DENİZ TAŞIMACILIĞINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE TALEP ve TESLİM SÜRESİ TAHMİNİ

Lara J M NASER

T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DENİZ TAŞIMACILIĞINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE TALEP ve TESLİM SÜRESİ TAHMİNİ

Lara J M NASER

Doç. Dr. Seval ENE YALÇIN
(Danışman)

YÜKSEK LİSANS TEZİ
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI

BURSA – 2025

ÖZET

Yüksek Lisans Tezi

DENİZ TAŞIMACILIĞINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE TALEP ve TESLİM SÜRESİ TAHMİNİ

Lara J M NASER

Bursa Uludağ Üniversitesi
Fen Bilimleri Enstitüsü
Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Seval ENE YALÇIN

Değişken küresel ticaret dinamiklerinin yaşandığı bir çağda, deniz taşımacılığı uluslararası lojistiğin temel bir bileşenidir. Ancak deniz taşımacılığı, dalgalanan talep, operasyonel kısıtlar ve çevresel belirsizliklerden kaynaklanan kesintilere maruz kalmaktadır. Tez çalışmasının temel amacı, Türk denizcilik sektöründe operasyonel karar alma süreçlerini iyileştirmek üzere, konteyner iş hacmi talebi ve teslim süresi tahmininde yüksek doğruluk sağlayan, veri odaklı ve pratik uygulanabilir makine öğrenmesi (ML) modelleri geliştirmektir. Gerçek dünya verilerine dayalı olarak, çalışma, konteyner iş hacmi talebini ve teslim süresini etkileyen karmaşık, doğrusal olmayan etkileşimleri yakalamak için hibrit mimariler dahil olmak üzere bir dizi tahmin modeli geliştirmekte, eğitmekte ve değerlendirmektedir.

Tez çalışmasında, denetimli öğrenme metodolojileri benimsenmiş ve çeşitli ML algoritmaları karşılaştırılmıştır. Temel etkileyici faktörleri belirlemek için gri ilişkisel analiz yöntemi kullanılmıştır. Önceki çalışmalardan farklı olarak, bu çalışmada, konteyner iş hacmi çıktısı ve sevkiyat teslim sürelerini eş zamanlı olarak tahmin etmek için entegre bir yaklaşım sunulmuştur. Bu iki odaklı yaklaşım, mevcut literatürde nadiren birlikte ele alınmıştır.

Önerilen sistem, geleneksel istatistiksel yöntemlere kıyasla tahmin doğruluğunda önemli iyileştirmeler göstererek, nakliye operatörleri, liman otoriteleri ve tedarik zinciri paydaşları için eyleme dönüştürülebilir içgörüler sağlamaktadır. Sonuçlar, önerilen modellerin literatürdeki yaklaşımlara kıyasla daha yüksek tahmin doğruluğu sunduğunu ortaya koymuş, ayrıca, geliştirilen kullanıcı dostu arayüz sayesinde modellerin saha uygulamalarında etkin biçimde kullanılabileceği anlaşılmıştır.

Bu tez, yalnızca denizcilik lojistiği tahminlemesinin akademik anlayışını ilerletmekle kalmayıp, aynı zamanda Türk ticaret altyapısının bölgesel karmaşıklıklarına uyarlanmış ölçeklenebilir araçlar da sunmaktadır.

Anahtar Kelimeler: Deniz taşımacılığı, makine öğrenmesi, konteyner iş hacmi tahmini, teslim süresi tahmini
2025, xix + 94 sayfa.

ABSTRACT

MSc Thesis

DEMAND AND LEAD TIME PREDICTION WITH MACHINE LEARNING METHODS IN MARITIME SHIPPING

Lara J M NASER

Bursa Uludağ University
Graduate School of Natural and Applied Sciences
Department of Industrial engineering

Supervisor: Assoc. Prof. Dr. Seval ENE YALÇIN

In an era of volatile global trade dynamics, maritime transport is a fundamental component of international logistics. However, maritime transport is subject to disruptions caused by fluctuating demand, operational constraints, and environmental uncertainties. The primary objective of this thesis is to develop data-driven and practically applicable machine learning (ML) models that provide high accuracy in predicting container volume demand and delivery times, with the aim of improving operational decision-making processes in the Turkish maritime sector. Based on real-world data, the study develops, trains, and evaluates a series of prediction models, including hybrid architectures, to capture the complex, non-linear interactions that influence container cargo volume demand and delivery times.

The thesis adopts supervised learning methodologies and compares various ML algorithms. The grey relational analysis method has been used to identify the key influencing factors. Unlike previous studies, this study presents an integrated approach to simultaneously predict container cargo volume output and shipment delivery times. This dual-focused approach is rarely addressed together in the existing literature.

The proposed system provides actionable insights for shipping operators, port authorities, and supply chain stakeholders, demonstrating significant improvements in prediction accuracy compared to traditional statistical methods. The results demonstrate that the proposed models offer higher prediction accuracy compared to approaches in the literature, and the developed user-friendly interface makes them usable for field applications.

This thesis not only advances the academic understanding of maritime logistics forecasting but also provides scalable tools adapted to the regional complexities of Turkey's trade infrastructure.

Key words: Maritime shipping, Machine learning, Container throughput prediction, Lead time prediction
2025, xix + 94 pages.

ÖNSÖZ VE/VEYA TEŞEKKÜR

Bu çalışma, öğrenim sürecim boyunca edindiğim bilgi ve deneyimleri akademik bir çerçevede değerlendirme ve uygulama fırsatı sunmuştur. Çalışmanın her aşamasında kazandığım deneyimler, mesleki gelişimime önemli katkılar sağlamıştır.

Araştırmanın yürütülmesi ve tamamlanması sürecinde bilgi birikimi, rehberliği ve yapıcı katkılarıyla bana yol gösteren değerli danışmanım **Doç. Dr. Seval Ene Yalçın**'a en içten teşekkürlerimi sunarım. Kendisinin akademik desteği, çalışmanın şekillenmesinde ve bilimsel niteliğinin güçlenmesinde önemli rol oynamıştır.

Ayrıca her koşulda yanımda olan ve manevi desteklerini hiçbir zaman esirgemeyen sevgili aileme; **annem Şohra Naser'e, babam Jamal Naser'e, kardeşlerime ve arkadaşlarıma** teşekkürü bir borç bilirim. Varlıkları bana her zaman güç ve motivasyon kaynağı olmuştur.

Bu süreçte emeği geçen tüm kişi ve kurumlara da ayrıca teşekkür ederim.

Lara J M NASER
28/07/2025

İÇİNDEKİLER

Sayfa

ÖZET vii
ABSTRACT viii
ÖNSÖZ VE/VEYA TEŞEKKÜR ix
SİMGELER ve KISALTMALAR DİZİNİ xii
ŞEKİLLER DİZİNİ xviii
ÇİZELGELER DİZİNİ xix
1. GİRİŞ 1
2. KAYNAK ARAŞTIRMASI 4
2.1. Makine Öğrenimi 4
2.2. Talep Tahmini 6
2.3. Teslim Süresi Tahmini 9
3. MATERYAL VE YÖNTEM 17
3.1. Tahmin Modelleri 17
3.1.1. Rastgele orman 18
3.1.2. Çoklu doğrusal regresyon 19
3.1.3. Destek vektör regresyonu 19
3.1.4. Aşırı gradyan güçlendirme 21
3.1.5. Yapay sinir ağları 21
3.1.6. Yinelemeli sinir ağları 22
3.1.7. Uzun kısa süreli bellek modeli 23
3.1.8. CatBoost regresyonu 24
3.1.9. En küçük mutlak küçültme ve seçim operatörü regresyonu 25
3.1.10. Hafif gradyan güçlendirme makinesi 25
3.1.11. Evrişimsel sinir ağları ve uzun kısa süreli bellek hibrit modeli 26
3.1.12. Destek vektör regresyonu ve yapay sinir ağı hibrit modeli 27
3.1.2. Model performans metrikleri 29
3.2.1. Konteyner iş hacminin talebi için geliştirilen tahmin modeli 30
3.2.1.1 Veri toplama 31
3.2.1.2. Veri ön işleme 33
3.2.1.3. Eksik veri işleme 33
3.2.1.4. Modelleme ve değerlendirme 33
3.2.1.5. Modellerin doğrulanması 34
3.2.2. Teslim süresi için geliştirilen tahmin modeli 36
3.2.2.1. Veri toplama 36
3.2.2.2. Veri ön işleme 42
3.2.2.3. Eksik veri işleme 43
3.2.2.4. Modelleme ve değerlendirme 43
3.2.2.5. Modellerin doğrulanması 44
4. BULGULAR ve TARTIŞMA 47
4.1. Konteyner İş Hacmi Talebinin Tahmini için Geliştirilen Modellerin Sonuçları 47
4.1.1. İstatistiksel analiz 47
4.1.2. Gri ilişki analizi 48
4.1.3. Makine öğrenmesi modellerinin sonuçları 50
4.1.4. Modellerin daha önce görmediği verilerin tahmini 55
4.1.5. 2025 yılı konteyner iş hacmi talep tahmini 56
4.2. Teslim Süresinin Tahmini için Geliştirilen Modellerin Sonuçları 59

4.2.1.İstatistiksel analiz 59
4.2.2. Gri ilişki analizi 63
4.2.3. Makine öğrenmesi modellerinin sonuçları 65
4.2.3.1. Düşük akış veri seti için modellerin sonuçları 65
4.2.3.2. Yüksek akış veri seti için modellerin sonuçları 67
4.2.4. Modellerin daha önce görmediği verilerin tahmini 68
4.2.4.1. Düşük akış veri setinde modellerin daha önce görmediği verilerin tahmini 69
4.2.4.2. Yüksek akış veri setinde modellerin daha önce görmediği verilerin tahmini 70
4.3. Kullanıcı Dostu Arayüz 72
4.3.1. Kullanıcı arayüzünün temel özellikleri 73
4.3.1.1. Esnek veri girişi yöntemleri 73
4.3.1.2. ML tabanlı tahmin sistemi 73
4.3.2. Kullanıcı arayüzünün tasarımı 74
4.3.3. Kullanıcı arayüzünün katkısı 74
5. SONUÇ 76
KAYNAKLAR 79
EKLER 86
ÖZGEÇMİŞ 94

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler	Açıklama
x_g	Girdi örneği
y_g	Hedef çıktı
$\hat{y}_g$	Tahmin edilen çıktı
K	Rastgele orman'da ağaç sayısı
$L(h_k)$	Bootstrap örneği
$T_{L(h_k)}$	K. ağaçtan gelen tahmin
MSE_{OOB}	Rastgele orman'da Torba-Dışı (OOB) örnekler için ortalama kare hatası
$Var(y)$	Hedef değişkeninin varyansı
Y_i	Gözlem i için yanıt değişkeni
α	Regresyon modellerinde sabit terim
β	Regresyon modellerinde katsayı vektörü
e_i	Gözlem i için hata terimi
w	Destek vektör regresyonunda ağırlık vektörü
C	Destek vektör regresyonunda düzenlileştirme parametresi
ξ_i^* , ξ_i	Destek vektör regresyonunda gevşek değişkenler
ε	Destek vektör regresyonunda hata toleransı
x_i	Destek vektör regresyonunda girdi örneği
y_i	Destek vektör regresyonunda hedef çıktı
N	Toplam eğitim örneği sayısı
ϕ	Aşırı gradyan artırmada kayıp fonksiyonu
y_i	Aşırı gradyan artırmada i. eğitim örneğinin gerçek değeri
$\hat{y}_i^t$	Aşırı gradyan artırmada t artırma iterasyonunda i örneği için tahmin edilen değer
$l(y_i, \hat{y}_i^t)$	y_i ve $\hat{y}_i^t$ karşılaştıran kayıp fonksiyonu
f_k	Aşırı gradyan artırmada k. regresyon ağacı
$\Omega(f_k)$	Aşırı gradyan artırmada k. ağaç için düzenlileştirme terimi
$f_t(x_i)$	Aşırı gradyan artırmada i. ağaçtan gelen tahmin
γ	Yaprak sayısını cezalandıran düzenlileştirme parametresi
T	Ağaçtaki yaprak sayısı
λ	L2 düzenlileştirme terimi katsayısı
ω	Yaprak ağırlıkları vektörü
$\|\omega\|^2$	Yaprak ağırlıklarının kare L2 normu
$\hat{y}_i^{(t-1)}$	Önceki t-1 iterasyonunda i örneği için tahmin
$f_t(x_i)$	Örnek i için yeni t. regresyon ağacının çıktısı
z	Giriş ağırlıklı toplamı artı bias
n	Toplam giriş özelliği veya nöron sayısı
b	Bias terimi
y	Yapay sinir ağlarında aktivasyon fonksiyonu uygulandıktan sonra nöronun çıktısı
$\varphi(z)$	z 'ye uygulanan aktivasyon fonksiyonu
z^{-1}	Tekrarlayan sinir ağlarında birim-zaman gecikme operatörü
$F^{(t)}$	t İterasyonundaki tahmin fonksiyonu

$F^{(t-1)}$	Önceki t iterasyonundan gelen tahmin fonksiyonu
α	Öğrenme oranı
h_t	t İterasyonundaki regresyon ağacı
$\hat{x}_k^i$	k-inci örnekteki i. kategorik özellik değeri
x_k^i	j Örneğindeki i. kategorik özellik değeri
$\sigma(j)$	j Örneğinin rastgele permütasyon konumu
$\sigma(k)$	Mevcut k örneğinin aynı permütasyondaki konumu
$1(x_j^i = x_k^i)$	Gösterge fonksiyonu
b	Düzgünleştirme parametresi
p	Önsel, hedef değişkenin genel ortalaması
$\min_{\beta}$	En küçük mutlak büzülme ve seçim operatörü regresyonunda ifadeyi minimize eden katsayıları kümesini bulma optimizasyon hedefi
x_{ij}	En küçük mutlak büzülme ve seçim operatörü regresyonunda i. gözlem için j. öngörücü değeri
β_j	En küçük mutlak büzülme ve seçim operatörü regresyonunda j öngörücüsüne atanan katsayı (ağırlık)
n	En küçük mutlak büzülme ve seçim operatörü regresyonunda gözlem sayısı
p	En küçük mutlak büzülme ve seçim operatörü regresyonunda öngörücü sayısı
$L1$	Ceza terimi
λ	$L1$ Ceza teriminin ağırlığını kontrol eden düzenlileştirme parametresi
A_i	i. gözlem için gerçek değer
P_i	i. gözlem için tahmin edilen değer
E_i	i. gözlem için hata terimi
n	Toplam gözlem sayısı
$A_{ortalama}$	Gerçek değerlerin ortalaması
K	Çapraz doğrulama sürecindeki kat sayısı
β_0	Sabit terim
β_1	j. bağımsız değişken için katsayı
ϵ_i	i gözlemi için hata terimi
α_0	Yardımcı regresyonda sabit terim
α_j	Yardımcı modelde bağımsız değişkeni için katsayı
u_i	Yardımcı regresyonda hata terimi
R_{aux}^2	Yardımcı regresyondan belirleme katsayısı
χ^2	Ki-kare dağılımı
p	Modeldeki bağımsız değişken sayısı

Kısaltmalar	Açıklama
1D CNN	Tek Boyutlu Evrişimli Sinir Ağı (One-Dimensional Convolutional Neural Network)
AdaBoost	Uyarlamalı Artırma (Adaptive Boosting)
AI	Yapay Zekâ (Artificial Intelligence)
AIS	Otomatik Tanımlama Sistemi (Automatic Identification System)
ANN	Yapay Sinir Ağı (Artificial Neural Network)
AR	Otoregresif (Autoregressive)
ARIMA	Otoregresif Bütünleşik Hareketli Ortalama (Autoregressive Integrated Moving Average)
ARIMA-BP	Otoregresif Bütünleşik Hareketli Ortalama Geri Yayılım (Autoregressive Integrated Moving Average Backpropagation)
BiLSTM	Çift Yönlü Uzun Kısa Süreli Bellek (Bidirectional Long Short Term Memory)
BN	Bayesyen ağ (Bayesian network)
BP	Breusch–Pagan (Breusch–Pagan)
BPNN	Geri Yayılım Sinir Ağı (Backpropagation Neural Network)
CatBoost	Kategorik Artırma (Categorical Boosting)
CD	Korentropik Ayırt Edici (Correntropy Discriminative)
CD–LSSVR	Korentropik Ayırt Edici En Küçük Kareler Destek Vektör Regresyonu (Correntropy Discriminative Least Squares Support Vector Regression)
CSAPSO	Kaos Senkronizasyon Algoritması-Parçacık Sürü Optimizasyonu (Chaos Synchronization Algorithm-Particle Swarm Optimization)
CSV	Virgülle Ayrılmış Değerler (Comma-separated Values)
DNN	Derin Sinir Ağı (Deep Neural Network)
DT	Karar Ağacı (Decision Tree)
DWT	Ölü Ağırlık Tonajı (Deadweight Tonnage)
DWT-LSTM	Ayrık Dalgacık Dönüşümü Uzun Kısa Süreli Bellek (Discrete Wavelet Transform Long Short-Term Memory)
EEMD–BELM–ECS	Topluluk Ampirik Mod Ayrıştırması-Çift Yönlü Aşırı Öğrenme Makinesi-Ampirik Mod Ayrıştırması (Ensemble Empirical Mode Decomposition-Bidirectional Extreme Learning Machine-Empirical Mode Decomposition)
EFB	Özel Özellik Paketleme (Exclusive Feature Bundling)
ELM	Aşırı Öğrenme Makinesi (Extreme Learning Machine)
ETS	Üstel Düzgünleştirme Durum Uzayı (Exponential Smoothing State Space)
EUR	Euro (Euro)
Extra Tree	Aşırı Rastgele Ağaçlar (Extremely Randomized Trees)
FEU	Kırk Feet Eşdeğer birim (Forty Feet Equivalent unit)
GB	Gradyan Artırma (Gradient Boosting)
GBDT	Gradyan Artırma Karar Ağacı (Gradient Boosting Decision Tree)

GM (1,1)	Gri Model (Grey Model)
GNB	Gaussian Naif Bayes (Gaussian Naïve Bayes)
GOSS	Gradyan-tabanlı Tek Taraflı Örnekleme (Gradient-based One-Side Sampling)
GRA-BiLSTM	Gri İlişki Analizi ve Çift Yönlü Uzun Kısa Süreli Bellek (Grey Relation Analysis and Bi-Directional Long Short Term Memory)
GRA-GRU	Gri İlişki Analizi-Kapılı Tekrarlayan Birim (Grey Relational Analysis-Gated Recurrent Unit)
GRA-LSTM	Gri İlişki Analizi Uzun Kısa Süreli Bellek (Grey Relational Analysis Long Short Term Memory)
GRA-REG	Gri İlişki Analizi Regresyonu (Grey Relational Analysis Regression)
GRA-SVR	Gri İlişki Analizi Destek Vektör Regresyonu (Grey Relational Analysis Support Vector Regression)
Grey-CNN	Gri Evrişimli Sinir Ağı (Grey Convolutional Neural Network)
GRU	Kapılı Tekrarlayan Birim (Gated Recurrent Unit)
Huber-R	Huber Regresyonu (Huber Regression)
IMO	Uluslararası Denizcilik Örgütü (International Maritime Organization)
ISO	Uluslararası Standardizasyon Örgütü (International Organization for Standardization)
Kn	(Knot) Denizcilikte ve havacılıkta hız birimi olarak kullanılır ve saatte deniz mili
KNN	K-En Yakın Komşu (K-Nearest Neighbors)
LASSO-R	En Küçük Mutlak Büzülme ve Seçim Operatörü Regresyonu (Least Absolute Shrinkage and Selection Operator Regression)
LCV	Olası Sınıflandırma Oylaması (Likely Classification Voting)
LightGBM	Hafif Gradyan-Artırma Makinesi (Light Gradient-Boosting Machine)
LR	Doğrusal Regresyon (Linear Regression)
LSSVR	En Küçük Kareler Destek Vektör Regresyonu (Least Squares Support Vector Regression)
LSTM	Uzun Kısa Süreli Bellek (Long Short Term Memory)
MA	Hareketli Ortalama (Moving Average)
MAE	Ortalama Mutlak Hata (Mean Absolute Error)
MAPE	Ortalama Mutlak Yüzde Hatası (Mean Absolute Percentage Error)
MARS	Çok Değişkenli Uyarlamalı Regresyon Splineları (Multivariate Adaptive Regression Splines)
MBPNN	Değiştirilmiş Geri Yayılım Sinir Ağı (Modified Backpropagation Neural Network)
MCP-GDP	Gayri Safi Yurtiçi Hasılanın Aylık Cari Fiyatları (Monthly Current Prices of Gross Domestic Products)
m-GRU	Çok Değişkenli Kapılı Tekrarlayan Birim (Multivariate Gated Recurrent Unit)
ML	Makine Öğrenmesi (Machine Learning)

MLP	Çok Katmanlı Algılayıcı (Multilayer Perceptron)
MLR	Çoklu Doğrusal Regresyon (Multiple Linear Regression)
m-LSTM	Çok Değişkenli Uzun Kısa Süreli Bellek (Multivariate Long Short Term Memory)
MMSI	Denizcilik Mobil Hizmet Kimliği (Maritime Mobile Service Identity)
MRSVR-CSAPSO	Değiştirilmiş Tekrarlayan SVR-Kaos Senkronizasyon Algoritması-Parçacık Sürü Optimizasyonu (Modified Recurrent SVR-Chaos Synchronization Algorithm-Particle Swarm Optimization)
MSE	Ortalama Kare Hatası (Mean Squared Error)
MVFP	Aylık Gemi Yakıt Fiyatları (Monthly Vessels Fuel Prices)
NB	Naif Bayes (Naïve Bayes)
PSO	Parçacık Sürü Optimizasyonu (Particle Swarm Optimization)
R^2	Korelasyonun Karesi (Square of Correlation)
ReLU	Düzeltilmiş Doğrusal Birim (Rectified Linear Unit)
RF	Rastgele Orman (Random Forest)
RMSE	Kök Ortalama Kare Hatası (Root Mean Squared Error)
RNN	Tekrarlayan Sinir Ağı (Recurrent Neural Network)
RR	Ridge Regresyonu (Ridge Regression)
RSVR	Tekrarlayan Destek Vektör Regresyonu (Recurrent Support Vector Regression)
SARIMA	Mevsimsel Otoregresif Bütünleşik Hareketli Ortalama (Seasonal Autoregressive Integrated Moving Average)
SARIMA–LSSVR	Mevsimsel Otoregresif Bütünleşik Hareketli Ortalama En Küçük Kareler Destek Vektör Regresyonu (Seasonal Autoregressive Integrated Moving Average Least Squares Support Vector Regression)
SARIMAX	Açıklayıcı Değişkenli Mevsimsel Otoregresif Bütünleşik Hareketli Ortalama (Seasonal Autoregressive Integrated Moving Average with Explanatory Variable)
SD–LSSVR	Tekil Değer Ayrıştırması En Küçük Kareler Destek Vektör Regresyonu (Singular Value Decomposition Least Squares Support Vector Regression)
SES	Basit Üstel Düzgünleştirme (Simple Exponential Smoothing)
Std	Standart Sapma (Standard Deviation)
SVR	Destek Vektör Regresyonu (Support Vector Regression)
TBATS	Trigonometrik Box-Cox Otoregresif Hareketli Ortalama Trend Mevsimsel (Trigonometric Box-Cox Autoregressive Moving Average Trend Seasonal)
TE	Toplam İhracat (Total Export)
TEU	Yirmi Feet Eşdeğer birim (Twenty Feet Equivalent unit)
TI	Toplam İthalat (Total Import)
TRY	Türk lirası (Turkish lira)
u-CNN	Tek Değişkenli Evrişimli Sinir Ağı

	(Univariate Convolutional Neural Network)
u-GRU	Tek Değişkenli Kapılı Tekrarlayan Birim (Univariate Gated Recurrent Unit)
UI	Kullanıcı Arayüzü (User Interface)
u-LSTM	Tek Değişkenli Uzun Kısa Süreli Bellek (Univariate Long Short Term Memory)
USD	Amerika Birleşik Devletleri doları (United States dollar)
UTC	Evrensel Zaman Koordinatı (Universal Time Coordinate)
u-Transformer	Tek Değişkenli Dönüştürücü (Univariate Transformer)
VMD-BELM	Varyasyonel Mod Ayrıştırması ve Kelebek Aşırı Öğrenme Makinesi (Variational Mode Decomposition and Butterfly Extreme Learning Machine)
VMD–BELM	Varyasyonel Mod Ayrıştırması-Çift Yönlü Aşırı Öğrenme Makinesi (Variational Mode Decomposition-Bidirectional Extreme Learning Machine)
VMD–BELM–ECS	Varyasyonel Mod Ayrıştırması-Çift Yönlü Aşırı Öğrenme Makinesi-Ampirik Mod Ayrıştırması (Variational Mode Decomposition-Bidirectional Extreme Learning Machine-Empirical Mode Decomposition)
VTS	Gemi Trafik Hizmeti (Vessel Traffic Service)
XGBoost	Aşırı Gradyan Artırma (Extreme Gradient Boosting)

ŞEKİLLER DİZİNİ

		Sayfa
Şekil 3.1.	Tek yapay nöron işlevi (Cabello, 2022).	22
Şekil 3.2.	A) Kendine geri besleme döngüleri ve gizli nöronları olmayan yinelemeli ağ B) Gizli nöronlara sahip yinelemeli ağ (Haykin, 2009)	23
Şekil 3.3.	CNN-LSTM hibrit modeli akış şeması	27
Şekil 3.4.	SVR-ANN hibrit modeli akış şeması	29
Şekil 3.5.	Türkiye'de toplam konteyner iş hacmi (TEU)	32
Şekil 3.6.	Konteyner iş hacminin talebi tahmin yaklaşımı akış şeması	35
Şekil 3.7.	Teslim süresi tahmin yaklaşımı akış şeması	39
Şekil 3.8.	Konteyner iş hacmi ve teslim süresi tahmini için önerilen yaklaşıma ait akış şeması	46
Şekil 4.1.	Konteyner iş hacmi tahmini veri seti için gri ilişki analizi	49
Şekil 4.2.	Senaryo 1 ve senaryo 2 için eğitim veri setine ait performans parametrelerinin sonuçları	52
Şekil 4.3.	Model tahminlerinin tarihsel görülmemiş veri seti ile karşılaştırılması	56
Şekil 4.4.	2025 yılı için en iyi performans gösteren modellerin tahmin sonuçları	57
Şekil 4.5.	Artık-Kalıntılar (Residuals) ve tahmin edilenler değişen varyans tespit testi	61
Şekil 4.6.	Teslim süresi tahmini veri seti için gri ilişki analizi	64
Şekil 4.7.	Modellerin daha önce görmediği düşük akış teslim süresi veri seti için gerçek tarihsel veriler ile modellerin tahmin sonuçlarının karşılaştırılması	70
Şekil 4.8.	Modellerin daha önce görmediği yüksek akış teslim süresi veri seti için gerçek tarihsel veriler ile modellerin tahmin sonuçlarının karşılaştırılması	72

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 2.1. Konteyner iş hacmi tahminine ilişkin literatür araştırması 13
Çizelge 2.2. Farklı süre terimi tahminine ilişkin literatür araştırması 15
Çizelge 3.1. Performans değerlendirme metrikleri 30
Çizelge 3.2 Teslim süresi tahmin modeli girdi özelliklerinin tanımları..... 40
Çizelge 4.1. Veri setinde yer alan bağımlı ve bağımsız değişkenlerin istatistiksel analizi..... 48
Çizelge 4.2. XGBoost, LSTM ve ANN modellerinin hiper parametreleri 51
Çizelge 4.3. Modellerin Senaryo 1 için test veri seti performans karşılaştırması... 53
Çizelge 4.4. Modellerin Senaryo 2 için test veri seti performans karşılaştırmas… 54
Çizelge 4.5. Daha önce görülmemiş veri seti için en iyi performans gösteren modellerin tahmin sonuçları..... 56
Çizelge 4.6. Katmanlaşma öncesi ve sonrası değişen varyans testi sonuçları 62
Çizelge 4.7. Farklı Katmanlaştırma oranları için sonuçlar 63
Çizelge 4.8. Düşük akış veri seti için modellerin performans değerlendirme sonuçları 66
Çizelge 4.9. Yüksek akış veri seti için modellerin performans değerlendirme ölçüm sonuçları 68
Çizelge 4.10. En iyi performans gösteren modellerin daha önce görülmemiş düşük akış teslim süresi veri seti için tahmin sonuçları 70
Çizelge 4.11. En iyi performans gösteren modellerin daha önce görülmemiş yüksek akış teslim süresi veri seti için tahmin sonuçları 72

1. GİRİŞ

Deniz taşımacılığı uluslararası ticaretin temel direği olup, küresel ticaretin hacim bazında %80'inden fazlasını oluşturmaktadır. Küreselleşmenin hızlanması ve tedarik zincirlerinin kıtalara yayılmasıyla birlikte, denizcilik lojistiğinin stratejik önemi artmış, bu da akademik çevreler ile sektör tarafından bu alandaki dinamiklerin anlaşılması konusundaki ilgiyi arttırmıştır. Diğer taşımacılık sektörlerinden farklı olarak deniz taşımacılığı, ekonomik döngüler, jeopolitik değişimler, liman altyapısı sınırlamaları ve değişken talep kalıpları tarafından şekillenen değişken bir ortamda faaliyet göstermektedir.

Uluslararası ticaretin büyümesiyle birlikte, 1961 yılında Uluslararası Standardizasyon Örgütü (International Organization for Standardization: ISO), Yirmi Ayak Eşdeğer Birimi (Twenty Feet Equivalent unit: TEU) ve Kırk Ayak Eşdeğer Birimi (Forty Feet Equivalent unit: FEU) olarak bilinen standart konteyner boyutlarını tanıtmıştır. Bu standardizasyon, standart konteyner boyutlarına uygun taşıma sistemleri ve elleçleme ekipmanlarının oluşturulmasına yol açarak gemiler, kamyonlar ve trenler arasında sorunsuz geçişleri kolaylaştırmıştır. Konteynerlerin standardizasyonu, yük işlemlerini kolaylaştırarak ve başlangıç noktasından nihai varış noktasına kadar kesintisiz, çok modlu taşımayı mümkün kılarak küresel lojistiği dönüştürmüştür (Christiansen vd., 2007).

Türk Deniz Ticaret Odası'nın 2024 raporuna göre, Türk armatörler toplamda 47,9 milyon net ağırlık (DWT, Deadweight Tonnage) büyüklüğünde bir filoyu kontrol etmektedir. Türk Bayraklı Filo 6,0 milyon DWT, Yabancı Bayraklı Filo ise 41,9 milyon DWT'dir. Türk Ticaret Filosu, dünya sıralamasında 12. sıradadır. Türkiye'nin dış ticaretinin %87,5'i deniz taşımacılığı ile gerçekleştirilmektedir (Maritime Sector Report, 2024).

Türk Deniz Ticaret Odası, Türkiye'nin deniz taşımacılığıyla yapılan konteyner iş hacminin 2014'te 8,3 milyon TEU'dan 2023'te 12,5 milyon TEU'ya yükseldiğini rapor etmiştir. Aynı dönemde, ithalat konteyner iş hacmi 3,4 milyon TEU'dan 4,9 milyon TEU'ya, ihracat konteyner iş hacmi ise 3,5 milyon TEU'dan 4,8 milyon TEU'ya artmış ve bu durum ülkenin deniz ticaretindeki sürdürülebilir büyümeyi göstermiştir. 2023 yılında dünya genelindeki konteyner hacmi sıralamasına göre Türkiye, %2,1 yıllık TEU

değişimi ile dünya genelinde 18. sırada yer almıştır (Maritime Sector Report, 2024). Türkiye'de deniz ticaretinin önemi göz önüne alındığında, bu tez çalışmasının ilk aşamasında 2019'dan 2024'e kadar konteyner iş hacmi analizi ve tahmini üzerine odaklanılmıştır (Maritime Sector Report, 2024).

Talep değişkenliğinin lojistik üzerindeki en önemli sonuçlarından biri, teslim süresine olan etkisidir. Teslim süresi, bir siparişin verilmesi ile teslim alınması arasında geçen toplam süre olarak tanımlanmaktadır. Bu süre, sipariş iletimi, sipariş işleme, sipariş hazırlığı ve taşıma için gereken süreyi veya malların başlangıç noktasından varış noktasına kadar taşınması için geçen toplam süreyi içermektedir. Artan talep, liman altyapısını, taşıma kapasitesini ve rota kullanılabilirliğini zorlayarak daha uzun teslim sürelerine ve hizmet güvenilirliğinin azalmasına yol açabilmektedir. Bu ilişkinin anlaşılması, tedarik zinciri performansını optimize etmek ve taşımacılık ihtiyaçlarını doğru bir şekilde tahmin etmek için önemlidir. Bu ilişki, tez çalışmasının önemini ve orijinalliğini vurgulamaktadır. Tez çalışmasının ikinci aşamasında teslim süresi tahmini üzerine odaklanılmıştır.

