

# Makine Öğrenimi Teknikleri ile Kredi Risk Tahmininde Yeniden Örnekleme Yöntemlerinin Karşılaştırılması

*Comparison of Machine Learning Techniques and Different Resampling Methods in Credit Risk Estimation*

**Gülçin KENDİRKİRAN**

Arş. Gör., Ankara Hacı Bayram Veli Üniversitesi  
gulcin.kendirikiran@gmail.com

<https://orcid.org/0000-0003-3146-0192>

**Seyyide DOĞAN**

Dr. Öğr. Üyesi, Karamanoğlu Mehmetbey Üniversitesi  
dogans@kmu.edu.tr

<https://orcid.org/0000-0001-7835-7905>

Makale Başvuru Tarihi: 13.11.2024

Makale Kabul Tarihi: 20.12.2024

Makale Türü: Araştırma Makalesi

## Anahtar Kelimeler:

Makine Öğrenmesi,  
Dengesiz Veri,  
Kredi Puanlama

## ÖZET

Makine öğrenmesi performansını etkileyen önemli hususların başında sınıf dengesizliği sorunları gelmektedir. Gerçek dünya problemlerinde sıklıkla karşılaşılabilen bu durumda öğrenme sürecinde azınlık sınıfın etkisi ihmal edilerek çoğunluk sınıfına doğru kayan yanlı bir tahmin elde edilir. Bu çalışma UCI (University of California, Irvine) veri kaynağı içerisinde yer alan popüler kredi puanlama (Australian ve German) veri seti üzerinde sınıf dengesizliği sorunuyla başa çıkmak için 4 farklı yeniden örneklem yönteminin değerlendirilmesi sunulmaktadır. Banka müşterilerinin kredibilitesinin tahmin edildiği bu problemde iyi ve kötü olarak sınıflandırılan ve dengesiz dağılan iki müşteri sınıfı vardır. Riskli müşterilerin tahmininde Destek Vektör Makineleri (SVM), Rassal Ormanlar (RF), Extra Artırma (XGBoost), CatBoost olmak üzere farklı makine öğrenmesi teknikleri kullanılmış ve bu algoritmalar sınıf dengesizliği sorununu gidermek için Random Oversampling (ROS), Random Undersampling (RUS), SMOTE ve Tomek Bağlantıları yeniden örneklem yaklaşımları ile birleştirilmiştir. Deneysel sonuçlara göre, yeniden örneklem yöntemlerinin yöntem performansını iyileştirmede etkili olduğu ve SMOTE yaklaşımı ile CatBoost sınıflandırıcısının; ROS yaklaşımı ile RF sınıflandırıcısının daha iyi sonuçlar ürettiği gözlenmiştir.

## Keywords:

Machine Learning,  
Imbalance Data,  
Credit Scoring

## ABSTRACT

One of the most important issues affecting machine learning performance is class imbalance problems. In this case, which is frequently encountered in real-world problems, the effect of the minority class is ignored in the learning process and a biased estimation is obtained that shifts towards the majority class. This study presents the evaluation of 4 different resampling methods to cope with the class imbalance problem on the popular credit scoring dataset (Australian and German) in the UCI (University of California, Irvine) data repository. In this problem, where the credibility of bank customers is estimated, there are two customer classes classified as good and bad and are imbalanced. Different machine learning techniques such as Support Vector Machines (SVM), Random Forests (RF), Extra Boosting (XGBoost), CatBoost have been used in the prediction of risky customers and these algorithms have been combined with resampling approaches such as Random Oversampling (ROS), Random Undersampling (RUS), SMOTE and Tomek Linkages to eliminate the class imbalance problem. According to the experimental results, resampling methods are effective in improving the method performance and the SMOTE approach and the CatBoost classifier; It has been observed that ROS approach produces better results than RF classifier.

## 1. GİRİŞ

Veri madenciliği ve makine öğrenmesi alanının ilgilendiği gerçek dünya problemlerinde sıkça karşılaşılan ek bir sorun veri setlerinin dengesizliğidir. Sınıf dengesizliği bir veri setinde bulunan sınıfların birinin (ya da birkaçının) çoğunluk ve azınlık olmak üzere eşit olmayan bir şekilde dağıldığı durumu ifade etmektedir. Bu sorun öğrenme süreci sırasının daha fazla eğitim örneği içeren sınıfa verilmesi anlamına gelir ve tahminin çoğunluk sınıfa doğru yanlı tahmini ile sonuçlanır (Fernández vd., 2018; Saini & Susan, 2023). Kullanılan algoritma öncelikli olarak çoğunluk sınıfın hatalarını minimize etmeye uğraştığından yeterince temsil edilmeyen sınıfın tahmininde düşük performans göstermektedir (Krawczyk, 2016).

Sınıf dengesizliği, dolandırıcılık tahmini, kredi risk tahmini, kanser teşhisi, tıbbi görüntü analizi, itiraz tespiti, saldırı tespiti, görüntü tanıma gibi uygulamaların performansını etkileyen önemli bir sorundur (Johnson & Khoshgoftaar, 2019; Thabtah vd., 2020). Günümüzde, büyük kredi portföylerinin yönetimine olan yüksek talep ve düzenleyici reçeteler nedeniyle, kredi risk tahmin modelleri bankaların rutin olarak kullandıkları modeller haline gelmiş ve doğru tahminlerin değeri giderek daha fazla önem kazanmıştır. Kredi risk modelleri, gelir ve kredi ödeme geçmişi gibi özelliklere ve daha önce kabul edilen ve reddedilen başvuru sahiplerine ilişkin verilere göre oluşturulmaktadır (Chen & Huang, 2003).

