

Turbofan Motor Kalan Ömrü Tahmini için Güven Aralığı Sağlayan Hibrit Derin Öğrenme Modeli

Caner Yıldırım - Erva Kansu
Yıldız Teknik Üniversitesi
caneryldrm@gmail.com - ervakansu@gmail.com

Abstract—Bu makale, turbofan motorlarının kalan faydalı ömrünü (RUL) yüksek doğrulukla tahmin etmeyi ve bu tahminler için güvenilir güven aralıkları sunmayı amaçlayan yenilikçi bir hibrit derin öğrenme modelini sunmaktadır. NASA C-MAPSS FD001 veri seti üzerinde yapılan deneylerde, beş farklı derin öğrenme mimarisi (LSTM, GRU, 1D-CNN, CNN+LSTM, CNN+BiLSTM) ve zaman tabanlı özellik mühendisliği teknikleri (kaydırılmış ortalama, standart sapma, varyans, fark) uygulanmıştır. Güven aralığı tahmini için eğitim sonrası Quantile Regression yöntemi kullanılmış ve model belirsizliği değerlendirilmiştir. Model performansını değerlendirmek için Kök Ortalama Kare Hatası (RMSE), Ortalama Mutlak Hata (MAE), R-Kare (R^2), Kapsama Olasılığı (Coverage Probability) ve Aralık Genişliği (Interval Width) metrikleri kullanılmıştır. Çalışma, farklı model mimarilerini, özellik mühendisliği yaklaşımlarını ve güven aralığı tahmin yöntemlerini karşılaştırarak en iyi uygulamaları belirlemeyi ve modelin tahmin belirsizliğini kapsamlı bir şekilde analiz etmeyi hedeflemektedir. En iyi performans gösteren model, test verisi üzerinde 14.75 RMSE ve 10.63 MAE değerleri elde etmiş, güven aralığı tahminleri için de 0.94 Kapsama Olasılığı ve 44.64 Aralık Genişliği değerlerine ulaşmıştır. Bu çalışma, güvenilir RUL tahminleri sağlayarak motor bakım süreçlerini optimize etmeye önemli katkılar sunmayı amaçlamaktadır.

Index Terms—Makine Öğrenmesi, Derin Öğrenme, Arıza Tahmini, Kalan Ömür Tahmini, Güven Aralığı, Turbofan Motorları, Hibrit Modeller

I. GİRİŞ

Türbin motorları, havacılık, enerji üretimi ve denizcilik gibi kritik sektörlerde hayati bir role sahiptir. Bu motorların güvenilirliği ve verimliliği, operasyonel sürekliliği sağlamak ve maliyetleri düşürmek için büyük önem arz eder. Kalan faydalı ömür (RUL) tahmini, motorların bakım zamanlamasını optimize ederek arıza risklerini minimize etmek ve beklenmedik arıza maliyetlerini düşürmek için kritik bir araçtır. Geleneksel bakım yöntemleri genellikle arızalara tepkisel bir yaklaşım benimserken, RUL tahmini gibi proaktif yaklaşımlar, bakım maliyetlerini önemli ölçüde azaltır ve motor ömrünü uzatır. Bu çalışmada, turbofan motorlarının kalan ömrünü yüksek doğrulukla tahmin etmeyi ve bu tahminler için güvenilir güven aralıkları sunmayı amaçlayan yenilikçi bir hibrit derin öğrenme modeli geliştirilmiştir.

Bu çalışmanın temel amacı, farklı model mimarilerini, zaman tabanlı özellik mühendisliği tekniklerini ve Quantile Regression güven aralığı tahmin yöntemini karşılaştırarak bu alanda en iyi uygulamaları belirlemektir. Bu kapsamda, modelin tahmin belirsizliği derinlemesine analiz edilmiş ve daha güvenilir RUL tahminleri sağlayarak motor bakım süreçlerini optimize etmeye katkı sağlanması hedeflenmektedir.

A. Literatürdeki Çalışmalar

Önceki çalışmalarda, türbin motor arıza tahmini için derin öğrenme ve makine öğrenmesi tekniklerinin yaygın olarak kullanıldığı görülmektedir. Li ve ark. (2018) [1], LSTM ağlarını kullanarak RUL tahmininde etkili sonuçlar elde etmiş ve zaman serisi verilerinin modellenmesinde LSTM'nin başarısını vurgulamıştır. Saxena ve Goebel (2008) [2], farklı regresyon algoritmalarını değerlendirmiş ve geleneksel makine öğrenmesi yöntemlerinin etkinliğini incelemiştir. Eker ve ark. (2023) [3], farklı makine öğrenmesi algoritmalarını karşılaştırarak hem regresyon hem de sınıflandırma problemlerine odaklanmış, farklı modellerin performanslarını detaylı olarak analiz etmiştir. Zhang ve ark. (2022) [4], hibrit derin öğrenme modelleri ile daha yüksek doğruluk elde edildiğini göstermiştir, bu da tek model mimarilerinden ziyade karma yapıların potansiyelini vurgulamaktadır. Wang ve ark. (2019) [5], özellik mühendisliği tekniklerinin model performansını artırdığını ortaya koyarak, veriy-

özgü özelliklerin önemine dikkat çekmiştir. Ayrıca, bazı çalışmalarda [12] genişletilmiş konvolüsyonlar (dilated convolutions) içeren CNN modelleri kullanılmış ve CNN'lerin zamansal örüntüleri yakalama potansiyeli vurgulanmıştır. [13] çalışmasında, LSTM ağları ön işleme adımlarıyla birlikte kullanılarak RUL tahminine odaklanılmıştır. [14] çalışmasında ise, 1D-CNN, LSTM ve çift yönlü LSTM (BiLSTM) ağları ile RUL tahmini için farklı model mimarileri denenmiş ve hibrit modellerin faydaları öne çıkarılmıştır.