Talep dalgalanmaları, liman tıkanıklığı, gemi programlama verimsizlikleri ve dışsal aksaklıklar, öngörülemeyen teslim sürelerine neden olarak tedarik zinciri risklerini arttırmaktadır. İstatistiksel zaman serisi modelleri ve ekonometrik yaklaşımlar gibi geleneksel tahmin yöntemleri, genellikle taşımacılık verilerindeki doğrusal olmayan, yüksek boyutlu desenleri yakalamakta zorlanmaktadır. Bu sınırlama, yanlış talep tahminlerine, alt optimal filo dağılımına ve verimsiz liman operasyonlarına yol açarak lojistik maliyetlerini arttırmakta ve hizmet güvenilirliğini azaltmaktadır.

Bu tez çalışması, deniz taşımacılığı talebi ve teslim süresi tahminlerini iyileştirmek için makine öğrenimi (ML, Machine Learning) tekniklerinden yararlanmayı amaçlamaktadır. Çalışma, Türk deniz taşımacılığı operasyonlarında konteyner iş hacmi ve teslim süreleri için ML tabanlı tahmin modeli geliştirmeyi, farklı ML algoritmalarının performansını karşılaştırmayı ve talep ve teslim sürelerini etkileyen temel değişkenleri belirlemeyi hedeflemektedir. Bu araştırma, ileri düzey ML tekniklerini entegre ederek deniz lojistiği verimliliğini artırmak için veri odaklı bir yaklaşım sunmaktadır. Doğru talep ve teslim

süresi tahminleri, nakliye şirketlerinin filo dağılımını optimize etmelerine, liman otoritelerinin altyapı planlamalarını iyileştirmelerine ve tedarik zinciri yöneticilerinin gecikmeleri azaltmalarına katkı sağlayacaktır. Bu çalışma, ekonometrik taşımacılık modelleri ile modern tahmin analitiklerini birleştirerek mevcut literatüre katkıda bulunmakta ve gerçek dünya deniz lojistiği optimizasyonu için ölçeklenebilir bir çerçeve sunmaktadır.

2. KAYNAK ARAŞTIRMASI

Bu bölümün ilk kısmı, ML ve denizcilik sektöründeki uygulamalarını tanıtacaktır; ikinci kısmı ise konteyner iş hacmi tahmini ve talep segmentini etkileyen faktörleri araştırarak denizcilik talep tahmini üzerine mevcut araştırmaları kapsayacaktır. Bölümün son kısmında, deniz taşımacılığında teslim süresi terimlerinin tahminine odaklanılacak ve teslim süresi tahmini konusu ile ilgili literatürdeki mevcut çalışmalar özetlenecektir. Literatürdeki önceki çalışmalar, Türkiye'deki denizcilik sektörü için talep ve teslim süresinin tahmini çalışmasının temelini oluşturmakla birlikte literatürde, özellikle Türkiye'de, talep ve teslim süresi tahminini eş zamanlı olarak ele alan bir araştırmaya rastlanmamıştır. Bu da tez çalışmasının literatüre olan önemini ve katkısını artırmaktadır.

2.1. Makine Öğrenimi

ML, büyük veri kümelerindeki değişkenler arasındaki desenleri ve ilişkileri otomatik olarak algılamaya ve çıkarmaya odaklanan yapay zekâ (AI: Artificial Intelligence) alt alanıdır (Biggio ve Roli, 2018). ML teknikleri, üretim, operasyonlar, sağlık hizmetleri ve konut gibi çeşitli alanlarda uygulanabilmektedir. Ayrıca, ML tedarik zinciri yönetiminin farklı bileşenlerini ve işlevlerini denetlemek ve optimize etmek için yaygın bir şekilde kullanılmaktadır. Ni vd. (2020) çalışmasında, ML tekniklerini kullanan en önemli endüstrilerin sırasıyla market ve gıda endüstrileri olduğunu, ardından otomotiv ve moda endüstrilerinin geldiğini ortaya koymuştur. Ni vd. (2020) çalışmasında ML algoritmaları ile tedarik zinciri yönetimi faaliyetleri arasındaki güçlü bağlantıyı gösteren en önemli alanları listelemiş ve bu faaliyetlerin; talep/satış tahmini, satın alma ve tedarik yönetimi, üretim, envanter ve depolama, taşımacılık ve dağıtım ile tedarik zinciri iyileştirme olduğunu belirtmiştir.

Tirkolaee vd. (2021) çalışmasında kavramsal bir çerçeve geliştirilmiştir. Bu çalışma ML tekniklerinin tedarikçi seçimi ve segmentasyonu, tedarik zinciri risklerinin tahmini, talep ve satış tahmini, üretim, envanter yönetimi, taşımacılık ve dağıtım, sürdürülebilir kalkınma ve döngüsel ekonomi alanlarındaki katkılarını tanımlamaktadır. Son olarak,

çalışmanın temel sınırlılıklar ve zorluklar üzerindeki etkileri tartışılmakta, ardından yönetimsel içgörüler ve gelecekteki araştırma yönleri sunulmaktadır.

Younis vd. (2022) çalışmasında, ML araştırmalarının çoğunun özellikle talep tahmini, olası geri sipariş tahmini ve talep değişkenliğinin derin öğrenme yönlerine odaklandığını belirtmişlerdir. Bu uygulamalar, ML'nin tedarik zinciri görünürlüğünü, verimliliğini ve duyarlılığını artırmadaki dönüştürücü potansiyelini kolektif olarak göstermektedir. ML modellerinin uygulaması denizcilik sektöründe yaygın olarak kullanılmaktadır. Shin ve Yang (2025) çalışmalarında, Busan Limanı'ndan Kasım 2018- Temmuz 2024 dönemine ait veriler ile hem çevresel hem de Gemi Trafik Hizmetleri (VTS Vessel Traffic Service) özelliklerini kullanarak deniz kazalarını tahmin etmek için ML modelleri geliştirmiş ve değerlendirmiştir. Fan vd. (2024) çalışmalarında, gemi yakıt tüketimi tahmin modellerini oluşturmak için altı ML yöntemi kullanmış ve gemi enerji tüketimini tahmin etmek için bir çalışma geliştirmiştir.

Liu vd. (2021) çalışmasında, deniz kazalarının nedenlerini belirlemek ve denizcilik otoritelerine ve denizcilere kapsamlı öneriler sunmak için veri odaklı bir Bayes Ağı (BN: Bayesian network) modeli oluşturulmuştur. Bu model, Çin'in kıyı sularındaki önemli kaza kayıtlarını ileri düzey bir ML tekniği ile analiz etmektedir. Ayrıca, Li vd. (2024) tarafından konteyner nakliye ağına nedensellik analizi uygulanarak konteyner gemisi gecikmelerinin yayılma riskleri araştırılmıştır.

Bu geniş uygulamalara rağmen, deniz taşımacılığı sektörü, ML araştırmalarında nispeten az keşfedilmiş benzersiz zorluklar ve fırsatlar sunmaktadır. Çok sayıda çalışma, ML uygulamalarını genel tedarik zinciri bağlamlarında incelemiş olsa da birkaç çalışma denizcilik sektörünün küresel ölçek, yasal mevzuat karmaşıklığı ve dış şoklara karşı kırılganlık gibi kendine özgü özelliklerini ele almıştır. Mevcut denizcilik odaklı ML araştırmaları, talep tahmini veya teslim süresi tahminini izole bir şekilde incelemekte olup, operasyonel karar alma süreçlerini önemli ölçüde etkileyen bu faktörler arasındaki kritik etkileşimi kaçırmaktadır. Ayrıca, denizcilik bağlamlarında ML yaklaşımlarının geleneksel tahmin yöntemleriyle karşılaştırıldığı araştırmalar sınırlıdır; özellikle nakliye

şirketlerinden ve limanlardan alınan gerçek operasyonel veriler kullanılarak yapılan çalışmalar eksiktir.

Bu çalışma, deniz taşımacılığında talep ve teslim süresi tahminini eş zamanlı olarak gerçekleştirmek için entegre bir ML tabanlı tahmin sistemi geliştirerek bu boşlukları doldurmayı amaçlamakta ve bölgesel olarak ilgili öngörülere katkıda bulunmak için Türk ticaret verilerine özel olarak odaklanmaktadır. Böylece, ML'nin pazar dalgalanmaları ve aksaklıklar karşısında deniz lojistiği verimliliğini ve dayanıklılığını artırma potansiyelini genişletmeyi hedeflemektedir.

2.2. Talep Tahmini

Denizcilik sektöründe talep tahmini konusu, doğru tahminlerin değişken ticaret modelleri, mevsimsel dalgalanmalar ve öngörülemeyen aksaklıkları hesaba katması gerektiği için endüstrinin karmaşık zorluklarını ele almak üzere önemli ölçüde gelişmiştir. Geleneksel istatistiksel yöntemler temel oluşturmakla birlikte, konteyner iş hacminin doğrusal olmayan dinamiklerini yakalamada sınırlamalar göstermiş ve araştırmacıları daha sofistike ML yaklaşımlarını keşfetmeye yönlendirmiştir. Bu değişim, hem operasyonel verilerin giderek artan erişilebilirliğini hem de liman yönetimi ve lojistik planlamada gerçek zamanlı karar alma süreçlerini destekleyebilecek daha uyarlanabilir tahmin araçlarına duyulan ihtiyacı yansıtmaktadır. Son yıllarda yapılan çalışmalar, çeşitli metodolojik yenilikler aracılığıyla bu ilerlemeyi göstermekte ve araştırmacılar, tahmin doğruluğunun arttırmak için giderek daha karmaşık algoritmaları çeşitli denizcilik veri setleri üzerinde test etmektedir.

Literatürdeki çeşitli çalışmalar, denizcilik talep tahmininde ML'nin etkinliğini göstermiştir. Örneğin, Moscoso-López vd. (2021) Algeciras Körfezi Limanı'ndaki bozulabilir kargo akışı tahmini için tatil ve mevsimsel göstergeler gibi zamansal özellikleri kargo ağırlık verileriyle birleştirerek yenilikçi bir ML sistemi geliştirmiştir. Jin vd. (2023) Ningbo Zhoushan Limanı'ndaki günlük konteyner iş hacmi akışı tahminini liman kapı operasyonları verilerini analiz ederek gerçekleştirmiştir.

Nieto vd. (2017) ise San Pedro limanlarında çoklu kargo hacmi tahmini için zaman serisi ve ML modellerinin kapsamlı bir karşılaştırmasını yapmıştır. Ayrıca, Ubaid vd. (2021) Avustralya ithalat talebi tahmini için Mevsimsel Otoregresif Entegre Hareketli Ortalama (SARIMA: Seasonal Autoregressive Integrated Moving Average), Holt-Winters ve Prophet modellerinin performansını değerlendirerek, özellikle kısa ve uzun vadeli tahmin yaklaşımları açısından metodolojik tartışmayı zenginleştirmiştir.

İncelenen çalışmalar, denizcilik sektörü talep tahmini araştırma alanının basit tek değişkenli modellerden hem zamansal desenleri hem de dış ekonomik göstergeleri entegre eden karmaşık çok değişkenli tahmin modellerine doğru ilerlediğini kolektif olarak göstermektedir.

Konteyner iş hacmi tahmini, literatürde karşılaşılan son çalışmalarda sunulan çeşitli modelleme yaklaşımları incelendiğinde karmaşıklığını açıkça ortaya koymaktadır ve Çizelge 2.1'de özetlenmiştir. Gelişmiş hibrit modeller özellikle etkili olarak ortaya çıkmış Zeng ve Xu (2024) çalışmalarında Gri Evrişimsel Sinir Ağı (Grey-CNN: Grey Convolutional Neural Network) kombinasyonlarının üstünlüğünü göstermiştir.

Çizelge 2.1'de sunulan literatür araştırması özeti, önemli bir eğilimi ortaya koymaktadır; geleneksel Otoregresif Entegre Hareketli Ortalama (Autoregressive Integrated Moving Average: ARIMA) varyantları bazı çalışmalarda rekabetçi kalırken (Yao, 2021; Rashed vd. 2017), en iyi performans gösteren modellerin çoğu ya Uzun Kısa Süreli Bellek (LSTM: Long Short Term Memory) ve Kapılı Yinelemeli Birim (GRU: Gated Recurrent Unit) gibi derin öğrenme unsurlarını ya da Varyasyonel Mod Ayrıştırma ve Kelebek Aşırı Öğrenme Makinesi (VMD-BELM: Variational Mode Decomposition and Butterfly Extreme Learning Machine) gibi karmaşık hibrit yaklaşımları içermektedir. Bu eğilim, denizcilik sektörü talep tahmini araştırma alanının hem zamansal dizileri hem de karmaşık özellik etkileşimlerini aynı anda işleyebilen mimarilere doğru ilerlediğini göstermektedir.

Çalışmalar arasında girdilerin özelliklerinin kritik bir şekilde incelenmesi, model geliştirme için önemli hususları ortaya koymaktadır. Ubaid vd. (2021) çalışmalarında

hem ithalat hem de ihracat için toplam dolu gelen ve giden konteynerlerin yanı sıra boş konteynerlerin sayısını içeren bir veri seti ele almıştır.

Ayrıca, Moscoso-López vd. (2021) çalışmalarında, Ocak 2010-Aralık 2017 dönemine ait taze meyve ve sebzelerin ağırlığını (kilogram cinsinden) temsil eden iki ana zaman serisinden oluşan veri setini kullanmıştır. Her bir ürün için dört türetilmiş versiyon bulunmaktadır: orijinal seri, üstel hareketli ortalama kullanılarak düzeltilmiş versiyon ve bu serilerin hem log-dönüştürülmüş hem de birinci dereceden fark alınmış versiyonları. Ek olarak, hafta sonları, resmî tatiller ve mevsimsellik göstergeleri gibi ek özellikler dahil edilmiştir. Bu işlenmiş özellikler, kargo akışını yedi gün önceden tahmin etmek için birden fazla tahmin modelini eğitmek ve test etmek amacıyla kullanılmıştır.

Tez çalışması konteyner iş hacmi talebini tahmin etmeye odaklandığı için literatürde konteyner iş hacmi talebini tahmin etmek için kullanılan girdi özelliklerine yoğunlaşılmıştır. Awah vd. (2021) çalışmalarında konteyner iş hacmini tahmin etmek için gemi dönüş süresi, konteyner bekleme süresi, ortalama gemi draftı, rıhtım verimliliği, konteyner depolama kapasitesi ve gümrük beyan süresi girdi değişkenlerini kullanmışlardır. Yanıt değişkeni olarak ise aylık konteyner iş hacmi tanımlanmıştır. Ayrıca, Lee ve Bang (2024) çalışmalarında Singapur Limanı'ndaki konteyner iş hacmini etkileyen faktörleri ham petrol fiyatları, Çin'in ihracat hacmi ve COVID-19 vaka sayısı olarak belirlemiştir.

Zeng ve Xu (2024) çalışmalarında ekonomik ölçütleri kapsamlı bir şekilde ele almış ve kısa vadeli liman hacmi tahminini iyileştirmeyi amaçlayan bir dizi girdi özelliği seçmiştir. Bunlar, birincil zaman serisi girdisi olarak geçmiş hacim değerleri ile birlikte makroekonomik göstergeler (örneğin, ithalat/ihracat endeksleri), limana özgü operasyonel göstergeler (gemi varış verileri ve rıhtım doluluk oranları gibi) ve zamansal özellikler (ay, çeyrek ve tatil göstergeleri gibi) gibi çeşitli dış açıklayıcı değişkenleri içermektedir.

Çalışma, makroekonomik göstergeler, liman operasyonel verileri ve zamansal işaretçileri birleştirerek kapsamlı bir özellik seti geliştirmiştir. Özellik mühendisliğindeki bu gelişim,

doğru talep tahmininin sadece geçmiş eğilimleri değil, aynı zamanda deniz taşımacılığı akışlarını şekillendiren daha geniş ekonomik ve operasyonel bağlamı da yakalaması gerektiği yönündeki artan farkındalığı yansıtmaktadır.

Bu ilerlemelere rağmen, deniz taşımacılığı talep tahmininde önemli zorluklar devam etmektedir. Farklı tahmin ufukları ve liman özellikleri için optimal model mimarileri konusunda ortak bir görüş geliştirilememiştir ve çalışmalar, farklı bölge ve zaman dönemleri arasında değişen performans sonuçlarını göstermektedir. Ayrıca, birçok çalışma tek liman tahminlerine odaklanırken, küresel tedarik zincirlerinde birbirine bağlı limanlar arasındaki ağ etkilerini ele alan çok az çalışma bulunmaktadır. Çalışmalar arasında özellik setlerindeki değişkenlik, değişken seçimi için daha sistematik yaklaşımlara ihtiyaç olduğunu da göstermektedir. Bu boşluklar, özellikle bölgesel özellikleri hesaba katarken farklı deniz taşımacılığı bağlamlarında sağlam performansı koruyabilecek uyarlanabilir tahmin yapıları geliştirme konusunda gelecekteki araştırmalar için önemli fırsatlara işaret etmektedir. Tez çalışması, Türkiye'nin küresel ticaret yollarındaki stratejik konumuna odaklanarak hem geleneksel hem de yeni girdi özelliklerinin bu önemli gelişen pazardaki tahmin doğruluğunu nasıl etkilediğini inceleyerek bu ilerlemeye katkıda bulunmaktadır.

2.3. Teslim Süresi Tahmini

Denizcilik lojistiğinde talep tahmini geniş çapta incelenmiş olsa da teslim sürelerinin doğru bir şekilde öngörülmesi, tedarik zinciri güvenilirliği ve operasyonel verimlilik açısından doğrudan etkili, aynı derecede önemli ancak daha karmaşık bir zorluk olarak öne çıkmaktadır. Teslim süresi tahmini, liman yoğunluğu, gemi özellikleri ve hava koşulları gibi çeşitli değişkenleri içeren ve her biri potansiyel gecikmeler ve belirsizlikler yaratan, gemi kalkışından varışına kadar geçen toplam süreyi tahmin eder. Talep tahmininin esas olarak hacim projeksiyonlarıyla ilgilenmesinin aksine, teslim süresi tahmini, sevkiyat sürelerinde doğrusal olmayan, genellikle öngörülemez değişiklikler yaratan dinamik operasyonel faktörler ve dış aksaklıkları hesaba katmalıdır.

Bu bölüm, denizcilik lojistiğinde zaman tahmini ve teslim süresi üzerine geliştirilen metodolojileri incelemekte, tahmin doğruluğunu artırmak amacıyla ML ve hibrit yaklaşımlar kullanan önemli çalışmaları vurgulamaktadır. Bu içgörüler birleştirilerek, teslim süresi tahmininin talep tahminiyle nasıl entegre edilebileceği ve denizcilik lojistiği optimizasyonuna yönelik daha bütüncül bir yaklaşımın nasıl geliştirilebileceği üzerine bir temel oluşturmaktadır.

Son yıllarda yapılan çalışmalar, varış zamanı tahmini için statik gemi özelliklerini dinamik operasyonel verilerle birleştirmenin etkinliğini göstermiştir. Örneğin, Arbabkhah vd. (2024), gemi tipi, ölü ağırlık gibi statik özellikleri hız, konum gibi dinamik otomatik tanımlama sistemi (AIS: Automatic Identification System) verileri ile entegre ederken, Evmides vd. (2024) tahminleri tahmini varış zamanı gibi gerçek zamanlı seyir ölçümlerini kullanarak geliştirmiştir. Benzer şekilde, Kolley vd. (2023) ve El Mekkaoui vd. (2022), gerçek zamanlı seyir ilerlemesini modellemek için liman çiftleri, varış noktasına mesafe gibi bağlamsal özelliklerin önemini vurgulamıştır. Bu çalışmalar, gemiye özgü özellikleri dinamik navigasyon verileriyle harmanlayan hibrit yaklaşımların, yapısal ve operasyonel değişkenliği yakalayarak daha yüksek doğruluk sağladığını ortaya koymaktadır.

Literatürde karşılaşılan çalışmalar, doğru varış zamanı tahmini için statik gemi özellikleri ve dinamik navigasyon verilerinin entegrasyonunun önemini vurgularken, farklı araştırmacılar arasında özellik seçimi ve hesaplama metodolojilerindeki değişimi göstermektedir.

Literatürdeki önemli katkılara rağmen, teslim süresi tahmini araştırmalarında dikkate değer boşluklar devam etmektedir. İlk olarak, birçok çalışma, ara rota geçişleri sırasında (örneğin, demirleme beklemeleri ve ara duraklar) geçen hareketsiz süreleri içeren kapsamlı teslim süresi yerine varış zamanına odaklanmaktadır. İkinci olarak, hava koşulları, yük durumu (dolu/boş) ve gerçek zamanlı liman yoğunluğu gibi kritik değişkenler genellikle ihmal edilmekte, bu da modelin tahmin gücünü sınırlamaktadır.

Denizcilik taşımacılığı operasyonlarında teslim süresi, bir geminin kalkış limanından hareket ettiği andan varış limanına ulaştığı ana kadar geçen toplam süreyi ifade eder ve aktif seyir ile hareketsiz süreleri içerir. Teslim süresi, aktif seyir, demirleme yerlerinde bekleme veya yolculuğun alt rotaları boyunca ara duraklar gibi tüm geçen süreleri içerir. Teslim süresini doğru bir şekilde tahmin etmek, rıhtım planlaması, envanter yönetimi ve tedarik zinciri güvenilirliği gibi operasyonel planlamalar için giderek daha önemli hale gelmiştir. Farklı çalışmalarda teslim süresi, sipariş verme süresi ve liman süresini de içerecek şekilde daha geniş tutulsa da tez çalışmasında tahmin sürecini daha belirli hale getirmek ve tahmin doğruluğunu artırmak için teslim süresine odaklanılarak sipariş verme süresi ve liman süresi dahil edilmemiştir. Ayrıca, seyahat edilen mesafe arttıkça, yüksek belirsizlik nedeniyle teslim süresini tahmin etme karmaşıklığı da artmaktadır.

Literatürde, ML modelleri kullanılarak yapılan teslim süresi tahmin çalışmaları Çizelge 2.2'de verilmektedir ve bu bölümde özetlenmiştir. Hathikal vd. (2020) tarafından yapılan çalışmada, tahmin edilen teslim süresi, bir rezervasyonun yapılmasından kargonun varış limanında teslim edilmesine kadar geçen toplam süreyi kapsamaktadır. Bu tanım, rezervasyon onayı, konteyner alma, yükleme, gümrükleme, liman kapısından giriş, gemi kalkışı, transit, varış limanında boşaltma ve son gümrükleme ve konteyner teslimi gibi birden fazla aşamayı içermektedir. Önerilen model, operasyonel ve gerçek geçmiş perspektiflerden elde edilen verileri birleştirerek, denizcilik lojistik zincirlerinde toplam teslim süresine ilişkin gerçekçi bir tahmin sağlamayı amaçlamaktadır.

Botnen vd. (2021) çalışmalarında denizcilik lojistiğinde teslim süresini tahmin etmeye odaklanarak, teslim süresini, rezervasyon oluşturulmasından konteynerin varış limanında boşaltılmasına kadar geçen süre olarak tanımlamaktadır. Buna göre teslim süresi rezervasyon, yükleme, gümrükleme, liman kapısından giriş, gemi kalkışı, okyanus geçişi ve boşaltma gibi aşamaları içermektedir. Bu kapsamlı özellik seti, gerçek dünya lojistik süreçleriyle uyumlu doğru ve pratik teslim süresi tahmini yapılmasını sağlamaktadır.

Öte yandan, Lilit (2023) tarafından hazırlanan tez çalışmasında, deniz taşımacılığında teslim süresi tahmini için liman varış noktası, gönderi ağırlığı, hacmi, miktarı ve geçmiş verilerden türetilen tahmini teslim süresi gibi özellikler kullanılarak öngörü modelleri

geliştirilmiştir. Çalışmada, Rastgele Orman (RF: Random Forest) ve Aşırı Gradyan Güçlendirme (XGBoost: Extreme Gradient Boosting) gibi topluluk yöntemleri uygulanmış ve XGBoost yönteminin özellikler arasındaki doğrusal olmayan ilişkilerden yararlanarak başarılı bir performans gösterdiği belirtilmiştir. Özellikle, liman ve gönderi özelindeki değişkenlerin dahil edilmesi, rota karmaşıklığı ve kargo özelliklerinden kaynaklanan değişkenliği ele almıştır.

Teslim süresi tahmini, denizcilik lojistiğinde operasyonel verimlilik ve müşteri memnuniyeti açısından kritik bir rol oynamaktadır. Bu tahmin, gemilerin yapı özelikleri, gemilerin hızı, hava koşulları, rota uzunluğu ve liman trafiği gibi dinamik faktörlere bağlı olarak bir geminin kaynaktan hedefe ulaşması için gereken toplam süreyi ifade eder. ML modelleri, gerçek geçmiş operasyon verilerinden yararlanarak bu tahmini gerçekleştirir.

Gerçekleştirilen kapsamlı literatür araştırması, denizcilik sektörü talebi ve teslim süresi tahminine yönelik ML uygulamalarındaki gelişmeleri ortaya koymakta ve hibrit modeller, özellik mühendisliği ve gerçek zamanlı operasyonel verilere yönelik artan ilgiye vurgu yapmaktadır. Ancak, mevcut çalışmalarda talep ve teslim süresi tahminini çoğunlukla ayrı problemler olarak ele alınmakta ve genellikle bu iki sürecin entegrasyonuna yeterince odaklanılmamaktadır. Ayrıca, özellikle gelişmekte olan denizcilik pazarlarında, rezervasyon, liman işlemleri ve deniz taşımacılığını içeren uçtan uca teslim süresi tanımlarını içeren çalışmalar oldukça sınırlıdır. Söz konusu boşlukları ele alan mevcut tez çalışması, stratejik öneme sahip ancak literatürde yeterince incelenmemiş bir bölge olan Türk denizcilik sektöründe, konteyner iş hacmi talebi ve teslim süresini eşzamanlı olarak tahmin eden bütünleşik bir ML tabanlı tahmin sistemi önermektedir. Önerilen entegre yaklaşım, tahmin doğruluğunu artırmakta ve belirsizlik altında denizcilik lojistik operasyonlarını optimize etmek için uygulanabilir içgörüler sunmaktadır.

Çizelge 2.1. Konteyner iş hacmi tahminine ilişkin literatür araştırması

Kaynak	Talep Terimi	Geliştirilen Modeller	En İyi Performans Gösteren Model
(Zeng ve Xu, 2024)	Konteyner İş Hacmi	NB, ARIMA, ARIMAX, u-LSTM, m-LSTM, u-GR U, m-GRU, u-CNN, u-Transformatör, Grey-CNN.	Grey-CNN
(Lee ve Bang, 2024)	Konteyner İş Hacmi	RFR, LR, SVR, SARIMA, Prophet, LSTM, HWM, LASSO-R, RR, XGBoost, ECM, SARIMAX, LightGBM	SARIMAX
(B. Liu vd., 2023)	Kargo İş Hacmi	LSTM, BiLSTM, GRA-SVR, GRA-REG, GRA-GRU, GRA-LSTM, GRA-BiLSTM	GRA-BiLSTM
(Xu vd., 2022)	Konteyner İş Hacmi	NB, MA, AR, ARIMA, MLP, GRU, LSTM, CNN, Transformatör	GRU
(Cuong vd., 2022)	Liman İş Hacmi	SARIMA, RNN, LSTM, GRU, DWT-LSTM	DWT-LSTM
(Awah vd., 2021)	Konteyner İş Hacmi	RFR, LR, SVR, SARIMA, ARIMA, DT, MLP, LSSVR	RF
(Tan vd., 2021)	Liman İş Hacmi	Gri Tahmin GM (1,1), Yankı durumu ağı	Yankı durumu ağı
(Ubaid vd., 2021)	İthalat ve İhracat	SARIMA, Prophet, Seasonal Holt–Winters	Prophet

Çizelge 2.1. Konteyner iş hacmi tahminine ilişkin literatür araştırması (Devamı)

Kaynak	Talep Terimi	Geliştirilen Modeller	En İyi Performans Gösteren Model
(Yang ve Chang, 2020)	Konteyner İş Hacmi	RFR, LR, AdaBoost, SVR, CNN-LSTM	CNN-LSTM
(Y. Zhang vd., 2020)	Konteyner İş Hacmi	ARIMA, BPNN, ARIMA-BP optimum ağırlıklar, ARIMA-BP	ARIMA-BP
(Shankar vd., 2020)	Konteyner İş Hacmi	LSTM, ARIMA, SES, HW, ETS, TBATS, NN, ARIMA + ANN	LSTM
(Du vd., 2019)	Konteyner İş Hacmi	ELM, LSSV, SARIMA, ARIMA, EEMD–BELM–ECS, VMD–BELM, VMD–BELM–ECS	VMD–BELM–ECS
(Chan vd., 2019)	Konteyner İş Hacmi	SVR, ANN, MA, MARS, GM (1, 1), ARIMA	SVR
(Rashed vd., 2017)	Konteyner İş Hacmi	ARIMA, ARIMAX	ARIMA
(Geng vd., 2015)	Liman İş Hacmi	ARIMA, RSVR, PSO, CSAPSO, MARS, MRSVR-CSAPSO	MRSVR-CSAPSO
(Xie vd., 2013)	Konteyner İş Hacmi	ARIMA, SARIMA, CD, BPNN, SVR, LSSV, SARIMA–LSSVR, SD–LSSVR, CD–LSSVR	CD–LSSVR

Naive Bayes(NB), Univariate- LSTM (u-LSTM), Multivariate-LSTM (m-LSTM), Univariate Gated Recurrent Unit (u-GRU), Multivariate Gated Recurrent Unit (m-GRU), Univariate Convolutional Neural Network (u-CNN), Univariate Transformer (u-Transformer), Least Absolute Shrinkage and Selection Operator Regression (LASSO-R), Ridge Regression (RR), Error Correction Model (ECM), SARIMA with Explanatory Variable (SARIMAX), Light Gradient Boosting Machine (LightGBM), Bidirectional LSTM (BiLSTM), Grey Relational Analysis-SVR (GRA-SVR), Grey Relational Analysis Regression (GRA-REG), Grey Relational Analysis-Gated Recurrent Unit (GRA-GRU), Grey Relational Analysis-LSTM (GRA-LSTM), Grey Relational Analysis-Bidirectional-LSTM (GRA-BiLSTM), Moving Average (MA), Autoregressive (AR), Multilayer Perceptron (MLP), Discrete Wavelet Transform-LSTM (DWT-LSTM), Decision Tree (DT), Least Squares SVR (LSSVR), Adaptive Boosting (AdaBoost), Backpropagation Neural Network (BPNN), ARIMA-Backpropagation Optimal Weights (ARIMA-BP), ARIMA -Backpropagation (ARIMA-BP), Simple Exponential Smoothing (SES), Exponential Smoothing State Space (ETS), Trigonometric Box-Cox ARMA Trend Seasonal (TBATS), Extreme Learning Machine (ELM), Ensemble Empirical Mode Decomposition-Bidirectional Extreme Learning Machine-Empirical Mode Decomposition (EEMD–BELM–ECS), Variational Mode Decomposition-Bidirectional Extreme Learning Machine (VMD–BELM), Variational Mode Decomposition-MELM-Empirical Mode Decomposition (VMD–BELM–ECS), Multivariate Adaptive Regression Splines (MARS), Grey Model GM (1,1), Modified BPNN (MBPNN), Recurrent SVR (RSVR), Particle Swarm Optimization (PSO), Chaos Synchronization Algorithm- PSO (CSAPSO), Modified Recurrent SVR-Chaos Synchronization Algorithm- PSO (MRSVR-CSAPSO), Correntropy Discriminative (CD), SARIMA-Least Squares SVR (SARIMA–LSSVR), Singular Value Decomposition-LSSVR (SD–LSSVR), Correntropy Discriminative- LSSVR (CD–LSSVR).

