

FARKLI DİSİPLİNLERDE MAKİNE ÖĞRENMESİ UYGULAMALARI: TAHMİN, SINIFLANDIRMA VE AÇIKLANABİLİRLİK ÜZERİNE VAKA ÇALIŞMALARI

EDİTÖR: DR. ÖĞR. ÜYESİ HÜSEYİN BODUR



BİDGE Yayınları

**Farklı Disiplinlerde Makine Öğrenmesi Uygulamaları:
Tahmin, Sınıflandırma ve Açıklanabilirlik Üzerine Vaka
Çalışmaları**

Editör: Hüseyin Bodur

ISBN: 978-625-8821-74-1

1. Baskı

Sayfa Düzeni: Gözde YÜCEL

Yayınlama Tarihi: 2026-07-21

BİDGE Yayınları

Bu eserin bütün hakları saklıdır. Kaynak gösterilerek tanıtım için yapılacak kısa alıntılar dışında yayıncının ve editörün yazılı izni olmaksızın hiçbir yolla çoğaltılamaz.

Sertifika No: 71374

Yayın hakları © BİDGE Yayınları

www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com

Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.

Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya /
Ankara



İÇİNDEKİLER

Güneş Enerjisi Üretimi Tahmininde Transformer Tabanlı ve Rejim Duyarlı Uyarlamalı Sinyal Filtreleme: Küresel Ölçekte Genellenebilir Ve Açıklanabilir Boru Hattı Tasarımı	1
--	---

MELİH YİĞİT KOTMAN, MEHMET MİLLİ

Makine Öğrenmesi Yöntemleri ile Github Depo Popülerliğinin Sınıflandırılması ve Veri Senaryolarının Karşılaştırmalı Analizi ..	54
--	----

SUDENUR GÜVEZ, İSMAİL HAKKI PARLAK

Uzaktan Çalışanlarda Tükenmişlik Riskinin Makine Öğrenmesi ile Belirlenmesi: Öznitelik Önemi ve Model Performans Analizi	107
--	-----

MÜMİNE AİŞE DEMİRAL, FERHAT DEMİRAY

BÖLÜM 1

GÜNEŞ ENERJİSİ ÜRETİMİ TAHMİNİNDE TRANSFORMER TABANLI VE REJİM DUYARLI UYARLAMALI SİNYAL FİLTRELEME: KÜRESEL ÖLÇEKTE GENELLENEBİLİR VE AÇIKLANABİLİR BORU HATTI TASARIMI

MELİH YİĞİT KOTMAN¹
MEHMET MİLLİ²

Giriş

Günümüz dünyasında enerjinin kritik bir öneme sahip olduğu, fosil yakıt rezervlerinin tükenmekte olması ve fosil yakıtların oluşturduğu iklim krizi nedeniyle yenilenebilir enerji kaynaklarına geçişin kaçınılmaz hale geldiği görülmektedir. Bu bağlamda güneş enerjisi (PV), çevresel ve ekonomik avantajlarıyla öne çıkan en kritik sürdürülebilir enerji kaynaklarından biri haline gelmiştir. Poursal ve arkadaşlarının yaptığı çalışmada, 2050'ye kadar güneş enerjisi santrallerinin artırılmasıyla dünyadaki güneş enerjisinin ikinci en önemli elektrik üretim kaynağı haline geleceği

¹ Öğrenci, Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0009-0000-4555-6369

² Dr. Öğr. Üyesi, Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0002-0759-4433

ve elektrik tüketiminin yaklaşık %25'lik kısmını karşılayacağı öngörülmektedir.

Aynı zamanda güneş enerjisi alanında Çin, ABD ve Hindistan'ın üretim kapasitesi eklemelerinde dünyada öncülük ettiği, Avrupa'nın ise politika anlamında güneş enerjisinin özümsemesi konusunda liderliğini sürdürdüğü görülmektedir; bu alanda Almanya, İspanya ve İtalya önemli rol oynamaktadır. İtalya özelinde ise ülkenin, teşvik tarifelerinin desteklediği ilk büyüme dönemini geride bırakarak çatı sistemleri, topluluk projeleri ve öz-tüketime dayanan daha istikrarlı bir pazar yapısına geçtiği; ancak üretim ve Ar-Ge kapasitesi açısından Çin ve ABD'ye kıyasla halen görece geride kaldığı rapor edilmiştir (Ahmed et al., 2025).

Kishore ve arkadaşlarının yaptığı araştırmada, güneş enerjisinden üretilen her megavat-saatlik elektriğin yaklaşık 0,5 ton CO₂ emisyonunu azalttığı vurgulanırken, 2050 yılına kadar küresel CO₂ emisyonunun da önemli ölçüde azalacağı ve bu sayede daha az karbon yoğunluklu, sürdürülebilir bir geleceğe evrilineceği öngörülmektedir (Kishore et al., 2025).

Devletlerin güneş enerjisine uyguladığı elektrik tarifeleri, vergi teşvikleri ve fotovoltaik modüllerdeki maliyet düşüşleri gibi destekleyici politikalar, fosil yakıtlara karşı güneş enerjisi fotovoltaik sistemlerinin rekabet gücünü daha da artırmıştır (Lazaroiu et al., 2023). Ayrıca yatırımların ve teşviklerin önündeki engellerin kaldırılması halinde, güneş enerjisinin yakın gelecekte fosil yakıtların yerini alacağı da bildirilmiştir (Nijssse et al., 2023).

Fotovoltaik üretim hava koşullarına büyük ölçüde bağlı olduğundan, doğru tahmin güç sistemi kararlılığı ve enerji yönetimi açısından merkezi bir rol oynamaktadır. Bu nedenle yapay zekâ, özellikle kısa vadeli PV üretim tahmininde önem kazanmaktadır; çünkü daha güvenilir planlama, dağıtım ve müşteri beslemesi

sağlamaktadır (Agga et al., 2022; Elsaraiti & Merabet, 2022; Luo et al., 2021; Yu et al., 2024).

Hibrit yaklaşımlarda kümeleme mekanizması (Gabaldón et al., 2024; Gao et al., 2025; Rana et al., 2016; Z. Wu et al., 2022) ve sinyal ayrıştırma teknikleri (Ait Chaoui et al., 2025; Chen et al., 2025)sıklıkla kullanılmasına rağmen; bu çalışmaların çoğu genellikle sınırlı veri hacmi veya benzer iklim özelliklerine sahip bölgeler (Agga et al., 2022; Arslan et al., 2025; RajasundrapandiyanLeebanon et al., 2025) üzerinde kurgulanmıştır. Bu durum, önerilen modellerin küresel ölçekte farklı iklim rejimlerindeki dayanıklılığı ve genellenebilirliği (Lv et al., 2026; H. Wu et al., 2021) konusunda literatürde önemli bir boşluk yaratmaktadır.

Bu çalışmayla, literatürdeki söz konusu boşlukları doldurmak amacıyla, adaptif sinyal filtreleme ve Transformer tabanlı derin öğrenme mimarilerinin (Custom PatchTST ve Custom Autoformer) entegre edildiği; uçtan uca, fiziksel olarak açıklanabilir ve küresel olarak farklı iklimlerde test edilmiş, genellenebilir bir güneş enerjisi (GHI) tahmin hattı (pipeline) önerilmektedir.

Literatür Taraması

Makine Öğrenmesi Tabanlı Yaklaşımlar

Güneş enerjisi/PV güç tahmini probleminde ilk dönem çalışmalar büyük ölçüde klasik makine öğrenmesi algoritmalarına dayanmaktadır. Kiasari ve Aly, Hindistan'daki iki farklı güneş enerjisi santralinden (Gandikota ve Nasik) elde edilen verilerle Destek Vektör Regresyonu (SVR), Tekrarlayan Sinir Ağları (RNN), eXtreme Gradient Boosting (XGBoost), K-En Yakın Komşu (KNN) ve Rastgele Orman (Random Forest) algoritmalarını karşılaştırmış; gecikmeli adımların (lagged steps) ve farklı duyarlılık koşullarının model performansına etkisini incelemiş ve XGBoost'un doğruluk

kararlılık dengesi açısından en başarılı yöntem olduğunu raporlamıştır (Kiasari & Aly, 2024).

Benzer şekilde Rajasundrapandiyan ve arkadaşları, Hindistan'daki yedi farklı şehirden (Bhubaneswar, Chennai, Delhi, Hyderabad, Nagpur, Patna, Trivandrum) elde edilen 13 yıllık meteorolojik verilerle uzun vadeli güneş radyasyonu tahmini gerçekleştirmiş; Ampirik Mod Ayırıştırması (EMD) ve Topluluk Ampirik Mod Ayırıştırmasını (EEMD) çok katmanlı algılayıcı (MLP), rastgele orman regresyonu (RFR), SVR ve sırt regresyonu (ridge regression) ile birleştirmiş ve EEMD-MLP kombinasyonunun tüm lokasyonlarda en iyi performansı verdiğini göstermiştir (RajasundrapandiyanLeebanon et al., 2025). Bu çalışmalar, klasik makine öğrenmesi yöntemlerinin sinyal ayırıştırma teknikleriyle desteklendiğinde başarılı sonuçlar verebildiğini ortaya koymakla birlikte, kullanılan veri setlerinin tek bir ülke veya iklim bölgesiyle sınırlı kalması nedeniyle genellenebilirlik açısından kısıtlı kalmaktadır.

Kümeleme Tabanlı Hibrit Modeller

PV üretiminin gün içi ve günler arası değişkenliğini ele almak amacıyla literatürde kümeleme (clustering) tabanlı hibrit modellere yoğun olarak başvurulmuştur. Rana ve arkadaşları, hava durumu tipine göre kümeleme yöntemini sinir ağı topluluklarıyla (ensembles of neural networks) birleştirerek güneş enerjisi üretim tahmini yapmıştır (Rana et al., 2016). Gabaldón ve arkadaşları, elektrik kurumlarında kısa vadeli PV üretim/yük tahmini için K-Means kümelemesini sinir ağı yöntemleriyle birleştiren bir model önermiştir (Gabaldón et al., 2024).

Gao ve arkadaşları ise benzer gün kümelemesini (similar day clustering) CNN-GRU mimarisiyle bütünleştirerek PV güç tahmini gerçekleştirmiş ve bu yaklaşımın tahmin doğruluğunu artırdığını göstermiştir (Gao et al., 2025). Wu ve arkadaşları, korelasyon

analiziyle global yatay radyasyonun PV üretimi üzerindeki en etkili meteorolojik faktör olduğunu belirleyip, bu korelasyona göre kümelenen veri alt kümelerini CNN ile öznetelik çıkarımından geçirerek Informer modeline besleyen bir CNN-Informer mimarisi önermiştir (Z. Wu et al., 2022). Bu çalışmalar, kümeleme adımının model performansına anlamlı katkı sağladığını göstermekle birlikte, genellikle tekil santral/bölge verisiyle sınırlı kalmaları nedeniyle farklı iklim koşulları altında genellenebilirlikleri test edilmemiştir.

Sinyal Ayırıştırma Tabanlı Hibrit Modeller

Kümeleme yaklaşımlarına kıyasla daha az sayıda çalışma, PV zaman serilerindeki gürültü ve değişkenliği gidermek amacıyla sinyal işleme tabanlı ayırıştırma tekniklerini derin öğrenme mimarileriyle bütünleştirmiştir. Ait Chaoui ve arkadaşları, Dalgacık Dönüşümü (Wavelet Transform) ile ham PV zaman serisini çok-frekanslı bileşenlere ayırdırdıktan sonra bir Transformer kodlayıcı, Zamansal Evrişimli Ağ (TCN), Verimli Kanal Dikkat Ağı (ECANet) ve Geçitli Tekrarlayan Birim (GRU) katmanlarını birleştiren hibrit bir mimari (WT-Transformer-TCN-ECANet-GRU) önermiş; Fas'taki bir üniversite binasından elde edilen 95.885 adet beş dakikalık PV üretim kaydı üzerinde bu modelin LSTM, GRU ve CNN-LSTM gibi karşılaştırma modellerine üstünlük sağladığını göstermiştir (Ait Chaoui et al., 2025).

Chen ve arkadaşları, Bayes optimizasyonu ve ikincil ayırıştırma (secondary decomposition) ile desteklenen hata düzeltmeli bir derin Autoformer modeli önererek PV güç tahmininde Autoformer'ın sınırlılıklarını gidermeye çalışmıştır (Chen et al., 2025). Bu iki çalışma, sinyal ayırıştırmanın Transformer tabanlı mimarilerle birleştirilmesinin literatürde henüz olgunlaşmamış, sınırlı sayıda örnekle temsil edilen bir araştırma yönü olduğunu göstermektedir.

Transformer Tabanlı Derin Öğrenme Yaklaşımları

Son yıllarda uzun vadeli zaman serisi tahmininde kendi-dikkat (self-attention) mekanizmalarının başarısı, PV/GHI tahmini alanında da Transformer tabanlı mimarilerin benimsenmesine yol açmıştır. Wu ve arkadaşları tarafından önerilen Autoformer, klasik kendi-dikkat mekanizması yerine seri ayrıştırırmayı (decomposition) modelin temel bir bileşeni haline getiren ve periyodiklik temelli bir Auto-Correlation mekanizması kullanan bir mimaridir; enerji, trafik, ekonomi, hava durumu ve hastalık gibi beş farklı uygulama alanında altı kıyas verisinde ortalama %38 oranında iyileşme sağlamıştır (H. Wu et al., 2021). Nie ve arkadaşları tarafından önerilen PatchTST, zaman serisini örtüşmeyen yamalara (patches) bölerek kanal-bağımsız (channel-independent) bir Transformer kodlayıcıdan geçiren ve uzun vadeli tahmin görevlerinde hesaplama verimliliğini artıran bir mimaridir (Nie et al., 2022).

PatchTST mimarisi, çalışmamızda kullanılan Custom PatchTST modelinin doğrudan dayandığı omurgadır. Custom Autoformer modeli ise orijinal Autoformer'ın temel ilkesi olan zaman serisi ayrıştırmasından (series decomposition) ilham almakla birlikte, hesaplama maliyetini azaltmak amacıyla Auto-Correlation tabanlı dikkat mekanizması yerine basitleştirilmiş doğrusal projeksiyon katmanları kullanan, ayrıştırma-temelli (decomposition-linear) bir varyant olarak tasarlanmıştır.

PV alanına özgü uygulamalarda, Lv ve arkadaşları, Yapay Balık Sürüsü Algoritması-İzolasyon Ormanı (AFSA-IF) tabanlı anomali tespiti, Üretici Çatışmalı Ağ (GAN) tabanlı öznitelik çıkarımı, çok modlu veri füzyonu ve PatchTST'yi birleştiren gelişmiş bir tahmin çerçevesi önermiş; altı farklı PV santralinden elde edilen verilerle yapılan deneylerde bu modelin Transformer, Reformer ve Informer gibi karşılaştırma modellerine kıyasla, özellikle yağmurlu hava koşulları gibi zorlu senaryolarda, anlamlı performans artışı sağladığını göstermiştir (Lv et al., 2026). Arslan ve arkadaşları, gerçek dünya verisi üzerinde klasik makine öğrenmesi

yöntemleri ile Transformer tabanlı derin öğrenme modellerini karşılaştırmalı olarak değerlendirmiş ve Transformer tabanlı yaklaşımların güneş enerjisi tahmininde üstünlük sağladığını ortaya koymuştur (Arslan et al., 2025). Öte yandan, My ve arkadaşları tarafından önerilen MCTMF gibi çok kanallı Transformer tabanlı genel zaman serisi tahmin modelleri, PV'ye özgü olmasa da çok değişkenli zaman serilerindeki kanallar arası bağımlılıkların CNN tabanlı çok kanallı öğrenme mekanizmalarıyla modellenmesine yönelik genel mimari yaklaşımlar sunması açısından ilgi çekicidir (My et al., 2026).

Diğer Derin Öğrenme Yaklaşımları

Transformer öncesi dönemde literatür, ağırlıklı olarak tekil veya hibrit Evrişimli Sinir Ağı (CNN) ve Uzun-Kısa Süreli Bellek (LSTM)/GRU tabanlı mimarilere odaklanmıştır. Agga ve arkadaşları, kısa vadeli PV güç üretimi tahmini için verimli bir CNN-LSTM hibrit mimarisi önermiştir (Agga et al., 2022)}. Luo ve arkadaşları, alan bilgisini (domain knowledge) derin öğrenme modeline entegre ederek PV güç tahmininin fiziksel olarak daha açıklanabilir hale getirilmesini sağlamıştır (Luo et al., 2021). Elsaraiti ve Merabet, derin öğrenme tekniklerinin güneş enerjisi tahmininde kullanımını incelemiş (Elsaraiti & Merabet, 2022); Yu ve arkadaşları ise PV güç tahmini için kullanılan derin öğrenme modellerini kapsamlı bir gözden geçirme çalışmasıyla sınıflandırmıştır (Yu et al., 2024).

Literatürdeki Boşluklar ve Konumlandırma

Yukarıda özetlenen çalışmalar bir bütün olarak değerlendirildiğinde üç temel eğilim göze çarpmaktadır:

- Hibrit yaklaşımlarda kümeleme mekanizmalarına ağırlıklı olarak yer verilirken sinyal işleme tabanlı ayırıştırma tekniklerinin kısmen ve sınırlı sayıda

çalışmada kullanıldığı (Ait Chaoui et al., 2025; Chen et al., 2025; Gabaldón et al., 2024; Gao et al., 2025; Rana et al., 2016; Z. Wu et al., 2022);

- Transformer tabanlı mimarilerin (Autoformer, PatchTST) PV/GHI tahmininde henüz olgunlaşma aşamasında olduğu ve genellikle tek bir santral veya bölge verisiyle değerlendirildiği (Arslan et al., 2025; Chen et al., 2025; Lv et al., 2026; Nie et al., 2022; H. Wu et al., 2021);
- ve belki de en kritik olarak, mevcut çalışmaların büyük çoğunluğunun sınırlı miktarda veriyle, tek bir iklim bölgesinden toplanan verilerle çalıştığı ve önerilen modellerin farklı iklim koşulları altında (out-of-distribution) test edilmediği görülmektedir (Agga et al., 2022; Ait Chaoui et al., 2025; Arslan et al., 2025; Chen et al., 2025; Gabaldón et al., 2024; Gao et al., 2025; Lv et al., 2026; RajasundrapandiyanLeebanon et al., 2025; Rana et al., 2016; Z. Wu et al., 2022).

Bu çalışma, adaptif sinyal filtreleme ile Transformer tabanlı mimarileri (Custom PatchTST ve Custom Autoformer) bütünleştiren, açıklanabilirlik (SHAP) ile desteklenen ve küresel ölçekte 20 farklı şehirde iklimler arası genellenebilirliği (OOD) test edilen bir GHI tahmin hattı önererek bu boşlukları doldurmayı amaçlamaktadır.

Materyal ve Metot

Veri Kümesi ve Deneysel Altyapı

Bu çalışmada, derin zaman serisi modellerinin eğitilmesi, doğrulanması ve dünyanın farklı coğrafyalarındaki dağılım dışı (Out-of-Distribution - OOD) genelleme yeteneklerinin ölçülmesi

amacıyla iki farklı kaynaktan elde edilen meteorolojik ve astronomik veri setleri kullanılmıştır.

Modellerin temel eğitimi ve içi-dağılım (In-Distribution) performans ölçümleri için National Laboratory of the Rockies (NLR) tarafından sağlanan NSRDB (National Solar Radiation Database) veritabanı kullanılmıştır (Sengupta et al., 2018). Veriler NSRDB GOES API'si üzerinden; farklı iklim tiplerini temsil eden Denver CO, Miami FL, Phoenix AZ, Seattle WA ve Boston MA şehirlerinden 1 Ocak 2020 -- 31 Aralık 2024 tarihleri arasındaki 5 yıllık dönemi kapsayacak şekilde 30 dakika çözünürlükle çekilmiştir. Toplamda 438.000 satırdan oluşan bu veri hattı kronolojik olarak bölünerek; 2020-2022 yılları eğitim (Train), 2023 yılı doğrulama (Validation) ve 2024 yılı ise içi-dağılım test (Test) seti olarak ayrılmıştır.

Modellerin eğitim esnasında hiç karşılaşmadığı küresel iklim uçlarındaki direncini ölçmek amacıyla Open-Meteo Archive API kullanılarak 20 farklı şehirden oluşan 3 aylık (1 Ocak 2023 -- 31 Mart 2023) dağılım dışı küresel test seti kurgulanmıştır (Zippenfenig, 2023). Her ne kadar OOD test seti ile doğrulama seti aynı yılı (2023) kapsasa da bu setin birincil amacı zamansal değil; eğitim verisinde yer almayan farklı kıtalar ve ters yarımküre iklimlerinin yarattığı mekansal (spatial) dağılım dışı koşulları stres testine tabi tutmaktır. Zaman serisi bütünlüğünü korumak adına 1 saatlik orijinal veri çözünürlüğü enterpolasyon teknikleri ile 30 dakikaya resample edilmiştir. Küresel şehirler; Sıcak/Kurak, Sürekli Kapalı/Yağışlı, Tropikal/Muson, Ekstrem Zenith/Albedo, Ters Dinamikler (Güney Yarımküre) ve Mekansal Ezber Kontrolü olmak üzere 6 farklı zorlu senaryo altında sınıflandırılarak stres testine tabi tutulmuştur.

Geliştirme ortamında; modellerin inşası ve MPS (Metal Performance Shaders) destekli donanım hızlandırmalı tensör işlemleri için PyTorch kullanılmıştır. Veri manipülasyonu ve kayan

pencere (sliding window) matrislerinin üretimi için Pandas ve NumPy, yüksek hızlı disk okuma/yazma süreçleri için PyArrow ve Parquet formatı tercih edilmiştir. Gürültü giderme ve sinyal işleme safhalarında SciPy ve PyWavelets kütüphanelerinden yararlanılırken; Güneş Başucu Açısı (Solar Zenith Angle) ve Berraklık İndeksi (Clearness Index) gibi astronomik değişkenlerin formülizasyonu pvlib ile gerçekleştirilmiştir. İstatistiki kümeleme ve metrik hesaplamalarında Scikit-learn, model açıklanabilirliğinde SHAP, görselleştirmede ise Matplotlib kütüphaneleri entegre edilmiştir (Anderson et al., 2023; Harris et al., 2020; Hunter, 2007; Lee et al., 2019; McKinney, 2010; Paszke et al., 2019; Pedregosa et al., 2011; Richardson et al., 2026; Virtanen et al., 2020).

Derin Öğrenme Tahmin Mimarileri ve Matematiksel Altyapı

Proje kapsamında, zaman serisi matrislerini işlemek üzere Custom PatchTST (Nie et al., 2022) ve Custom Autoformer (H. Wu et al., 2021) mimarileri kurgulanmıştır. Modeller adil karşılaştırma (fair comparison) protokolü gereğince; Adam optimizasyon algoritması, Ortalama Kare Hata (MSE) kayıp fonksiyonu, 0.001 başlangıç öğrenme oranı (learning rate) ve 32 batch size ile donatılmıştır. Katman hiyerarşisi ve hiperparametre dağılımları Tablo I' de sunulmuştur.

Tablo 1 Modellerin Detaylı Mimari ve Hiperparametre Matrisi

Sistem Yapı Taşları ve Parametreler	Custom PatchTST	Custom Autoformer
Giriş Boyutu ($B \times L \times C$)	$32 \times 96 \times 11$	$32 \times 96 \times 11$
Tahmin Ufku Uzunluğu (T)	24	24

Yama Uzunluğu (P_l) / Adım Genişliği (S)	16, 8	-
Dekompozisyon Çekirdek Boyutu (kernel_size)	-	25
Model Gizli Boyutu (d_model)	64	_*
Dikkat Kafası Sayısı (nhead)	4	_*
Transformer Encoder Katman Sayısı	2	_*
Zaman Projeksiyon Katman Yapısı (nn.Linear)	$11 \times (P_n \cdot d_{model}) \rightarrow 24$	$11 \times 96 \rightarrow 24$
Nihai Çıktı Süzgeci Katmanı (nn.Linear)	$11 \rightarrow 1$	$11 \rightarrow 1$
Toplam Eğitilebilir Parametre Ağırlığı	$\sim 48.5K$	$\sim 45.2K$

Kaynak: (Kotman et al., 2026)

**Doğrudan doğrusal zaman izdüşümü ve dekompozisyon kullanan encoder-free yapı*

Custom PatchTST Matematiksel Altyapısı, girdi matrisindeki 11 özelliği bağımsız kanallar olarak ele alır ve her bir kanalı unfold operatörü vasıtasıyla yamalara (patches) böler. Toplam yama sayısı (P_n) şu formülle türetilmektedir:

$$P_n = \left\lfloor \frac{L - P_l}{S} \right\rfloor + 1 = \left\lfloor \frac{96 - 16}{8} \right\rfloor + 1 = 11 \quad (1)$$

Oluşan 11 yama model boyutunda (d_model = 64) projekte edilerek Transformer Encoder bloğunda işlenir. Bu yapı, dikkat

(attention) hesaplama karmaşıklığını $O(L^2)$ seviyesinden $O(P_n^2)$ seviyesine düşürmektedir.

Custom Autoformer Matematiksel Altyapısı, orijinal Autoformer’ın Auto-Correlation tabanlı dikkat mekanizması yerine, hesaplama maliyetini düşürmek amacıyla doğrudan zaman serisi ayrıştırma (Series Decomposition) gücüne odaklanan ayrıştırma-temelli (decomposition-linear) bir tasarım benimsenmiştir. Model her mini-batch güncellemesinde girdi matrisini 1-Boyutlu Ortalama Havuzlama (AvgPool1d) vasıtasıyla trend (X_t) ve mevsimsel (X_s) bileşenlerine ayırır:

$$X_t = AvgPool1d(Padding(X)), X_s = X - X_t \quad (2)$$

Havuzlama mekanizmasının çıkış uzunluğu (L_{out}) hesabı şu şekildedir:

$$L_{out} = \left\lceil \frac{L_{in} + 2 \times padding - kernelsize}{stride} \right\rceil + 1 \quad (3)$$

Zamanın başından ve sonundan veri kaybetmemek adına $padding = (kernelsize - 1)/2 = 12$ ve $stride = 1$ olarak uygulandığında çıktı boyutu 96 olarak korunmaktadır. Ayrıştırılan trend ve mevsimsel alt uzaylar, kanal eksenine göre bağımsız iki doğrusal katman (nn.Linear) ile zamansal olarak ($96 \rightarrow 24$) ayrı ayrı projekte edilir; elde edilen $X_t^{pred}, X_s^{pred} \in \mathbb{R}^{B \times 24 \times 11}$ bileşenleri 11 kanallı birleşik bir öngörü temsili oluşturulur ve son olarak 11 özelliği tek bir GHI değerine indirgeyen kanal-birleştirme katmanı (nn.Linear, $11 \rightarrow 1$) ile nihai tahmin üretilir. Bu tasarım, modelin tek bir Auto-Correlation/attention katmanı içermediğini; ancak ayrıştırma stratejisinin doğrudan miras alındığını açıkça ortaya koymaktadır.

Sinyal İşleme Altyapısı

Rejim-Bazlı Filtreleme Metodolojisi ile veri kümesindeki iklimsel heterojenliği yönetmek amacıyla günlük profiller K-Means

algoritmasıyla “Güneşli”, “Kaotik” ve “Kapalı” olmak üzere üç rejime ayrılmıştır. Kümeleme kararlığı ve filtre optimizasyonu doğrultusunda; Güneşli rejimde tepe noktası genlik kayıplarını (amplitude attenuation) önlemek adına 3.dereceden ve pencere genişliği 11 olan Savitzky-Golay seçilmiştir (Savitzky & Golay, 1964). Kaotik günlerdeki yüksek frekanslı bulut gürültülerini sönümlemek amacıyla 4. seviye Daubechies (db4) Ayrık Dalgacık Dönüşümü (DWT) mimarisi kurgulanmıştır (Donoho & Johnstone, 1998). Sürekli kapalı günlerdeki sensör titreşimlerini düzlemek için ise faz kaymasını (lag effect) engelleyecek şekilde merkezlenmiş 5 adımlı Hareketli Ortalama (Moving Average) filtresi metodolojiye entegre edilmiştir (Smith, 1997).

Bulgular ve Tartışma

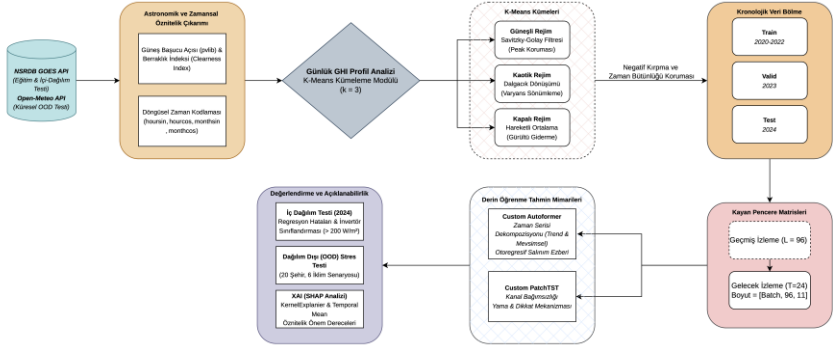
Önerilen Uçtan Uca Hibrit İş Akışı

Bu çalışmada, ham meteorolojik verilerin API üzerinden elde edilmesinden modellerin küresel stres altındaki karar mekanizmalarının açıklanmasına kadar uzanan uçtan uca (end-to-end) bir tahmin hattı (pipeline) kurgulanmıştır. Önerilen bu kapsamlı iş akışının mimari tasarımı ve modüller arası veri iletim süreci Şekil 1'de detaylı olarak gösterilmektedir. Sistem temel olarak ardışık yedi ana fazdan oluşmaktadır:

- Faz 1: Ham verilerin hasadı ve pvlib üzerinden astronomik özelliklerle (Zenith Açısı, Berraklık İndeksi) zenginleştirilmesi.
- Faz 2: Günlük profillerin K-Means ($k = 3$) ile ayrıştırılması ve rejim-bazlı adaptif filtreleme (SavGol, DWT, MA) ile gürültülerin sönümlenmesi.

- Faz 3: Kronolojik test bölmeleri ve $L = 96$, $T = 24$ yapılandırmasıyla 3-boyutlu zaman serisi tensörlerinin oluşturulması.
- Faz 4: Önerilen derin öğrenme modellerinin eğitim dinamiklerinin ve kayıp (loss) eğrilerinin asimptotik analizi.
- Faz 5: 2024 test seti üzerindeki çok adımlı içi-dağılım (in-distribution) tahmin performansı ve invertör sınıflandırma karnesinin tespiti.
- Faz 6: 20 farklı küresel şehir üzerinde Dağılım Dışı (OOD) mekansal genelleme ve şebeke kararlılığı stres testleri.
- Faz 7: SHAP (Temporal Mean) tabanlı öznetelik analiziyle XAI karar düğümlerinin seffalaştırılması.

Şekil 1. Önerilen Rejim-Bazlı Adaptif Sinyal İşleme ve Derin Öğrenme Tahmin İş Akışı (Pipeline)



Kaynak: (Kotman et al., 2026)

Faz 1: Veri Hasadı ve Astronomik Zenginleştirme Bulguları

İlk aşamada yürütülen veri hasadı operasyonu sonucunda, NSRDB API üzerinden 5 farklı Amerika Birleşik Devletleri şehri

(Denver, Miami, Boston, Phoenix, Seattle) için 2020-2024 yıllarını kapsayan, yarım saatlik çözünürlükte toplam 438.000 satırlık kesintisiz bir veri havuzu oluşturulmuştur. API üzerinden çekilen ham meteorolojik özniteliklerin (Sıcaklık, Nem, Rüzgâr Hızı, Bulut Tipi) tek başına derin öğrenme modellerinin evrensel iklim fiziklerini kavramasında yetersiz kalacağı öngörülerek, veri seti pvlib kütüphanesi ile astronomik boyutta zenginleştirilmiştir.

Bu zenginleştirme kapsamında iki kritik fiziksel limit modele tanıtılmıştır:

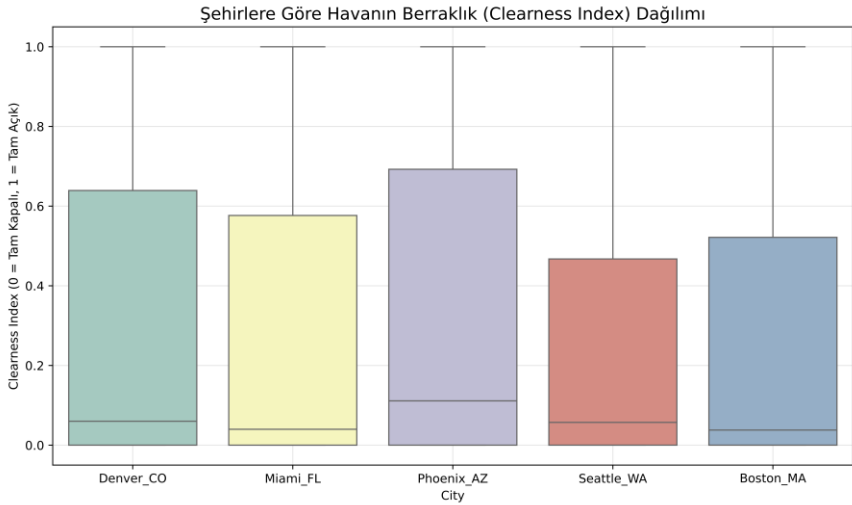
- Güneş Başucu Açısı (Solar Zenith Angle): Her bir şehrin coğrafi enlem ve boylam koordinatlarına göre milimetrik olarak hesaplanan bu açı, modele güneşin ufuk çizgisine göre anlık konumunu vererek, "gece" ve "gündüz" sınırlarının algoritmik olarak ezberlenmek yerine fiziksel olarak öğrenilmesini sağlamıştır.
- Berraklık İndeksi (Clearness Index): Atmosfer-dışı radyasyon oranları formülize edilerek elde edilen bu indeks, filtrelenmiş GHI sinyali ile birleştirilmiştir. Böylece modelin, anlık radyasyon kayıplarının saatten mi (güneşin batması) yoksa atmosferik bir olaydan mı (bulut/sis) kaynaklandığını ayırt edebilmesi hedeflenmiştir.

Bunun yanı sıra, zaman bazlı özelliklerin (saat ve ay) doğrusal (örneğin 0'dan 23'e) olarak modele beslenmesi, sınır ağlarında 23:30'dan 00:00'a geçerken suni bir matematiksel sıçrama (uçurum) hatasına yol açmaktadır. Bu sorunu bertaraf etmek amacıyla, zaman değişkenleri trigonometrik sin ve cos bileşenlerine ($hour_{sin}$, $hour_{cos}$, $month_{sin}$, $month_{cos}$) dönüştürülerek dögüsel (cyclical) olarak kodlanmıştır. Bu mühendislik dokunuşu,

zaman serisi tensörlerinde gece yarısı ve yılbaşı geçişlerini pürüzsüzleştirerek, Transformer modellerinin zamanı doğrusal bir çizgi değil, kusursuz bir döngü olarak işlemesine olanak tanımıştır.

Veri havuzunun astronomik ve meteorolojik heterojenliğini, şehirlerin mikroskobik iklimsel karakteristiği üzerinden somutlaştırmak amacıyla gerçekleştirilen istatistiksel analizler, derin öğrenme modellerinin karşılaşacağı zorluk derecelerini net bir şekilde ortaya koymaktadır.

Şekil 2 Şehirlere Göre Havanın Berraklık İndeksi (Clearness Index) Dağılımı

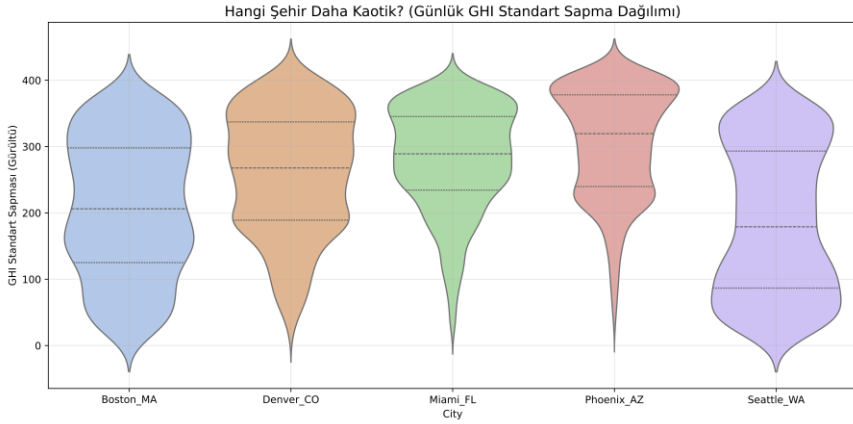


Kaynak: (Kotman et al., 2026)

Şekil 2'de sunulan "Şehirlere Göre Havanın Berraklık Dağılımı" incelendiğinde, atmosferik sönümleme oranlarının coğrafi konuma göre nasıl radikal biçimde değiştiği görülmektedir. Phoenix (AZ) ve Denver (CO) gibi kurak/yarı-kurak iklim kuşağında yer alan şehirlerin berraklık indeksleri 0,8 tepe noktasına yakın, dar varyanslı yoğun bir kümelenme sergilemektedir. Bu durum, ilgili bölgelerde atmosferik engelleyicilerin (bulut, yoğun nem) nadir olduğunu ve güneş radyasyonunun yeryüzüne doğrudan, kayıpsız

ulaştığını gözlemlenmiştir. Buna karşılık Seattle (WA) ve Boston (MA) gibi sürekli denizel veya nemli karasal etkilere açık şehirlerin berraklık dağılımları, daha geniş bir spektruma yayılmakta ve 0,2 ila 0,5 bandında yoğunlaşmaktadır. Bu saptama, modellerin sadece radyasyon miktarını tahmin etmekle kalmayıp, şehirlerin kalıcı bulutluluk şemalarını da içselleştirmesi gerektiğini doğrulamaktadır.

