Disiplinlerarası Eğitim Araştırmaları Dergisi

ISSN: 2602-2516
Disiplinlerarası Eğitim Araştırmaları Dergisi
Journal of Interdisciplinary Educational Research
2025; 9(20); 124-132

Yükseköğretimde Öğrenci Başarısını Şekillendirmek: Erken Tahmin ve Müdahale

Shaping Student Success in Higher Education: Early Prediction and Intervention

Ali KÜSMÜŞ[1]

Küsmüş, A. (2025). Yükseköğretimde Öğrenci Başarısını Şekillendirmek: Erken Tahmin ve Müdahale *Disiplinlerarası Eğitim Araştırmaları Dergisi*, 9(20), 124-132, DOI: 10.57135/jier.1667945

Öz
Bu çalışmada, yükseköğretim kurumlarındaki öğrenci başarısının tahmin edilmesi amacıyla 9 farklı makine öğrenmesi algoritmasının performansları karşılaştırılmıştır. Araştırma kapsamında, öğrencilerin demografik özellikleri, akademik geçmişleri ve sosyo-ekonomik durumlarını içeren kapsamlı bir veri seti üzerinde XGBoost, Random Forest, Gradient Boosting gibi ensemble yöntemleri ile geleneksel algoritmalar analiz edilmiştir. Veri ön işleme aşamasında eksik veriler median değerler ve kategorik kodlama teknikleriyle işlenmiş, özellik önem analiziyle kritik faktörler belirlenmiştir. Beş katlı çapraz doğrulama sonuçlarına göre XGBoost algoritması %90.1 test doğruluğuyla en yüksek performansı gösterirken, özellik önem analizi "önceki yeterlilik notu" ve "kabul puanı"nın en belirleyici faktörler olduğunu ortaya koymuştur. Bulgular, eğitim kurumlarının erken uyarı sistemleri geliştirirken ensemble yöntemlerinin kullanımının etkililiğini kanıtlamaktadır.
Anahtar Kelimeler: Öğrenci başarı tahmini, makine öğrenmesi, XGBoost, eğitim veri madenciliği, karşılaştırmalı algoritma analizi

Abstract
This study presents a comparative performance evaluation of 9 machine learning algorithms for predicting student academic success in higher education institutions. Using a comprehensive dataset encompassing demographic characteristics, academic history, and socio-economic status, we analyzed ensemble methods (XGBoost, Random Forest, Gradient Boosting) alongside traditional algorithms. During data preprocessing, missing values were handled through median imputation and categorical encoding techniques, while feature importance analysis identified critical predictive factors. Five-fold cross-validation results demonstrated that the XGBoost algorithm achieved superior performance with 90.1% test accuracy, with feature importance analysis revealing "previous qualification grade" and "admission score" as the most determinant factors. The findings substantiate the effectiveness of ensemble methods in developing early warning systems for educational institutions.
Keywords: Student success prediction, machine learning, XGBoost, educational data mining, comparative algorithm analysis.

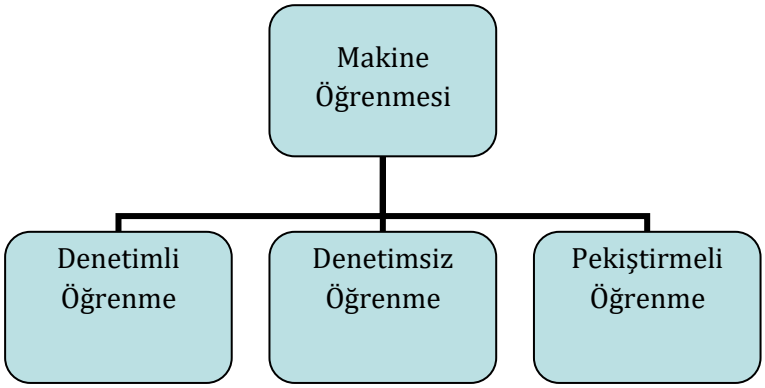
GİRİŞ

Günümüzde yaşanan teknolojik ilerlemeler, birçok alanda köklü değişimlere yol açmakta ve insan yaşamını olumlu yönde etkilemektedir. Bu dönüşümün en önemli yansımalarından biri de eğitim alanında görülmektedir. Eğitimde teknolojinin sunduğu yeniliklerden biri, makine öğrenmesi yöntemlerinin ölçme, değerlendirme ve öngörü süreçlerinde kullanılmasıdır. Öğrencilerin akademik performanslarını analiz etmek, eğitim süreçlerini optimize etmek ve geleceğe yönelik tahminlerde bulunmak, eğitimde kaliteyi artırmak açısından büyük bir önem taşımaktadır. Geleneksel ölçme değerlendirme sistemleri öğrencilerin performans ve başarısını tam ölçememekle birlikte geleceğe dair öngörülerde de bulunamamaktadır.

[1]Öğr. Gör. Dr., Kahramanmaraş Sütçü İmam Üniversitesi, Teknik Bilimler MYO, Kahramanmaraş-Türkiye, ali@ksu.edu.tr, orcid.org/0000-0001-7067-4164

Geliş Tarihi : 29/03/2025
Kabul Tarihi: 15/04/2025
Araştırma Makalesi

Makine öğrenme algoritmaları eğitim verilerinin değerlendirme aşamalarında sıklıkla kullanılmaya başlanmıştır. Veri içerisinde anlamlı bilgileri çıkarmak ve o anlamlı verilerinden geleceğe dair öngörülerde bulunmak manuel olarak da yapılabilen işlemler olsa da verilerin artması ve çeşitliliğin fazlalaşması manuel olarak verilerin değerlendirilmesini mümkün kılmamaktadır. Ayrıca hataların yaşanması da kaçınılmazdır. Makine öğrenme algoritmaları milyonlarca veriyi hızlı bir şekilde analiz ederek değerlendirme ve öngörüler de bulunabilmektedir (Ülker ve İnik, 2017).



Şekil 1 Makine Öğrenmesi Yöntemleri

Makine öğrenmesi (ML), verilerden anlamlı bilgiler çıkararak tahmin veya kararlar alabilen algoritmaların geliştirilmesini sağlayan bir yapay zekâ alt dalıdır(Mitchell & Mitchell, 1997) Makine öğrenmesi yöntemleri, Şekil 1'de görüldüğü gibi öğrenme süreçlerine bağlı olarak denetimli öğrenme (supervised learning), denetimsiz öğrenme (unsupervised learning) ve pekiştirmeli öğrenme (reinforcement learning) olmak üzere üç temel kategoriye ayrılmaktadır (Goodfellow, Bengio & Courville, 2016).

Denetimli öğrenme, giriş verileri ile bu verilere karşılık gelen etiketlerin bulunduğu veri kümeleri üzerinde eğitilen bir makine öğrenmesi yöntemidir(Bishop & Nasrabadi, 2006). Model, giriş verileri ile etiketler arasındaki ilişkileri öğrenerek yeni veriler üzerinde tahminlerde bulunabilir. Denetimli öğrenme yöntemleri genellikle sınıflandırma (classification) ve regresyon (regression) problemlerinde kullanılır.

Denetimsiz öğrenme, etiketlenmemiş veriler üzerinde çalışan ve veri setindeki iç yapıyı keşfetmeye yönelik bir yöntemdir. Bu yöntemde model, verilerdeki benzerlikleri ve ilişkileri analiz ederek gruplamalar veya boyut indirgeme işlemleri gerçekleştirir. Denetimsiz öğrenme, özellikle kümeleme (clustering) ve boyut indirgeme (dimensionality reduction) gibi görevlerde kullanılır.