Çizelge 2.2. Farklı süre terimi tahminine ilişkin literatür araştırması

Kaynak	Süre Terimi	Geliştirilen Modeller	En İyi Performans Gösteren Model
(Choi ve Yoon, 2025)	Bekleme Süresi	RF, XGBoost, SVR, LSTM, LightGBM, GB, LR, Ridge-R,	XGBoost
(Evmides vd., 2024)	Varış Süresi	RF, XGBoost, DT, DNN, KNN, GNP	RF
(Dinh vd., 2024)	Dönüş Süresi	RF, XGBoost, DT	XGBoost
(Chua vd., 2024)	Varış Süresi	RF, XGBoost, ANN, LSTM, Stacking, LightGBM	Stacking
(Arbabkhah vd., 2024)	Varış Süresi	XGBoost	XGBoost
(Rao vd., 2024)	Kalış ve Gecikme Süresi	RF, XGBoost, DT, KNN, ANN, GB, LR, LightGBM, AdaBoost, Extra Tree, Bayesian Ridge, Ridge-R, Lasso-R, Huber-R	RF
(Lei vd., 2024)	Varış Süresi	RF, XGBoost, Stacking, LightGBM, ANN, LR	LightGBM- based stacking
(Yoo ve Kim, 2023)	Transit Süresi	XGBoost, MLP, LightGBM, GB, 1D CNN	1D CNN
(Lilit, 2023)	Teslim süresi	RF, XGBoost, DT, LCV,	XGBoost

Çizelge 2.2. Farklı süre terimi tahminine ilişkin literatür araştırması (Devamı)

Kaynak	Süre Terimi	Geliştirilen Modeller	En İyi Performans Gösteren Model
(Kolley vd., 2023)	Varış Süresi	DT, KNN, ANN, LR, Extra Tree	KNN
(Yoon vd., 2023)	Bekleme Süresi	XGBoost	XGBoost
(Valero vd., 2022)	Varış Süresi	RF, XGBoost, DT, LightGBM, LG, AdaBoost, Extra Tree	KNN
(El Mekkaoui vd., 2022)	Varış Süresi	RF, GNB, ANN, LR, Ridge-R	ANN
(Botnen vd., 2021)	Teslim süresi	MLR, LASSO, Bootstrap	LASSO
(Stepec vd., 2020)	Dönüş Süresi	SVR, LR, AdaBoost	CatBoost
(Hathikal vd., 2020)	Teslim süresi	DT, MLP, SVR	MLP
(Servos vd., 2020)	Seyahat süresi	SVR, XGBoost, Extra Tree	SVR
(Jonquais ve Krempl, 2019)	Gönderim Süresi	RF, ANN, LR	RF

Gradient Boosting (GB), Deep Neural Network (DNN), Huber Regression (Huber-R), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Extremely Randomized Trees (Extra Tree), One-Dimensional Convolutional Neural Network (1D CNN), Likely Classification Voting (LCV), Ridge Regression (Ridge-R).

3. MATERYAL VE YÖNTEM

Bu bölüm, tez çalışmasında benimsenen metodolojik çerçeveye genel bir bakış sunmaktadır.

3.1. Tahmin Modelleri

Tez çalışmasında tahmin problemi, sürekli hedef değişkenler için tahmin modellerini öğrenmek amacıyla geçmiş gözlemlerin kullanıldığı denetimli makine öğrenmesi paradigması aracılığıyla ele alınmaktadır. Çeşitli regresyon algoritmalarının karşılaştırmalı analizi, en doğru ve genelleştirilebilir modeli belirlemektedir. Metodoloji, pratik uygulanabilirliği sağlamak amacıyla model geliştirme, ayarlama ve değerlendirmeyi entegre etmektedir.

ML, sistemlerin verilerden öğrenmesini ve açıkça programlanmadan performansını iyileştirmesini sağlayan çeşitli metodolojileri kapsamaktadır. Verinin doğasına ve amaçlanan öğrenme hedefine bağlı olarak, ML algoritmaları genel olarak çeşitli paradigmalar altında sınıflandırılmaktadır (Nasteski, 2017) :

1. Denetimli öğrenme,
2. Denetimsiz öğrenme,
3. Yarı-denetimli öğrenme,
4. Pekiştirmeli öğrenme,
5. Transdüksiyon öğrenmesi,
6. Öğrenmeyi öğrenme.

Her paradigma bir model geliştirme ve veri yorumlama yaklaşımı kullanarak, farklı alanlarda benzersiz uygulamalar sunmaktadır (Nasteski, 2017).

Tez çalışması, bir algoritmanın tahmin veya sınıflandırma yapmak için açıklamalı eğitim verilerini kullanarak girdi özelliklerinden etiketli çıktılara eşleştirme fonksiyonunu öğrendiği temel bir ML paradigması olan denetimli öğrenmeyi benimsemektedir. Süreç tipik olarak iki ana aşamayı içermektedir: modelin etiketli örneklerden öğrendiği eğitim aşaması ve modelin genelleme yeteneğinin görülmemiş veriler üzerinde değerlendirildiği

test aşamasıdır. Denetimli öğrenme teknikleri esas olarak hedef çıktının kesikli olduğu sınıflandırma ve çıktının sürekli olduğu regresyon olarak kategorize edilmektedir. Standart algoritmalar DT, SVR, MLR, LR ve NB gibi olasılıksal modelleri içermektedir. Bu yöntemler, yüksek tahmin doğruluğu ve biyoinformatik, metin sınıflandırması ve görüntü tanıma gibi çeşitli alanlarda pratik uygulanabilirlik nedeniyle önemli bir popülarite kazanmıştır. Denetimli öğrenmenin etkinliği, gizli kalıpları çıkarmak ve güvenilir gelecek tahminleri yapmak için geçmiş verileri kullanma yeteneğinden kaynaklanmakta ve böylece karmaşık sistemlerde optimize edilmiş karar verme süreçlerini desteklemektedir (Nasteski, 2017).

3.1.1. Rastgele orman

RF, eğitim sırasında çoklu karar ağaçları oluşturan ve regresyon için ortalama tahmini ya da sınıflandırma için çoğunluk oyunu çıktı olarak veren güçlü bir topluluk öğrenme yöntemidir. Özellikle değişkenler arasındaki karmaşık, doğrusal olmayan ilişkileri modellemek için uygundur ve aşırı öğrenmeye karşı dayanıklılık, ölçek değişmezliği ve eksik verilere tolerans sunmaktadır. Breiman (2001) tarafından tanıtıldığı şekliyle, RF modeli her ağacı veri kümesinden bir önyükleme örneği kullanarak oluşturmakta ve her bölünme düğümünde rastgele bir tahmin değişkenleri alt kümesi seçerek ek rastgelelik sağlamaktadır (Auret ve Aldrich, 2012). Bir $x_g$örneği için nihai tahmin $\hat{y}_g$, topluluktaki tüm K ağaçlarının çıktısının ortalaması ile verilmektedir ve denklem (3.1) ile matematiksel olarak ifade edilmiştir:

$$\hat{y}_g = \frac{1}{K} \sum_{k=1}^{K} T_{L(h_k)}(x_g) \tag{3.1}$$

$T_{L(h_k)}$, önyükleme örneği $L(h_k)$ üzerinde eğitilen k. ağaçtan tahmini ifade etmektedir. Modelin genelleme hatası, çanta dışı (Out-of-Bag) örnekler kullanılarak tahmin edilebilir ve bu da denklem (3.2)'de matematiksel olarak ifade edilen sözde-R^2performans ölçüsüne yol açmaktadır:

$$R^2 = 1 - \frac{MSE_{OOB}}{Var(y)} \tag{3.2}$$

Tahmine ek olarak, RF, değişken önem ölçütleri ve kısmi bağımlılık grafikleri aracılığıyla yorumlanabilirlik sağlamaktadır. Değişken önemi, her özelliği permütasyona tabi tutarak ve bunun sonucunda model hatasındaki artışı değerlendirerek ölçülmekte, böylece bir tahmin edicinin yanıta olan ilgisi yakalanmaktadır. Bu yetenekler, RF modelini, endüstriyel süreçlerde hata teşhisi veya performans optimizasyonu gibi değişken etkisini anlamanın tahmin doğruluğu kadar kritik olduğu sistemlerde değerli kılmaktadır (Auret ve Aldrich, 2012).

3.1.2. Çoklu doğrusal regresyon

Bir regresyon modeli birden fazla bağımsız değişken içerdiğinde, bu durum çoklu regresyon olarak adlandırılmaktadır. Çoklu doğrusal regresyon, basit doğrusal regresyonu genelleştirmekte olup, basit modelden kaynaklanan birçok ilke hâlâ geçerliliğini korumaktadır. Saçılım grafikleri, korelasyon analizi ve en küçük kareler tahmin tekniği gibi araçlar, çoklu regresyon bağlamında temel niteliğini sürdürmektedir. MLR modeli denklem (3.3)'de matematiksel olarak gösterilmiştir:

$$Y_i = \alpha + x_{i,1}\beta_1 + x_{i,2}\beta_2 + \ldots + x_{i,p}\beta_p + e_i = \alpha + x_i^T\beta + e_i = \alpha + \beta^T x_i e_i \qquad (3.3)$$

Burada, $i = 1, \ldots, N$, Y_i yanıt değişkenini, x_i önemsiz olmayan tahmin edicilerden oluşan p × 1 vektörünü, α bilinmeyen sabiti, β bilinmeyen katsayılardan oluşan p×1 vektörünü ve e_i hata olarak adlandırılan rastgele değişkeni temsil etmektedir (Shen, 2009).

3.1.3. Destek vektör regresyonu

SVR genel performansı artırmak için yalnızca eğitim hataları yerine, genelleme hatalarını minimize etmeye odaklanmaktadır. Girdi verilerinin doğrusal olmayan bir fonksiyon aracılığıyla haritalandığı yüksek boyutlu özellik uzayında doğrusal regresyon fonksiyonunu hesaplayarak çalışmaktadır (Zhang ve O'Donnell, 2019). Regresyon fonksiyonu, denklem (3.4)- (3.7)'de verildiği gibi bir optimizasyon modeli olarak ifade

edilebilmektedir. Amaç fonksiyonu, *w* vektörünün büyüklüğü arttıkça daha fazla sayıda değerinin sıfırdan farklı hale geldiğini ve bu durumun daha karmaşık çözümlere yol açtığını göstermektedir. C, bir düzenleme parametresidir, dolayısıyla bu çok amaçlı optimizasyon probleminde düzlemselliği azaltmaya veya hatayı minimize etmeye daha fazla vurgu yapan ayarlanabilir bir parametredir. Örneğin, daha büyük bir C değeri hatayı minimize etmeye daha fazla ağırlık vermektedir (Awad ve Khanna, 2015). SVR'de kullanılan yaklaşıma benzer şekilde yumuşak-marj yaklaşımı benimsendiğinde, aykırı değerlere karşı koruma sağlamak için modele ξ_i^* , ξ_i gevşek değişkenleri eklenebilmektedir. ε hata tolerans değeridir. Daha küçük bir ε değeri hata için daha düşük tolerans göstermekte ve destek vektörlerinin sayısını ve dolayısıyla çözümün seyrekliğini etkilemektedir.

$$min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i + \xi_i^* \tag{3.4}$$

Kısıtlar

$$y_i - w^T x_i \leq \varepsilon + \xi_i^* \quad i = 1, \ldots, N. \tag{3.5}$$

$$w^T x_i - y_i \leq \varepsilon + \xi_i \quad i = 1, \ldots, N. \tag{3.6}$$

$$\xi_i^*, \xi_i \geq 0 \quad i = 1, \ldots, N. \tag{3.7}$$

Burada:

- w: Ağırlık vektörü
- C: Hatayı minimize etme ile model karmaşıklığı arasındaki dengeyi kontrol eden düzenleme parametresi
- ξ_i^*, ξ_i : Hatayı minimize etme ile model karmaşıklığı arasındaki dengeyi kontrol eden düzenleme parametresi
- ε: Hata toleransı
- x_i: Girdi veri noktaları
- y_i: Hedef değerleri.
- N: Toplam eğitim örneği sayısı

SVR, zaman serisi tahmini, finansal tahminler, karmaşık mühendislik yaklaşımları ve konveks kuadratik programlama gibi alanlarda yaygın olarak kullanılmaktadır.

3.1.4. Aşırı gradyan güçlendirme

XGBoost, tahmin doğruluğu ve hesaplama verimliliği nedeniyle geniş popülarite kazanan, ölçeklenebilir ve verimli bir gradyan güçlendirme uygulamasıdır. XGBoost, düzenleme, küçültme, seyreklik farkında algoritmalar ve paralel hesaplama yoluyla performansını artırmaktadır. Bu modelde tahminler, her yeni ağacın düzenlenmiş amaç fonksiyonunu minimize etmek üzere eğitildiği bir dizi regresyon ağacının çıktılarının toplamı ile yapılmaktadır. XGBoost algoritmasının toplam kayıp fonksiyonu denklem (3.8)'de verilmiştir.

$$L(\phi) = \sum_{i=1}^{n} l\left(y_i, \hat{y_i}^{t}\right) + \sum_{k=1}^{t} \Omega(f_k) \tag{3.8}$$

L Türevlenebilir bir kayıp fonksiyonudur (örneğin, kare hata). $\hat{y}_i^t,\ t$ İterasyonundaki i örneği için yapılan tahmindir. $\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2$ İse her f_k ağacının karmaşıklığını cezalandıran bir düzenleme terimidir. Burada T yaprak sayısını ve ω yaprak ağırlıklarını temsil etmektedir. Model, tahminleri denklem (3.9) ile güncellemektedir:

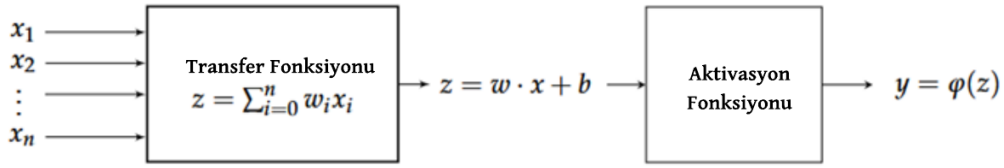
$$\hat{y}_i^t = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{3.9}$$

Kayıp fonksiyonunun ikinci dereceden Taylor açılımını ve açgözlü ağaç yapılandırma algoritmalarını kullanarak, XGBoost model karmaşıklığı ile tahmin gücü arasında etkili bir denge kurmaktadır. Bu yaklaşım aşırı öğrenmeyi önlemeye yardımcı olmakta ve seyrek girdi ile eksik verileri desteklemesi, modeli gerçek dünya tahminleri için güçlü kılmaktadır (Chen ve Guestrin, 2016).

3.1.5. Yapay sinir ağları

ANN'ler, katmanlar halinde düzenlenmiş yapay nöronlar adı verilen temel unsurları kullanarak insan beyninin bilgiyi işleme şeklini taklit etmek için oluşturulan doğrusal olmayan hesaplama modelleridir. Bu ağlar, yinelemeli ve evrişimli tipler de dahil olmak üzere farklı biçimlerde ortaya çıkmaktadır. Her bir yapay nöron, $i = 1\ldots,n$ için x_i

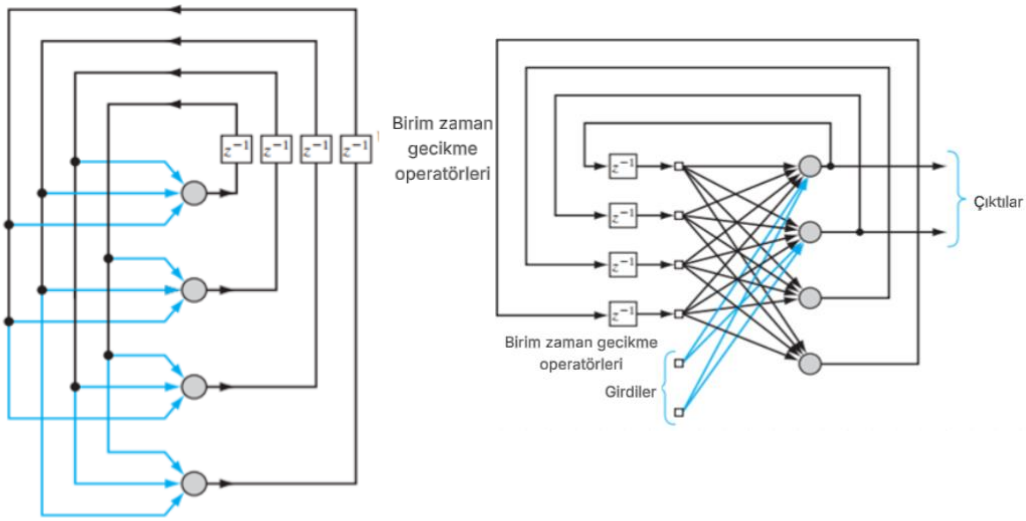
girdileri üzerinde hem transfer fonksiyonunu (doğrusal) hem de aktivasyon fonksiyonunu (doğrusal olmayan) zincirleyerek y çıktısını elde eden matematiksel bir fonksiyon olarak görülebilir. Şekil 3.1 tek bir yapay nöronun işleyişini göstermektedir (Cabello, 2022). Burada b sapma olarak adlandırılır (b, verilen verileri gerçek ihtiyaçlara uydurmaya yardımcı olan insan filtresini taklit etme girişimidir). w_i İle gösterilen ağırlıklar $(i = 1\ldots, n)$ bir z değeri üretir, bu z değeri de φ aktivasyon fonksiyonu aracılığıyla nihai çıktı y'yi oluşturur: $y = \varphi(z)$.



Şekil 3.1. Tek yapay nöron işlevi (Cabello, 2022).

3.1.6. Yinelemeli sinir ağları

RNN, 1990'larda araştırma ve geliştirmenin önemli bir odak noktası olmuştur. Bu ağlar, sıralı veya zaman içinde değişen kalıpları öğrenmek üzere tasarlanmıştır. RNN teknikleri çeşitli problemlere uygulanmıştır(Lu vd., 2016). RNN, en az bir geri beslenme döngüsüne sahip olması nedeniyle ileri beslemeli sinir ağından farklılık göstermektedir. Örneğin, yinelemeli bir ağ, her nöronun çıkış sinyalini diğer tüm nöronların girdilerine geri gönderdiği tek nöron katmanından oluşabilir; bu durum Şekil 3.2.A'de mimari diyagramda gösterilmektedir. Bu konfigürasyonda kendine geri beslenme döngüleri bulunmamaktadır, yani bir nöronun çıkışı kendi girdisine geri beslenmemektedir. Gösterilen yinelemeli ağ aynı zamanda gizli nöronlardan yoksundur. Şekil 3.2.B'de, gizli nöronları içeren başka bir yinelemeli ağ türü sunulmaktadır. Bu durumda, geri beslenme bağlantıları gizli ve çıkış nöronlarından gelmektedir. Geri beslenme döngülerinin varlığı ağın öğrenme yeteneği ve performansı üzerinde derin bir etkiye sahiptir. Ayrıca, geri beslenme döngüleri, birim-zaman gecikme elemanlarından (z^{-1} ile gösterilir) oluşan özel dalları içermekte ve sinir ağının doğrusal olmayan birimler içerdiği varsayımı altında doğrusal olmayan dinamik davranışa yol açmaktadır (Kussul vd., 2010).



Şekil 3.2. A) Kendine geri besleme döngüleri ve gizli nöronları olmayan yinelemeli ağ B) Gizli nöronlara sahip yinelemeli ağ (Haykin, 2009)

3.1.7. Uzun kısa süreli bellek modeli

LSTM mimarisi, hafıza blokları olarak adlandırılan yinelemeli şekilde birbirine bağlı alt-ağlardan oluşur. Bu bloklar, dijital bilgisayarlardaki hafıza çiplerine benzer şekilde ancak türevlenebilir bir biçimde işlev görmektedir. Her blok, bir veya daha fazla kendi kendine bağlı hafıza hücresini ve bu hücrelere uygulanan yazma, okuma ve sıfırlama işlemlerinin sürekli analogları olarak hizmet eden üç temel birim olan girdi, çıktı ve unutma kapılarını içermektedir (Graves, 2012).

LSTM ağı, standart RNN'ya benzemekle birlikte gizli katmandaki toplama birimlerinin yerini hafıza bloklarının alması ile farklılaşmaktadır. LSTM blokları geleneksel toplama birimleri ile birleştirilebilse de bu genellikle gereksizdir. Standart RNN'lerde kullanılan aynı çıktı katmanları, LSTM ağlarında da uygulanabilmektedir. Çarpımsal kapılar, LSTM hafıza hücrelerinin uzun süreler boyunca bilgiyi muhafaza etmesini ve geri çağırabilmesini sağlayarak, kaybolan gradyan problemini çözmektedir (Graves, 2012).

3.1.8. CatBoost regresyonu

CatBoost, geleneksel uygulamalarda hedef sızıntısının neden olduğu tahmin kayması problemini ortadan kaldırmak için tasarlanmış özelleşmiş bir gradyan güçlendirme algoritmasıdır. Bu algoritma iki temel yenilik sunar: sıralı güçlendirme ve sıralı hedef istatistikleri. Sıralı güçlendirme tekniği, eğitim veri kümesinin rastgele permütasyonuna dayanmakta ve model eğitimi sırasında her tahmin yalnızca permütasyondaki kendisinden önce gelen elemanlar kullanılarak hesaplanmaktadır. Bu yaklaşım, örneklem-dışı senaryoyu simüle ederek her model güncellemesinin yansız olmasını sağlamaktadır (Prokhorenkova vd., 2018). Model yinelemeli olarak eğitilmekte, her t adımında denklem (3.10)'da verilen tahmin fonksiyonu güncellenmektedir:

$$F^{(t)} = F^{(t-1)} + \alpha h_t \tag{3.10}$$

h_t kayıp fonksiyonunun negatif gradyanını yaklaşık olarak temsil eden regresyon ağacını, α ise öğrenme oranını ifade eder. Hedef bilgisini sızdırmadan kategorik değişkenleri etkili bir şekilde kodlamak için CatBoost, sıralı hedef istatistiklerini kullanmaktadır. Burada, belirli bir kategorik değer x_k^i için kodlama, permütasyondaki önceki gözlemler kullanılarak hesaplanır. Kodlama denklem (3.11)'de tanımlanmaktadır:

$$\hat{x}_k^i = \frac{\sum_{j:\sigma(j)<\sigma(k)} 1\left(x_j^i = x_k^i\right) y_i + bp}{\sum_{j:\sigma(j)<\sigma(k)} 1\left(x_j^i = x_k^i\right) + b} \tag{3.11}$$

Burada p öncül bir değerdir (genellikle global ortalama), b aşırı öğrenmeyi önleyen bir yumuşatma parametresidir, i gözlem, j öngörücü değeri ve k örnek olak tanımlanmıştır. Bu yenilikler CatBoost'un son derece doğru ve yansız tahminler elde etmesini sağlayarak onu karmaşık ve yüksek kardinaliteli kategorik özellikler içeren regresyon görevleri için güçlü bir araç haline getirmektedir (Prokhorenkova vd., 2018).

3.1.9. En küçük mutlak küçültme ve seçim operatörü regresyonu

LASSO-R, Tibshirani (1996) tarafından tanıtılan ve aynı anda değişken seçimi ve katsayı küçültmesi yaparak doğrusal regresyon modellerinin tahmin doğruluğunu ve yorumlanabilirliğini artıran bir düzenleme tekniğidir. Model parametreleri üzerinde herhangi bir kısıtlama olmaksızın artık kareler toplamını en aza indiren geleneksel en küçük kareler regresyonunun aksine, LASSO-R regresyon katsayılarının büyüklüğüne bir $L1$ cezası uygular. LASSO-R regresyon için optimizasyon problemi denklem (3.12)'de tanımlanmaktadır:

$$\min_{\beta}\{\sum_{i=1}^{n}\ (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2\ \} \tag{3.12}$$

Kısıtlar $\sum_{j=1}^{p}|\beta_j| \leq t$, ve $t \geq 0$ ceza gücünü kontrol etmek için bir ayar parametresidir. Bu problem benzer şekilde, denklem (3.13)'deki gibi de ifade edilebilmektedir:

$$\min_{\beta}\{\sum_{i=1}^{n}\ (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2\ +\ \lambda\ \sum_{j=1}^{p}|\beta_j|\} \tag{3.13}$$

Burada λ bir düzenleme parametresidir. λ arttıkça, $L1$ cezası bazı katsayıları tam olarak sıfıra zorlar ve ilgili değişkenlerin modelden çıkarılmasını sağlar. Bu özellik LASSO-R'nu özellikle yüksek boyutlu ortamlarda veya çoklu doğrusallığın mevcut olduğu durumlarda kullanışlı hale getirir. LASSO-R tahmin edici değişkenlerin seyrek bir alt kümesi seçerek iyi bir genelleme performansı sağlar. $L1$ kısıtının dışbükey yapısı da hesaplama fizibilitesini sağlar. Yorumlanabilir modeller üretme yeteneği LASSO-R'nu istatistiksel öğrenme ve regresyon modellemesinde temel bir araç haline getirmiştir Tibshirani (1996).

3.1.10. Hafif gradyan güçlendirme makinesi

LightGBM, Ke vd. (2019) tarafından önerilen, eğitim hızını ve bellek tüketimini optimize etmek üzere tasarlanmış, özellikle büyük ölçekli ve yüksek boyutlu veriler için son derece verimli bir gradyan güçlendirme karar ağacı (GBDT: Gradient Boosting Decision Tree)

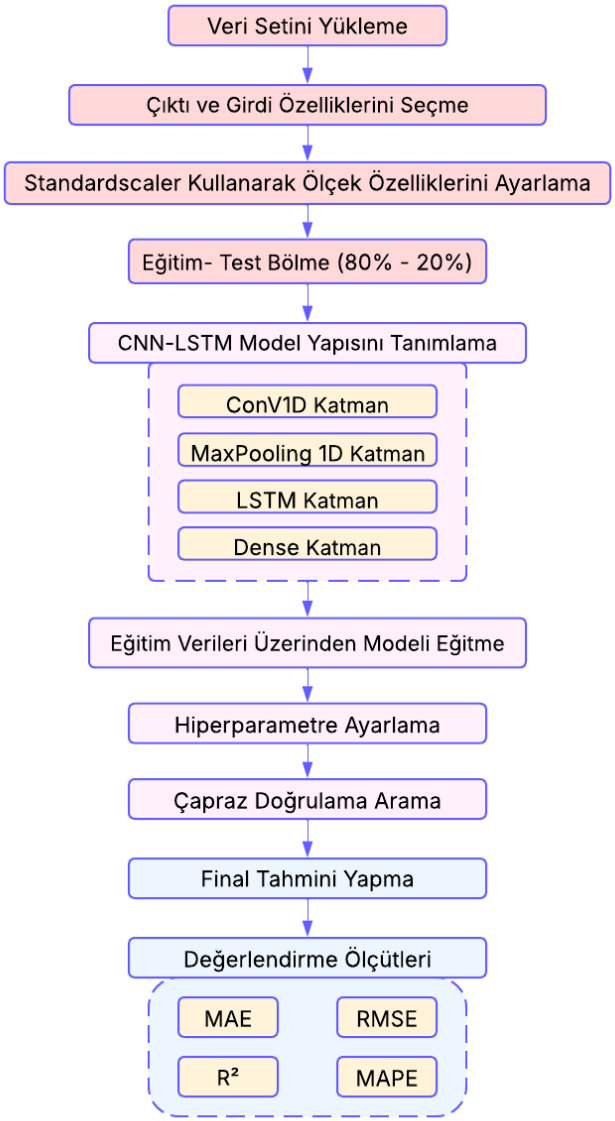
mimarisidir. Geleneksel GBDT uygulamalarının aksine, LightGBM iki yeni teknik sunmaktadır: Gradyan Tabanlı Tek Taraflı Örnekleme (GOSS: Gradient-based One-Side Sampling) ve Özel Özellik Demetleme (EFB: Exclusive Feature Bundling).

GOSS, bilgi kazancı hesaplamasında en etkili olan büyük gradyanlara sahip örnekleri seçici olarak muhafaza ederken, küçük gradyanlara sahip örneklerden rastgele örnekleme yapmaktadır. Bu strateji, model doğruluğundan ödün vermeden hesaplama yükünü önemli ölçüde azaltmaktadır. Teorik olarak da GOSS'un bilgi kazancına düzgün rastgele örneklemeden daha doğru bir şekilde yaklaştığı doğrulanmaktadır (Ke vd., 2019). EFB, nadiren eş zamanlı olarak sıfırdan farklı değerler alan, karşılıklı olarak dışlayıcı özellikleri tek bir bileşik özellik halinde demetleyerek özellik seyrekliğini ele almaktadır. EFB, açgözlü grafik renklendirme yaklaşımı yoluyla etkin özellik sayısını azaltarak, bölünme kararlarının doğruluğunu etkilemeden hesaplama maliyetini düşürmektedir (Ke vd., 2019).

3.1.11. Evrişimsel sinir ağları ve uzun kısa süreli bellek hibrit modeli

CNN-LSTM hibrit modeli, CNN ve LSTM ağlarını entegre etmektedir. CNN anahtar kalıpları tanımlayarak özellik çıkarımında üstün performans sergilemekte, LSTM ise zamansal bağımlılıkları yakalayarak zaman serisi tahmini için uygun hale gelmektedir. Modelin ana yapısı, giriş katmanı, tek boyutlu evrişim katmanı, havuzlama katmanı, LSTM gizli katmanı ve tam bağlantı katmanını içeren CNN ve LSTM'den oluşmaktadır (Lu vd., 2020).

Model, CNN ve LSTM'nin özelliklerine göre inşa edilmektedir. Özellik çıkarımı için Conv1D katmanı, boyut indirgeme için MaxPooling1D katmanı ve zaman bağımlılıklarını yakalamak için LSTM katmanını içeren Keras Sequential API kullanılarak CNN-LSTM tabanlı bir tahmin modeli oluşturulmaktadır. CNN-LSTM modeli, nihai çıktıyı, tam bağlantılı Dense katmanları kullanarak üretmektedir (Lu vd., 2020). Şekil 3.3, CNN-LSTM hibrit modelinin akış şemasını göstermektedir.



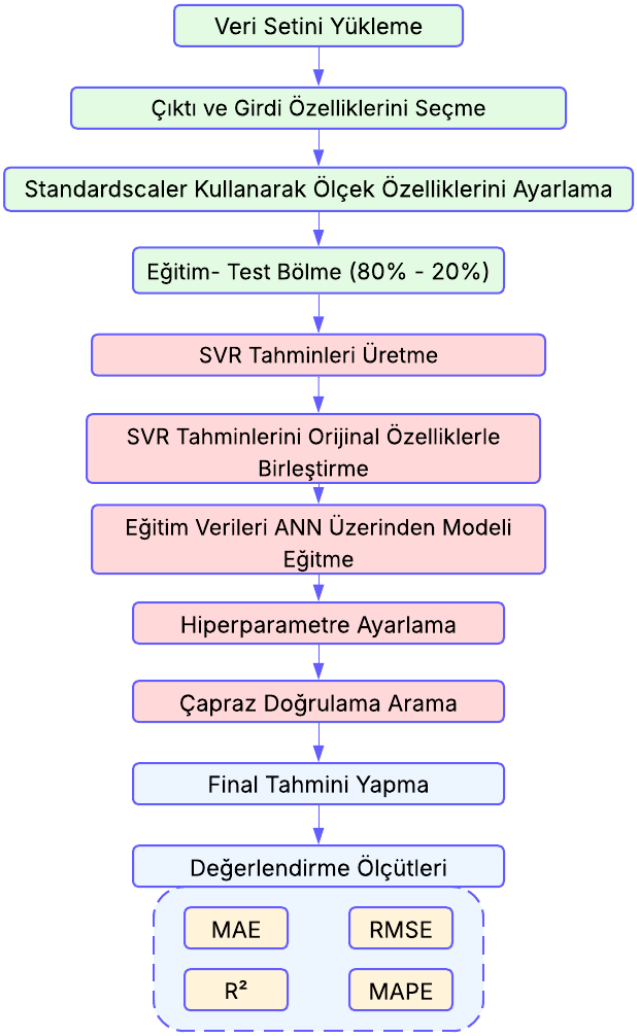
Şekil 3.3. CNN-LSTM hibrit modeli akış şeması

3.1.12. Destek vektör regresyonu ve yapay sinir ağı hibrit modeli

SVR-ANN hibrit modeli, daha yüksek doğruluk elde etmek için SVR ve ANN'nin güçlü yanlarını birleşik olarak kullanmaktadır. SVR modülü öncelikle değişkenlerin doğrusal olmayan ilişkilerini tespit etmek için özellik çıkarıcı olarak kullanılmaktadır. Diğer taraftan, ANN bölümü daha derin düzeydeki özellikleri öğrenerek tahminlerin iyileştirilmesini gerçekleştirmektedir. Bu hibrit yöntemin kombinasyonu ile genelleme yeteneği artırılmakta ve daha güçlü tahminler elde edilmektedir (Xiang vd., 2018).

İlk olarak, girdi özellikleri kullanılarak bir SVR modeli eğitilmektedir. Tahmin performansını iyileştirmek için düzenleme parametresi (C), çekirdek (kernel) türü (doğrusal, radyal baz fonksiyonu, sigmoid ve poly) ve hata tolerans payı (ε) dahil olmak üzere temel SVR hiper parametrelerini optimize etmek amacıyla ızgara araması gerçekleştirilmektedir (Xiang vd., 2018).

Eğitilmiş en iyi SVR modeli daha sonra orijinal veri kümesi ile birlikte ek özellikler olarak kullanılan tahminler üretmektedir. İkinci aşamada, SVR tarafından türetilmiş bu özellikler ölçeklendirilmiş orijinal girdi özellikleri ile birleştirilerek ANN modeline beslenmektedir. ANN, düzeltmeli doğrusal birim (ReLU: Rectified Linear Unit**)** aktivasyonuna sahip birden çok yoğun(dense) katmandan oluşmakta, kayıp fonksiyonu olarak ortalama karesel hata (MSE: Mean Squared Error) kullanmakta ve Adam optimizasyon algoritması kullanılarak eğitilmektedir. Hibrit model, tahmin hatasını minimize etmek ve tahmin yeteneğini artırmak için 100 dönem (epoch) boyunca eğitilmekte ve doğrulanmaktadır. Şekil 3.4, SVR-ANN hibrit modeline ait akış şemasını göstermektedir.



