

BİREYSEL KREDİ SATIŞ PERFORMANSI OPTİMİZASYONU

Teknik Uygulama ve Metodoloji Raporu

Bengü ÖZSUBAŞI PEK, Yavuz Selim PEK

KTUN Endüstri Mühendisliği Yüksek Lisans / Proje Yönetimi

Aralık, 2025

1. ÖZET

Dijitalleşen bankacılık sektöründe, doğru müşteriye doğru ürünü sunmak, operasyonel verimlilik açısından oldukça önem arz eder. Analiz sürecinde kullanılan veri setinde mevcut Bireysel Kredi satış süreçlerinde yaşanan verimsizliği gidermek amacıyla başlatılan bu projede, 5.000 müşteriye ait demografik ve finansal veriler, veri bilimi ve makine öğrenmesi yöntemleriyle analiz edilmiştir.

Çalışma sonucunda, mevcut durumda %9,6 seviyesinde olan satış dönüşüm oranı, geliştirilen Random Forest algoritması ile %99,1 tahmin başarısına ulaştırılmıştır. Müşteri havuzunun tükenme hızını ve segment bazlı riskleri hesaba katan 30 günlük dinamik simülasyonda; model odaklı stratejinin günlük ortalama 329.467 TL ciro yarattığı hesaplanmıştır. Bu sonuç, mevcut rastgele arama sisteminin (96.000 TL) 3,5 katı üzerinde bir performansa ve yıllık bazda milyonlarca liralık potansiyel gelir artışına işaret etmektedir.

2. GİRİŞ

Günümüz rekabetçi bankacılık ekosisteminde, müşteri edinimi maliyetlerini (CAC) düşürmek ve çapraz satış oranlarını artırmak, kârlılığın en kritik parametrelerinden biri olarak değerlendirilmektedir. Veri setindeki mevcut pazarlama stratejisi, hedef kitle ayrimı gözetmeksiz yapılan Soğuk Arama yöntemine dayanmaktadır. Ancak veriler incelendiğinde, bu yöntemin %90,4 oranında başarısızlıkla sonuçlandığı anlaşılmaktadır. Bu durum, satış ekiplerinin eforunun büyük kısmını ilgisiz müşteriler üzerinde harcamasına ve operasyonel süreçlerin maliyetinin artmasına yol açmaktadır.

2.2. PROJENİN AMACI VE KAPSAMI

Bu projenin temel amacı, Soğuk Arama yerine satın alma olasılığı yüksek müşteri kitlesına satış yapma prensibine dayalı analistik bir hedefleme modeli geliştirmektir. Proje kapsamında: Müşterilerin kredi alma kararını etkileyen faktörlerin istatistiksel analizler ile belirlenmesi, makine öğrenmesi algoritmaları ile potansiyel müşterilerin yüksek doğrulukla

tahmin edilmesi, geliştirilen modelin finansal etkisinin simüle edilerek somut bir yol haritası sunulması hedeflenmiştir.

3. VERİ SETİ VE METODOLOJİ

Çalışmada kullanılan veri seti, müşterilerin yaşı, geliri, deneyimi, eğitim düzeyi, aile büyüğlüğü gibi demografik bilgilerini ve kredi kartı harcaması, vadeli mevduat gibi bankacılık ürün sahiplerini içeren 14 değişkenden oluşmaktadır.

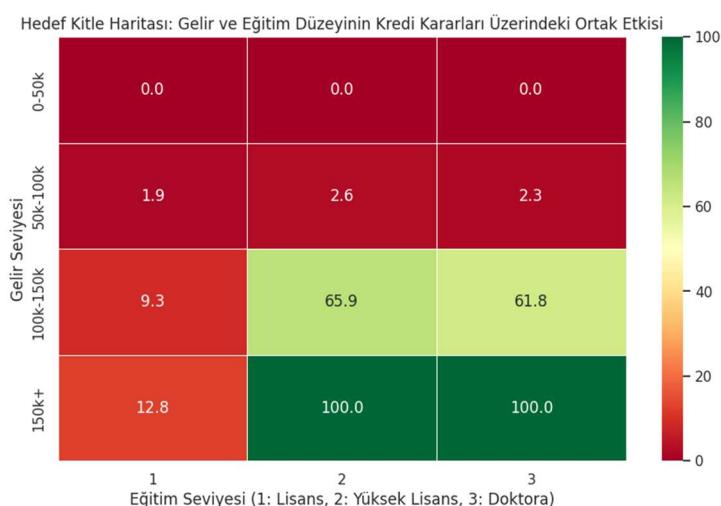
Analiz sürecinde şu adımlar izlenmiştir:

- Veri Ön İşleme: Hatalı (negatif) veriler düzeltilmiş, eksik veriler kontrol edilmiştir.
- Keşifçi Veri Analizi (EDA): Değişkenler arası korelasyonlar ve dağılımlar görselleştirilmiştir.
- Hipotez Testleri: Kategorik değişkenlerin hedef değişken üzerindeki etkisi Ki-Kare testi ile doğrulanmıştır.
- Modelleme: Lojistik Regresyon ve Random Forest algoritmaları karşılaştırmalı olarak test edilmiştir.

4. VERİ ANALİZİ VE BULGULAR

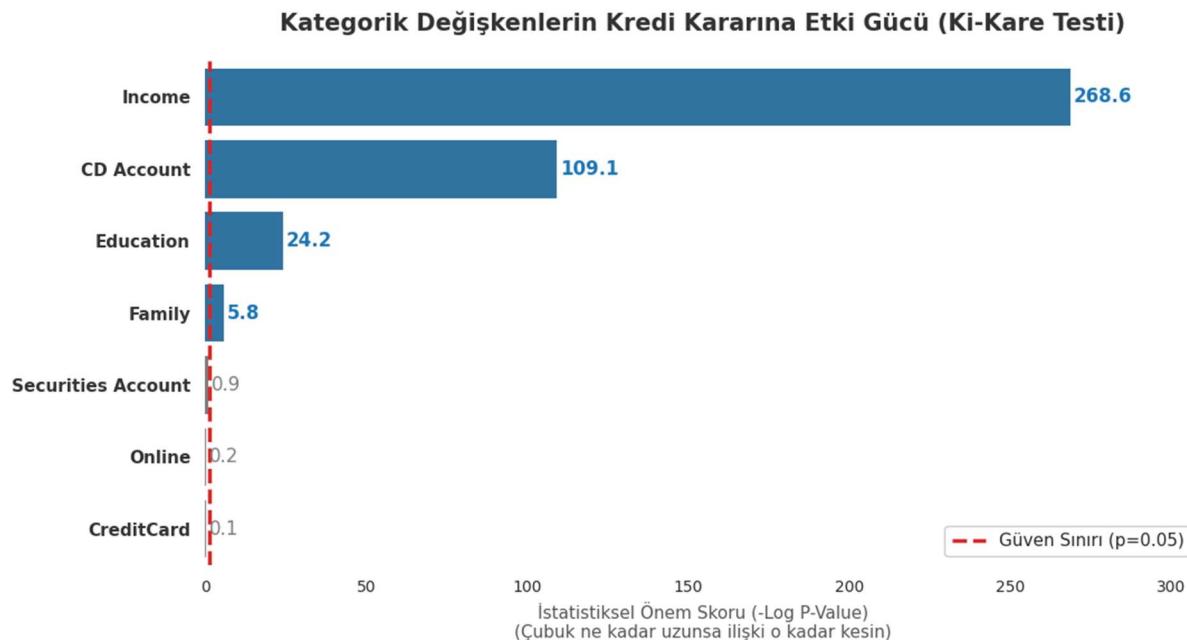
4.1. KRİTİK DEĞİŞKENLER ARASI İLİŞKİLER

Veri setindeki kritik sayısal değişkenlerin birbirleriyle ve bireysel kredi değişkeni ile olan ilişkisi korelasyon matrisi ile incelenmiştir (Şekil 1). Ayrıca Ki-Kare testi ile de değişkenlerin kredi kararına etki gücü ortaya konmuştur (Şekil 2).



Şekil 1. İşı Haritası – Değişkenler Arası İlişkilerin Analizi

Matris incelendiğinde; kredi kullanım kararı ile en yüksek pozitif ilişkinin Gelir (Income), Kredi Kartı Harcaması (CCAvg) ve Vadeli Mevduat (CD Account) arasında olduğu netleşmiştir. Geliri yüksek olan ve diğer banka ürünlerini (Kredi Kartı, Mevduat) aktif kullanan müşteriler, krediye en yatkın segmenti oluşturmaktadır.

