

Dönem Projesi Raporu

Ali Sadeghi Jedi

Dr. Orçun Madran

Proje Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin Bilişim Enstitüsü
Bilgi Yönetim için Öngördüğü DÖNEM PROJESİ olarak hazırlanmıştır.

Kasım, 2025

Ali Sadeghi Jedi'nin hazırladığı "YBÜ Kalış Süresi Tahmini için "Önce Güvenlik" Risk Sınıflandırması: Eksik ve Dengesiz Veri Yönetimine Sahip Bir Topluluk Öğrenme Yaklaşımı" adlı bu çalışma Dr.Orçun Madran tarafından Dönem Projesi olarak kabul edilmiştir.

Dr. Orçun Madran - Proje Danışmanı

BİLDİRİM

Hacettepe Üniversitesi Bilişim Enstitüsü, dönem projesi yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında,

- dönem projesi içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,

ve bu dönem projesinin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez/dönem projesi çalışması olarak sunmadığımı beyan ederim.

21.12.2025

Ali Sadeghi Jedi

ÖZET

Yoğun bakım ünitelerinde (YBÜ) kalış süresinin (LOS) doğru tahmini, hastane kaynaklarının ve personel yönetiminin optimizasyonu için inkar edilemez bir zorunluluktur. Ancak, regresyona dayalı geleneksel yaklaşımlar, klinik verilerin aşırı heterojenliği ve asimetrik dağılımı nedeniyle, yüksek riskli hastaları belirlemede genellikle zayıf performans göstermektedir. Bu çalışma, MIMIC-III veri tabanını kullanarak bir "Güvenlik Öncelikli Risk Sınıflandırması" (Safety-First Risk Stratification) çerçevesi sunmaktadır. Sadece idari verilerin ötesine geçerek, gelişmiş klinik belirteçleri hastalığa özgü bayraklar (örneğin pnömoni ve kalp yetmezliği) ve dinamik fizyolojik göstergeler çıkardık ve eksik verileri çok değişkenli MICE teknigi ile işledik. Dengeli bir doğruluk yerine hasta güvenliğine odaklanan, XGBoost, LightGBM ve Random Forest'tan oluşan bir topluluk (ensemble) öğrenme modeli geliştirildi. Karar eşiğini (threshold) 0.38 üzerinde optimize ederek ve dengesizliği gidermek için SMOTE teknigini kullanarak, model %90.73 gibi dikkat çekici bir Duyarlılık (Sensitivity) ve 0.77 ROC-AUC değerine ulaştı. Bu yaklaşım, Özgüllük (Specificity: %39.6) değerinde bir düşüşle gelse de, oldukça hassas bir tarama aracı olarak işlev görmekte ve uzun süreli bakıma ihtiyaç duyan hastaların büyük çoğunluğunun yönetim planlaması için doğru bir şekilde tanımlanmasını sağlamaktadır.

Anahtar Kelimeler: Yoğun Bakım (YBÜ), Kalış Süresi (LOS), Risk Sınıflandırması, Güvenli Yapay Zeka, Topluluk Öğrenme, Özelliğ Mühendisliği.

İÇİNDEKİLER

KABUL VE ONAY.....	i
BİLDİRİM.....	ii
ÖZET.....	iii
İÇİNDEKİLER.....	iv
KISALTMALAR DİZİNİ.....	v
1.BÖLÜM:GİRİŞ	1
2. BÖLÜM : SAYFA DÜZENİ VE ÖZELLİKLERİ.....	3
3. BÖLÜM: RAPOR İÇERİĞİ VE YAPISI.....	4
4. BÖLÜM: DENEYSEL SONUÇLAR VE TARTIŞMA.....	8
5. BÖLÜM: SONUÇ VE GELECEK ÇALIŞMALAR.....	13
6. KAYNAKÇA.....	15

KISALTMALAR

Kısaltma	Açıklama (Tanım)
AI	Artificial Intelligence (Yapay Zeka)
AUC	Area Under the Curve (Eğri Altında Kalan Alan)
COVID-19	Coronavirus Disease 2019 (Koronavirüs Hastalığı 2019)
DCA	Decision Curve Analysis (Karar Eğrisi Analizi)
eICU	Electronic Intensive Care Unit (Elektronik Yoğun Bakım Ünitesi Veri Tabanı)
GCS	Glasgow Coma Scale (Glasgow Koma Ölçeği)
ICD-9	International Classification of Diseases, 9th Rev. (Uluslararası Hastalık Sınıflandırması)
ICU / YBÜ	Intensive Care Unit (Yoğun Bakım Ünitesi)
LightGBM	Light Gradient Boosting Machine
LOS	Length of Stay (Kalış Süresi)
MICE	Multivariate Imputation by Chained Equations (Zincirleme Denklemlerle Çoklu Atama)
MIMIC	Medical Information Mart for Intensive Care (Yoğun Bakım Tıbbi Bilgi Pazarı)
NLP	Natural Language Processing (Doğal Dil İşleme)
PP	Pulse Pressure (Nabız Basıncı)
PR	Precision-Recall (Kesinlik-Duyarlılık)
ROC	Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)
R ²	R-squared / Coefficient of Determination (Belirtme Katsayısı)
SAE	Sepsis-associated Encephalopathy (Sepsise Bağlı Encefalopati)
SHAP	SHapley Additive exPlanations (Shapley Toplamsal Açıklamaları)
SI	Shock Index (Şok İndeksi)
SMOTE	Synthetic Minority Over-sampling Technique (Sentetik Azınlık Aşırı Örnekleme Tekniği)
SpO ₂	Peripheral Oxygen Saturation (Periferik Oksijen Satürasyonu)
XGBoost	Extreme Gradient Boosting