Şekil 3.4. SVR-ANN hibrit modeli akış şeması

3.1.2. Model performans metrikleri

Tez çalışmasında kullanılan performans değerlendirme metrikleri ve bu metriklerin matematiksel ifadeleri Çizelge 3.1'de gösterilmiştir. Ortalama Mutlak Hata (MAE: Mean Absolute Error), iki sürekli değişken arasındaki farkı ölçmekte; Ortalama Karesel Hata Kökü (RMSE: Root Mean Squared Error), ortalama karesel hatanın karekökünü ifade etmekte; Ortalama Mutlak Yüzde Hata (MAPE: Mean Absolute Percentage Error), hatanın yüzdelik olarak büyüklüğünü ölçmekte; Belirleme Katsayısı ise korelasyon katsayısının karesidir (R^2: Square of Correlation) (Naser ve Alavi, 2023).

Çizelge 3.1. Performans değerlendirme metrikleri

Performans Metrikleri	Formülasyon
MAE	$\frac{\sum_{i=1}^{n}\lvert E_i\rvert}{n}$
RMSE	$\sqrt{\frac{\sum_{i=1}^{n} {E_i}^2}{n}}$
MAPE	$\frac{100}{n}\sum_{i=1}^{n}\lvert E_i\rvert / \lvert A_i\rvert$
R^2	$1 - \sum_{i=1}^{n}(P_i - A_i)^2 / \sum_{i=1}^{n}(A_i - A_{ortalama})^2$

Notasyon

- A: Gerçek ölçümler,
- P: Tahminler,
- n: Veri noktası sayısı,
- E: Hata ($E = A - P$)

3.2. Geliştirilen Tahmin Modelleri

Tez çalışmasında, önerilen tahmin sistemini geliştirmek ve değerlendirmek için iki temel veri seti kullanılmaktadır: Konteyner iş hacmi veri seti ve teslim süresi veri seti. Her iki veri seti de denizcilik lojistiğinde operasyonel verimlilik modellemesi ve performans tahmini açısından kritik öneme sahiptir.

3.2.1. Konteyner iş hacminin talebi için geliştirilen tahmin modeli

Konteyner iş hacmi veri seti, Türkiye'deki aylık konteyner iş hacmini içermekte olup, zaman içindeki dinamik talep kalıplarını ve liman performansını yansıtmaktadır.

3.2.1.1 Veri toplama

Bu bölümde, konteyner iş hacmi talebi için veri toplama stratejisi açıklanmaktadır. Türkiye'nin konteyner iş hacmi talebini ve bunun temel etkileyici faktörlerini bir arada ele alan mevcut bir veri seti bulunmadığından, literatürde yaygın olarak kullanılan gerekli özellikler entegre edilmiştir.

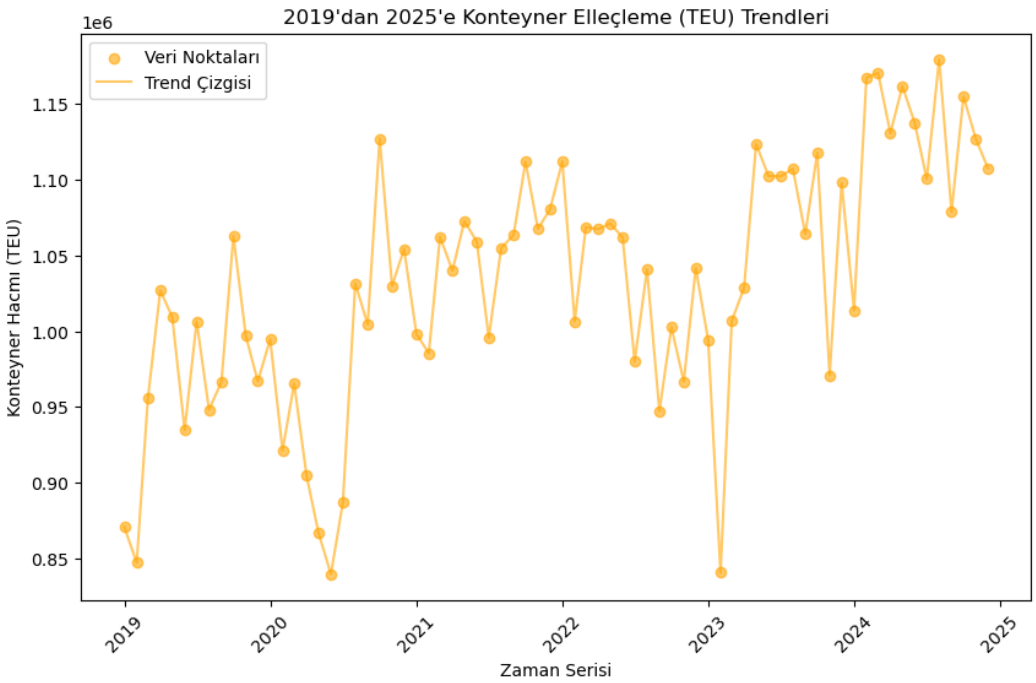
Analizin temeli, Türkiye Cumhuriyeti Ulaştırma ve Altyapı Bakanlığı Denizcilik Genel Müdürlüğü'nün resmî web sitesinde aylık olarak yayımlanan ve TEU cinsinden ifade edilen yüksek frekanslı konteyner iş hacmi verilerinin doğrudan temin edilmesiyle oluşturulmuştur.

Bu çalışmada, konteyner iş hacmi bağımlı değişken olarak belirlenmiştir. Bağımsız değişkenler arasında, deniz ticaretini temsil eden toplam ithalat (TI: Total Import) ve toplam ihracat (TE: Total Export) değerleri yer almaktadır. Buna ek olarak, ekonomik özelliklerden biri olan aylık cari fiyatlarla gayri safi yurt içi hasıla (MCP-GDP: Monthly Current Prices of Gross Domestic Products), Türkiye İstatistik Kurumu'ndan bin Türk Lirası cinsinden elde edilmiştir.

Diğer önemli göstergeler ise farklı güvenilir veri tabanlarından sağlanmıştır. Bunlar; Aylık Gemi Yakıt Fiyatları (MVFP: Monthly Vessels Fuel Prices), Türk Lirası'nın Amerikan Doları karşısındaki değeri (TRY_USD) ve Euro karşısındaki değeridir (TRY_EUR).

Türk Deniz Ticaret Odası tarafından yayımlanan verilere göre, Türkiye'nin denizyolu konteyner iş hacmi, 2014 yılında 8,3 milyon TEU iken 2023 yılı itibarıyla 12,5 milyon TEU'ya yükselmiştir, Şekil 3.5'te 2019'dan 2024'e konteyner iş hacmi (TEU) trendleri gösterilmiştir. Aynı dönemde, ithalat konteyner iş hacminin 3,4 milyon TEU'dan 4,9 milyon TEU'ya; ihracat konteyner iş hacminin ise 3,5 milyon TEU'dan 4,8 milyon TEU'ya ulaştığı bildirilmiştir. Bu artış, Türkiye'nin deniz ticaretinde gösterdiği sürekli büyümeyi ortaya koymaktadır. Küresel konteyner taşımacılığı verilerine dayanarak,

Türkiye'nin 2023 yılı itibarıyla %2,1'lik yıllık TEU değişim oranıyla dünya genelinde 18. sırada konumlandığı tespit edilmiştir (Maritime Sector Report, 2024).



Şekil 3.5. Türkiye'de toplam konteyner iş hacmi (TEU)

Toplam ithalat ve ihracat talebi, Türk bayrağı altında taşınan (boş ve dolu) konteyner iş hacimleri ile yabancı bayraklar altında taşınan (boş ve dolu) konteyner iş hacimlerinin toplamı şeklinde tanımlanabilir. Bu ilişki Denklem (3.14) ile ifade edilmektedir:

$$\text{İ}thalat\ Talebi_{Toplam} = T\ddot{u}rk\ Bayrağı\ Talebi_{Toplam} + Yabancı\ Bayrak\ Talebi_{Toplam} \quad (3.14)$$

Benzer şekilde, toplam ihracat talebi; Türk bayrağı taşıyan konteynerlerin (boş ve dolu) hacmi ile yabancı bayraklı konteynerlerin (boş ve dolu) hacimlerinin toplamı alınarak hesaplanabilmektedir. Bu durum ise Denklem (3.15) ile gösterilmektedir:

$$\text{İ}hracat\ Talebi_{Toplam} = T\ddot{u}rk\ Bayrağı\ Talebi_{Toplam} + Yabancı\ Bayrak\ Talebi_{Toplam} \quad (3.15)$$

3.2.1.2. Veri ön işleme

Veri toplama sürecinde, özelliklerin zaman ölçekleri farklılık göstermekteydi; bazı veriler günlük, bazıları aylık, bazıları ise üç aylık periyotlarla kaydedilmiştir. Yanıt değişkeni, aylık olarak kaydedilen ve TEU cinsinden ölçülen konteyner iş hacmi olduğundan, tüm girdilerin aynı zaman ölçeğine uyarlanabilmesi için çeşitli dönüşüm teknikleri uygulanmıştır. Örneğin, yakıt fiyatları günlük olarak sağlandığından, aylık ortalama değerleri alınarak bu değişken aylık forma dönüştürülmüştür. Bununla birlikte, MCP-GDP verileri üç aylık olarak yayımlandığı için ve bu değişkenin zincirlenmiş cari fiyatları yansıtması nedeniyle, aylık değerlere ulaşmak amacıyla TRY_USD değişkeni referans alınarak interpolasyon yöntemi kullanılmıştır. Öte yandan, diğer giriş değişkenlerinin verileri hâlihazırda aylık formatta olduğundan herhangi bir dönüşüm işlemine gerek duyulmamıştır.

3.2.1.3. Eksik veri işleme

Konteyner iş hacmi talebi verilerinde kayıtlar arasında herhangi bir eksik gözlem bulunmamıştır. Bu durum, veri setinin güvenilirliğini artırmakta ve ML modellerinin, önyargı veya gürültü yaratabilecek veri tamamlama ya da veri temizleme işlemlerine ihtiyaç duyulmaksızın eğitilmesine olanak tanımaktadır. Böylece, modellerin altta yatan gerçek desenleri daha etkin bir şekilde öğrenebilmesi mümkün olmakta ve bu da tahmin performansının doğruluğunu artırmaktadır.

3.2.1.4. Modelleme ve değerlendirme

Veri ön işleme adımlarının tamamlanmasının ardından, tahmin modellerinin oluşturulmasına yönelik modelleme süreci gerçekleştirilmiştir. Bu süreçte ilk olarak, değişkenlerin farklı ölçekte olmalarının model performansı üzerindeki olumsuz etkilerini gidermek amacıyla tüm sayısal değişkenlere normalizasyon uygulanmıştır. Bu işlemde, her bir özelliğin ortalaması sıfır ve standart sapması bir olacak şekilde dönüştürülmesini sağlayan StandardScaler yöntemi tercih edilmiştir. Model eğitimi için veri seti rastgele olacak şekilde %80 eğitim ve %20 test veri kümesine bölünmüştür. Eğitim verisi modelin

öğrenmesini sağlarken, test verisi modeli görmediği veriler üzerinde değerlendirmek amacıyla ayrılmıştır. Model performansını değerlendirebilmek amacıyla çeşitli regresyon başarı metrikleri kullanılmıştır. Bu kapsamda, MAE, RMSE, MAPE ve R^2 metrikleri dikkate alınmıştır. Böylece, her bir modelin teslim süresini ne ölçüde doğru tahmin edebildiği, sayısal olarak karşılaştırılabilir hale getirilmiştir.

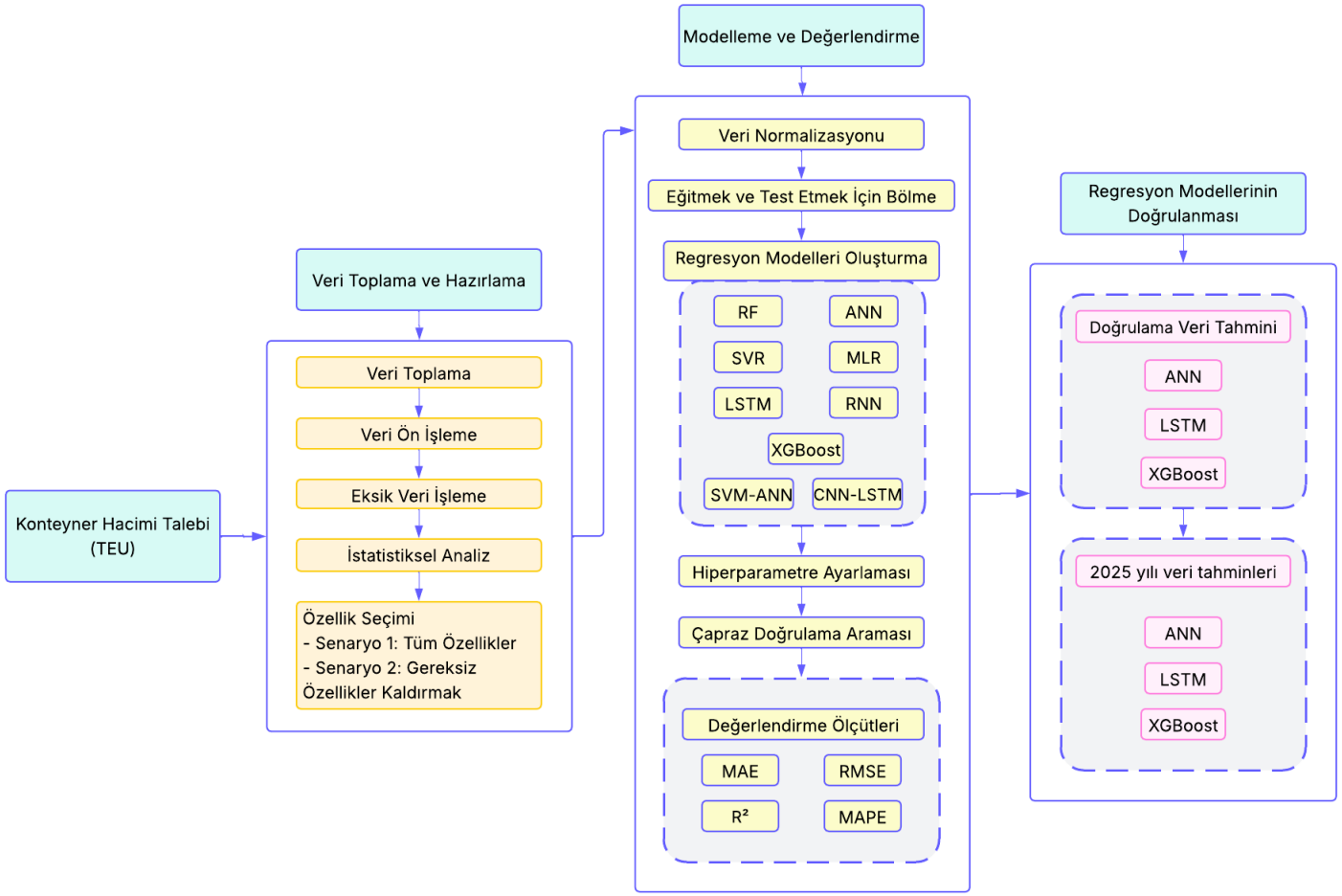
Her bir model için hiperparametre ayarları Izgara Arama (Grid Search) yöntemiyle optimize edilmiştir. Bu süreçte, örneğin SVR için kernel tipi, C ve gamma parametreleri; RF için ağaç sayısı ve maksimum derinlik; ANN için katman sayısı, nöron sayısı, aktivasyon fonksiyonu ve öğrenme oranı gibi parametreler sistematik biçimde test edilmiştir.

3.2.1.5. Modellerin doğrulanması

Geliştirilen tahmin modellerinin güvenilirliğini ve genellenebilirliğini test etmek amacıyla çok katmanlı doğrulama stratejileri izlenmiştir. İlk olarak, eğitim ve test veri kümelerine ayrılmış veri seti üzerinde, eğitim sırasında modelin aşırı öğrenme veya yetersiz öğrenme eğilimi gösterip göstermediğini anlamak için karşılaştırmalı analizler yapılmıştır.

Ardından, daha sağlam bir doğrulama süreci için k-katlı çapraz doğrulama (k-fold cross-validation) uygulanmıştır. Bu tez çalışmasında genellikle k=5 olarak belirlenmiş, yani veri seti beş alt kümeye ayrılmıştır. Her iterasyonda dört parça eğitimde, bir parça ise doğrulamada kullanılarak modelin farklı alt kümelerdeki kararlılığı test edilmiştir. Böylece veri kümesine özel öğrenme yerine daha genellenebilir modeller elde edilmeye çalışılmıştır.

Doğrulama sürecinde, modellerin doğrulama veri seti üzerinde test edilmesine de özen gösterilmiştir. Bu amaçla, ana veri setinden tamamen bağımsız olarak ayrılmış yeni verileri kullanılarak modellerin tahmin gücü değerlendirilmiş ve bu sayede gerçek dünya koşullarına yakın senaryolarda da başarılı olup olmadıkları kontrol edilmiştir. Şekil 3.6 konteyner iş hacmi tahmin modelinin akışını göstermektedir



Şekil 3.6. Konteyner iş hacminin talebi tahmin yaklaşımı akış şeması

3.2.2.Teslim süresi için geliştirilen tahmin modeli

Teslim süresi veri seti, gemi kalkışı ile varışı arasındaki sevkiyat sürelerine ilişkin ayrıntılı kayıtların yanı sıra bu süreci etkileyen çeşitli özellikleri de içermektedir. Bu veri yapısı, nakliye rotaları boyunca hizmet güvenilirliği ve transit süresindeki değişkenlik hakkında kritik içgörüler sunmaktadır. Şekil 3.7'de teslim süresi tahminine yönelik geliştirilen model açıklanmaktadır.

3.2.2.1. Veri toplama

Geminin, kalkıştan varışa kadar olan tüm duraklarını ve alt turlarını içeren hareketlerinin izlenebilmesi amacıyla, AIS teknolojisini kullanan bir nakliye platformundan teslim süresi verileri toplanmıştır.

Teslim süresi veri seti, modellerde doğrudan kullanılan ya da uygun biçimlere dönüştürülerek kullanılan çeşitli özellikler içermektedir. Bu özellikler, genel gemi nitelikleri, zamansal faktörler, mekânsal öğeler ve deniz rotası karakteristikleri gibi özellik kategorileri altında sınıflandırılmış ve analiz edilmiştir.

Genel gemi nitelikleri kapsamında, Deniz Mobil Hizmet Kimliği (MMSI: Maritime Mobile Service Identity) ve Uluslararası Denizcilik Örgütü numarası (IMO: International Maritime Organization) her satırın benzersiz bir sefere ait olduğunu gösteren tekil sefer tanımlayıcıları olarak kullanılmaktadır. Gemi alt tipi değişkeni ise, tanker tipi belirtilmeksizin iki kategoriye ayrılmıştır: biri kargo gemilerini, diğeri ise tanker gemilerini kapsamaktadır. Ayrıca, elleçleme özelliklerini ve operasyonel kısıtlamaları etkileyen ölü ağırlık tonajı (DWT) ve gemi boyutları da bu kategoriye dâhil edilmiştir. Daha büyük gemiler, kütlelerinden dolayı hızlanma ve yavaşlama işlemleri için daha fazla zamana ihtiyaç duymakta, bu da doğrudan teslim süresini etkileyebilmektedir. Örneğin, kargo gemileri ile tankerlerin, sefer sürelerini etkileyebilecek kendilerine özgü performans ve elleçleme karakteristikleri bulunmaktadır. Bu çalışmada, söz konusu

özelliklerin modele dâhil edilerek, farklı operasyonel kısıtlamaların hesaba katılması yoluyla teslim süresi tahminlerinin doğruluğunun artırılması önerilmektedir. Buna ek olarak, bu kategori içinde geminin draftı ve yük koşulları gibi ilave faktörler de yer almaktadır.

Deniz operasyonlarının dinamik yapısını yakalamak için bir dizi zamansal özellik çıkarılmış ve tahmin modellerine dahil edilmiştir. Bu özellikler, seyir davranışını ve teslim süresi üzerindeki etkisini anlamak için önemli olan zaman temelli değişkenleri içermektedir. Her seyir sırasında kaydedilen ortalama ve maksimum hızlar, operasyonel verimlilik ve program sapmaları hakkında bilgi sağlar ki bu gecikmelerin göstergesi olmaktadır. Ancak veri setinde bu özelliğin kullanılmasının temel nedeni, doğrudan teslim süresini etkilemesidir. Ayrıca, gerçek kalkış zamanı ve varış zamanı, teslim süresinin kesin hesaplanmasını sağlayan ve aynı zamanda liman tıkanıklığı ile transit değişkenliğinin mevsimsel veya günün saatine bağlı etkilerinin analizine olanak tanıyan kritik zamansal özelliklerdir. Karşılıklı olarak, bu özellikler tahmin modellerinin zamansal temelini oluşturarak, karmaşık zamansal örüntüleri öğrenme ve tahmin doğruluğunu artırma becerilerini güçlendirmektedir.

Zamansal değişkenlere ek olarak, deniz operasyonlarının coğrafi boyutlarını dikkate alabilmek amacıyla çeşitli mekânsal özellikler de tahmin modellerine entegre edilmiştir. Kalkış ve varış limanlarının enlem ve boylam koordinatları, rota geometrisi gibi mekânsal desenlerin çıkarılmasını mümkün kılan doğrusal mekânsal göstergeler olarak işlev görmektedir. Bu koordinatlar, coğrafi konumun seyir davranışları üzerindeki etkisini analiz etmede kritik bir rol üstlenmektedir. Ayrıca, sefer mesafesi, güzergâhın fiziksel uzunluğunu nicel olarak tanımlayarak modellerde kullanılan merkezi bir mekânsal değişken olarak değerlendirilmektedir.

Tahmin modellerinin kestirim yeteneklerinin daha da geliştirilmesi amacıyla, seyir sırasında karşılaşılan meteorolojik koşulları yansıtan çevresel/deniz rotası özellikleri de modele entegre edilmiştir. Bu kapsamda modele dâhil edilen başlıca değişkenler arasında rüzgâr hızı ve rüzgâr yer almaktadır. Söz konusu parametrelerin, gemi performansı ile seyir güvenilirliği üzerinde doğrudan etkiler oluşturabildiği bilinmektedir. Örneğin,

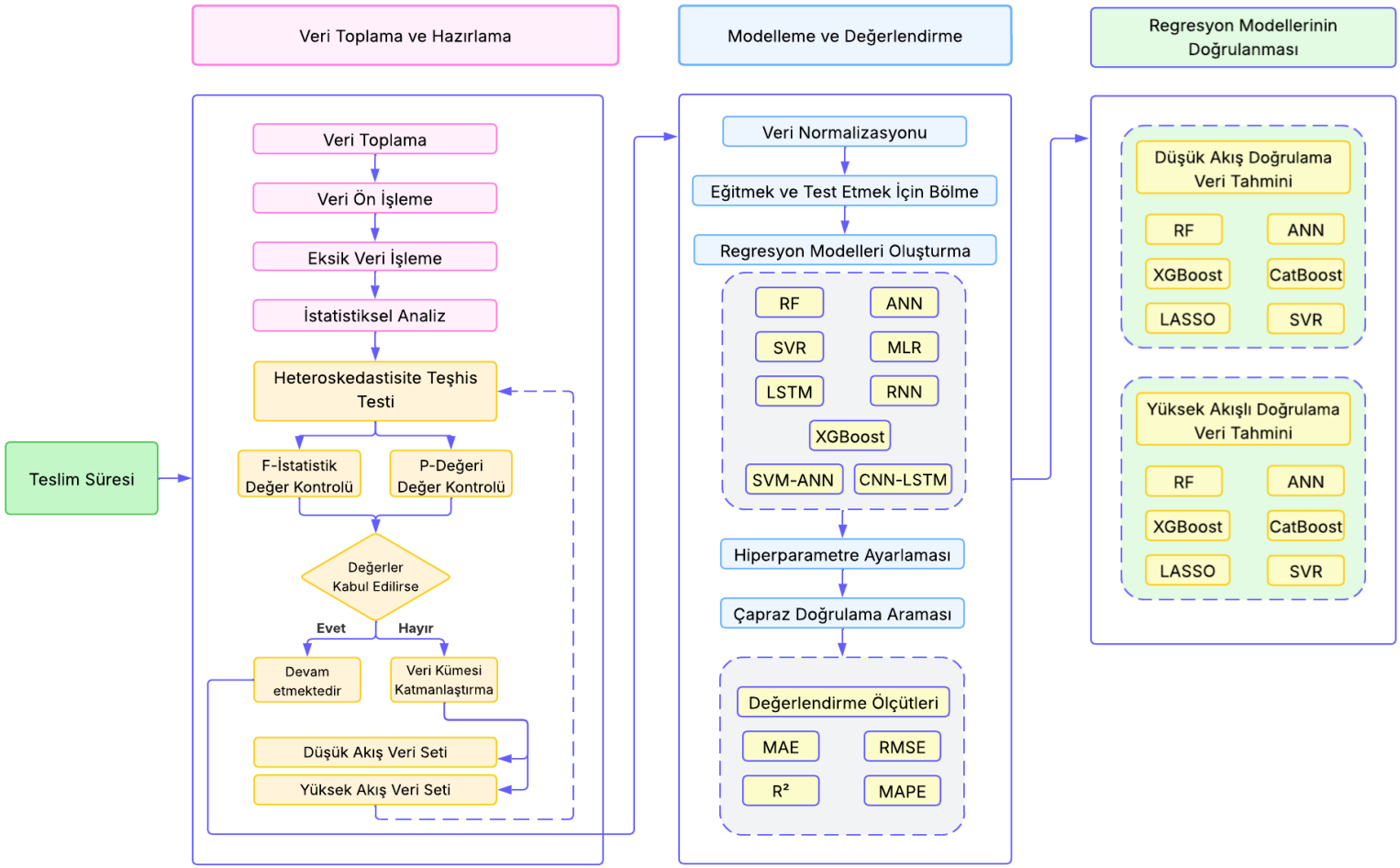
rüzgâr hızı, özellikle açık deniz koşullarında geminin direncini artırarak, yüksek rüzgâr durumlarında teslim süresinin uzamasına yol açabilmektedir. Rüzgâr açısı (rüzgârın gemi rotasına göre yönü), karşı veya yan rüzgâr etkisiyle gemi hızını ve seyrin stabilitesini önemli ölçüde değiştirebilmektedir. Bu çevresel değişkenlerin modelde dikkate alınması, gerçek zamanlı meteorolojik kesintilerin modellenmesine olanak tanımakta ve değişken deniz koşulları altında teslim süresi tahminlerinin doğruluğunu artırmaktadır. Böylece, daha sağlam ve güvenilir denizcilik tahminlerinin yapılabilmesi mümkün hâle gelmektedir.

Gemiye ait teknik özelliklerin, zamansal ve mekânsal verilerin ve çevresel faktörlerin kapsamlı şekilde entegre edilmesiyle oluşturulan bu veri seti, deniz taşımacılığında teslim süresi tahmini için sağlam bir analitik temel sunmaktadır. Her bir özellik kategorisi, gerçek dünya nakliye operasyonlarının teknik spesifikasyonlar, seyir davranışları, rota karakteristikleri ve meteorolojik koşullar gibi çok boyutlu karmaşıklıklarını özgün bir biçimde yansıtmakta ve böylece modelin bütüncül bir tahmin kapasitesine ulaşmasını mümkün kılmaktadır.

Tez çalışması kapsamında, yanıt değişkeni olarak daha önce giriş bölümlerinde tanımlandığı şekilde gemi teslim süresi kullanılmıştır. Teslim süresi, gemilerin kalkış noktasından varış noktasına ulaşması için gereken toplam süre olarak tanımlanmıştır. Buna göre, saat cinsinden teslim süresi, denklem (3.17) ile hesaplanmıştır:

$$Hesaplanan \text{ teslim süresi (saat)} = 24 * (Varış\ Zamanı\ (gün) - Kalkış\ Zamanı\ (gün)) \quad (3.17)$$

Teslim süresi tahmini için geliştirilen modelde kullanılan girdi özelliklerinin her birinin daha iyi şekilde anlaşılması için, Çizelge 3.2'de özelliklerin tanımları verilmiştir.



Şekil 3.7. Teslim süresi (saat) tahmin yaklaşımı akış şeması

Çizelge 3.2. Teslim süresi (saat) tahmin modeli girdi özelliklerinin tanımları

Sınıflandırma	Özellik	Açıklama
Genel Gemi Özellikleri	Gemi Alt Türü	Belirli yapısal, işlevsel veya operasyonel özelliklerle ayrıştırılan, daha geniş bir gemi kategorisi içinde yer alan ikincil sınıflandırmadır.
	MMSI	AIS gibi deniz navigasyon sistemlerinde, bir gemiye dijital haberleşme ve takip amacıyla atanan benzersiz dokuz haneli sayısal tanımlayıcıdır (International Telecommunication Union, 2016).
	IMO	Gemi tanımlama şemasına göre seyreden bir gemiye kalıcı olarak atanan benzersiz yedi haneli tanımlayıcıdır (International Maritime Organization, 1987).
	DWT (Ton)	Bir geminin ton cinsinden ifade edilen toplam kaldırma kapasitesi, 2.240 pound'tur. Bu değer hafif deplasman ile yüklü deplasman arasındaki farktır. Kargo DWT, toplam DWT üzerinden yakıt, su, yük, ekip, yolcular ve yolculukta gerekli diğer öğelerin ağırlığının çıkarılmasıyla belirlenmektedir (Monroe ve Stewart, 2005).
	Uzunluk (m)	Bir geminin baş tarafının en ön noktasından kıç tarafının en son noktasına kadar olan doğrusal mesafedir (Monroe ve Stewart, 2005).
	Genişlik (m)	Geminin en geniş noktasında ölçülen maksimum yatay genişliğidir ve metre cinsinden ifade edilir. Bu değer geminin dengesini, rıhtıma yanaşma ve kanal geçiş uygunluğunu etkilemektedir (Ghosh, 2022).
	Su çekimi (Draft)	Bir geminin yüzdüğü su derinliği, basitçe geminin tabanından su hattına olan mesafedir (Monroe ve Stewart, 2005).

Çizelge 3.2. Teslim süresi (saat) tahmin modeli girdi özelliklerinin tanımları (Devamı)

Sınıflandırma	Özellik	Açıklama
Zamansal Özellikler	Ortalama hızı	Belirli bir yolculuk boyunca bir geminin ortalama hızı, toplam kat edilen mesafenin toplam zamana (seyir duraklamaları veya operasyonel duraklamalar dahil) bölünmesiyle hesaplanmaktadır.
	Maksimum hızı	Belirli bir yolculuk sırasında, geçerli koşullar altında bir geminin ulaştığı en yüksek kaydedilmiş hızdır.
	Kalkış zamanı	Bir geminin yolculuğuna başladığı, tüm bağlama halatlarının ayrılması veya bir rıhtımdan, demir yerinden veya limandan tahrik gücünün başlatılmasıyla işaretlenen kayıtlı anıdır.
	Varış zamanı	Bir gemi, seferini bir rıhtımda bağlama halatlarını bağlayarak, varış noktasında demir atarak veya yerel makamlarca belirlenen liman sınırlarına girerek tamamlar.
Mekânsal Özellikler	Enlem	Meridyenindeki kuzey veya güney konumunun ekvatordan derece cinsinden ölçülen açısal uzaklığıdır (25 derece kuzey enlemindeki bir gemi) (Monroe ve Stewart, 2005).
	Boylam	Standart Greenwich meridyeninin doğusunda veya batısında ekvator üzerindeki konumun 180 derece doğu veya batıya kadar açısal uzaklığıdır (Monroe ve Stewart, 2005).
	Sefer mesafesi	Bir geminin kalkış noktası (liman, demirleme yeri veya ara nokta) ile varış noktası arasında kat ettiği toplam seyir uzunluğu, genellikle deniz mili (NM: Nautical Miles) cinsinden ölçülüdür.
Deniz Yolu Özellikleri	Rüzgâr Hızı	Rüzgar vektörünün yatay bileşeninin büyüklüğünü ifade eder ve havanın Dünya yüzeyi boyunca ne kadar hızlı hareket ettiğini göstermektedir (Wallace ve Hobbs, 2006).
	Rüzgâr Açısı	Rüzgarın geldiği pusula yönünü ifade eder. Genellikle gerçek kuzey referans alınarak saat yönünde derece cinsinden ölçülür (Hobbs ve Wallace, 2006).

3.2.2.2. Veri ön işleme

Yukarıda açıklanan tüm özelliklerin tahmin modellerinde girdi olarak kullanılması planlanmamıştır. Örneğin, MMSI numarası, veri setinde yalnızca geminin kimliğini temsil eden bir tanımlayıcı olarak yer almakta olup, modelleme sürecine doğrudan dâhil edilmemektedir.

