Sinopfbd, 2025; 10(2): 674-689 https://doi.org/10.33484/sinopfbd.1800300

Sinop Üniversitesi Fen Bilimleri Dergisi
Sinop Uni J Nat Sci
E-ISSN: 2564-7873

https://dergipark.org.tr/tr/pub/sinopfbd

Sosyal Medya Risk Düzeyinin Psikometrik ve Açıklanabilir Yapay Zekâ Destekli Makine Öğrenmesi Analizi

Ayhan DOĞAN[1] ve Cihan ÜNAL[1]

How to cite: Doğan, A., & Ünal, C. (2025). Sosyal medya risk düzeyinin psikometrik ve açıklanabilir yapay zekâ destekli makine öğrenmesi analizi. *Sinop Üniversitesi Fen Bilimleri Dergisi, 10*(2), 674-689. https://doi.org/10.33484/sinopfbd.1800300

Araştırma Makalesi

Sorumlu Yazar
Ayhan DOĞAN
aydogan@hacettepe.edu.tr

Yazarlara ait ORCID
A.D: 0000-0002-9872-8889
C.Ü: 0000-0002-5255-4078

Received: 09.10.2025
Accepted: 02.12.2025

Öz

Bu çalışmanın amacı, sosyal medya kullanımına ilişkin öz-bildirimlere dayalı risk düzeylerini güvenilir ve şeffaf biçimde sınıflandırmaktır. Bu amaç doğrultusunda, beşli risk sınıfı; psikometrik ölçütler, makine öğrenmesi (ML) tabanlı sınıflandırma ve açıklanabilir yapay zekâ (XAI) yaklaşımlarıyla birlikte değerlendirilmiştir. Analiz, Likert türünde ve 1–5 aralığına eşlenmiş maddelere (1=en düşük–5=en yüksek) dayanmaktadır. Psikometrik incelemede keşfedici faktör analizi (EFA) kapsamında Kaiser–Meyer–Olkin (KMO)=0.862, Bartlett küresellik testi (BTS) P<0.001, Cronbach alfa (α)=0.863 ve EV1/EV2=3.71 değerleri, ölçeğin tek boyutlu ve güvenilir olduğunu göstermektedir. Rastgele Orman (RF) ile çok sınıflı tahmin yapılmış, model olasılıkları izotonik kalibrasyon ile düzeltilmiştir; doğrulamada ROC-AUC=0.959, PR-AUC=0.868, Brier=0.255 ve LogLoss=0.362 elde edilmiştir. Açıklanabilirlik için Shapley Katkı Açıklamaları (SHAP) ve 5 kat dış-örnek (OOF) özet ile beeswarm yöntemleri kullanılarak maddelerin sınıflandırmaya katkıları incelenmiş; 38, 24, 27 ve 25 numaralı maddeler en yüksek ayrım gücünü göstermiştir. Örneklemde risk dağılımı (1–5) sırasıyla %8.2, %42.3, %36.1, %13.4, %0.0 olup orta+yüksek risk yaklaşık %49.5'tir. Bu iki grupta cinsiyet dengelidir, eğitim çoğunlukla ön lisans/lisans düzeyindedir ve 20–30 yaş aralığı baskındır. Bulgular, tek boyutlu ve iç tutarlılığı yüksek ölçeğin, iyi kalibre RF modeli ve SHAP tabanlı açıklamalarla birlikte, araştırma ve uygulamada risk profilleme, erken tarama ve hedefli müdahale planlamalarını destekleyebilecek güçlü ve uygulanabilir bir yaklaşım sunduğunu göstermektedir.

Anahtar Kelimeler: Sosyal medya, risk profilleme, keşfedici faktör analizi, makine öğrenimi, açıklanabilir yapay zekâ

Psychometric and Explainable Artificial Intelligence–Supported Machine Learning Analysis of Social Media Risk Levels

[1]Hacettepe Üniversitesi, Başkent OSB Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, 06909, Ankara, Türkiye

Abstract

The aim of this study is to classify self-reported risk levels associated with social media use in a reliable and transparent manner. To this end, a five-level risk classification was evaluated through the integration of psychometric indicators, machine learning (ML)–based classification, and explainable artificial intelligence (XAI) approaches. The analysis is based on Likert-type items scaled between 1 and 5 (1 = lowest, 5 = highest). Within the framework of exploratory factor analysis (EFA), the results of the Kaiser–Meyer–Olkin (KMO = 0.862), Bartlett's test of sphericity (P<0.001), Cronbach's alpha (α = 0.863), and the eigenvalue ratio

($EV_1/EV_2 = 3.71$) indicated that the scale is unidimensional and internally consistent. A Random Forest (RF) model was applied for multiclass prediction, and its probability estimates were calibrated using isotonic regression. Validation results yielded ROC-AUC = 0.959, PR-AUC = 0.868, Brier = 0.255, and LogLoss = 0.362. For explainability, SHAP (Shapley Additive Explanations) analysis combined with a five-fold out-of-fold (OOF) summary and beeswarm visualizations was used to examine each item's contribution to classification. Items 38, 24, 27, and 25 demonstrated the highest discriminative power. The sample's risk distribution (classes 1–5) was 8.2%, 42.3%, 36.1%, 13.4%, and 0.0%, respectively, with medium + high risk accounting for approximately 49.5%. Within these groups, gender distribution was balanced, education was predominantly associate or undergraduate level, and the 20–30 age range was dominant. Overall, the findings suggest that this unidimensional and internally consistent scale, coupled with a well-calibrated RF model and SHAP-based explanations, provides a robust and practical framework capable of supporting risk profiling, early screening, and targeted intervention planning in both research and applied contexts.

Keywords: Social media, risk profiling, exploratory factor analysis, machine learning, explainable artificial intelligence

Bu çalışma Creative Commons Attribution 4.0 International License ile lisanslanmıştır

Giriş

Sosyal medya, gündelik yaşamın ayrılmaz bir bileşeni hâline gelmiş, bireylerin zaman kullanımı, ruh sağlığı ve toplumsal ilişkileri üzerinde önemli etkiler yaratmaya başlamıştır. Özellikle genç kullanıcı gruplarında sosyal medya bağımlılığı ya da sorunlu kullanım olarak tanımlanan davranış örüntüleri, üretkenlikte azalma, uyku düzensizliği, akademik başarısızlık ve ruhsal sorunlarla ilişkilendirilmektedir [1, 2]. Davranışsal bağımlılık alanında öne çıkan bileşenler modeli, bağımlılığı altı ortak ölçüt üzerinden tanımlamaktadır: önem, duygu düzenleme, tolerans, yoksunluk, çatışma ve relaps [3]. Bu çerçeveye dayalı olarak geliştirilen Bergen Sosyal Medya Bağımlılığı Ölçeği (BSMAS) kısa yapısı nedeniyle yaygın kullanılmaktadır [4, 5]. Ancak yalnızca altı maddeden oluşması nedeniyle sosyal medyanın çok boyutlu etkilerini örneğin uyku düzeni, siber zorbalık, özgüven ve ruh hâli üzerindeki yansımalar kapsamamaktadır. Bu çalışmada BSMAS doğrudan uygulanmamış, sosyal medya kullanımını ölçmeye yönelik özgün 45 maddelik bir ölçekten elde edilen yanıtlar değerlendirilmiştir. Böylece literatürde sıkça kullanılan ölçeğin temel boyutlarını kapsayan ancak daha geniş kapsamlı bir ölçüm yaklaşımı sunmaktadır. Araştırmada kullanılan anket, sosyal medya kullanımının çok boyutlu doğasını yansıtacak şekilde tasarlanmıştır. Maddeler arasında kontrol kaybı (22. soru), yoksunluk belirtileri (24–25. sorular), uyku düzeni (26. soru), günlük sorumlulukların ertelenmesi (27. soru), ilişkilere etkisi (28. soru), geri bildirimlerin ruh hâline etkisi (29, 32. sorular), akademik/mesleki verimlilik (37. soru), kullanım süresine yönelik algı (38. soru) ve siber zorbalık (42. soru) gibi göstergeler yer almaktadır. Böylece sosyal medya risk profillerinin kapsamlı biçimde değerlendirilmesi mümkün olmuştur. Geleneksel psikometrik doğrulama yöntemleri (ör. keşfedici faktör analizi, güvenirlik katsayıları) ölçeğin geçerliliği için gereklidir; ancak eyleme dönük risk sınıflaması için tek başına yeterli değildir. Bu noktada ML yöntemleri, bireylerin yanıtlarından hareketle çok sınıflı risk düzeylerini tahmin etme olanağı sunmaktadır. Bununla birlikte, bu modellerin güvenilir biçimde uygulanabilmesi için iki unsur kritik