Kredi risk tahmininde veri setleri genellikle çok sayıda geri ödemesiz müşteri yerine, geri ödeyen müşteriye içermektedir. Son yıllarda, kredi talebinde bulunan müşteriler arasından en uygun profilin belirlenmesi ve bankaların risk yönetimini güçlendirmesi amacıyla makine öğrenimi teknikleri kullanımı yaygınlaşmaktadır. Bu teknikler, bankaların kredi risklerini azaltarak kârlılıklarını artırmalarına önemli katkı sağlamaktadır. Ancak, kredi riskini modellemede karşılaşılan bu sınıf dengesizliği sorunu model performansını olumsuz etkilemektedir. Dengesiz veri etkilerini azaltmak için fazla temsil edilen gözlemleri azaltmaya ya da az temsil edilen gözlemleri çoğaltmaya yönelik stratejiler ile yetersiz örnekleme veya aşırı örnekleme yöntemleri kullanılabilir. Diğer etkili yaklaşımlar: Sentetik Azınlık Aşırı Örnekleme Tekniği (SMOTE, Chawla vd., 2002), Tomek Bağlantıları ile temizlenmiş SMOTE-Tomek (Tomek, 1976)'tir.

Bu çalışmada, dengesiz olarak dağılmış “iyi” ve “kötü” olarak nitelenen kredi müşterilerinin tahmininde hangi sınıf dengeleme yönteminin makine öğrenimi tekniklerinin tahmin performansı üzerinde daha iyi bir etkiye sahip olduğu araştırılmıştır. Bu amaçla, kredi geri ödemesine yönelik riskin tespiti amacıyla tek sınıflandırıcı ve topluluk makine öğrenme algoritmalarının tahmin performansını artırmaya yönelik sınıf dengeleme çabaları için Rassal Aşırı Örnekleme (Random OverSampling, ROS), Rassal Alt Örnekleme (Random UnderSampling, RUS), Sentetik Azınlık Sınıfta Örneklem Artırma Yöntemi (Synthetic Minority Over-sampling Technique, SMOTE) ve Tomek Bağlantıları tercih edilmiştir. Bilgimiz dahilinde ilgilendiğimiz ana konu itibarı ile dengesiz sınıflara sahip olan kredi skorlamasında CatBoost algoritmasının performansı araştırılmamıştır. Özellikle kategorik verileri işleyebilme kabiliyeti ile ön plana çıkan bu yeni topluluk makine öğrenmesi algoritmasının performansının çok sayıda kategorik veriye sahip olan German ve Australian benchmark veri setleri üzerindeki tahmin performansının araştırılması ve hangi sınıf dengeleme yöntemi ile daha iyi tahmin doğruluğu elde edeceğinin ortaya konulması alan yazınımıza katkı sağlayacaktır.

Çalışmanın takip eden bölümlerinde, dengesiz veri setleri üzerinde yeniden örnekleme yöntemlerini kullanan geçmiş çalışmalara ait bulgular; makine öğrenimi teknikleri ve yeniden örnekleme yöntemlerinin teorik alt yapısı, uygulama sonuçları ve değerlendirmeler sunulmuş ve çalışma, araştırma bulgularının özetlendiği bir bölümle sonlandırılmıştır.

## 2. KAVRAMSAL ÇERÇEVE

Makine öğrenimi tekniklerinin kullanıldığı çalışmalarda sınıflandırıcıların en iyi performanslarına ulaşmak için ana unsur verilerin kalitesi olmaktadır. Verilerin iyi bir kaliteye sahip olması, veri dengesizliklerinin olmaması ve gürültülü verilerin işlenebilmesine bağlı olmaktadır. Literatürde bu sorunlara eğilen ve çözüm üreten farklı yeniden örnekleme yöntemlerine rastlamak mümkündür (Han vd., 2019; Zhang, Yang ve Zhang, 2021; Oreski, 2023).

Sınıflandırıcılar için yeniden örnekleme yöntemleri birçok farklı alanda kullanılmaktadır. Örneğin, Mooijman vd. (2023), Hollanda sığır endüstrisinde yasadışı hormon kullanım vakalarını doğru şekilde belirleyebilen bir sınıflandırma modeli üzerinde çalışmışlardır. Çalışmalarında veri ön işleme süreçlerinde niteliklerin standardize edildiği durumla beraber logaritmik dönüşümleri de ele almışlardır. Sınıflandırmada veri seti dengelenmediği için seçilen sınıflandırıcı üzerinde aşırı ve yetersiz örnekleme yöntemleri değerlendirilmiştir. Elde ettikleri sonuçlara göre lojistik regresyon yüksek performansa sahip olurken SMOTE, ADASYN ve Borderline-SMOTE

yöntemleri model performansını yükseltmiştir. Glučina vd. (2023), Serviks kanseri olarak adlandırılan ve kadınlarda görülen üçüncü en yaygın kanser türünün incelenmesi için çalışmalarında model tamin sonuçlarını değerlendirmişlerdir. Sınıf dengeleme yöntemleri kullanılmadan düşük sınıflandırma performanslarına ulaşmaktadırlar. Ancak özellikle çok katmanlı algılayıcı ve k-en yakın komşu sınıflandırıcılarında SMOTE-ENN ve SMOTE-Tomek yöntemleri kullanıldığında model performanslarının arttığı sonucunu gözlemlemişlerdir.