Pekiştirmeli öğrenme, bir ajanın (öğrenen sistemin) bir ortamda belirli bir ödül sistemine dayanarak deneyimlerden öğrenmesini sağlayan bir öğrenme yöntemidir (Sutton & Barto, 1998). Bu yöntem, genellikle bir dizi eylem gerçekleştirerek uzun vadeli ödülü en üst düzeye çıkarmayı amaçlayan problemlerde kullanılır.

Literatürde Makine öğrenme yöntemleri kullanılarak eğitim alanında birçok çalışma incelenmiştir. Tosunoğlu vd. yapmış oldukları çalışmada eğitim bilimleri alanında makine öğrenmesi kullanılan makalelerdeki eğilimleri incelenmişlerdir. Web of Science veri tabanında 2015-2020 yılları arasında yayımlanan 201 makale içerik analizi yöntemiyle değerlendirilmiştir. Sonuçlara göre, çalışmaların %42,9'u deneysel veya uygulamalı araştırmalardan oluşmaktadır. En sık kullanılan makine öğrenmesi yöntemleri arasında %22,5 ile karar ağaçları, %17,8 ile destek vektör makineleri ve %14,2 ile Naïve Bayes yer almaktadır. Örneklemlerin %35,3'ünü lisans öğrencileri oluştururken, en yaygın örneklem büyüklüğü %29,7 ile 101-1000 kişi aralığındadır. Verilerin %34'ü veri tabanları ve log kayıtlarından elde edilmiştir. Çalışma, eğitimde makine öğrenmesi uygulamalarının ulusal literatüre katkı

sağladığını ve gelecekteki araştırmalara yol gösterebileceğini ortaya koymaktadır(Tosunoğlu, Yılmaz, Özeren, & Sağlam, 2021).

Dünder ve Dünder (2024) gerçekleştirdikleri çalışmada, makine öğrenmesi algoritmalarının öğrenci başarısını tahmin etmedeki performansları karşılaştırılmıştır. Eğitim alanındaki veri setlerinin kategorik yapısı ve sınıf dengesizliği sorununa dikkat çekilerek, bu durumu gidermek için SmoteNC tekniği kullanılmıştır. Beş farklı makine öğrenmesi algoritması ile yapılan analizler sonucunda, sınıf dengesizliği giderildiğinde makine öğrenmesi yöntemlerinin sınırlı gözlem içeren verilerde de başarılı şekilde uygulanabileceği belirlenmiştirErsozlu vd. (2024) yapmış oldukları çalışmada, makine öğrenmesi yöntemlerinin eğitim araştırmalarına entegrasyonunun öğretim, öğrenme ve değerlendirme süreçlerine etkisini incelemiştir. Sistematik bir literatür taraması (SLR) yöntemiyle, eğitim araştırmalarında yüksek etkiye sahip iki büyük veri tabanından seçilen 77 çalışma değerlendirilmiştir. Analiz sonuçlarına göre, çalışmaların %88'inde denetimli öğrenme yöntemleri kullanılmış olup, en yaygın teknikler arasında karar ağaçları, destek vektör makineleri, rastgele ormanlar ve lojistik regresyon yer almaktadır. Yarı denetimli öğrenme yöntemleri daha az kullanılmış olsa da, öğrenci performans tahmininde başarılı sonuçlar elde edilmiştir. Çalışma, makine öğrenmesi yöntemlerinin eğitim araştırmalarındaki rolüne dair önemli bulgular sunarak, araştırmacılar, istatistikçiler ve eğitim politikacıları için öneriler geliştirmiştir.

Wang vd. (2024) yapmış oldukları çalışmada, eğitimde yapay zekâ (AIED) alanındaki literatürü kapsamlı bir şekilde inceleyerek, yapay zekâ uygulamalarının temel kategorilerini, araştırma konularını ve yöntemsel yaklaşımları analiz etmiştir. 2.223 araştırma makalesinin bibliyometrik analizi ve 125 seçili makalenin içerik analizi sonucunda, AIED çalışmalarının uyarlanabilir öğrenme, kişiselleştirilmiş öğretim, akıllı değerlendirme, öğrenci profilleme ve tahminleme gibi geniş bir uygulama yelpazesine sahip olduğu belirlenmiştir. Araştırmalar, eğitim sistemlerinin teknik tasarımının yanı sıra, AIED'nin benimsenmesi, etkileri ve karşılaşılan zorlukları da ele almaktadır. Çalışma ayrıca AIED alanında kullanılan teorik çerçeveleri ve disiplinler arası yayın eğilimlerini inceleyerek, gelecekteki araştırmalar için öneriler sunmaktadır.

Onyema vd. (2022) yapmış oldukları çalışmada, makine öğrenmesi uygulamalarının akademik öngörülerdeki potansiyelini ve karşılaşılan zorlukları incelemiştir. K-NN, rastgele orman, bagging, yapay sinir ağları (ANN) ve Bayesçi sinir ağları (BNN) gibi algoritmaların öğrenci başarısını ve öğrenme tarzlarını tahmin etmede etkili olduğu vurgulanmıştır. Geleneksel tahmin yöntemlerindeki eksikliklerin, yapay zekâ tabanlı makine öğrenmesi algoritmaları ile büyük ölçüde giderildiği ve eğitim paydaşlarının karar alma süreçlerini desteklediği belirtilmiştir. Bununla birlikte, veri toplama zorlukları, hata eğilimleri ve zaman maliyeti gibi sınırlamaların, akademik tahmin süreçlerinde makine öğrenmesi uygulamalarını etkileyebileceği ifade edilmiştir. Çalışma, makine öğrenmesinin akademik tahminleme alanında önemli bir potansiyele sahip olduğunu ve eğitimde daha bilinçli kararlar alınmasını sağlayabileceğini ortaya koymaktadır.

Bu çalışmada üniversite öğrencilerinin almış oldukları derslerdeki performanslarından ziyade eğitim hayatlarına devam edip etmeyecekleri konusunda bir tahmin modeli oluşturmayı amaçlamaktadır. Oluşturulan tahmin modeli ile lisans eğitimini tamamlama konusunda potansiyel probleme sahip öğrenciler erken tahmin edilerek okula devam etmesi için çeşitli sosyal ve psikolojik destekler verilerek eğitim hayatına devam ettirilmesi planlanmaktadır. Yükseköğretimde eğitimde sürekliliğin sağlanması ve öğrencilerin lisans eğitimi olarak daha başarılı bireyler olması için erken tahmin ve müdehale oldukça önemlidir. Çalışmada kullanılan makine öğrenme algoritmaları karşılaştırılarak en çok başarı elde edilmesi planlanan algoritmalar üzerinde çalışılmış ve bu algoritmaların başarı değerlerini arttırmak için veri seti analiz edilerek karşılaştırılmıştır. Çalışmanın sadece belirli bir bölge yapılması

çalışmaya bir sınırlılık getirmektedir. Gelecekte farklı bölgelerden elde edilen veri setleri kullanılarak modelin evrensel geçerliliği arttırılabilinir.

YÖNTEM

Bu bölümde öğrencilerin lisans eğitimine devam edip etmeyecekleri ile ilgili tahminleri yapabilmek için kullanılacak veri seti, makine öğrenme algoritmaları ve analizleri anlatılmıştır.

