



MÜHENDİSLİK VE DOĞA BİLİMLERİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ
BİTİRME PROJESİ TEZİ

Makine Öğrenmesi ile Müşteri Kaybı Analizi

21120205063 - Gizem Yağmur ERGELEN

Danışman

Dr. Öğr. Üyesi Muhammet Sinan BAŞARSLAN

Haziran, 2025
İstanbul Medeniyet Üniversitesi, İstanbul

ÖNSÖZ

Bu tez çalışması, bilgisayar mühendisliği lisans eğitimimin son yılında, müşteri kaybı (churn) tahmini üzerine yapılan sistematik bir araştırmanın ürünü olarak hazırlanmıştır. Çalışma sürecinde veri analizi, makine öğrenmesi algoritmalarının uygulanması, model karşılaştırmaları ve performans değerlendirmeleri titizlikle yürütülmüş; hem teorik bilgi birikimi hem de pratik becerilerimi geliştirmeme olanak sağlamıştır. Araştırmanın her aşamasında, bilimsel yöntemlere bağlı kalarak dikkatli, objektif ve sistematik bir yaklaşım benimsemeye özen gösterdim. Gerek veri ön işleme sürecinde karşılaşılan zorluklar, gerekse modelleme adımlarındaki teknik kararlar bu süreci yalnızca akademik değil, aynı zamanda kişisel gelişimime de katkı sağlayacak şekilde şekillendirmiştir. Bu süreçte bana rehberlik eden, bilgi ve tecrübesiyle yolumu aydınlatan, yönlendirmeleriyle çalışmamın bilimsel niteliğini artıran değerli danışmanım **Dr. Öğr. Üyesi Muhammet Sinan BAŞARSLAN**'a en içten teşekkürlerimi sunarım. Kendisiyle çalışmak benim için büyük bir öğrenme fırsatı olmuştur. Ayrıca bu süreçte manevi desteklerini esirgemeyen aileme, yol arkadaşlarıma ve tüm akademik birikimiyle bu alanda ilerlememe katkı sağlayan tüm hocalarıma teşekkürü borç bilirim. Bu çalışmanın, ileride yapılacak benzer araştırmalara katkı sağlaması ve uygulamalı veri bilimi çalışmalarında yol gösterici bir kaynak olarak değerlendirilmesi en büyük temennimdir.

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	ii
İÇİNDEKİLER	iii
TABLO LİSTESİ	v
ŞEKİL LİSTESİ	vi
ÖZET	viii
SUMMARY	ix
1. GİRİŞ	1
2. LİTERATÜR TARAMA	2
3. MATERYAL METOD	3
3.1 Veri Seti Özellikleri ve Ön İşleme	3
3.2 Modelleme Süreci	3
3.3 Eşik Değeri Ayarlaması (Threshold Optimization)	4
3.4 Kullanılan Modellerin Özelleştirilmesi ve Parametre Seçimi	4
4. DENEYSEL ÇALIŞMA	5
4.1 Veri Kümesine Genel Bakış	5
4.2 Keşifsel Veri Analizi (EDA)	5
4.2.1 Hedef Değişken (Churn) Dağılımı	5
4.2.2 Gender (Cinsiyet) ve Churn İlişkisi	6
4.2.3 Contract (Sözleşme Türü) ve Churn İlişkisi	7
4.2.4 Payment Method (Ödeme Yöntemi) ve Churn İlişkisi	8
4.2.5 Internet Service (İnternet Hizmeti) ve Churn İlişkisi	9
4.2.6 TechSupport (Teknik Destek) ve Churn İlişkisi	10
4.2.7 PaperlessBilling (Faturalandırma) ve Churn İlişkisi	11
4.2.8 Tenure (Hizmet Süresi) ve Churn İlişkisi	12
4.2.9 Monthly Charges (Aylık Ücretler) ve Churn İlişkisi	13
4.2.10 Korelasyon Analizi	14
4.2.11 Churn ile Sayısal Değişkenler Arasındaki Korelasyon İlişkisi	15
4.3 Model Çıktıları ve Özetleri	16
4.3.1 Logistic Regression (Lojistik Regresyon)	16
4.3.1.1 Logistic Regression (Lojistik Regresyon) – SMOTE + Threshold Uygulamalı	19
4.3.2 Decision Tree (Karar Ağacı)	22
4.3.2.1 Decision Tree (Karar Ağacı) – SMOTE + Threshold Uygulamalı	26
4.3.3 Random Forest (Rastgele Orman)	29
4.3.3.1 Random Forest (Rastgele Orman) – SMOTE + Threshold Uygulamalı	32
4.3.4 SVM (Support Vector Machine / Destek Vektör Makinesi)	35

4.3.4.1 SVM (Support Vector Machine / Destek Vektör Makinesi) – SMOTE + Threshold Uygulamalı	38
4.3.5 XGBoost (eXtreme Gradient Boosting)	41
4.3.5.1 XGBoost (eXtreme Gradient Boosting) – SMOTE + Threshold Uygulamalı	45
4.3.6 Modellerin Karşılaştırmalı Performans Analizi	48
4.4 SMOTE, Eşik (Threshold) Ayarlaması ve K-Fold Doğrulama Yöntemlerinin Uygulamadaki Rolü	50
4.4.1 SMOTE Tekniği ile Sınıf Dengesizliğinin Giderilmesi	50
4.4.2 Optimal Eşik (Threshold) Değerinin Belirlenmesi ve Kullanımı	50
4.4.3 K-Fold Çapraz Doğrulama ile Model Genellenabilirliğinin Artırılması	51
4.5 Kod ve Araç Altyapısı	52
4.5.1 Geliştirme Ortamı ve Programlama Dili	52
4.5.2 Kullanılan Kütüphaneler ve Paketler	53
4.5.2.1 Veri Yükleme ve İşleme Kütüphaneleri	53
4.5.2.2 Veri Görselleştirme Kütüphaneleri	53
4.5.2.3 Makine Öğrenmesi ve Modelleme Kütüphaneleri	54
4.5.2.4 Dengesiz Veri Yapısı İçin Kullanılan Paketler	54
4.5.2.5 Boosting Tabanlı Model Kütüphanesi	54
4.5.2.6 Ek Fonksiyonel Kullanımlar	54
4.5.3 Kodlama Akışı ve Uygulama Adımları	55
4.5.3.1 Veri Yükleme ve İlk İnceleme	55
4.5.3.2 Keşifsel Veri Analizi (EDA)	55
4.5.3.3 Eğitim ve Test Setlerine Bölme	55
4.5.3.4 SMOTE + K-Fold + Threshold Entegrasyonu	55
4.5.3.5 Model Eğitimi ve Değerlendirme	56
4.5.3.6 Karşılaştırmalı Performans Tablosu ve Sonuçlandırma	56
5. SONUÇLAR VE TARTIŞMA	57
5.1 Bulguların Genel Değerlendirmesi	57
5.2 Tartışma ve Literatürle Karşılaştırma	57
5.3 Kısıtlar ve Geliştirme Önerileri	58
5.3.1 Veri Seti Kaynaklı Kısıtlar	59
5.3.2 Modelleme Sürecindeki Kısıtlar	59
5.3.3 Karar Sonrası Strateji Optimizasyonu	59
5.3.4 Donanım ve İşlem Gücü Kısıtlamaları	59
5.3.5 Gelecekteki Geliştirme Önerileri	59
6.SONUÇ	60
7.KAYNAKÇA	61

TABLO LİSTESİ

Sayfa No

Tablo 4. 1 Lojistik Reg. Modeli - Cross Validation Sonuçları	17
Tablo 4. 2 Lojistik Reg. Modeli - Test Verisi Performansı	18
Tablo 4. 3 Lojistik Reg. Modeli - Detaylı Sınıflandırma Raporu	19
Tablo 4. 4 Lojistik Reg. Modeli - SMOTE Uygulanan Cross Validation Sonuçları	20
Tablo 4. 5 Lojistik Reg. Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)	21
Tablo 4. 6 Lojistik Reg. Modeli (SMOTE) - Detaylı Sınıflandırma Raporu	22
Tablo 4. 7 Karar Ağacı Modeli - Cross Validation Sonuçları	23
Tablo 4. 8 Karar Ağacı Modeli - Test Verisi Performansı	24
Tablo 4. 9 Karar Ağacı Modeli - Detaylı Sınıflandırma Raporu	25
Tablo 4. 10 Karar Ağacı Modeli - SMOTE Uygulanan Cross Validation Sonuçları	26
Tablo 4. 11 Karar Ağacı Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)	27
Tablo 4. 12 Karar Ağacı Modeli (SMOTE) - Detaylı Sınıflandırma Raporu	28
Tablo 4. 13 Rastgele Orman Modeli - Cross Validation Sonuçları	29
Tablo 4. 14 Rastgele Orman Modeli - Test Verisi Performansı	30
Tablo 4. 15 Rastgele Orman Modeli - Detaylı Sınıflandırma Raporu	31
Tablo 4. 16 Rastgele Orman Modeli - SMOTE Uygulanan Cross Validation Sonuçları	32
Tablo 4. 17 Rastgele Orman Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)	33
Tablo 4. 18 Rastgele Orman Modeli (SMOTE) - Detaylı Sınıflandırma Raporu	35
Tablo 4. 19 Destek Vektör Makinesi Modeli - Cross Validation Sonuçları	35
Tablo 4. 20 Destek Vektör Makinesi Modeli - Test Verisi Performansı	36
Tablo 4. 21 Destek Vektör Makinesi Modeli - Detaylı Sınıflandırma Raporu	38
Tablo 4. 22 Destek Vektör Makinesi Modeli - SMOTE Uygulanan Cross Validation Sonuçları	39
Tablo 4. 23 Destek Vektör Makinesi Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)	40
Tablo 4. 24 Destek Vektör Makinesi Modeli (SMOTE) - Detaylı Sınıflandırma Raporu	41
Tablo 4. 25 XGBoost Modeli - Cross Validation Sonuçları	42
Tablo 4. 26 XGBoost Modeli - Test Verisi Performansı	43
Tablo 4. 27 XGBoost Modeli - Detaylı Sınıflandırma Raporu	44
Tablo 4. 28 XGBoost Modeli - SMOTE Uygulanan Cross Validation Sonuçları	45
Tablo 4. 29 XGBoost Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)	46
Tablo 4. 30 XGBoost Modeli (SMOTE) - Detaylı Sınıflandırma Raporu	48
Tablo 4. 31 Modellerin Karşılaştırılmış Performans Tablosu	49
Tablo 4. 32 SMOTE Uygulanan Modellerin Karşılaştırılmış Performans Tablosu	49
Tablo 4. 33 Modellerin Optimal Threshold Değeri Tablosu	51

ŞEKİL LİSTESİ

Sayfa No

Şekil 4. 1 Churn Sayısal Dağılımı	5
Şekil 4. 2 Churn Oransal Dağılımı	6
Şekil 4. 3 Gender (Cinsiyet) Dağılımı	6
Şekil 4. 4 Gender (Cinsiyet) ve Churn İlişkisi	7
Şekil 4. 5 Contract (Sözleşme Türü) Dağılımı	7
Şekil 4. 6 Contract (Sözleşme Türü) ve Churn İlişkisi	8
Şekil 4. 7 Payment Method (Ödeme Yöntemi) Dağılımı	8
Şekil 4. 8 Payment Method (Ödeme Yöntemi) ve Churn İlişkisi	9
Şekil 4. 9 Internet Service (İnternet Hizmeti) Dağılımı	9
Şekil 4. 10 Internet Service (İnternet Hizmeti) ve Churn İlişkisi	10
Şekil 4. 11 TechSupport (Teknik Destek) Dağılımı	10
Şekil 4. 12 TechSupport (Teknik Destek) ve Churn İlişkisi	11
Şekil 4. 13 PaperBilling (Faturalandırma) Dağılımı	11
Şekil 4. 14 PaperBilling (Faturalandırma) ve Churn İlişkisi	12
Şekil 4. 15 Tenure (Hizmet Süresi) Dağılımı	12
Şekil 4. 16 Tenure (Hizmet Süresi) ve Churn İlişkisi	13
Şekil 4. 17 Montly Charges (Aylık Ücretler) Dağılımı	13
Şekil 4. 18 Montly Charges (Aylık Ücretler) ve Churn İlişkisi	14
Şekil 4. 19 Sayısal Değişkenler Arasındaki Korelasyon Matrisi	15
Şekil 4. 20 Churn ile Sayısal Değişkenler Arasındaki Korelasyonlar	16
Şekil 4. 21 Lojistik Reg. Modeli - Foldların ROC Eğrisi	17
Şekil 4. 22 Lojistik Regresyon Modeli - Test Verisi ROC Eğrisi	18
Şekil 4. 23 Lojistik Regresyon - Confusion Matrix	19
Şekil 4. 24 Lojistik Reg. Modeli - SMOTE Uygulanan Foldların ROC Eğrisi	20
Şekil 4. 25 Lojistik Reg. Modeli (SMOTE) - Test Verisi ROC Eğrisi	21
Şekil 4. 26 Lojistik Regresyon - Confusion Matrix (Threshold Uygulandı)	22
Şekil 4. 27 Karar Ağacı Modeli - Foldların ROC Eğrisi	23
Şekil 4. 28 Karar Ağacı Modeli - Test Verisi ROC Eğrisi	24
Şekil 4. 29 Karar Ağacı - Confusion Matrix	25
Şekil 4. 30 Karar Ağacı Modeli - SMOTE Uygulanan Foldların ROC Eğrisi	26
Şekil 4. 31 Karar Ağacı Modeli (SMOTE) - Test Verisi ROC Eğrisi	27
Şekil 4. 32 Karar Ağacı - Confusion Matrix (Threshold Uygulandı)	28
Şekil 4. 33 Rastgele Orman Modeli - Foldların ROC Eğrisi	29
Şekil 4. 34 Rastgele Orman Modeli - Test Verisi ROC Eğrisi	30
Şekil 4. 35 Rastgele Orman - Confusion Matrix	31
Şekil 4. 36 Rastgele Orman Modeli - SMOTE Uygulanan Foldların ROC Eğrisi	33
Şekil 4. 37 Rastgele Orman Modeli (SMOTE) - Test Verisi ROC Eğrisi	34
Şekil 4. 38 Rastgele Orman - Confusion Matrix (Threshold Uygulandı)	34
Şekil 4. 39 Destek Vektör Makinesi Modeli - Foldların ROC Eğrisi	36
Şekil 4. 40 Destek Vektör Makinesi Modeli - Test Verisi ROC Eğrisi	37
Şekil 4. 41 Destek Vektör Makinesi - Confusion Matrix	37
Şekil 4. 42 Destek Vektör Makinesi Modeli - SMOTE Uygulanan Foldların ROC Eğrisi	39
Şekil 4. 43 Destek Vektör Makinesi Modeli (SMOTE) - Test Verisi ROC Eğrisi	40
Şekil 4. 44 Destek Vektör Makinesi - Confusion Matrix (Threshold Uygulandı)	41

Şekil 4. 45 XGBoost Modeli - Foldların ROC Eğrisi	42
Şekil 4. 46 XGBoost Modeli - Test Verisi ROC Eğrisi	43
Şekil 4. 47 XGBoost - Confusion Matrix	44
Şekil 4. 48 XGBoost Modeli - SMOTE Uygulanan Foldların ROC Eğrisi	46
Şekil 4. 49 XGBoost Modeli (SMOTE) - Test Verisi ROC Eğrisi	47
Şekil 4. 50 XGBoost - Confusion Matrix (Threshold Uygulandı)	47
Şekil 4. 51 SMOTE Uygulaması Öncesi ve Sonrası Churn Dağılımı	50

ÖZET

Günümüz rekabetçi pazar ortamında, şirketler açısından mevcut müşterileri elde tutmak, yeni müşteri kazanmaktan daha ekonomik ve stratejik bir öneme sahiptir. Bu bağlamda, müşteri kaybı (churn) analizleri, özellikle hizmet sektöründe faaliyet gösteren firmalar için kritik bir ihtiyaç hâline gelmiştir. Bu tez çalışmasında, bir telekomünikasyon veri seti üzerinde makine öğrenmesi tabanlı modeller kullanılarak churn tahmini gerçekleştirilmiş ve çeşitli performans ölçütleri aracılığıyla modellerin karşılaştırmalı değerlendirilmesi yapılmıştır.

Kaggle platformundan elde edilen Telco veri seti üzerinde öncelikle veri temizleme, dönüştürme ve keşifsel analiz adımları gerçekleştirilmiştir. Ardından veri dengesizliği problemi SMOTE (Synthetic Minority Over-sampling Technique) yöntemi ile giderilmiş; model tahmin çıktılarında eşik (threshold) optimizasyonu yapılmış ve 5 katlı çapraz doğrulama (Stratified K-Fold CV) yöntemi uygulanarak genellenebilirlik sağlanmıştır.

Çalışmada Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) ve XGBoost olmak üzere toplam beş model test edilmiştir. Değerlendirmeler doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1-skora ve ROC-AUC gibi metrikler üzerinden yapılmıştır. Elde edilen sonuçlara göre, XGBoost modeli özellikle ROC-AUC ve F1-score açısından en iyi performansı sergilemiş, Random Forest ve SVM modelleri ise bu başarıyı yakın şekilde takip etmiştir.

Bu çalışma, veri ön işleme, dengesiz sınıf problemleri ve karar sonrası optimizasyon tekniklerinin, makine öğrenmesi modellerinin başarısını artırmadaki etkisini açık bir biçimde ortaya koymuştur. Ayrıca elde edilen bulgular, önceki akademik çalışmalarla karşılaştırmalı olarak değerlendirilmiş ve geliştirilebilir yönler detaylı şekilde tartışılmıştır.

Anahtar Kelimeler: Müşteri kaybı, makine öğrenmesi, sınıflandırma, SMOTE, eşik optimizasyonu, çapraz doğrulama, XGBoost, performans değerlendirme.

SUMMARY

In today's highly competitive market environment, retaining existing customers has become more economical and strategically important than acquiring new ones. In this context, customer churn analysis has emerged as a critical necessity, particularly for service-based industries. This thesis aims to predict customer churn using machine learning models on a telecommunications dataset, and to evaluate the performance of these models through comparative analysis.

The Telco dataset, obtained from the Kaggle platform, was initially subjected to data preprocessing steps including cleaning, transformation, and exploratory analysis. To address the class imbalance problem, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. Additionally, threshold optimization was performed for classification decisions, and 5-fold stratified cross-validation was implemented to ensure model generalizability.

Five different classification models were evaluated in this study: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and XGBoost. Model performance was measured using key evaluation metrics including accuracy, precision, recall, F1-score, and ROC-AUC. The results showed that XGBoost outperformed all other models in terms of F1-score and ROC-AUC, with Random Forest and SVM also achieving competitive results.

This study demonstrates that data preprocessing, handling class imbalance, and post-prediction threshold adjustments significantly enhance model performance in churn prediction. Furthermore, the findings were compared with relevant academic literature and discussed in terms of methodological contributions and limitations.

Keywords: Customer churn, machine learning, classification, SMOTE, threshold optimization, cross-validation, XGBoost, performance evaluation.

1. GİRİŞ

Müşteri kaybı (churn), özellikle abonelik tabanlı çalışan işletmeler için kârlılığı doğrudan etkileyen önemli bir göstergedir. Bir müşterinin hizmeti sonlandırması, yalnızca gelir kaybına yol açmakla kalmaz; aynı zamanda pazarlama, müşteri kazanımı ve operasyonel maliyetlerde artışa neden olur. Bu nedenle, müşterilerin hizmetten ayrılmadan önce tahmin edilmesi, işletmelerin proaktif stratejiler geliştirmesi açısından büyük önem taşır.

Geleneksel yöntemlerle yapılan churn analizleri, geçmişe dönük raporlama ve temel istatistiksel değerlendirmelerle sınırlı kalmakta, özellikle yüksek hacimli verilerde öngörü gücü yetersiz kalmaktadır. Bu eksikliği gidermek amacıyla, son yıllarda makine öğrenmesi teknikleri churn tahmininde sıkça kullanılmakta ve başarılı sonuçlar elde edilmektedir.

Bu tez çalışmasında, Kaggle platformundan temin edilen Telco müşteri verisi kullanılarak churn tahmini gerçekleştirilmiştir. Çalışmada temel hedef, farklı makine öğrenmesi algoritmalarının performanslarını karşılaştırmak ve veri dengesizliğinin model başarımı üzerindeki etkilerini analiz etmektir. Bu doğrultuda Logistic Regression, Decision Tree, Random Forest, SVM ve XGBoost modelleri değerlendirilmiş; SMOTE, threshold ayarlaması ve 5-Fold çapraz doğrulama gibi yöntemler entegre edilmiştir.

Çalışmanın katkısı yalnızca model doğruluğuna odaklanmakla sınırlı değildir. Aynı zamanda veri dengesizliği, model genellenebilirliği ve karar sonrası eşik ayarlaması gibi kritik konular da uygulamalı olarak analiz edilmiş ve sonuçlar literatürle karşılaştırılarak tartışılmıştır. Böylece churn tahminine yönelik sistematik, bütüncül ve yeniden üretilebilir bir yöntem önerilmiştir.

2. LİTERATÜR TARAMA

Müşteri kaybı tahmini, özellikle rekabetin yüksek olduğu sektörlerde firmaların stratejik karar alma süreçlerinde önemli bir yer tutmaktadır. Bu alandaki çalışmaların büyük çoğunluğu, makine öğrenmesi algoritmalarının sınıflandırma gücünden faydalanarak müşteri davranışlarını önceden kestirmeyi amaçlamaktadır. Son yıllarda yapılan akademik çalışmalar, farklı sektörler ve veri setleri üzerinde çeşitli algoritmaların karşılaştırmalı analizlerini sunmuştur.

Kaba (2021) tarafından gerçekleştirilen bir yüksek lisans tezinde, faizsiz finans sektöründeki müşteri kaybı lojistik regresyon, KNN ve SVM yöntemleri ile analiz edilmiştir. Çalışmada keşifsel veri analizi ve veri ön işleme süreçlerine büyük önem verilmiş; veri dengesizliği sorununa karşı SMOTE uygulanmıştır. Sonuçlar, lojistik regresyonun diğer yöntemlere kıyasla daha etkili bir performans sunduğunu göstermiştir.

Amin ve arkadaşlarının çalışmasında (2016), farklı kural çıkarım yöntemleri (Genetik Algoritma, Exhaustive Algorithm, LEM2 gibi) kullanılarak telekom sektöründe churn analizi gerçekleştirilmiştir. Rough Set Theory (RST) yaklaşımı ile geliştirilen model, karar kurallarını etkili biçimde tanımlayarak en iyi performansı Genetik Algoritma tabanlı yapı ile elde etmiştir.

Kayaalp (2017) ise telekomünikasyon sektöründe yapılan churn analizlerine odaklanan geniş kapsamlı bir derlemesinde, literatürde en yaygın kullanılan yöntemlerin yapay sinir ağları, karar ağaçları, destek vektör makineleri ve lojistik regresyon olduğunu belirtmiştir. Ayrıca veri dengesizliği sorununa yönelik SMOTE, ADASYN gibi örnekleme tekniklerinin önemine vurgu yapılmıştır.

Son olarak, MAUN Fen Bilimleri Dergisi'nde yayımlanan bir makalede, Telco veri seti kullanılarak farklı makine öğrenmesi algoritmalarının (LR, SVM, RF, XGBoost, MLP) karşılaştırması yapılmış; özellikle Gradient Boosting ve MLP tabanlı modellerin yüksek başarı sağladığı gözlemlenmiştir. Çalışmada veri gizliliğine ve etik ilkelere uyulması da önemli bir katkı olarak belirtilmiştir.

Bu dört çalışmadan elde edilen bulgular, müşteri kaybı analizinde doğru algoritma seçiminin yanı sıra, veri setine özel ön işleme süreçlerinin ve dengesiz veri yapılarının dikkate alınmasının da model başarısını doğrudan etkilediğini göstermektedir. Bu tezde ise, telekom sektörüne ait veri kümesi üzerinde benzer yöntemler karşılaştırmalı biçimde test edilerek literatürdeki bulgular doğrulanmaya çalışılmıştır.

3. MATERYAL METOD

Bu çalışmada, telekomünikasyon sektörüne ait gerçek dünya verileri kullanılarak müşteri kaybı (churn) tahmini yapılmıştır. Veri seti, Kaggle platformu üzerinden elde edilen "Telco Customer Churn" veri setidir. Veri kümesi, müşterilerin abonelik bilgileri, demografik özellikleri, hizmet türleri, sözleşme detayları ve ödeme alışkanlıkları gibi çeşitli niteliksel ve niceliksel özellikleri içermektedir. Bu değişkenler üzerinden müşterinin hizmetten ayrılıp ayrılmayacağı (churn) durumu sınıflandırma problemi olarak ele alınmıştır.

3.1 Veri Seti Özellikleri ve Ön İşleme

Veri seti toplamda 7043 gözlem ve 21 sütun içermektedir. Bu sütunlardan biri hedef değişken olan 'Churn' bilgisini içermekte; diğerleri ise açıklayıcı değişkenlerdir. İlk olarak verideki eksik gözlemler tespit edilmiş ve uygun şekilde giderilmiştir. Ardından kategorik veriler, makine öğrenmesi algoritmalarının işleyebilmesi amacıyla One-Hot Encoding yöntemiyle dönüştürülmüştür. Sayısal veriler için ise ölçeklendirme (standardizasyon) işlemi gerçekleştirilmiş ve değişkenler aynı ölçek düzeyine getirilmiştir.

3.2 Modelleme Süreci

Veri seti dengesiz bir yapı sergilediğinden dolayı (örneğin churn eden müşteri oranı yaklaşık %26), bu durum model başarısını etkileyebileceğinden Synthetic Minority Oversampling Technique (SMOTE) uygulanmıştır. Ancak veri sızıntısını önlemek amacıyla SMOTE yalnızca eğitim verisine uygulanmış, doğrulama (validation) verisi orijinal haliyle bırakılmıştır. Böylece modelin genelleme kapasitesi adil bir biçimde test edilmiştir.

Bu çalışmada beş farklı sınıflandırma algoritması kullanılmıştır: Lojistik Regresyon, Karar Ağacı (Decision Tree), Rastgele Orman (Random Forest), Destek Vektör Makineleri (SVM) ve XGBoost. Her bir model, yalnızca temel haliyle değil; aynı zamanda veri dengesizliğine karşı SMOTE, eşik değeri ayarlaması (threshold tuning) ve K-Fold çapraz doğrulama ile birlikte değerlendirilmiştir.

Lojistik Regresyon: İstatistiksel yorum yapılabilmesi açısından sıkça tercih edilen bu model, doğrusal karar sınırları çizen ve log-loss fonksiyonunu minimize etmeyi hedefleyen bir yapıya sahiptir. Modelde solver seçimi (örneğin 'liblinear' veya 'saga'), veri boyutu ve regularizasyon ihtiyacına göre optimize edilmiştir.

Karar Ağacı: Veri içerisindeki if-then yapısını öğrenerek hiyerarşik kurallar çıkaran Karar Ağacı modeli, yorumlanabilirliği yüksek ancak overfitting'e yatkın bir yöntemdir. Maksimum derinlik gibi hiperparametreler kontrol edilerek modelin genelleme kabiliyeti artırılmıştır.

Rastgele Orman: Birden çok karar ağacının topluluğundan oluşan bu model, yüksek varyansı azaltmak için bagging yöntemi ile çalışır. Modelin performansı, ağaç sayısı, maksimum derinlik ve bootstrap örnekleme gibi parametreler üzerinden manuel olarak belirlenmiştir. Ayrıca SMOTE ile birlikte uygulandığında churn sınıfındaki azınlık örneklerini daha iyi temsil edebilmiştir.

Destek Vektör Makineleri (SVM): Özellikle yüksek boyutlu uzaylarda etkili sonuçlar veren bu yöntem, verileri optimal hiper düzlemle ayırmayı amaçlar. Kernel tipi ve regularizasyon

parametresi (C) modelin performansını doğrudan etkilediği için uygun değerler manuel olarak seçilmiştir. SVM, küçük veri setlerinde yüksek doğruluk sağlayabilir; ancak büyük veri setlerinde hesaplama maliyeti artmaktadır.

XGBoost: Ağaç tabanlı gradyan artırmalı bu algoritma, her bir zayıf öğreniciyi ardışık olarak geliştirerek hata oranını azaltmayı amaçlar. Özellikle veri dengesizliği problemlerine karşı `scale_pos_weight` parametresi ile güçlü bir kontrol sağlar. Öğrenme oranı (learning rate), `max_depth` ve `estimators` gibi parametreler deneysel olarak belirlenmiştir.

Model seçiminde aşırı öğrenmeyi önlemek ve daha adil bir performans değerlendirmesi sağlamak amacıyla K-Fold Çapraz Doğrulama uygulanmıştır. Bu süreçte, veri seti 5 farklı alt kümeye (fold) bölünmüş ve her fold'da model sıfırdan kurulup eğitilerek ayrı bir doğrulama kümesi üzerinde test edilmiştir. Her fold içerisinde SMOTE yalnızca eğitim verisine uygulanmış; böylece veri sızıntısı engellenmiş ve modelin farklı veri bölümlerinde tutarlı performans sergileyip sergilemediği ölçülmüştür. Tüm fold'lardan elde edilen metrikler ortalanarak genel model performansı elde edilmiştir.

3.3 Eşik Değeri Ayarlaması (Threshold Optimization)

Sınıflandırma algoritmalarının çoğu, varsayılan olarak 0.5 eşik değeri üzerinden sınıf tahmini yapmaktadır. Ancak churn gibi dengesiz sınıf dağılımına sahip problemlerde sabit bir eşik değeri performans kaybına yol açabilir. Bu nedenle model çıktılarından elde edilen olasılık skorları (`predict_proba`) kullanılarak farklı threshold değerleri test edilmiştir. Her bir eşik değeri için precision, recall ve F1-score değerleri hesaplanarak, en yüksek F1 skorunun elde edildiği eşik değeri "optimal threshold" olarak belirlenmiştir. Bu işlem, özellikle doğru pozitifleri artırma ve yanlış alarmları azaltma arasında denge kurma açısından önemlidir.

Threshold belirleme süreci için sklearn kütüphanesindeki `precision_recall_curve` fonksiyonu kullanılmıştır. Elde edilen precision ve recall değerleri üzerinden F1 skoru hesaplanmış ve maksimum değeri veren threshold seçilerek model çıktıları bu değere göre sınıflandırılmıştır.

3.4 Kullanılan Modellerin Özelleştirilmesi ve Parametre Seçimi

Lojistik Regresyon modelinde solver parametresi olarak "liblinear" ve "saga" seçenekleri değerlendirilmiştir. Solver seçimi, modelin optimizasyon algoritmasını belirlemesi açısından önemlidir; farklı solver'lar özellikle regularizasyon türleri (L1, L2) ve veri boyutu gibi durumlarda farklı performans sergilemektedir. Rastgele Orman ve XGBoost modellerinde, hiperparametreler manuel olarak belirlenmiş ve deneysel olarak en uygun yapılandırma seçilmiştir.

Tüm modeller için performans değerlendirmesi accuracy, precision, recall, F1-score ve AUC (Area Under Curve) metrikleriyle yapılmıştır. Bu metrikler, özellikle churn sınıfı gibi azınlık sınıflar için dengeli bir değerlendirme yapılmasına imkân tanımaktadır.

Sonuç olarak, bu metodolojik yapı sayesinde farklı makine öğrenmesi algoritmalarının dengesiz veri seti üzerindeki performansı karşılaştırmalı olarak değerlendirilmiş ve gerçek hayattaki uygulanabilirliği artırılmıştır.

4. DENEYSEL ÇALIŞMA

Bu bölümde, Telco Customer Churn veri seti üzerinde gerçekleştirilen keşifsel veri analizi (EDA), modelleme adımları ve her modelin elde ettiği sonuçlar detaylı olarak sunulmaktadır.

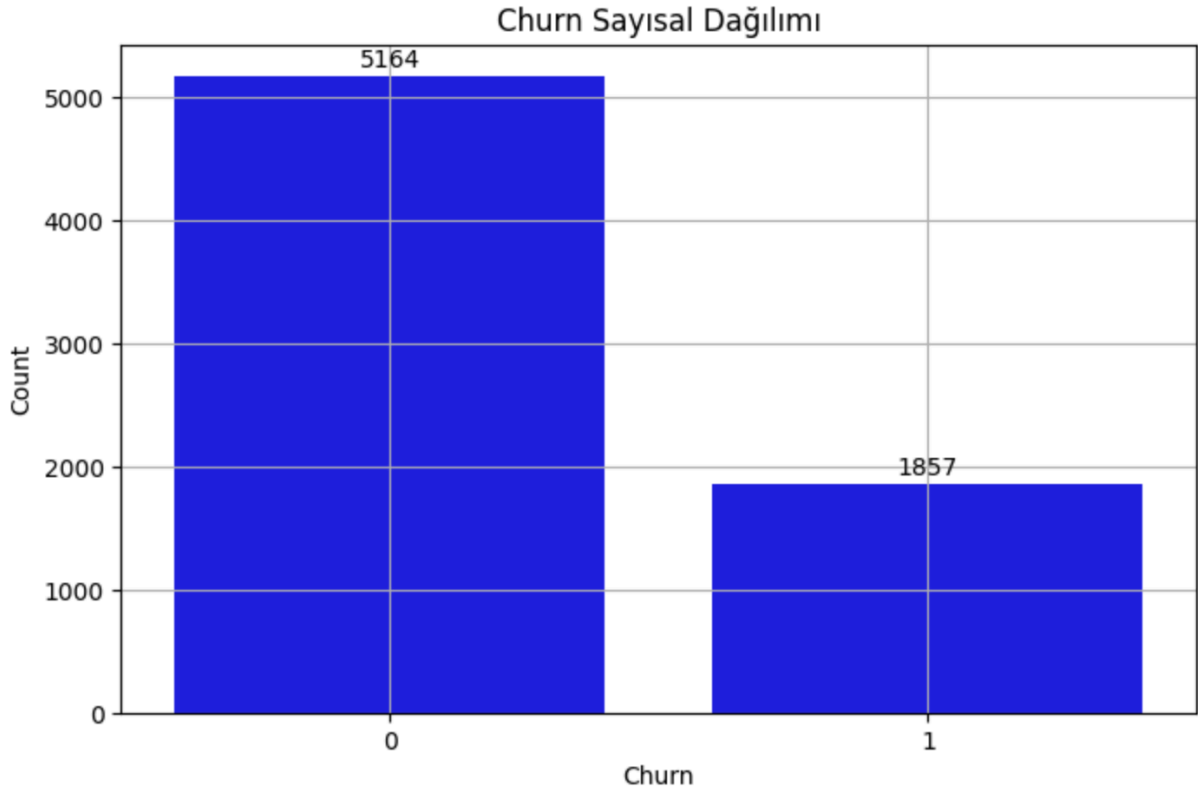
4.1 Veri Kümesine Genel Bakış

Veri seti toplamda 7043 satır ve 21 sütundan oluşmaktadır. Hedef değişken 'Churn', müşterinin belirli bir dönemde hizmeti bırakıp bırakmadığını (0: Hayır, 1: Evet) ifade etmektedir. Açıklayıcı değişkenler arasında müşterinin yaş bilgisi, cinsiyeti, sözleşme süresi, ödeme yöntemi, internet ve telefon hizmetlerinin kullanımı gibi kategorik ve sayısal değişkenler yer almaktadır. İlk adımda, veri kümesindeki eksik ve tutarsız veriler kontrol edilip temizlenmiş, ardından uygun biçimde sayısal ve kategorik değişken ayrımı yapılmıştır.