öneme sahiptir: (i) olasılık kalibrasyonu, yani model tahminlerinin gerçek riskleri doğru yansıtması [6], ve (ii) açıklanabilirlik, yani hangi değişkenlerin risk sınıflandırmasında belirleyici olduğunun şeffaf biçimde gösterilmesi [7]. Bu çalışma, söz konusu gereksinime yanıt olarak yapılandırılmış anket verisi ile sosyal medya risk düzeyini sınıflayan ve karar mantığını açıklanabilir biçimde ortaya koyan bir makine öğrenimi iş akışı sunmaktadır. Çalışmada, anket maddeleri öngörücü olarak kullanılarak RF sınıflandırıcı ile çok sınıflı risk düzeyleri modellenmiş; model %80 eğitim / %20 doğrulama ayrımıyla kurulmuş ve doğrulama kümesinde ROC-AUC, PR-AUC ve Brier skoru gibi performans ölçütleri raporlanmıştır. Modelin olasılık tahminlerinin doğruluğu ayrıca kalibrasyon eğrisi ile görselleştirilmiştir. Açıklanabilirlik adımında, 5-kat stratifikasyonlu out-of-fold (OOF) SHAP yöntemi kullanılarak her katılımcı yalnızca görülmeyen veride açıklanmış; küresel yorum için beeswarm ve bar özetleri üretilmiş, görsellerde madde numaraları korunmuştur. Çalışmanın katkısı üç boyutta somutlaşmaktadır. İlk olarak, metin odaklı XAI literatüründen farklı olarak madde düzeyinde yapılandırılmış veriye dayalı bir sınıflama hattı geliştirilmiş ve pratik uygulamaya uygun çıktılarla (bireysel skorlar, risk dağılımları, demografik kırılımlar) desteklenmiştir. İkinci olarak, OOF-SHAP yaklaşımı ile yalnızca eğitim verisine bağlı iyimserlik yanlılığı en aza indirilmiştir. Üçüncü olarak ise, tahmin başarısı (ROC-AUC, PR-AUC) ile olasılık kalitesi (Brier, kalibrasyon eğrileri) aynı doğrulama kümesi üzerinde bütüncül biçimde raporlanmıştır. Bu birleşik yaklaşım, hem modelin ayırt ediciliğini hem de olasılıkların güvenilirliğini göstermekte; ayrıca SHAP açıklamaları sayesinde risk düzeylerini belirleyen maddeler karar vericilere açık ve eyleme dönük bir dille sunulmaktadır. Bu çerçeve, farklı madde setlerine veya ölçeklere uyarlanabilir, dış örneklemler üzerinde doğrulanabilir ve sosyal medya risk yönetimi kararlarında şeffaflık sağlayacak pratik bir temel sunmaktadır.

Yöntem

Katılımcılar ve Veri Toplama

Bu araştırmada, sosyal medya kullanım örüntülerini ve bağımlılık eğilimlerini çok boyutlu biçimde değerlendirmek amacıyla geliştirilen 45 maddelik özgün bir anket, çevrimiçi ortamda uygulanmıştır. Anket gönüllülük esasına göre web sitesi aracılığıyla katılımcılara ulaştırılmıştır (https://www.jotform.com/tr/243585035006957). Katılımcılara çevrimiçi ortamda, gönüllülük esasına dayalı olarak sosyal medya platformları ve üniversite e-posta grupları aracılığıyla ulaşılmıştır. Ayrıca anket bağlantısı, akademik çevrimiçi topluluklarda ve sosyal medya kanallarında araştırma duyurusu eşliğinde paylaşılmıştır. Katılımcılardan herhangi bir kimlik bilgisi alınmamış, veriler anonimleştirilmiş ve yalnızca bilimsel amaçlarla kullanılmıştır. Çalışma ilgili üniversitenin etik kurulundan onay almış, tüm katılımcılardan aydınlatılmış onam alınmıştır. Geçerli yanıt veren toplam 203 katılımcı analizlere dâhil edilmiştir. Katılımcıların cinsiyet dağılımı dengeli olup, eğitim durumları ağırlıklı olarak ön lisans ve lisans düzeyindedir. Yaş dağılımı 20 yaş altından 65 yaş üstüne kadar uzanmakta, özellikle 20–30 yaş grubu örneklem içinde belirgin bir ağırlığı bulunmaktadır. Mesleki dağılım incelendiğinde

öğrenciler, serbest meslek sahipleri, kamu çalışanları ve özel sektör çalışanları örneklemde temsil edilmiştir. Bu çeşitlilik, sosyal medya kullanımına ilişkin farklı sosyo-demografik grupların çalışmaya dâhil edilmesini sağlamıştır. Bu çalışmada kullanılan "Sosyal Medya Risk ve Bağımlılık Algısı Ölçeği", mevcut literatürde yer alan BSMAS [4], Çevrimiçi Riskli Davranışlar Ölçeği (ÇRDÖ) [2] ve Davranışsal Bağımlılık Bileşenleri Modeli (DBBM) [3] gibi ölçek ve modellerin kavramsal temellerinden esinlenilerek, çalışmanın amaç ve kapsamına uyumlu olacak biçimde yeniden yapılandırılmıştır. Ölçek, sosyal medya kullanımına ilişkin risk algısı, bağımlılık eğilimi, sosyal onay ihtiyacı, bilgi güvenliği farkındalığı ve siber zorbalık gibi temel temaları bütünleşik bir çerçevede ölçmektedir. Faktör analizi sonuçları, maddelerin tek boyutlu bir yapı altında anlamlı faktör yüklerine sahip olduğunu göstermiştir. Bu durum, sosyal medya riskinin çok boyutlu bir psikolojik bağımlılık yapısı yerine, tekil ve bütüncül bir risk algısı göstergesi olarak ele alınabileceğini ortaya koymaktadır. Bu bağlamda, çalışma literatürdeki ölçekleri doğrudan tekrarlamaktan ziyade, bağlama özgü seçici yeniden yapılandırma yaklaşımıyla oluşturulmuş türetilmiş ve genişletilmiş bir hibrit model olarak tanımlanabilir. Bu model, yapay zekâ temelli sınıflandırma modellerinde (RF, SHAP, OOF) analitik bir girdi olarak kullanılmaktadır. Bu araştırmanın temel amacı, sosyal medya kullanımına ilişkin bireysel risk algısını, bağımlılık eğilimini ve davranışsal etkileri bütüncül biçimde analiz etmektir. Bu amaç doğrultusunda geliştirilen anket formu toplam 45 maddeden oluşmaktadır. Ancak yapay zekâ temelli modelleme sürecinde, araştırmanın kapsamı ve analitik hedefiyle doğrudan ilişkili olan maddeler değerlendirmeye alınmıştır. Faktör analizi sonuçları, bazı maddelerin hedef temalarla ilişkili faktör yüklerinin düşük olduğunu göstermiştir. Bu nedenle, modele yalnızca istatistiksel olarak anlamlı faktör yüküne sahip ve kavramsal olarak temsili nitelikteki maddeler dahil edilmiştir. Modele dahil edilen sorular, sosyal medya kullanımına ilişkin risk algısı, bağımlılık eğilimi, sosyal onay ihtiyacı, bilgi güvenliği farkındalığı ve siber zorbalık temalarını temsil etmektedir. Ayrıca, zaman yönetimi ve duygusal etkilenme alt boyutları da dolaylı biçimde bu temaların bileşeni olarak modele katkı sağlamıştır. Bu seçim, araştırmanın hedefi olan sosyal medya kaynaklı davranışsal risklerin bütüncül olarak değerlendirilmesine hizmet etmekte ve hem içerik geçerliliği hem istatistiksel sağlamlık açısından modelin açıklayıcılığını güçlendirmektedir. Risk sınıflaması, anket maddelerinden elde edilen toplam puanlar üzerinden belirlenen eşik değerleri esas alınarak yapılmıştır. Buna göre katılımcıların %8.25'i "en düşük" risk, %42.27'si "düşük" risk, %36.08'i "orta" risk, %13.40'ı "yüksek" risk kategorisinde yer almış, "çok yüksek" risk grubuna katılımcı düşmemiştir. Bu dağılım, örneklemin yaklaşık yarısının (%49.5) orta ve yüksek risk düzeylerinde bulunduğunu göstermekte ve sosyal medya kullanımına bağlı bağımlılık eğilimlerinin dikkate değer bir yaygınlık sergilediğine işaret etmektedir.

