



# **İST480 Araştırma Yöntemleri Dersi Final Ödevi**

**Doç.Dr.NİLGÜN ÖZGÜL ÇELEBİ**

**Konu: Makine Öğrenimi Algoritmaları ile Kronik Böbrek Hastalığı olan  
Kritik Hastalarda Hastane-içi Mortaliteyi Tahmin Etmek**

Ecem ÇIRAKOĞLU - 21821665

Oğuzcan ERTÜL - 21821814

Hacettepe Üniversitesi, İstatistik Bölümü

Haziran 4, 2023

## 1. Makalenin Yayınlandığı Dergi Hakkında Genel Bilgiler

## 2. Makale İncelemesi

- 2.1 Makale Hakkında Genel Bilgiler
- 2.2 Makalenin Metodolojisi
- 2.3 Sonuç ve Tartışma
- 2.4 Makale Hakkında Görüşler
  - 2.4.1 Makalenin Artı Yönleri
  - 2.4.2 Makalenin Eksi Yönleri

## 3. Kaynaklar

# Makalenin Yayınlandığı Dergi Hakkında Genel Bilgiler

---

**Dergi Adı:** Renal Failure

**Dergi İndeksi:** Web of Science-Science Citation Index Expanded(SCIE)

**Dergi Hakkında:**Renal Failure, akut böbrek hasarı ve sonuçları üzerine yayın yapan açık erişimli bir dergidir.Derginin birincil odak noktası akut böbrek hasarıdır.Kritik hastalarda akut böbrek hasarı sonuçlarını iyileştirdiğini gösteren yöntemleri ve ilgili araştırmaları yayınlamayı amaçlamaktadır. Renal Failure ayrıca kronik böbrek yetmezliği, hipertansiyon ve renal transplantasyon alanlarındaki gelişmeleri de yayınlamaktadır.

Dergi, böbrek yetmezliğinin klinik ve deneysel yönlerini bir araya getirerek çeşitli konularda güncel ve pratik bilgiler sunmayı amaçlamaktadır.Bu konulardan bazıları:

- Akut böbrek yetmezliğinin patolojisi ve patofizyolojisi
- İlaçların ve diğer maddelerin nefrotoksitesi
- Böbrek yetmezliğinin önlenmesi, tedavisi ve terapisi

**Derginin Erişim Linki:** <https://www.tandfonline.com/journals/irnf20>

# Makale Hakkında Genel Bilgiler

---

## Makale Adı

Kronik Böbrek Hastalığı olan Kritik Hastalarda Hastane-içi Mortaliteyi Tahmin Etmek için Makine Öğrenimi Algoritması

**Yazarlar:** Xunliang Li, Yuyu Zhu, Wenman Zhao, Rui Shi, Zhijuan Wang, Haifeng Pan ve Deguang Wang

**Makale Dili:** İngilizce

**Yayınlanma Tarihi:** 19 Mayıs 2023

**Makalenin Konusu:** Kronik böbrek hastalığı (KBH) olan kritik hastalarda hastane-içi mortalitenin tahmin edilebilmesi için makine öğrenimi (ML) modellerinin geliştirilmesi ve doğrulanmasıdır.

**Makalenin Amacı:** Geliştirilen ML modellerinin yüksek riskli KBH hastalarında mortaliteyi azaltmak için erken müdahalelerin yönetilmesi ve uygulanmasında klinisyenlere yardımcı olmaktadır.

# Makale Hakkında Genel Bilgiler

---

**Makalenin Önemi:** Kronik böbrek hastalığı (KBH), yetişkin nüfusun yaklaşık % 15'ini etkileyerek dünya çapında önde gelen 10 ölüm nedeni arasında yer almaktadır.[1]. Bu nedenle çalışma, yoğun bakımda yatan kritik hastaların mortalite riskini erken tahmin etmenin önemini vurgulamaktadır. Erken tahmin, uygun önleyici tedbirlerin alınmasını ve tedavi stratejilerinin optimize edilmesini sağlayabilir. Böylece hastaların sonuçlarını iyileştirebilir ve mortalite oranlarını azaltabilir. Bu doğrultuda elektronik sağlık kayıtlarındaki büyük miktarda veriyi analiz ederek hasta sonuçlarını iyileştirmek için makine öğrenimi (ML) algoritmaları geliştirilmiştir.