Son yıllarda kredi puanlama modellerinde makine öğreniminin kullanımı oldukça ilgi görmektedir. Finans ve makine öğrenimi alanının bu ortak problemi üzerinde artan sayıda çalışmanın kapsamlı incelemesi Leo vd. (2019) tarafından sunulmuştur. Chen vd. (2024) ise hem regresyon hem sınıflama hem de derin öğrenmede sınıf dengesizlik sorununa sahip veri setlerinden öğrenme alanında yönetim bilimlerinden mühendisliğe kadar son yıllarda kaydedilen gelişmeleri özetlemekte ve ortaya çıkan yeni zorlukları kapsamlı olarak sunmaktadırlar.

Çalışma kapsamında değerlendirilen veri setlerinden birisi olan German veri seti daha önce Hussin Adam Khatir ve Bee (2022) tarafından değerlendirilmiştir. Çalışmalarında yapay sinir ağları, saf Bayes, karar ağaçları, RF, k-en yakın komşu ve düzenlenmiş lojistik regresyon sınıflandırıcılarını değerlendirmişlerdir. SMOTE, SMOTE-Tomek ve rassal aşırı örnekleme ise kullanılan yeniden örnekleme yöntemleridir. Elde ettikleri sonuca göre RF sınıflandırıcısının rassal aşırı örnekleme ile birleştirilmesinde en iyi sınıflandırıcı olduğu sonucuna ulaşmışlardır. Rofik vd. (2024) ise aynı veri setinde RF, SVM, ekstra ağaç ve XGboost'tan oluşan sınıflandırıcıları kullanmaktadırlar. Daha sonra SMOTE kullanarak dengesiz sınıfları ele almaktadırlar. RF, SVM ve ekstra ağaç topluluk öğrenme sınıflandırıcılarının modelde iyi bir performans ürettiğini ve performansı iyileştirdiklerini belirtmektedirler.

Bir başka çalışma ise Milli vd. (2024) tarafından sunulmaktadır. Çalışmalarında, German ve Australian veri setlerinde özellikle ağaç tabanlı sınıflandırıcılarda SMOTE-Tomek ve SMOTE-ENN yöntemlerinin yüksek performans özelliklerine sahip olduğu görülmektedir. Anis ve Ali (2017), German ve Australian veri setinde SMOTE yöntemi kullanıldığında RF sınıflandırıcısının en iyi performansa sahip olduğu sonucuna ulaşmışlardır. Benzer sonuca German veri seti üzerinde Dong, Liu ve Tham (2024) da ulaşmıştır. Esenogho vd. (2022) German veri setine SMOTE-ENN yöntemi uygulandığında uzun kısa süreli bellek (long short-term memory) sinir ağı ile en yüksek performansa ulaşıldığını belirtmişlerdir.

Zhou vd. (2021) German ve Australian veri setlerine SMOTE yöntemi uygulandığında SVM sınıflandırıcısının, Louis vd. (2024) ise XGBoost sınıflandırıcısının en yüksek performansa sahip olduğu sonucuna ulaşmışlardır. Aynı veri setine SMOTE yönteminin uygulandığı bir başka çalışma Kumari ve Mishra (2019) tarafından yapılmıştır. Çalışmalarında German veri seti için RF sınıflandırıcısının ve Australian veri seti için regresyon sınıflandırıcısının en yüksek performansa sahip olduğunu gözlemlemişlerdir.

Alam vd. (2020) German veri setinde, gradyan arttırmalı karar ağaçları (Gradient Boosted Decision Tree) sınıflandırıcısının performansını diğer sınıflandırıcılar ile karşılaştırmışlardır. Performans değerlendirmesinde çeşitli aşırı ve yetersiz örnekleme yöntemlerini kullanmışlardır. Gradyan arttırmalı karar ağaçları sınıflandırıcısı hem aşırı hem de yetersiz örnekleme yöntemlerinde iyi performans göstermiştir. Yetersiz örnekleme yöntemleri arasında sadece RF sınıflandırıcısında küme merkezi yeniden örnekleme yöntemi yüksek performans göstermiştir. Aşırı örnekleme yöntemlerinde de gradyan arttırmalı karar ağaçları sınıflandırıcısı en yüksek performansa sahipken sadece RF sınıflandırıcısında ADASYN yeniden örnekleme yöntemi yüksek performans göstermiştir. Diğer taraftan, Aruleba ve Sun (2024), SMOTE-ENN yöntemi kullanılarak German veri setinde XGBoost, Australian veri setinde ise RF sınıflandırıcısının en iyi performansa sahip olduğunu belirtmişlerdir. Seera vd. (2024) hem German hem de Australian veri setlerinde yetersiz örnekleme yöntemlerini kullanmışlar ve German veri seti için doğrusal regresyon sınıflandırıcısının, Australian veri seti için gradyan arttırmalı karar ağaçları sınıflandırıcısının performanslarının en yüksek olduğu sonucuna ulaşmışlardır.

Aly vd. (2022), Australian veri setine SMOTE-Tomek yönteminin Torbalama (Bagging) sınıflandırıcısı ile German veri setinin ise aşırı örnekleme yönteminde XGBoost sınıflandırıcısının en iyi performansa sahip olduğu sonucuna ulaşmışlardır. Xiao vd. (2021) German veri setine SMOTE yöntemi uygulandığında RF, Australian veri setine ise ROS yöntemi uygulandığında Torbalama sınıflandırıcısının en yüksek model performansına sahip olduğu sonucuna ulaşmışlardır. Bu sonuçlar gösteriyor ki literatür incelendiğinde aşırı örnekleme yöntemleri kulanıldığında ağaç tabanlı, özellikle RF sınıflandırıcısı yüksek performansa sahip olmaktadır.

Bu çalışmada literatürdeki diğer çalışmalar temel alınarak hem aşırı hem de yetersiz örnekleme yöntemleri ile çeşitli sınıflandırıcıların performansları hiperparametre optimizasyonu ile değerlendirilmektedir.

