Fırat Üniversitesi Müh. Bil. Dergisi
38(1), 285-298, 2026

*Araştırma Makalesi*
https://doi.org/10.35234/fumbd.1805121

# Üniversite Öğrencilerinde Stres Düzeyinin Makine Öğrenmesiyle Tahmini: Sentetik Veri Destekli Optimizasyon Yaklaşımı

**Emre ÇANAYAZ**[1*]
[1] Elektronik ve Otomasyon Bölümü, Teknik Bilimler Meslek Yüksekokulu, Marmara Üniversitesi, İstanbul, Türkiye
[*1] emre.canayaz@marmara.edu.tr

**(Geliş/Received: 16/10/2025; Kabul/Accepted: 22/02/2026)**

**Öz:** Üniversite öğrencileri, akademik baskı ve sosyal zorluklar nedeniyle yüksek stres riski taşımaktadır. Bu çalışma, 2.000 öğrenciden elde edilen öz-bildirim temelli yaşam tarzı verilerini kullanarak stres düzeylerini makine öğrenmesi ile tahmin etmeyi amaçlamaktadır. Veri setindeki sınıf dengesizliği, eğitim aşamasında Koşullu Tablo Üretici Üretken Karşıt Ağ yöntemiyle üretilen sentetik verilerle giderilmiş; modeller tabakalı çapraz doğrulama ile test edilmiştir. Karşılaştırılan beş algoritma arasında Rastgele Orman modeli hem dengeli hem dengesiz veri setlerinde tüm örnekleri doğru sınıflandırarak en yüksek performansa ulaşmıştır (Doğruluk = 1.00; AUC = 1.00; Makro F1 = 1.00). Modelin karar mekanizması Shapley analizi ile incelenmiş; "Günlük Çalışma Süresi" ve "Uyku Süresi" en belirleyici faktörler olarak saptanmıştır. Elde edilen yüksek başarımın veri sızıntısı ya da rastlantısal uyumdan kaynaklanma olasılığı, bağımsız sınama kümesinin süreç boyunca ayrı tutulması ve sağlamlık kontrolleri olarak uygulanan ablasyon (özellik çıkarma) ile etiket permütasyon testi bulguları ile desteklenmiştir. Sonuçlar, sentetik veriyle desteklenen ve açıklanabilir yapay zeka ile doğrulanan modellerin, öğrencilerin stres düzeylerinin erken tespitinde güvenilir bir araç olduğunu göstermektedir.

**Anahtar kelimeler:** Stres düzeyi tahmini, makine öğrenmesi, sentetik veri, eğitim psikolojisi

# Predicting Stress Levels in University Students Using Machine Learning: An Optimization Approach Based on Synthetic Data

**Abstract:** University students are at high risk of stress due to academic pressure and social challenges. This study aims to predict stress levels with machine learning using self-report-based lifestyle data from 2,000 students. The class imbalance in the dataset was eliminated in the training phase with synthetic data generated by the Conditional Table Generating Generative Adversarial Network method, and the models were tested with stratified cross-validation. Among the five algorithms compared, the Random Forest model achieved the highest performance (Accuracy = 1.00; AUC = 1.00; Macro F1 = 1.00), correctly classifying all instances in both balanced and imbalanced data sets. The decision mechanism of the model was analyzed by Shapley analysis; “Daily Working Time” and “Sleeping Time” were found to be the most determining factors. The possibility that the high performance obtained was due to data leakage or random fit was supported by the fact that the independent test set was kept separate throughout the process and the findings of the ablation (feature extraction) and label permutation tests applied as robustness checks. The results suggest that models supported by synthetic data and validated by explainable artificial intelligence are a reliable tool for early detection of students' stress levels.

**Keywords:** Stress level prediction, machine learning, synthetic data, educational psychology

## 1. Giriş

Üniversite öğrencileri arasında ruhsal sağlık sorunlarının artan yaygınlığı, yükseköğretim kurumları için kritik bir sorun alanı haline gelmiştir. Bu durum, geniş ölçekte uygulanabilecek erken müdahale yaklaşımlarının geliştirilmesini gerekli kılmaktadır. Stres, psikolojik rahatsızlıkların temel tetikleyicilerinden biri olarak tanımlanmakta; kaygı ve depresyon gibi birçok bozukluğun şiddetini ve seyrini olumsuz etkileyerek öğrencilerin akademik performansını, sosyal etkileşimini ve genel ruhsal iyilik halini azaltmaktadır [1]. Epidemiyolojik bulgular, yaşam boyu görülen ruhsal bozuklukların büyük çoğunluğunun 24 yaşından önce başladığını göstermekte; bu durum, geleneksel üniversite öğrencisi yaş aralığı ile örtüşmektedir [2].

Pandemi dönemlerinde uygulanan kapanma önlemleri ve eğitimde yaşanan hızlı geçiş süreçleri gibi koşullarda yürütülen uluslararası araştırmalar, öğrencilerde ruhsal sıkıntı düzeylerinin belirgin biçimde arttığını ortaya koymuştur [3]. Örneğin El Morr ve ark. [4], ABD’de yaptıkları çalışmada pandemi sürecinde öğrencilerin %48,14’ünde orta şiddetli depresyon, %38,48’inde kaygı saptandığını; %71,26’sının stres düzeyinde artış bildirdiğini raporlamıştır. Bu tür bulgular, öğrencilerde stres yükünün dönemsel olarak artabildiğini ve akademik süreçlerle yakından ilişkili olabileceğini düşündürmektedir. Nitekim stres etkenleri çoğu zaman akademik aşırı

[*] Sorumlu yazar: emre.canayaz@marmara.edu.tr. Yazarların ORCID Numarası: [1] 0000-0002-3695-3642

yüklenme, ekonomik baskılar ve iş özerkliğinin sınırlılığı gibi kurum kaynaklı faktörlerle ilişkilendirilmektedir [5].

Stresin yaygın ve çok boyutlu doğası dikkate alındığında, otomatik stres tespiti yapan modeller akademik ortamlarda erken uyarı sistemlerinin tasarımı açısından umut verici bir yaklaşım sunmaktadır. Makine öğrenmesi (ML) yöntemleri, bu tür sistemlerin geliştirilmesinde önemli bir rol oynamaktadır. Stres ve diğer ruhsal sağlık göstergeleri; uyku süresi, akademik eğilim, internet kullanımı süresi ve davranışsal veriler gibi çok sayıda öznitelik üzerinden modellenebilmektedir.

Literatürde, üniversite öğrencilerinde stres ve ilişkili ruhsal sağlık göstergelerini tahmin etmeye yönelik farklı ML yaklaşımlarının değerlendirildiği çalışmalar bulunmaktadır. Ahuja ve Banga [6], sınav dönemlerinde öğrencilerin zihinsel stresini tespit etmek amacıyla Destek Vektör Makineleri (SVM), Rastgele Orman (RF), Naive Bayes (NB) ve Lojistik Regresyon (LR) gibi algoritmaları karşılaştırmış; SVM'in %85,71 doğruluk ile en başarılı sonuç verdiğini raporlamıştır. Benzer şekilde Nayan ve ark. [7], COVID-19 pandemisinin ilk dalgasında öğrencilerde depresyon ve kaygı düzeylerini tahmin etmek için RF, SVM, LR, k-En Yakın Komşu (k-NN) ve NB modellerini değerlendirmiş; RF ve SVM'nin sırasıyla %89 ve %91,49 doğruluk ile daha yüksek performans gösterdiğini bildirmiştir.

Medikonda [8], çevrimiçi anketlerle toplanan gerçek dünya verileri üzerinden algılanan stres düzeylerini tahmin etmeye yönelik bir ML modeli geliştirmeyi amaçlamış; Başlıca Bileşen Analizi (PCA) ve Ki-kare testi gibi özellik seçimi yöntemlerini, Izgara Arama Çapraz Doğrulama (Grid Search Cross-Validation) ve Genetik Algoritmalar ile birlikte kullanmıştır. Çalışmada %80,5 doğruluk, %1,000 kesinlik, %0,826 duyarlılık ve %0,890 F1-skoru elde edilmiş; bu bulgular özellik seçimi ile optimizasyonun model başarımına katkısını vurgulamıştır. Aynı yıl Baba ve Bunji [9], öğrenci anket verilerini yanıt süreleri ile birlikte kullanarak ruhsal sağlık problemlerini tahmin etmeye çalışmış; LightGBM (LGBM), XGBoost (XGB), ElasticNet ve LR modelleri ile zihinsel sağlık verilerindeki karmaşık örüntülerin yakalanabileceğini göstermiştir.

Ratul ve ark. [10], davranışsal ve sosyal değişkenleri içeren bir veri kümesi üzerinden stres düzeylerini sınıflandırmaya yönelik bir yaklaşım sunmuş; MLP, RF, Gradyan Artırma Sınıflandırıcısı (GBC), XGB ve AdaBoost modellerini değerlendirmiştir. Çalışmada MLP ve RF'nin sırasıyla %80,5 doğruluk ve %0,89 F1-skoru ile daha iyi sonuçlar verdiği bildirilmiştir. El Morr ve ark. [4] ise COVID-19 pandemisi sürecinde Lübnanlı öğrencilerin Depresyon, Anksiyete ve Stres Ölçeği (DAS) skorlarını tahmin etmek amacıyla RF, NB, MLP, AdaBoost, XGB, k-NN ve LR modellerini karşılaştırmış; RF'nin %78,27 AUC değeriyle en başarılı model olduğunu göstermiştir.

