Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 9, S. 177-183, 8, 2025

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 9, pp. 177-183, 8, 2025

Research Article

Makine Öğrenmesi Tabanlı Potansiyel Müşteri Tahmini: Tasarruf Finansmanı Sektöründen Bir Uygulama

Cengiz SERTKAYA[1*]

[1] *BT Uygulama Çözümleri Bölümü, Eminevim, Türkiye*

[*] *(cengiz.sertkaya@eminevim.com.tr)*

(Received: 11 August 2025, Accepted: 16 August 2025)

(6th International Conference on Engineering, Natural and Social Sciences ICENSOS 2025, August 10-11, 2025)

ATIF/REFERENCE: Sertkaya, C. (2025). Makine Öğrenmesi Tabanlı Potansiyel Müşteri Tahmini: Tasarruf Finansmanı Sektöründen Bir Uygulama, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(8), 177-183.

Özet **–** Finansal hizmetler sektörü, dijital dönüşüm ve veri temelli karar destek sistemlerinin yaygınlaşmasıyla birlikte hızlı bir evrim geçirmektedir. Bu dönüşüm, müşteri ilişkileri yönetiminden pazarlama stratejilerine kadar pek çok alanda yapay zekâ ve makine öğrenimi tabanlı yaklaşımların önemini artırmıştır. Türkiye'de son yıllarda büyüme gösteren tasarruf finansmanı sektörü de bu teknolojik gelişmelere adapte olarak, müşteri kazanım süreçlerini daha etkin yönetme ihtiyacı duymaktadır. Mevcut müşteri kayıtlarından (kişi kartları) elde edilen demografik, sosyoekonomik ve davranışsal verilerin analizi, potansiyel müşteri öngörüsü açısından önemli fırsatlar sunmaktadır.

Bu çalışmada, tasarruf finansmanı sektöründe faaliyet gösteren bir firmanın mevcut müşteri verilerine dayanarak potansiyel müşterileri tahmin edebilecek bir sınıflandırma modeli geliştirilmiştir. Bu amaçla, makine öğrenimi algoritmalarından Random Forest yöntemi tercih edilmiştir. Random Forest, yüksek doğruluk oranı, değişken önem düzeylerini belirleyebilme yeteneği ve aşırı öğrenmeye karşı dayanıklılığı sayesinde müşteri tahmin modellerinde etkin bir çözüm sunmaktadır. Çalışma sonucunda geliştirilen modelin, hedefli pazarlama faaliyetlerinde kullanılabilir olduğu ve sektörün dijitalleşme sürecine katkı sağlayacağı düşünülmektedir.

Anahtar Kelimeler: Tasarruf Finansmanı, Müşteri Veri Analizi, Potansiyel Müşteri Tahmini, Yapay Zekâ, Random Forest.

I. GİRİŞ

Gelişen bilgi teknolojileri, finansal hizmetler sektörünü derinden dönüştürmekte ve veri odaklı karar alma süreçlerini hızla yaygınlaştırmaktadır. Özellikle müşteri ilişkileri yönetimi (CRM), risk analizi, satış tahmini ve pazarlama stratejilerinin belirlenmesinde yapay zekâ (YZ) ve makine öğrenimi (MÖ) tekniklerinin kullanımı büyük bir artış göstermiştir[1], [2]. Bu gelişmeler, sadece geleneksel bankacılık ve sigortacılık alanlarında değil, alternatif finans sistemlerinde de müşteri yönetim yaklaşımlarını yeniden şekillendirmektedir.

Tasarruf finansmanı sektörü, Türkiye gibi gelişmekte olan ülkelerde faizsiz, katılım esaslı yapısıyla önemli bir finansal çözüm alternatifi sunmaktadır. Eminevim gibi bu sektörde faaliyet gösteren firmalar, özellikle konut ve araç edinimine yönelik çözümleriyle geniş bir müşteri kitlesine ulaşmaktadır[3], [4]. Ancak sektörün hızlı büyümesi, müşteri kazanımı süreçlerinin daha verimli, hedefli ve veri temelli bir yapıya kavuşmasını gerekli kılmaktadır. Bu bağlamda, mevcut müşteri kartlarından elde edilen bilgilerin analiziyle potansiyel müşterilerin tespit edilmesi, stratejik bir rekabet avantajı sunabilir [5].