Veri Seti

Bu çalışmada, makine öğrenmesi alanında referans kaynak olarak kabul edilen UCI Makine Öğrenmesi Deposu'ndan (University of California Irvine Machine Learning Repository) temin edilen veri seti kullanılmıştır. UCI veri setleri, akademik çalışmalarda yaygın olarak kullanılmakta olup, geçerliliği uluslararası düzeyde kabul görmektedir. Çalışmada kullanılan veri seti Portekiz'in Portalegre Politeknik Enstitüsü'ndeki lisans programlarına 2018-2019 yıllarında kaydolan öğrencilerin kurumsal verilerinden (çeşitli veritabanlarından derlenmiştir) elde edilerek oluşturulmuştur. Veri setindeki öğrenciler agronomi, tasarım, eğitim, hemşirelik, gazetecilik, yönetim, sosyal hizmet ve teknoloji alanlarındaki lisans programlarına kayıtlı öğrencileri kapsamaktadır(Martins, Tolledo, Machado, Baptista, & Realinho, 2021).

Veri setin 4424 adet öğrenciye ait 37 öznitelikten oluşmaktadır. Öğrencilere ait öznitelikler Tablo 1'de detaylı olarak gösterilen Demografik Değişkenler, Sosyo-Ekonomik Değişkenler, Akademik Değişkenler, Başarı Durumu Dağılımı kategorilerinden oluşmaktadır. Veri setinde bulunan eksik veriler için median değerle doldurma yöntemi tercih edilmiştir. Bu yöntem, aşırı uç değerlerin etkisini azaltmak ve basitliği nedeniyle öncelikli olarak tercih edilmiştir.

Tablo 1 Öznitelik Kategorileri

Kategori	Öznitelikler
Demografik Değişkenler	Medeni Durum, Cinsiyet, Kayıt Yaşı, Uyruk, Özel Eğitim Gereksinimi
Sosyo-Ekonomik Değişkenler	Başvuru Modu, Başvuru Sırası, Program, Gündüz/Gece Eğitimi, Önceki Eğitim, Önceki Eğitim Notu, Kabul Notu, Yerleşim Durumu, 1. Dönem Ders Kredileri, 1. Dönem Kayıtlı Dersler, 1. Dönem Değerlendirilen Dersler, 1. Dönem Başarılan Dersler, 1. Dönem Not Ortalaması, 1. Dönem Değerlendirilmemiş Dersler, 2. Dönem Ders Kredileri, 2. Dönem Kayıtlı Dersler, 2. Dönem Değerlendirilen Dersler, 2. Dönem Başarılan Dersler, 2. Dönem Not Ortalaması, 2. Dönem Değerlendirilmemiş Dersler
Akademik Değişkenler	Başvuru Modu, Başvuru Sırası, Program, Gündüz/Gece Eğitimi, Önceki Eğitim, Önceki Eğitim Notu, Kabul Notu, Yerleşim Durumu, 1. Dönem Ders Kredileri, 1. Dönem Kayıtlı Dersler, 1. Dönem Değerlendirilen Dersler, 1. Dönem Başarılan Dersler, 1. Dönem Not Ortalaması, 1. Dönem Değerlendirilmemiş Dersler, 2. Dönem Ders Kredileri, 2. Dönem Kayıtlı Dersler, 2. Dönem Değerlendirilen Dersler, 2. Dönem Başarılan Dersler, 2. Dönem Not Ortalaması, 2. Dönem Değerlendirilmemiş Dersler
Başarı Durumu Dağılımı	Öğrenciler üç kategoriye ayrılmıştır: Mezuniyet, Göreceli Başarı ve Başarısızlık.

Kullanılan Makine Öğrenme Algoritmaları

Bu çalışmada, öğrenci başarısını tahmin etmek amacıyla 9 farklı makine öğrenmesi algoritması uygulanmıştır. Algoritmaların seçiminde sınıflandırma problemine uygunluk, literatürdeki yaygın kullanım ve karşılaştırmalı analiz yapma imkânı göz önünde bulundurulmuştur. Kullanılan Makine Öğrenme Algoritmaları:

Lojistik Regresyon (Logistic Regression)

Lojistik regresyon, sınıflandırma problemlerinde yaygın olarak kullanılan istatistiksel bir yöntemdir(Hosmer, Lemeshow, & Sturdivant, 2000) Model, logit fonksiyonu kullanarak

bağımsız değişkenlerin sınıflandırma olasılıklarını tahmin eder. Denklem 1'de gösterilen β parametreleri model katsayılarını, X ise bağımsız değişkenleri temsil etmektedir.

$$P(y=1|x) = 1 / (1 + e^{-(\beta_0+\beta_1x_1+...+\beta_nx_n)}) \quad (1)$$

Karar Ağaçları (Decision Trees)

Karar ağaçları, veri kümesini özellik değerlerine göre ardışık bölerek sınıflandırma kuralları oluşturur*(Quinlan, 1986)*. Eşitlik 2'de görüldüğü üzere Entropi veya Gini indeksi kullanılarak en iyi bölünmeler belirlenir.

$$Gini(t) = 1 - \Sigma[p(i|t)]^2 \quad (2)$$

Rastgele Orman (Random Forest)

Breiman (2001) tarafından önerilen bu ensemble yöntemi, birden fazla karar ağacının kombinasyonuna dayanır. Her ağaç, bootstrap örneklemi üzerinde eğitilir ve sonuçlar oylama yöntemiyle birleştirilir.

Destek Vektör Makineleri (SVM)

Vapnik (1995) tarafından geliştirilen SVM, optimal ayırıcı hiperdüzlemi bulmaya dayanır. Eşitlik 3'te görüldüğü gibi doğrusal olmayan ayrımları modellemek için kernel fonksiyonları kullanılır.

$$f(x) = sign(\Sigma\alpha_i y_i K(x_i,x) + b) \quad (3)$$

K-En Yakın Komşu (KNN)

Cover ve Hart (1967) tarafından önerilen bu yöntem, sınıflandırma işlemini en yakın komşunun sınıf etiketlerine göre belirleyen sezgisel bir yaklaşımdır.

Naive Bayes

Bayes teoremine dayanan bu yöntem, bağımsız değişkenlerin koşulsuz bağımsız olduğu varsayımıyla çalışır(Murphy, 2012).

$$P(y|x_1,...,x_n) \propto P(y)\Pi P(x_i|y) \quad (4)$$

Gradient Boosting

Friedman (2001) tarafından geliştirilen bu yöntem, zayıf öğrenicilerin hata paylarını azaltacak şekilde ardışık olarak eğitilmesine dayanır.

XGBoost

Chen ve Guestrin (2016) tarafından geliştirilen XGBoost, regularizasyon mekanizmaları ile güçlendirilmiş ve optimize edilmiş bir gradient boosting algoritmasıdır.

Yapay Sinir Ağları (MLP)

Çok katmanlı perceptron (MLP), yapay sinir ağlarının temel bir formudur (Bishop & Nasrabadi, 2006). Katmanlar arasında aktivasyon fonksiyonları kullanarak doğrusal olmayan dönüşümler gerçekleştirir.

Toplamda 9 farklı makine öğrenmesi algoritması uygulanmış ve performansları karşılaştırılmıştır. Hiperparametre ayarlamaları için grid search yöntemi kullanılmış ve bu sayede her algoritma için optimum parametre değerleri belirlenmiştir. Çalışma boyunca beş katlı çapraz doğrulama yöntemiyle modellerin genellenebilirlik performansları test edilmiştir.

