000	16.500000	1.060000	21.620000	0.740000
75%	250.000000	1.000000	0.790000	0.000000	1.000000	0.000000	0.150000	0.000000	1.000000	1.600000	0.400000	0.300000	18.310000	1.100000	22.450000	0.771000
max	333.000000	1.000000	1.000000	1.000000	1.100000	1.000000	1.000000	0.330000	1.000000	2.400000	0.950000	0.800000	22.000000	1.300000	30.560000	0.840000

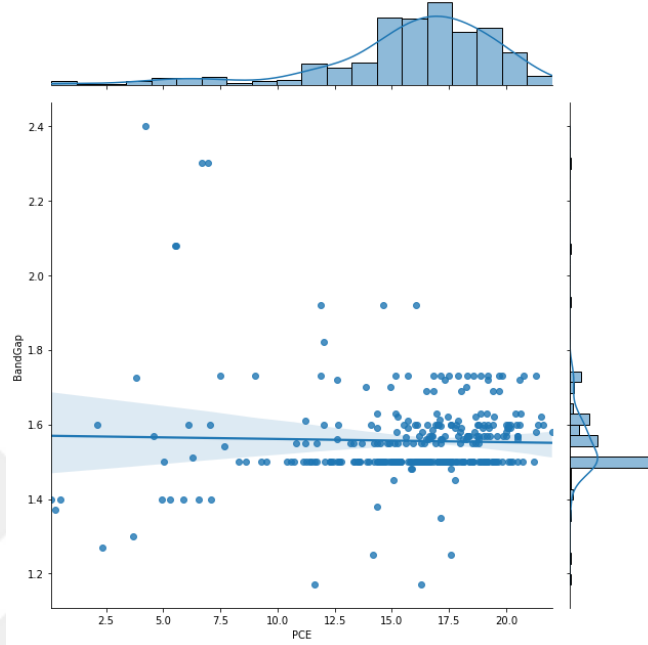
Şekil 6.2 Parametrelerin ortalama, standart sapma, minimum, Q1, medyan, Q3 ve maksimum değerleri

ML uygulamalarında verinin görselleştirilmesi, modelin oluşturulması öncesinde parametrelerin arasındaki ilişkinin ortaya konulabilmesi için uygulanan yöntemlerden biri olduğundan çalışmanın bu aşamasında veri görselleştirilmesi yapılmaktadır. Parametrelerin birbirleriyle olan ilişkileri hakkında yorum yapabilmek adına oluşturulmuş olan korelasyon matrisi Şekil 6.3'te gösterilmektedir. Korelasyon matrisinde, hedef değişken olarak belirlenmiş olan "BandGap" parametresinin I (ters yönlü) ve Br değerleriyle yüksek korelasyona sahip olduğu görülmektedir.



Şekil 6.3 Korelasyon matrisi

PCE ve BandGap parametrelerinin veri setindeki dağılımlarını gösteren histogramlar ve bu parametrelerin birbiriyle ilişkisini ortaya koyan grafikler Şekil 6.4'te yer almaktadır. Grafik ayrıca, BandGap parametresinin yoğunlaştığı değerlerin farklı PCE değerleriyle eşleştiğini göstermektedir.



Şekil 6.4 PCE ve BandGap parametreleri arasındaki ilişkinin görselleştirilmesi

Modelin eğitilmesi sırasında veri seti, %80 eğitim verisi %20 test verisi olacak şekilde ikiye ayrılmaktadır. Eğitim verisi kullanılarak geliştirilmiş modelin test verisi üzerindeki tahminleri gerçek değerlerle karşılaştırılıp başarı oranları hesaplanmaktadır. Sonuçta makalede elde edilen sonuçlarla, geliştirilmiş olan ML modellerinin performans ölçümleri karşılaştırılıp Tablo 6.1'de gösterilmektedir.

Tablo 6.1 Beş farklı ML modelinin performans ölçüm sonuçlarının karşılaştırılması

	Li et al., 2019		Geliştirilen model sonuçları	
	RMSE	Pearson	RMSE	Pearson
LR	0.115	0.88	0.078	0.925
KNN	0.104	0.98	0.048	0.948
SVR	0.076	0.97	0.039	0.963
RF	0.076	0.98	0.035	0.977
ANN	0.060	0.97	0.056	0.940

6.2. Perovskite Veri Tabanı Projesi

Jacobsson et al., tarafından 2021 yılında PSC alanında nitelikli verinin önemine vurgu yapılan çalışmada, PSC'ler üzerine yaklaşık 15.000 hakemli makale incelenmekte ve her biri için 100'e yakın parametreye sahip 42.460 aygıt verisinden oluşan bir veri seti oluşturulmaktadır. Parametreler; çalışmaların yapıldığı makalelerin verileri, hücre yapısının özellikleri (kullanılan mimari, boyut, esneklik, şeffaflık, modül özellikleri vb.), cihaz yapısı (altlık, ETL, perovskite, HTL, metal kontak vb.), kullanılan kimyasalların elde edilme yöntemleri ve performans parametreleri şeklinde gruplanmaktadır. Perovskite Database Project olarak adlandırılan çalışmanın veri setine erişim talebi olumlu sonuçlanmış ve elde edilen veri seti üzerinde çalışmalar gerçekleştirilmiştir. Bn

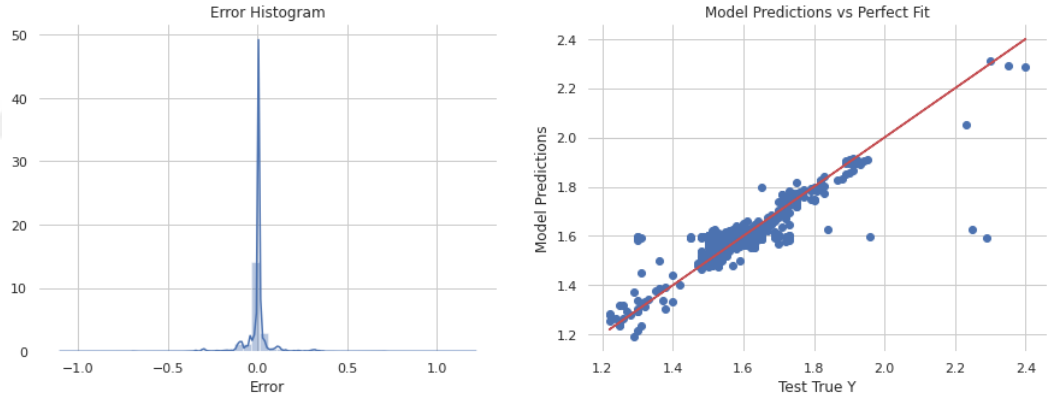
Veri setinin oluşturulması aşamasında, ilk olarak üzerinde çalışılacak olan parametrelerin seçimi yapılarak diğer parametreler veri setinden çıkarılmıştır. Ardından yüksek verim değerlerine sahip yapılar üzerinde çalışmak için PCE parametresi %10'dan büyük verilerden oluşacak şekilde belirlenmiştir. Sonrasında eksik değerlerin yer aldığı verilerin de veri setinden çıkarılması işlemi yapılmıştır. Aynı performans parametrelerine ve ABX_3 yapısına sahip tekrar eden veriler, daha tutarlı tahminler elde edebilmek adına veri setinden çıkarılmış ve nihayetinde veri seti son haline ulaşmıştır. Mevcut veri seti 19.247 farklı veriden ve perovskite yapısını oluşturan ABX_3 formundaki A (MA, FA, Cs), B (Pb ve Sn) ve X yapısını oluşturan (Br, Cl, I) malzemelerin oranlarıyla bant boşluğu, PCE, V_{oc} , J_{sc} ve FF şeklinde 14 farklı parametreden oluşmaktadır.

Veri setinde yer alan “Perovskite_band_gap” parametresinin tahmin edilebilmesi için LR, KNN, SVR, RF, ANN, XGBoost ve CatBoost ML algoritmaları kullanılmaktadır. Elde edilen modellerin performans ölçümlerinde RMSE ve Pearson korelasyon katsayısı metriklerinden yararlanılmaktadır. Modellerin eğitilmesi sırasında veri seti %80 eğitim verisi, %20 test verisi olacak şekilde ikiye ayrılmıştır. Bunun ardından MinMaxScaler dönüşümü kullanılarak normalizasyon işlemi gerçekleştirilmiştir. Elde edilen modellerin sonuçları Tablo 6.2'de gösterilmektedir.

Tablo 6.2 ML modellerinin performans ölçüm sonuçlarının karşılaştırılması

	RMSE	Pearson korelasyon katsayısı
LR	0.036	0.814
KNN	0.033	0.842
SVR	0.033	0.845
RF	0.033	0.850
ANN	0.035	0.833
XGBoost	0.030	0.859
CatBoost	0.030	0.856

Performans ölçümlerinde en başarı sonuç değerine sahip olan XGBoost modelinin, hata ve tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafikleri Şekil 6.5'te gösterilmektedir.

**Şekil 6.5** XGBoost modelinin hata ve tahmin edilen değer – gerçek değer uyum grafikleri

6.3. Empedans Spektroskopisi Sonuçlarının Tahmin Edilmesi

Parikh et al., tarafından 2023 yılında yapılan çalışmada, halojenür perovskit tek kristallerinin elektrokimyasal empedans spektroskopisi ölçüm verileri kullanılarak, düşük frekanslı tepkisini tahmin etmek için ML modeli geliştirilmektedir. EIS kullanılarak PSC'lerde iyon göçü ve taşıyıcı birikimi hakkındaki önemli bilgiler elde edilmesine karşın düşük frekanslarda uzun ölçüm

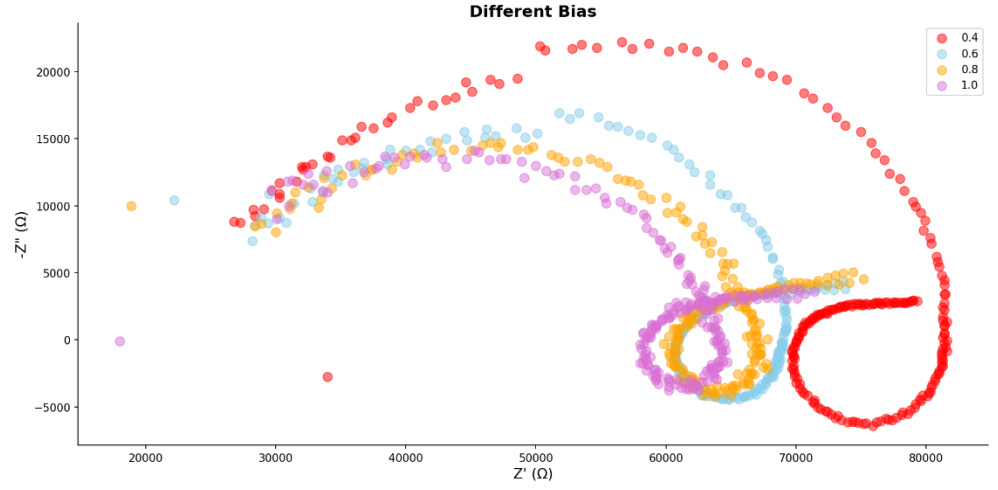
süresi ve uzun süre ışığa ve ön gerilime maruz kalma nedeniyle malzeme bozulması gözlenmektedir. Çalışmada, farklı ışınlam (88, 250, 382, 490 ve 575 W/m²) ve ön gerilim (0, 0.2, 0.4, 0.6, 0.8 ve 1V) değerlerinde 8741 veri noktasından oluşan veri seti kullanılmaktadır.