Ayrıca, veri setinde rüzgârla ilgili bilgiler —örneğin rüzgâr hızı ve yönü— ayrı sütunlarda yer almaktadır. 271050337 MMSI numaralı gemi için bu değerler 8 Kn/ GB (243°) şeklinde kaydedilmiştir. Rüzgârın hız ve yön etkilerini birbirinden bağımsız olarak değerlendirebilmek amacıyla, bu veriler ayrıştırılmıştır.

Teslim süresinin doğru şekilde hesaplanabilmesi için, gerçek kalkış ve gerçek varış zamanlarına ait veri tiplerinin ayrıştırılması gerekmiştir. Örneğin, 271050337 MMSI numaralı gemi için bu zaman bilgileri, kalkışta 2025-05-03 12:43 (UTC+3) ve varışta 2025-05-05 07:54 (UTC+3) olarak kaydedilmiştir. Yapılan ayrıştırma işlemi sonucunda bu zamanlar, sırasıyla 3/5/2025 09:43 ve 5/5/2025 04:54 biçiminde dönüştürülmüştür.

Bu dönüşüm, zaman farklarının hassasiyetle hesaplanabilmesini sağlamak amacıyla Python programlama dili kullanılarak gerçekleştirilmiştir; zira gerçek kalkış ve varış zamanları, teslim süresi üzerinde doğrudan etkili olan kritik değişkenlerdir. Bu temelde, varış zamanından kalkış zamanının çıkarılmasıyla teslim süresi saat cinsinden hesaplanmış ve böylece her bir gemi seferi için gerçek teslim süresi değeri elde edilmiştir.

Veri setinde, henüz tamamlanmamış olan bazı seyir kayıtlarının da yer aldığı tespit edilmiştir. Ancak, tez çalışmasının kapsamı yalnızca rotasını tamamlamış gemilere ait gerçek geçmiş verilerle sınırlandırıldığından, tamamlanmamış seyirler filtrelenerek analiz dışında bırakılmıştır.

Buna ek olarak, teslim süresi değişkeninin birkaç farklı faktörden etkilenebilecek yapıda olması nedeniyle, veri kalitesinin korunması amacıyla IMO veya MMSI numarası tanımlanmamış gemilere ait kayıtlar da veri setinden çıkarılmıştır.

3.2.2.3. Eksik veri işleme

Eksik veriler söz konusu olduğunda, her bir özellik ayrı ayrı ele alınmıştır. Örneğin, maksimum hız için eksik olan değerler, ilgili seferin ortalama hız değeri ile eşitlenmiştir. Gemi özelliklerine ait eksik veriler (yapım yılı, uzunluk, genişlik ve DWT kapasitesi gibi) ise, açık erişime sahip tanınmış gemi navigasyon web sitelerinden temin edilerek tamamlanmıştır.

3.2.2.4. Modelleme ve değerlendirme

Modelleme aşamasında, değişkenler arasındaki ölçek farklarının modelin doğruluğunu olumsuz etkilemesini önlemek amacıyla tüm sayısal özellikler üzerinde normalizasyon işlemi gerçekleştirilmiştir. Bu işlem için, her bir değişkenin ortalamasını sıfır ve standart sapmasını bir olacak şekilde dönüştüren StandardScaler yöntemi kullanılmıştır. Böylece, özellikle mesafeye duyarlı algoritmaların daha tutarlı sonuçlar üretmesi sağlanmıştır.

Veri seti, modelin eğitilmesi ve test edilmesi amacıyla rastgele şekilde %80 eğitim ve %20 test olacak biçimde iki alt kümeye ayrılmıştır. Eğitim verisi modelin öğrenme sürecinde kullanılırken, test verisi modele daha önce gösterilmemiş örneklerle performansının değerlendirilmesi için ayrılmıştır.

Modellerin başarımını ölçmek adına çeşitli regresyon değerlendirme metrikleri kullanılmıştır. Bu kapsamda; MAE, RMSE, MAPE ve R^2 değerleri dikkate alınmıştır. Bu metrikler sayesinde farklı algoritmaların teslim süresi tahminlerindeki doğrulukları sayısal olarak karşılaştırılabilir hale getirilmiştir.

Ayrıca, her bir modelin performansını en üst düzeye çıkarmak için Grid Search yöntemi ile hiperparametre optimizasyonu gerçekleştirilmiştir. Bu süreçte, algoritmalara özgü parametre kombinasyonları sistematik olarak test edilerek en iyi sonuçları veren ayarlar belirlenmiştir.

3.2.2.5. Modellerin doğrulanması

Geliştirilen tahmin modellerinin doğruluğunu ve farklı veri kümeleri üzerindeki tutarlılığını değerlendirmek amacıyla kapsamlı bir doğrulama yaklaşımı benimsenmiştir. İlk aşamada, veri seti eğitim ve test alt kümelerine ayrılarak, modellerin öğrenme sürecinde aşırı uyum ya da yetersiz öğrenme eğilimi gösterip göstermediği karşılaştırmalı olarak analiz edilmiştir.

Modelin genellenebilirliğini daha güçlü biçimde test edebilmek adına, doğrulama sürecine k-katlı çapraz doğrulama (k-fold cross-validation) yöntemi de dâhil edilmiştir. Bu çalışmada k=5 olarak seçilmiş ve veri seti beş eşit parçaya bölünmüştür. Her döngüde dört alt küme eğitim için, bir alt küme ise doğrulama için kullanılarak modelin farklı veri bölümlerindeki performansı sistematik biçimde ölçülmüştür. Bu yöntem sayesinde, modele özgü değil, veri kümesine genellenebilir yapılar oluşturulması hedeflenmiştir.

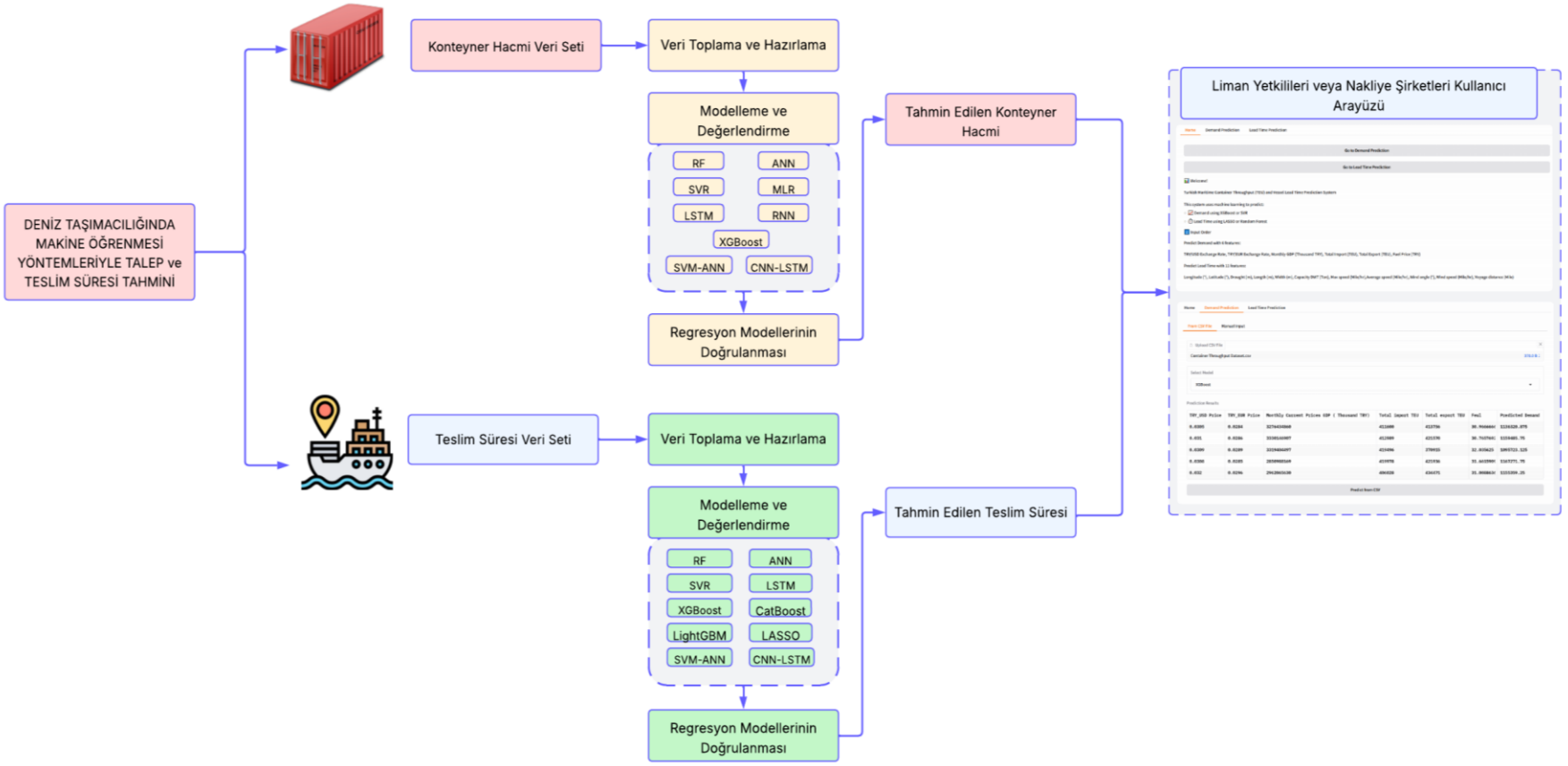
Ayrıca, doğrulama süreci yalnızca çapraz doğrulama ile sınırlı tutulmamış; modellerin gerçek kullanım senaryolarında nasıl bir performans sergileyeceğini değerlendirmek amacıyla, daha önce eğitim sürecinde yer almamış ve tamamen yeni doğrulama verileri üzerinde de test edilmiştir. Bu adım, modellerin görülmemiş veriler karşısındaki tahmin başarısını ölçerek, gerçek dünya uygulamaları için uygunluklarını ortaya koymuştur.

3.3. Konteyner İş Hacmi ve Teslim Süresi için Önerilen Tahmin Yaklaşımı

Tez çalışması, deniz lojistiği operasyonlarında konteyner iş hacmi ve teslim süresi tahminine yönelik önerilen bir kestirim çerçevesi sunmaktadır. Bu çerçeve, özellik bağımlılıklarını ve operasyonel dinamikleri etkili bir şekilde yakalayabilen gelişmiş ML modellerini kapsamaktadır. Modeller, tahmin doğruluğunu artırmak amacıyla çeşitli özelliklerle zenginleştirilmiş gerçek geçmiş sevkiyat verileri üzerinde eğitilmiştir.

Eğitim ve doğrulama sürecini takiben, tahmin modelleri pratik kullanımı kolaylaştırmak amacıyla kullanıcı dostu bir arayüz aracılığıyla uygulanmaktadır. Önerilen arayüz, armatörler, liman otoriteleri ve lojistik planlayıcılar gibi paydaşların, temel operasyonel

parametreleri analiz edebilmelerine ve konteyner iş hacmi ile beklenen teslim süresi için anlık tahminleri uygulayabilmelerine olanak tanımaktadır. Karmaşık model çıktılarının anlaşılır içgörülere dönüştürülmesi sayesinde, önerilen sistem denizcilik tedarik zinciri genelinde daha bilinçli ve zamanında karar alma süreçlerini desteklemektedir. Şekil 3.8'de konteyner iş hacmi ve teslim süresi önerilen tahmin yaklaşımına ait akış şeması sunulmaktadır.



Şekil 3.8. Konteyner iş hacmi (TEU) ve teslim süresi (saat) tahmini için önerilen yaklaşıma ait akış şeması

4. BULGULAR ve TARTIŞMA

Bu bölümde, gerçekleştirilen ML yöntemleri neticesinde elde edilen sonuçların detaylı analizi yapılarak ulaşılan bulgular sistematik bir şekilde ortaya konulmuştur. Tez çalışmasında geliştirilen makine öğrenmesi modelleri Python programlama dili kullanılarak kodlanmıştır. Modelleme süreçlerinde Jupyter Notebook üzerinde çalışılmış, veri işleme, model eğitimi ve değerlendirme aşamalarında ağırlıklı olarak pandas, numpy, scikit-learn, xgboost, keras ve tensorflow gibi açık kaynaklı Python kütüphaneleri kullanılmıştır.

4.1. Konteyner İş Hacmi Talebinin Tahmini için Geliştirilen Modellerin Sonuçları

Bu bölümde, konteyner iş hacmi talebinin tahmin edilmesi amacıyla geliştirilen ML modellerinin sonuçları detaylı bir şekilde sunulmaktadır. Farklı algoritmaların performansları, önceden belirlenen doğruluk ölçütleri (MAE, RMSE, R^2 ve MAPE) kullanılarak karşılaştırmalı olarak analiz edilmiştir.

4.1.1. İstatistiksel analiz

Konteyner iş hacmi talebi analizinde kullanılan temel değişkenler Çizelge 4.1'de sunulmakta olup, veri setinin operasyonel ve ekonomik dinamikleri ortaya konmaktadır. Veri setinin istatiksel analizi, bağımlı değişken olan konteyner iş hacmi, ortalama 1 022 943 TEU ve 80 422 standart sapma değeri ile zaman içinde orta düzeyde dalgalanma gösterdiğine işaret etmektedir. Kayıtlı değerler, 839 568 TEU'dan başlayıp en yüksek kapasite olan 1 170 670 TEU'ya kadar uzanmakta; bu da 330 000 TEU'nin üzerinde bir farkla deniz taşımacılığı faaliyetlerinde belirgin bir değişkenliğe işaret etmektedir.

Veri setindeki bağımsız değişkenlerden biri olan döviz kurları (TRY/USD ve TRY/EUR), önemli ölçüde dalgalanma göstermektedir. TRY/USD, 0,0305 ile 0,1936 arasında; TRY/EUR ise 0,0284 ile 0,1691 arasında değişmektedir. Bu tür kur dalgalanmalarının, ticaret maliyetleri ve sevkiyat talebi üzerinde etkili olabileceği değerlendirilmektedir.

Öte yandan, veri setinde yer alan MCP-GDP, TI ve TE gibi daha geniş ölçekli ekonomik göstergeler, konteyner akışlarını etkileyen makroekonomik güçleri yansıtmaktadır. MCP-GDP değişkeni ortalama 1,15 milyar Türk Lirası seviyesindeyken, en yüksek değeri 3,3 milyarı aşmakta ve bu da ekonomik faaliyetin yoğunlaştığı dönemleri göstermektedir.

Yakıt fiyatları ise MVFP değişkeni ile temsil edilmekte olup, 3,23 TRY ile 32,04 TRY arasında değişerek talep değişkenliğine katkı sağlamaktadır. Yakıtın taşımacılık maliyetleri üzerindeki doğrudan etkisi göz önünde bulundurulduğunda, bu dalgalanmaların ticaret lojistiğinde kritik bir rol oynadığı düşünülmektedir.

İstatistiksel göstergeler birlikte ele alındığında, temel değişkenlerin ölçeği ve dağılımı nicel olarak ortaya konmakta; bu sayede daha derinlemesine modelleme çalışmaları için zemin hazırlanmaktadır. Söz konusu desenlerin belirlenmesiyle birlikte, talep değişimlerinin daha isabetli şekilde öngörülmesi ve analitik yaklaşımın daha hassaslaştırılması mümkün hale gelmektedir.

Çizelge 4.1. Veri setinde yer alan bağımsız değişkenlerin istatistiksel analizi.

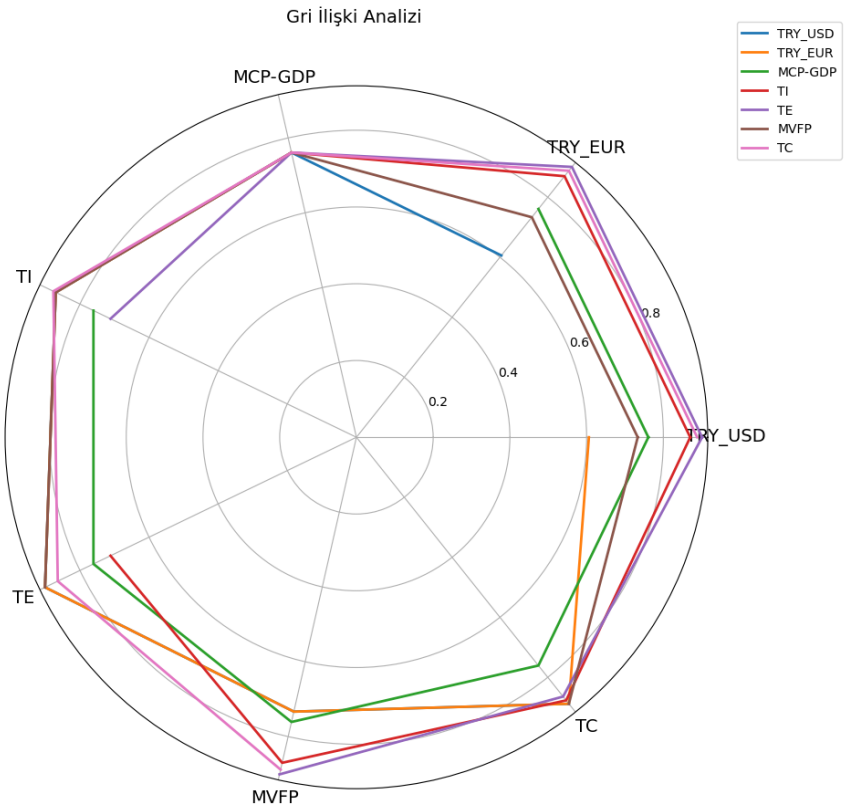
Açıklama	TRY_USD	TRY_EUR	MCP-GDP	TI	TE	MVFP
Veri Sayısı	72	72	72	72	72	72
Ortalama	0,1000	0,0888	1154483000	391665	393197	12,70
Std	0,0537	0,0467	937852500	26601	27768	9,46
Min	0,0305	0,0284	296626100	310495	330628	3,22
Maks	0,1936	0,1691	3330147000	450379	454785	32,03

4.1.2. Gri ilişki analizi

Tez çalışmasında, birden fazla değişken ile toplam konteyner iş hacmi arasındaki ilişkileri değerlendirmek amacıyla Gri İlişki Analizi (GRA: Grey Relational (Correlation) Analysis) uygulanmıştır. Analiz sonuçları Şekil 4.1'de paylaşıldığı üzere, her ekseni farklı bir özelliği temsil eden örümcek grafiği ile görselleştirilmiştir. Grafikte çizgilerin dış kenarlara yakınlığı, ilgili özelliğin toplam konteyner iş hacmi ile olan ilişki düzeyini yansıtmaktadır. Analiz sonuçlarına göre, Toplam İthalat (TEU) ile Toplam İhracat (TEU)

arasında güçlü bir korelasyon tespit edilmiştir; bu durum, ticaret faaliyetleri arasında yüksek düzeyde bir karşılıklı bağımlılık olduğunu göstermektedir. Ayrıca, MCP-GDP göstergelerinin ticaret hacmi ile yüksek derecede ilişkili olduğu belirlenmiştir. Bu da ekonomik faktörlerin deniz ticareti üzerindeki etkisini ortaya koymaktadır.

Yakıt fiyatları ile konteyner iş hacmi arasında ise orta düzeyde bir ilişki bulunmuştur. Bu da yüksek yakıt maliyetlerinin deniz taşımacılığı faaliyetlerini dolaylı olarak etkileyebileceğini düşündürmektedir. Buna karşılık, döviz kurları (TRY/USD, TRY/EUR) ile konteyner iş hacmi arasında ters yönlü bir ilişki gözlemlenmiştir. Bu durum, döviz kurlarındaki dalgalanmaların lojistik hizmetlerin sağlanması ve performansı üzerinde doğrudan etkili olmayabileceğini, ancak maliyetler üzerinde etkili olabileceğini göstermektedir. Elde edilen bulgular, makroekonomik faktörler ile deniz ticareti arasındaki etkileşimi açıklamakta ve daha güçlü tahmine dayalı modelleme ve karar verme süreçlerine duyulan gereksinimi desteklemektedir.



Şekil 4.1. Konteyner iş hacmi tahmini veri seti için gri ilişki analizi

4.1.3. Makine öğrenmesi modellerinin sonuçları

Bu bölümde, geleneksel tahmin modellerinin etkinliği, hibrit modellerin entegrasyonu yoluyla değerlendirilmiştir. Analiz kapsamında dokuz ML modeli incelenmiştir: RF, XGBoost, SVR, MLR, RNN, LSTM, ANN, CNN-LSTM ve SVM-ANN. Bu modeller, iki farklı senaryo çerçevesinde uygulanmıştır.

- Senaryo 1'de, tüm mevcut değişkenler herhangi bir azaltma işlemi uygulanmaksızın kullanılmıştır.
- Senaryo 2'de ise, TRY_USD, TRY_EUR ve MCP-GDP değişkenleri veri setinden çıkarılarak, kalan değişkenler ile model performansı değerlendirilmiştir.

Her bir modelin konteyner iş hacmi talebini tahmin etmedeki başarısını en üst düzeye çıkarmak amacıyla hiperparametre ayarlamaları gerçekleştirilmiştir. Çizelge 4.2'de XGBoost, LSTM ve ANN modelleri için elde edilen en uygun parametre değerleri özetlenmektedir.

Örneğin, XGBoost modeli, colsample_bytree parametresi (0,8), öğrenme oranı (0,09), maksimum derinlik 6 ve 150 tahminleyici ile birlikte L1 ve L2 düzenlileştirme değerlerinin 1 olarak ayarlanmasıyla en iyi performansa ulaşmıştır. Bu optimal değerler, dengeli özellik seçimi ve aşırı öğrenme riskinin en aza indirilmesini sağlamıştır.

Benzer şekilde, LSTM modeli, iki katmanlı bir mimariyle yapılandırılmış, 0,05 öğrenme oranı, 0,1 bırakma (dropout) oranı, 32 partı boyutu (batch size) ve 50 eğitim dönemi (epoch) kullanılarak sıralı verilerdeki uzun vadeli bağımlılıklar etkili bir şekilde yakalanmıştır.

ANN modeli ise, biraz daha derin bir üç katmanlı mimari ile tasarlanmış, 0,07 öğrenme oranı ve 0,3 bırakma oranı kullanılarak özellik çıkarımının iyileştirilmesi hedeflenmiştir. Model, 32 partı boyutu ve 50 eğitim dönemi ile tutarlılık sağlayacak şekilde eğitilmiştir.

Çizelge 4.2. XGBoost, LSTM ve ANN modellerinin hiper parametreleri

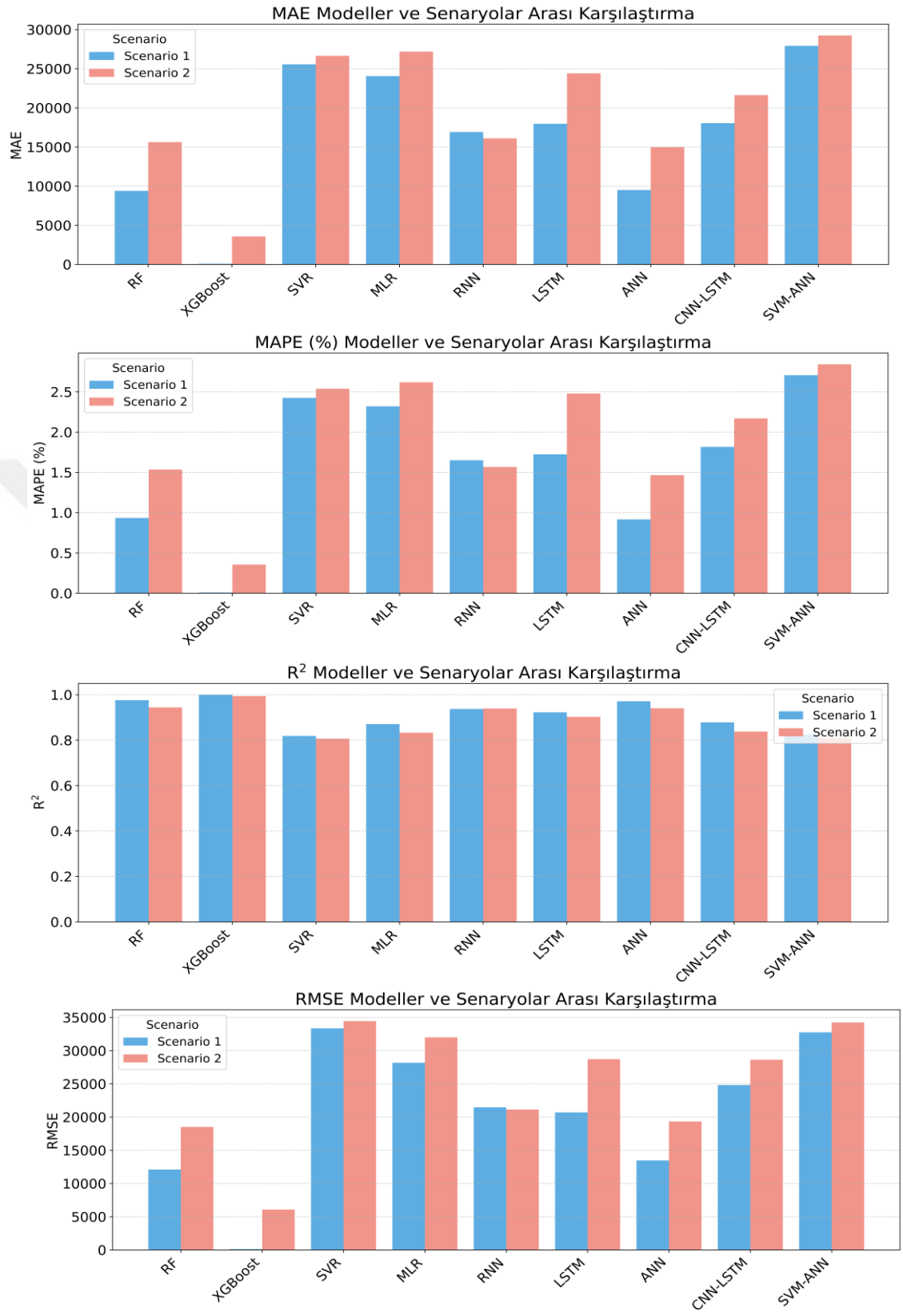
XGBoost		LSTM		ANN	
Parametre	En İyi Değer	Parametre	En İyi Değer	Parametre	En İyi Değer
Colsample_bytree	0,8	Num_layers	2	Num_layers	3
Learning_rate	0,09	Learning_rate	0,05	Learning rate	0,07
Max_depth	6	Dropout_rate	0,1	Dropout_rate	0,3
N_estimators	150	Batch_size	32	Batch_size	32
Reg_alpha	1	Epochs	50	Epochs	50
Reg_lambda	1	Neurons_num	64	Neurons_num	64
Subsample	0,8	Activation_fun.	Relu	Activation_fun.	Relu

Şekil 4.2'de her iki senaryo için eğitim veri setine ait performans parametrelerinin sonuçları gösterilmektedir. Eğitim veri seti sonuçlarına göre, LSTM, XGBoost ve ANN modelleri, hata oranları açısından diğer modellere kıyasla daha iyi sıralamalar elde etmiştir.

Senaryo 1 kapsamında, en iyi performans LSTM modeli tarafından sergilenmiş olup, bu sonuç, modelin sıralı veri örüntülerini öğrenme konusundaki başarısını ortaya koymaktadır. Senaryo 2'de ise, daha az sayıda özellikle çalışılmasına rağmen XGBoost modeli daha yüksek sınıflandırma doğruluğu elde etmiştir. Ayrıca, ANN modeli, her iki senaryoda da göreli olarak istikrarlı sonuçlar vermiştir.

Daha basit yapıya sahip modeller, örneğin SVR ve MLR, daha yüksek hata oranları ile çalışmıştır; bu da bu modellerin karmaşık veri örüntülerini modellemedeki yetersizliğini göstermektedir.

Hibrit modelin performansı ise, farklı yapıların birleştirilmesinin potansiyel taşıdığını; ancak optimum sonuçlara ulaşmak için dikkatli hiperparametre ayarlarının gerekli olduğunu ortaya koymaktadır.



Şekil 4.2. Senaryo 1 ve senaryo 2 için eğitim veri setine ait performans parametrelerinin sonuçları

Modellerin test veri seti için performansı Çizelge 4.3'te Senaryo 1 için, Çizelge 4.4'te ise Senaryo 2 için sunulmuş olup, modeller arasındaki performans farklılıkları daha ayrıntılı şekilde ortaya konmuştur.

Çizelge 4.3. Modellerin Senaryo 1 için test veri seti performans karşılaştırması

Model	MAE	RMSE	(R^2)	MAPE (%)
RF	31584,756	40844,678	0,7686	3,2116
XGBoost	20897,545	27454,675	0,8054	1,9992
SVR	23526,924	28931,301	0,8839	2,2731
MLR	24782,973	28943,350	0,8838	2,4409
RNN	22181,784	27465,397	0,6344	2,2255
LSTM	21144,634	26229,507	0,9045	2,1166
ANN	21676,592	26780,447	0,8643	2,0994
CNN-LSTM	46673,822	54926,718	0,6157	4,6453
SVM-ANN	25003,625	30532,836	0,8707	2,4634

Senaryo 1'in değerlendirilmesi, test edilen dokuz model arasında belirgin performans kalıpları ortaya koymaktadır. LSTM modeli, tüm temel metriklerde olağanüstü performans sergileyerek açık bir lider olarak öne çıkmaktadır. 0,9046'lik üstün R^2 değeri ile en yüksek açıklayıcı güce ulaşırken, aynı zamanda tüm modeller arasında en düşük RMSE (26229,50720) değerini korumaktadır. Bu kombinasyon, LSTM'in hem varyansı açıklama hem de tahmin hatalarını minimize etme konusundaki üstün yeteneğini göstermektedir.

XGBoost, nokta tahminlerinde kayda değer kesinlik göstermekte, en düşük MAE (20897,544643) ve MAPE (% 1,99922) değerlerini kaydederek minimal mutlak hata gerektiren uygulamalar için önemini ortaya koymaktadır. SVR ve MLR modelleri güçlü bir ikinci kademe oluşturmakta, her ikisi de 0,88'in üzerinde R^2 değerlerine ulaşmakta, ancak LSTM ve XGBoost'a kıyasla biraz daha yüksek hata metrikleri sergilemektedir.

ANN modeli dengeli metriklerle (R^2: 0,864, MAPE: % 2,099) sağlam performans sunmaktadır, RF ise yeterli ancak sıra dışı olmayan sonuçlar göstermektedir. Dikkat

çekici biçimde, CNN-LSTM tüm metriklerde, özellikle hata ölçümlerinde (MAE: 46673,822, RMSE: 54926,718) önemli ölçüde zorlanmakta, bu da bu mimarinin söz konusu tahmin görevi için uygun olmayabileceğini düşündürmektedir.

Çizelge 4.4. Modellerin Senaryo 2 için test veri seti performans karşılaştırması

Model	MAE	RMSE	(R^2)	MAPE (%)
RF	30833,428	38189,728	0,7977	3,135
XGBoost	21259,826	27267,160	0,8550	2,134
SVR	26125,002	31417,758	0,8631	2,534
MLR	27059,880	32013,347	0,8578	2,641
RNN	23946,587	29445,836	0,5798	2,401
LSTM	24695,710	29722,336	0,8775	2,436
ANN	26298,628	31712,459	0,8040	2,581
CNN-LSTM	56437,870	61063,154	0,5250	5,654
SVM-ANN	27718,857	33042,445	0,8485	2,744

Senaryo 2'de, rekabet ortamı bazı tutarlı kalıpları korurken kısmen değişiklik göstermektedir. XGBoost, en düşük MAE (21259,825), RMSE (27267,1597) ve MAPE (% 2,134) değerlerini elde ederek sağlamlığını kanıtlamakta ve hata minimizasyonu için en iyi performans gösterici olarak kendini konumlandırmaktadır. Bununla birlikte, LSTM açıklayıcı güçteki üstünlüğünü en yüksek R^2 (0,8775) ile korumakta, ancak XGBoost'a kıyasla biraz yüksek hata metrikleri sergilemektedir.

SVR modeli güçlü R^2 değeri (0,8631) ve makul hata metrikleri ile iyi performans göstermeye devam etmektedir. MLR ise yüksek R^2 değeri (0,8578) sağlamakla birlikte daha yüksek tahmin hataları sergilemektedir. Bu senaryoda en iyi ve en kötü performans göstericiler arasındaki fark genişlemiş, CNN-LSTM daha da belirgin bir şekilde düşük performans göstermiştir (MAE: 56437, 8702, R^2: 0,5250).

Dokuz ML modelinin iki farklı senaryo kapsamında gerçekleştirilen kapsamlı değerlendirmesi, LSTM'nin en dengeli model olarak öne çıktığını ortaya koymaktadır.