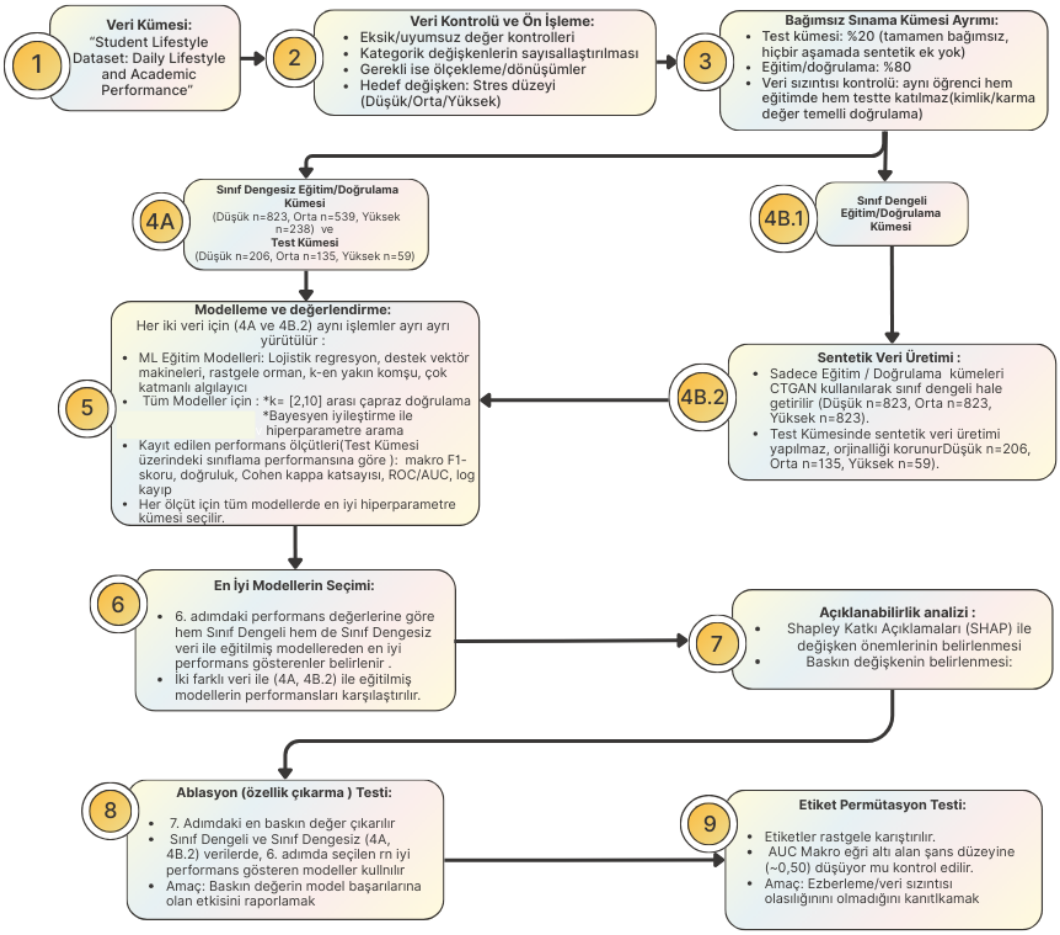
Zhang ve ark. [11], tedarik zinciri yönetimi (SCM) öğrencilerinde stres düzeylerini tahmin etmek için Transformer mimarisi tabanlı bir derin öğrenme modeli önermiş; bu yaklaşımın klasik makine öğrenmesi yöntemlerine kıyasla daha düşük hata ve daha yüksek açıklayıcılık sağladığını raporlamıştır. Dudić [12], çalışma ve uyku süresi gibi değişkenlere dayalı olarak öğrenci stresini tahmin etmek için karar ağaçları (DT) ve LR modellerini karşılaştırmış; karar ağaçlarının tüm stres düzeylerinde mükemmel doğruluk ve duyarlılıkla daha iyi performans gösterdiğini belirtmiştir. Ghara ve ark. [13], sınav dönemlerinde öğrencilerin stres düzeylerini tespit etmek için uyku düzeni ve çalışma alışkanlığı gibi yaşam tarzı özelliklerini içeren bir yaklaşım geliştirmiş ve RF algoritmasının daha başarılı olduğunu bildirmiştir. Bastos ve ark. [14], COVID-19 pandemisi sırasında lisansüstü öğrencilerde akademik stres faktörlerinin depresif belirtileri ne ölçüde etkilediğini incelemiş; Epsilon-Destek Vektör Regresyonu (Ɛ-SVR) kullanarak $r = 0,51$ ve $R^2 = 0,26$ elde etmiş, akademik stresörlerin depresif belirtiler üzerindeki belirleyici rolünü vurgulamıştır. Tuan ve ark. [15], RFE/ET/Boruta tabanlı özellik seçimi ile birden çok sınıflandırıcıyı entegre eden bir çerçeve önermiş; değerlendirmede 10-kat çapraz doğrulama ve TOPSIS'i kullanmıştır. Çalışmada RF, DT, MLP ve Gradient Boosting %100 doğruluk elde ederken, RF-Boruta konfigürasyonu en yüksek TOPSIS skoru ile öne çıkmıştır. Rahunathan ve ark. [16], ise PSS ve DASS-21 tabanlı iki veri kümesinde NB ve k-NN'i karşılaştırmış; NB'nin her iki kümede de daha yüksek doğruluk verdiğini raporlamıştır.

Bu çalışma, üniversite öğrencilerinde stres düzeyi sınıflandırmasına yönelik yaklaşımları iki yönden genişletmektedir. İlk olarak, özgün veri kümesinde gözlenen sınıf dengesizliğini gidermek için eğitim kümesi üzerinde koşullu sentetik veri üretimi uygulanmış; böylece dengelenmemiş özgün veri ile sentetik olarak dengelenmiş veri üzerinde eğitilen modellerin başarımı doğrudan ve karşılaştırmalı biçimde değerlendirilmiştir. İkinci olarak, model başarımı tek bir doğrulama düzenine indirgenmeden farklı kat sayılarında çapraz doğrulama ($k = 2–10$) ile sınanmış; makro F1-skoru temel ölçüt olmak üzere doğruluk, Cohen kappa katsayısı, eğri altı alan ve log kayıp gibi tamamlayıcı ölçütlerle yöntemsel sağlamlık ortaya konmuştur. Ayrıca, çok yüksek başarım elde edilen durumlarda olası veri sızıntısı veya ezberleme olasılığı açıklanabilir yapay zeka temelli değişken katkı analizi ile incelenmiş; baskın değişkenin etkisi çıkarım deneyiyle doğrulanmış ve etiket permütasyon testi ile başarımın şans düzeyine gerilediği gösterilerek bulguların veri sızıntısından değil öğrenmeden kaynaklandığı

desteklenmiştir. Bu yönleriyle çalışma, öğrenci stresi tahmini literatüründe sınıf dengesizliğinin sentetik veri ile ele alınması ve bulguların açıklanabilirlik ile sağlamlık testleri üzerinden doğrulanmasına dayalı bütüncül bir değerlendirme çerçevesi sunmaktadır.

## 2. Metot

Bu bölümde, öğrencilerin yaşam tarzı ve akademik alışkanlıklarından türetilen çok değişkenli verilerin, makine öğrenmesi temelli stres düzeyi tahmin modellerine dönüştürülme süreci ayrıntılı biçimde sunulmaktadır. Şekil 1, önerilen sistemin genel iş akışını göstermekte olup; veri ayrımı (bağımsız test kümesinin en başta ayrılması), ön işleme, sentetik veri üretimi, model eğitimi, doğrulama, açıklanabilirlik ve sağlamlık kontrolleri adımlarını bütüncül bir yapıda özetlemektedir.




**Şekil 1:** Üniversite öğrencilerinde stres düzeyi tahminine yönelik önerilen makine öğrenmesi iş akış diyagramı.

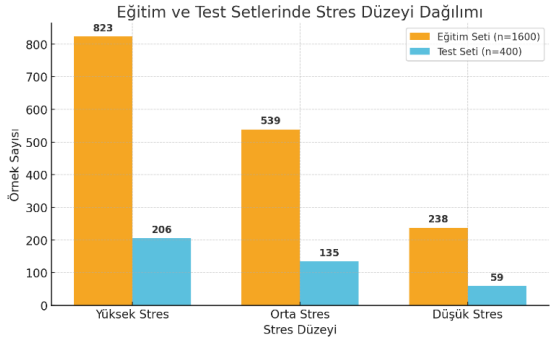
Çalışmanın metodolojik çerçevesi beş temel aşamadan oluşmaktadır: (1) kamuya açık kaynaktan sağlanan veri kümesinin [17] hazırlanması ve tanımlayıcı istatistiklerinin raporlanması, (2) veri dengesizliğinin giderilmesi amacıyla yalnızca eğitim/doğrulama kümesi üzerinde Koşullu Tablo Üretici Ağ (CTGAN) tabanlı sentetik veri üretimi ve bağımsız test kümesine hiçbir aşamada sentetik örnek eklenmemesi, (3) dengelenmiş ve dengelenmemiş veri kümeleri üzerinde Bayesyen optimizasyonla desteklenmiş gözetimli makine öğrenmesi modellerinin eğitimi ve çoklu ölçütlerle değerlendirilmesi, (4) modelin karar mekanizmasının açıklanması için Shapley katkı değerleri (SHAP) analizi ile değişken katkılarının incelenmesi ve baskın değişkenin etkisinin çıkarım deneyiyle sınanması, (5) etiket permütasyon testi ile başarımın şans düzeyine gerilediğinin gösterilerek ezberleme veya veri sızıntısı olasılığının dışlanması. Bu yapı sayesinde hem veri kalitesi hem de model güvenilirliği sistematik olarak test edilmekte; ayrıca geliştirilen yaklaşımın genellenebilirliği ve açıklanabilirliği ampirik olarak doğrulanmaktadır.

### 2.1 Veri kümesinin tanımı

Bu çalışmada kullanılan birincil veri kümesi, kamuya açık olarak erişilebilen *"Daily Lifestyle and Academic Performance of Students"* adlı veri setidir [17]. Bu çalışma kapsamında veri kümesi, Kaggle platformunda kamuya açık biçimde paylaşılan kaynaktan edinilmiş; veri toplama süreci çalışmanın araştırmacısı tarafından yürütülmemiştir. Veri seti, 2023 Ağustos–2024 Mayıs akademik yılı boyunca çevrim içi Google Form anketi aracılığıyla toplanan 2.000 lisans öğrencisinin öz-bildirime dayalı yanıtlarından oluşmaktadır (Tablo 1). Katılımcıların büyük çoğunluğunun Hindistan'daki üniversitelerde öğrenim gördüğü bildirilmektedir; bu yönüyle veri seti, gelişmekte olan bir ülkedeki geniş bir öğrenci kitlesinden elde edilmiş gerçek dünya örneklemini temsil etmektedir. Veri setinde her bir satır tek bir öğrenciye karşılık gelmekte olup, aynı öğrenciye ait birden fazla kayıt bulunmamaktadır. Veri kümesi, tahmin değişkenleri olarak kullanılan altı adet öz-bildirime dayalı yaşam tarzı ve akademik değişken içermektedir:

1. Günlük çalışma saati,
2. Ders dışı etkinlik (kulüp, hobi vb.) süresi,
3. Uyku süresi,
4. Sosyal etkileşim süresi,
5. Fiziksel aktivite süresi ve
6. Kümülatif not ortalaması (GPA).