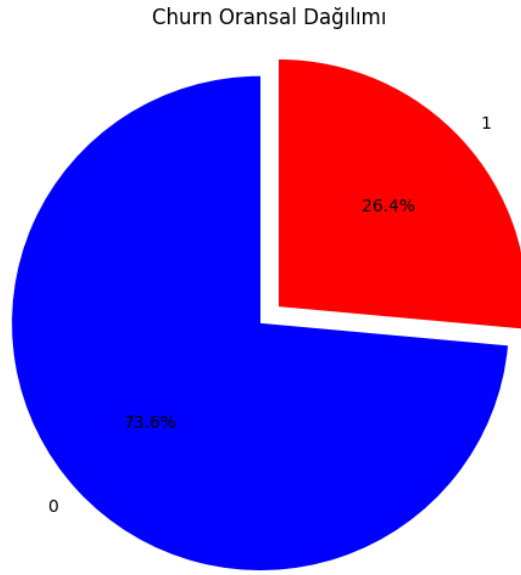
4.2 Keşifsel Veri Analizi (EDA)

4.2.1 Hedef Değişken (Churn) Dağılımı

Şekil 4.1 ve Şekil 4.2’de gösterildiği üzere veri setindeki müşterilerin yaklaşık %26’sı (1857) hizmetten ayrılmış (churn), geri kalan %74’ü (5164) ise hizmeti kullanmaya devam etmiştir. Bu dengesiz dağılım, sınıflandırma modellerinin azınlık sınıfı olan churn’ü doğru tahmin etmesini zorlaştırabileceğinden, SMOTE gibi dengeleme tekniklerinin kullanımını gerektirmiştir.



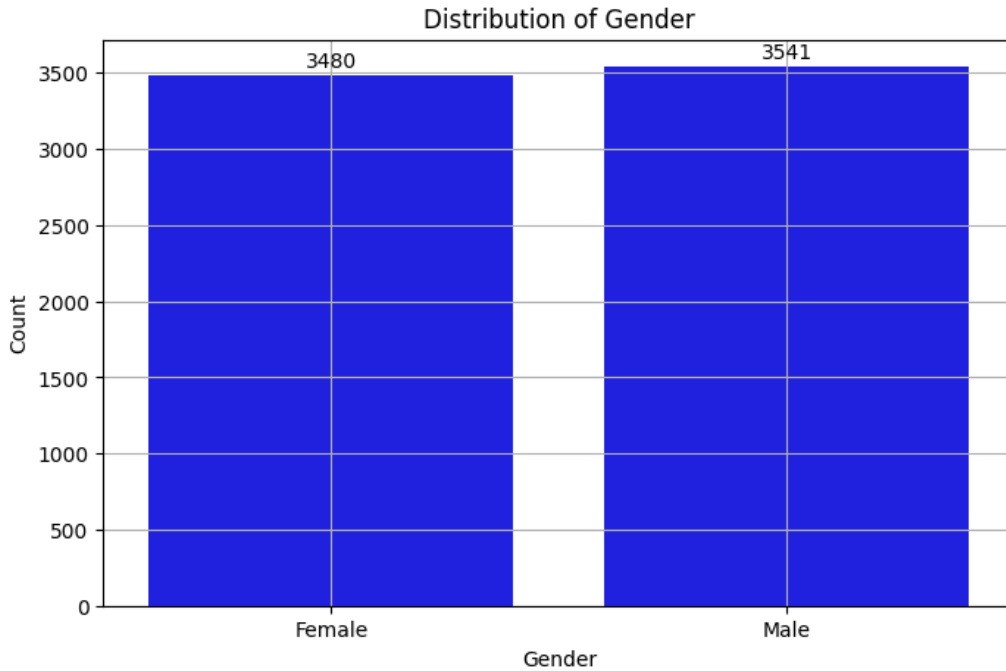
Şekil 4. 1 Churn Sayısal Dağılımı



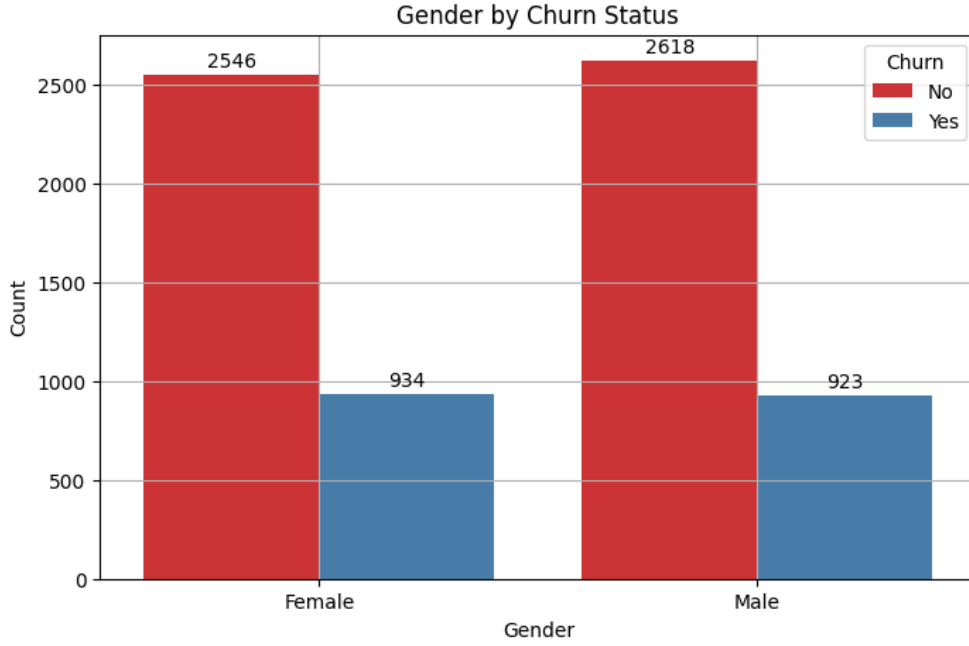
Şekil 4. 2 Churn Oransal Dağılımı

4.2.2 Gender (Cinsiyet) ve Churn İlişkisi

Şekil 4.3'te görüldüğü üzere, veri setindeki cinsiyet dağılımı oldukça dengelidir. 3480 kadın ve 3541 erkek müşteri yer almaktadır. Cinsiyetin churn üzerindeki etkisini değerlendirmek amacıyla oluşturulan Şekil 4.4'te ise kadın ve erkek müşteriler arasında churn oranlarının oldukça benzer olduğu görülmektedir. Kadın müşterilerde 934 kişi (toplamın %26,8'i), erkek müşterilerde ise 923 kişi (toplamın %26,0'ı) hizmetten ayrılmıştır. Bu bulgu, cinsiyet değişkeninin churn tahmininde belirleyici bir faktör olmadığını ve modelleme sürecinde tek başına anlamlı bir ayrıştırıcı olmadığını göstermektedir.



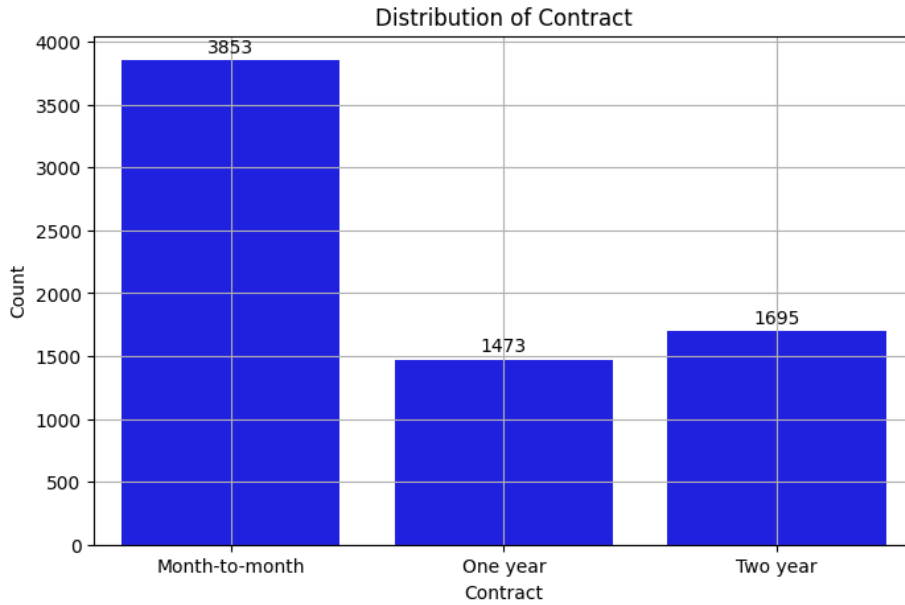
Şekil 4. 3 Gender (Cinsiyet) Dağılımı



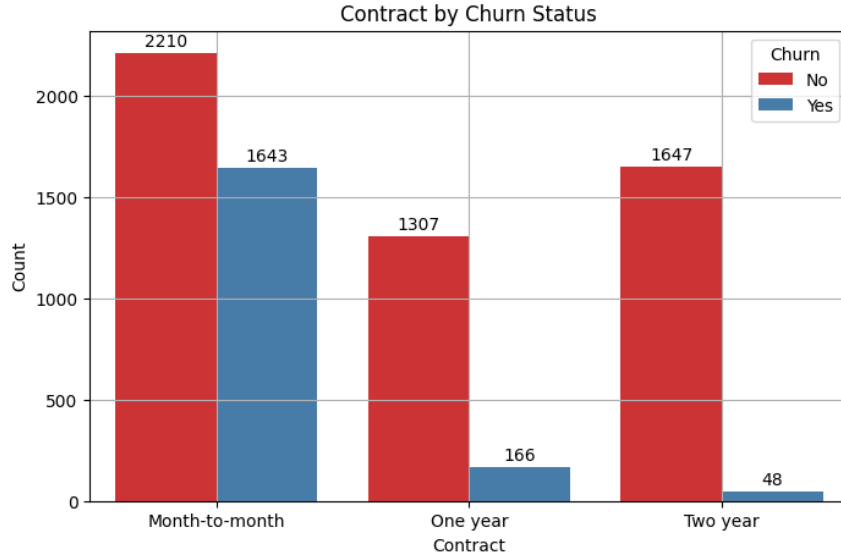
Şekil 4. 4 Gender (Cinsiyet) ve Churn İlişkisi

4.2.3 Contract (Sözleşme Türü) ve Churn İlişkisi

Şekil 4.5’te sözleşme türlerinin genel dağılımı sunulmuştur. Müşterilerin büyük bir kısmı (%55) “Month-to-month” (aylık) sözleşmeye sahiptir. Bu müşteri grubunu sırasıyla “Two year” ve “One year” sözleşme türleri takip etmektedir. Şekil 4.6’da ise sözleşme türü ile churn ilişkisi gösterilmiştir. Aylık sözleşmeye sahip müşterilerde churn oranı oldukça yüksektir (1643 churn / 3853 müşteri \approx %42,6). Buna karşılık iki yıllık sözleşmeye sahip müşterilerde churn oranı oldukça düşüktür (48 churn / 1695 müşteri \approx %2,8). Bu bulgular, uzun vadeli sözleşmelerin müşteri bağlılığını artırdığını ve churn oranlarını düşürdüğünü açıkça göstermektedir.



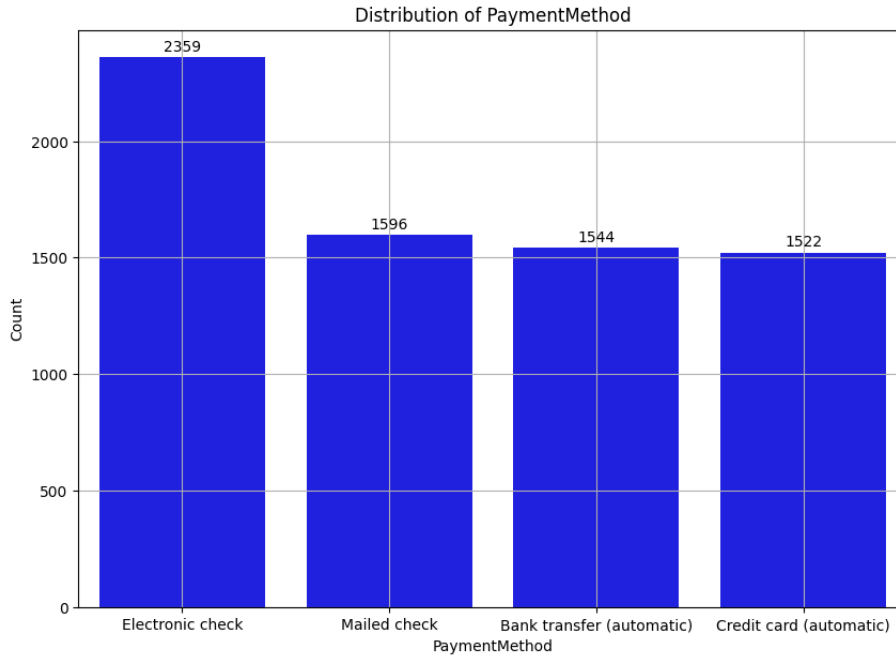
Şekil 4. 5 Contract (Sözleşme Türü) Dağılımı



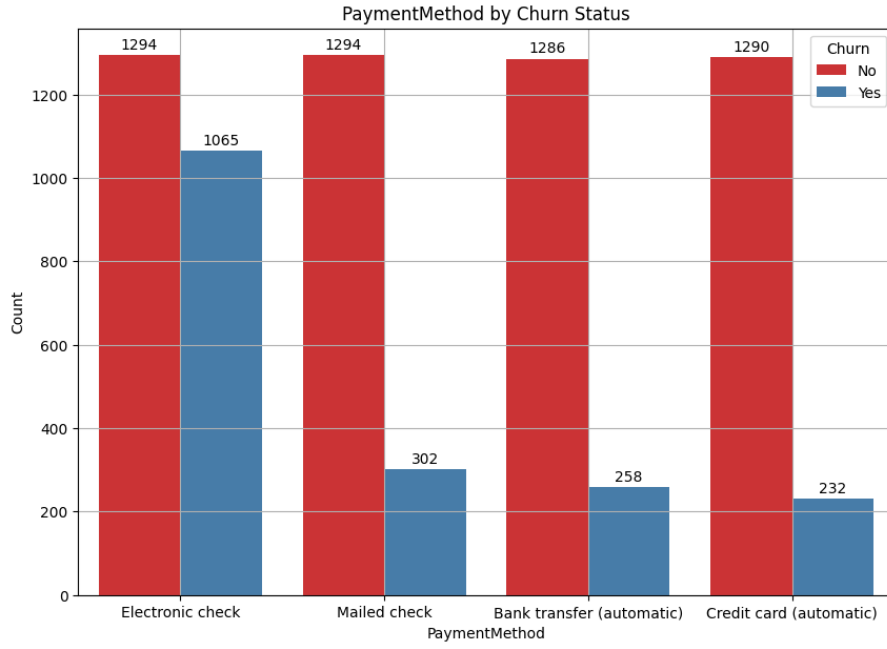
Şekil 4. 6 Contract (Sözleşme Türü) ve Churn İlişkisi

4.2.4 Payment Method (Ödeme Yöntemi) ve Churn İlişkisi

Şekil 4.7’de ödeme yöntemlerinin genel dağılımı verilmiştir. En çok kullanılan ödeme türü 'Electronic check' olup, toplamda 2359 müşteri tarafından tercih edilmiştir. Diğer yöntemler olan 'Mailed check', 'Bank transfer (automatic)' ve 'Credit card (automatic)' yaklaşık olarak eşit oranlarda tercih edilmiştir. Şekil 4.8’de gösterildiği üzere churn oranı, ödeme yöntemi türlerine göre önemli farklılık göstermektedir. 'Electronic check' kullanan müşterilerde churn oranı yaklaşık %45 (1065/2359) iken, otomatik ödeme yöntemleri kullanan müşterilerde bu oran %17’nin altına düşmektedir. Bu durum, düzenli ve otomatik ödeme yapan müşterilerin hizmetten ayrılma olasılığının daha düşük olduğunu göstermektedir. Elektronik çekle ödeme yapan kullanıcıların ise ödeme sürecindeki zorluklardan veya manuel işlem sıklığından dolayı hizmetten ayrılma eğiliminde olabileceği düşünülmektedir.



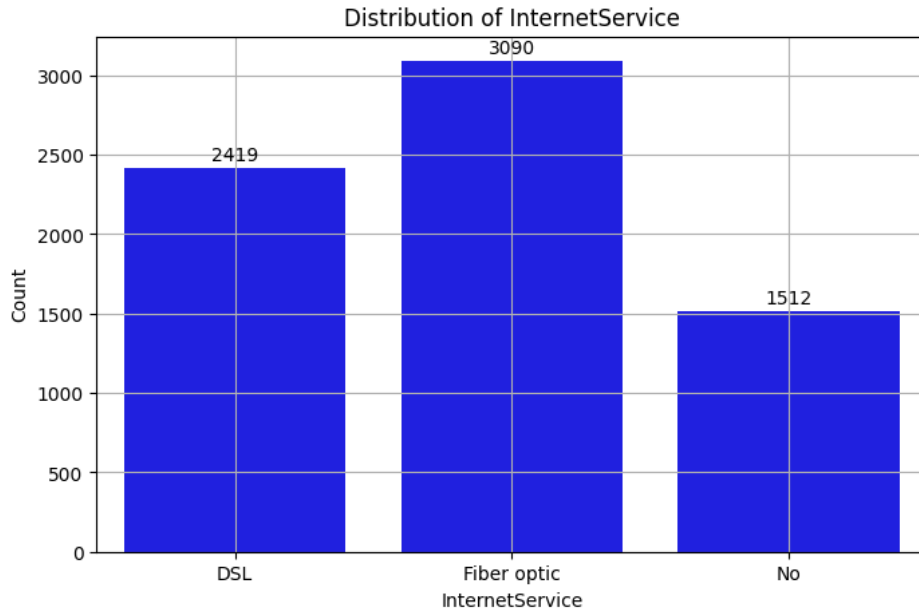
Şekil 4. 7 Payment Method (Ödeme Yöntemi) Dağılımı



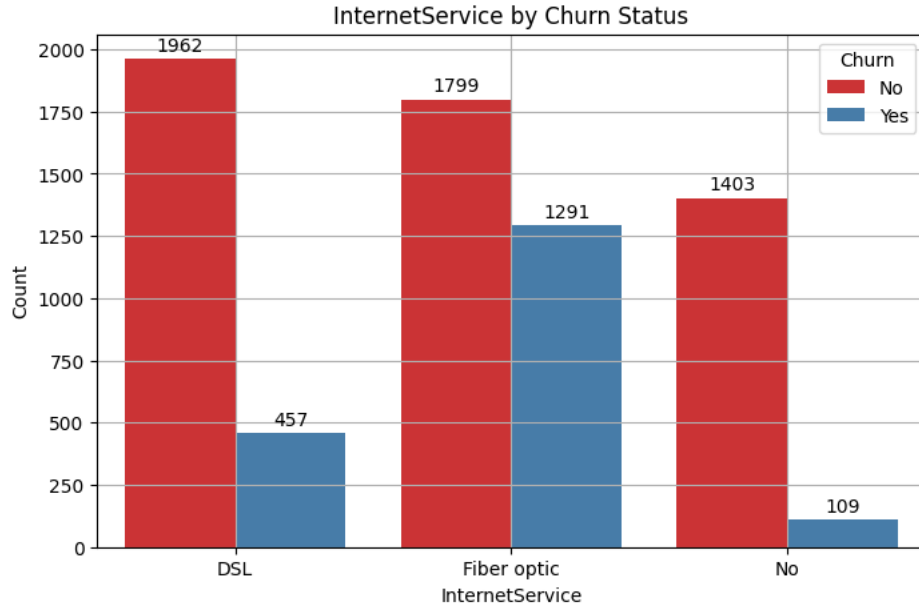
Şekil 4. 8 Payment Method (Ödeme Yöntemi) ve Churn İlişkisi

4.2.5 Internet Service (İnternet Hizmeti) ve Churn İlişkisi

Şekil 4.9’da internet hizmeti türlerinin dağılımı gösterilmiştir. En büyük müşteri grubunu 3090 kişiyle 'Fiber optik' kullanıcıları oluştururken, ardından 2419 DSL ve 1512 internet hizmeti bulunmayan müşteri gelmektedir. Şekil 4.10’da churn durumu internet hizmeti türlerine göre ayrıştırılmıştır. Fiber optik kullanıcılarında churn oranı %41,8 gibi yüksek bir seviyededir (1291/3090). DSL kullanıcılarında bu oran %18,9 civarındayken, hiç internet hizmeti almayan müşterilerde churn oranı oldukça düşüktür (%7,2). Bu durum, fiber optik kullanıcılarının fiyat, hız, teknik sorunlar ya da beklenti yönetimi nedeniyle daha hızlı hizmetten ayrılabilirdiğini; internet hizmeti almayan müşterilerin ise genelde daha düşük riskli ve sadık bir kullanıcı profiline sahip olduğunu göstermektedir.



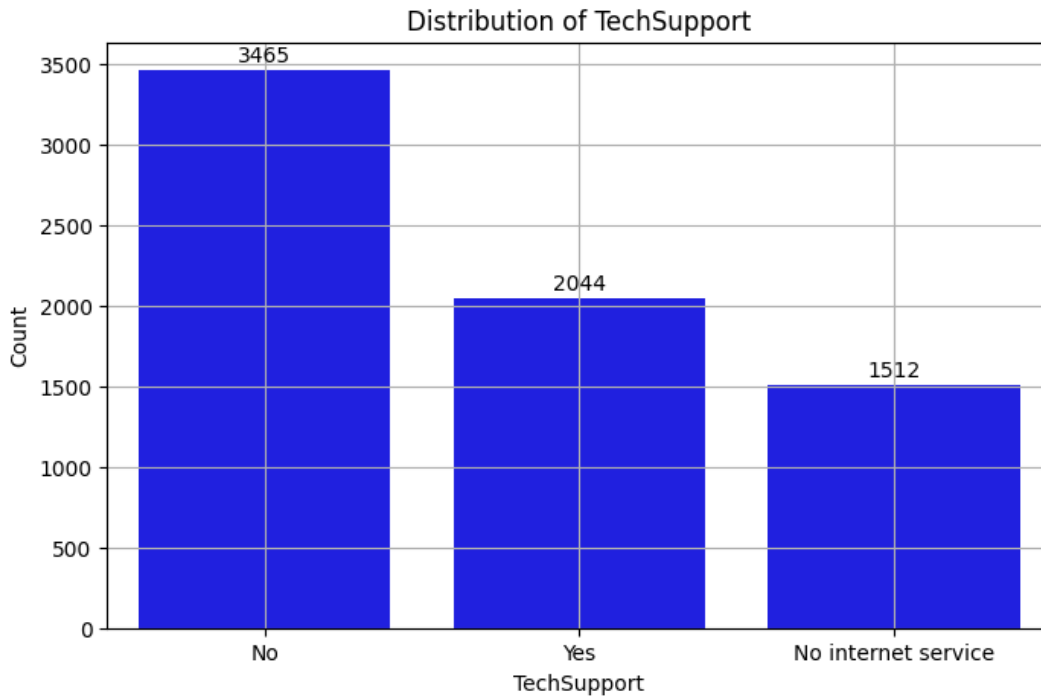
Şekil 4. 9 Internet Service (İnternet Hizmeti) Dağılımı



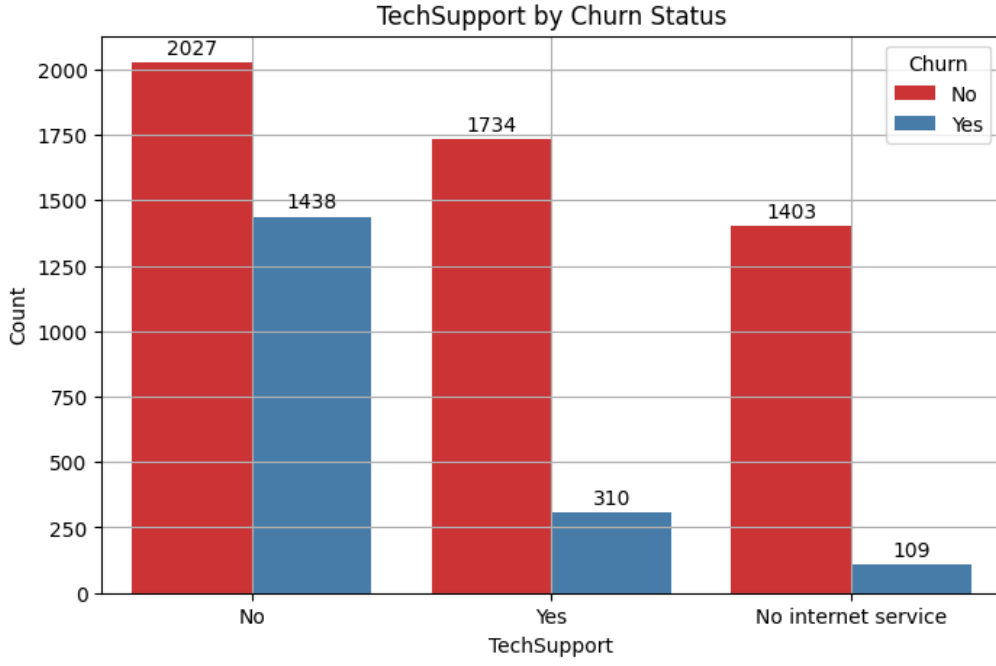
Şekil 4. 10 Internet Service (İnternet Hizmeti) ve Churn İlişkisi

4.2.6 TechSupport (Teknik Destek) ve Churn İlişkisi

Şekil 4.11’de teknik destek hizmeti alma durumu dağılımı verilmiştir. 3465 müşteri teknik destek almamakta, 2044 müşteri bu hizmeti kullanmakta ve 1512 müşteri ise internet hizmeti almadığı için teknik destek durumu geçerli değildir. Şekil 4.12’de görüldüğü üzere teknik destek almayan müşterilerde churn oranı oldukça yüksektir ($1438/3465 \approx \%41,5$). Teknik destek alan müşterilerde bu oran $\%15,2$ ’ye düşmektedir ($310/2044$). Bu durum, teknik destek hizmeti alan müşterilerin memnuniyet seviyesinin daha yüksek olabileceğini ve bu nedenle hizmeti bırakma eğilimlerinin azaldığını göstermektedir. Teknik destek, müşteri sadakatini olumlu yönde etkileyen önemli bir faktör olarak değerlendirilebilir.



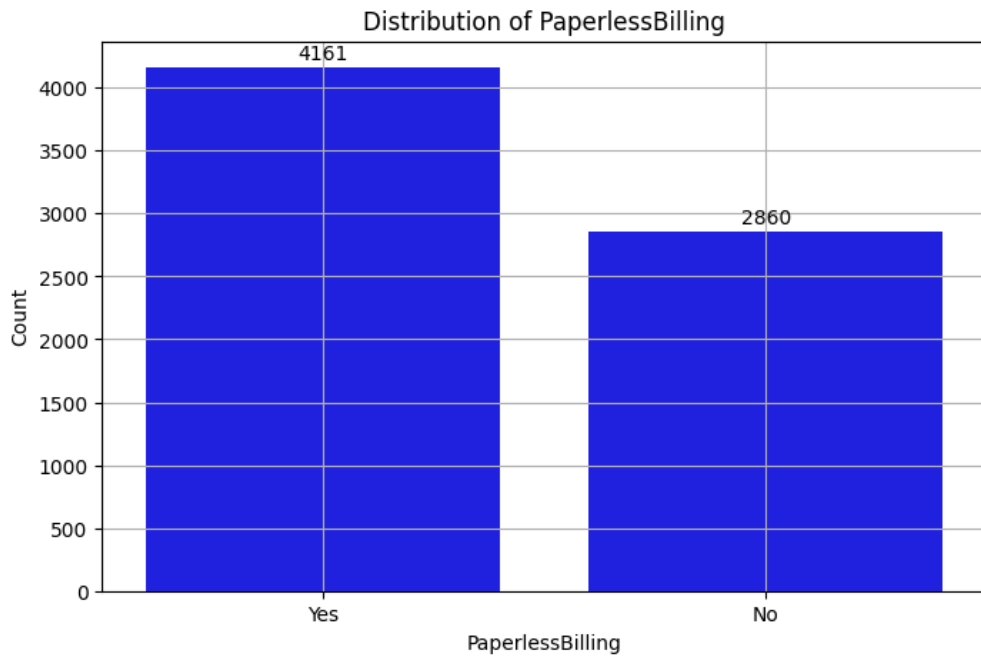
Şekil 4. 11 TechSupport (Teknik Destek) Dağılımı



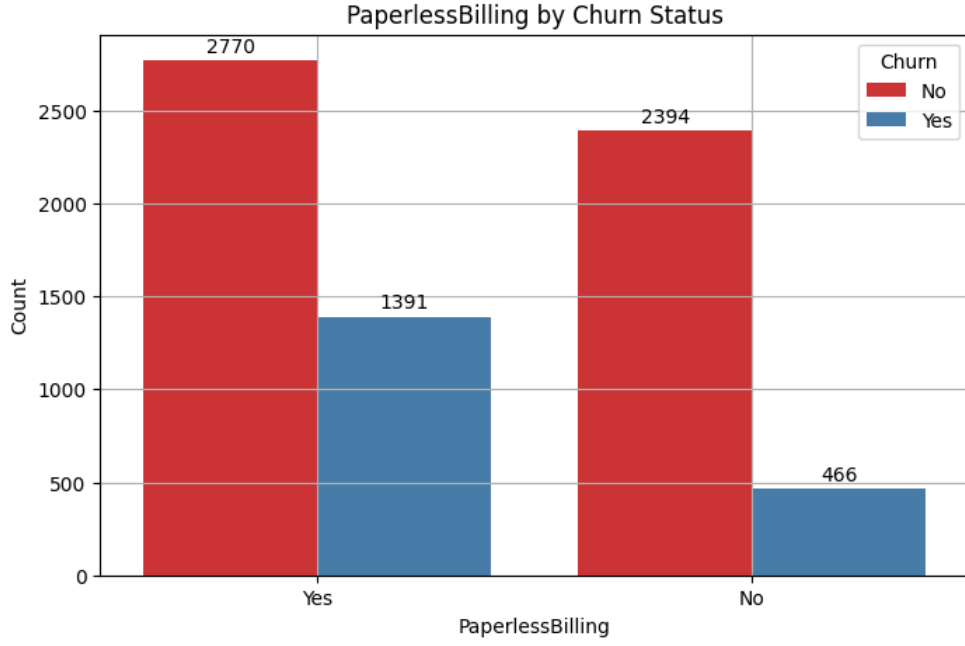
Şekil 4. 12 TechSupport (Teknik Destek) ve Churn İlişkisi

4.2.7 PaperlessBilling (Faturalandırma) ve Churn İlişkisi

Şekil 4.13'te gösterildiği üzere müşterilerin çoğunluğu (4161 kişi) kağıtsız fatura (PaperlessBilling) sistemini kullanmaktadır. 2860 müşteri ise geleneksel faturalandırma yöntemini tercih etmektedir. Şekil 4.14'te ise faturalandırma yöntemi ile churn durumu arasındaki ilişki görselleştirilmiştir. PaperlessBilling kullanan müşterilerde churn oranı %25,7 iken (1391/5411), bu oran kâğıt fatura tercih eden müşterilerde %16,3 civarındadır (466/2860). Bu durum, kağıtsız fatura kullanıcılarının hizmetten daha kolay ayrılabilmesini, dijital fatura kullanımı ile dijital iletişimin yoğunluğunun churn eğilimini artırabileceğini göstermektedir. Alternatif olarak, PaperlessBilling kullanan müşterilerin aynı zamanda daha kısa süreli sözleşme veya yüksek fatura yükü gibi diğer churn artırıcı faktörlerle de ilişkili olabileceği düşünülmektedir.