Literatürde, müşteri davranışlarını anlamaya ve sınıflandırmaya yönelik pek çok yapay zekâ temelli yaklaşım önerilmiştir. Makine öğrenimi teknikleri bu yaklaşımların merkezindedir. Sıklıkla kullanılan algoritmalar arasında karar ağaçları (decision trees), destek vektör makineleri (SVM), yapay sinir ağları (ANN), k-en yakın komşu (k-NN), kümeleme algoritmaları (K-Means, DBSCAN) ve ansambl öğrenme yöntemleri yer almaktadır[6],[7]. Bu yöntemler sayesinde büyük hacimli müşteri verileri üzerinde örüntüler tespit edilmekte, segmentler oluşturulmakta ve potansiyel müşteri davranışları yüksek doğrulukla tahmin edilebilmektedir[8].

Özellikle müşteri segmentasyonu, müşteri sadakati analizi ve potansiyel müşteri tahmini gibi konularda yapay zekâ uygulamaları önemli avantajlar sağlamaktadır. Bir sistematik derleme çalışmasında, müşteri davranışlarını analiz etmek için MÖ tekniklerinin bankacılıkta yaygınlaştığını ve müşteri yaşam boyu değerinin hesaplanmasında etkili olduğunu göstermiştir[9]. Benzer şekilde, bir çalışmada, veri odaklı karar verme süreçlerinin, müşteri edinim maliyetlerini düşürdüğünü ve dijital pazarlama kampanyalarının dönüşüm oranlarını artırdığını belirtmektedir[10].

Bununla birlikte, müşteri kazanımında kullanılan veri türlerinin çeşitliliği ve karmaşıklığı, klasik istatistiksel modellerin yetersiz kaldığı durumlarda yapay zekânın önemini artırmaktadır. Demografik, sosyoekonomik ve davranışsal verilerin aynı anda değerlendirilmesi ve aralarındaki ilişkilerin öğrenilmesi için çok katmanlı yapay sinir ağları veya hibrit modeller kullanılmaktadır[11]. Örneğin, önceki katılımcıların ödeme davranışları, finansman tercihleri, bölgesel dağılımı gibi veriler incelenerek, benzer eğilimleri gösterecek bireyler önceden tahmin edilebilmektedir[12].

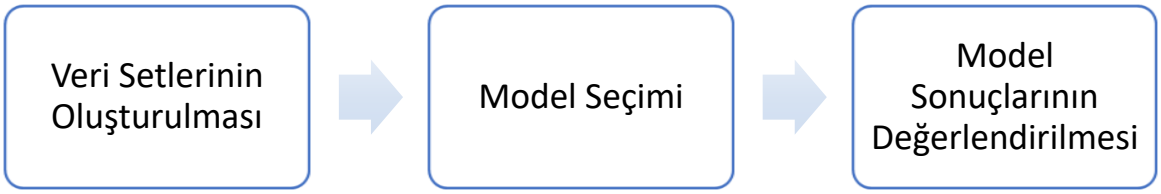
Yapay zekâ temelli tahmin sistemleri yalnızca öngörüde bulunmakla kalmaz, aynı zamanda hangi değişkenlerin (örneğin yaş, gelir, medeni durum) müşteri kararlarını daha fazla etkilediğini ortaya koyarak, firmalara stratejik içgörüler sunar[13]. Bu tür analizler, saha ekiplerinin kaynaklarını verimli kullanmasını, pazarlama mesajlarının kişiselleştirilmesini ve ürünlerin doğru hedef kitleye yönlendirilmesini sağlar[14].

Türkiye'deki tasarruf finansmanı sektörü, dijital dönüşüm açısından henüz başlangıç aşamasındadır. Sektöre dair yapılan çalışmalar, çoğunlukla yasal düzenlemeler veya finansal yapı analizleriyle sınırlıdır[15]. Dolayısıyla yapay zekâ temelli veri analitiği yöntemlerinin bu sektöre uygulanması hem akademik literatür hem de sektör uygulamaları açısından önemli bir boşluğu dolduracağı düşünülmektedir.

Bu çalışma, yapay zekâ temelli veri analizi yaklaşımlarının, tasarruf finansmanı sektöründe müşteri öngörüsü amacıyla uygulanabilirliğini test etmeyi hedeflemektedir. Mevcut müşteri kartlarındaki veriler kullanılarak benzer profildeki potansiyel müşterilerin tahmin edilmesi ve bu doğrultuda pazarlama stratejilerinin yönlendirilmesi amaçlanmaktadır. Bu yaklaşım, hem veri temelli yönetim kültürünün gelişmesine katkı sağlayacak hem de sektörün müşteri kazanım süreçlerini optimize etmesine olanak tanıyacaktır.