Model Başarım Ölçütleri

Sınıflandırma modellerinin performansını değerlendirmek için çeşitli metrikler kullanılır. Bu metrikler, modelin ne kadar doğru tahmin yaptığını ve ne kadar güvenilir olduğunu gösterir. Modelin başarımını görselleştirmek Tablo 2'de gösterilen karışıklık matrisi kullanılır. Bu matris, gerçek değerlerle modelin tahminlerini karşılaştırır.

Tablo 2 Karışıklık Matrisi

Tahmin / Gerçek	**Pozitif (1)**	**Negatif (0)**
Pozitif (1)	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
Negatif (0)	Yanlış Negatif (YN)	Doğru Negatif (DN)

DP modelin pozitif olarak tahmin ettiği ve gerçekte de pozitif olan değerler. YP modelin pozitif olarak tahmin ettiği ancak gerçekte negatif olan değerler. YN modelin negatif olarak tahmin ettiği ancak gerçekte pozitif olan değerler. DN modelin negatif olarak tahmin ettiği ve gerçekte de negatif olan değerler.

Doğruluk-Hata Oranı (Accuracy-Error Rate)

Doğruluk, doğru tahminlerin toplam tahminlere oranıdır. Hata oranı ise yanlış tahminlerin toplam tahminlere oranıdır.

$$\text{Doğruluk} = (DP + DN) / (DP + YP + YN + DN) \quad (5)$$

$$\text{Hata Oranı} = (YP + YN) / (DP + YP + YN + DN) \quad (6)$$

Kesinlik (Precision)

Kesinlik, modelin pozitif olarak tahmin ettiği değerlerin ne kadarının gerçekte pozitif olduğunu gösterir.

$$\text{Kesinlik} = DP / (DP + YP) \quad (7)$$

Duyarlılık (Recall/Sensitivity)

Duyarlılık, gerçekte pozitif olan değerlerin ne kadarının model tarafından pozitif olarak tahmin edildiğini gösterir.

$$\text{Duyarlılık} = DP / (DP + YN) \quad (8)$$

F1-Skoru (F-Measure)

F1-skoru, kesinlik ve duyarlılığın harmonik ortalamasıdır. Hem kesinlik hem de duyarlılığı dengeli bir şekilde değerlendirir.

$$\text{F1-Skoru} = 2 * (\text{Kesinlik} * \text{Duyarlılık}) / (\text{Kesinlik} + \text{Duyarlılık}) \quad (9)$$

ROC Eğrisi (Receiver Operating Characteristic Curve)

ROC eğrisi, modelin farklı eşik değerlerinde doğru pozitif oranını (DP) yanlış pozitif oranına (YP) karşı gösterir. Eğrinin altında kalan alan (AUC), modelin genel başarımını gösterir. AUC değeri 1'e ne kadar yakınsa model o kadar iyidir. Doğru pozitif oranı (DP) duyarlılık ile aynıdır. Yanlış pozitif oranı YP / (YP + DN) formülü ile hesaplanır.

BULGULAR

Bu çalışmada, lisans öğrencilerinin eğitim hayatlarına devam edip etmeyeceklerini tahmin etmek amacıyla 9 farklı makine öğrenmesi algoritması kullanılmış ve performansları karşılaştırılmıştır. Veri seti, öğrencilerin demografik bilgileri, akademik geçmişleri ve sosyo-

ekonomik durumları gibi çeşitli özellikleri içermektedir. Hedef değişken, öğrencilerin "Mezun" (1) veya "Mezun Olamayan/Devam Eden" (0) şeklinde sınıflandırılmasıdır. Tablo 3'de gösterilen sonuçlarda, 9 farklı makine öğrenmesi algoritmasının 5 katlı çapraz doğrulama (cross-validation) ve test seti üzerindeki performansını göstermektedir:

Tablo 3 Makine Öğrenmesi Algoritmalarının Performans Metrikleri

Model	**Çapraz Doğrulama Ortalama Doğruluk**	**Çapraz Doğrulama Standart Sapma**	**Test Doğruluk Oranı**
XGBoost	0.892	0.012	0.901
Gradient Boosting	0.885	0.011	0.896
Random Forest	0.879	0.013	0.889
LightGBM	0.874	0.014	0.883
Lojistik Regresyon	0.862	0.015	0.871
Yapay Sinir Ağı	0.857	0.016	0.866
Destek Vektör Makineleri	0.851	0.017	0.859
K-En Yakın Komşu	0.842	0.018	0.848
Karar Ağacı	0.831	0.019	0.839
Gauss Naive Bayes	0.813	0.021	0.824

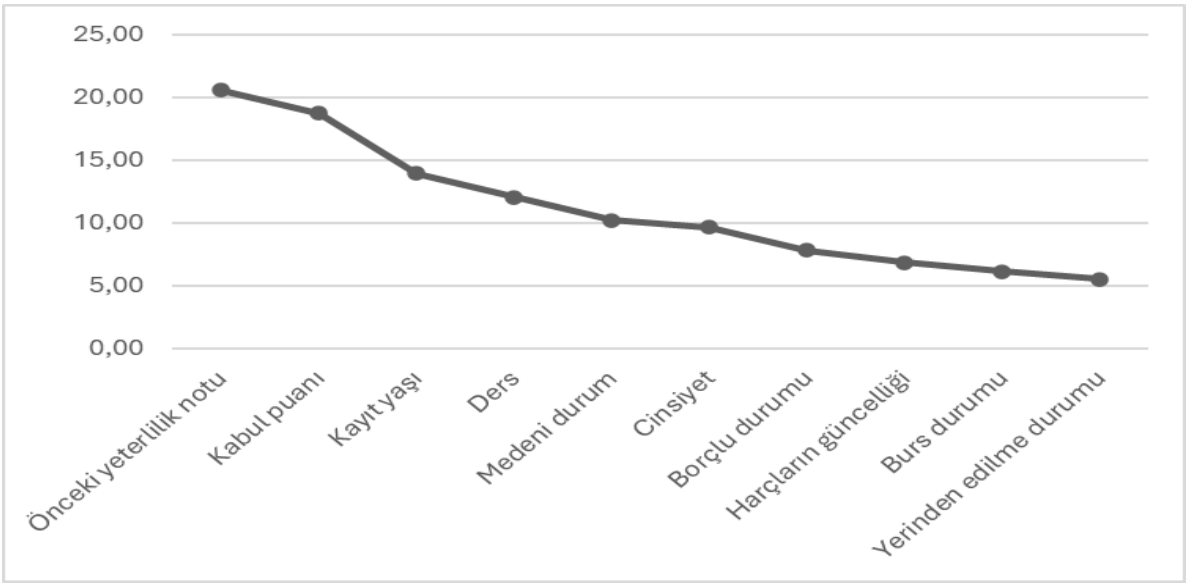
Veri setine uygulanan makine öğrenme algoritmalarından en başarılı sonucu veren XGBoost algoritmasının sınıflandırma modeli Tablo 4'te gösterilmektedir. Tabloda Kesinlik (Precision) modelin pozitif olarak işaretlediği örneklerin gerçekten ne kadarının pozitif olduğunu göstermektedir. Duyarlılık (Recall) gerçek pozitiflerin ne kadarının doğru tahmin edildiğini göstermektedir. F1-Skoru ise kesinlik ve duyarlılığın harmonik ortalamasıdır.