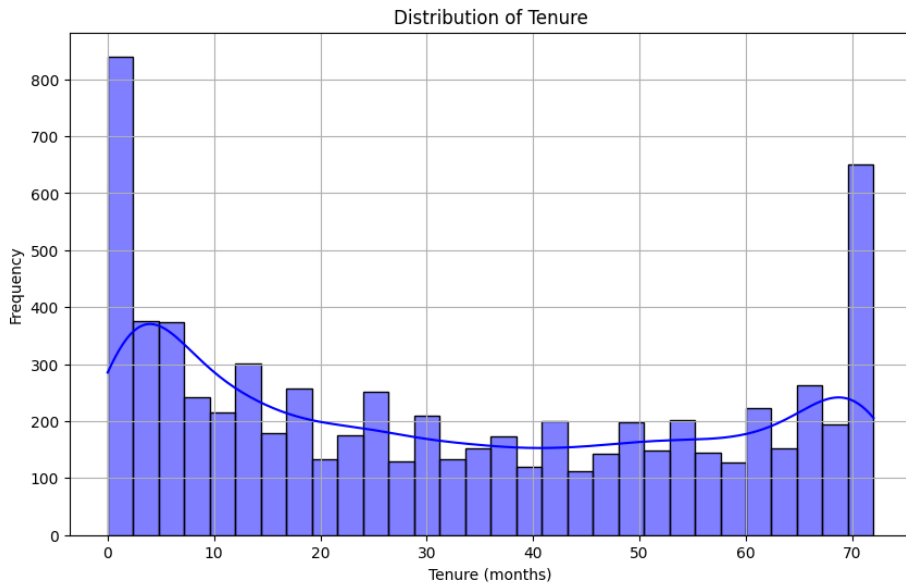
Şekil 4. 13 PaperBilling (Faturalandırma) Dağılımı



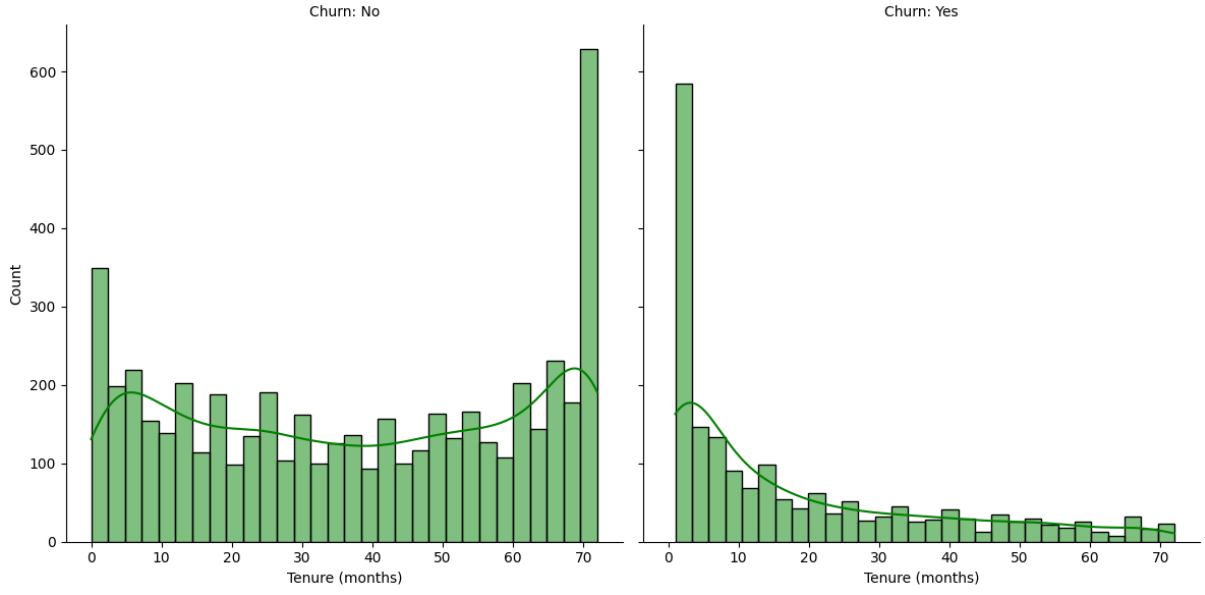
Şekil 4. 14 PaperBilling (Faturalandırma) ve Churn İlişkisi

4.2.8 Tenure (Hizmet Süresi) ve Churn İlişkisi

Şekil 4.15’te müşteri hizmet süresinin (tenure) dağılımı verilmiştir. Grafik incelendiğinde, çok kısa süreli (0–1 ay) ve çok uzun süreli (70 ay üzeri) müşterilerin sayısal olarak öne çıktığı görülmektedir. Bu durum, müşteri kitlesinin bir kısmının çok kısa süreli abonelikten sonra hizmeti bıraktığını, bir kısmının ise oldukça uzun süreli sadık müşteri profiline sahip olduğunu göstermektedir. Şekil 4.16’da churn durumuna göre tenure dağılımı iki ayrı grafikte sunulmuştur. Hizmeti bırakmayan müşterilerin çoğunluğu uzun hizmet süresine sahipken, churn eden müşterilerin hizmet süresinin genellikle ilk 0–15 ay arasında yoğunlaştığı gözlemlenmiştir. Bu bulgu, müşteri sadakatının zamanla arttığını ve ilk aylarda yaşanan deneyimin churn kararında belirleyici olabileceğini göstermektedir. Uzun vadeli müşteri memnuniyetinin artırılması churn oranını azaltmak açısından kritik öneme sahiptir.



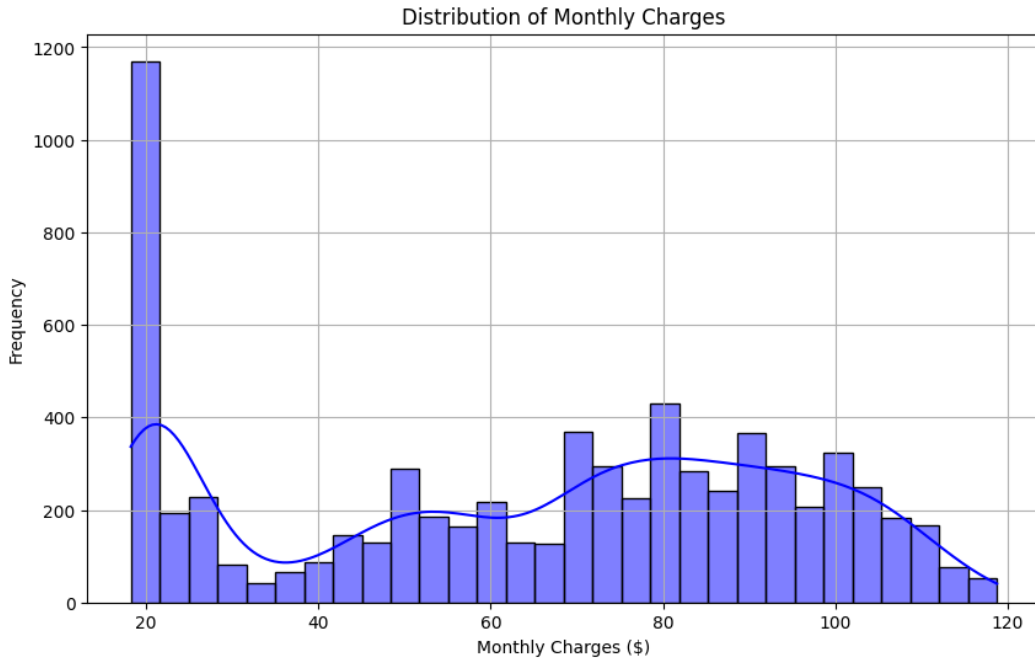
Şekil 4. 15 Tenure (Hizmet Süresi) Dağılımı



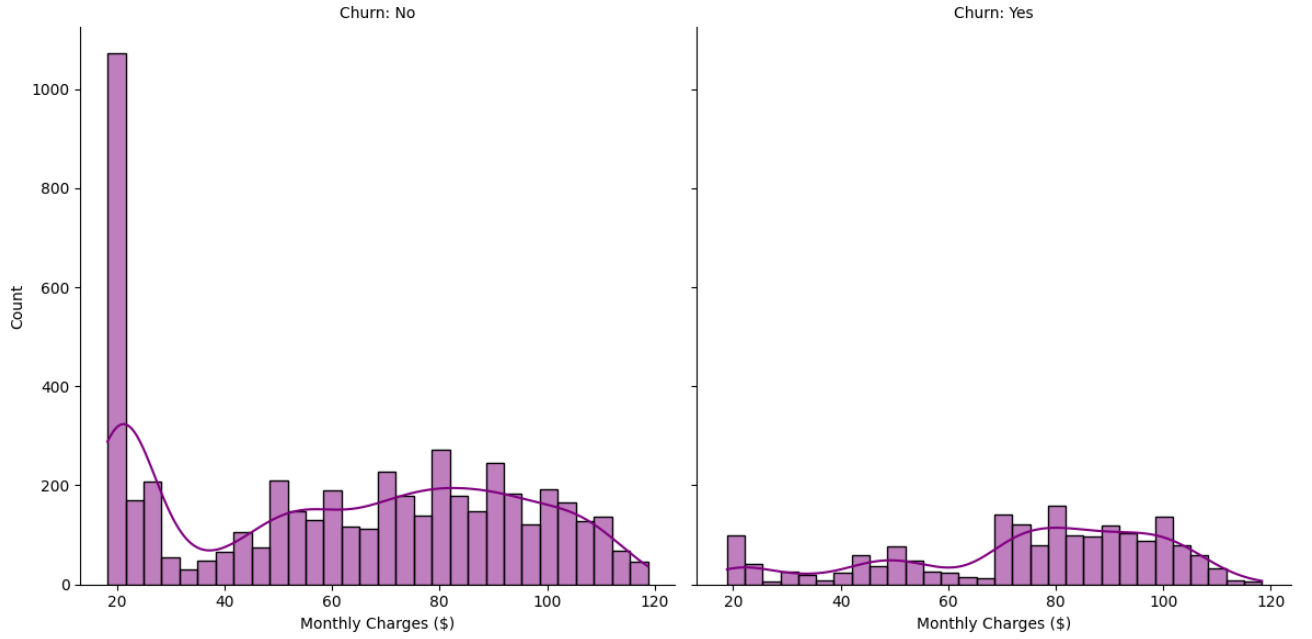
Şekil 4. 16 Tenure (Hizmet Süresi) ve Churn İlişkisi

4.2.9 Monthly Charges (Aylık Ücretler) ve Churn İlişkisi

Şekil 4.17’de müşteri başına aylık ücretlerin dağılımı gösterilmiştir. Grafik, düşük ücretli segmentte (özellikle 20 dolar civarında) müşteri yoğunluğunun yüksek olduğunu, ayrıca 70–100 dolar aralığında da belirgin bir yığılma olduğunu göstermektedir. Bu dağılım hem düşük hem de orta-yüksek ücretli paketlerin yoğunlukla tercih edildiğini ortaya koymaktadır. Şekil 4.18’de churn durumuna göre aylık ücret dağılımları görselleştirilmiştir. Churn eden müşterilerin çoğunlukla yüksek faturalı segmentte (70 dolar ve üzeri) yoğunlaştığı, buna karşın churn etmeyen müşterilerin önemli kısmının düşük ücretli grupta yer aldığı gözlemlenmiştir. Bu durum, yüksek fatura bedelinin müşteri memnuniyetini olumsuz etkileyebileceğini ve churn riskini artırabileceğini göstermektedir. Aylık ücretin, hizmete devam kararında belirleyici bir etmen olabileceği değerlendirilmiştir.



Şekil 4. 17 Montly Charges (Aylık Ücretler) Dağılımı

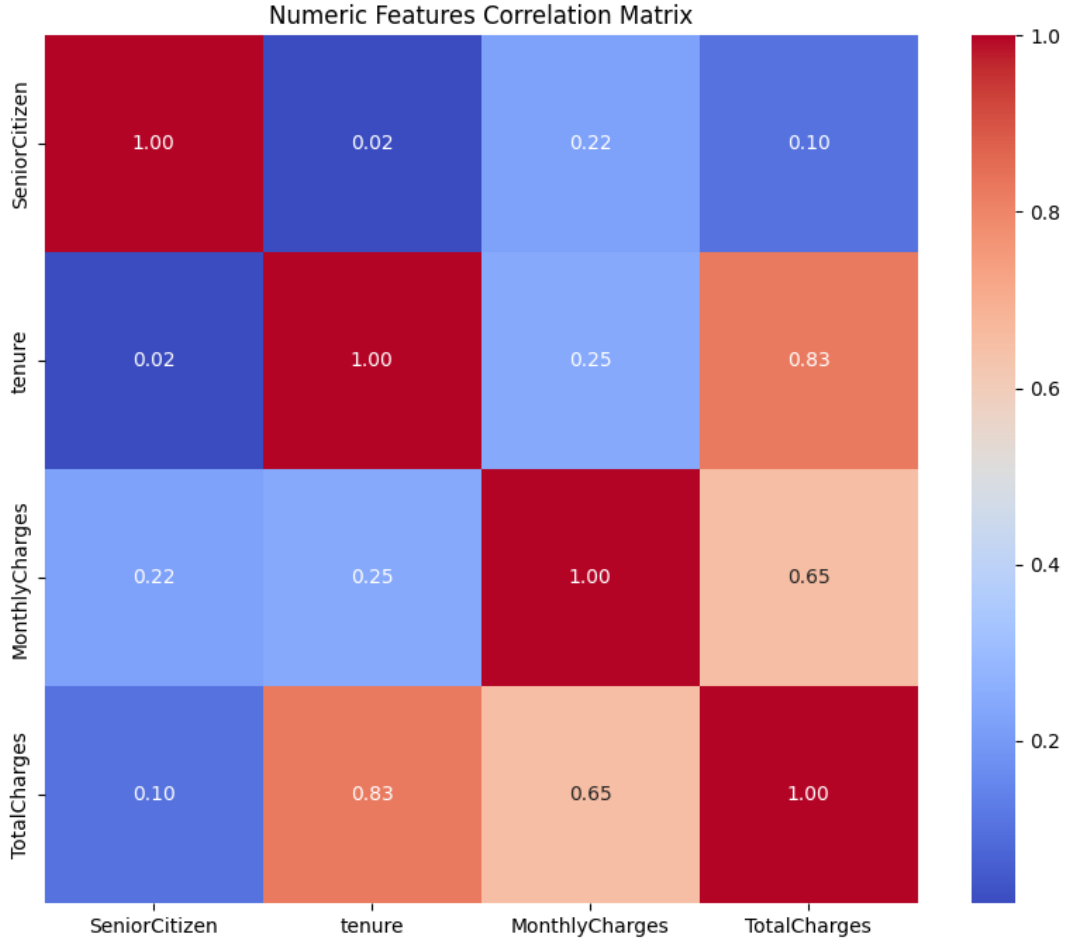


Şekil 4. 18 Montly Charges (Aylık Ücretler) ve Churn İlişkisi

4.2.10 Korelasyon Analizi

Şekil 4.19’da sayısal değişkenler arasındaki korelasyon matrisi görselleştirilmiştir. Bu matris, her bir sayısal değişken çiftinin arasındaki doğrusal ilişki düzeyini göstermekte olup, 1.00 değeri tam pozitif korelasyonu, 0 ise ilişkisizliği ifade etmektedir. Korelasyon analizi, veri setindeki değişkenlerin birbirini ne ölçüde açıkladığını görmek ve çoklu doğrusal bağlantı risklerini öngörmek amacıyla uygulanmıştır. Grafik incelendiğinde en yüksek korelasyonun tenure (hizmet süresi) ile TotalCharges (toplam ödeme) arasında olduğu ($r = 0.83$) gözlemlenmektedir. Bu oldukça güçlü pozitif ilişki, müşterinin hizmeti daha uzun süre kullanması durumunda doğal olarak toplam ödemesinin de arttığını göstermektedir. Bu iki değişkenin yüksek derecede bağlantılı olması, model oluşturma sürecinde dikkat edilmesi gereken bir noktadır. MonthlyCharges (aylık ücret) ile TotalCharges arasında da güçlü sayılabilecek bir korelasyon ($r = 0.65$) bulunmuştur. Bu ilişki, müşterilerin aylık faturalarının yüksek olmasının, belirli bir hizmet süresinde toplam ödemelerine de yansıdığını göstermektedir. Öte yandan SeniorCitizen değişkeni hem tenure hem de diğer sayısal değişkenlerle çok zayıf korelasyon göstermektedir (maksimum $r = 0.22$). Bu durum, yaş değişkeninin doğrudan bir finansal büyüklükle ilişki taşımadığını ve bağımsız bir etkene sahip olabileceğini düşündürmektedir.

Sonuç olarak, korelasyon analizi hem öznetelik mühendisliği sürecinde dikkat edilmesi gereken ilişkileri belirlemede, hem de çoklu doğrusal bağlantının (multicollinearity) olasılığını değerlendirmede önemli bir araç olarak kullanılmıştır.



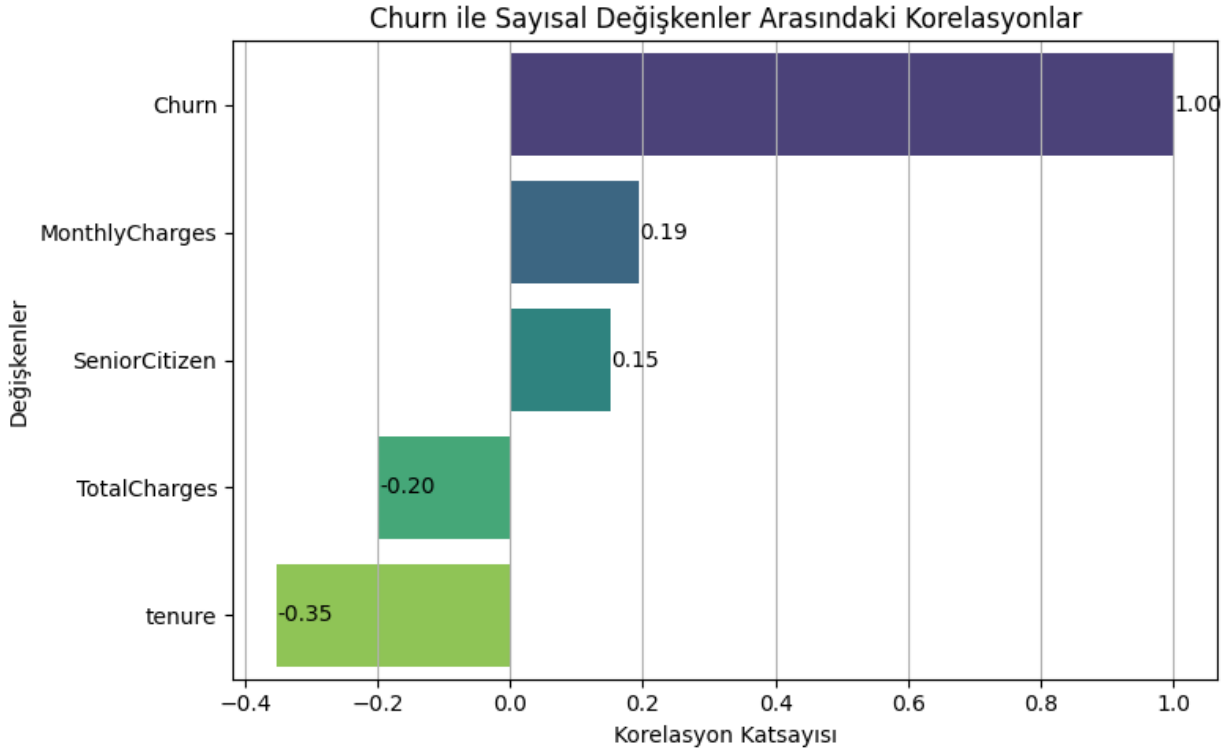
Şekil 4. 19 Sayısal Değişkenler Arasındaki Korelasyon Matrisi

4.2.11 Churn ile Sayısal Değişkenler Arasındaki Korelasyon İlişkisi

Şekil 4.20’de churn değişkeni ile diğer sayısal değişkenler arasındaki korelasyon katsayıları görselleştirilmiştir. Bu analiz, churn’ün doğrudan etkilenebileceği sayısal faktörleri belirlemek açısından önemlidir.

- **Tenure (hizmet süresi):** Churn ile en yüksek doğrusal ilişkiye sahip değişkendir. Korelasyon katsayısı-0.35 olarak bulunmuştur. Bu negatif ilişki, hizmet süresi arttıkça müşterinin hizmeti bırakma olasılığının azaldığını göstermektedir. Dolayısıyla sadık müşterilerin churn etme ihtimali daha düşüktür.
- **TotalCharges (toplam ödeme):** Bu değişkenin churn ile olan korelasyon katsayısı-0.20’dir. Yani, daha fazla toplam ödeme yapmış müşteriler churn etmeye daha az meyillidir. Bu bulgu da müşteri bağlılığıyla ilişkilidir.
- **SeniorCitizen:** Korelasyon katsayısı 0.15’tir ve pozitif yöndedir. Bu durum, yaşlı müşterilerin hizmetten ayrılma olasılığının bir miktar daha fazla olabileceğini işaret etmektedir. Ancak bu ilişkinin zayıf olduğu unutulmamalıdır.
- **MonthlyCharges (aylık ücret):** Bu değişkenin churn ile korelasyonu 0.19’dur. Aylık ücreti yüksek olan müşterilerin churn riski daha fazladır. Bu durum, yüksek faturalandırmanın müşteri memnuniyetini olumsuz etkileyebileceğini göstermektedir.

Sonuç olarak, churn ile doğrudan ilişkili olan en önemli sayısal faktörün tenure olduğu görülmektedir. Diğer değişkenlerle olan ilişkiler daha zayıf düzeyde olsa da churn tahmin modellerinde dikkate alınmalıdır.



Şekil 4. 20 Churn ile Sayısal Değişkenler Arasındaki Korelasyonlar

4.3 Model Çıktıları ve Özetleri

Bu bölümde, çalışmada kullanılan beş farklı sınıflandırma algoritmasının performans çıktıları detaylı şekilde sunulmaktadır. Değerlendirme sürecinde her bir model, SMOTE yöntemi ile sınıf dengesi sağlanmış, optimal threshold ayarı yapılmış ve 5 katlı çapraz doğrulama (K-Fold CV) yöntemi kullanılarak test edilmiştir. Böylece modellerin genelleme yetenekleri güvenilir bir şekilde ölçülmüştür.

Her modelin doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall), F1-Score ve AUC (Area Under the Curve) metrikleri elde edilmiş ve yorumlanmıştır. Son olarak tüm modellerin sonuçları karşılaştırmalı bir tablo halinde sunulmuş ve genel performans değerlendirmesi yapılmıştır.

4.3.1 Logistic Regression (Lojistik Regresyon)

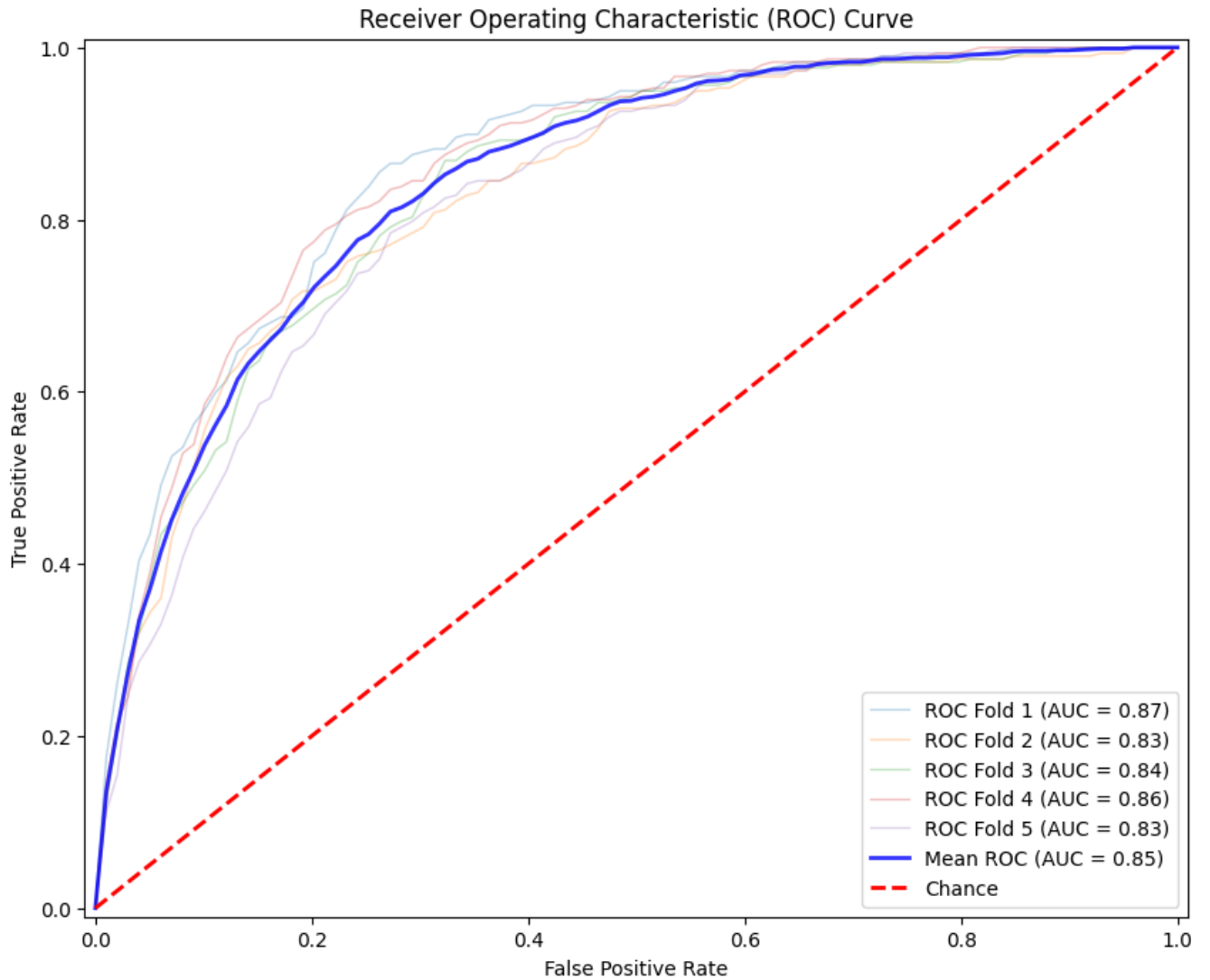
Bu bölümde, Logistic Regression modelinin SMOTE uygulanmadan önceki temel haliyle gösterdiği performans sonuçları detaylı olarak sunulmaktadır. Model, 5 katlı çapraz doğrulama (K-Fold Cross Validation) yöntemiyle değerlendirilmiş, ardından tüm eğitim verisiyle eğitilerek test verisi üzerindeki başarımı analiz edilmiştir.

Çapraz doğrulama sonuçları (Tablo 4.1) incelendiğinde, modelin ortalama doğruluk oranı %74,86, ortalama F1-Score değeri %62,92 ve ortalama AUC (Area Under Curve) skoru %85 olarak elde edilmiştir. Bu sonuçlar, Logistic Regression modelinin dengesiz veri yapısına rağmen sınıflandırma görevinde makul bir performans gösterdiğini ortaya koymaktadır. Özellikle ortalama recall değerinin %80,67 gibi yüksek bir seviyede olması, churn eden müşterilerin büyük bir kısmını doğru tahmin edebildiğini göstermektedir. Aynı zamanda modelin aşırı öğrenme (overfitting) yapmadığını foldlardaki sonuçların birbirine yakınlığından göstermektedir.

Tablo 4. 1 Lojistik Reg. Modeli - Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7696	0.7364	0.7337	0.7605	0.7427	0.7486
PRECİSİON	0.5404	0.5011	0.4979	0.5304	0.5087	0.5157
RECALL	0.8552	0.7710	0.7980	0.8215	0.7879	0.8067
F1-SCORE	0.6623	0.6074	0.6132	0.6446	0.6182	0.6292
AUC	0.87	0.83	0.84	0.86	0.83	0.85

ROC eğrisi (Şekil 4.21) üzerinden değerlendirildiğinde her bir fold için AUC skorları sırasıyla 0.87, 0.83, 0.84, 0.86 ve 0.83 şeklindedir. Bu da modelin genel ayırt edicilik gücünün istikrarlı olduğunu göstermektedir. Ortalama ROC eğrisi altında kalan alan %85 olarak hesaplanmıştır.

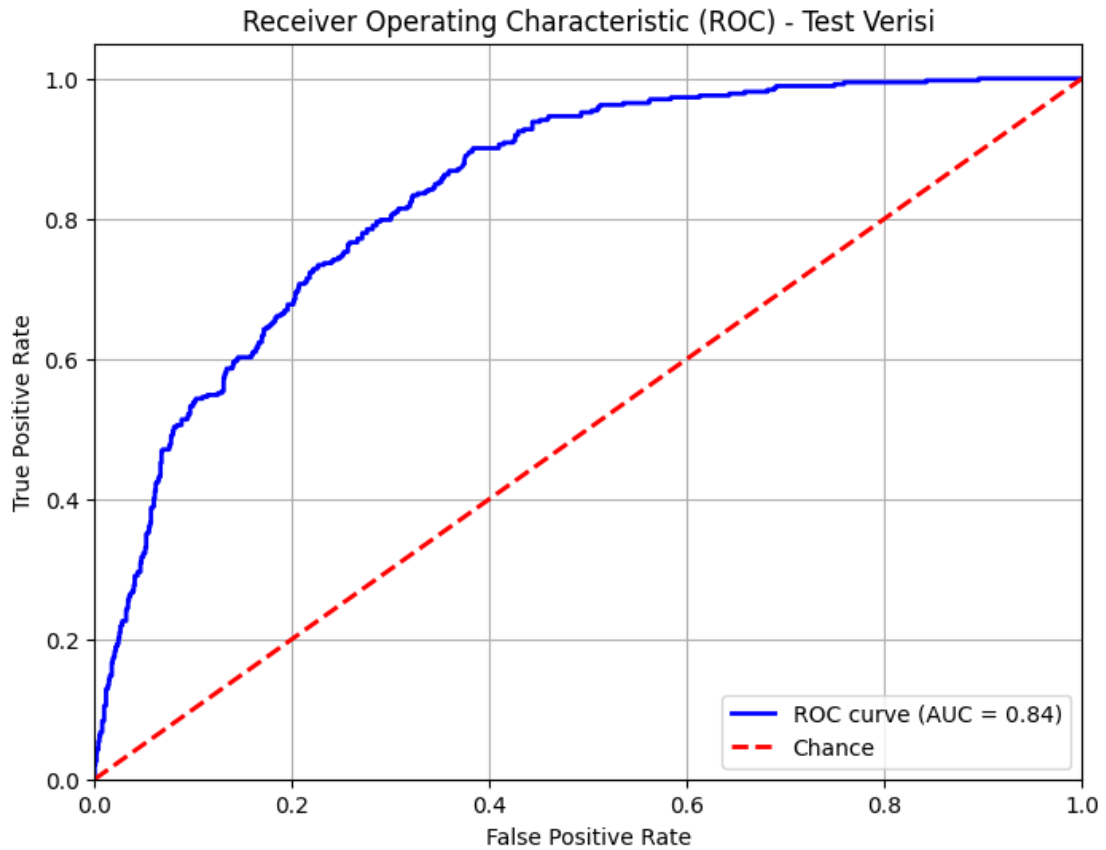


Şekil 4. 21 Lojistik Reg. Modeli - Foldların ROC Eğrisi

Test verisi üzerindeki deęerlendirmeye (Tablo 4.2) geildięinde modelin genel doęruluk oranı %74.09, precision deęeri %50,70, recall deęeri %77,96 ve F1-Score deęeri %61,44 olarak ölçölmüştür. Özellikle churn sınıfındaki yüksek recall deęeri modelin bu sınıfı yakalamakta etkili olduęunu göstermektedir. Bununla birlikte precision deęerinin görece düşük olması, modelin yanlış pozitif tahmin oranının yüksek olabileceğini düşöndürmektedir.

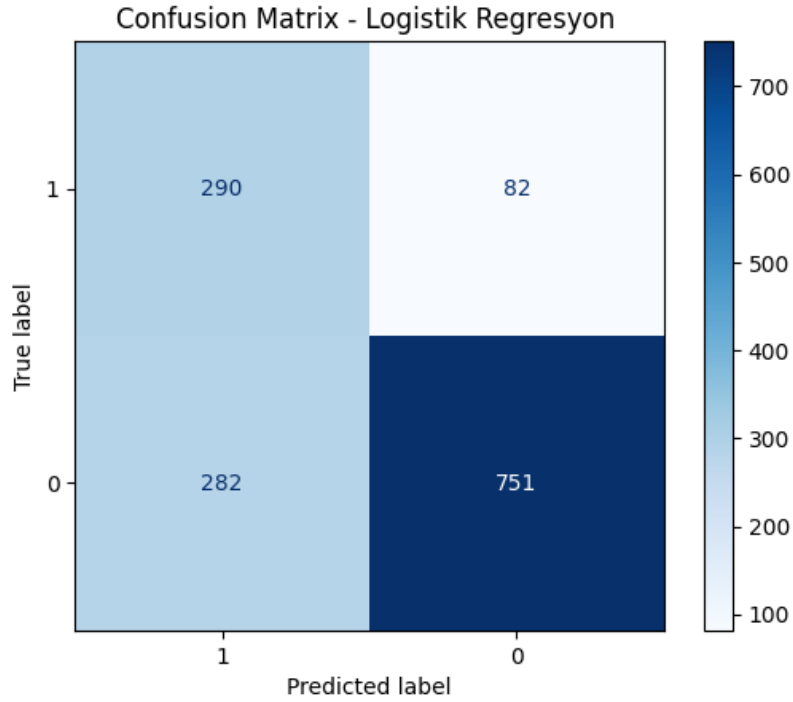
Tablo 4. 2 Lojistik Reg. Modeli - Test Verisi Performansı

METRİK	DEęER
ACCURACY	0.7409
PRECİSİÖN	0.5070
RECALL	0.7796
F1-SCORE	0.6144
AUC	0.84



Şekil 4. 22 Lojistik Regresyon Modeli - Test Verisi ROC Eęrisi

Karışıklık matrisi (confusion matrix) (Şekil 4.22) incelendięinde, modelin 290 churn müşterisini doęru tahmin ettięi, ancak 282 churn olmayan müşteriyi yanlış şekilde churn olarak tahmin ettięi görölmektedir. Bu, modelin churn tahmininde hassas (sensitive) olduęunu fakat false positive sayısının da dikkate alınması gerektiğini ortaya koymaktadır.



Şekil 4. 23 Lojistik Regresyon - Confusion Matrix

Ayrıca detaylı sınıflandırma raporunda (Tablo 4.3) churn olan sınıf için precision değeri 0.51, recall değeri 0.78, F1-Score ise 0,61 olarak görülmektedir. Bu değerler, dengesiz veri yapısında Logistic Regression modelinin, threshold ayarlaması yapılmadan da makul ölçüde iyi sonuçlar verebildiğini göstermektedir.

Tablo 4. 3 Lojistik Reg. Modeli - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.90	0.73	0.80	1033
Churned	0.51	0.78	0.61	372
ACCURACY			0.74	1405
Macro AVG	0.70	0.75	0.71	1405
Weighted AVG	0.80	0.74	0.75	1405

Sonuç olarak, Logistic Regression modeli churn sınıfını tanımlamada güçlü bir temel sağlamıştır. Ancak precision değerinin görece düşük olması, bu modelin daha dengeli bir sonuç sunabilmesi için SMOTE gibi veri dengeleme yöntemlerine ve threshold optimizasyonuna ihtiyaç duyabileceğini göstermektedir.

4.3.1.1 Logistic Regression (Lojistik Regresyon) – SMOTE + Threshold Uygulamalı

Bu bölümde, Logistic Regression modeline SMOTE uygulaması, optimal threshold belirleme ve 5 katlı çapraz doğrulama yöntemi entegre edilerek elde edilen sonuçlar analiz edilmiştir. Bu yapılandırma, veri dengesizliğini azaltmak, sınıflandırma kararlarında eşik esnekliği sağlamak ve modelin genellenebilirliğini artırmak için tasarlanmıştır.

