Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

VERİ MADENCİLİĞİ YAKLAŞIMI İLE BANKACILIK KAMPANYA BAŞARISI ANALİZİ: XGBOOST VE RASTGELE ORMAN TABANLI UYGULAMALI BIR MODEL

Analyzing Banking Campaign Success through a Data Mining Approach: A Practical Model Based on XGBoost and Random Forest

Kamil Abdullah EŞİDİR[1]

MAKALE BİLGİSİ

Araştırma Makalesi

Makale Geliş Tarihi : 10/05/2025
Makale Kabul Tarihi : 13/06/2025

ÖZ

Çalışmada, bankacılık sektöründe yürütülen doğrudan pazarlama kampanyalarının başarısını tahmin etmek için, makine öğrenmesi tabanlı iki farklı sınıflandırma modelinin karşılaştırmalı analizi yapılmıştır. UCI Machine Learning Repository'de yayımlanan "Bank Marketing" veri seti kullanılarak, müşterilerin mevduat aboneliği kararları tahmin edilmiştir. Veri seti, demografik, sosyoekonomik ve geçmiş kampanya etkileşimlerine dair çok boyutlu değişkenler içermektedir. Modelleme sürecinde, karar ağaçları tabanlı iki güçlü makine öğrenmesi modeli olan XGBoost ve Rastgele Orman kullanılmıştır. Veri setindeki sınıf dengesizliği problemi göz önünde bulundurularak doğruluk, F1 skoru, kesinlik, duyarlılık ve ROC-AUC gibi çok boyutlu performans metrikleri ile değerlendirmeler yapılmıştır. Analizlerde, XGBoost genel sınıflandırmada ve ayrıştırma açısından daha başarılı olmuştur. Rastgele Orman ise pozitif sınıfı tanımada daha iyi sonuçlar sunmuştur. Kampanya başarısını etkileyen başlıca değişkenler olarak yaş, konut ve kişisel kredi durumu ile önceki kampanya yanıtları öne çıkmıştır. Kredi yükümlülüğü olmayan bireylerin ve geçmiş dönemlerde yapılmış kampanyalara olumlu yanıt vermiş müşterilerin abonelik eğilimlerinin daha yüksek olduğu belirlenmiştir. Bilgi sızıntısı riski taşıyan "duration" değişkeni analiz dışı bırakılmış ve özellik mühendisliği süreçleriyle modelin etik ve metodolojik güvenilirliği sağlanmıştır. Elde edilen bulgular, kişiselleştirilmiş pazarlama stratejilerinin geliştirilmesi ve yapay zekâ temelli karar destek sistemlerinin bankacılık uygulamalarına entegrasyonu açısından önemli katkılar sunmaktadır. Çalışma, hem akademik literatüre hem de uygulamalı finans ve pazarlama alanlarına katkılarda bulunmaktadır.

Anahtar Kelimeler: Veri Madenciliği, Makine Öğrenmesi, XGBoost, Bankacılık Kampanyası, Abonelik Tahmini.

ARTICLE INFORMATION

Research Article

Submission Date : 10/05/2025
Accepted Date : 13/06/2025

ABSTRACT

In this study, a comparative analysis of two different machine learning-based classification models is conducted to predict the success of direct marketing campaigns in the banking sector. Using the "Bank Marketing" dataset published in the UCI Machine Learning Repository, customers' deposit subscription decisions are predicted. The dataset contains multidimensional variables on demographic, socioeconomic and past campaign interactions. Two powerful machine learning models based on decision trees, XGBoost and Random Forest, are used in the modeling process. Considering the class imbalance problem in the dataset, multidimensional performance metrics such as accuracy, F1 score, precision, sensitivity and ROC-AUC were evaluated. In the analysis, XGBoost was more successful in terms of overall classification and discrimination. Random Forest provided better results in recognizing the positive class. The main variables affecting campaign success were age, housing and personal loan status, and previous campaign responses. It was found that individuals with no credit obligations and customers who responded

[1] Dr., Fırat Kalkınma Ajansı, e-posta: abdullahesidir@yahoo.com, ORCID: 0000-0002-8106-1758 (Correspondent Author/ Sorumlu Yazar)

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

positively to previous campaigns had a higher propensity to subscribe. The "duration" variable, which carries the risk of information leakage, was excluded from the analysis and the ethical and methodological reliability of the model was ensured through feature engineering processes. The findings make important contributions to the development of personalized marketing strategies and the integration of artificial intelligence-based decision support systems into banking applications. The study contributes to both academic literature and applied finance and marketing fields.
Keywords: Data Mining, Machine Learning, XGBoost, Banking Campaign, Subscriber Prediction.

1. Giriş

Hızla gelişen dijital dünyada bankacılık sektörü, müşteri ilişkilerini veri temelli stratejiler güderek yeniden yapılandırmaktadır. Özellikle bireysel bankacılık hizmetlerinin yürüttüğü kampanyalar istatistiksel modelleme ve yapay zekâ tabanlı analizler ile değerlendirilmektedir. Gerçekleşen dijital dönüşüm, müşteri davranışlarının sadece geçmiş işlemler üzerinden değil, aynı zamanda sosyo-demografik ve davranışsal öngörüler yoluyla modellenmesini gerekli kılmaktadır. Büyük veri analitiği ve makine öğrenmesi modelleri, bankacılık kampanyalarının başarısını tahmin etmeye yönelik gelişmiş karar destek araçları sunmaktadır (Davenport ve Harris, 2017).

Makine öğrenmesi, veri kümeleri içerisinde yer alan karmaşık örüntüleri ve ilişkileri geleneksel istatistiksel modellere kıyasla daha esnek biçimde ortaya çıkarabilen, denetimli ve denetimsiz algoritmalar bütünü olarak tanımlanmaktadır. Bu algoritmalar, özellikle sınıflandırma problemlerinde yüksek başarı oranları ile öne çıkmaktadır. Müşteri davranışlarını doğru tahmin edebilme potansiyeli sayesinde pazarlama analitiği alanında giderek daha fazla tercih edilmektedir (Aslan, 2025).

Ancak literatürde, bankacılık sektörüne yönelik pazarlama kampanyalarının tahmine dayalı analizi konusunda yapılan çalışmaların büyük bölümü ya sınırlı sayıda değişkenle çalışmakta ya da model performansını sadece doğruluk oranı üzerinden değerlendirmektedir. Bunun yanı sıra, birçok çalışmada değişkenlerin kampanya başarısı üzerindeki etkileri ayrıntılı biçimde analiz edilmemektedir. Bu durum, bankacılık kampanyalarının dinamik doğasının ve müşteri yanıtlarını etkileyen çok boyutlu faktörlerin göz ardı edilmesine neden olabilmektedir (Choi & Choi, 2023).

Portekiz'de faaliyet gösteren bir bankaya ait gerçek pazarlama verileri analiz edilmiştir. Veri seti, UCI Machine Learning Repository'de yer almakta ve kamuya açık olarak sunulmaktadır (Aydın & Yalçın, 2024). Çalışmada, müşteri abonelik kararlarını tahmin etmek için Rastgele Orman ve XGBoost modelleri kullanılmıştır. Ayrıca kampanya başarısını etkileyen sosyo-demografik değişkenler istatistiksel ve görsel olarak incelenmiştir. Veri sızıntısı riski taşıyan değişkenler analizden çıkarılmıştır. Sınıf dengesizliği problemi ise uygun

DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

yöntemler ile giderilmiştir. Bu yönleri ile çalışma hem yöntemsel hem de pratik uygulama açısından literatüre katkıda bulunmaktadır.

2. Literatür Taraması

Bankacılık sektöründe müşteri davranışlarını analiz eden çalışmalar, son yıllarda dijitalleşmenin artması ile beraber veri odaklı modellere ve çözümlere yönelmiştir. Geleneksel istatistiksel modeller, çok değişkenli ve doğrusal olmayan ilişkilerin yoğun kullanıldığı büyük veri yapılarında yetersiz kalmaktadır. Karşılaşılan bu durum, daha esnek ve güçlü tahminleme yeteneğine sahip modellerin kullanılması gerekliliğini ortaya çıkarmaktadır (Speer, 2021). Makine öğrenmesi modelleri, bankacılık benzeri verileri yoğun kullanan sektörlerde, müşteri davranışlarının modellenmesi açısından birçok avantajlar sunmaktadır (Wu, 2023).

Veri odaklı karar verme süreçlerinin yaygınlaşması ile birlikte, CRM uygulamalarında veri madenciliği ve makine öğrenmesi modellerinin kullanımı giderek artmaktadır. Bu modeller; müşteri segmentasyonu, davranış tahmini ve kişiselleştirilmiş kampanya yönetimi gibi alanlar için önemli katkılar sunmaktadır. Ngai, Xiu ve Chau (2009) tarafından gerçekleştirilen kapsamlı bir literatür taramasında, veri madenciliği tekniklerinin CRM'in dört temel boyutunda nasıl uygulandığı sistematik biçimde sınıflandırılmıştır. Müşteri tutma stratejilerinde sınıflandırma ve birliktelik kurallarının sıklıkla tercih edildiği görülmüştür. Lemmens ve Croux (2009) ise müşteri kaybı (churn) tahminine yönelik çalışmalarında, bagging ve boosting gibi makine öğrenmesi tabanlı modellerin, geleneksel logit modellerine kıyasla daha yüksek sınıflandırma başarısı sunduğunu göstermiştir. Ayrıca, nadir olayların tahmininde dengeli örnekleme (balanced sampling) ve önyargı düzeltme tekniklerinin (örneğin intercept düzeltmesi) model performansını artırdığı vurgulanmıştır. Geleneksel modellerin sınırlı tahminleme kapasitesi, özellikle dengesiz sınıf dağılımları söz konusu olduğunda, model performansının düşmesine neden olmaktadır. Bu nedenle, karar ağaçları temelli algoritmalar, daha yüksek genellenebilirlik ve değişkenler arası etkileşimleri modelleme kapasitesi ile tercih edilmektedir (Lemmens & Croux, 2006). Her iki çalışma da CRM süreçlerinin veri temelli karar destek sistemleriyle entegre edilmesinin, müşteri değeri yönetiminde stratejik bir avantaj sağladığını göstermektedir. Bu model, özellikle büyük ve karmaşık veri setlerinde özniteliklerin ayırt ediciliğini belirleyebilme ve aşırı öğrenmeye karşı dirençli olma özellikleriyle öne çıkmaktadır. Nitekim yapılan karşılaştırmalı çalışmalarda, XGBoost'un müşteri abonelik kararlarını öngörme konusunda lojistik regresyon ve rastgele orman gibi modellere kıyasla daha üstün performans sergilediği ortaya konmuştur (Moro vd., 2014).