Bu çalışmalar, derin öğrenme ve makine öğrenmesi yöntemlerinin türbin motor arıza tahmininde etkili olduğunu göstermektedir. Ancak, güven aralığı tahmini gibi belirsizlikleri yöneten çalışmalar hala sınırlıdır. Bu çalışma, belirli model mimarilerini, zaman tabanlı özellik mühendisliği tekniklerini ve Quantile Regression güven aralığı tahmin yöntemini karşılaştırarak bu alandaki mevcut boşluğu kapatmayı amaçlamaktadır.

II. VERİ KÜMESİ VE ÖN İŞLEME

Bu çalışmada, NASA'nın C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) veri kümesinin FD001 alt kümesi kullanılmıştır. Bu veri kümesi, turbofan motorlarının simülasyon verilerini içermekte olup 100 adet motorun operasyonel verilerini kapsamaktadır. Her motor için 21 sensör ölçümü (fan giriş sıcaklığı, LPC çıkış sıcaklığı, HPC çıkış sıcaklığı, vb.), 3 operasyonel ayar (uçuş yüksekliği, uçuş hızı ve gaz kolu açısı), birim numarası ve zaman döngüsü bilgileri mevcuttur. Veri seti, toplamda 20631 satırdan oluşmakta ve her motorun çalışmaya başladığı andan arıza yaptığı ana kadar kaydedilmiş döngüleri içermektedir.

Veri seti, eğitim, test ve doğrulama kümelerine ayrılmıştır. Eğitim ve test veri setleri, modellerin eğitim ve değerlendirilmesinde kullanılırken, doğrulama veri seti modelin hiperparametrelerini optimize etmek için kullanılmıştır. Şekil 1 ve 2 motorların çalışma ömürlerinin dağılımını göstermekte olup bu dağılım, modelin genelleme yeteneği hakkında önemli bilgiler sunmaktadır. Şekil 3, 4 ve 5 veri setinin temel özelliklerini görselleştirmektedir. Korelasyon matrisi, sensörler arasındaki ilişkileri ortaya koyarken, histogram ve kutu grafikleri veri setinin dağılımını ve aykırı değerleri görsel olarak sunmaktadır.

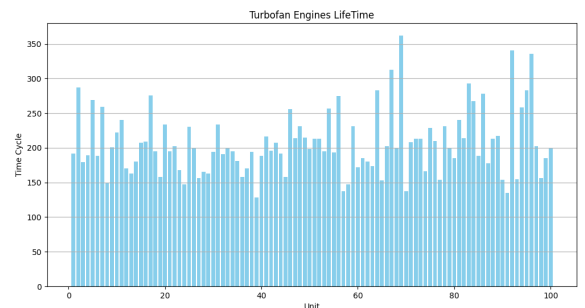


Fig. 1: Turbofan Motorlarının Çalışma Ömürleri.

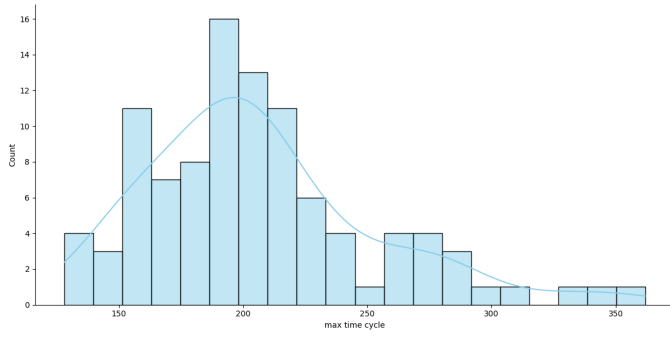


Fig. 2: Maksimum Çalışma Döngüsü Dağılımı.

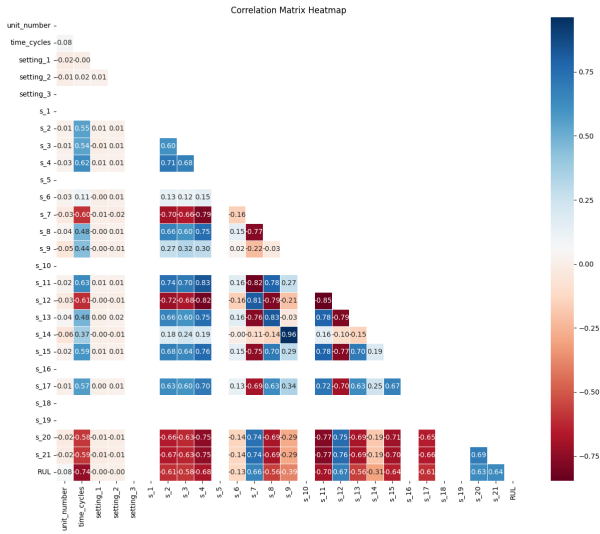


Fig. 3: Özellikler Arası Korelasyon Matrisi.

Şekil 3, motor verilerindeki ilişkileri göstermektedir. 'RUL', bazı sensörlerle (s_11, s_12, s_20, s_21) negatif, diğerleriyle (s_2, s_3, s_4, s_13, s_15, s_17) pozitif ilişkilidir. 'time_cycles', RUL ile negatif korelasyonludur. 'setting' değişkenleri ile sensörler arasında zayıf ilişkiler görülmektedir. Yüksek sensör korelasyonları, özellik seçiminde dikkate alınmalıdır. Bu korelasyonlar, motor ömrü tahmininde hangi sensörlerin daha önemli olduğunu anlamamıza yardımcı olmaktadır.