Model, en yüksek doğruluk değerine (Senaryo 1'de R^2: 0,9046) ulaşırken güçlü hata metrikleri ve yorumlanabilirlik gerektiren uygulamalar için özellikle uygun hale gelmektedir. XGBoost, tahmin hatalarını minimize etmede olağanüstü kesinlik sergilemekte, özellikle en düşük MAE, RMSE ve MAPE değerlerini kaydettiği Senaryo 2'de saf tahmin doğruluğu için optimal seçenek olarak kendini konumlandırmaktadır. SVR ve MLR tutarlı biçimde güvenilir performans sunmakta, güçlü R^2 değerleri ve makul hata oranları ile sağlam alternatifler sunmaktadır. Buna karşın, CNN-LSTM her iki senaryoda da önemli ölçüde düşük performans göstermekte, en yüksek hataları ve en düşük açıklayıcı gücü sergileyerek bu görevler için uygun olmadığını düşündürmektedir. Sonuçlar, senaryoya özgü model seçiminin önemini vurgulamakta; Senaryo 1 için LSTM, Senaryo 2 için XGBoost ve çok yönlü bir seçenek olarak SVR önerilmektedir. Bu bulgular, farklı uygulamaların gerektirdiği doğruluk ve yorumlanabilirlik dengesine dayalı model dağıtımı için değerli rehberlik sağlamaktadır.

4.1.4. Modellerin daha önce görmediği verilerin tahmini

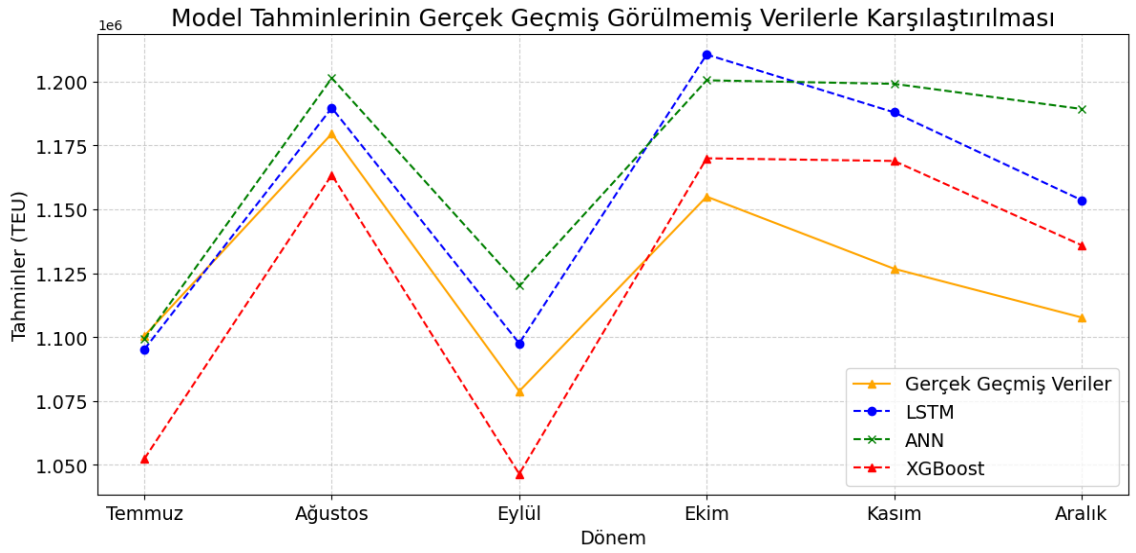
Tahmin modellerinin geçerliliğini değerlendirmek amacıyla, Temmuz – Aralık 2024 dönemine ait daha önce kullanılmamış gerçek geçmiş bir veri kümesi kullanılmıştır. Çizelge 4.5, Senaryo 1 kapsamında, geliştirilen XGBoost, LSTM ve ANN modellerinden elde edilen tahmin sonuçlarını, görülmemiş gerçek geçmiş veri kümesi ile karşılaştırmalı olarak sunmaktadır.

LSTM modeli, altı aydan üçünde (Temmuz, Eylül ve Kasım) en isabetli tahminleri üretmiş ve bu yönüyle en başarılı model olarak öne çıkmıştır. XGBoost modeli, Ekim ve Aralık aylarında en iyi sonuçları verirken, ANN modeli yalnızca Ağustos ayında üstünlük göstermiştir.

Bu bulgular doğrultusunda, gelecekteki konteyner iş hacimlerinin tahmini için LSTM modelinin birincil öneri olarak kullanılmasının uygun olacağı; XGBoost modelinin ise belirli senaryolarda alternatif bir seçenek olarak değerlendirilebileceği önerilmektedir. Şekil 4.3'te model tahminlerinin modellerin daha önce görmediği gerçek geçmiş veri seti ile karşılaştırılması gösterilmektedir.

Çizelge 4.5. Daha önce görülmemiş veri seti için en iyi performans gösteren modellerin tahmin sonuçları

Dönem	Gerçek Geçmiş Veriler	XGBoost	LSTM	ANN
Aralık/2024	1107691	1135951	1137915	1189231
Kasım/2024	1126736	1168896	1166708	1199057
Ekim/2024	1154943	1169935	1193425	1200445
Eylül/2024	1078841	1046485	1091836	1087006
Ağustos/2024	1179563	1163446	1172268	1173788
Temmuz/2024	1100466	1052387	1074900	1072342

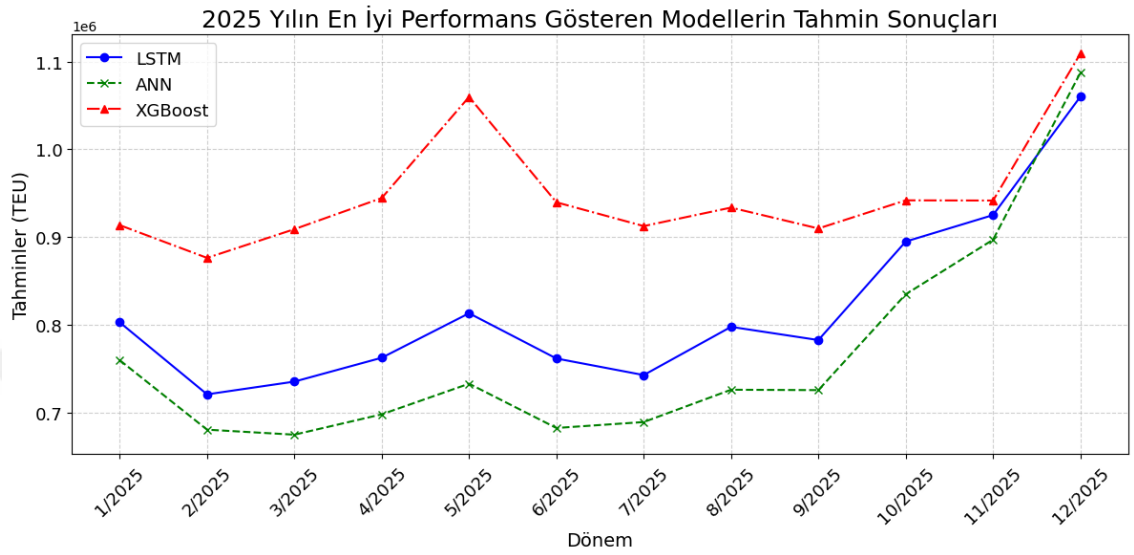


Şekil 4.3. Model tahminlerinin daha önce görülmemiş veri seti ile karşılaştırılması

4.1.5. 2025 yılı konteyner iş hacmi talep tahmini

İleriye dönük olarak, 2025 yılına yönelik kapsamlı bir tahmin süreci, ML yaklaşımları kullanılarak uygulanmıştır. Tahmin sonuçları, Şekil 4.4'te görsel olarak sunulmaktadır. Zaman serisi verilerindeki trend ve mevsimsellik bileşenlerini modelleyebilme yeteneği nedeniyle, ilk olarak Holt-Winters üssel düzeltme yöntemi kullanılarak birinci senaryo kapsamındaki girdi değişkenlerinin 2025 yılı değerlerinin tahmini elde edilmiştir.

Ardından LSTM, ANN ve XGBoost modelleri ile 2025 yılı için konteyner iş hacmi talebi tahmin edilmiştir.



Şekil 4.4. 2025 yılı için en iyi performans gösteren modellerin tahmin sonuçları

Şekil 4.4, 2025 yılına yönelik olarak LSTM, ANN ve XGBoost modellerinden elde edilen tahmin sonuçlarının, konteyner iş hacminde genel bir yukarı yönlü eğilimi başarıyla yakaladığını ve yıl sonuna doğru belirgin zirvelerin ortaya çıktığını göstermektedir.

Bu tahminlere dayanarak çeşitli ekonomik senaryolar türetilebilmekte; bu ilişkilerin detaylı biçimde incelenmesi, deniz lojistiği üzerindeki olası etkiler açısından geleceğe dair daha derin içgörülerin elde edilmesine olanak tanımaktadır.

Özellikle Türk Lirası (TRY) ile USD veya EUR arasındaki döviz kuru dalgalanmaları, ticaret rekabetçiliği üzerinde önemli etkilere sahiptir. Liranın değer kaybettiği durumlarda, Türk malları yabancı alıcılar için daha ucuz hâle gelmekte, bu da ihracat talebinde artışa ve ihracat hacminin yükselmesine neden olabilmektedir. Ancak bu durum, ithalat maliyetlerini artırmakta ve gelen konteyner iş hacminde azalma riskini beraberinde getirmektedir.

Eğer Türk Lirası'nın değer kaybı eğilimi sürerse, ihracat konteynerlerinde artış, buna karşılık ithalat konteynerlerinde azalma gözlemlenebilir; bu da konteyner sirkülasyonunda dengesizlik yaratabilecek bir senaryo ortaya koymaktadır.

Toplam ithalat ve ihracat verileri, ticaret hacmindeki mevsimsel dalgalanmaları ve büyüme eğilimlerini yansıtmaktadır. İthalat ve ihracatta sürdürülebilir artışlar, konteyner taşımacılığına yönelik talep değişimlerini işaret etmekte ve lojistik kapasitenin artırılması gerekliliğini vurgulamaktadır. Ancak, döviz kuru dalgalanmaları veya finansal belirsizlikler gibi faktörlerle tetiklenebilecek ticaret faaliyetlerindeki düşüşler, konteyner iş hacimlerini olumsuz etkileyebilir ve bu bağlamda denizcilik alanında dinamik planlamaya duyulan ihtiyacı öne çıkmaktadır.

Genel olarak, yakıt fiyatlarındaki artışlar, taşımacılık maliyetlerini yükselterek ticaret faaliyetlerini baskılayabilir veya taşıma yöntemlerinde değişiklik gerektirebilir. Yakıt maliyetlerinin artmaya devam etmesi durumunda, bu durum yalnızca Türkiye'nin denizcilik politikalarını etkilemekle kalmayacaktır, aynı zamanda tahmin modellerinin yeni ticaret desenlerine uyum sağlayacak şekilde yeniden yapılandırılmasını da zorunlu kılacaktır.

MCP-GDP eğilimi ise makroekonomik bir gösterge olarak genel ticaret hacmini etkilemektedir. Ticaret büyümesi, MCP-GDP'nın genişlemesine bağlı olarak artmakta ve bu da konteyner liman kapasitesine olan talebin yükselmesine neden olmaktadır. Buna karşılık, MCP-GDP büyümesinde bir yavaşlama ya da düşüş yaşanması durumunda, tüketici talebi ve ticaret akışları zayıflayabilir, bu da konteyner trafiğinde yavaşlamaya neden olmaktadır. Tüm bu değerlendirmeler ışığında, proaktif tahminleme ve stratejik planlama, liman verimliliğinin korunması ve tedarik zinciri dayanıklılığının artırılması açısından hayati önem taşımaktadır.

Bu bulgular, ekonomik dalgalanmalara erken uyarı sinyalleri verebilecek, AI destekli tahmin sistemlerine duyulan ihtiyacı açıkça ortaya koymaktadır. Gerçek zamanlı veri entegrasyonu ve gelişmiş tahmin doğruluğu sayesinde, bu modeller lojistik ekiplerin daha hazırlıklı hareket etmesini ve planlamalarını daha etkili yapabilmesini sağlayacaktır.

Böylece, Türkiye'nin denizcilik sektörünün belirsizlik dönemlerinde dahi rekabet gücünü korumasına katkı sunacaktır.

4.2.Teslim Süresinin Tahmini için Geliştirilen Modellerin Sonuçları

Bu bölümde, deniz taşımacılığı operasyonlarında kritik bir öneme sahip olan teslim süresi tahmini için geliştirilen makine öğrenmesi modellerinin deneysel sonuçları sistematik bir şekilde sunulmaktadır. Çalışma kapsamında, teslim süresini etkileyen operasyonel ve çevresel faktörler dikkate alınarak çeşitli tahmin modelleri geliştirilmiş ve bu modellerin performansları kapsamlı olarak değerlendirilmiştir.

4.2.1. İstatistiksel analiz

İstatistiksel analiz, anlamlı içgörülerin elde edilmesi ve varsayımların doğrulanması için temel bir zemin sunmaktadır. Bu sürecin önemli bir yönü, regresyon varsayımlarının olası ihlallerinin teşhis edilmesini içermektedir. Özellikle değişen varyans gibi ihlaller, tahminlerin güvenilirliğini tehdit edebileceğinden, bu tür sorunların tanımlanması ve ele alınması, elde edilen bulguların geçerliliğini artırmakta ve modellerin tahmin doğruluğunu geliştirmektedir.

Değişen varyans, regresyon modellerinde hata terimlerinin varyansının sabit olduğu varsayımını ihlal etmektedir. Daha basit bir ifadeyle, artıkların saçılımı veya değişkenliği, tahmin edilen sonuca ya da bağımsız değişkenlere bağlı olarak sistematik biçimde değişmektedir (Amemiya, 1977).

Değişen varyans teşhis edilmesinde en yaygın kullanılan yöntemlerden biri olan Breusch–Pagan testi (BP testi) uygulanabilir. Bu test, hata terimlerinin varyansının açıklayıcı değişkenlere bağlı olup olmadığını inceleyerek, regresyon modellerinde değişen varyansı tespit etmeye yarayan istatistiksel bir testtir (Amemiya, 1977).

Denklemler (4.1)- (4.3)'de verilen regresyon modeli için, hata teriminin (ϵ_i) kareleri ile bağımsız değişkenler $X_{1i}, \ldots, X_{pi}$ arasında doğrusal bir ilişki olup olmadığına bakılan bir

denklem formu kullanılmaktadır. Burada, Y_i bağımlı değişkeni $X_{1i}, \dots, X_{pi}$ bağımsız değişkenleri temsil etmektedir.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \epsilon_i \quad (4.1)$$

$$\epsilon_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \cdots + \alpha_p X_{pi} + u_i \quad (4.2)$$

$$BP = n \, . \, R_{aux}^2 \sim \chi^2 \, (p) \quad (4.3)$$

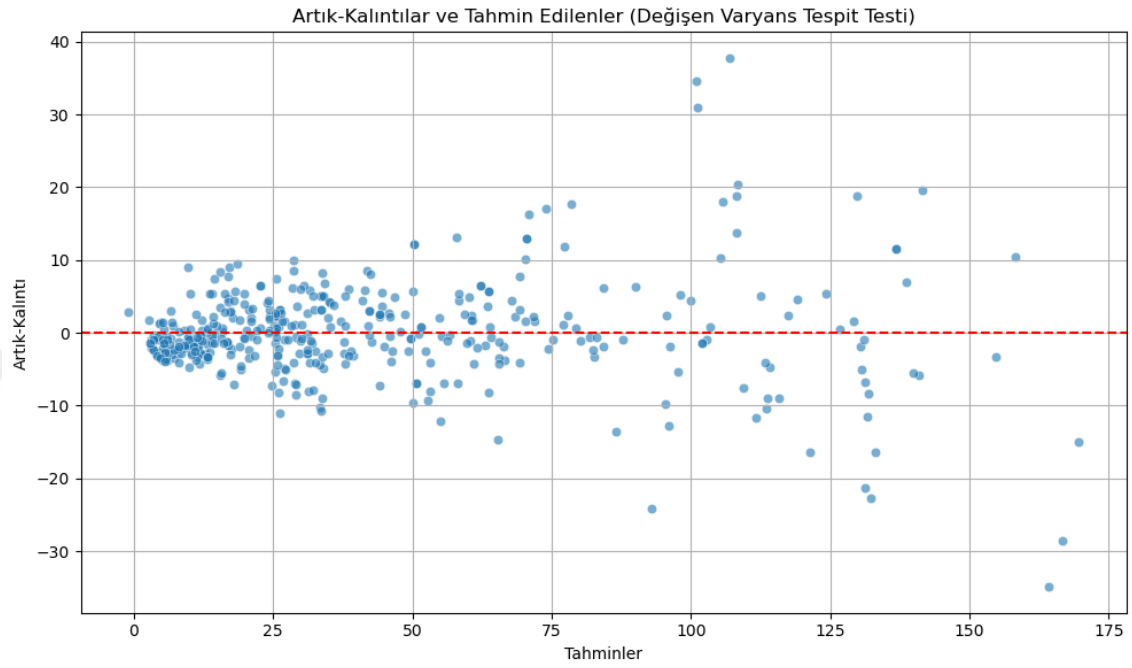
Test süreci, ilk olarak orijinal regresyon modelinin tahmin edilmesi ve elde edilen hata terimlerinin karelerinin hesaplanmasıyla başlatılmaktadır. Daha sonra, bu kareler, yardımcı bir regresyon modeli kapsamında orijinal bağımsız değişkenlere karşı regresyona tabi tutulmaktadır. Test istatistiği (BP), örneklem büyüklüğü ile bu yardımcı regresyon modelinden elde edilen R^2 değerinin çarpımı olarak hesaplanmakta olup, değişen varyans sıfır hipotezi (H_0: $\alpha_1 = \alpha_2 = \dots = \alpha_p = 0$) altında ki-kare dağılımı izlemektedir. Anlamlı bir test sonucu, hata varyansının bağımsız değişkenlere göre değiştiğini ve dolayısıyla değişen varyans varsayımının ihlal edildiğini göstermektedir.

BP testi, değişen varyansın doğrusal biçimlerinin tespiti açısından etkili olup, uygulanması oldukça basit bir yöntem olarak literatürde yaygın şekilde tercih edilmektedir (Astivia ve Columbia, 2019).

Tez çalışmasında, teslim süresi veri seti üzerinden değişen varyans tanısı koymak amacıyla Python kodlaması kullanılarak BP testi uygulanmıştır. Elde edilen sonuçlar Çizelge 4.6'da sunulmuştur. Test sonucuna göre, BP F-istatistiği 12,707 ve p-değeri 0,0003 olarak hesaplanmıştır. P-değeri <0,05 olduğu için, sabit varyans varsayımını ifade eden sıfır hipotezi reddedilmiştir. Bu durum, model artıklarının sabit varyansa sahip olmadığını, yani tahmin hatalarının tahmin aralığı boyunca değişkenlik gösterdiğini ortaya koymaktadır.

Şekil 4.5'te sunulan grafik, farklı aralıklarda değişkenliğin sabit olmadığını ortaya koyarak değişen varyans varlığını desteklemektedir. Veri setini ve tahmin performansını

iyileştirmek için logaritmik dönüşüm gibi çeşitli çözümler uygulanabilmektedir. Ancak, bu tür dönüşümlerin belirli bir keyfîlik taşıdığı belirtilmiştir (Astivia ve Columbia, 2019). Bu bağlamda, önerilen yöntemlerden biri de veri setinin katmanlara ayrılmasıdır.



Şekil 4.5. Artık-Kalıntılar (Residuals) ve tahmin edilenler değişen varyans tespit testi

BP F-testi sonuçları, ana veri setinde anlamlı düzeyde değişen varyans bulunduğunu (F = 12,707, p = 0,003) göstermektedir. F-istatistiği, hata terimlerinin karelerinin orijinal bağımsız değişkenlere karşı regresyona tabi tutulması ve açıklanan ile açıklanamayan varyansın karşılaştırılması yoluyla elde edilmiştir. Buna karşılık, ana veri seti düşük ve yüksek akış olarak alt kümelemeye tabi tutulduğunda, düşük akışlı veri alt kümesinde (F = 1,3522, p = 0,1706) ve yüksek akışlı veri alt kümesinde (F = 0,6847, p = 0,7945) anlamlı bir değişen varyans tespit edilmemiştir.

Bu bulgu, veri setinin alt kümelere ayrılmasının artık terimlerin davranışında daha istikrarlı bir yapı sağladığını ve model varsayımlarının daha sağlıklı bir şekilde karşılandığını doğrulamaktadır.

Çizelge 4.6, veri seti yüksek ve düşük teslim süresi aralıkları şeklinde ikiye ayrılarak ayrı modeller eğitilmeden önce ve sonra uygulanan değişen varyans test sonuçlarını göstermektedir. Bu yaklaşım, modelleme sürecinde daha homojen varyans dağılımlarına ulaşılmasını amaçlamaktadır.

Çizelge 4.6. Katmanlaşma öncesi ve sonrası değişen varyans testi sonuçları

Veri Seti	MAE	RMSE	R^2	F-istatistik	P-değeri
Ana Veri Seti	4,4128	6,7589	0,9681	12,707	0,003
Düşük Akış Veri Seti	1,9571	2,6671	0,9631	1,3522	0,1706
Yüksek Akış Veri Seti	5,1226	6,6942	0,9673	0,6847	0,7945

Çizelge 4.7'de sunulan çeşitli oranlar test edilerek sistematik bir katmanlaştırma yaklaşımı uygulanmıştır. En uygun katmanlaştırma oranının belirlenebilmesi amacıyla, MAE, RMSE ve R^2 gibi temel performans metrikleri kullanılarak titiz bir değerlendirme gerçekleştirilmiştir.

Çizelge 4.7'de yüksek akışlı veri seti üzerinde farklı bölme oranları ile ANN modelinin test sonuçları sunulmaktadır. Bu analiz, en etkili veri bölme stratejisinin belirlenmesini amaçlamaktadır. Elde edilen sonuçlar, %67–%33'lük oranının en iyi performansı sağladığını göstermektedir; bu oranla MAE değeri 5,38, RMSE değeri 6,59, MAPE oranı % 6,01 ve oldukça yüksek bir R^2 değeri (0,9561) elde edilmiştir. Bu durum, modelin aşırı uyum ya da yetersiz öğrenme yaşamadan iyi bir genelleme yapabildiğine işaret etmektedir. Buna karşılık, %80–%20 ve %75–%25 bölme oranları daha yüksek hata metriklerine ve daha düşük R^2 değerlerine yol açmış, bu da söz konusu oranların model performansı açısından daha az etkili olduğunu ortaya koymuştur.

Çizelge 4.7. Farklı katmanlaştırma oranları için sonuçlar

Veri Seti Bölme Oranı		Yüksek Akış Performans Sonuçları			
Yüksek Akış	Düşük Akış	MAE	RMSE	R^2	MAPE%
50%	50%	5,883	9,411	0,930	8,027
67%	33%	5,379	6,589	0,956	6,006
75%	25%	7,766	11,231	0,879	8,362
80%	20%	10,946	14,286	0,741	10.417

Ana veri seti için geliştirilen ilk regresyon modelinde gözlemlenen değişen varyansı gidermek amacıyla, veri seti teslim süresi değerlerine göre iki alt kümeye ayrılarak katmanlaştırılmıştır: düşük akış ve yüksek akış. Katmanlaştırma sonrasında, düşük ve yüksek teslim süresi alt kümeleri için ayrı ayrı doğrusal regresyon modelleri eğitilmiştir. Kısa teslim sürelerini içeren düşük akış modeli, dikkate değer bir şekilde daha iyi bir doğruluk sağlamış; MAE değeri 1,96'ya, RMSE değeri ise 2,67'ye düşürülmüş ve R^2 = 0,963 ile güçlü bir açıklayıcılık korunmuştur. Bu model için yapılan BP testi, p-değeri = 0,171 ile anlamsız bulunmuş ve bu alt kümede değişen varyansın etkili şekilde giderildiğini göstermiştir.

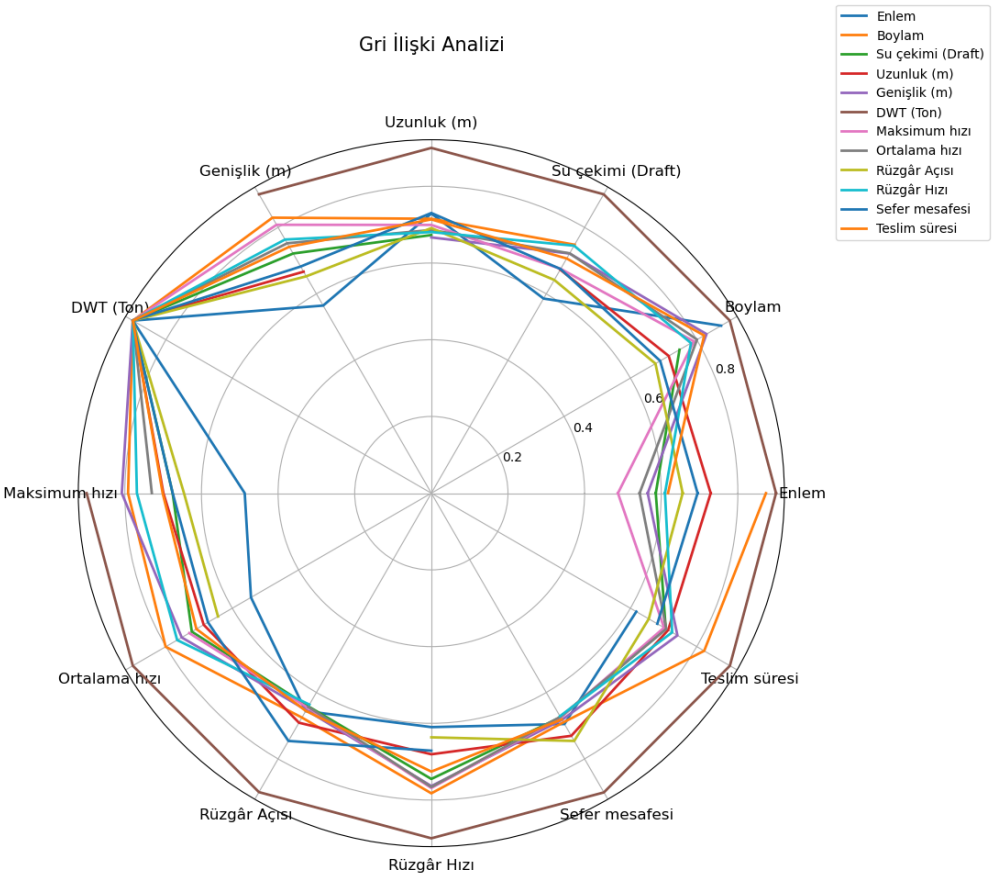
Benzer şekilde, yüksek akış modeli de R^2 = 0,967 ile benzer bir açıklayıcılığı sürdürmüş; RMSE = 6,69 düzeyinde kalmış, ancak MAE değeri 5,12 ile biraz daha yüksek bulunmuştur. Buna karşın, p-değeri = 0,795 olan BP testi, bu modelde de anlamlı bir değişen varyans olmadığını doğrulamıştır. Genel olarak, elde edilen bulgular, veri setini teslim süresine göre katmanlaştırmanın, tahmin doğruluğunu artırmak ve hata varyansını istikrara kavuşturmak için etkili bir yöntem olduğunu ortaya koymaktadır.

4.2.2. Gri ilişki analizi

Teslim süresi tahmini veri seti için gerçekleştirilen gri ilişki analizi Şekil 4.6'da verilmiştir. Teslim süresi ile en yüksek ilişkisel yakınlık gösteren değişken mil cinsinden toplam sefer mesafesidir. Bu değişkenin çizgisi çoğu boyutta teslim süresi eğrisinin şeklini ve büyüklüğünü yakından takip ederek teslim süresi üzerinde güçlü bir etkiye sahip olduğunu göstermektedir. Bu durum, mil cinsinden normalize edilmiş toplam sefer

mesafesinin bir sevkiyatın ne kadar süreceğini belirlemede merkezi bir rol oynadığını göstermektedir.

Ayrıca, Kapasite (DWT), Su Çekimi (draft) ve Uzunluk da teslim süresi deseni ile önemli ölçüde uyum göstererek, yükleme/boşaltma verimliliği, gemi kararlılığı ve izin verilen hızlar üzerindeki etkileri nedeniyle teslim süresini etkileyen önemli yapısal gemi faktörleri olduklarını göstermektedir. Rüzgâr Hızı ve Genişlik orta düzeyde uyum göstererek çevresel ve boyutsal etkilerin mevcut olduğuna ancak daha az baskın olduğuna işaret etmektedir. Bunun tersine, Enlem, Rüzgâr Açısı ve Ortalama Hız gibi değişkenler teslim süresi eğrisinden belirgin sapma göstererek düşük korelasyon veya muhtemelen dolaylı ilişkilere işaret etmektedir.



Şekil 4.6. Teslim süresi tahmini veri seti için gri ilişki analizi

4.2.3. Makine öğrenmesi modellerinin sonuçları

Bu bölümde, teslim süresini tahmin etmek amacıyla ikiye ayrılmış veri seti (Düşük Akış ve Yüksek Akış) üzerinde uygulanan çeşitli ML modellerinin karşılaştırmalı değerlendirmesi sunulmaktadır. Modeller, öncelikli olarak R^2 ve MAPE temelinde değerlendirilmiştir. Bu iki metrik birlikte ele alındığında, her bir modelin hem açıklayıcıyı gücünü hem de göreli tahmin doğruluğunu yansıtmaktadır.

4.2.3.1. Düşük akış veri seti için modellerin sonuçları

Modellerin düşük akış teslim süresi veri seti için eğitim ve test veri setleri bazında performans değerlendirme sonuçları Çizelge 4.8'de verilmiştir. Çizelge 4.8'de sunulan bulgulara göre, Düşük Akış veri setinde en iyi performansı sergileyen model LASSO-R modeli olmuştur.

Bu model, 0,9940'lık en yüksek test R^2 değeri ve % 5,19'luk en düşük MAPE skorlarından biriyle, yüksek doğruluk ve güçlü genelleme kapasitesi göstermiştir. RF modeli, 0,9936 R^2 ve % 5,53 MAPE değerleri ile tahmin tutarlılığı açısından LASSO-R'yi yakından takip etmiştir.

Diğer yüksek performans gösteren modeller arasında, XGBoost ve ANN yer almaktadır; bu modeller sırasıyla 0,9928 ve 0,9923 R^2 değerlerine ulaşmıştır. Ancak, bu modellerin %8,75 ve %5,66 oranlarındaki MAPE değerleri, her ne kadar yüksek varyans açıklama gücüne sahip olsalar da göreli hata oranlarının LASSO-R ve RF modellerine kıyasla bir miktar yüksek olduğunu göstermektedir. SVR-ANN hibrit modeli de 0,9921 R^2 ve %7,29 MAPE ile rekabetçi bir performans ortaya koymuştur.

Buna karşılık, CatBoost ve CNN-LSTM gibi modeller, eğitim aşamasında yüksek R^2 değerleri göstermiş olsalar da test aşamasındaki 0,9888 ve 0,9904 R^2 ile birlikte, % 9,88 ve % 10,30 MAPE skorlarıyla aşırı öğrenme belirtileri göstermiştir. Benzer şekilde, LSTM ve LightGBM modelleri, sırasıyla 0,9699 ve 0,9916 test R^2 değerleri ve % 7,92

MAPE ile, kabul edilebilir ancak LASSO-R ve RF modellerine göre daha az başarılı performans göstermiştir.

Bu sonuçlar, LASSO-R ve RF modellerinin, Düşük Akış operasyonel senaryolarında varyans açıklama gücü ile göreli tahmin hatası arasında en iyi dengeyi sunduğunu ortaya koymaktadır.

Çizelge 4.8. Düşük akış veri seti için modellerin performans değerlendirme sonuçları

Model	Veri Seti	MAE	RMSE	R^2	MAPE (%)
ANN	Eğitim	0,6191	0,950	0,995	3,792
	Test	0,8165	1,220	0,992	5,658
CatBoost	Eğitim	0,580	0,775	0,997	5,2553
	Test	1,044	1,478	0,988	9,8784
LSTM	Eğitim	0,959	1,307	0,991	7,3746
	Test	1,245	2,416	0,970	7,9186
RF	Eğitim	0,381	0,550	0,998	3,8241
	Test	0,692	1,117	0,994	5,5315
XGBoost	Eğitim	0,453	0,627	0,998	4,7369
	Test	0,789	1,190	0,993	8,7535
LightGBM	Eğitim	0,414	0,566	0,998	3,3667
	Test	0,862	1,281	0,992	7,9299
LASSO-R	Eğitim	0,692	1,034	0,995	5,1746
	Test	0,683	1,069	0,994	5,1950
SVR	Eğitim	0,763	1,106	0,994	5,5906
	Test	0,750	1,149	0,993	7,4626
SVR-ANN	Eğitim	0,662	0,962	0,995	4,4380
	Test	0,822	1,231	0,992	7,2938
CNN-LSTM	Eğitim	0,495	0,657	0,998	6,8189
	Test	1,156	1,525	0,990	10,2963

4.2.3.2. Yüksek akış veri seti için modellerin sonuçları

Modellerin yüksek akış teslim süresi veri seti için eğitim ve test veri setleri bazında performans değerlendirme sonuçları Çizelge 4.9'da verilmiştir. Çizelge 4.9'da yüksek akış veri setinde, LASSO-R modeli yine en iyi performans gösteren modellerden biri olarak öne çıkmıştır; 0,9595 test R^2 değeri ve %5,32 MAPE skoru ile, yüksek varyans koşullarında dahi güçlü tahmin doğruluğunu koruduğu görülmektedir. Benzer şekilde, SVR modeli de 0,9595 R^2 ve %5,72 MAPE değerleri ile neredeyse aynı düzeyde sonuç vermiştir ve yüksek hacimli teslim süresi değişkenliğini modellemede etkili olduğunu ortaya koymuştur. ANN modeli de rekabetçi sonuçlar sunmuş; 0,9561 R^2 ve % 6,01 MAPE değerleri ile dikkat çekmiştir.