Veri madenciliği ve makine öğrenmesi yöntemlerinin yalnızca müşteri odaklı pazarlama stratejilerinde değil, aynı zamanda kurumsal kaynak planlaması, insan kaynakları yönetimi ve organizasyonel karar destek sistemlerinde de yaygın biçimde kullanıldığı görülmektedir.

DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Nitekim Zhu, Sawhney ve Upreti (2016), departmanlar düzeyinde çalışan gönüllü ayrılmalarının nedenlerini analiz ettikleri çalışmalarında, tahmine dayalı modelleme tekniklerinin insan kaynakları süreçlerinde nasıl etkili biçimde uygulanabileceğini ortaya koymuşlardır. Bu çalışma, veri temelli karar destek sistemlerinin disiplinler arası uygulamalarla geniş bir yelpazede değerlendirilebileceğine işaret etmektedir.

Makine öğrenmesi tabanlı modellerin pazarlama ve müşteri ilişkileri yönetimi (CRM) alanında kullanımında yaşanan ilerlemeler, özellikle karar ağaçları temelli topluluk yöntemlerinin öne çıkmasına neden olmuştur. Geliştirilen XGBoost modeli, yüksek doğruluk oranı, hızlı öğrenme kapasitesi ve büyük veri setlerine uyumlu yapısı ile öne çıkmaktadır. Chen ve Guestrin (2016) tarafından geliştirilen model, düzenlileştirilmiş (regularized) hedef fonksiyonu, ikinci dereceden türev bilgisiyle optimize edilen öğrenme süreci ve seyrek veri yapısına duyarlı bölünme algoritması sayesinde klasik ağaç tabanlı modellerden ayrılmaktadır. Modelin etkinliği, çeşitli Kaggle yarışmalarında elde ettiği birinciliklerle de desteklenmektedir. XGBoost, yalnızca akademik çalışmalarda değil, aynı zamanda endüstriyel ölçekte uygulamalarda da yaygın olarak tercih edilmektedir. Müşteri davranış tahmini ile kampanya yönetimi süreçlerinde avantaj sağlamaktadır (Chen ve Guestrin, 2016).

Yerli literatürde ise Calp (2025), Kaggle kaynaklı 10.127 gözlem içeren kredi kartı veri seti üzerinden sekiz farklı makine öğrenmesi modeli ile müşteri kaybını (churn) tahmin etmiştir. En yüksek başarıyı Gradient Boosting modelinde %98,7 AUC değeriyle elde etmiştir. AdaBoost, Rastgele Orman ve Yapay Sinir Ağları gibi modelleri ile de yüksek doğruluk değerleri elde edilmiştir. Sonuçlar, makine öğrenmesi tabanlı modellerin kredi kartı müşteri davranışlarını tahmin etmede etkin biçimde kullanılabildiğini göstermiştir. Model değerlendirmelerinde ROC-AUC, F1 skoru, duyarlılık ve özgüllük benzeri performans ölçütleri de dikkate alınmıştır.

Çalışma, literatürdeki iki temel boşluğu ele almaktadır. İlk olarak, sınıf dengesizliği içeren veri yapılarında Rastgele Orman ve XGBoost modellerinin performansları; sadece doğruluk değil, F1 skoru, duyarlılık ve kesinlik gibi metrikler üzerinden değerlendirilmiştir. İkinci olarak ise yaş, eğitim düzeyi, meslek, kredi durumu ve önceki kampanya etkileşimleri gibi sosyo-demografik değişkenlerin kampanya başarısına etkisi, görsel ve istatistiksel analizlerle ortaya konulmuştur. Böylece, hem modelleme düzeyinde hem de karar destek sistemlerine katkı açısından literatüre katkı sunulması hedeflenmiştir.

3. Metodoloji

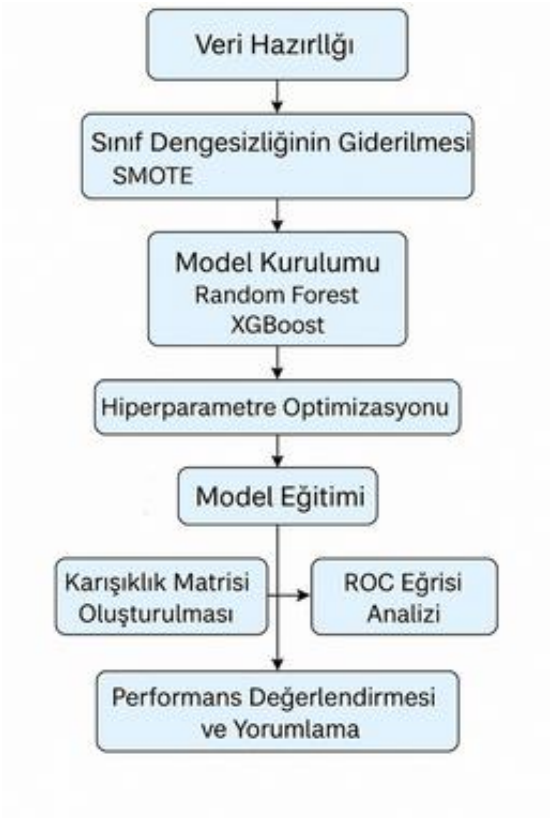
Büyük veri setleri; yüksek hacim, çeşitlilik ve hız gibi özelliklerinden ötürü, makine öğrenmesi uygulamalarında veri temizliği, ön işleme ve model karmaşıklığı benzeri alanlarda

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

çeşitli zorluklar ortaya çıkarmaktadır (Eşidir, 2025a). Çalışmada, bankacılık sektörüne yönelik pazarlama kampanyalarının başarısını tahmin etmek amacıyla, karar ağaçları tabanlı iki güçlü makine öğrenmesi modeli olan Rastgele Orman ve XGBoost modelleri kullanılmıştır. Modelleme sürecinde veri setinde çeşitli ön işleme adımları uygulanmıştır. Kategorik değişkenler uygun biçimde kodlanmış, sayısal değişkenler ise standardize edilmiştir. Ayrıca, bilgi sızıntısı riski taşıyan "duration" değişkeni analiz dışı bırakılmıştır. Eğitim ve test veri kümeleri ayrılarak modellerin aşırı öğrenme riskleri minimize edilmiştir.

Özellik seçimi aşamasında, değişkenler hem istatistiksel anlamlılık hem de bağlamsal uygunluk açısından değerlendirilmiş ve yalnızca model performansına katkı sağlayan özniteliklere yer verilmiştir. Performans değerlendirmesi, yalnızca doğruluk oranı ile değerlendirilmemiştir. F1 skoru, hassasiyet ve duyarlılık gibi metrikler de kullanılmıştır. Rastgele Orman ve XGBoost modellerinin güçlü ve zayıf yönleri karşılaştırılmış, elde edilen sonuçların pazarlama stratejilerinde kullanılabilirliği artırılmıştır. Şekil 1'de akış diyagramı ve izlenen adımlar özetle gösterilmiştir.

Şekil 1: Akış Diyagramı ve İzlenen Adımlar



DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

3.1. Yazılım ve Donanım Altyapısı

Çalışmanın veri analizi ve modelleme süreçleri, Python 3.12 programlama dili ile gerçekleştirilmiştir. Gerekli veri işleme, modelleme ve değerlendirme işlemleri için NumPy, Pandas, Scikit-learn ve XGBoost gibi yaygın kullanılan açık kaynak kütüphanelerden yararlanılmıştır. Görselleştirme işlemleri ise Matplotlib ve Seaborn kütüphaneleri aracılığıyla gerçekleştirilmiştir. Analizler, kullanıcı dostu arayüzü ve bütünleşik çalışma ortamı sunan Jupyter Notebook 7.2 platformu üzerinde yürütülmüştür. Donanım altyapısı olarak Intel Core i7 işlemcili, 16 GB RAM kapasitesine sahip, Windows 11 Pro işletim sistemli bir bilgisayar kullanılmıştır. Donanımsal özellikler, veri setinin verimli biçimde işlenmesi ve modelleme süreçlerinin uygun sürelerde tamamlanması açısından yeterli düzeydedir.

3.2. Veri Seti

Araştırmada kullanılan veri seti, UCI Machine Learning Repository tarafından yayımlanan "Bank Marketing" isimli veri setidir. Portekiz'de faaliyet gösteren bir bankanın doğrudan pazarlama kampanyalarına ilişkin müşteri yanıtlarını içermektedir. Veri seti, telefon görüşmelerine dayalı kampanya sürecinde elde edilen demografik, sosyoekonomik ve geçmiş etkileşim verilerini kapsamaktadır. Toplamda 45.211 gözlem birimi ve 12 açıklayıcı değişken ile 1 ikili hedef değişken bulunmaktadır. Hedef değişken (y), müşterinin kampanya sonucunda bir mevduat hesabı açıp açmadığını belirtmektedir. Hedef değişken "yes" (abone) ve "no" (abone değil) şeklinde sınıflandırılmıştır.

Bağımsız değişkenler arasında yaş (age), meslek (job), medeni durum (marital), eğitim düzeyi (education), kredi durumu (housing, loan), kampanya zamanı (month, day_of_week), önceki kampanya etkileşim bilgileri (pdays, previous, poutcome) ve iletişim şekli (contact) yer almaktadır. Hedef değişkenin dağılımı incelendiğinde, yalnızca %11,7'lik bir kesimin "abone" olduğu görülmektedir. Bu durum, veri setinde ciddi bir sınıf dengesizliğinin olduğunu göstermektedir. Modelleme sürecinde söz konusu dengesizlik göz önünde bulundurularak yalnızca doğruluk oranına odaklanmak yerine, sınıflandırma başarısını daha bütüncül bir şekilde değerlendiren F1 skoru, duyarlılık ve ROC-AUC gibi performans metriklerine ağırlık verilmiştir.