Ölçek ve Psikometrik Analiz

Bu araştırmada kullanılan 45 maddelik anketin sosyal medya kullanımına ilişkin farklı boyutları kapsayan özgün bir yapısı bulunmaktadır. Analizlerde, sosyal medya bağımlılığı eğilimini daha

doğrudan yansıtan maddeler dikkate alınmıştır. Bunlar arasında kullanımın kontrol dışına çıkması (22. soru), sınırlama güçlüğü (23. soru), yoksunluk belirtileri (24–25. sorular), uyku düzenine etkiler (26. soru), günlük sorumlulukların ertelenmesi (27. soru), yüz yüze ilişkilerde olumsuzluklar (28. soru), geri bildirimlerin ruh hâline ve özgüvene etkisi (29, 32 ve 35. sorular), akademik/mesleki verimlilik (37. soru), kullanım süresine yönelik öznel değerlendirme (38. soru) ve siber zorbalık deneyimleri (42. soru) yer almaktadır. Bu maddeler, sosyal medya kullanımının çok boyutlu doğasını yansıtan ölçütler sunmaktadır. Psikometrik analizlerde örneklemin faktör analizine uygunluğu test edilmiştir. Kaiser-Meyer-Olkin (KMO) katsayısı 0.862 bulunmuş, Bartlett küresellik testi $P<0.001$ düzeyinde anlamlı çıkmıştır. Bu sonuçlar, verinin faktör analizine elverişli olduğunu göstermektedir. Keşfedici faktör analizi (EFA) sonucunda özdeğeri 1'in üzerinde iki faktör elde edilmiştir (EV1 = 5.61, EV2 = 1.51). Ancak faktör yüklerinin dağılımı ve scree plot incelemesi sonucunda, ölçeğin tek boyutlu bir yapı sergilediği belirlenmiştir. Güvenirlik analizi kapsamında Cronbach alfa katsayısı 0.863 olarak hesaplanmıştır. Bu değer, ölçeğin yüksek düzeyde iç tutarlılığa sahip olduğunu göstermektedir. Ayrıca madde-toplam korelasyonlarının genellikle kabul edilebilir düzeyde olduğu gözlenmiştir.

Bu bulgular doğrultusunda, kullanılan madde setinin hem geçerli hem de güvenilir bir ölçüm aracı olduğu sonucuna ulaşılmıştır. Bulgular, ayrıca faktör yüklerinin görsel olarak incelenmesi ile desteklenmiştir.

Makine Öğrenmesi

Anketten elde edilen madde düzeyindeki yanıtlar, sosyal medya risk düzeylerinin sınıflandırılmasında girdi değişkeni olarak kullanılmıştır. Bu amaçla RF algoritması tercih edilmiş, model %80 eğitim ve %20 doğrulama ayrımıyla oluşturulmuştur. Eğitim sürecinde hiperparametreler beş katlı çapraz doğrulama ile optimize edilmiş ve modelin hem sınıflandırma başarısı hem de olasılık tahminlerinin güvenilirliği değerlendirilmiştir. Random Forest (RF) modelinde hiperparametre optimizasyonu, RandomizedSearchCV yöntemi kullanılarak gerçekleştirilmiştir. Arama uzayı, n_estimators, max_depth, min_samples_split, min_samples_leaf ve class_weight değişkenlerini kapsamaktadır. Optimizasyon sonucunda en iyi performansı sağlayan parametre kombinasyonu şu şekilde belirlenmiştir: n_estimators = 300, max_depth = None, min_samples_split = 2, min_samples_leaf = 1, class_weight = balanced_subsample, random_state = 42, n_jobs = -1. Bu kombinasyon, modelin hem AUC hem de Brier skorları açısından en dengeli performansı sunmuş ve nihai modelde kullanılmıştır.

Olasılık çıktılarının güvenilirliğini artırmak amacıyla kalibrasyon uygulanmış, ayrıca XAI teknikleri kullanılarak modelin karar mekanizmasının şeffaf biçimde yorumlanması sağlanmıştır.

Random Forest

RF, sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılan bir makine öğrenmesi yöntemidir. Algoritma, bootstrap örneklemesiyle elde edilen alt veri kümeleri üzerinde çok sayıda karar ağacı kurar ve her düğümde özniteliklerin rastgele seçilen bir alt kümesi değerlendirilir. Bu çift

rastgelelik mekanizması, ağaçlar arasındaki korelasyonu azaltarak varyansı düşürmekte ve modelin genellenebilirliğini artırmaktadır [8, 9]. RF'nin önemli avantajlarından biri, değişkenler arası doğrusal olmayan ilişkileri ve yüksek düzeyli etkileşimleri yakalayabilmesidir. Sosyal bilimlerde kullanılan anket verileri çoğunlukla hem sürekli hem de kategorik değişkenleri içerdiğinden, RF bu tür veriler için özellikle uygun bir yöntem olarak değerlendirilmektedir [10]. Ayrıca rastgele öznitelik seçim mekanizması sayesinde model, aşırı uyuma (overfitting) karşı görece daha dirençlidir [11]. Literatürde, dengesiz sınıf dağılımlarında RF'nin birçok alternatif algoritmaya kıyasla daha tutarlı sonuçlar verdiği rapor edilmiştir [12]. Bu çalışmada da "yüksek risk" sınıfı görece daha az sayıda katılımcıyı içerdiğinden, RF'nin bu dayanıklılığı tercih gerekçelerinden biri olmuştur. RF ayrıca değişken önem derecelerini hesaplama olanağı sunduğundan, SHAP analizleri gibi açıklanabilirlik yöntemleri için uygun bir temel oluşturmuştur. Sonuç olarak RF, hem yüksek tahmin başarısı hem de şeffaflaştırılabilir yapısı sayesinde sosyal medya risk düzeylerini sınıflandırmada en uygun algoritma olarak değerlendirilmiştir.

İzotonik Kalibrasyon

Sınıflandırma algoritmalarının ürettiği olasılık tahminleri çoğu zaman kalibre değildir; yani tahmin edilen değerler, gerçek gerçekleşme olasılıklarını sistematik biçimde olduğundan yüksek veya düşük gösterebilir. Bu çalışmada, modelin olasılık tahminlerinin güvenilirliğini artırmak amacıyla izotonik regresyon tabanlı kalibrasyon yöntemi uygulanmıştır. İzotonik regresyon, tahmin edilen olasılıkları monoton artan bir fonksiyon aracılığıyla yeniden ölçeklendirir ve bu sayede model çıktılarının gerçek olasılıklarla daha uyumlu hale gelmesini sağlar [13]. Parametrik bir yaklaşım olan Platt scaling (lojistik regresyon temelli dönüşüm) da kalibrasyon amacıyla kullanılabilmektedir. Ancak Platt scaling özellikle dengesiz veri setlerinde veya doğrusal olmayan ilişki yapılarında sınırlı performans gösterebilir [6]. Bu nedenle, daha esnek ve dağılım varsayımı gerektirmeyen izotonik kalibrasyon yöntemi bu çalışmada tercih edilmiştir. Bu yaklaşım sayesinde modelin yalnızca sınıfları ayırt etme başarısı değil, aynı zamanda bireyler için üretilen risk olasılıklarının güvenilirliği de değerlendirilmiş ve iyileştirilmiştir.