II. MATERYAL VE YÖNTEM

Bu çalışmada, tasarruf finansmanı sektöründe faaliyet gösteren bir firmanın kayıtlı müşteri verileri kullanılarak, potansiyel müşterilerin tahminine yönelik bir makine öğrenimi modeli geliştirilmiştir. Bu amaçla uygulanan temel adımlar veri setlerinin oluşturulması, yapay zeka algoritma seçimi ve değerlendirme methodolojisinin seçimi gibi adımlardan oluşmaktadır. Bu adımlar Şekil 1'de gösterilmektedir.



Şekil 1. Model Geliştirme Adımları

A. Veri Setlerinin Oluşturulması

Verisetinin oluşturulması aşamasındaki analiz süreçleri; veri toplama, ön işleme, özellik seçimi ve nihayi veri setlerinin oluşturulması adımlarına dayanmaktadır. Veri seti, bir tasarruf finansmanı şirketinin müşteri bilgi sistemi üzerinden temin edilen, anonimleştirilmiş kişi kartı verilerinden oluşmaktadır. Söz konusu veriler; yaş, cinsiyet, medeni durum, eğitim durumu, meslek, gelir aralığı, ikamet ili, başvuru kanalı, kampanya tercihi, taksit tutarı, vade süresi gibi niteliksel ve niceliksel değişkenleri içermektedir. Veri seti, mevcut kayıtlı müşterilere ait yaklaşık 800.000 gözlem içermektedir ve veri gizliliği ilkeleri doğrultusunda kimlik bilgileri çalışmaya dâhil edilmemiştir

Makine öğrenimi modellerinde başarının en belirleyici aşamalarından biri, veri ön işleme sürecidir. Ham veriler sıklıkla eksiklik, tutarsızlık ve yorumlanabilirlik açısından sorunlar içerebilir. Bu nedenle, model eğitiminden önce uygulanan ön işleme adımları, modelin genelleme gücünü doğrudan etkilemektedir[16]. Bu çalışmada, veri ön işleme süreci üç temel başlık altında yürütülmüştür: sayısallaştırma, çatı değer analizi ile eksik veri analizine dayalı alan çıkarımı şeklindedir. Veri ön işleme sürecinde uygulanan adımlar Tablo 1'de gösterilmektedir.

Tablo 1. Veri Ön İşleme Adımları

Adım	Süreç Adı	Açıklama	Kullanılan Yöntem/İşlem
1	Sayısallaştırma	Kategorik verilerin modele uygun hale getirilmesi	Label Encoding, One-Hot Encoding
2	Aykırı Değer Temizleme	Gerçek dışı değerlerin veri setinden çıkarılması	IQR yöntemi, Mantıksal kontrol
3	Eksik Veri Analizi	Yetersiz/gürültülü veri alanlarının çıkarılması	Eksik veri oranı > %40, sistematik analiz

Veri setindeki birçok alan, niteliksel (kategorik) özellikler içermektedir. Bu tür değişkenlerin doğrudan bir makine öğrenimi algoritmasına girdi olarak verilmesi mümkün değildir. Bu nedenle, etiketleme (label encoding) ve ikili vektörleme (one-hot encoding) gibi sayısallaştırma teknikleri uygulanmıştır. Örneğin, "medeni durum" veya "eğitim seviyesi" gibi değişkenler, sıralı kategorik yapıda oldukları için etiketleme yöntemiyle sayısallaştırılmıştır. Buna karşın, sıralı olmayan "şehir", "kampanya türü" gibi değişkenler one-hot encoding yöntemiyle vektörleştirilmiştir[17]. Bu yaklaşım, algoritmanın bu kategorileri birbirine üstünlük sırası atfetmeden işleyebilmesini sağlamaktadır.