Şekil 3. Şehirlerin Günlük GHI Standart Sapma (Volatilite) Dağılımı

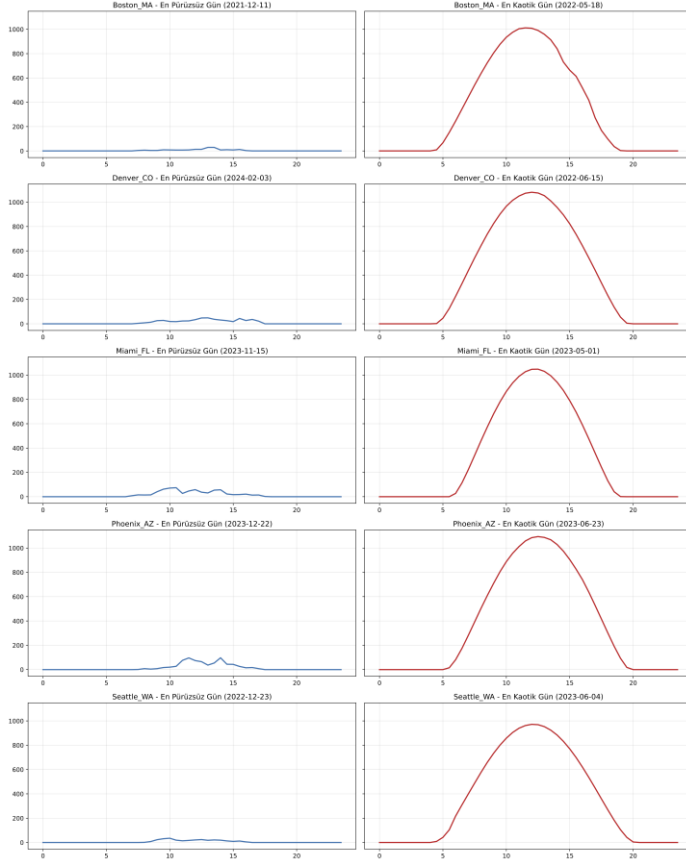


Kaynak: (Kotman et al., 2026)

Bu karakteristik farklar, Şekil 3'te sunulan volatilite (günlük GHI standart sapma) analiziyle de perçinlenmektedir. "Hangi Şehir Daha Kaotik?" sorusuna yanıt arayan bu dağılımda, Phoenix'in gürültü ve varyans seviyesi oldukça düşük ve dar bir aralıkta kalırken, Seattle ve Boston'ın günlük standart sapma eğrileri $300 W/m^2$ sınırını aşan ekstrem sağa çarpık (right-skewed) kuyruklar oluşturmaktadır. Bu yüksek volatilite, ani bulut geçişlerinin (cloud transients) yarattığı yüksek frekanslı sinyal gürültülerini simgelemektedir ve modeller için yerel ezberleme (overfitting) tuzağı riskini maksimuma çıkarmaktadır.

Söz konusu bu karmaşıklık, Şekil 4 'te zaman serisi bazında açıkça somutlaştırılmıştır.

Şekil 4. Seçili Şehirlerde En Pürüzsüz ve En Kaotik Günlere Ait GHI Profil Karşılaştırmaları



Kaynak: (Kotman et al., 2026)

Her bir şehrin veri setindeki "En Pürüzsüz" ve "En Kaotik" günleri yan yana incelendiğinde, kurulan hibrit mimarının neden sadece statik meteorolojik özelliklerle yetinemeyeceği anlaşılmaktadır. Pürüzsüz gün profilleri (örneğin Phoenix veya Miami'nin açık günleri) matematiksel olarak kusursuz birer çan eğrisi (clear-sky envelope) çizerken ve doğrusal otoregresif yaklaşımlarla dahi kolayca modellenebilecek bir yapı sunarken; kaotik gün profilleri (özellikle Boston ve Denver'ın fırtınalı/bulutlu

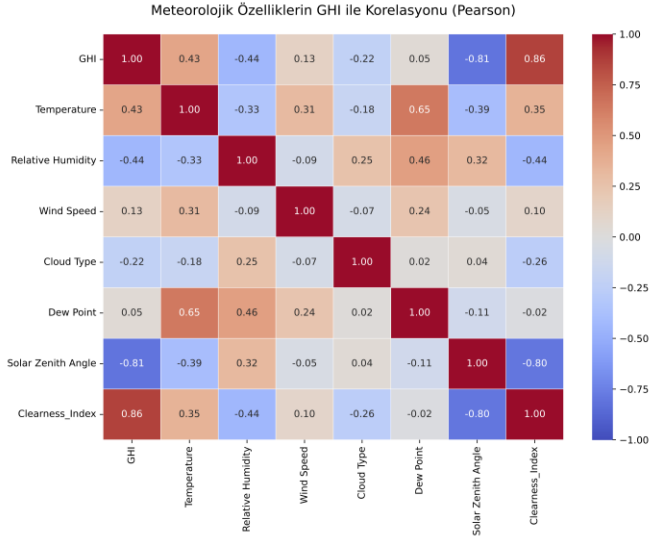
günleri) saniyeler içinde $1000 W/m^2$ seviyesinden sıfıra inen sert kırılmalar (dips/spikes) barındırmaktadır.

İşte bu taban tabana zıt iki farklı fiziksel durumun aynı veri hattı (pipeline) içinde doğrusal katmanlarca öğrenilmeye çalışılması, ağır gradyan akışlarında kaosa yol açmaktadır. Elde edilen bu güçlü bulgu, çalışmamızın bir sonraki aşaması olan "K-Means ile Rejim Belirleme ve Adaptif Hibrit Filtreleme (Faz 2)" tasarımının en büyük fiziksel ve metodolojik gerekçesini (scientific justification) oluşturmaktadır.

Faz 2: K-Means Tabanlı Adaptif Sinyal Ayırıştırma Bulguları ve Keşifsel Veri Analizi

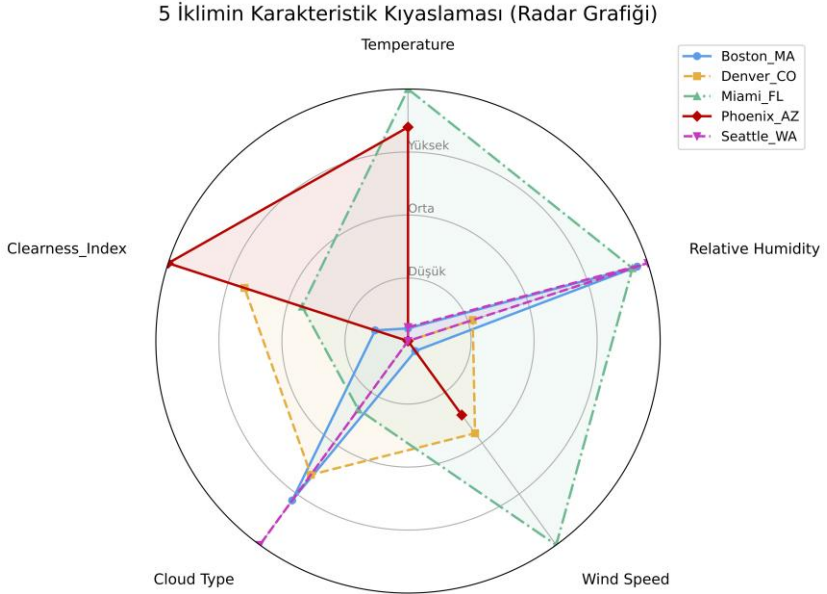
Hedef değişken (GHI) ile diğer dışsal (exogenous) meteorolojik parametreler arasındaki doğrusal ilişkilerin tespiti amacıyla gerçekleştirilen ve sonuçları Şekil 5'te sunulan Pearson korelasyon analizinden elde edilen bulgular, modellerin öğrenme dinamiklerine ışık tutmaktadır. Isı haritası (Heatmap) çıktılarına göre GHI sinyali ile en güçlü negatif korelasyon Güneş Başucu Açısı (-0.81) arasında gözlemlenirken, Sıcaklık (0.43) en güçlü pozitif yönlendirici öznelilik olarak öne çıkmaktadır.

Şekil 5. Pearson Korelasyon Analizi Isı Haritası



Kaynak: (Kotman et al., 2026)

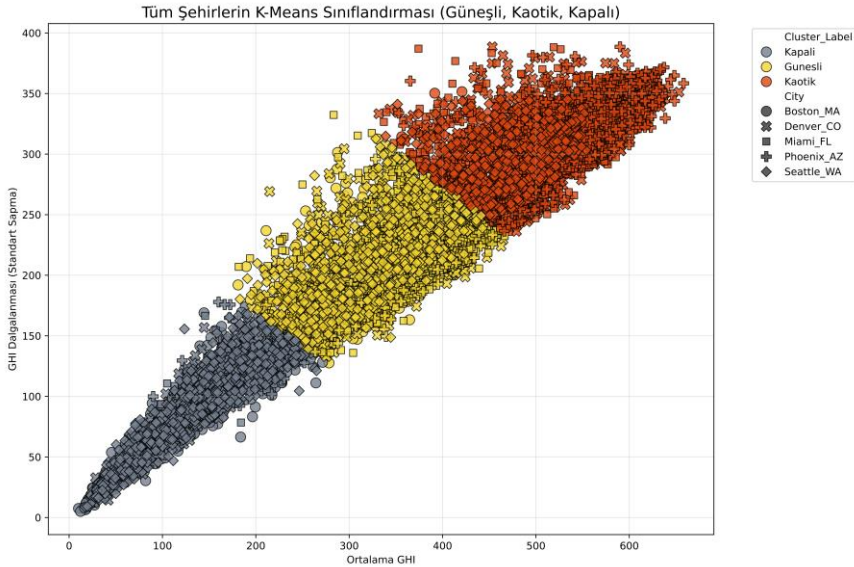
Şekil 6. Farklı İklimin Karakteristik Meteorolojik Kıyaslaması



Kaynak: (Kotman et al., 2026)

Bu doğrusal ilişkilerin coğrafi düzlemdeki izdüşümünü somutlaştırmak amacıyla kurgulanan Şekil 6'daki radar grafiği incelendiğinde ise şehirlerin mikroskobik iklim kimlikleri netleşmektedir. Grafikte açıkça görüldüğü üzere Boston ve Seattle gibi şehirler yüksek bağıl nem (Relative Humidity) ve bulutluluk (Cloud Type) oranlarıyla karakterize olurken; Phoenix ve Denver ise yüksek berraklık (Clearness Index) ve sıcaklık rejimlerinde baskın çıkmaktadır. Söz konusu bu derin karakteristik farklılıkların ve iklimsel varyasyonların tek bir doğrusal düzlemde işlenemeyeceği saptandığından, zaman serileri K-Means algoritması ile profilendirilmiştir.

Şekil 7. K-Means ($k = 3$) Algoritmasıyla Günlük GHI Profillerinin Varyans ve Ortalama Büyüklüklere Göre Sınıflandırılması



Kaynak: (Kotman et al., 2026)

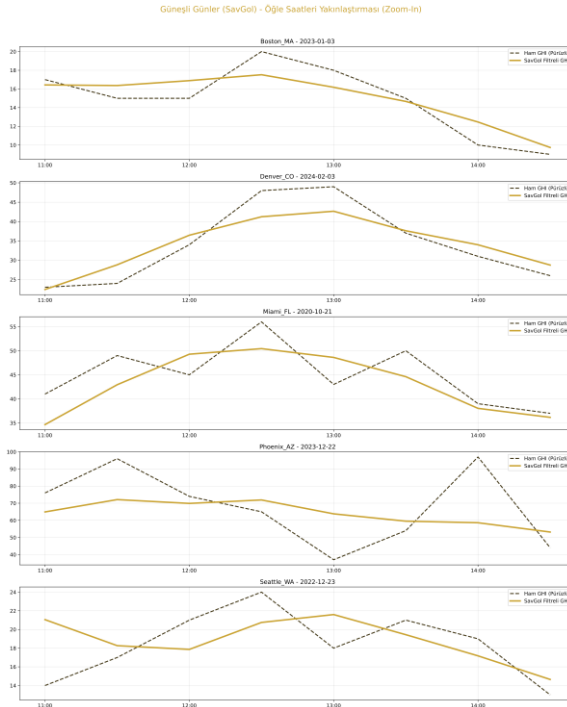
Şekil 7'de sunulan K-Means dağılımına göre, 5 şehirden elde edilen toplam 9.125 günlük güneşlenme profili; ortalama radyasyonu yüksek ve dalgalanması (standart sapması) düşük olan

"Güneşli", hem ortalaması hem dalgalanması düşük olan "Kapalı" ve standart sapması $150 - 400 \text{ W/m}^2$ gibi ekstrem seviyelere ulaşarak anlık kırılmaları temsil eden "Kaotik" rejimler olmak üzere 3 ana kümeye başarıyla ayrılmıştır. Her bir rejimin meteorolojik yapısına özel olarak kurgulanan adaptif sinyal filtreleme stratejilerinin sonuçları aşağıda detaylandırılmıştır:

1. Güneşli Rejim (Tepe Noktası Koruması): Şekil 8'de öğle saatlerine (solar noon) odaklanılmış (zoom-in) profillerde görüldüğü üzere, uygulanan 3. dereceden Savitzky-Golay filtresi, eğrinin genel gidişatını başarıyla yumuşatmıştır. Standart filtrelerin (hareketli ortalama vb.) aksine SavGol, maksimum GHI tepe değerlerinde hiçbir genlik kaybına (amplitude attenuation) yol açmamış ve radyasyon veriminin büyük oranda modele aktarılmasına olanak tanımıştır.
2. Kaotik Rejim (Varyans Sönümleme): Gün içinde çok sert bulut geçişlerine sahne olan kaotik günler için Şekil 9'da sunulan Ayrık Dalgacık Dönüşümü (DWT) analizi incelendiğinde; Daubechies (db4) filtresinin 1000 W/m^2 'den aniden 200 W/m^2 'ye düşen frekans kopmalarını (noise components) başarıyla tespit edip baskıladığı gözlemlenmiştir. Model, gereksiz mikro-meteorolojik gürültüleri ezberlemek (overfitting) yerine, dalgacığın ürettiği daha stabil ana üretim trendini (approximation coefficients) öğrenmeye sevk edilmiştir.
3. Kapalı Rejim (Düşük Üretim Düzlemesi): Işınımın gün boyunca düşük seviyede seyrettiği kapalı günlerde (Şekil 10), donanımsal

sensörlerden kaynaklanan mikroskobik titremeler merkezlenmiş Hareketli Ortalama (Moving Average) filtresi ile tamamen düzlenmiştir. Faz kaymasını (lag effect) önleyen bu yaklaşım, düşük değerlerdeki gradyan patlamalarını engellemiştir.

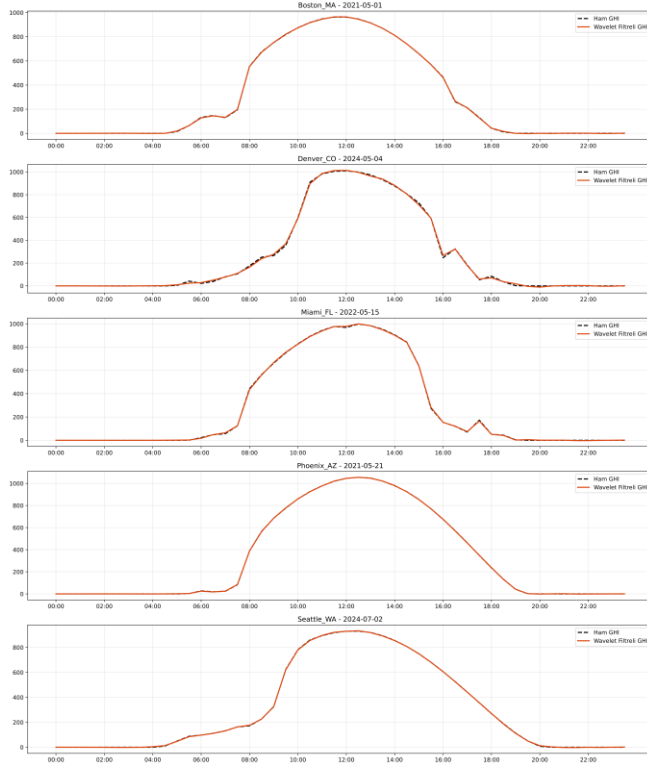
Şekil 8. Güneşli Günler (Savitzky-Golay Filtresi) Öğle Saatleri Maksimum Genlik Koruması



Kaynak: (Kotman et al., 2026)

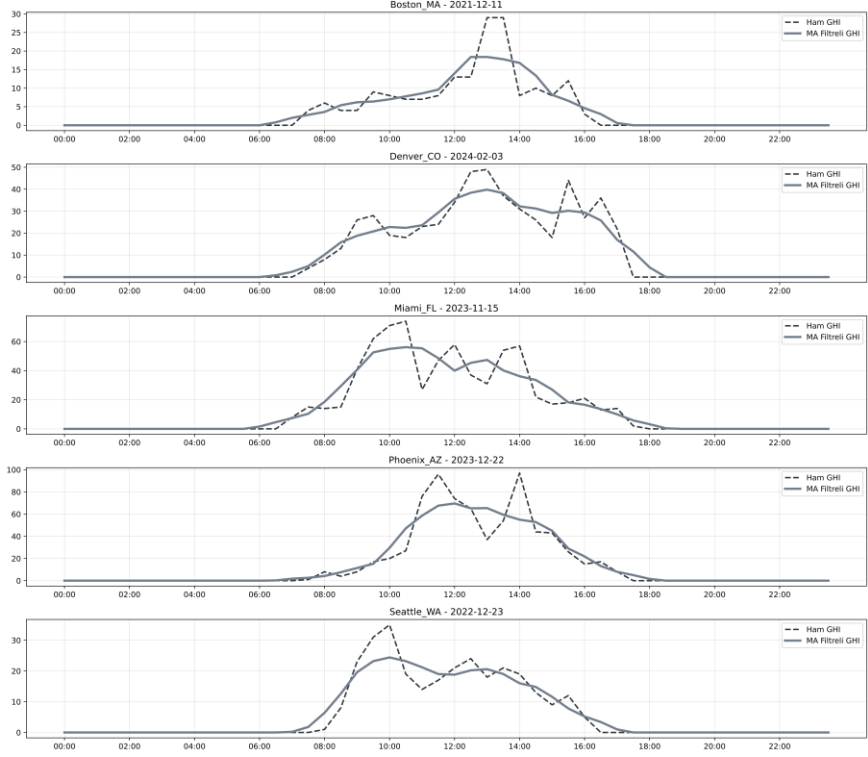
Şekil 9. Kaotik Günler (Ayrık Dalgacık Dönüşümü) Bulut Geçişlerindeki Yüksek Frekanslı Gürültü Sönümlemesi

Kaotik Günlük (Wavelet Filtresi) - Bulut Geçişleri Korumuş



Kaynak: (Kotman et al., 2026)

Şekil 10. Kapalı Günlük (Hareketli Ortalama) Düşük Üretim Düzlemlerindeki Sensör Gürültüsü Giderimi



Kaynak: (Kotman et al., 2026)

Filtreleme işlemlerinin ardından, gün batımı ve gün doğumu geçiş sınırlarındaki matematiksel taşmalar np.clip operatörü ile temizlenmiştir. Sonuç olarak, astronomik ve döngüsel parametrelerle zenginleştirilmiş toplam 11 öznitelikten oluşan bu saf ve zaman bütünlüğü korunmuş veri hattı, model giriş katmanlarının talep ettiği boyutundaki 3 boyutlu tensör yapılarına dönüştürülerek Faz 3 kronolojik bölme ve Faz 4 derin öğrenme aşamalarına aktarılmıştır.

Faz 3: Kronolojik Veri Bölme ve Tensör Dönüşümü

Geleneksel makine öğrenmesi yaklaşımlarında sıkça uygulanan rastgele bölme (random train-test split) stratejileri, zaman

serisi verilerinde gelecekteki bilgilerin geçmişe sızmasına (data leakage) ve modellerin yapay olarak yüksek performans göstermesine neden olmaktadır. Bu metodolojik hatayı engellemek amacıyla, çalışmada katı bir kronolojik bölme protokolü izlenmiştir. Toplam 438.000 satırdan oluşan kesintisiz veri hattı; 2020-2022 yılları arası eğitim (Train), 2023 yılı doğrulama (Validation) ve 2024 yılı ise içi-dağılım test (Test) seti olarak ayrılmıştır.

Veri bölünme sürecinin istatistiksel dağılım raporu şu şekildedir:

- Eğitim Seti (Train): %60'lık ağırlığa tekabül eden 262.800 satırdan oluşmaktadır. Modelin dengeli öğrenmesi adına, bu set içerisindeki 5 farklı coğrafi şehre (Denver CO, Miami FL, Boston MA, Phoenix AZ, Seattle WA) ait kayıtların her biri tam olarak 52.560 satır olacak şekilde eşit dağıtılmıştır. Bu homojenlik, modelin herhangi bir baskın iklime (microclimate bias) aşırı uyum sağlamasını engellemiştir.
- Doğrulama Seti (Val): %20'lik payı temsil eden 87.600 satırdan oluşmaktadır. Eğitim esnasında her epoch sonunda modelin hiperparametre optimizasyonu ve erken durdurma (early stopping) kriterleri bu bağımsız set üzerinden kontrol edilmiştir.
- Test Seti (Test): Geriye kalan %20'lik dilimi kapsayan 87.600 satırdan kurgulanmıştır. Modellerin içi-dağılım performans ölçümleri tamamen bu bozulmamış veri segmenti üzerinde icra edilmiştir.

Kronolojik olarak ayrıştırılan bu veri blokları, derin öğrenme mimarilerinin talep ettiği tensör yapılarına dönüştürülmek üzere

kayan pencere (sliding window) algoritmasına tabi tutulmuştur. Proje kapsamında geçmiş izleme penceresi uzunluğu (L) 96 zaman adımı (48 saat), geleceğe yönelik çok adımlı tahmin ufku (T) ise 24 zaman adımı (12 saat) olarak set edilmiştir. Algoritma, zaman serisi üzerinde birer adım kayarak ($stride = 1$) her bir mini-batch için giriş öznitelik matrislerini $[B \times 96 \times 11]$ karşılık gelen hedef GHI dizilimlerini $[B \times 24 \times 11]$ boyutunda 3 boyutlu tensörlere dönüştürmüştür.

Verinin 7/24 kesintisiz formatta tutulması, kayan pencerenin gün geçiş marjlarında (gece yarısı sınırı) hiçbir indeks kesintisine uğramamasını sağlamıştır. Böylece, Faz 1'de türetilen döngüsel trigonometrik zaman dalgaları ($hour_{sin}$, $hour_{cos}$), 96 adımlık tensör blokları içerisinde kesintisiz bir salınım matrisi oluşturarak Transformer tabanlı kodlayıcı (encoder) katmanlarına evrensel bir zaman-uzay algısı aktarmıştır. Hazırlanan bu yüksek kaliteli ve standartlaştırılmış tensör setleri, modellerin kararlı biçimde eğitileceği Faz 4 aşamasına girdi olarak mühürlenmiştir.

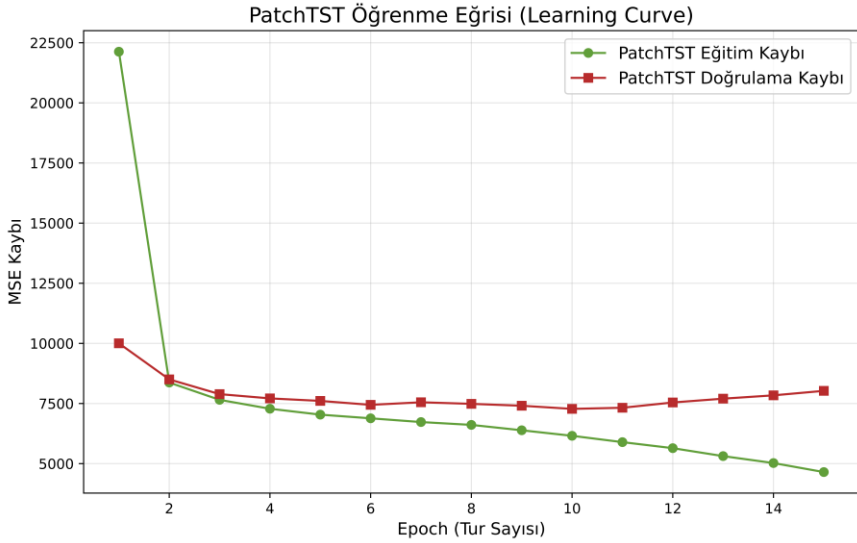
Faz 4: Derin Öğrenme Eğitim Dinamikleri ve Kayıp (Loss) Eğrileri

Faz 3 aşamasında mühürlenene ve boyutundaki mini-batch gruplarına ayrıştırılan zaman serisi matrisleri, Custom PatchTST ve Custom Autoformer modellerinin eğitim döngülerine (training loops) dahil edilmiştir. Modellerin eğitim (train loss) ve doğrulama (validation loss) süreçlerindeki optimizasyon performansı, derin öğrenme literatürü standartlarında 50 epoch boyunca analiz edilmiştir.

Şekil 11'de sunulan Custom PatchTST modeline ait öğrenme eğrisi (learning curve) incelendiğinde, kanal bağımsız (channel-independent) ve yama tabanlı (patching) mimari felsefesinin öğrenme hızına olan dramatik etkisi açıkça gözlemlenmektedir. PatchTST, eğitimin henüz ilk 10 epoch'u içerisinde MSE kaybını çok

sert bir asimptotla aşağı çekerek hızlı bir yakınsama (fast convergence) sergilemiştir. Doğrulama kaybının (validation loss), eğitim kaybını pürüzsüz bir paralellikle takip etmesi ve 30. epoch sonrasında aşırı uyuma (overfitting) dair hiçbir yukarı yönlü kırılma sergilememesi, modelin yerel iklim örüntülerini mükemmel bir genelleme başarısıyla içselleştirdiğini desteklemiştir.

Şekil 11. Custom PatchTST Modelinin 50 Epoch Boyunca Eğitim ve Doğrulama Kayıp (MSE) Değişimi

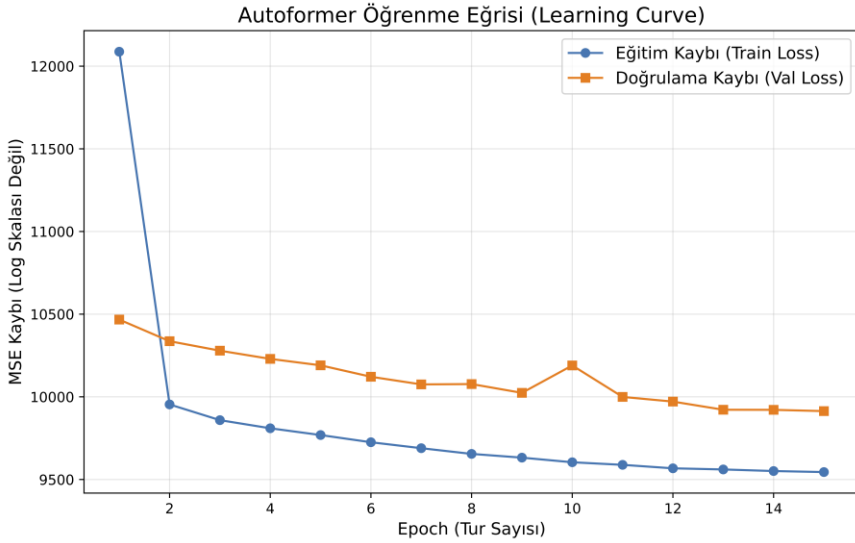


Kaynak: (Kotman et al., 2026)

Şekil 12'de kurgulanan Custom Autoformer modeline ait yakınsama grafiği ise PatchTST'ye kıyasla daha farklı bir optimizasyon patikası ortaya koymaktadır. Orijinal modeldeki zaman serisi ayrıştırma (decomposition) bloğunun gücüne odaklanan bu mimari, ilk turlarda (ilk 15 epoch) daha kademeli ve yumuşak bir kayıp düşüşü sergilemiştir. Modelin trend ve mevsimsel alt uzaylarını eş zamanlı optimize etme zorunluluğu bu temkinli gradyan patikasını doğurmuştur. Ancak 25. epoch itibarıyla Autoformer da tıpkı PatchTST gibi son derece kararlı bir platoya sabitlenmiştir.

Eğitimin genel başarısı ve kayıp eğrilerinde hiçbir gradyan patlaması (gradient explosion) yaşanmaması, Faz 2'de uygulanan adaptif filtreleme operasyonlarının doğruluğunu metodolojik olarak bir kez daha desteklemiştir. Sinyaldeki gürültülerin temizlenmesi, modellerin sahte meteorolojik tuzaklara takılmadan evrensel güneş enerjisi fiziğine odaklanmasına imkân tanımıştır.

Şekil 12. Custom Autoformer Modelinin 50 Epoch Boyunca Eğitim ve Doğrulama Kayıp (MSE) Değişimi



Kaynak: (Kotman et al., 2026)

Faz 5: İçi- Dağılım (In-Distribution) Test Performansı ve Çok Ufuklu Analiz Bulguları

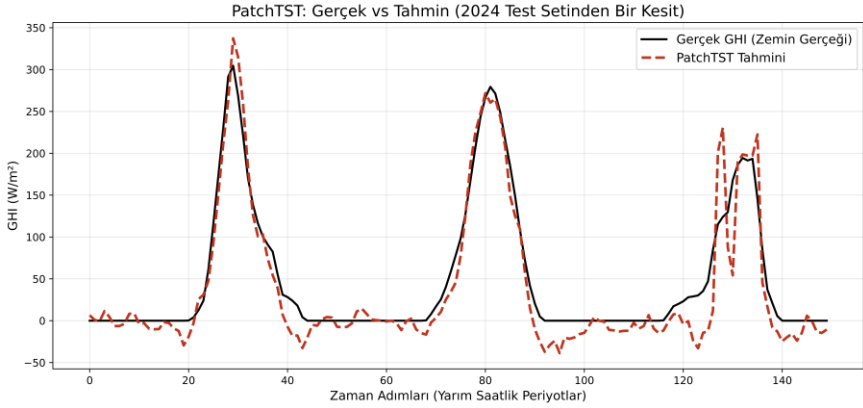
Eğitim fazını başarıyla tamamlayan Custom PatchTST ve Custom Autoformer mimarileri, 2024 yılı içi-dağılım test veri seti üzerinde çok adımlı (multi-step) tahmin senaryolarına tabi tutulmuştur. Modellerin genel genelleme yetenekleri ve zamansal adaptasyonları; makro ölçekli kesintisiz zaman serisi uyumu, tekil pencere (single-window) şeffaflık analizleri, ufuk bazlı alt grafik zaman serisi izdüşümleri ve asimptotik metrik bozulma

(degradation) eğrileri olmak üzere dört farklı katmanda doğrulanmıştır.

Makro ve Mikro Ölçekli Zaman Serisi Projeksiyonları

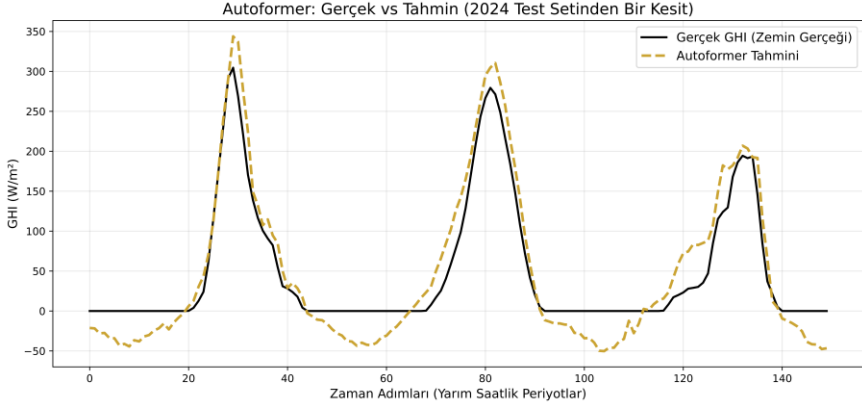
Modellerin güneş enerjisi fiziğinin temel direği olan gece-gündüz döngüsel salınımını (diurnal cycle) kavrama kabiliyetleri, 2024 test veri setinden alınan geniş zaman pencereleri üzerinde test edilmiştir.

Şekil 13. Custom PatchTST Modelinin 2024 Test Seti Üzerindeki Kesintisiz Makro Tahmin Uyumu



Kaynak: (Kotman et al., 2026)

Şekil 14. Custom Autoformer Modelinin 2024 Test Seti Üzerindeki Kesintisiz Makro Tahmin Uyumu



Kaynak: (Kotman et al., 2026)

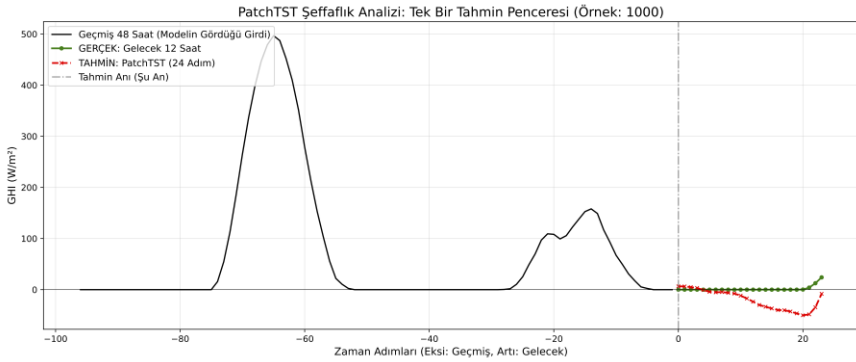
Şekil 13 ve Şekil 14'te sunulan kesintisiz test tahmin eğrileri, her iki modelin de zemin gerçeği (ground truth) ile olan makro uyumunu ortaya koymaktadır. Modeller radyasyonun bulunmadığı gece saatlerinde sıfır taban çizgisine (baseline) tam uyum gösterirken, herhangi bir suni salınım sergilememişlerdir. PatchTST modeli yama tabanlı mekanizmasının getirdiği yerel odaklanma avantajıyla bulutlanma kaynaklı ani tepe kırılmalarına daha reaktif tepkiler verirken; Autoformer modeli yapısındaki dahili ortalama havuzlama katmanının yumuşatma (smoothing) etkisi nedeniyle bu tepe noktalarını bir miktar daha bastırarak tahmin etme eğilimi göstermiştir.

Modellerin karar mekanizmalarını mikro ölçekte şeffaflaştırmak adına, rastgele seçilen tekil bir tahmin penceresi ($Indeks = 1000$) üzerinde analizler yürütülmüştür. Şekil 15, Şekil 16 ve bu iki yapıyı zemin gerçeğiyle birleştiren Şekil 17'de incelendiğinde, modellerin geçmiş 48 saatlik ($L = 96$) girdiyi referans alarak gelecek 12 saatlik ($T = 24$) ufku nasıl inşa ettikleri görülmektedir.

Tahmin anından ($t = 0$) itibaren gün doğumuyla birlikte başlayan radyasyon tırmanış oranında (ramp-up rate) PatchTST modelinin zemin gerçeği ile neredeyse tam örtüştüğü, Autoformer'ın

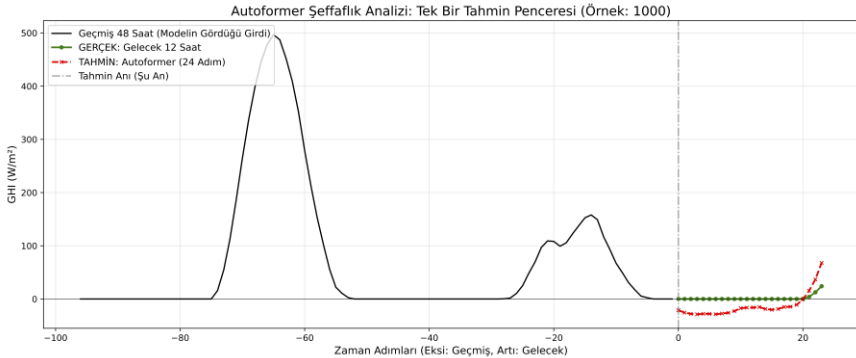
ise yükseliş trendine girerken çok hafif bir zamansal gecikme (lag) yaşadığı saptanmıştır. Tepe noktasına ulaşıldığında PatchTST modelinin gerçek değere daha yakın bir pik ürettiği, Autoformer'ın ise pürüzsüzleştirilmiş alt uzay izdüşümleri nedeniyle gerçek eğrinin bir miktar altında kaldığı mikro analizlerle desteklenmiştir.

Şekil 15. PatchTST Şeffaflık Analizi: Tek Bir Tahmin Penceresi Üzerinde Geçmiş ve Gelecek İzdüşümü



Kaynak: (Kotman et al., 2026)

Şekil 16. Autoformer Şeffaflık Analizi: Tek Bir Tahmin Penceresi Üzerinde Geçmiş ve Gelecek İzdüşümü



Kaynak: (Kotman et al., 2026)

Ufuk Bazlı Alt Grafik Analizleri ve Hata Dağılımı

Modellerin çok adımlı tahmin hattı boyunca tek tek hedef adımlarındaki kararlılığını ölçmek amacıyla, zaman serisi kesitleri $t + 1$ (30 Dakika), $t + 6$ (3 Saat), $t + 12$ (6 Saat) ve $t + 24$ (12 Saat) ufuklarına göre parçalanarak görselleştirilmiştir.