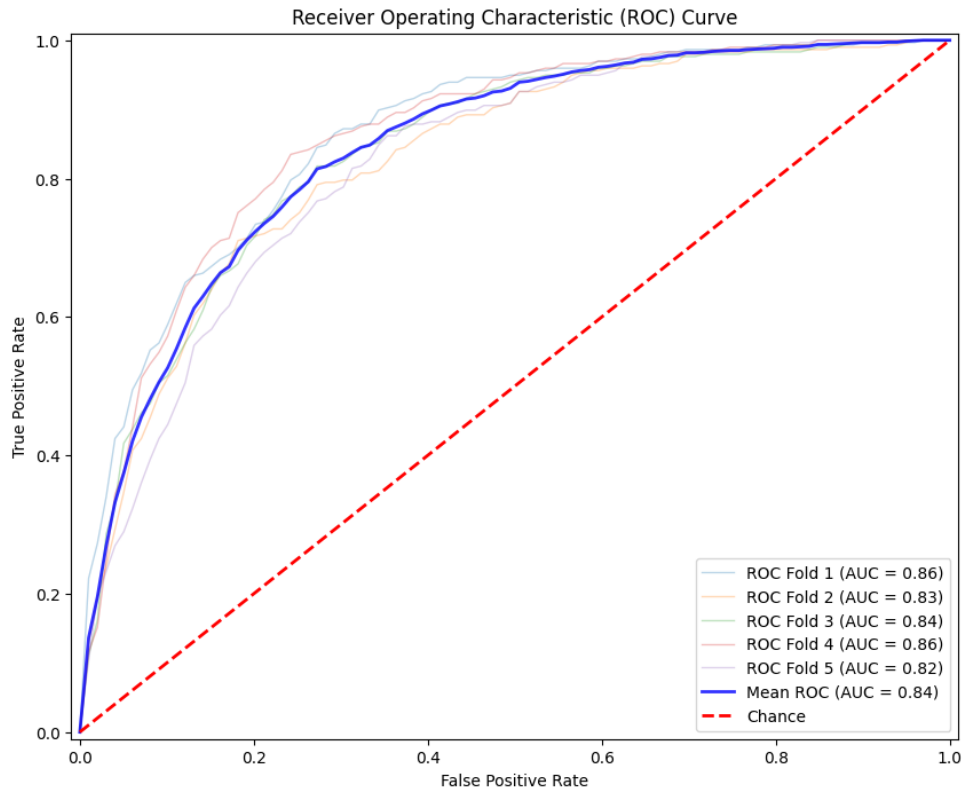
Model, her bir fold'da eğitim verisine SMOTE uygulayarak sınıf dağılımını dengeli hale getirmiştir (churn = 0 ve churn = 1 örnek sayısı eşitlenmiştir). Ardından, tahmin edilen olasılık skorları üzerinden farklı eşik değerleri denenmiş ve her fold için en iyi F1-Score'u veren optimal threshold değeri belirlenmiştir. Fold'lar arası optimal threshold değerleri 0.38 ile 0.53 arasında değişmiştir. Bu değişkenlik, modelin veri bölünmesine duyarlılığını ve sınıflandırma kararlarının eşik ayarına bağlılığını göstermektedir.

Çapraz doğrulama sonuçları (Tablo 4.4) değerlendirildiğinde; modelin ortalama doğruluğu %78,92, ortalama precision değeri %58,67, ortalama recall değeri %71,52, ortalama F1-Score değeri %64,23 ve ortalama AUC skoru %84 olarak ölçülmüştür. Bu metrikler, temel modele kıyasla daha dengeli bir performans sunulduğunu göstermektedir.

Tablo 4. 4 Lojistik Reg. Modeli - SMOTE Uygulanan Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.8185	0.7890	0.7676	0.7988	0.7720	0.7892
PRECİSİON	0.6587	0.5824	0.5425	0.5932	0.5568	0.5867
RECALL	0.6498	0.7138	0.7744	0.7609	0.6768	0.7152
F1-SCORE	0.6542	0.6415	0.6380	0.6667	0.6109	0.6423
AUC	0.86	0.83	0.84	0.86	0.82	0.84
OPTİMAL THRESHOLD	0.5374	0.4340	0.3845	0.4224	0.4361	0.44287941428029554

ROC eğrisi (Şekil 4.23) altında kalan alanlar fold'lara göre 0.82–0,86 arasında değişmekte olup genel ortalama yine %84 olarak elde edilmiştir. Bu durum, modelin ayırt edici gücünün korunduğunu göstermektedir.

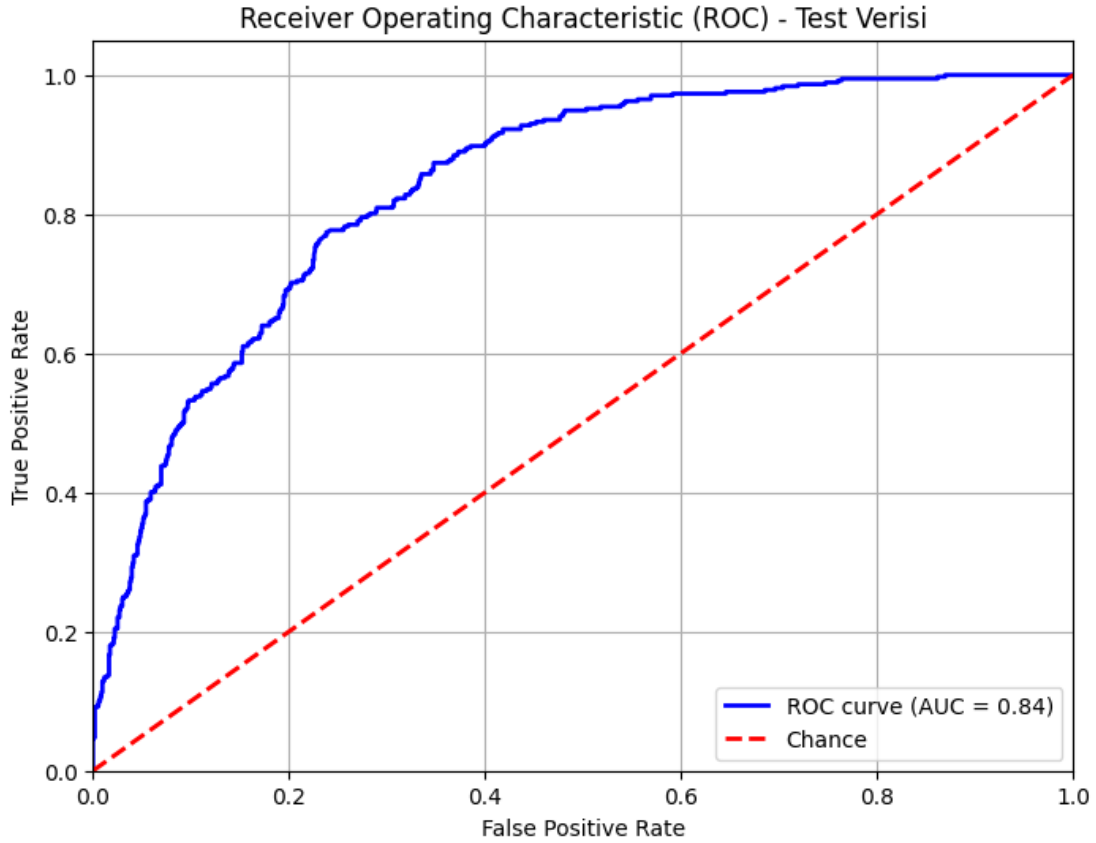


Şekil 4. 24 Lojistik Reg. Modeli - SMOTE Uygulanan Foldların ROC Eğrisi

Test verisi üzerindeki performansa (Tablo 4.5) bakıldığında modelin doğruluk oranı %77.01, precision değeri %55,43, recall değeri %67,20, F1-Score değeri %60,75 ve AUC skoru %84 olarak gözlemlenmiştir. Bu sonuçlar, modelin test verisi üzerinde de oldukça tutarlı ve dengeli çalıştığını kanıtlamaktadır. Optimal threshold bu aşamada 0.4428 olarak belirlenmiş ve bu eşik ile sınıflandırma gerçekleştirilmiştir.

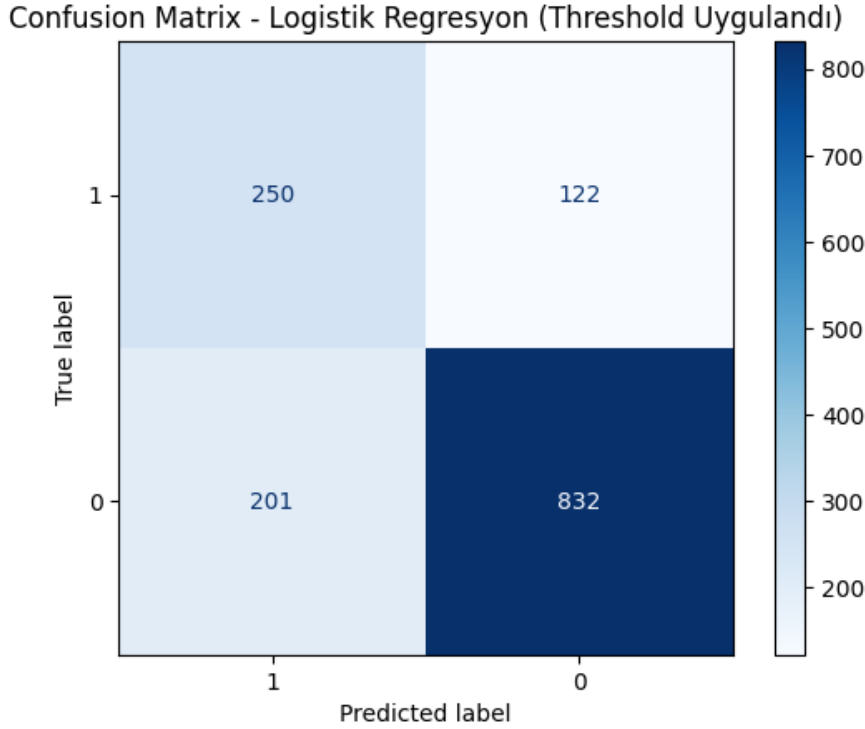
Tablo 4. 5 Lojistik Reg. Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)

METRİK	DEĞER
ACCURACY	0.7701
PRECİSİON	0.5543
RECALL	0.6720
F1-SCORE	0.6075
AUC	0.84
OPTİMAL THRESHOLD	0.44287941428029554



Şekil 4. 25 Lojistik Reg. Modeli (SMOTE) - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.24), modelin 250 churn eden müşteriyi doğru tahmin ettiği, ancak 122 churn eden müşteriyi gözden kaçırdığı (false negative) ve 201 churn etmeyen müşteriyi yanlış churn olarak sınıflandırdığı (false positive) anlaşılmaktadır. Bu durum, modelin recall'ı artırmak uğruna bazı yanlış pozitiflere tolerans gösterdiğini, ancak yine de aşırı hata üretmediğini göstermektedir.



Şekil 4. 26 Lojistik Regresyon - Confusion Matrix (Threshold Uygulandı)

Detaylı sınıflandırma raporu (Tablo 4.6) incelendiğinde churn sınıfı için precision değeri %55, recall %67 ve F1-score %61 olarak elde edilmiştir. SMOTE uygulanmamış versiyona kıyasla precision ve F1-score değerlerinde iyileşme gözlemlenmiştir. Ayrıca non-churn sınıfındaki precision ve recall değerlerinin de yüksek kalması, modelin genel başarısının sınıf dengesinden olumsuz etkilenmediğini ortaya koymaktadır.

Tablo 4. 6 Lojistik Reg. Modeli (SMOTE) - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.87	0.81	0.84	1033
Churned	0.55	0.67	0.61	372
ACCURACY			0.77	1405
Macro AVG	0.71	0.74	0.72	1405
Weighted AVG	0.79	0.77	0.78	1405

Sonuç olarak, SMOTE + Threshold + K-Fold yapısı Logistic Regression modelinin performansını genel olarak iyileştirmiştir. Overfitting riski ise düşük seviyede kalmıştır; zira eğitim süreci sırasında farklı fold'larda modelin metrikleri birbirine yakın çıkmış ve test seti performansı da bu başarıyı teyit etmiştir. Modelin bu versiyonu, churn tahmini için daha kararlı ve güvenilir bir alternatif sunmaktadır.

4.3.2 Decision Tree (Karar Ağacı)

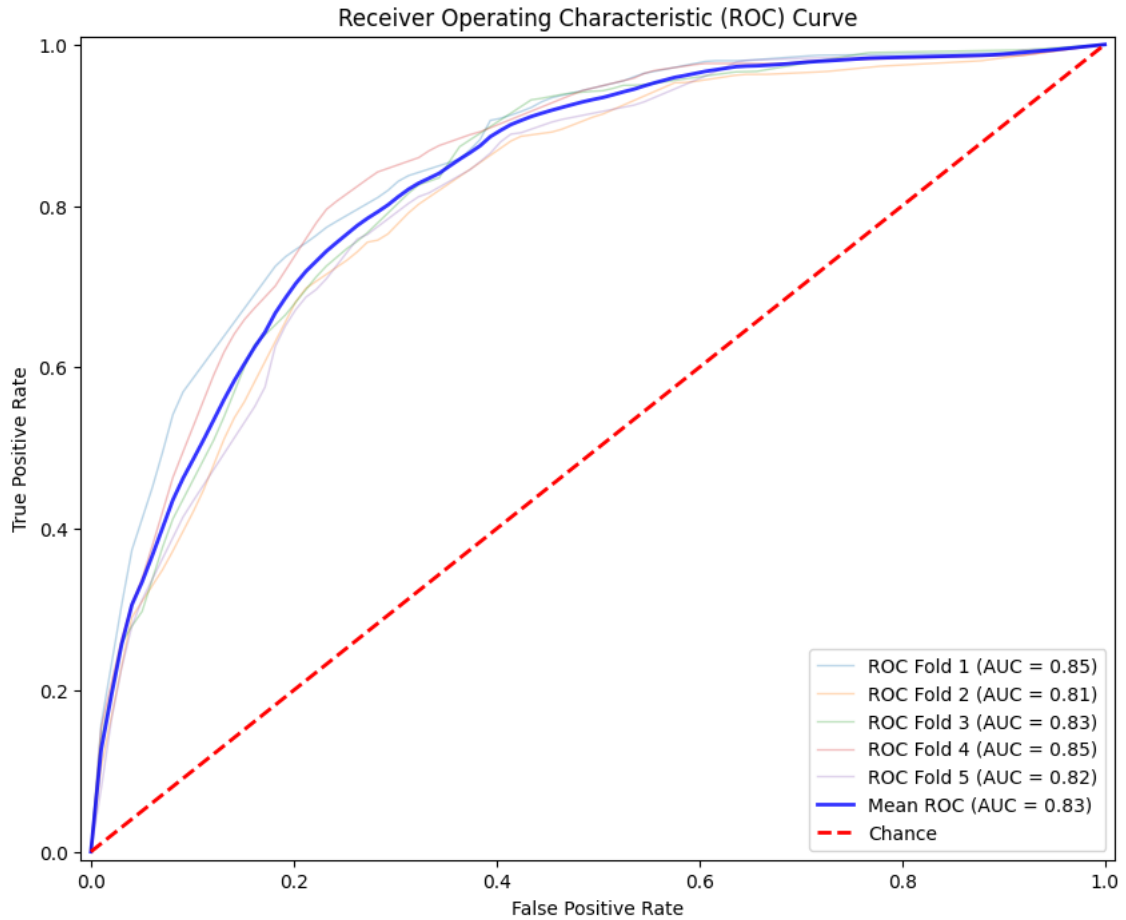
Bu bölümde, Decision Tree (Karar Ağacı) algoritmasının temel haliyle, yani SMOTE ve threshold ayarlaması yapılmadan elde edilen performans sonuçları değerlendirilmektedir. Model 5 katlı çapraz doğrulama yöntemiyle test edilmiştir. Değerlendirme kriteri olarak accuracy, precision, recall, F1-score ve AUC metrikleri kullanılmıştır.

Çapraz doğrulama sonuçları (Tablo 4.7) incelendiğinde modelin ortalama doğruluk oranı %74,79, ortalama F1-Score değeri %62.06 ve ortalama AUC değeri %83 olarak ölçülmüştür.

Tablo 4. 7 Karar Ağacı Modeli - Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7687	0.7124	0.7409	0.7747	0.7427	0.7479
PRECİSİON	0.5437	0.4740	0.5068	0.5512	0.5090	0.5169
RECALL	0.7744	0.7980	0.7576	0.7980	0.7609	0.7778
F1-SCORE	0.6389	0.5947	0.6073	0.6520	0.6100	0.6206
AUC	0.85	0.81	0.83	0.85	0.82	0.83

ROC eğrisi (Şekil 4.27) fold'lara göre 0.81 ila 0,85 arasında değişmiş olup, genel ortalaması %83 olarak hesaplanmıştır. Bu sonuçlar modelin istikrarlı bir ayrıştırma yeteneğine sahip olduğunu göstermektedir.

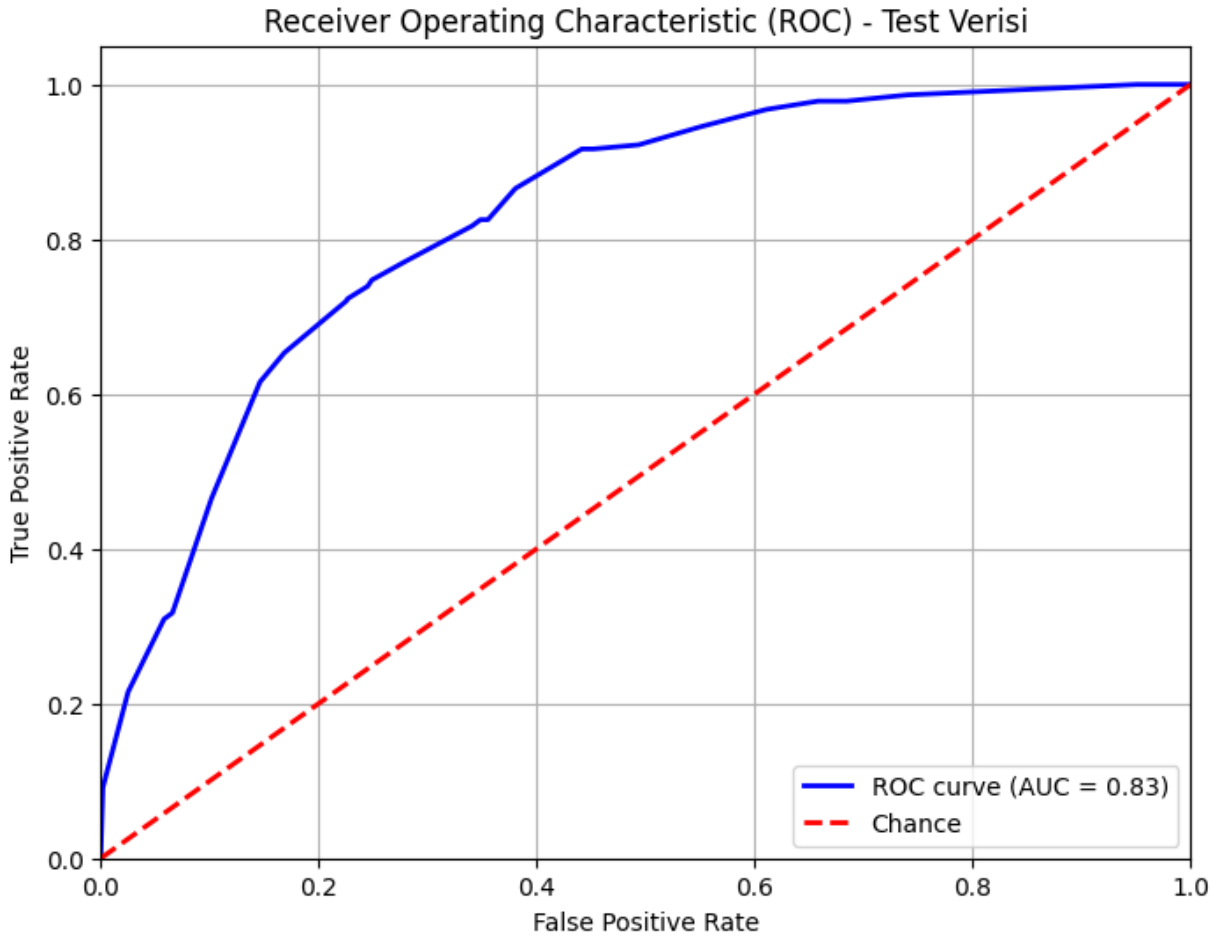


Şekil 4. 27 Karar Ağacı Modeli - Foldların ROC Eğrisi

Test verisi üzerindeki değerlendirme sonuçlarına göre doğruluk oranı %73,38, precision %49,83, recall %77,15, F1-Score %60,55 ve AUC değeri %83 olarak elde edilmiştir (Tablo 4.8). Bu metrikler, özellikle yüksek recall değeri sayesinde churn eden müşterilerin büyük çoğunluğunun başarıyla tahmin edildiğini göstermektedir.

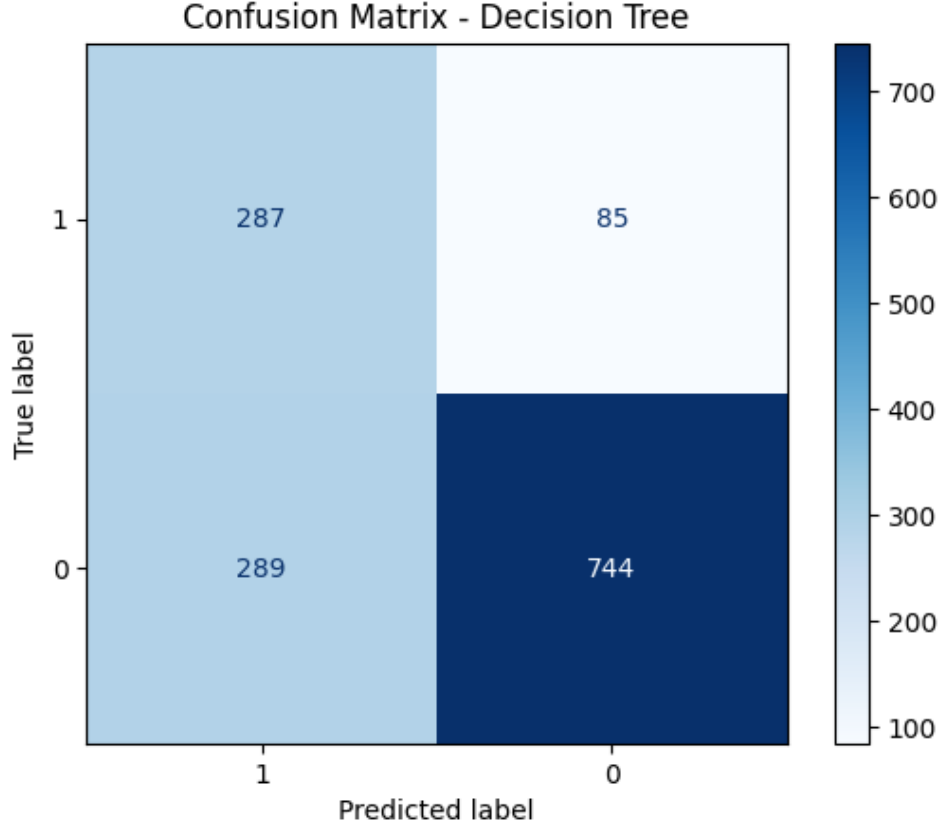
Tablo 4. 8 Karar Ağacı Modeli - Test Verisi Performansı

METRİK	DEĞER
ACCURACY	0.7338
PRECİSİON	0.4983
RECALL	0.7715
F1-SCORE	0.6055
AUC	0.83



Şekil 4. 28 Karar Ağacı Modeli - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.29) analizine göre model, 287 churn eden müşteriyi doğru tahmin etmiş, 85 churn eden müşteriyi gözden kaçırmış ve 289 churn etmeyen müşteriyi hatalı şekilde churn olarak sınıflandırmıştır. Bu, modelin churn sınıfına karşı duyarlı (yüksek recall), ancak precision açısından temkinli değerlendirilmesi gerektiğini ortaya koymaktadır.



Şekil 4. 29 Karar Ağacı - Confusion Matrix

Detaylı sınıflandırma raporuna (Tablo 4.9) göre churn sınıfı için precision %50, recall %77 ve F1-Score %61 düzeyindedir. Bu sonuçlar, karar ağacı modelinin churn eden müşterileri doğru tahmin etme (recall) yeteneğinin güçlü olduğunu, fakat bazı yanlış pozitif tahminlerin precision değerini düşürdüğünü göstermektedir. Non-churn sınıfı için precision değeri %90 ve F1-Score değeri %80 olup modelin bu sınıfta oldukça isabetli çalıştığını göstermektedir.

Tablo 4. 9 Karar Ağacı Modeli - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.90	0.72	0.80	1033
Churned	0.50	0.77	0.61	372
ACCURACY			0.73	1405
Macro AVG	0.70	0.75	0.70	1405
Weighted AVG	0.79	0.73	0.75	1405

Genel olarak Decision Tree modeli, veri setinde sınıf dengesizliği varken dahi churn tahmininde istikrarlı ve kabul edilebilir sonuçlar vermiştir. Ancak churn sınıfı için precision değerinin görece düşük olması, veri dengeleme stratejileriyle daha optimize edilebilir bir model altyapısına ihtiyaç duyulduğunu ortaya koymaktadır.

4.3.2.1 Decision Tree (Karar Ağacı) – SMOTE + Threshold Uygulamalı

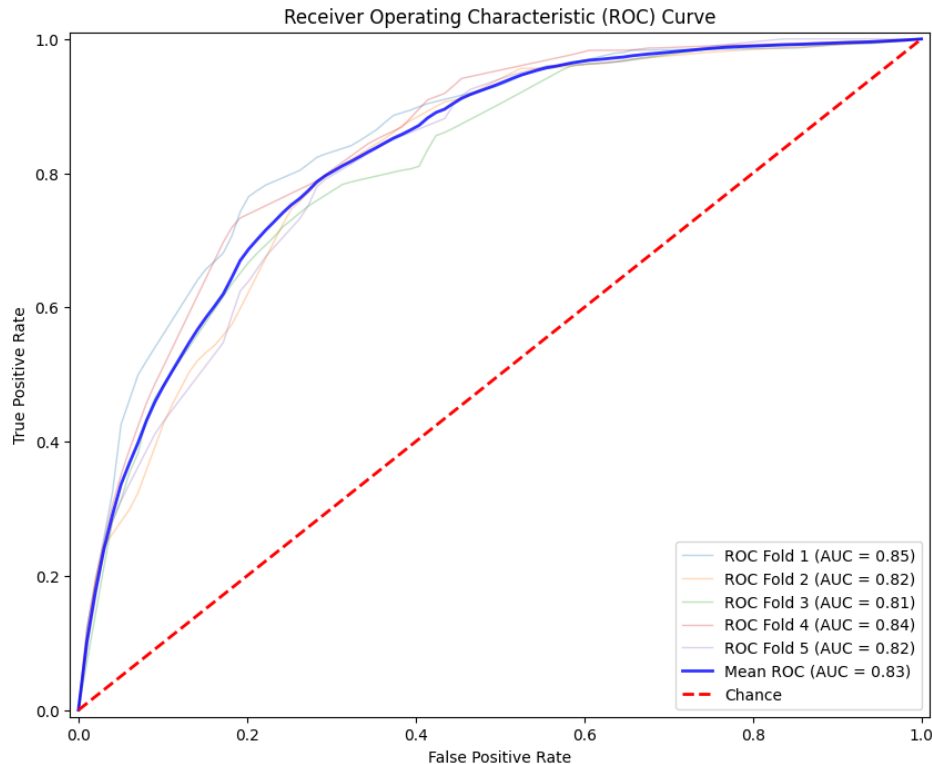
Bu bölümde, Decision Tree algoritmasına SMOTE uygulaması, optimal threshold ayarlaması ve 5 katlı çapraz doğrulama entegre edilerek oluşturulan yapılandırmanın performansı analiz edilmektedir. Bu yaklaşımla, sınıf dengesizliği azaltılmış, modelin karar eşiği optimize edilmiş ve sonuçların genellenebilirliği test edilmiştir.

Her bir fold'da SMOTE ile churn = 0 ve churn = 1 sınıflarının sayıları eşitlenmiş, ardından eğitim verisi ile model oluşturulmuş ve farklı threshold değerleri denenerek en uygun eşik belirlenmiştir. Optimal threshold değerleri 0.42 ile 0.63 arasında değişmekte olup genel ortalama eşik değeri 0.5096 olarak hesaplanmıştır. Bu varyasyon, fold'lara göre karar eşiklerinin değişebileceğini ve modelin bu eşik ayarlarına duyarlı olduğunu göstermektedir. Çapraz doğrulama sonuçları (Tablo 4.10) değerlendirildiğinde modelin ortalama doğruluk skoru %75,94, precision %53,52, recall %75,76, F1-score %62,58 ve AUC %83 olarak ölçülmüştür.

Tablo 4. 10 Karar Ağacı Modeli - SMOTE Uygulanan Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7918	0.7320	0.7489	0.7907	0.7337	0.7594
PRECISION	0.5810	0.4958	0.5182	0.5833	0.4979	0.5352
RECALL	0.7609	0.7946	0.7172	0.7306	0.7845	0.7576
F1-SCORE	0.6589	0.6106	0.6017	0.6487	0.6092	0.6258
AUC	0.85	0.82	0.81	0.84	0.82	0.83
OPTİMAL THRESHOLD	0.5366	0.4202	0.4894	0.6386	0.4636	0.509661031106949

ROC eğrisine (Şekil 4.27) bakıldığında fold'lar arasında AUC değerlerinin 0.81 ile 0,85 arasında değiştiği görülmekte olup ortalama AUC skoru %83'tür. Bu değerler modelin ayırıştırma yeteneğini genel olarak koruduğunu göstermektedir.

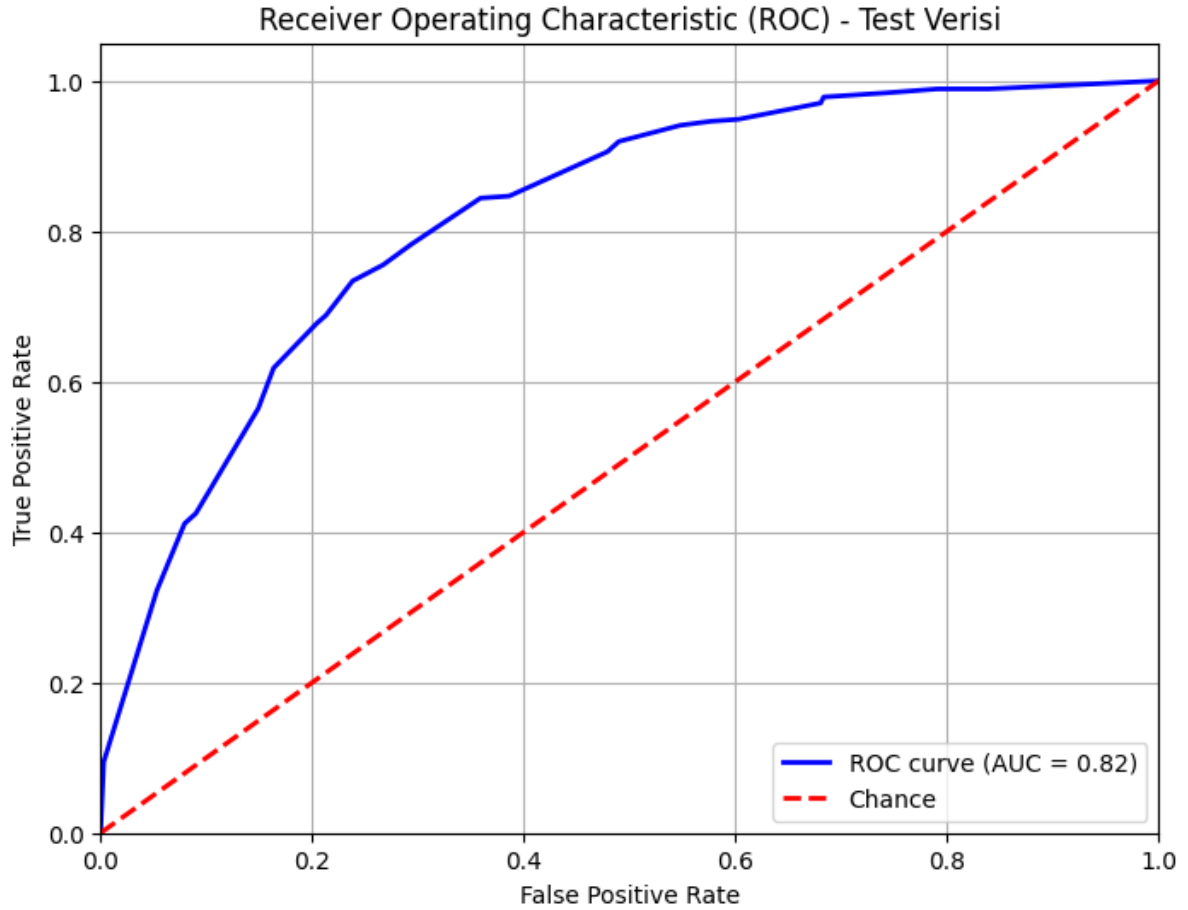


Şekil 4. 30 Karar Ağacı Modeli - SMOTE Uygulanan Foldların ROC Eğrisi

Test verisi üzerindeki deęerlendirmeye (Tablo 4.11) gre modelin doęruluk skoru %75,44, precision %52,60, recall %73,39, F1-score %61,28 ve AUC deęeri %82 olarak llmřtr. Bu performans, apraz doęrulama sonularıyla olduka uyumlu olup overfitting riskinin dřk olduęunu gstermektedir.