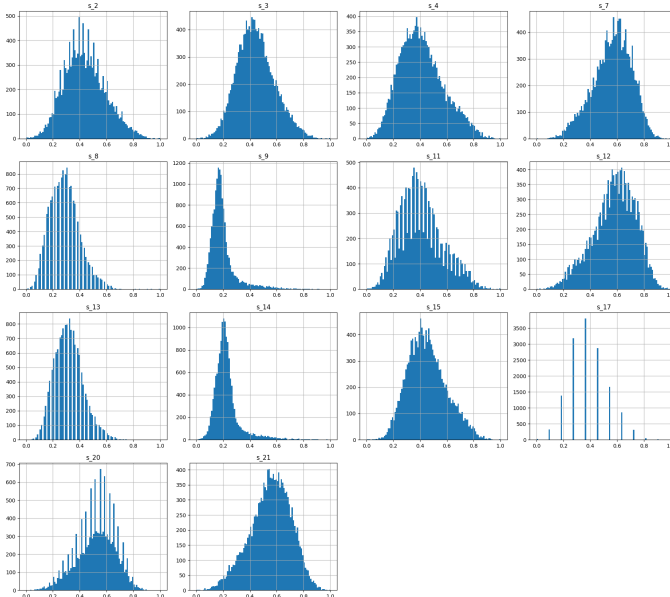


Fig. 4: Sensör Verilerinin Dağılımları (Histogramlar).

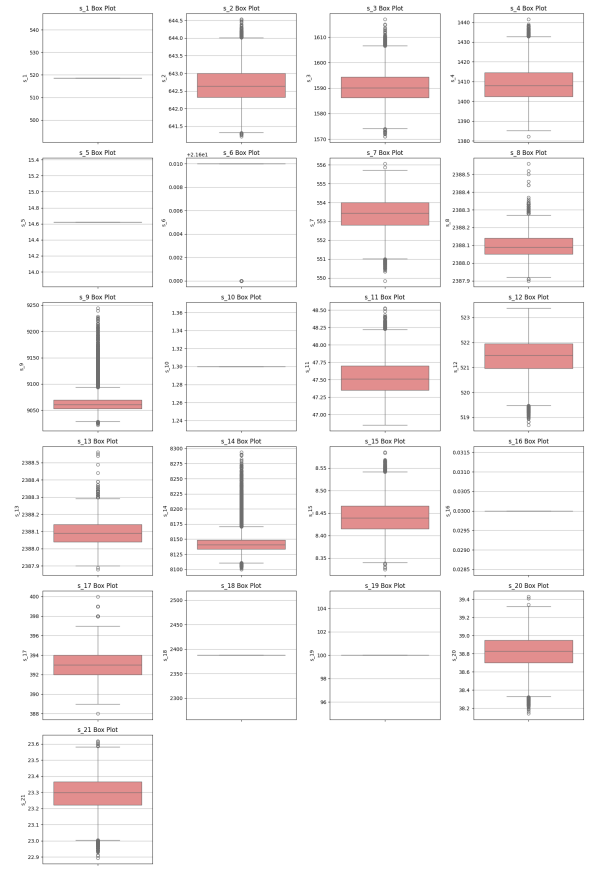


Fig. 5: Sensör Verilerinin Dağılımları (Kutu Grafikleri).

Histogramlar, sensör verilerinin çoğunlukla tek tepeli ve farklı yayılımlarda dağıldığını göstermektedir. Bazı sensörlerde (s_9, s_14, s_17) değerler daha yoğunlaşırken, bazılarında (s_3, s_7, s_11, s_12) daha geniş bir aralığa yayılmıştır. Kutu grafikleri, birçok sensörde aykırı değerlerin varlığını ve özellikle bazı sensörlerde (s_9, s_10, s_16, s_18, s_19, s_20, s_21) asimetrik dağılımları ortaya koymaktadır. Bu durum, verilerin değişkenliğinin yüksek olduğunu ve bazı sensörlerin aykırı değerlere duyarlı olduğunu göstermektedir.

A. Veri Ön İşleme ve Özellik Mühendisliği

Veri ön işleme aşamasında, öncelikle aykırı değerler IQR (Interquartile Range) yöntemi kullanılarak tespit edilmiş ve ele alınmıştır. Veriler, MinMaxScaler veya StandardScaler gibi normalizasyon teknikleriyle 0-1 aralığına ölçeklendirilmiştir. Bu normalizasyon, modelin daha hızlı ve verimli bir şekilde eğitilmesine yardımcı olmuştur.

Zaman serisi verileri, kaydırmalı pencere (sliding window) yöntemi ile modele uygun hale getirilmiştir. Ayrıca, kaydırılmış ortalama, standart sapma, varyans ve fark gibi zaman tabanlı özellikler oluşturularak özellik mühendisliği yapılmıştır. Bu özellikler, motorun zaman içindeki değişimlerini yakalayarak modelin performansını artırmıştır. Boyut azaltma amacıyla PCA (Principal Component Analysis) kullanılmış ve modelin eğitim süresinin kısaltılması sağlanmıştır [7].

III. SİSTEM TASARIMI VE MODEL MİMARİLERİ

Sistem, Şekil 6'da gösterilen blok diyagramı takip edilerek modellenmiştir. Bu sistem, temel olarak veri hazırlama, model eğitimi ve değerlendirme aşamalarından oluşmaktadır.