Geleneksel modelleme tekniklerinin aksine, makine öğrenimi yöntemleri daha esnektir ve değişkenlerin birbirinden bağımsız olmasını gerektirmek yerine istatistiksel anlamlılıklarına veya klinik önemlerine göre kullanabilir. Çalışmanın sonucunda, geliştirilen ML modellerinin özellikle de XGBoost modelinin, KBH'li kritik hastalarda mortaliteyi etkili bir şekilde erken tahmin edebileceğini vurgulanmaktadır.

# Makalenin Metodolojisi

---

## **Çalışma Popülasyonu ve Örneklemi**

Medical Information Mart for Intensive Care IV (MIMIC-IV) veritabanı, Massachusetts teknoloji enstitüsü tarafından onaylanmış kapsamlı, anonimleştirilmiş bir klinik veri tabanıdır [2]. MIMIC-IV veritabanı, 2008 ve 2019 yılları arasındaki tüm Beth Israel Deaconess Tıp merkezi yoğun bakım ünitesindeki hastalara ilişkin verileri içermektedir.

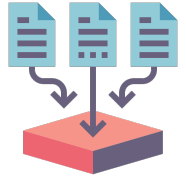
Veritabanındaki tüm hastalar anonim olduğundan ve klinik karar verme üzerinde hiçbir etkisi olmadığından, hasta onayı ve etik onay gerekmemiştir. Yazarlardan biri (Xunliang Li), Ulusal Sağlık Enstitüleri'nin İnsan Araştırma Katılımcılarını Koruma sınavını geçerek bu veritabanına izinli erişim elde etmiştir. Bu nedenle bu araştırma MIMIC-IV'e kayıtlı olan ve KBH tanısı konmuş hastaları kapsamaktadır.

# Makalenin Metodolojisi

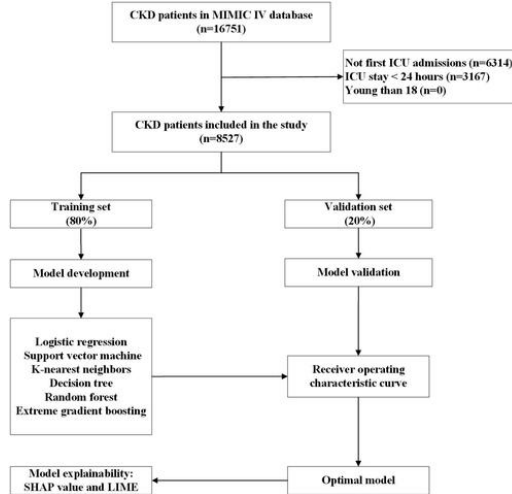
---

## Veri Toplama Yöntemi

- Navicat Premium[3] kullanılarak MIMIC-IV veri tabanından demografik ve klinik veriler çekilmiştir. Yaş, cinsiyet, kilo, etnik köken ve kabul tipi gibi demografik faktörler; konjestif kalp yetmezliği, peptik ülser hastalığı, miyokard enfarktüsü, periferik vasküler hastalık, diyabet, kronik pulmoner hastalık, romatizmal hastalık, serebrovasküler hastalık, kanser, parapleji, karaciğer hastalığı ve edinilmiş immün yetmezlik sendromu gibi tıbbi faktörler toplanmıştır.
- Toplam 16.751 kişinin KBH'si olduğu bulunmuştur ancak 6314 hasta ilk yoğun bakıma kabul edilmediği için ve 3167 hasta da yoğun bakımda 24 saatten az bulunması sebebiyle diskalifiye edilmiş ve Şekil 1'de gösterildiği gibi 8527 hasta çalışma için uygun bulunmuştur.