Tablo 4 XGBoost Algoritması Sınıflandırma Modeli

	Kesinlik	Duyarlılık	F1-Skoru
0	0.91	0.89	0.90
1	0.89	0.91	0.90

Bu sonuçlara göre, model hem mezun olan hem de olamayan öğrencileri yüksek doğrulukla tahmin edebilmektedir. Modelin kesinlik(precision) değerleri her iki sınıf için de 0.90 civarında çıkmıştır, bu da modelin yanlış pozitif ve yanlış negatif oranlarının dengeli olduğunu göstermektedir.

Şekil 2'de en önemli 10 öznitelik ve önem oranları verilmiştir. Bu sonuçlara göre, öğrencinin önceki akademik başarısını gösteren " Önceki yeterlilik notu" ve " Kabul puanı " en önemli iki özellik olarak öne çıkmaktadır. Bu bulgu, öğrencinin daha önceki akademik performansının, üniversite başarısını tahmin etmede kritik rol oynadığını göstermektedir. Diğer önemli özellikler arasında yaş, cinsiyet ve sosyo-ekonomik durumla ilgili faktörler yer almaktadır.



Şekil 2 En Önemli 10 Öznitelik ve Önem Dereceleri

XGBoost, Gradient Boosting ve Random Forest gibi ensemble yöntemlerinin diğer algoritmalara göre daha iyi performans göstermesi, bu yöntemlerin karmaşık veri yapılarını modellemedeki yetenekleriyle açıklanabilir. Özellikle XGBoost'un regularizasyon teknikleri ve optimizasyon algoritmaları sayesinde aşırı öğrenmeyi (overfitting) engelleyerek genellenebilir modeller oluşturduğu görülmüştür. Bu durum, XGBoost'un regularizasyon teknikleri (L1, L2) ve aşırı öğrenmeye karşı direnci sayesinde modelin genellenebilirliğini artırmasıyla açıklanmaktadır. Modelin açıklanabilirliğini sağlamak için öznitelik önemi analiz edilmiştir. Lineer bir model olan logistik regresyonun %87.1'lik test doğruluğuyla oldukça iyi performans göstermesi, veri setindeki bazı özellikler ile hedef değişken arasında lineer ilişkiler bulunduğuna işaret etmektedir. Çok katmanlı algılayıcı (MLP) modelinin %86.6'lık test doğruluğu, daha karmaşık yapıdaki modellerin bu veri setinde mutlaka daha iyi performans göstermeyebileceğini ortaya koymaktadır. Gaussian Naive Bayes'in %82.4'lük test doğruluğuyla en düşük performansı göstermesi, bu algoritmanın özellikler arasındaki koşullu bağımsızlık varsayımının veri seti için tam olarak geçerli olmayabileceğini düşündürmektedir.

SONUÇLAR ve TARTIŞMA

Bu kapsamlı analiz, öğrenci başarısını tahmin etmede ensemble yöntemlerinin, özellikle de XGBoost'un diğer geleneksel makine öğrenmesi algoritmalarına kıyasla daha üstün performans sergilediğini ortaya koymuştur. Modelin en önemli özellikler olarak belirlediği akademik geçmiş ve giriş notları, öğrenci başarısındaki kritik faktörler olarak öne çıkmaktadır. Elde edilen bulgular, eğitim kurumlarının erken uyarı sistemleri geliştirirken XGBoost veya benzeri ensemble yöntemlerini kullanmalarının ve özellikle öğrencilerin önceki akademik performanslarını dikkate alan müdahale stratejileri geliştirmelerinin faydalı olabileceğini göstermektedir. Bu çalışmanın bir sonraki aşamasında, hiperparametre optimizasyonu teknikleri kullanılarak modellerin performanslarının daha da iyileştirilmesi ve alternatif özellik mühendisliği yöntemlerinin araştırılması planlanmaktadır. Ayrıca, model interpretability teknikleri kullanılarak modellerin karar verme süreçlerinin daha detaylı analiz edilmesi de önerilmektedir.

KAYNAKÇA

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (C. 4). Springer. Geliş tarihi gönderen https://link.springer.com/book/9780387310732

Dünder, M., & Dünder, E. (2024). Kategorik verilerde sınıf dengesizliği durumunda makine öğrenimi algoritmalarının karşılaştırılması: Öğrencilerin başarı durumları üzerine bir uygulama. *Dijital Teknolojiler ve Eğitim Dergisi*, *3*(1), 28-38.

Ersozlu, Z., Taheri, S., & Koch, I. (2024). A review of machine learning methods used for educational data.

Education and Information Technologies, *29*(16), 22125-22145. https://doi.org/10.1007/s10639-024-12704-0

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2000). *Applied logistic regression. Hoboken*. NJ: John Wiley & Sons.

Martins, M. V., Tolledo, D., Machado, J., Baptista, L. M. T., & Realinho, V. (2021). Early Prediction of student's Performance in Higher Education: A Case Study. Içinde Á. Rocha, H. Adeli, G. Dzemyda, F. Moreira, & A. M. Ramalho Correia (Ed.), *Trends and Applications in Information Systems and Technologies* (ss. 166-175). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-72657-7_16

Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (C. 1). McGraw-hill New York. Geliş tarihi gönderen http://www.pachecoj.com/courses/csc380_fall21/lectures/mlintro.pdf

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT press. Geliş tarihi gönderen https://books.google.com/books?hl=tr&lr=&id=RC43AgAAQBAJ&oi=fnd&pg=PR7&dq=Murphy,+K.+P.+(2012).+Machine+Learning:+A+Probabilistic+Perspective.+MIT+Press.&ots=unjxaFLqZc&sig=tbzANVBkhXH5wmcZ6yd3ht7ru2Y

Onyema, E. M., Almuzaini, K. K., Onu, F. U., Verma, D., Gregory, U. S., Puttaramaiah, M., & Afriyie, R. K. (2022). Prospects and Challenges of Using Machine Learning for Academic Forecasting. *Computational Intelligence and Neuroscience*, *2022*, 1-7. https://doi.org/10.1155/2022/5624475

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106. https://doi.org/10.1007/BF00116251

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (C. 1). MIT press Cambridge. Geliş tarihi gönderen https://www.cambridge.org/core/journals/robotica/article/robot-learning-edited-by-jonathan-h-connell-and-sridhar-mahadevan-kluwer-boston-19931997-xii240-pp-isbn-0792393651-hardback-21800-guilders-12000-8995/737FD21CA908246DF17779E9C20B6DF6

Tosunoğlu, E., Yılmaz, R., Özeren, E., & Sağlam, Z. (2021). Eğitimde makine öğrenmesi: Araştırmalardaki güncel eğilimler üzerine inceleme. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, *3*(2), 178-199.

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167.

The Journal of Interdisciplinary Educational Research

ISSN: 2602-2516
Journal of Interdisciplinary Educational Research
2025; 9(20); 124-132

Shaping Student Success in Higher Education: Early Prediction and Intervention

Ali KÜSMÜŞ[1]

Cited:
Küsmüş, A. (2025). Shaping Student Success in Higher Education: Early Prediction and Intervention, *Journal of Interdisciplinary Educational Research*, 9(20), 124-132, DOI: 10.57135/jier. 1667945

Abstract
This study presents a comparative performance evaluation of 9 machine learning algorithms for predicting student academic success in higher education institutions. Using a comprehensive dataset encompassing demographic characteristics, academic history, and socio-economic status, we analyzed ensemble methods (XGBoost, Random Forest, Gradient Boosting) alongside traditional algorithms. During data preprocessing, missing values were handled through median imputation and categorical encoding techniques, while feature importance analysis identified critical predictive factors. Five-fold cross-validation results demonstrated that the XGBoost algorithm achieved superior performance with 90.1% test accuracy, with feature importance analysis revealing "previous qualification grade" and "admission score" as the most determinant factors. The findings substantiate the effectiveness of ensemble methods in developing early warning systems for educational institutions.