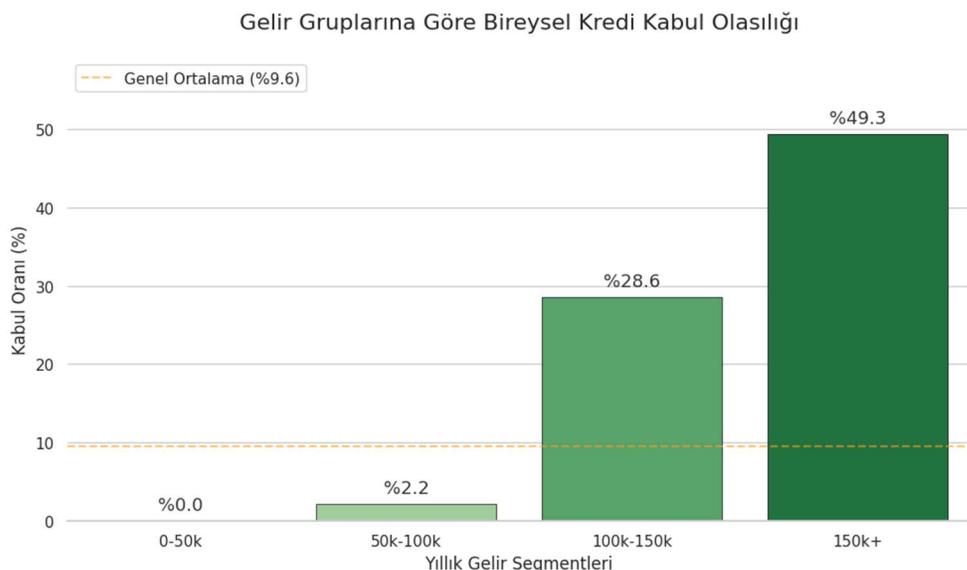


Şekil 2 Ki-Kare Testi

Ki-Kare testi ile yapılan analiz incelendiğinde kredi kararına etki gücü en yüksek olan faktör Gelir faktörü olarak tespit edilmiştir. Vadeli Mevduat ve Eğitim ise onu izleyen güçlü faktörler olarak yer almaktadır.

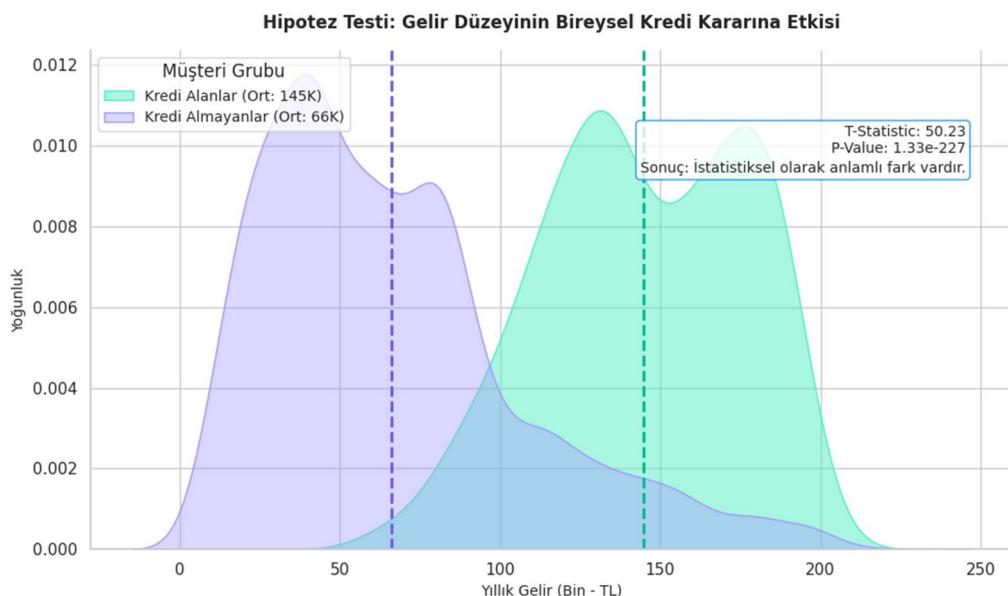
4.2. GELİR GRUBU FAKTÖRÜNÜN ANALİZİ

Kredi kabul eden ve etmeyen müşterilerin gelir dağılımları karşılaştırıldığında, iki grup arasında keskin bir ayrışma tespit edilmiştir. Kredi kullanan ve kullanmayan müşterilerin gelir dağılımları incelenmiş ve gelir seviyesinin kredi alma ihtimali (olasılığı) üzerindeki etkisi görselleştirilmiştir (Şekil 3). Aynı zamanda hipotez testi yapılarak anlamlı bir ilişkinin varlığı kanıtlanmıştır (Şekil 4).