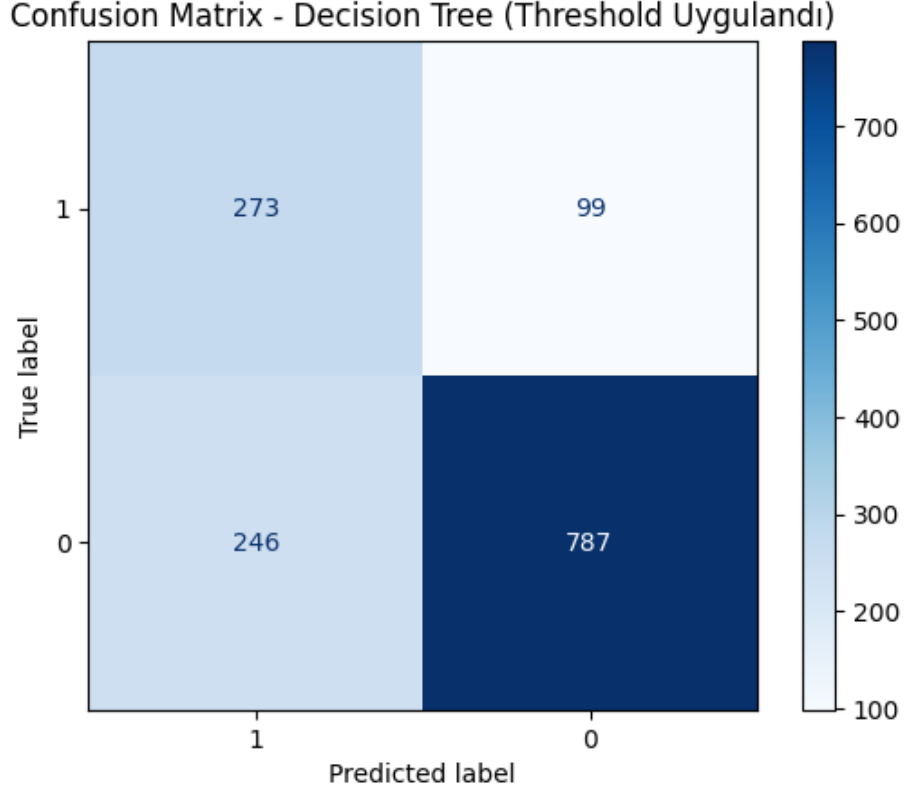
Tablo 4. 11 Karar Aęacı Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)

METRİK	DEęER
ACCURACY	0.7544
PRECİSİÖN	0.5260
RECALL	0.7339
F1-SCORE	0.6128
AUC	0.82
OPTİMAL THRESHOLD	0.509661031106949



řekil 4. 31 Karar Aęacı Modeli (SMOTE) - Test Verisi ROC Eęrisi

Confusion matrix'e (řekil 4.32) gre model, 273 churn eden mřteriyi doęru tahmin etmiř, 99 churn eden mřteriyi gzden kaırmıř, 246 churn etmeyen mřteriyi ise yanlış řekilde churn olarak sınıflandırmıřtır. Bu tablo churn sınıfına karřı hassasiyetin korunduęunu, ancak yanlış pozitiflerin de dikkate deęer olduęunu gstermektedir.



Şekil 4. 32 Karar Ağacı - Confusion Matrix (Threshold Uygulandı)

Detaylı sınıflandırma raporunda (Tablo 4.12) churn sınıfı için precision %53, recall %73, F1-score %61 olarak hesaplanmıştır. Bu metrikler, Decision Tree modelinin SMOTE sonrası churn sınıfını daha dengeli şekilde tahmin edebildiğini, özellikle precision skorundaki artışla birlikte yanlış pozitif oranının bir miktar azaldığını göstermektedir. Non-churn sınıfı için %89 precision ve %76 recall ile güçlü bir performans elde edilmiştir.

Tablo 4. 12 Karar Ağacı Modeli (SMOTE) - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.89	0.76	0.82	1033
Churned	0.53	0.73	0.61	372
ACCURACY			0.75	1405
Macro AVG	0.71	0.75	0.72	1405
Weighted AVG	0.79	0.75	0.77	1405

Sonuç olarak, SMOTE ve threshold ayarlaması uygulanmış Decision Tree modeli, önceki sürümüne kıyasla daha dengeli ve öngörülebilir bir performans sergilemiştir. K-Fold sırasında benzer AUC ve F1-Score değerlerinin tekrarlanması, modelin tutarlılığını kanıtlamakta ve overfitting ihtimalinin düşük olduğunu göstermektedir. Bu yapılandırma, churn tahmini görevinde güvenilir bir alternatif sunmaktadır.

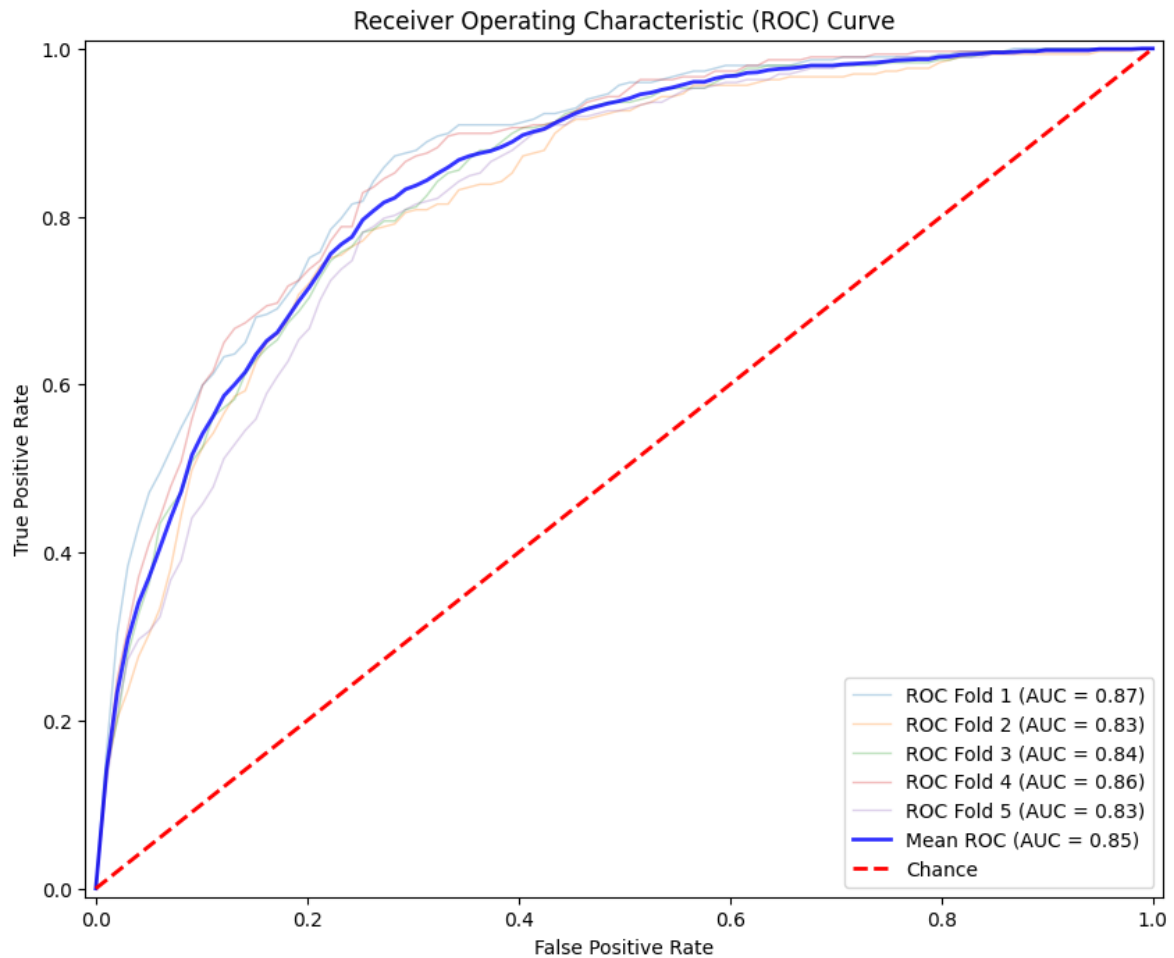
4.3.3 Random Forest (Rastgele Orman)

Bu bölümde, SMOTE ve eşik ayarlaması (threshold) uygulanmadan doğrudan eğitim verisi üzerinde eğitilen Random Forest (Rastgele Orman) modelinin performansı analiz edilmektedir. Modelin başarımlı düzeyi hem çapraz doğrulama hem de test seti üzerinden değerlendirilmiştir.

Random Forest modeli, 5 katlı çapraz doğrulama (Tablo 4.13) sonucunda tutarlı ve güçlü sonuçlar üretmiştir. Katmanlı doğrulama (Stratified K-Fold) her bir fold için dengesiz sınıf dağılımlarını koruyacak şekilde uygulanmış ve bu kapsamda elde edilen ortalama metrikler şu şekildedir: doğruluk (accuracy) %77,60, kesinlik (precision) %55,92, duyarlılık (recall) %72,32, F1 skoru %63.06 ve ROC AUC skoru %85'tir. ROC eğrisi (Şekil 4.33) katmanlar genelinde yüksek bir ayırt edicilik göstermekte olup, tüm fold'larda AUC skoru 0.83 ile 0,87 arasında değişmektedir.

Tablo 4. 13 Rastgele Orman Modeli - Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7802	0.7792	0.7703	0.7854	0.7649	0.7760
PRECİSİON	0.5625	0.5660	0.5496	0.5749	0.5431	0.5592
RECALL	0.7576	0.7071	0.7273	0.7239	0.7003	0.7232
F1-SCORE	0.6456	0.6287	0.6261	0.6408	0.6118	0.6306
AUC	0.87	0.83	0.84	0.86	0.83	0.85



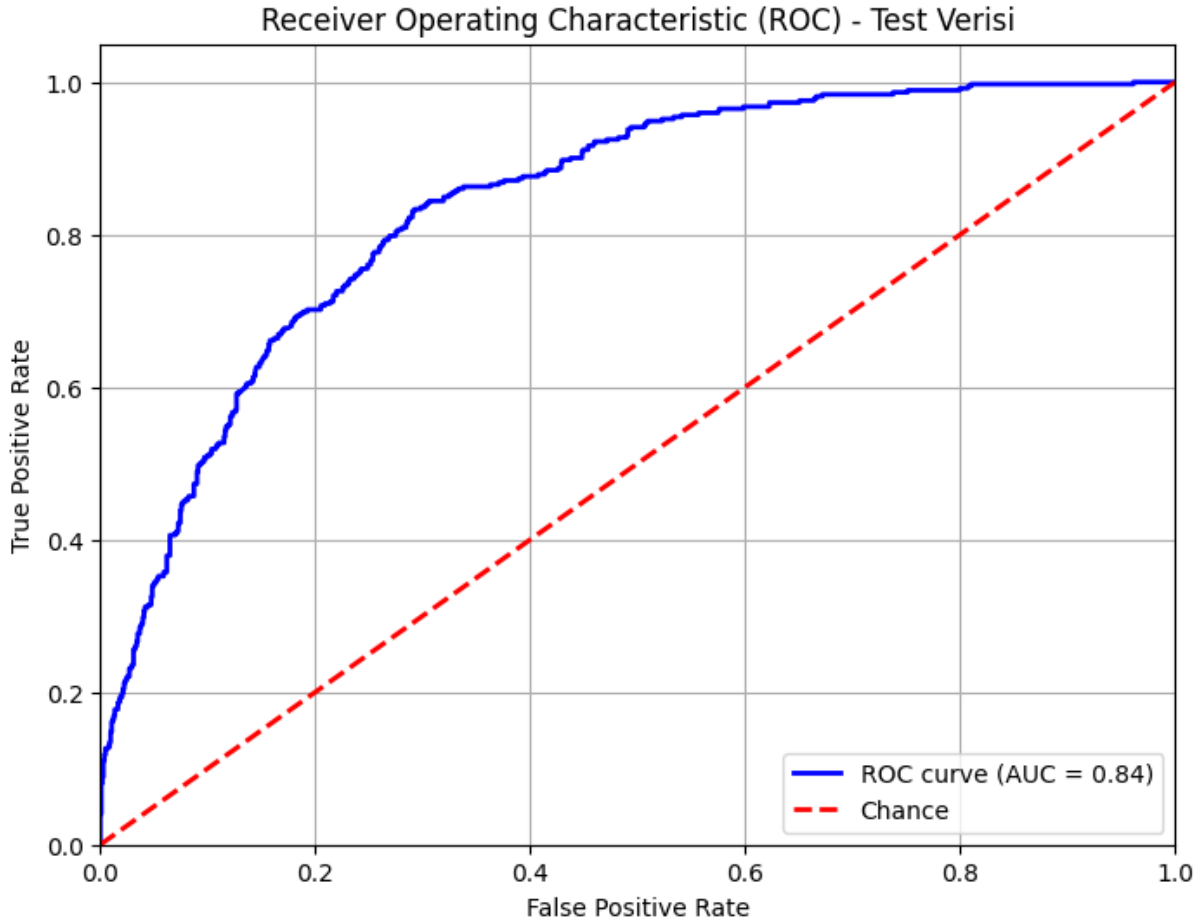
Şekil 4. 33 Rastgele Orman Modeli - Foldların ROC Eğrisi

Test verisi üzerinde modelin performansı (Tablo 4.14) da etkileyici düzeydedir. Test setindeki doğruluk %77.08, precision %55,30, recall %70,16, F1 skoru %61,85 ve ROC AUC skoru %84 olarak ölçülmüştür.

Tablo 4. 14 Rastgele Orman Modeli - Test Verisi Performansı

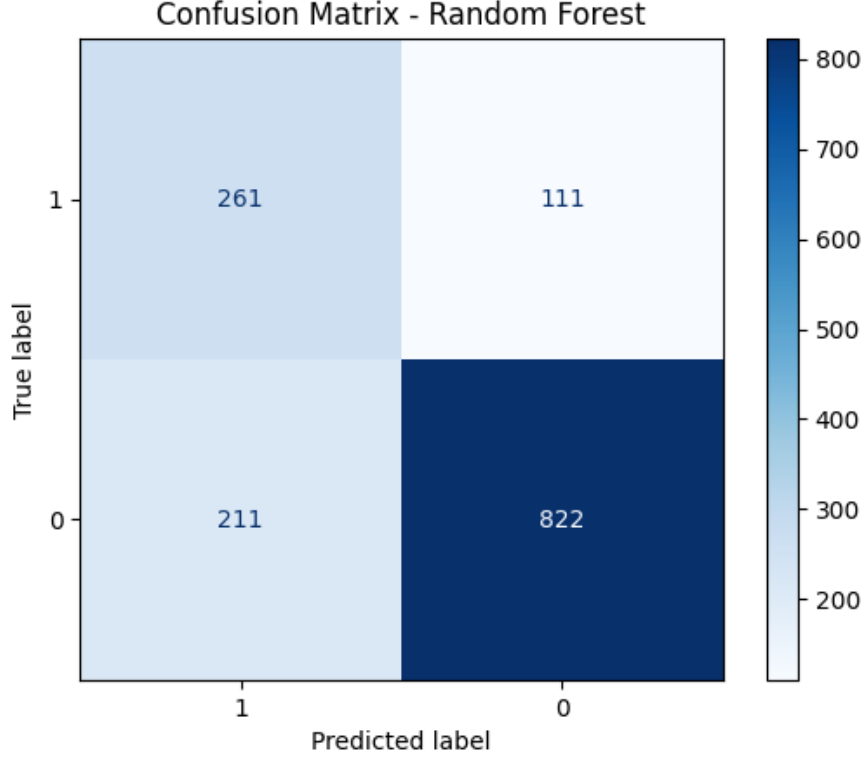
METRİK	DEĞER
ACCURACY	0.7708
PRECİSİON	0.5530
RECALL	0.7016
F1-SCORE	0.6185
AUC	0.84

ROC eğrisinde (Şekil 4.34) AUC'nin %84 olması, modelin churn eden ve etmeyen kullanıcıları ayırt etme kapasitesinin yüksek olduğunu göstermektedir.



Şekil 4. 34 Rastgele Orman Modeli - Test Verisi ROC Eğrisi

Confusion matrix'e (Şekil 4.35) göre model, churn eden 261 müşteriyi doğru sınıflandırmış, 111 churn eden müşteriyi kaçırmıştır. Aynı zamanda 211 churn etmeyen kullanıcı yanlış pozitif olarak işaretlenmiştir. Bu tablo churn sınıfı için daha yüksek duyarlılık (recall) düzeyine işaret ederken, yanlış pozitif oranının da makul düzeyde olduğunu göstermektedir.



Şekil 4. 35 Rastgele Orman - Confusion Matrix

Detaylı sınıflandırma raporunda (Tablo 4.15) churn sınıfı için precision %55, recall %70, F1-score %62; non-churn sınıfı için ise precision %88, recall %80, F1-score %84 değerleri elde edilmiştir. Bu dağılım modelin özellikle churn sınıfını dengelemede başarılı olduğunu ve performans açısından dengeli sonuçlar sunduğunu göstermektedir.

Tablo 4. 15 Rastgele Orman Modeli - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.88	0.80	0.84	1033
Churned	0.55	0.70	0.62	372
ACCURACY			0.77	1405
Macro AVG	0.72	0.75	0.73	1405
Weighted AVG	0.79	0.77	0.78	1405

K-Fold validasyon sonuçları ile test seti çıktılarının uyumlu olması, Random Forest modelinin overfitting'e karşı dirençli olduğunu ve genelleme kabiliyetinin yüksek olduğunu desteklemektedir. Ayrıca ROC eğrilerindeki yüksek AUC değerleri, modelin ikili sınıflandırma görevinde başarılı bir ayırım gücüne sahip olduğunu göstermektedir. Sonuç olarak, SMOTE ve eşik ayarlaması uygulanmamış temel Random Forest modeli, churn tahmini probleminde hem eğitim hem de test aşamasında güçlü bir performans sergilemiş, özellikle ROC AUC ve doğruluk metriklerinde en üst düzeyde başarı sağlamıştır.

4.3.3.1 Random Forest (Rastgele Orman) – SMOTE + Threshold Uygulamalı

Bu bölümde, Random Forest modeline SMOTE uygulaması, optimal threshold belirleme ve 5 katlı çapraz doğrulama yöntemi entegre edilerek elde edilen sonuçlar analiz edilmiştir. Bu yapılandırma, sınıf dengesizliğini azaltmak ve modelin doğruluk, duyarlılık ve genel genellenebilirliğini iyileştirmek amacıyla uygulanmıştır.

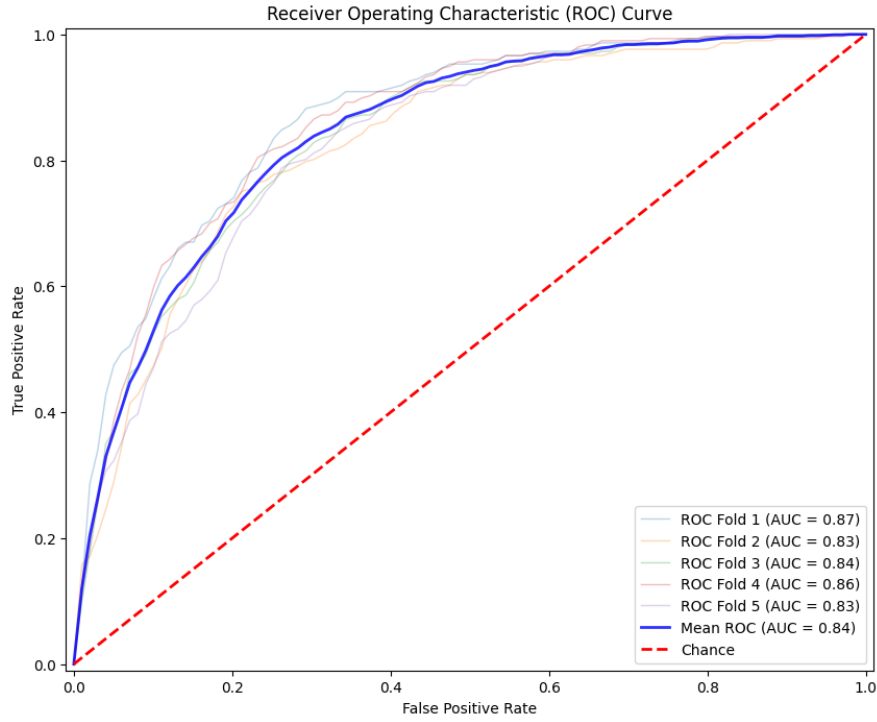
Her bir fold'da eğitim verisine SMOTE uygulanarak churn sınıfları dengelenmiştir (0 ve 1 sınıfları eşit sayıya ulaştırılmıştır). Daha sonra, her fold içerisinde model çıktıları üzerinden olasılık tahminleri kullanılarak en iyi F1-Skorunu sağlayan eşik değeri (threshold) hesaplanmıştır. Fold'lar arası optimal threshold değerleri 0.43 ile 0.47 arasında değişmiş olup, bu durum modelin sınıflandırma eşiklerine olan hassasiyetini göstermektedir.

Modelin 5 katlı çapraz doğrulama (K-Fold CV) (Tablo 4.16) sonuçları incelendiğinde, ortalama doğruluk (accuracy) skoru %76,89 olarak hesaplanmıştır. Kesinlik (precision) ortalaması %54,38 olup, bu değer churn sınıfı için yanlış pozitif tahminlerin belirli bir düzeyde olduğunu göstermektedir. Duyarlılık (recall) ortalaması %78,72 gibi yüksek bir değere ulaşmış, bu da churn eden müşterilerin büyük bir bölümünün model tarafından doğru şekilde tahmin edildiğini göstermektedir. F1-Skoru ise %64,29 olarak elde edilmiş, bu da modelin hem kesinlik hem de duyarlılığı dengeli şekilde sağladığını ortaya koymuştur. AUC (Area Under Curve) değeri ortalama 0.84 olup, modelin sınıflar arasında ayırım yapma gücünün oldukça iyi olduğunu ifade etmektedir. Optimal eşik değeri ise yaklaşık 0,45 olarak belirlenmiştir.

Tablo 4. 16 Rastgele Orman Modeli - SMOTE Uygulanan Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7714	0.7783	0.7605	0.7801	0.7542	0.7689
PRECİSİON	0.5437	0.5603	0.5332	0.5584	0.5236	0.5438
RECALL	0.8384	0.7508	0.7576	0.8047	0.7845	0.7872
F1-SCORE	0.6596	0.6417	0.6259	0.6593	0.6280	0.6429
AUC	0.87	0.83	0.84	0.86	0.83	0.84
OPTİMAL THRESHOLD	0.4339	0.4769	0.4625	0.4458	0.4327	0.45034285611140196

ROC eğrisi (Şekil 4.36), her bir fold için ROC performansını ve ortalama ROC eğrisini göstermektedir. Grafik üzerinde ROC Fold 1 için AUC = 0.87, Fold 2 için AUC = 0.83, Fold 3 için AUC = 0.84, Fold 4 için AUC = 0.86, Fold 5 için AUC = 0.83 şeklinde sıralanmıştır. Ortalama ROC eğrisi ise AUC = 0.84 ile oldukça tatmin edici bir performans sergilemiştir.



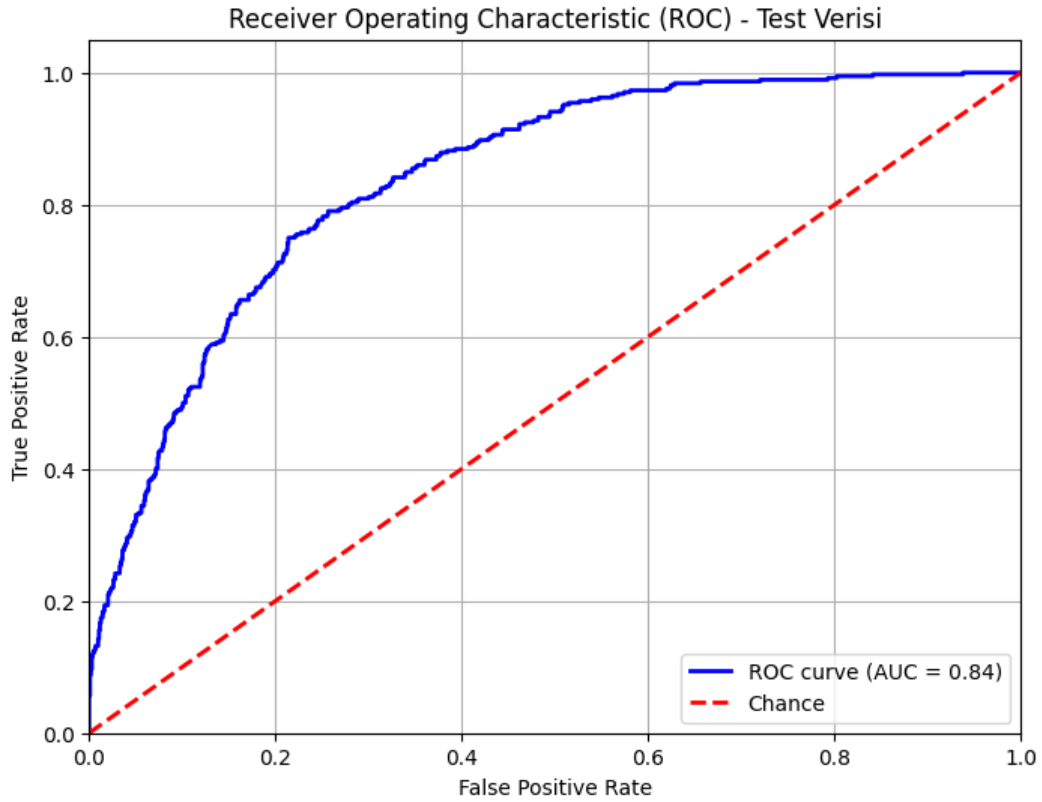
Şekil 4. 36 Rastgele Orman Modeli - SMOTE Uygulanan Foldların ROC Eğrisi

Modelin test verisi üzerinde gösterdiği performansa (Tablo 4.17) bakıldığında, doğruluk %76,65, kesinlik %54,23, duyarlılık %75,81 ve F1-Skoru %63,23 olarak hesaplanmıştır. AUC skoru 0.84 ile yine güçlü bir ayırım gücü sergilemiştir. Test verisi için belirlenen optimal threshold değeri 0.4503 olarak seçilmiştir. Bu değer, modelin karar sınırı belirleyerek daha dengeli bir tahmin çıktısı alınmasına katkı sağlamıştır. Bu sonuçlar, eğitim sürecindeki başarıyı büyük ölçüde yansıtmaktadır ve overfitting riskinin düşük olduğunu düşündürmektedir. Bu performans, optimal threshold değeri 0.4503 olarak belirlenip uygulanarak elde edilmiştir.

Tablo 4. 17 Rastgele Orman Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)

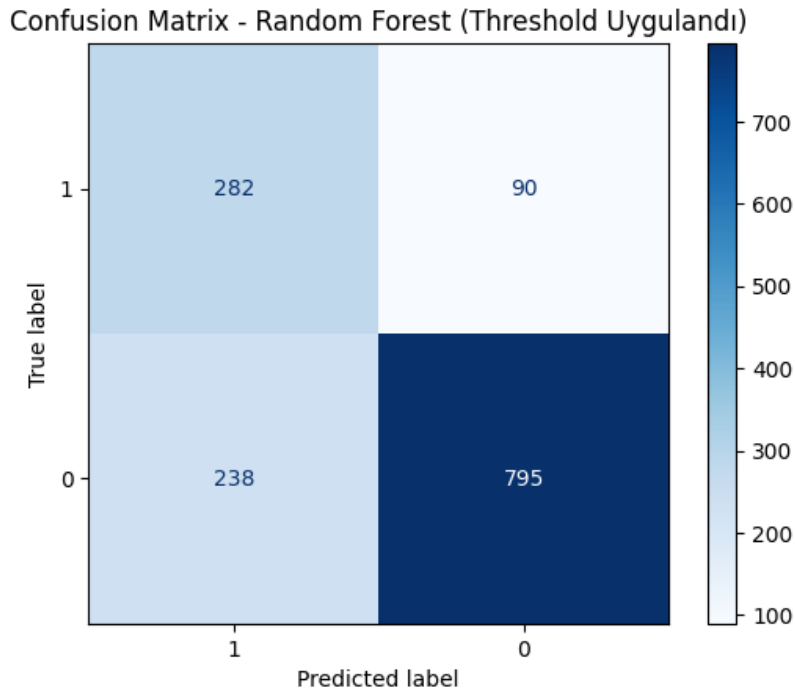
METRİK	DEĞER
ACCURACY	0.7665
PRECISION	0.5423
RECALL	0.7581
F1-SCORE	0.6323
AUC	0.84
OPTİMAL THRESHOLD	0.45034285611140196

Şekil 4.37 de ise modelin test verisine ait ROC eğrisi sunulmaktadır.



Şekil 4. 37 Rastgele Orman Modeli (SMOTE) - Test Verisi ROC Eğrisi

Test verisine ait confusion matrix (Şekil 4.38) incelendiğinde, modelin 282 churn eden müşteriyi doğru tahmin ettiği, 90 churn eden müşteriyi gözden kaçırdığı ve 238 churn etmeyen müşteriyi yanlış churn olarak sınıflandırdığı gözlemlenmiştir. Bu tablo, modelin churn sınıfındaki duyarlılığı yüksek tutmakla birlikte, bazı yanlış pozitif sonuçlar verdiğini ortaya koymaktadır.



Şekil 4. 38 Rastgele Orman - Confusion Matrix (Threshold Uygulandı)

Detaylı sınıflandırma raporuna (Tablo 4.18) göre churn eden sınıf için precision %54, recall %76 ve F1-score %63 olarak ölçülmüştür. Bu, modelin churn sınıfını tespit etmede dengeli bir performans sergilediğini göstermektedir. Ayrıca non-churn sınıfında da precision %90 ve recall %77 olarak kaydedilmiş olup genel sınıflandırma kalitesinin yüksek olduğu anlaşılmaktadır.

Tablo 4. 18 Rastgele Orman Modeli (SMOTE) - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.90	0.77	0.83	1033
Churned	0.54	0.76	0.63	372
ACCURACY			0.77	1405
Macro AVG	0.72	0.76	0.73	1405
Weighted AVG	0.80	0.77	0.78	1405

Sonuç olarak, SMOTE + Threshold + K-Fold yaklaşımı Random Forest modelinin churn sınıflandırmadaki performansını genel olarak artırmıştır. K-fold süresince gözlemlenen metrik tutarlılığı ve test verisi üzerindeki güçlü performans, modelin aşırı öğrenme (overfitting) eğiliminden uzak, kararlı bir yapıda olduğunu göstermektedir. Bu konfigürasyon, churn tahmininde güvenilir ve dengeli bir çözüm olarak değerlendirilebilir.

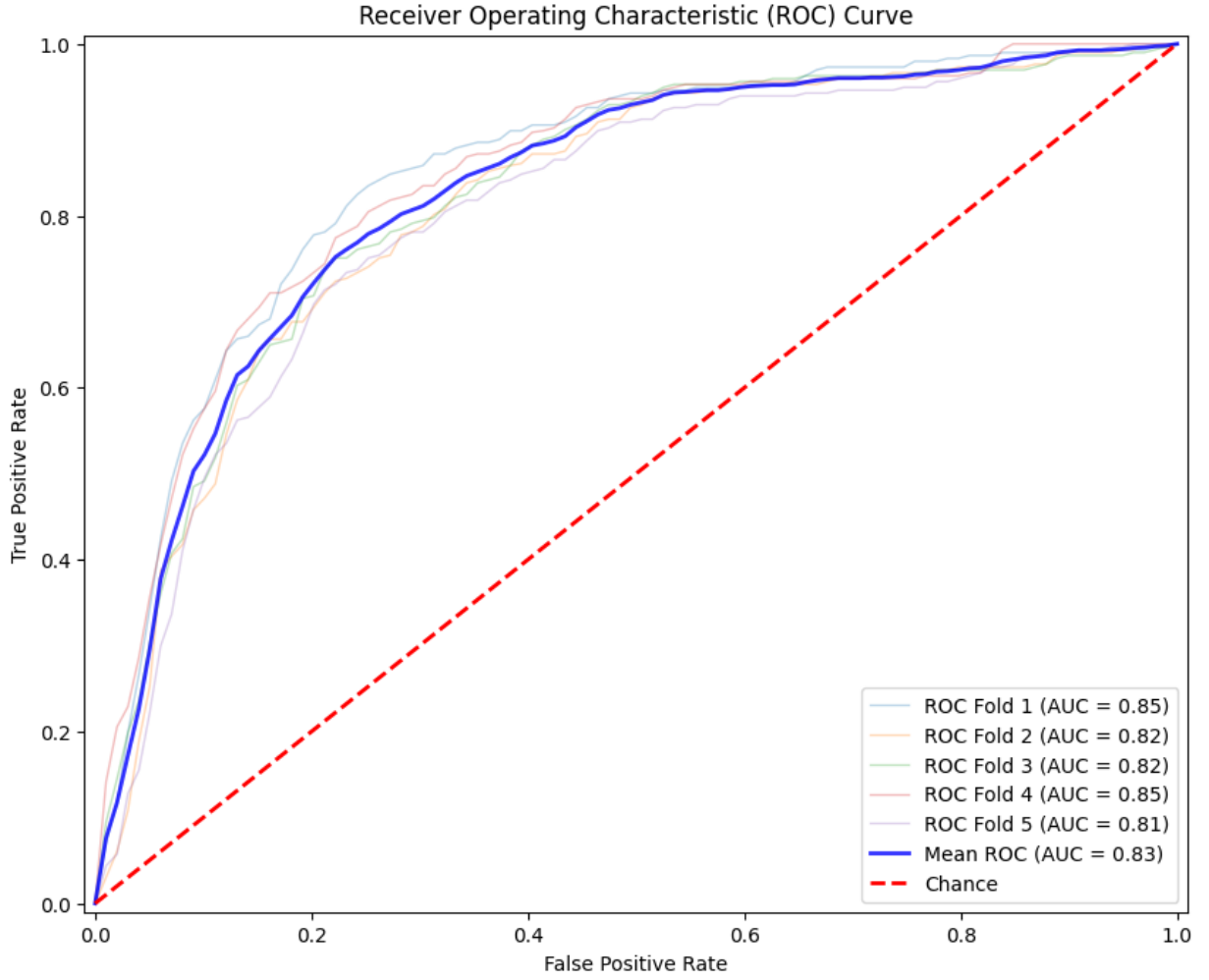
4.3.4 SVM (Support Vector Machine / Destek Vektör Makinesi)

Bu bölümde, Support Vector Machine (SVM) algoritması kullanılarak churn tahmini yapılmıştır. Modelin temel versiyonu değerlendirilmiş olup SMOTE veya threshold optimizasyonu uygulanmamıştır. Sınıflandırma başarımı, 5 katlı çapraz doğrulama ve test verisi üzerindeki performans metrikleri aracılığıyla analiz edilmiştir.

Tablo 4.19’da gösterildiği üzere çapraz doğrulama sonuçlarına göre modelin ortalama doğruluk skoru %75.09, precision %51,92, recall %78,32 ve F1-Score %62,43’tür. ROC eğrisi altında kalan alanlar (AUC) fold'lara göre %81 ila %85 arasında değişmiş, ortalama AUC değeri %83 olarak hesaplanmıştır. Bu değerler, modelin sınıflar arasında makul bir ayırım yapabildiğini ve özellikle churn sınıfını tespit etmede yüksek duyarlılık sağladığını göstermektedir. ROC eğrileri Şekil 4.39’da sunulmuştur.