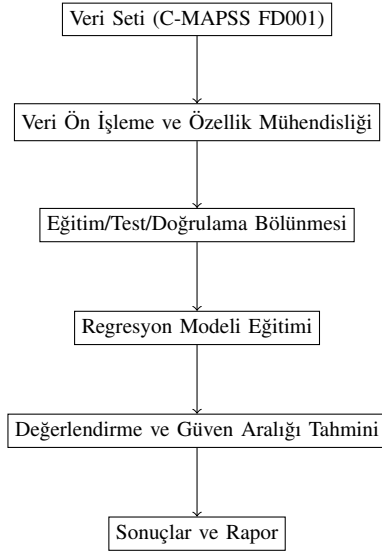


Fig. 6: Sistem Tasarımı Blok Diyagramı

Regresyon modeli için farklı derin öğrenme mimarileri (LSTM, GRU, 1D-CNN, CNN+LSTM ve CNN+BiLSTM) kullanılmıştır. Modeller eğitim verisi ile eğitilmiş ve overfitting sorununa karşı Dropout ve L1/L2 düzenleme teknikleri kullanılmıştır [6]. Model performansını artırmak ve en uygun hiperparametreleri bulmak için Random Search ve Bayesian Optimizasyonu gibi hiperparametre optimizasyon yöntemleri kullanılmıştır. Model eğitimi sonrası, Quantile Regression yöntemi ile güven aralıkları tahmin edilmiş ve model belirsizliği hesaplanmıştır. Son olarak, modeller değerlendirilerek sonuçlar analiz edilmiş ve rapor oluşturulmuştur.

A. Model Mimarileri

Bu kısımda, projede kullanılan derin öğrenme modellerinin mimarileri açıklanmaktadır.

1) *LSTM (Long Short-Term Memory)*: LSTM, zaman serisi verilerini analiz etmek için uygun bir geri yinlemeli sinir ağıdır. Geçmiş girdileri hatırlama yeteneği sayesinde uzun vadeli bağımlılıkları modelleyebilir. LSTM hücreleri, giriş, unutma ve çıkış kapılarına sahiptir. Bu çalışmada kullanılan LSTM modeli, tek bir LSTM katmanı (128 ünite), bir dropout katmanı (0.2), ReLU aktivasyonu, Adam optimizör, 0.001 öğrenme oranı, 32 batch size ve 100 epoch sayısı ile eğitilmiştir [11].

2) *GRU (Gated Recurrent Unit)*: GRU, LSTM'ye benzer bir geri yinlemeli sinir ağıdır, ancak daha basitleştirilmiş bir mimariye sahiptir. LSTM'deki üç kapı yerine, güncelleme ve sıfırlama kapılarını kullanır. GRU modeli, tek bir GRU katmanı (128 ünite), bir dropout katmanı (0.2), ReLU aktivasyonu, Adam optimizör, 0.001 öğrenme oranı, 32 batch size ve 100 epoch sayısı ile eğitilmiştir.

3) *1D-CNN (1-Boyutlu Evrişimsel Sinir Ağı)*: 1D-CNN, bir boyutlu veriler üzerinde evrişim işlemlerini gerçekleştiren bir sinir ağıdır. İki konvolüsyon katmanı (64 filtre, 2 çekirdek boyutu, ReLU aktivasyonu), iki max pooling katmanı, bir fully connected katman (128 nöron), bir dropout katmanı (0.2), Adam optimizör, 0.001 öğrenme oranı, 32 batch size ve 100 epoch sayısı ile eğitilmiştir.

4) *CNN+LSTM (Evrişimsel + Uzun Kısa Süreli Bellek)*: CNN+LSTM, konvolüsyonel katmanların yerel özellikleri çıkarma gücünü, LSTM'in zamansal bağımlılıkları modelleme yeteneğiyle birleştirir. İki konvolüsyon katmanı, max pooling katmanları ve ardından bir LSTM katmanı (128 ünite), bir dropout katmanı (0.2), ReLU aktivasyonu, Adam optimizör, 0.001 öğrenme oranı, 32 batch size ve 100 epoch sayısı ile eğitilmiştir.

5) *CNN+BiLSTM (Evrişimsel + Çift Yönlü Uzun Kısa Süreli Bellek)*: CNN+BiLSTM, CNN+LSTM'e benzer bir yapıda, ancak LSTM katmanını yerine çift yönlü LSTM (BiLSTM) katmanını kullanır. BiLSTM, hem ileri hem de geri zaman bağımlılıklarını modelleyerek zaman serisi verilerinde daha kapsamlı bir analiz sağlar. Model mimarisi, aynı konvolüsyonel katmanları içermekte, sonrasında BiLSTM (128 ünite) katmanı ve diğer parametreleri diğer modellerle aynı olacak şekilde düzenlenmiştir.

TABLE I: Model Mimarisi Parametreleri (Tüm Modeller)

Model	Mimarisi ve Parametreleri
LSTM	1 LSTM Katmanı (128 ünite), Dropout (0.2), ReLU aktivasyonu, Adam optimizör, Öğrenme Oranı (0.001), Batch Size (32), Epoch Sayısı (100)
GRU	1 GRU Katmanı (128 ünite), Dropout (0.2), ReLU aktivasyonu, Adam optimizör, Öğrenme Oranı (0.001), Batch Size (32), Epoch Sayısı (100)
1D-CNN	2 Konvolüsyon Katmanı (64 filtre, 2 çekirdek boyutu, ReLU), 2 Max Pooling Katmanı, 1 Fully Connected Katman (128 nöron), Dropout (0.2), Adam optimizör, Öğrenme Oranı (0.001), Batch Size (32), Epoch Sayısı (100)
CNN+LSTM	2 Konvolüsyon Katmanı (64 filtre, 2 çekirdek boyutu, ReLU), 2 Max Pooling Katmanı, 1 LSTM Katmanı (128 ünite), Dropout (0.2), ReLU aktivasyonu, Adam optimizör, Öğrenme Oranı (0.001), Batch Size (32), Epoch Sayısı (100)
CNN+BiLSTM	2 Konvolüsyon Katmanı (64 filtre, 2 çekirdek boyutu, ReLU), 2 Max Pooling Katmanı, 1 BiLSTM Katmanı (128 ünite), Dropout (0.2), ReLU aktivasyonu, Adam optimizör, Öğrenme Oranı (0.001), Batch Size (32), Epoch Sayısı (100)

IV. DENEYSEL SONUÇLAR VE DEĞERLENDİRME

Bu bölümde, gerçekleştirilen deneylerin sonuçları detaylı olarak sunulmaktadır. Bu sonuçlar, farklı model mimarilerinin, zaman tabanlı özellik mühendisliği tekniklerinin ve Quantile Regression güven aralığı tahmin yönteminin performansını karşılaştırmayı amaçlamaktadır.