Makalenin ilave bilgiler bölümünden veri tabanı temin edilmiş ve bu veriler üzerinde EIS grafiklerini tahmin edebilmek için çalışma yapılmıştır. Veri tabanı toplam altı öznitelik verisinden oluşmaktadır. Bunlar; ön gerilim, ışınlam, frekans, akım ve EIS grafiğini oluşturan gerçek (Z') ve sanal (-Z'') değerlerdir. Hedef değişkenler olarak EIS grafiğini oluşturan gerçek ve sanal değerler belirlenmiştir. Çalışmada XGBoost ML algoritmasıyla birlikte hedef değişken birden fazla olduğu için MultiOutputRegressor algoritması kullanılmıştır. Modellerin eğitilmesi sırasında veri seti %80 eğitim verisi, %20 test verisi olacak şekilde ikiye ayrılmıştır. Model başarımının ölçülmesinde tercih edilen RMSE sonuçları, tahmin edilen iki öznitelik için Tablo 6.3'te gösterilmektedir.

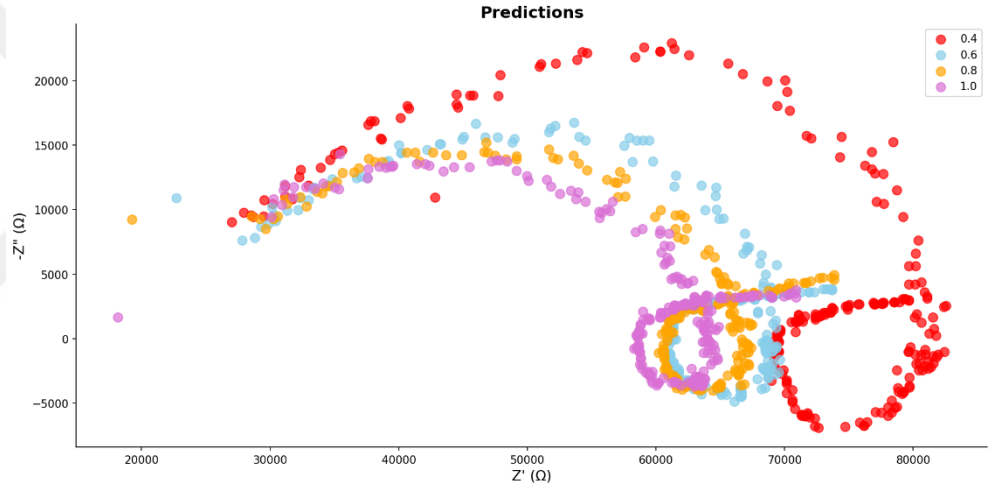
Tablo 6.3 Z' ve -Z'' değerli için elde edilen RMSE, ortalama ve hata yüzdesi sonuçları

	RMSE	Ortalama	Hata Yüzdesi
Z' (Gerçek)	11.886,19	380.856,50	% 3,12
-Z'' (Sanal)	5.265,32	51.776,48	% 10,17

Geliştirilmiş olan model çıktısına göre elde edilen 575W/m² ışınlam ve farklı ön gerilim değerlerindeki, deneysel ölçüm ve tahmin edilen empedans spektroskopisi grafikleri sırasıyla Şekil 6.6'da ve Şekil 6.7'de gösterilmektedir. Daha rahat okunabilmesi için 0.4, 0.6, 0.8 ve 1V ön gerilim değerlerinin sonuçları grafikte yer almaktadır.



Şekil 6.6 Farklı ön gerilim değerleri için deneysel EIS grafiği



Şekil 6.7 Farklı ön gerilim değerleri için tahmin edilen EIS grafiği

7. BULGULAR VE TARTIŞMA

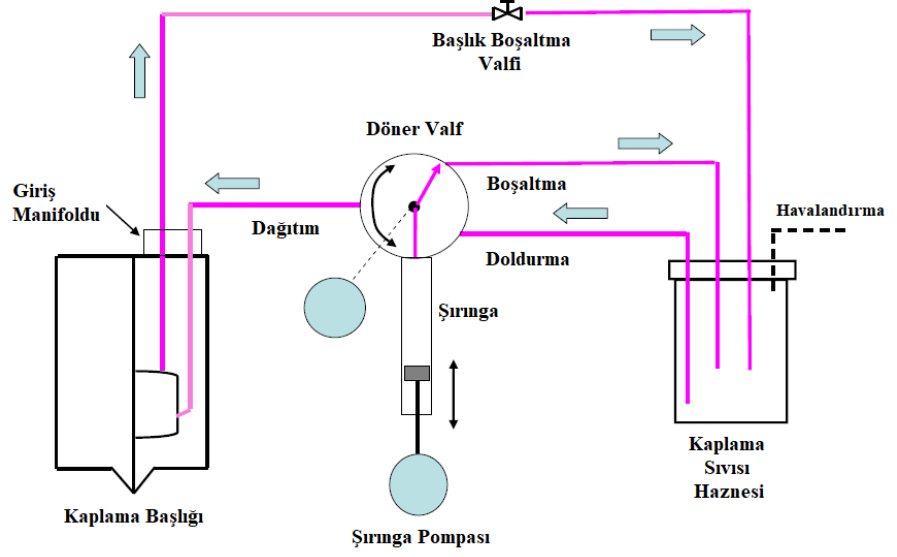
Güneş Enerjisi Enstitüsü bünyesinde faaliyet gösteren Yeni Nesil Fotovoltaik ve Enerji Depolama Teknolojileri Laboratuvarı'nda, 5x5 cm² boyutlarındaki perovskit güneş hücrelerinin üretimi başarıyla gerçekleştirilmiştir. Bu üretim süreci, hassas ve kontrollü bir ortam sağlayan temiz oda koşullarında yürütülmüştür. Üretim sırasında, büyük alan PSC üretimi için gerekli parametrelerin optimize edilmesi amacıyla slot-die kaplama tekniği ve bu tekniğe entegre edilmiş air-knife cihazı kullanılmıştır. Slot-die kaplama, perovskit tabakasının homojen bir şekilde uygulanmasını sağlarken air-knife cihazı, kaplama işlemi sırasında azot gazı aracılığıyla fazla çözgeni uzaklaştırmak için kullanılmaktadır. Bu cihazlar, üretim sürecinin verimliliğini ve ürün kalitesini artırmak için kritik öneme sahiptir. Kullanılan ekipmanlar Şekil 7.1'de görsel olarak sunulmaktadır.



Şekil 7.1 Slot-die ve air-knife cihazları

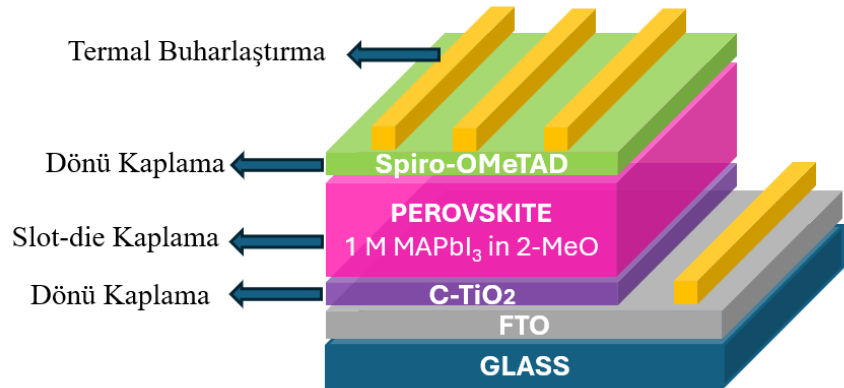
Slot-die cihazının çalışma prensibi, Şekil 7.2'de şematik olarak gösterilmiştir. Bu şemada, slot-die kaplama tekniğinin nasıl işlediği ve cihazın farklı bileşenlerinin nasıl etkileşimde bulunduğu görsel olarak sunulmaktadır. Slot-die cihazı, ince film

kaplama işlemlerinde kullanılan bir yöntemdir ve bu şematik gösterim, cihazın temel bileşenlerini ve bu bileşenlerin kaplama sürecindeki rollerini açıklamaktadır.