### 3. YÖNTEM

#### 3.1. Makine Öğrenmesi Yöntemleri

Çalışmada, makine öğreniminde kredi risk tahminlerinin Destek Vektör Makinaları (Support Vector Machine, SVM), Rassal Ormanlar (Random Forest, RF), XgBoost (eXtreme Gradian Boosting) ve CatBoost (Categorical Boosting) sınıflandırıcıları ile karşılaştırmalı sonuçları değerlendirilmektedir.

##### 3.1.1. Destek Vektör Makinleri (Support Vector Machine – SVM)

Yüksek genelleme yetenekleri ile bilinen SVM'ler karmaşık ve yüksek boyutlu gerçek dünya problemlerinde regresyon ve sınıflandırma görevlerini yerine getirmek için sıklıkla tercih edilmektedir (Moguerza & Muñoz, 2006). Klasik algoritmaların deneysel riski azaltma amacının aksine, yapısal riski minimize etmeyi hedefleyen, istatistiksel öğrenme teorisine dayalı bir makine öğrenmesi yöntemidir. Örüntü tanıma işlemlerinde maksimum marjı belirlemek için hiper düzlemi kullanılmakta ve problemi ikinci dereceden programlama olarak ele almaktadır (Kecman, 2005). SVM, girdi verilerinin doğrusal olmayan örüntülerini polinom, sigmoid ve radyal, simetrik işlevlere sahip uygun bir çekirdek fonksiyonu yardımıyla veriyi yüksek boyutlu bir uzaya taşıyarak eşleme işlevleriyle öğrenebilmektedir.

##### 3.1.2. Rassal Ormanlar Algoritması (Random Forest- RF)

Rastgele ormanlar, iyi sınıflandırma performansı, ölçeklenebilirliği ve kullanım kolaylığı nedeniyle son on yılda makine öğrenimi uygulamalarında büyük popülerlik kazanmıştır (Sebastian Raschka (2015). Leo Breiman (2001) tarafından önerilen RF, sıklıkla aşırı uyum hatası ile sonuçlanan karar ağaçları gibi bir dizi zayıf öğreniciyi güçlü bir öğreniciye dönüştüren topluluk öğrenme algoritmasıdır. Karar ağaçlarının her birinde rastgeleliği artırmak için yeniden örnekleme yöntemiyle (bootstrap) alt kümeler oluşturulur. Dahası RF, her karar ağacında en iyi ayrımı yapmak için tüm özelliklerin yerine onların bir alt kümesi olan özellikleri rastgele seçerek işleme dahil eder ve birbiri ile daha az korelasyonsuz tahminler elde edilmesini sağlar. Bu aşamada belirli sayıda ağaç eğitildikten sonra nihai tahmin, sınıflama problemi için tüm ağaçlarda tahminlerin çoğunluk oyları toplanarak; regresyon problemi için ortalaması alınarak oluşturulur. Bu prosedür (bootstrap agregation ya da bagging) ile RF genelleme yapma yeteneği daha yüksek ve aşırı öğrenme hatasına daha az duyarlı güçlü bir öğrenici haline gelmektedir.

##### 3.1.3. Aşırı Gradyan Artırma (eXtreme Gradian Boosting-XGBoost)

Büyük veri setlerinde hızlı ve ölçeklenebilir bir model olarak öne çıkan XGBoost, oldukça etkili ve yaygın kullanılan bir makine öğrenmesi algoritmasıdır. XGBoost'un daha yüksek bir tahmin performansını sağlamak için geliştirdiği temel fikir karar ağaçları ve doğrusal modele dayalı bir gradyan artırma topluluğu algoritmasını kullanmasıdır. Gradyan artırma karar ağacı algoritmaları her ağacın  $n - 1$  sayıdaki ağaçtan öğrenerek aşağı doğru gradyan yaptığı birden fazla karar ağacından oluşur. Nihai tahmin tüm ağacın ürettiği yanıtların toplanması ile bulunur. Chen ve Guestrin (2016) tarafından bu konsept geliştirilerek önerilen XGBoost her tahmin modelinin hatası bir sonraki modele eklenerek tahmin doğruluğu iyileştirilmektedir. Bununla birlikte ağaç karmaşıklığını düzenleyen bir terim ekleyerek aşırı öğrenmeyi önlemeye çalışmaktadır (Pall vd., 2023).

##### 3.1.4. Kategorik Artırma (Categorical Boosting-CatBoost)

Prokhorenkova vd. (2018) ve Dorogush vd. (2018) tarafından önerilen CatBoost oldukça yeni bir topluluk makine öğrenme algoritmasıdır. CatBoost, temel tahmin edici olarak ikili karar ağaçlarını kullanır ve gradyan artırma algoritmalarında bulunan hedef sızıntısı sorununun neden olduğu tahmin kayması ile mücadele eden yenilikçi bir yaklaşım önerir. Bu yeni yaklaşım, gradyan artırmada permutasyon odaklı bir alternatif olan sıralı artırmanın uygulanması ve kategorik verileri minimum bilgi kaybı ile işleyebilmesidir (Ostroumova vd., 2017). Bu işleme adımında orijinal değişken ortadan kaldırılır ve her kategori için yeni bir ikili özellik eklenir. Aynı kategori değerine sahip veri örnekleri için simetrik karar ağaçları oluşturulur. Bu sayede aşırı uyum sorununun kolaylıkla üstesinden gelebilmektedir (Prokhorenkova vd., 2018).