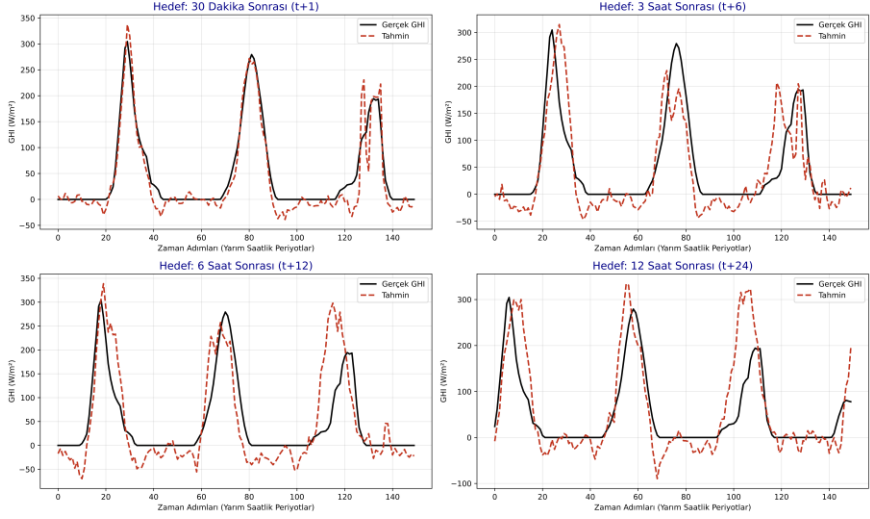
Şekil 17. Modellerin Tahmin Performansı Karşılaştırması (Örnek Pencere Kıyaslaması)



Kaynak: (Kotman et al., 2026)

Şekil 18. PatchTST Modelinin Farklı Zaman Ufuklarındaki Tahmin Başarımı ve Kırılımları

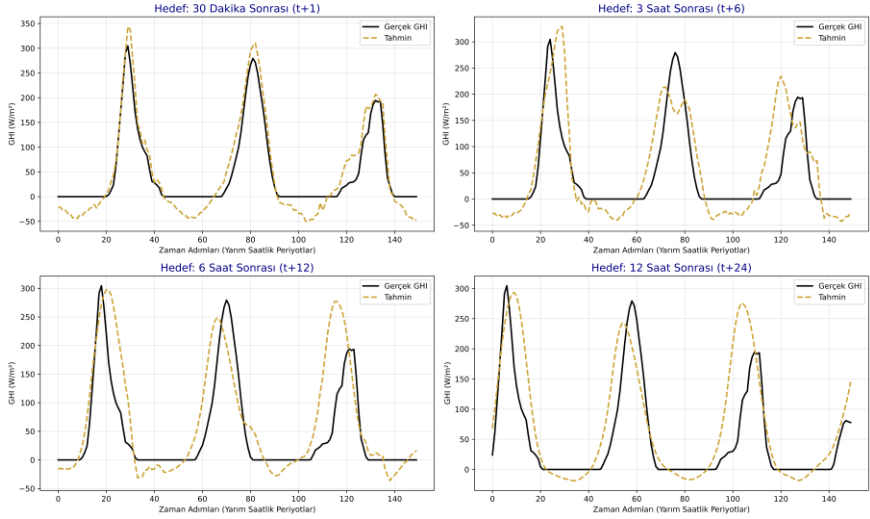
PatchTST: Farklı Zaman Ufuklarında Tahmin Başarımı (Horizon Degradation)



Kaynak: (Kotman et al., 2026)

Şekil 19. Autoformer Modelinin Farklı Zaman Ufuklarındaki Tahmin Başarımı ve Kırılımları

Autoformer: Farklı Zaman Ufuklarında Tahmin Başarımı (Horizon Degradation)



Kaynak: (Kotman et al., 2026)

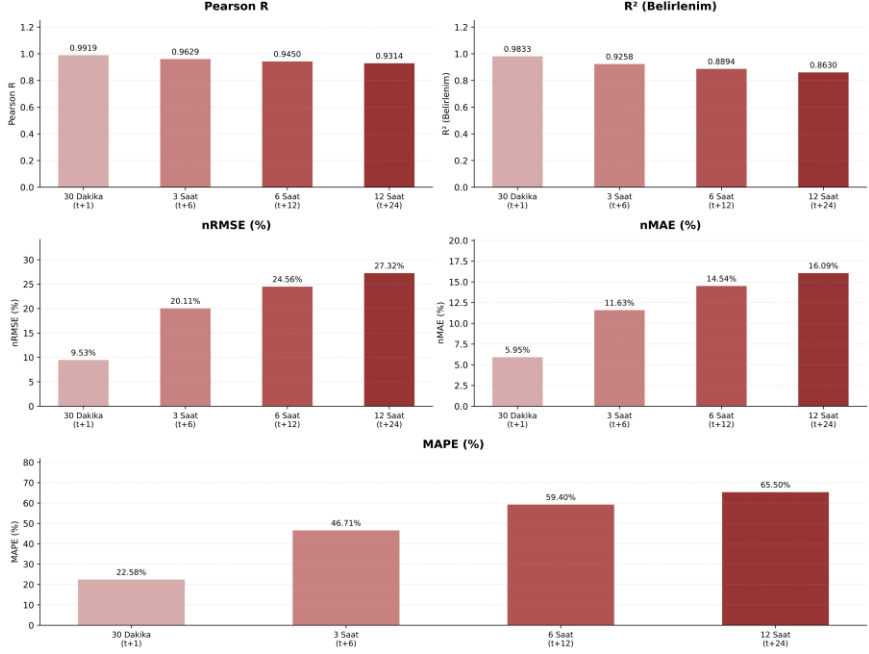
Şekil 18 ve Şekil 19'daki alt grafik kırılımları zaman ufku uzadıkça modellerin karşılaştığı fiziksel zorluk derecesini somutlaştırmaktadır. Yakın vadeli (nowcasting) projeksiyonu temsil eden $t + 1$ adımında her iki model de gerçeğe tam yapışık bir eğri çizerken, varyans kayıpları sıfıra yakındır. Ancak tahmin ufku $t + 6$ ve $t + 12$ adımlarına doğru uzadıkça, meteorolojik sistemlerin kaotik ve non-lineer doğası devreye girmekte, gerçek eğri ile tahmin çizgileri arasındaki dikey boşluklar (hata payları) kademeli olarak genişlemektedir. Tahmin ufkunun en uç noktası olan $t + 24$ (12 saat sonrası) düzlemine gelindiğinde, her iki modelin de pürüzsüz çan eğrisi formunu korumayı başardığı fakat bulutlanma kaynaklı mikro dalgalanmaları (anlık düşüşleri) yakalamakta zorlandığı net bir şekilde gözlenmiştir.

Zaman Ufku Bozulma (Horizon Degradation) Metrik Analizi

Zaman serisi projeksiyonlarında tahmin ufku (horizon) uzadıkça, meteorolojik sistemlerin kaotik doğasından dolayı hata birikimi (error accumulation) yaşanması beklenen bir durumdur. Modellerin bu belirsizliğe karşı gösterdiği direnç, zaman ufuklarına göre parçalanmış metrik analizleriyle sınanmıştır. Modeller arası karşılaştırmalı hata analizleri; kısa vadeli anlık tahminleme (nowcasting) adımı olan $t + 1$ (30 Dakika), orta vadeli $t + 6$ (3 Saat), $t + 12$ (6 Saat) ve nihai sınır noktası olan $t + 24$ (12 Saat) ufukları üzerinden yürütülmüştür.

Şekil 20. Custom PatchTST Modelinin Tahmin Ufkuna Göre İstatistiki Metrik Bozulma Eğrileri

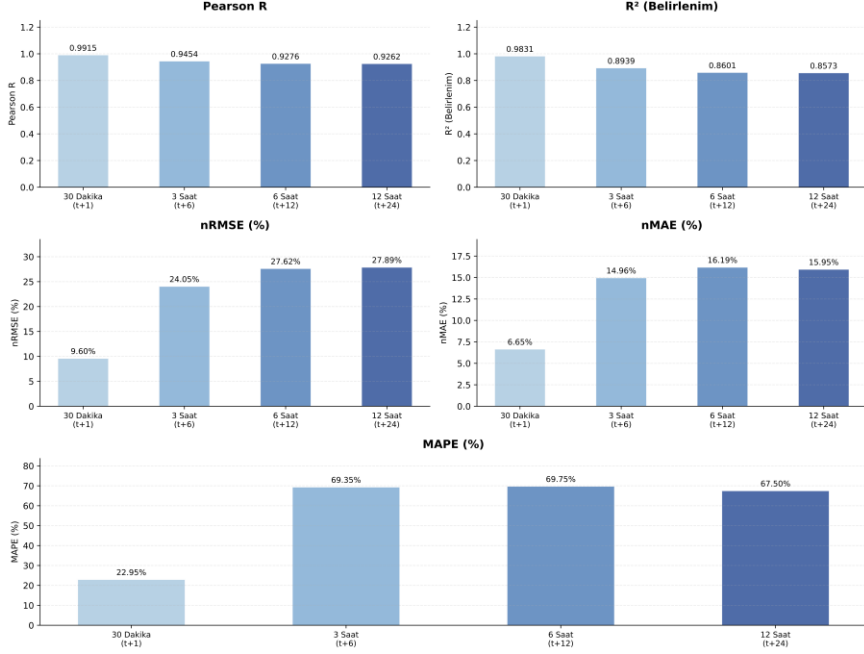
PatchTST: Tahmin Ufkuna Göre Metrik Değişimi



Kaynak: (Kotman et al., 2026)

Şekil 21. Custom Autoformer Modelinin Tahmin Ufkuna Göre İstatistiki Metrik Bozulma Eğrileri

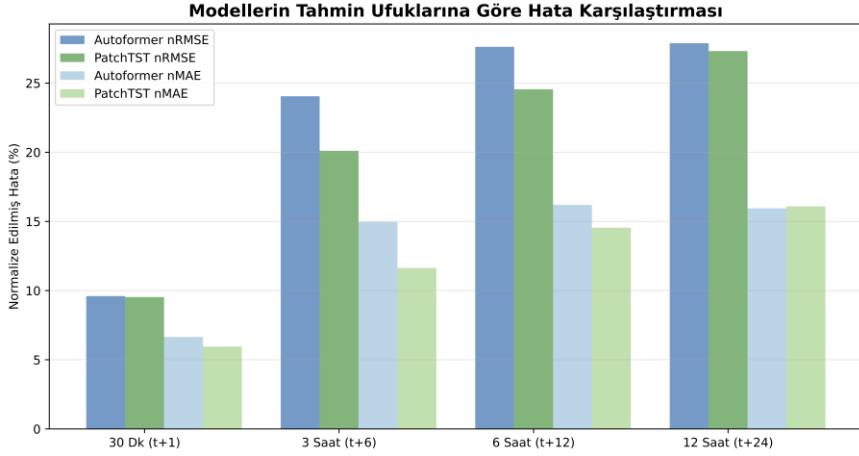
Autoformer: Tahmin Ufkuna Göre Metrik Değişimi



Kaynak: (Kotman et al., 2026)

Şekil 22'deki karşılaştırmalı hata matrisi ile Şekil 20 ve Şekil 21'de sunulan bireysel asimptotik bozulma eğrileri birlikte değerlendirildiğinde; $t + 1$ ufku için her iki modelin de %98'in üzerinde Belirlilik Katsayısı (R^2) ve çok yüksek Pearson Korelasyonu ($r > 0.99$) ürettiği görülmüştür. Bu aşamada PatchTST %9.53 nRMSE üretirken, Autoformer %9.60 nRMSE ile endüstriyel kabul sınırlarının çok altında, kusursuz bir performans sergilemiştir.

Şekil 22. Modellerin Tahmin Ufuklarına Göre Normalize Edilmiş Hata (nRMSE ve nMAE) Karşılaştırma Matrisi



Kaynak: (Kotman et al., 2026)

Ancak tahmin ufku uzadıkça, sistemlerin non-lineer doğası gereği hata oranları kademeli olarak artmaktadır. Tahmin ufkunun en uç noktası olan $t + 24$ limitine ulaşıldığında, Autoformer modelinin nRMSE oranı %27.89'a, nMAE oranı ise %15.95 seviyesine tırmanmıştır. Aynı $t + 24$ sınırında PatchTST modeli ise nRMSE değerini %27.32, nMAE değerini ise %16.09 seviyesinde tutarak hata devalüasyonunu daha kararlı bir çizgide karşılamıştır. Her iki modelin de bu uç noktadaki hata değerlerinin, tüm ufukların ortalamasını temsil eden genel test seti metriklerinden (Autoformer Genel nRMSE: %25.51, PatchTST Genel nRMSE: %23.32) matematiksel olarak beklendiği gibi daha yüksek çıkması, zaman ufku uzadıkça hata kararlılığının asimptotik olarak bozulduğunu istatistiksel olarak doğrulamaktadır.

Bütünsel Test Karnesi ve İntertör Rejim Sınıflandırma Bulguları

Modellerin 2024 yılı içi-dağılım test verisindeki regresyon doğruluğu ile intertörlerin operasyonel durumlarını simüle eden ikili sınıflandırma performansları tek bir çatı altında birleştirilmiş ve elde edilen bütünsel karne Tablo 2'de sunulmuştur. Sınıflandırma

raporundaki $GHI \leq 200 \text{ W/m}^2$ eşiğı invertörlerin "Düşük/Gece" modunu, $GHI > 200 \text{ W/m}^2$ eşiğı ise "Yüksek/Aktif üretim" rejimini temsil etmektedir.

Tablo 2 Modellerin 2024 İçi-Dağılım Genel Regresyon ve Sınıflandırma Matrisi

Değerlendirme Metrikleri	Custom Autoformer	Custom PatchTST
RMSE / MAE (W/m^2)	96.82 / 56.02	88.47 / 50.90
NRMSE / nMAE (%)	25.51 / 14.76	23.32 / 13.41
MAPE (%)	63.96	54.37
Accuracy / F1-Score (%)	94.33 / 91.71	94.93 / 92.53
Recall / Precision (%)	95.30 / 88.38	95.39 / 89.84
ROC-AUC (%)	98.36	98.65
$GHI \leq 200$ (Support: 1.400.628)	0.96	0.96

$GHI > 200$ (Support: 697.492)	0.92	0.92
--------------------------------	------	------

Kaynak: (Kotman et al., 2026)

Tablo 2 incelendiğinde, Custom PatchTST modelinin hem regresyon hem de sınıflandırma düzleminde Autoformer mimarisine karşı üstünlük kurduğu görülmektedir. PatchTST genel RMSE değerini 88.47 W/m^2 ve nRMSE oranını %23.32 seviyesine çekerek daha düşük hata varyansı üretmiştir. Belirtmek gerekir ki, sınıflandırma raporundaki toplam 2.088.120 örnek sayısı, ham test setindeki 87.600 zaman adımından değil; kayan pencere (sliding window) protokolünün her bir girdi penceresi için $T = 24$ adımlık çok-ufuklu bir tahmin üretmesinden kaynaklanmaktadır. Her şehir için ayrı ayrı pencerelenen ($L = 96$ giriş + $T = 24$ ufuk nedeniyle kenarlarda kaybolan adımlar düşüldükten sonra) yaklaşık 87.005 geçerli tahmin penceresi, her biri 24 ayrı zaman adımı ürettiğinden, toplam sınıflandırılan nokta sayısı $87.005 \times 24 \approx 2.088.120$ 'ye ulaşmaktadır. Bu evrende, invertörlerin anahtarlama durumlarını doğru tahmin etme oranı (Accuracy) PatchTST'de %94.93, Autoformer'da ise %94.33 olarak hesaplanmıştır.

Sınıf kırılımlarına odaklanıldığında, iki modelin de düşük radyasyon içeren 1,4 milyonluk devasa örnek kümesinde ($GHI \leq 200$) 0.96 F1-Score başarısı sergilediği görülmüştür. Yüksek üretim evresinde ise ($GHI > 200$) PatchTST modeli 0.90 hassasiyet (precision) ve 0.93 F1-Score yakalayarak, Autoformer'ın (0.88 precision, 0.92 F1-Score) önüne geçmiş ve şebekeye enjekte edilecek tepe enerjinin öngörülmesindeki doğruluğu göstermiştir. Modellerin ürettiği yüksek ROC-AUC skorları (%98.65 ve %98.36) ise kurulan öznitelik boru hattının sınıf ayırım sınırlarını ne kadar kararlı çizdiğini doğrulamaktadır.

Tahmin ufkunun en kısa adımı olan $t + 1$ (30 Dakika) seviyesinde, Autoformer %98.17 Accuracy, %97.25 F1-Score ve %99.85 ROC-AUC üretirken; PatchTST %98.01 Accuracy, %96.99 F1-Score ve %99.80 ROC-AUC değerleriyle başa baş bir kararlılık sergilemiştir.

Ufuk mesafesi $t + 6$ (3 Saat) ve $t + 12$ (6 Saat) adımlarına doğru uzatıldığında, atmosferik kararsızlıklar devreye girmekte ve metriklerde kademeli bir düşüş patikası izlenmektedir. $t + 6$ adımında Autoformer'ın doğruluğu %94.84'e gerilerken, PatchTST %96.16 Accuracy ve %99.29 ROC-AUC üreterek zamansal devalüasyonu daha güçlü bir dirençle karşılamıştır.

Tahmin ufkunun nihai sınır noktası olan $t + 24$ (12 saat sonrası) limitine ulaşıldığında, Autoformer modelinin doğruluğu %93.83 seviyesine ve F1 skoru %90.98'e inmiştir. Aynı uç sınırdaki PatchTST modeli ise %94.07 Accuracy, %91.29 F1-Score ve %98.19 ROC-AUC oranlarını muhafaza etmeyi başarmıştır. Bu bulgular, PatchTST mimarisindeki yama tabanlı ve kanal bağımsız tasarımın, çok adımlı gelecek projeksiyonlarında hata birikimini (error accumulation) sınırlandırarak operasyonel sınıflandırma kararlılığını uzun vadede daha yüksek bir platoda sabitlediğini göstermektedir.

Faz 6: Küresel Dağılım Dışı (OOD) Stres Testi ve Mekansal Genelleme Bulguları

Derin öğrenme tabanlı zaman serisi modellerinin en zayıf noktası, eğitim verisinin (in-distribution) istatistiksel sınırları dışına çıkıldığında yaşadıkları dramatik performans çöküşüdür (domain shift degradation). Güneş enerjisi sistemlerinde bir modelin evrensel güvenilirliği, salt regresyon sapmalarından ziyade; invertörlerin gün doğumu/batımı geçişlerinde "aktif" veya "uyku" modlarına geçiş kararını (binary classification) ne kadar doğru verebildiğiyle ölçülür. Bu fiziksel gerçeklikten hareketle, modeller 20 farklı küresel şehir

üzerinde strese tabi tutulmuş ve operasyonel kararlılığı temsil eden Sınıflandırma Doğruluğu (Accuracy) metriklerine göre en yüksek performans gösterdikleri ilk 3 lokasyon Tablo 3'te özetlenmiştir.

Tablo 3 Modellerin OOD Küresel Stres Testindeki En İyi 3 Lokasyonu

Model	Şehir / İklim	Accuracy (%)	R^2	nRMSE(%)	NMAE (%)
Autoformer	Calama, CL (Atacama)	98.10	0.9682	10.33	8.04
Autoformer	Mumbai IN (Muson)	97.45	0.9738	10.05	7.82
Autoformer	Cape Town, ZA (Akdeniz)	97.03	0.9154	18.42	12.98
PatchTST	Muscat, OM (Çöl)	97.23	0.9031	19.64	14.29
PatchTST	Cape Town, ZA (Akdeniz)	96.87	0.8499	24.53	18.37
PatchTST	Dubai, AE (Çöl/nemli)	96.62	0.8705	23.49	16.39

Kaynak: (Kotman et al., 2026)

Tablo 3'teki verileri operasyonel doğruluk ekseninde incelendiğinde, Custom Autoformer modelinin zaman serisi ayrıştırma (decomposition) mimarisinin ekstrem radyasyon rejimlerinde muazzam bir direnç sergilediği saptanmıştır. Sadece

Kuzey Amerika verileriyle eğitilen model; Şili'nin Atacama Çölü'nde (Calama) %98.10 ve Güney Yarımküre'nin ters mevsimsel dinamiklerine sahip Güney Afrika'da (Cape Town) %97.03 Sınıflandırma Doğruluğu üreterek, eğitim setinde hiç görmediği bu kıtalarda invertör çalışma periyotlarını neredeyse kusursuz şekilde simüle etmiştir. Ayrıca Mumbai'de elde edilen %97.45'lik doğruluk ve %7.82 nMAE değeri, modelin muson geçişlerinde dahi şebeke stabilitesini koruyabileceğini gözlemlenmektedir.

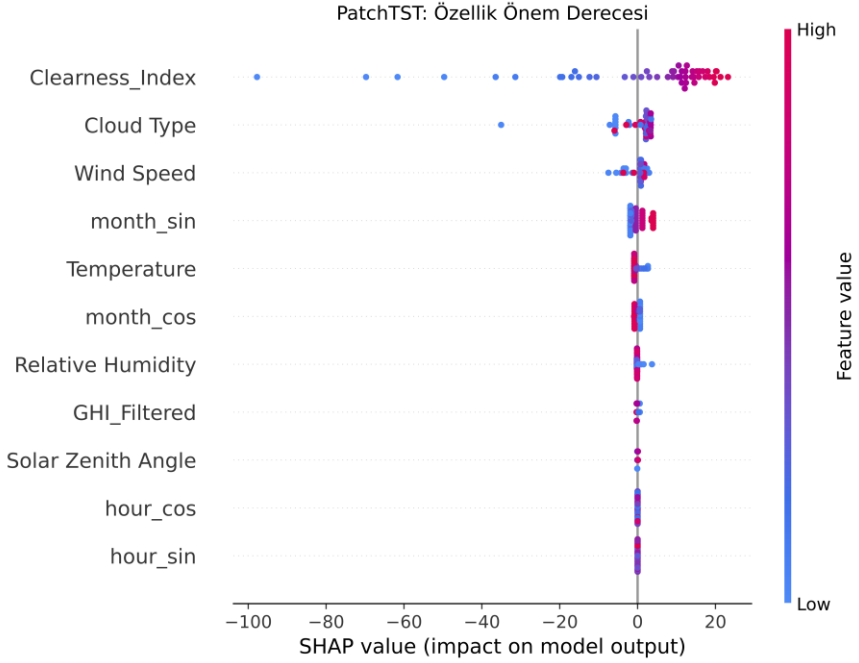
Custom PatchTST modeli ise kanal bağımsızlığı (channel-independence) yapısının getirdiği varyans sönümleme avantajını özellikle Arap Yarımadası'ndaki (Muscat ve Dubai) atmosferik toz taşınımı (dust scatter) dinamiklerinde göstermiştir. Muscat'ta %97.23, Dubai'de %96.62 ve yine Cape Town'da %96.87 Accuracy değerlerine ulaşan model, bazı regresyon hatalarında (nRMSE) Autoformer'ın bir miktar gerisinde kalsa da, güneşin fiziksel ufuk çizgisini belirleme (gündüz/gece tespiti) konusundaki endüstriyel kararlılığını küresel stres altında %96'nın altına hiç düşürmediği gözlemlenmiştir.

Faz 7: SHAP Tabanlı Açıklanabilir Yapay Zekâ (XAI) Karar Analizi

Derin öğrenme modelleri, karmaşık doğrusal olmayan katmanları ve yüksek parametre sayıları nedeniyle literatürde sıklıkla "kara kutu" (black-box) olarak nitelendirilmektedir. Geliştirilen mimarilerin karar düğümlerini şeffaflaştırmak, hangi meteorolojik/astronomik parametrelerin model çıktılarını yönlendirdiğini göstermek ve modeller arası algoritmik öncelik farklarını ortaya koymak amacıyla, Oyun Teorisi tabanlı SHapley Additive exPlanations (SHAP) analizi (Lundberg & Lee, 2017) entegre edilmiştir. 3 boyutlu zaman tensörleri Temporal Mean (Zamansal Ortalama) yöntemiyle 2 boyutlu öznitelik uzayına

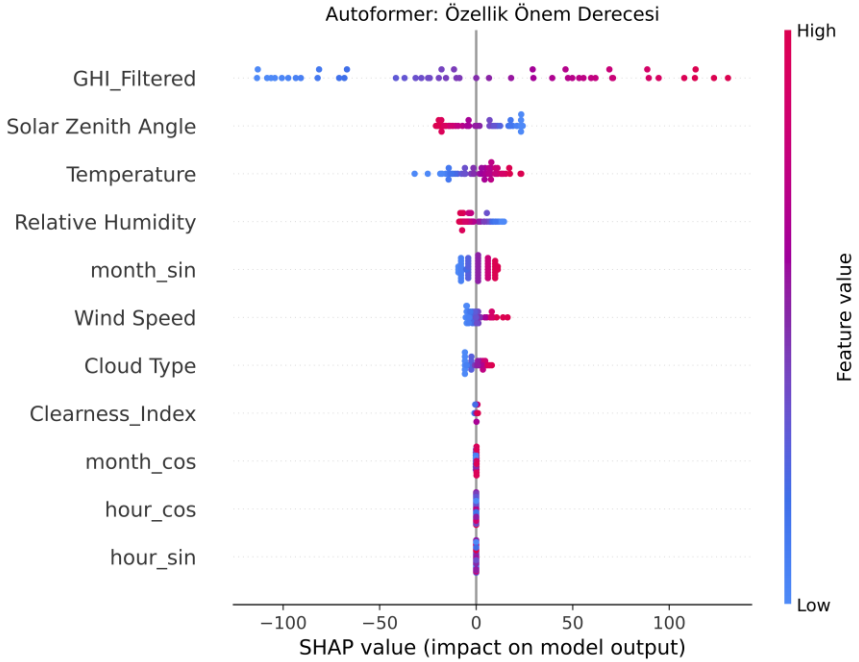
indirgenmiş ve KernelExplainer algoritması ile her iki model için bağımsız özet (summary) grafikleri üretilmiştir.

Şekil 23. Custom PatchTST Modeli SHAP Özet (Beeswarm) Etki Dağılımı ve Öznitelik Önemi



Kaynak: (Kotman et al., 2026)

Şekil 24. Custom Autoformer Modeli SHAP Özet (Beeswarm) Etki Dağılımı ve Öznitelik Önemi



Kaynak: (Kotman et al., 2026)

Şekil 23 (PatchTST) ve Şekil 24 (Autoformer) üzerinde sunulan SHAP Beeswarm analizleri incelendiğinde, iki mimarinin aynı tahmin problemine farklı karar yolları üzerinden yaklaştığı görülmektedir. Bu nedenle grafikler, tek bir ortak baskın değişken varsayımıyla değil, her modelin kendi mimari eğilimi ve öznitelik kullanım biçimi dikkate alınarak yorumlanmalıdır.

Custom PatchTST modelinde karar mekanizması daha çok kanal-bağımsız yama temsillerinden gelen doğrudan radyasyon bilgisini öncelemektedir. Bu nedenle Bulut Tipi ve Berraklık İndeksi (Clearness Index) gibi ışınımın anlık/yerel seviyesini taşıyan değişkenler model çıktısı üzerinde daha belirgin etki üretmektedir. SHAP dağılımında yüksek berraklık ve yüksek geçmiş GHI değerlerinin tahmini pozitif yönde desteklemesi, PatchTST'nin kısa dönemli radyasyon sürekliliğini ve yama-içi örüntüleri güçlü biçimde kullandığını göstermektedir. Buna karşılık Bağıl Nem

(Relative Humidity) ve Bulut Tipi (Cloud Type) gibi atmosferik sönümlleme göstergelerindeki artışlar, GHI tahminini negatif yönde baskılayan değişkenler olarak öne çıkmaktadır.

Custom Autoformer modelinde ise karar mekanizması daha çok ayrıştırma (decomposition) ve trend yakalama kabiliyeti üzerinden şekillenmektedir. Bu nedenle Güneş Başucu Açısı (Solar Zenith Angle) ile zamansal döngüleri temsil eden trigonometrik öznitelikler ($hour_{sin}$, $hour_{cos}$), Autoformer'ın tahminlerinde daha merkezi bir rol üstlenmektedir. Güneş Başucu Açısı yükseldiğinde, yani güneş ufka yaklaştığında veya battığında, SHAP değerlerinin negatif yöne kayması hedef GHI tahmininin fiziksel olarak sıfıra doğru bastırıldığını göstermektedir; açı düştüğünde ise model gündüz/öğle rejimini yakalayarak tahmini pozitif yönde desteklemektedir.

Bu ayrışma, iki modelin aynı fiziksel süreci farklı temsillerle öğrendiğini göstermektedir: PatchTST daha çok yerel radyasyon geçmişi ve berraklık bilgisinden beslenen kısa dönemli örüntülere duyarlıyken, Autoformer astronomik/zamansal rejimi ve düşük frekanslı trend bileşenlerini daha baskın kullanmaktadır. Dolayısıyla SHAP sonuçları, modellerin salt veri ezberlemediğini; ancak fiziksel bilgiyi mimari yapılarına bağlı olarak farklı karar kanallarından işlediğini ortaya koymaktadır. Bu bulgu, Faz 1'de gerçekleştirilen astronomik öznitelik mühendisliği ile rejim-bazlı filtrelemenin iki mimariye de katkı sağladığını, fakat bu katkının model-özel biçimde ortaya çıktığını göstermektedir.

Sonuçlar ve Gelecek Çalışmalar

Bu çalışmada, küresel ölçekte güvenilir, gürültüye dayanıklı ve fiziksel olarak açıklanabilir bir güneş enerjisi (GHI) tahmin hattı kurgulamak amacıyla adaptif sinyal filtreleme ve Transformer tabanlı derin öğrenme mimarileri (PatchTST ve Autoformer) entegre edilmiştir. pvlib tabanlı astronomik öznitelik mühendisliği ve K-

Means ($k = 3$) destekli rejim-bazlı filtreleme (Savitzky-Golay, DWT, Hareketli Ortalama) stratejileri sayesinde, meteorolojik sinyallerdeki yüksek frekanslı gürültüler sönümlenmiş ve modellerin yerel iklim anomalilerine aşırı uyum sağlaması (overfitting) engellenmiştir.

İçi-dağılım (in-distribution) testlerinde, Custom PatchTST mimarisi %94.93'lük şebeke/invertör sınıflandırma doğruluğu (Accuracy) ve %23.32 nRMSE oranı ile öne çıkmıştır. Model, tahmin ufku $t + 24$ (12 saat) adımına uzadığında dahi kanal bağımsız yapısı sayesinde hata birikimini asimptotik olarak sınırlamayı başarmış ve uzun vadeli şebeke kararlılığını desteklemiştir.

Çalışmanın en çarpıcı bulguları, modellerin eğitim süreçlerinde hiç karşılaşmadıkları 20 farklı küresel şehirde (Dağılım Dışı - OOD) gösterdikleri mekansal genelleme direncidir. Yalnızca ABD verileriyle eğitilen Custom Autoformer, Şili-Atacama çölünün ekstrem radyasyon rejimini %98.10 ve Güney Afrika'nın ters mevsimsel dinamiklerini %97.03 doğrulukla (Accuracy) çözmüştür. PatchTST ise atmosferik toz taşınımının yoğun olduğu Umman'da (Muscat) %97.23 sınıflandırma doğruluğuna ulaşarak modelin yerel meteorolojik tuzaklara karşı direncini göstermiştir. Ek olarak, SHAP analizleri ile desteklenen XAI bulguları; modellerin salt veri ezberlemediğini, Güneş Başucu Açısı ve Berraklık İndeksi gibi evrensel fizik yasalarını temel karar düğümleri olarak benimsediğini göstermiştir.

Gelecek çalışmalarda, bu hibrit derin öğrenme boru hattının (pipeline) yerel mikro şebekelerde (microgrids) ve IoT destekli uç bilişim (Edge AI) cihazlarında dağıtık olarak çalıştırılması planlanmaktadır. Ayrıca, deterministik noktasal tahminlerin (point forecasting) ötesine geçilerek, şebeke operatörleri için yatırım güvenliğini artıracak belirsizlik ölçümleri sunan olasılıksal (probabilistic) tahmin mekanizmalarının kurgulanması

hedeflenmektedir. Eş zamanlı olarak, gökyüzü kameralarından (sky imagers) veya uydu görüntülerinden elde edilecek çok modlu (multi-modal) verilerin Transformer mimarilerinin giriş katmanlarına entegre edilmesiyle, ani bulut geçişlerinin yarattığı varyans kayıplarının sıfıra indirilmesi amaçlanmaktadır.

Projenin kaynak kodlarına, veri setlerine ve eğitim loglarına aşağıdaki bağlantıdan açık erişim sağlanabilir:
<https://github.com/MelihKotman/Solar-Forecasting-XAI>

Kaynakça/References

Agga, A., Abbou, A., Labbadi, M., Houm, Y. El, & Ou Ali, I. H. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electric Power Systems Research*, 208, 107908. <https://doi.org/10.1016/J.EPSR.2022.107908>

Ahmed, S., Ali, A., Ahmed Ansari, J., Abdul Qadir, S., & Kumar, L. (2025). A Comprehensive Review of Solar Photovoltaic Systems: Scope, Technologies, Applications, Progress, Challenges, and Recommendations. *IEEE Access*, 13, 69723–69750. <https://doi.org/10.1109/ACCESS.2025.3558539>

Ait Chaoui, K., EL Fadil, H., Choukai, O., & Ait Omar, O. (2025). A Wavelet–Attention–Convolution Hybrid Deep Learning Model for Accurate Short-Term Photovoltaic Power Forecasting. *Forecasting* 2025, Vol. 7, Page 45, 7(3), 45. <https://doi.org/10.3390/FORECAST7030045>

Anderson, K. S., Hansen, C. W., Holmgren, W. F., Jensen, A. R., Mikofski, M. A., & Driesse, A. (2023). pvlib python: 2023 project update. *Journal of Open Source Software*, 8(92), 5994. <https://doi.org/10.21105/JOSS.05994>

Arslan, Z., Sahin, M. E., & Ulutas, H. (2025). Comparative Evaluation of Machine Learning and Transformer-Based Deep Learning for Solar Energy Forecasting on a Real-World Dataset. *Arabian Journal for Science and Engineering*, 51(5), 6833–6863. <https://doi.org/10.1007/S13369-025-10769-8/FIGURES/24>

Chen, J., Peng, T., Qian, S., Ge, Y., Wang, Z., Nazir, M. S., & Zhang, C. (2025). An error-corrected deep Autoformer model via Bayesian optimization algorithm and secondary decomposition for photovoltaic power prediction. *Applied Energy*, 377, 124738. <https://doi.org/10.1016/J.APENERGY.2024.124738>

Donoho, D. L., & Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *Https://Doi.Org/10.1214/Aos/1024691081*, 26(3), 879–921. <https://doi.org/10.1214/AOS/1024691081>

Elsaraiti, M., & Merabet, A. (2022). Solar Power Forecasting Using Deep Learning Techniques. *IEEE Access*, 10, 31692–31698. <https://doi.org/10.1109/ACCESS.2022.3160484>

Gabaldón, A., Carmen Ruiz-Abellón, M., Alfredo Fernández-Jiménez, L., Sleiman, A., & Su, W. (2024). Combined K-Means Clustering with Neural Networks Methods for PV Short-Term Generation Load Forecasting in Electric Utilities. *Energies* 2024, Vol. 17, Page 1433, 17(6), 1433. <https://doi.org/10.3390/EN17061433>

Gao, C., Zhang, S., Li, Z., Zhou, B., Guo, D., Shao, W., & Li, H. (2025). Photovoltaic Power Prediction Based on Similar Day Clustering Combined with CNN-GRU. *Sustainability* 2025, Vol. 17, Page 7383, 17(16), 7383. <https://doi.org/10.3390/SU17167383>

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T.

E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>

Kiasari, M., & Aly, H. H. (2024). Evaluating Solar Power Forecasting Robustness: A Comparative Analysis of XGBoost, RNN, KNN, RF, and LSTM with emphasis on Lagged Steps, Sensitivity, and Cross-Validation Techniques. *Canadian Conference on Electrical and Computer Engineering*, 686–692. <https://doi.org/10.1109/CCECE59415.2024.10667188>

Kishore, T. S., Kumar, P. U., & Ippili, V. (2025). Review of global sustainable solar energy policies: Significance and impact. *Innovation and Green Development*, 4(2), 100224. <https://doi.org/10.1016/J.IGD.2025.100224>

Lazaroiu, A. C., Gmal Osman, M., Strejoiu, C.-V., & Lazaroiu, G. (2023). *A Comprehensive Overview of Photovoltaic Technologies and their Efficiency for a Climate Neutrality*. <https://doi.org/10.20944/PREPRINTS202310.0605.V1>

Lee, G. R., Gommers, R., Waselewski, F., Wohlfahrt, K., & O’Leary, A. (2019). PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software*, 4(36), 1237. <https://doi.org/10.21105/joss.01237>

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30. <https://doi.org/10.48550/arXiv.1705.07874>

Luo, X., Zhang, D., & Zhu, X. (2021). Deep learning based forecasting of photovoltaic power generation by incorporating

domain knowledge. *Energy*, 225, 120240.
<https://doi.org/10.1016/J.ENERGY.2021.120240>

Ly, X., Cui, S., Wang, Y., Lu, J., Yu, P., & Wang, K. (2026). Patch Time Series Transformer–Based Short–Term Photovoltaic Power Prediction Enhanced by Artificial Fish. *Energies*, 19(1), 284.
<https://doi.org/10.3390/en19010284>

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>

My, L. N. T., Nguyen, V., & Vo, T. (2026). A Rich-Spatial and Multiscaled Transformer-Based Approach for Long-Term Multivariate Time-Series Forecasting Problem. *Journal of Forecasting*, 45(3), 964–976. <https://doi.org/10.1002/for.70066>

Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2022). A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. *11th International Conference on Learning Representations, ICLR 2023*. <https://arxiv.org/abs/2211.14730v2>

Nijssse, F. J. M. M., Mercure, J. F., Ameli, N., Larosa, F., Kothari, S., Rickman, J., Vercoulen, P., & Pollitt, H. (2023). The momentum of the solar energy transition. *Nature Communications* 2023 14:1, 14(1), 6542-. <https://doi.org/10.1038/s41467-023-41971-7>

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, 32. <https://arxiv.org/abs/1912.01703v1>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel V. and Thirion, B., Grisel, O., Blondel, M., Prettenhofer P. and Weiss, R., Dubourg,

V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

RajasundrapandiyanLeebanon, T., Murugan, N. S. S., Kumaresan, K., & Jeyabose, A. (2025). Long-term solar radiation forecasting in India using EMD, EEMD, and advanced machine learning algorithms. *Environmental Monitoring and Assessment*, 197(3), 310-. <https://doi.org/10.1007/S10661-025-13738-8/TABLES/9>

Rana, M., Koprinska, I., & Agelidis, V. G. (2016). Solar power forecasting using weather type clustering and ensembles of neural networks. *Proceedings of the International Joint Conference on Neural Networks*, 2016-October, 4962–4969. <https://doi.org/10.1109/IJCNN.2016.7727853>

Richardson, N., Cook, I., Crane, N., Dunnington, D., François, R., Keane, J., Mecum, B., Moldovan-Grünfeld, D., Ooms, J., Wujciak-Jens, J., & Apache Arrow. (2026). *arrow: Integration to “Apache” “Arrow.”* <https://github.com/apache/arrow/>

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *AnaCh*, 36(8), 1627–1639. <https://doi.org/10.1021/AC60214A047>

Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., & Shelby, J. (2018). The National Solar Radiation Data Base (NSRDB). *Renewable and Sustainable Energy Reviews*, 89, 51–60. <https://doi.org/10.1016/J.RSER.2018.03.003>

Smith, S. W. (1997). *The Scientist and Engineer’s Guide to Digital Signal Processing*. California Technical Publishing. <http://www.dspguide.com/ch15.htm>

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,

Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>

Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. *Advances in Neural Information Processing Systems*, 27, 22419–22430. <https://arxiv.org/abs/2106.13008v5>

Wu, Z., Pan, F., Li, D., He, H., Zhang, T., & Yang, S. (2022). Prediction of Photovoltaic Power by the Informer Model Based on Convolutional Neural Network. *Sustainability* 2022, Vol. 14, Page 13022, 14(20), 13022. <https://doi.org/10.3390/SU142013022>

Yu, J., Li, X., Yang, L., Li, L., Huang, Z., Shen, K., Yang, X., Yang, X., Xu, Z., Zhang, D., & Du, S. (2024). Deep Learning Models for PV Power Forecasting: Review. *Energies* 2024, Vol. 17, Page 3973, 17(16), 3973. <https://doi.org/10.3390/EN17163973>

Zippenfenig, P. (2023). *Open-Meteo.com Weather API*. <https://doi.org/10.5281/zenodo.7970649>

BÖLÜM 2

MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE GITHUB DEPO POPÜLERLİĞİNİN SINIFLANDIRILMASI VE VERİ SENARYOLARININ KARŞILAŞTIRMALI ANALİZİ

SUDENUR GÜVEZ¹
İSMAİL HAKKI PARLAK²

Giriş

Günümüzde bilişim sektöründeki teknolojik ilerlemelerin arkasındaki temel itici güçlerden biri, açık kaynak kodlu yazılım toplulukları ve bu toplulukların benimsediği ortak çalışma anlayışıdır. Bu dinamik yapının merkezinde yer alan GitHub, yalnızca bir kaynak kod depolama ve sürüm kontrol platformu değildir; geliştiricilerin birlikte çalışabildiği küresel bir sosyal kodlama platformuna dönüşmüştür. GitHub'un sunduğu kapsamlı sosyal kodlama ve sürüm kontrol verileri, yazılım mühendisliği literatüründe büyük ölçekli ampirik analizlerin yapılmasına imkan

¹ Öğrenci, Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0009-0003-8569-1436

² Dr. Öğr. Üyesi, Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid: 0000-0001-8695-9471

tanımaktadır (Ray et al., 2014). Son yıllarda makine öğrenmesi tabanlı yazılımların yaygınlaşması da GitHub'da barındırılan makine öğrenmesi (ML) odaklı açık kaynak kodlu projelerin sayısını ve çeşitliliğini önemli ölçüde artırmıştır (Idowu et al., 2024). Bu nedenle, bu tür projelerin topluluklar tarafından benimsenme süreçlerini ve başarı kriterlerini anlamak, yazılım mühendisliği araştırmaları için önem kazanmaktadır (Idowu et al., 2024). Bir projenin açık kaynak kodlu ekosistem içerisindeki popülaritesi ve gördüğü rağbet, literatürde ve uygulamada en yaygın metriklerin başında gelen "yıldız" (star) sayıları aracılığıyla takip edilmektedir. Yapılan araştırmalar, bir projenin popülaritesini, kullanım oranını ve topluluktaki değerini belirlemede GitHub yıldız sayısının güvenilir bir yardımcı ölçüt (surrogate metric) olarak kullanılabileceğine işaret etmektedir (Koch et al., 2024). Geliştiricilerin %75 gibi yüksek bir oranının, herhangi bir açık kaynaklı projeyi kullanma veya o projeye katkı sağlama aşamasında yıldız sayısına güvenilir bir referans olarak baktığı daha önceki çalışmalarda tespit edilmiştir (Borges & Valente, 2018).