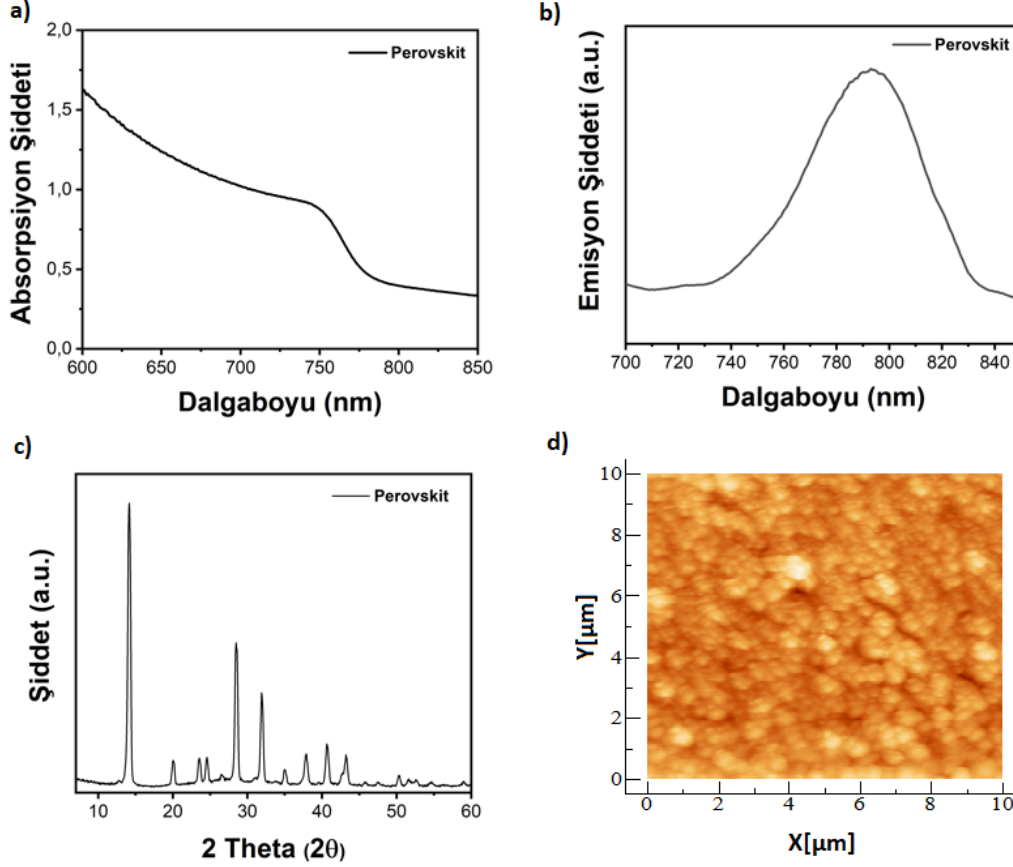
Şekil 7.2 Slot-die cihazının şematik olarak gösterimi

Üzerinde çalışılan perovskit güneş hücresi mimarisi; cam yüzey üzerine kaplanmış flor katkılı kalay oksit tabaka, elektron transfer katmanı olarak kompakt TiO_2 (titanyum dioksit), perovskite yapısı olarak 2-MeO (2-Metoksietanol) çözgeni içinde hazırlanmış 1M MAPbI_3 tabakası, boşluk transfer katmanı olarak Spiro-OMeTAD ve metal kontak olarak altından oluşmaktadır. Üretilen PSC aygıt mimarisi Şekil 7.3'te gösterilmektedir.



Şekil 7.3 PSC aygıt mimarisi

Perovskit tabakasının slot-die cihazı kullanılarak kaplanması sonrasında elde edilen filmler hakkında daha ayrıntılı bilgi sahibi olabilmek adına farklı karakterizasyon çalışmaları gerçekleştirilmiştir. Yapılan karakterizasyon işlemlerinden elde edilen sonuçlar Şekil 7.4'te gösterilmektedir.

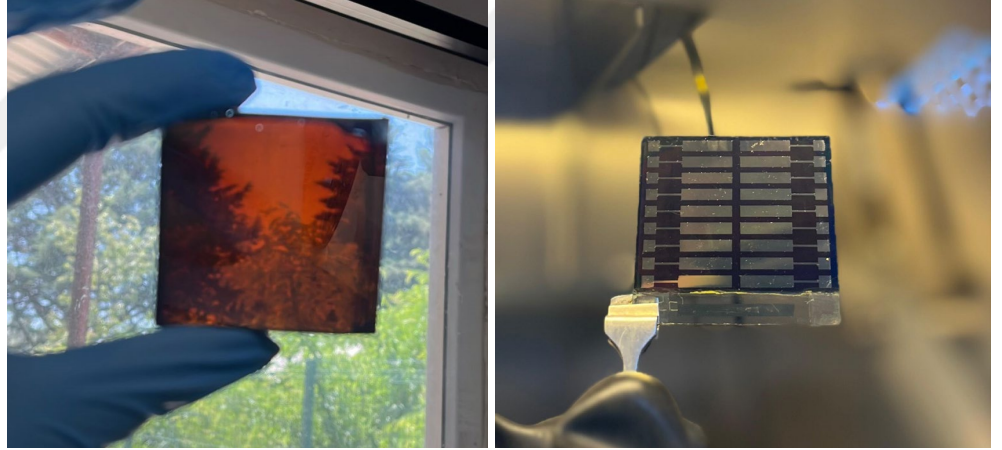


Şekil 7.4 Karakterizasyon çalışması sonuçları a) Absorpsiyon grafiği, b) Emisyon grafiği c) XRD grafiği ve d) AFM görüntüsü

Hazırlanan filmlerin ışığı absorplama yetenekleri ve bant boşluğu değerlerinin belirlenmesi amacıyla absorpsiyon ölçümleri gerçekleştirilmiştir. Şekil 7.4 (a)'da gösterildiği gibi 750 nm civarında görülen omuz piki, perovskit siyah fazın karakteristik piki olup film oluşumunun başarılı bir şekilde gerçekleştiğini göstermektedir. Absorpsiyon spektrumu üzerinden hesaplanan bant boşluğu değeri 1,59eV olarak bulunmuştur. Şekil 7.4 (b)'de perovskit ince filmler üzerinden alınan emisyon grafiği gösterilmiştir. 480nm dalga boyunda uyarılarak elde edilen emisyon dalga boyu, literatürdeki perovskit için belirtilen dalga boyu ile uyumluluk göstermektedir (Do-Kyoung et al., 2019). Şekil 7.4 (c)'de perovskit kaplı ince

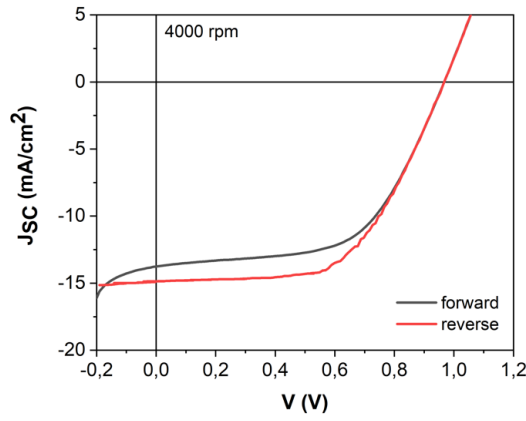
filmlerin XRD grafikleri yer almaktadır. Perovskit siyah faza ait karakteristik pik olan 14.1’de keskin ve yüksek şiddetli bir pik elde edilmiştir. Ayrıca 12.6’da kurşun piki ve 11.6’da sarı faza ait olan pik gözlemlenmemiştir. Bu sonuç literatürle uyumlu olup yapıda fazla kurşun bulunmadığını, faz dönüşümünün tamamlandığını ve kristallliği yüksek bir film elde edildiğini göstermektedir. Şekil 7.4 (d)’de perovskit filmlerdeki morfoloji farklılıklarını gözlemlemek amacıyla AFM ölçümleri gerçekleştirilmiştir. Yapılan kaplamada homojen ve küçük taneciklere sahip bir morfoloji oluşumu görülmektedir. Pürüzlülük RMS değerlerine incelendiğinde filmin RMS değerinin 65.91 nm olduğu belirlenmiştir.

Toplamda 5x5cm² büyüklüğünde 52 farklı aygıt elde edilmiş olup elde edilen aygıtların her bir kontağının performans ölçümleri (I-V) yapılmıştır. Slot-die cihazıyla kaplanmış film ve tamamlanmış aygıtlardan birinin görüntüsü Şekil 7.5’te gösterilmektedir.



Şekil 7.5 Slot-die cihazıyla kaplanmış film ve tamamlanmış aygıt görüntüsü

Herhangi bir optimizasyon çalışması yapılmamışken üretim süreci sonunda elde edilen aygıtlardan en yüksek verim değerine sahip olanın, düz ve ters performans ölçümleriyle I-V grafikleri Şekil 7.6’da gösterilmiştir.

Aktif alan: 0,39 cm²

Film	Jsc (mA/cm ²)	Voc (mV)	FF	PCE (%)
4000 rpm – F	13,75	970	0,57	7,65
4000 rpm – R	14,86	973	0,57	8,30

Şekil 7.6 En yüksek verim değerine sahip olan aygıtın performans ölçümleriyle I-V grafikleri

PSC'lerin üretim sürecindeki parametrelerle aygıt performans parametrelerinin yer aldığı veri seti oluşturulmuştur. Veri setinin bir bölümünün görüntüsü Şekil 7.7'de gösterilmektedir.