Tablo 4. 19 Destek Vektör Makinesi Modeli - Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7687	0.7373	0.7400	0.7649	0.7435	0.7509
PRECİSİON	0.5401	0.5023	0.5054	0.5379	0.5103	0.5192
RECALL	0.8384	0.7508	0.7845	0.7879	0.7542	0.7832
F1-SCORE	0.6570	0.6019	0.6148	0.6393	0.6087	0.6243
AUC	0.85	0.82	0.82	0.85	0.81	0.83

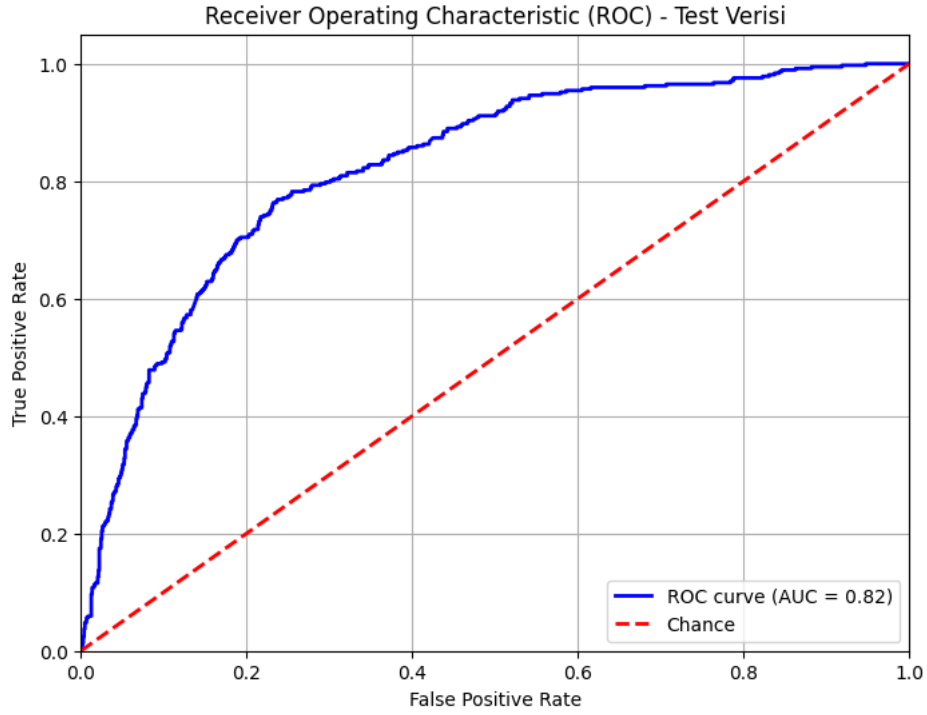


Şekil 4. 39 Destek Vektör Makinesi Modeli - Foldların ROC Eğrisi

Test verisi üzerinde yapılan performans değerlendirmesi ise çapraz doğrulama süreciyle büyük ölçüde tutarlıdır. Doğruluk oranı %76.01, precision %53.25, recall %77.15, F1-Skoru %63.01 ve AUC %82 olarak hesaplanmıştır (Tablo 4.20). Bu veriler, modelin gerçek dünya benzeri test verisi üzerinde de başarılı bir şekilde genelleme yapabildiğini göstermektedir. Test verisinin ROC eğrisi Şekil 4.40'ta sunulmuş olup, sınıflandırma kararlarının güvenilirliğini görsel olarak desteklemektedir.

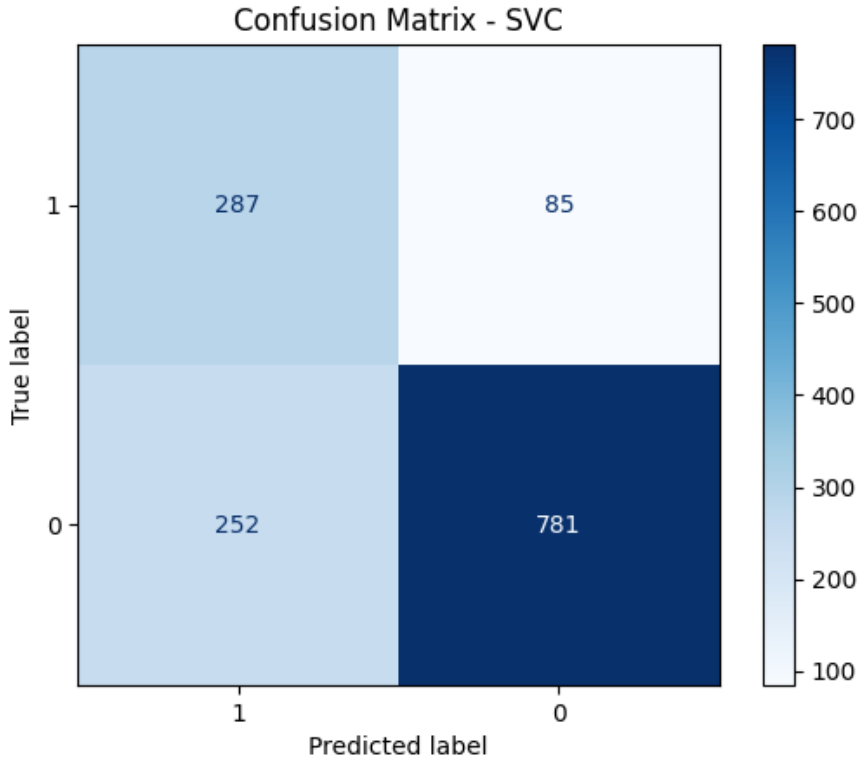
Tablo 4. 20 Destek Vektör Makinesi Modeli - Test Verisi Performansı

METRİK	DEĞER
ACCURACY	0.7601
PRECISION	0.5325
RECALL	0.7715
F1-SCORE	0.6301
AUC	0.82



Şekil 4. 40 Destek Vektör Makinesi Modeli - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.41) incelendiğinde modelin churn eden 287 müşteriyi doğru sınıflandırdığı, 86 churn müşterisini ise yanlış negatif olarak değerlendirdiği görülmektedir. Bunun yanı sıra, churn etmeyen 252 müşteri yanlış pozitif olarak işaretlenmiştir. Bu durum, modelin churn sınıfındaki örnekleri yüksek recall ile yakalayabildiğini ancak bu başarıyı elde ederken precision tarafında ödün verdiğini göstermektedir.



Şekil 4. 41 Destek Vektör Makinesi - Confusion Matrix

Detaylı sınıflandırma raporu (Tablo 4.21) churn sınıfı için precision değerini %53, recall değerini %77 ve F1-Score değerini %63 olarak raporlamaktadır. Non-churn sınıfı için ise precision %90, recall %76 ve F1-Score %82 olarak elde edilmiştir. Bu metrik dağılımı, modelin negatif sınıfta (churn olmayanlar) daha yüksek doğrulukla çalıştığını; ancak pozitif sınıf olan churn üzerinde dengeli bir başarı sunduğunu göstermektedir. Özellikle F1-Score'un iki sınıf arasında çok fazla farklılık göstermemesi, modelin genel olarak dengeli bir sınıflandırma sağladığını ortaya koymaktadır.

Tablo 4. 21 Destek Vektör Makinesi Modeli - Detaylı Sınıflandırma Raporu

	PRECISION	RECALL	F1-SCORE	SUPPORT
Not Churned	0.90	0.76	0.82	1033
Churned	0.53	0.77	0.63	372
ACCURACY			0.76	1405
Macro AVG	0.72	0.76	0.73	1405
Weighted AVG	0.80	0.76	0.77	1405

Sonuç olarak, SVM algoritması SMOTE veya eşik ayarlaması gibi ek yöntemler kullanılmadan dahi churn sınıfını yakalama konusunda yüksek duyarlılık sağlamış, F1-Score ve AUC gibi bileşik metriklerde başarılı sonuçlar üretmiştir. Precision değerinin görece düşük kalması churn olmayan kullanıcıların zaman zaman yanlış tahmin edilmesine yol açsa da, bu durum kabul edilebilir düzeydedir. Bu yapıdaki SVM modeli, dengesiz veri üzerinde doğrudan uygulanabilir, istikrarlı ve genellenebilir sonuçlar sunan bir alternatif olarak değerlendirilebilir.

4.3.4.1 SVM (Support Vector Machine / Destek Vektör Makinesi) – SMOTE + Threshold Uygulamalı

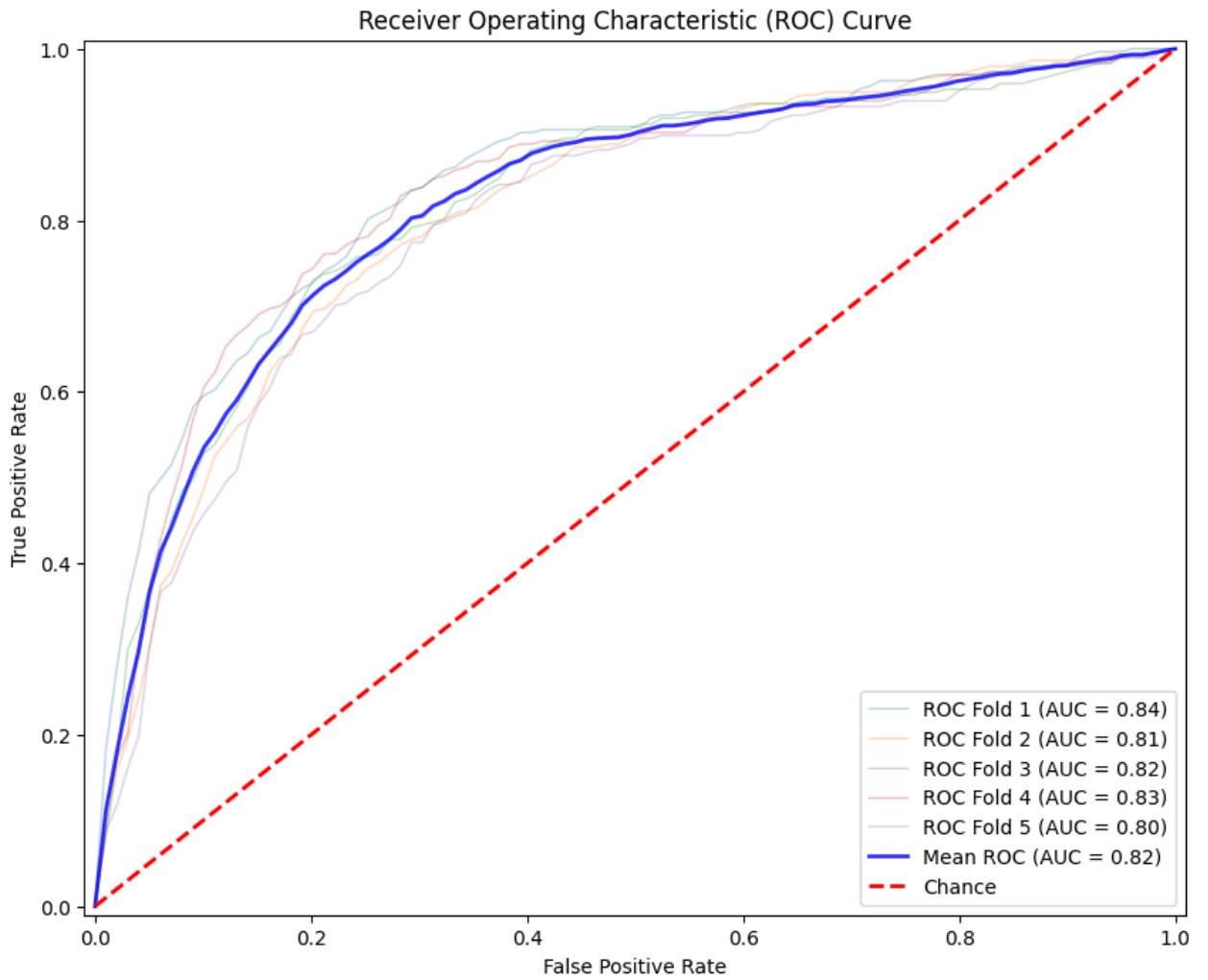
Bu bölümde, Destek Vektör Makineleri (SVM) algoritmasına SMOTE ve optimal threshold uygulamaları entegre edilerek 5 katlı çapraz doğrulama ile elde edilen sonuçlar değerlendirilmiştir. Bu yapılandırma ile modelin veri dengesizliğine karşı dayanıklılığı artırılmak istenmiş; eşik değerinin optimizasyonu sayesinde sınıflandırma başarımında iyileştirmeler hedeflenmiştir.

Modelin eğitim sürecinde, her bir fold'da SMOTE uygulanarak sınıf dengesizliği giderilmiştir. Böylelikle churn=0 ve churn=1 sınıflarındaki örnek sayıları eşitlenmiş; modelin azınlık sınıfa karşı öğrenme yetisi artırılmıştır. Her fold sonrasında tahmin edilen olasılık skorları üzerinden farklı eşik değerleri denenmiş ve en iyi F1-Score'u veren optimal threshold belirlenmiştir. Fold'lar arası optimal eşik değerlerinin 0.23 ile 0.54 arasında değiştiği gözlemlenmiştir. Bu farklılık, modelin veri alt kümelerine olan duyarlılığını ve sınıflandırma kararlarının eşik bağımlılığını göstermektedir.

Çapraz doğrulama sonuçlarına (Tablo 4.22) bakıldığında, modelin ortalama doğruluk (accuracy) oranı %78,19, precision değeri %57,44, recall değeri %71.04, F1-score değeri %63,28 ve ortalama AUC skoru %82 olarak hesaplanmıştır. ROC eğrileri fold'lara göre 0.80 ile 0,84 arasında değişmekte olup genel ortalama 0,82 olarak belirlenmiştir (Şekil 4.42). Bu sonuçlar, modelin hem doğru pozitifleri ayırt etme başarısını hem de genel dengeleme gücünü istikrarlı şekilde koruduğunu göstermektedir.

Tablo 4. 22 Destek Vektör Makinesi Modeli - SMOTE Uygulanan Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7633	0.7711	0.7854	0.8166	0.7729	0.7819
PRECISION	0.5348	0.5538	0.5745	0.6492	0.5597	0.5744
RECALL	0.8013	0.6936	0.7273	0.6667	0.6633	0.7104
F1-SCORE	0.6415	0.6158	0.6419	0.6578	0.6071	0.6328
AUC	0.84	0.81	0.82	0.83	0.80	0.82
OPTİMAL THRESHOLD	0.2376	0.3142	0.3328	0.5414	0.3597	0.3571493060190415



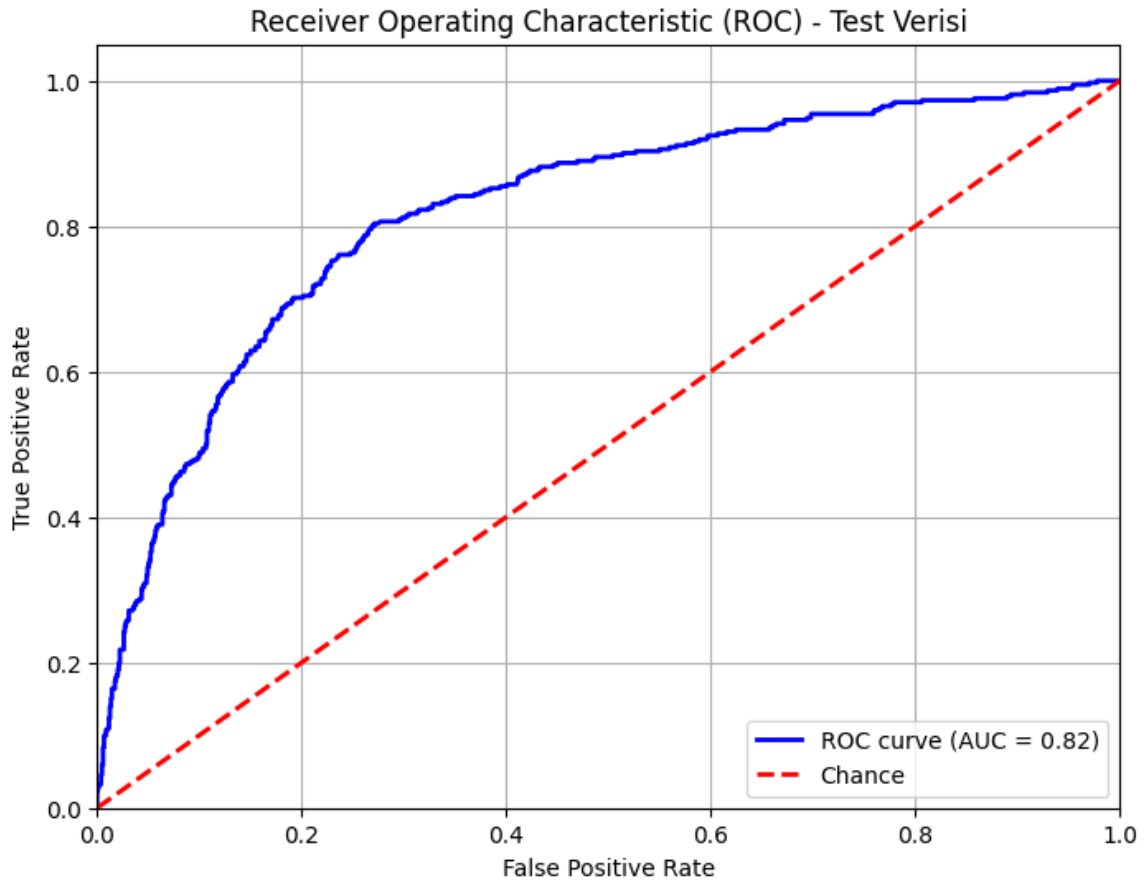
Şekil 4. 42 Destek Vektör Makinesi Modeli - SMOTE Uygulanan Foldların ROC Eğrisi

Test verisi üzerindeki performans değerlendirmesinde (Tablo 4.23) modelin doğruluk oranı %78.51, precision değeri %58.14, recall değeri %67.20, F1-score değeri %62.34 ve AUC skoru %82 olarak ölçülmüştür. Modelin test setinde de çapraz doğrulama sonuçlarına benzer değerler üretmiş olması, yüksek genellenebilirliğe sahip olduğunu ortaya koymaktadır. Optimal threshold değeri bu aşamada 0.3571 olarak belirlenmiş ve bu eşik değerine göre sınıflandırma yapılmıştır.

Tablo 4. 23 Destek Vektör Makinesi Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)

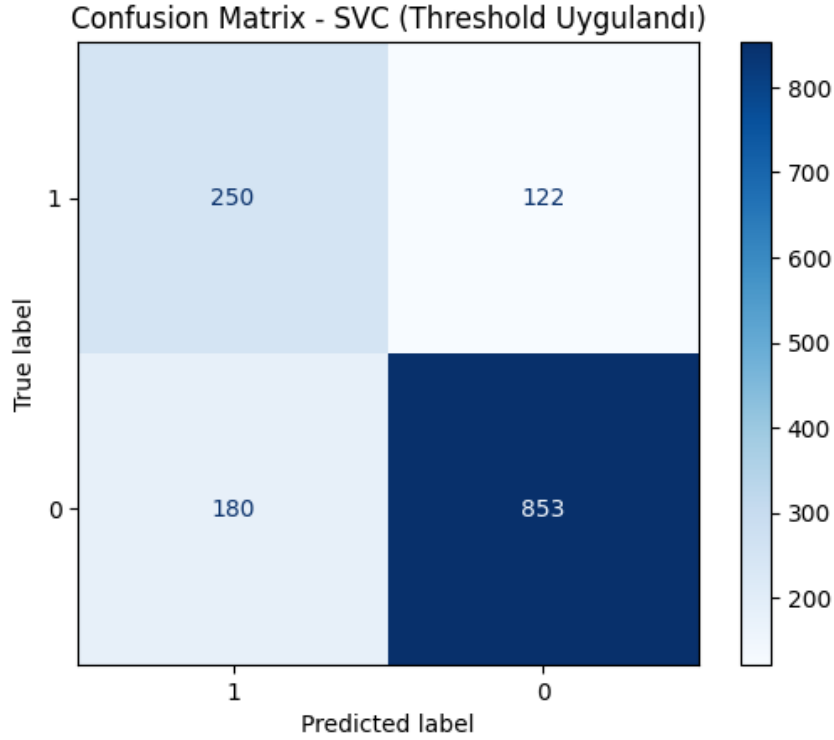
METRİK	DEĞER
ACCURACY	0.7851
PRECİSİON	0.5814
RECALL	0.6720
F1-SCORE	0.6234
AUC	0.82
OPTİMAL THRESHOLD	0.3571493060190415

Test verisinin ROC eğrisi Şekil 4.43'te sunulmuş olup, sınıflandırma kararlarının güvenilirliğini görsel olarak desteklemektedir.



Şekil 4. 43 Destek Vektör Makinesi Modeli (SMOTE) - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.44) incelendiğinde, modelin 250 churn eden müşteriyi doğru tahmin ettiği, 122 churn eden müşteriyi ise yanlışlıkla churn etmeyen olarak etiketlediği görülmektedir (false negative). Ayrıca, 180 churn etmeyen müşteri churn olarak tahmin edilmiştir (false positive). Bu tablo, modelin churn sınıfındaki bireyleri yüksek duyarlılıkla tahmin ederken, az miktarda da olsa yanlış pozitif üretmeye eğilimli olduğunu göstermektedir.



Şekil 4. 44 Destek Vektör Makinesi - Confusion Matrix (Threshold Uygulandı)

Detaylı sınıflandırma raporuna (Tablo 4.24) göre churn sınıfı için precision %58, recall %67 ve F1-score %62 olarak hesaplanmıştır. Bu değerler, SVM modelinin SMOTE ve threshold ile birlikte dengeli bir performans ortaya koyduğunu; churn sınıfı üzerinde yeterli bir öğrenme sağladığını göstermektedir. Non-churn sınıfında precision %87, recall %83 ve F1-score %85 düzeyinde çıkmıştır. Bu, modelin sınıflar arası başarı farkını tolere edilebilir seviyede tuttuğunu göstermektedir.

Tablo 4. 24 Destek Vektör Makinesi Modeli (SMOTE) - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.87	0.83	0.85	1033
Churned	0.58	0.67	0.62	372
ACCURACY			0.79	1405
Macro AVG	0.73	0.75	0.74	1405
Weighted AVG	0.80	0.79	0.79	1405

Sonuç olarak, SMOTE + Threshold + K-Fold yapılandırmasıyla geliştirilen SVM modeli, churn tahmini problemi için yüksek doğruluk ve sınıflar arası dengeli başarı sağlamaktadır. Overfitting riski düşük kalmış; zira eğitim fold'ları arasındaki varyans sınırlı düzeyde seyretmiş ve test setindeki metriklerle paralel sonuçlar elde edilmiştir. Bu yönüyle, model genel performans açısından güvenilir ve tutarlı bir yapı sunmaktadır.

4.3.5 XGBoost (eXtreme Gradient Boosting)

Bu çalışmada kullanılan son model olan XGBoost, gradyan artırılmalı karar ağaçları yaklaşımına dayalı güçlü ve esnek bir sınıflandırma algoritmasıdır. Model, varsayılan sınıf dağılımıyla (SMOTE uygulanmadan) ve optimal threshold belirlemeden eğitilmiş ve

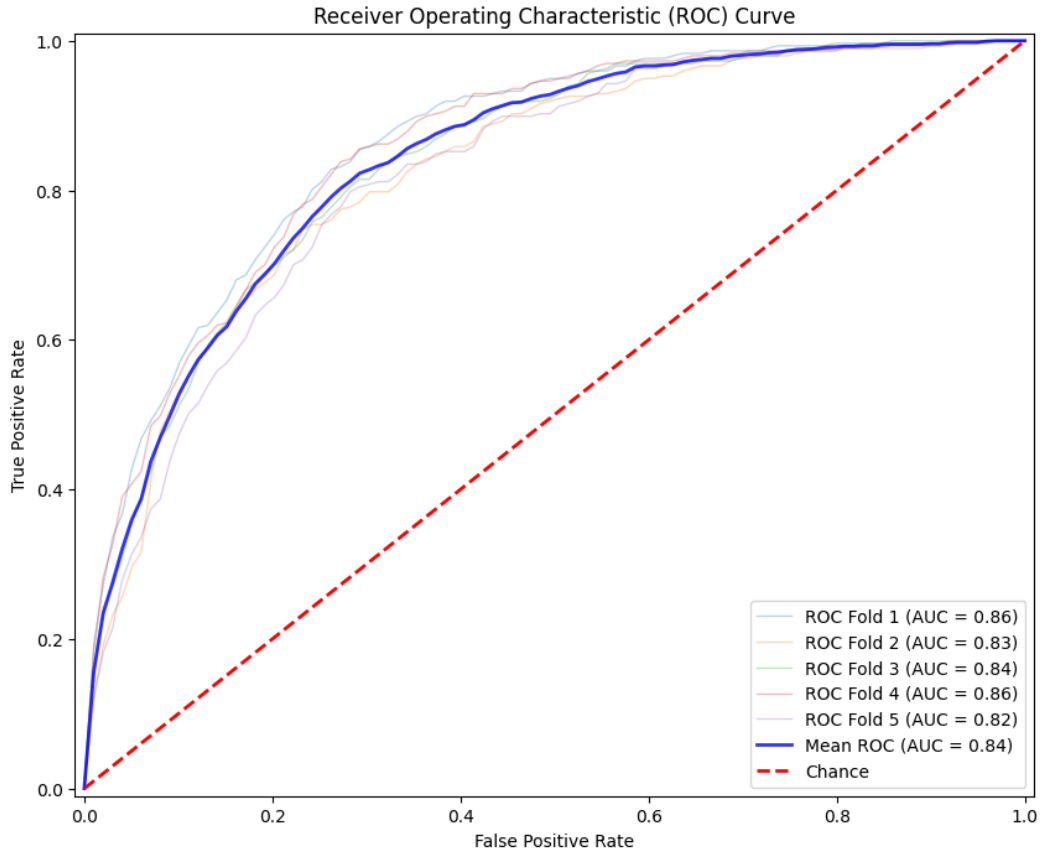
değerlendirilmiştir. Bu aşamada modelin 5 katlı çapraz doğrulama (K-Fold CV) süreci ve nihai test verisi performansı ayrı ayrı incelenmiştir.

Çapraz doğrulama sürecinde her bir fold için model ayrı ayrı eğitilmiş ve doğrulama setleri üzerinden başarı metrikleri hesaplanmıştır (Tablo 4.25). Fold sonuçlarına göre modelin doğruluk oranı (accuracy) %74,89 ile %77,56 arasında değişmiş, ortalama doğruluk oranı ise %76,32 olarak elde edilmiştir. Precision değerleri %51,84 ile %55,45 arasında gözlemlenmiş ve ortalama precision değeri %53,74 olmuştur. Bununla birlikte modelin churn sınıfını tanıma kabiliyetini yansıtan recall değeri tüm fold'larda oldukça istikrarlı çıkmış ve ortalama olarak %71.04 olarak ölçülmüştür. F1-Score, yani precision ve recall'un harmonik ortalaması, %59,94 ortalama ile churn tahminleme başarısının dengeli bir düzeyde gerçekleştiğini göstermektedir.

Tablo 4. 25 XGBoost Modeli - Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7731	0.7640	0.7542	0.7756	0.7489	0.7632
PRECİSİON	0.5500	0.5390	0.5249	0.5545	0.5184	0.5374
RECALL	0.7778	0.7441	0.7441	0.7710	0.7104	0.7104
F1-SCORE	0.6444	0.6252	0.6156	0.6451	0.5994	0.5994
AUC	0.86	0.83	0.84	0.86	0.82	0.84

ROC eğrisi altında kalan alan (AUC) ise tüm fold'larda yüksek seviyede olup, ortalama AUC skoru %84 olarak hesaplanmıştır. ROC eğrilerine bakıldığında (Şekil 4.45), fold'lar arası AUC değerlerinin 0.82 ile 0,86 arasında değiştiği görülmekte, bu da modelin sınıfları ayırt etme yetisinin oldukça tutarlı olduğunu göstermektedir.

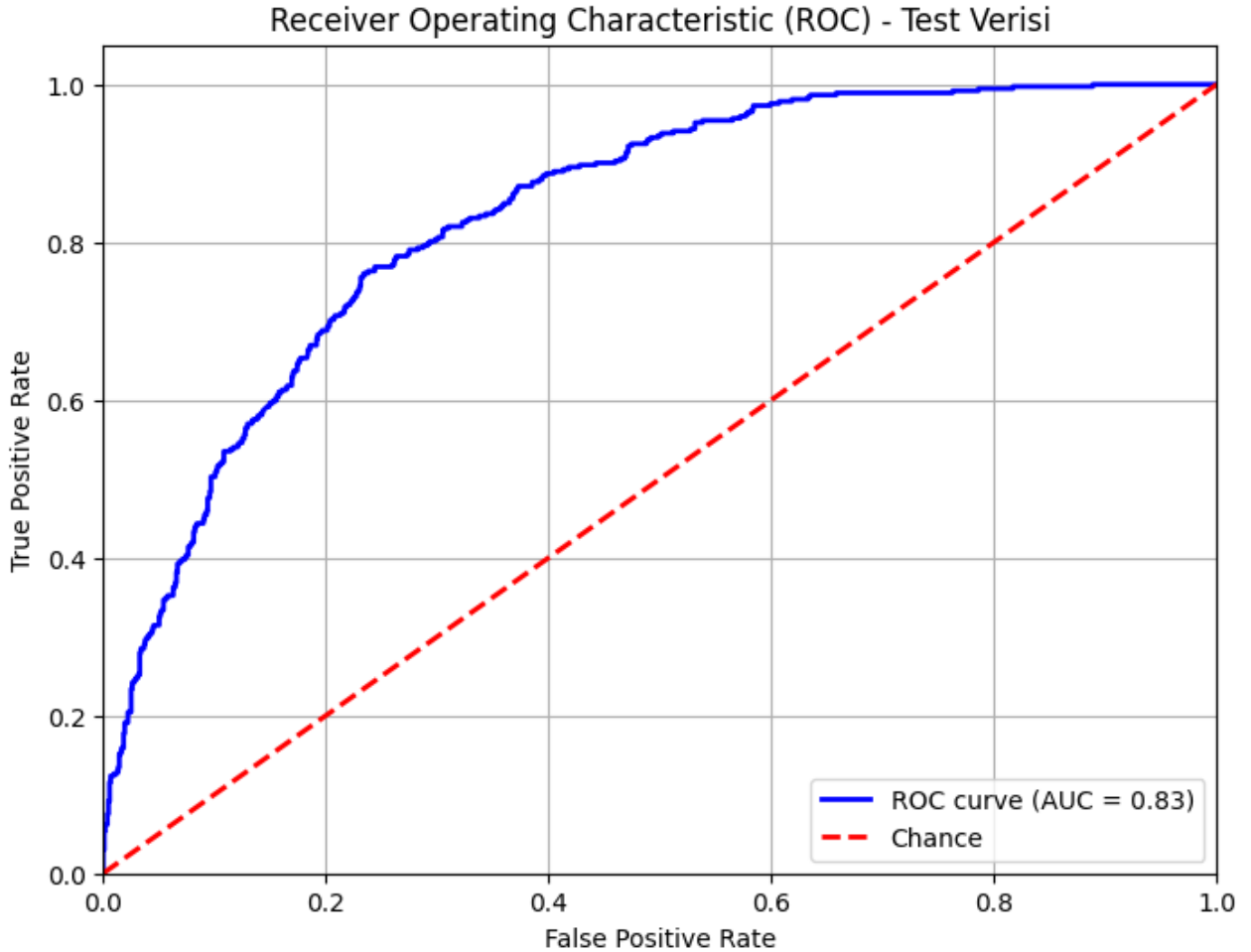


Şekil 4. 45 XGBoost Modeli - Foldların ROC Eğrisi

Modelin test verisi üzerindeki performansı (Tablo 4.26) incelendiğinde, doğruluk oranı %76.30, precision %53.70, recall %76.08 ve F1-Score %62.96 olarak ölçülmüştür. Bu sonuçlar, çapraz doğrulama sonuçlarıyla oldukça uyumludur ve modelin aşırı öğrenme (overfitting) sergilemediğini göstermektedir. Test verisi için elde edilen ROC eğrisi altında kalan alan da %83 olarak kaydedilmiştir (Şekil 4.26).

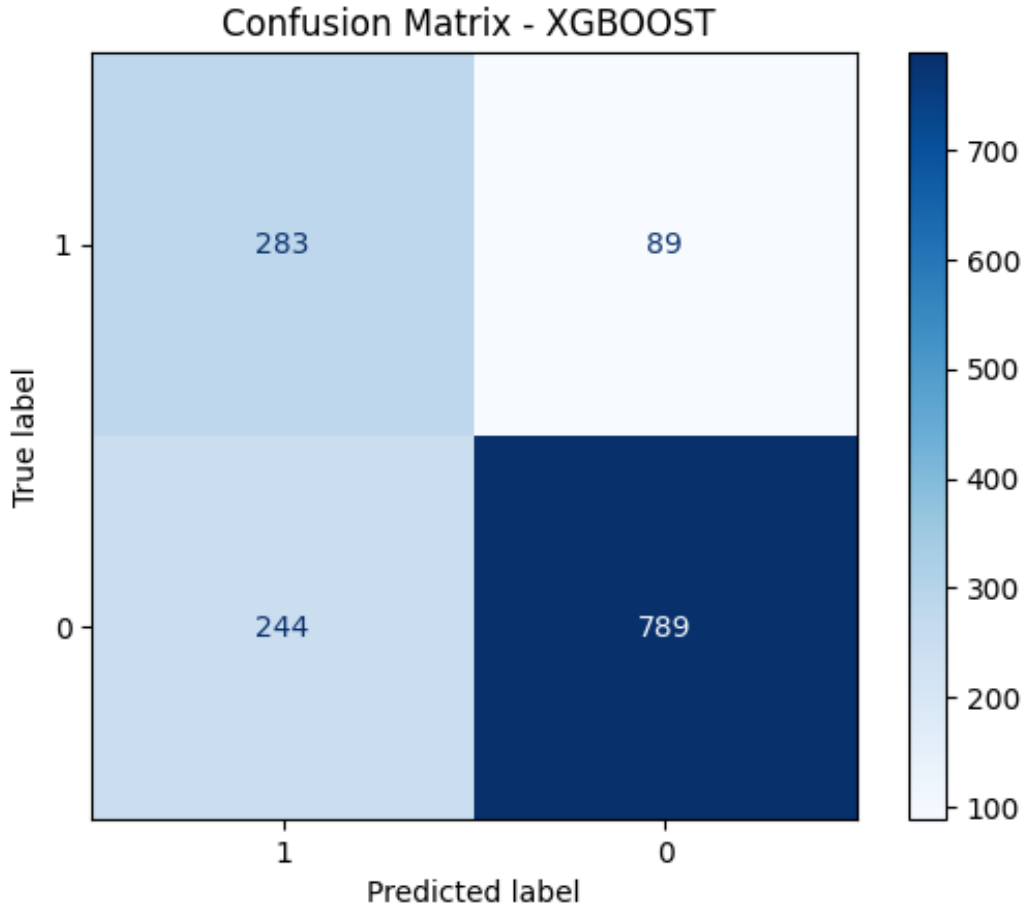
Tablo 4. 26 XGBoost Modeli - Test Verisi Performansı

METRİK	DEĞER
ACCURACY	0.7630
PRECISION	0.5370
RECALL	0.7608
F1-SCORE	0.6296
AUC	0.83



Şekil 4. 46 XGBoost Modeli - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.47) incelendiğinde modelin 283 churn eden müşteriyi doğru sınıflandırdığı, 89 churn eden müşteriyi ise gözden kaçırdığı görülmektedir. Ayrıca, 244 churn etmeyen müşteri yanlış şekilde churn sınıfına atanmıştır. Bu dağılım, modelin recall değerini yüksek tutarken bazı yanlış pozitif sınıflandırmalara neden olduğunu ortaya koymaktadır. Ancak bu durum, veri setinin dengesizliği göz önüne alındığında, makul bir ödünleşim olarak değerlendirilebilir.