Veri setinde yer alan sayısal ve kategorik değişkenlerde, analiz öncesi aşamada anormal veya gerçek dışı değerlerin tespiti ve temizlenmesi önem taşımaktadır. Bu çalışmada, özellikle sayısal değişkenlerde gözlemlenen çatı değerler (outliers), model başarısını olumsuz etkileyebileceği için çıkarılmıştır. Aykırı değerlerin belirlenmesinde interquartile range (IQR) yöntemi uygulanmış ve 1.5 * IQR sınırlarının dışındaki gözlemler anormal değer olarak kabul edilmiştir[18]. Ayrıca, kategorik değişkenlerde mantıksal tutarsızlık içeren ve kayıt hatası olarak değerlendirilebilecek gerçek dışı değerler de temizlenmiştir. Örneğin, yaş değişkeninde "150" gibi biyolojik olarak mümkün olmayan değerler çıkarılmıştır. Bu işlemler, modelin aşırı sapmalardan etkilenmeden daha sağlıklı öğrenmesini sağlamıştır[19].

Bazı alanlarda eksik veri oranı %40'ın üzerine çıkmıştır. Özellikle geçmişe yönelik bilgi içeren "referans bilgileri" ve "kişisel not" gibi nitelikler, gözlemleme sıklığının düşüklüğü nedeniyle modellemede anlamlı

katkı sunamayacak düzeydedir. Bu tür alanlar, modelin performansını düşürebileceği, aşırı öğrenmeye (overfitting) neden olabileceği ve gürültü yaratabileceği için analiz dışı bırakılmıştır. Eksik verilerin sistematik olup olmadığı kontrol edilmiş, rastgele eksiklik taşıyan alanlar çıkarılırken herhangi bir önyargı (bias) yaratılmamasına özen gösterilmiştir[20]. Veri kalitesi açısından düşük olan bu değişkenlerin çıkarılması, modelin daha dengeli bir şekilde eğitilmesini sağlamıştır.

Veri ön işleme aşamasının son adımı olarak, işlenmiş veri seti, modelin öğrenme ve değerlendirme aşamalarında kullanılması için ikiye ayrılmıştır. Veri setinin %70'i eğitim (training), %30'u ise test (testing) veri seti olarak belirlenmiştir. Bu yöntem, modelin genel performansının değerlendirilmesi ve aşırı öğrenmenin önüne geçilmesi açısından yaygın olarak kullanılmaktadır[21]. Model, eğitim verisi üzerinde öğrenirken test verisi ile gerçek dünya performansı ölçülmüştür.

B. Model Seçimi

Veri setlerinin hazırlanmasının ardından yapay zeka model çalışmalarına başlanmıştır. Bu aşamada, Leo Breiman tarafından önerilen, ansambl öğrenme yöntemlerinden biri olan güçlü bir sınıflandırma ve regresyon algoritması olarak literatürde geçen Random Forest algoritması kullanılmıştır[22]. Temel mantığı, çok sayıda karar ağacının (decision tree) bağımsız olarak eğitilmesi ve bu ağaçların toplu kararının çoğunluk oyuyla sınıflandırma yapılmasıdır. Her bir karar ağacı, eğitim verisinin farklı bir rastgele alt kümesi (bootstrap örneklemesi) kullanılarak oluşturulur ve her düğümde sadece rastgele seçilen özellik alt kümesi üzerinden bölünmektedir. Bu yöntem, ağaçlar arasındaki korelasyonu azaltarak modelin genelleme yeteneğini artırmakta ve aşırı öğrenme riskini minimize etmektedir[23].

Random Forest algoritmasının tercih edilmesindeki önemli nedenler arasında gürültülü veri ve eksik değerler karşısında dayanıklılık göstermesi, değişken önemini ölçme imkânı sağlaması ve çok yüksek boyutlu veri setlerinde iyi performans göstermesi şeklinde ifade edilebilir[22][24].

Modelin uygulanmasında, New Zealand Üniversitesi tarafından geliştirilen ve makine öğrenimi algoritmalarını kolayca uygulamak için kullanılan açık kaynaklı bir yazılım paketi olan WEKA uygulaması kullanılmıştır[25]. Java tabanlı olan WEKA, kullanıcı dostu grafik arayüzü ve komut satırı arayüzü ile veri madenciliği süreçlerinin hızlı prototiplenmesini sağlamaktadır. Sınıflandırma, regresyon, kümeleme, özellik seçimi gibi birçok makine öğrenimi algoritmasını bünyesinde barındırmaktadır.

Ayrıca WEKA'nın Random Forest uygulaması, kullanıcıların model parametrelerini kolayca ayarlamasına olanak tanımaktadır. Bunlar arasında ağaç sayısı (numTrees), her ağacın maksimum derinliği (maxDepth), rastgele özellik sayısı (numFeatures) ve bootstrap örneklemesi seçenekleri bulunmaktadır. Ayrıca, model performansı farklı çapraz doğrulama yöntemleriyle test edilebilmektedir.