Veri setinde eksik gözlem bulunmamaktadır. Analiz öncesinde anlamlılık açısından değerlendirme yapılmış ve sonuçsal nitelikteki "duration" değişkeni dışlanmıştır. Bu değişken, müşteriyle yapılan telefon görüşmesinin süresini saniye cinsinden ifade etmekte, ancak değeri müşterinin abonelik kararına doğrudan bağlı olarak oluşmaktadır. Bu bağlamda, modelin tahmin etmeye çalıştığı çıktıya ilişkin bilgiyi önceden içerdiği için veri sızıntısı (data leakage) riski taşımaktadır. Literatürde bu tür değişkenlerin modelleme dışı bırakılması önerilmektedir (Kaufman vd., 2012). Çalışmada da benzer bir yaklaşım benimsenmiştir.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Tablo 1: Veri Setinden Alınan Örnek Kayıtlar

age	job	marital	education	balance	housing	loan	month	campaign	pdays	previous	poutcome	y
40	blue-collar	married	secondary	580	yes	no	may	1	-1	0	unknown	no
47	services	single	secondary	3644	no	no	jun	2	-1	0	unknown	no
25	student	single	tertiary	538	yes	no	apr	1	-1	0	unknown	no
42	management	married	tertiary	1773	no	no	apr	1	336	1	failure	no
56	management	married	tertiary	217	no	yes	jul	2	-1	0	unknown	no

Tablo 1'de, Bank Marketing veri setinden rastgele seçilmiş beş örnek müşteri kaydı sunulmaktadır. Bu kayıtlar, modelin eğitileceği temel değişkenlerin nitelikleri hakkında genel bir çerçeve sunmakta ve veri setinin yapısal çeşitliliğini gözler önüne sermektedir. Örneklerde yaş aralığı 25 ile 56 arasında değişmektedir. Pazarlama kampanyasının geniş bir yaş kitlesine hitap ettiği örneklerden anlaşılmaktadır.

Müşteri profilleri incelendiğinde, farklı meslek gruplarına (ör. blue-collar, services, student, management), medeni durumlara (single, married), ve eğitim düzeylerine (secondary, tertiary) sahip bireylerin yer aldığı görülmektedir. Bu çeşitlilik, modelin heterojen bir demografik yapıya dayanarak genellenebilir çıkarımlar yapmasına olanak sağlar. Ekonomik değişkenler açısından değerlendirildiğinde, bireylerin banka hesap bakiyelerinde ciddi farklılıklar mevcuttur (örneğin, 217 € ile 3.644 € arasında değişmektedir). Bu durum, bireysel finansal durumun kampanya sonuçlarını etkileyebileceği varsayımını destekleyebilir. Ayrıca, housing ve loan gibi kredi durumları da çeşitlilik göstermektedir; bu değişkenlerin, abonelik kararları üzerindeki etkisi modelleme sürecinde dikkatle ele alınmalıdır. Kampanya geçmişine ilişkin değişkenler olan campaign, pdays, previous ve poutcome, müşterilere daha önce ulaşılıp ulaşılmadığını ve önceki kampanyaların başarı durumunu yansıtmaktadır. Bu değişkenler, özellikle zaman temelli stratejik pazarlama analizlerinde önemli rol oynamaktadır.

3.3. Veri Ön İşleme

Veri ön işleme süreci, ham verinin analize ve modellemeye uygun hale getirilmesini hedefleyen kritik bir aşamadır. Bu süreçte uygulanan dönüşüm, temizleme ve yapılandırma işlemleri, makine öğrenmesi modellerinin doğruluğu ile genellenebilirliğini doğrudan etkileyerek model performansını belirleyici biçimde şekillendirmektedir (Eşidir, 2025a). Veri ön işleme aşamasında, kategorik değişkenler sayısal değerlere dönüştürülmüştür. Ardından, veri seti eğitim (%80) ve test (%20) olarak ayrılmıştır. Sayısal veriler, Standart Ölçekleyici kullanılarak ölçeklendirilmiştir. Özellik ölçekleme, makine öğrenmesi modellerinde eğitim sürecinin kararlılığını artırmakta ve değişkenler arası karşılaştırmayı daha anlamlı hale getirmektedir. Zira farklı ölçek aralıklarına sahip değişkenlerin birlikte modele dâhil edilmesi, modelin öğrenme sürecini olumsuz yönde etkileyebilmektedir. Bu nedenle, standartlaştırma

DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

veya normalizasyon gibi ölçekleme teknikleri kullanılarak değişkenler ortak bir ölçekte temsil edilmektedir. Böylece modelin hem öğrenme dinamikleri dengelenmekte hem de eğitim performansı iyileştirilmektedir (Tafralı, 2022). Sınıf dengesizliğini gidermek amacıyla eğitim verisi üzerinde SMOTE (Synthetic Minority Over-sampling Technique) algoritması uygulanmıştır. Bu yöntem, azınlık sınıfa ait sentetik örnekler oluşturarak sınıf dağılımını dengelemekte ve modelin pozitif sınıfı daha etkin şekilde öğrenmesini sağlamaktadır.

Veri sızıntısı, modelin tahmin etmeye çalıştığı çıktıya ilişkin bilgileri eğitim sürecinde öğrenmesi anlamına gelmektedir. Bu türden bilgiler, modelin eğitim aşamasındaki performansını yapay olarak artırmakta, ancak model gerçek hayattaki veriyle karşılaştığında bu bilgiye erişemeyeceğinden dolayı doğru tahminleme başarısız olmaktadır. Kaufman (2012), modelin yalnızca gerçek dünyada erişilebilecek bilgilerle eğitilmesi gerektiğini savunmaktadır. Modelleme aşamasında, "duration" değişkeni analiz dışı bırakılmıştır. Söz konusu değişken, müşterilerle yapılan telefon görüşmesinin süresini saniye cinsinden ifade etmektedir. Ancak, görüşme süresi doğrudan müşteri yanıtına bağlı olarak oluştuğu için, hedef değişkenle sonuçsal bir ilişki içermektedir. Bu tür bir değişkenin modelde yer alması, öğrenme sürecinde veri sızıntısı (data leakage) riski doğurarak modelin genellenebilirliğini olumsuz etkileyebilir. Veri ön işleme süreci, model başarısında doğrudan etkili olan kritik bir aşamadır.

3.4. Performans Değerlendirme Kriterleri

Makine öğrenmesi modellerinin performanslarını nesnel biçimde değerlendirebilmek amacıyla çeşitli metrikler ve analiz yöntemleri kullanılmaktadır. Modellerin başarısı; kesinlik, duyarlılık, F1 skoru ve doğruluk olmak üzere dört temel performans metriği ile analiz edilmiştir. Söz konusu ölçütler, modelin farklı senaryolardaki güçlü ve zayıf yönlerinin ortaya konmasına olanak tanımaktadır. Kesinlik, modelin pozitif sınıf için yaptığı tahminlerin ne ölçüde doğru olduğunu göstermektedir. Duyarlılık, gerçek pozitif örneklerin ne kadarının doğru şekilde sınıflandırıldığını belirtmektedir (Eşidir, 2025b). F1 Skoru, bu iki metriği dengeli biçimde birleştirerek sınıflandırma performansını tek bir ölçüt altında özetler. Doğruluk ise modelin genel başarı oranını ifade eder ve tüm sınıflar üzerindeki doğru tahminlerin oranına karşılık gelir. Bu metrikler, özellikle sınıf dağılımının dengesiz olduğu veri setlerinde, modelin pozitif sınıfı ayırt etme becerisini daha gerçekçi biçimde değerlendirmek açısından kritik öneme sahiptir.

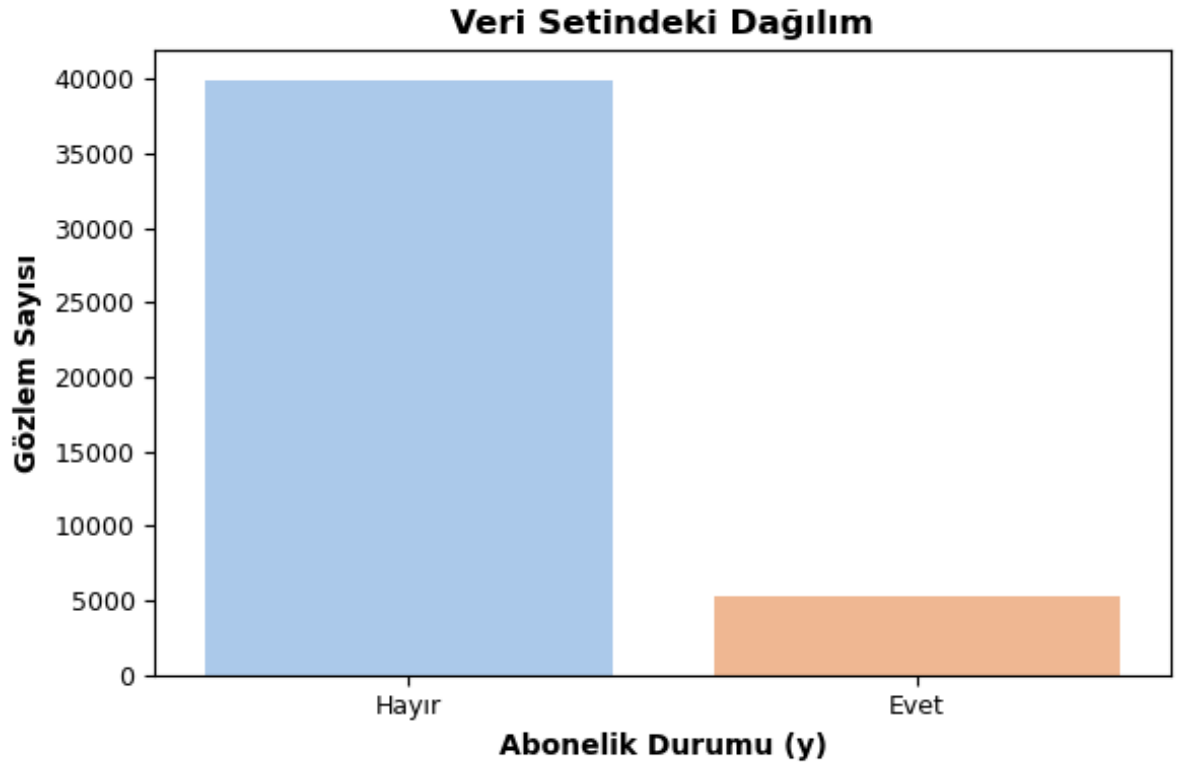
4. Görsel İstatistiksel Analizler ve Bulgular

Bu bölümde, çalışmada kullanılan veri setinin betimsel istatistikleri ve değişkenlerin hedef değişken (abonelik durumu) üzerindeki etkileri görsel analizlerle desteklenerek

DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

sunulmuştur. Amaç, kampanya başarısını etkileyen temel değişkenleri tespit etmek ve bu değişkenlerin dağılımları üzerinden modelleme sürecine dair sezgisel çıkarımlar elde etmektir.