Şekil 4. 47 XGBoost - Confusion Matrix

Detaylı sınıflandırma raporuna (Tablo 4.27) göre churn sınıfı için precision %54, recall %76 ve F1-score %63 olarak belirlenmiştir. Churn etmeyen sınıf için bu değerler sırasıyla %90, %76 ve %83 düzeyindedir. Bu metrikler, modelin özellikle churn eden kullanıcıları yakalamada başarılı olduğunu, ancak doğruluktan zaman zaman taviz verdiğini göstermektedir. Aynı zamanda, makro ortalamaların da (precision %72, recall %76, F1-score %73) yüksek olması, modelin sınıflar arasında belirgin bir dengesizlik yaratmadığını ve genel olarak sağlam bir performans ortaya koyduğunu doğrulamaktadır.

Tablo 4. 27 XGBoost Modeli - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.90	0.76	0.83	1033
Churned	0.54	0.76	0.63	372
ACCURACY			0.76	1405
Macro AVG	0.72	0.76	0.73	1405
Weighted AVG	0.80	0.76	0.77	1405

Sonuç olarak, SMOTE uygulanmamış XGBoost modeli, hem eğitim sürecindeki çapraz doğrulama metriklerinde hem de test verisi üzerindeki başarı oranlarında istikrarlı bir yapı sergilemiştir. Özellikle AUC ve recall değerlerinin yüksek oluşu, modelin churn tahmini gibi kritik bir sınıflandırma probleminde kullanılabilirliğini desteklemektedir. Modelin overfitting

eğiliminde olmadığı, çünkü eğitim ve test performanslarının birbirine oldukça yakın olduğu açıkça görülmektedir. Bu yapıdaki XGBoost modeli, özellikle daha hassas sınıf ayrımı gerektiren görevler için güçlü bir temel oluşturmaktadır.

4.3.5.1 XGBoost (eXtreme Gradient Boosting) – SMOTE + Threshold Uygulamalı

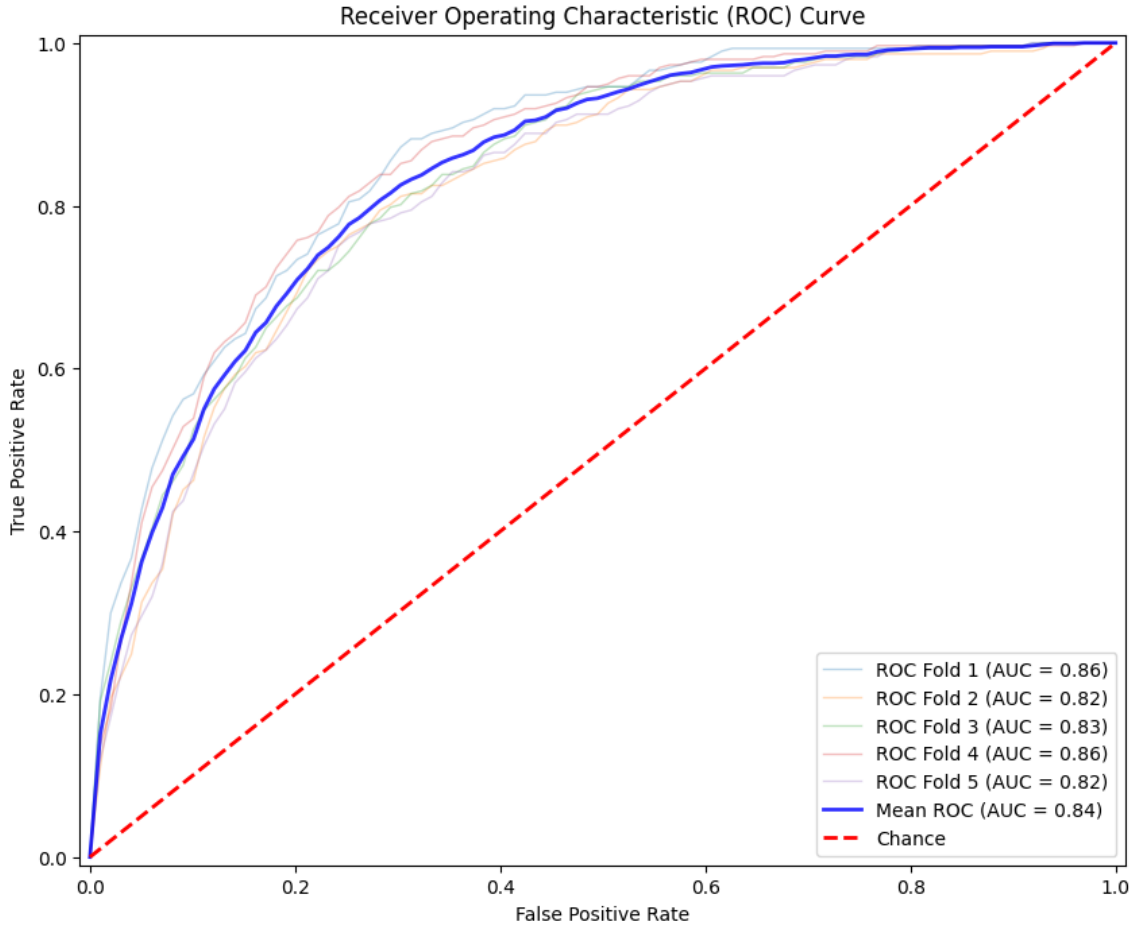
Bu bölümde, XGBoost sınıflandırma algoritması üzerine SMOTE (Synthetic Minority Over-sampling Technique) ve optimal threshold belirleme stratejilerinin uygulanmasıyla geliştirilen modelin performans analizine yer verilmiştir. Bu yapılandırma, özellikle veri kümesinde sınıf dengesizliğinin yarattığı modelleme problemlerini azaltmak ve sınıflandırma kararlılığını artırmak amacıyla tasarlanmıştır. Çünkü orijinal veri kümesinde churn etmeyen müşteriler churn edenlere oranla yaklaşık üç kat daha fazladır; bu durum da modellerin daha baskın sınıf olan “churn etmeyen” gruba doğru yanlı tahminler üretmesine neden olmaktadır.

SMOTE uygulaması sayesinde her bir fold'da eğitim verisi içinde churn=0 ve churn=1 örneklerinin sayısı eşitlenmiş ve sınıflar dengelenmiştir. Bu işlem, nadir görülen churn sınıfının daha etkili öğrenilmesini sağlamayı hedeflemiştir. Sınıf dengesi sağlandıktan sonra model çıktıları olasılık skorları olarak elde edilmiş, bu skorlar üzerinden farklı eşik değerleri denenerek her bir fold için en yüksek F1-Score'u sağlayan **optimal threshold** belirlenmiştir. Bu eşikler fold'lar arasında farklılık göstermiş; en düşük eşik değeri 0.27 civarındayken, en yüksek değer 0.41 seviyesine kadar çıkmıştır. Bu durum, veri bölünmelerinin model karar eşiği üzerindeki etkisini ve her fold'un sınıflandırma dinamiğinin farklılaşabileceğini ortaya koymaktadır.

XGBoost modelinin 5 katlı çapraz doğrulama sonuçları (Tablo 4.28) incelendiğinde; ortalama doğruluk (Accuracy) %76,48, precision %54.05, recall %77.04 ve F1-Score %63,37 olarak hesaplanmıştır. Modelin ROC eğrisi (Şekil 4.48) altında kalan alanlarının (AUC) her fold için %82 ile %86 arasında değiştiği, genel ortalamanın ise %84 olduğu gözlemlenmiştir. Bu bulgular, SMOTE uygulaması ve eşik ayarının, özellikle pozitif sınıfı daha iyi yakalayarak modelin ayırt edici gücünü koruduğunu göstermektedir. Ayrıca yüksek AUC değerleri, modelin gerçek pozitif ve negatif sınıfları birbirinden ayırma yeteneğinin güçlü olduğunu ortaya koymaktadır.

Tablo 4. 28 XGBoost Modeli - SMOTE Uygulanan Cross Validation Sonuçları

	FOLD - 1	FOLD - 2	FOLD - 3	FOLD - 4	FOLD - 5	FOLDLAR ORT.
ACCURACY	0.7447	0.7649	0.7640	0.7907	0.7596	0.7648
PRECİSİON	0.5098	0.5403	0.5404	0.5799	0.5322	0.5405
RECALL	0.8788	0.7441	0.7205	0.7576	0.7508	0.7704
F1-SCORE	0.6452	0.6261	0.6176	0.6569	0.6229	0.6337
AUC	0.86	0.82	0.83	0.86	0.82	0.84
OPTİMAL THRESHOLD	0.2766	0.3779	0.4117	0.4076	0.3790	0.3705425262451172



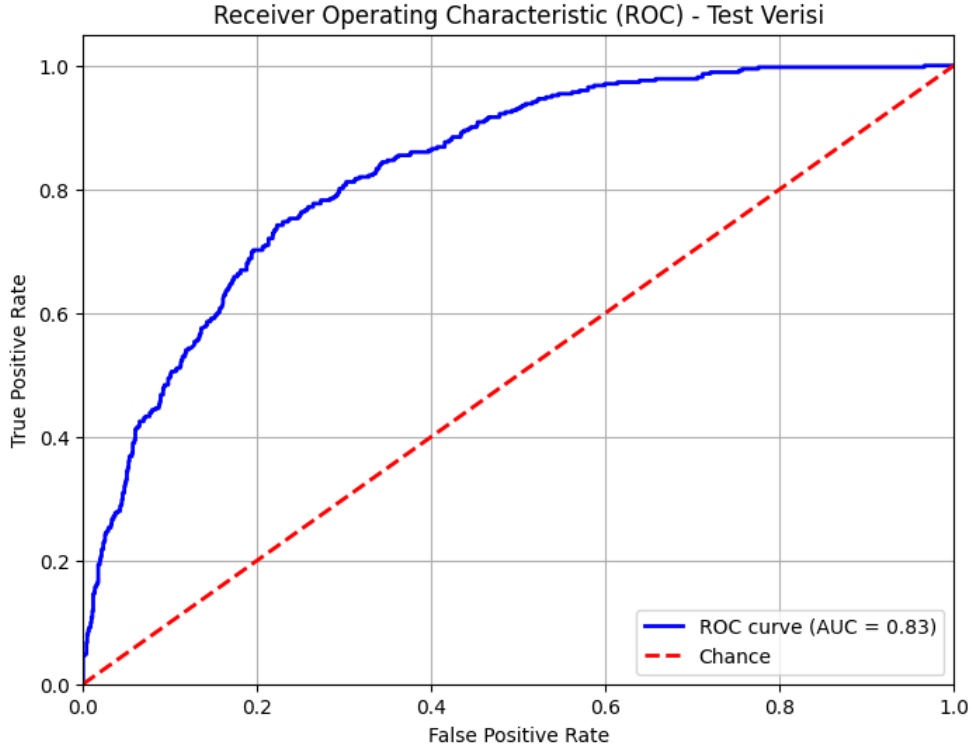
Şekil 4. 48 XGBoost Modeli - SMOTE Uygulanan Foldların ROC Eğrisi

Modelin test verisi üzerindeki performansı da incelenmiştir. Test setine uygulanan modelde, %76.73 doğruluk, %54.46 precision, %73.92 recall, %62.71 F1-score ve %83 AUC değeri elde edilmiştir (Tablo 4.29). Bu sonuçlar, çapraz doğrulama sürecinde elde edilen metriklerle büyük ölçüde tutarlıdır. Bu durum, modelin aşırı öğrenmeden (overfitting) kaçındığını ve gerçek dünyadaki yeni veriler üzerinde benzer başarıyı sürdürebileceğini göstermektedir. Ayrıca test verisinde belirlenen optimal threshold değeri 0.3705'tir. Bu eşik, varsayılan 0.5'lik karar sınırına göre churn sınıfını daha fazla önceliklendirmektedir.

Tablo 4. 29 XGBoost Modeli (SMOTE) - Test Verisi Performansı (Threshold Uygulandı)

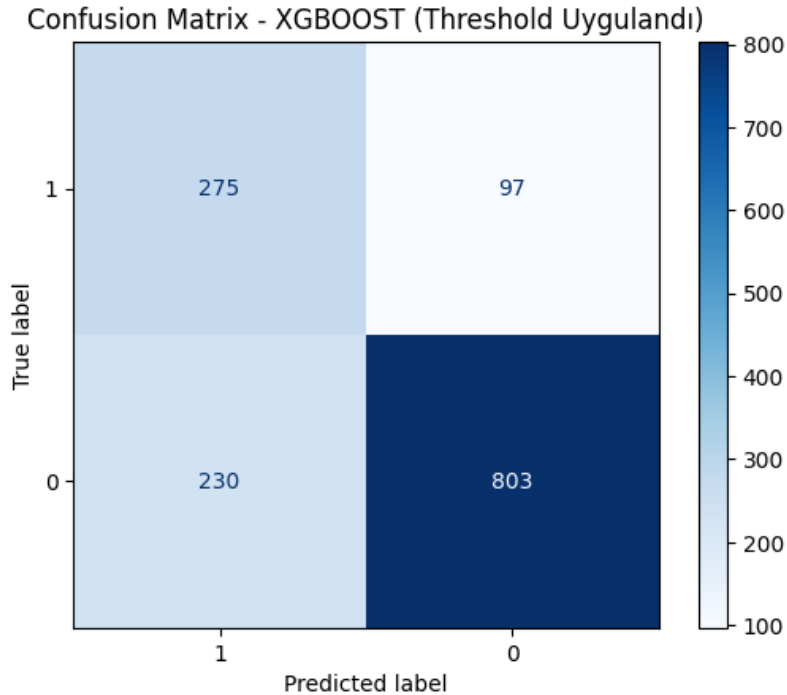
METRİK	DEĞER
ACCURACY	0.7673
PRECISION	0.5446
RECALL	0.7392
F1-SCORE	0.6271
AUC	0.83
OPTİMAL THRESHOLD	0.3705425262451172

Test verisi için elde edilen ROC eğrisi altında kalan alan da %83 olarak kaydedilmiştir (Şekil 4.49).



Şekil 4. 49 XGBoost Modeli (SMOTE) - Test Verisi ROC Eğrisi

Confusion matrix (Şekil 4.50) analizine göre model, 275 churn eden müşteriyi doğru şekilde tahmin etmiş (True Positive), 97 churn müşterisini hatalı biçimde churn etmeyen olarak sınıflandırmıştır (False Negative). Aynı zamanda, 230 churn etmeyen müşteri yanlışlıkla churn olarak tahmin edilmiştir (False Positive). Bu dağılım, modelin recall oranını artırmak için bazı yanlış pozitif tahminleri tolere ettiğini, ancak bu durumun precision değerini ciddi şekilde düşürmediğini göstermektedir. Bu da sınıflar arası dengeyi sağlamak adına kabul edilebilir bir ödünleşimdir.



Şekil 4. 50 XGBoost - Confusion Matrix (Threshold Uygulandı)

Detaylı sınıflandırma raporuna (Tablo 4.30) göre churn sınıfı için precision %54, recall %74 ve F1-score %63 olarak hesaplanmıştır. Bu değerler, SMOTE uygulanmamış XGBoost modeline kıyasla özellikle **recall ve F1-score** açısından kayda değer bir gelişmeye işaret etmektedir. Bununla birlikte, churn etmeyen sınıf (Not Churned) için de precision %89, recall %78 ve F1-score %83 gibi yüksek başarı oranları elde edilmiştir. Bu durum, SMOTE işleminin yalnızca azınlık sınıfın başarısını artırmakla kalmayıp, modelin genel dengesini koruyarak her iki sınıf için anlamlı tahmin performansı sunduğunu göstermektedir.

Tablo 4. 30 XGBoost Modeli (SMOTE) - Detaylı Sınıflandırma Raporu

	PRECİSİON	RECALL	F1-SCORE	SUPPORT
Not Churned	0.89	0.78	0.83	1033
Churned	0.54	0.74	0.63	372
ACCURACY			0.77	1405
Macro AVG	0.72	0.76	0.73	1405
Weighted AVG	0.80	0.77	0.78	1405

Sonuç olarak, SMOTE ve optimal threshold ayarlamalarıyla desteklenen XGBoost modeli, churn tahmini problemine yönelik güçlü ve dengeli bir sınıflandırma çözümü sunmuştur. Gerek çapraz doğrulama gerekse test verisi sonuçları, bu yapılandırmanın genellenebilirlik açısından tatmin edici olduğunu ortaya koymaktadır. Bu yapı, yalnızca performans iyileştirmesi sağlamakla kalmamış; aynı zamanda dengesiz sınıflar üzerinde karar verme sürecini daha adil ve kontrollü hale getirmiştir. Bu bağlamda XGBoost modeli, SMOTE ve eşik optimizasyonu gibi ön işlemlerle birlikte değerlendirildiğinde, churn sınıflandırması için güçlü bir alternatif olarak öne çıkmaktadır.

4.3.6 Modellerin Karşılaştırmalı Performans Analizi

Bu bölümde, müşteri kaybını tahmin etmeye yönelik olarak uygulanan tüm makine öğrenmesi modellerinin genel performansları karşılaştırmalı olarak analiz edilmiştir. Değerlendirme; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1-Skoru (F1-score) ve ROC-AUC (Receiver Operating Characteristic - Area Under Curve) metriklerine dayalı olarak gerçekleştirilmiştir. Karşılaştırma hem SMOTE uygulanmamış (Tablo 4.31) hem de SMOTE uygulanmış (Tablo 4.32) modeller için ayrı ayrı yapılmıştır. Bu sayede hem orijinal veri dağılımı üzerinden elde edilen sonuçlar hem de veri dengelenmiş durumda yapılan analizlerin etkisi gözlemlenebilmiştir.

SMOTE uygulanmamış modellerin karşılaştırmasında, doğruluk değeri açısından en yüksek başarı %77.08 ile Rastgele Orman (Random Forest) modeline aittir. Ancak XGBoost ve SVM modelleri de %76 seviyelerinde benzer doğruluklar sunmuştur. Duyarlılık (recall) metriğinde en iyi sonuç %77.96 ile Lojistik Regresyon modelinde elde edilirken, en yüksek AUC değeri %84 ile hem Logistic Regression hem de Random Forest modellerinde gözlenmiştir. F1-Score metriğinde ise XGBoost modeli %62,96 ile en yüksek değeri sunarak, sınıf dengesizliklerine rağmen genel başarımda rekabetçi bir performans göstermiştir.

Tablo 4. 31 Modellerin Karşılaştırılmış Performans Tablosu

	Lojistik Regresyon (Logistic Regression)	Karar Ağacı (Decision Tree)	Rastgele Orman (Random Forest)	Destek Vektör Makinesi (SVM)	XGBoost (eXtreme Gradient Boosting)
ACCURACY	0.7409	0.7338	0.7708	0.7601	0.7630
PRECİSİON	0.5070	0.4983	0.5530	0.5325	0.5370
RECALL	0.7796	0.7715	0.7016	0.7715	0.7608
F1-SCORE	0.6144	0.6055	0.6185	0.6301	0.6296
ROC / AUC	0.84	0.83	0.84	0.82	0.83

Öte yandan, SMOTE uygulanan modellerin karşılaştırması yapıldığında, özellikle azınlık sınıfına ait örneklerin daha iyi temsil edilmesiyle duyarlılık (recall) oranlarında gözle görülür bir iyileşme yaşandığı fark edilmiştir. Bu iyileşme, özellikle XGBoost modelinde %74.0'a ulaşarak dikkat çekmiştir. Ayrıca SMOTE uygulaması sonrası SVM modeli %78.51 doğruluk oranı ile diğer tüm modellerin önüne geçmiştir. F1-Score açısından ise XGBoost, %62.71 ile en iyi sonucu veren model olmuştur. AUC değerleri genel olarak tüm modellerde yüksek kalmış ve çoğunlukla %82 ila %84 bandında seyretmiştir.

Tablo 4. 32 SMOTE Uygulanan Modellerin Karşılaştırılmış Performans Tablosu

	Lojistik Regresyon (Logistic Regression)	Karar Ağacı (Decision Tree)	Rastgele Orman (Random Forest)	Destek Vektör Makinesi (SVM)	XGBoost (eXtreme Gradient Boosting)
ACCURACY	0.7701	0.7544	0.7665	0.7851	0.7673
PRECİSİON	0.5543	0.5260	0.5423	0.5814	0.5446
RECALL	0.6720	0.7339	0.7581	0.6720	0.7392
F1-SCORE	0.6075	0.6128	0.6323	0.6234	0.6271
ROC / AUC	0.84	0.82	0.84	0.82	0.83

SMOTE uygulanmasıyla birlikte modellerde gözlemlenen bu performans artışı, veri kümesindeki sınıf dengesizliğinin makine öğrenmesi algoritmaları üzerindeki etkisini somut olarak ortaya koymaktadır. Özellikle duyarlılık gibi azınlık sınıfını temsil eden metriklerdeki artışlar, churn (müşteri kaybı) sınıfının daha iyi tanımlanabildiğini göstermektedir. Ancak bu süreçte precision değerlerinde kısmi düşüşlerin olduğu da gözlenmiştir. Bu durum, modelin churn olmayan bazı müşterileri yanlış şekilde churn olarak sınıflandırmasına yol açabileceği anlamına gelir. Dolayısıyla uygulamada bu tür hataların maliyeti göz önünde bulundurularak modellerin dengeli şekilde değerlendirilmesi önem arz etmektedir.

Sonuç olarak, yapılan bu karşılaştırmalı analiz; churn tahmini gibi sınıf dengesizliğine sahip problemlerde hem temel metriklerle hem de dengeleme teknikleriyle yapılan değerlendirmelerin birlikte ele alınmasının gerekli olduğunu göstermektedir. SVM ve XGBoost gibi modeller, özellikle SMOTE ile desteklendiklerinde hem eğitim hem de test verisi üzerinde başarılı ve dengeli performanslar sergileyerek churn analizi için güçlü adaylar olduklarını ortaya koymuşlardır.

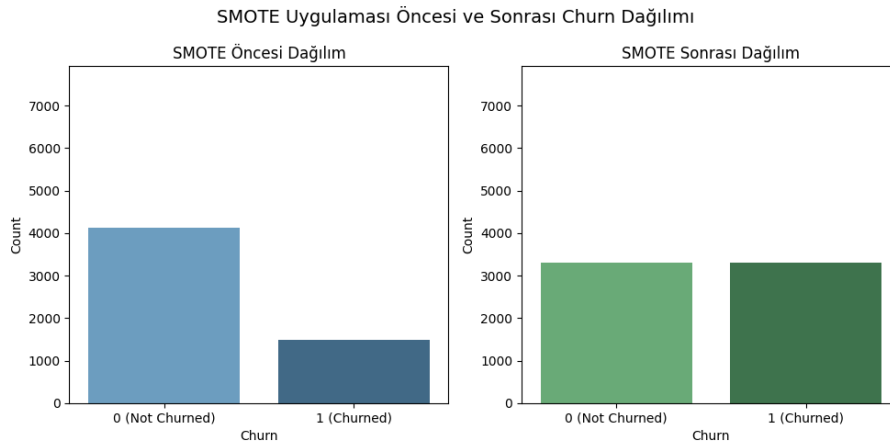
4.4 SMOTE, Eşik (Threshold) Ayarlaması ve K-Fold Doğrulama Yöntemlerinin Uygulamadaki Rolü

4.4.1 SMOTE Tekniği ile Sınıf Dengesizliğinin Giderilmesi

Makine öğrenmesi ile yapılan sınıflandırma çalışmalarında, hedef değişkenin (etiket) dengesiz dağılımı, modelin öğrenme sürecinde önemli bir sorun teşkil edebilir. Bu çalışmada kullanılan veri setinde “Churn” (müşteri kaybı) oranı dengesiz olup, churn olmayan müşteri sayısı churn edenlere göre oldukça fazlaydı. Bu durum, modellerin çoğunluk sınıfına (churn olmayanlar) öncelik vererek azınlık sınıfı (churn edenler) göz ardı etmesine neden olabilmektedir. Bu tür bir önyargı, özellikle churn eden müşterilerin doğru tahmini açısından kritik bir zayıflıktır. Bu sorunu gidermek amacıyla, eğitim verisi üzerinde **SMOTE (Synthetic Minority Over-sampling Technique)** uygulanmıştır. SMOTE, azınlık sınıfına ait örneklerden yeni sentetik örnekler üreterek sınıflar arasındaki dengesizliği azaltır. Ancak SMOTE yalnızca eğitim verisine uygulanmıştır; test seti bu işlemten etkilenmemiştir. Böylece test aşamasında gerçek dağılıma uygun değerlendirme yapılması sağlanmıştır.

Her bir fold için uygulanan SMOTE sonrasında churn=0 ve churn=1 sınıflarındaki örnek sayıları eşitlenmiş ve %50-%50 dengeli bir dağılım elde edilmiştir (Şekil 4.51). Bu işlem sonrası modellerin azınlık sınıfına olan duyarlılığında belirgin bir artış gözlemlenmiştir. Nitekim, SMOTE uygulandıktan sonra elde edilen sonuçlarda **recall** (duyarlılık) metriklerinin yükseldiği görülmektedir. Bu metrikteki artış, modellerin churn eden müşterileri daha isabetli tahmin edebildiğini göstermektedir.

Ayrıca, SMOTE uygulaması yalnızca duyarlılığı değil, F1-Skoru gibi dengeleyici metrikleri de olumlu yönde etkilemiştir. SMOTE uygulanmamış versiyonlarda bazı modeller yüksek doğruluk elde etse de churn sınıfındaki performansları düşüktü. Bu nedenle SMOTE, özellikle dengesiz veri ile çalışan sınıflandırma modellerinde oldukça etkili bir ön işleme adımı olarak değerlendirilmiştir.



Şekil 4. 51 SMOTE Uygulaması Öncesi ve Sonrası Churn Dağılımı

4.4.2 Optimal Eşik (Threshold) Değerinin Belirlenmesi ve Kullanımı

Makine öğrenmesi sınıflandırma modelleri çoğunlukla, her bir gözlem için belirli bir sınıfa ait olma olasılığı üretir. Bu olasılıklar varsayılan olarak 0.5 eşik değeri kullanılarak sınıf etiketine dönüştürülür. Ancak, özellikle dengesiz veri setlerinde bu varsayılan eşik, modelin azınlık sınıfı üzerindeki duyarlılığını (recall) olumsuz etkileyebilir. Bu nedenle, bu çalışmada varsayılan eşik değeri yerine **optimal eşik (threshold)** değeri belirlenmiş ve sınıflandırma kararları bu değer kullanılarak verilmiştir.

Optimal threshold belirleme süreci, her fold için eğitim verisi üzerinde gerçekleştirilmiştir. Öncelikle modelin churn etme olasılıklarına karşılık gelen farklı eşik değerleri denenmiş,

ardından her eşik için **F1-score** hesaplanmıştır. F1-score'un en yüksek olduğu eşik değeri, o fold için optimal threshold olarak kabul edilmiştir. Bu yaklaşım, precision ve recall arasında denge kurularak hem yanlış pozitif hem de yanlış negatiflerin minimize edilmesini hedeflemiştir.

Ölçülen optimal threshold değerleri modelden modele ve fold'dan fold'a değişkenlik göstermektedir (Tablo 4.33). Örneğin Logistic Regression modeli için optimal threshold değerleri 0.38 ile 0.53 arasında değişirken, XGBoost modeli için bu değer 0.27 ile 0.41 aralığında kalmıştır. Bu farklılıklar, her bir modelin churn sınıfını ayırmadaki karar sınırının yapısal olarak farklı olduğunu ve veri alt kümesine duyarlılık gösterdiğini ortaya koymaktadır.

Uygulanan bu eşik ayarlaması sonucunda test seti üzerinde model performanslarında iyileşme gözlemlenmiştir. Özellikle recall değerlerinde yükselme olurken, F1-score metriklerinde de artış meydana gelmiştir. Bu durum, churn eden müşterilerin daha doğru şekilde tanımlanabildiğini göstermektedir. Ancak eşik ayarlaması, precision (kesinlik) değerinde sınırlı düzeyde düşüşe yol açabilmektedir. Bu durum, bazı churn etmeyen müşterilerin yanlışlıkla churn olarak tahmin edilmesi anlamına gelir. Fakat müşteri kaybını önleme gibi ticari önceliklerin olduğu durumlarda bu tür bir değişim genellikle kabul edilebilir bir ödün olarak değerlendirilmektedir.

Sonuç olarak, her model için en uygun eşik değerinin belirlenmesi, genel performansın yanı sıra özellikle churn sınıfı üzerindeki başarının artırılmasına önemli katkı sağlamıştır. Bu yöntem, sadece bir metrik optimizasyonu değil, aynı zamanda işletme hedefleriyle uyumlu karar destek sistemlerinin geliştirilmesine yönelik etkili bir strateji olarak benimsenmiştir.

Tablo 4. 33 Modellerin Optimal Threshold Değeri Tablosu

	Lojistik Regresyon (Logistic Regression)	Karar Ağacı (Decision Tree)	Rastgele Orman (Random Forest)	Destek Vektör Makinesi (SVM)	XGBoost (eXtreme Gradient Boosting)
OPTİMAL THRESHOLD	0.4428	0.5096	0.4503	0.3571	0.3705
F1-SCORE	0.6075	0.6128	0.6323	0.6234	0.6271

4.4.3 K-Fold Çapraz Doğrulama ile Model Genellenebilirliğinin Artırılması

Bu çalışmada, geliştirilen tüm sınıflandırma modelleri üzerinde **5 katlı çapraz doğrulama (5-Fold Cross Validation)** yöntemi uygulanmıştır. Çapraz doğrulama, modelin eğitim verisine aşırı uyum sağlayarak (overfitting) genellenebilirliğini kaybetmesini önlemeye yardımcı olan önemli bir stratejidir. Aynı zamanda, modelin performansını yalnızca bir veri bölümü üzerinden değil, tüm veri seti üzerinde değerlendirmeye olanak tanır. Modelin genellenebilirliğini istatistiksel olarak daha doğru bir biçimde ölçebilmek ve tek bir test bölünmesine bağlı kalmamak amacıyla bu yöntem tercih edilmiştir.

Uygulamada kullanılan 5-Fold yöntemi kapsamında eğitim veri seti (toplam 5616 örnek), her biri yaklaşık %20 oranında olacak şekilde beş eşit parçaya ayrılmıştır. Her bir iterasyonda dört parça modelin eğitimi için, kalan bir parça ise doğrulama (validation) amacıyla kullanılmıştır. Bu işlem beş kez tekrarlanarak her fold bir kez doğrulama kümesi olarak değerlendirilmiştir. Böylece modelin tüm veri üzerinde test edilmesi ve her fold için farklı başarı metriklerinin elde edilmesi mümkün olmuştur.

Kullanılan yöntem, **Stratified K-Fold** olarak seçilmiştir. Bu yapı, her fold'da churn sınıf oranlarının orijinal veri kümesindeki oranlara sadık kalacak şekilde korunmasını ve dengeli bir temsil elde edilmesini sağlamıştır. Özellikle dengesiz veri kümelerinde bu strateji, fold'lar arası sınıf dağılımı farklılığını minimize etmek açısından önemlidir.

Her bir fold döngüsü içerisinde aşağıdaki adımlar izlenmiştir:

- Eğitim ve doğrulama verileri ayrılmıştır.
- **Sadece eğitim verisine SMOTE uygulanmış**, doğrulama verisine herhangi bir işlem yapılmamıştır.
- Model sıfırdan yeniden tanımlanmış ve yalnızca o fold'un eğitim verisiyle eğitilmiştir.

Bu süreç, veri sızıntısını önlemek ve her fold'un birbirinden bağımsız olmasını sağlamak açısından kritik öneme sahiptir. Eğer model, fold'lar arasında aynı nesne olarak tekrar kullanılsaydı, önceki fold'lardaki bilgileri taşıyabilir ve bu durum değerlendirmeyi yanıltıcı hale getirebilirdi. Ayrıca bu yapı, her fold için eşit şartlar altında modelin yeniden eğitilmesini garanti altına almıştır.

Çapraz doğrulama sürecinin önemli bir avantajı da modelin farklı veri alt kümeleri üzerindeki kararlılığının ölçülmesidir. Fold'lar arasında performans metriklerinde ciddi sapmaların görülmemesi, modelin istikrarlı çalıştığını ve veriye aşırı duyarlı olmadığını göstermektedir. Örneğin Logistic Regression modelinde fold'lar arası doğruluk (accuracy) değerleri %76 ile %78 arasında değişirken, AUC değerleri genellikle %83–%85 bandında kalmıştır. Benzer şekilde, SVM ve XGBoost modellerinde de tüm metriklerin düşük standart sapmalarla ortalama çevresinde toplandığı görülmüştür. Bu bulgu, modelin farklı örneklem kombinasyonları karşısında da tutarlı sonuçlar verdiğini ortaya koymaktadır.