Random Forest modelinin uygulanmasında kullanılan model parametreleri Tablo 2'de gösterilmektedir.

Tablo 2. Random Forest Model Parametreleri

Parametre Adı	Değer	Açıklama
numTrees	100	Oluşturulacak toplam karar ağacı sayısı
maxDepth	20	Her ağacın maksimum derinliği
numFeatures	Otomatik (√n)	Her düğümde değerlendirilen özellik sayısı
seed	1	Rastgelelik için kullanılan başlangıç değeri
bagSizePercent	100	Bootstrap örneklemesi yüzdesi
crossValidationFold	10	10 katlı çapraz doğrulama ile test

C. Model Sonuçlarının Değerlendirilmesi

Makine öğrenimi modellerinin performanslarını güvenilir şekilde değerlendirmek, modelin pratikteki geçerliliğini ve genelleme yeteneğini ortaya koymak açısından büyük önem taşımaktadır. Bu kapsamda,

sınıflandırma problemlerinde yaygın olarak kabul görmüş değerlendirme metriklerinden yararlanılmıştır. Bu amaçla Random Forest algoritması ile geliştirilen modelin, test verisi üzerindeki performansı çeşitli istatistiksel ölçütler ile değerlendirilmiş ve sonuçlar çapraz doğrulama ile desteklenmiştir. Kullanılan metrikler sırasıyla doğruluk (accuracy), kesinlik (precision) ve duyarlılık (recall) metrikleridir. Bu metriklerin hesaplanmasında sırasıyla; TP (True Positive): Pozitif olan ve doğru tahmin edilenlerin, TN (True Negative): Negatif olan ve doğru tahmin edilenlerin, FP (False Positive): Negatif olan ama pozitif tahmin edilenlerin ve FN (False Negative): Pozitif olan ama negatif tahmin edilenlerin sayısını ifade eden değerler kullanılmaktadır.

Doğruluk, modelin doğru tahmin ettiği gözlem sayısının, toplam gözlem sayısına oranı olarak Denklem 1'de gösterildiği şekilde tanımlanmaktadır[26].

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Kesinlik, modelin pozitif olarak sınıflandırdığı örneklerden ne kadarının gerçekten pozitif olduğunu ölçmektedir ve Denklem 2'de gösterildiği şekilde tanımlanmaktadır[27].

$$Kesinlik = \frac{TP}{TP+FP} \quad (2)$$

Duyarlılık, modelin gerçek pozitifleri ne ölçüde doğru tahmin ettiğini gösteren bir ölçü olarak Denklem 3'de gösterildiği şekilde tanımlanmaktadır[17].

$$Duyarlılık = \frac{TP}{TP+FN} \quad (3)$$

III. BULGULAR

Random Forest algoritmasıyla geliştirilen sınıflandırma modeli, test verisi üzerinde %94 doğruluk, %91 precision ve %93 recall gibi yüksek başarı ölçütlerine ulaşmıştır. Bu bulgular, Random Forest'ın sınıflandırma görevlerinde üstün başarısını ortaya koyan güncel çalışmalarla örtüşmektedir. Örneğin, yapılan bir çalışmada, e-ticaret müşterilerinin churn davranışlarını tahmin etmek üzere geliştirilen modelde Random Forest %95'e yakın doğruluk oranı ile en iyi performansı göstermiştir [28]. Aynı şekilde, finansal teknoloji alanında yapılan bir çalışmada, Random Forest'ın SVM ve Naive Bayes'e kıyasla daha yüksek F1 skoru elde ettiği ortaya koyulmuştur [29].

Modelin başarısı, eğitim ve test setleri arasında anlamlı bir performans farkının olmamasıyla da desteklenmektedir. Bu çalışmada, aşırı öğrenmenin (overfitting) önüne geçildiğini ve modelin genelleme yeteneğinin güçlü olduğunu göstermektedir [23], [22].

Random Forest algoritmasının sağladığı "özellik önem derecesi" analizi sonucunda, potansiyel müşterileri tahmin etmede en etkili değişkenlerin sırasıyla gelir seviyesi, yaş, ikamet ili, taksit tutarı, meslek grubu ve kampanya tercihi olduğu belirlenmiştir. Bu bulgular, literatürde yapılan ve demografik–davranışsal verilerin kredi başvuru skorlamasında kullanımını inceleyen çalışmayla paralellik göstermektedir [30].