Şekil 2: Veri Setindeki Dağılımı

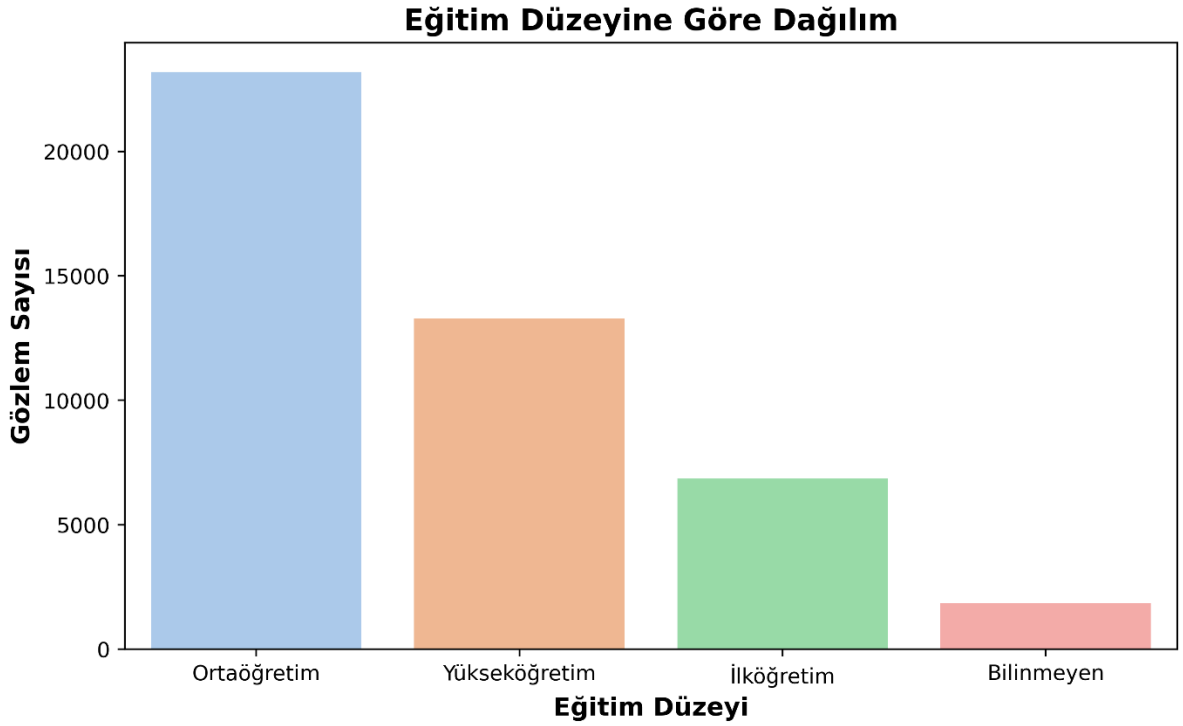


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 2'de veri setindeki hedef değişken olan y (abonelik durumu) için sınıf dağılımı görselleştirilmiştir. Veri setinde yer alan hedef değişken olan abonelik durumu, müşterilerin belirli bir bankacılık kampanyasına abone olup olmadığını göstermektedir. Veri setinde abone olmayan bireylerin sayısı 39.922 (%88,3) kişi iken, abone olan bireylerin sayısı yalnızca 5.289 (%11,7)'dir. Grafikte açıkça görüldüğü üzere, "Hayır" (abone olmayanlar) sınıfı baskın durumdadır.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Şekil 3: Eğitim Düzeyine Göre Dağılım

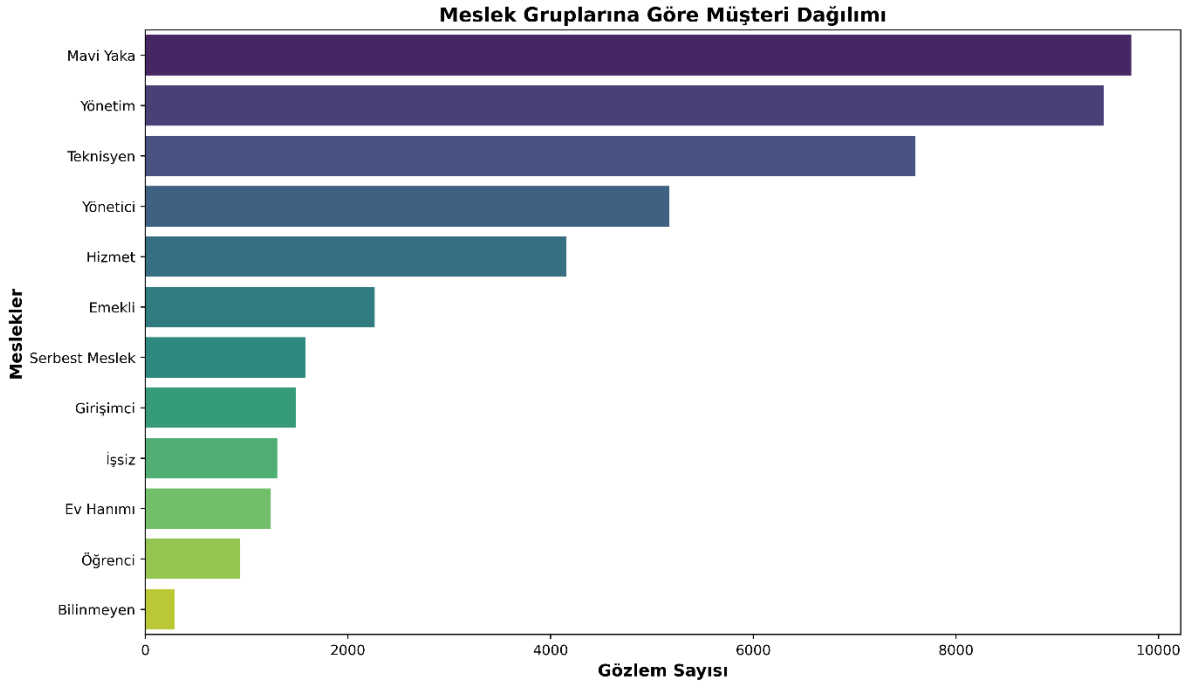


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 3'te, veri setindeki bireylerin eğitim düzeylerine göre dağılımı görselleştirilmiştir. Eğitim düzeyine göre yapılan analizlerde, bireylerin %51,3'ünün ortaöğretim, %29,4'ünün yükseköğretim ve %15,1'inin ilköğretim mezunudur. %4,1'lik bir kesimin ise eğitim durumu bilinmemektedir. Kampanya başarı oranları eğitim düzeyine göre değişkenlik göstermektedir. Özellikle yükseköğretim düzeyinde daha yüksek abonelik oranları gözlenmiştir. Bireylerin eğitim düzeyi ile finansal ürünlere olan ilgisi ve karar verme davranışları arasında anlamlı bir ilişki bulunmaktadır.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Şekil 4: Meslek Gruplarına Göre Müşteri Dağılımı

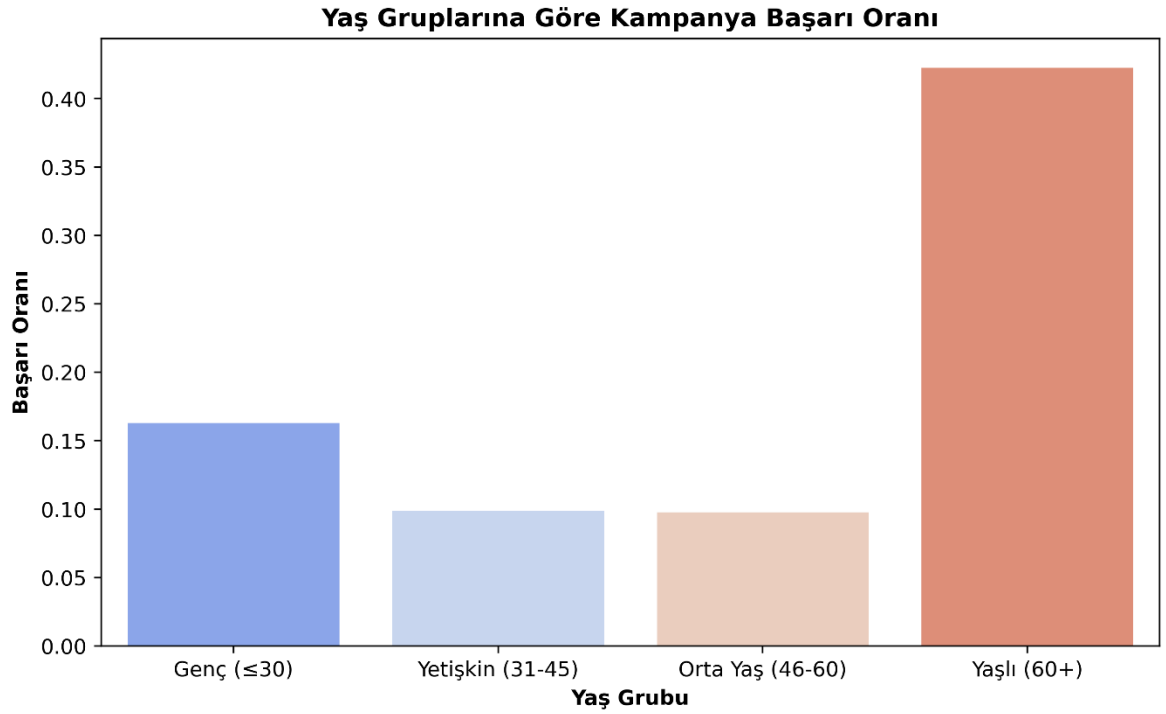


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 4'te meslek gruplarına göre müşteri dağılımı görülmektedir. Mavi yaka çalışanlar (%21,5) ve yönetici pozisyonundaki bireyler (%20,9) oranları ile başı çekmektedir. Bu grupları teknisyenler (%16,8) ve büro personeli (%11,4) takip etmektedir. Öğrenciler, işsiz bireyler ve ev hanımları ise daha düşük oranlardadır. Yönetici ve beyaz yaka çalışanların abonelik eğilimlerinin diğer gruplara kıyasla daha yüksek olduğu tespit edilmiştir. Gelir seviyesi ve finansal bilinç düzeyi ile kampanya yanıtı arasında dolaylı bir ilişkinin bulunduğu anlaşılmaktadır.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Şekil 5: Yaş Gruplarına Göre Kampanya Başarı Oranı