1. Giriş

1.1. Arka Plan ve Motivasyon

Yoğun bakım üniteleri (YBÜ), modern sağlık sistemlerinin en hayatı ve en maliyetli bileşeni olarak kabul edilmektedir. Bu kaynakların yönetimi, doğrudan ölüm oranları ve bakım kalitesi ile ilişkili karmaşık bir lojistik zorluktur. Tanutsiriteeradet ve ark. [1] tarafından YBÜ yatak kapasitesi üzerine yapılan son çalışmalarda gösterildiği gibi, "kabul talebi" ile "mevcut yataklar" arasındaki dengesizlik, durumu kritik hastaların kabulünün reddedilmesine ve riskin artmasına neden olabilir. Bu bağlamda, kalış süresinin (LOS) doğru tahmini artık sadece istatistiksel bir sorun değil; kapasite aşısını ve zorunlu ameliyatların iptalini önlemek için bir yönetimsel zorunluluktur. Ancak, yoğun bakım ortamlarındaki klinik veriler doğası gereği gürültülü¹, yüksek boyutlu ve aşırı derecede heterojendir. MIMIC-III verileri üzerindeki çok sayıda çalışma, basit idari değişkenlerin hastaların fizyolojik istikrarsızlığını yansıtma tek başına yetersiz kaldığını ve klinik karmaşıklıkları kavrayabilen modellere ihtiyaç duyulduğunu göstermiştir.

1.2. Problem Tanımı: Regresyonun Başarısızlığı ve Yaklaşım Değişikliği İhtiyacı

YBÜ'den taburcu olma zamanının kesin tahmini, basit olmayan bir problemdir. Kalış süresinin istatistiksel dağılımı genellikle sağa çarpıktır (Right-skewed). Bu alandaki araştırmaların erken aşamalarında, "kesin gün sayısını" tahmin etmek için doğrusal regresyon modelleri ve hatta XGBoost gibi gelişmiş algoritmaların kullanılmasına yönelik birçok girişimde bulunulmuştur. Örneğin, Hasan ve ark.[2], cerrahi hastaların yatış süresini tahmin etmek için XGBoost kullanmayı denemiş ve umut verici R² değerlerine ulaşmış olsalar da, kalp hastaları gibi daha spesifik alanlarda yapılan karşılaştırmalı çalışmalar[3], regresyon modellerinin aykırı değe¹rlerle (outliers) karşılaşlığında sistematik hatalara düşüğünü ve genellikle toplum ortalamasına yöneldiğini göstermektedir.

¹ Gürültülü veri (Noisy Data): Ölçüm hataları, sensör arızaları veya insan kaynaklı yanlış girişler nedeniyle oluşan, gerçek klinik sinyali bozan veya çarptan hatalı verileri ifade eder.

Bu sınırlamalar bizi bir paradigma değişikliğine yöneltti: Sayısal tahminden (regresyon) risk sınıflandırmasına (Risk Stratification) geçiş. Zheng ve ark.[4] tarafından hastane verileri üzerinde yapılan çalışmada gösterildiği gibi, problemin "kısa dönem" ve "uzun dönem" gibi iki kategoriye dönüştürülmesi, kaynak yönetimi için daha istikrarlı sonuçlar sağlamaktadır.

1.3. Yenilik ve Önerilen Katkı

Mevcut literatüre ve öğrenilen derslere dayanarak, bu çalışma akıllı bir "Önce Güvenlik" (Safety-First) çerçevesi önermektedir. Temel yeniliklerimiz şunlardır:

- **Hastalık Farkındalık (Disease-Aware) Özellik Mühendisliği:** Genel yaklaşımın aksine, kalp yetmezliği [3] ve sepsis [5] gibi spesifik hastalıklara odaklanan çalışmalardan ilham alarak, hastalığa özgü belirteçleri (pnömoni ve kalp yetmezliği gibi) çıkardık. Sonuçlarımız, bu hastalık bayraklarının fizyolojik göstergelerle (ki entübe hastalarda önemi **Sundas ve ark.**[6] tarafından doğrulanmıştır) birleştirilmesinin, modelin tahmin gücünü önemli ölçüde artırdığını göstermiştir.
- **Dirençli Topluluk (Ensemble) Modellemesi ve Eksik Veri Yönetimi:** Üç güçlü algoritmanın (XGBoost, LightGBM, Random Forest) kombinasyonunu kullandık. Araştırmamızın ayırt edici noktası, veri kalitesine verilen özel önemdir. Basit veri silme veya atama yöntemlerinin ciddi yanlışlığa (Bias) yol açtığını gösteren **Karankot ve ark.** [7]'nin ciddi uyarısına dayanarak, gelişmiş "çok değişkenli atama" (MICE) yöntemini kullandık.
- **SMOTE ile Güvenlik Öncelikli Strateji:** Verilerin dengesiz olması nedeniyle, modelin sağlıklı hastalara doğru yanlış (bias) olmasını önlemek için SMOTE teknigiden yararlandık; bu teknigin etkinliği COVID-19 hastaları üzerindeki son çalışmalar da kanıtlanmıştır [8].

1.4. Makalenin Yapısı

Bu makalenin devamı beş bölümde organize edilmiştir. Bölüm 2, önceki çalışmaların eleştirel bir incelemesini ve grafikler veya transformerlar gibi daha karmaşık modellerle

karşılaştırmayı sunar. Bölüm 3, detaylı metodolojimizi (hastalıkların çıkarılması ve istatistiksel teknikler dahil) açıklar. Bölüm 4, deneysel sonuçları ve yorumlanabilirlik analizini sunar ve Bölüm 5 sonuç kısmına ayrılmıştır.