Ayrıca, yapılan bir çalışmada da eğitim düzeyi, ödeme gücü ve başvuru kanalı gibi özniteliklerin müşteri dönüşüm oranları üzerinde anlamlı etkisi olduğu vurgulanmıştır. Bu bağlamda, tasarruf finansmanı sektöründe de benzer örüntülerin gözlemlenmesi, modelin iç tutarlılığını ve alana özgü uygulanabilirliğini desteklemektedir[31].

IV. SONUÇLAR

Bu çalışma, Türkiye'de faaliyet gösteren bir tasarruf finansmanı firmasının müşteri veri seti kullanılarak, potansiyel müşteri tahmini amacıyla geliştirilen Random Forest modelinin hem teknik başarı düzeyini hem de stratejik faydasını ortaya koymaktadır. Geliştirilen modelin doğruluk, precision ve recall gibi temel sınıflandırma ölçütlerinde yüksek performans göstermesi, modelin saha uygulamalarına uygun olduğunu göstermektedir.

Literatürdeki çalışmalarda, müşteri kazanımında veri temelli tahmin sistemlerinin pazarlama etkinliğini %20'ye varan oranlarda artırabildiğini ortaya koyulmuştur. Bu da çalışmanın pratik yansımalarının yalnızca akademik değil, operasyonel olarak da değerli olduğunu göstermektedir[32], [33].

Müşteri tahmini gibi kritik kararlarda modelin şeffaflığı önem arz etmektedir. Bu anlamda ileriki çalışmalarda, SHAP (SHapley Additive exPlanations) veya LIME gibi açıklanabilir yapay zekâ araçlarıyla, hangi değişkenin kararı ne ölçüde etkilediği görselleştirilebilir. Bu yaklaşım, hem regülasyon uyumluluğu hem de pazarlama ekiplerinin model sonuçlarını yorumlaması açısından faydalı olacaktır[34].

Ayrıca ileriki çalışmalarda, Random Forest dışında, XGBoost, LightGBM ve yapay sinir ağları gibi farklı modellerle karşılaştırmalı analizler yapılarak en uygun algoritma belirlenebilir. Bir çalışmada, XGBoost algoritmasının yüksek veri hacminde Random Forest'a kıyasla daha kısa sürede benzer doğruluk oranı sunabileceği önerilmektedir[35].

Ayrıca, müşteri davranışlarının zamansal değişimi dikkate alınarak, örneğin ödeme düzeni veya başvuru tarihleri gibi zaman serisi verileriyle geliştirilmiş yeni modeller oluşturulabilir. Bunun, özellikle churn veya erken dönem müşteri tahmini için etkili olabileceği literatürde belirtilmiştir [36].

Bunun dışında, segmentasyon ile farklı müşteri kümelerine göre ayrı stratejiler geliştirilebilir. Özellikle gelir seviyesi, yaş grubu veya kampanya tercihi gibi değişkenler üzerinden segment temelli kişiselleştirilmiş öneri sistemleri kurulması literatürde önerilmektedir[14] [13].

KAYNAKLAR

[1] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," Expert Syst. Appl., vol. 36, no. 2, pp. 2592–2602, 2009.

[2] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4th ed. Morgan Kaufmann, 2016.

[3] M. A. Yülek, "Türkiye'de Alternatif Finansal Sistemler: Katılım Bankacılığı ve Tasarruf Finansman Modelleri," Finans. Araştırmalar ve Çalışmalar Derg., vol. 12, no. 23, pp. 45–62, 2020.

[4] S. G. Akın, "Türkiye'de Alternatif Finansal Modellerin Gelişimi: Katılım Finans ve Tasarruf Finansman Şirketleri," Finans. Araştırmalar ve Çalışmalar Derg., vol. 13, no. 25, pp. 55–74, 2021.

[5] V. Kumar and A. Petersen, Statistical Methods in Customer Relationship Management. Hoboken, NJ: John Wiley \& Sons, 2012.

[6] B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, "Benchmarking state-of-the-art classification algorithms for credit scoring," J. Oper. Res. Soc., vol. 54, no. 6, pp. 627–635, 2009.

[7] K. Tsiptsis and A. Chorianopoulos, Data Mining Techniques in CRM: Inside Customer Segmentation. Wiley, 2009.