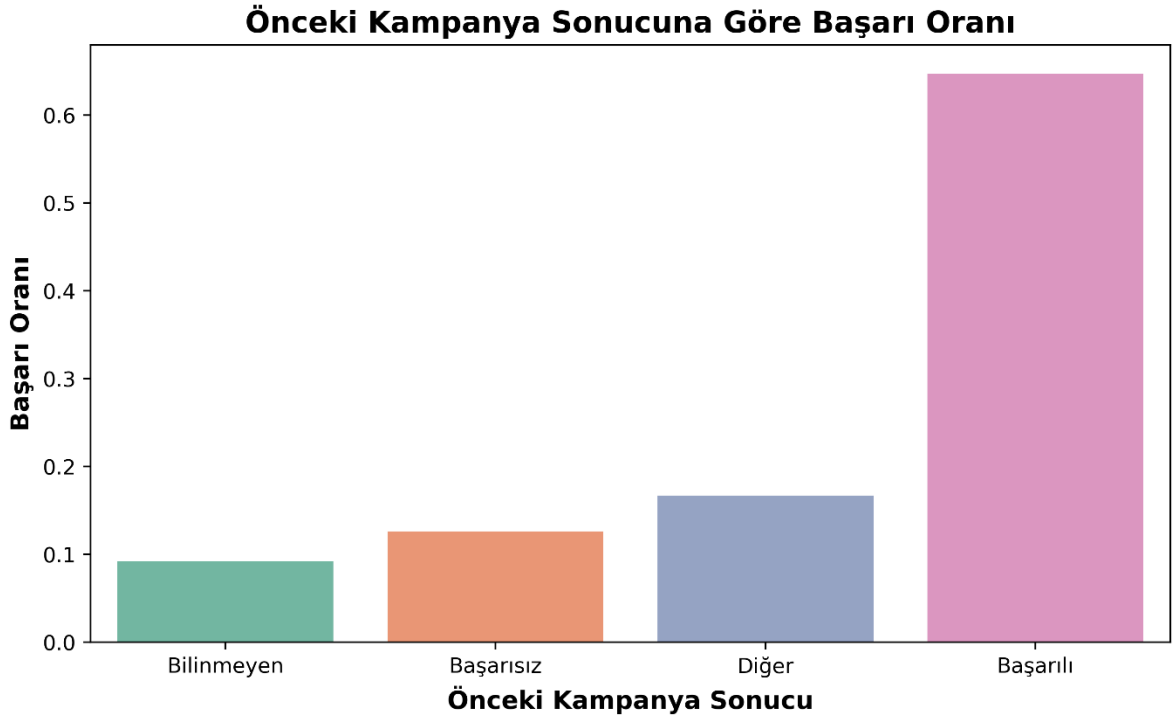


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 5'te yaş gruplarına göre kampanya başarı oranları görülmektedir. Yaş değişkeni kategorilere ayrılarak analiz edilmiştir. Sonuçlar, 60 yaş ve üzeri bireylerin kampanya tekliflerine daha duyarlı olduğunu ortaya koymuştur. Bu yaş grubundaki abonelik başarı oranı %42,26 ile diğer tüm grupların üzerinde yer almaktadır. 30 yaş altı genç bireylerde bu oran %16,29, 31–45 yaş grubunda %9,88, 46–60 yaş grubunda ise %9,78 seviyesindedir. İleri yaş bireyleri uzun vadeli tasarruf ve yatırım ürünlerine daha fazla ilgi göstermektedir.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Şekil 6: Önceki Kampanya Sonucuna Göre Başarı Oranı

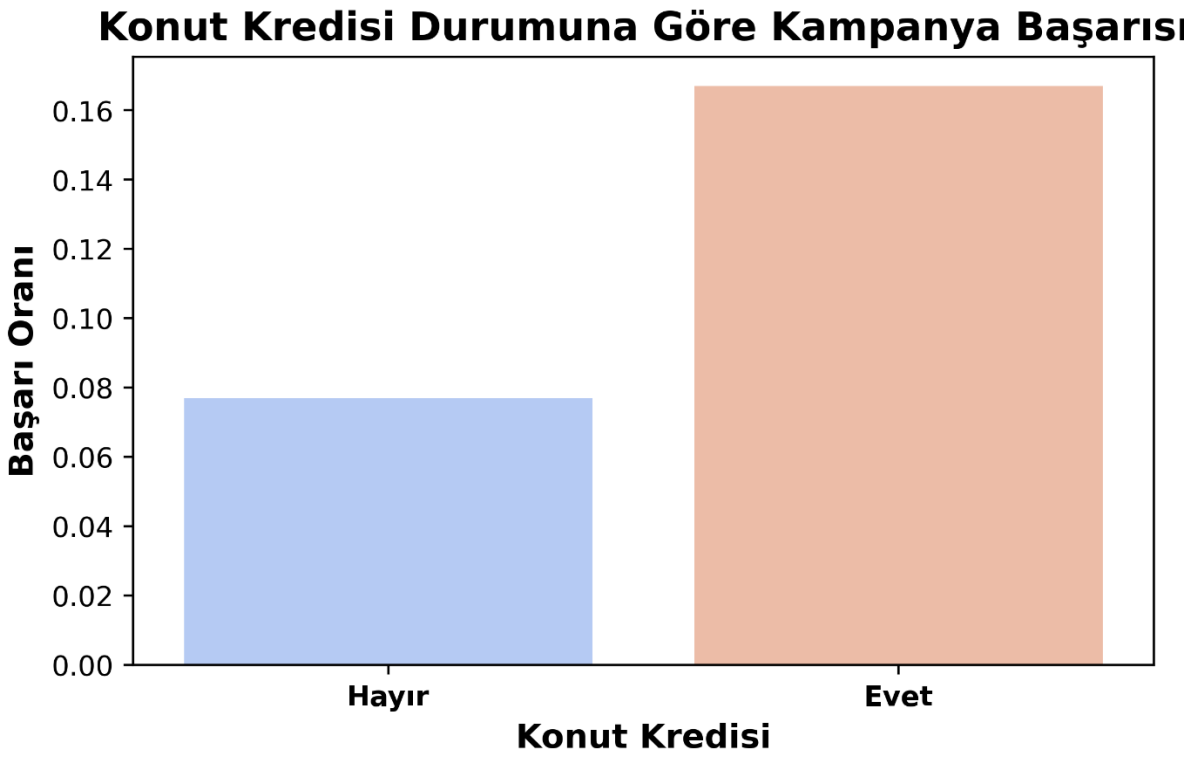


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 6'da önceki kampanya sonucuna göre başarı oranları verilmiştir. Müşterilerin daha önceki kampanyalardaki geri bildirimleri ile mevcut kampanyaya verdikleri yanıtlar arasındaki ilişki anlamlı düzeydedir. "Başarılı" (success) olarak kodlanan önceki kampanya deneyimine sahip bireylerin mevcut kampanyada abonelik gerçekleştirme oranı %64,73'tür. Buna karşılık, daha önce başarısız (failure) geri bildirim veren bireylerde ise bu oran yalnızca %12,61'dir. Önceki olumlu deneyimlerin müşteri güveni ve tekrar etkileşim eğilimini artırdığını gözlemlenmektedir. Davranışsal hedeflemenin kampanya başarısı açısından önemi ortaya konulmuştur.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