Şekil 3. Gelir Gruplarına Göre Bireysel Kredi Olasılık Analizi

Kredi alanların büyük çoğunluğu, yıllık 100.000 TL ve üzeri gelir bandında yoğunlaşmaktadır. Gelir seviyesi arttıkça, müşterinin kredi alma olasılığı doğrusal olmayan bir şekilde artmaktadır. Düşük gelir grubunda kredi kullanım ihtimali neredeyse sıfır yakındır. Bu görsel, satış aramalarında "Gelir > 100k" filtresinin uygulanmasının, başarı şansını matematiksel olarak katlayacağını kanıtlamaktadır.

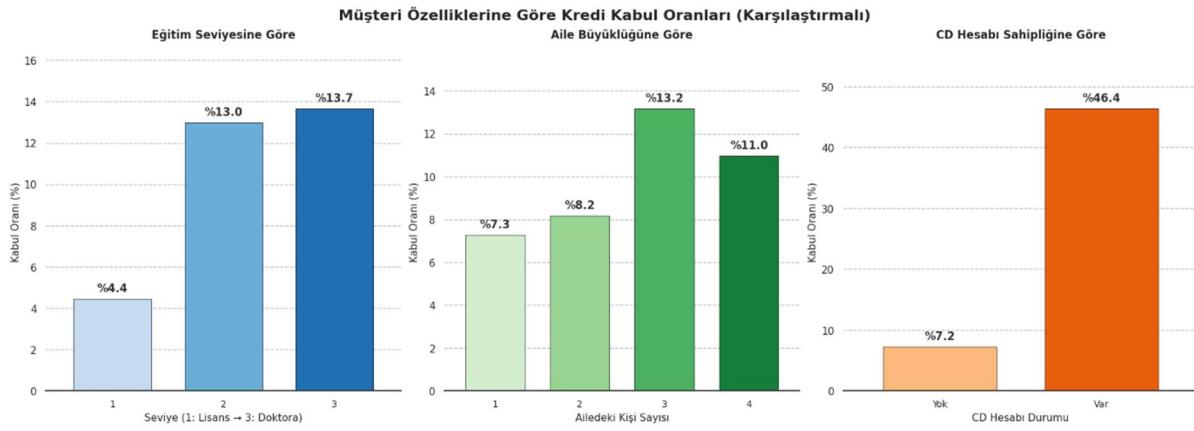


Şekil 4. Hipotez Testi: Gelir Düzeyinin Bireysel Kredi Kararına Etkisi

4.3. KRİTİK FAKTÖRLER: EĞİTİM VE VADELİ MEVDUAT

Geliri yüksek olsa bile, eğitim seviyesi Lisans (Seviye 1) olan müşterilerin kredi alma oranı beklenenin çok altındadır. Buna karşın, "Yüksek Lisans" ve "Doktora" eğitim seviyesine

sahip müşterilerde bu oran artmaktadır. Ayrıca Vadeli Mevduat sahipliği de kararı oldukça fazla etkilemektedir. Aile büyüklüğü faktörünün ise anlamlı bir etken olmadığı tespit edilmiştir (Şekil 5).