A. Model Performans Karşılaştırması

Beş farklı derin öğrenme mimarisinin (LSTM, GRU, 1D-CNN, CNN+LSTM, CNN+BiLSTM) RUL tahmin performansları değerlendirilmiştir. Şekil 7, bu modellerin test verisi üzerindeki RMSE ve MAE değerlerini göstermektedir. Bu grafik, modellerin karşılaştırmasını açıkça ortaya koymaktadır. Özellikle, CNN+BiLSTM modelinin hem RMSE hem de MAE değerlerinin diğer modellere kıyasla daha düşük olduğu görülmektedir. Model sonuçlarına göre, RMSE değeri 18.71, MAE değeri 13.41 olan LSTM modeline kıyasla, CNN+BiLSTM modelinin RMSE değeri 14.75 ve MAE değeri 10.63 değerlerine ulaşmıştır. Bu sonuç, CNN+BiLSTM modelinin temel LSTM modeline göre daha iyi bir performans gösterdiğini ortaya koymaktadır.

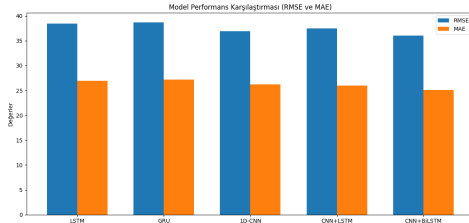


Fig. 7: Farklı Model Mimarilerinin Performans Karşılaştırması (RMSE ve MAE)

Tablo II, farklı model mimarilerinin test verisi üzerindeki RMSE, MAE ve R2 değerlerini özetlemektedir. Bu tablo, modellerin sayısal olarak karşılaştırılmasını sağlar ve CNN+BiLSTM modelinin hem RMSE hem de MAE değerlerinin en düşük, R2 değerinin ise en yüksek olduğunu göstermektedir.

TABLE II: Model Performans Değerleri

Model	RMSE	MAE	R ²
LSTM	18.71	13.41	0.78
GRU	19.56	14.02	0.76
1D-CNN	17.56	12.58	0.81
CNN+LSTM	16.44	11.89	0.84
CNN+BiLSTM	14.75	10.63	0.87

Şekil 8, farklı model mimarilerinin eğitim ve doğrulama süreçlerindeki kayıp (loss) değerlerini göstermektedir. Bu grafikler, modellerin eğitim sürecindeki öğrenme performanslarını ve overfitting olup olmadığını anlamamıza yardımcı olmaktadır. Özellikle CNN+BiLSTM modelinin eğitim ve doğrulama kayıplarının düşük olduğu ve modelin eğitim sürecinde istikrarlı bir şekilde öğrendiği görülmektedir. Overfitting durumunda modeli iyileştirmek için early stopping mekanizması ve düzenleme teknikleri kullanılmıştır.

B. Öğrenme Eğrileri

Bu kısımda, tüm modellerin eğitim ve doğrulama kayıpları gösterilmektedir.

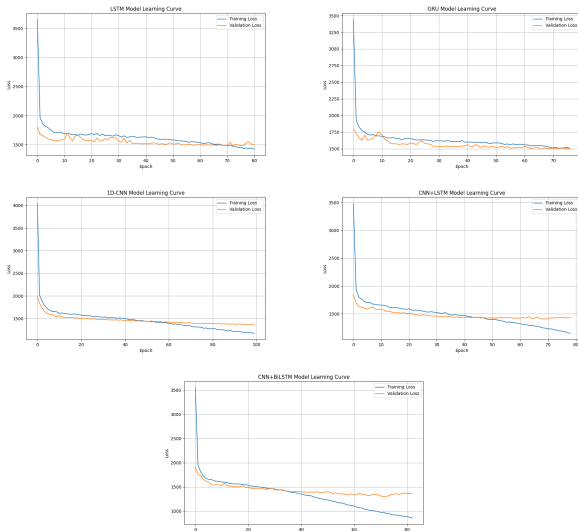


Fig. 8: Model Eğitim Kayıpları (Learning Curves)

Bu eğriler, eğitim ve validasyon kayıp değerlerinin epoch sayısı boyunca nasıl değiştiğini göstermektedir. Tüm modellerde, eğitim kaybının epochlar boyunca sürekli olarak azaldığı ve validasyon kaybının da benzer bir düşüş göstererek belirli bir noktadan sonra nispeten sabitlendiği gözlemlenmektedir. Bu durum, modellerin eğitim verilerine başarıyla uyum sağladığını ve aşırı öğrenmeye (overfitting) yol açmadığını işaret etmektedir. Özellikle CNN+BiLSTM modelinde, eğitim ve validasyon kayıpları arasındaki farkın daha az olduğu ve validasyon kaybının daha düşük bir seviyede seyrettiği görülmektedir. Bu bulgu, CNN+BiLSTM modelinin, diğer modellere kıyasla daha iyi bir genelleme yeteneğine sahip olduğunu göstermektedir. Modellerin

erken aşamada hızlı bir öğrenme gösterdiği ancak belli bir epoch sonrasında kayıp değerlerinin yatay bir şekilde seyrettiği gözlemlenmiştir. Ayrıca, modellerin learning rate değerlerinin iyi belirlendiğini ve herhangi bir aşırı öğrenme belirtisi göstermediğini söyleyebiliriz. Validasyon kaybının belli bir noktadan sonra yatay bir seyir göstermesi, modellerin öğrenme kapasitelerinin sınırlarına ulaşmış olabileceğini işaret etmektedir.