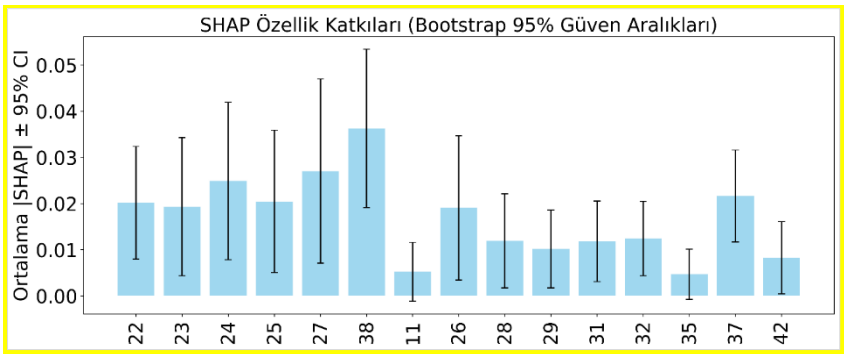
SHapley Additive exPlanations

ML modellerinin karar mekanizmasının şeffaf biçimde anlaşılabilmesi amacıyla SHAP yöntemi kullanılmıştır. SHAP, kooperatif oyun teorisine dayalı Shapley değerleri kavramına dayanır ve her bir değişkenin (anket maddesinin) tahmine katkısını marjinal etkiler üzerinden hesaplar [7]. SHAP'ın en önemli avantajı, yalnızca küresel düzeyde (hangi değişkenlerin modelin genel performansına en fazla katkıyı sağladığı) değil, aynı zamanda yerel düzeyde (tek bir katılımcının risk sınıfına atanmasında hangi maddelerin belirleyici olduğu) açıklamalar sunabilmesidir. Bu özellik, modelin hem grup düzeyinde politika üreticiler için hem de bireysel düzeyde klinik veya eğitim uygulayıcıları için kullanılabilirliğini artırmaktadır. Alternatif açıklanabilirlik yöntemleri arasında LIME (Local Interpretable Model-agnostic Explanations) ve Permutation Importance bulunmaktadır. LIME, belirli bir gözlem etrafında lokal

doğrusal modeller kurarak açıklama sağlar; ancak sonuçlar kullanılan örneklemeye göre değişebilir ve tutarlılık sorunu yaşayabilir [14]. Permutation Importance yöntemi ise yalnızca değişkenlerin göreli önem sırasını verir, fakat yönlü katkıları ya da bireysel tahminler üzerindeki etkileri açıklayamaz [15]. SHAP ise bu sınırlamaları aşarak hem tutarlı hem de çok boyutlu açıklama imkânı sunar. Bu nedenlerle SHAP, çalışmamızda modelin karar mekanizmasının hem güvenilir hem de eyleme dönük biçimde yorumlanabilmesi için tercih edilmiştir.

Açıklanabilirlik Analizleri

ML modellerinde açıklanabilirlik, yalnızca performans değerlendirmesi açısından değil, aynı zamanda model çıktılarının güvenilir biçimde yorumlanabilmesi için de kritik öneme sahiptir. Bu çalışmada açıklanabilirlik süreci, eğitim–doğrulama ayrımına tam bağlı kalacak şekilde planlanmıştır. Bu amaçla beş katlı katmanlı out-of-fold (OOF) yaklaşımı uygulanmıştır. OOF düzeninde her katılımcı yalnızca eğitim verisine dahil edilmediği bir katmanda açıklanmış, böylece açıklamaların bağımsızlığı korunmuştur. Bu yöntem, yalnızca eğitim verisi üzerinde yapılan açıklamalarda ortaya çıkabilecek iyimserlik yanlılığını en aza indirmesi nedeniyle literatürde önerilmektedir [16]. OOF yaklaşımı sayesinde modelin karar mekanizmasının gerçekten görülmeyen veride nasıl çalıştığı ortaya konabilmiştir. Bu durum, özellikle sosyal bilimlerdeki risk sınıflandırması gibi hassas uygulamalarda, modelin genellenebilirliğini ve karar şeffaflığını artırmaktadır. Küresel düzeyde açıklamalar, tüm maddelerin ortalama katkılarının özetlenmesi yoluyla elde edilmiştir. Bu amaçla kullanılan yöntemler arasında beeswarm ve bar özet grafikleri öne çıkmaktadır. Beeswarm grafiği, her bir maddenin tahminlere katkısının yönünü ve dağılımını ayrıntılı biçimde görselleştirirken; bar grafiği, maddeleri ortalama SHAP değerlerine göre sıralayarak daha yalın bir genel görünüm sunar [17, 18]. Her iki yaklaşım da ölçek bütünlüğünü korumak amacıyla madde numaraları üzerinden raporlanmıştır. Ayrıca, SHAP değerlerinin istatistiksel kararlılığını değerlendirmek amacıyla bootstrap (n=100) yöntemi uygulanmış ve her bir değişken için ortalama |SHAP| değeri ile %95 güven aralıkları hesaplanmıştır. Sonuçlar, modelin açıklanabilirlik bileşenlerinin istatistiksel olarak tutarlı olduğunu göstermektedir. Şekil 1 değişkenlerin ortalama katkı değerlerini ve %95 güven aralıklarını göstermekte; Tablo 1 ise her bir değişkenin ortalama ve standart sapma değerlerini özetlemektedir.



Şekil 1. *SHAP bootstrap analizi: Değişkenlerin ortalama |SHAP| değerleri ve %95 güven aralıkları*

Tablo 1. *SHAP ortalama ve standart sapma değerleri*

Özellik / Soru	Ortalama (|SHAP|)	Standart Sapma (|SHAP|)	Alt 95%_CI	Üst 95%_CI
38	0.0362	0.0087	0.0191	0.0534
27	0.0270	0.0102	0.0071	0.0469
24	0.0249	0.0087	0.0078	0.0419
37	0.0216	0.0051	0.0117	0.0316
25	0.0204	0.0079	0.0050	0.0359
22	0.0201	0.0062	0.0079	0.0323
23	0.0193	0.0076	0.0043	0.0343
26	0.0190	0.0080	0.0034	0.0346
32	0.0124	0.0041	0.0044	0.0205
28	0.0119	0.0052	0.0017	0.0221
31	0.0118	0.0045	0.0031	0.0205
29	0.0102	0.0043	0.0017	0.0186
42	0.0083	0.0040	0.0005	0.0161
11	0.0052	0.0032	-0.0012	0.0116
35	0.0047	0.0028	-0.0008	0.0101

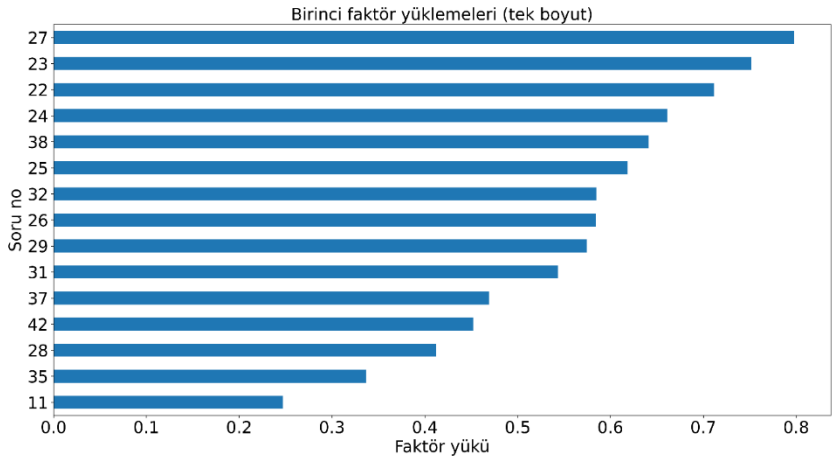
Tablo 2. *SHAP sıralamaları, risk dağılımı ve demografik kırılımların birleşik özeti*