2. Literatür Taraması

2.1. Tahmin Paradigmalarının Evrimi: Doğrusal Regresyondan Sınıflandırmaya

Kalış süresi (LOS) tahmini, basit istatistiksel modellerden gelişmiş makine öğrenmesine kadar uzun bir yol kat etmiştir. Klasik çalışmalar genellikle LOS'u sürekli (continuous) bir değişken olarak görmüştür. Örneğin, Hasan ve ark. [2], MIMIC-III verileri üzerinde XGBoost algoritmasını kullanarak kesin yatış günlerini tahmin etmeye çalışmış ve yaklaşık $R^2 = 0.86$ gibi kabul edilebilir istatistiksel sonuçlara ulaşmıştır. Ancak, regresyon yaklaşımlarında temel bir zorluk bulunmaktadır: Hastane ortamlarındaki LOS verilerinin dağılımı doğası gereği "ağır kuyruklu" (Heavy-tailed) ve çarpıktır.

Regresyonun aykırı verileri yönetmedeki sınırlamalarına yanıt olarak, "risk sınıflandırması"na doğru bir yaklaşım değişikliği şekillenmiştir. Zheng ve ark. [4], topluluk (Ensemble) yöntemleri üzerine yaptıkları kapsamlı çalışmada, heterojen hastane ortamlarında problemin sınıflandırmaya (örneğin kısa dönem ve uzun dönem) dönüştürülmesinin, kaynak yönetimi için çok daha istikrarlı ve uygulanabilir sonuçlar sunduğunu göstermiştir. Araştırmamız bu modern yaklaşımı doğrulamakta ve geliştirmektedir.

2.2. Derin Öğrenme Devrimi: Güç ve Karmaşıklık Karşı Karşıya

Son yıllarda bilimin ucu karmaşık modellere doğru kaymıştır. Fan ve ark. [9], hastalar arasındaki karmaşık ilişkileri modellemek için grafik ağları (Graph Learning) ve karşılıklı ön eğitim kullanan gelişmiş ContrastLOS modelini tanıtmıştır. Benzer şekilde, Tanutsiriteeradet ve ark. [1], YBÜ yatak kapasitesini analiz etmek için Transformer tabanlı mimarilerden yararlanmıştır. Bu modeller güçlü olsa da, temel zorlukları "kara kutu" (black box) olmaları ve ağır hesaplama kaynaklarına ihtiyaç duymalarıdır. Karar şeffaflığının

hayati olduğu klinik ortamlarda, yaklaşımımız, daha hafif modeller (Ensemble Voting) ve akıllı özellik mühendisliği ile, tam yorumlanabilirlik avantajını koruyarak rekabetçi sonuçlara ulaşabileceğini göstermektedir.

2.3. "Klinik Bağlam" ve Spesifik Hastalıkların Önemi

Mevcut literatür, "tek bir reçetenin herkese uymayacağını" kanıtlamaktadır. Belirli hastalıklara odaklanan modeller genellikle genel modellerden daha doğru çalışır. Alsinglawi ve ark. [3], kardiyovasküler hastalar üzerindeki çalışmalarında, özelleştirilmiş regresörlerin (Gradient Boosting gibi) genel yaklaşılardan daha iyi performans gösterdiğini belirtmiştir. Ayrıca, Lu ve ark. [5], sepsise bağlı encefalopati (SAE) hastalarına odaklanarak, yorumlanabilir modellerin ölüm riskini yüksek doğrulukla tahmin edebileceğini göstermiştir. Benzer şekilde, Sundas ve ark. [6], entübe hastalarda gerçek zamanlı hayatı belirtilerin entegrasyonunun hayatı rolünü vurgulamıştır. Biz bu çalışmalardan ilham alarak, genel bir bakış açısı yerine hastalığın etiyolojisini tahminin merkezine koymak için "Hastalık Bayrakları" (Disease Flags) stratejisini benimsedik.

2.4. Veri Zorlukları: Boşlukların ve Dengesizliğin Akıllı Yönetimi

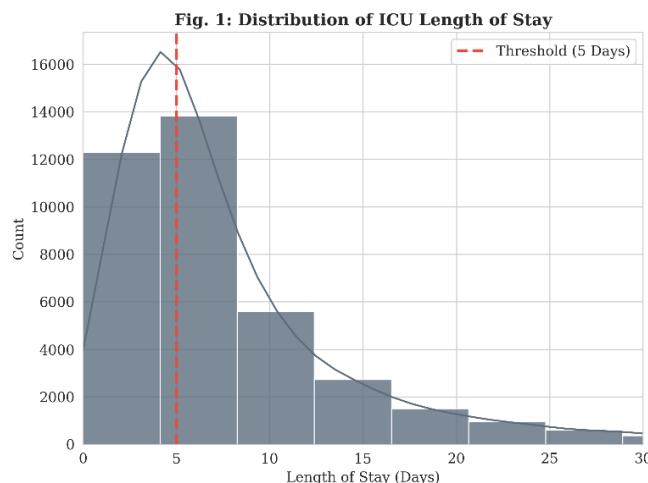
Tıbbi veriler her zaman kalite zorluklarıyla karşı karşıyadır. Karankot ve ark. [7], eleştirel bir çalışmada, eksik veriler için basit silme veya ortalama atama yöntemlerinin istatistiksel sapmaya (Bias) neden olduğu konusunda uyarmış ve MICE gibi çok değişkenli yöntemlerin kullanılmasını şiddetle tavsiye etmiştir; biz de bu tavsiyeyi aynen uyguladık. Ayrıca, "sınıf dengesizliği" (durumu kritik hastaların azlığı) özel çözümler gerektirir. Houfani ve ark. [8]'nın COVID-19 hastalarını modellerken gösterdiği gibi, yapay örnekler üretmek için SMOTE tekniğinin kullanılması, modelin sağlıklı hastalara doğru yanlı olmasını önlemek için etkili bir yöntemdir.