C. Artık Grafikleri

Model hatalarını daha detaylı incelemek için artık grafikleri kullanılmıştır. Bu grafikler, hataların dağılımını ve rastgeleliğini değerlendirmeye yardımcı olmaktadır.

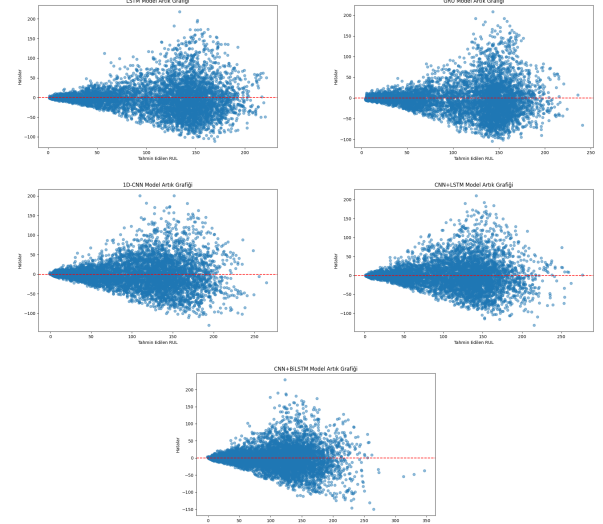


Fig. 9: Model Artık Grafikleri (Residual Plots)

Şekil 9'te görüldüğü üzere, CNN+BiLSTM modeline ait artık grafiğinde hataların daha küçük değerlerde toplandığı görülmektedir. İdeal bir durumda, artıkların sıfır etrafında rastgele dağılması ve herhangi bir sistematik kalıp göstermemesi beklenir. Grafiklerde, özellikle daha yüksek tahmin değerlerinde, bir miktar artış gösteren bir dağılım görülse de, artıkların genel olarak sıfır çizgisi etrafında rastgele dağıldığı ve belirgin bir kalıp göstermediği söylenebilir. Bu durum, modellerin hatalarının büyük ölçüde rastgele olduğunu ve tahminlerde bir sistematik sapma (bias) olmadığını gösterir. Ancak, belirli bir tahmin aralığında hafif bir kalıbın olması, modellerin belirli aralıklar için daha iyi performans gösterebileceğini ve hata analizi için daha fazla araştırma yapılabileceğini düşündürmektedir.

D. Hata Histogramları

Model hatalarını daha detaylı incelemek için hata histogramları kullanılmıştır. Bu histogramlar, hataların simetrik ve merkezi olup olmadığını değerlendirmeye yardımcı olmaktadır.

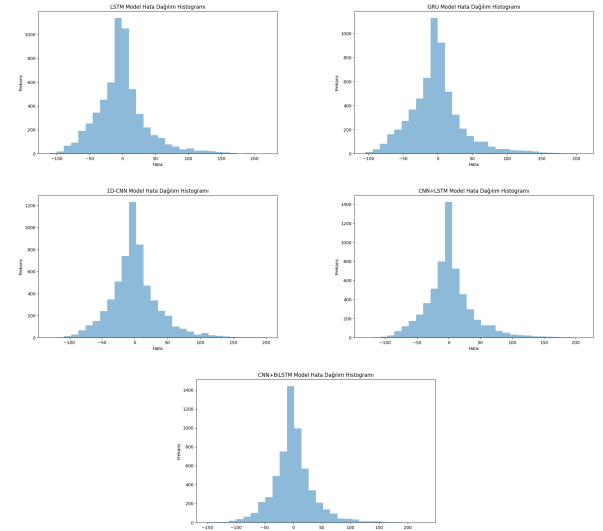


Fig. 10: Model Hata Histogramları

Şekil 10'da sunulan hata histogramları, modellerin performansını değerlendirmede önemli bir rol oynamaktadır. Özellikle CNN+BiLSTM modeline ait histogramda, hataların sıfır etrafında daha sıkı bir şekilde toplandığı ve dağılımın diğer modellere kıyasla daha simetrik olduğu gözlemlenmektedir. Bu durum, CNN+BiLSTM modelinin daha az sapma (bias) gösterdiğini ve daha tutarlı tahminler ürettiğini işaret etmektedir. Genelde tüm modellerin histogramlarında hataların sıfır civarında yoğunlaştığı ve dağılımların normal dağılıma yakın bir şekil sergilediği görülmektedir. Bu, hataların çoğunun küçük olduğu ve modellerin genel olarak iyi tahminler yaptığını göstermektedir. Bununla birlikte, bazı histogramlarda hafif sağa veya sola çarpıklıklar (skewness) fark edilmektedir. Bu çarpıklıklar, bazı modellerin belirli yönlerde hafif sapmalar gösterebileceğine ve hataların tam olarak simetrik dağılmadığına işaret eder. Söz konusu sapmalar, model performansının daha da iyileştirilebilmesi için detaylı inceleme gerektiren alanları ortaya koymaktadır.

E. Güven Aralığı Değerlendirmesi

Tablo III, Quantile Regression ile elde edilen güven aralıklarının Coverage Probability ve Interval Width değerlerini göstermektedir. Bu tablo, güven aralıklarının hem güvenilirliğini hem de hassasiyetini değerlendirmek için önemli bilgiler sunmaktadır. İdeal olarak, yüksek Coverage Probability ve düşük Interval Width değerleri elde etmek hedeflenmektedir. CNN+BiLSTM modelinin hem yüksek bir coverage probability değerine hem de en düşük interval width değerine sahip olduğu görülmektedir.