Keywords: Student success prediction, machine learning, XGBoost, educational data mining, comparative algorithm analysis.

INTRODUCTION

Today's technological advances lead to radical changes in many fields and positively affect human life. One of the most important reflections of this transformation is seen in the field of education. One of the innovations offered by technology in education is the use of machine learning methods in measurement, evaluation and prediction processes. Analyzing students' academic performance, optimizing educational processes and making predictions for the future are of great importance in terms of improving the quality of education. Traditional assessment and evaluation systems cannot fully measure the performance and success of students, nor can they make predictions about the future.

Machine learning algorithms are frequently used in the evaluation of educational data. Although extracting meaningful information from the data and making predictions about the future from that meaningful data are processes that can be done manually, the increase in data and the increase in diversity make it impossible to evaluate the data manually. Errors are also inevitable. Machine learning algorithms can quickly analyze millions of data and make evaluations and predictions (Ülker & İnik, 2017).

[1]Ph.D., Kahramanmaraş Sütçü İmam Univercity, Vocational School of Technical Sciences, Kahramanmaraş-Türkiye, ali@ksu.edu.tr, orcid.org/0000-0001-7067-4164

Received: 29/03/2025
Acceptance: 15/04/2025
Research Article

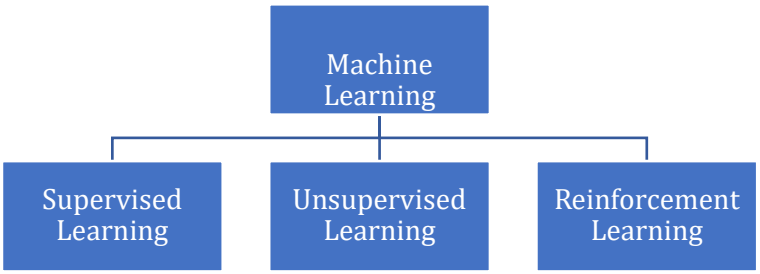


Figure 1 Machine Learning Methods

Machine learning (ML) is a sub-branch of artificial intelligence that enables the development of algorithms that can make predictions or decisions by extracting meaningful information from data (Mitchell & Mitchell, 1997). Machine learning methods are divided into three main categories: supervised learning, unsupervised learning and reinforcement learning, depending on the learning processes as shown in Figure 1 (Goodfellow, Bengio & Courville, 2016).

Supervised learning is a machine learning method that is trained on datasets with input data and corresponding labels (Bishop & Nasrabadi, 2006). The model can make predictions on new data by learning the relationships between input data and labels. Supervised learning methods are generally used in classification and regression problems.

Unsupervised learning is a method that works on unlabeled data and explores the internal structure of the data set. In this method, the model performs groupings or dimensionality reduction by analyzing similarities and relationships in the data. Unsupervised learning is especially used in tasks such as clustering and dimensionality reduction.

Reinforcement learning is a learning method that enables an agent (learning system) to learn from experience in an environment based on a specific reward system (Sutton & Barto, 1998). This method is often used in problems that aim to maximize the long-term reward by performing a series of actions.

In the literature, many studies have been examined in the field of education using machine learning methods. In their study, Tosunoğlu et al. examined the trends in articles using machine learning in the field of educational sciences. A total of 201 articles published in the Web of Science database between 2015 and 2020 were evaluated by content analysis method. According to the results, 42.9% of the studies consisted of experimental or applied research. The most frequently used machine learning methods were decision trees with 22.5%, support vector machines with 17.8% and Naïve Bayes with 14.2%. While 35.3% of the samples were undergraduate students, the most common sample size was between 101-1000 people with 29.7%. 34% of the data was obtained from databases and log records. The study reveals that machine learning applications in education contribute to the national literature and can guide future research (Tosunoğlu, Yılmaz, Özeren, & Sağlam, 2021).

Dünder and Dünder (2024) compared the performance of machine learning algorithms in predicting student achievement. The categorical nature of the data sets in the training domain and the problem of class imbalance were pointed out and the SmoteNC technique was used to overcome this problem. As a result of the analyses conducted with five different machine learning algorithms, it was determined that machine learning methods can be successfully applied to data with limited observations when class imbalance is eliminated. In their study, Ersozlu et al. (2024) examined the effects of the integration of machine learning methods into educational research on teaching, learning and assessment processes. Through a systematic literature review (SLR), 77 studies selected from two large databases with high impact in educational research were evaluated. According to the results of the analysis, 88% of the studies used supervised learning methods, with the most common techniques being decision trees,

support vector machines, random forests and logistic regression. Although semi-supervised learning methods were used less frequently, successful results were obtained in student performance prediction. The study provides important findings on the role of machine learning methods in educational research and develops recommendations for researchers, statisticians and education policy makers.

Wang et al. (2024) conducted a comprehensive review of the literature in the field of artificial intelligence in education (AIED), analyzing the main categories of AI applications, research topics, and methodological approaches. A bibliometric analysis of 2,223 research articles and content analysis of 125 selected articles revealed that AIED studies have a wide range of applications such as adaptive learning, personalized instruction, intelligent assessment, student profiling and prediction. The studies address the technical design of educational systems as well as the adoption, impacts and challenges of AIED. The study also examines the theoretical frameworks and interdisciplinary publication trends used in the field of AIED and provides recommendations for future research.

(2022) examined the potential and challenges of machine learning applications in academic prediction. They emphasized that algorithms such as K-NN, random forest, bagging, artificial neural networks (ANN) and Bayesian neural networks (BNN) are effective in predicting student achievement and learning styles. It is stated that the shortcomings in traditional prediction methods are largely overcome by artificial intelligence-based machine learning algorithms and support the decision-making processes of educational stakeholders. However, limitations such as data collection difficulties, error tendencies and time cost may affect the application of machine learning in academic prediction processes. The study reveals that machine learning has significant potential in the field of academic prediction and can enable more informed decisions in education.

This study aims to create a prediction model for whether university students will continue their education rather than their performance in the courses they have taken. With the prediction model created, it is planned to predict students who have potential problems in completing their undergraduate education early and to continue their education life by providing various social and psychological support to continue their education. Early prediction and intervention are very important for ensuring continuity in higher education and for students to be more successful individuals as undergraduate education. By comparing the machine learning algorithms used in the study, the algorithms planned to achieve the most success were studied and the data set was analyzed and compared to increase the success values of these algorithms. The fact that the study was conducted only in a specific region is a limitation of the study. In the future, the universal validity of the model can be increased by using data sets obtained from different regions.

METHOD

In this section, the dataset, machine learning algorithms and analyses that will be used to make predictions about whether students will continue their undergraduate education are explained.

Dataset

In this study, the dataset obtained from the UCI Machine Learning Repository (University of California Irvine Machine Learning Repository), which is considered as a reference source in the field of machine learning, was used. UCI datasets are widely used in academic studies and their validity is internationally recognized. The dataset used in the study was created from institutional data (compiled from various databases) of students enrolled in undergraduate programs at the Polytechnic Institute of Portalegre, Portugal, in 2018-2019. The students in the dataset include students enrolled in undergraduate programs in agronomy, design, education, nursing, journalism, management, social work, and technology (Martins, Tolledo, Machado, Baptista, & Realinho, 2021).