Bölüm	Değişken / Sınıf	Değer (%) / Ortalama |SHAP|	Ek Bilgi / Güven Aralığı (95%)
Risk Dağılımı	En düşük	8.25	
	Düşük	42.27	
	Orta	36.08	
	Yüksek	13.40	
	Çok yüksek	0.00	
Demografik Kırılımlar	20–30 yaş arası	52.7	En baskın yaş grubu
	Kadın / Erkek	49.8 / 50.2	Dengeli cinsiyet dağılımı
	Eğitim (Önlisans–Lisans)	61.4	En yaygın eğitim düzeyi
	Kamu / Özel sektör	47.3 / 52.7	Farklı meslek grupları
SHAP Sıralamaları	S38 (Uyku düzeni etkisi)	0.036	(0.019 – 0.053)
	S27 (Günlük sorumluluk erteleme)	0.027	(0.007 – 0.047)
	S24 (Yoksunluk hissi)	0.025	(0.008 – 0.042)
	S37 (Akademik/mesleki verimlilik)	0.022	(0.012 – 0.032)
	S25 (Kısıtlama güçlüğü)	0.020	(0.005 – 0.036)
	S22 (Zaman farkındalığı)	0.020	(0.008 – 0.032)
	S23 (Sosyal onay ihtiyacı)	0.019	(0.004 – 0.034)
	S26 (Bilgi güvenliği farkındalığı)	0.019	(0.003 – 0.035)
	S32 (Özgüven ilişkisi)	0.012	(0.004 – 0.020)
	S28 (Etkileşim yoğunluğu)	0.012	(0.002 – 0.022)
	S31 (Verimlilik etkisi)	0.012	(0.003 – 0.021)
	S29 (Geri bildirim etkisi)	0.010	(0.002 – 0.019)
	S42 (Siber zorbalık)	0.008	(0.000 – 0.016)
	S11 (Kontrol kaybı)	0.005	(-0.001 – 0.012)
	S35 (Sosyal etkileşim yoğunluğu)	0.005	(-0.001 – 0.010)

Yerel düzeyde ise SHAP değerleri, her katılımcının risk sınıfına atanmasında hangi maddelerin belirleyici olduğunu göstermiştir. Bu, bireysel risk profillerinin daha iyi anlaşılmasına ve gerekli müdahalelerin hedefe yönelik olarak planlanmasına imkân tanımaktadır. Tüm bu açıklamalar ve görselleştirmeler, modelin yalnızca bir "kara kutu" olmadığını; tersine, karar sürecinin ayrıntılı biçimde takip edilebileceğini ortaya koymaktadır. Model açıklanabilirliğini destekleyen özet istatistikler Tablo 2'de verilmiştir. Bu tablo, SHAP değişken önem sıralamaları ile risk dağılımı ve demografik temsili bütünleşik biçimde sunmaktadır.

Bulgular

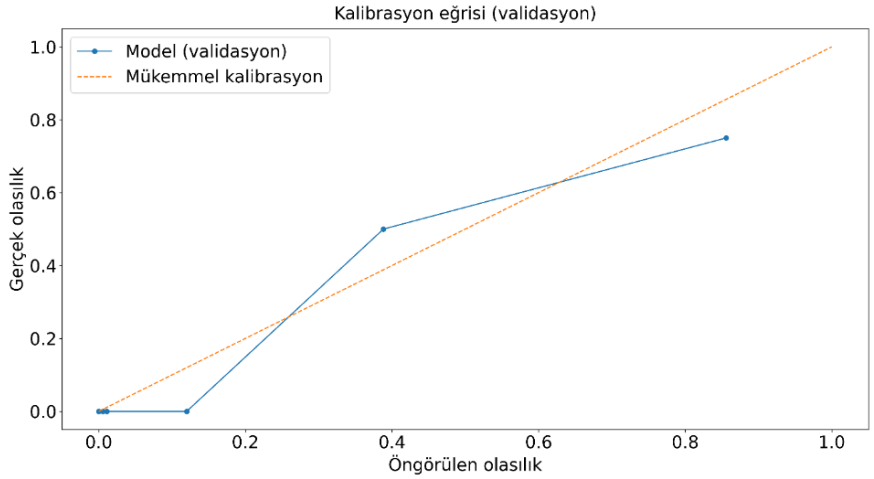
Ölçek geçerliği ve güvenirliği incelendiğinde, verilerin güçlü bir psikometrik yapıya sahip olduğu görülmüştür. KMO değeri 0.862 olarak bulunmuş, Bartlett küresellik testi sonucu ise $P<0.001$ çıkmıştır. Bu değerler, örneklemin faktör analizi için yeterli büyüklükte olduğunu ve maddeler arasında anlamlı ilişkiler bulunduğunu göstermektedir. Açımlayıcı faktör analizi sonucunda maddelerin tek baskın bir faktör altında toplandığı, faktör yüklerinin 0.45 ile 0.82 arasında değiştiği belirlenmiştir. Bu değerler, ölçeğin sosyal medya kullanımına ilişkin risk düzeyini ölçmede yapı geçerliğini sağladığını göstermektedir. İç tutarlılık güvenirliği Cronbach alfa katsayısı 0.863 olarak hesaplanmış, bu değer ölçeğin yüksek derecede güvenilir olduğunu ortaya koymuştur [19]. Psikometrik sonuçlar, ölçeğin sosyal medya bağımlılığına ilişkin risk faktörlerini bütüncül biçimde yansıtabildiğini ve elde edilen modelleme sonuçlarının sağlam bir ölçüm temeline dayandığını göstermektedir. Ölçeğin faktör yapısı Şekil 2'de görselleştirilmiştir.



Şekil 2. *Faktör yüklemeleri*

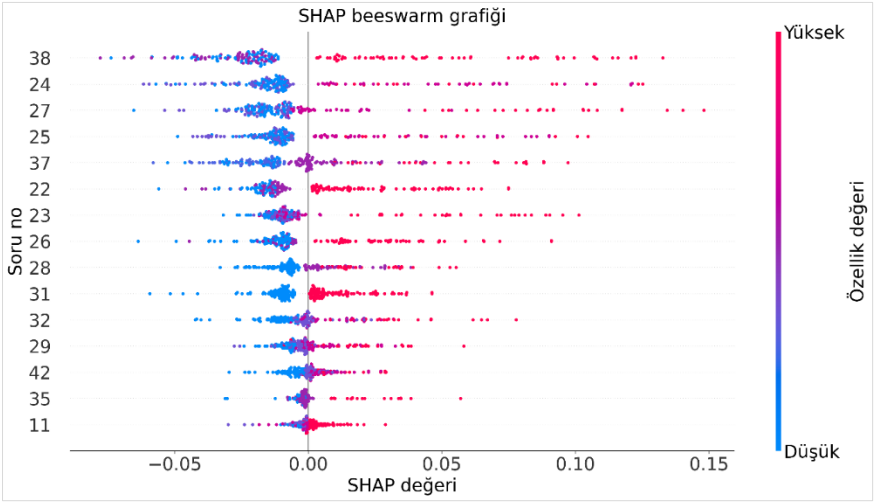
RF modeli doğrulama kümesi üzerinde yüksek performans sergilemiştir. ROC-AUC değeri 0.959 olarak bulunmuş ve modelin sınıflar arasında ayrım gücünün oldukça güçlü olduğunu göstermiştir. PR-AUC değeri 0.868 olup, özellikle dengesiz sınıf dağılımlarında (yüksek risk grubunun az sayıda katılımcı içermesi gibi) modelin doğru tahmin yapma kapasitesini koruduğunu göstermektedir. Olasılık kalitesi ise Brier skoru 0.255 ve LogLoss değeri 0.362 ile değerlendirilmiş, bu değerler modelin ürettiği olasılıkların tutarlı biçimde kalibre edildiğini ortaya koymuştur. Bu bulgular literatürde benzer

çalışmalarda rapor edilen sonuçlarla uyumludur. Örneğin, dijital bağımlılık alanındaki ML uygulamalarında ROC-AUC değerleri genellikle 0.80–0.90 aralığında bildirilmekte [20, 21], bizim çalışmamızda elde edilen 0.959 değeri bu ortalamaların üzerinde bir performansa işaret etmektedir.