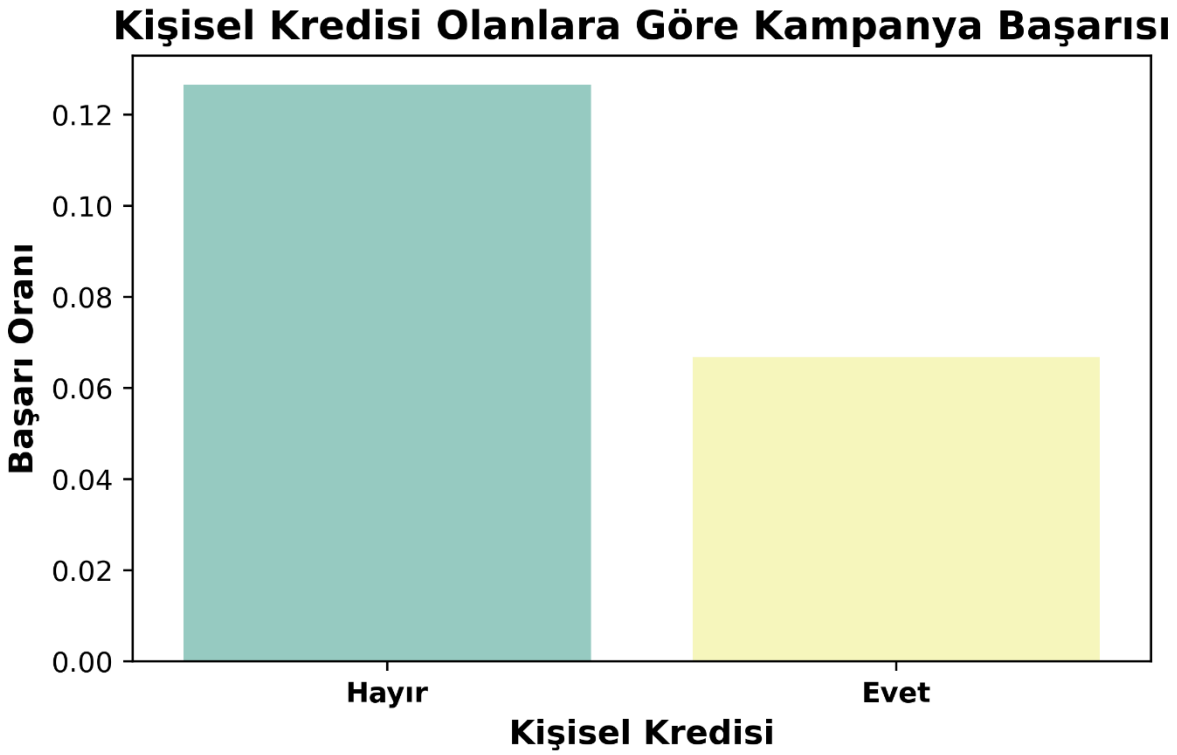
Şekil 7: Konut Kredisi Durumuna Göre Kampanya Başarısı



Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 7'de konut kredisi durumuna göre kampanya başarısı ifade edilmiştir. Konut kredisi olmayan bireylerde abonelik oranı %16,7 iken, konut kredisi bulunanlarda bu oran %7,7'ye düşmektedir. Geleneksel yaklaşımlarda kredi sahipliğinin müşteri ile banka arasındaki bağı güçlendirdiği varsayılmasına rağmen, bu bulgu ters yönde bir ilişkiye işaret etmektedir. Borç yükü altındaki bireylerin yeni finansal ürünlere yönelik daha temkinli bir tutum geliştirmesi, kampanya yanıtlarını olumsuz yönde etkilemiş olabilir.

Şekil 8: Kişisel Kredisi Olanlara Göre Kampanya Başarısı



Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 8'de kişisel kredisi olanlara göre kampanya başarısı görülmektedir. Kampanya başarısının kişisel kredi durumuna göre farklılaştığı gözlemlenmiştir. Krediye sahip olmayan bireylerde abonelik oranı %12,66 iken, kişisel kredi borcu bulunan bireylerde bu oran %6,68'e düşmektedir. Bu durum, bireysel borçluluğun yeni finansal tekliflere karşı isteksizlik yarattığını ve risk algısının müşteri davranışları üzerindeki etkisini ortaya koymaktadır. Literatürde de kredi yükü altındaki bireylerin tasarruf eğilimlerinin azaldığı ve riskten kaçınma davranışının arttığı yönünde bulgulara rastlanmaktadır.

5. Makine Öğrenmesi Modelleri

Makine öğrenmesi, bilgisayar sistemlerinin geçmiş verilere dayalı örüntüleri öğrenerek bu bilgiler doğrultusunda karar verebilme ve tahminde bulunabilme yetisi kazanmasını amaçlayan, istatistiksel ve hesaplamalı temellere dayanan bir yapay zekâ alt alanıdır (Eşidir, 2025c). Bu bağlamda, problemi tanımlayan veri kümesinin yapısal özellikleri dikkate alınarak uygun model seçimi ve optimizasyonu, makine öğrenmesi uygulamalarının başarısında belirleyici rol oynamaktadır.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

5.1. XGBoost (Aşırı Gradyan Artırma- Extreme Gradient Boosting) Makine Öğrenmesi Modeli

XGBoost, karar ağaçlarına dayalı gradyan artırma (gradient boosting) algoritmasının optimize edilmiş bir türevi olarak geliştirilmiştir. Model, art arda inşa edilen zayıf öğreniciler (çoğunlukla karar ağaçları) aracılığıyla her iterasyonda bir önceki öğrenicinin tahmin hatalarına odaklanmakta ve bu hataları minimize ederek nihayetinde güçlü bir sınıflandırıcı elde etmeyi hedeflemektedir (Aslan, 2025). Bu yaklaşım, öğrenme sürecini kademeli olarak iyileştirmekte ve yüksek doğruluk düzeyine ulaşılmasını sağlamaktadır. XGBoost, geniş veri setleriyle yüksek işlem verimliliği sunması, overfitting riskine karşı dayanıklı olması ve paralel işlemleme desteği sayesinde modern makine öğrenmesi uygulamalarında yaygın biçimde tercih edilmektedir (Chen & Guestrin, 2016). Öte yandan, bu model boosting algoritmaları ailesine dâhil bir topluluk (ensemble) öğrenme yöntemi olup; bagging, stacking gibi diğer topluluk tekniklerinden farklı olarak öğreniciler arası art arda hata düzeltmeye dayalı sıralı bir yapı izlemektedir (Wang vd., 2020). XGBoost'un öne çıkan avantajları arasında yüksek tahmin başarımı, parametrik esneklik, eksik veri ile çalışabilme yeteneği ve işlem maliyetinin optimize edilebilirliği yer almaktadır. Bu nitelikler sayesinde, alternatif modellere kıyasla sınıflandırma problemlerinde daha iyi genelleme kabiliyeti sunduğu ve sektörel uygulamalarda etkili sonuçlar verdiği çeşitli çalışmalarda ortaya konmuştur (Abar, 2020; Liang vd., 2020).

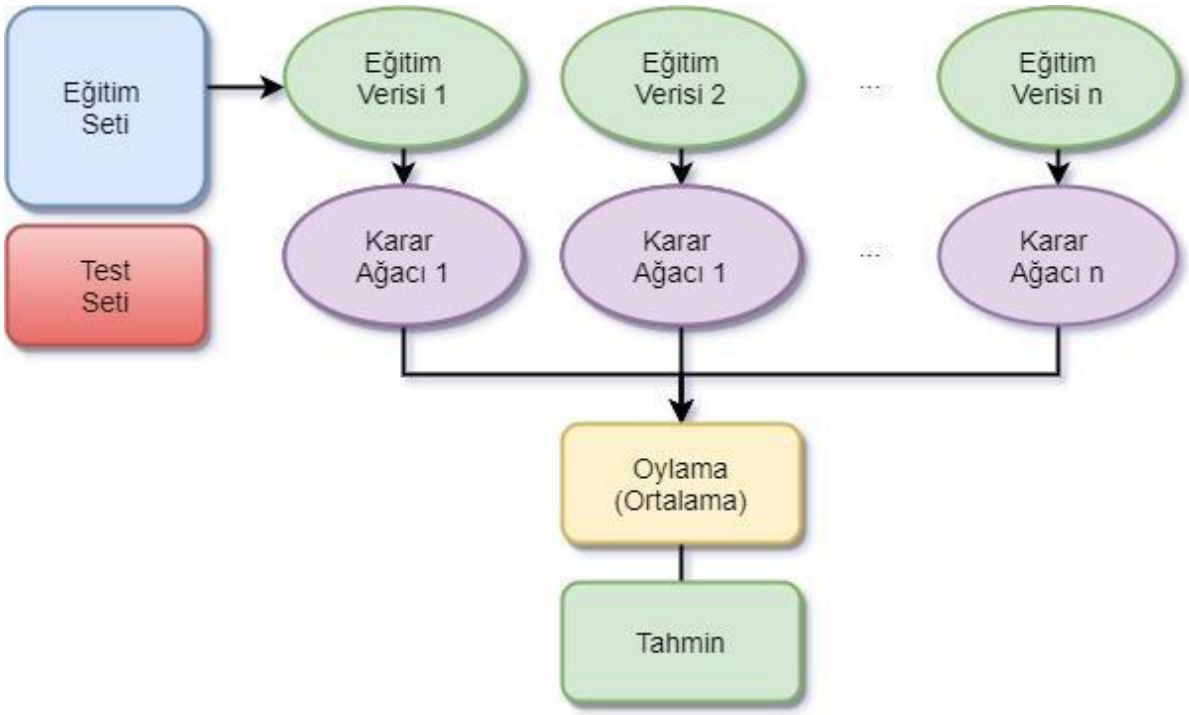
5.2. Rastgele Orman (Random Forest) Makine Öğrenmesi Modeli

Rastgele Orman, farklı alt veri setlerinde oluşturulan birden fazla karar ağacının birleşiminden meydana gelen, topluluk öğrenme (ensemble learning) metoduna dayalı güçlü bir makine öğrenmesi modelidir. Sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılmakta olup, birden fazla karar ağacını bir araya getirerek her bir ağacın bireysel zayıflıklarını dengelemekte ve daha kararlı, güçlü ve genelleştirilebilir bir model oluşturmaktadır. Bu model, bootstrap örnekleme yöntemiyle eğitim veri setinden birçok alt örneklem (subsample) oluşturur. Her bir bootstrap örneği, orijinal veri kümesinden rastgele seçilen örnekleri içermekte olup, bu örnekler kullanılarak bağımsız karar ağaçları inşa edilir. Bu süreçte, her karar ağacı veri setinin farklı bir alt kümesi üzerinde eğitilir. Ayrıca, her düğümdeki bölünme işlemi tüm özellikler yerine rastgele seçilen bir alt küme üzerinden gerçekleştirilir. Bu yöntem, modelin varyansını düşürerek aşırı öğrenmeyi (overfitting) önlemekte ve farklı karar ağaçları arasındaki korelasyonu azaltarak modelin genelleme performansını artırmaktadır. Sınıflandırma problemlerinde, nihai tahmin genellikle oylama (majority voting) yöntemiyle belirlenmektedir (Oguine ve Oguine, 2021). Algoritmanın şeması Şekil 9'da gösterilmiştir. Veri kümesinin sınıfını tahmin etmek için birden çok ağacı

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

birleştirildiğinden, bazı karar ağaçlarının doğru çıktıyı tahmin etmesi, bazılarının ise tahmin etmemesi mümkündür. Ancak birlikte, tüm ağaçlar doğru çıktıyı tahmin eder.