Hedef değişken Stres Düzeyi olup öğrencilerin öz değerlendirmelerine göre düşük, orta ve yüksek stres olmak üzere üç sınıfta tanımlanmıştır. Bu çalışma kapsamında hedef değişken, yararlanılan özgün veri setinde sunulduğu biçimiyle doğrudan kategorik etiketler olarak ele alınmıştır. Dolayısıyla, herhangi bir puanlama üzerinden tarafımızca gerçekleştirilen bir sınıf türetme, yeniden sınıflandırma ya da belirli eşik değerlere göre kategorize etme süreci uygulanmamıştır. Veri seti, katılımcıların doğrudan bu üç kategorik seviyeden birini seçtiği orijinal yapısıyla analizlere dahil edilmiştir. Veri setinin özgün sınıf dağılımı belirgin biçimde dengesizdir; katılımcıların %51,4'ü (n = 1029) yüksek stres, %33,7'si (n = 674) orta stres ve %14,8'i (n = 297) düşük stres düzeyinde olduklarını bildirmiştir (Şekil 2).




**Şekil 2:** Veri setindeki hedef değişkenin (stres düzeyi) sınıflara göre dağılımı.

**Tablo 1.** ML eğitimi için kullanılan verilerini istatistiksel özeti

| Değişken Adı | İstatistik Özeti |
|---|---|
| Günlük çalışma süresi (saat) | Ortalama = 7.48 (Standart Sapma = 1.42), Medyan = 7.40, Min–Max = 5.00–10.00, n = 2000 |
| Ders dışı etkinlik süresi (saat) | Ortalama = 2.33 (Standart Sapma = 0.96), Medyan = 2.20, Min–Max = 1.00–5.00, n = 2000 |
| Uyku süresi (saat) | Ortalama = 6.90 (Standart Sapma = 1.20), Medyan = 7.00, Min–Max = 4.00–10.00, n = 2000 |
| Sosyal etkileşim süresi (saat) | Ortalama = 2.66 (Standart Sapma = 1.01), Medyan = 2.60, Min–Max = 1.00–5.00, n = 2000 |
| Fiziksel aktivite süresi (saat) | Ortalama = 2.52 (Standart Sapma = 0.99), Medyan = 2.50, Min–Max = 1.00–5.00, n = 2000 |
| Kümülatif not ortalaması (GPA) | Ortalama = 7.16 (Standart Sapma = 0.87), |

| | |
|---|---|
| | Medyan = 7.20, Min–Max = 5.00–10.00, n = 2000 |
| Stres düzeyi | Düşük n=297, Orta n=674, Yüksek n=1029 |

Model geliştirme ve değerlendirme sürecinde, veri kümesi stres düzeyi sınıfları korunacak biçimde iki parçaya ayrılmıştır: verilerin %80'i eğitim/doğrulama kümesi, %20'si ise bağımsız sınama (test) kümesi olarak belirlenmiştir (Şekil 1). Veri sızıntısını önlemek amacıyla her bir öğrenci kaydı yalnızca tek bir kümeye atanmış; aynı kaydın eş zamanlı olarak birden fazla kümede yer almasına izin verilmemiştir (Şekil 1). Bu ayrım sonucunda eğitim/doğrulama kümesinde sınıf sayıları yüksek (n = 823), orta (n = 539) ve düşük (n = 238) olarak oluşmuştur. Bağımsız sınama kümesi ise veri setinin özgün dağılımını yansıtacak şekilde yüksek (n = 206), orta (n = 135) ve düşük (n = 59) örnek içermektedir.

Bu çalışmada, sınıf dengesizliğinin model başarımına etkisini doğrudan inceleyebilmek için iki ayrı veri düzeni tanımlanmıştır. Birinci veri düzeni (dengesiz veri), yukarıdaki ayrım sonrası elde edilen ve sınıf dağılımı korunmuş eğitim/doğrulama kümesidir. İkinci veri düzeni (dengeli veri) ise yalnızca eğitim/doğrulama kümesi üzerinde Koşullu Tablo Üretici Üretken Karşıt Ağ (CTGAN) yaklaşımı [18] kullanılarak sınıfların dengelendiği kümedir. Bu yaklaşımın temel gerekçesi, dengeleme işleminin yalnızca eğitim/doğrulama aşamasında uygulanması ve böylece gerçek dünyaya daha yakın koşulları temsil eden bağımsız sınama kümesinin tarafsız biçimde korunmasıdır. Bu nedenle, bağımsız sınama kümesine hiçbir aşamada sentetik örnek eklenmemiş, sınama kümesi özgün haliyle bırakılmıştır; amaç, sınama başarımının sentetik örneklerin oluşturabileceği iyimser yanlılıktan etkilenmesini engellemektir.

Sınıf dengesizliğini gidermek ve sınıflar arası temsil sorununu azaltmak amacıyla CTGAN modeli, sayısal ve kategorik veri tiplerini birlikte işleyebilen ve değişkenler arasındaki karmaşık ortak dağılımları koruyarak yeni örnekler üretebilen derin üretici bir model olarak kullanılmıştır. Bu kapsamda CTGAN, yalnızca eğitim/doğrulama kümesi üzerinde 1000 eğitim dönemi (epoch) boyunca eğitilmiş; stres düzeyi değişkeni koşul değişkeni olarak modele verilmiştir (Şekil 3). Veri sızıntısı riskini en aza indirmek için (i) sentetik veri üretimi sürecinde bağımsız sınama kümesi hiçbir biçimde kullanılmamış ve model bu örnekleri ne eğitim ne de sentetik üretim sırasında "görmemiştir", (ii) her bir kaydın tek bir öğrenciye ait olması nedeniyle aynı öğrencinin birden fazla kümeye düşmesi engellenmiş ve ayrım sonrası bu durum ayrıca kontrol edilmiştir (Şekil 1).

**Algorithm 1:** CTGAN Kullanarak Verinin Ön İşlenmesi ve Sentetik Örnek Benzerliğinin Değerlendirilmesi

**Girdi:** Girdi değişkenlerini (öznitelikler) ve hedef değişkeni içeren orijinal veri kümesi $D$

**Çıktı:** Eşit sınıf temsiline sahip, dengelenmiş ve doğrulanmış sentetik benzerlik içeren veri kümesi $D_{balanced}$

$D$ veri kümesini $X$ (girdi değişkenleri) ve $y$ (hedef değişken) olarak ayır;
$(X, y)$ çiftini katmanlı ayırma (stratified split) yöntemiyle böl:
$D_{train}$(%80), $D_{test}$(%20);
Kategorik değişken olarak *Stres Düzeyi*'ni belirle;
$D_{train}$ üzerinde CTGAN modelini 1000 yineleme (epoch) boyunca eğit;
Gerekli sentetik örnek sayılarını belirle:

- Orta (Moderate): $n = 823 - 539 = 284$;
- Düşük (Low): $n = 823 - 238 = 585$;

**while** *Orta veya Düşük örnek sayısı* $< 823$ **do**
  **if** *örnek sınıfı Orta ise* **then**
    CTGAN ile 1000 sentetik örnek üret ve 284 örneğe ulaşana kadar "Orta" havuzuna ekle;
  **end**
  **else if** *örnek sınıfı Düşük ise* **then**
    CTGAN ile örnek üret ve 585 örneğe ulaşana kadar "Düşük" havuzuna ekle;
**end**
Gerçek ve sentetik verileri birleştir: $D_{balanced} = D_{real} + D_{synthetic}$;
Sentetik veri kalitesini değerlendir:

- Sentetik ve gerçek değişkenler arasındaki değişken bazlı korelasyonu (Column Correlation) hesapla;
- Ortalama özellik bazlı korelasyonu (Mean Feature Correlation) hesapla;
- 1− Ortalama Mutlak Yüzde Hata (MAPE) değerini hesapla;
- Genel benzerlik skorunu (Overall Similarity Score) hesapla;

$D_{balanced}$ veri kümesini model eğitimi için dışa aktar;

**Şekil 3.** CTGAN ile birincil veri ön işleme ve sentetik veri güvenilirliği değerlendirme algoritması

CTGAN eğitimi tamamlandıktan sonra, sınıflar arası dengeyi sağlamak üzere toplam 869 sentetik örnek üretilmiştir. Bu örneklerin 284'ü orta stres ve 585'i düşük stres sınıfına eklenmiş; böylece eğitim/doğrulama

kümesinde üç sınıfın her biri n = 823 olacak şekilde dengeli bir yapı elde edilmiştir (düşük = 823, orta = 823, yüksek = 823). Üretilen sentetik verinin gerçek veriye benzerliğini değerlendirmek amacıyla birden fazla istatistiksel benzerlik ölçütü kullanılmıştır. Bu kapsamda değişken bazlı korelasyon katsayısı 0,91 ve ortalama özellik bazlı korelasyon 0,92 olarak hesaplanmıştır. Hata temelli benzerlik göstergesi olarak ortalama mutlak yüzde hata (MAPE) da raporlanmıştır; burada MAPE değeri yüzde yerine 0–1 aralığında oran olarak hesaplanmış ve yorum kolaylığı için benzerlik puanı 1–MAPE biçiminde ifade edilmiştir. Buna göre (1–MAPE) = 0,84 bulunmuş; tüm ölçütlerin birleşik değerlendirmesini yansıtan genel benzerlik skoru 0,92 olarak ölçülmüştür. Bu değerler, ilgili çalışmalarda önerilen eşiklerle karşılaştırıldığında yüksek uyum düzeyine işaret etmektedir [19-21]. Sonuç olarak, sentetik örnekler yalnızca eğitim/doğrulama aşamasında kullanılarak sınıf dengesizliği giderilmiş; dengeli ve dengesiz veri düzenleri üzerinden modellerin başarımı karşılaştırmalı olarak incelenebilecek iki ayrı veri yapısı oluşturulmuştur.