#### 3.2. Yeniden Örnekleme Yöntemleri

Bir veri setinde sınıf etiket sayıları birbirine eşit değilse dengesiz olmaktadır. Yeniden örnekleme yöntemleri, sınıflardaki etiket sayıları arasında aşırı bir orantısızlık olduğunda kullanılmaktadır. Veri setindeki bu dengesizliklerin dengeli hale getirilmesinden sonra çeşitli sınıflandırıcı türlerinin nihai performansı büyük ölçüde iyileştirilmektedir. Aynı zamanda, çalışma kapsamında kullanılan yeniden örnekleme yöntemleri sınıflandırıcıdan bağımsızdır, kolay uygulanabilmekte ve veri setine herhangi bir algoritmanın uyarlanmasına ihtiyaç duyulmamaktadır (Waititu vd., 2020).

### 3.2.1. Yetersiz Örneklem (UnderSampling) Algoritmaları

Rassal Alt Örneklem (Random UnderSampling, RUS), alt örneklem yöntemlerinde kullanılan en temel yöntemlerden birisidir. Yöntemde, sınıf dağılımı dengelenene kadar çoğunluk sınıfındaki örnekler kaldırılmaktadır. RUS'un kullanımında bazı zorluklar bulunmaktadır. Öncelikle çoğunluk sınıfından kaldırılan örnekler bir karar sınırı için yararlı, hatta kritik olabilmektedir. Diğer taraftan, çoğunluk sınıfındaki örnekler düzgün bir şekilde dağılmamışsa, rastgele seçilen örnekleri kaldırmak istatistiklerin temellerini değiştirebilmekte hatta sınıflandırma performansını bozabilmektedirler. Ancak sıklıkla kullanılan bir yöntemdir (Wang vd., 2023).

### 3.2.2. Aşırı Örneklem (OverSampling) Algoritmaları

Aşırı örnekleme yöntemlerinden ilki olan Rassal Aşırı Örneklem (Random OverSampling, ROS), sınıf dağılımı dengelenene kadar azınlık örneklerini çoğaltan aşırı örnekleme yöntemlerinin en basiti olarak tanımlanmaktadır. Bu yöntem, herhangi bir yeni bilgi eklemeyen azınlık verilerini aşırı örnekleme ve bu da aşırı uyumla sonuçlanabileceği için sıklıkla eleştirilmektedir. Ancak RUS'ta olduğu gibi ROS yöntemi de performans değerlendirmelerinde sıklıkla kullanılan yöntemlerden birisidir (Wang vd., 2023).

### 3.2.3. SMOTE (Synthetic Minority OverSampling Technique)

Makine öğreniminde sınıf dengesizliklerini gidermek için eğitim setinin ya da orijinal veri setinin yeniden örneklemeyle ele alındığı yöntemler bulunmaktadır. Yaygın olarak kullanılan yöntemler genelde orijinal veri setinden sağlanan bilgilerden öğrenmeye devam eden yeniden örnekleme yoluyla yapılmaktadır. Söz konusu yöntemlerden Sentetik Azınlık Sınıfında Örneklem Artırma Yöntemi (Synthetic Minority OverSampling Technique, SMOTE), çoğunluk sınıfının yetersiz örneklemeyle azınlık sınıfının aşırı örneklemeyle birleştirilmesidir. Azınlık sınıfındaki gözlemlere aşırı örnekleme yapılırken sentetik örnekler oluşturulmaktadır. Böylece gözlemler ya da gözlem değerleri değiştirilmeden çoğaltılmaktadır (Chawla vd., 2002). Sentetik örneklerin oluşturulmasında aşağıdaki adımlar izlenmektedir:

1. Rastgele bir azınlık sınıfı ve ona en yakın  $k$  azınlık sınıfı komşusu seçilir.  $k$  değeri, ihtiyaç duyulan aşırı örnekleme miktarına göre belirlenir.
2. Seçilen azınlık vektörü ile en yakın komşularından birinin vektörü arasındaki fark hesaplanır.
3. Elde edilen fark daha sonra 0 ile 1 arasında rastgele bir sayı ile çarpılır. Elde edilen sonuç, seçilen azınlık vektörüne eklenir. Bunu yaparak, iki vektörü birleştiren çizgi boyunca yeni bir rastgele nokta üretilmiş olmaktadır (Waititu vd., 2020).

### 3.2.4. Tomek Bağlantıları

Tomek Bağlantıları yetersiz örnekleme için kullanılmakla birlikte en klasik alt örnekleme yöntemlerinden biri olduğu söylenebilmektedir. Tomek Bağlantıları kısaca aşağıdaki gibi özetlenebilmektedir.

a ve b örnekleri, aşağıdaki durumlarda bir Tomek Bağlantısı tanımlar:

1. a örneğinin en yakın komşusu b ise,
2. b örneğinin en yakın komşusu a ise
3. a ve b örnekleri farklı sınıflara ait ise.

Tomek Bağlantısını bulmaya yönelik bu adımlar, örneklerin en yakın komşulara sahip olmasını sağlayabilmektedir (Wang & Liu, 2021).

Tomek bağlantılarının çiftlerindeki örnekler, sınır çizgisi veya gürültü olarak sınıflandırılır ve örtüşen bir bölgede yani en yakın komşuda bulundurulurlar. Yöntemde bir veri setindeki yetersiz örnekleme için Tomek Bağlantıları'nda yalnızca çoğunluk sınıfındaki örnekler silinirken azınlık sınıfındaki örnekler tutulmaktadır (Bunkhumpornpat ve Sinapiromsaran, 2014).