Şekil 9: Rastgele Orman Algoritması



Kaynak: Avcu, 2022.

6. Makine Öğrenmesi Modellerinin Performansları ve Bulgular

Bankacılık sektöründe yürütülen pazarlama kampanyalarına müşteri yanıtlarını tahmin etmek amacıyla, karar ağaçları tabanlı XGBoost ve Rastgele Orman modelleri kullanılmıştır. Bu modeller, veri setinde yer alan demografik özellikler, sosyoekonomik durumlar ve geçmiş kampanya etkileşimlerine ilişkin çok boyutlu değişkenleri dikkate alarak, müşterilerin mevduat aboneliği kararlarını sınıflandırmak üzere eğitilmiştir. XGBoost modeli, ardışık biçimde kurulan karar ağaçları üzerinden her bir modelleme adımında hata oranını minimize etmeyi amaçlayan gradyan artırma yaklaşımını esas almaktadır (Chen & Guestrin, 2016). Rastgele Orman modeli, çok sayıda bağımsız karar ağacından oluşan bir topluluk modeli olup, sınıflandırma kararını bu ağaçların çoğunluk oyuna göre vermktedir. Bu şekilde, modelin varyansını azaltılarak daha istikrarlı ve genellenebilir sonuçlar elde edilebilmektedir. Model performanslarının karşılaştırmalı değerlendirmesi Tablo 2'de sunulmuştur.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Tablo 2: Makine Öğrenmesi Modellerinin Performans Metrikleri

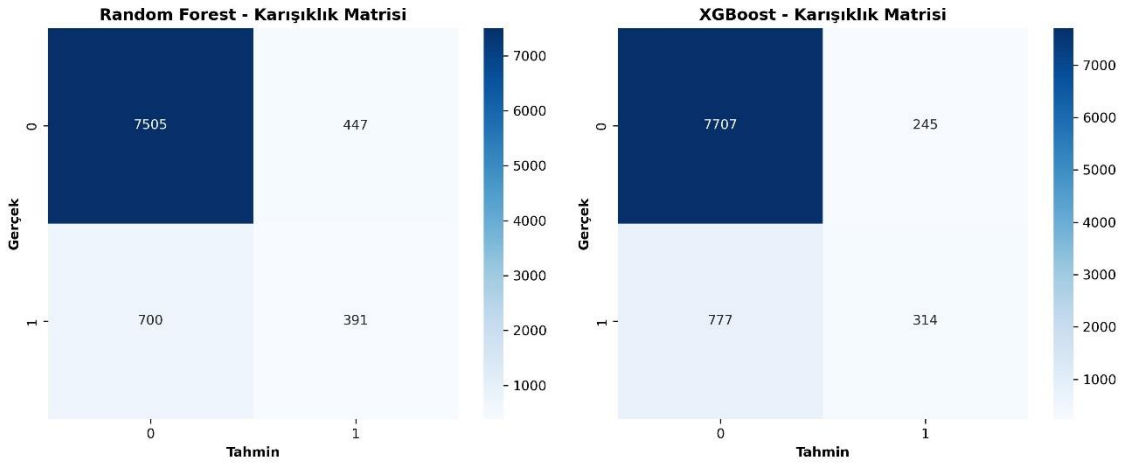
Modeller	Doğruluk (Eğitim)	Doğruluk (Test)	Ezberleme Farkı	Kesinlik	Duyarlılık	F1 Skoru	ROC-AUC
Rastgele Orman	0,918	0,873	0,045	0,467	0,358	0,405	0,747
XGBoost	0,942	0,887	0,055	0,562	0,288	0,381	0,761

Kaynak: Yazar tarafından Python ortamında hesaplanmıştır.

Tablo 2 incelendiğinde, her iki modelin hem eğitim hem de test setlerindeki yüksek doğruluk oranları (XGBoost için %94,2 ve %88,7; Rastgele Orman için %91,8 ve %87,3) modellerin genellenebilirliğinin güçlü olduğunu ve ezberleme riskinin düşük olduğunu göstermektedir. Ancak yüksek doğruluk değerleri, veri setindeki belirgin sınıf dengesizliği dikkate alındığında tek başına yeterli bir gösterge değildir. Pozitif sınıfa ilişkin başarım ölçütleri daha derinlemesine analiz edildiğinde, XGBoost modeli %56,2 kesinlik ile doğru pozitif tahminlerde Rastgele Orman'a (%46,7) kıyasla üstünlük sağlamaktadır. Ancak duyarlılık değeri açısından XGBoost (%28,8), Rastgele Orman'ın (%35,8) gerisinde kalmakta; bu durum, XGBoost'un pozitif sınıf örneklerini tespit etmede daha seçici ancak eksik kaldığını göstermektedir. F1 skorları incelendiğinde, Rastgele Orman %40,5 ile XGBoost'un (%38,1) üzerinde bir bütünleşik performans sunmaktadır. Buna rağmen, ROC-AUC değeri açısından XGBoost (%76,1), sınıflar arasında daha güçlü bir ayrım yapabilme yeteneği göstermektedir. Bu metrik, modelin pozitif ve negatif sınıfları ayırt etme kapasitesini ortaya koyması açısından kritiktir ve XGBoost'un genel ayrıştırma gücünün daha yüksek olduğunu işaret etmektedir.

Genel değerlendirmede, XGBoost modeli doğruluk ve ayrıştırma kapasitesi açısından daha başarılı iken, Rastgele Orman pozitif sınıfı daha dengeli tanımlamasıyla öne çıkmaktadır. Bu bulgular, sınıf dengesizliğinin var olduğu ortamlarda model performansının çok boyutlu metriklerle değerlendirilmesi gerektiğini açıkça ortaya koymaktadır.

Şekil 10: Karışıklık Matrisleri

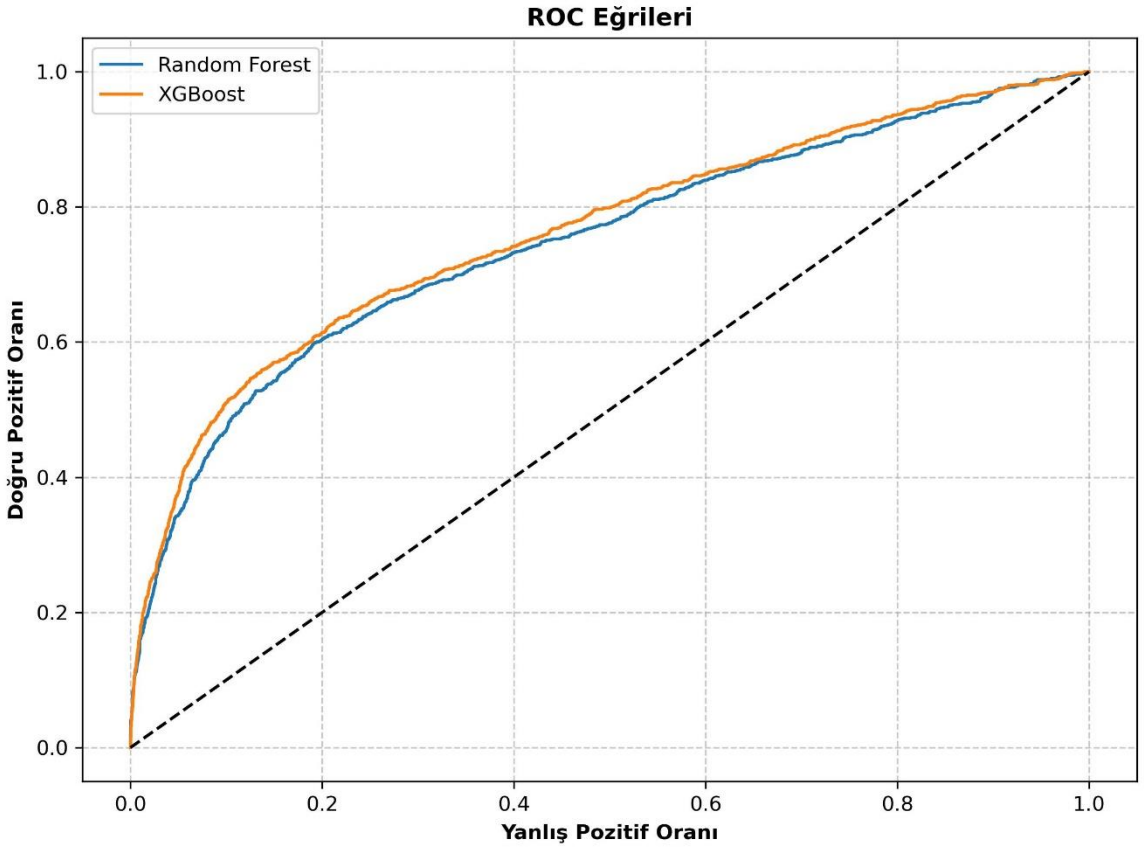


Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Giresun Üniversitesi İktisadi ve İdari Bilimler Dergisi
Giresun University Journal of Economics and Administrative Sciences
https://dergipark.org.tr/tr/pub/guiibd
DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

Şekil 10'da sunulan karışıklık matrislerine göre, XGBoost modeli negatif sınıfı daha doğru şekilde sınıflandırarak (7.707 doğru negatif) yüksek hassasiyet sağlamıştır. Ancak pozitif sınıfta yalnızca 314 doğru tahminle düşük bir duyarlılık sergilemiştir. Buna karşılık, Rastgele Orman modeli pozitif sınıfta 391 doğru tahminle daha yüksek duyarlılık sağlamış; ancak 447 adet yanlış pozitif üretmiştir. Bu durum, XGBoost'un yanlış alarm üretme olasılığını azalttığını, Rastgele Orman'ın ise pozitif sınıfı daha kapsayıcı biçimde tanımladığını ortaya koymaktadır. Dolayısıyla, model tercihi uygulamanın hedeflerine göre yapılmalıdır. Örneğin, yanlış pozitiflerin kritik olduğu alanlarda XGBoost tercih edilmeliyken, pozitif sınıfın kaçırılmasının riskli olduğu durumlarda Rastgele Orman daha uygun bir seçenek olarak öne çıkmaktadır (Tunca, 2025).