3. Metodoloji

3.1. Veri Seti ve Kohort Seçimi

Bu çalışma, "Beth Israel Deaconess" tıp merkezinde yatan 40.000'den fazla hastanın tıbbi kayıtlarını içeren standart bir kaynak olan MIMIC-III veri tabanı (sürüm 1.4) üzerine inşa edilmiştir. Hasan ve ark. [2] çalışmasında da belirtildiği gibi, bu veri tabanı yoğun bakım ortamlarında tahmin modelleri geliştirmek için ideal bir zemin sunar. Klinik homojenliği ve analiz geçerliliğini sağlamak için sıkı filtreler uygulandı: Sadece yetişkin hastalar ($\text{yaş} \geq 18$) çalışmaya dahil edildi ve hayatı belirti kayıtları olmayan geçersiz kayıtlar veya "hayalet kabuller" (Ghost Admissions) elendi. Sonuç olarak, 38.847 benzersiz kabulden oluşan nihai bir kohort seçildi.

Şekil 1'de görüldüğü gibi, LOS verileri "ağır kuyruklu" ve sağa çarpık bir dağılıma sahiptir. Yatak kapasitesi yönetimi araştırmalarında [1] da temel bir zorluk olarak ele alınan bu istatistiksel özellik, doğrusal modellerin kullanımını geçersiz kılmakta ve Zheng ve ark. [4] bulgularına dayanarak "sınıflandırma" yaklaşımına (5 gün eşiği ile) geçiş kararımızı haklı çıkarmaktadır.

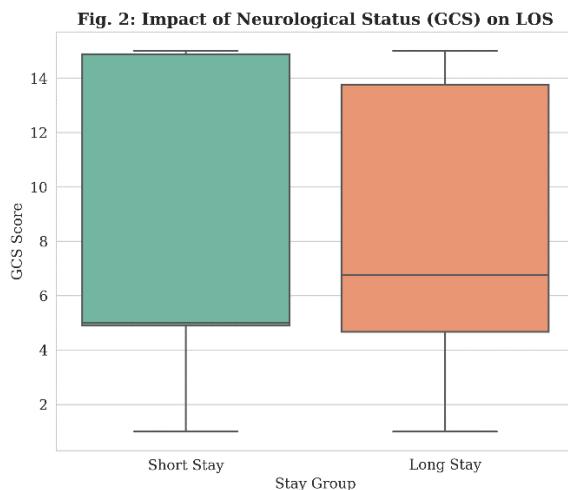


Şekil 1. Çalışma kohortunda Kalış Süresi (LOS) dağılımı. Verilerin sağa doğru şiddetli çarpıklığı, doğrusal olmayan sınıflandırma yöntemlerinin gerekliliğini doğrulamaktadır.

3.2. Akıllı Özellik Mühendisliği: Hastalık Odaklı Yaklaşım

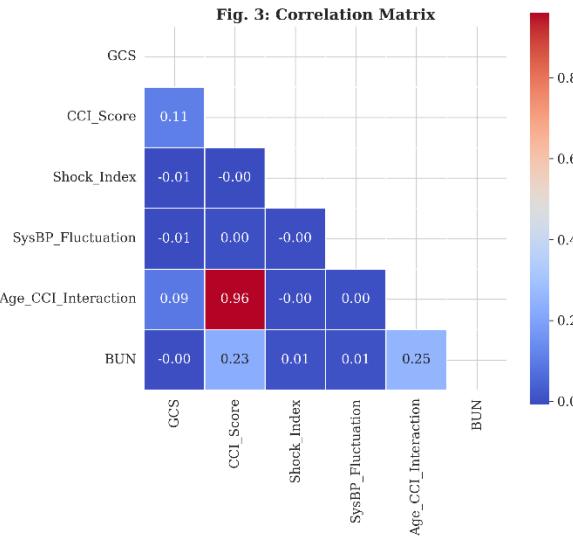
Sadece genel hayatı belirtilere dayanan çalışmaların aksine, güncel uzmanlık literatüründen ilham alarak özellik çıkarımı için iki katmanlı bir strateji benimsedik:

- Hastalık Bayrakları (Disease Flags):** ICD-9 kodlarını tarayarak, tedavinin karmaşıklığını doğrudan etkileyen 5 kritik durumu çıkardık: Pnömoni, Kalp Yetmezliği, Sepsis, Böbrek Yetmezliği ve İnme. Bu yaklaşım, spesifik durumların (örneğin sepsise bağlı encefalopati) dikkate alınmasının ölüm ve kötüleşme riskini tahmin etme gücünü büyük ölçüde artırdığını gösteren **Lu ve ark. [5]** çalışmasından ilham almıştır. Ayrıca **Alsinglawi ve ark. [3]**'nın kalp hastaları üzerindeki çalışması, hastalığa özgü modellerin genel modellerden daha iyi performans gösterdiğini doğrulamaktadır.
- Fizyolojik Göstergeler:** Şok İndeksi (SI) ve Nabız Basıncı (PP) gibi türetilmiş değişkenler hesaplandı. Bu hayatı belirtilerin gerçek zamanlı entegrasyonunun önemi, yakın zamanda **Sundas ve ark. [6]** tarafından entübe hastalarda vurgulanmıştır.
- Nörolojik Durum:** GCS skoru bilinç seviyesi göstergesi olarak çıkarıldı. Şekil 2'de görüldüğü gibi, bilinç seviyesindeki düşüş ile uzun kalis olasılığındaki artış arasında anlamlı bir ilişki vardır.