## 4. DENEYSEL ANALİZ

### 4.1. Veri ve Veri Önsüreci

Çalışma kapsamında Güney Alman Kredi (South German Credit, bundan sonra German olarak bahsedilecektir) ve Avustralya Kredi Onayı (Australian Credit Approval, bundan sonra Australian olarak bahsedilecektir) veri setleri kullanılmakta olup özellikleri Tablo 1'de özetlenmiştir. Bu veri setleri banka hesapları, avans oranları,

menkul kıymet ölçümleri, mevcut kredi sayısı gibi müşterilerin finansal durumları ve yaş, cinsiyet medeni durum gibi demografik özellikleri hakkında bilgi veren bir dizi özelliği ve kredi vermeye “uygun” , “uygun olmayan” şeklinde ifade edilen hedef değişkeninin kredi riskini içermektedir.

**Tablo 1.** German ve Australian Veri Setlerinin Açıklamaları

|            | Örnek Sayısı | Nitelik Sayısı | Sınıflar  | Veri Türü          | Kayıp Gözlem |
|------------|--------------|----------------|-----------|--------------------|--------------|
| German     | 1000         | 20             | (700/300) | Kategorik, Sürekli | Yok          |
| Australian | 690          | 14             | (383/307) | Kategorik, Sürekli | Yok          |

Her iki veri seti de ikili sınıf etiketlerine ve dengelenmemiş gözlemlere sahiptir. German veri setinde toplam 1000 gözlem olmakla birlikte 20 tane niteliğe sahiptir. Gözlemlerde 700 iyi kredi örneği ve 300 kötü kredi örneği bulunmaktadır. Australian veri setinde ise 383 iyi kredi örneği ile 307 kötü kredi örneği vardır. Toplam 690 gözleme ve 14 tane niteliğe sahiptir (Chen, vd., 2017). Veri setlerine , UCI Machine Learning Repository veri tabanından ulaşılabilmektedir.

Veri setleri dengeli hale getirildikten sonra gözlemler standardize edilmektedir. Standardizasyon, her bir nitelik dağılımının, ortalaması sıfır ve standart sapması bir olacak şekilde ölçeklenmesidir. Makine öğrenimi modelleri uygulanırken verileri standardize etmek faydalı olmaktadır. Her bir nitelik, gözlemlerden ortalaması ve standart sapması hesaplanarak aşağıda yer alan eşitlik ile ölçeklendirilmektedir (Saleh & Fleyeh, 2022).

$$x' = \frac{x - \text{ort}(x)}{\sigma(x)} \quad (1)$$

Gözlemler ölçeklendirildikten sonra veri seti eğitim ve test setine ayrılmaktadır. Sınıflandırma için veri modelini üretirken kullanılan eğitim seti %70 ve veri modelini değerlendirmek için kullanılan test seti %30 oranlarında bölünmüştür.

#### 4.2. Model Parametrelerinin Ayarlanması

Makine öğreniminde en iyi performansa ulaşmak için sınıflandırıcı içindeki tahmin edilen parametreler dışındaki diğer parametrelerin ayarlanması gerekmektedir. Buna, hiperparametre optimizasyonu denmektedir. Makine öğreniminde, model performansının değerlendirilmesinde hiperparametrelerin önemli etkisi bulunmaktadır. Sıklıkla tercih edilen optimize edicilerden birisi ızgara arama (grid search) algoritmasıdır (Shen vd., 2020).

Izgara aramasını bir alt küme incelemesi olarak değerlendirmek mümkündür. Hiperparametrelerin en iyi değerini bulmak için bu alt kümelerde eşit aralıklara sahip ızgaralar hazırlanmakta ve uygun ızgara için tüm olasılıklar değerlendirilmektedir. Izgara aramasının çalışma şekli şu şekildedir:

1. Verili aralıktaki parametrenin tüm değerleri modelde yerine koyulur,
2. Tüm parametre değerlerinin birleşimi döngüye alınır,
3. Eğitim verileri üzerinde makine öğrenimi algoritması ile eğitim gerçekleştirilir,
4. Elde edilen sonuçlar test verileriyle değerlendirilir,
5. En az hata ve en iyi parametre değerinin birleşimi depolanır (Anggoro ve Mukti, 2021).

Buna göre hem German hem de Australian veri setleri için ızgara arama ile 10-katlı çapraz doğrulama sonucunda hiperparametre optimizasyonuna ait seçilen hiperparametre değerlerine Ek 1’de yer verilmektedir. Bu değerler, en yüksek model performansına elde edilen hiperparametre değerleri ile ulaşılabileceğini ifade etmektedir.

#### 4.3. Performans Değerlendirmesi

Bir sınıflandırıcının performansı değerlendirilirken doğruluk (accuracy), kesinlik (precision), duyarlılık (sensitivity/recall), matthews korelasyon katsayısı (MCC) ve F1 puanı hesaplamaları kullanılmaktadır. Performans ölçütleri, karmaşıklık matrisi değerlerinden hesaplanmaktadır. Karmaşıklık matrisinin bir tanımı Tablo 2 ile gösterilebilmektedir.