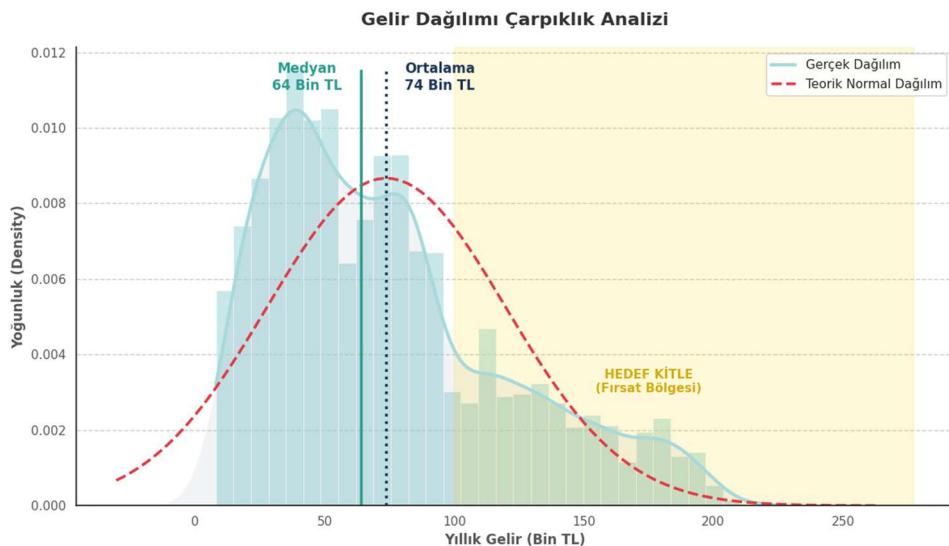
Şekil 5. Müşteri Özelliklerine Göre Kredi Kabul Oranlarının Karşılaştırılması

4.4. STRATEJİK SEGMENTASYON

Bu çalışma kapsamında, hedef kitleyi rastgele seçmek yerine, veri analizi bulgularına dayalı analitik kurallarla segmentasyonun satış performansı üzerindeki etkisi incelenmiştir. Aşağıdaki analizler, oluşturulan hedefleme stratejilerinin istatistiksel geçerliliğini ve ticari potansiyelini doğrulamaktadır.

4.4.1. Gelir Dağılımı ve Kritik Eşik Değer Analizi

Bireysel kredi kullanım kararı ile müşterinin yıllık geliri arasındaki ilişki analiz edilmiş ve grafik haline getirilmiştir (Şekil 6).

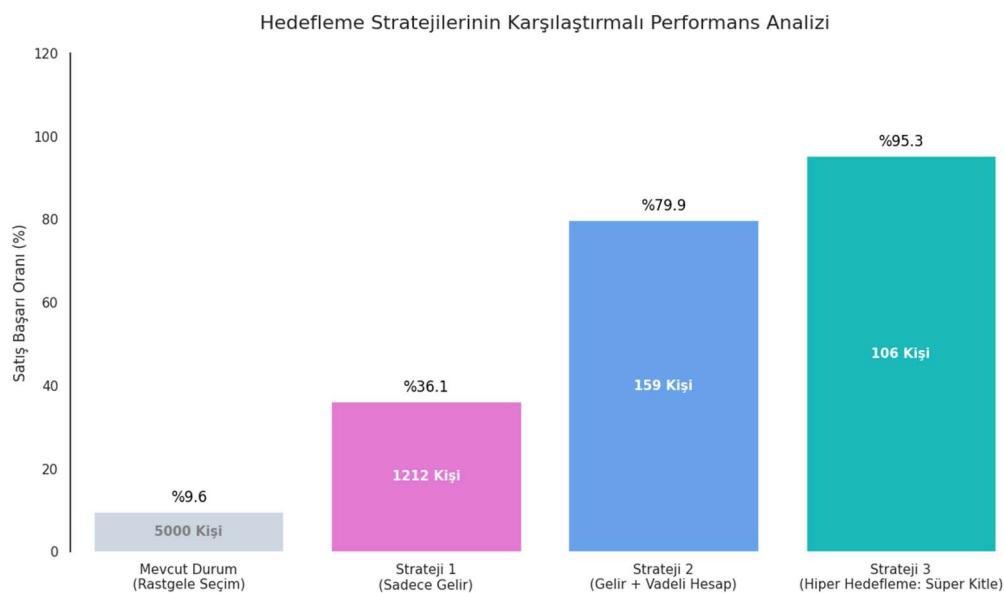


Şekil 6. Gelir Dağılımı Çarpıklığı Analizi

Kredi kabul eden kitlenin alt çeyrek (Q1) değeri, 100.000 TL seviyesine tekabül etmektedir. Bu istatistiksel kanıt ışığında, segmentasyon stratejimizin ilk ve en güçlü filtresi olarak ‘Gelir > 100k’ kuralı belirlenmiştir.