### 2.2. Model eğitimi ve değerlendirilmesi

Bu çalışmada, sınıf dengesizliğinin model başarımına etkisini değerlendirebilmek amacıyla iki ayrı veri düzeni üzerinde modelleme yapılmıştır: (1) dengesiz veri düzeni (sınıf dengesiz hedef değişken ayrımı sonrası eğitim/doğrulama kümesinin özgün sınıf dağılımı korunur) ve (2) dengeli veri düzeni (yalnızca eğitim/doğrulama kümesinin CTGAN ile dengelenmesi). Sınıf dengesizliğini gidermek amacıyla çoğunluk sınıflarının örneklem sayısını en küçük sınıfa *(n=238)* indirgeyen örnek eksiltme stratejisi yerine, tüm sınıfları eğitim kümesindeki en büyük örneklem sayısına *(n=823)* tamamlayan CTGAN tabanlı sentetik çoğaltma yaklaşımı benimsenmiştir. Örnek eksiltme yönteminin tercih edilmemesinin temel nedeni, çoğunluk sınıflarındaki gerçek dünya verilerinin önemli bir kısmının analiz dışı bırakılmasıyla oluşacak potansiyel bilgi kaybını önleme arzusudur. Dengesiz veri seti üzerinde yapılan ön değerlendirmelerin (Tablo 2), modellerin azınlık sınıfından dahi yeterli ayırt edici sinyali alabildiğini göstermesi, mevcut verilerin korunarak sentetik örneklerle desteklenmesinin hem istatistiksel gücü artıracağı hem de modelin genellenebilirliğini optimize edeceği öngörüsünü doğrulamıştır. Her iki düzende de bağımsız test kümesi en başta ayrılmış, eğitim, doğrulama, hiperparametre arama ve sentetik veri üretimi adımlarının hiçbirinde kullanılmamış; böylece nihai değerlendirme gerçek gözlemler üzerinde ve tarafsız biçimde yapılabilecek şekilde test kümesinin özgün yapısı korunmuştur (Şekil 1).

Modelleme aşamasında literatürde yaygın kullanılan beş gözetimli ML algoritması eğitilmiş ve karşılaştırılmıştır: Lojistik Regresyon (LR), Destek Vektör Makineleri (SVM), Rastgele Orman (RF), k-En Yakın Komşu (k-NN) ve Çok Katmanlı Algılayıcı (MLP). Ölçek duyarlı yöntemlerde öğrenmeyi tutarlı hale getirmek için sayısal değişkenler z-skoru (z-score) normalizasyonu ile standartlaştırılmıştır. Veri sızıntısını önlemek amacıyla bu standartlaştırma, katmanlı çapraz doğrulama (CV) döngüsü içinde yalnızca eğitim katı üzerinde öğrenilmiş; doğrulama katına ve daha sonra test kümesine aynı dönüşüm parametreleri uygulanmıştır. Böylece doğrulama/test verisinden eğitim aşamasına doğrudan ya da dolaylı bilgi taşınması engellenmiştir.

Hiperparametre ayarı, örnek verimliliği yüksek ve kara-kutu amaç fonksiyonlarında etkin bir yöntem olan Bayesyen optimizasyon ile yürütülmüştür. Arama süreci hem dengesiz hem de dengeli veri düzeni için ayrı ayrı işletilmiş; her model için belirlenen hiperparametre uzayında amaç fonksiyonu CV tabanlı değerlendirme ile iyileştirilmiştir. Model başarımı, sınıflar arası temsil dengesizliklerine daha duyarlı olduğu için öncelikle Makro F1 üzerinden izlenmiş; bununla birlikte değerlendirme bütünlüğünü artırmak amacıyla Doğruluk, Cohen kappa, Makro AUC ve Log kayıp ölçütleri de her deneyde hesaplanarak kaydedilmiştir. Ayrıca her bir ölçüt için en iyi hiperparametre kümeleri ayrı ayrı belirlenerek, farklı ölçütlerin model seçimine etkisi sistematik biçimde karşılaştırılabilir hale getirilmiştir.

Bayesyen optimizasyon sürecinde üretilen her bir hiperparametre adayı, k = 2–10 aralığındaki katmanlı çapraz doğrulama düzeneklerinde ayrı ayrı değerlendirilmiş; her k değeri için katlamalar üzerinden elde edilen performans değerleri kaydedilmiştir. Modellerin kararlılığını ve genellenebilirliğini en üst düzeyde test etmek amacıyla özellikle 10-katlı tabakalı çapraz doğrulama yöntemi esas alınmış ve sonuçlar bu katlamaların ortalaması olarak raporlanmıştır. Nihai hiperparametre seçimi, bu tarama içinde Makro F1 değerini en yüksek yapan yapı esas alınarak yapılmış; Doğruluk, Makro AUC, Cohen kappa ve Log kayıp ölçütleri ise destekleyici olarak raporlanmıştır. Katmanlı çapraz doğrulama sonuçları, her bir k değeri için katlamalar üzerinden ortalama ± standart sapma biçiminde hesaplanmıştır. Eğitim aşamasındaki bu ortalama performans bildirimlerinden farklı olarak, bağımsız test kümesi sonuçları, optimize edilen nihai modelin daha önce hiç görülmemiş veriler üzerindeki tekil başarısını yansıtmaktadır.

Modelin karar mekanizmasını yorumlanabilir kılmak amacıyla SHAP analizi uygulanmış ve değişkenlerin tahminlere katkısı nicel olarak incelenmiştir. SHAP analizlerinde ağaç tabanlı modeller için özelleşmiş açıklayıcı (Tree Explainer) yöntemi kullanılarak modellerin olasılık ölçeği üzerindeki çıktıları esas alınmıştır. Bu sayede her

bir özniteliğin, modelin ilgili sınıf için ürettiği olasılık puanını (0-1 aralığı) baz değerden ne kadar saptırdığı hesaplanmıştır. SHAP bulgularında belirgin biçimde baskın katkı verdiği gözlenen değişkenin etkisini sınamak için, yöntem bilimsel olarak tek öznitelik hariç tutma (leave-one-feature-out-LOFO) yaklaşımına dayanan bir ablasyon (özellik çıkarma) [22] analizi gerçekleştirilmiştir. Bu analizde, Williamson ve ark. (2023) tarafından çerçevelenen algoritma-bağımsız değişken önemi yaklaşımı esas alınarak, en baskın öznitelik veri setinden çıkarılmış ve modellerin performansındaki değişim nicel olarak doğrulanmıştır."

Son olarak, modelin hedef etiket ile kurduğu ilişkinin rastlantısal bir örüntüden kaynaklanmadığını göstermek üzere etiket permütasyon testi [23] uygulanmıştır. Bu doğrultuda, eğitim/doğrulama kümesindeki etiketler *n=30* kez rastgele karıştırılarak model her tekrarda orijinal hiperparametrelerle yeniden eğitilmiştir. Bu işlem sonucunda elde edilen Makro AUC değerleri üzerinden bir boş hipotez dağılımı oluşturulmuştur. Gerçek verilerle elde edilen model başarımı ile bu rastgele dağılım arasındaki istatistiksel fark, ampirik p-değeri aşağıdaki Denklem 1'deki formüle göre hesaplanmıştır:

$$p = (C + 1)/(n + 1) \quad (1)$$

Burada *C*, rastgele başarımı gerçek başarıdan büyük olan tekrar sayısıdır. Makro AUC değerinin şans düzeyine (AUC≈0,50) gerilemesi ve p-değerinin anlamlılık eşiğinin *(p <0,05)* altında kalması, modelin verideki anlamlı örüntüleri öğrendiğini kanıtlayan bir negatif kontrol olarak kullanılmıştır.

## 3. Sonuçlar

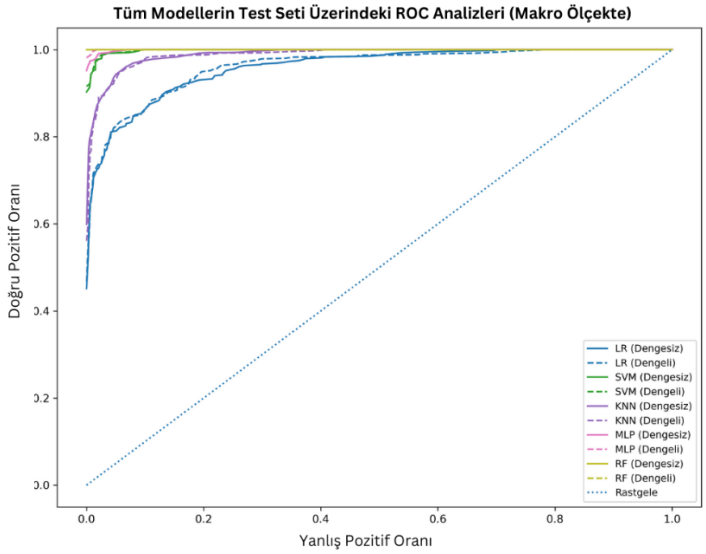
Bağımsız sınama (test) kümesi üzerindeki performans değerleri, seçilen en iyi model ve hiperparametre yapısı kullanılarak tek bir nihai değerlendirme sonucunda elde edilmiştir; ortalama ± standart sapma raporlaması ise yalnızca katmanlı çapraz doğrulama (CV) sonuçları için kullanılmıştır. Veri sızıntısını önlemek amacıyla veri bölme aşamasında stres düzeyi sınıfları korunacak biçimde katmanlı ayırma uygulanmış; test kümesi en başta ayrılarak tüm süreç boyunca değiştirilmeden korunmuştur. Modeller iki ayrı eğitim düzeniyle karşılaştırılmıştır: (i) dengesiz eğitim/doğrulama veri kümesi (özgün sınıf dağılımı korunmuş) ve (ii) dengeli eğitim/doğrulama veri kümesi (yalnızca eğitim/doğrulama kümesinin CTGAN ile dengelenmesi). Performans; doğruluk, Makro F1-skoru (Macro F1), Makro AUC (Macro AUC), Cohen kappa katsayısı, log kayıp ve sınıf bazlı F1-skorları ile raporlanmıştır (Tablo 2 ve Tablo 3).