R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
AK_CG(µm)	AK_V(mm/s)	AK_Pressure	AK_Heat	Spiro(rpm)	Jsc_fw	Voc_fw	FF_fw	PCE_fw	Jsc_rev	Voc_rev	FF_rev	PCE_rev	Device_N	Active_Area	Device_N	Di	Temperature	Humidity
280	3	2,5	0	4000	14,47	980	0,50	7,15	14,6	934	0,48	6,56	f7	0,39	2-2	9 May 24	21,5	49
280	3	2,5	0	4000	14,13	970	0,55	7,47	14,84	973	0,57	8,28	f7	0,39	3-1	9 May 24	21,5	49
280	3	2,5	0	4000	13,44	970	0,57	7,46	14,73	973	0,58	8,29	f7	0,39	3-2	9 May 24	21,5	49
280	3	2,5	0	4000	13,36	970	0,58	7,47	14,73	973	0,58	8,24	f7	0,39	3-3	9 May 24	21,5	49
280	3	2,5	0	4000	13,30	970	0,58	7,49	14,88	973	0,57	8,26	f7	0,39	3-4	9 May 24	21,5	49
280	3	2,5	0	4000	13,40	970	0,58	7,52	14,71	973	0,57	8,2	f7	0,39	3-5	9 May 24	21,5	49
280	3	2,5	0	4000	13,35	970	0,58	7,52	14,72	963	0,58	8,15	f7	0,39	3-6	9 May 24	21,5	49
280	3	2,5	0	4000	13,33	960	0,59	7,53	14,77	973	0,57	8,19	f7	0,39	3-7	9 May 24	21,5	49
280	3	2,5	0	4000	13,59	960	0,58	7,52	14,72	963	0,58	8,22	f7	0,39	3-8	9 May 24	21,5	49
280	3	2,5	0	4000	13,42	960	0,58	7,53	14,76	963	0,57	8,16	f7	0,39	3-9	9 May 24	21,5	49
280	3	2,5	0	4000	13,44	970	0,58	7,59	14,81	973	0,57	8,25	f7	0,39	3-10	9 May 24	21,5	49
280	3	2,5	0	4000	13,45	970	0,58	7,59	14,88	973	0,58	8,33	f7	0,39	3-11	9 May 24	21,5	49
280	3	2,5	0	4000	13,68	970	0,58	7,64	14,81	973	0,57	8,28	f7	0,39	3-12	9 May 24	21,5	49
280	3	2,5	0	4000	13,53	970	0,58	7,63	14,9	973	0,57	8,28	f7	0,39	3-13	9 May 24	21,5	49
280	3	2,5	0	4000	13,58	970	0,58	7,63	14,83	973	0,57	8,28	f7	0,39	3-14	9 May 24	21,5	49
280	3	2,5	0	4000	13,75	970	0,57	7,65	14,86	973	0,57	8,3	f7	0,39	3-15	9 May 24	21,5	49
280	3	2,5	0	4000	13,99	930	0,49	6,41	14,23	944	0,53	7,15	f7	0,39	4-1	9 May 24	21,5	49
280	3	2,5	0	4000	13,01	940	0,53	6,48	14,11	954	0,54	7,2	f7	0,39	4-2	9 May 24	21,5	49
280	3	2,5	0	4000	12,88	940	0,54	6,54	14,01	954	0,54	7,28	f7	0,39	4-3	9 May 24	21,5	49
280	3	2,5	0	4000	12,75	950	0,55	6,62	13,98	954	0,55	7,35	f7	0,39	4-4	9 May 24	21,5	49
280	3	2,5	0	4000	14,43	930	0,50	6,72	14,71	944	0,53	7,42	f7	0,39	5-1	9 May 24	21,5	49
280	3	2,5	0	4000	13,51	940	0,54	6,81	14,43	944	0,55	7,45	f7	0,39	5-2	9 May 24	21,5	49
280	3	2,5	0	4000	13,32	940	0,55	6,93	14,34	954	0,55	7,56	f7	0,39	5-3	9 May 24	21,5	49
280	3	2,5	0	4000	13,16	950	0,56	6,96	14,38	954	0,55	7,6	f7	0,39	5-4	9 May 24	21,5	49
280	3	2,5	0	4000	13,17	950	0,56	7,00	14,88	963	0,55	7,87	f7	0,39	5-6	9 May 24	21,5	49
280	3	2,5	0	4000	13,71	960	0,54	7,16	14,66	963	0,56	7,88	f7	0,39	5-7	9 May 24	21,5	49
280	3	2,5	0	4000	13,44	960	0,55	7,14	14,61	963	0,56	7,88	f7	0,39	5-8	9 May 24	21,5	49
280	3	2,5	0	4000	14,47	910	0,45	5,87	15,21	904	0,49	6,8	f7	0,39	6-1	9 May 24	21,5	49
280	3	2,5	0	4000	14,39	890	0,47	6,01	15,01	894	0,5	6,76	f7	0,39	6-2	9 May 24	21,5	49
280	3	2,5	0	4000	14,08	900	0,48	6,08	14,65	894	0,5	6,5	f7	0,39	6-3	9 May 24	21,5	49