# Makalenin Metodolojisi



Şekil 1. Hasta seçimi akış şeması. **MIMIC IV**: Yoğun Bakım IV Tıbbi Bilgi Mortu; **ICU**: yoğun bakım ünitesi; **KBH**: kronik böbrek hastalığı; **LIME**: Yerel Yorumlanabilir Model-Agnostik Açıklamalar



# Makalenin Metodolojisi

---

## Uygulanan Yöntemler

- Çalışmada eksik veriler, çoklu interpolasyon yöntemi kullanılarak ele alınmıştır.
- Sürekli değişkenler medyan ve çeyrekler arası aralık (IQR) olarak ifade edilmiş ve normal dağılım göstermemeleri nedeniyle gruplar arasındaki farklılıkları belirlemek için Mann-Whitney testi kullanılmıştır.
- Kategorik değişkenler sayı ve yüzde olarak ifade edilmiş ve grup karşılaştırmaları, duruma göre Ki-kare testi veya Fisher'in kesin testi kullanılarak yapılmıştır.



İstatistiksel analizler, R yazılımı ve Python kullanılarak yapılmıştır.  
 $p < 0.05$  değeri istatistiksel olarak anlamlı kabul edilmiştir.

# Makalenin Metodolojisi

---

## Makine Öğrenimi

Makine öğrenimi algoritmalarını kullanarak kritik hastalarda ölüm riski modelleri geliştirmeyi ve doğrulamayı amaçlamaktadır. Beş katlı bir çapraz doğrulama yöntemi kullanılarak eğitim ve doğrulama veri setleri oluşturulmuştur. Lojistik regresyon, destek vektör makinesi (SVM), k-en yakın komşu (KNN), karar ağacı, rastgele orman (RF) ve eXtreme Gradient Boosting (XGBoost) gibi çeşitli algoritmalar kullanılmıştır.

Çalışma, her bir algoritmanın varsayılan hiperparametreleriyle başlamış ve daha sonra manuel olarak hiperparametreleri ayarlamak için bir grid arama yöntemi kullanmıştır. Optimal hiperparametre ayarları, on kat çapraz doğrulama kullanılarak belirlenmiştir. Modellerin performansı doğruluk, eğri altındaki alan (AUC), duyarlılık ve seçicilik gibi metrikler kullanılarak değerlendirilmiştir.

# Makalenin Metodolojisi

---

Sonuçlar, en iyi performans gösteren modelin doğruluk ve AUC metriklerine göre belirlendiğini göstermektedir. Ayrıca, Brier puanları, kalibrasyon eğrileri ve yeniden sınıflandırma iyileştirmesi gibi değerlendirme yöntemleri kullanılarak modeller arasında karşılaştırmalar yapılmıştır. SHAP ve LIME algoritmalarıyla en iyi modelin yorumlanabilirliği analiz edilmiştir.



# Sonuç ve Tartışma

---

## Sonuç ve Tartışma

Yoğun bakım ünitesine yatırılan KBH hastaları arasında hastane içi ölüm oranı %16 olarak hesaplanmıştır. Bu hastaların ortalama yaşı 75,1 yıl ve %61,7'si erkektir. Konjestif kalp yetmezliği (%53,3), diabetes mellitus (%49,5) ve sepsis (%41,2) en yaygın görülen ilk üç komorbidite olmuştur.

Klinik değişkenler bağımsız değişken olarak kullanılarak lojistik regresyon, SVM, KNN, karar ağacı, RF ve XGBoost dahil olmak üzere altı ML modeli geliştirilmiş ve diğer ML modelleriyle karşılaştırıldığında, XGBoost modeli 0.860 AUC ve 0.860 doğruluk oranı ile en iyi performansı göstermiştir.(Şekil 2)

SOFA ve SAPS II skorları gibi geleneksel ciddiyet skorum sistemlerini ML modellerine kıyasla düşük performans göstermektedir(Şekil 3).