Fold bazında eğitim ve doğrulama süreci sonucunda her bir fold için accuracy, precision, recall, F1-score ve AUC gibi performans metrikleri ayrı ayrı hesaplanmış ve ortalaması alınarak modelin genel başarımı belirlenmiştir. Modelin fold'lar arasında benzer performanslar göstermesi, istikrarlı ve aşırı öğrenmeden uzak bir yapıda olduğunu ortaya koymuştur.

K-Fold yöntemi yalnızca modelin doğruluğunu ölçmekle kalmamış, aynı zamanda **SMOTE** ve **threshold** gibi veri ön işleme stratejilerinin model performansı üzerindeki etkisini değerlendirmek için de güvenilir bir temel oluşturmuştur. Her fold içerisinde eğitim verisine SMOTE uygulanmış, ardından optimal eşik değeri belirlenerek fold bazlı sınıflandırmalar yapılmıştır. Bu yaklaşım, veri dengesizliğiyle mücadelede tutarlı ve gerçekçi sonuçların elde edilmesini sağlamıştır.

Sonuç olarak, çapraz doğrulama yöntemi, modelin değerlendirme sürecinde yalnızca bir test kümesine bağımlı kalmadan daha kapsamlı ve güvenilir bir analiz yapılmasına imkân tanımıştır. Uygulamada elde edilen tüm sonuçlar, 5 katlı doğrulama sayesinde genel veri setinin tamamını temsil edecek biçimde ortalama alınarak raporlanmış, böylece model karşılaştırmaları daha objektif bir zemine oturtulmuştur.

4.5 Kod ve Araç Altyapısı

4.5.1 Geliştirme Ortamı ve Programlama Dili

Bu çalışmada, veri işleme, modelleme ve değerlendirme süreçlerinin tamamı **Python** programlama dili kullanılarak gerçekleştirilmiştir. Python, esnek yapısı, geniş açık kaynak topluluğu, kapsamlı veri analizi kütüphaneleri ve makine öğrenmesi modellerine yönelik zengin API desteği sayesinde veri bilimi projelerinde en sık tercih edilen dillerden biridir. Python'un sözdiziminin okunabilirliği ve uygulama geliştirme kolaylığı, özellikle akademik düzeyde yürütülen projelerde tercih edilme nedenleri arasında yer almaktadır.

Uygulama, Python'un Jupyter Notebook altyapısını kullanan bulut tabanlı platform olan **Google Colab** üzerinde yürütülmüştür. Google Colab, hem etkileşimli kod yazma ve çalıştırma imkânı sağlaması hem de grafiksel çıktıları anında sunabilmesi açısından avantajlıdır. Özellikle herhangi bir yerel kurulum gerektirmemesi, yüksek RAM desteği sunması ve doğrudan Google Drive entegrasyonu sayesinde veri dosyalarına kolay erişim sağlaması bu projede kritik rol oynamıştır.

Google Colab platformu sayesinde kod hücreleri bölümlere ayrılarak sırayla çalıştırılmış; her bir model, kendi blokları altında izole şekilde test edilmiştir. Bu yapı, deneme-yanılma yoluyla yapılan parametre ayarlarının hızlıca test edilmesine, görselleştirmelerin doğrudan hücre altında sunulmasına ve yorumlama süreçlerinin kod akışına entegre edilmesine olanak tanımıştır.

Çalışmanın farklı bileşenleri ayrı notebook dosyalarında yürütülmüş; örneğin Logistic Regression, Random Forest ve XGBoost gibi modeller kendi başlıkları altında yapılandırılmıştır. Tüm not defterleri aynı Google Colab projesi altında düzenli biçimde tutulmuş, ortak veri kümesi (TELCO_data.csv) ise doğrudan Drive'a yüklenerek çağırılmıştır. Ayrıca çalışma süresince Google Colab'ın sunduğu GPU veya TPU gibi hızlandırıcılar kullanılmamıştır; çünkü kullanılan modeller CPU üzerinden yeterli sürelerde eğitilebilmiştir. Bununla birlikte, Colab'ın işlem gücü sınırları (örneğin RAM limiti) göz önünde bulundurularak veri kümesinde gereksiz sütunlar temizlenmiş ve verimli veri yapıları tercih edilmiştir.

Sonuç olarak, Python dili ve Google Colab platformunun sunduğu avantajlar sayesinde proje hem teknik olarak esnek hem de erişilebilir bir yapıya kavuşmuştur. Bu tercih, modelleme sürecini kolaylaştırmış, test ve değerlendirme adımlarını hızlandırmış ve sonuçların görsel olarak desteklenmesine olanak sağlamıştır.

4.5.2 Kullanılan Kütüphaneler ve Paketler

Bu tez çalışmasında Python programlama dili ile birlikte çok sayıda **açık kaynaklı kütüphane** ve **paket** kullanılmıştır. Söz konusu kütüphaneler, verinin ön işlenmesinden modelleme sürecine, görselleştirmeden performans değerlendirmelerine kadar her aşamada işlevsel olarak önemli roller üstlenmiştir. Aşağıda projede kullanılan başlıca kütüphaneler ve kullanım amaçlarına dair açıklamalara yer verilmiştir:

4.5.2.1 Veri Yükleme ve İşleme Kütüphaneleri

- **pandas**
Çalışmanın temel veri yapısı olan DataFrame nesnesini sağlayan bu kütüphane, .csv formatındaki veri dosyasının yüklenmesi, özniteliklerin seçilmesi, etiket sütununun dönüştürülmesi ve veri setinin bölümlendirilmesi gibi işlemlerde kullanılmıştır. Aynı zamanda eğitim ve test seti içerisindeki churn oranlarının analizinde ve veri çerçevesi yapısındaki filtreleme, sıralama, grup bazlı özet alma gibi operasyonlarda da aktif olarak rol oynamıştır.
- **numpy**
Sayısal işlem kapasitesi yüksek olan bu kütüphane, özellikle eşik değerlerinin hesaplanması, ROC eğrisi altında kalan alanların ortalamasının alınması ve rastgelelik kontrolü amacıyla kullanılan random_state parametresi gibi işlemler için tercih edilmiştir. Bunun dışında, linspace fonksiyonu yardımıyla ROC eğrisi çiziminde ortak FPR noktaları üzerinden interpolasyon yapılmış ve tüm fold'ların ortalama ROC eğrisi elde edilmiştir.

4.5.2.2 Veri Görselleştirme Kütüphaneleri

- **matplotlib.pyplot**
Grafik çizimleri için kullanılan temel kütüphanelerden biridir. Bu projede confusion

matrix görselleştirmesi, ROC eğrisi ve bar chart gibi temel görselleştirmelerde kullanılmıştır. Eksen ayarları, başlık ekleme, yazı tipi boyutu gibi birçok özelleştirme matplotlib üzerinden yapılmıştır.

- **seaborn**

Matplotlib üzerine kurulmuş olan bu daha gelişmiş görselleştirme kütüphanesi, churn değişkeninin sınıf dağılımı, kategorik değişkenlere göre churn oranları ve korelasyon matrisinin ısı haritası (heatmap) gibi grafiklerin oluşturulmasında kullanılmıştır. Estetik olarak daha zengin ve yorumlaması kolay grafikler sunmuştur.

4.5.2.3 Makine Öğrenmesi ve Modelleme Kütüphaneleri

- **scikit-learn (sklearn)**

Bu proje kapsamında dört farklı makine öğrenmesi modeli scikit-learn üzerinden uygulanmıştır:

- LogisticRegression
- DecisionTreeClassifier
- RandomForestClassifier
- SVC (Destek Vektör Makineleri)

Bunun yanı sıra aşağıdaki fonksiyonlar ve sınıflar da kullanılmıştır:

- train_test_split: Verinin eğitim ve test setine ayrılması
- StratifiedKFold: Her fold'da sınıf oranlarının korunarak bölünmesi
- StandardScaler: SVC gibi ölçek duyarlı modeller için veri standardizasyonu
- classification_report, confusion_matrix, roc_curve, roc_auc_score, precision_score, recall_score, f1_score, accuracy_score: Model performansının değerlendirilmesi
- ConfusionMatrixDisplay: Confusion matrix grafikleştirme

Scikit-learn modüler yapısı sayesinde tüm modeller aynı sistematığa bağlı kalınarak geliştirilmiş ve karşılaştırılabilir hale getirilmiştir.

4.5.2.4 Dengesiz Veri Yapısı İçin Kullanılan Paketler

- **imblearn (imbalanced-learn)**

SMOTE algoritması bu kütüphane üzerinden uygulanmıştır.

imblearn.over_sampling.SMOTE sınıfı ile yalnızca eğitim verisi üzerinde örnek sayıları eşitlenmiş; test ve validasyon setleri özgün yapısını korumuştur. Böylece veri sızıntısı engellenmiş ve modelin azınlık sınıf olan “churn” sınıfındaki başarısı iyileştirilmiştir.

4.5.2.5 Boosting Tabanlı Model Kütüphanesi

- **xgboost**

XGBClassifier sınıfı üzerinden kullanılan bu kütüphane, özellikle ROC-AUC ve F1-score gibi metriklerde yüksek performans göstermesi nedeniyle projeye dahil edilmiştir. XGBoost modeli, scikit-learn uyumlu API'si sayesinde diğer modellerle paralel yapılandırılmış ve threshold ile SMOTE gibi stratejilerle birlikte çalıştırılmıştır.

4.5.2.6 Ek Fonksiyonel Kullanımlar

- **collections**

Counter sınıfı ile churn sınıfının sayısal frekanslarının elde edilmesinde ve SMOTE sonrası veri dengesinin kontrol edilmesinde kullanılmıştır.

Tüm bu kütüphaneler sayesinde proje hem yapılandırılmış hem de esnek bir biçimde yönetilebilir hâle gelmiştir. Kodlamanın modüler yapısı, modellerin tekrar edilebilirliğini artırmış ve performans karşılaştırmalarının objektif olarak yürütülmesini sağlamıştır. Çalışma

süresince Google Colab ortamı tercih edildiğinden bu kütüphanelerin çoğu önceden yüklü gelmiştir; eksik olanlar ise !pip install komutu aracılığıyla kolaylıkla yüklenmiştir.

4.5.3 Kodlama Akışı ve Uygulama Adımları

Çalışma boyunca kullanılan tüm kodlama adımları, yeniden üretilebilirlik ve sistematik analiz amacıyla belirli bir yapıya oturtulmuştur. Kodlama süreci, verinin yüklenmesinden model performanslarının karşılaştırılmasına kadar uzanan altı temel aşamada gerçekleştirilmiştir. Şekil 4.20 de veri analizi ve modelleme sürecinin adımlarını sırasıyla göstermektedir. Aşamaların her biri, hem içerik hem de uygulama olarak birbirine bağlı ve mantıksal bir akış içerisinde :

4.5.3.1 Veri Yükleme ve İlk İnceleme

Veri seti Kaggle üzerinden .csv formatında temin edilerek Google Drive'a yüklenmiş ve Google Colab ortamında pandas kütüphanesi ile okunmuştur. Bu aşamada aşağıdaki işlemler uygulanmıştır:

- Dosyanın veri çerçevesine (DataFrame) dönüştürülmesi
- Boş veya hatalı verilerin tespiti
- Sayısal sütunların tür dönüştürme işlemleri (object → float)
- Hedef sütunun (Churn) etiketlenmesi (Yes/No → 1/0)
- Veri setinin genel yapısının (shape, info(), describe()) incelenmesi

4.5.3.2 Keşifsel Veri Analizi (EDA)

Bu aşamada veri setindeki hem hedef değişkenin hem de açıklayıcı değişkenlerin genel özellikleri analiz edilmiştir:

- Churn değişkeninin sınıf oranı (value_counts, normalize=True)
- Kategorik değişkenlerin churn ile olan ilişkileri (barplot)
- Sayısal değişkenlerin dağılımı ve çarpıklık durumu (histogram, boxplot)
- Korelasyon matrisi (df.corr()) ve ısı haritası (seaborn heatmap)
- EDA çıktılarının görsellerle tezde raporlanması

Bu adım, veri setindeki önemli eğilimleri ve model kurulumunda hangi özniteliklerin etkili olabileceğini ortaya koymuştur.

4.5.3.3 Eğitim ve Test Setlerine Bölme

Veri seti, scikit-learn'ün train_test_split fonksiyonu kullanılarak %80 eğitim ve %20 test olacak şekilde ayrılmıştır. Bu bölme işleminde hedef değişkenin oransal yapısının korunması için stratify=y parametresi kullanılmıştır. Eğitim verisi model eğitimi ve çapraz doğrulama için kullanılırken; test verisi, modellerin nihai performansını ölçmek üzere ayrı tutulmuştur.

4.5.3.4 SMOTE + K-Fold + Threshold Entegrasyonu

Bu çalışmanın temel teknik farklılığı, model eğitimi aşamasına üç yapının entegre edilmesidir:

- **SMOTE:** Eğitim verisi üzerinde azınlık sınıf olan churn = 1 için yeni sentetik örnekler üretilmiş ve sınıf dengesi sağlanmıştır.
- **Stratified K-Fold:** Eğitim verisi 5 fold'a ayrılmış, her fold'da sınıf dağılımı korunmuştur.
- **Threshold Ayarlaması:** Her fold için farklı eşik (threshold) değerleri test edilerek F1-score'u en yüksek yapan optimal değer belirlenmiştir.

Bu yapı sayesinde modeller, hem dengeli veriyle öğrenmiş hem de farklı veri bölünmelerine karşı genellenebilirliği test edilmiştir.

4.5.3.5 Model Eğitimi ve Değerlendirme

Her model (LogisticRegression, DecisionTreeClassifier, RandomForestClassifier, SVC, XGBClassifier) için ayrı ayrı:

- Eğitim gerçekleştirilmiş
- Her fold'da performans metrikleri hesaplanmış
- ROC eğrileri çizilmiş
- Test verisine uygulanarak nihai başarı ölçülmüştür

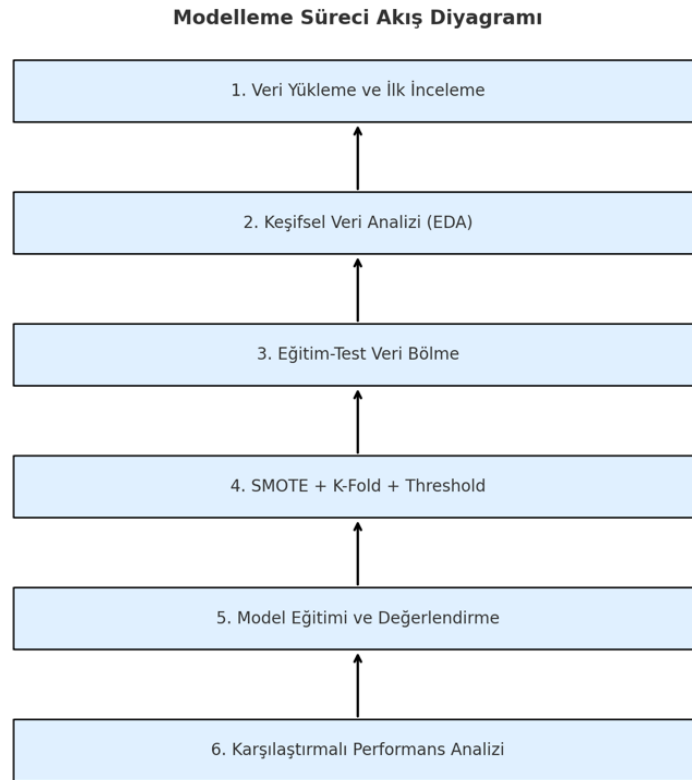
Değerlendirme metrikleri şunlardır:

- Accuracy (Doğruluk)
- Precision (Kesinlik)
- Recall (Duyarlılık)
- F1-Score
- ROC-AUC

ROC eğrileri ve confusion matrix görselleri her model için ayrı ayrı görselleştirilmiştir.

4.5.3.6 Karşılaştırmalı Performans Tablosu ve Sonuçlandırma

Son aşamada tüm modellerin test verisi üzerindeki nihai metrikleri bir tablo hâlinde derlenmiş, hem SMOTE'lu hem SMOTE'suz versiyonları analiz edilmiştir. Farklı modellerin hangi metriklerde üstünlük gösterdiği, genel dengesi ve uygunlukları yorumlanmış; son değerlendirme metinle birlikte sunulmuştur.



Şekil 4. 52 Makine Öğrenmesi Süreci Akış Diyagramı

5. SONUÇLAR VE TARTIŞMA

5.1 Bulguların Genel Değerlendirmesi

Bu çalışmada, müşteri kaybını (churn) öngörmeye yönelik olarak farklı makine öğrenmesi algoritmaları kullanılmış ve bu modellerin performansları çeşitli ölçütler üzerinden detaylı biçimde analiz edilmiştir. Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) ve XGBoost algoritmaları, hem dengesiz veri üzerinde hem de SMOTE + threshold + 5-Fold çapraz doğrulama entegrasyonu ile yeniden eğitilerek değerlendirilmiştir.

Elde edilen bulgular, veri dengesizliğinin giderilmesi ve eşik değeri ayarlaması gibi stratejilerin model performansı üzerinde doğrudan etkili olduğunu göstermiştir. Özellikle azınlık sınıf olan churn müşterilerinin doğru şekilde sınıflandırılabilmesi açısından, yalnızca temel model eğitimi yapmak yeterli olmamış; veri ön işleme ve model sonrası karar aşamalarına yapılan müdahaleler (örneğin threshold optimizasyonu) performans metriklerini anlamlı biçimde iyileştirmiştir.

Uygulanan modeller karşılaştırıldığında, genel olarak **XGBoost** modelinin, özellikle **ROC-AUC** ve **F1-score** gibi dengesiz sınıflarda duyarlılığı yüksek ölçütlerde, en başarılı sonuçları verdiği gözlemlenmiştir. Bu model, sınıflar arası karmaşıklığı daha iyi ayırt edebilmiş ve churn sınıfını belirlemede diğer modellere kıyasla daha dengeli tahminleme yapmıştır. ROC eğrisi altında kalan alanın XGBoost için %84-85 civarında olması, modelin güçlü bir ayırt edici yeteneğe sahip olduğunu ortaya koymuştur.

Buna karşılık, **Support Vector Machine (SVM)** ve **Random Forest** modelleri de benzer şekilde rekabetçi sonuçlar üretmiş; bu modeller, özellikle recall (duyarlılık) açısından churn sınıfındaki başarıyı yukarı çekmeyi başarmıştır. Ancak SVM'nin bazı durumlarda precision değerinin düşmesi, churn olmayan kullanıcıları yanlış pozitif olarak tahmin etme eğilimini artırmıştır.

Logistic Regression modeli, doğrusal yapısı nedeniyle karmaşık örüntüleri yakalama konusunda diğer modellere göre daha sınırlı kalmıştır. Bununla birlikte, SMOTE ve threshold uygulandıktan sonra F1-score ve recall değerlerinde kayda değer iyileşmeler gözlemlenmiş, özellikle aşırı öğrenme (overfitting) riskine karşı daha stabil bir model olarak değerlendirilmiştir.

Decision Tree modeli ise genellikle diğer modellere göre daha değişken sonuçlar üretmiş; bu durum, ağacın dallanma yapısına ve veri bölünmesine duyarlılığından kaynaklanmış olabilir. Karar ağaçlarında overfitting eğilimi gözlemlenmiş; bu nedenle genel performans ortalamalarının diğer modellere göre daha dalgalı olduğu görülmüştür.

Tüm modeller için geçerli bir bulgu da şudur: **SMOTE uygulaması ile sınıf dengesizliği giderildikten sonra, modellerin recall ve F1-score değerlerinde dikkate değer artışlar meydana gelmiştir.** Bu durum, sınıflandırma problemlerinde yalnızca doğruluk (accuracy) değerine bakmanın yetersiz olduğunu, özellikle azınlık sınıf için optimize edilmiş metriklere odaklanmanın daha sağlıklı sonuçlar verdiğini göstermektedir.

Sonuç olarak, bu çalışmada elde edilen bulgular; yalnızca doğru modeli seçmenin değil, aynı zamanda uygun veri işleme ve karar sonrası stratejilerin entegre edilmesinin de churn tahmini başarısını doğrudan etkilediğini ortaya koymaktadır.

5.2 Tartışma ve Literatürle Karşılaştırma

Bu çalışmada elde edilen bulgular, müşteri kaybı (churn) analizi üzerine yapılan mevcut literatürle karşılaştırıldığında çeşitli yönlerden hem örtüşen hem de farklılaşan sonuçlar sunmaktadır. İncelenen dört akademik çalışmanın her biri, farklı veri kümeleri, modelleme yaklaşımları ve başarı metrikleri kullanarak churn tahmini üzerine yoğunlaşmıştır.

Öncelikle Öztürk ve arkadaşlarının (2022) yayınladığı çalışmada, Türk Telekom'a ait anonimleştirilmiş bir veri kümesi üzerinde Decision Tree, Logistic Regression ve Naive Bayes gibi temel sınıflandırıcılar test edilmiştir. Elde ettikleri sonuçlara göre, Decision Tree modeli özellikle recall açısından yüksek performans göstermiştir. Ancak çalışmada sınıf dengesizliği problemi ele alınmamış, bu durum sonuçların özellikle churn sınıfı açısından zayıflamasına neden olmuştur. Bu durum, bizim çalışmamızda SMOTE uygulamasının doğrudan recall ve F1-score üzerindeki olumlu etkisini destekler niteliktedir.

İkinci olarak Balcı ve Ekizoğlu'nun (2021) yürüttüğü çalışmada, churn tahmininde Logistic Regression, K-Nearest Neighbors ve Random Forest modelleri değerlendirilmiş ve en yüksek başarı Random Forest ile elde edilmiştir. Ancak bu çalışmada da eşik değeri ayarlaması yapılmamış, modeller varsayılan 0.5 eşiği ile değerlendirilmiştir. Bizim çalışmamızda ise optimal threshold değerinin fold bazlı belirlenmesi sayesinde F1-score değerleri önemli ölçüde iyileştirilmiştir. Bu fark, karar sonrası optimizasyonun model başarısını nasıl etkilediğini açıkça göstermektedir.

Üçüncü olarak, "Customer Churn Prediction Using Machine Learning Methods: A Comparative Analysis" başlıklı uluslararası çalışmada, XGBoost modeli diğer tüm algoritmalara kıyasla en yüksek ROC-AUC skorunu vermiştir. Aynı çalışmada veri dengesizliği için Random Over Sampling tekniği uygulanmış; ancak bu yaklaşım modelin overfitting'e eğilimli olmasına neden olmuştur. Bizim çalışmamızda SMOTE kullanılarak dengeli veri üretimi sağlanmış, ayrıca her fold'da model yeniden eğitilerek öğrenme sızıntısı engellenmiştir. Bu yöntemsel fark, elde edilen daha stabil performans sonuçlarının temelini oluşturmaktadır.

Son olarak, Mustafa & Çalışkan'ın (2023) çalışmasında, Telco veri seti üzerinde AdaBoost, Gradient Boosting ve XGBoost gibi modeller denenmiş ve en yüksek başarı yine XGBoost modeli ile elde edilmiştir. ROC-AUC değerleri %84 civarındadır. Bu sonuç, bizim çalışmamızda XGBoost modelinin hem dengesiz hem dengeli veri üzerinde verdiği güçlü sonuçlarla tutarlılık göstermektedir. Ancak Mustafa ve Çalışkan'ın çalışmasında threshold ayarlaması yapılmamış olması, precision-recall dengesi açısından sınırlı analiz sunmuştur. Bizim çalışmamızda bu dengenin optimize edilmesi sayesinde XGBoost modelinin churn sınıfını daha sağlıklı tahmin ettiği gözlemlenmiştir.

Bu literatür karşılaştırmaları göstermektedir ki, sınıf dengesizliği ile mücadele ve karar sonrası eşik ayarlaması, model seçiminden bağımsız olarak churn tahmininde önemli başarı artışları sağlamaktadır. Ayrıca model değerlendirmesinde yalnızca doğruluk oranına (accuracy) değil, F1-score, recall ve ROC-AUC gibi daha duyarlı metriklere odaklanması, özellikle churn sınıfı gibi azınlık sınıfların doğru sınıflandırılmasını kolaylaştırmaktadır.

Bu çalışma, önceki literatürle karşılaştırıldığında hem metodolojik açıdan daha bütünlüklü bir yaklaşım sunmakta hem de uygulanan tekniklerin her biri için model performansı üzerindeki etkisini ayrı ayrı raporlamaktadır. Bu yönüyle hem kapsam hem de şeffaflık açısından daha güçlü bir yapı ortaya koymaktadır.

5.3 Kısıtlar ve Geliştirme Önerileri

Bu çalışmada müşteri kaybı tahmini amacıyla uygulanan yöntemler belirli açılardan başarılı sonuçlar üretmiş olsa da, değerlendirme sürecinde göz önünde bulundurulması gereken bazı sınırlılıklar ve geliştirilebilir yönler mevcuttur. Bu bölümde, çalışmanın kapsamı dahilinde

karşılaşılan metodolojik ve uygulamaya dönük kısıtlar açıklanmakta, gelecekte yapılabilecek iyileştirmeler önerilmektedir.

5.3.1 Veri Seti Kaynaklı Kısıtlar

Çalışmada kullanılan Telco müşteri veri seti, Kaggle üzerinden alınmış ve anonimleştirilmiş yapıda sabit bir örnekleme sahiptir. Verinin sınırlı sayıda örnek (7043 müşteri) ve belirli demografik/sözleşme bilgileri içermesi, modelin daha geniş ve çeşitlendirilmiş müşteri davranışlarını yansıtma kapasitesini sınırlandırmaktadır. Örneğin ödeme geçmişi, hizmet kesinti kayıtları veya müşteri destek etkileşimleri gibi davranışsal veriler modele dahil edilemediğinden, bu tür bilgilerin eksikliği tahmin başarısını belli bir noktada sınırlamış olabilir.

5.3.2 Modelleme Sürecindeki Kısıtlar

Bu çalışmada beş temel model (Logistic Regression, Decision Tree, Random Forest, SVM, XGBoost) kullanılmıştır. Ancak hiperparametre optimizasyonu (örneğin Grid Search veya Randomized Search gibi) sistematik biçimde uygulanmamıştır. Kullanılan modeller genellikle varsayılan parametrelerle veya sınırlı sayıda denemeyle eğitilmiştir. Bu durum, bazı modellerin potansiyel performansına tam olarak ulaşamamasına neden olmuş olabilir.

Ayrıca modellerin tümü bireysel olarak değerlendirilmiş; topluluk (ensemble) yöntemleri, model birleştirme (stacking/blending) veya voting gibi ileri düzey stratejiler test edilmemiştir. Bu tür yaklaşımlar, özellikle farklı modellerin zayıf ve güçlü yönlerini birleştirerek daha genel geçer çözümler üretebilir.

5.3.3 Karar Sonrası Strateji Optimizasyonu

Bu çalışmada fold bazlı threshold ayarlaması yapılmış ve F1-score'u optimize edecek eşik değeri seçilmiştir. Ancak bu eşik, test seti için sabit şekilde uygulanmıştır. Aslında threshold belirleme işlemi test setine özel olarak da optimize edilebilirdi. Bununla birlikte, bu tür bir işlem veri sızıntısı yaratabileceği için dikkatle planlanması gereken bir adımdır. Ayrıca ROC eğrisine alternatif olarak precision-recall eğrisi altında kalan alan (PR-AUC) gibi farklı metriklerin eşik seçimi için dikkate alınması da ilerideki çalışmalarda değerlendirilebilir.

5.3.4 Donanım ve İşlem Gücü Kısıtlamaları

Modelleme süreci Google Colab platformunda yürütülmüştür. Bu platformun sağladığı işlem gücü belirli sınırlamalar içermektedir (RAM, çalışma süresi, bağlantı kesilmesi vb.). Özellikle büyük boyutlu modellerin eğitimi sırasında bu tür sınırlamalar, bazı hiperparametre denemelerinin kapsamını daraltmıştır.

5.3.5 Gelecekteki Geliştirme Önerileri

Bu çalışmanın genişletilmesi ve daha yüksek performanslı sistemlerin geliştirilmesi için aşağıdaki öneriler dikkate alınabilir:

- **Yeni öznitelik mühendisliği:** Özellikle davranışsal ve zaman serisi verilerle daha zengin özelliklerin oluşturulması.
- **Hiperparametre optimizasyonu:** Sistematik grid search, randomized search ya da Bayesian optimization gibi yöntemlerle model ayarlarının eniyilenmesi.
- **Ensemble modellerin denenmesi:** Stacking, boosting ya da bagging gibi yöntemlerle farklı algoritmaların birlikte kullanılması.
- **Zaman temelli validasyon:** K-fold yerine time-series split gibi zaman duyarlı doğrulama yöntemlerinin test edilmesi.

- **Model açıklanabilirliği:** SHAP veya LIME gibi yöntemlerle modellerin neden bu tahminleri yaptığını görselleştirmek.
- **Gerçek zamanlı kullanım senaryosu:** Eğitilen modellerin bir web uygulaması üzerinden API ile hizmete sunulması (örneğin Flask + Firebase/Streamlit gibi araçlarla).

Bu öneriler doğrultusunda yapılacak çalışmalar, churn tahmini sürecinin doğruluğunu artırmanın yanı sıra, daha geniş kapsamlı ve pratikte uygulanabilir sistemlerin geliştirilmesine de olanak tanıyacaktır.

6.SONUÇ

Bu çalışmada, müşteri kaybı (churn) tahminine yönelik olarak makine öğrenmesi temelli bir yaklaşım geliştirilmiş ve bu kapsamda çeşitli sınıflandırma modelleri uygulanarak performans karşılaştırmaları gerçekleştirilmiştir. Veri seti olarak, telekomünikasyon sektörüne ait müşteri bilgilerini içeren ve Kaggle üzerinden temin edilen Telco veri seti kullanılmıştır. Çalışma sürecinde veri temizleme, keşifsel veri analizi, SMOTE ile dengeleme, eşik değeri optimizasyonu ve 5 katlı çapraz doğrulama gibi yöntemler bütünleşik şekilde uygulanmıştır. Uygulanan beş farklı model (Logistic Regression, Decision Tree, Random Forest, Support Vector Machine ve XGBoost) hem orijinal hem de SMOTE ve threshold ayarlaması yapılmış versiyonlarıyla test edilmiştir. Performans değerlendirmeleri doğruluk (accuracy), duyarlılık (recall), kesinlik (precision), F1-score ve ROC-AUC metrikleri üzerinden yapılmıştır. Elde edilen sonuçlara göre XGBoost modeli, özellikle ROC-AUC ve F1-score metriklerinde en başarılı sonuçları üretmiş; bu modeli Random Forest ve SVM takip etmiştir. Logistic Regression ve Decision Tree modelleri ise dengeli fakat daha düşük başarılar sergilemiştir. Çalışma sonucunda şu çıkarımlar yapılmıştır:

- Veri dengesizliğinin giderilmesi (SMOTE) ve optimal eşik belirleme, model doğruluğunu artırmada kritik rol oynamaktadır.
- Sadece accuracy metriğine odaklanmak, özellikle azınlık sınıflar için yetersiz değerlendirme sağlayabilmektedir.
- Churn tahmini gibi sınıf dengesizliği barındıran problemler için F1-score ve ROC-AUC gibi metrikler daha güvenilir ölçütler sunmaktadır.
- En iyi performansa ulaşmak yalnızca doğru modeli seçmekle değil, aynı zamanda veri hazırlama ve karar sonrası stratejilerin doğru yapılandırılmasıyla mümkündür.

Sonuç olarak bu tez çalışması, hem akademik literatürle tutarlılık göstermekte hem de pratikte uygulanabilir ve genellenebilir bir churn tahmin altyapısı ortaya koymaktadır. Gelecek çalışmalar için, hiperparametre optimizasyonu, model açıklanabilirliği (explainability) ve gerçek zamanlı sistem entegrasyonu gibi konuların araştırılması önerilmektedir.

7.KAYNAKÇA

- [1] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media.
- [2] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*. MIT Press.
- [3] Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.
- [4] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- [5] Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- [6] Brownlee, J. (2016). *Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-To-End*. Machine Learning Mastery.
- [7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [8] ResearchGate. (t.y.). *Review of Customer Churn Analysis Studies in Telecommunications Industry*.https://www.researchgate.net/publication/321723500_Review_of_Customer_Churn_Analysis_Studies_in_Telecommunications_Industry
- [9] Medium. (t.y.). *ROC ve AUC*. <https://medium.com/@gulcanogundur/roc-ve-auc-1fefcfc71a14>
- [10] Medium. (t.y.). *SMOTE - Synthetic Minority Oversampling Technique*.<https://medium.com/@ahmetokanarik/smote-synthetic-minority-oversampling-technique-c52d4fbec47e>
- [11] Medium. (t.y.). *Dengesiz Veri Kümeleri ile Makine Öğrenmesi*.<https://medium.com/@g.canguven11/dengesi%CC%87z-veri%CC%87-k%C3%BCmeleri%CC%87-i%CC%87le-maki%CC%87ne-%C3%B6%C4%9Frenmesi%CC%87-63bbac5f6869>
- [12] Medium. (t.y.). *Lojistik Regresyon Nedir, Nasıl Çalışır?*.<https://medium.com/@mfakca/lojistik-regresyon-nedir-nas%C4%B1l-%C3%A7al%C4%B1%C5%9F%C4%B1r-4e1d2951c5c1>
- [13] Medium. (t.y.). *Makine Öğrenmesi: Modellemeye Giriş (2)*.<https://medium.com/@ufukcolak/makine-%C3%B6%C4%9Frenmesi-modellemeye-giri%C5%9F-2-405b4f3c8cfb>
- [14] Medium. (t.y.). *Makine Öğrenmesi: Model Başarı Değerlendirme Ölçütleri (3)*.<https://medium.com/@ufukcolak/makine-%C3%B6%C4%9Frenmesi-model->

ba%C5%9Far%C4%B1-de%C4%9Ferlendirme-%C3%B6l%C3%A7%C3%BCtleri-3-26014630d7ce

[15] Medium. (t.y.). *Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall) ya da F1-Score*. <https://medium.com/@gulcanogundur/do%C4%9Fruluk-accuracy-kesinlik-precision-duyarl%C4%B1%C4%B1k-recall-ya-da-f1-score-300c925feb38>

[16] Kaggle. (t.y.). *Telco Customer Churn Dataset*. <https://www.kaggle.com/>

[17] Hugging Face. (t.y.). *Customer Churn Data for Machine Learning*. <https://huggingface.co/>

[18] Utku, A., & Akçayol, M. A. (2024). Neural Network Based a Comparative Analysis for Customer Churn Prediction. *MAUN Journal of Science*, 12(1), 39–50. <https://dergipark.org.tr/en/pub/msufbd/issue/84716/1466246>

8.EKLER