**Tablo 2.** Dengesiz eğitim/doğrulama veri kümesi ile eğitilen modellerin test kümesi üzerindeki başarımı

| Model | Doğruluk | Makro AUC | Makro F1 | Cohen kappa | Log kayıp | F1 (Düşük) | F1 (Orta) | F1 (Yüksek) |
|---|---|---|---|---|---|---|---|---|
| LR | 0.8475 | 0.9589 | 0.8447 | 0.7445 | 0.3813 | 0.8667 | 0.7893 | 0.8783 |
| SVM | 0.9750 | 0.9983 | 0.9759 | 0.9582 | 0.0787 | 0.9828 | 0.9668 | 0.9782 |
| KNN | 0.9300 | 0.9893 | 0.9115 | 0.8828 | 0.2579 | 0.8621 | 0.9111 | 0.9614 |
| MLP | 0.9775 | 0.9993 | 0.9759 | 0.9623 | 0.0497 | 0.9744 | 0.9701 | 0.9831 |
| **RF** | **1.0000** | **1.0000** | **1.0000** | **1.0000** | **0.0696** | **1.0000** | **1.0000** | **1.0000** |

**Tablo 3.** Dengeli eğitim/doğrulama veri kümesi ile eğitilen modellerin test kümesi üzerindeki başarımı

| Model | Doğruluk | Makro AUC | Makro F1 | Cohen kappa | Log kayıp | F1 (Düşük) | F1 (Orta) | F1 (Yüksek) |
|---|---|---|---|---|---|---|---|---|
| LR | 0.8400 | 0.9608 | 0.8413 | 0.7373 | 0.3945 | 0.8661 | 0.7956 | 0.8622 |
| SVM | 0.9750 | 0.9984 | 0.9764 | 0.9583 | 0.0729 | 0.9831 | 0.9706 | 0.9756 |
| KNN | 0.9125 | 0.9881 | 0.8950 | 0.8570 | 0.2620 | 0.8421 | 0.8955 | 0.9474 |
| MLP | 0.9900 | 0.9999 | 0.9882 | 0.9833 | 0.0324 | 0.9831 | 0.9888 | 0.9928 |
| **RF** | **1.0000** | **1.0000** | **1.0000** | **1.0000** | **0.0003** | **1.0000** | **1.0000** | **1.0000** |



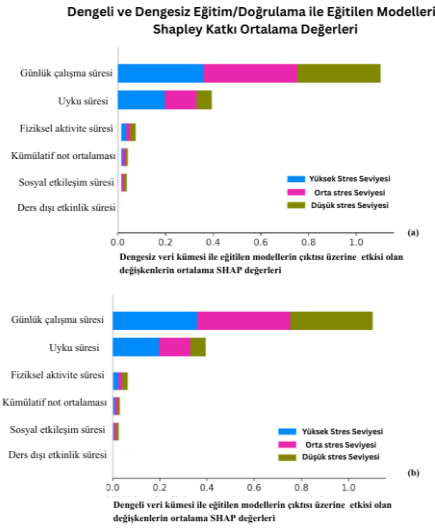

**Şekil 4.** Dengesiz eğitim/doğrulama veri kümesi ve dengeli eğitim/doğrulama veri kümesi ile eğitilmiş en iyi performans gösteren modellerin alıcı işletim karakteristik analizleri

Tüm modellerin bağımsız test kümesi üzerindeki ayırt etme başarımı, sınıflar arası dengesizlikten etkilenmemesi için makro ortalama yaklaşımı ile hesaplanan Alıcı İşletim Karakteristiği (ROC) eğrileri üzerinden değerlendirilmiş ve Şekil 4'de sunulmuştur. Dengesiz eğitim/doğrulama veri kümesi üzerinde eğitilen modellerin bağımsız test kümesindeki karşılaştırmalı sonuçları incelendiğinde, RF modeli en yüksek başarımı göstermiştir (Doğruluk = 1.0000). RF'yi, yüksek ve dengeli performans sergileyen MLP (Doğruluk = 0.9775) ve SVM (Doğruluk = 0.9750) izlemiştir. k-NN orta düzeyde başarı sağlamış (Doğruluk = 0.9300), LR) ise görece daha düşük doğrulukla (Doğruluk = 0.8475) en zayıf performansı sergilemiştir.

Dengeli eğitim/doğrulama veri kümesi (CTGAN ile dengelenmiş) üzerinde eğitilen modellerde de en yüksek başarım yine RF ile elde edilmiştir **(**Doğruluk = 1.0000**)**. Bu düzenekte MLP modelinin doğruluğunun arttığı görülmüştür (Doğruluk = 0.9900). SVM modelinin doğruluk düzeyi korunmuş **(**Doğruluk = 0.9750), buna karşın LR **(**Doğruluk = 0.8400) ve k-NN **(**Doğruluk = 0.9125**)** modellerinde sınırlı düzeyde düşüş gözlenmiştir.

İki veri düzeni birlikte değerlendirildiğinde, en iyi model her iki durumda da RF olmuş ve test kümesinde değişmeden en başarılı sınıflandırma performansını göstermiştir **(**Doğruluk = 1.0000). CTGAN ile dengelenmiş eğitim/doğrulama veri kümesine geçişin en belirgin katkısı MLP üzerinde görülmüş; MLP'nin test doğruluğu 0.9775'ten 0.9900'a yükselmiştir. SVM doğruluğunu korurken **(**0.9750 → 0.9750), LR ve k-NN modellerinde küçük ölçekli düşüşler gözlenmiştir. Bu bulgular, dengeleme etkisinin modele bağlı olduğunu; bazı modellerde genellenebilirliği artırabildiğini, bazı modellerde ise performansı sınırlı ölçüde değiştirdiğini gözlemlenmiştir.

Şekil 5'de, RF modeli için SHAP analizi sonuçları sunulmuştur. Şekil 5(a), Dengesiz eğitim/doğrulama veri kümesi ile eğitilen RF modeline; Şekil 5(b) ise CTGAN ile dengelenmiş Dengeli eğitim/doğrulama veri kümesi ile eğitilen RF modeline aittir. Analiz edilen özniteliklerin ortalama mutlak SHAP değerleri, modellerin olasılık ölçeği üzerinden hesaplanmış olup yığılmış bar grafikler halinde sunulmuştur. Grafiklerde 1,00 eşiğinin üzerinde gözlenen değerler, ilgili özniteliğin düşük, orta ve yüksek stres sınıflarının tamamı üzerindeki toplam mutlak katkısını ($\sum |SHAP|$) temsil etmektedir. Her iki eğitim düzeninde de tüm alt sınıflar için modellerin karar verme sürecinde en yüksek ortalama katkı Günlük çalışma süresi değişkeninden gelmektedir. Bunu uyku süresi. Diğer değişkenlerin (fiziksel aktivite süresi, kümülatif not ortalaması, sosyal etkileşim süresi ve ders dışı etkinlik süresi) ortalama katkıları görece düşük düzeyde kalmıştır. Bu bulgular hem dengeli hem de dengesiz eğitim/doğrulama düzeninde stres düzeyi tahmininin ağırlıklı olarak çalışma ve uyku alışkanlıkları etrafında şekillendiğine işaret etmektedir. İki SHAP değerleri arasındaki benzerlik, CTGAN'nin değişkenler arası ilişki yapısını (ortak dağılımları) koruyup yalnızca sınıf dağılımını dengelemesi ve her iki düzende de günlük çalışma süresi değişkeninin belirgin biçimde baskın kalmasının etkisi olduğu gözlemlenmiştir.




**Şekil 5.** Dengesiz eğitim/doğrulama veri kümesi ve CTGAN ile dengelenmiş eğitim/doğrulama veri kümesi ile eğitilen RF modeli için modellerin ağaç açıklayıcı olasılık ölçeği üzerinden hesaplanmış SHAP özet sonuçları. Yatay eksen, değişkenlerin model çıktısı üzerindeki ortalama etkisini gösteren ortalama mutlak SHAP değerini ifade eder.
(a) Dengesiz eğitim/doğrulama veri kümesi ile eğitilen modelin değişken katkıları.
(b) CTGAN ile dengelenmiş eğitim/doğrulama veri kümesi ile eğitilen modelin değişken katkıları.

SHAP analizinde *Günlük Çalışma Süresi* değişkeninin belirgin biçimde baskın katkı verdiği görüldüğünden, bu değişkenin etkisini nicel olarak ölçmek ve yüksek başarımın tek bir değişkene aşırı bağımlılıktan kaynaklanıp kaynaklanmadığını sınamak amacıyla ablasyon (özellik çıkarma) deneyi gerçekleştirilmiştir. Ablasyon analizi, her iki veri düzeninde de en yüksek test başarımını veren modelin RF olması nedeniyle bu modeller üzerinde yürütülmüş; hem Dengesiz eğitim/doğrulama veri kümesi hem de Dengeli eğitim/doğrulama veri kümesi için aynı eğitim-değerlendirme akışı korunarak yalnızca *Günlük Çalışma Süresi* öznitelik kümesinden çıkarılmış ve bağımsız test kümesinde tekrar değerlendirilmiştir (Tablo 4). Sonuçlar, değişken çıkarıldığında performansın her iki düzende de belirgin biçimde düştüğünü göstermiştir: dengesiz düzende Doğruluk 1.0000 → 0.7950, Makro F1 1.0000 → 0.7007 ve Makro AUC 1.0000 → 0.9264; dengeli düzende ise Doğruluk 1.0000 → 0.8100, Makro F1 1.0000 → 0.7604 ve Makro AUC 1.0000 → 0.9337 seviyesine gerilemiştir. Sınıf bazlı ölçütler incelendiğinde düşüşün özellikle düşük stres sınıfında daha belirgin olduğu görülmüştür (dengesizde F1: 1.0000 → 0.4651, dengelide F1: 1.0000 → 0.6379); bu bulgular, *Günlük Çalışma Süresi* değişkeninin RF modelinin kararlarında kritik rol oynadığını ve genel başarım üzerinde belirleyici düzeyde etkisi bulunduğunu göstermektedir.