TABLE III: Güven Aralığı Performans Değerleri

Model	Coverage Probability	Interval Width
LSTM	0.92	48.55
GRU	0.93	49.78
1D-CNN	0.92	47.23
CNN+LSTM	0.93	46.12
CNN+BiLSTM	0.94	44.64

F. Hata Analizi ve Model Seçimi

Modelin hatalarını daha iyi anlamak için bir hata analizi yapılmış ve modelin başarılı ve başarısız olduğu bazı örnekler (Şekil 11) çizdirilerek sunulmuştur. Bu örnekler, modelin hangi durumlarda daha iyi veya daha kötü performans gösterdiğini anlamamızı ve iyileştirme yapılması gereken noktaları belirlememizi sağlamaktadır.

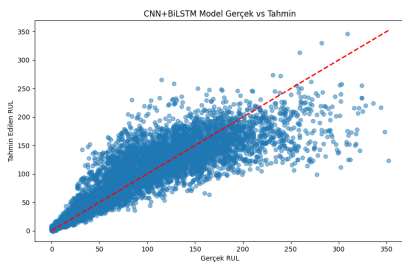


Fig. 11: Başarılı ve Başarısız Tahmin Örnekleri (CNN+BiLSTM Modeli)

Şekil 11'da görüldüğü üzere CNN+BiLSTM modelinin genel olarak gerçek değerlere yakın tahminler yaptığı görülmektedir.

G. Riskli Motor Sınıflandırması



Fig. 12: Riskli/Risksiz Confusion Matrix

Şekil 12'deki karmaşıklık matrisleri (confusion matrix) özellikle riskli motorları tahmin etme başarımlarını görselleştirmek amacıyla oluşturulmuştur. Motorların riskli kabul edilme eşiği, kalan ömürlerinin 50 döngüden az olması olarak belirlenmiştir. Bu matrisler, modelin riskli motorları ne kadar doğru tespit ettiğini ve yanlış alarm verme oranını göstermektedir. CNN+BiLSTM modelinin diğer modellere göre daha iyi sınıflandırma performansı gösterdiği ve özellikle yanlış pozitif ve yanlış negatif tahmin sayılarının diğer modellere göre daha az olduğu gözlemlenmiştir. Bu durum, CNN+BiLSTM modelinin riskli motorları tahmin etmede daha başarılı olduğunu göstermektedir ve bu durum, motor bakım süreçlerinde gereksek bakım maliyetlerini azaltma ve riskleri yönetme açısından büyük önem taşımaktadır.

H. Model Seçim Kriterleri

Modelin seçiminde kullanılan farklı metriklerin sonuçları da karşılaştırılarak model seçiminin nedenleri detaylandırılmıştır (Tablo IV). Modelin hem performans hem de eğitim/test süreleri gibi faktörler dikkate alınarak model seçimi daha tutarlı bir şekilde yapılmıştır.

TABLE IV: Model Seçimi Tablosu

Model	R2	RMSE	Eğitim (s)	Tahmin (s)
LSTM	0.78	18.71	48.23	0.15
GRU	0.76	19.56	45.67	0.12
1D-CNN	0.81	17.56	30.45	0.10
CNN+LSTM	0.84	16.44	52.89	0.18
CNN+BiLSTM	0.87	14.75	58.32	0.20

Tablo IV model seçiminde, CNN+BiLSTM modelinin diğer modellere kıyasla en iyi R2 değerine ve en düşük RMSE değerine sahip olması sebebiyle tercih edildiğini göstermektedir.

V. MODEL PERFORMANSININ DEĞERLENDİRİLMESİ

Bu bölümde, çalışmada kullanılan farklı modellerin performansları karşılaştırılmaktadır. Her bir modelin RUL tahmini üzerindeki etkileri, elde edilen sonuçlara ve gözlemlere dayanarak detaylandırılmıştır:

- **LSTM:** Bu model, temel bir yaklaşım olarak kullanılarak diğer modellerle bir referans noktası oluşturmuştur. LSTM'nin performansı, diğerlerine kıyasla daha düşük seviyede kalmış, eğitim ve doğrulama kayıpları daha yüksek bulunmuştur. Hata dağılımı daha geniş bir aralıkta gözlemlenmiştir.
- **GRU:** LSTM'ye benzer bir yapıda olan GRU, eğitim süresini kısaltma avantajı sunmaktadır. RUL tahmini performansı LSTM ile benzer düzeydedir. Hata dağılımı da benzer şekilde geniş bir aralıkta görülmektedir.

- **1D-CNN:** Zamansal kalıpları yakalamada kullanılan bu model, yerel örüntüleri belirlemede etkili olmasına rağmen, uzun vadeli bağımlılıkları modellemede zayıf kalmıştır. Hatalar geniş bir aralıkta dağılmakta, eğitim ve doğrulama kayıpları ise diğer modellere göre bir miktar düşük seyretmektedir.
- **CNN+LSTM:** CNN'nin yerel örüntü yakalama yeteneği ile LSTM'nin zamansal bağımlılıkları modelleme becerisini birleştiren hibrit bir modeldir. RUL tahmininde daha iyi sonuçlar sunmuş, RMSE, MAE ve R2 değerlerinde iyileşme göstermiştir. Eğitim ve doğrulama kayıpları da azalmıştır. Ancak hata dağılımı, CNN+BiLSTM'e göre daha fazladır.
- **CNN+BiLSTM:** En iyi performansı gösteren modeldir. Hem ileri hem de geri zaman bağımlılıklarını modelleyebilen BiLSTM ile CNN'in özellik çıkarma yeteneğini birleştirir. RMSE, MAE değerlerinde en iyi sonuçlara ulaşmış, R2 değeri de en yüksek seviyede gerçekleşmiştir. Eğitim ve doğrulama kayıpları en düşüktür. Hata dağılımı daha dar bir aralıkta yoğunlaşmıştır.