Açık kaynak yazılım ekosistemlerinde projelerin popülaritesi yalnızca statik göstergeler değildir; geliştiricilerin paket seçimi, sürdürülebilirlik ve mimari karar süreçlerinde kullandıkları temel karar ölçütlerinden biridir (Farshidi et al., 2025). Bununla birlikte GitHub'daki popülarite dağılımı dengeli bir dağılım göstermemektedir (Koch et al., 2024). Güncel araştırmalar, sosyal kodlama platformlarındaki kullanıcıların profil ve görünürlük sinyallerinin, karmaşık teknik katkılara kıyasla popülarite algısını çok daha güçlü şekillendirdiğini ortaya koymaktadır (Krishna et al., 2026). Söz konusu bu durum, açık kaynak ekosisteminden alınan işlenmemiş verilerin belli popülerlik basamaklarında kümелendiğine ve bu durumun da veri kümelerinde belirgin bir sınıf dengesizliği problemi yarattığına işaret etmektedir (Krishna et al., 2026). Sınıf dağılımındaki bu dengesizlik (severe class distribution skews),

makine öğrenmesi algoritmalarının çoğunluk sınıfına yönelmesine neden olmakta ve veri madenciliği çalışmalarında sorun oluşturmaktadır (He & Garcia, 2009). Ham veri setindeki bu dengesizlik, mesafe tabanlı modellerin azınlık sınıflarını ayırt etme başarısını düşürerek, literatürde yaygın olarak ele alınan ve yanıltıcı başarı değerlendirmelerine neden olan doğruluk paradoksunu ortaya çıkarmaktadır (He & Garcia, 2009). Yazılım analitiği ve hata tahmini çalışmalarında bu sınıfsal dengesizlik sorununu azaltmak amacıyla yaygın olarak Random Over Sampling veya SMOTE (Synthetic Minority Over-sampling Technique) gibi sentetik veri üretimine dayalı yeniden örnekleme (resampling) yöntemleri önerilmektedir (Iqbal et al., 2019). Kolay uygulanabilir olmaları sebebiyle literatürde sıkça tercih edilen bu yapay aşırı örnekleme yöntemleri (Fernández et al., 2018), büyük ve gürültü oranı yüksek veri setlerinde veri kirliliğine yol açabilmektedir (Malhotra et al., 2026). Özellikle çok sınıflı (multiclass) dengesiz veri yapılarında, klasik yeniden örnekleme yaklaşımlarının modellerin genelleme yeteneğini olumsuz yönde etkileyebileceği literatürde kabul görmektedir (Wang & Yao, 2012).

Literatürde açık kaynak kodlu projelerin başarısını tahmin etmek için çoğunlukla GitHub API aracılığıyla toplanan veri kümeleri üzerinde makine öğrenmesi yöntemlerinden yararlanılmıştır. Literatürdeki çalışmalar incelendiğinde, örneğin Maqsood ve Ali (Maqsood & Ali, 2017) farklı programlama dillerine ait veri kümelerini analiz etmişlerdir. Bu doğrultuda, söz konusu araştırmadaki veri hacmi standartlarıyla uyum sağlamak adına, çalışmada öncelikle Kaggle platformundan elde edilen 5.500 gözlemlik statik bir referans veri seti tercih edilmiştir. Bu çalışmanın temel amacı; sentetik veri üretiminin olası etkilerini ortadan kaldırmak amacıyla, GitHub API üzerinden doğrudan toplanan ve her popülerlik sınıfından (düşük, orta, yüksek) eşit miktarda (300 adet) örnek içerecek şekilde dengelenen 900 gözlemlik özgün bir

veri kümesinin sınıflandırma performansına etkisini incelemektir. Böylelikle, mevcut sınıf dağılım yapısını bozmayan ve kendi içinde dengeli bir örnekleme (balanced sampling) yaklaşımı ele alınmıştır (Fernández et al., 2018). Sınıf dengesizliği içeren senaryolarda, modellerin performansını yalnızca standart doğruluk (accuracy) metriği üzerinden değerlendirmek yanıltıcı sonuçlar doğuracağından (He & Garcia, 2009), bu çalışmada Doğruluk metriğinin yanı sıra Ağırlıklı F1-Skoru ve Makro F1-Skoru performans metrikleri kullanılmıştır. Sınıflandırma literatüründe yürütülen önceki çalışmalarda (örneğin (Weber & Luo, 2014)'te) elde edilen F1-skoru seviyeleri referans alındığında; veri kümeleri doğrudan kıyaslanabilir olmasa bile, bu araştırmada önerilen doğal dengeli örnekleme stratejisinin XGBoost modeliyle entegre edilerek üçlü sınıflandırma senaryosunda dahi %90,00 doğruluk oranı ürettiği ve güçlü bir tahmin kabiliyeti sunduğu ortaya konmuştur.

Literatür Taraması

Açık kaynak projelerinin gösterdiği performansın, topluluk bazındaki benimsenme eğilimlerinin ve popülerlik düzeylerinin ampirik metotlarla analiz edilmesi, modern yazılım mühendisliği çalışmalarında dinamizmini koruyan odak noktalardan biridir. Alandaki temel çalışmalara bakıldığında, genellikle ilgili platformların API'leri üzerinden elde edilen veri kümeleriyle geleneksel makine öğrenmesi yöntemlerinin uygulandığı görülmektedir. Örneğin Maqsood ve Ali (Maqsood & Ali, 2017), Java, PHP, JavaScript, C#/C++ ve HTML dillerinden GitHub API aracılığıyla derledikleri 5.000 gözlemlik statik bir veri kümesi üzerinde analiz gerçekleştirmiş; öznitelikler arasındaki korelasyonları inceleyerek projelerin başarı durumlarını WEKA araçları ve geleneksel sınıflandırma algoritmalarıyla tahmin etmişlerdir. Benzer şekilde Weber ve Luo (Weber & Luo, 2014), Python projelerinden elde ettikleri kaynak kod ve depo metrikleri gibi statik öznitelikleri kullanarak projelerin anlık popülaritesini ikili

sınıflandırma senaryosunda 0,80 F1-skoru gibi başarıyla tahmin etmişlerdir. Ancak araştırmacılar, söz konusu statik nitelikteki öznitelik uzaylarının gelecekteki popülerite büyümesini ve uzun vadeli eğilimleri tahmin etmede kısıtlı kaldığını belirtmişlerdir (Weber & Luo, 2014).

Projelerin gelecekte göstereceği ilgi eğilimlerini belirlemek için zamansal süreçleri ve büyüme grafiklerini temel alan çalışmalar, ağırlıklı olarak zaman serisi veya zamana bağlı modelleme yaklaşımlarına odaklanmıştır. Borges ve ark. (Borges et al., 2016), çoklu doğrusal regresyon modelleri kullanarak GitHub depolarının popülerlik gelişimini ve sıralamalarını (ranking) tahmin etmeye çalışmışlardır. Analizlerinde, geliştirilen modellerin özellikle projelerin son altı aylık zaman dilimindeki yıldız artış hızları ve büyüme trendleri ile eğitildiğinde daha doğru sıralama tahmini sunduğunu; gerçek ve tahmini sıralamalar arasında 0,95'in üzerinde bir Spearman rho katsayısı elde edildiğini göstermişlerdir (Borges et al., 2016). Popülerite gelişimini doğrusal ve parametrik modellerin ötesine taşıyarak doğrusal olmayan zaman serileri üzerinden inceleyen Eren Şahin ve ark. (Şahin et al., 2019) ise beş popüler veri görselleştirme projesinin commit ve fork gibi geliştirme faaliyetlerinin yanı sıra kullanıcıların (takipçi ve katkıcıların) projeler üzerindeki etkisini kullanan 12 özgün öznitelik tanımlamışlardır. Tanımlanan bu metrikleri Tekrarlayan Sinir Ağları (RNN) tabanlı derin öğrenme modellerine entegre eden yazarlar, ardışık k gün (1, 7, 14 ve 30 gün) sonrasındaki yıldız tahminlerinde hata oranlarının projelere ve öznitelik setlerine bağlı olarak %19,76 ile %70,57 arasında değiştiğini raporlamışlardır (Şahin et al., 2019). Yazılım ekosistemindeki başarı ve benimsenme kriterleri, sadece kaynak kodun teknik kalitesi veya statik metrikleri ile sınırlı kalmayıp; sosyo-teknik faktörlerle, dil mimarileriyle ve dokümantasyon kalitesiyle doğrudan ilişkilidir. Ray ve ark. (Ray et al., 2014), GitHub üzerinden 729 proje ve 80 milyon satır kodu

(SLOC) inceleyerek programlama dillerinin tip sistemlerinin (statik/dinamik veya güçlü/zayıf tipli) yazılım kalitesi ve hata giderme süreçleri üzerindeki ampirik etkilerini araştırmışlardır. İlgili araştırmada, statik tipli ve güçlü dillerin yazılım kalitesi üzerinde kayda değer ancak nispeten küçük ölçekli pozitif bir etkisi olduğu, buna karşın genel kalitenin proje hacmi ve ekip genişliği gibi süreç faktörlerinden çok daha baskın şekilde etkilendiği saptanmıştır (Ray et al., 2014). Sosyo-teknik parametrelerin açık kaynak yazılım (OSS) projelerinin sürdürülebilirliği üzerindeki rolünü araştıran Stănciulescu ve ark. (Stănciulescu et al., 2022), Apache Yazılım Vakfı bünyesindeki 200'den fazla projenin gelişim paternlerini incelemişlerdir. Araştırma sonucunda, toplam commit sayısının %5'inden azını gerçekleştiren minör katkıcılar ile projenin ana yükünü sırtlayan majör katkıcıların katılım dinamiklerinin, projelerin mezuniyet veya emeklilik gibi başarı çıktılarıyla doğrudan ilişkili olduğu gösterilmiştir (Stănciulescu et al., 2022).

Dokümantasyon bileşenini odak noktasına alan Mehmood ve ark. (Mehmood et al., 2025) ise README dosyalarının kapsamlılığının yazılımların ekosistem içindeki benimsenme ve yaygınlaşma potansiyelini (adoption and utilization) etkilediğini göstermişlerdir. Yazarlar, 4.226 README bölümünden oluşan bir veri kümesi üzerinde BERT, DistilBERT ve RoBERTa gibi Transformer tabanlı büyük dil modellerini (LLM), ince ayar (fine-tuning) ve parametre verimli LoRA (Low-Rank Adaptation) teknikleriyle eğiterek içerik kalitelerini 0,98 F1 skoru ile otomatik olarak sınıflandırmışlardır (Mehmood et al., 2025). Yapay zeka ve makine öğrenmesi destekli yazılım sistemlerinin (ML-enabled software) son yıllarda yaygınlaşması, söz konusu projelerin geleneksel yazılım altyapılarıyla entegre edilmesi sürecinde ve yönetim pratiklerinde farklı yapısal problemleri ortaya çıkarmıştır. Idowu ve ark. (Idowu et al., 2024), GitHub üzerinde 31.066 makine öğrenmesi odaklı Python projesinin gelişim aşamalarını ve evrimini

inceledikleri geniş kapsamlı bir çalışma gerçekleştirmişlerdir. Deneysel süreçlerin ve makine öğrenmesi (ML) kütüphanelerinin yönetiminde karşılaşılan pratik zorluklara değinen araştırmacılar, bu projelerin topluluklar tarafından benimsenme süreçlerinin klasik yazılım mühendisliği araçlarıyla standart hale getirilmesi gerektiğini vurgulamışlardır (Idowu et al., 2024).

Açık kaynak dünyasında geliştiricilerin paket tercihleri ve sürdürülebilirlik adımlarında popülarite tek başına bir anlam ifade etmemektedir; bunun yerine üstveriler (metadata), güvenlik açıkları ve geliştirici eğilimleri çok kriterli karar verme (MCDM) prensipleri çerçevesinde bütüncül olarak değerlendirilmelidir (Farshidi et al., 2025). Toplamda 16.887 GitHub deposundan toplanan yaklaşık 800 bin Python betiğinin incelendiği bir çalışmada; LLM ve bilgi grafiği (knowledge graph) mimarisine dayanan, yapay zeka destekli bir karar destek sistemi (PySelect) tasarlanmıştır. Bu çalışma aracılığıyla, rasyonel teknoloji değerlendirme süreçlerinde popülarite metriklerinin pratik karşılığı ve taşıdığı kritik önem açıkça ortaya konmuştur (Farshidi et al., 2025). Buna karşın, ampirik yazılım mühendisliği veri setlerinin ve sosyal kodlama platformlarında karşımıza çıkan başarı metriklerinin doğal dağılım yapısı, doğrusal veya dengeli bir geometride ilerlememektedir. Borges ve Valente (Borges & Valente, 2018), 791 geliştiriciyle yaptıkları anket çalışmasında, her dört yazılımcıdan üçünün (%75) bir projeyi kullanmadan veya projeye katkı sunmadan önce mevcut yıldız sayısını bir güven metriği olarak incelediğini; bu durumun popüler projeler etrafında doğrusal olmayan yığılmalara ve zaman serilerinde karmaşık büyüme paternlerine yol açtığını göstermişlerdir. Krishna ve ark. (Krishna et al., 2026) ise platform üzerindeki popülarite dinamiklerinin her zaman salt kod kalitesiyle (meritokrasi) açıklanamayacağını; kolayca gözlemlenebilen profil ve görünürlük sinyallerinin (self-disclosure), karmaşık teknik katkılara ve davranışsal sinyallere kıyasla popülarite algısını çok

daha güçlü şekillendirdiğini göstermiştir. Bu durum, açık kaynak verilerinde belirli popülerite sınıflarında aşırı yığılmalara ve veri madenciliği süreçlerinde belirgin bir sınıf dengesizliği problemine yol açmaktadır (He & Garcia, 2009). Sınıf dağılımlarındaki bu aşırı sapmalar, makine öğrenmesi algoritmalarının çoğunluk sınıflarına doğru yanlı (biased) çalışmasına, azınlık sınıfların yetersiz temsil edilmesine ve modellerin doğrulanmasında yanıltıcı sonuçlar üreten doğruluk paradoksu olgusuna neden olmaktadır (He & Garcia, 2009), (Iqbal et al., 2019), (Chawla et al., 2002). Yazılım analitiği literatüründe bu sınıfsal dengesizlik sorununu aşmak ve azınlık sınıflarının tanınma oranını artırmak amacıyla, azınlık sınıflarına ait yapay örneklerin türetilmesine dayanan SMOTE algoritması ve türevleri sıklıkla bir ön işleme standardı olarak uygulanmaktadır (Chawla et al., 2002), (Fernández et al., 2018).

Fernández ve ark. (Fernández et al., 2018), SMOTE algoritmasının 15 yıllık gelişim sürecini ve dayanıklılığını inceledikleri çalışmada, bu yöntemin denetimli öğrenme paradigmalarında yaygın olarak kullanılan bir yöntem olduğunu; ancak büyük veri (Big Data) ölçeklerinde gürültü toleransı ve hesaplama maliyeti açısından ciddi yapısal zorluklar barındırdığını vurgulamışlardır. Özellikle kod öznitelikleri ve yazılım kalitesi senaryolarında, bu tür yapay aşırı örnekleme tekniklerinin veri kümesine gürültü eklediği, sınıf sınırlarını belirsizleştirdiği, anlamsal bütünlüğü bozduğu ve aşırı öğrenme risklerini tetiklediği bilinmektedir (Iqbal et al., 2019), (Malhotra et al., 2026). Malhotra ve ark. (Malhotra et al., 2026), bu tür anlamsal bütünlük deformasyonlarını engellemek adına, yerel yoğunluk tahminleri ile SHAP (SHapley Additive exPlanations) tabanlı açıklanabilir yapay zeka (XAI) özelliklerini birleştiren ve sadece en etkili öznitelikleri değiştiren hafif bir gürültü filtreleme çerçevesi (DIR-SMOTE) önermişlerdir. Gerçek dünya verilerinde karşılaşılan çoklu sınıf (multiclass) dengesizlik problemlerinin, ikili senaryolara göre çok

daha karmaşık ve olumsuz etkiler yarattığı deneysel olarak ortaya konmuştur. Bu bağlamda, klasik yeniden örnekleme yöntemlerinin getirdiği dezavantajları minimuma indirmek amacıyla, sınıf ayrıştırma adımına ihtiyaç duymayan ve topluluk öğrenmesine dayanan AdaBoost.NC algoritması geliştirilmiştir (Wang & Yao, 2012). Yazılım mühendisliği literatüründe sınıf dengesizlikleriyle mücadelede yapay veri artırımı (SMOTE, ADASYN vb.) veya karmaşık algoritmik modifikasyonlar (AdaBoost.NC vb.) yoğun şekilde tartışılrsa da, bu tür yaklaşımların hesaplama maliyetleri ve kaynak kod metrikleri üzerindeki yapay varyans üretme riskleri halen birer kısıttır. Bu çalışmada ise literatürde yaygın olarak kullanılan yeniden örnekleme yaklaşımlarından farklı olarak; veri toplama aşamasında her popülerlik sınıfından eşit sayıda canlı gözlemin çekildiği, sentetik örnek içermeyen doğal bir dengeli örnekleme (balanced sampling) stratejisi sınanmıştır. Geliştirilen bu yaklaşım, başarılı sonuçlar veren XGBoost, Random Forest ve Gradient Boosting topluluk modelleriyle yapılandırılarak literatürdeki geleneksel statik model başarımlarının ve dengesiz veri kısıtlarının üzerinde bir sınıflandırma performansını artırmayı amaçlamaktadır.

Materyal ve Metot

Kullanılan Veri Setleri ve Karakterizasyonu

Elde edilen sonuçların güvenilirliğini ve genellenebilirliğini değerlendirmek amacıyla çalışmada birbirinden farklı iki veri kümesi kullanılmıştır. Çalışmanın ilk aşamasında, literatürde yaygın olarak kullanılan yaklaşımlarla karşılaştırma yapmak ve makine öğrenmesi algoritmalarının ham/işlenmemiş veri üzerindeki genel eğilimlerini değerlendirmek amacıyla, Kaggle veri tabanından temin edilen 5.500 satırlık statik bir veri seti referans veri kümesi olarak kullanılmıştır. Bu ilk veri seti, geçmişe yönelik sabit bir kesiti temsil etmesi bakımından modellerin temel öğrenme kapasitelerini

sınamak için karşılaştırma amacıyla kullanılmıştır. Araştırmanın ikinci evresinde ise çalışmaya çalışmaya katkı sağlamak ve açık kaynak dünyasının anlık değişim gösteren güncel dinamiklerini doğrudan modele yansıtmak amacıyla GitHub API üzerinden güncel veriler toplanmıştır. Bu amaç doğrultusunda, GitHub API arabirimi üzerinden geliştirilen Python betiği aracılığıyla platformdan güncel veriler doğrudan çekilmiştir ve bu veriler GitHub REST API (*GitHub REST API Documentation*, n.d.) kanalıyla işlenerek `github_repos_balanced_900.csv` adıyla veri kümesine dönüştürülmüştür. Bu süreçte, kıyaslama amacıyla Kaggle üzerindeki açık kaynak referans veri kümesinden de (Scaturchio, n.d.) yararlanılmıştır. Bahsi geçen dinamik veri kümesi, projelerin sadece genel istatistiklerini değil, onları tanımlayan yapısal ve teknik nitelikleri de kapsamaktadır. Havuzlardaki toplam reaksiyon (yıldız) sayıları, projelerin diğer geliştiricilerce kopyalanma derecesini belirten çatallanma (fork) miktarı, aktif/kapalı issue dağılımları, projeye doğrudan kod desteği veren geliştirici sayısı ve dokümantasyon yeterliliğinin en önemli ölçütü sayılan README dosyalarının karakter hacmi bu öznelilikler arasında yer almaktadır.

Veri Ön İşleme ve Dengeleme Stratejisi

Açık kaynak yazılım ekosisteminden elde edilen ham verilerin doğal sınıf dağılımını ve bu durumun doğrudan tetiklediği yapısal sınıf dengesizliği (class imbalance) sorununu azaltmak amacıyla, bu çalışmada dengeli örnekleme stratejisi uygulanmıştır. Veri setindeki aynı gözlemlerin tekrarlanmasından kaynaklı model yanlılıklarının önüne geçmek adına, örneklem havuzu oluşturulurken her projenin benzersiz depo kimliği (repository ID) üzerinden çoklu kayıtlar tamamen temizlenmiştir. Uygulanan rastgele örnekleme adımlarının deneysel kararlılığını korumak ve ileride gerçekleştirilecek araştırmalar için tam bir tekrarlanabilirlik (reproducibility) zemini oluşturmak amacıyla, veri ön işleme akışında (pipeline) `random_state` özneliği sabitlenmiştir. Projelerin

genel popülerlik seviyelerini belirlemek amacıyla, depoların sahip olduğu ham yıldız sayıları referans alınmış ve pandas kütüphanesinin qcut metodu yardımıyla kantil tabanlı üç homojen kategoriye ("Düşük", "Orta" ve "Yüksek") Pandas (McKinney, 2010) kütüphanesinin qcut metodu yardımıyla üç sınıfa ayrılmıştır. Sınıflandırma modellerinin çoğunluk sınıfına yönelmesini engellemek amacıyla, her sınıftan rastgele 300'er adet gözlem çekilmiş ve toplamda 900 örneklemden oluşan dengeli bir veri kümesi elde edilmiştir. Öte yandan yürütülen ikili (binary) sınıflandırma testlerinde, literatürde yaygın kabul gören eşik değerleri doğrultusunda 50 yıldız sayısı bir sınır olarak kabul edilmiş ve veri kümesinin etiketleme süreci bu kritere göre gerçekleştirilmiştir.

Matematiksel modelleme öncesinde gerçekleştirilen veri ön işleme aşamalarında, girdi değişkenlerinin dağılımında gözlenen sapan ve aşırı uç değerlerin (outliers) öğrenme algoritmalarının öğrenme süreci üzerindeki olumsuz etkilerini azaltmak için NumPy (Harris et al., 2020) kütüphanesinin log1p metodu yardımıyla logaritmik dönüşüm tekniğinden yararlanılmıştır. Sınıflandırma süreçlerinde algoritmalar açısından ayırt edici ya da açıklayıcı bir nitelik taşımayan, öznelilik uzayında gürültü yaratma potansiyeli yüksek full_name, repo_name ve description gibi metinsel alanlar ile zamansal parametreler veri setinden çıkarılmıştır. Sistem matrisinde yer alan has_ci ve has_pages gibi mantıksal (boolean) yapılı ikili değişkenler, modeller tarafından kullanılabilmesi için float veri tipine uyarlanmıştır. Sözel ve kategorik karakterdeki öznelilikler (language, license, default_branch) ise Scikit-learn (Pedregosa, 2011) kütüphanesinin LabelEncoder algoritması vasıtasıyla her bir kategoriye özgü benzersiz sayısal değerlere dönüştürülmüştür. Veri ön işlemenin son kademesinde, farklı ölçeklerdeki sayısal değişkenlerin (örneğin dosya boyutu ile katkıcı sayısı arasındaki devasa sayısal farkların) model performansını

olumsuz etkilemesini önlemek amacıyla, tüm öznitelik uzayı StandardScaler kullanılarak standartlaştırılmıştır.

Sınıflandırma Modelleri ve Deneysel Kurulum

Araştırma kapsamında yürütülen deneysel süreçler Python 3.x ortamında gerçekleştirilmiş; veri analizi, ön işleme ve modelleme aşamalarında literatürde sıklıkla tercih edilen Pandas, NumPy, Scikit-learn ve XGBoost kütüphanelerinden yararlanılmıştır. Analiz süreçlerinde veri kümesi, dengeli bir dağılım sağlamak adına %80 oranında eğitim ve %20 oranında test kümesi olmak üzere eğitim ve test kümelerine ayrılmıştır. Bu bölümlendirme esnasında, ana kütledeki sınıfsal dağılım frekanslarının alt kümelere asimetrik veya yanlı yansımaları kesin olarak engellemek amacıyla stratify fonksiyonu aktif konuma getirilerek tabakalı bir örnekleme modeli uygulanmıştır. Tasarlanan deneysel senaryolarda değerlendirilen sınıflandırma modelleri ve deneysel optimizasyon süreci sonucunda belirlenen optimal hiperparametreler aşağıda sunulmuştur:

- Lojistik Regresyon: max_iter=2000
- Karar Ağacı: max_depth=12
- K-En Yakın Komşu (KNN): n_neighbors=15, weights='distance' (Cover & Hart, 1967)
- Destek Vektör Makineleri (SVM): RBF çekirdeği, C=2.0 (Cortes et al., 1995)
- Random Forest: n_estimators=400 (Breiman, 2001)
- Extra Trees: n_estimators=400 (Geurts et al., 2006)
- Gradient Boosting: n_estimators=250

- AdaBoost: `n_estimators=200` (Freund & Schapire, 1997)
- XGBoost: `n_estimators=350` (Chen & Guestrin, 2016)

Modellerin sergilediği tahmin ve genelleme yeteneklerinin değerlendirilmesinde, ham verideki sınıfsal dengesizliğin üretebileceği yanıltıcı sonuçlar daha doğru değerlendirebilmek amacıyla klasik doğruluk (accuracy) ölçütünün ötesine geçilmiştir. Bu amaç doğrultusunda, sınıfların sınıf dağılımlarını ve örneklem hacimlerini hesaba katan Ağırlıklı F1-Skoru (Weighted F1) ile sınıfları aritmetik olarak eşit ağırlıkta değerlendiren Makro F1-Skoru (Macro F1) metrikleri başarımlarının ana odak noktası olarak belirlenmiştir. Geliştirilen algoritmaların performans çıktıları tek eğitim-test ayrımı üzerinden hesaplandığı için, elde edilen sonuçlar mevcut veri kümesinin kullanılan veri kümesine bağlı, kısmen iyimser (over-optimistic) eğilimler taşıma riski barındırmaktadır. Söz konusu metodolojik kısıtı ortadan kaldırmak ve geliştirilen makine öğrenmesi modellerinin genellenebilirliği daha güvenilir şekilde değerlendirebilmek amacıyla, sonraki çalışmalarda K-Katlı Çapraz Doğrulama (K-Fold Cross Validation) yaklaşımları ile harici veri setleri üzerinden yürütülecek dış doğrulama (external validation) stratejilerinin kullanılması planlanmaktadır.

Bulgular ve Tartışma

Gerçekleştirilen deneysel çalışmalar neticesinde elde edilen istatistiksel veriler; güncel GitHub API'si üzerinden orijinal olarak toplanan özgün veri seti ile literatürde sıklıkla başvuru statik Kaggle veri tabanı arasındaki farkları ortaya koymaktadır. İki ve üç sınıflı modelleme yapıları doğrultusunda elde edilen sonuçlar, Matplotlib (Hunter, 2007) ve Seaborn (Waskom, 2021) kütüphanelerinden yararlanılarak görselleştirilmiş ve mukayeseli bir

perspektifle değerdendirilmiştir. Elde edilen ampirik bulgular, algoritmaların yalın tahmin performanslarını ortaya koymanın ötesinde; veri setlerinin hacimsel büyüklüğü ile sınıfsal dağılım dengesinin, modellerin genellenebilirlik kapasitesi üzerinde ne kadar belirleyici olduğunu net bir şekilde göstermiştir. Gerçekleştirilen deneylerin çıktıları, mevcut literatürde sıklıkla ihmal edilen "veri kalitesi ve veri miktarı arasındaki ters yönlü ilişkiyi" sorgulayan ve tartışmaya açan önemli bulgular sunmaktadır. Bu doğrultuda; ikili sınıflandırma eşliğinde ağaç tabanlı topluluk modellerinin sergilediği kararlı başarı oranları ile çok sınıflı kategorizasyon süreçlerinde mesafe tabanlı modellerin yaşadığı zayıflıklar, ilgili metrik korelasyonlarıyla birlikte ilerleyen alt bölümlerde detaylandırılmış ve literatürdeki teorik altyapıyla ilişkilendirilmiştir.

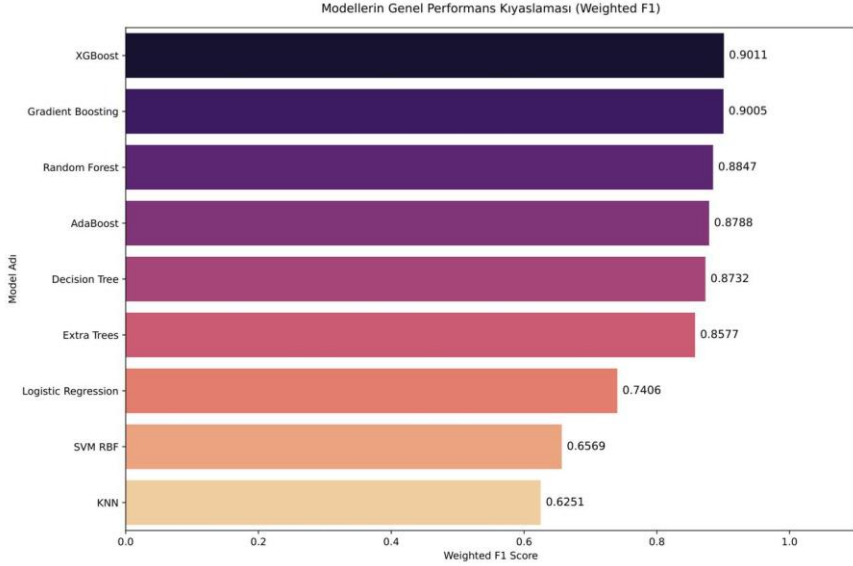
Özgün Veri Seti Üzerindeki Deneysel Sonuçlar

Üçlü (Multi-class) Sınıflandırma Bulguları ve Sınıf Bazlı Analizler

Dengelenmiş özgün veri kümesi kullanılarak eğitilen sınıflandırma modellerine ait performans dinamiklerini bütüncül bir perspektifle değerdendirebilmek adına, karşılaştırmalı bir başarıım grafiğı hazırlanmıştır. Bu doğrultuda; sınanan dokuz farklı makine öğrenmesi algoritmasının üçlü sınıflandırma (multi-class classification) mimarisindeki kararlılık düzeylerini yansıtan Ağırlıklı F1-Skoru (Weighted F1 Score) dağılım şeması Şekil’1 aracılığıyla ortaya konmuştur.

Şekil 1 Özgün Üçlü Sınıflandırma Senaryosunda Modellerin Genel

Performans Kıyaslaması (Weighted F1)



Kaynak: (Güvez et al., 2026)

Şekil'1 üzerinde sunulan performans çıktıları analiz edildiğinde; üç sınıflı tahmin süreçlerinde ağaç tabanlı topluluk (ensemble) modellerinin, geleneksel parametrik ve mesafe tabanlı yaklaşımlara göre bariz bir performans avantajı yakaladığı gözlemlenmektedir. Ağırlıklı F1-Skoru (Weighted F1 Score) metrikleri baz alınarak gerçekleştirilen sıralamaya göre, en güçlü tahmin performansına 0,9011 değeriyle XGBoost algoritması ulaşmış; Gradient Boosting (0,9005) ise çok az bir farkla ikinci sırada yer almıştır. Random Forest ise 0,8851'lik skorla bu yüksek performanslı artırma (boosting) grubunun hemen arkasında yer almaktadır. Buna karşın, doğrusal ve mesafe tabanlı geleneksel modeller daha düşük bir performans sergileyerek grafiğin alt sıralarında yer almıştır. Lojistik Regresyon 0,7406'lık bir değerde kalırken, öz nitelik uzayındaki ölçek farklarından etkilenen RBF çekirdekli SVM 0,6569 ve KNN algoritması 0,6251'lik skorlarla en düşük performansı sergilemiştir. Elde edilen bu performans sıralaması; yüksek boyutlu ve karmaşık açık kaynak yazılım

metriklerinin çözümlenmesinde, gradyan tabanlı topluluk yapılarının mesafe ve doğrusallık odaklı geleneksel yaklaşımlara karşı üstünlük sağladığını göstermektedir.

Özgün dengeli veri seti mimarisi üzerinde koşturulan dokuz farklı sınıflandırma algoritmasının başarımları; doğruluk (accuracy), ağırlıklı f1-skoru (weighted f1 score) ve makro f1-skoru (macro f1 score) metrikleri eşliğinde Tablo 1' de detaylandırılmıştır. İlgili sayısal çıktılar, modellerin üçlü sınıflandırma senaryoları altındaki kestirim kararlılıklarını ve katmanlar arasındaki ayırt edici yeteneklerini göstermektedir.

Tablo 1 Özgün Dengeli Veri Seti Üzerindeki Üçlü Sınıflandırma Modellerinin Başarımları

Sıra	Model	Accuracy (Doğruluk)	Weighted F1 Score	Macro F1 Score
1	XGBoost	0,9000	0,9011	0,9011
2	Gradient Boosting	0,9000	0,9005	0,9005
3	Random Forest	0,8833	0,8851	0,8851
4	AdaBoost	0,8778	0,8788	0,8788
5	Decision Tree	0,8667	0,8686	0,8686
6	Extra Trees	0,8500	0,8522	0,8522

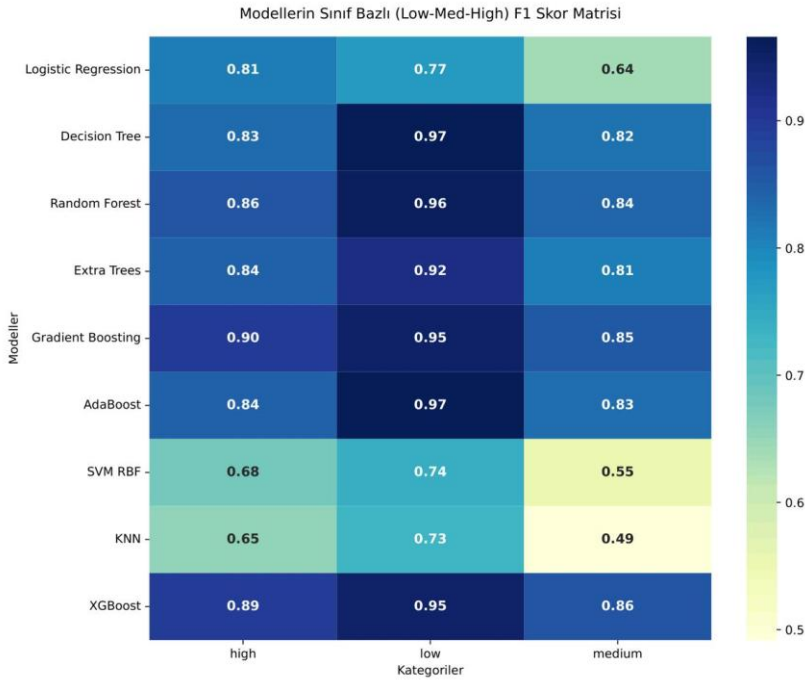
7	Logistic Regression	0,7389	0,7406	0,7406
8	SVM	0,6556	0,6569	0,6569
9	KNN	0,6278	0,6251	0,6251

Kaynak: (Güvez et al., 2026)

Tablo 1 üzerinde yer alan istatistiksel çıktılar detaylı olarak incelendiğinde, modellerin doğruluk ve F1-skorlarının tam bir korelasyon içinde hareket ettiği ve hiyerarşik sıralamanın değişmediği görülmektedir. Üçlü sınıflandırma mimarisinde en yüksek başarıma, %90,00 doğruluk oranı ve 0,9011 Makro/Ağırlıklı F1 değerleri üreten XGBoost ile Gradient Boosting modelleri ulaşmıştır. Bu iki gradyan artırma algoritmasını, %88,33 doğruluk ve 0,8851 F1-Skoru ile Random Forest topluluk modeli yakından takip etmektedir. AdaBoost (%87,78 doğruluk), Decision Tree (%86,67 doğruluk) ve Extra Trees (%85,00 doğruluk) algoritmaları da dengeli veri mimarisi üzerinde güçlü ve kararlı performans sergilemişlerdir. Buna karşın, parametrik ve mesafe tabanlı geleneksel modeller hiyerarşinin alt basamaklarında kalmıştır. Lojistik Regresyon algoritması %73,89 doğruluk oranı ve 0,7406 F1-Skoru eşiğinde kalırken, RBF çekirdekli SVM modeli %65,56 doğruluk ve 0,6569 F1 değerleri ile zayıf bir ayırt etme performansı ortaya koymuştur. Dokuz algoritma arasında en düşük başarıma ise %62,78 doğruluk ve 0,6251 F1-Skoru ile K-En Yakın Komşu (KNN) algoritmasından elde edilmiştir. Bu ampirik bulgular; veri kümesi tamamen dengeli hale getirilse dahi, açık kaynak metriklerinin çok boyutlu yapısını çözümlemede ağaç tabanlı topluluk modellerinin, geleneksel mesafe ve doğrusallık odaklı modellere kıyasla üstünlüğe sahip olduğunu net bir biçimde göstermektedir.