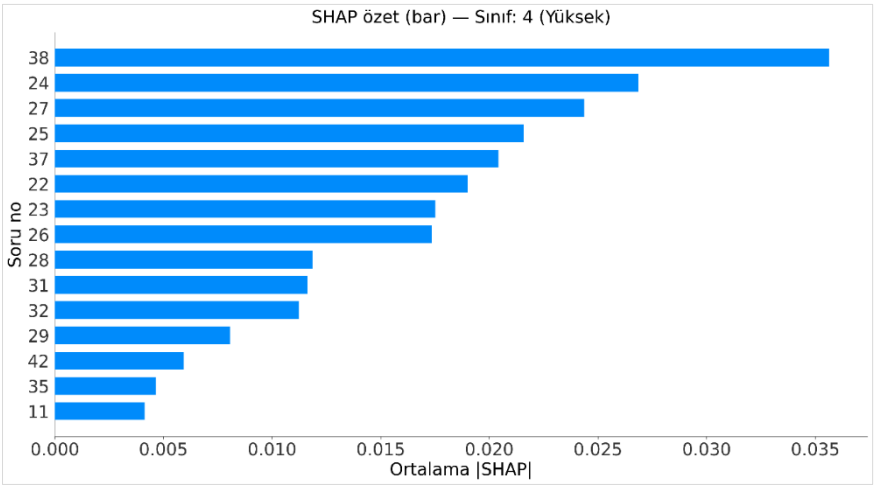
Şekil 3. *Kalibrasyon Eğrisi*

Kalibrasyon başarısı ise Şekil 3'de sunulan eğri aracılığıyla görselleştirilmiş, tahmin edilen olasılıkların gerçek risk olasılıklarıyla uyumlu olduğu açık biçimde görülmüştür. Kalibrasyon eğrileri, modelin tahmin ettiği olasılıkların gerçek sınıf dağılımıyla ne kadar örtüştüğünü göstermekte; izotonik kalibrasyon eğrisinin doğrusal referans çizgisine yakın seyretmesi, modelin güvenilir olasılık tahminleri ürettiğini açık biçimde ortaya koymaktadır. SHAP yöntemiyle elde edilen bulgular, modelin karar mekanizmasının şeffaf biçimde yorumlanmasını sağlamıştır. Küresel düzeyde yapılan analizlerde 38. maddenin (sosyal medya kullanım süresine ilişkin öznel değerlendirme) en güçlü katkıyı sunduğu, 24. ve 25. maddelerin (yoksunluk ve sınırlama güçlüğü) ile 27. maddenin (sorumlulukları erteleme) yüksek risk grubunu ayırmada kritik önemde olduğu görülmüştür. Ayrıca, 26. madde (uyku düzeni) ve 42. madde (siber zorbalık deneyimleri) de risk sınıflamasına önemli katkılar sağlamıştır. Bu durum, sosyal medya risk düzeyinin yalnızca davranışsal bağımlılık boyutu ile değil, aynı zamanda psikososyal etkilerle de belirlendiğini göstermektedir. Faktör analizinde yüksek yük değerleri veren maddelerle SHAP değerlerinin tutarlı çıkması, hem ölçüm aracının geçerliliğini hem de modelin karar süreçlerinin güvenilirliğini desteklemektedir.



Şekil 4. *SHAP beeswarm grafiği*

Küresel katkılar beeswarm ve bar özet grafikleriyle görselleştirilmiş olup, sırasıyla Şekil 4 ve Şekil 5'te sunulmuştur. Beeswarm grafiği, her bir değişkenin tahmin üzerindeki etkisinin yönünü ve büyüklüğünü gösterirken; bar grafiği, değişkenlerin ortalama mutlak SHAP değerlerine göre sıralamasını ortaya koymaktadır. Bu grafikler, modelin hangi maddelere daha fazla ağırlık verdiğini açık biçimde görselleştirerek açıklanabilirliği güçlendirmektedir. Yerel düzeydeki SHAP analizleri ise her bir katılımcının risk profiline ilişkin ayrıntılı bir tablo ortaya koymuş, böylece bireysel müdahale ve politika önerilerinin hangi risk faktörleri üzerinden geliştirilebileceğine dair ipuçları sağlamıştır.



Şekil 5. *SHAP bar özet grafiği*

Bulgular bir arada ele alındığında üç temel sonuç öne çıkmaktadır. İlk olarak, ölçek tek baskın faktörlü ve yüksek güvenilirlik düzeyine sahip bir yapıya işaret etmiş, böylece psikometrik açıdan sağlam bir ölçüm aracı elde edilmiştir. İkinci olarak, RF modeli yüksek ayrım gücü ve güvenilir olasılık tahminleri ile sosyal medya risk sınıflamasında güçlü bir performans sergilemiştir. Üçüncü olarak, SHAP açıklamaları risk sınıflamasında hem davranışsal (kontrol kaybı, yoksunluk, kullanım süresi, erteleme) hem de psikososyal (uyku sorunları, siber zorbalık) faktörlerin belirleyici olduğunu ortaya koymuş, bu

durum ölçek ve model bulgularının tutarlılığını güçlendirmiştir. Bu bütüncül tablo, çalışmanın yalnızca istatistiksel olarak değil, aynı zamanda pratik ve yorumlanabilir çıktılar sunduğunu göstermektedir. Bulgular, sosyal medya risk düzeyinin hem bireysel davranış kalıpları hem de sosyal etkileşim boyutları üzerinden anlaşılması gerektiğini ortaya koymakta ve sonraki bölümde tartışılacak teorik ve uygulamalı katkıların temelini oluşturmaktadır.

Tartışma

Bu çalışmanın bulguları, sosyal medya kullanımına ilişkin risk düzeylerinin hem psikometrik hem de ML temelli yaklaşımlarla güvenilir biçimde değerlendirilebileceğini göstermektedir. Tek baskın faktörlü yapı ve yüksek iç tutarlılık katsayısı, ölçeğin ölçüm gücünü desteklerken, bu bulgu sosyal medya bağımlılığı ölçekleri literatüründe daha önce bildirilen sonuçlarla uyumludur. Örneğin Andreassen ve ark. [4] tarafından geliştirilen Bergen Sosyal Medya Bağımlılığı Ölçeği'nin farklı kültürel bağlamlarda da benzer şekilde tek faktörlü yapılar sergilediği rapor edilmiştir. Bu durum, çalışmamızın elde ettiği psikometrik yapının yalnızca yerel değil, küresel düzeyde de geçerli olabileceğini göstermektedir. RF modeliyle elde edilen performans sonuçları literatürdeki benzer çalışmalara kıyasla oldukça güçlüdür. ROC-AUC değerinin 0.959 olması, modelin sınıflar arasında ayırt edici gücünün yüksek olduğunu, PR-AUC değerinin 0.868 olması ise sınıf dengesizliklerine rağmen modelin güvenilir tahminler üretebildiğini göstermektedir. Önceki araştırmalarda dijital bağımlılık sınıflama çalışmalarında AUC değerleri genellikle 0.80–0.90 aralığında rapor edilmiştir [20, 21]. Dolayısıyla bu çalışmanın performans metrikleri, mevcut literatürün üzerinde bir başarıya işaret etmektedir. Bunun yanı sıra izotonik regresyonla yapılan kalibrasyon, tahmin edilen olasılıkların güvenilirliğini artırmış ve modelin yalnızca ayırt edici değil, aynı zamanda olasılık açısından da doğru tahminler üretebilmesini sağlamıştır. Bu nokta, önceki çalışmalarda çoğu zaman göz ardı edilen ancak karar verme süreçleri için kritik olan bir katkıyı temsil etmektedir. Açıklanabilirlik açısından elde edilen bulgular, sosyal medya bağımlılığının çok boyutlu yapısını gözler önüne sermektedir. Özellikle sosyal medya kullanım süresine ilişkin öznel değerlendirmeler, yoksunluk belirtileri, kullanım sınırlama güçlüğü ve sorumlulukların ertelenmesi gibi maddeler yüksek risk grubunun belirlenmesinde kritik rol oynamıştır. Bunun yanı sıra uyku düzenine ilişkin sorunlar ve siber zorbalık deneyimleri de risk sınıflamasına anlamlı katkılar sağlamıştır. Bu sonuçlar, literatürde sosyal medya kullanımının yalnızca davranışsal bağımlılık boyutuyla değil, aynı zamanda psikososyal etkilerle de ele alınması gerektiğini vurgulayan çalışmalarla örtüşmektedir [22-24]. Özellikle SHAP yönteminin küresel ve yerel düzeyde sunduğu açıklamalar, ölçek maddeleriyle modelin karar mekanizması arasında güçlü bir tutarlılık ortaya koymuştur. Bu durum, hem ölçüm aracının geçerliliğini hem de kullanılan ML yaklaşımının güvenilirliğini desteklemektedir. Yöntemsel açıdan değerlendirildiğinde, bu çalışma birkaç önemli katkı sunmaktadır. İlk olarak, RF algoritmasının kullanımı, çok boyutlu anket verilerinde değişkenler arası etkileşimlerin etkili biçimde yakalanabilmesine olanak tanımıştır. İkinci olarak, izotonik kalibrasyonun