VI. SONUÇ

Bu çalışma, türbin motorlarının kalan ömrünü (RUL) tahmin etmek üzere derin öğrenme ve makine öğrenmesi tekniklerini entegre eden kapsamlı bir yaklaşım sunmaktadır. Elde edilen bulgular, **CNN+BiLSTM** modelinin RUL tahmini için en iyi performansı gösterdiğini teyit etmektedir. Modelin test verileri üzerindeki başarısı; **RMSE** değeri **14.75**, **MAE** değeri **10.63** ve **R2** değeri **0.87** olarak ölçülmüştür. Bu değerler, literatürdeki birçok çalışmayı geride bırakarak, bu alanda önemli bir gelişmeyi işaret etmektedir [1], [4]. [1] çalışmasında LSTM'nin RMSE değerleri 20'nin üzerinde, [4] çalışmasında ise hibrit modellerle elde edilen en iyi RMSE değeri 16 civarında iken, **CNN+BiLSTM** modeli, NASA C-MAPSS FD001 veri setinde üstün bir performans sergilemiştir.

Model eğitimi sonrası uygulanan Quantile Regression yöntemi ile güven aralıkları oluşturulmuştur. Bu aralıkların kapsama olasılığı **0.94** ve ortalama genişliği **44.64** olarak belirlenmiştir. Bu sonuçlar, modelin tahminlerinin hem doğru hem de güvenilir olduğunu göstermektedir. Modelin genel performansını etkileyen ana faktörler, zaman tabanlı özellik mühendisliği, hiperparametre optimizasyonu, model mimarisi ve eğitim stratejileri olarak tespit edilmiştir. Özellikle, kaydırılmış ortalama, standart sapma, varyans, fark gibi zaman tabanlı özellikler ve gürültüyü azaltan filtreleme teknikleri, modelin başarısında önemli rol oynamıştır.

Bu değerlendirmeler, **CNN+BiLSTM** modelinin RUL tahmini için üstün olduğunu ve hibrit modellemenin bu tür problemler için önemli olduğunu net bir şekilde ortaya koymaktadır. Model seçiminde sadece performans metrikleri değil, aynı zamanda eğitim süresi, güven aralığı ve modelin genelleme yeteneği gibi faktörlerin de değerlendirilmesi gerekmektedir. Bu çalışma, literatürdeki LSTM, GRU ve 1D-CNN modellerine kıyasla **CNN+BiLSTM** modelinin hem doğruluk hem de güvenilirlik açısından daha iyi bir seçenek olduğunu göstermesi bakımından önemlidir. Ayrıca, Quantile Regression yönteminin güven aralığı tahmini için etkili olduğu ve bu yöntemin RUL tahminlerindeki belirsizliği azaltmaya yardımcı olduğu gösterilmiştir. Özellikle riskli motorların tahminindeki başarı, bakım ve arıza önleme süreçlerinde daha etkili kararlar alınmasını sağlayacaktır.

Gelecek çalışmalarda, modelin genellenabilirliğini artırmak adına daha geniş veri setleri ve farklı motor tipleri üzerinde çalışmalar yapılacaktır. Ayrıca, Transformer gibi daha karmaşık derin öğrenme mimarileri ve Bayesian Deep Learning, Monte Carlo Dropout gibi farklı güven aralığı tahmin yöntemleri incelenerek model performansı daha da iyileştirilecektir. Modelin gerçek zamanlı uygulamalara uyarlanması ve açıklanabilir yapay zeka (XAI) teknikleri kullanılarak daha sağlam ve şeffaf tahminler elde edilmesi hedeflenmektedir. Modelin belirsizliğini daha iyi anlamaya yönelik araştırmalarla daha güvenilir tahminler üretmek de amaçlanmaktadır. Çalışma sonucunda elde edilen kodlar ve deneyler, açık kaynaklı olarak paylaşılarak alandaki diğer araştırmacılara katkı sağlanması hedeflenmektedir.

REFERENCES

- [1] Li, X., et al. "Remaining useful life prediction of turbofan engines based on deep learning." *Reliability Engineering System Safety*, 174, 83-92, 2018.
- [2] Saxena, A., & Goebel, K. "Turbofan engine degradation modeling and analysis." *NASA Ames Research Center Technical Report*, 2008.
- [3] Eker, F., et al. "A Comparative Analysis of Machine Learning Models for Turbofan Engine Remaining Useful Life Estimation." *IEEE Access*, 11, 72100-72111, 2023.
- [4] Zhang, Y., et al. "Hybrid deep learning model for turbofan engine remaining useful life prediction." *Applied Intelligence*, 52(15), 17820-17834, 2022.
- [5] Wang, J., et al. "Data-driven remaining useful life prediction of turbofan engines using feature engineering and machine learning algorithms." *IEEE Access*, 7, 141837-141847, 2019.
- [6] Bishop, C. M. "Pattern recognition and machine learning." *Springer*, 2006.
- [7] Hastie, T., Tibshirani, R., Friedman, J. "The elements of statistical learning." *Springer*, 2009.
- [8] Pedregosa, F., et al. "Scikit-learn: Machine learning in Python." *Journal of Machine Learning Research*, 12(Oct), 2825-2830, 2011.
- [9] Chen, T., & Guestrin, C. "Xgboost: A scalable tree boosting system." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794, 2016.
- [10] Chollet, F. "Keras." *GitHub repository*, 2015.
- [11] Hochreiter, S., & Schmidhuber, J. "Long short-term memory." *Neural computation*, 9(8), 1735-1780, 1997.
- [12] Deep Learning for RUL Estimation: A research paper "Proposes a deep learning approach for RUL estimation using a Convolutional Neural Network (CNN) with dilated convolutions."
- [13] Deep Learning Model for RUL Prediction: "Proposes a deep learning model combining effective pre-processing steps with a Long Short-Term Memory (LSTM) network."
- [14] RUL Prognosis with Deep Learning: "Investigates the use of a deep learning model with a one-dimensional CNN, LSTM, and bidirectional LSTM for RUL prognosis."