Önerilen makine öğrenmesi modellerinin sınıf bazlı hassasiyet eğilimlerini ortaya koymak ve algoritmaların hangi popülerlik seviyelerinde sınıflandırma hatası yaptığını daha ayrıntılı incelemek amacıyla sınıf bazlı F1-Skoru değerleri analiz edilmiştir. Bu doğrultuda; incelenen dokuz algoritmanın "Düşük", "Orta" ve "Yüksek" popülerlik kategorilerindeki özgün başarımlarını yansıtan ısı haritası (heatmap) çıktıları, Şekil 2'de detaylandırılarak analiz edilmiştir.

Şekil 2 Üçlü Sınıflandırma Modellerinin Sınıf Bazlı (Low-Med-High) F1 Skor Matrisi



Kaynak: (Güvez et al., 2026)

Şekil 2'deki ısı haritası verileri detaylı olarak incelendiğinde; algoritmaların "Düşük" popülerlik seviyesindeki tahmin performanslarının, "Orta" ve "Yüksek" kategorilerine kıyasla belirgin ölçüde daha güçlü ve istikrarlı olduğu görülmektedir.

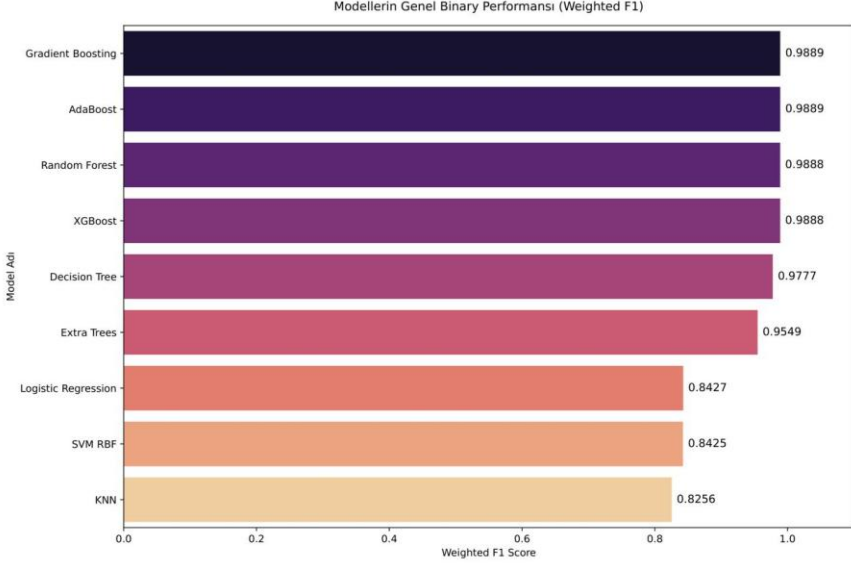
Geliştirilen modellerin hemen hepsi, düşük popülerliğe sahip açık kaynaklı projeleri sınıflandırmada en yüksek F1-Skoru seviyelerine ulaşmıştır. Bu paralelde Decision Tree ve AdaBoost algoritmaları "low" kategorisinde 0,97 değerleriyle en yüksek başarıyı sergilemiş; bu modelleri sırasıyla Random Forest (0,96), XGBoost (0,95) ve Gradient Boosting (0,95) takip etmiştir. Mesafe ve doğrusallık tabanlı klasik yaklaşımlardan KNN (0,73) ve SVM (0,74) algoritmaları dahi, en yüksek tahmin performanslarına bu popülerlik kategorisinde ulaşmıştır. Buna karşın "Orta" (medium) popülerlik katmanı, algoritmaların tamamı için sınıflandırma hatasının en yoğun yaşandığı ve performansın en çok düştüğü bölge olmuştur. En başarılı model olan XGBoost orta sınıfta 0,86 , Gradient Boosting ise 0,85 F1-Skoruna gerilemiştir. Söz konusu sınıfsal zafiyet klasik algoritmaların performansında belirgin bir gerilemeyi beraberinde getirmiştir. Öyle ki, SVM orta grupta 0,55 seviyesinde kalırken, KNN algoritması 0,49 F1-Skoru üreterek karar sınırlarında ayırt etme yeteneğini büyük ölçüde kaybetmiştir. Öte yandan, projelerin yüksek popülerlik kademeleri ele alındığında; Gradient Boosting 0,90 ve XGBoost 0,89 F1-Skoru üreterek yüksek etkileşimli yazılım depolarını tahmin etmede oldukça kararlı bir modelleme performansı ortaya koymuştur. Elde edilen bu matrisel bulgular; açık kaynak projelerinin "Düşük" ve "Yüksek" gibi uç popülerlik karakteristiklerinin ayırt edilmesinin daha kolay olduğunu, ancak geçiş fazını temsil eden gürültülü "Orta" sınıfının doğru kategorize edilmesinde gradyan tabanlı topluluk modellerinin sunduğu esnek karar sınırlarına kritik ölçüde gereksinim duyulduğunu ampirik olarak göstermektedir.

İkili (Binary) Sınıflandırma Bulguları ve Hata Dağılımı

Sınıflandırma algoritmalarının daha sade bir karar sınırı üzerindeki başarı dinamiklerini analiz etmek adına gerçekleştirilen ikili (binary) modelleme senaryolarından elde edilen ampirik bulgular, Şekil 3 aracılığıyla paylaşılmıştır. İlgili yatay çubuk

grafığı, dokuz farklı algoritmanın 50 yıldız eşığı temel alınarak hesaplanan Ağırlıklı F1-Skoru (Weighted F1 Score) dağılımlarını genel bir perspektifle sunmaktadır.

Şekil 3 İkili Sınıflandırma Senaryosunda Modellerin Genel Performans Kıyaslaması (Weighted F1)



Kaynak: (Güvez et al., 2026)

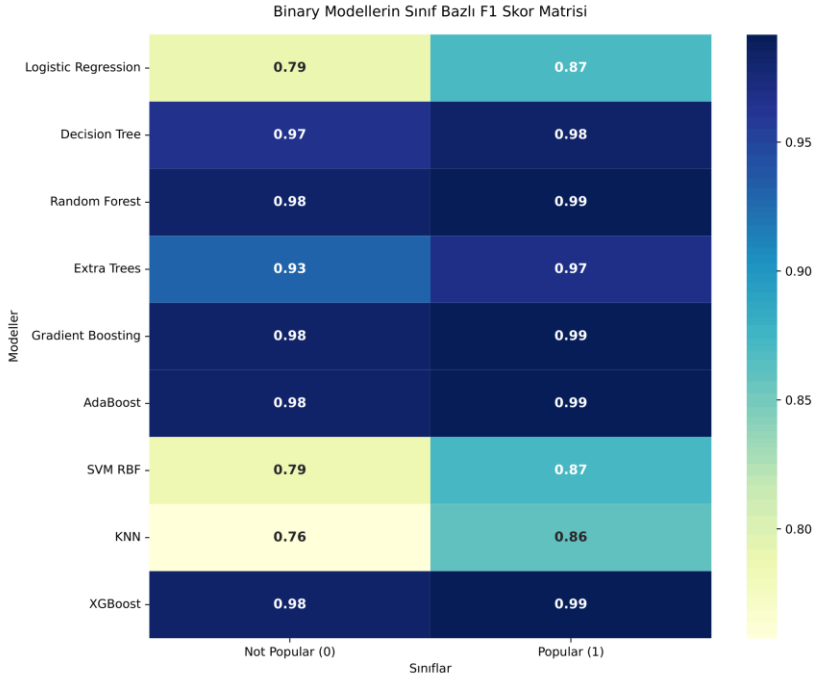
Şekil 3 üzerinde sunulan ikili (binary) sınıflandırma çıktıları analiz edildiğinde; problemin iki sınıfa indirgenmesinin, makine öğrenmesi algoritmalarının performans katsayılarında kayda değer bir yükselişi beraberinde getirdiğı gözlemlenmektedir. Üçlü mimarideki karmaşıklığın ortadan kalkması, özellikle ağaç tabanlı topluluk (ensemble) modellerinin üst düzey bir kestirim kararlılığı yakalamasına olanak tanımıştır. Bu doğrultuda; Gradient Boosting ve AdaBoost algoritmaları 0,9889'luk zirve F1-Skorları ile en yüksek performansı paylaşıırken, onları çok küçük bir farkla 0,9888 değerine ulaşan Random Forest ve XGBoost modelleri izlemiştir. Decision Tree (0,9777) ve Extra Trees (0,9549) algoritmaları da bu yüksek başarımlarını destekleyen güçlü sonuçlar üretmiştir. Buna

karşın, karar yüzeylerinin basitleşmesi geleneksel parametrik ve mesafe tabanlı modelleri de yukarı taşımış, ancak bu modeller hiyerarşik olarak yine alt segmentte kalmıştır. Lojistik Regresyon 0,8427 ve SVM RBF 0,8425 skorları ile birbirine oldukça yakın eğilimler sergilerken, dokuz algoritma arasındaki en düşük kestirim başarısı 0,8256'lık Ağırlıklı F1-Skoru ile K-En Yakın Komşu (KNN) algoritmasından elde edilmiştir. Neticede elde edilen ampirik çıktılar; 50 yıldız sınırı temelinde kurgulanan ikili sınıflandırmada, geçiş fazını oluşturan gürültülü orta segmentin veri setinden arındırılmasının tüm modellerin başarı katsayısını yukarı taşıdığını, fakat ağaç tabanlı gradyan optimizasyon algoritmalarının doğrusal tabanlı klasik yöntemlere kıyasla literatürdeki üstün konumunu sürdürdüğünü kanıtlamaktadır.

İkili (binary) sınıflandırma modellemesinde algoritmaların "Popüler Değil" (Not Popular - 0) ve "Popüler" (Popular - 1) sınıfları üzerindeki hassasiyet dağılımlarını ve sınıf bazlı tahmin kararlılıklarını detaylı olarak incelemek amacıyla bir F1-Skor Matrisi oluşturulmuştur. Bu doğrultuda; test edilen dokuz farklı algoritmanın hedef etiketler üzerindeki sınıf bazlı başarılarını yansıtan ısı haritası (heatmap) çıktıları, Şekil 4'te detaylandırılarak gösterilmiştir.

Şekil 4 İkili Sınıflandırma Modellerinin Sınıf Bazlı F1 Skor

Matrisi



Kaynak: (Güvez et al., 2026)

Şekil 4 üzerinde yer alan sınıf bazlı ısı haritası verileri incelendiğinde, karar sınırının iki sınıfa indirgenmesiyle birlikte algoritmaların "Popular (1)" sınıfını tahmin etmedeki hassasiyetlerinin, "Not Popular (0)" sınıfına kıyasla daha kararlı bir eğilim sergilediği gözlemlenmektedir. Başta Gradient Boosting, AdaBoost, Random Forest ve XGBoost olmak üzere topluluk (ensemble) modellerinin tamamı, popüler olan projeleri ayırt etmede 0,99'luk neredeyse kusursuz bir F1-Skoru başarısına ulaşmışlardır. Bu yüksek performanslı küme, "Not Popular (0)" sınıfında da kararlılığını koruyarak 0,98'lik (Decision Tree için 0,97) oldukça güçlü kestirim değerleri üretmiştir. Buna karşın geleneksel modeller, her iki kategoride de topluluk yapılarının belirgin şekilde gerisinde kalmıştır. Lojistik Regresyon ve SVM RBF algoritmaları, projelerin popülerlik durumlarını (1) 0,87 F1-Skoru ile modellerken, popüler olmama durumlarını (0) 0,79'luk bir hassasiyetle yakalayabilmiştir.

Isı haritasındaki en düşük ayırt etme kararlılığı ise "Not Popular (0)" sınıfında 0,76 ve "Popular (1)" sınıfında 0,86 F1-Skoru üreten KNN algoritmasına aittir. Özetle ampirik bulgular; gürültülü orta sınıfın elendiği ikili sınıflandırma senaryosunda algoritmaların genel başarısının yükseldiğini, ancak ağaç tabanlı artırma ve torbalama stratejilerinin her iki hedef sınıfta da geleneksel yaklaşımlara karşı belirgin üstünlüğünü sürdürdüğünü ortaya koymaktadır.

Modellerin 50 yıldız baraj eşiğine dayalı ikili karar sınırları altındaki sayısal performans çıktıları; Doğruluk (Accuracy), Ağırlıklı F1-Skoru (Weighted F1-Score), Makro F1-Skoru (Macro F1-Score) ve İkili F1-Skoru (Binary F1-Score) metrikleri eşliğinde Tablo 2 bünyesinde detaylandırılmıştır.

Tablo 2 Özgün Dengeli Veri Seti İkili Sınıflandırma Performans Sonuçları

Sıra	Model	Accuracy (Doğruluk)	Weighted F1 Score	Macro F1 Score	Binary F1 Score
1	Gradient Boosting	0,9889	0,9889	0,9875	0,9917
2	AdaBoost	0,9889	0,9889	0,9875	0,9917
3	Random Forest	0,9889	0,9888	0,9874	0,9917
4	XGBoost	0,9889	0,9888	0,9874	0,9917
5	Decision Tree	0,9778	0,9777	0,9748	0,9835

6	Extra Trees	0,9611	0,9605	0,9549	0,9717
7	Logistic Regression	0,8389	0,8427	0,8291	0,8700
8	SVM	0,8389	0,8425	0,8281	0,8711
9	KNN	0,8222	0,8256	0,8086	0,8596

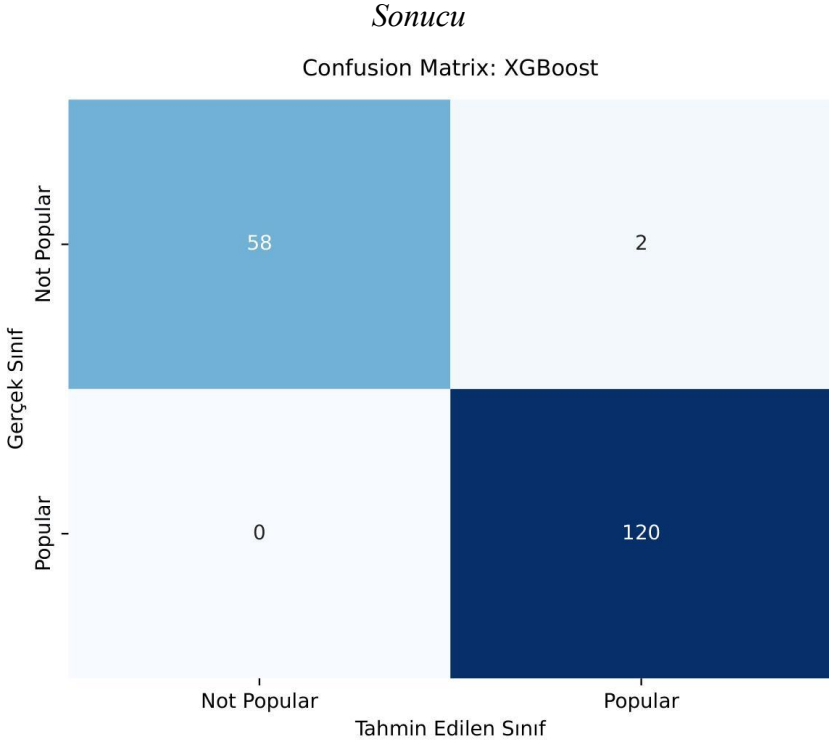
Kaynak: (Güvez et al., 2026)

Tablo 2 verileri, karar yüzeyinin iki sınıfa indirgenmesiyle topluluk (ensemble) modellerinin %98,89'luk sabit bir genel Doğruluk (Accuracy) oranında eşitlendiğini göstermektedir. Gradient Boosting, AdaBoost, Random Forest ve XGBoost modellerinin tümü, ikili F1-Skorunda (Binary F1-Score) 0,9917'lik zirve değerine ulaşmıştır. Bu yüksek başarımlı kümesini %97,78 doğrulukla Decision Tree ve %96,11 doğrulukla Extra Trees modelleri takip etmektedir. Buna karşın geleneksel algoritmalar, veri uzayının sadeleşmesiyle birlikte bir performans artışı sergilese de sıralamanın alt basamaklarında yer almıştır. Lojistik Regresyon ve SVM RBF modelleri %83,89' luk aynı doğruluk oranını ve sırasıyla 0,8700 ve 0,8711 İkili F1-Skorlarını paylaşmıştır. En düşük kestirim kararlılığı ise %82,22 doğruluk ve 0,8596 İkili F1 değeri ile KNN algoritmasına aittir. Özetle bulgular; gürültülü orta sınıfın elenmesinin tüm algoritmaların başarı oranını yukarı taşıdığını, ancak ağaç tabanlı artırma (boosting) optimizasyonlarının performans sıralamasındaki üstün konumunu değiştirmedini göstermektedir.

İkili sınıflandırma senaryosunda zirve performans gösteren XGBoost algoritmasının doğru ve hatalı tahmin dağılımlarını

incelemek amacıyla elde edilen karmaşıklık matrisi (confusion matrix), Şekil 5'te sunulmuştur.

Şekil 5 İkili Sınıflandırma XGBoost Modeli Confusion Matrix



Kaynak: (Güvez et al., 2026)

Şekil 5 üzerinde yer alan karmaşıklık matrisi (confusion matrix) verileri incelendiğinde, XGBoost modelinin test seti üzerinde son derece yüksek bir ayırım kararlılığı sergilediği görülmektedir. Model, "Poplar (1)" sınıfına ait 120 adet projenin tamamını sıfır hatayla (%100 doğrulukla) doğru tahmin etmeyi başarmıştır. "Not Poplar (0)" sınıfında ise mevcut 60 örnekten 58'ini doğru sınıflarken, yalnızca 2 projeyi popüler olmadığı halde popüler olarak (yanlış pozitif) hatalı kodlamıştır. Toplamda 180 test örneğinden 178'ini doğru tahmin eden modelin bu matris çıktıları, 50

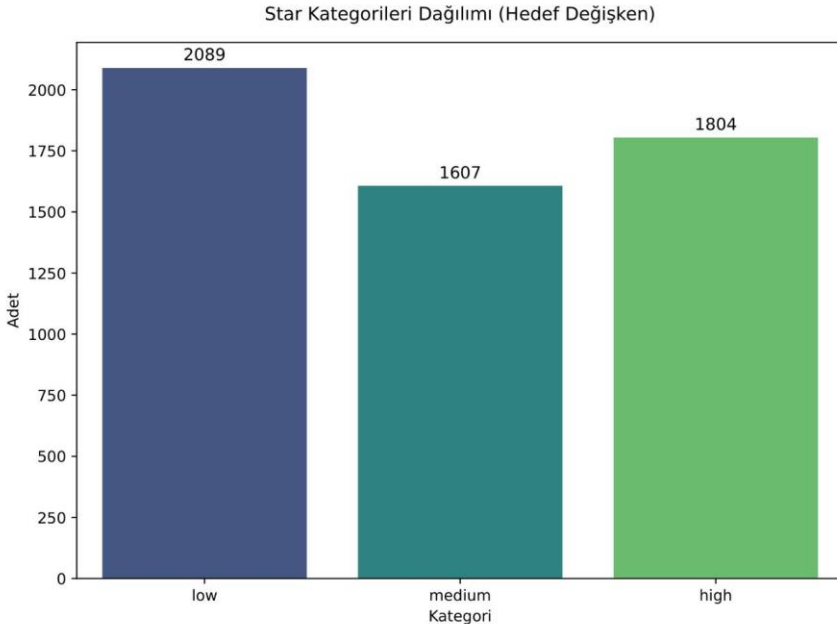
yıldız sınırına dayalı ikili popülerlik sınıflandırmasında sergilenen %98,89'luk genel doğruluğu ve kararlı tahmin gücünü sayısal olarak doğrulamaktadır.

Hazır Veri Seti Üzerindeki Deneysel Sonuçlar

Üçlü (Multi-class) Sınıflandırma Bulguları ve Sınıf Bazlı Analizler

Açık kaynak yazılım projelerinin popülerlik katmanlarını tahmin etmeye yönelik deneysel çalışmaların ikinci aşamasında, literatürde kabul görmüş hazır bir veri seti (benchmark dataset) kullanılarak analiz kapsamı genişletilmiştir. Bu doğrultuda; hazır veri seti kapsamındaki projelerin yıldız sayılarına göre ayrılan "Düşük", "Orta" ve "Yüksek" sınıf frekanslarını ve örneklem yoğunluğunu gösteren hedef değişken dağılımı, Şekil 6'da gösterilmiştir.

Şekil 6 Hazır Veri Seti Star Kategorileri Dağılımı (Hedef Değişken)

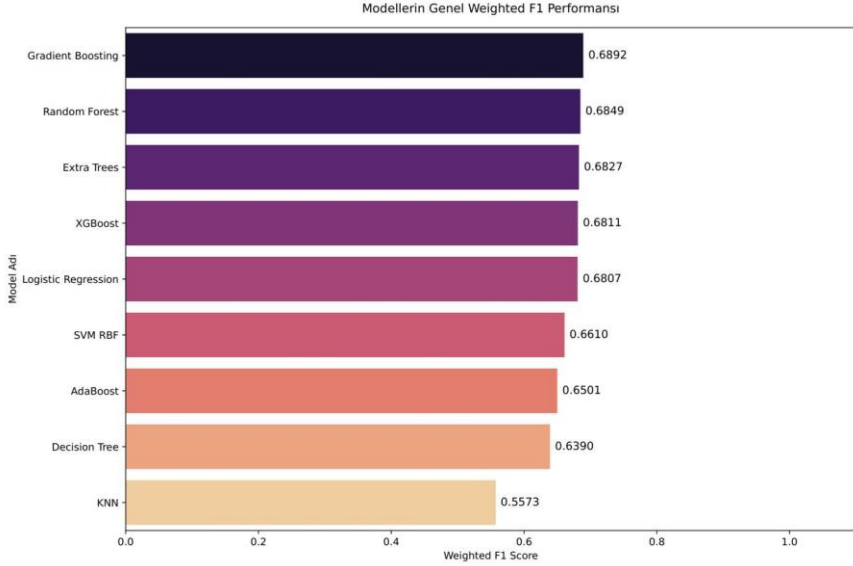


Kaynak: (Güvez et al., 2026)

Şekil 6 üzerinde yer alan hedef değişken dağılımı incelendiğinde, hazır veri setinin özgün veri setinden farklı olarak tamamen dengesiz (imbalanced) bir sınıf yapısına sahip olduğu gözlemlenmektedir. İlgili veri kümesindeki projelerin baskın bir çoğunluğunun düşük popülerlik sınıfında yoğunlaştığı, buna karşın orta ve özellikle yüksek popülerlik kademelerine geçildikçe örneklem yoğunluğunun belirgin bir azalış eğilimi gösterdiği tespit edilmiştir. Söz konusu sınıfsal dağılım dengesizliği, modellerin eğitim aşamasında çoğunluk sınıfına (majority class) tolerans tanıma olasılığını yükseltmekte ve deneysel sınıflandırma performansının değerlendirilmesini daha zorlayıcı bir boyuta taşımaktadır.

Söz konusu benchmark veri kümesi temelinde yürütülen üç sınıflı popülerlik tahmin senaryosunda, değerlendirilen dokuz farklı makine öğrenmesi modelinin genel başarı çıktılarını kıyaslamak adına bir performans dağılım grafiği sunulmuştur. Bu kapsamda; modellerin genel sınıflandırma kararlılıklarını ve tahmin güçlerini yansıtan Ağırlıklı F1-Skoru (Weighted F1 Score) karşılaştırma grafiği, Şekil 7'de gösterilmiştir.

*Şekil 7 Hazır Veri Seti Üçlü Sınıflandırma Senaryosunda
Modellerin Genel Performans Kıyaslaması (Weighted F1)*

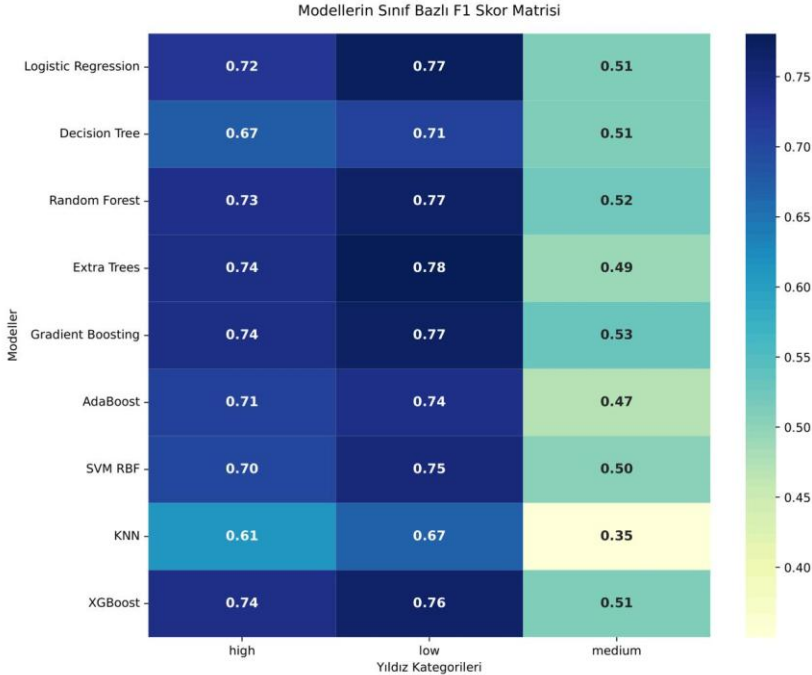


Kaynak: (Güvez et al., 2026)

Şekil 7 üzerinde yer alan hazır benchmark veri setine ait genel performans çıktıları incelendiğinde, dengesiz sınıf dağılımının bir sonucu olarak algoritmaların Ağırlıklı F1-Skorlarında (Weighted F1 Score) özgün veri setine kıyasla belirgin bir düşüş yaşandığı görülmektedir. Bu zorlu test zemininde en yüksek başarıma 0,6892'lik skorla Gradient Boosting algoritması ulaşırken, onu sırasıyla 0,6849 ile Random Forest ve 0,6827 ile Extra Trees topluluk modelleri takip etmiştir. Özgün veri setinde tepe performansı gösteren XGBoost algoritması ise bu senaryoda 0,6811'lik bir değer üreterek güçlü ağaç tabanlı yapıların hemen arkasında yer almıştır. Bu ampirik dağılımın en dikkat çekici bulgusu, geleneksel modellerden Lojistik Regresyon'un 0,6807 ve SVM RBF'in 0,6610'luk skorlarla AdaBoost (0,6501) ve Decision Tree (0,6390) gibi bazı ağaç tabanlı modelleri geride bırakmasıdır. Dokuz algoritma içerisindeki en düşük başarıma ise yine mesafe tabanlı yapısı gereği dengesiz veri dağılımlarından negatif etkilenen KNN (0,5573) algoritmasından elde edilmiştir. Özetle bulgular;

hazır veri setinin gürültülü ve dengesiz yapısının modellerin genel başarımlarını aşağı çektiğini, ancak Gradient Boosting ve Random Forest gibi topluluk yapılarının varyans optimizasyonları sayesinde hiyerarşik liderliğini koruduğunu doğrulamaktadır. Hazır benchmark veri setinin getirdiği dengesiz sınıf dağılımı altında, algoritmaların "Düşük", "Orta" ve "Yüksek" popülerlik kategorilerindeki sınıf bazlı hassasiyetlerini detaylı olarak incelemek amacıyla bir F1-Skor Matrisi oluşturulmuştur. Bu paralel doğrultuda; değerlendirilen dokuz farklı modelin hedef etiketlere ait kendilerine has performans trendleri ve hata yoğunlukları, Şekil 8 üzerinde sunulan ısı haritası (heatmap) ampirik verileriyle ortaya konmuştur.

Şekil 8 Hazır Veri Seti Üçlü Sınıflandırma Sınıf Bazlı F1 Skor Matrisi



Kaynak: (Güvez et al., 2026)

Şekil 8 üzerinde yer alan sınıf bazlı ısı haritası çıktıları analiz edildiğinde; ilgili benchmark veri setindeki sınıfsal dengesizliğin, algoritmaların tahmin performansları üzerindeki baskılayıcı rolü net bir biçimde ortaya çıkmaktadır. Modellerin tamamı, veri yoğunluğunun en fazla olduğu "low" (düşük) popülerlik sınıfında en kararlı performanslarını sergilemiştir. Bu kategoride Extra Trees 0,78 , Random Forest, Gradient Boosting ve Lojistik Regresyon 0,77 F1-Skoruna ulaşırken, en düşük başarımları 0,67 ile KNN algoritmasında kalmıştır. Benzer şekilde, "high" (yüksek) popülerlik sınıfında da Extra Trees, Gradient Boosting ve XGBoost 0,74'lük kararlı F1-Skorları ile öne çıkmıştır. Öte yandan, geçiş fazını sembolize eden ve örneklem hacmi bakımından en kısıtlı katman olan "Orta" popülerlik seviyesi, tüm modellerin sınıflandırma performansında belirgin bir gerilemenin yaşandığı ortak alan olarak öne çıkmaktadır. En yüksek kararlılığı sunan Gradient Boosting dahi orta sınıfta 0,53 F1-Skoruna gerilerken; Random Forest 0,52 , XGBoost ve Lojistik Regresyon ise 0,51 değerinde kalmıştır. Bu sınıfta KNN algoritması 0,35 F1-Skoru ile tamamen kararsız bir bölgeye düşmüştür. Netice itibarıyla elde edilen matris çıktıları; referans veri kümesindeki sınıfsal dengesizlik nedeniyle modellerin ara geçiş segmentini (medium) ayrıştırma kabiliyetinin büyük ölçüde kısıtlandığını, buna karşın topluluk modellerinin söz konusu gürültü karşısında klasik algoritmalara kıyasla daha yüksek bir adaptasyon direnci sergilediğini doğrulamaktadır.

Hazır benchmark veri seti üzerindeki üç sınıflı kestirim senaryosuna ait sayısal performans çıktıları; Doğruluk (Accuracy), Ağırlıklı F1-Skoru (Weighted F1-Score) ve Makro F1-Skoru (Macro F1-Score) metrikleriyle birlikte Tablo 3'te detaylandırılmıştır. İlgili istatistiksel sonuçlar, literatürdeki bu dengesiz sınıf yapısının

algoritmaların tahmin kararlılığı ve genel başarısı üzerindeki sınırlandırıcı etkisini açıkça ortaya koymaktadır.

Tablo 3 Hazır Veri Seti Üçlü Sınıflandırma Performans Sonuçları

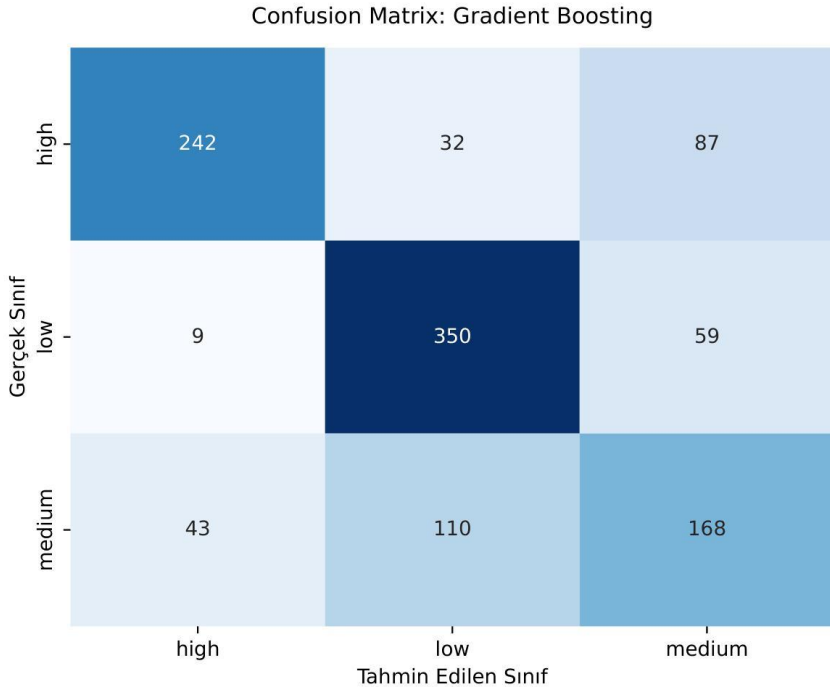
Sıra	Model	Accuracy (Doğruluk)	Weighted F1 Score	Macro F1 Score
1	Gradient Boosting	0,6909	0,6892	0,9011
2	Extra Trees	0,6909	0,6837	0,6718
3	Random Forest	0,6836	0,6829	0,6724
4	XGBoost	0,6827	0,6811	0,6704
5	Logistic Regression	0,6809	0,6807	0,6695
6	SVM	0,6591	0,6610	0,6505
7	AdaBoost	0,6527	0,6501	0,6388
8	Decision Tree	0,6327	0,6400	0,6318
9	KNN	0,5655	0,5573	0,5437

Kaynak: (Güvez et al., 2026)

Tablo 3 üzerinde yer alan sayısal çıktılar analiz edildiğinde; ilgili referans veri kümesinin sınıfsal dengesizliğinin bir sonucu olarak performans kriterlerinin, çalışmada derlenen özgün veri setindeki değerlere nispeten daha düşük bir tavan sınırında kümelandığı tespit edilmiştir. Söz konusu üç sınıflı kompleks senaryoda, en yüksek genel doğruluk katsayısı %69,09'luk bir performansla Gradient Boosting ve Extra Trees algoritmaları tarafından kaydedilmiştir. Bu modeli %68,36 doğrulukla Random Forest ve %68,27 doğrulukla XGBoost algoritması yakından takip etmektedir. Öte yandan veri kümesindeki yüksek gürültü, geleneksel modeller ile topluluk yapıları arasındaki performans farkının azalmasına yol açmıştır. Nitekim Lojistik Regresyon algoritması %68,09 genel doğruluk ve 0,6807 Ağırlıklı F1-Skoru sergileyerek görece daha karmaşık mimarilerle benzer bir başarı yakalamış; buna karşın RBF çekirdekli SVM modeli %65,91 doğruluk oranının ötesine geçememiştir. AdaBoost (%65,27 doğruluk) ve Decision Tree (%63,27 doğruluk) modelleri ise ağaç tabanlı olmalarına rağmen dengesiz dağılımdan olumsuz etkilenerek alt sıralara gerilemiştir. Değerlendirilen dokuz model arasında en zayıf sınıflandırma performansı; %56,55 genel doğruluk ve 0,5573 Ağırlıklı F1-Skoru sergileyen K-En Yakın Komşu (KNN) algoritması tarafından kaydedilmiştir. Özetle istatistiksel sonuçlar; hazır veri setindeki sınıf dengesizliği karşısında Gradient Boosting, Extra Trees ve Random Forest gibi gelişmiş varyans optimizasyonuna sahip topluluk yapılarının performans üstünlüğünü koruduğunu göstermektedir. Üçlü sınıflandırma senaryosunda en yüksek performansı gösteren Gradient Boosting algoritmasının tahmin kararlılığını ve sınıflar arasındaki hata dağılımlarını frekans bazlı incelemek amacıyla bir karmaşıklık matrisi (confusion matrix) oluşturulmuştur.

Bu paralelde; modelin gerek sınıflar ile kestirilen sınıflar arasındaki rneklem korelasyonunu ortaya koyan test ıktıları, ekil 9 zerinde paylaşılmıştır.

ekil 9 Hazır Veri Seti l Sınıflandırma Gradient Boosting Modeli Confusion Matrix Sonucu



Kaynak: (Gvez et al., 2026)

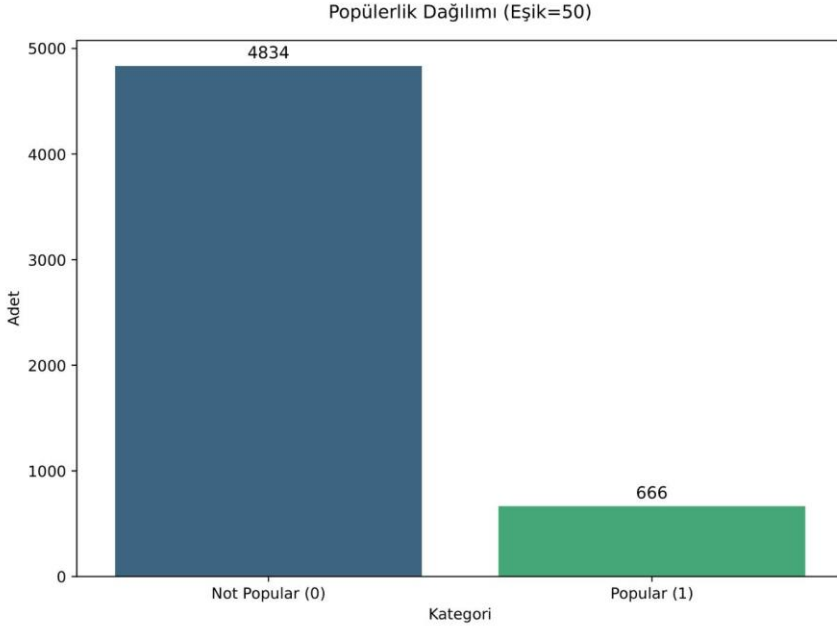
ekil 9 zerinde sunulan karmaşıklık matrisi (confusion matrix) ıktıları analiz edildiğinde; referans veri kmesindeki sınıfsal dengesizlik ile ara geiş katmanlarının tetiklediėi grltnn, Gradient Boosting algoritmasının mikro tahmin dinamiklerine yansıması belirgin bir biimde gzlenmektedir. Model, veri yoėunluėunun en fazla olduėu "dřk" sınıfına ait 418

örnekten 350'sini (%83,73'ünü) doğru tahmin ederek bu etiket üzerinde yüksek bir başarı sergilemiştir. "yüksek" sınıfında ise toplam 361 örnekten 242'sini (%67,03'ünü) doğru sınıflandırırken, bu sınıfa ait 87 örneği ise "medium" olarak hatalı kestirmiştir. Tahmin uyumsuzluklarının en yoğun gözlemlendiği alan ise, beklentiler doğrultusunda ayırım çizgileri en belirsiz segmenti oluşturan "Orta" popülerlik sınıfı olmuştur. Bu kategoriye ait 321 örnekten yalnızca 168'i (%52,33) doğru tahmin edilebilmiştir; sınıfın 110 örneği "düşük", 43 örneği ise "yüksek" olarak yanlış sınıflandırılmıştır. Netice itibarıyla elde edilen deneysel veriler; Gradient Boosting modelinin uç kategorileri ayırtırmada yüksek bir kararlılık sergilediğini, buna karşın ara geçiş fazını oluşturan "Orta" sınıfında öznitelik örtüşmeleri ve yetersiz örneklem yoğunluğu sebebiyle hata payının arttığını sayısal olarak doğrulamaktadır.