**Tablo 4.** SHAP analizine göre her iki veri kümesinde de en baskın özellik olan *Günlük Çalışma Süresi*'nin çıkarılması ile gerçekleştirilen ablasyon testi sonuçları

| Eğitim Verisi | Ayar | Doğruluk | Makro F1 | Makro AUC | Cohen kappa | Log kayıp | F1 (Düşük) | F1 (Orta) | F1 (Yüksek) |
|---|---|---|---|---|---|---|---|---|---|
| **Dengesiz eğitim/doğrulama veri kümesi** | Tüm değişkenler | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0664 | 1.0000 | 1.0000 | 1.0000 |
| **Dengesiz eğitim/doğrulama veri kümesi** | Günlük çalışma Süresi çıkarıldı | 0.7950 | 0.7007 | 0.9264 | 0.6473 | 0.4707 | 0.4651 | 0.7230 | 0.9139 |
| **Dengeli eğitim/doğru** | Tüm değişkenler | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **lama veri kümesi** | | | | | | | | | |
| **Dengeli eğitim/doğrulama veri kümesi** | Günlük çalışma Süresi çıkarıldı | 0.8100 | 0.7604 | 0.9337 | 0.6835 | 0.4297 | 0.6379 | 0.7292 | 0.9140 |
| **Referans Model** | Sadece Günlük Çalışma Süresi ile eğitilmiştir | 0.7525 | 0.7674 | 0.9050 | 0.5901 | 0.5826 | 0.8571 | 0.6742 | 0.7707 |

Etiket permütasyon testi, yalnızca CTGAN ile dengelenmiş Dengeli eğitim/doğrulama veri kümesi üzerinde seçilen RF modeli için 30 tekrar ile uygulanmış ve bir negatif kontrol olarak, dengeli ve dengesiz eğitim/doğrulama düzeneklerinde gözlenen yüksek başarımın etiketlerle rastlantısal eşleşme ya da veri sızıntısı gibi yöntemsel sorunlardan kaynaklanıp kaynaklanmadığını değerlendirmek amacıyla kullanılmıştır (Tablo 5). Bu kapsamda, modellerin ayırt etme gücünün makro AUC açısından şans düzeyine (0,50 civarı) gerileyip gerilemediği izlenmiştir. Sonuçlar, etiketler rastgele karıştırıldığında bağımsız test kümesinde başarımın şans düzeyine indiğini göstermiştir: Makro AUC = 0,504 ± 0,052, Makro F1 = 0,337 ± 0,043, Cohen kappa = 0,014 ± 0,071 ve log kayıp = 1,091 ± 0,034. Üç sınıflı bir problemde bu düzeyler, modelin rastlantısal tahmin davranışı ile uyumlu olup, yüksek başarımın veri sızıntısı veya rastlantısal uyum ile açıklanmadığını desteklemektedir; bu bulgu, SHAP ve ablasyon sonuçlarıyla birlikte değerlendirildiğinde *Günlük Çalışma Süresi* değişkeninin modelin karar verme sürecindeki belirleyici rolüyle tutarlı bir çerçevede sunduğu gözlemlenmiştir.

**Tablo 5.** Etiket permütasyon testi sonuçları (Dengeli eğitim/doğrulama; Rastgele Orman Modeli, 30 tekrar): bağımsız sınama kümesi başarımı

| Ölçüt | Dengeli eğitim/doğrulama veri kümesi ile eğitilmiş RF Yöntemi'nin Test Kümesi üzerindeki Performans Ortalaması |
|---|---|
| Tekrar sayısı | 30 |
| Makro AUC (ortalama±ss) | 0,504±0,052 |
| Makro F1 (ortalama±ss) | 0,337±0,043 |
| Doğruluk (ortalama±ss) | 0,390±0,044 |
| Cohen kappa (ortalama±ss) | 0,014±0,071 |
| Log kayıp (ortalama±ss) | 1,091±0,034 |
| Ampirik p-değeri * | (p <0,033) |

***** Ampirik p-değeri, $p = (C + 1)/(n + 1)$ formülüyle hesaplanmıştır; burada *C*, rastgeleleştirilmiş etiketlerle elde edilen AUC değerinin orijinal modelin AUC değerinden (AUS=1.00) büyük veya eşit olduğu durumların sayısıdır ($C=0$).

## 4. Tartışma

Bu bölümde, elde edilen bulgular iki ana eksen altında tartışılmaktadır. İlk olarak, Bulguların özeti ve model karşılaştırması başlığı altında dengesiz ve CTGAN ile dengelenmiş eğitim/doğrulama veri kümeleri üzerinde eğitilen modellerin test başarımı bütüncül biçimde ele alınmakta; ardından yüksek başarımın yöntemsel güvenilirliğini destekleyen açıklanabilirlik (SHAP), ablasyon (özellik çıkarma) ve etiket permütasyon testi bulguları yorumlanmaktadır. Böylece sonuçlar yalnızca performans düzeyinde değil, aynı zamanda modelin karar verme mantığı ve sağlamlık kontrolleri çerçevesinde değerlendirilerek literatürdeki benzer çalışmalarla ilişkilendirilmektedir.

### 4.1 Bulguların özeti ve model karşılaştırması

Bu çalışmada, üniversite öğrencilerinin stres düzeylerini tahmin etmek üzere LR, SVM, RF, k-NN ve MLP modelleri; doğruluk, Makro AUC, Makro F1, Cohen kappa, log kayıp ve sınıf bazlı F1 ölçütleri ile iki ayrı eğitim düzeni altında karşılaştırılmıştır (Tablo 2, Tablo 3). Her iki eğitim düzeninde de en yüksek performans RF modeli ile elde edilmiştir (Doğruluk = 1,000). Dengesiz eğitim/doğrulama veri kümesinde MLP (Doğruluk = 0,9775) ve

SVM (Doğruluk = 0,9750) RF'yi izlerken, dengeli eğitim/doğrulama veri kümesinde MLP'nin doğruluğu 0,9900'a yükselmiş, SVM ise benzer düzeyde kalmıştır (Doğruluk = 0,9750). Buna karşın LR (0,8475 → 0,8400) ve k-NN (0,9300 → 0,9125) modellerinde dengeli eğitim/doğrulama düzenine geçişle birlikte sınırlı düşüşler gözlenmiştir.

Dengeli ve dengesiz veri düzenleri arasındaki bu yüksek performans benzerliği, tanımlayıcı değişkenlerin (özellikle günlük çalışma süresi ve uyku süresi) hedef değişken üzerindeki ayırt edici gücünün oldukça yüksek olduğunu ortaya koymaktadır. Azınlık sınıflarındaki mevcut örneklem miktarının dahi modellerin temel karar sınırlarını belirlemesi için yeterli bilgi yüküne sahip olması, sentetik dengeleme işleminin genel başarım üzerinde marjinal bir etki yaratmasıyla sonuçlanmıştır.

Bu durum, modellerin orijinal veri setindeki örüntüleri öğrenme noktasında doygunluğa ulaştığını ve elde edilen başarının sentetik veri desteğinden ziyade, ham verideki güçlü sinyal-gürültü oranından kaynaklandığını doğrulamaktadır. CTGAN ile dengelemenin etkisinin modele bağlı olduğu; dengeleme işleminin özellikle veri hacmine ve sınıf temsiline daha duyarlı olan MLP modelinde sınırlı bir iyileşme sağladığı, diğer modellerde ise mevcut yüksek başarımı stabilize ettiği gözlemlenmiştir.

### 4.2 Kusursuz doğruluk bulgusunun güvenilirliği: veri sızıntısı ve ezberleme olasılığının dışlanması

Dengeli ve dengesiz eğitim/doğrulama veri kümeleri ile eğitilen RF modellerinin bağımsız test kümesinde Doğruluk = 1,0000 düzeyine ulaşması, yöntemsel olarak yanlılık kaynaklarını azaltmaya yönelik ek kontrollerle desteklenmiştir. Öncelikle, bağımsız test kümesi en başta ayrılmış ve tüm süreç boyunca değiştirilmeden korunmuştur; CTGAN ile sentetik üretim dahil tüm işlemler yalnızca eğitim/doğrulama kümesi üzerinde yürütülmüştür. Ayrıca, ölçekleme gibi ön işleme adımları katmanlı çapraz doğrulama döngüsü içinde yalnızca eğitim katında öğrenilmiş; doğrulama ve test verisine aynı dönüşüm parametreleri uygulanarak olası veri sızıntısı riski azaltılmıştır. Bu çerçeve, test başarımının sentetik örneklerle yapay biçimde kolaylaşması veya eğitim bilgisinin test aşamasına taşınması gibi yanlılıkları sınırlayacak şekilde kurgulanmıştır. Performans değerlerinin dengeli ve dengesiz veri düzenleri arasında radikal bir farklılık sergilememesi, ham veri setindeki özellikle günlük çalışma ve uyku süresi özniteliklerinin stres düzeyini ayırt etmede oldukça güçlü ve net bir sinyal taşıdığını kanıtlamaktadır. Bu durum, modellerin orijinal veri yapısındaki doğal örüntüleri öğrenme noktasında erkenden doygunluğa ulaştığını ve sınıf dengesizliğinin öğrenme sürecini kısıtlayacak düzeyde bir bilgi eksikliği yaratmadığını metodolojik olarak doğrulamaktadır.

Veri kümesi incelendiğinde, Günlük Çalışma Süresi değişkeni ile yüksek stres düzeyi arasında gözlenen belirgin örtüşme, problemin doğrusal ve kesin kurallı (deterministik) bir yapıya sahip olabileceği yönünde bir izlenim uyandırabilir. Söz konusu olasılığı ampirik olarak değerlendirmek amacıyla, ek bir kontrol basamağı olarak yalnızca 'Günlük Çalışma Süresi' özniteliğiyle eğitilen temel bir referans modeli oluşturulmuştur. Bu modelin bağımsız sınama kümesinde sergilediği 0,7525 doğruluk oranının, tüm öznitelikleri içeren RF modelinin başarım düzeyinin gerisinde kaldığı gözlemlenmiştir (Tablo 4). Stres düzeyinin tek boyutlu bir eşik kuralından ziyade, değişkenler arası etkileşimlere dayalı daha karmaşık bir örüntü üzerinden şekillendiğine işaret etmektedir.