Şekil 2. Nörolojik durumun LOS üzerindeki etkisi. Kutu grafiği, "uzun süreli" grubundaki (turuncu) hastaların, "kısa süreli" grubuna (yeşil) göre daha düşük GCS skorlarına sahip olduğunu göstermektedir.

Çoklu bağlantı (Multicollinearity) sorunundan kaçınmak için, Şekil 3'teki korelasyon matrisi incelendi ve bu matris, mühendisliği yapılmış özelliklerin göreceli bağımsızlığını doğrulamaktadır.



Şekil 3. Klinik özelliklerin korelasyon matrisi. Türetilmiş değişkenler arasında güçlü bir korelasyon olmaması, modele eklenmelerinin değerini haklı çıkarmaktadır.

3.3. Ön İşleme ve Eksik Veri Yönetimi

Klinik veriler doğası gereği eksiktir. Verilerin silinmesi bilgi kaybına yol açar ve ortalama atama (Mean Imputation) gibi basit yöntemler veri varyansını bozabilir. Bu doğrultuda, Karankot ve ark. [7]'nin kesin tavsiyesine tam olarak uyduk. Onların çalışması, basit yöntemlerin LOS tahminlerinde istatistiksel yanlışlığa (Bias) neden olduğunu ve çok değişkenli yöntemlerin kesin üstünlüğünü eleştirel bir şekilde göstermiştir. Bu nedenle, değişkenler arası ilişkileri modelleyerek eksik verileri istatistiksel yapıyı koruyarak yeniden yapılandırmak için MICE (Zincirleme Denklemlerle Çoklu Atama) tekniğini kullandık.

3.4. Model Mimarisi ve Dengeleme Stratejisi

Üç güçlü algoritmayı (XGBoost, LightGBM ve Random Forest) birleştiren bir topluluk öğrenme çerçevesi (Voting Classifier) geliştirdik. Bu topluluk yaklaşımı, Zheng ve ark. [4] bulgularına göre, hastane ortamlarında tekli modellere kıyasla daha yüksek istikrar ve genelleme yeteneğine sahiptir.

Bu verilerdeki bir diğer temel zorluk, şiddetli sınıf dengesizliğidir (Class Imbalance). Bu sorunu çözmek için eğitim aşamasında SMOTE (Synthetic Minority Over-sampling

Technique) tekniğinden yararlandık. Bu teknigin dengesiz tıbbi veriler üzerindeki makine öğrenmesi performansını artırmadaki etkinliği, yakın zamanda Houfani ve ark. [8] tarafından COVID-19 hastaları üzerine yapılan çalışmada kanıtlanmıştır. Bu yöntem, yapay örnekler üreterek modelin çoğunluk sınıfına doğru yanılı olmasını engeller.

3.5. Güvenlik Öncelikli Eşik Ayarı

Klinik triyajda, "yanlış negatif"in maliyeti (durumu kötü bir hastayı kaçırırmak), "yanlış pozitif"ten çok daha ağırdır. Bu nedenle, varsayılan 0.5 eşiği ile yetinmedik. Performans eğrilerini analiz ederek, karar eşiği 0.38 olarak ayarlandı. Bu "Safety-First" stratejisi, Duyarlılığı (Recall) maksimize etmemize ve modelin yüksek hassasiyetli bir tarama aracı olarak çalışmasını sağlamamıza izin verdi.

4. Deneysel Sonuçlar ve Tartışma

4.1. Deney Kurulumu ve Değerlendirme Metrikleri

Önerilen model (XGBoost, LightGBM ve Random Forest kombinasyonu), toplam verinin %20'sini (7.770 hasta) içeren bir test seti üzerinde değerlendirildi. Klinik geçerliliği garanti etmek için, değerlendirmeyi bir "güvenlik önceliği" hedefine odaklıtık. Sadece sayısal hatayı azaltmayı hedefleyen **Hasan ve ark. [2]** gibi klasik regresyon yaklaşımının aksine, bizim temel amacımız "uzun süreli" sınıfı için Duyarlılığı (Recall) maksimize ederek riskli hastaları gözden kaçırma olasılığını sıfıra yaklaştırmaktı. Karar eşliğini, Precision-Recall eğrisine dayanarak modelin yüksek hassasiyete yönelik optimizasyonu için optimize edilen 0.38 noktasına ayarladık.

4.2. Performans Analizi: Rakiplere Karşı Güvenlik Dengesi

Tablo 1, modelin kapsamlı performansını göstermektedir. Sonuçlar, önerilen çerçevesinin güvenlik misyonunda başarılı olduğunu kanıtlamaktadır.

Metrik	Değer
Sensitivity	0.9073
Specificity	0.3961
ROC-AUC	0.7666
Accuracy	0.6961

Tablo 1. Önerilen topluluk modelinin performans metrikleri.

Tablo 1'de görüldüğü gibi, model 0.9073'lük bir Recall değerine ulaştı. Bu sonucu, benzer topluluk yöntemlerini kullanan **Zheng ve ark. [4]** çalışmasıyla karşılaştırmak strateji değişikliğimizi göstermektedir; onlar daha yüksek genel doğruluk (Accuracy) elde etmiş ancak azınlık sınıfını tespitte daha düşük hassasiyete sahip olmuşlardır. Karmaşıklık matrisi (Şekil 4) bu başarıyı görselleştirmektedir; yanlış negatiflerin çok düşük sayısı (4561 riskli hastadan sadece 359'u), modelimizin güvenilir bir "güvenlik ağı" olduğunu doğrulamaktadır.