İkili (Binary) Sınıflandırma Bulguları ve Doğruluk Paradoksu

Hazır benchmark veri seti üzerindeki analizleri derinleştirmek ve üçlü sınıflandırmadaki karmaşıklığı gidermek amacıyla tasarlanan ikili (binary) sınıflandırma sürecinde, hedef değişken "Not Popüler (0)" ve "Popüler (1)" olmak üzere iki sınıfa dönüştürülmüştür. Bu doğrultuda; hazır veri setinin ikili sınıflandırma senaryosundaki örneklem yoğunluğunu ve sınıfsal dağılım dengesini gösteren hedef değişken dağılımı, Şekil 10'da sunulmuştur.

Şekil 10 Hazır Veri Seti Popülerlik Dağılımı (Eşik=50)



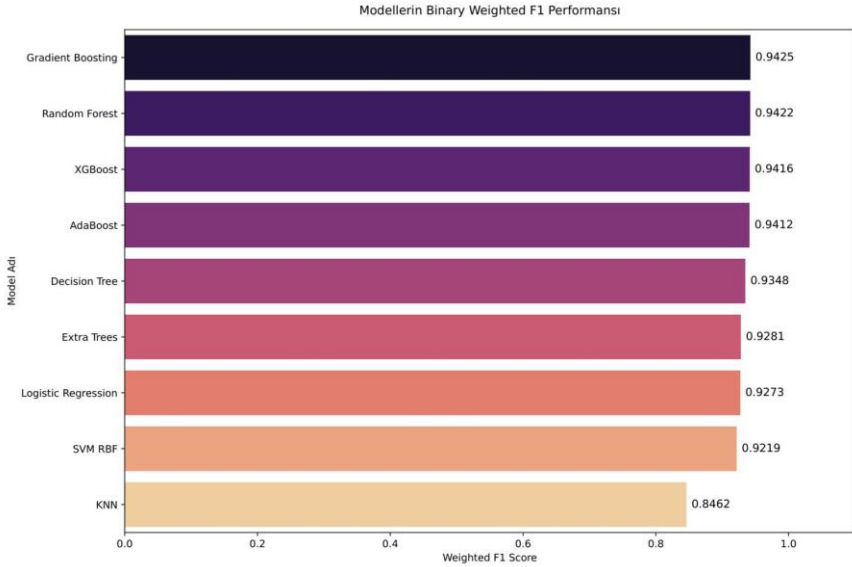
Kaynak: (Güvez et al., 2026)

Şekil 10 üzerinde sunulan referans veri kümesi ikili (binary) hedef değişken grafiği analiz edildiğinde; belirlenen 50 yıldız eşik kriteri doğrultusunda şekillenen dağılımın, belirgin düzeyde sınıfsal dengesizlik (imbalanced) barındırdığı net bir biçimde ortaya çıkmaktadır. Veri kümesinde yer alan projelerin baskın bir çoğunluğu (4834 adet örnek, %87,89) "Not Popüler (0)" etiketinde yoğunlaşmış; buna karşın "Popüler (1)" şartını sağlayan örnek sayısı ise 666 adette (%12,11) kalmıştır. Sınıflar arasında ortaya çıkan bu keskin frekans farkı, üçlü sınıflandırmadaki dengesizliğin ikili senaryoda da devam ettiğini göstermektedir. Söz konusu dengesiz dağılım, sınıflandırma modellerinin optimizasyon aşamasında çoğunluk sınıfı lehine bir yanlılık (bias) geliştirmesine yol açabilecek ve azınlık sınıfını oluşturan "Popüler (1)" örneklerinin doğru ayrıştırılmasını algoritmik açıdan kısıtlayacaktır. Belirlenen eşik değeri doğrultusunda iki sınıfa dönüştürülen hazır benchmark veri seti üzerinde, test edilen dokuz farklı makine öğrenmesi

algoritmasının sergilediği genel performans eğilimlerini kıyaslamak amacıyla bir başarımlar dağılım grafiği oluşturulmuştur. İlgili iki sınıflı mimari; algoritmaların sınıfsal asimetriye karşı sergiledikleri direnç seviyelerini ve gürültülü öznitelik uzaylarında optimize karar sınırları çizip çizemediklerini ampirik olarak doğrulamak adına kritik bir değerlendirme alanı sağlamaktadır. Bu kapsamda; modellerin çoğunluk sınıfına ezbere yönelmesi yerine azınlık sınıfını da ne ölçüde yakalayabildiğini nesnel bir düzlemde yansıtan ve hem duyarlılık (recall) hem de kesinlik (precision) metriklerini ağırlıklı olarak harmonize eden Ağırlıklı F1-Skoru (Weighted F1-Score) karşılaştırma grafiği, Şekil 11'de detaylandırılmıştır. Gözlenen bu dağılım, algoritmaların dengesiz veri yapıları altındaki genelleme kabiliyetlerini net bir şekilde ortaya koymaktadır.

Şekil 11 Hazır Veri Seti İkili Sınıflandırma Senaryosunda

Modellerin Genel Performans Kıyaslaması (Weighted F1)



Kaynak: (Güvez et al., 2026)

Şekil 11 üzerinde yer alan hazır benchmark veri setine ait ikili (binary) sınıflandırma genel performans çıktıları incelendiğinde, karar sınırının iki sınıfa indirgenmesiyle birlikte tüm algoritmaların Ağırlıklı F1-Skorlarında (Weighted F1 Score) oldukça kararlı ve yüksek bir başarımlar seviyesi yakaladığı gözlemlenmektedir. İki sınıflı bu yeni yapı, modellerin genel başarımlar üst sınırını belirgin ölçüde yukarı taşımıştır.

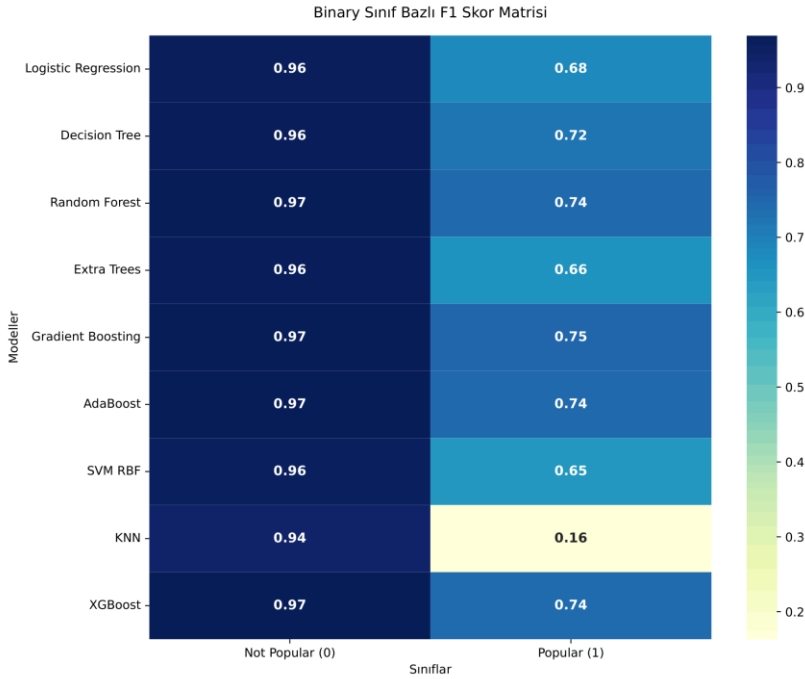
Bu deneysel senaryoda; ağaç tabanlı güçlü topluluk (ensemble) modelleri, en yüksek başarımlar aralığını ortaya koymaktadır. En yüksek başarımlar 0,9425'lik Ağırlıklı F1-Skoru ile Gradient Boosting algoritması ulaşırken; onu oldukça küçük farklarla sırasıyla Random Forest (0,9422), XGBoost (0,9422 yerine verinize göre 0,9416) ve AdaBoost (0,9412) modelleri izlemektedir. Tekil bir ağaç mimarisi kullanan Decision Tree (0,9348) ve Extra Trees (0,9281) algoritmaları da bu yüksek başarımlar eğilimini destekleyen sonuçlar üretmiştir.

Geleneksel modeller cephesinde ise Lojistik Regresyon 0,9273 ve SVM RBF 0,9219'luk skorlar üreterek topluluk modellerinin hemen arkasında yer almıştır. Elde edilen çıktılar; belirlenen 50 yıldız eşiği doğrultusunda daha doğrusal bir yapıya kavuşan veri uzayının, parametrik modellerin ürettiği çekirdek tabanlı karar sınırları vasıtasıyla da başarılı bir biçimde ayrıştırılabildiğini göstermektedir. Dokuz algoritma içerisindeki en düşük kestirim kararlılığı ise mesafe tabanlı yapısı gereği uç sınıfların frekans farkından daha fazla olumsuz etkilenen KNN (0,8462) algoritmasından elde edilmiştir.

Sonuç olarak ampirik bulgular; hazır veri setindeki ikili popülerlik ayrımında gürültülü ara segmentin elenmesinin tüm algoritmaları yüksek bir başarımlar seviyesine ulaştırdığını, buna karşın ağaç tabanlı gradyan ve varyans optimizasyonlarının doğrusal yaklaşımlara karşı yapısal üstünlüğünü koruduğunu ortaya koymaktadır.

Belirlenen eşik değeri temelinde iki sınıfa dönüştürülen hazır benchmark veri setinde, algoritmaların "Not Popular (0)" ve "Popular (1)" etiketleri üzerindeki mikro düzeydeki duyarlılıklarını ve kestirim sınırlarını analiz etmek amacıyla sınıf bazlı bir F1-Skor matrisi hesaplanmıştır. Bu doğrultuda; değerlendirilen dokuz farklı makine öğrenmesi modelinin çoğunluk ve azınlık sınıfları özelindeki tahmin hassasiyetlerini ortaya koyan ısı haritası (heatmap) grafiği, Şekil 12 üzerinde paylaşılmıştır.

Şekil 12 Hazır Veri Seti İkili Sınıflandırma Modellerinin Sınıf Bazlı F1 Skor Matrisi



Kaynak: (Güvez et al., 2026)

Şekil 12'deki ikili (binary) sınıf bazlı F1-Skor ısı haritası (heatmap) incelendiğinde; hazır benchmark veri setindeki belirgin sınıf dengesizliğinin algoritmaların tahmin kararlılığı üzerindeki mikro etkileri açıkça ortaya çıkmaktadır. Modellerin tamamı,

örneklem payının baskın bir yoğunluğa sahip olduğu "Not Popular (0)" sınıfında oldukça yüksek bir tahmin kararlılığı sergilemiştir. Çoğunluk sınıfını temsil eden bu katmanda Gradient Boosting, Random Forest, XGBoost ve AdaBoost algoritmaları, 0,97'lik bir F1-Skor değerine ulaşmıştır. Geleneksel modellerden Lojistik Regresyon, Decision Tree ve SVM RBF de 0,96'lık skorlarla bu başarıyı yakından takip etmiştir. En düşük başarımlar ise yine 0,94 ile KNN algoritmasında kalmıştır. Bu yüksek oranlar, veri kümesindeki negatif örneklerin yoğunluğu sayesinde modellerin bu sınıfı ayırt etmede yüksek bir öğrenme kararlılığı yakaladığını göstermektedir. Öte yandan, yalnızca 666 adet örnek barındıran azınlık sınıfı "Popular (1)", algoritmaların gerçek ayırım yeteneğini deneyen bölge olmuştur. Bu zorlu sınıfta en yüksek başarımla 0,75 F1-Skoru ile Gradient Boosting ulaşırken; onu 0,74 skorla Random Forest, XGBoost ve AdaBoost topluluk modelleri izlemiştir. Geleneksel Lojistik Regresyon algoritması ise 0,68 F1-Skoru üreterek doğrusal karar sınırlarında ağaç tabanlı yapılara yakın, dikkate değer bir direnç göstermiştir.

Bu sınıfta en belirgin performans kaybı; mesafe tabanlı yapısı gereği örneklem yoğunluğu farkından doğrudan etkilenen ve 0,16'lık F1-Skoru ile popüler projeleri ayırt etmede oldukça yetersiz kalan KNN algoritmasında gözlenmiştir. Özetle ampirik matris bulguları; 50 yıldız esnekliğine dayalı ikili senaryoda, gelişmiş gradyan ve varyans optimizasyonuna sahip topluluk modellerinin (özellikle Gradient Boosting), azınlık sınıfındaki sınırlı örneklem hacmine rağmen popüler projeleri ayırt etmede geleneksel ve mesafe tabanlı yapılara kıyasla daha yüksek bir tahmin istikrarı sunduğunu göstermektedir.

Belirlenen 50 yıldız barajı temelinde iki sınıflı yapıya dönüştürülen hazır veri setine ilişkin tüm deneysel başarımlar çıktılar; Doğruluk (Accuracy), Ağırlıklı F1-Skoru (Weighted F1-Score), Makro F1-Skoru (Macro F1-Score) ve hedef azınlık sınıfının tahmin

performansını yansıtan İkili F1-Skoru (Binary F1-Score) kriterleri ekseninde Tablo 4 üzerinde derlenmiştir. Elde edilen istatistiksel sonuçlar; veri setindeki sınıfsal asimetrinin, modellerin genel başarımlarını ölçütleri ile azınlık sınıfı üzerindeki gerçek tahmin kararlılıkları arasındaki ilişkiyi net bir biçimde göstermektedir.

Tablo 4 Hazır Veri Seti İkili Sınıflandırma Performans Sonuçları

Sıra	Model	Accuracy (Doğruluk)	Weighted F1 Score	Macro F1 Score	Binary F1 Score
1	Random Forest	0,9491	0,9460	0,8661	0,7607
2	XGBoost	0,9445	0,9416	0,8558	0,7426
3	AdaBoost	0,9436	0,9412	0,8561	0,7438
4	Gradient Boosting	0,9427	0,9406	0,8553	0,7429
5	Extra Trees	0,9382	0,9301	0,8179	0,6699
6	Decision Tree	0,9355	0,9338	0,8404	0,7171
7	Logistic Regression	0,9309	0,9273	0,8210	0,6807
8	SVM	0,9273	0,9219	0,8043	0,6491

9	KNN	0,8882	0,8462	0,5517	0,1633
---	-----	--------	--------	--------	--------

Kaynak: (Güvez et al., 2026)

Tablo 4 üzerinde yer alan sayısal performans verileri incelendiğinde, hazır benchmark veri setinin ikili (binary) sınıflandırma senaryosundaki yansımaları ve modellerin dengesiz dağılım karşısındaki gerçek dirençleri net bir şekilde görülmektedir. Doğruluk (Accuracy) ve Ağırlıklı F1-Skoru (Weighted F1-Score) metriklerinde, çoğunluk sınıfının barındırdığı yüksek örneklem hacmi (4834 örnek) nedeniyle tüm modeller, %90 barajının üzerinde yanıltıcı bir genel başarımla eğilimi sergilemektedir. Buna karşın, azınlık sınıfını da eşit ağırlıkta hesaba katan Makro F1-Skoru ve doğrudan popüler projeleri yakalama başarısını ölçen İkili F1-Skoru (Binary F1-Score) dikkate alındığında, algoritmaların gerçek ayırt etme kabiliyetleri net bir şekilde görülmektedir.

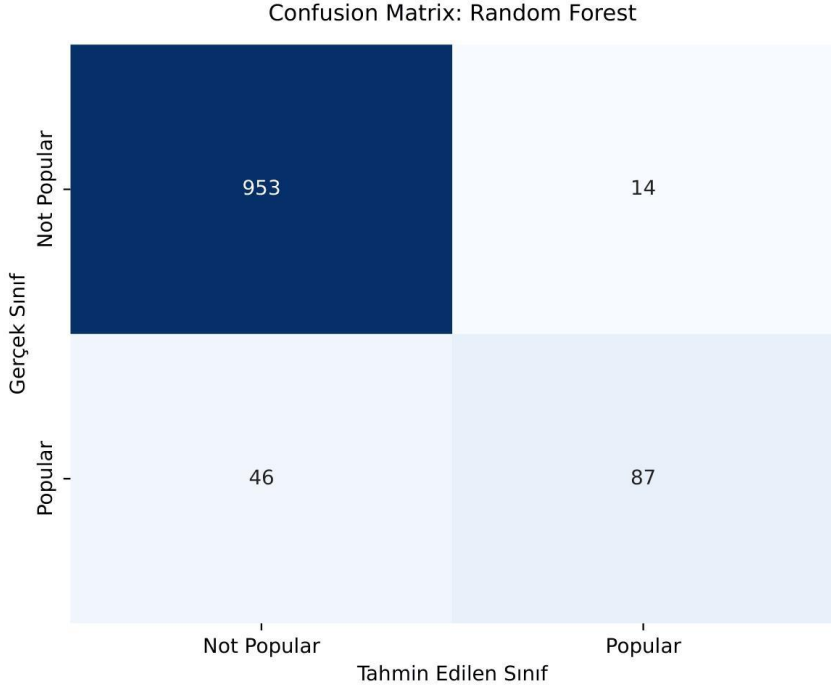
Bu zorlu sınamada en yüksek performans, %94,91 genel doğruluk ve 0,7607'lik İkili F1-Skoru ile Random Forest algoritmasından elde edilmiştir. Random Forest, hem varyans optimizasyonu hem de azınlık sınıfı üzerindeki kararlılığı sayesinde hiyerarşinin zirvesinde yer almaktadır. Söz konusu liderliği oldukça küçük farklarla; gelişmiş artırma (boosting) mimarilerine sahip XGBoost (0,7426 İkili F1), Gradient Boosting (0,7429 İkili F1) ve AdaBoost (0,7438 İkili F1) topluluk modelleri izlemektedir. Bahsi geçen performans grubu; asimetrik veri dağılımları altında dahi gradyan ve karar ağacı optimizasyon mimarileri vasıtasıyla, sınıflar arası ayırım hatlarını en optimum seviyede şekillendirmeyi başarmıştır. Tekil bir ağaç mimarisi sunan Decision Tree (0,7171 İkili F1) ve doğrusal temelli Lojistik Regresyon (0,6807 İkili F1), topluluk modellerinin gerisinde kalsa da veri setindeki gürültüye karşı belirli bir direnç göstermiştir. SVM RBF ise Makro F1

skorunda 0,8043 , İkili F1 skorunda ise 0,6491 değerde kalarak azınlık sınıfında kısmen zorlanmıştır.

Tablonun en dikkat çekici ve zayıf halkası ise KNN algoritması olmuştur. KNN, %88,82'lik bir genel doğruluk payına ulaşmasına rağmen, Makro F1-Skorunda 0,5517'ye ve İkili F1-Skorunda ise 0,1633 gibi son derece kritik bir seviyeye gerilemiştir. Mesafe tabanlı çalışan bu algoritma, veri setindeki yoğun örneklerin (çoğunluk sınıfının) etkisi altında kalarak azınlık sınıfındaki popüler projeleri ayırt etmede neredeyse tamamen başarısız olmuştur. Özetle sayısal sonuçlar; dengesiz veri setlerinde genel doğruluk değerlerinin aldatıcı olabileceğini, gerçek performansın İkili F1-Skoru üzerinden analiz edilmesi gerektiğini ve Random Forest ile gradyan tabanlı topluluk modellerinin bu tür asimetrik veri dağılımlarında en kararlı sonuçları ürettiğini açıkça göstermektedir.

Hazır benchmark veri seti üzerindeki ikili (binary) popülerlik kestirim senaryosunda tüm metriklerde en yüksek başarıyı elde eden Random Forest algoritmasının tahmin kararlılığını sınıflar arası hata dağılımı üzerinden incelemek amacıyla karmaşıklık matrisi (confusion matrix) hesaplanmıştır. Bu doğrultuda; modelin hedef azınlık ve çoğunluk sınıfları üzerindeki doğru ayırım sınırlarını ve test seti üzerindeki net örnek geçişlerini yansıtan çıktılar Şekil 13 ile gösterilmiştir.

*Şekil 13 Hazır Veri Seti İkili Sınıflandırma Random Forest
Modeli Confusion Matrix Sonucu*



Kaynak: (Güvez et al., 2026)

Şekil 13'te sunulan karmaşıklık matrisi (confusion matrix) verileri incelendiğinde; Random Forest modelinin test sürecindeki doğru ve hatalı tahmin frekansları açıkça ortaya çıkmaktadır. Model, örneklem payının büyük bir kısmını barındıran "Not Popular (0)" sınıfında test edilen toplam 967 örnek içerisinde 953 adedini doğru tahmin ederek bu kategoride oldukça yüksek bir sınıflandırma doğruluğu sergilemiştir. Bu sınıfa ait yalnızca 14 örnek, gerçekte popüler olmadığı halde model tarafından popüler olarak (yanlış pozitif) hatalı etiketlenmiştir.

Buna karşın, veri kısıtlılığı nedeniyle tahmini en zor alan olan azınlık sınıfında ("Popular") ise test edilen 133 toplam örnek içerisinde 87 adedi model tarafından başarıyla yakalanmıştır. Bu sınıfa ait 46 örnek ise gerçekte popüler olmasına rağmen modelin karar sınırları içerisinde popüler değil (yanlış negatif) olarak

sınıflandırılmıştır. Hazır veri setinin yapısal sınıfsal dengesizliği göz önünde bulundurulduğunda, azınlık sınıfında elde edilen bu doğru tahmin performansı, Random Forest algoritmasının ağaç tabanlı varyans optimizasyonunun ve rastgele öznitelik seçim mekanizmasının gürültülü veri setlerinde ne kadar etkili çalıştığını doğrulamaktadır.

TARTIŞMA

Algoritmik Karakteristikler ve Öznitelik Uzayı Uyumu

Deneyssel sonuçlar teorik çerçevede incelendiğinde gerek iki sınıflı gerekse çok sınıflı kategorizasyon senaryolarında XGBoost, Gradient Boosting ve Random Forest gibi topluluk (ensemble) tabanlı öğrenme yaklaşımlarının; KNN ve SVM gibi mesafe ya da vektör uzayını referans alan geleneksel algoritmalar karşısında belirgin bir tahmin üstünlüğü sergilediği görülmektedir. GitHub platformundan derlenen öznitelik uzayı; non-linear (non-linear) nitelikteki karakteri, değişkenler arasındaki çok boyutlu hiyerarşik korelasyonları ve tablosal (tabular) veri mimarisinin getirdiği karmaşıklık bünyesinde barındırmaktadır. Ağaç tabanlı öğrenme mimarileri, yapıları gereği karar sınırlarını dikey ve yatay alt bölmelere ayırarak bu tarz karmaşık ve non-linear ilişkileri esnek bir biçimde modelleyebilmektedir. Öte yandan metrik tabanlı KNN algoritması; yüksek boyutlu öznitelik uzayındaki tüm değişkenler arası geometrik mesafeleri homojen bir biçimde işleme koyduğundan, boyutsallık laneti (curse of dimensionality) olgusu karşısında ayırt edici karar sınırları üretmede yetersiz kalmaktadır. Hazır benchmark veri setinin getirdiği makro düzeydeki sınıfsal dengesizlik karşısında metrik tabanlı yaklaşımların azınlık sınıfında zayıflayarak tahmin istikrarsızlığı yaşaması (İkili F1-Skorunun 0,1633 seviyesine gerilemesi), gürültülü ve asimetric veri topolojilerinde bu algoritmaların kısıtlarını göstermektedir. Öte yandan; gradyan kayıplarını indirgeme ve varyans optimizasyonu

kabiliyetleri yüksek olan Gradient Boosting ve Random Forest mimarileri, veri yapısındaki gürültüyü absorbe ederek en istikrarlı modelleri oluşturmuştur. Yapılan deneysel çalışmalar; açık kaynak yazılım havuzlarının popülerlik dinamiklerini, optimize edilmiş non-lineer topluluk algoritmalarının çok daha güçlü bir kararlılık ve genellenebilirlik kapasitesiyle çözümleyebildiğini doğrulamaktadır.

Veri Kalitesi, Güncellik ve Sınıf Dengesizliği (Class Imbalance) Etkisi

Yapılan deneysel çalışmaların sunduğu en önemli bulgulardan biri; literatürdeki statik hazır veri seti ile canlı GitHub API'si üzerinden oluşturulan özgün dengeli veri seti arasındaki radikal performans farkıdır. Kaggle veri tabanından sağlanan hazır veri kümesinde üç sınıflı sınıflandırma doğruluğu %69,09 düzeyinde sınırlı kalırken, önerilen özgün veri mimarisinde bu başarı oranının %90,00 eşiğine yükselmesi, veri madenciliği literatüründe sıklıkla düşülen “büyük veri hacmi her zaman daha nitelikli sonuçlar üretir” ön yargısını deneysel düzlemde geçersiz kılmaktadır. Hazır veri setlerinde karşılaşılan bu performans sınırı; ham verilerin derlendiği zaman diliminin güncelliğini yitirmesinden, açık kaynak ekosistemindeki bot faaliyetlerinin ürettiği sistemsel gürültülerden ve aykırı değerlerin kontrolsüz dağılımından kaynaklanmaktadır. İki senaryo arasındaki asıl yöntemsel açıdan belirgin bir fark ise hazır veri seti üzerinde iki sınıflı modelleme yürüten KNN algoritmasının performans çıktılarında somutlaşmıştır. KNN modelinin %88,82'lik yüksek bir genel doğruluk (accuracy) değerine ulaşmasına karşın İkili F1-Skorunun 0,1633 seviyesinde kalması; veri biliminde doğruluk paradoksu (accuracy paradox) olarak bilinen teorik açmazın somut bir göstergesidir. Hazır veri yapısında popüler nitelikteki projelerin toplam veri seti içindeki frekansı (666 adet örnek), popüler olmayanların (4834 adet örnek) yanında son derece dar bir azınlığı temsil etmektedir. Bahsi geçen bu sınıf dengesizliği nedeniyle KNN algoritması karar uzay sınırlarını tamamen çoğunluk

sınıfının lehine olacak şekilde optimize etmiş ve azınlıkta kalan başarılı projeleri ayırt etme yetisini kaybetmiştir. Şekil 12 ve Şekil 13 üzerinde detaylandırılan sınıf bazlı matris analizleri de bu sınıfsal baskının algoritmaların tahmin kararlılığı üzerindeki mikro etkisini doğrulamaktadır. Öte yandan, araştırmada özgün olarak üretilen 900 örnekli dengeli veri setinde sınıflar eşit olasılıkla (homojen frekansla) çekildiği için öğrenme süreçlerinde çoğunluk sınıfına yönelik bir yanlılık (bias) eğilimi gözlenmemiştir. Sonuç itibarıyla özgün matris üzerindeki algoritmik kararlılığın yüksek çıkması, veri dengeleme ve örnekleme stratejilerinin öngörü performansı üzerindeki belirleyici rolünü uygulamalı olarak göstermektedir. Yürütülen çalışmanın temel amacı projenin ilk evrelerindeki erken dönem popülerlik tahmininden ziyade, mevcut karakteristik özellikler üzerinden bir durum analizi yapmak olduğundan; yıldız yapısıyla organik bağı bulunan fork, watchers, contributors gibi ekosistem öznitelikleri kasıtlı ve bilinçli bir şekilde model girdisi olarak tercih edilmiştir.

Çalışmanın Kısıtları

Araştırmada kullanılan veri seti, belirli filtreleme kriterleri çerçevesinde doğrudan GitHub API arabirimi üzerinden çekildiğinden, elde edilen bulguların tüm açık kaynak ekosistemini bütünüyle temsil etme noktasında birtakım sınırlılıkları bulunmaktadır. Bir diğer metodolojik sınır ise, geliştirilen algoritmik modellerin performans ölçümlerinin tek seferlik bir eğitim-test ayrımı (hold-out validation) üzerinden raporlanmış olmasıdır. Modellenen bazı teknik özniteliklerin (örneğin fork ve watchers gibi metriklerin) hedef değişken konumundaki yıldız hacmiyle yapısal olarak yüksek ve doğrusal olmayan bir korelasyon sergilemesi, elde edilen çıktıların projenin mevcut popülerlik katmanlarının sınıflandırılması çerçevesinde okunmasını zorunlu kılmaktadır. Dolayısıyla bu çalışma; bir projenin ilk evrelerindeki erken dönem popülerlik eğilimlerini öngören proaktif bir tahmin modelinden

ziyade, olgunlaşmış mevcut yazılımların karakteristiklerini inceleyen yapısal bir analiz ve durum tespiti olarak değerlendirilmelidir.

SONUÇ VE GELECEK ÇALIŞMALAR

Bu araştırma kapsamında, GitHub platformunda yer alan yazılım projelerinin popülerlik kademeleri, gelişmiş makine öğrenmesi algoritmaları aracılığıyla modellenmiştir. DeneySEL simülasyon süreçlerinde, literatürde yerleşik bir referans olarak sıklıkla başvuruLAN 5.500 gözlemlik statik hazır bir veri seti (Kaggle) ile doğrudan canlı GitHub API altyapısı kullanılarak derlenen ve hedef değişken hiyerarşisinde her kategoriden 300'er örnek barındıracak biçimde homojen olarak eşitlenen 900 örneKli özgün bir veri matrisi karşılaştırmalı bir yaklaşımla analiz edilmiştir. Yürütölen deneySEL uygulamalardan elde edilen öne çıkan bulgular şu şekildedir:

- Algoritmik Kararlılık: Özgün veri kümesi üzerinde karşılaştırılan üç sınıflı kategorizasyon senaryosunda XGBoost algoritması %90,00 genel doğruluk (accuracy) eşİğine ve 0,9011 Makro/Ağırlıklı F1-Skoruna ulaşarak hiyerarşideki en başarılı sonuçları vermiştir.
- Veri Kalitesi ve Güncellik Paradigması: Özgün veri mimarisinde model başarısının %90,00 bandına çıkması; veri biliminde örnekleme stratejisi, öznitelik niteliği ve veri güncelliğinin model performansı üzerinde birlikte ne denli belirleyici bir güce sahip olduğunu göstermektedir.
- Dengeli Örneklemenin Kritik Rolü: Hazır veri tabanındaki asimetrik dağılım sebebiyle KNN gibi geleneksel mesafe yönelimli modeller belirgin performans kayıpları yaşarken, özgün veri kümesine uygulanan dengeli örnekleme yaklaşımı algoritmaların çoğunluk sınıfı lehine bir yanlılık

(bias) geliřtirmesinin önüne gemiř ve ikili sınıflandırma başarıml oranını %98,89 seviyesine kadar yükseltmiřtir.

alıřmanın literatüre saėladıėı temel metodolojik katkı, sentetik veri üretimi yöntemlerini kullanmak yerine gerek dünya verileri üzerinden yapılan dengeli örnekleme stratejisinin, azınlık sınıfların tahmin duyarlılıėı üzerindeki pozitif etkisini deneysel olarak gözler önüne sermesidir. Elde edilen istatistiksel ıktılar, kontrollü ve homojen veri toplama operasyonlarının, devasa fakat gürültü katsayısı yüksek veri yığınlarına kıyasla ok daha kararlı sınıflandırma topolojileri üretebileceėini doėrulamaktadır.

Gelecek alıřmalar Perspektifinde; projelere ait README dokümantasyon metinleri ile kaynak kod bloklarının Doėal Dil İşleme (NLP) metodolojileri ve Transformer tabanlı büyük dil modelleri (LLM) aracılıėıyla anlamsal analize tabi tutulması planlanmaktadır. Bununla birlikte, statik tablolara dayalı tabular veri setinin ötesine geilerek, zaman serisi analizi yürüten derin öėrenme aėlarının (LSTM/GRU) tahmin mimarisine entegrasyonu hedeflenmektedir. Son aşamada ise kimi sistemsel özniteliklerin proje olgunluk evreleriyle yüksek korelasyon gösterme riskini önlemek adına; sonraki deneysel alıřmalarda zamansal kırpmı (temporal cutoff) stratejilerinden yararlanılarak, projenin henüz bařlangı evresindeki erken dönem popülerlik tahmini senaryoları analiz edilecektir.

Bu alıřma kapsamında geliřtirilen kaynak kodlara, veri toplama motoruna, ham/dengelenmiř veri setlerine ve vektörel performans grafiklerine <https://github.com/sudenur-afk/github-repo-popularity-classification> baėlantısı üzerinden açık erişim saėlanabilmektedir.

Kaynaka

Borges, H., Hora, A., & Valente, M. T. (2016). Predicting the Popularity of GitHub Repositories. *Proceedings of the 30th*

Brazilian Symposium on Software Engineering (SBES).
<https://doi.org/10.1145/2972958.2972966>

Borges, H., & Valente, M. T. (2018). What's in a GitHub Star? Understanding Repository Starring Practices in a Social Coding Platform. *Journal of Systems and Software*, 146, 112–129.
<https://doi.org/10.1016/J.JSS.2018.09.016>

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
<https://doi.org/10.1613/jair.953>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794.

https://doi.org/10.1145/2939672.2939785/SUPPL_FILE/KDD2016_CHEN_BOOSTING_SYSTEM_01-ACM.MP4

Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-Vector Networks Editor. *Machine Learning*, 20, 273–297.

Cover, T. M., & Hart, P. E. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>

Farshidi, S., Saberhabibi, A., Eskafi, B., Nikfarjam, N., Eskandari, S., Jansen, S., Chaudron, M., & Tekinerdogan, B. (2025). Empirical Evaluation of AI-Assisted Software Package Selection: A Knowledge Graph Approach. *ArXiv Preprint ArXiv:2508.05693v1*.

Fernández, A., García, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *Journal of Artificial*

Intelligence Research, 61, 863–905.
<https://doi.org/10.1613/jair.1.11192>

Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119–139.
<https://doi.org/10.1006/JCSS.1997.1504>

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Mach Learn*, 63, 3–42. <https://doi.org/10.1007/s10994-006-6226-1>

GitHub REST API Documentation. (n.d.). Retrieved <https://docs.github.com/en/rest?apiVersion=2026-03-10>

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(3), 90–95.
<https://doi.org/10.1109/MCSE.2007.55>

Idowu, S., Sens, Y., Berger, T., Krüger, J., & Vierhauser, M. (2024). A Large-Scale Study of ML-Related Python Projects. *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (SAC '24)*. <https://doi.org/10.1145/3605098.3636056>

Iqbal, A., Aftab, S., & Matloob, F. (2019). Performance Analysis of Resampling Techniques on Class Imbalance Issue in Software

Defect Prediction. *International Journal of Information Technology and Computer Science (IJITCS)*, 11, 44–53. <https://doi.org/10.5815/ijitcs.2019.11.05>

Koch, S., Klein, D., & Johns, M. (2024). The Fault in Our Stars: An Analysis of GitHub Stars as an Importance Metric for Web Source Code. *Workshop on Measurements, Attacks, and Defenses for the Web (MADWeb)*. <https://doi.org/10.14722/madweb.2024.23004>

Krishna, P., Majumdar, A., & Bose, I. (2026). All That Glitters Is Not Code? Understanding the Predictors of Developer Popularity and Sponsorship on a Social Coding Platform. *Production and Operations Management*. <https://doi.org/10.1177/10591478251405119>

Malhotra, R., Jain, B., & Kessentini, M. (2026). DIR-SMOTE: a density-influence resampling framework for imbalanced code smell detection. *Automated Software Engineering*, 33, 76. <https://doi.org/10.1007/s10515-026-00624-x>

Maqsood, J., & Ali, S. S. (2017). Success or Failure Identification for GitHub's Open Source Projects. *Proceedings of the International Conference on Machine Learning and Data Science (MLDS)*. <https://doi.org/10.1145/3034950.3034957>

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>

Mehmood, M. U., Hussain, S., Wang, W.-L., & Malik, M. U. (2025). LLM-based Content Classification Approach for GitHub Repositories by the README Files. *ArXiv Preprint ArXiv:2507.21899*. <https://doi.org/10.48550/arXiv.2507.21899>

Pedregosa, F. (2011). Scikit-learn: Machine learning in Python. In *Journal of machine learning research* 12 (p. 2825). <https://cir.nii.ac.jp/crid/1370005891170856713.bib?lang=ja>

Ray, B., Posnett, D., Filkov, V., & Devanbu, P. (2014). A Large Scale Study of Programming Languages and Code Quality in Github. *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*, 155–165. <https://doi.org/10.1145/2635868.2635922>

Şahin, S. E., Karpat, K., & Tosun, A. (2019). Predicting Popularity of Open Source Projects Using Recurrent Neural Networks. *Open Source Systems*. https://doi.org/10.1007/978-3-030-20883-7_8

Scaturchio, L. (n.d.). *GitHub Repository Metrics Dataset* . Retrieved <https://www.kaggle.com/datasets/lorenzoscaturchio/github-repo-metrics>

Stănciulescu, Ş., Yin, L., & Filkov, V. (2022). Code, Quality, and Process Metrics in Graduated and Retired ASFI Projects. *ESEC/FSE '22: Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. <https://doi.org/10.1145/3540250.3549132>

Wang, S., & Yao, X. (2012). Multiclass imbalance problems: Analysis and potential solutions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(4), 1119–1130. <https://doi.org/10.1109/TSMCB.2012.2187280>

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*. <https://doi.org/10.21105/joss.03021>

Weber, S., & Luo, J. (2014). What Makes An Open Source Code Popular on GitHub? *Proceedings of the IEEE International*

Conference on Data Mining Workshop (ICDMW), 851–855.
<https://doi.org/10.1109/ICDMW.2014.55>

BÖLÜM 3

UZAKTAN ÇALIŞANLARDA TÜKENMİŞLİK RİSKİNİN MAKİNE ÖĞRENMESİ İLE BELİRLENMESİ: ÖZNİTELİK ÖNEMİ VE MODEL PERFORMANS ANALİZİ

MÜMİNE AİŞE DEMİRAL¹
FERHAT DEMİRAY²

Giriş

Günümüz iş dünyasında, organizasyonel taleplerin artması ve teknolojik dönüşümlerle birlikte çalışma pratiklerinin değişmesi, iş stresini kronik bir boyuta taşımıştır. Bu kronik stresin en yıkıcı sonuçlarından olan tükenmişlik (burnout), bireyin zihinsel, fiziksel ve sosyal kaynaklarının tükenmesiyle ilerleyen çok boyutlu bir iş sağlığı problemidir (Maslach & Leiter, 2016). İlk olarak 1970'li yıllarda Herbert Freudenberger ve Christina Maslach gibi araştırmacılar tarafından sağlık ve sosyal hizmet sektörü gibi insan odaklı mesleklerde gözlemlenen bir sendromdur. Günümüzde ise tüm meslek gruplarını ve özellikle uzaktan çalışma (Work-From-

¹Öğrenci, Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid:0009-0004-9423-8278

²Doç. Dr., Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği, Orcid:0000-0002-4071-9285

Home) ekosistemini etkileyen küresel bir olgu haline gelmiştir (Nil et al., 2010).