Buna karşılık, SVR-ANN hibrit modeli, test R^2 değerinde 0,9415'e hafif bir düşüş göstermiş ve % 6,57 MAPE ile sonuçlanmıştır. RF modeli sıralamada biraz daha düşük olsa da 0,934 R^2 ve % 6,26 MAPE değerleriyle tatmin edici bir performans göstermiştir.

Bununla birlikte, CatBoost modeli, eğitim aşamasında neredeyse kusursuz bir 0,9999 R^2 değeri elde etmesine rağmen, test aşamasında 0,823 R^2 ve % 8,96 MAPE ile ciddi bir aşırı öğrenme sorunu göstermiştir.

Benzer şekilde, CNN-LSTM ve LightGBM modelleri de yüksek eğitim başarımına karşın zayıf genelleme yetenekleri sergilemiştir; bu modellerin test R^2 değerleri sırasıyla 0,9189 ve 0,8769, MAPE değerleri ise % 7,56 ve %8,32 olarak gözlemlenmiştir. XGBoost ise daha orta düzeyde bir performans göstermiş; 0,9248 test R^2 ve % 7,30 MAPE değerlerine ulaşmıştır.

Genel olarak, yüksek akış teslim süresi veri seti sonuçları, LASSO-R gibi doğrusal modellerin ve SVR gibi çekirdek-tabanlı modellerin, yüksek varyansa sahip karmaşık koşullarda daha kararlı ve genellenebilir olduklarını ortaya koymuştur. Buna karşılık, ağaç tabanlı güçlendirme veya derin öğrenme temelli modeller, eğitimde yüksek doğruluk göstermelerine rağmen, test aşamasında aşırı öğrenme eğilimi sergileyerek daha zayıf genelleme yeteneği göstermiştir.

Çizelge 4.9. Yüksek akış veri seti için modellerin performans değerlendirme ölçüm sonuçları

Model	Veri Seti	MAE	RMSE	R^2	MAPE (%)
ANN	Eğitim	5,029	6,705	0,9672	5,7488
	Test	5,3791	6,5894	0,9561	6,0058
CatBoost	Eğitim	1,0516	1,2978	0,9999	0,3022
	Test	8,5925	13,414	0,8228	8,955
LSTM	Eğitim	6,9888	9,6960	0,9315	7,9851
	Test	8,6632	10,4777	0,8890	9,5725
RF	Eğitim	2,6628	3,5194	0,9874	2,9791
	Test	6,0082	8,1946	0,9339	6,2618
XGBoost	Eğitim	1,1799	1,7618	0,9977	1,2441
	Test	6,7802	8,6259	0,9248	7,2963
LightGBM	Eğitim	1,6151	2,1512	0,9953	3,3667
	Test	7,8986	11,1808	0,8769	8,3213
LASSO-R	Eğitim	5,2708	7,1416	0,9628	5,8682
	Test	4,9524	6,3320	0,9595	5,3168
SVR	Eğitim	4,8426	6,9112	0,9652	5,4445
	Test	5,1436	6,3284	0,9595	5,7179
SVR-ANN	Eğitim	5,3747	7,2218	0,9620	5,7857
	Test	6,0259	7,6052	0,9415	6,5686
CNN-LSTM	Eğitim	2,5940	3,6611	0,9901	2,6040
	Test	7,1183	8,8780	0,9189	7,5642

4.2.4. Modellerin daha önce görmediği verilerin tahmini

En iyi performans gösteren modellerin—LASSO-R, RF, ANN, CatBoost, SVR ve XGBoost—tahmin yetenekleri, eğitim ve test aşamalarının ötesindeki genellenebilirliklerini değerlendirmek amacıyla, modellerin daha önce görmediği bir veri kümesi üzerinde ayrıca doğrulanmıştır. Bu bağlamda, gerçek geçmiş veriler, her bir modelin tahmin doğruluğunu kıyaslamak için referans olarak kullanılmıştır. Bu yöntemle,

modellerin yalnızca öğrenme süreçlerindeki başarıları değil, gerçek dünya koşullarına ne ölçüde uyum sağladıkları da ortaya konmuştur.

4.2.4.1. Düşük akış veri setinde modellerin daha önce görmediği verilerin tahmini

Geliştirilen modellerin genellenebilirlik yeteneklerini değerlendirmek amacıyla, bu modeller modellerin daha önce görmediği bir düşük akışlı veri kümesi üzerinde test edilmiştir. Bu veri kümesi, daha düşük yoğunluk ve değişkenliğe sahip konteyner akış desenlerini temsil etmekte olup, modellerin tahmin doğruluğu ve dayanıklılığı açısından gerçek geçmiş teslim süresi değerlerine karşı kıyaslama yapılmıştır.

Çizelge 4.10'da sunulan karşılaştırmalı analiz sonuçları, ANN, XGBoost ve CatBoost modellerinin, veri aralığının büyük kısmında gerçek geçmiş verilere göreceli olarak yakın tahminler ürettiğini göstermiştir. Özellikle ANN modeli, orta ve yüksek teslim süresi değerlerinde düşük hata payı ile istikrarlı sonuçlar sunarak, düşük hacimli denizcilik operasyonlarında doğrusal olmayan ilişkileri öğrenme kapasitesini ortaya koymuştur. Buna karşılık, LASSO-R modeli sürekli olarak teslim süresini fazla tahmin etmiş ve bu durum, modelin doğrusal yapısından kaynaklanan sınırlamalarını gözler önüne sermiştir.

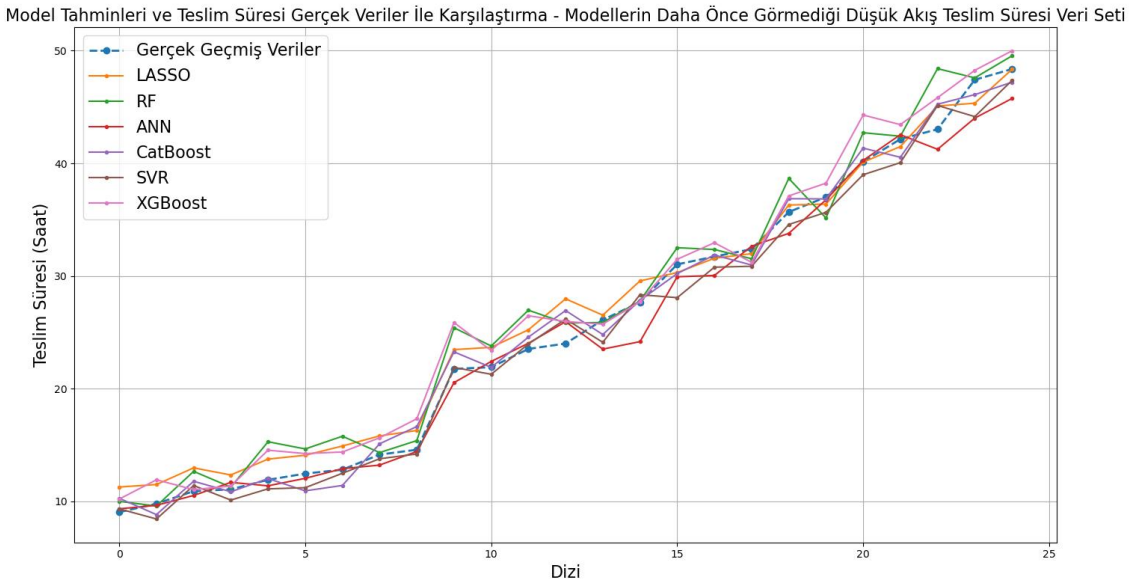
Veri kümesinin erken aşamalarında (~9 ila ~15 saat arası teslim sürelerinde), SVR ve ANN modelleri gerçek geçmiş eğilimi en iyi izleyen modeller olarak öne çıkarken, RF ve XGBoost bazı noktalarda sapma ve aşırı tahmin göstermiştir. Bununla birlikte, XGBoost ve CatBoost modelleri, 30 saat üzerindeki yüksek teslim sürelerinde tahmin doğruluğunu arttırarak, karmaşık ilişkileri modellemedeki güçlerini ortaya koymuştur.

Genel olarak tüm seri boyunca değerlendirildiğinde, XGBoost ve ANN modelleri, gerçek geçmiş değerlerden daha az sapma ile en dengeli tahmin performansını sergilemiştir; özellikle riskin yüksek olduğu üst uç teslim sürelerinde bu fark daha belirgin olmuştur. Bu bulgular, hesaplama maliyetine rağmen ANN modelinin, modellerin daha önce görmediği düşük yoğunluklu konteyner akışı senaryolarında teslim süresi tahmini için güçlü bir aday olduğunu göstermektedir. Modellerin daha önce görmediği tüm veri setine ait tahmin sonuçları düşük akış için EK 1'de yer almaktadır. Şekil 4.7'de modellerin daha

önce görmediği düşük akış teslim süresi veri seti için gerçek geçmiş veriler ile modellerin tahmin sonuçları karşılaştırmalı olarak gösterilmektedir.

Çizelge 4.10. En iyi performans gösteren modellerin daha önce görülmemiş düşük akış teslim süresi (saat) veri seti için tahmin sonuçları

Gerçek Geçmiş Veriler	LASSO-R	RF	ANN	CatBoost	SVR	XGBoost
9,03	11,26	9,99	9,32	10,26	9,32	10,19
14,57	16,28	15,38	14,42	16,63	14,19	17,32
21,75	23,46	25,40	20,53	23,26	21,88	25,86
32,38	31,98	31,53	32,64	30,97	30,86	31,21
48,37	48,34	49,54	45,76	47,20	47,35	49,99



Şekil 4.7. Modellerin daha önce görmediği düşük akış teslim süresi (saat) veri seti için gerçek veriler ile modellerin tahmin sonuçlarının karşılaştırılması

4.2.4.2. Yüksek akış veri setinde modellerin daha önce görmediği verilerin tahmini

Yüksek akışlı modellerin daha önce görmediği veri kümesi üzerinde yapılan model performansı değerlendirmesi, yoğun konteyner akışı ve yüksek teslim süresi değişkenliğinin hâkim olduğu koşullarda tahmin güvenilirliğini test etmeyi amaçlamıştır.

Bu veri kümesi, yüksek işlem hacmi ve operasyonel karmaşıklığın söz konusu olduğu gerçek dünya senaryolarını yansıtarak, doğru tahminlerin kaynak tahsisi ve risk azaltımı açısından kritik önem taşıdığı durumları temsil etmektedir.
Çizelge 4.11'de özetlenen sonuçlara göre, CatBoost ve XGBoost modellerinin yaklaşık 50 ila 90 saat aralığında, gerçek geçmiş teslim süreleriyle yüksek düzeyde uyumlu tahminler ürettiği ortaya konmuştur. Bu modellerin temelini oluşturan gradyan arttırmalı mimari, doğrusal olmayan ilişkileri ve özellikler arası etkileşimleri ileri düzeyde işleyebilme yeteneği sayesinde, yüksek hacimli denizcilik operasyonlarında ortaya çıkan karmaşık desenleri başarıyla modelleyebilmiştir.

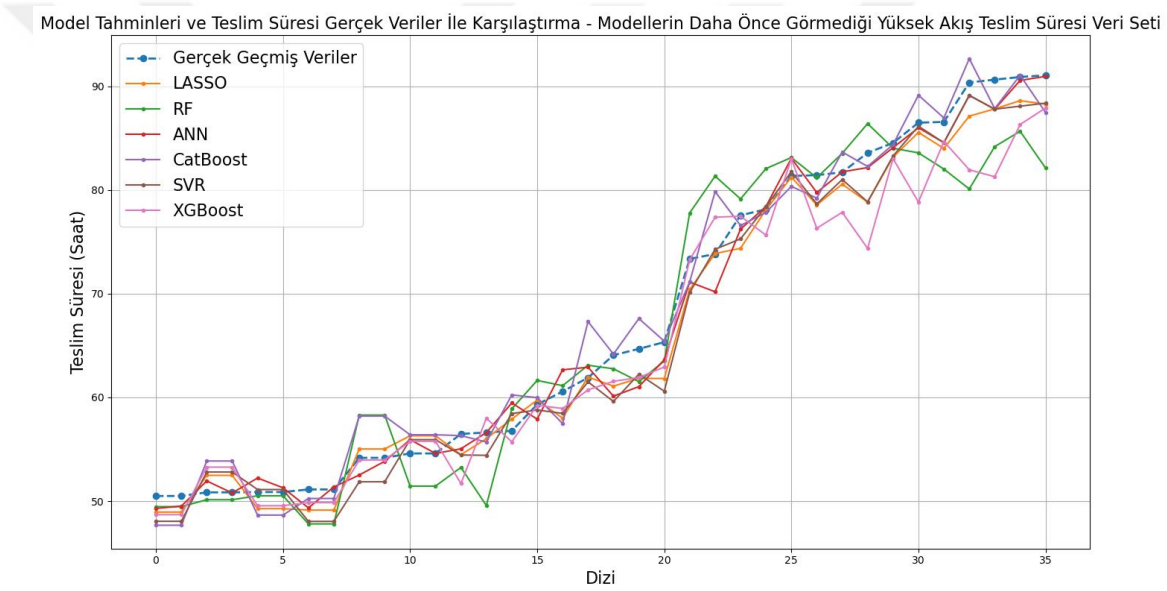
ANN modeli de özellikle orta teslim süresi aralığında (yaklaşık 60–80 saat) rekabetçi bir performans sergilemiş, ancak yüksek teslim sürelerine doğru zaman zaman fazla tahmin eğilimi göstermiştir.

Buna karşılık, LASSO-R modeli, özellikle teslim süresinin üst çeyreğinde belirgin sapmalar göstererek, doğrusal yapısının yüksek akışlı veri setindeki karmaşık örüntüleri modelleme konusundaki yetersizliğini ortaya koymuştur. SVR modeli ise orta düzeyde etkili olmakla birlikte, uç değerlerde tutarsızlık göstermiştir. RF modeli daha kararlı ancak genelde konservatif tahminler üretmiş; özellikle yüksek teslim sürelerinde aşırı genelleştirme eğilimi gözlenmiştir.

Bu değerlendirme, CatBoost ve XGBoost modellerinin, yüksek hacimli denizcilik senaryolarında üstün doğruluk ve dayanıklılık sunduğunu doğrulamaktadır. Aynı zamanda ANN, hesaplama kaynaklarının yeterli olduğu ve esneklik gerektiren durumlarda uygulanabilir bir alternatif olarak öne çıkmaktadır. Elde edilen bulgular, karmaşık ve veri yoğun lojistik ortamlarında teslim süresi tahmini için ileri düzey topluluk ve derin öğrenme modellerinin kullanılmasının önemini vurgulamaktadır. Modellerin daha önce görmediği tüm veri seti için elde edilen tahmin sonuçları EK 2'de bulunmaktadır. Şekil 4.8'de modellerin daha önce görmediği yüksek akış teslim süresi veri seti için gerçek geçmiş veriler ile modellerin tahmin sonuçları karşılaştırmalı olarak gösterilmektedir.

Çizelge 4.11. En iyi performans gösteren modellerin daha önce görülmemiş yüksek akış teslim süresi (saat) veri seti için tahmin sonuçları

Gerçek Geçmiş Veriler	LASSO-R	RF	ANN	CatBoost	SVR	XGBoost
50,50	48,95	49,47	49,30	47,68	48,06	48,70
61,90	61,96	63,10	62,91	67,31	61,50	60,72
73,35	70,40	77,78	71,12	71,14	70,11	73,27
81,42	78,55	81,17	79,75	79,21	78,65	76,30
91,05	88,27	82,11	90,94	87,46	88,36	77,90



Şekil 4.8. Modellerin daha önce görmediği yüksek akış teslim süresi (saat) veri seti için gerçek veriler ile modellerin tahmin sonuçlarının karşılaştırılması

4.3. Kullanıcı Dostu Arayüz

Türkiye'deki denizcilik lojistiği sektörü küresel ticarette hayati bir öneme sahip olup, büyük limanlar konteyner trafiği için kilit merkezler olarak hizmet vermektedir. Verimli talep tahmini ve gemi teslim süresi öngörüsü, liman operasyonlarının optimize edilmesi, yoğunluğun minimize edilmesi ve genel tedarik zinciri performansının artırılması açısından önemlidir.

Bu zorlukları ele almak amacıyla, iki temel metrik için doğru ve uygulanabilir tahminler sunan kullanıcı dostu bir arayüz (UI: User Interface) geliştirilmiştir: konteyner iş hacmi (TEU) ve gemi teslim süresi. Sistem, hem dosya yükleme (CSV) hem de manuel veri girişi seçeneklerini destekleyecek şekilde tasarlanmış olup, farklı teknik uzmanlık seviyelerine sahip kullanıcılar için esneklik sağlamaktadır.

Ayrıca, arayüz iki yüksek performanslı ML modelini bünyesinde barındırmakta ve kullanıcıların tahminleri karşılaştırarak en güvenilir seçeneği tercih etmelerine olanak tanımaktadır. Bir geri bildirim mekanizması, gerçek dünya sonuçlarından öğrenerek model doğruluğunu zamanla daha da artırmaktadır.

4.3.1. Kullanıcı arayüzünün temel özellikleri

4.3.1.1. Esnek veri girişi yöntemleri

Arayüz, farklı kullanıcı tercihlerine hitap etmek amacıyla çoklu veri girişi yöntemlerini desteklemektedir. Kullanıcılar, liman trafiği, gemi özellikleri, hava durumu koşulları ve ekonomik göstergelerin geçmiş kayıtlarını içeren yapılandırılmış veri dosyalarını (CSV: Comma-separated values) yükleyebilmektedir. Alternatif olarak, form tabanlı manuel girişi seçeneği, kullanıcıların hızlı ve isteğe bağlı tahminler için bireysel veri noktalarını girmelerine olanak tanımaktadır. Bu ikili giriş yaklaşımı hem veri analisti hem de operasyonel personel için erişilebilirlik sağlamaktadır.

4.3.1.2. ML tabanlı tahmin sistemi

Tahmin doğruluğunu maksimize etmek amacıyla, arayüz XGBoost, RF, SVR ve LASSO-R gibi optimize edilmiş ML modellerini entegre etmektedir. Kullanıcılar tercih ettikleri modeli seçebilmekte veya belirli kullanım durumları için en uygun seçeneği belirlemek amacıyla tüm modellerin çıktılarını karşılaştırabilmektedir. MAE, RMSE, R^2 ve MAPE dahil olmak üzere performans metrikleri, şeffaflık sağlamak ve kullanıcıların güvenilirliği değerlendirmelerine yardımcı olmak amacıyla görüntülenmektedir.

4.3.2. Kullanıcı arayüzünün tasarımı

Gradio kullanılarak oluşturulan arayüz, denizcilik lojistiği tahmini için optimize edilmiş web tabanlı bir ön yüz sağlamaktadır. Arayüze giriş yapıldığında, kullanıcılar iki ana girişi seçeneği ile karşılanmaktadır: geçmiş gemi ve konteyner özellik verilerinin toplu işlenmesi için sürükle-bırak dosya yükleme kutusu (CSV) ve gerekli her özellik için açılır menüler ve sayısal alanlar içeren manuel girişi paneli.

Veriler sağlandıktan sonra, kullanıcılar konteyner iş hacmi talebi veya gemi teslim süresi tahminleri oluşturmak için "Tahmin Et" düğmesine tıklamakta ve sonuçlar anında yapılandırılmış bir tablo ve özet bölümünde görüntülenmektedir.

Arayüz ayrıca, kullanıcıların üst düzey performans gösteren algoritmalara—RF, XGBoost, SVR ve LASSO-R—geçiş yapmalarına olanak tanıyan ve tahmin sonuçlarının gerçek zamanlı karşılaştırılmasını sağlayan bir model seçici özelliği de barındırmaktadır.

Genel tasarım, basitlik, duyarlılık ve netliği öne çıkararak, gelişmiş tahmin analitiğini minimal teknik uzmanlığa sahip liman operatörleri, nakliye hattı plancıları ve lojistik analistleri için erişilebilir kılmaktadır. Arayüzün temel tasarımı Ek 3'te, konteyner iş hacminin talebi için geliştirilen arayüz Ek 4'de ve teslim süresi için geliştirilen arayüz Ek 5'te bulunmaktadır.

4.3.3. Kullanıcı arayüzünün katkısı

Bu kullanıcı arayüzü, Türkiye'deki denizcilik lojistiği alanına çeşitli şekillerde katkı sağlamaktadır:

- Operasyonel Optimizasyon: Doğru, veri odaklı tahminler sunarak, liman operatörleri ve lojistik plancıları çizelgeleme işlemlerini optimize edebilmekte, gemi dönüş sürelerini azaltabilmekte ve rıhtım tahsisini iyileştirebilmektedir.

- Veri Erişilebilirliği: Yüklenen veri veya manuel girişin kullanılması, geçmiş veri tabanlarına sahip büyük nakliye işletmeleri ile gelişmiş veri altyapısına sahip olmayan küçük operatörler veya karar vericiler için aracı erişilebilir kılmaktadır.

- Karar Desteği: Gerçek zamanlı tahminler, yöneticilerin kaynak tahsisi, rota planlama ve acil durum müdahale yönetimi gibi bilinçli stratejik ve taktik kararlar almasını sağlamaktadır.

- Eğitimsel Değer: Araştırmacılar, analistler ve lojistik stajyerleri için modeller arasındaki geçiş, farklı algoritmaların aynı veri girişleri ile nasıl performans gösterdiğini gözlemlemelerine olanak tanıyan bir öğrenme fırsatı sunmaktadır.

Bununla birlikte bu arayüz, Türkiye'deki denizcilik ve lojistik ekosistemindeki çeşitli paydaşlara fayda sağlayacak konumdadır:

- Liman Otoriteleri: Liman kapasitesini optimize etmek, yoğunluğu azaltmak ve liman genişleme stratejilerini desteklemek amacıyla kullanılabilecektir.
- Nakliye Şirketleri: Daha iyi filo çizelgelemesi ve gecikmelerin minimize edilmesi için kullanılabilecektir.
- Lojistik Hizmet Sağlayıcıları: Tedarik zinciri koordinasyonunu ve teslimat doğruluğunu iyileştirmek amacıyla kullanılabilecektir.
- Devlet Kurumları ve Planlama Uzmanları: Stratejik altyapı planlama, sürdürülebilirlik öngörüsü ve kriz yönetimi için kullanılabilecektir.
- Akademik ve Araştırma Kurumları: Denizcilik tahmin uzmanları, lojistikte yapay zekâ ve veri bilimi uygulamaları alanlarında uygulamalı araştırma ve öğretim için kullanabilecektir.

Gelişmiş tahmin modellerini basit ancak güçlü bir arayüz aracılığıyla erişilebilir kılarak, bu çözüm, Türk denizcilik lojistiğinin dijital dönüşümüne katkı sağlamakta, verimlilik, şeffaflık ve veri odaklı yeniliği teşvik etmektedir.

5. SONUÇ

Bu çalışma kapsamında, konteyner iş hacmi talebi ve teslim sürelerinin tahmin edilmesine yönelik olarak ML modelleri kullanılarak geniş kapsamlı bir analiz gerçekleştirilmiştir.

Çalışmanın ilk aşamasında, temel ekonomik ve operasyonel göstergelere ilişkin istatistiksel çözümler yapılmış ve bu değişkenlerin konteyner iş hacmi faaliyetleri üzerindeki etkileri değerlendirilmiştir. GRA aracılığıyla, toplam konteyner hareketleri ile makroekonomik parametreler arasındaki ilişkiler sayısal olarak ortaya konmuş; ithalat ve ihracat hacimleri ile MCP-GDP göstergelerinin belirleyici etkenler olduğu saptanmıştır.

Modelleme sürecinde hem geleneksel hem de hibrit ML yaklaşımları uygulanmış; modellerin performansı farklı senaryolar altında karşılaştırmalı olarak incelenmiştir. Senaryo 1'de tüm değişkenler kullanılmış, Senaryo 2'de ise özellik seçimi yoluyla daha sade veri yapısına ulaşılmıştır. Her iki senaryoda hiperparametre ayarlamalarıyla optimize edilen modellerin, özellikle LSTM, XGBoost ve ANN gibi ileri seviye algoritmalar aracılığıyla başarılı tahminler ürettiği ortaya konmuştur.

Senaryo 1 dahilinde en yüksek performans LSTM modeli tarafından sergilenmiş olup, model, zaman serilerindeki uzun vadeli bağlılıkları öğrenme kapasitesiyle öne çıkmıştır. Senaryo 2'de ise daha az sayıda özellik kullanılmasına rağmen, XGBoost modeli en iyi genel performansı ortaya koymuştur. Bu bulgu, özellik seçiminin model hassasiyetine katkı sağladığını ve bazı durumlarda veri boyutunun azaltılmasının genel başarıya zarar vermediğini göstermektedir.

Teslim süresine yönelik modellerin değerlendirilmesinde, veri seti teslim sürelerine göre "Düşük Akış" ve "Yüksek Akış" olmak üzere iki alt küme şeklinde katmanlandırılmış ve bu yöntemin, model varsayımlarının daha sağlıklı karşılanmasına katkı sağladığı görülmüştür. BP testleri ile desteklenen bu yaklaşım, değişen varyans etkisinin veri alt kümelemesi yoluyla azaltılabileceğini ortaya koymuştur. Ayrıca, katmanlandırılmış veri setleri üzerinde eğitilen modeller, daha homojen hata dağılımı sergilemiş ve genelleme başarısı yüksek olmuştur.

Düşük akış senaryosunda, LASSO-R ve RF modellerinin en düşük hata oranları ve en yüksek açıklama gücü ile öne çıktığı; buna karşılık CNN-LSTM ve CatBoost gibi daha karmaşık modellerin test verisi üzerinde aşırı öğrenme gösterdiği tespit edilmiştir. Yüksek akış senaryosunda da benzer bir durum gözlemlenmiş; LASSO-R, SVR ve ANN gibi modellerin daha stabil ve tutarlı olduğu, buna karşılık derin öğrenme tabanlı mimarilerin çoğunlukla test verilerinde daha düşük genelleme kabiliyeti sunduğu anlaşılmıştır.

Daha önce görülmemiş veri kümesi üzerinde yapılan testler, LSTM ve ANN modellerinin özellikle yüksek akış teslim sürelerinde çarpıcı başarı gösterdiğini, LASSO-R modelinin ise sıklıkla tahmin değerini aşırı yüksek belirlediğini ortaya koymuştur. Bu durum, modellerin yapısal farklılıklarının belirli teslim süreleri aralıklarında öngörü performansını önemli ölçüde etkilediğini göstermektedir.

2025 yılına yönelik öngörüler ise genel olarak konteyner iş hacmi talebinde yükselen bir trendin sürmekte olduğuna işaret etmektedir. Bu artışın, döviz kuru oynaklıkları, ekonomik genleşme ve yakıt fiyatlarındaki değişkenliğe duyarlı olduğu belirlenmiştir. Özellikle, ihracatı teşvik eden kur seviyeleri, konteyner hareketliliğini olumlu etkilerken; ithalat maliyetlerini arttırarak toplam konteyner dengesini bozabilecek riskler taşımaktadır.

Sonuç olarak, bu çalışma, deniz lojistiği alanında talep ve teslim sürelerinin tahminine yönelik veri odaklı karar verme süreçlerinin geliştirilmesi adına katkı sağlamaktadır. Uygulanan ML modelleri, operasyonel planlamanın iyileştirilmesi, liman verimliliğinin arttırılması ve stratejik öngörülerin desteklenmesi noktasında potansiyel sunmaktadır. Katmanlandırılmış veri yapıları ve hibrit model yaklaşımları sayesinde, belirsizliklere karşı daha dirençli ve esnek bir tahmin altyapısı geliştirilmiştir.

Türkiye'nin deniz lojistiği sektörüne odaklanan ilk kapsamlı akademik çalışma olma niteliğini taşıyan bu araştırma, sektöre yönelik talep odaklı ve teslim süresi tahminine dair veri temelli bir metodolojik çerçeve sunarak literatürdeki önemli bir boşluğu doldurmaktadır. Mevcut çalışmalar çoğunlukla uluslararası düzeyde rota optimizasyonu,

maliyet analizi veya genel tedarik zinciri performansına odaklanmakta; ancak Türkiye'nin özgün lojistik yapısına uygun, teslim süresi öngörüsüne yönelik ampirik modellerin eksikliği dikkat çekmektedir. Bu bağlamda sunulan model, yalnızca yerel sektör paydaşlarına karar destek sağlamayı hedeflemekle kalmayıp, aynı zamanda benzer yapısal özellikler taşıyan gelişmekte olan pazarlara da uygulanabilir bir çerçeve sunmaktadır.

Bununla birlikte, gelecekteki çalışmalar için bazı iyileştirme alanlarını da gündeme getirmektedir. Özellikle, daha yüksek öngörü doğruluğu ve uygulama geçerliliği sağlamak adına, veri setlerinin kapsamı ve ayrıntı düzeyinin artırılması, sipariş alımından varış limanı operasyonlarına kadar tüm lojistik yaşam döngüsünü kapsayacak biçimde zamansal segmentasyonun dahil edilmesi önerilmektedir. Ayrıca, demiryolu/karayolu gecikmeleri, gümrük işlem süreleri ve aktarma darboğazları gibi intermodal taşımacılık değişkenlerinin entegrasyonu, tedarik zinciri boyunca oluşabilecek sistemik gecikmelerin modellenmesini mümkün kılacaktır. Bu tür genişletmeler yalnızca model dayanıklılığını artırmakla kalmayacak, aynı zamanda lojistik ağın her bir düğümündeki verimsizliklerin tanımlanmasına da olanak tanıyacaktır.

KAYNAKLAR

Amemiya, T. (1977). A note on a heteroscedastic model. Içinde *Journal of Econometrics* (C. 6, Sayı 3, ss. 365-370). https://doi.org/https://doi.org/10.1016/0304-4076(77)90006-9

Arbabkhah, H., Sedaghat, A., Jafari Kang, M., & Hamidi, M. (2024). Automatic identification system-based prediction of tanker and cargo estimated time of arrival in narrow waterways. *Journal of Marine Science and Engineering*, *12*(2), 215. https://doi.org/10.3390/jmse12020215

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research & Evaluation*, *24*(1), n1.

Auret, L., & Aldrich, C. (2012). Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering*, *35*, 27-42. https://doi.org/10.1016/j.mineng.2012.05.008

Awad, M., Khanna, R., Awad, M., & Khanna, R. (2015). Support vector regression. *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*, 67-80. https://doi.org/https://doi.org/10.1007/978-1-4302-5990-9_4

Awah, P. C., Nam, H., & Kim, S. (2021). Short term forecast of container throughput: New variables application for the Port of Douala. *Journal of marine Science and Engineering*, *9*(7), 720. https://doi.org/10.3390/jmse9070720

Biggio, B., & Roli, F. (2018, October). Wild patterns: Ten years after the rise of adversarial machine learning. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (pp. 2154-2156). https://doi.org/10.1145/3243734.3264418

Botnen, F. A., & Lam, H. H. A. (2021). *Short term analysis of lead time movements in dry bulk shipping A linear approach to explore the relationship between lead time and other order-specific determinants of historical voyage orders* (Master's thesis).

García Cabello, J. (2022). Mathematical neural networks. *Axioms*, *11*(2), 80. https://doi.org/10.3390/axioms11020080

Chan, H. K., Xu, S., & Qi, X. (2019). A comparison of time series methods for forecasting container throughput. *International journal of logistics research and applications*, *22*(3), 294-303. https://doi.org/10.1080/13675567.2018.1525342

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). https://doi.org/10.1145/2939672.2939785

Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2007). Maritime transportation. *Handbooks in operations research and management science*, *14*, 189-284. https://doi.org/10.1016/S0927-0507(06)14004-9

Chu, Z., Yan, R., & Wang, S. Vessel Arrival Time to Port Prediction: A Uniform Approach Integrating Port Call Records with Ais Data. *Available at SSRN 5004783*.