# Sonuç ve Tartışma

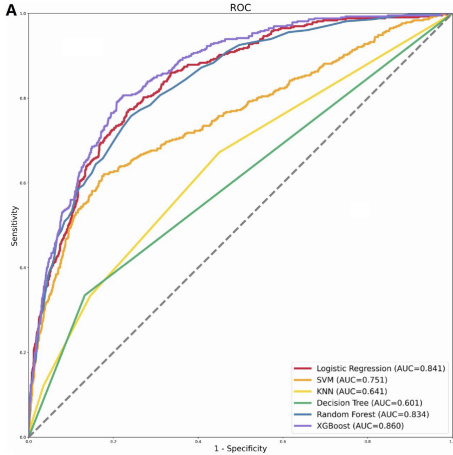


Figure: Şekil 2

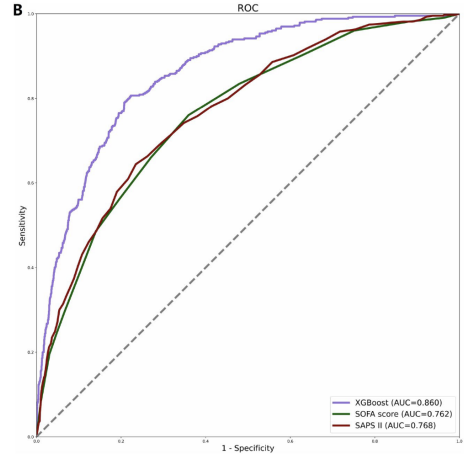


Figure: Şekil 3

# Sonuç ve Tartışma

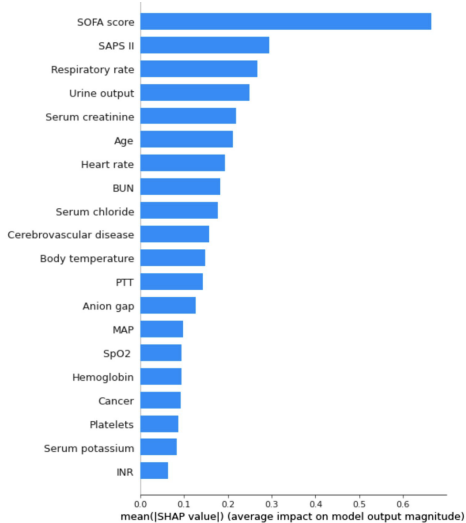


Figure: Şekil 4

Oluşturulan SHAP özet grafiğine göre SOFA skoru, SAPS II, solunum hızı ve idrar çıkışı gibi değişkenler, KBH'li kritik hastalarda mortaliteyi öngörmeye katkı sağlayan en önemli dört faktör olarak belirlenmiştir(Şekil 4).

Not: SOFA skoru, organ disfonksiyonunun varlığını tanımlayan bir araçtır. Altı organ sisteminin her birine organ yetmezliğinin ciddiyetine göre 1 ila 4 arasında günlük bir puan verir ve yüksek değerler daha ciddi organ işlev bozukluğuna işaret eder[4]. Benzer şekilde SAPS II skoru da hastalığın ciddiyetini gösterir, yüksek skorlar hastalığın ciddiyetini ifade eder.

## Sonuç ve Tartışma

---

SOFA ve SAPS II skorları, XGBoost modeline göre zayıf performans göstermektedir. Bu durum, sistemlerin, KBH'li kritik hastalarda tek başlarına ölümü öngörmek için güvenilir araçlar olmayabileceğini göstermektedir. SOFA skoru, SAPS II, solunum hızı ve idrar çıkışı gibi değişkenler, KBH'li kritik hastalarda mortaliteyi öngörmeye katkı sağlayan en önemli dört faktör olarak belirlenmiştir. Bu bilgi, hastaların klinik durumlarının değerlendirilmesi ve tedavi planlarının yapılması için önemli bir rehber sağlayabilmektedir.