[8] P. C. Verhoef, W. J. Reinartz, and M. Krafft, "Customer Engagement as a New Perspective in Customer Management," J. Serv. Res., vol. 13, no. 3, pp. 247–252, 2010.

[9] S. Chatterjee, N. P. Rana, K. Tamilmani, and A. Sharma, "Exploring the roles of customers in sustainable banking: A systematic literature review," Int. J. Bank Mark., vol. 37, no. 6, pp. 1361–1389, 2019.

[10] E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton \& Company, 2014.

[11] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras \& TensorFlow, 2nd ed. O'Reilly Media, 2019.

[12] A. Verikas, A. Gelzinis, and M. Bacauskiene, "Mining data with random forests: A survey and results of new tests," Pattern Recognit., vol. 44, no. 2, pp. 330–349, 2011.

[13] V. Kotu and B. Deshpande, Data Science: Concepts and Practice, 2nd ed. Morgan Kaufmann, 2019.

[14] P. Kotler and K. L. Keller, Marketing Management, 15th ed. Pearson, 2015.

[15] Ö. Küçük and G. Özer, “Katılım Bankacılığı ve Tasarruf Finansman Şirketleri: Yeni Finansal Araçlara İhtiyacın Bir Yansıması,” Katılım Ekon. ve Finans Derg., vol. 7, no. 1, pp. 55–68, 2021.
[16] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, “Data Preprocessing for Supervised Learning,” Int. J. Comput. Sci., vol. 1, no. 2, pp. 111–117, 2006.
[17] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011.
[18] S. Chawla and A. Gionis, “k-means--: A unified approach to clustering and outlier detection,” in Proceedings of the 2013 SIAM International Conference on Data Mining, 2013, pp. 189–197.
[19] C. C. Aggarwal, Outlier Analysis. Springer, 2017.
[20] P. D. Allison, Missing Data. Sage Publications, 2001.
[21] P. Refaeilzadeh, L. Tang, and H. Liu, “Cross-Validation,” in Encyclopedia of Database Systems, Springer, 2018, pp. 532–538. doi: 10.1007/978-1-4899-7993-3_191.
[22] L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
[23] A. Liaw and M. Wiener, “Classification and regression by randomForest,” R News, vol. 2, no. 3, pp. 18–22, 2002.
[24] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot, “Variable selection using random forests,” Pattern Recognit. Lett., vol. 31, no. 14, pp. 2225–2236, 2010.
[25] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: An update,” ACM SIGKDD Explor. Newsl., vol. 11, no. 1, pp. 10–18, 2009.
[26] F. Provost and T. Fawcett, Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O’Reilly Media, 2013.
[27] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” Inf. Process. \& Manag., vol. 45, no. 4, pp. 427–437, 2009.
[28] L. Yin, R. Zhang, and Y. Guo, “Predictive analytics for e-commerce customer retention using ensemble learning,” J. Bus. Res., vol. 158, p. 113647, 2023.
[29] S. Choudhury and R. Singh, “Comparative performance of machine learning classifiers in FinTech applications,” Appl. Soft Comput., vol. 125, p. 109278, 2022.
[30] Y. Li, Q. Wang, and S. Li, “Feature importance in credit scoring using RF and SHAP,” Expert Syst. Appl., vol. 237, p. 119011, 2024.
[31] M. A. Siddiqui and M. Farooq, “Socioeconomic determinants of conversion in digital finance,” Inf. Syst. Front., vol. 25, no. 1, pp. 45–60, 2023.
[32] L. Zhou, Y. Fang, and Y. Wu, “Enhancing marketing decisions with AI: The case of customer acquisition,” Decis. Support Syst., vol. 157, p. 113790, 2022.
[33] M. Ghasemaghaei, “Big data analytics capability and firm performance: The mediating role of customer relationship management,” Inf. \& Manag., vol. 58, no. 6, p. 103478, 2021.
[34] C. Molnar, Interpretable Machine Learning, 2nd ed. Leanpub, 2022.
[35] V. Singh, A. Jaiswal, and N. Kumar, “A benchmarking study on XGBoost and RF in marketing analytics,” Comput. Econ., vol. 62, no. 1, pp. 155–170, 2023.
[36] M. Nasir, M. A. Qureshi, and A. Ali, “Time-aware prediction of customer behavior using temporal machine learning,” J. Retail. Consum. Serv., vol. 75, p. 103510, 2023.