**Tablo 2.** İkili sınıflama için Karmaşıklık Matrisi

|        |     | Gerçek              |                     |
|--------|-----|---------------------|---------------------|
|        |     | (+)                 | (-)                 |
| Tahmin | (+) | Doğru Pozitif (TP)  | Yanlış Pozitif (FP) |
|        | (-) | Yanlış Negatif (FN) | Doğru Negatif (TN)  |

Bazı durumlarda, tek bir ölçüt değerinden hareketle çıkarım yapmak doğru sonuca götürmemektedir. Yapılan çalışmaların amaçlarına göre sınıflandırma hatalarının maliyetleri değişkenlik gösterebilmektedir. Örneğin, yüksek bir kesinlik değeri az sayıda yanlış pozitif tahminini ifade ederken yüksek geri çağırma değeri daha az sayıda yanlış negatif tahminini göstermektedir. Çok sayıda yanlış pozitif tahmininde kesinlik iyi bir ölçüt olabilmektedir. Diğer taraftan çok sayıda yanlış negatif tahminde geri çağırma ölçütü tercih edilebilmektedir. Yanlış negatif ve yanlış pozitif tahminler aynı olduğunda doğruluk bir performans metriği olarak kullanılabilir. Ancak farklı olduğunda, kesinlik veya duyarlılık veya her ikisi de kullanılabilir. F1 puanı, kesinlik ve duyarlılığın ağırlıklı ortalaması olduğundan dolayı yüksek F1 puanı hem yanlış pozitif hem de yanlış negatif durumlarında daha iyi bir performansı ifade etmektedir. Performans ölçütlerinin eşitlikleri aşağıda verilmektedir (Gupta ve Goel, 2023).

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

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

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$MCC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}} \quad (5)$$

$$F1 - Score = \frac{2*TP}{2*TP+FP+FN} \quad (6)$$

#### 4.4. Deneysel Analiz Sonuçları

Çalışmanın bu bölümünde optimal hiperparametreler ile sınıflandırıcıların performansları değerlendirilmektedir. Her değerlendirme aşamasında yeniden örnekleme yöntemlerine göre ayrı bir şekilde ele alınmaktadır. Tablo 3 ve Tablo 4 sırasıyla German ve Australian veri setlerinde, dört güçlü sınıflandırma algoritması (SVM, RF, XGBoost, CatBoost) ve farklı veri dengeleme yöntemlerinin (SMOTE, ROS, RUS, TOMK) kombinasyonu ile elde edilen tahmin sonuçlarını göstermektedir. Tahmin performansı Doğruluk, Kesinlik, Recall, MCC (Matthews Correlation Coefficient) ve F1 puanı gibi farklı ölçütler ile değerlendirilmiştir.

**Tablo 3.** German Veri Seti için Performans Ölçütleri Tablosu

|       |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|-------|----------|---------------|----------|------------|--------|----------|
| DÜZEY | SVM      | 0.7800        | 0.7500   | 0.6905     | 0.4364 | 0.7067   |
|       | RF       | 0.7833        | 0.7541   | 0.6960     | 0.4464 | 0.7124   |
|       | XgBoost  | 0.8000        | 0.7813   | 0.7143     | 0.4910 | 0.7333   |
|       | CatBoost | 0.8000        | 0.7843   | 0.7111     | 0.4899 | 0.7309   |
| SMOTE |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.8405        | 0.8414   | 0.8405     | 0.6819 | 0.8404   |
|       | RF       | 0.8333        | 0.8338   | 0.8333     | 0.6672 | 0.8333   |
|       | XgBoost  | 0.8357        | 0.8357   | 0.8357     | 0.6714 | 0.8357   |
|       | CatBoost | 0.8548        | 0.8548   | 0.8548     | 0.7096 | 0.8548   |
| ROS   |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.8619        | 0.8659   | 0.8619     | 0.7278 | 0.8615   |
|       | RF       | <b>0.8905</b> | 0.8922   | 0.8905     | 0.7827 | 0.8904   |
|       | XgBoost  | 0.8524        | 0.8550   | 0.8524     | 0.7074 | 0.8521   |

|       |          |               |          |            |        |          |
|-------|----------|---------------|----------|------------|--------|----------|
|       | CatBoost | <b>0.8929</b> | 0.8933   | 0.8929     | 0.7862 | 0.8928   |
| RUS   |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.7611        | 0.7611   | 0.7611     | 0.5222 | 0.7611   |
|       | RF       | 0.7611        | 0.7651   | 0.7611     | 0.5262 | 0.7602   |
|       | XgBoost  | 0.7722        | 0.7764   | 0.7722     | 0.5486 | 0.7714   |
|       | CatBoost | 0.7389        | 0.7389   | 0.7389     | 0.4778 | 0.7389   |
| TOMEK |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.7753        | 0.7509   | 0.7322     | 0.4828 | 0.7395   |
|       | RF       | 0.7640        | 0.7377   | 0.7183     | 0.4555 | 0.7255   |
|       | XgBoost  | 0.7828        | 0.7637   | 0.7324     | 0.4951 | 0.7430   |
|       | CatBoost | 0.7978        | 0.7921   | 0.7382     | 0.5276 | 0.7536   |

German veri setinde elde edilen sonuçlara göre dengelenmemiş versiyonu temsil eden Düzey’de en yüksek tahmin doğruluk oranı XgBoost ve CatBoost algoritması ile elde edilmiştir. Özellikle bu sınıflandırıcıların, tüm sınıflar için doğru tahminlerin daha iyi bir ölçütünü ortaya koyan F1 puanı ve MCC performans değerleri göz önüne alındığında da iyi birer tahminci olduğu gözlenmektedir. SMOTE yönteminin uygulanarak sınıf dengesizliğinin giderilmesi ile tüm sınıflandırıcılar için tahminlerin kayda değer ölçüde iyileştiği görülmektedir. SMOTE yönteminin genelleme kabiliyetine en çok katkıda bulunduğu yöntemin CatBoost olduğu söylenebilir. Bu sayede tahmin doğruluk oranı %80’den %85’e yükselmiştir. Aşırı örnekleme yöntemlerinden bir diğeri olan ROS yöntem ise sınıflandırıcıların tahmin performansına en fazla katkıyı sağlayan yöntem olmuştur. ROS yöntemi ile RF ve CatBoost sınıflandırıcının beraber uygulanması en yüksek doğruluk oranının (%89.05 ve %89.29) elde edilmesini sağlamıştır. Sınıfları dengelemek için tercih edilen RUS ve TOMEK yönteminin en iyi sonuç sergilediği yöntemler sırasıyla XgBoost ve CatBoost algoritmaları olmasına karşılık, şaşırtıcı bir sonuç bu yöntemlerin dengesiz veri seti üzerindeki tahminlerden daha kötü bir performansa neden olmasıdır.