Bununla birlikte, ablasyon (özellik çıkarma) analizi de bu değerlendirmeyi destekleyen tamamlayıcı bulgular sunmaktadır (Tablo 4). En baskın değişken olan Günlük Çalışma Süresi veri setinden çıkarıldığında dahi RF modelinin test kümesinde %79,50–81,00 aralığında bir performans sergileyebilmesi; stres düzeyinin tek bir faktöre dayalı basit bir karar kuralı ile değil, uyku süresi ve kümülatif not ortalaması gibi diğer yaşam tarzı ve akademik değişkenlerin birleşik etkisiyle açıklandığını destekler niteliktedir. Sonuç olarak problem, bütünüyle deterministik bir türetme süreci olarak değil; değişkenler arası etkileşimin yüksek olduğu ve olasılıksal bir ilişki yapısının baskın olduğu bir örüntü öğrenimi çerçevesinde değerlendirilmektedir.

Bu yüksek başarımın rastlantısal bir uyumdan kaynaklanmadığını desteklemek amacıyla ayrıca etiket permütasyon testi negatif kontrol olarak uygulanmıştır (Tablo 5). 30 tekrarlı etiket permütasyon testi bulguları, rastgele bir tahmin mekanizması için beklenen şans düzeyini (AUC ≈ 0,50) yansıtan bir boş hipotez dağılımı ortaya koymuştur. Analiz kapsamındaki rastgele denemelerin hiçbiri, asıl modelin performans seviyesine ulaşamamıştır. Elde edilen ampirik p-değeri *($p < 0,033$),* gözlenen başarının tesadüfi faktörlerle açıklanmasının güç olduğunu ve modelin veri setindeki anlamlı ilişkileri başarıyla yakaladığını destekler niteliktedir. Nitekim Makro AUC = 0,504 ± 0,052, Makro F1 = 0,337 ± 0,043, Cohen kappa = 0,014 ± 0,071 ve log kayıp = 1,091 ± 0,034 değerleri, üç sınıflı bir problemde rastlantısal tahmin davranışı ile uyumlu bir performansa işaret etmektedir. Bu bulgular, kusursuz test başarımının etiketlerle rastlantısal eşleşme ya da yöntemsel bir sızıntı ile açıklanmasından ziyade, veri içinde taşınan anlamlı örüntülerin öğrenilmesiyle ilişkili olduğunu desteklemektedir.

### 4.3 Değişken katkıları: SHAP ve ablasyon bulgularının yorumu

Model kararlarının hangi değişkenler üzerinden şekillendiğini ortaya koymak üzere RF modeli için SHAP analizi yapılmıştır (Şekil 5). Hem dengesiz hem de CTGAN ile dengelenmiş eğitim/doğrulama düzeninde, en yüksek ortalama katkının Günlük Çalışma Süresi değişkeninden geldiği, bunu Uyku Süresi değişkeninin izlediği görülmüştür. Bu bulgular, stres düzeyinin ağırlıklı olarak çalışma ve uyku düzeni etrafında biçimlendiğine işaret etmektedir. Ayrıca iki SHAP grafiğinin birbirine yakın görünmesi, CTGAN'nin değişkenler arası ilişki yapısını (ortak dağılımları) koruyup yalnızca sınıf dağılımını dengelemesi ve her iki düzende de Günlük Çalışma Süresi değişkeninin baskın rolünü sürdürmesiyle açıklanabilir.

Günlük Çalışma Süresi değişkeninde 8 saat civarında gözlenen keskin ayrışma, ilk bakışta sınıfların basit bir eşik kuralı ile ayrıldığı izlenimi doğurabilir. Bununla birlikte, bu tür bir kırılma, akademik literatürde uzun süreli çalışma yükünün zihinsel yorgunluk ve tükenmişlik riskini artırdığı ve belirli bir yoğunluk düzeyinden sonra stres tepkisinin doğrusal olmayan biçimde hızlandığı yönündeki bulgularla uyumludur [24]. Bu nedenle gözlenen eşik etkisi, hedef değişkenin kestirim değişkenlerden kısmen türetilmiş olmasına dayalı bir deterministik etiketleme varsayımından ziyade, üniversite öğrencisi popülasyonunda akademik yük ile stres arasındaki ilişkinin kritik bir noktada doğrusal olmayan bir kırılma gösterebileceğine işaret eden ampirik bir desen olarak değerlendirilmelidir.

Bu baskınlığın performansa etkisini nicel olarak sınamak amacıyla ablasyon (özellik çıkarma) deneyi uygulanmıştır (Tablo 4). Her iki veri düzeninde de en iyi test başarımını veren model RF olduğundan, en baskın özellik olan Günlük Çalışma Süresi çıkarılarak RF yeniden eğitilmiş ve bağımsız test kümesinde tekrar değerlendirilmiştir. Değişken çıkarıldığında performansın her iki düzende de belirgin biçimde düştüğü görülmüştür (dengesiz düzende Doğruluk 1,0000 → 0,7950, Makro F1 1,0000 → 0,7007; dengeli düzende Doğruluk 1,0000 → 0,8100, Makro F1 1,0000 → 0,7604). Bu düşüş, yüksek başarımın tek başına bir "eşik kuralı" ya da ezberleme ile açıklanmadığını; aksine Günlük Çalışma Süresi'nin güçlü ancak tekil olmayan bir belirleyici olduğunu ve modelin diğer değişkenlerden (özellikle uyku süresi ve GNO) de anlamlı sinyal ürettiğini göstermektedir. Nitekim çalışma süresinin artışıyla stres düzeyinin yükselmesi, akademik yük, performans baskısı ve zaman yönetimi güçlükleriyle ilişkilendirilebilir ve bu yorum literatürle tutarlıdır [24]. Benzer biçimde uyku süresinin katkısı, uyku azalmasının öznel stres ve bilişsel tükenmişlikle ilişkili olabileceğine işaret eden bulgularla paraleldir [25]. GNO'nun özellikle yüksek stres düzeyinde etkili olması ise akademik başarı baskısının stres üzerinde dolaylı bir rolü olabileceğini düşündürmektedir [26].

### 4.4 Literatürle karşılaştırma ve çalışmanın özgün katkısı

Bu çalışmada elde edilen yüksek tahmin başarımı, özellikle karar ağacı temelli yaklaşımlar kullanılarak çok yüksek doğruluk değerleri bildirilen çalışmalarla benzer bir çizgi göstermektedir [12]. Bununla birlikte, mevcut çalışma literatürde sık görülen *tek veri düzeni–tek model–tek doğrulama şeması* yaklaşımının ötesine geçerek, stres düzeyi tahminini karşılaştırmalı, çoklu doğrulama şemalarıyla sınanmış ve açıklanabilirlik ile sağlamlık kontrolleriyle gerekçelendirilmiş bir çerçevede ele almaktadır. Bu yönüyle çalışma, Dudić ve ark. [12] başta olmak üzere benzer çalışmalarla aynı hedef problem alanında konumlansa da, yöntemsel kapsam ve doğrulama derinliği açısından belirgin biçimde ayrışmaktadır.

Çalışmanın literatüre başlıca katkıları aşağıda özetlenmiştir;

- Dengeli–dengesiz veri düzeneklerinin doğrudan karşılaştırılması: Sınıf dengesizliği açık bir problem olarak ele alınmış; Dengesiz eğitim/doğrulama veri kümesi ile yalnızca eğitim/doğrulama üzerinde CTGAN kullanılarak oluşturulan Dengeli eğitim/doğrulama veri kümesi altında modellerin başarımı karşılaştırmalı biçimde değerlendirilmiştir. Bu karşılaştırma, aynı bağımsız test kümesi temelinde yapılarak dengelemenin genellenebilirlik üzerindeki etkisinin doğrudan gözlemlenmesine olanak sağlamıştır.
- Doğrulama derinliği ve kararlılık analizi: Değerlendirme tek bir çapraz doğrulama şemasına indirgenmemiş; k = 2–10 aralığında farklı kat sayılarında katmanlı çapraz doğrulama ile performans izlenmiştir. Bu yaklaşım, sonuçların belirli bir doğrulama düzenine özgü kalmasını engelleyerek performans kararlılığını daha geniş bir çerçevede sınamıştır.
- Çoklu performans ölçütleriyle dengeli değerlendirme: Sadece doğruluk veya AUC gibi genel ölçütlerle sınırlı kalınmamış; Makro F1 temel ölçüt olmak üzere doğruluk, Makro AUC, Cohen kappa, log kayıp ve sınıf bazlı F1 birlikte raporlanmıştır. Böylece düşük–orta–yüksek stres düzeylerinde model davranışı daha ayrıntılı biçimde değerlendirilmiştir.
- Açıklanabilirlik ve sağlamlık kontrollerinin bütünleşik kullanımı: Model kararları SHAP (Shapley Katkı Açıklamaları) ile yorumlanmış; belirleyici değişkenin etkisi ablasyon (özellik çıkarma) ile

nicel olarak sınanmış; ek olarak etiket permütasyon testi negatif kontrol olarak kullanılarak modelin ayırt etme gücünün etiketler rastgeleleştirildiğinde şans düzeyine gerilediği gösterilmiştir. Bu bütünleşik yaklaşım, yüksek başarım bulgularının yöntemsel olarak daha güçlü biçimde gerekçelendirilmesini sağlamaktadır.

Özetle, Dudić ve ark. [12] gibi çalışmalarda yüksek doğruluk bildirimleri bulunmakla birlikte, bu çalışmanın ayırt edici yönü; CTGAN ile dengeleme–dengesiz durum karşılaştırması, k = 2–10 aralığında doğrulama kararlılığı, ve SHAP + ablasyon + etiket permütasyonu kombinasyonuyla model davranışının yalnızca başarım düzeyinde değil, aynı zamanda yorumlanabilirlik ve sağlamlık düzeyinde de bütüncül biçimde desteklenmesidir. Bu çerçeve, benzer problem alanlarında çalışan araştırmacılar için uygulanabilir bir yöntemsel yol haritası sunmaktadır.