Sonuç olarak, kullanılan XGBoost modeli, diğer modellere ve geleneksel ciddiyet skora sistemlerine kıyasla daha iyi performans göstermektedir. Böylece klinisyenlerin erken müdahaleleri doğru bir şekilde yönetmesine ve uygulamasına yardımcı olabilecek en etkili ML modelidir, bu da yüksek ölüm riski olan kritik KBH hastalarında mortaliteyi azaltabilir.

# Makale Hakkında Görüşler

---

## Makalenin Artı Yönleri:

- Makale, yoğun bakımdaki KBH'li kritik hastalarda hastane içi mortaliteyi öngörmek için altı farklı makine öğrenimi modeli kullanarak kapsamlı bir analiz sunmaktadır. Bu, KBH'de ölüm riskini tahmin etmek için farklı yaklaşımların karşılaştırılmasını sağlıyor.
- Araştırma, XGBoost modelindeki özelliklerin önem sıralamasını sunarak, SOFA skoru, SAPS II, solunum hızı ve idrar çıkışının mortaliteyi öngörmek için en önemli değişkenler olduğunu gösteriyor. Bu, kritik hastalarda organ disfonksiyonu, hastalık şiddeti ve solunum ve böbrek fonksiyonları gibi faktörlerin önemini vurguluyor.
- Sonuçlar, önceki çalışmaların sonuçlarıyla uyumludur ve XGBoost modelinin diğer makine öğrenimi modellerine göre KBH hastalarında mortaliteyi tahmin etmede daha iyi performans gösterdiğini doğrulamaktadır.



# Makale Hakkında Görüşler

---

## Makalenin Eksi Yönleri:

- Retrospektif veri analizi, belirli bir veri tabanındaki mevcut verilere dayanır ve bu veriler üzerinde yapılan seçimler analizin sonuçlarını etkileyebilir. Bu tür bir seçim yanlılığı, sonuçların genellemesini sınırlayabilir.
- Çalışma, MIMIC IV veritabanındaki verilere dayandığı için bu verilerin kapsamı ve doğruluğu sınırlamalara tabi olabilir. Veri setindeki eksiklikler veya yanlılık, sonuçların güvenilirliğini etkileyebilir.
- Bu araştırma tek bir merkezde yürütülmüştür. Sonuçların genellemesi için birden fazla merkezde yapılan çalışmalara ihtiyaç olabilir. Farklı merkezlerde farklı hasta popülasyonları ve uygulamalar olabileceğinden, sonuçların evrensel olarak geçerli olup olmadığı belirsizlik gösterebilir.
- Yoğun bakımdaki KBH hastalarında hastane içi mortaliteyi tahmin etmek için bir ML modeli oluşturmayı amaçlamaktadır. Bu nedenle, modelin diğer hasta grupları veya sağlık koşulları için genelleştirilebilirliği belirsiz olabilir.

# Kaynaklar

---

- [1] Hypoxia-inducible factor signaling in vascular calcification in chronic kidney disease patients J Nephrol. 2022;35(9):2205–2213.
- [2] MIMIC-IV database (2021) A retrospective analysis from MIMIC-IV database Ann intensive care. 2021;11(1):42.
- [3] Navicat Premium'a bu adresten ulaşabilirsiniz:  
<https://www.navicat.com/en/products/navicat-premium>.
- [4] Evaluation of SOFA-based models for predicting mortality in the ICU(2008)
- [5] İncelenen Makaleye şu adresten ulaşabilirsiniz:<https://www.tandfonline.com/doi/full/10.1080/0886022X.2023.2212790>.

# Teşekkürler!