Cuong, T. N., Kim, H. S., You, S. S., & Nguyen, D. A. (2022). Seaport throughput forecasting and post COVID-19 recovery policy by using effective decision-making strategy: A case study of Vietnam ports. *Computers & Industrial Engineering*, *168*, 108102. https://doi.org/10.1016/j.cie.2022.108102

Dinh, G. H., Pham, H. T., Nguyen, L. C., Dang, H. Q., & Pham, N. D. K. (2024). Leveraging artificial intelligence to enhance port operation efficiency. *Polish Maritime Research*, *31*(2), 140-155. https://doi.org/10.2478/pomr-2024-0030

Du, P., Wang, J., Yang, W., & Niu, T. (2019). Container throughput forecasting using a novel hybrid learning method with error correction strategy. *Knowledge-Based Systems*, *182*, 104853. https://doi.org/10.1016/j.knosys

El Mekkaoui, S., Benabbou, L., & Berrado, A. (2022). Machine learning models for efficient port terminal operations: Case of vessels' arrival times prediction. *IFAC-PapersOnLine*, *55*(10), 3172-3177. https://doi.org/10.1016/j.ifacol.2022.10.217

Evmides, N., Aslam, S., Ramez, T. T., Michaelides, M. P., & Herodotou, H. (2024). Enhancing Prediction Accuracy of Vessel Arrival Times Using Machine Learning. *Journal of Marine Science & Engineering*, *12*(8). https://doi.org/10.3390/jmse12081362

Fan, A., Wang, Y., Yang, L., Tu, X., Yang, J., & Shu, Y. (2024). Comprehensive evaluation of machine learning models for predicting ship energy consumption based on onboard sensor data. *Ocean & Coastal Management*, *248*, 106946. https://doi.org/10.1016/j.ocecoaman.2023.106946

Geng, J., Li, M. W., Dong, Z. H., & Liao, Y. S. (2015). Port throughput forecasting by MARS-RSVR with chaotic simulated annealing particle swarm optimization algorithm. *Neurocomputing*, *147*, 239-250. https://doi.org/10.1016/j.neucom.2014.06.070

Ghosh, S. (2022). *Understanding the beam of a ship*. Marine Insight. https://www.marineinsight.com/naval-architecture/understanding-the-beam-of-a-ship/

Graves, A., & Graves, A. (2012). Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, 37-45.

Hathikal, S., Chung, S. H., & Karczewski, M. (2020). Prediction of ocean import shipment lead time using machine learning methods. *SN Applied Sciences*, *2*(7), 1272. https://doi.org/10.1007/s42452-020-2951-5

Haykin, S. (2009). *Neural networks and learning machines, 3/E*. Pearson Education India.

Hobbs, P. V., & Wallace, J. M. (1977). *Atmospheric science: an introductory survey*. Academic press.

International Maritime Organization. (1987). *Resolution A.600(15): IMO ship identification number scheme. In Resolutions and other decisions (Resolutions 596–635): Assembly, fifteenth session, 9–20 November 1987* (Resolution A.600(15)). International Maritime Organization.

International Telecommunication Union. (2016). *Manual for use by the maritime mobile and maritime mobile-satellite services* (Edition 2016). Retrieved from https://www.itu.int

Jin, J., Ma, M., Jin, H., Cui, T., & Bai, R. (2023). Container terminal daily gate in and gate out forecasting using machine learning methods. *Transport Policy*, *132*, 163-174. https://doi.org/10.1016/j.tranpol.2022.11.010

Jonquais, A., & Krempl, F. (2019). Predicting shipping time with machine learning.

Machado, M. R., Karray, S., & De Sousa, I. T. (2019, August). LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. In *2019 14th international conference on computer science & education (ICCSE)* (pp. 1111-1116). IEEE. https://doi.org/10.1109/ICCSE.2019.8845529

Kolley, L., Rückert, N., Kastner, M., Jahn, C., & Fischer, K. (2023). Robust berth scheduling using machine learning for vessel arrival time prediction. *Flexible services and manufacturing journal*, *35*(1), 29-69. https://doi.org/10.1007/s10696-022-09462-x

Kussul, E., Baidyk, T., & Wunsch, D. C. (2010). *Neural networks and micromechanics* (pp. IX-1). Berlin: Springer. https://doi.org/10.1007/978-3-642-02535-8

Lee, G. C., & Bang, J. Y. (2024). Forecasting Container Throughput of Singapore Port Considering Various Exogenous Variables Based on SARIMAX Models. *Forecasting*, *6*(3), 748-760. https://doi.org/10.3390/forecast6030038

Lei, J., Chu, Z., Wu, Y., Liu, X., Luo, M., He, W., & Liu, C. (2024). Predicting vessel arrival times on inland waterways: A tree-based stacking approach. *Ocean Engineering*, *294*, 116838. https://doi.org/10.1016/j.oceaneng.2024.116838

Li, J., Zhang, S., & Xu, B. (2024). Containership delay propagation risk analysis based on effective multivariate transfer entropy. *Ocean Engineering*, *298*, 117077.

https://doi.org/10.1016/j.oceaneng.2024.117077

Liu, B., Wang, X., & Liang, X. (2023). Neural network-based prediction system for port throughput: A case study of Ningbo-Zhoushan Port. *Research in Transportation Business & Management*, *51*, 101067. https://doi.org/10.1016/j.rtbm.2023.101067

Liu, K., Yu, Q., Yuan, Z., Yang, Z., & Shu, Y. (2021). A systematic analysis for maritime accidents causation in Chinese coastal waters using machine learning approaches. *Ocean & Coastal Management*, *213*, 105859. https://doi.org/10.1016/j.ocecoaman.2021.105859

Liu, L., & Özsu, M. T. (Eds.). (2009). *Encyclopedia of database systems* (Vol. 6). New York, NY, USA: Springer. https://doi.org/10.1007/978-0-387-39940-9_542

Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-based model to forecast stock prices. *Complexity*, *2020*(1), 6622927. https://doi.org/10.1155/2020/6622927

Nieto, M. R., Carmona-Benitez**,** R. B., & Martinez, J. N. (2021). Comparing models to forecast cargo volume at port terminals. *Journal of applied research and technology*, *19*(3), 238-249. https://doi.org/10.1016/j.jart.2017.02.005

İMEAK Deniz Ticaret Odası – Turkish Chamber of Shipping. (2024). *Maritime Sector Report 2023 [PDF]. Retrieved from https://www.denizticaretodasi.org.tr/media/SharedDocuments/sektorraporu/2024/maritime%20sector%20report_2023_web2.pdf*

Monroe, J. W., & Stewart, R. (2005). *Dictionary of maritime and transportation terms*. Cornell Maritime Press.

Moscoso-López, J. A., Urda, D., Ruiz-Aguilar, J. J., Gonzalez-Enrique, J., & Turias, I. J. (2021). A machine learning-based forecasting system of perishable cargo flow in maritime transport. *Neurocomputing*, *452*, 487-497. https://doi.org/10.1016/j.neucom.2019.10.121

Naser, M. Z., & Alavi, A. H. (2023). Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences. *Architecture, Structures and Construction*, *3*(4), 499-517. https://doi.org/10.1007/s44150-021-00015-8

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, *4*(51-62), 56. https://doi.org/10.20544/horizons.b.04.1.17.p05

Ni, D., Xiao, Z., & Lim, M. K. (2020). A systematic review of the research trends of machine learning in supply chain management. *International Journal of Machine Learning and Cybernetics*, *11*, 1463-1482. https://doi.org/10.1007/s13042-019-01050-0

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, *31*.

Rajendra Rao, A., Wang, H., & Gupta, C. (2024). Predictive Analysis for Optimizing Port Operations. *arXiv e-prints*, arXiv-2401. http://arxiv.org/abs/2401.14498

Rashed, Y., Meersman, H., Van de Voorde, E., & Vanelslander, T. (2017). Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp. *Maritime Economics & Logistics*, *19*(4), 749-764. https://doi.org/10.1057/mel.2016.8

Servos, N., Liu, X., Teucke, M., & Freitag, M. (2019). Travel time prediction in a multimodal freight transport relation using machine learning algorithms. *Logistics*, *4*(1), 1. https://doi.org/10.3390/logistics4010001

Shankar, S., Ilavarasan, P. V., Punia, S., & Singh, S. P. (2020). Forecasting container throughput with long short-term memory networks. *Industrial management & data systems*, *120*(3), 425-441. https://doi.org/10.1108/IMDS-07-2019-0370

Shin, G. H., & Yang, H. (2025). Maritime accident prediction in busan port using machine learning: An integrated approach with maritime accident reports and VTS data. *Ocean Engineering*, *316*, 119968. https://doi.org/10.1016/j.oceaneng.2024.119968

Štepec, D., Martinčič, T., Klein, F., Vladušič, D., & Costa, J. P. (2020, June). Machine learning based system for vessel turnaround time prediction. In *2020 21st IEEE international conference on mobile data management (MDM)* (pp. 258-263). IEEE. https://doi.org/10.1109/MDM48529.2020.00060

Tan, N. D., Yu, H. C., Long, L. N. B., & You, S. S. (2021). Time series forecasting for port throughput using recurrent neural network algorithm. *Journal of International Maritime Safety, Environmental Affairs, and Shipping*, *5*(4), 175-183. https://doi.org/10.1080/25725084.2021.2014245

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tirkolaee, E. B., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., & Aeini, S. (2021). Application of machine learning in supply chain management: a comprehensive overview of the main areas. *Mathematical problems in engineering*, *2021*(1), 1476043. https://doi.org/10.1155/2021/1476043

Lilit, T. (2023). *Freight and Lead Time Estimator* (Master's thesis, Universidade NOVA de Lisboa (Portugal)).

Ubaid, A., Hussain, F., & Saqib, M. (2021). Container shipment demand forecasting in the Australian shipping industry: A case study of Asia–Oceania trade lane. *Journal of Marine Science and Engineering*, *9*(9), 968. https://doi.org/10.3390/jmse9090968

Valero, C. I., Martínez, Á., Oltra-Badenes, R., Gil, H., Boronat, F., & Palau, C. E. (2022). Prediction of the Estimated Time of Arrival of container ships on short-sea shipping: A pragmatical analysis. *IEEE Latin America Transactions*, *20*(11), 2354-2362. https://doi.org/10.1109/TLA.2022.9904760

Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric science: an introductory survey* (Vol. 92). Elsevier. https://books.google.com.tr/books?id=HZ2wNtDOU0oC

Xiang, Y., Gou, L., He, L., Xia, S., & Wang, W. (2018). A SVR–ANN combined model based on ensemble EMD for rainfall prediction. *Applied Soft Computing*, *73*, 874-883. https://doi.org/10.1016/j.asoc.2018.09.018

Xie, G., Wang, S., Zhao, Y., & Lai, K. K. (2013). Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study. *Applied Soft Computing*, *13*(5), 2232-2241. https://doi.org/10.1016/j.asoc.2013.02.002

Xu, S., Zou, S., Huang, J., Yang, W., & Zeng, F. (2022). Comparison of different approaches of machine learning methods with conventional approaches on container throughput forecasting. *Applied Sciences*, *12*(19), 9730. https://doi.org/10.3390/app12199730

Yang, C. H., & Chang, P. Y. (2020). Forecasting the demand for container throughput using a mixed-precision neural architecture based on CNN–LSTM. *Mathematics*, *8*(10), 1784. https://doi.org/10.3390/math8101784

Yao, W. (2021, August). Prediction of container throughput of Dalian Port Based on factor analysis and ARIMA model. In *IOP Conference Series: Earth and Environmental Science* (Vol. 831, No. 1, p. 012046). IOP Publishing. https://doi.org/10.1088/1755-1315/831/1/012046

Yoo, S. L., & Kim, K. I. (2023). Deep learning-based prediction of ship transit time. *Ocean Engineering*, *280*, 114592. https://doi.org/10.1016/j.oceaneng.2023.114592

Yoon, J. H., Kim, S. W., Jo, J. S., & Park, J. M. (2023). A comparative study of machine learning models for predicting vessel dwell time estimation at a terminal in the Busan new port. *Journal of Marine Science and Engineering*, *11*(10), 1846. https://doi.org/10.3390/jmse11101846

Younis, H., Sundarakani, B., & Alsharairi, M. (2022). Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions. *Journal of Modelling in Management*, *17*(3), 916-940. https://doi.org/10.1108/JM2-12-2020-0322

Zeng, F., & Xu, S. (2024). A hybrid container throughput forecasting approach using bi-directional hinterland data of port. *Scientific Reports*, *14*(1), 25502. https://doi.org/10.1038/s41598-024-77376-9

Zhang, F., & O'Donnell, L. J. (2020). Support vector regression. In *Machine learning* (pp. 123-140). Academic Press. https://doi.org/10.1016/B978-0-12-815739-8.00007-9

Zhang, Y., Fu, Y., & Li, G. (2020, September). Research on container throughput forecast based on ARIMA-BP neural network. In *Journal of Physics: Conference Series* (Vol. 1634, No. 1, p. 012024). IOP Publishing. https://doi.org/10.1088/1742-6596/1634/1/012024

EKLER

EK 1 Düşük akış doğrulama veri tabanlı en iyi performans gösteren modeller teslim süresi tahmini (saat).

EK 2 Yüksek akış doğrulama veri tabanlı en iyi performans gösteren modeller teslim süresi tahmini (saat).

EK 3 Arayüzün ana giriş sayfası ve sistem genel bilgeleri

EK 4 Konteyner iş hacminin talebi için geliştirilen arayüzü

EK 5 Teslim süresi için geliştirilen arayüzü

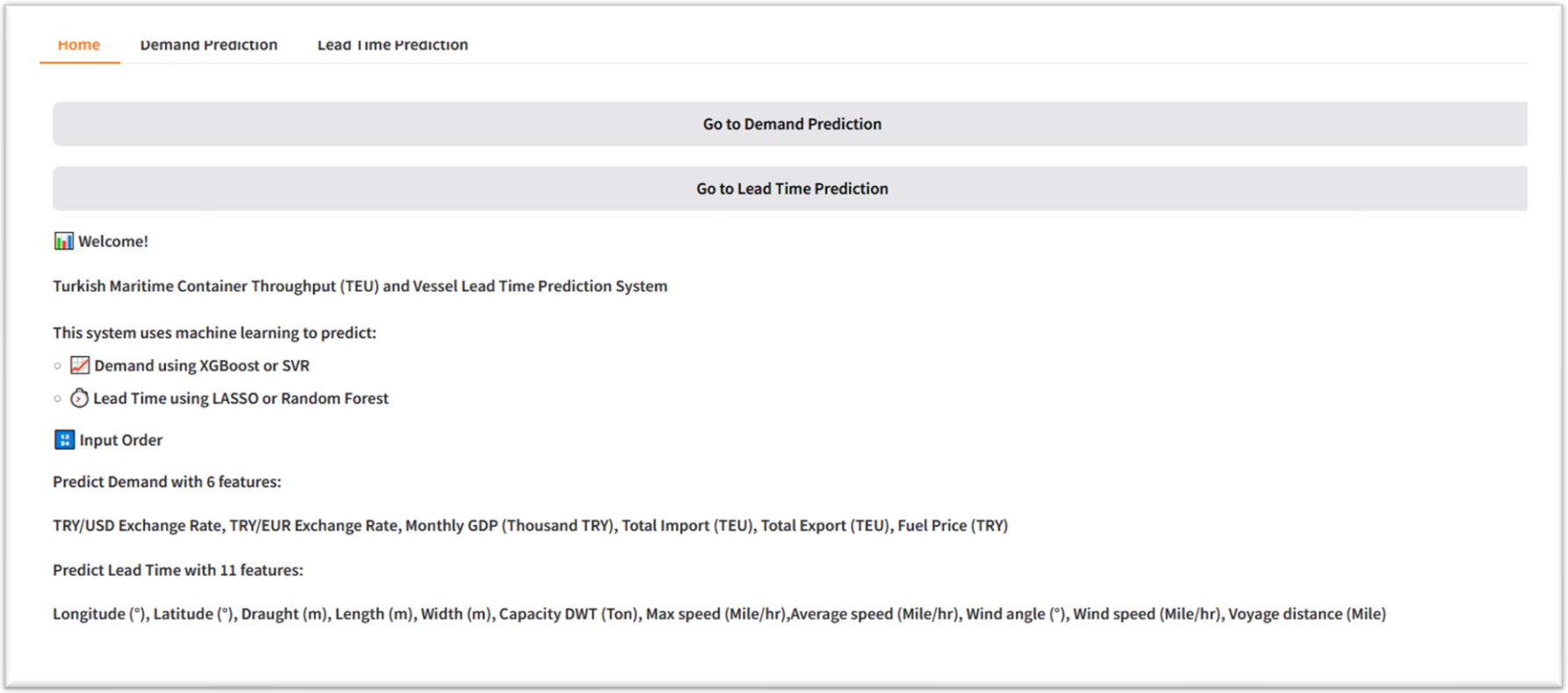
EK 1 Düşük akış doğrulama veri tabanlı en iyi performans gösteren modeller teslim süresi tahmini (saat).

Gerçek Geçmiş Veriler	LASSO	RF	ANN	CatBoost	SVR	XGBoost
9.03	11.26	9.99	9.32	10.26	9.32	10.19
9.78	11.48	9.57	9.64	8.81	8.42	11.90
10.89	12.96	12.65	10.51	11.77	11.39	11.03
11.05	12.33	11.25	11.67	10.84	10.10	11.35
11.90	13.74	15.28	11.37	12.01	11.10	14.53
12.45	14.08	14.64	12.03	10.92	11.20	14.23
12.80	14.90	15.77	12.88	11.40	12.48	14.37
14.14	15.80	14.32	13.21	15.10	13.77	15.63
14.57	16.28	15.38	14.42	16.63	14.19	17.32
21.75	23.46	25.40	20.53	23.26	21.88	25.86
21.90	23.66	23.78	22.41	21.88	21.27	23.38
23.52	25.22	26.96	24.02	24.57	23.96	26.47
24.00	27.97	25.82	25.94	26.93	26.17	25.96
26.10	26.52	25.87	23.50	24.80	24.11	25.72
27.67	29.56	27.78	24.17	27.83	28.34	27.71
31.03	30.30	32.51	29.93	30.21	28.07	31.48
31.72	31.57	32.35	30.05	31.86	30.78	32.95
32.38	31.98	31.53	32.64	30.97	30.86	31.21
35.68	36.30	38.66	33.78	36.86	34.57	37.11
37.00	36.38	35.15	36.76	36.84	35.65	38.24
40.13	40.11	42.72	40.26	41.35	38.99	44.29
42.15	41.49	42.41	42.54	40.54	40.06	43.44
43.03	45.09	48.41	41.25	45.25	45.14	45.83
47.42	45.34	47.59	44.01	46.10	44.15	48.26
48.37	48.34	49.54	45.76	47.20	47.35	49.99

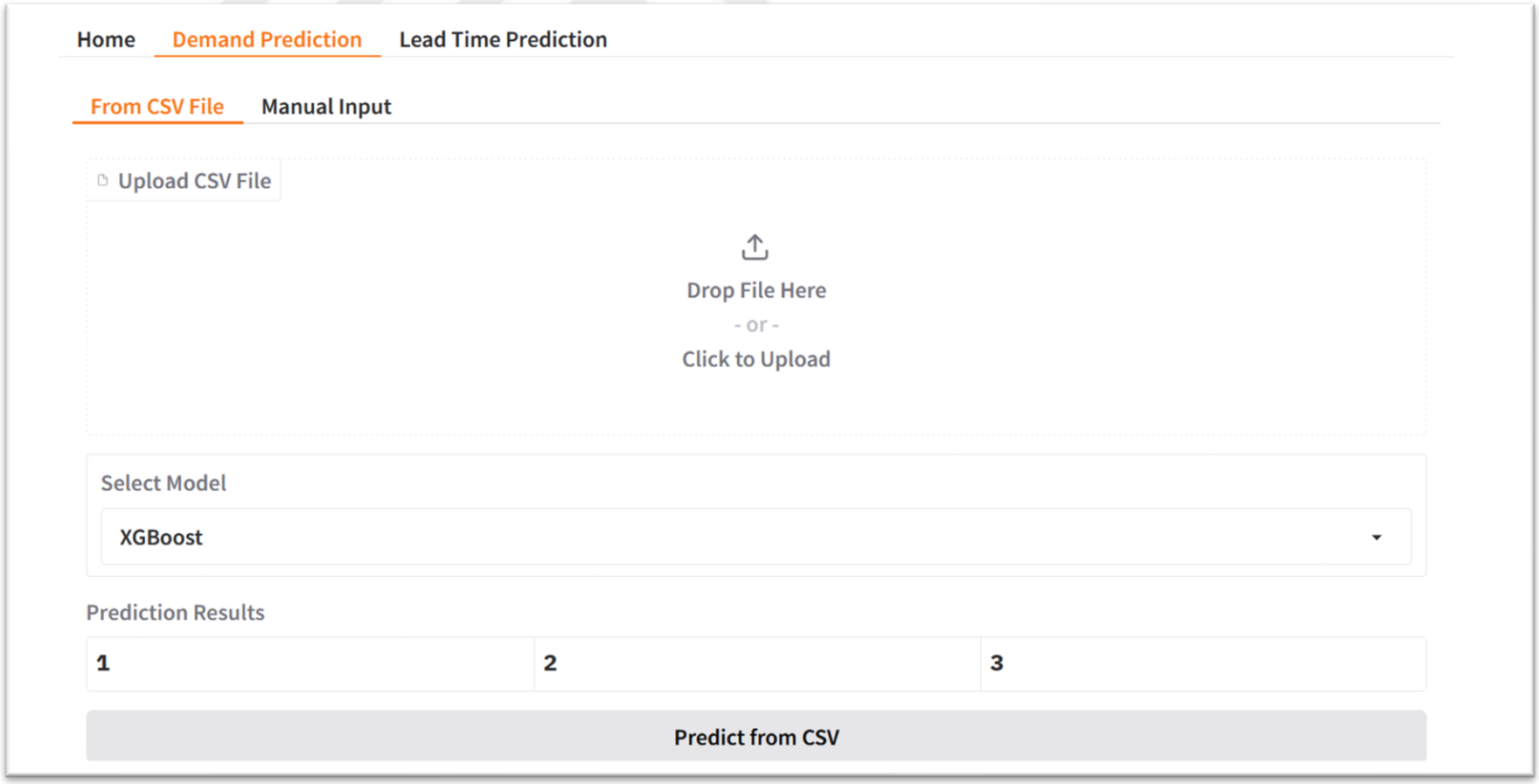
EK 2 Yüksek akış doğrulama veri tabanlı en iyi performans gösteren modeller teslim süresi tahmini (saat).

Gerçek Geçmiş Veriler	LASSO	RF	ANN	CatBoost	SVR	XGBoost
50.50	48.95	49.47	49.30	47.68	48.06	48.70
50.85	52.49	50.14	51.95	53.87	52.81	53.28
50.88	49.28	50.52	52.23	48.66	51.12	49.56
51.13	49.14	47.81	49.35	50.26	48.05	49.86
54.18	55.03	58.30	52.52	58.20	51.87	53.97
54.60	56.30	51.45	55.93	56.40	55.93	55.77
56.48	54.48	53.25	55.05	56.32	54.45	51.70
56.62	56.02	49.58	56.61	55.70	54.42	57.99
56.78	57.91	58.89	59.46	60.24	58.44	55.72
59.32	59.75	61.64	57.91	59.98	58.79	59.24
60.57	58.00	61.13	62.66	57.48	58.46	58.93
61.90	61.96	63.10	62.91	67.31	61.50	60.72
64.07	61.09	62.76	60.12	64.17	59.61	61.55
64.68	61.84	61.52	61.04	67.60	62.24	61.92
65.33	61.82	63.51	63.63	65.42	60.60	62.93
73.35	70.40	77.78	71.12	71.14	70.11	73.27
73.82	73.87	81.33	70.18	79.84	74.28	77.37
77.55	74.37	79.13	76.22	76.57	75.28	77.46
78.13	78.09	82.05	78.40	77.83	78.39	75.64
81.35	81.23	83.12	83.05	80.34	81.75	83.00
81.42	78.55	81.17	79.75	79.21	78.65	76.30
81.70	80.55	83.52	81.74	83.65	81.00	77.84
83.57	78.82	86.39	82.15	82.26	78.85	74.36
84.53	83.16	84.02	84.07	84.37	83.27	83.01
86.48	85.54	83.57	85.99	89.13	86.14	78.83
86.55	84.00	82.01	84.55	86.94	84.55	84.67
90.35	87.11	80.09	89.11	92.65	89.11	81.94
90.63	87.81	84.17	87.78	87.83	87.78	81.27
90.88	88.60	85.65	90.55	91.05	88.07	86.30
91.05	88.27	82.11	90.94	87.46	88.36	77.90
50.50	48.95	49.47	49.52	47.68	48.06	48.70
50.85	52.49	50.14	50.79	53.87	52.81	53.28
50.88	49.28	50.52	51.29	48.66	51.12	49.56
51.13	49.14	47.81	51.37	50.26	48.05	49.86
54.18	55.03	58.30	53.84	58.20	51.87	53.97
54.60	56.30	51.45	54.60	56.40	55.93	55.77

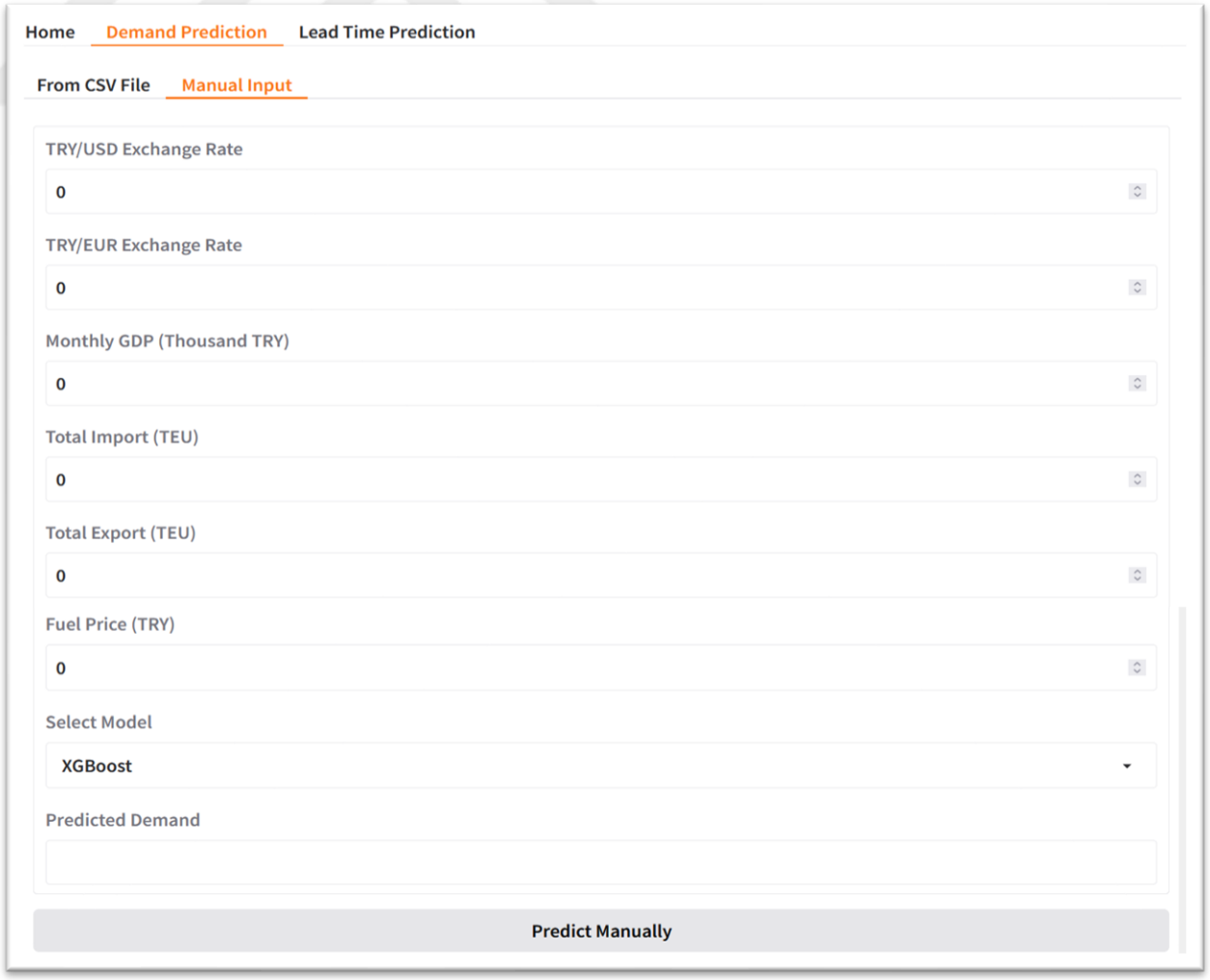
EK 3 Arayüzün ana giriş sayfası ve sistem genel bilgileri



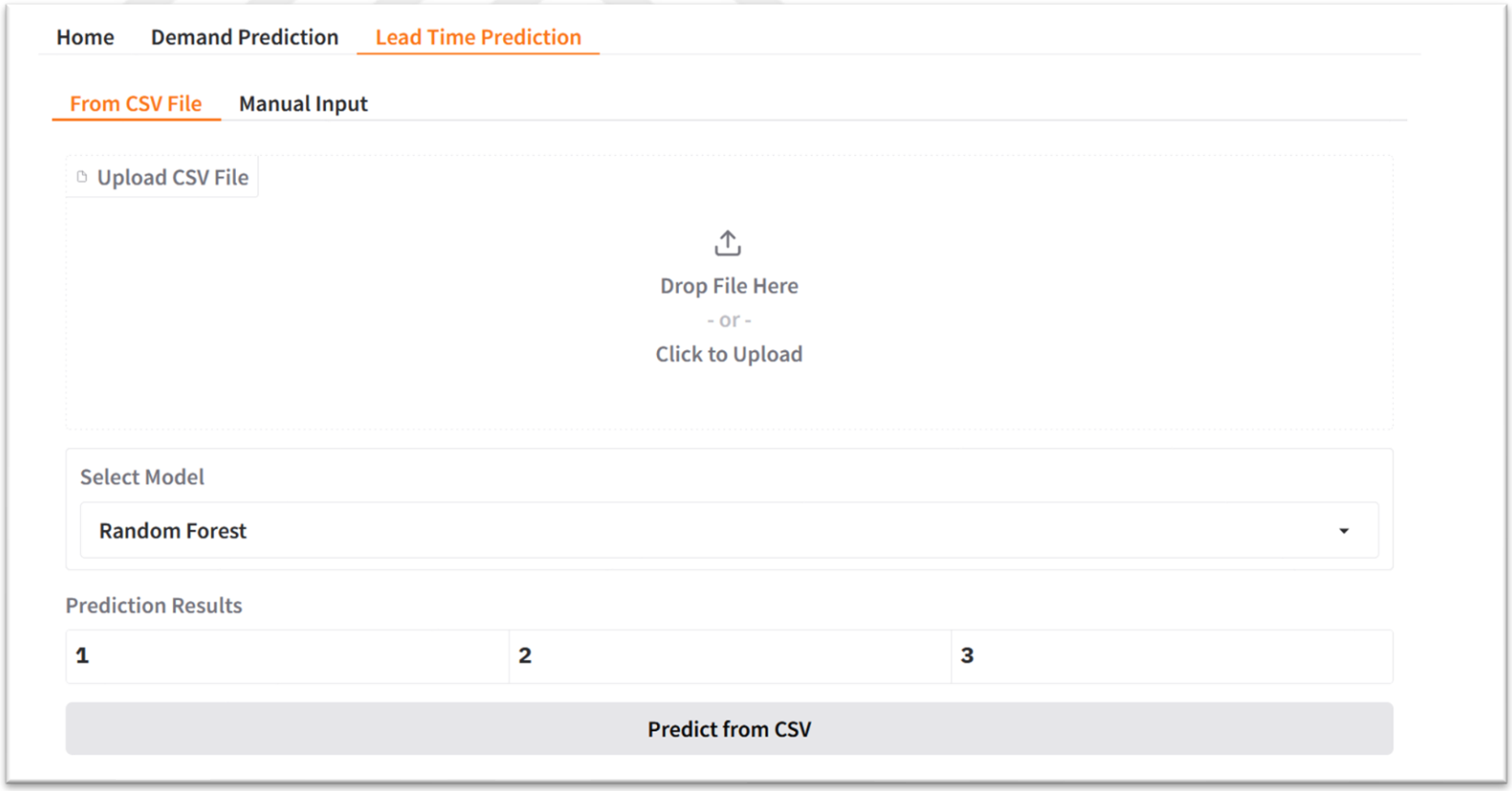
EK 4 Konteyner iş hacminin talebi için geliştirilen arayüz



EK 4 Konteyner iş hacminin talebi için geliştirilen arayüz (Devamı)



EK 5 Teslim süresi için geliştirilen arayüz



EK 5 Teslim süresi için geliştirilen arayüz (Devamı)