Şekil 7.7 Oluşturulmuş olan veri setinin bir bölümünün görüntüsü

Toplamda 829 farklı veriden ve 36 parametreden oluşan bir veri seti elde edilmiştir. Veri seti oluşturulurken belirlenmiş olan parametreler Şekil 7.8'de gösterilmektedir.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 829 entries, 0 to 828
Data columns (total 37 columns):
#   Column                Non-Null Count  Dtype
---  -
0   index                 829 non-null    int64
1   PP_LEP(mm)           829 non-null    int64
2   PP_DSP(mm)           829 non-null    int64
3   PP_EoPP              829 non-null    int64
4   PP_PG(μm)            829 non-null    int64
5   PP_SV(mm/s)          829 non-null    int64
6   PD_DR(μl/s)          829 non-null    int64
7   PD_DD(ms)            829 non-null    int64
8   CP_LEP(mm)           829 non-null    int64
9   CP_DSO(mm)           829 non-null    int64
10  CP_EoCP(mm)          829 non-null    int64
11  CP_CG(μm)            829 non-null    int64
12  CP_SV(mm/s)          829 non-null    float64
13  CD_DR(μl/s)          829 non-null    float64
14  CD_DD(ms)            829 non-null    int64
15  CD_RR(μl/s)          829 non-null    int64
16  CD_PRD(ms)           829 non-null    int64
17  AK_CG(μm)            829 non-null    int64
18  AK_SV(mm/s)          829 non-null    float64
19  AK_Pressure          829 non-null    float64
20  AK_Heat              829 non-null    int64
21  Spiro(rpm)           829 non-null    int64
22  Jsc_fw               829 non-null    float64
23  Voc_fw               829 non-null    int64
24  FF_fw               829 non-null    float64
25  PCE_fw               829 non-null    float64
26  Jsc_rev              829 non-null    float64
27  Voc_rev              829 non-null    int64
28  FF_rev               829 non-null    float64
29  PCE_rev              829 non-null    float64
30  Device_id            829 non-null    object
31  Active_Area          829 non-null    float64
32  Device_Num           829 non-null    object
33  Date                 829 non-null    object
34  Temperature          818 non-null    float64
35  Humidity              818 non-null    float64
36  Notes                 0 non-null      float64
dtypes: float64(14), int64(20), object(3)
memory usage: 239.8+ KB

```

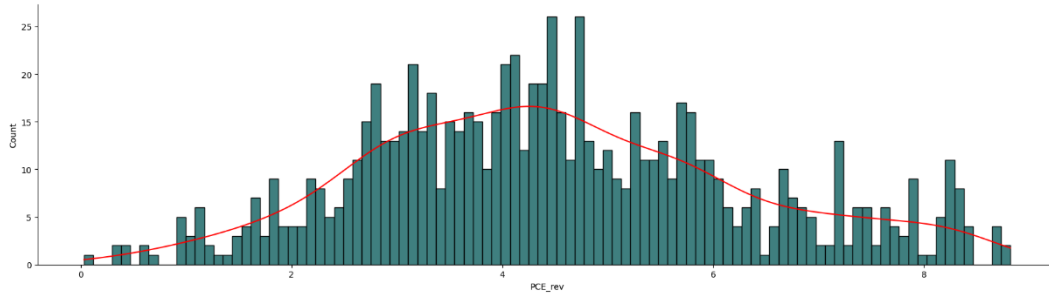
Şekil 7.8 Belirlenmiş olan parametreler

Veri seti daha sonradan gerçekleştirilmesi planlanan farklı çalışmalar da düşünülerek olabildiğince fazla parametre yer alacak şekilde oluşturulmuştur. Tez çalışmasında üzerinde çalışılacak parametreler olarak 10 tanesi seçilmiş ve seçilen parametreler Şekil 7.9’da gösterilmektedir.

	count	mean	std	min	25%	50%	75%	max
CP_CG(μ m)	829.0	224.921592	13.147248	200.00	220.00	220.00	220.00	260.00
CP_SV(mm/s)	829.0	3.835947	0.392700	3.00	4.00	4.00	4.00	4.50
CD_DR(μ l/s)	829.0	5.968637	0.663202	3.50	6.00	6.00	6.00	8.00
AK_CG(μ m)	829.0	278.238842	8.669224	250.00	280.00	280.00	280.00	300.00
AK_SV(mm/s)	829.0	3.232811	0.393282	3.00	3.00	3.00	3.50	4.00
AK_Pressure	829.0	2.506273	0.174412	1.50	2.50	2.60	2.60	2.70
Spiro(rpm)	829.0	3925.211098	321.084906	3000.00	4000.00	4000.00	4000.00	4500.00
Voc_fw	829.0	794.921592	98.775840	260.00	740.00	810.00	850.00	980.00
FF_fw	829.0	0.396526	0.069151	0.13	0.36	0.40	0.44	0.59
PCE_rev	829.0	4.506309	1.783760	0.03	3.20	4.38	5.68	8.82

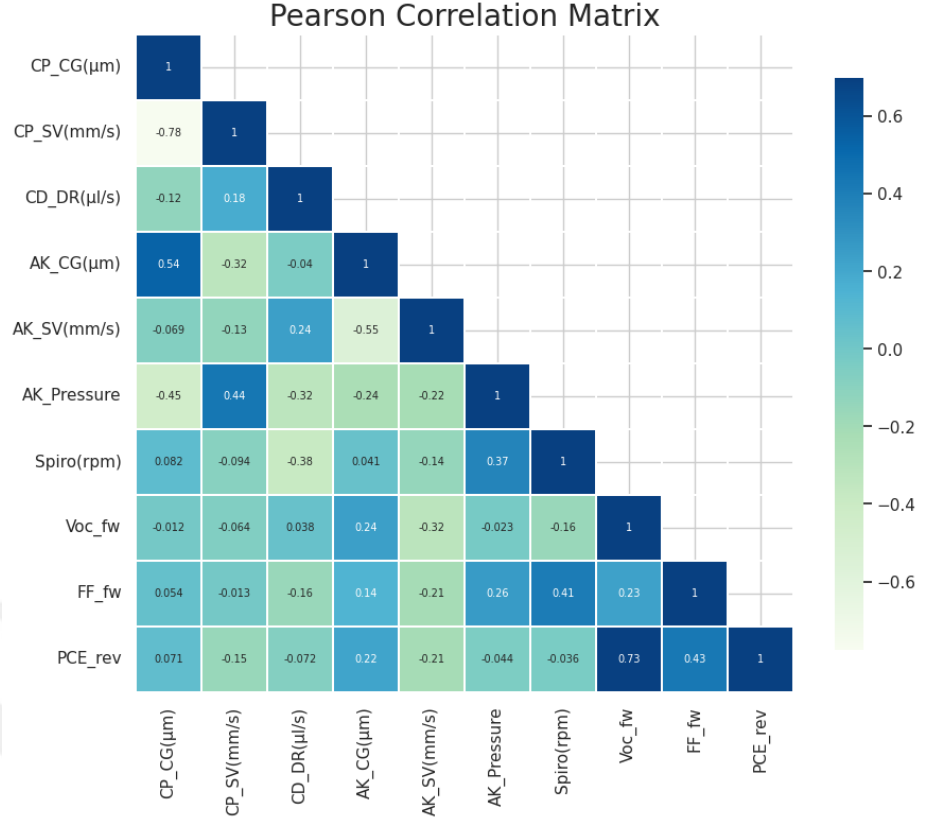
Şekil 7.9 ML modellerini eğitmek için seçilmiş olan parametreler

Veri setinde yer alan ve hedef değişken olarak belirlenmiş “PCE_rev” parametresinde yer alan değerlerin dağılımından oluşan grafik Şekil 7.10’da gösterilmektedir.



Şekil 7.10 PCE_rev parametresinde yer alan değerlerin dağılım grafiği

Veri setinde yer alan parametrelerin birbirleriyle ilişkisinin yer aldığı Pearson korelasyon matrisi Şekil 7.11’de gösterilmektedir.



Şekil 7.11 Korelasyon matrisi

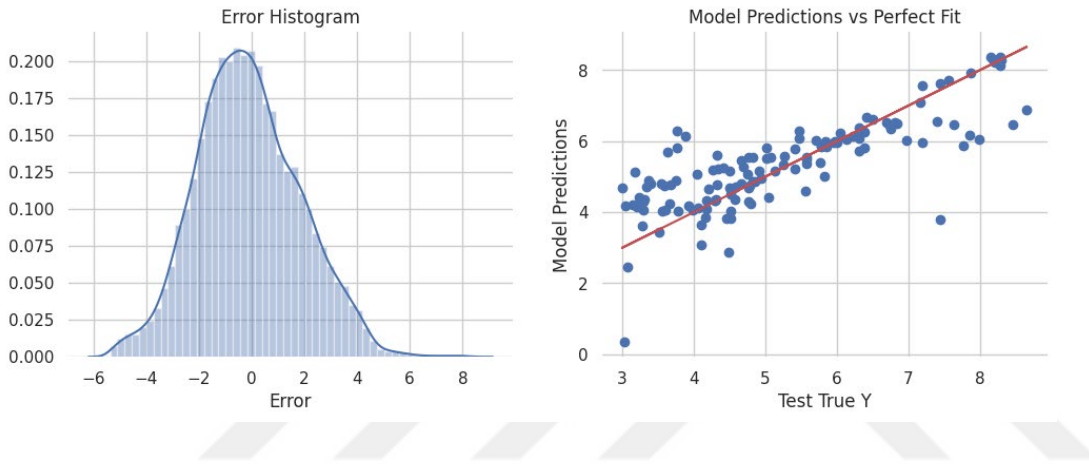
Veri setinde yer alan “PCE_rev” parametresi için 3’ten büyük olacak şekilde bir sınır değeri belirlenmiştir. Belirlenen değer sonrasında, veri setinde bu koşulu sağlayan 662 veri bulunmaktadır. Modellerin eğitilmesi sırasında veri seti %80 eğitim verisi, %20 test verisi olacak şekilde ikiye ayrılmıştır. Bunun ardından MinMaxScaler dönüşümü kullanılarak normalizasyon işlemi gerçekleştirilmiştir. Hedef değişken olarak belirlenen “PCE_rev” parametresinin tahmin edilebilmesi için LR, KNN, SVR, RF, XGBoost ve CatBoost makine öğrenmesi algoritmaları kullanılmıştır.

Kullanılan yöntemler sonucunda elde edilen modellerin performans ölçümlerinde ise RMSE ve Pearson korelasyon katsayısı metriklerinden yararlanılmıştır.

Geliştirilmiş olan LR modelinin performans ölçümleri Tablo 7.1’de, aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.12’de gösterilmiştir.

Tablo 7.1 LR modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.9092
Pearson Korelasyon Katsayısı	0.7981

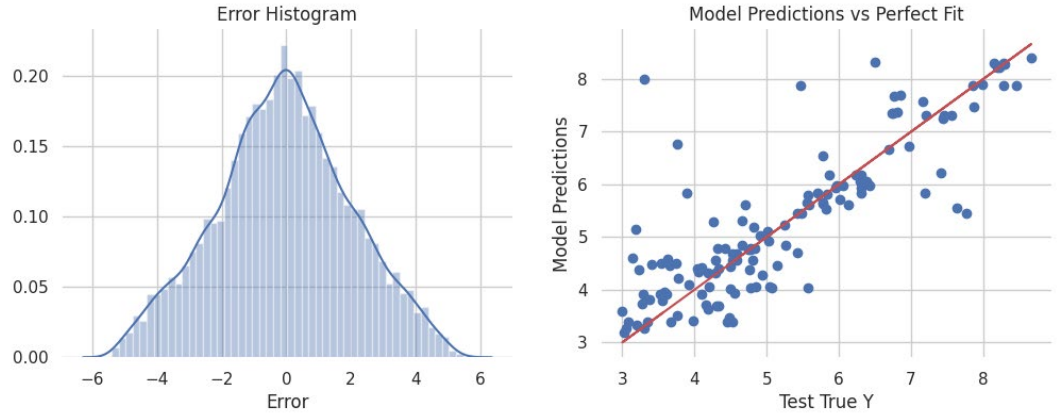


Şekil 7.12 LR modelinin tahmin edilen değer – gerçek değer uyum grafikleri

Geliştirilmiş olan KNN modelinin performans ölçümleri Tablo 7.2’de, aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.13’te gösterilmiştir.

Tablo 7.2 KNN modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.8353
Pearson Korelasyon Katsayısı	0.8407



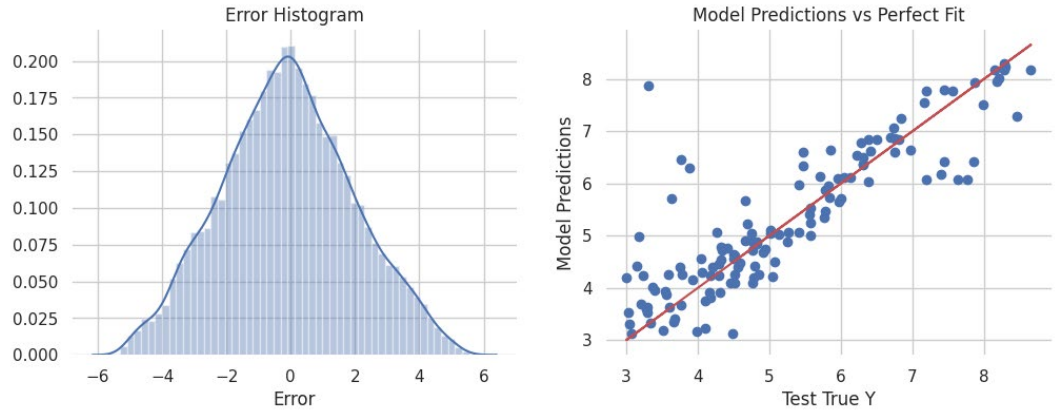
Şekil 7.13 KNN modelinin tahmin edilen değer – gerçek değer uyum grafikleri

KNN modelinin hiper parametre seçiminde `{'n_neighbors': np.arange(2, 150)}` kullanılmış olup en iyi değeri 2 olarak bulunmuştur.

Geliştirilmiş olan SVR modelinin performans ölçümleri Tablo 7.3'te, aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.14'te gösterilmiştir.

Tablo 7.3 SVR modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.7665
Pearson Korelasyon Katsayısı	0.8632



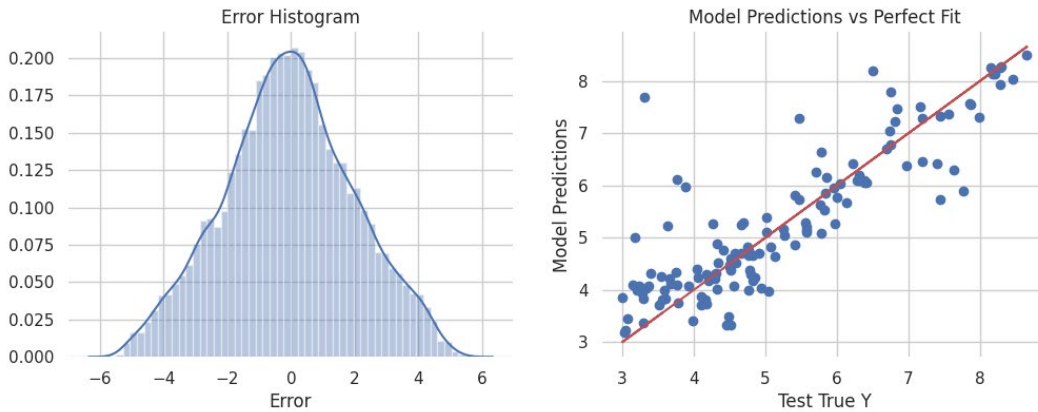
Şekil 7.14 SVR modelinin tahmin edilen değer – gerçek değer uyum grafikleri

SVR modelinin hiper parametre seçiminde {'kernel': ['linear', 'poly', 'rbf', 'sigmoid'], 'degree': [3, 5, 7, 9], 'C': [0.1, 0.2, 0.3, 1], 'epsilon': [0.01, 0.1, 1, 10]} kullanılmış olup en iyi değerler {kernel: poly, 'degree': 5, 'C': 0.1, 'epsilon': 0.1} olarak bulunmuştur.

Geliştirilmiş olan RF modelinin performans ölçümleri Tablo 7.4'te, aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.15'te gösterilmiştir.

Tablo 7.4 RF modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.7642
Pearson Korelasyon Katsayısı	0.8631



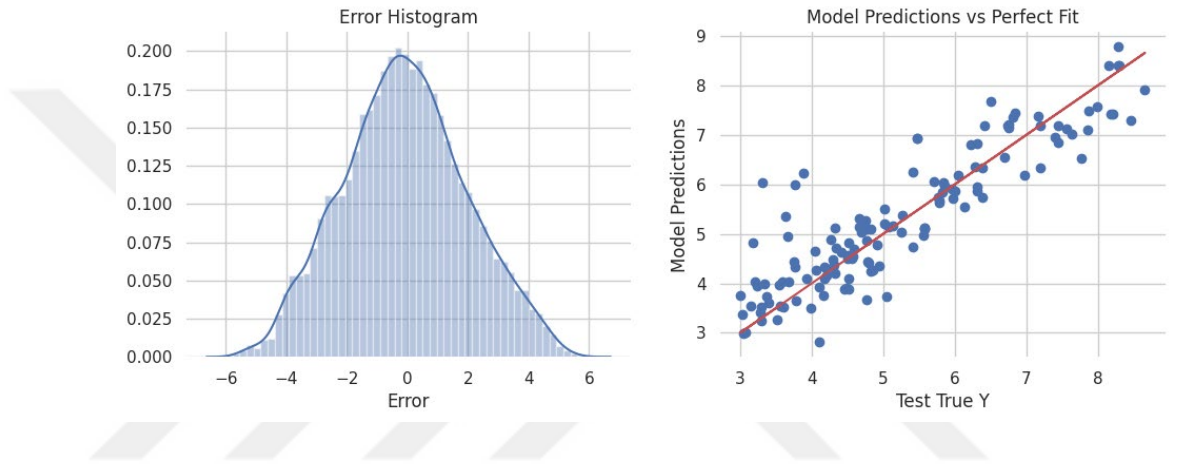
Şekil 7.15 RF modelinin tahmin edilen değer – gerçek değer uyum grafikleri

RF modelinin hiper parametre seçiminde {'n_estimators': [100, 500, 1000], 'max_features': ['sqrt'], 'max_depth': [10, 60, 110], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4], 'bootstrap': [True]} kullanılmış olup en iyi değerler {'bootstrap': True, 'max_features': ['sqrt'], 'max_depth': 110, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 1000} olarak bulunmuştur.

Geliştirilmiş olan XGBoost modelinin performans ölçümleri Tablo 7.5'te, aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.16'da gösterilmiştir.

Tablo 7.5 XGBoost modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.6737
Pearson Korelasyon Katsayısı	0.8952



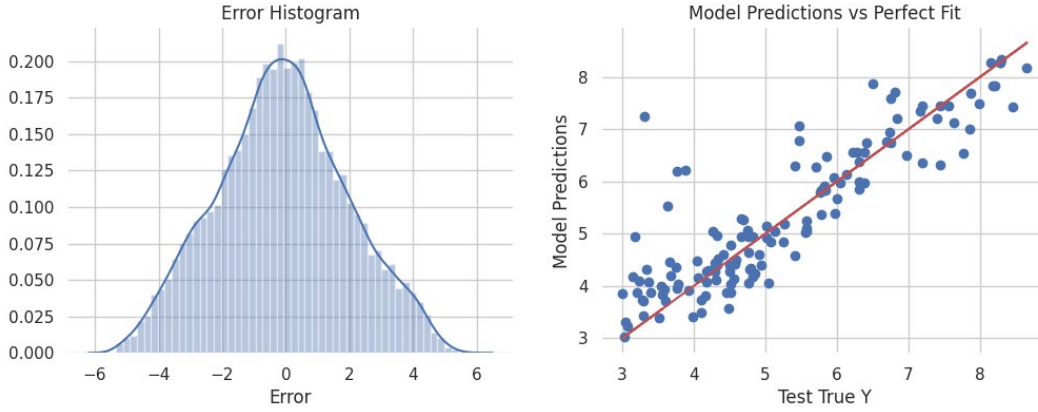
Şekil 7.16 XGBoost modelinin tahmin edilen değer – gerçek değer uyum grafikleri

XGBoost modelinin hiper parametre seçiminde optimizasyon için optuna kullanılmış olup en iyi değerler {random_state = 23, tree_method = "hist", device = "cuda", n_estimators = 5000, learning_rate = 0.2941104226302799, reg_lambda = 0.09501163254567294, reg_alpha = 0.21071622085976102, subsample = 0.531997365758499, colsample_bytree = 0.5513509843282024 ve max_depth = 1} olarak bulunmuştur.