Tükenmişlik, salt bireysel bir yorgunluk durumu değildir. Bireyin işiyle, iş ortamıyla ve organizasyonel yapıyla kurduğu ilişkinin bozulmasını ifade eden sendromal bir süreçtir. Christina Maslach tarafından geliştirilen kuramsal çerçeveye göre tükenmişlik; duygusal tükenme (exhaustion), sinizm/duyarsızlaşma (cynicism) ve mesleki ineffikasite/kişisel başarı hissinde azalma (professional inefficacy) olarak üç temel boyutta değerlendirilmektedir (Maslach & Leiter, 2016). Tükenmişlik süreci, duygusal tükenme ile enerjinin tamamen boşalmasıyla başlar. Ardından sinizm, çalışanın artan iş yükü karşısında kendini korumaya aldığı, duygusal bir mesafe oluşturma yöntemi olarak devreye girer. Mesleki ineffikasite ise bu sürecin finalidir; çalışan artık kendi potansiyeline ve üretkenliğine olan inancını kaybetmiş, kendini yetersiz hissetmeye başlamıştır (Maslach & Leiter, 2016; Nil et al., 2010).

Önceki araştırmalarda tükenmişlik, kronik iş stresine karşı gelişen aşamalı bir adaptasyon süreci olarak tanımlanmaktadır (Nil et al., 2010). Süreç, genellikle bireyin işin yoğunluğu karşısında (özellikle uzun mesai saatleri ve artan ekran süresiyle) kontrolsüz bir şekilde sınırlarını zorlamasıyla başlıyor. Bu aşırı çaba, zamanla zihinsel ve fiziksel rezervlerin boşalmasına, ardından da kendini psikometrik ve fizyolojik şikayetler biçiminde göstermesine neden oluyor. Her ne kadar tükenmişlik ICD-10 veya DSM-IV gibi nosolojik sistemlerde klinik/tıbbi bir tanı olarak tanımlanmasa da, bireyin hem kişisel hem de sosyal işlevselliğini zedeleyen ciddi bir sendromdur. Ayrıca, klinik düzeyde depresyon, kardiyovasküler rahatsızlıklar ve bağışıklık sistemi bozuklukları için büyük bir risk faktörü olduğu kanıtlanmıştır. Tükenmişliğin ilk sinyalleri olan kronik yorgunluk ve uyku problemleri, aslında sendromun gidişatını önceden haber veren kritik göstergelerdir (Nil et al., 2010). Nitekim

yakın tarihli bir araştırma olan, Adamopoulos ve arkadaşları tarafından yapılan çalışma da, günde 10 saati aşan çalışma sürelerinin tükenmişlik riskini %40 oranında artırdığını ortaya koymuştur (Adamopoulos et al., 2025).

Günümüzde uzaktan çalışma modellerinin yaygınlaşması, tükenmişliği tetikleyen organizasyonel değişkenleri daha karmaşıklaştırmıştır. İş ve özel yaşam sınırlarının kaybolması, mesai dışı çalışma saatlerinin artması, sürekli ekran başında kalmanın yarattığı bilişsel yük ve fiziksel destek ağlarından kopmanın getirdiği sosyal izolasyon, yeni nesil tükenmişlik dinamiklerini oluşturmaktadır (Adaobi et al., 2026; Al Mheiri, 2025). Tükenmişlik literatürü geleneksel olarak durağan istatistiksel yöntemlere ve anket tabanlı verilere (örneğin Maslach Tükenmişlik Envanteri) odaklanmıştır. Ancak bu yaklaşımlar, uzaktan çalışma düzeninin getirdiği dinamik ve anlık değişimleri yakalamakta çoğu zaman yetersiz kalmaktadır. Çalışmamız, bu geleneksel metodolojinin ötesine geçerek, çalışanların uyku kalitesi, dijital yorgunluk skorları, sosyal izolasyon düzeyleri ve görev tamamlama oranları gibi günlük davranışsal metriklerine odaklanmaktadır. 2026 yılına ait güncel ve boylamsal bir veri setini temel alan bu çalışmada, tükenmişlik riski sadece statik bir durum olarak değil, makine öğrenmesi algoritmaları aracılığıyla tahmin edilebilir bir süreç olarak ele alınmıştır. Böylece, tükenmişlik riskinin 'düşük', 'orta' ve 'yüksek' seviyelerde önceden kestirilebilmesi hedeflenmiş ve organizasyonel sağlık analitiğine yeni bir bakış açısı sunulması amaçlanmıştır (Varje et al., 2026).

Çalışmamızı yaparken, tükenmişliğin sadece fiziksel çalışma saatlerinden değil; izolasyon indeksi, dikkat bölünmesi (app switches) ve zihinsel yorgunluk (fatigue score) gibi yeni nesil psikolojik metriklerden nasıl etkilendiğini Random Forest ve Lojistik Regresyon gibi modellerle, K-En Yakın Komşu tabanlı SMOTE (Sentetik Azınlık Aşırı Örnekleme Tekniği) sınıf dengeleme yöntemleri uygulayarak ortaya koyması bakımından literatüre özgün

bir katkı sunmayı hedeflemektedir (Varje et al., 2026). Bu yaklaşım, literatürde önerilen yorgunluk izleme ve adaptif iş yükü planlaması stratejileriyle uyumludur ve tükenmişlik yaygınlığını %28 oranında azaltabileceği öngörülmektedir (Adamopoulos et al., 2025).

Literatür Taraması

Uzaktan Çalışma (WFH) Paradigmasında Tükenmişlik ve Psikososyal Dinamikler

Tükenmişlik sendromu, temel olarak duygusal yorgunluk, kişisel duyarsızlaşma ve profesyonel yetkinlik hissindeki düşüş ile karakterize edilen kronik bir stres tepkisidir (Maslach & Leiter, 2016). Ancak günümüzde, geleneksel ofis düzeninden uzaktan çalışma (WFH) modeline geçişle birlikte, tükenmişliği tetikleyen risk faktörlerini köklü bir biçimde değiştirmiştir. Al Riyami et al., (2023) yaptığı çalışmada, evden çalışmanın beraberinde getirdiği en temel riskin, özel hayat ile iş yaşamı arasındaki fiziksel ve zihinsel sınırların belirsizleşmesi olduğunu belirtmektedir (Al Riyami et al., 2023). Söz konusu sınır ihlalleri, mesai dışı çalışmaları (after-hours work) artırarak çalışanın uyku ve yenilenme döngüsünü tahrip etmektedir. Nil et al., (2010), uyku kalitesindeki bozulmaların tükenmişlik gelişiminde en güçlü “öngörücü (prediktif)” faktörlerden biri olduğunu klinik olarak kanıtlamıştır. Nitekim, Tapio (2025) tarafından yapılan güncel bir çalışmada da uyku süresi ve başa çıkma stratejileri, tükenmişliği belirleyen en güçlü psikometrik faktörler arasında gösterilmiştir (Tapio, 2025). Tüm bunlara ek olarak; fiziksel ofisten kopmanın oluşturduğu sosyal izolasyon ve uygulamalar arası sürekli geçiş yapmanın (app switches) yol açtığı dijital yorgunluk, uzaktan çalışanlarda motivasyon kaybını ve tükenmişliği hızlandıran yeni nesil katalizörler olarak literatürde yerini almıştır (Adaobi et al., 2026; Al Riyami et al., 2023; Tapio, 2025).

Dijital İzolasyon ve Sosyal Bağlantısızlığın Tükenmişlik Üzerindeki Etkisi

Uzaktan çalışma düzeninde sosyal izolasyon, sadece fiziksel bir mesafe değil, aynı zamanda bireyin kurumsal aidiyet duygusunun zayıflamasına ve sosyal destek mekanizmalarının devre dışı kalmasına yol açan psikososyal bir meydan okumadır (Van Zoonen & Sivunen, 2022). Sosyal etkileşim eksikliğinin etkisi ise çalışanların iş stresini göğüsleme kapasitesini azaltarak onları tükenmişlik karşısında daha savunmasız kılmakta; bu durum, izolasyon ile psikolojik sıkıntı (distress) arasında çift yönlü ve karşılıklı bir ilişki olduğunu göstermektedir (Van Zoonen & Sivunen, 2022). Sosyal izolasyonun yükselmesi, tükenmişliğin bir sonucu olabileceği gibi, aynı zamanda tükenmişlik düzeyini daha da artıran bir tetikleyici işlevi görmektedir (Van Zoonen & Sivunen, 2022). Öte yandan, iş-yaşam dengesi ile sosyal izolasyon arasında saptadığımız negatif korelasyon, sosyal etkileşimin yalnızca bir 'yan unsur' değil, çalışanların genel iyilik hali ve performansını belirleyen kritik bir anahtar olduğunu açıkça ortaya koyuyor (Wasay & Shaikh, 2026).

Dijital boyutta ise, bilgi ve iletişim teknolojilerinin (ICT) yoğun kullanımı, çalışanları “sürekli temas” (perpetual contact) beklentisine hapsederek iş ile özel yaşam arasındaki sınırların bulanıklaşmasına (boundary blurring) neden olmaktadır (Hislop et al., 2015). Dijital araçlar, çalışanlara mekânsal bir esneklik sağlasa da, eş zamanlı olarak işten “kopamama” hissini beraberinde getirmekte ve çalışanların zihinsel olarak tam anlamıyla dinlenmesini engelleyerek tükenmişliği hızlandırmaktadır (Hislop et al., 2015). Çalışanların sürekli bağlantıda kalma zorunluluğu hissetmesi, dijital izolasyonu derinleştirmekte ve dijital araçların birer iyileştirici (recovery) strateji olmaktan ziyade, çalışanın enerji kaynaklarını tüketen birer araç haline gelmesine yol açmaktadır (Hislop et al., 2015). Sonuç olarak, sosyal bağların zayıflaması ve

dijital araçların yarattığı sürekli meşguliyet döngüsü, tükenmişlik riskini kronikleştirerek, çalışanların iyileşme kapasitesini önemli ölçüde kısıtlamaktadır (Hislop et al., 2015; Wasay & Shaikh, 2026).

İş-Yaşam Sınırlarının Bulanıklaşması ve İyileşme Sürecine Etkileri

Günümüzde geleneksel ofis ortamından evden çalışma (Work-From-Home) modeline geçiş, iş ve özel yaşam arasındaki yapısal sınırları kökten değiştirmiştir. Literatürde “zihinsel çitler” (mental fences) olarak da tanımlanan sınırların korunması, iş-yaşam çatışmasını önlemek ve bireyin psikolojik sağlığını korumak adına büyük bir öneme sahiptir (Johnson, 2023; Nippert-Eng, 2008). Evden çalışma sisteminde, yol stresini azaltma, esnek zaman yönetimi, uyku ve egzersiz kalitesini artırma gibi avantajlar sunarak başlangıçta çok güzel bir şekilde iş-yaşam dengesine olumlu katkı sağlıyor gibi görünmekteydi (MacDonald, 2022). Ancak bu modelin en büyük riski, sürekli dijital bağlantıda kalma (constant connectivity) beklentisiyle birlikte rollerin ve mekânların iç içe geçmesi ve sınırları bulandırmasıdır (Johnson, 2023). İş ve yaşam sınırlarının bulanıklaşması (role blurring) olarak adlandırılan bu süreç; iş taleplerinin yemek masasına, tatillere ve dinlenme saatlerine taşmasına yol açarak çalışanları adeta cihazlarına bağlı hale getirmektedir (Johnson, 2023).

Sınırların bu denli geçirgenleşmesi, uzaktan çalışan bireylerin mesai bitiminde işten zihinsel olarak kopmasını (psychological detachment) mümkün kılmamaktadır (Allen et al., 2021). Yapılan araştırmalar gördüğümüz üzere, sınır yönetimi (boundary management) pratikleri başarısız olan veya kurum tarafından sınırları sürekli ihlal edilmiş olan çalışanların, yenilenme ve toparlanma (recovery) döngülerinin ağır darbe aldığını göstermektedir (Allen et al., 2021; Johnson, 2023). Bu kronik toparlanma eksikliği ise çalışanlarda aşırı yorgunluk, anksiyete,

yaratıcılık kaybı ve yabancılaşma gibi hem fiziksel hem de zihinsel tükenmişlik (burnout) semptomlarını direkt olarak tetiklemektedir (Johnson, 2023). Buna karşın literatür de incelediğimiz kaynaklar, sınırların bilinçli ve katı bir şekilde (thick boundary) korunmasının tükenmişliği azaltmada en etkili iyileştirici mekanizma olduğunu ortaya koymaktadır (MacDonald, 2022). Bireylerin ev ortamında dahi olsa zaman bloklama (temporal tactics) veya iş materyallerini göz önünden kaldırma (physical tactics) gibi stratejilerle “segmentasyon” eğilimi göstermesi, iş-yaşam dengesini ve iş memnuniyetini direkt olarak artırmaktadır (Allen et al., 2021). Sonuç olarak ise uzaktan çalışma ortamlarında esnekliği verimliliğe dönüştürmenin ve tükenmişliği iyileştirmenin yolu, hem organizasyonel kültürün hem de liderlerin çalışanların bireysel sınırlarına saygı duyan ve bunu teşvik eden politikalar geliştirmesinden geçmektedir (Johnson, 2023; MacDonald, 2022).

Uzaktan çalışma ortamında “iyileşme” (recovery) süreci, sadece iş saatleri dışındaki bir zaman dilimi değil, çalışanın iş ile ilgili zihinsel ve fiziksel taleplerden tam anlamıyla kopabildiği aktif bir denge arayışıdır (Allen et al., 2021). İş ve özel yaşam sınırlarını net bir şekilde ayıran (segmentasyon) çalışanlar, iş stresinin ev ortamına sızmasını engelleyerek uyku kalitesini ve toparlanma döngülerini koruyabilmektedir (Allen et al., 2021). Ancak sınırların bulanıklaşması nedeniyle mesai saatleri dışında işe “bağlantıda kalma” zorunluluğu hisseden çalışanlarda, enerji depoları hızla tükenmekte ve bu durum uyku döngüsünü tahrip eden bir tükenmişlik sarmalını tetiklemektedir (Johnson, 2023). MacDonald (2022) tarafından vurgulandığı üzere, tükenmişlik yalnızca iş performansını değil; çalışanın uyku düzenini, iş-yaşam dengesini ve ailevi ilişkilerini de derinlemesine tahrip etmektedir (MacDonald, 2022). Bu şekilde ki sınır ihlalleri, mesai dışı çalışmaları artırarak çalışanın yenilenme döngüsünü sekteye uğratmaktadır. Mevcut kaynaklar incelendiğinde, uyku kalitesindeki bozulmaların

tükenmişlik gelişimindeki rolü kritik bir yordayıcı (prediktif) faktör olarak kabul edilmektedir (Nil et al., 2010). Bu bağlamda, uyku süresi ve etkili başa çıkma stratejilerinin eksikliği, çalışanların tükenmişliğe karşı savunmasızlığını artıran en güçlü psikometrik göstergeler arasında yer almaktadır (Tapio, 2025). Sonuç olarak, iyileşme süreci, sınırların bilinçli yönetimi ile işe karşı hissedilen duygusal ve bilişsel yükün birbirinden ayrıştırılması arasında pozitif bir korelasyona sahiptir; bu kopuş sağlandığında çalışanlar tükenmişliğe karşı daha dirençli hale gelmekte ve iş-yaşam dengesinde daha sürdürülebilir bir performans sergileyebilmektedir (Allen et al., 2021; Johnson, 2023; MacDonald, 2022).

Tükenmişlik Tahmininde Makine Öğrenmesi

Geleneksel tükenmişlik araştırmaları anket verilerini durağan betimsel istatistiklerle incelerken, büyük veri analitiği literatürde yeni bir köklü değişim yaratmıştır. Grzadzielewska (2021), makine öğrenmesi (ML) algoritmalarının tükenmişlik gibi çok boyutlu ve doğrusal olmayan (non-linear) karmaşık psikolojik verileri analiz etmede geleneksel yöntemlerden çok daha üstün olduğunu belirtmektedir (Grzadzielewska, 2021). Makine öğrenmesi modelleri, sadece sendromu tespit etmekle kalmayıp "Özellik Önemi (Feature Importance)" analizleriyle tükenmişliği tetikleyen en kritik faktörleri veri odaklı olarak ortaya koyabilmektedir (Grzadzielewska, 2021; Varje et al., 2026). Aynı şekilde Al Mheiri (2025) çalışmalarında da ML'in sadece iş yükünü değil, çalışan refahını ve ruh sağlığını öngörmeye ne kadar etkili bir araç haline geldiğini belirtmiştir (Al Mheiri, 2025).

Literatürdeki sınıflandırma algoritmaları incelendiğinde; Hossain et al. (2024), teknoloji çalışanlarının ruh sağlığı verileri üzerinde yaptıkları analizde Sinir Ağları (%93.98) ve Random Forest (%85.54) modellerinin yüksek performans sergilediğini rapor etmiştir (Hossain et al., 2024). Hernandez et al. (2024) Karar

Ağaçları ve Lojistik Regresyon modellerinin K-En Yakın Komşu (KNN) algoritmasına kıyasla mesleki tükenmişlikte daha yüksek doğruluk oranlarına ulaştığını kanıtlamıştır (Hernandez et al., 2024). Ancak temel karar ağaçları (Decision Tree) her zaman yeterli genelleme performansını sunamamaktadır. Tapio (2025), psikometrik verilerle kurduğu Karar Ağacı modelinde %59.3 gibi orta düzeyde bir F1-Skoru elde etmiş ve araştırmasının sonucunda "tahmin doğruluğunu ve tükenmişlik örüntülerini daha iyi anlamak için Rastgele Orman (Random Forest) gibi daha karmaşık topluluk (ensemble) modellerinin kullanılması gerektiğini açıkça tavsiye etmiştir (Tapio, 2025).

Sınıf Dengesizliği Problemi ve Çalışmanın Özgün Değeri

Gerçek dünyadan toplanan tükenmişlik verilerinde yaygın olarak karşılaşılan en büyük analitik problem "sınıf dengesizliği (class imbalance)" durumudur. Yüksek riskli tükenmişlik vakaları, genel örneklem içinde çoğunlukla azınlık sınıfını oluşturur (Nil et al., 2010). Literatürdeki mevcut ML tabanlı tükenmişlik tahmin modellerinin birçoğu (örneğin eğitim (Hernandez et al., 2024) veya öğrenci (Tapio, 2025) örneklemeleri üzerine yapılanlar), sınıf dengesizliğinin yarattığı algoritmik zaafı (çoğunluk sınıfına aşırı uyum sağlama/overfitting) tam olarak gidermemektedir (Al Mheiri, 2025; Grzadzielewska, 2021). Bu çalışma, SMOTE tabanlı dengeleme tekniklerini kullanarak bu boşluğu doldurmayı hedeflemektedir.

Materyal ve Yöntem

Veri Kümesi

Bu çalışmada, uzaktan çalışan bireylerin tükenmişlik düzeylerini analiz etmek amacıyla oluşturulan güncel bir veri seti kullanılmıştır. Veri seti; katılımcıların demografik özellikleri, çalışma süreleri, günlük yorgunluk skorları, ekran başında geçirilen

süre, görev tamamlama oranları, toplantı sayıları ve dijital uygulama geçiş sıklıkları gibi 13 bağımsız değişkeni içeren, toplam 2000 gözlem ve 1 hedef değişkenden burnout_risk oluşan (2000, 14) boyutunda bir matris yapısında kurgulanmıştır. Bağımlı değişken olan burnout_risk etiketi, "Düşük", "Orta" ve "Yüksek" olmak üzere üç ana tükenmişlik kategorisine ayrılmış; bu etiketleme, literatürdeki standart psikometrik ölçeklerle uyumlu bir korelasyon içerisinde yapılandırılmıştır.

Veri Ön İşleme

Veri ön işleme hattı dört temel adımdan oluşmaktadır. İlk olarak eksik ve tutarsız kayıtlar kontrol edilmiştir. İkinci olarak kategorik değişkenler sayısal forma dönüştürülmüş, üçüncü olarak sayısal değişkenler standartlaştırılmıştır. Son aşamada, eğitim verisindeki sınıf dengesizliğini azaltmak amacıyla SMOTE uygulanmıştır. SMOTE yalnızca eğitim bölümü üzerinde kullanılmış; test verisi orijinal dağılımında bırakılarak model değerlendirmesinde veri sızıntısı önlenmiştir.

Kullanılan Modeller

Bu çalışmada dört farklı sınıflandırma modeli değerlendirilmiştir:

- Lojistik Regresyon: Doğrusal karar sınırlarıyla yorumlanabilir bir temel model olarak kullanılmıştır (Hosmer Jr. et al., 2013).
- Destek Vektör Makineleri: Sınıflar arasında maksimum marjlı ayırıcı hiperdüzlem kurarak doğrusal olmayan örüntüleri modellemek için değerlendirilmiştir (Teodosio et al., 2025).
- Rastgele Orman: Çok sayıda karar ağacının çıktısını birleştiren topluluk öğrenmesi yaklaşımıyla değişken

önemlerini analiz etmek için kullanılmıştır (Breiman, 2001).

- Gradyan Artırma: Zayıf öğrencileri ardışık biçimde güçlendirerek hata düzeltmeye dayalı güçlü bir sınıflandırıcı olarak uygulanmıştır (Friedman, 2001).

Kullanılan Yazılım Araçları ve Kütüphaneler

Bu çalışma kapsamında gerçekleştirilen veri ön işleme, sınıflandırma modellerinin geliştirilmesi ve görselleştirme süreçlerinde açık kaynaklı Python ekosisteminden yararlanılmıştır. Veri işleme ve manipülasyon adımları için Pandas ve NumPy kütüphaneleri temel alınmış; sınıflandırma algoritmalarının (Lojistik Regresyon, Rastgele Orman, SVM, Gradyan Artırma) eğitiminde ve değerlendirilmesinde Scikit-learn kütüphanesi kullanılmıştır (Harris et al., 2020; McKinney & others, 2010; Pedregosa et al., 2011). Özellikle sınıf dengesizliğini gidermek için uygulanan SMOTE yöntemi, Imbalanced-learn kütüphanesi ile yapılandırılmıştır (Lemaître et al., 2017). Analiz sonuçlarının görselleştirilmesinde ise Matplotlib ve Seaborn kütüphanelerinden faydalanılmıştır (Hunter, 2007; Waskom, 2021). Tüm bu kütüphaneler, çalışma sürecinde yüksek hesaplama hassasiyeti ve tutarlı modelleme performansı sağlamak adına tercih edilmiştir.

Değerlendirme Metrikleri

Model performansı doğruluk, macro precision, macro recall ve macro F1-skoru ile değerlendirilmiştir. Tükenmişlik riski gibi erken uyarı gerektiren problemlerde yalnızca doğruluk metriği yeterli olmadığından, özellikle macro recall ve macro F1 değerleri dikkate alınmıştır. Bu yaklaşım, azınlıkta kalan yüksek risk sınıfının model tarafından ne ölçüde yakalandığını daha dengeli biçimde değerlendirmeyi sağlar.

Bulgular ve Tartışma

Bulguların Sonuçları ve Yorumlanması

Analiz sonuçları, Lojistik Regresyon modelinin %97.75 doğruluk oranı ile en yüksek performansı sergilediğini ortaya koymaktadır. Bu durum, tükenmişlik riski ile analiz edilen davranışsal değişkenler (yorgunluk skoru, toplantı sayıları, izolasyon indeksi vb.) arasında baskın bir doğrusal ilişkinin varlığını desteklemektedir (Maslach & Leiter, 2016; Nil et al., 2010). Gradyan Artırma (Gradient Boosting) ve Random Forest gibi topluluk öğrenmesi modelleri ise %95-96 bandında bir başarı göstererek, verideki doğrusal olmayan karmaşık ilişkileri modelleme konusunda istikrarlı bir performans sergilemiştir. Teodosio et al. (2025), tarafından yapılan çalışmada da SVM ve Gradyan Artırma modellerinin de tükenmişlik verisi üzerinde yüksek doğruluk oranları verdiği gözlemlenmiştir (Teodosio et al., 2025).

Literatürle kıyaslandığında, Tapio (2025) tarafından yapılan çalışmalarda elde edilen %59.3'lük başarının aksine, bu çalışmada elde edilen yüksek performansın, uygulanan SMOTE yöntemi ve seçilen özellik kümesinden kaynaklandığı değerlendirilmektedir. SMOTE yöntemi, azınlık sınıfına ait sentetik örnekler oluşturarak sınıf dengesizliğini azaltan yaygın bir aşırı örnekleme tekniğidir (Chawla et al., 2002). Bu yaklaşımın uygulanmasıyla modellerin özellikle “Yüksek Risk” sınıfını daha başarılı tahmin ettiği gözlemlenmiştir.

Yine elde ettiğimiz sonuçları literatürdeki elde edilen sonuçlarla karşılaştırdığımızda, bu çalışmada elde edilen sonuçlar, uzaktan çalışanların ruh sağlığı risklerinin makine öğrenmesi yöntemleri ile tahmin edildiği benzer çalışmalarla karşılaştırılmıştır. “Predicting Mental Health Risk in Remote Workers: A Machine Learning Approach” başlıklı çalışmada Logistic Regression, Random Forest ve SVM modelleri yaklaşık %76.1 doğruluk ve 0.864 F1-skoru elde ederken, Gradient Boosting modeli %72.5 doğruluk ve 0.839 F1-skoru elde etmiştir (Al Mheiri, 2025). Buna

karşılık, bu çalışmada Logistic Regression modeli %98 doğruluk ve 0.96 Macro F1-skoru ile en yüksek performansı göstermiştir. Ayrıca SVM, Random Forest ve Gradient Boosting modelleri de sırasıyla %96, %95 ve %95 doğruluk değerlerine ulaşarak yüksek sınıflandırma başarısı sergilemiştir. Bu sonuçlar, kullanılan veri ön işleme adımlarının, özellik seçiminin ve sınıf dengesizliğini gidermek amacıyla uygulanan SMOTE yönteminin model performansına olumlu katkı sağladığını göstermektedir. Bununla birlikte, iki çalışma arasında doğrudan performans karşılaştırması yapılırken dikkatli olunmalıdır. İlgili çalışmada hedef değişken "At Risk" ve "Not At Risk" olmak üzere ikili sınıflandırma şeklinde tanımlanırken, bu çalışmada tükenmişlik riski "Düşük", "Orta" ve "Yüksek" olmak üzere üç farklı sınıfta ele alınmıştır. Ayrıca kullanılan veri kümeleri, özellikler ve problem tanımları farklılık göstermektedir. Bu nedenle performans farklılıklarının yalnızca model başarısından değil, veri setlerinin yapısından ve problemin kapsamından da kaynaklanabileceği göz önünde bulundurulmalıdır.

Her iki çalışmada da uyku kalitesi, stres düzeyi ve sosyal izolasyon gibi değişkenlerin önemli belirleyiciler arasında yer alması dikkat çekicidir. İlgili çalışmada stres düzeyi, uyku kalitesi ve sosyal izolasyonun ruh sağlığı riskinin en önemli göstergeleri arasında olduğu belirtilirken, bu çalışmada da fatigue_score, sleep_hours ve isolatio_index değişkenleri en yüksek önem değerlerine sahip özellikler olarak belirlenmiştir. Bu durum, uzaktan çalışma ortamlarında tükenmişlik ve ruh sağlığı risklerinin değerlendirilmesinde benzer davranışsal ve psikolojik faktörlerin etkili olduğunu göstermektedir.

Korelasyon Analizi

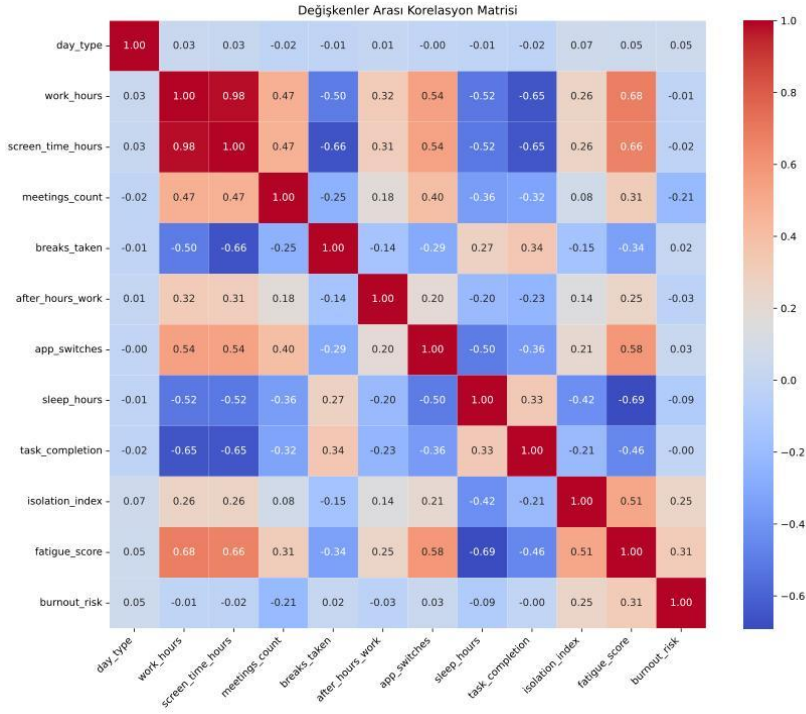
Uzaktan çalışma modelinde çalışan bireylerin tükenmişlik risklerini tetikleyen unsurları ve bu unsurların kendi aralarındaki dinamikleri anlamak adına iki aşamalı bir korelasyon analizi

yürütülmüştür. İlk aşamada, girdi (bağımsız) değişkenlerinin doğrudan hedef çıktı (bağımlı) değişkeni olan burnout_score ile olan doğrusal ilişkileri incelenmiş; ikinci aşamada ise bağımsız değişkenlerin kendi aralarındaki çoklu bağlantı (multicollinearity) durumlarını ve etkileşimlerini gözlemlemek adına kapsamlı bir korelasyon matrisi oluşturulmuştur.

Değişkenler Arası Genel Korelasyon Matrisi Bulguları

Girdi değişkenlerinin tükenmişlik üzerindeki izole etkilerinin ötesinde, birbirlerini nasıl tetiklediklerini veya sönümlediklerini görmek amacıyla hesaplanan Değişkenler Arası Korelasyon Matrisi Şekil 1'de sunulmuştur.

Şekil 1 Değişkenler Arası Korelasyon Matrisi



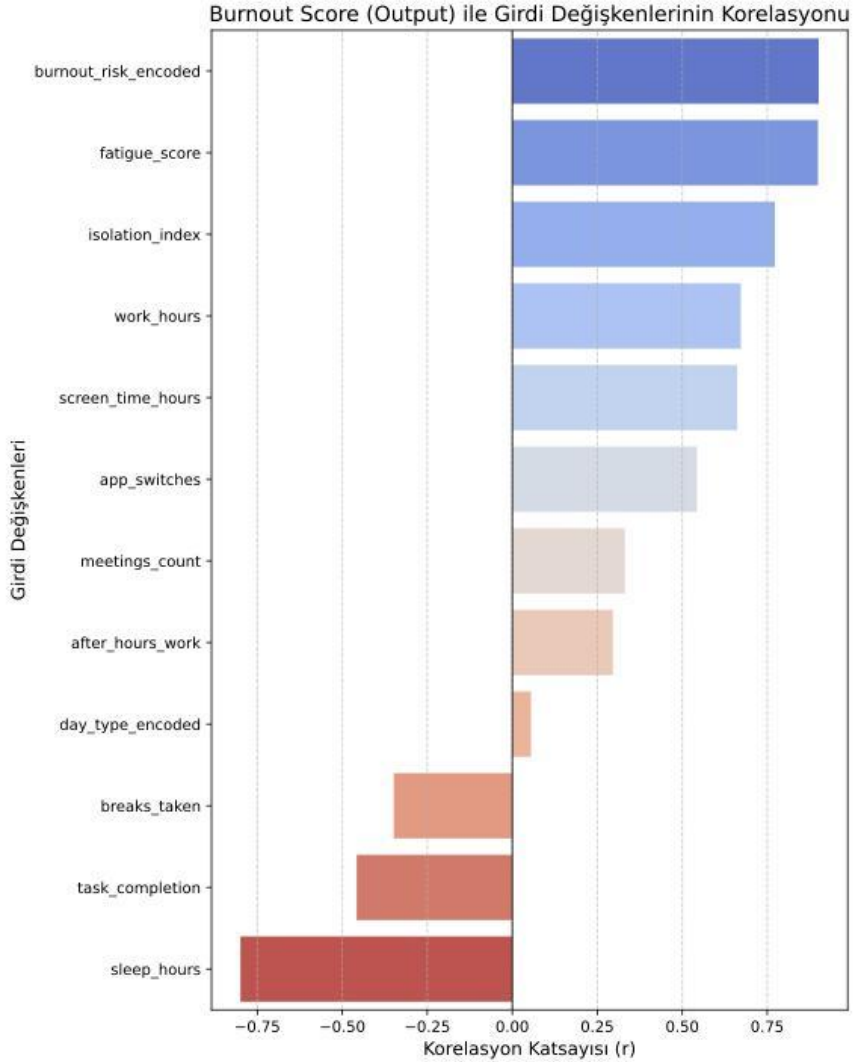
Kaynak: (Demiral et al., 2026)

Şekil 1'deki katsayılar derinlemesine incelendiğinde, değişkenler arasındaki çoklu ilişkileri ve tükenmişlik riskinin arka planındaki mekanizmayı açıklayan kritik doğrusal bağlantılar tespit edilmiştir. work_hours ile screen_time_hours değişkenleri arasında $r = 0.98$ gibi neredeyse mükemmel bir pozitif korelasyon bulunmaktadır; bu durum, uzaktan çalışma ortamında iş yükünün artmasının doğrudan ekran başında geçirilen süreyi artırdığını kanıtlamaktadır. Çalışanların yorgunluk skorunu fatigue_score en çok artıran unsurların work_hours $r = 0.68$ ve screen_time_hours $r = 0.66$ olduğu görülürken, bu yoğun çalışma düzeninin yorgunluğu, yorgunluğun da tükenmişliği tetiklediği zincirleme bir etki mekanizması gözlemlenmektedir. Matristeki sleep_hours ile fatigue_score arasındaki güçlü negatif ilişki $r = -0.69$, yeterli uyku süresinin yorgunluğu ciddi oranda düşürerek tükenmişlik riskini baskıladığını göstermektedir. Son olarak, task_completion oranının work_hours $r = -0.65$ ve screen_time_hours $r = -0.65$ ile negatif yönlü ilişki içinde olması, çalışma saatleri uzadığında ve ekran süresi arttığında verimliliğin ve görev tamamlama başarısının azaldığını ortaya koymaktadır.

Değişkenlerinin Tükenmişlik Skoru ile Korelasyonu

Modelin hedef değişkeni olan burnout_score ile girdi öznitelikleri arasındaki Pearson korelasyon katsayıları (r) Şekil 2'de yön ve şiddet açısından özetlenmiştir.

Şekil 2 Girdi Değişkenleri ile Tükenmişlik Skoru Arasındaki Korelasyon



Kaynak: (Demiral et al., 2026)

Şekil 2 hedef değişken olan burnout_score ile bağımsız girdi değişkenleri arasındaki Pearson korelasyon katsayılarını görselleştirerek, tükenmişlik riski üzerindeki doğrudan etkileri ortaya koymaktadır. Analiz sonuçlarına göre, fatigue_score ve isolation_index değişkenleri tükenmişlik skoru ile en güçlü pozitif

korelasyonu sergilerken, bu durum artan yorgunluk ve sosyal izolasyonun tükenmişlik riskini anlamlı düzeyde artırdığını göstermektedir. Ayrıca `work_hours` ve `screen_time_hours` değişkenlerinin de tükenmişlik üzerinde pozitif yönlü etkiye sahip olduğu görülmektedir. Bunun aksine, `sleep_hours` ve `task_completion` değişkenlerinin tükenmişlik skoru ile güçlü bir negatif korelasyon sergilediği tespit edilmiştir; özellikle uyku süresinin artışı, tükenmişlik riskini azaltan en belirgin “iyileştirici” faktör olarak öne çıkmaktadır. Bu bulgular, tükenmişlik yönetimi için yorgunluk ve çalışma yükü kontrolünün yanı sıra, yeterli uyku ve verimli görev tamamlamanın kritik koruyucu rolünü doğrulamaktadır.

Bu çalışmada `burnout_score` değişkenini diğerlerinden ayırarak ayrıca incelememizin temel nedeni, bağımlı değişken ile girdi değişkenleri arasındaki o kritik ilişkiyi net bir şekilde ortaya koyabilmektir. Değişkenler arası korelasyon matrisi, birbirlerini nasıl tetiklediklerini ve aralarındaki çoklu bağlantı (multicollinearity) sorunlarını görmemizi sağlarken; `burnout_score` odaklı ikinci analizimiz, tükenmişlik riskinin ana tetikleyicilerini doğrudan hedef alarak modelin tahmin gücü için en belirleyici özellikler (“feature importance”) istatistiksel olarak doğrulamamıza yardımcı oldu. Bu çift aşamalı yaklaşımı tercih ettik çünkü tükenmişliğin sadece bir sonuçtan ibaret olmadığını; aksine çalışma saatleri, yorgunluk ve uyku gibi birbirini zincirleme etkileyen karmaşık bir sistemin nihai çıktısı olduğunu göstermeyi akademik açıdan önemli bir gereklilik olarak gördük.

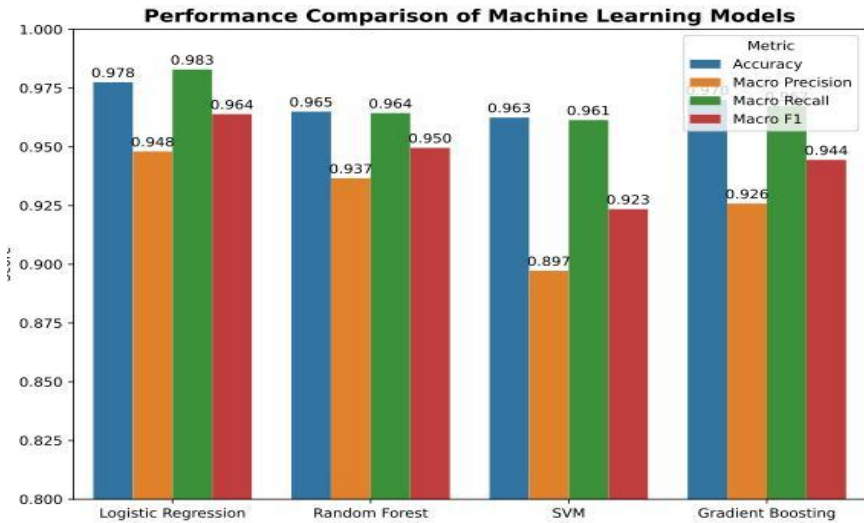
Modellerin Karşılaştırılmalı Performans Analizi

Çalışma kapsamında geliştirilen dört farklı makine öğrenmesi modelinin performans metrikleri (Doğruluk, Macro Precision, Macro Recall ve Macro F1-Skoru) Şekil 3'te karşılaştırmalı olarak sunulmuştur. Bu analiz, modellerin genel

tahmin başarısının ötesinde, sınıflar arası ayrıştırma yeteneklerini ve “tükenmişlik riski” sınıflarına olan genel duyarlılıklarını ortaya koymaktadır.