Şekil 11: ROC Eğrileri



Kaynak: Yazar tarafından Python kullanılarak Jupyter Notebook ortamında oluşturulmuştur.

Şekil 11'de sunulan ROC eğrileri, XGBoost ve Rastgele Orman modellerinin pozitif ve negatif sınıfları ayırt etme kapasitelerini karşılaştırmalı olarak ortaya koymaktadır. ROC eğrisi altında kalan alan (AUC) değeri bakımından değerlendirildiğinde, XGBoost modelinin

DOI: 10.46849/guiibd.1678452
GUİİBD-JEAS, 2025, 11(1): 111-133 ISSN: 2149-3391

eğrisinin daha üst bölgede konumlandığı görülmektedir. Bu durum, modelin hem yanlış pozitifleri düşük düzeyde tutarken hem de doğru pozitifleri daha yüksek doğrulukla tahmin etme becerisine sahip olduğunu göstermektedir. Diğer bir ifadeyle, XGBoost modeli, sınıflar arası ayrım gücü açısından Rastgele Orman modeline kıyasla daha yüksek bir ayırt edicilik sergilemiştir. Özellikle sınıf dengesizliğinin belirgin olduğu veri setlerinde, ROC-AUC değeri modellerin genel sınıflandırma başarımını yansıtmak açısından kritik bir metrik olarak öne çıkmaktadır. Bu bağlamda, XGBoost'un yüksek AUC performansı, model seçimi açısından anlamlı bir üstünlük sağlamaktadır (Oktay vd., 2023).

7. Sonuç ve Değerlendirme

Çalışmada, bankacılık sektöründe doğrudan pazarlama kampanyalarına verilen müşteri yanıtlarını tahmin etmek için, Rastgele Orman ve XGBoost modelleri ile analizler gerçekleştirilmiştir. Kaggle platformundan elde edilen Portekiz menşeli "Bank Marketing" isimli veri seti kullanılmıştır. Gerçekleştirilen modelleme sürecinde, müşterilerin demografik, sosyoekonomik ve geçmiş kampanya etkileşimlerine dair bilgileri temel alan çok değişkenli analizler yapılmıştır. Kampanya başarısının belirleyicileri görsel ve istatistiksel yöntemler ile açığa çıkarılmıştır.

Sonuçlar, her iki modelin de kampanya başarısını tahmin etmede güçlü yönlere sahip olduğunu göstermiştir. XGBoost modeli, %88,7 doğruluk ve %76,1 ROC-AUC değeri ile sınıflar arasında yüksek ayrıştırma başarısı sergilemiştir. %56,2 kesinlik ile doğru pozitif tahminlerde başarılı olan model, %28,8 duyarlılık oranıyla bazı pozitif sınıf örneklerini atlamıştır. Model seçici fakat dar kapsamlı tahminlerde bulunmaktadır. Rastgele Orman modeli ise %87,3 doğruluk ve %35,8 duyarlılık oranıyla pozitif sınıfı daha iyi bir biçimde tanımlamıştır. Ancak kesinlik değeri %46,7 ile XGBoost'un gerisinde kalmıştır ve yanlış pozitif tahminler daha yüksektir. F1 skoru açısından Rastgele Orman (%40,5), XGBoost'a (%38,1) kıyasla daha iyi bir performans sunmuştur.

Analiz bulguları, kampanya yanıtlarını etkileyen temel değişkenlerin başında müşterilerin geçmiş kampanya etkileşimlerinin geldiğini göstermektedir. Daha önce olumlu yanıt vermiş bireylerin yeniden abone olma eğiliminin yüksek olduğu belirlenmiştir. Bu durum, müşteri sadakati ve davranış sürekliliğinin kampanya başarısı üzerindeki kritik etkisine işaret etmektedir. 60 yaş ve üzeri bireyler kampanya tekliflerine daha duyarlıdır. Konut kredisi ve kişisel kredi yükü taşımayan müşterilerin ise kampanyalara daha olumlu yanıt verme eğiliminde oldukları saptanmıştır. Elde edilen bulgular, müşteri profiline dayalı kişiselleştirilmiş pazarlama stratejilerinin geliştirilmesinin gerekliliğini ortaya koymaktadır.

Genel olarak değerlendirildiğinde, çalışma her iki modelin de bankacılık kampanya başarısını tahminlemede etkin karar destek araçları olarak kullanılabileceğini göstermektedir.

Aynı zamanda sosyoekonomik ve davranışsal faktörlerin müşteri yanıtları üzerindeki etkileri çok boyutlu olarak analiz edilmektedir. İlerleyen çalışmalarda veri dengeleme tekniklerinin entegrasyonu, farklı modeller ile karşılaştırmalı analizlerin yapılması ve model çıktılarının SHAP gibi açıklayıcı yöntemlerle yorumlanması önerilmektedir. LightGBM ve CatBoost gibi diğer gradyan artırma modelleri ile karşılaştırmalı analizler yapılarak model çeşitliliği artırılabilir. Bu modellerin yüksek boyutlu ve dengesiz veri yapılarındaki performansları değerlendirilebilir.

Kaynakça

Abar, H. (2020). XGBoost ve Mars yöntemleriyle altın fiyatlarının kestirimi. *EKEV Akademi Dergisi, 83*, 427-446.

Aslan, K. (2025). Yapay Zekâ Makine Öğrenmesi ve Veri Bilimi Kursu, *Sınıfta Yapılan Örnekler ve Özet Notlar,* C ve Sistem Programcıları Derneği, İstanbul.

Avcu, F. M. (2022). Az Veri Setli Çalışmalarında Derin Öğrenme ve Diğer Sınıflandırma Algoritmalarının Karşılaştırılması: Agonist Ve Antagonist Ligand Örneği. *İnönü Üniversitesi Sağlık Hizmetleri Meslek Yüksek Okulu Dergisi,* 10(1), 356-371. https://doi.org/10.33715/inonusaglik.1022065

Aydın, G. & Yalçın, M. (2024). Banka Sektöründe Otomatik Makine Öğrenme Yöntemi (DataBlender) ile Vadeli Mevduat Talep Tahmini. In: Akoğul, S. & Tuna, E. (eds.), *Güncel Ekonometrik ve İstatistiksel Uygulamalar ile Akademik Çalışmalar.* Özgür Yayınları. DOI: https://doi.org/10.58830/ozgur.pub518.c2128

Calp, M. H. (2025). Predicting of Credit Card Customer Churn Using Machine Learning Methods, *Gazi Journal of Engineering Sciences,* vol.11, no.1, p p. 16-34, doi:10.30855/gmbd.0705N02

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Choi, Y., & Choi, J. W. (2023). Assessing the Predictive Performance of Machine Learning in Direct Marketing Response. *International Journal of E-Business Research*, 19(1), 1–12. https://doi.org/10.4018/IJEBR.321458

Davenport, T. H., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press.

Eşidir, K. A. (2025a). Makine Öğrenimi Modelleri ile Yetişkin Eğitimi Analizi: Modellerin Karşılaştırmalı Performansı. *Elektronik Sosyal Bilimler Dergisi,* 24(2), 946-964. https://doi.org/10.17755/esosder.1589887

Eşidir, K. A. (2025b). TÜİK Mikro Verileri ile Çocuk İşgücü Tahmini: Makine Öğrenimi Modellerinin Performans Analizleri. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 37(1), 453-466. https://doi.org/10.35234/fumbd.1607609

Eşidir, K. A. (2025c). Türkiye'nin Kimyasal Madde İthalatının Gelecek Tahmini: Makine Öğrenmesi ve Topluluk Öğrenme Yöntemleri Performans Analizi. *Firat University Journal of Social Sciences,* 35(1), 261-278. https://doi.org/10.18069/firatsbed.1580620

Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 6(4), 1-21.

Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, 43(2), 276–286. https://doi.org/10.1509/jmkr.43.2.276

Liang, W., Luo, S., Zhao, G., & Wu, H. (2020). Predicting hard rock pillar stability using gbdt, XGBoost, and LightGBM algorithms. *Mathematics,* 8(5), 765. https://doi.org/10.3390/math8050765

Moro, S., Rita, P., & Cortez, P. (2014). Bank Marketing [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5K306.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36(2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Oguine OC ve Oguine MB. (2021). Comparative analysis and forecasting on the death rate of COVID-19 patients in Nigeria using Random Forest and Multinomial Bayesian epidemiological models, *Journal of Clinical Case Studies Reviews & Reports,* 1-7.

Oktay, S., Bakır, H., & Tabaru, T. E. (2023). Makine öğrenmesi teknikleri kullanılarak bankalardaki potansiyel müşterilerin sınıflandırılması. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 12(2), 22–41.

Speer AB. (2021). Empirical attrition modelling and discrimination: Balancing validity and group differences, *Human Resource Management Journal*, 34 (1), 1-19.

Tafralı, S. (2022). *Veri Biliminde Özellik Ölçeklendirme (Feature Scaling)*. Medium. Erişildiği adres: https://serdartafrali.medium.com/veri-biliminde-%C3%B6zellik-%C3%B6l%C3%A7eklendirme-feature-scaling-e05d9f3e96ce

Tunca, S. (2025). Bankacılıkta Müşteri Verilerini Anlama: CRISP-DM Yaklaşımı ve Makine Öğrenimi Uygulamaları. *AJIT-E: Academic Journal of Information Technology,* 16(1), 68-87. https://doi.org/10.5824/ajite.2025.01.004.x

Wang, L., Wang, X., Chen, A., Jin, X., & Che, H. (2020). Prediction of type 2 diabetes risk and its effect evaluation based on the XGBoost model. *Healthcare,* 8(3), 247. https://doi.org/10.3390/healthcare8030247

Wu Y. (2023). Job embeddedness review: Presentation, measurement and development. *Advances in Economics, Management and Political Sciences,* 47 (1), 169- 174.

Zhu X, Sawhney R, Upreti G. (2016). Determinates of employee voluntary turnover and forecasting in departments: A case study, *Studies in Engineering and Technology*, 3(1):64-73.