Geliştirilmiş olan CatBoost modelinin performans ölçümleri Tablo 7.6'da aynı modelin tahmin edilen değerlerle gerçek değerlerin uyumlarının karşılaştırılmış olduğu grafiği ise Şekil 7.17'de gösterilmiştir.

Tablo 7.6 CatBoost modelinin performans ölçüm sonuçları

Kök Ortalama Kare Hatası	0.7172
Pearson Korelasyon Katsayısı	0.8818

**Şekil 7.17** CatBoost modelinin tahmin edilen değer – gerçek değer uyum grafikleri

CatBoost modelinin hiper parametre seçiminde optimizasyon için optuna kullanılmış olup en iyi değerler $\{\text{loss_function} = \text{"RMSE"}, \text{random_state} = 0, \text{task_type} = \text{"GPU"}, \text{n_estimators} = 5000, \text{learning_rate} = 0.024639898961148685, \text{max_depth} = 3, \text{min_data_in_leaf} = 54 \text{ ve } \text{l2_leaf_reg} = 1.0\}$ olarak bulunmuştur.

Modellerin eğitim aşamasının ardından elde edilen en başarılı model üzerinde, slot-die ve air-knife cihazlarındaki belirlenmiş olan farklı parametre değerleri için güç dönüşüm verimi değerini tahmin etme çalışması gerçekleştirilmiştir. Tahmin sonucundaki en yüksek verim değerinin elde edilebileceği yapı Tablo 7.7’de gösterilmektedir.

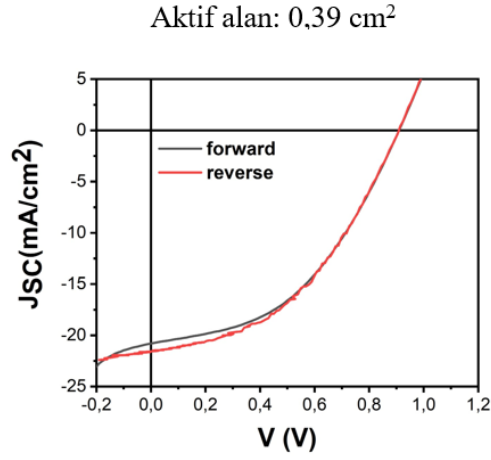
Tablo 7.7 ML modeli kullanılarak elde edilen en yüksek verim değerin tahmin sonuçları

Parametreler	CP_CG (μm)	240
	CP_SV (mm/s)	3
	CD_DR ($\mu\text{l/s}$)	5
	AK_CG (μm)	280
	AK_SV (mm/s)	3
	AK_Pressure	2,5
	Spiro (rpm)	4000
Tahmin Edilen	PCE (%)	9,01

Tahmin sonucunda elde edilmiş olan sonuçlar, veri setinin oluşturulma aşamasındaki PSC üretim çalışmalarında daha önceden kullanılmamış ve veri setinde yer almayan parametrelerdir. Bu sonucun ardından karşılaştırma yapılabilmesi adına elde edilmiş olan parametreler kullanılarak PSC aygıt üretimi gerçekleştirilmiştir. Kullanılan parametreler sonucunda üretimi gerçekleştirilmiş olan aygıt Şekil 7.18’de gösterilmektedir.

**Şekil 7.18** Tahmin sonuçlarına göre üretimi gerçekleştirilen PSC

Üretim prosesi sonunda performans ölçümleri elde edilen sonuçlar Şekil 7.19'da gösterilmektedir.



Film	J _{sc} (mA/cm ²)	V _{oc} (mV)	FF	PCE (%)
4000 rpm – F	20,81	910	0,45	8,52
4000 rpm – R	21,52	914	0,44	8,71

Şekil 7.19 Tahmin sonuçlarına göre üretilmiş olan aygıtın performans ölçümleriyle I-V grafikleri

Tahmin edilen parametrelerin etkinliğini değerlendirmek ve karşılaştırma yapabilmek amacıyla, bu parametreler kullanılarak yeni bir PSC aygıtı üretilmiştir. Üretim süreci, tahmin edilen parametrelerin fiziksel olarak uygulanabilirliğini ve performansını test etmeyi amaçlamıştır. Sonuç olarak, tahmin edilen parametreler kullanılarak üretimi gerçekleştirilen perovskit güneş hücresinde, düz ve ters tarama sonucunda elde edilen PCE değerlerinin her ikisinde de önceki deneysel sonuçlarla karşılaştırıldığında artış elde edilmiştir.

8. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, Ege Üniversitesi Güneş Enerjisi Enstitüsü bünyesinde yürütülen araştırmaların sonuçları kullanılarak kapsamlı bir veri seti oluşturulmuştur. Bu veri seti, perovskit güneş hücrelerinin üretim sürecinde kullanılan çeşitli parametreler ile aygıt performans parametrelerini içerecek şekilde tasarlanmıştır. Veri seti, toplamda 36 farklı parametreye sahip 829 veri noktasından oluşmaktadır. Bu parametreler, malzeme bileşenlerinden üretim koşullarına, cihaz yapısından performans ölçümlerine kadar geniş bir yelpazeyi kapsamaktadır.

Veri setinin oluşturulmasındaki temel amaç, perovskit güneş hücrelerinin büyük ölçekli üretiminde en yüksek verimin elde edilebilmesi için üretim yapısının optimizasyonunda makine öğrenmesi tekniklerinin kullanılmasını sağlamaktır. Bu bağlamda, çeşitli ML algoritmaları değerlendirilmiştir. Seçilen algoritmalar arasında Doğrusal Regresyon, K-En Yakın Komşu, Destek Vektör Makinesi, Rastgele Orman, XGBoost ve CatBoost yer almaktadır. Bu algoritmalar, farklı veri yapıları ve problem türleri için uygunlukları nedeniyle tercih edilmiştir. ML modellerinin performansını değerlendirmek için kök ortalama kare hatası ve Pearson korelasyon katsayısı gibi metrikler kullanılmıştır. RMSE, modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkın büyüklüğünü ölçerken, Pearson korelasyon katsayısı, tahmin edilen ve gerçek değerler arasındaki doğrusal ilişkiyi değerlendirmektedir. Bu metrikler, modellerin doğruluğunu ve güvenilirliğini belirlemek için kritik öneme sahiptir.

Performans ölçümlerinde, XGBoost algoritması kullanılarak eğitilmiş olan model, RMSE: 0,6737 ve Pearson korelasyon katsayısı: 0.8952 değerleriyle en başarılı sonucu vermiştir. Bu sonuçlar, modelin tahmin doğruluğunun yüksek olduğunu ve verilerle güçlü bir doğrusal ilişki kurduğunu göstermektedir. XGBoost'un bu başarısı, algoritmanın karmaşık veri setleri üzerinde etkili bir şekilde çalışabilme yeteneğinden kaynaklanmaktadır.

XGBoost modeli, slot-die ve air-knife cihazlarındaki belirlenmiş olan farklı parametre değerleri için güç dönüşüm verimini tahmin etmek amacıyla kullanılmıştır. Slot-die kaplama tekniği, ince film tabakalarının homojen bir şekilde

kaplanmasını sağlarken, air-knife teknolojisi ise kaplama sonrası fazla çözgenin uzaklaştırılmasına yardımcı olmaktadır. Bu iki teknik, özellikle büyük alan perovskit güneş hücreleri üretiminde kritik öneme sahiptir. Model tarafından tahmin edilen parametreler kullanılarak PSC aygıt üretimi gerçekleştirilmiştir. Üretim sürecinde, belirlenen parametrelerin optimize edilmesiyle, PSC aygıtının performans ölçümlerinde düz ölçüm sonucunda %8,52 ve ters ölçüm sonucunda %8,71 PCE değeri elde edilmiştir. Bu değerler, deneysel çalışmalar sürecinde üretilen en yüksek PCE değerine sahip PSC için sırasıyla %11,37'lik ve %4,94'lük bir artış sağlandığını göstermektedir.

Bu çalışmanın sonuçları, XGBoost modelinin doğru parametre tahminleri yaparak üretim sürecini optimize etme potansiyelini açıkça ortaya koymaktadır. XGBoost, gradyan artırma tekniğini kullanan güçlü bir makine öğrenmesi algoritmasıdır ve özellikle karmaşık veri setleri üzerinde yüksek doğrulukla tahminler yapabilme yeteneğine sahiptir. Bu özellikleri sayesinde, PSC üretim süreçlerinde kullanılan parametrelerin optimize edilmesinde etkili bir araç olarak öne çıkmaktadır. Perovskit güneş hücreleri, yüksek verimlilikleri ve düşük üretim maliyetleri nedeniyle güneş enerjisi alanında büyük bir potansiyel göstermektedir. Ancak, bu teknolojinin ticari uygulamalara geçişi için verimlilik ve kararlılık gibi parametrelerin daha da iyileştirilmesi gerekmektedir. XGBoost gibi ML algoritmaları, bu iyileştirmelerin gerçekleştirilmesinde kritik bir rol oynayabilir.

Perovskit katmanının farklı kimyasallar kullanılarak oluşturulması yanında, her bir katman da farklı malzemeler ve teknikler kullanılarak elde edilebilmektedir. Bu katmanlarda yapılan her bir değişiklik güç dönüşüm verimi veya kararlılık gibi parametrelere olumlu/olumsuz etki etmektedir. Gelecekte, daha fazla parametre ve farklı üretim teknikleri üzerinde çalışarak, PSC aygıtlarının verimliliğinin daha da artırılması hedeflenmektedir. Bu, farklı malzeme bileşimlerinin, katman kalınlıklarının ve üretim koşullarının sistematik bir şekilde incelenmesini ve çalışma sonunda elde edilmiş olan veri setine dahil edilmesini içermektedir. Makine öğrenmesi algoritmaları, bu tür çok boyutlu optimizasyon problemlerinde hızlı ve etkili çözümler sunmak için özellikle son dönemde çok farklı araştırma alanında kullanılmaktadır. Nitelikli verinin her geçen gün öneminin arttığı, özellikle PSC