4.4.2. Çok Katmanlı Hedefleme Stratejilerinin Performans Karşılaştırması

Tek değişkenli (Gelir) analizin yetersiz kaldığı durumlarda, modele Eğitim Seviyesi ve Vadeli Mevduat değişkenlerinin eklenmesiyle oluşturulan kademeli bir strateji başarısı aşağıdaki grafik ile kıyaslanmıştır (Şekil 7).



Şekil 7. Çok Katmanlı Hedefleme Stratejilerin Karşılaştırılması

Bu karşılaştırma tablosu, demografik ve finansal davranış değişkenlerinin birlikte kullanılmasının, hedefleme doğruluğunu mevcut sisteme kıyasla 10 kata kadar artırabileceğini matematiksel olarak doğrulamaktadır.

5. MODELLEME VE PERFORMANS DEĞERLENDİRMESİ

Veri analizi aşamasında tespit edilen ilişkilerin (Gelir, Eğitim vb.) bir tahmin mekanizmasına dönüştürülmesi amacıyla makine öğrenmesi algoritmaları kullanılmıştır. Çalışmada, temel kıyaslama modeli olarak Lojistik Regresyon ve nihai model olarak Random Forest algoritmaları karşılaştırmalı olarak test edilmiştir.

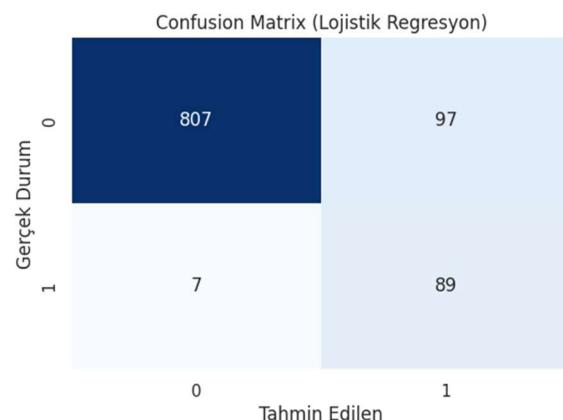
5.1. YÖNTEM SEÇİMİ VE KARŞILAŞTIRMALI ANALİZİ

Veri setindeki değişkenler arasındaki ilişkilerin doğrusal olmaması nedeniyle "Ensemble Learning" prensibiyle çalışan Random Forest algoritması, Lojistik Regresyon'a göre belirgin bir üstünlük sağlamıştır.

Aşağıda, her iki modelin operasyonel verimliliği doğrudan etkileyen hata matrisleri (Confusion Matrix) karşılaştırılmıştır.

5.1.1. Lojistik Regresyon Modeli

İlk aşamada kurulan Lojistik Regresyon modeli %90 genel doğruluk oranına ulaşmış olsa da bankacılık operasyonları için kritik olan "Hatalı Pozitif" oranında yetersiz kalmıştır.



Şekil 8. Lojistik Regresyon Hata Matrisi

Matrisin sağ üst köşesindeki 97 değeri, modelin aslında kredi almayacak olan 97 kişiyi hatalı bir şekilde "Potansiyel Müşteri" olarak işaretlediğini göstermektedir (Şekil 8). Bu durum, kaynakların israf edilmesine neden olacağı için model operasyonel açıdan verimsiz bulunmuştur.

5.1.2. Random Forest Modeli

Random Forest algoritmasına geçildiğinde, performans metriklerinde oldukça iyi bir iyileşme kaydedilmiştir.