### 4.5 Sınırlılıklar

Kullanılan veri kümesi öz-bildirime dayalı olduğundan stres düzeyi ve yaşam tarzı değişkenleri ölçüm hatası ve yanıt yanlılığı içerebilir. Örneklemin ağırlıklı olarak tek bir ülke bağlamından gelmesi bulguların farklı kültürel ve kurumsal bağlamlara genellenebilirliğini sınırlayabilir. Çalışma kesitsel nitelikte olduğundan değişkenler arasındaki ilişkiler nedensellik biçiminde yorumlanmamalıdır. Ayrıca analizlerin tek bir kamuya açık veri kümesi üzerinde yürütülmüş olması, bulguların dış doğrulama gereksinimini artırmaktadır.

## 5. Genel Sonuçlar ve Gelecek Planı

Bu araştırma, üniversite öğrencilerinin stres düzeylerinin yaşam tarzı ve akademik değişkenler kullanılarak yüksek doğrulukla tahmin edilebildiğini göstermektedir. Çalışmada sınıf dengesizliğinin olası etkisini değerlendirmek amacıyla iki farklı eğitim düzeni kullanılmıştır: Dengesiz eğitim/doğrulama veri kümesi (özgün dağılım korunarak) ve yalnızca eğitim/doğrulama verisi üzerinde CTGAN ile dengeleme uygulanarak oluşturulan Dengeli eğitim/doğrulama veri kümesi. Her iki düzende de bağımsız test kümesi en başta ayrılmış ve süreç boyunca değiştirilmeden korunmuştur. Karşılaştırmalı değerlendirme sonucunda RF modeli her iki eğitim düzeninde de en yüksek performansı sergilemiş ve test kümesinde kusursuz sınıflandırma sağlamıştır (Doğruluk = 1.0000, Makro AUC = 1.0000, Makro F1 = 1.0000).

Model davranışının yorumlanabilirliği ve bulguların yöntemsel güvenilirliği, açıklanabilirlik ve sağlamlık analizleriyle desteklenmiştir. SHAP sonuçları, stres düzeyi tahmininde en belirleyici değişkenlerin *Günlük Çalışma Süresi* ve *Uyku Süresi* olduğunu; daha sonra *Genel Ağırlıklı Not Ortalaması* değişkeninin katkı sunduğunu göstermiştir. Bu baskın etkinin nicel doğrulaması için uygulanan ablasyon (özellik çıkarma) deneyinde günlük çalışma süresi değişkeni çıkarıldığında performansta belirgin düşüş gözlenmiş; böylece yüksek başarımın önemli ölçüde bu değişkenin taşıdığı bilgiyle ilişkili olduğu ortaya konmuştur. Ek olarak, etiket permütasyon testi negatif kontrol olarak kullanılarak etiketler rastgeleleştirildiğinde modelin ayırt etme gücünün şans düzeyine gerilediği gösterilmiş; bu bulgu yüksek başarımın rastlantısal uyum veya yöntemsel sorunlarla açıklanmadığını desteklemiştir.

Gelecek çalışmalarda, geliştirilen yaklaşımın farklı örneklemler üzerinde doğrulanması ve yerel bağlamlara uyarlanması hedeflenmektedir. Bu kapsamda, başta Marmara Üniversitesi olmak üzere Türkiye'de farklı üniversitelerden daha geniş bir örneklemle veri toplanarak modelin dış doğrulamasının yapılması ve gerekirse yeniden kalibre edilmesi planlanmaktadır. Türkiye bağlamına özgü akademik ve sosyal stres etmenlerini daha iyi yansıtacak ek değişkenlerin dahil edilmesiyle modelin kültürel duyarlılığının artırılması amaçlanmaktadır. Elde edilen çıktılar temelinde, üniversite birimlerinin kullanabileceği, risk düzeyini izleyen ve uygun yönlendirmeleri destekleyen ölçeklenebilir bir erken uyarı yaklaşımının tasarlanması da uygulamaya dönük bir sonraki adım olarak değerlendirilmektedir.

### Teşekkür

Bu çalışmanın fikir aşamasından tasarımına, veri toplama ve analiz süreçlerinden makalenin kaleme alınmasına kadar tüm aşamaları E.Ç. tarafından yürütülmüştür. Çalışma kurgusu itibarıyla etik kurul onayı gerektirmemekte olup, herhangi bir çıkar çatışması bulunmamaktadır. Ayrıca, bu araştırma kapsamında herhangi bir kurum veya kuruluştan finansal destek ya da ödenek alınmamıştır.

## Kaynaklar


[1] Lazarou E, Exarchos TP. Predicting stress levels using physiological data: Real-time stress prediction models utilizing wearable devices. AIMS neuroscience 2024; 11(2): 76.

[2] Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of general psychiatry 2005; 62(6): 593-602.

[3] Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, Chen-Li D, Iacobucci M, ve diğerleri. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. Journal of affective disorders 2020; 277: 55-64.

[4] El Morr C, Jammal M, Bou-Hamad I, Hijazi S, Ayna D, Romani M, Hoteit R. Predictive machine learning models for assessing lebanese university students' depression, anxiety, and stress during COVID-19. Journal of Primary Care & Community Health 2024; 15: 21501319241235588.

[5] Sieverding M, Schmidt LI, Obergfell J, Scheiter F. Stress und studienzufriedenheit bei bachelor-und diplom-psychologiestudierenden im vergleich. Psychologische Rundschau 2013; 64(2): 94-100.

[6] Ahuja R, Banga A. Mental stress detection in university students using machine learning algorithms. Procedia Computer Science 2019; 152: 349-53.

[7] Nayan MIH, Uddin MSG, Hossain MI, Alam MM, Zinnia MA, Haq I, Rahman MM, Ria R, ve diğerleri. Comparison of the performance of machine learning-based algorithms for predicting depression and anxiety among University Students in Bangladesh: A result of the first wave of the COVID-19 pandemic. Asian Journal of Social Health and Behavior 2022; 5(2): 75-84.

[8] Medikonda J. A clinical and technical methodological review on stress detection and sleep quality prediction in an academic environment. Computer methods and programs in biomedicine 2023; 235: 107521.

[9] Baba A, Bunji K. Prediction of mental health problem using annual student health survey: machine learning approach. JMIR Mental Health 2023; 10: e42420.

[10] Ratul IJ, Nishat MM, Faisal F, Sultana S, Ahmed A, Al Mamun MA. Analyzing perceived psychological and social stress of university students: A machine learning approach. Heliyon 2023; 9(6): e17004.

[11] Zhang J, Tee M, Lin C, Huili S. ZJ-EduFormer: Predicting Supply Chain Student Stress Using Transformer. In: 2024 5th International Conference on Information Science and Education (ICISE-IE); 2024; Rhodes, Greece. New York, NY, USA: IEEE. pp. 307-311.

[12] Đokić A, Stefanović H, Dudić D. Comparative analysis of classification model performance in predicting stress levels in students. In: Annual Conference on Challenges of Contemporary Higher Education (ACCHE); 3 February 2025; Belgrade, Serbia. pp. 61-66.

[13] Ghara A, Khan A, Mahata P. Lifestyle-Based Student Stress Detection with Real-Time ML Recommendations. International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET) 2025; 14(6): 5.

[14] Bastos AF, Fernandes-Jr O, Liberal SP, Pires AJL, Lage LA, Grichtchouk O, Cardoso AR, de Oliveira L, ve diğerleri. Academic-related stressors predict depressive symptoms in graduate students: A machine learning study. Behavioural Brain Research 2025; 478: 115328.

[15] Tuan T, Loan D, Buddhahai B. Classification models combined with optimized features for mental stress prediction. International Journal of Data and Network Science 2025; 9: 737-50.

[16] Rahunathan L, A SN, G A, P R, S SL, N SV. Machine Learning Model for Student Stress Level Prediction. In: 2025 4th OPJU International Technology Conference (OTCON); 9-11 April 2025; Raigarh, India. pp. 1-7.

[17] Kumar S. Student Lifestyle Dataset: Daily Lifestyle and Academic Performance of Students. Kaggle; 2025.

[18] Xu L, Skoularidou M, Cuesta-Infante A, Veeramachaneni K. Modeling Tabular data using Conditional GAN. In: Advances in Neural Information Processing Systems 32; 8-14 December 2019; Vancouver, Canada. Red Hook, NY, USA: Curran Associates, Inc. pp. 7335-7345

[19] SDMetrics. SDMetrics 2025 [Available from: https://github.com/sdv-dev/SDMetrics].

[20] Esteban C, Hyland SL, Rätsch G. Real-valued (medical) time series generation with recurrent conditional GANs. arXiv preprint 2017; arXiv:1706.02633.

[21] Papadaki E, Vrahatis AG, Kotsiantis S. Exploring innovative approaches to synthetic tabular data generation. Electronics 2024; 13(10): 1965.

[22] Williamson BD, Gilbert PB, Simon NR, Carone M. A general framework for inference on algorithm-agnostic variable importance. Journal of the American Statistical Association 2023; 118(543): 1645-58.

[23] Ojala M, Garriga GC. Permutation tests for studying classifier performance. J Mach Learn Res 2010; 11: 1833-1863.

[24] Tran D-S, Nguyen D-T, Nguyen T-H, Tran C-T-P, Duong-Quy S, Nguyen T-H. Stress and sleep quality in medical students: a cross-sectional study from Vietnam. Frontiers in psychiatry 2023; 14: 1297605.

[25] Ban DJ, Lee TJ. Sleep duration, subjective sleep disturbances and associated factors among university students in Korea. Journal of Korean Medical Science 2001; 16(4): 475.

[26] Bozorgmehr A, Weltermann B. Prediction of chronic stress and protective factors in adults: development of an interpretable prediction model based on XGBoost and SHAP using national cross-sectional DEGS1 data. JMIR AI 2023; 2: e41868.