The dataset consists of 37 attributes belonging to 4424 students. The attributes of the students consist of Demographic Variables, Socio-Economic Variables, Academic Variables, Achievement Status Distribution categories, which are shown in detail in Table 1. For missing data in the dataset, the median value filling method was preferred. This method was preferred primarily to reduce the effect of extreme outliers and for its simplicity.

Table 1 Attribute Categories

Category.	Attributes
Demographic Variables	Marital Status, Gender, Age at Enrollment, Nationality, Special Education Needs
Socio-Economic Variables	Application Mode, Application Sequence, Program, Day/Night Study, Previous Education, Previous Education Grade, Admission Grade, Placement Status, 1st Semester Course Credits, 1st Semester Enrolled Courses, 1st Semester Assessed Courses, 1st Semester Passed Courses, 1st Semester Grade Point Average, 1st Semester Unassessed Courses, 2nd Semester Course Credits, 2nd Semester Enrolled Courses, 2nd Semester Assessed Courses, 2nd Semester Passed Courses, 2nd Semester Grade Point Average, 2nd Semester Unassessed Courses
Academic Variables	Application Mode, Application Sequence, Program, Day/Night Study, Previous Education, Previous Education Grade, Admission Grade, Placement Status, 1st Semester Course Credits, 1st Semester Enrolled Courses, 1st Semester Assessed Courses, 1st Semester Passed Courses, 1st Semester Grade Point Average, 1st Semester Unassessed Courses, 2nd Semester Course Credits, 2nd Semester Enrolled Courses, 2nd Semester Assessed Courses, 2nd Semester Passed Courses, 2nd Semester Grade Point Average, 2nd Semester Unassessed Courses
Success Status Distribution	Students are divided into three categories: Graduation, Relative Success and Failure.

Machine Learning Algorithms Used

In this study, 9 different machine learning algorithms were applied to predict student achievement. The algorithms were selected based on their suitability for the classification problem, widespread use in the literature, and the possibility of comparative analysis. Machine Learning Algorithms Used:

Logistic Regression

Logistic regression is a statistical method widely used in classification problems (Hosmer, Lemeshow, & Sturdivant, 2000). The model estimates the classification probabilities of independent variables using a logit function. The parameters β shown in Equation 1 represent the model coefficients and X represents the independent variables.

$$P(y=1|x) = 1 / (1 + e^{-(\beta_0+\beta_1x_1+...+\beta_nx_n)}) \quad (1)$$

Decision Trees

Decision trees create classification rules by successively partitioning the dataset according to feature values (Quinlan, 1986). As seen in Equation 2, the best partitions are determined using Entropy or Gini index.

$$Gini(t) = 1 - \Sigma[p(i|t)]^2 \quad (2)$$

Random Forest

This ensemble method, proposed by Breiman (2001), is based on a combination of multiple decision trees. Each tree is trained on a bootstrap sample and the results are combined by voting.

Support Vector Machines (SVM)

SVM, developed by Vapnik (1995), is based on finding the optimal separating hyperplane. As seen in Equation 3, kernel functions are used to model nonlinear separations.

$$f(x) = \text{sign}(\Sigma\alpha_i y_i K(x_i,x) + b) \quad (3)$$

K-Nearest Neighbor (KNN)

This method, proposed by Cover and Hart (1967), is a heuristic approach that determines the classification based on the class labels of the nearest neighbor.

Naive Bayes

This method, based on Bayes' theorem, works on the assumption that independent variables are unconditionally independent (Murphy, 2012).

$$P(y|x_1,...,x_n) \propto P(y)\Pi P(x_i|y) \quad (4)$$

Gradient Boosting

This method, developed by Friedman (2001), is based on training weak learners sequentially in a way that reduces their margin of error.

XGBoost

XGBoost, developed by Chen and Guestrin (2016), is a gradient boosting algorithm augmented and optimized with regularization mechanisms.

Artificial Neural Networks (MLP)

The multilayer perceptron (MLP) is a basic form of artificial neural network (Bishop & Nasrabadi, 2006). It performs non-linear transformations between layers using activation functions.

In total, 9 different MLP algorithms were implemented and their performance was compared. The grid search method was used for hyperparameter tuning and the optimum parameter values were determined for each algorithm. Throughout the study, the generalizability performance of the models was tested with a five-fold cross-validation method.

Model Achievement Criteria

Various metrics are used to evaluate the performance of classification models. These metrics indicate how accurately the model predicts and how reliable it is. To visualize the performance of the model, the confusion matrix shown in Table 2 is used. This matrix compares the predictions of the model with the actual values.

Table 2 Confusion Matrix

Estimate / Actual	Positive (1)	Negative (0)
Positive (1)	True Positive (TP)	False Positive (FP)
Negative (0)	False Negative (FN)	True Negative (TN)

DP values that are predicted positive by the model and are positive in reality. FC values that the model predicts as positive but are actually negative. Values that the YN model predicts as negative but are actually positive. Values that the DN model predicts as negative and are actually negative.

Accuracy-Error Rate

Accuracy is the ratio of correct forecasts to total forecasts. Error rate is the ratio of wrong forecasts to total forecasts.

$$\text{Accuracy} = (TP + TN) / (TP + FP + FN + TN) \quad (5)$$

Error Rate = (FP + FN) / (TP + FP + FN + TN) (6)

Precision

Precision indicates how much of the model's predicted positive values are actually positive.

Precision = TP / (TP + FP) (7)

Recall/Sensitivity

Sensitivity shows how much of the values that are actually positive are predicted to be positive by the model.

Sensitivity = TP / (TP + FN) (8)

F-Measure

The F1-score is the harmonic mean of precision and sensitivity. It assesses both precision and sensitivity in a balanced way.

F1-Score =2*(Precision*Sensitivity) / (Precision + Sensitivity) (9)

ROC Curves (Receiver Operating Characteristic Curve)

The ROC curve shows the true positive rate (TP) of the model against the false positive rate (FP) at different thresholds. The area under the curve (AUC) shows the overall performance of the model. The closer the AUC value is to 1, the better the model is. The true positive rate (TP) is the same as the sensitivity. False positive rate is calculated by the formula FP/ (FP+ TN).

RESULTS

In this study, 9 different machine learning algorithms are used to predict whether undergraduate students will continue their education and their performances are compared. The dataset includes various characteristics such as demographic information, academic background and socio-economic status of the students. The target variable is the classification of students as “Graduated” (1) or “Not Graduated/Continuing” (0). The results shown in Table 3 show the performance of 9 different machine learning algorithms on the 5-fold cross-validation and test set:

Table 3 Performance Metrics of Machine Learning Algorithms

Model	Cross Validation Average Accuracy	Cross Validation Standard Deviation	Test Accuracy Rate
XGBoost	0.892	0.012	0.901
Gradient Boosting	0.885	0.011	0.896
Random Forest	0.879	0.013	0.889
LightGBM	0.874	0.014	0.883
Logistic Regression	0.862	0.015	0.871
Artificial Neural Network	0.857	0.016	0.866
Support Vector Machines	0.851	0.017	0.859
K-Nearest Neighbor	0.842	0.018	0.848
Decision Tree	0.831	0.019	0.839

Model	Cross Validation Average Accuracy	Cross Validation Standard Deviation	Test Accuracy Rate
Gauss Naive Bayes	0.813	0.021	0.824

The classification model of the XGBoost algorithm, which gives the most successful result among the machine learning algorithms applied to the dataset, is shown in Table 4. In the table, Precision shows how many of the samples marked as positive by the model are actually positive. Recall shows how many of the true positives are correctly predicted. F1-Score is the harmonic mean of precision and recall.