uygulanmasıyla modelin olasılık tahminleri güvenilir hale getirilmiş, bu da risk düzeylerinin pratik karar süreçlerinde daha doğru kullanılabilmesini sağlamıştır. Üçüncü olarak, SHAP yöntemi yalnızca modelin küresel düzeyde hangi maddelere daha çok ağırlık verdiğini göstermekle kalmamış, aynı zamanda bireysel katılımcı düzeyinde risk profillerinin ayrıntılı biçimde anlaşılmasına imkân tanımıştır. Ayrıca, OOF (out-of-fold) yaklaşımıyla açıklamaların yalnızca eğitim verisine bağlı kalmaması sağlanarak yanlılık riski en aza indirilmiştir. Çalışmanın teorik katkısı, sosyal medya bağımlılığının yalnızca bireysel kullanım alışkanlıklarıyla değil, aynı zamanda psikososyal etkileşimlerle de şekillenen bütüncül bir olgu olduğunu göstermesidir. Uygulamalı açıdan bakıldığında, elde edilen sonuçlar eğitim kurumlarının, ailelerin ve sağlık profesyonellerinin bireylerin risk düzeylerini daha doğru ve şeffaf biçimde belirleyebilmelerine katkı sunmaktadır. Politika yapıcılar açısından ise bu tür risk sınıflandırmaları, sosyal medya kullanımına ilişkin düzenlemeler ve önleme stratejilerinin geliştirilmesinde veri temelli bir dayanak sağlayabilir. Elde edilen model sonuçları, yalnızca akademik araştırmalarda değil, aynı zamanda uygulamalı ortamlarda da kullanılabilir niteliktedir. Özellikle eğitim kurumlarında öğrencilerin sosyal medya bağımlılığı risk düzeylerinin erken dönemde belirlenmesi, farkındalık eğitimlerinin kişiye özel planlanması ve dijital okuryazarlık programlarının geliştirilmesi için bu model pratik bir karar destek aracı olarak kullanılabilir. Ayrıca, psikolojik danışmanlık ve rehberlik birimlerinde bireylerin dijital davranış örüntülerine dayalı risk analizlerinin yapılmasına da olanak tanımaktadır. Bu çalışma, sosyal medya bağımlılığı araştırmalarında ilk kez yapay zekâ temelli bütüncül bir açıklanabilirlik yaklaşımını (OOF–SHAP–ensemble) sosyal psikolojik ölçeklerle birleştirerek, bireysel risk profillerinin hem nicel hem de yorumlanabilir biçimde modellenmesini sağlamaktadır. Bu yönüyle çalışma, geleneksel istatistiksel analizleri aşarak hem yöntemsel hem kavramsal düzeyde yenilik sunmaktadır. Bununla birlikte, çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle, örneklemin 203 katılımcı ile sınırlı olması, sonuçların genellenebilirliğini kısıtlamaktadır. Ayrıca, verilerin öz-bildirim (self-report) yöntemiyle toplanmış olması, yanıtların doğruluğu üzerinde potansiyel yanlılıklar yaratabilir. Ölçeğin yalnızca tek bir kültürel bağlamda uygulanmış olması da sonuçların farklı coğrafi veya kültürel gruplarda tekrar doğrulanmasını gerekli kılmaktadır. Gelecek araştırmalar açısından, bu çalışma yapay zekâ tabanlı açıklanabilir modellerin sosyal psikoloji ölçekleriyle bütünleştirilmesi için bir temel sunmaktadır. İlerleyen çalışmalarda, derin öğrenme (deep learning) tabanlı modellerin (örneğin LSTM, BERT veya TabNet) RF temelli yaklaşımlarla karşılaştırmalı olarak değerlendirilmesi, model performansının farklı veri boyutları ve örneklem çeşitliliği altında nasıl değiştiğine ışık tutabilir. Ayrıca, zamana bağlı veri setlerinin kullanılması, bireylerin sosyal medya bağımlılığı dinamiklerinin nedensel yönlerini daha derin biçimde analiz etmeye olanak sağlayacaktır. Sonuç olarak, bu çalışmanın bulguları hem teorik hem de uygulamalı düzeyde önemli katkılar sunmaktadır. Ölçeğin psikometrik açıdan güvenilirliği, RF modelinin yüksek performansı ve SHAP açıklamalarının sağladığı şeffaflık bir araya geldiğinde, sosyal medya risk düzeylerinin daha güvenilir, anlaşılır ve uygulanabilir biçimde sınıflandırılabileceği görülmektedir.

Gelecek Araştırmalar İçin Öneriler

Bu çalışmada elde edilen bulgular, sosyal medya risk düzeylerinin sınıflandırılmasında güçlü bir çerçeve sunmakla birlikte, anket yoluyla veri toplamanın katılımcı sayısını artırmada güçlükler barındırdığı da göz önünde bulundurulmalıdır. Öncelikle, örneklemin büyütülmesi ve farklı yaş, sosyoekonomik ve kültürel grupları kapsayacak şekilde çeşitlendirilmesi, sonuçların genellenebilirliğini artıracaktır. Ayrıca, boylamsal veri toplama yöntemleri kullanılarak sosyal medya risk düzeylerinin zaman içerisindeki değişiminin incelenmesi önerilmektedir. Böylece yalnızca mevcut durum değil, aynı zamanda riskin gelişim süreçleri de anlaşılabilir hale gelecektir.