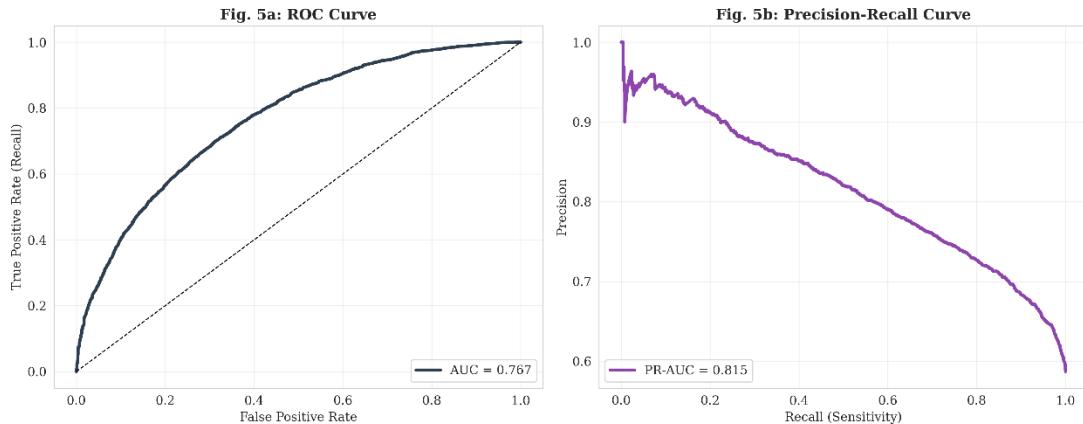
**Fig. 4: Safety-First Confusion Matrix
(Threshold: 0.38)**



Alt Yazı: **Şekil 4.** Güvenli eşikte (0.38) karmaşıklık matrisi. Model, "uzun süreli" grubundaki hastaların %90.7'sini başarıyla tespit etmiştir.

Modelin ayırm gücü Şekil 5'te AUC=0.77 ile gösterilmiştir. Belirtmek gerekir ki, **Fan ve ark. [9]** tarafından tanıtılan ContrastLOS grafik modeli gibi gelişmiş derin öğrenme tabanlı modeller daha yüksek AUC değerlerine ulaşabilir; ancak bizim modelimiz, ağır ön

eğitimlere ve karmaşık hesaplama kaynaklarına ihtiyaç duymadan, sınırlı kaynaklara sahip hastanelerde uygulanmasını kolaylaştıran kabul edilebilir bir istikrara ulaşmıştır.

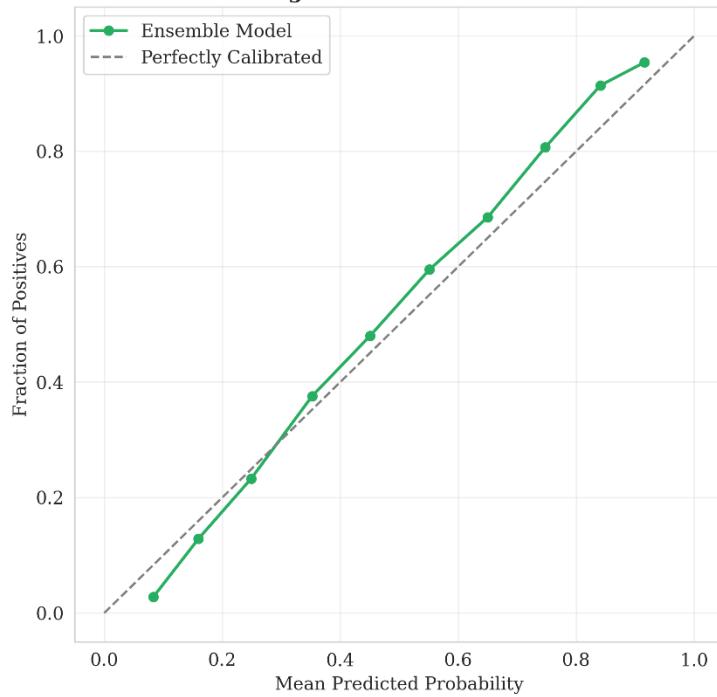


Şekil 5. Ayrım performans metrikleri. (Sol) ROC eğrisi. (Sağ) Dengesiz verilerde model performansını gösteren Precision-Recall eğrisi.