üretimi gibi karmaşık süreçlerden oluşan akademik çalışmalarda ML kullanılarak çok daha hızlı şekilde hedeflenen sonuçlara ulaşılabilceđi düşünölmektedir.

Bu tez çalışması hem akademik araştırmalar hem de endüstriyel uygulamalar için önemli bir referans noktası oluşturmaktadır. Akademik alanda, PSC'lerin büyük alan üretim süreçlerinin daha iyi anlaşılmasında, endüstriyel alanda ise üretim süreçlerinin daha verimli hale getirilmesinde, ML algoritmalarının özellikle nitelikli verinin varlığında çok önemli ve güçlü bir araç olarak kullanılabileceđini göstermektedir. Bu, yenilenebilir enerji sektöründe dijital dönüşümün bir parçası olarak değerlendirilebilir ve gelecekte daha sürdürülebilir enerji çözümlerinin geliştirilmesine olanak tanıyabilir. Yenilenebilir enerji kaynaklarının daha geniş çapta benimsenmesi ve enerji üretiminde sürdürülebilirliđin sağlanması için bu tür yenilikçi yaklaşımlar kritik öneme sahiptir.



KAYNAKLAR DİZİNİ

- Abadi, E.A.J., Sahu, H., Javadpour, S.M. and Goharimanesh, M., (2022).** “Interpretable machine learning for developing high-performance organic solar cells”, *Materials Today Energy*, 25, 100969.
- Alpaydm, E., (2020).** *Introduction to Machine Learning*, The MIT Press.
- Almora, O., Baran, D., Bazan, G. C., Berger, C., Cabrera, C. I., Catchpole, K. R., Erten-Ela, S., Guo, F., Hauch, J., Ho-Baillie, A. W. Y., Jacobsson, T. J., Janssen, R. A. J., Kirchartz, T., Kopidakis, N., Li, Y., Loi, M. A., Lunt, R. R., Mathew, X., McGehee, M. D., Min, J., Mitzi, D. B., Nazeeruddin, M. K., Nelson, J., Nogueira, A. F., Paetzold, U. W., Park, N., Rand, B. P., Rau, U., Snaith, H. J., Unger, E., Vaillant-Roca, L., Yip, H. and Brabec, C.J., (2021).** “Device Performance of Emerging Photovoltaic Materials (Version 2)”, *Advanced Energy Materials*, 2102526.
- Breeze, P., (Ed.), (2016).** *Solar Power Generation*, Academic Press, London, 100p.
- Chen, C., Maqsood, A., & Jacobsson, T. J., (2023).** The role of machine learning in perovskite solar cell research. *Journal of Alloys and Compounds*, Volume 960, 170824. <https://doi.org/10.1016/j.jallcom.2023.170824>
- Chen, T., & Guestrin, C., (2016).** XGBoost: A Scalable Tree Boosting System, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://api.semanticscholar.org/CorpusID:4650265>
- Chollet, F., (2017).** *Deep Learning with Python*, Manning Publications.
- David, T.W., Anizelli, H., Jacobsson, T.J., Gray, C., Teahan, W. and Kettle, J., 2020,** “Enhancing the stability of organic photovoltaics through machine learning”, *Nano Energy*, 78, 105342.

KAYNAKLAR DİZİNİ (devam)

- Dhimish, M. (2021).** Defining the best-fit machine learning classifier to early diagnose photovoltaic solar cells hot-spots. *Case Studies in Thermal Engineering*, 25. <https://doi.org/10.1016/j.csite.2021.100980>
- Do-Kyoung, L., Dong-Nyuk, J., Tae, K. A., & Nam-Gyu, P., (2019).** Precursor Engineering for a Large-Area Perovskite Solar Cell with >19% Efficiency, *ACS Energy Letters*, 4 (10), 2393-2401pp. <https://doi.org/10.1021/acsenenergylett.9b01735>
- Fraas, L. M. (2015).** Photovoltaics: Comparison of PV Cell Types. *Handbook of Clean Energy Systems*. <https://doi.org/10.1002/9781118991978.hces063>
- Fu, F., Feurer, T., Weiss, T. P., Pisoni, S., Avancini, E., Andres, C., Buecheler, S., & Tiwari, A. N. (2016).** High-efficiency inverted semi-transparent planar perovskite solar cells in substrate configuration. *Nature Energy*, 2(1), 16190. <https://doi.org/10.1038/nenergy.2016.190>
- Geron., A., (2019).** Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Publications.
- Graniero, P., Khenkin, M., Köbler, H., Hartono, N. T. P., Schlatmann, R., Abate, A., Unger, E., Jacobsson, T. J., & Ulbrich, C., (2023).** The challenge of studying perovskite solar cells' stability with machine learning. *Frontiers in Energy Research*, Volume 11 – 2023. <https://doi.org/10.3389/fenrg.2023.1118654>
- Guo, Z. and Lin, B., (2021).** “Machine learning stability and band gap of lead-free halide double perovskite materials for perovskite solar cells”, *Solar Energy*, 228, 689-699pp.
- Hancock, J. T., & Khoshgoftaar, T. M., (2020).** CatBoost for big data: an interdisciplinary review. *Journal of Big Data*, 7, 94pp. <https://doi.org/10.1186/s40537-020-00369-8>

KAYNAKLAR DİZİNİ (devam)

- Hao, T., Leng, S., Yang, Y., Zhong, W., Zhang, M., Zhu, L., Song, J., Xu, J., Zhou, G., Zou, Y., Zhang, Y. and Liu, F., (2021).** “Capture the high-efficiency non-fullerene ternary organic solar cells formula by machine-learning-assisted energy-level alignment optimization”, *Patterns*, 2, 100333.
- Hartono, N. T. P., Köbler, H., Graniero, P., Khenkin, M., Schlattmann, R., Ulbrich, C., & Abate, A., (2023).** Stability follows efficiency based on the analysis of a large perovskite solar cells ageing dataset. *Nature Communications*, 14, Article number: 4869 (2023). <https://doi.org/10.1038/s41467-023-40585-3>
- Im, J., Lee, S., Ko, T., Kim, H.W., Hyon, Y. and Chang, H., (2019).** “Identifying Pb-free perovskites for solar cells by machine learning”, *npj Computational Materials*, 5:37.
- Jacobsson, T. J., Hultqvist, A., García-Fernández, A., Anand, A., Al-Ashouri, A., Hagfeldt, A., Crovetto, A., Abate, A., Ricciardulli, A. G., Vijayan, A., Kulkarni, A., Anderson, A. Y., Darwich, B. P., Yang, B., Coles, B. L., Perini, C. A. R., Rehermann, C., Ramirez, D., Fairen-Jimenez, D., ... Unger, E. (2022).** An open-access database and analysis tool for perovskite solar cells based on the FAIR data principles. In *Nature Energy* (Vol. 7, Issue 1, pp. 107–115). *Nature Research*. <https://doi.org/10.1038/s41560-021-00941-3>
- Kirman, J., Johnston, A., Kuntz, D.A., Askerka, M., Gao, Y., Todorovic, P., Ma, D., Prive, G.G. and Sargent, E.H., (2020).** “Machine-Learning-Accelerated Perovskite Crystallization”, *Matter*, 2, 1-10pp.
- Kumar, R. E., Tiihonen, A., Sun, S., Fenning, D. P., Liu, Z., & Buonassisi, T. (2022).** Opportunities for machine learning to accelerate halide-perovskite commercialization and scale-up. In *Matter* (Vol. 5, Issue 5, pp. 1353–1366). *Cell Press*. <https://doi.org/10.1016/j.matt.2022.04.016>

KAYNAKLAR DİZİNİ (devam)

- Le Corre, V.M., Sherkar, T.S., Koopmans, M. and Koster, L.J.A., (2021).** “Identification of the dominant recombination process for perovskite solar cells based on machine learning”, *Cell Reports Physical Science*, 2, 100346.
- Li, F., Peng, X., Wang, Z., Zhou, Y., Wu, Y., Jiang, M. and Xu, M., (2019).** “Machine Learning (ML)-Assisted Design and Fabrication for Solar Cells”, *Energy & Environmental Materials*, 280-291pp.
- Li, J., Pradhan, B., Gaur, S. and Thomas, J., (2019).** “Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells”, *Advanced Energy Materials*, 1901891. <https://doi.org/10.1002/aenm.201901891>
- Li, Y., Lu, Y., Huo, X., Wei, D., Meng, J., Dong, J., Qiao, B., Zhao, S., Xu, Z., & Song, D., (2021).** Bandgap tuning strategy by cations and halide ions of lead halide perovskites learned from machine learning, *RSC Advances*, Volume 11, Issue 26, 15688-15694pp. <https://doi.org/10.1039/d1ra03117a>
- Liu, Y., Tan, X., Liang, J., Han, H., Xiang, P., & Yan, W., (2023).** Machine Learning for Perovskite Solar Cells and Component Materials: Key Technologies and Prospects. *Advanced Functional Materials*, Volume33, Issue17, 2214271. <https://doi.org/10.1002/adfm.202214271>
- Liu, Y., Yan, W., Han, S., Zhu, H., Tu, Y., Guan, L., & Tan, X., (2022b),** How Machine Learning Predicts and Explains the Performance of Perovskite Solar Cells, *Solar RRL*, Volume 6, Issue 6, 2100794pp. <https://doi.org/10.1002/solr.202101100>