Sonuç

Bu çalışma, sosyal medya kullanımına ilişkin risk düzeylerini sınıflandırmada ölçek temelli anket verilerini ML ile bütünleştiren yenilikçi bir çerçeve sunmaktadır. Katılımcılardan elde edilen yanıtlar, RF modeli aracılığıyla yüksek doğrulukla sınıflandırılmış ve olasılık tahminleri izotonik kalibrasyon ile güvenilir hale getirilmiştir. Doğrulama kümesinde ulaşılan performans değerleri (ROC-AUC = 0.959, PR-AUC = 0.868, Brier skoru = 0.255, LogLoss = 0.362) modelin hem ayrım gücü hem de olasılık kalitesi açısından güçlü bir performans ortaya koyduğunu göstermektedir. Açıklanabilirlik analizleri, risk sınıflamasının yalnızca kullanım alışkanlıklarına değil, aynı zamanda uyku sorunları ve siber zorbalık gibi psikososyal faktörlere de dayandığını göstermiştir. Bu bulgu, sosyal medya bağımlılığının çok boyutlu doğasını desteklemekte ve bireylerin risk profillerinin davranışsal ve sosyal etkileşim kalıplarıyla birlikte şekillendiğini ortaya koymaktadır. Faktör analizi bulguları ile SHAP katkılarının tutarlılığı, kullanılan ölçüm aracının ve modelin güvenilirliğini pekiştirmektedir. Çalışmanın katkısı üç düzeyde öne çıkmaktadır. Teorik düzeyde, sosyal medya bağımlılığına ilişkin risklerin davranışsal ve psikososyal boyutlarının bütünleştirilmesi sağlanmıştır. Metodolojik düzeyde, RF, izotonik kalibrasyon ve SHAP'ın OOF yaklaşımıyla entegrasyonu, sosyal bilimlerde anket verilerinin işlenmesi için şeffaf ve güçlü bir analitik iş akışı sunmuştur. Uygulamalı düzeyde ise elde edilen sonuçlar, eğitimciler, aileler, klinik uygulayıcılar ve politika yapıcılar için bireylerin risk düzeylerini daha doğru ve veri temelli biçimde değerlendirme olanağı sağlamaktadır. Sonuç olarak, bu çalışma sosyal medya kullanımına ilişkin risk sınıflamasında hem yüksek doğruluk hem de açıklanabilirlik sağlayan bir yaklaşım ortaya koymuş, farklı örneklemler ve kültürel bağlamlarda yapılacak araştırmalar için sağlam bir metodolojik temel sunmuştur. Bu temel, özellikle eğitim kurumlarında erken tarama, sağlık politikalarında önleme programları ve sosyal medya regülasyonlarında risk tabanlı stratejilerin geliştirilmesine katkı sağlamaktadır.

Teşekkür -

Fon/Finansman Bilgileri: Bu çalışma için herhangi bir kurum ve/veya kuruluştan destek alınmamıştır.

Etik Kurul Onayı ve İzinler: Bu araştırma, Hacettepe Üniversitesi Sosyal ve Beşeri Bilimler Araştırma Etik Kurulu'nun 21 Ocak 2025 tarih ve E-66777842-300-00004014095 karar numaralı onayı ile yürütülmüştür. Tüm katılımcılardan araştırmaya gönüllü olarak katılım sağladıklarına ilişkin bilgilendirilmiş onam alınmıştır.

Çıkar Çatışmaları/Çatışan Çıkarlar: Yazarlar arasında herhangi bir çıkar çatışması yoktur.

Yazarların Katkısı: Yazarlar çalışmaya eşit oranda katkı sağlamıştır

Kaynaklar

[1] Andreassen, C. S. (2015). Online social network site addiction: A comprehensive review. *Current Addiction Reports, 2*(2), 175-184. https://doi.org/10.1007/s40429-015-0056-9

[2] Kuss, D. J., & Griffiths, M. D. (2017). Social networking sites and addiction: Ten lessons learned. international *Journal of Environmental Research and Public Health, 14*(3), 1-17. https://doi.org/doi:10.3390/ijerph14030311

[3] Griffiths, M. (2005). A components model of addiction within a biopsychosocial framework. *Journal of Substance Use, 10*(4), 191-197. https://doi.org/10.1080/14659890500114359

[4] Andreassen, C. S., Pallesen, S., & Griffiths, M. D. (2017). The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey. *Addictive Behaviors, 64*(2017), 287-293. https://doi.org/https://doi.org/10.1016/j.addbeh.2016.03.006

[5] Lin, C.-Y., Broström, A., Nilsen, P., Griffiths, M. D., & Pakpour, A. H. (2017). Psychometric validation of the Persian Bergen Social Media Addiction Scale using classic test theory and Rasch models. *Journal of behavioral addictions, 6*(4), 620-629. https://doi.org/10.1556/2006.6.2017.071

[6] Niculescu-Mizil, A., & Caruana, R. (2005, August 7-11). Predicting good probabilities with supervised learning [Conference presentation]. Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany. https://doi.org/10.1145/1102351.1102430

[7] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017).* https://arxiv.org/abs/1705.07874

[8] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

[9] Doğan, A. (2024). Prediction of major earthquakes that may affect İzmir using machine learning methods. *Yerbilimleri, 45*(2), 93-106. https://doi.org/10.17824/yerbilimleri.1402618

[10] Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research, 15*(1), 3133-3181.

[11] Biau, G. & Scornet, E. (2016). A random forest guided tour. *TEST, 25*(2), 197-227. https://doi.org/10.1007/s11749-016-0481-7

[12] Cutler, D. R., Edwards Jr, T.C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology, 88*(11) 2783-2792. https://doi.org/10.1890/07-0539.1

[13] Silva Filho, T., Song, H., Perello-Nieto, M., Santos-Rodriguez, R., Kull, M., & Flach, P. (2023). Classifier calibration: a survey on how to assess and improve predicted class probabilities. *Machine Learning, 112*(9), 3211-3260. https://doi.org/10.1007/s10994-023-06336-7

[14] Vimbi, V., Shaffi, N., & Mahmud, M. (2024). Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. *Brain Informatics, 11*(1), 10. https://doi.org/10.1186/s40708-024-00222-1

[15] Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics, 26*(10), 1340-1347. https://doi.org/10.1093/bioinformatics/btq134

[16] Arévalo-Cordovilla, F. E., & Peña, M. (2025). Evaluating ensemble models for fair and interpretable prediction in higher education using multimodal data. *Scientific Reports, 15*(29420) 1-14. https://doi.org/10.1038/s41598-025-15388-9

[17] Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. *Clinical and Translational Science, 17*(e70056), 1-15. https://doi.org/10.1111/cts.70056

[18] Zhou, W., Yan, Z., & Zhang, L. (2024). A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction. *Scientific Reports, 14*(5905), 1-16. https://doi.org/10.1038/s41598-024-55243-x

[19] Doğan, A., & Ünal, C. (2024). Üniversite yerleşme başarısını etkileyen faktörlerin yapay zekâ yöntemleriyle araştırılması. *Türk Eğitim Bilimleri Dergisi, 22*(3), 1678-1698. https://doi.org/10.37217/tebd.1438947

[20] Gan, Y., Kuang, L., Xu, X.-M., Ai, M., He, J.-L., Wang, W., Hong, S., Chen, J.m., Cao, J., & Zhang, Q. (2025). Application of machine learning in predicting adolescent Internet behavioral addiction. *Frontiers in Psychiatry, 15*(2025), 1-8. https://doi.org/10.3389/fpsyt.2024.1521051

[21] Liu, J., Chen, L., Chen, Y., Luo, J., Yu, K., Fan, L., Yong, C., He, H., Liao, S., Ge, Z., & Jiang, L. (2025). Explainable machine learning prediction of internet addiction among Chinese primary and middle school children and adolescents: A longitudinal study based on positive youth development data (2019–2022). *Frontiers in Public Health, 13*(2025), 1-13. https://doi.org/10.3389/fpubh.2025.1590689

[22] Zarate, D., Hobson, B. A., March, E., Griffiths, M. D., & Stavropoulos, V. (2023). Psychometric properties of the Bergen Social Media Addiction Scale: An analysis using item response theory. *Addictive Behaviors Reports, 17*(100473), 1-8. https://doi.org/https://doi.org/10.1016/j.abrep.2022.100473

[23] Yu, D. J., Wing, Y. K., Li, T. M. H., & Chan, N. Y. (2024). The impact of social media use on sleep and mental health in youth: a scoping review. *Current Psychiatry Reports, 26*(3), 104-119. https://doi.org/10.1007/s11920-024-01481-9

[24] Peprah, P., Oduro, M. S., Atta-Osei, G., Addo, I. Y., Morgan, A. K., & Gyasi, R. M. (2024). Problematic social media use mediates the effect of cyberbullying victimisation on psychosomatic complaints in adolescents. *Scientific Reports, 14*(9773), 1-12. https://doi.org/10.1038/s41598-024-59509-2