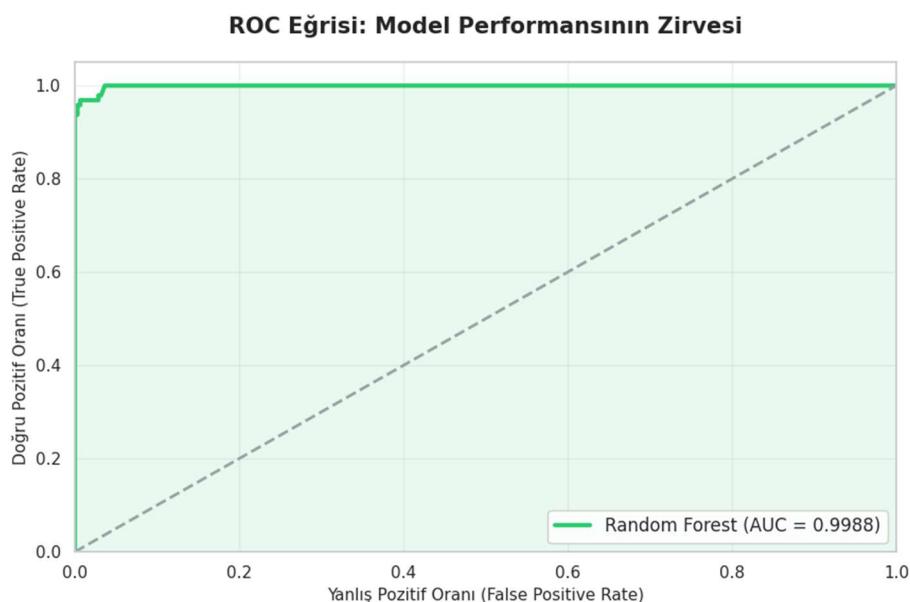


Şekil 9. Random Forest Hata Matrisi

Lojistik Regresyon modelindeki 97 hatalı tahmin, Random Forest ile sadece 3'e düşürülmüştür. Bu, hatalı aramaların %97 oranında elimine edildiği anlamına gelir. Model, test setindeki potansiyel 96 müşterinin 90'ını (%94) başarıyla tespit etmiştir. %99,1'lik genel doğruluk oranı, modelin mükemmel yakını bir hassasiyetinde çalıştığını kanıtlamaktadır (Şekil 9).

5.2. Modelin Ayırt Edicilik Gücü (ROC - AUC Analizi)

Modelin sınıflandırma başarısı, endüstri standartı olan ROC eğrisi ile test edilmiştir (Şekil 10).



Şekil 10. ROC Eğrisi

0.9988 AUC skoru ile modelin rastgele tahminden tamamen ayırttığını ve ayırt edicilik gücünün oldukça güçlü olduğunu göstermektedir.

5.3. KALİTE KONTROL: MODEL GÜVENİLİRLİĞİ TESTİ

Elde edilen %99,1'lik başarının veri setinin o anki dağılımına bağlı bir şans eseri (overfitting) olup olmadığını test etmek amacıyla 5-Katlı Çapraz Doğrulama uygulanmıştır (Şekil 11).

5 Katlı Çapraz Doğrulama: [0.98875 0.98625 0.98 0.98 0.985]
Ortalama Başarı: %98.40
Standart Sapma: 0.0035

GENEL YORUM: TEMİZ. Ezberleme tespit edilmedi.

Şekil 11. 5-Katlı Çapraz Doğrulama

Ekran çıktısında görülen test skorlarının birbirine çok yakın olması ve Standart Sapma değerinin 0.0035 olması, modelin kararlı olduğunu kanıtlamaktadır.

6. FİNANSAL ETKİ VE SÜRDÜRÜLEBİLİRLİK SİMÜLASYONU

Modelin %99,1'lik teknik başarısının ticari hayatı arasındaki karşılığını ölçmek ve yönetime gerçekçi bir projeksiyon sunmak amacıyla, "Havuz Problemi" prensibine dayalı bir simülasyon modeli kurgulanmıştır. Bu simülasyon ile sonsuz sayıda ideal müşteri varsayımları yerine, zamanla havuz tükenir ve verim zamanla azalır gerektiğini matematiksel olarak modellemek hedeflenmiştir.