4.3. Güvenilirlik ve Klinik Fayda

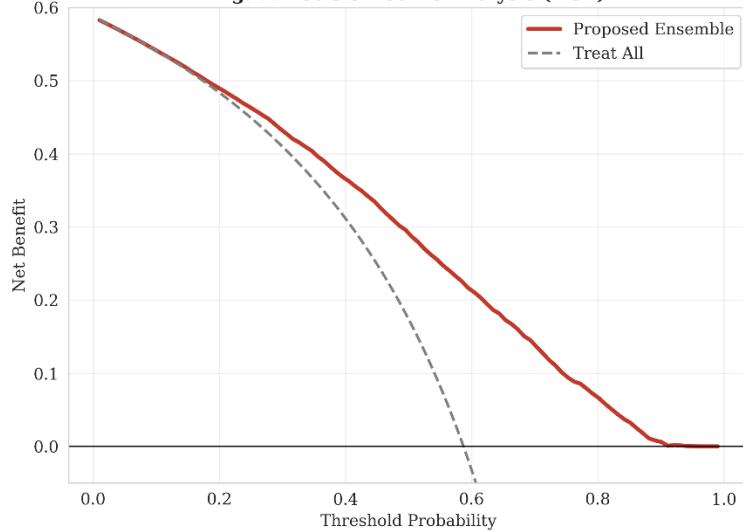
Doğru bir model, güvenilir değilse klinikte başarısız olur. Veri boşluklarını **Karankot ve ark.** [7] tavsiyesi üzerine MICE yöntemiyle yönettiğimiz için, model eksik verilerden kaynaklanan sapmalara maruz kalmamıştır. Şekil 6, modelimiz tarafından tahmin edilen olasılıkların gerçekle iyi bir uyum içinde olduğunu (iyi kalibrasyon) göstermektedir. İstatistiksel doğruluğun ötesinde, modelin ekonomik değerini Şekil 7'deki Karar Eğrisi Analizi (DCA) ile ölçtüük. Kırmızı çizgi (bizim modelimiz), klinik risk aralıklarında varsayılan stratejilerin ("herkesi tedavi et" veya "hiç kimseyi tedavi etme") üzerinde yer almaktadır. Bu, modelin triyaj için kullanılmasının, hastane için geleneksel yargılara kıyasla daha yüksek Net Fayda (Net Benefit) sağladığı anlamına gelir.

Fig. 6: Calibration Plot



Şekil 6. Kalibrasyon grafiği. Tahmin edilen olasılıkların gerçekle uyumu, modelin güvenilirliğini göstermektedir.

Fig. 7: Decision Curve Analysis (DCA)



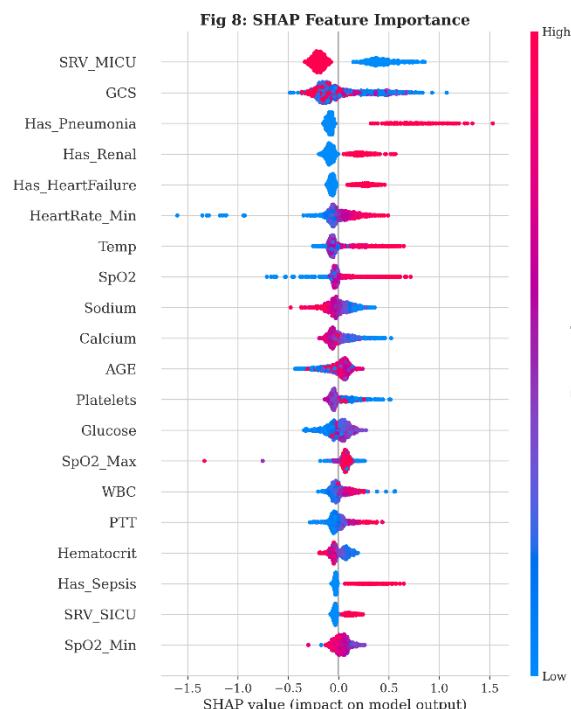
Şekil 7. Karar Eğrisi Analizi (DCA). Kırmızı çizgi, önerilen modelin geleneksel stratejilere kıyasla daha yüksek Net Fayda sağladığını göstermektedir.

4.4. Yorumlanabilirlik: "Hastalık Türü"nın Hayatı Belirtilere Üstünlüğü

Yaklaşımımızın, **Tanutsiriteeradet [1]** tarafından önerilen Transformerlar veya karmaşık derin ağlar [8] gibi "kara kutu" modellere göre temel avantajlarından biri şeffaflığıdır. SHAP analizi (Şekil 8) kullanarak modelin karar mantığını ortaya çıkardık. Bu yaklaşım, sepsiste ölüm faktörlerini keşfetmek için yorumlanabilirliği kullanan **Lu ve ark. [5]** metodolojisiyle uyumludur. Analiz sonuçlarımız şunları ortaya koydu:

1. **Pnömoni (Has_Pneumonia)**: En güçlü tahmin edici olarak belirlendi.
2. **Oksijen Seviyesi (SpO2_Max)**: İkinci sırada yer aldı.
3. **Kalp ve Böbrek Yetmezliği**: Sonraki sıralarda yer aldı.

Bu sıralama, hayatı belirtilerin önemli olmasına rağmen (ki **Sundas ve ark. [6]** buna işaret etmiştir), modelimizin "hastalık etiyolojisini" ana itici güç olarak tanımlayabildiğini göstermektedir; bu bulgu, **Alsinglawi ve ark. [3]**'nın kalbe özgü değişkenlerin önemi hakkındaki sonuçlarıyla tamamen örtüşmektedir.



Şekil 8. SHAP özet grafiği. Özellikle sıralaması, Pnömoni (Has_Pneumonia) ve oksijen seviyesinin en etkili faktörler olduğunu göstermekte ve modelin klinik geçerliliğini doğrulamaktadır.