Table 4 XGBoost Algorithm Classification Model

	Precision	Sensitivity	F1-Score
0	0.91	0.89	0.90
1	0.89	0.91	0.90

According to these results, the model can predict both graduating and non-graduating students with high accuracy. The precision values of the model are around 0.90 for both classes, indicating that the false positive and false negative rates of the model are balanced.

Figure 2 shows the top 10 most important attributes and their importance ratios. According to these results, “Previous qualification grade” and “Admission score” are the two most important attributes indicating the student's previous academic performance. This finding shows that the student's previous academic performance plays a critical role in predicting university success. Other important characteristics include factors related to age, gender and socio-economic status.

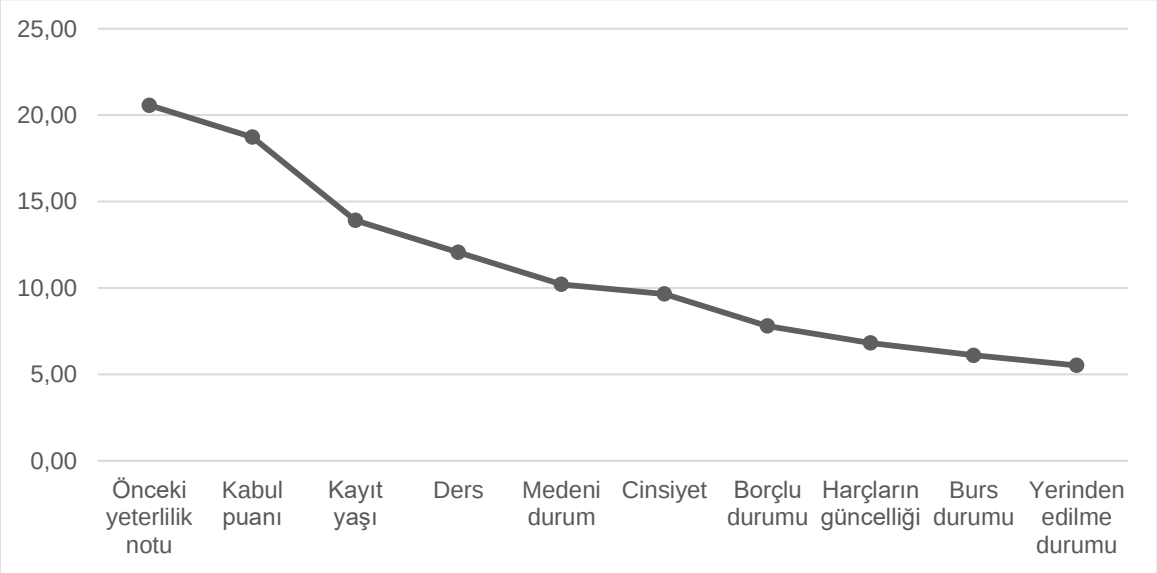


Figure 2 Top 10 Most Important Attributes and Their Importance

The better performance of ensemble methods such as XGBoost, Gradient Boosting and Random Forest compared to other algorithms can be explained by their ability to model complex data structures. In particular, it has been observed that XGBoost creates generalizable models by preventing overfitting thanks to its regularization techniques and optimization algorithms. This is explained by the fact that XGBoost increases the generalizability of the model thanks to its regularization techniques (L1, L2) and resistance to overlearning. Attribute importance was analyzed to ensure the explainability of the model. Logistic regression, a linear model, performed very well with a test accuracy of 87.1%, indicating that there are linear relationships between some features in the dataset and the target variable. The 86.6% test accuracy of the multilayer perceptron (MLP) model suggests that more complex models may not necessarily

perform better on this dataset. Gaussian Naive Bayes performed the poorest with a test accuracy of 82.4%, suggesting that this algorithm's assumption of conditional independence between features may not be fully valid for the dataset.

RESULTS and DISCUSSION

This comprehensive analysis reveals that ensemble methods, especially XGBoost, outperform other traditional machine learning algorithms in predicting student success. Academic background and entrance grades, which the model identifies as the most important features, stand out as critical factors in student success. The findings suggest that it may be beneficial for educational institutions to use XGBoost or similar ensemble methods when developing early warning systems and to develop intervention strategies that take into account students' previous academic performance. In the next phase of this study, we plan to further improve the performance of the models using hyperparameter optimization techniques and investigate alternative feature engineering methods. It is also proposed to analyze the decision-making processes of the models in more detail using model interpretability techniques.

KAYNAKÇA

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (C. 4). Springer. Date of arrival sender https://link.springer.com/book/9780387310732

Dünder, M., & Dünder, E. (2024). Kategorik verilerde sınıf dengesizliği durumunda makine öğrenimi algoritmalarının karşılaştırılması: Öğrencilerin başarı durumları üzerine bir uygulama. *Dijital Teknolojiler ve Eğitim Dergisi*, *3*(1), 28-38.

Ersozlu, Z., Taheri, S., & Koch, I. (2024). A review of machine learning methods used for educational data. *Education and Information Technologies*, *29*(16), 22125-22145. https://doi.org/10.1007/s10639-024-12704-0

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2000). *Applied logistic regression. Hoboken*. NJ: John Wiley & Sons.

Martins, M. V., Tolledo, D., Machado, J., Baptista, L. M. T., & Realinho, V. (2021). Early Prediction of student's Performance in Higher Education: A Case Study. Içinde Á. Rocha, H. Adeli, G. Dzemyda, F. Moreira, & A. M. Ramalho Correia (Ed.), *Trends and Applications in Information Systems and Technologies* (ss. 166-175). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-72657-7_16

Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (C. 1). McGraw-hill New York. Geliş tarihi gönderen http://www.pachecoj.com/courses/csc380_fall21/lectures/mlintro.pdf

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT press. Geliş tarihi gönderen https://books.google.com/books?hl=tr&lr=&id=RC43AgAAQBAJ&oi=fnd&pg=PR7&dq=Murphy,+K.+P.+(2012).+Machine+Learning:+A+Probabilistic+Perspective.+MIT+Press.&ots=unjxaFLqZc&sig=tbzANVBkhXH5wmcZ6yd3ht7ru2Y

Onyema, E. M., Almuzaini, K. K., Onu, F. U., Verma, D., Gregory, U. S., Puttaramaiah, M., & Afriyie, R. K. (2022). Prospects and Challenges of Using Machine Learning for Academic Forecasting. *Computational Intelligence and Neuroscience*, *2022*, 1-7. https://doi.org/10.1155/2022/5624475

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106. https://doi.org/10.1007/BF00116251

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (C. 1). MIT press Cambridge. Date of arrival sender https://www.cambridge.org/core/journals/robotica/article/robot-learning-edited-by-jonathan-h-connell-and-sridhar-mahadevan-kluwer-boston-19931997-xii240-pp-isbn-0792393651-hardback-21800-guilders-12000-8995/737FD21CA908246DF17779E9C20B6DF6

Tosunoğlu, E., Yılmaz, R., Özeren, E., & Sağlam, Z. (2021). Eğitimde makine öğrenmesi: Araştırmalardaki güncel eğilimler üzerine inceleme. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, *3*(2), 178-199.

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167.