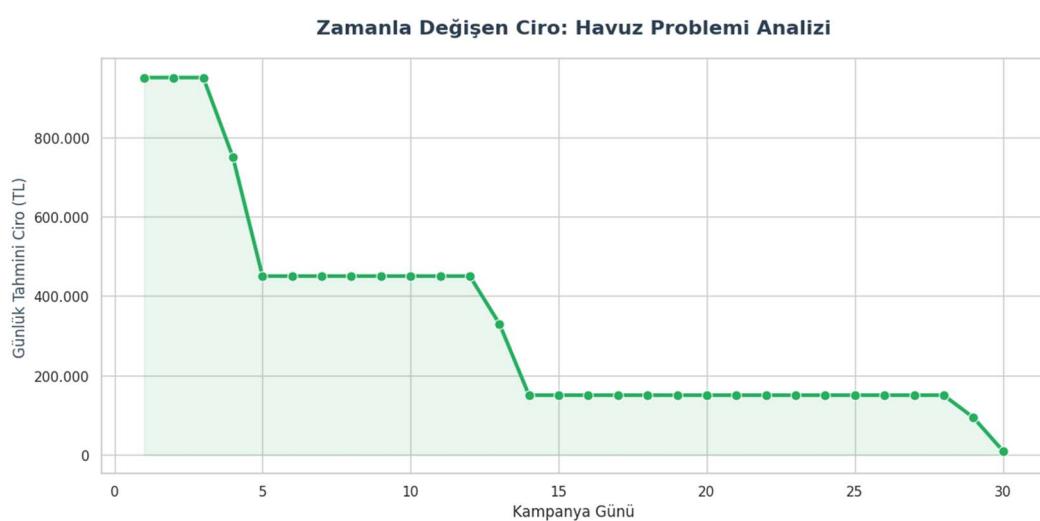
6.1. SİMÜLASYON METODOLOJİSİ VE VARSAYIMLAR

Simülasyonda kullanılan başarı oranları Bölüm 4'teki Çok Katmanlı Hedefleme Stratejilerin Karşılaştırılması analizinden elde edilen verilere dayanırmış ve riski elemeye etmek adına daha düşük oranlar tercih edilmiştir.

Segment	Gelir	Vadeli Hesap	Eğitim	Satış Oranı
Platinum	100k +	Var	Lisansüstü	%95
Gold	100k +	Yok	Lisansüstü	%45
Silver	100k +	Yok	Lisans ve altı	%15
Bronze	0 – 100k	Yok	Lisans ve altı	%1

6.2. SİMÜLASYON

Günlük 50 arama kapasitesi ile çalıştırılan 30 günlük simülasyonun sonuçları aşağıdadır (Şekil 12):



Şekil 12. Havuz Problemi Analizi

Grafikteki basamaklı yapı, müşteri havuzlarının kalitesine göre operasyonun nasıl şekilleneceğini göstermektedir:

Bu veri setine göre yapılan analizde ilk dört gün Platinum segmentin arandığı günler olarak günlük ciro potansiyeli maksimum seviyededir. En nitelikli havuz tükendikten sonra Gold ve Silver segmentlere geçişle birlikte ciro, sürdürülebilir bir denge noktasına oturmaktadır.

Tüm bu iniş çıkışlar hesaba katıldığında, model destekli operasyonun Günlük Ortalama Cirosu 329.467 TL olarak analiz edilmiştir.

7. SONUÇ VE ÖNERİLER

Bu çalışma, bireysel kredi satış süreçlerinde analitik yöntemlere geçişin, sadece bir teknolojik bir tercih olmaktan ziyade finansal bir zorunluluk olduğunu ortaya koymuştur. Geliştirilen Random Forest modeli, potansiyel müşterileri %99,1 doğrulukla tespit etmiş ve boşuna yapılan aramaları %97 oranında azaltmıştır. Mevcut sistemde 96.000 TL olan günlük ortalama ciro, modelin devreye alınmasıyla 3.5 katına çıkarak 329.467 TL seviyesine ulaşma potansiyeli taşımaktadır. Satış ekiplerinin motivasyonunu düşüren negatif arama oranları minimize edilmiş ve operasyonel süreçlerdeki maliyet en aza indirilmiştir.

Modelin başarısını kalıcı hale getirmek için çağrı merkezi arama listelerinin, veri setine göre üretilen model doğrultusunda olasılık skorları bazında dinamik olarak sıralanması önerilmektedir. Hızlı karar alınması gereken durumlarda veri analizinde elde edilen kurallar, satış temsilcilerine önerilmelidir. Analizde en güçlü faktörlerden biri olan vadeli mevduat müşterilerine yönelik, krediyi teşvik edici faiz kampanyaları uygulanması da öneriler dahilindedir.

DİPNOT

Bu çalışmada yer alan tüm finansal değerler, gelir grupları ve simülasyon çıktıları Türk lirası (TL) cinsinden hesaplanmıştır. Analiz, Türkiye bankacılık sektöründeki gelir dağılımı ve harcama alışkanlıklarını baz alınarak kurgulanmış; bu doğrultuda "Yıllık 100k+ Gelir" gibi eşik değerler yerel ekonomik parametrelere göre belirlenmiştir.