5. Sonuç ve Gelecek Çalışmalar

5.1. Sonuç

Bu çalışma, YBÜ'de kalış süresini tahmin etme gibi hayatı bir zorluğu çözmek amacıyla, geleneksel regresyon modellemesinden (**Hasan ve ark.** [2] yaklaşımı gibi) "Güvenlik Öncelikli Risk Sınıflandırması" çerçevesine radikal bir geçiş önermiştir. İncelemelerimiz, doğrusal modellerin yoğun bakım verilerinin çarpık ve heterojen dağılımı karşısında yetersiz kaldığını doğrulamıştır. **Zheng ve ark.** [4] bulgularıyla uyumlu olarak, problemin sınıflandırmaya dönüştürülmesinin kaynak yönetimi için daha istikrarlı bir çözüm olduğunu gösterdik. Gelişmiş klinik bilginin özellikle akıllı hastalık bayraklarının çıkarılması ve fizyolojik göstergelerle (**Sundas ve ark.** [6] tarafından vurgulandığı üzere) birleştirilmesi entegrasyonu ile modelin tahmin gücünü artırılmıştır. Önerilen topluluk modeli, SMOTE tekniğinden (ki tıbbi verilerdeki etkinliği **Houfani ve ark.** [8] tarafından kanıtlanmıştır) yararlanarak ve eşiği 0.38 olarak ayarlayarak, %90.73 gibi parlak bir Duyarlılık (Recall) değerine ulaşmıştır. Bu başarı, sistemimizin güvenilir bir tarama aracı olarak çalışmasını ve uzun süreli bakıma ihtiyaç duyan neredeyse tüm hastaları tespit etmesini garanti eder. Ayrıca, yorumlanabilirlik analizimiz yeni bir içgörü sunmuştur: Yaygın kanının aksine, model "hastalık etiyolojisinin" (pnömoni gibi) uzun kalışın ana itici gücünü olduğunu göstermiştir; bu bulgu **Lu ve ark.** [5] ve **Alsinglawi ve ark.** [3]'nın hastalığa özgü değişkenlerin önemi konusundaki sonuçlarıyla tam bir uyum içindedir.

5.2. Sınırlamalar

Umut verici sonuçlara rağmen, bu çalışmanın kabul edilmesi gereken sınırlamaları vardır:

- **Tek Merkez Yanılılığı:** Modelimiz sadece MIMIC-III verileri kullanılarak geliştirilmiştir. **Fan ve ark.** [9] gibi öncü çalışmalar genelleştirilebilirliği sağlamak için birden fazla veri tabanı (MIMIC-IV ve eICU) kullanırken, mevcut modelimiz farklı coğrafi ortamlarda yeniden kalibrasyona ihtiyaç duyabilir.
- **Veri Doğası:** Ağırlıklı olarak yapılandırılmış (tablosal) verilere güvendik. Hemşire notlarında bulunan ve kötüleşmenin erken belirtilerini ortaya çıkarabilecek metinsel klinik nüanslar bu yinelemede göz ardı edilmiştir.

5.3. Gelecek Yollar

Gelecek araştırmalar iki ana eksene odaklanacaktır:

- 1. Doğal Dil İşleme (NLP) ile Entegrasyon:** Tanutsiriteeradet [1] çalışmasında YBÜ zaman serisi verilerinin analizinde Transformer mimarilerinin başarısı göz önüne alındığında, tıbbi raporlardan anlamsal sinyaller çıkarmak için bu mimarinin metinsel sürümlerini (ClinicalBERT gibi) kullanmayı planlıyoruz.
- 2. Klinik Dağıtım:** Modelin, hasta tedavisine doğrudan müdahale etmeden, doktorların karar alma akışı üzerindeki pratik etkisini değerlendirmek için hastane ortamında "Gölge Modu"nda (Shadow Mode) uygulanması.

6. References

- [1] S. Tanutsiriteeradet, W. Kumwilaisak, and K. Kumwilaisak, "ICU Bed Capacity Analysis with Transformer-based Length of Stay Prediction and Erlang Loss Formula," in *2024 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2024.
- [2] M. N. Hasan, S. Hamdan, S. Poudel, J. Vargas, and K. Poudel, "Prediction of Length-of-stay at Intensive Care Unit (ICU) Using Machine Learning based on MIMIC-III Database," in *2023 IEEE Conference on Artificial Intelligence (CAI)*, 2023, pp. 321-323.
- [3] B. Alsinglawi, F. Alnajjar, O. Mubin, M. Novoa, M. Alorjani, O. Karajeh, and O. Darwish, "Predicting Length of Stay for Cardiovascular Hospitalizations in the Intensive Care Unit: Machine Learning Approach," in *Proceedings of the 2022 IEEE International Conference on Digital Health (ICDH)*, 2022.
- [4] L. Zheng, J. Wang, A. Sheriff, and X. Chen, "Hospital Length of Stay Prediction with Ensemble Methods in Machine Learning," in *2021 International Conference on Cyber-Physical Social Intelligence (ICCSI)*, 2021.
- [5] X. Lu, J. Zhu, J. Gui, and Q. Li, "Prediction of All-cause Mortality with Sepsis-associated Encephalopathy in the ICU Based on Interpretable Machine Learning," in *2022 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2022.
- [6] A. Sundas, G. Singh, S. Badotra, A. Verma, and B. Kaur, "Optimizing Length of Stay Prediction After Intubation: An Advanced Machine Learning Model with Real-time Vital Sign Integration," in *2023 Seventh International Conference on Image Information Processing (ICIIP)*, 2023.
- [7] M. I. Karankot, M. Marceau, E. M. Glenn, R. P. Fowers, D. M. Hedges, B. Sheehey, and B. M. Whitaker, "Addressing the Challenge of Missing Medical Data in Healthcare Analytics: A Focus on Machine Learning Predictions for ICU Length of Stay," in *2024 Intermountain Engineering, Technology and Computing (IETC)*, 2024.
- [8] D. Houfani, S. Slatnia, O. Kazar, H. Saouli, H. S. Eddine, I. Remadna, and M. Zouai, "TabNet Based Prediction Model for ICU admission in Covid-19 patients," in *2022 International Symposium on iNnovative Informatics of Biskra (ISNIB)*, 2022.
- [9] G. Fan, A. Liu, and C. Zhang, "ContrastLOS: A Graph-Based Deep Learning Model With Contrastive Pre-Training for Improved ICU Length-of-Stay Prediction," *IEEE Access*, vol. 13, 2025.