- Liu, Y., Yan, W., Zhu, H., Tu, Y., Guan, Li. and Tan, X., (2021).** “Study on bandgap predications of ABX₃-type perovskites by machine learning”, *Organic Electronics*, 101, 106426.

KAYNAKLAR DİZİNİ (devam)

- Liu, Z., Rolston, N., Flick, A.C., Colburn, T.W., Ren, Z., Dauskardt, R.H. and Buonassisi, T., (2022a).** “Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing”, *Joule*, 834–849pp.
- Lu, Y., Wei, D., Liu, W., Meng, J., Huo, X., Zhang, Y., Liang, Z., Qiao, B., Zhao, S., Song, D., & Xu, Z. (2022).** Predicting the device performance of the perovskite solar cells from the experimental parameters through machine learning of existing experimental results. *Journal of Energy Chemistry*. <https://doi.org/10.1016/j.jechem.2022.10.024>
- Mammeri, M., Dehimi, L., Bencherif, H., & Pezzimenti, F. (2023).** Paths towards high perovskite solar cells stability using machine learning techniques. *Solar Energy*, 249, 651–660. <https://doi.org/10.1016/j.solener.2022.12.002>
- Messenger, R. A. and Ventre, J., (2004).** Photovoltaic Systems Engineering, CRC Press, London-New York -Washington DC, 21-45pp.
- Mellit, A., & Kalogirou, S. (2022a).** Handbook of Artificial Intelligence Techniques in Photovoltaic Systems. Elsevier. <https://doi.org/10.1016/C2019-0-00960-0>
- Mellit, A., & Kalogirou, S. (2022b).** Assessment of machine learning and ensemble methods for fault diagnosis of photovoltaic systems. *Renewable Energy*, 184, 1074–1090. <https://doi.org/10.1016/j.renene.2021.11.125>
- Moroney, L., (2020).** AI and Machine Learning for Coders_ A Programmer's Guide to Artificial Intelligence, O'Reilly Publications.
- Müller, A.C., & Guido, S., (2017).** Introduction to Machine Learning with Python, O'Reilly Publications.

KAYNAKLAR DİZİNİ (devam)

Nield, T., (2022). Essential Math for Data Science - Take Control of Your Data with Fundamental Linear Algebra, Probability, and Statistics, O'Reilly Publications.

NREL. (2024). Best Research-Cell Efficiencies:
<https://www.nrel.gov/pv/assets/pdfs/best-research-cell-efficiencies.pdf>
 (Erişim Tarihi: 8 Ağustos 2024)

Odabaşı, Ç. and Yıldırım, R., (2019). “Performance analysis of perovskite solar cells in 2013–2018 using machine-learning tools”, Nano Energy, 56, 770-791pp.

Parikh, N., Akin, S., Kalam, A., Prochowicz, D., & Yadav, P., (2023). Probing the Low-Frequency Response of Impedance Spectroscopy of Halide Perovskite Single Crystals Using Machine Learning. ACS Appl. Mater. Interfaces 2023, 15, 27801–27808. <https://doi.org/10.1021/acsami.3c00269>

Park, N. G., & Zhu, K. (2020). Scalable fabrication and coating methods for perovskite solar cells and solar modules. In Nature Reviews Materials (Vol. 5, Issue 5, pp. 333–350). Nature Research. <https://doi.org/10.1038/s41578-019-0176-2>

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., & Gulin, A., (2017). CatBoost: unbiased boosting with categorical features. Neural Information Processing Systems.
<https://api.semanticscholar.org/CorpusID:5044218>

Russell, S.J., & Norvig, P., (2010). Artificial Intelligence A Modern Approach Third Edition, Prentice Hall.

Sahu, H. and Ma, H., 2019, “Unraveling Correlations between Molecular Properties and Device Parameters of Organic Solar Cells Using Machine Learning”, The Journal of Physical Chemistry Letters, 10 (22), 7277-7284pp.

KAYNAKLAR DİZİNİ (devam)

- Sharifzadeh, M., Sikinioti-Lock, A., & Shah, N. (2019).** Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression. *Renewable and Sustainable Energy Reviews*, 108, 513–538. <https://doi.org/10.1016/j.rser.2019.03.040>
- Sun, S., Hartono, N.T.P., Ren, Z.D., Oviedo, F., Buscemi, A.M., Layurova, M., Chen, D.X., Ogunfunmi, T., Thapa, J., Ramasamy, S., Settens, C., DeCost, B.L., Kusne, A.G., Liu, Z., Tian, S.I.P., Peters, I.M., Correa-Baena, J.P. and Buonassisi, T., (2019).** “Accelerated Development of Perovskite-Inspired Materials via High-Throughput Synthesis and Machine-Learning Diagnosis”, *Joule*, 1437-1451pp.
- Tsoi, K. & Yerci, S., (2024).** Bayesian Optimization with Experience for Fast Development of Monolithic Tandem Solar Cells: SimulationCase Study, *Advanced Theory and Simulations*, (Vol. 7, Issue 4, 2301013), <https://doi.org/10.1002/adts.202301013>
- VanderPlas, J., (2017).** *Python Data Science Handbook Essential Tools for Working with Data*, O'Reilly Publications.
- Xu, W., Liu, Z., Piper, R. T., & Hsu, J. W. P., (2023).** Bayesian Optimization of photonic curing process for flexible perovskite photovoltaic devices, *Solar Energy Materials and Solar Cells*, (Vol. 249, 112055), <https://doi.org/10.1016/j.solmat.2022.112055>
- Wu, T. and Wang, J., (2019).** “Global discovery of stable and non-toxic hybrid organic-inorganic perovskites for photovoltaic systems by combining machine learning method with first principle calculations”, *Nano Energy*, 66, 104070.

KAYNAKLAR DİZİNİ (devam)

- Yan, W., Liu, Y., Zang, Y., Cheng, J., Wang, Y., Chu, L., Tan, X., Liu, L., Zhou, P., Li, W., Zhong, Z., (2022).** “Machine learning enabled development of unexplored perovskite solar cells with high efficiency”, Nano Energy, 99, 107394.
- Yılmaz, B., Odabaşı, Ç. and Yıldırım, R., (2022).** “Efficiency and Stability Analysis of 2D/3D Perovskite Solar Cells Using Machine Learning”, Energy Technology, 10, 2100948.
- Yılmaz, B. and Yıldırım, R., (2020).** “Critical review of machine learning applications in perovskite solar research”, Nano Energy, 80, 105546.
- Yin, L., Gao, Q., Zhao, L., Zhang, B., Wang, T., Li, S., & Liu, H. (2020).** A review of machine learning for new generation smart dispatch in power systems. Engineering Applications of Artificial Intelligence, 88. <https://doi.org/10.1016/j.engappai.2019.103372>
- Zhang, J., Liu, B., Liu, Z., Wu, J., Arnold, S., Shi, H., Osterrieder, T., Hauch, J. A., Wu, Z., Luo, J., Wagner, J., Berger, C. G., Stubhan, T., Schmitt, F., Zhang, K., Sytnyk, M., Heumueller, T., Sutter-Fella, C. M., Peters, I. M., Zhao, Y., & Brabec, C. J., (2023).** Optimizing Perovskite Thin-Film Parameter Spaces with Machine Learning-Guided Robotic Platform for High-Performance Perovskite Solar Cells. Advanced Energy Materials, Volume13, Issue 48, 2302594. <https://doi.org/10.1002/aenm.202302594>

TEŞEKKÜR

Doktora eğitimim ve tez çalışmam süresince bana rehberlik eden, bilgi ve tecrübelerini benimle paylaşan, her daim destek olan değerli hocam ve danışmanım Doç. Dr. Ahmet YILANCI'ya sonsuz teşekkürlerimi sunarım. Görüşlerinden sıkça faydalandığım, önerileriyle çalışmamı yönlendiren Prof. Dr. Ceylan ZAFER'e, kıymetli tecrübelerinden ve ihtiyaç duyduğumdaki desteklerinden ve değerli katkılarından dolayı Doç. Dr. Burak GÜLTEKİN ve Doç. Dr. Meriç ÇETİN'e teşekkürlerimi ve saygılarımı sunarım.

Bu süreçte yanımda olan ve ihtiyaç duyduğum her an sonsuz desteklerini esirgemeyen değerli arkadaşlarım Dilek ÇIRAK, Zeynep GÜNEL ve Başak TURGUT'a içten teşekkürlerimi iletmek isterim. Onların yardımları olmadan bu çalışmayı tamamlamak mümkün olmazdı.

Doktora tez sürecimde, 2016K12-2841 numaralı Yeni Nesil Fotovoltaik Teknolojilerinin Üretim ve Karakterizasyon Altyapısının Geliştirilmesi isimli T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı destekli altyapı projesinin sağlamış olduğu maddi desteklerden dolayı en içten teşekkürlerimi sunarım. Tez aşamasından itibaren beni desteklemeye değer bulan TÜBİTAK'a, BİDEB 2211-C Yurt İçi Öncelikli Alanlar Doktora Burs Programı kapsamında verdikleri destek için de ayrıca teşekkür ederim.

01/10/2024

Burak KAHRAMAN