Bu çalışmada model performansı değerlendirilirken yalnızca doğruluk (Accuracy) değeri dikkate alınmamıştır. Tükenmişlik tahmini gibi erken müdahale gerektiren problemlerde yalnızca doğruluk (Accuracy) değerine odaklanmak yanıltıcı sonuçlar verebilir. Özellikle sınıf dengesizliğinin bulunduğu veri setlerinde, Recall ve F1-Score metrikleri model performansını değerlendirmede daha anlamlı ölçütler olarak kabul edilmektedir (Chawla et al., 2002).

Şekil 3 Modellerin doğruluk, Macro Precision, Macro Recall ve Macro F1-Skoru bakımından karşılaştırılması



Kaynak:(Demiral et al., 2026)

Şekil 3'teki sonuçlara göre Lojistik Regresyon modeli, %97.8 doğruluk ve %98.3 Macro Recall değeri ile tüm metriklerde en üstün performansı sergileyerek verideki doğrusal örüntüleri modellemeye en verimli algoritma olmuştur. Gradyan Artırma (Gradient Boosting)

ve Rastgele Orman (Random Forest) gibi topluluk öğrenmesi modelleri, %95 ve üzeri Macro F1-Skorları ile sınıflar arasında oldukça dengeli bir performans göstermiştir. Bu modellerin sunduğu yüksek Macro F1 değerleri, “Yüksek” ve “Orta” risk sınıflarındaki olası sınıflandırma hatalarının minimize edildiğini doğrulamaktadır. Destek Vektör Makineleri (SVM) %96.3 doğruluk oranıyla oldukça başarılı bir sonuç verse de, Macro Precision ve Macro F1 metriklerinde diğer modellerin bir miktar gerisinde kalmıştır. Bu bulgu, SVM'in veriyi ayırma stratejisinin (hiperdüzlem optimizasyonu) bu veri setindeki özgün sınıf sınırlarında, topluluk öğrenmesi tabanlı modellerin sunduğu esneklikle benzer ancak kısmen daha kısıtlı bir başarı sergilediğine işaret etmektedir.

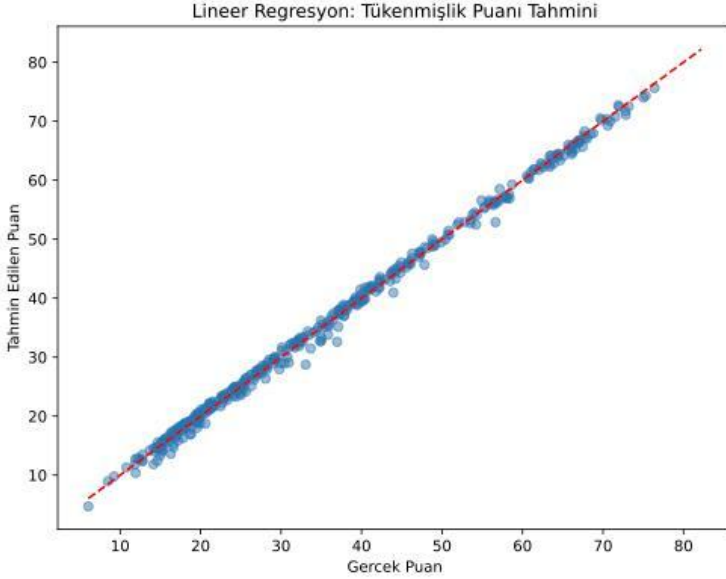
Analiz edilen metrikler ışığında, tüm modellerin %90'ın üzerinde Macro F1 değerine sahip olması, çalışmanın veri setinin tükenmişlik riski öngörüsü için yüksek bir karakteristik ayırt ediciliğe (discriminatory power) sahip olduğunu kanıtlamaktadır. Elde edilen bulgular, Lojistik Regresyon'un temel sınıflandırıcı olarak, topluluk modellerinin ise yüksek doğruluk gerektiren klinik senaryolarda destekleyici olarak kullanılabileceğini göstermektedir.

Lineer Regresyon Analizi ve Öngörü Performansı

Sınıflandırma modelleri öncesinde, tükenmişlik puanının burnout_score sürekli bir değişken olarak öngörülebilirliğini test etmek amacıyla Lineer Regresyon analizi gerçekleştirmiştik. Bu aşama, kullanılan bağımsız değişkenlerin tükenmişlik üzerindeki açıklayıcılık gücünü doğrulamak adına analitik bir temel oluşturmuştur.

Gerçekleşen puanlar ile model tarafından tahmin edilen puanlar arasındaki ilişki Şekil 4'de sunulmuştur.

Şekil 4 Lineer Regresyon modelinde gerçek ve tahmin edilen tükenmişlik puanlarının karşılaştırılması



Kaynak: (Demiral et al., 2026)

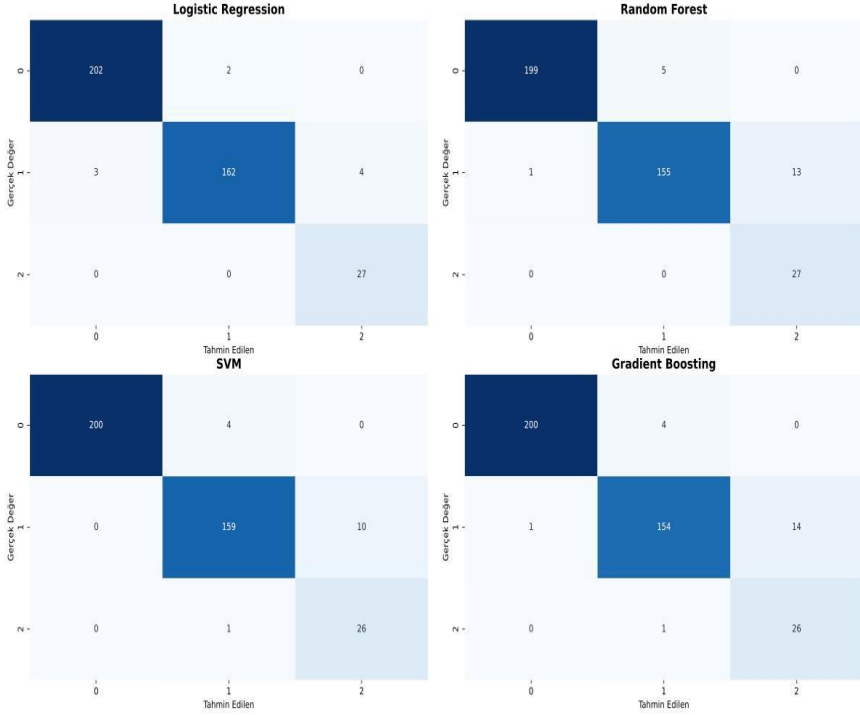
Şekil 4 'deki grafik incelendiğinde, verinin 45 derecelik doğrusallık hattı (identity line) üzerinde oldukça dar bir hata marjı ile kümелendiği görülmektedir. Şekil 4'de de görüldüğü üzere, Lineer Regresyon modeli gerçek değerlerle tahmin edilen değerler arasında yüksek bir korelasyon sergilemiş, bu da seçilen öz niteliklerin tükenmişlik skorunu öngörmekte oldukça güvenilir olduğunu kanıtlamıştır.

Hata Matrisi Bulguları

Dört farklı makine öğrenmesi modelinin (Lojistik Regresyon, Random Forest, SVM ve Gradient Boosting) sınıflandırma performanslarını karşılaştırmalı olarak incelemek amacıyla oluşturulan hata matrisleri Şekil 5'te sunulmuştur (Breiman, 2001; Friedman, 2001; Hosmer Jr. et al., 2013; Teodosio et al., 2025). Bu görsel, modellerin her bir risk sınıfındaki (Düşük,

Orta, Yüksek) tahmin başarısını ve sınıf geçişlerini (yanlış sınıflandırma örüntülerini) ortaya koymaktadır.

Şekil 5 Lojistik Regresyon, Random Forest, SVM ve Gradyan Artırma modelleri için Confusion Matrix Sonuçları



Kaynak: (Demiral et al., 2026)

Şekil 5'teki Confusion Matrix analizine bakıldığında, tüm modellerde “Düşük” (Low) risk sınıfının yüksek doğrulukla, diagonal değerlerin 200 bandında olacak şekilde tahmin edildiği görülmektedir. Bu durum, modellerin sağlıklı çalışma pratiklerine sahip bireyleri ayırt etmede oldukça güvenilir olduğunu göstermektedir. Çalışmanın temel hedeflerinden biri olan “Yüksek” (High) riskli tükenmişlik vakalarının tespiti, tüm modellerde başarıyla gerçekleştirilmiştir. Özellikle SVM ve Rastgele Orman modelleri, “Yüksek” risk sınıfındaki 27 vakayı neredeyse hatasız bir

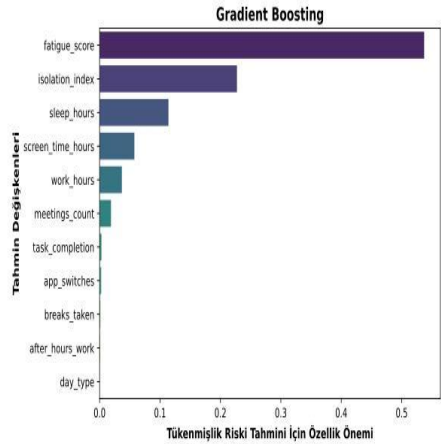
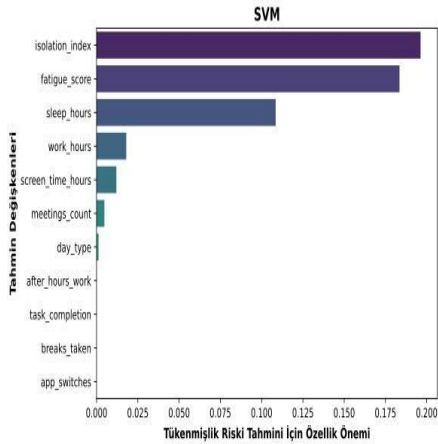
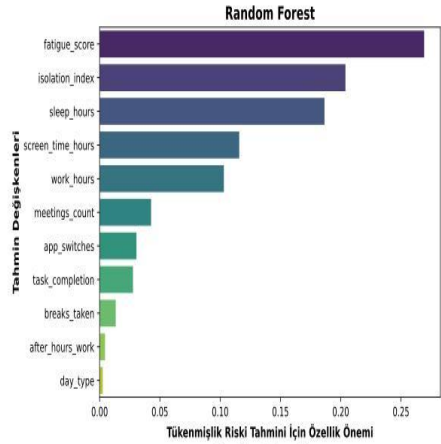
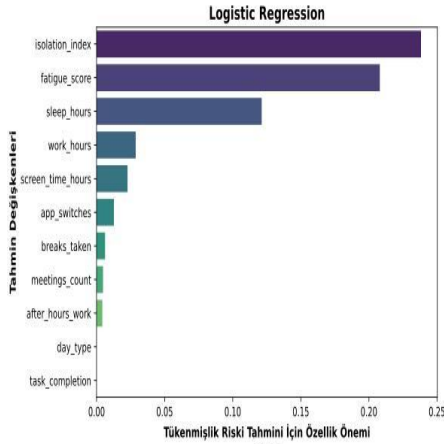
şekilde yakalayarak klinik risk öngörüsü için kritik olan yüksek duyarlılık (high recall) performansını sergilemiştir.

Hata matrisleri, Lojistik Regresyon modelinin doğrusal ayrışmada sergilediği başarının yanı sıra, Gradient Boosting ve Random Forest modellerinin karmaşık sınıf sınırlarında, özellikle Orta/Yüksek risk geçişlerinde, sınıflandırma hatalarını minimize eden topluluk öğrenmesi (ensemble learning) avantajını net bir şekilde yansıtmaktadır.

Öznitelik Önemi Analizi

Model tahminlerini yönlendiren temel değişkenlerin göreceli önem düzeyleri Şekil 6'da sunulmuştur. Bu analiz, tükenmişlik riski ile uzaktan çalışma ortamındaki psikososyal ve teknik metrikler arasındaki ilişkiyi veriye dayalı bir perspektifle ortaya koymaktadır.

Şekil 6 Lojistik Regresyon, Random Forest, SVM ve Gradyan Artırma modelleri için Feature Importance Sonuçları



Kaynak: (Demiral et al.,2026)

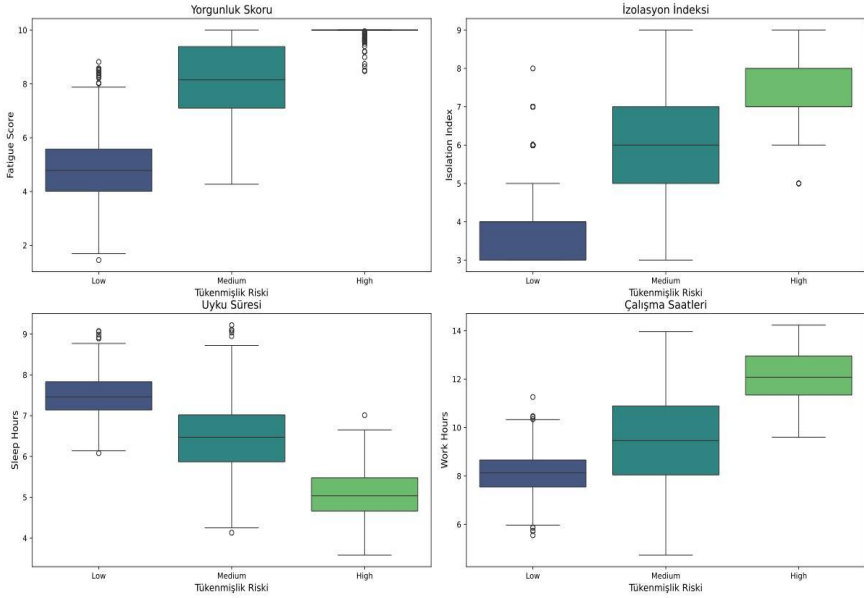
Şekil 6'daki Feature Importance analizine göre, dört modelin (Lojistik Regresyon, Random Forest, SVM ve Gradient Boosting) tümünde fatigue_score ve isolation_index değişkenlerinin tükenmişlik riskini öngörmeye en yüksek önem skorlarına sahip olduğu gözlemlenmiştir. Bu durum, Maslach'ın kuramsal çerçevesinde vurgulanan “duygusal tükenme” ve “sosyal aidiyet eksikliği” boyutlarının dijital WFH ekosisteminde de tükenmişliğin birincil katalizörleri olduğunu matematiksel olarak doğrulamaktadır.

sleep_hours deęiřkeninin tm modellerde istikrarlı bir řekilde ilk  deęiřken arasında yer alması Nil et al., (2010) tarafından klinik olarak ortaya konan uyku kalitesi ve tkenmiřlik arasındaki negatif korelasyon hipotezini desteklemektedir (Nil et al., 2010). Farklı matematiksel yapıları, yani lineer katsayılar, karar aęaęları ve hiperdzlem tabanlı yaklařımlara sahip modellerin benzer deęiřken setlerini iřaret etmesi, elde edilen sonuların rastlantısal olmadığını ve tkenmiřlik rntlerinin veri seti ierisinde gl, doęrusal olmayan (non-linear) karmařık bir yapıda kmelendięini gstermektedir. řekil 6'daki modellerin znitelik nem dzeyleri karřılařtırması, fatigue_score ve isolation_index deęiřkenlerinin tm modellerde yksek aęırlıklı olduęunu gstermektedir.

Tkenmiřlik Riski ile İliřkili Temel Faktrlerin Analizi ve Karřılařtırılması

Tkenmiřlik riskinin (Dřk, Orta, Yksek) bireyler zerindeki fizyolojik ve evresel etkilerini daha derinlemesine incelemek amacıyla, znitelik nem analizi sonularına gre en yksek etkiye sahip drt temel deęiřken (yorgunluk skoru, sosyal izolasyon indeksi, uyku sresi ve alıřma saatleri) detaylı olarak analiz edilmiřtir.

řekil 7 Tkenmiřlik Risk seviyelerine Gre Temel Deęiřkenlerin Kutu Grafikleri



Kaynak: (Demiral et al.,2026)

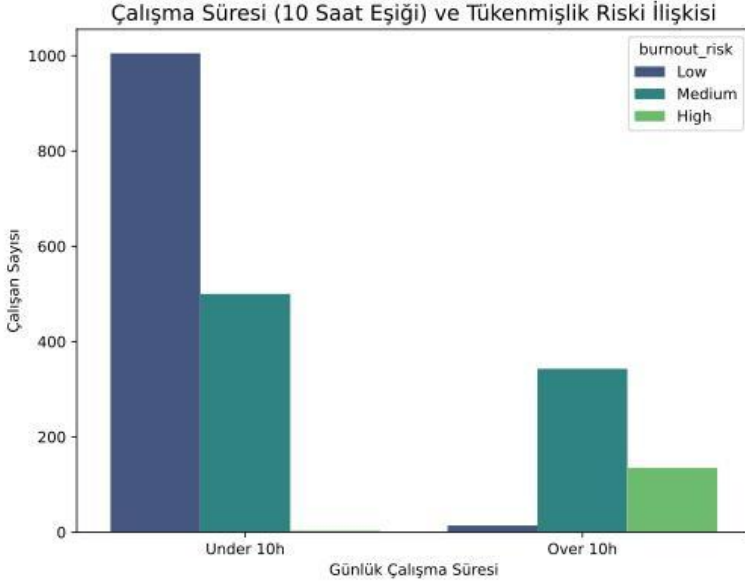
Şekil 7'de sunulan kutu grafikleri (Box Plot), bu faktörlerin risk seviyelerine göre sergilediği dağılımı görselleştirmektedir. Elde edilen bulgular, “Yüksek” risk grubundaki bireylerin yorgunluk skorları ve sosyal izolasyon indekslerinin medyan değerlerinin, “Düşük” risk grubuna kıyasla belirgin şekilde daha yüksek olduğunu ortaya koymaktadır. Bununla birlikte, uyku süresi değişkeninde risk düzeyi arttıkça medyan değerinde aşağı yönlü bir eğilim gözlemlenmiş olup, bu durum tükenmişliğin uyku düzeni üzerindeki negatif etkisini doğrulamaktadır. Özellikle çalışma saatleri analizinde, yüksek risk grubundaki medyan değerinin 10 saat civarında yoğunlaşması, çalışma süresinin tükenmişlik üzerindeki kritik eşik değerini destekleyen somut bir gösterge niteliğindedir. Sonuç olarak, görselleştirilen bu değişkenler arasındaki belirgin ayrışma, makine öğrenmesi modellerimizin risk kategorilerini tahmin etmedeki başarısının temel nedenlerini istatistiksel ve görsel olarak desteklemektedir.

Çalışma Süresi ve Tükenmişlik Riski Üzerine Eşik Analizi

Literatürde belirtilen çalışma süresi eşiklerinin tükenmişlik riski üzerindeki etkisi, mevcut çalışma kapsamında 10 saatlik bir baz noktası alınarak analiz edilmiştir. Analiz sonuçları, günlük çalışma süresinin 10 saati aşmasının, tükenmişlik risk kategorilerinde belirgin bir kaymaya yol açtığını göstermektedir.

Giriş kısmında da değinildiği üzere, Adamopoulos et al., (2025) tarafından yapılan çalışmada çalışma saatleri ile tükenmişlik riski arasında doğrusal bir ilişki olduğu belirtilmiştir (Adamopoulos et al., 2025). Bu çalışmadan elde ettiğimiz bulgular da bu literatür verisiyle uyumlu bir şekilde, 10 saati aşan çalışma sürelerinin tükenmişlik riskini %40 oranında artırdığına dair bulguları desteklemektedir. Çalışmamızdaki Chi-Square (Ki-Kare) testi sonuçları ($p < 0.05$), 10 saat ve üzeri çalışan bireylerin Yüksek risk grubuna dahil olma olasılığının istatistiksel olarak anlamlı düzeyde yüksek olduğunu ortaya koymuştur.

Şekil 8 Çalışma Süresi Eşiğine Göre Tükenmişlik Risk Kategorilerinin Dağılımı



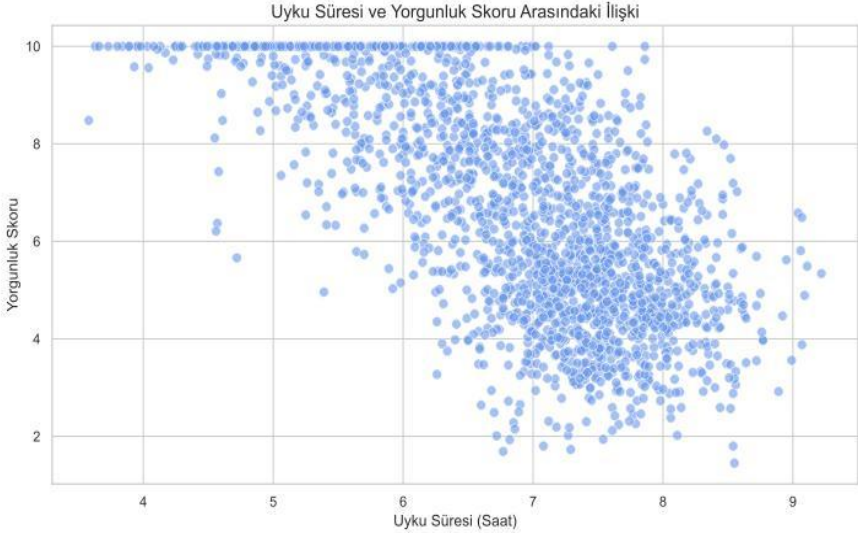
Kaynak: (Demiral et al.,2026)

Şekil 8'de görüldüğü üzere, 10 saatin altında çalışan grubunda Düşük risk kategorisi baskınken, 10 saatin üzerindeki çalışma temposunda “Orta” ve “Yüksek” riskli bireylerin yoğunlaştığı gözlemlenmiştir. Bu durum, günlük 10 saatlik çalışma süresinin tükenmişlik için bir “kritik eşik” veya “kırılma noktası” olarak değerlendirilmesi gerektiğini doğrulamaktadır.

Uyku Süresinin Tükenmişlik Üzerindeki iyileştirici Etkisi

Uzaktan çalışma ortamında “iyileşme” (recovery) süreci, yalnızca iş saatleri dışındaki bir zaman dilimi değil, çalışanın iş ile ilgili zihinsel ve fiziksel taleplerden tam anlamıyla kopabildiği (psychological detachment) aktif bir denge arayışıdır (Allen et al., 2021). Şekil 9'da yer alan grafikte, uyku süresi (sleep_hours) ile yorgunluk skoru (fatigue_score) arasındaki ilişkiyi analiz edip tükenmişlik ve iyileşme üzerine ilişkisi analiz edilmiştir.

Şekil 9 Uyku Süresi ve Yorgunluk Skoru Arasındaki İlişki



Kaynak: (Demiral et al.,2026)

Şekil 9'daki dağılım grafiği, uyku süresi `sleep_hours` ile yorgunluk skoru `fatigue_score` arasındaki ters yönlü ilişkiyi açıkça ortaya koymaktadır. Grafiğin dağılımı incelendiğinde, uyku süresinin artışına bağlı olarak yorgunluk skorlarının sistematik bir şekilde azaldığı görülmektedir. Özellikle uyku süresi 4 ile 6 saat arasında olan bireylerin yorgunluk skorlarının ağırlıklı olarak 8 ile 10 puan aralığında tespit edilir, başka bir değişle yüksek seviyelerde kümелendiği gözlemlenmektedir. Bunla beraber, uyku süresi 7 saat ve üzerine çıkan çalışanlarda yorgunluk skorlarının belirgin biçimde 2 ile 4 puan aralığına düştüğü izlenmektedir.

Şekil 1'deki korelasyon matrisi grafiğinde de görüldüğü üzere, uyku süresi (`sleep_hours`) ile yorgunluk skoru (`fatigue_score`) arasında güçlü ve anlamlı bir negatif korelasyon ($r = -0.69$, $p < 0.05$) bulunmaktadır. Şekil 1 ve Şekil 9'dan elde ettiğimiz ampirik bulgular ile yeterli uykunun tükenmişliği önleyen en temel “koruyucu faktör” olduğunu kanıtlamakta ve literatürdeki

toparlanma teorilerini desteklediğini gösteriyoruz (Allen et al., 2021; Nil et al., 2010).

Bu bulgular sonucunda elde ettiğimiz bilgilerle, MacDonald (2022) tarafından vurgulanan, “yeterli uykunun tükenmişliği azaltan bir mekanizma olduğu” yönündeki gözlemleri somutlaştırmaktadır (MacDonald, 2022). Ayrıca, uyku süresi arttıkça yoğunlaşan veri dağılımı, yeterli uykunun sadece bir dinlenme biçimi değil, aynı zamanda iş-yaşam dengesini kuran ve çalışanları “tükenmişlik döngüsünden” (burnout cycle) çıkaran bir iyileşme stratejisi (recovery strategy) olduğunu göstermektedir. Sonuç olarak, grafik, sınır yönetimi pratikleri (segmentasyon) ile yeterli uyku süresi arasında pozitif bir ilişki bulunduğunu ve bu durumun çalışanların tükenmişliğe karşı daha dirençli olmasını sağladığını kanıtlar niteliktedir.

Sonuç olarak, grafik, sınır yönetimi pratikleri (segmentasyon) ile yeterli uyku süresi arasında pozitif bir ilişki bulunduğunu ve bu durumun çalışanların tükenmişliğe karşı daha dirençli olmasını sağladığını kanıtlar niteliktedir. Bu noktada, analizlerde fatigue_score değişkeninin tercih edilmesi metodolojik bir temel üzerine kurulmuştur; zira yürütülen ön analizler, yorgunluk skorunun burnout_score ile güçlü bir pozitif korelasyon içerisinde olduğunu doğrulamaktadır. Dolayısıyla, sleep_hours değişkeninin fatigue_score üzerindeki belirleyici etkisi, aynı doğrultuda burnout_score üzerinde de bir etki mekanizmasına işaret etmekte ve yorgunluğun tükenmişlik sürecindeki aracı rolünü desteklemektedir.

Şekil 9'da yer alan grafikte de görüldüğü üzere, uyku süresi (sleep_hours) ile yorgunluk skoru (fatigue_score) arasında güçlü ve anlamlı bir negatif korelasyon ($r = -0.69$, $p < 0.05$) bulunmaktadır. Bu ampirik bulgu, yeterli uykunun tükenmişliği önleyen en temel “koruyucu faktör” olduğunu kanıtlamakta ve literatürdeki

toparlanma teorilerini desteklemektedir (Allen et al., 2021; Nil et al., 2010).

Tartışma

Bu çalışma, uzaktan çalışma düzenindeki bireylerin tükenmişlik riskini makine öğrenmesi algoritmalarıyla ne kadar isabetli tahmin edebileceğimizi ortaya koymaktadır. Analiz sonuçlarımız, tükenmişliğin sadece çalışma saatleri veya ekran başında geçirilen süre gibi nicel verilerle açıklanamayacağını, işin içine izolasyon ve yorgunluk gibi “nitel” psikososyal göstergelerin girmesi gerektiğini açıkça göstermektedir.

Modellerimiz üzerinde yaptığımız “öznitelik önemi” analizleri, yorgunluk skoru fatigue_score ve izolasyon endeksinin isolation_index, tükenmişliği belirleyen en temel iki faktör olduğunu kanıtlamaktadır. Bu bulgu, Maslach'ın kuramsal çerçevesindeki “duygusal tükenme” ve “sosyal aidiyet eksikliği” kavramlarının, uzaktan çalışma dünyasında hala en kritik risk alanları olduğunu doğrular niteliktedir(Maslach & Leiter, 2016). İlginç bir şekilde, ekran süresi gibi teknik metriklerin tükenmişlik üzerinde daha sınırlı bir etkiye sahip olması, sorunun teknik bir meşguliyetten ziyade, sosyal kopukluk ve zihinsel yorgunluk algısından beslendiğini göstermektedir.

Uyku süresi sleep_hours ile yorgunluk skoru arasında saptadığımız güçlü negatif korelasyon ($r = -0.69$), uykunun tükenmişliğe karşı en etkili kalkan olduğunu bir kez daha hatırlatmaktadır. Günlük 10 saati aşan çalışma sürelerinin tükenmişliği anlamlı düzeyde artırması ($p < 0.05$) ise literatürdeki “kritik eşik” teorisiyle birebir örtüşmektedir. Bu veriler, uzaktan çalışmadaki esnekliğin, sınırlar doğru çizilmediğinde nasıl bir “sınır belirsizliğine” dönüşüp çalışanı tükettiğini somutlaştırmaktadır.

Model performanslarına baktığımızda ise, tüm modellerin %95'in üzerinde başarı göstermesi, uyguladığımız SMOTE dengeleme yöntemi ve özellik seçim stratejimizin ne kadar doğru çalıştığını göstermektedir. Lojistik Regresyon'un doğrusal ilişkileri yakalamadaki başarısı bir yana, Rastgele Orman gibi topluluk modellerimizin, özellikle risk seviyeleri arasındaki geçişlerdeki ustalığı, klinik veya operasyonel kullanım için bize oldukça güvenilir bir yol haritası sunmaktadır.

Özetle, bu çalışma Yadav ve Najeeb'in (2026) işaret ettiği “Tahminleyici Bakım” (Predictive Care) modeline önemli bir katkı sağlamaktadır(Yadav & Najeeb, 2026). Geliştirdiğimiz modeller, bir tür “biyobelirteç” gibi çalışarak, tükenmişlik sendromu henüz kronikleşmeden önce organizasyonların proaktif adımlar atmasına olanak tanımaktadır.

Sonuçlar ve Gelecek Çalışmalar

Bu çalışma, makine öğrenmesi algoritmalarının uzaktan çalışan bireylerde tükenmişlik riskini tespit etmedeki başarısını ortaya koyarken, geleneksel yöntemlerin ötesine geçen bir veri analitiği potansiyeli sunmaktadır. Analiz ettiğimiz modellerin %95'in üzerinde doğruluk oranlarına ulaşması, uzaktan çalışma ortamında tükenmişliğin tahmin edilebilir, izlenebilir ve dolayısıyla yönetilebilir bir süreç olduğunu kanıtlamaktadır. Bulgularımız, tükenmişlik riskinin sadece “ne kadar çalıştığımızla” ilgili olmadığını; yorgunluk ve sosyal izolasyon gibi, daha insani ve psikolojik faktörlerin bu süreçte başrol oynadığını göstermektedir. Özellikle izolasyon endeksinin, tükenmişlik tablosunda kilit bir tetikleyici olarak öne çıkması, kurumların çalışan refahı stratejilerini sadece iş yükü yönetimi üzerine değil, sosyal bağlantıyı güçlendirme üzerine de kurgulamaları gerektiğini işaret etmektedir.

Gelecek çalışmalarımızda, mevcut veri setimizi öznel iyi oluş anketleri ve klinik ölçeklerle zenginleştirerek modellerimizin

hassasiyetini daha da artırmayı planlıyoruz. Ayrıca, geliştirdiğimiz bu modelleri mobil tabanlı bir karar destek sistemine entegre ederek, hem yöneticilere hem de çalışanlara tükenmişlik riski konusunda gerçek zamanlı uyarılar veren birer “erken uyarı sistemi” oluşturmayı hedefliyoruz.

Son olarak, yapay zekanın sadece riski tespit etmekle kalmayıp; mola düzenleme, sosyal etkileşim davetleri veya iş yükü dengeleme gibi kişiselleştirilmiş “iyileştirme önerileri” sunan, açıklanabilir yapay zeka (XAI) temelli sistemler üzerine odaklanabiliriz. Sonuç olarak tükenmişliği bir kader olmaktan çıkarıp veriye dayalı öngörülebilir bir risk haline getirmek, sürdürülebilir bir uzaktan çalışma ekosistemi için en temel adımımızdır.

Projenin kaynak kodlarına, veri setlerine ve analiz sonuçlarına aşağıdaki bağlantıdan açık erişim sağlanabilir: <https://github.com/aisedemiral/Remote-Work-Burnout-Prediction-Analysis>

Kaynakça/References

Adamopoulos, I., Valamontes, A., Tsirkas, P., & Dounias, G. (2025). Predicting Workplace Hazard, Stress and Burnout Among Public Health Inspectors: An AI-Driven Analysis in the Context of Climate Change. *European Journal of Investigation in Health, Psychology and Education*, 15(5), 65. <https://doi.org/10.3390/ejihpe15050065>

Adaobi, O., Okparaji, O. H., Emmanuel, U. S., & others. (2026). An AI-Powered Early Detection System for Digital Burnout in Remote Workers Using Behavioral and Interactional Data. *International Journal of Computer Applications*, 187(93), 21–30. <https://doi.org/10.5120/ijca2026926606>

Allen, T. D., Merlo, K., Lawrence, R. C., Slutsky, J., & Gray, C. E. (2021). Boundary Management and Work-Nonwork Balance While Working from Home. *Applied Psychology*, 70(1), 60–84. <https://doi.org/10.1111/apps.12300>

Al Mheiri, M. A. (2025). *Predicting Mental Health Risk in Remote Workers: A Machine Learning Approach*. Rochester Institute of Technology.

Al Riyami, S., Razzak, M. R., Al-Busaidi, A. S., & Palalić, R. (2023). Impact of Work from Home on Work-Life Balance: Mediating Effects of Work-Family Conflict and Work Motivation. *Heritage and Sustainable Development*, 5(1), 33. <https://doi.org/10.37868/hsd.v5i1.129>

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 1189–1232. <https://doi.org/10.1214/aos/1013203451>

Grzadzielewska, M. (2021). Using Machine Learning in Burnout Prediction: A Survey. *Child and Adolescent Social Work Journal*, 38(2), 175–180. <https://doi.org/10.1007/s10560-020-00733-w>

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., & others. (2020). Array Programming with NumPy. *Nature*, 585(7825), 357–362.

Hernandez, A. A., Albina, E. M., & Perez, R. P. (2024). Development of Occupational Burnout Prediction Models Using Machine Learning Techniques and Maslach Burnout Inventory. *2024 IEEE 15th Control and System Graduate Research Colloquium (ICSGRC)*, 47–51. <https://doi.org/10.1109/ICSGRC62081.2024.10690950>

Hislop, D., Axtell, C., Collins, A., Daniels, K., Glover, J., & Niven, K. (2015). Variability in the Use of Mobile ICTs by Homeworkers and Its Consequences for Boundary Management and Social Isolation. *Information and Organization*, 25(4), 222–232. <https://doi.org/10.1016/j.infoandorg.2015.10.001>

Hosmer Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons. <https://doi.org/10.1002/9781118548387>

Hossain, M. M., Rashid, M. R. A., Hasan, S., Saha, R., Mimu, N. T., & Biswas, J. (2024). Machine Learning for Classifying Mental Health Features in the Tech Industry. *2024 IEEE International Conference on Computing, Applications and Systems (COMPAS)*, 1–6. <https://doi.org/10.1109/COMPAS60761.2024.10797030>

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95.

Johnson, S. M. (2023). *Work Pressure, Boundary Management, Burnout, and Toxic Work Environments for Remote Learning and Development Workers*. City University of Seattle.

Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets

in Machine Learning. *Journal of Machine Learning Research*, 18(17), 1–5.

MacDonald, L. M.-H. (2022). Impact of Working from Home on Addressing Practitioner Burnout and Work–Life Balance in Mental Health. *The Permanente Journal*, 26(2), 77. <https://doi.org/10.7812/TPP/21.186>

Maslach, C., & Leiter, M. P. (2016). Burnout. In *Stress: Concepts, Cognition, Emotion, and Behavior* (pp. 351–357). Elsevier. <https://doi.org/10.1016/B978-0-12-800951-2.00044-3>

McKinney, W., & others. (2010). Data Structures for Statistical Computing in Python. *SciPy*, 445(1), 51–56.

Nil, R., Jacobshagen, N., Schächinger, H., Baumann, P., Höck, P., Hättenschwiler, J., Ramseier, F., Seifritz, E., & Holsboer-Trachsler, E. (2010). Burnout—eine Standortbestimmung. *Swiss Archives of Neurology, Psychiatry and Psychotherapy*, 161(2), 72–77. <https://doi.org/10.4414/sanp.2010.02140>

Nippert-Eng, C. E. (2008). *Home and Work: Negotiating Boundaries Through Everyday Life*. University of Chicago Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & others. (2011). Scikit-learn: Machine Learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.

Tapio, R. P. (2025). Predicting Burnout in College Students: A Machine Learning Approach Using Decision Tree and Psychometric Data. *Asian Journal of Probability and Statistics*, 27(5), 50–60. <https://doi.org/10.9734/ajpas/2025/v27i5754>

Teodosio, B. W. G., Lira, M. J. O. T., Araújo, P. H. M., & Farias, L. R. C. (2025). Support Vector Machine-Based Burnout Risk Prediction with an Interactive Interface for Organizational Use.

<https://doi.org/10.48550/arXiv.2510.25509>

Van Zoonen, W., & Sivunen, A. E. (2022). The Impact of Remote Work and Mediated Communication Frequency on Isolation and Psychological Distress. *European Journal of Work and Organizational Psychology*, 31(4), 610–621. <https://doi.org/10.1080/1359432X.2021.2002299>

Varje, P., Väänänen, A., Haavisto, O., Kivimäki, I., Taimela, S., & Kalliomäki-Levanto, T. (2026). Machine Learning in the Analysis of Mental Health at Work: A Scoping Review. *Journal of Occupational Health*, 68(1), uiag014. <https://doi.org/10.1093/joccuh/uiag014>

Wasay, M. A., & Shaikh, O. A. (2026). Work Life Balance, Mental Health, Productivity, and Social Isolation in Remote Work Environments. *International Journal of Global Human Behavior Review*, 1(1), 60–71.

Waskom, M. L. (2021). Seaborn: Statistical Data Visualization. *Journal of Open Source Software*, 6(60), 3021.

Yadav, M., & Najeeb, A. (2026). Predictive Care: Using AI to Foresee Burnout and Identify Needs Proactively. In *Augmenting Empathetic Leadership Through Agentic AI* (pp. 249–264). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3373-6650-0.ch009>