**Tablo 4.** Australian Veri Seti için Performans Ölçütleri Tablosu

|       |          |               |          |            |        |          |
|-------|----------|---------------|----------|------------|--------|----------|
|       |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
| DÜZEY | SVM      | 0.8792        | 0.8780   | 0.8772     | 0.7552 | 0.8776   |
|       | RF       | 0.9179        | 0.9173   | 0.9163     | 0.8336 | 0.9168   |
|       | XgBoost  | 0.8841        | 0.8826   | 0.8870     | 0.7696 | 0.8834   |
|       | CatBoost | 0.8841        | 0.8844   | 0.8804     | 0.7648 | 0.8821   |
| SMOTE |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.8870        | 0.8899   | 0.8870     | 0.7769 | 0.8867   |
|       | RF       | 0.8783        | 0.8787   | 0.8783     | 0.7570 | 0.8782   |
|       | XgBoost  | 0.8826        | 0.8840   | 0.8826     | 0.7666 | 0.8825   |
|       | CatBoost | 0.9130        | 0.9130   | 0.9130     | 0.8261 | 0.9130   |
| ROS   |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.8913        | 0.8916   | 0.8913     | 0.7829 | 0.8913   |
|       | RF       | <b>0.9261</b> | 0.9261   | 0.9261     | 0.8522 | 0.9261   |
|       | XgBoost  | 0.8826        | 0.8833   | 0.8826     | 0.7659 | 0.8826   |
|       | CatBoost | 0.9130        | 0.9132   | 0.9130     | 0.8262 | 0.9130   |
| TOMEK |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.9081        | 0.9085   | 0.9082     | 0.8167 | 0.9081   |
|       | RF       | 0.9189        | 0.9189   | 0.9189     | 0.8379 | 0.9189   |
|       | XgBoost  | 0.8919        | 0.8921   | 0.8918     | 0.7839 | 0.8919   |
|       | CatBoost | 0.8919        | 0.8967   | 0.8916     | 0.7883 | 0.8915   |
| TOMEK |          | Doğruluk      | Kesinlik | Duyarlılık | MCC    | F1 puanı |
|       | SVM      | 0.9167        | 0.9166   | 0.9166     | 0.8332 | 0.9166   |



|  |          |        |        |        |        |        |
|--|----------|--------|--------|--------|--------|--------|
|  | RF       | 0.8958 | 0.9032 | 0.8938 | 0.7970 | 0.8949 |
|  | XgBoost  | 0.9063 | 0.9062 | 0.9062 | 0.8123 | 0.9062 |
|  | CatBoost | 0.9010 | 0.9011 | 0.9008 | 0.8019 | 0.9009 |

Sınıf dengesizliğinin daha az olduğu Australian veri setinde sınıf dengeleme yöntemlerinin önemli katkı sunduğu gözlenmektedir. German veri seti ile uyumlu olarak en yüksek tahmin doğruluk oranını elde edildiği sınıflandırıcı ve yeniden örnekleme yöntemi kombinasyonu ROS-RF olmuştur. SMOTE yönteminin uygulandığı sınıflandırıcılar göz önünü alındığında en yüksek katkının CatBoost sınıflandırıcı ile gerçekleştiği ve tahmin doğruluğunun %88'ten %91'e yükseldiği görülmektedir. ROS yönteminde de tahmin performansının German veri setinde olduğu gibi RF ve CatBoost sınıflandırıcıları tarafından elde edildiği sonucuna ulaşılmaktadır. Öyle ki tahmin performansına RF ve CatBoost sınıflandırıcıları için sırasıyla %92 ve %91 katkı sağlamaktadır. Yetersiz örnekleme yöntemleri değerlendirildiğinde ise RUS yönteminde RF ve TOMEK yönteminde ise SVM sınıflandırıcılarının %91 oranında model tahminine katkı sağladığı anlaşılmaktadır. Elde edilen sonuçlara göre Australian veri setinde çalışmada kullanılan sınıflandırıcıların yeniden örnekleme yöntemleri ile yüksek model tahmin doğruluklarına sahip oldukları sonucuna ulaşılmaktadır.

Genel olarak sınıf dengesizliğini gidermek için tercih edilen yöntemlerin model performansı üzerinde olumlu bir etkisi vardır. Bu yöntemlerden özellikle az sayıda sınıfın temsil edildiği sınıfların sayısını artırmaya odaklanan SMOTE ve ROS yöntemlerinin her iki veri setinde de sonuçların iyileştirilmesine katkı sağladığı, CatBoost ve RF sınıflandırıcılarını desteklediği görülmektedir.

## 5. SONUÇ

Sınıf dengeleme sorunlarının araştırıldığı birçok çalışma, kredi risk ölçümü temel araştırmalarından biri haline gelmektedir. Bu çalışmada, iyi ve kötü olarak isimlendirilen kredi müşterilerini en verimli şekilde ayırt edebilecek makine öğrenimi sınıflandırıcıları kullanılmaktadır. Dengesiz veri yapısına sahip olan bu iki veri seti için dört farklı yeniden örnekleme yöntemi uygulanmakta ve sınıflandırma modellerindeki performans iyileşmeleri değerlendirilmektedir. Çalışmamızda, her iki veri setinde de RF sınıflandırıcısının rassal aşırı örnekleme (ROS) ile birleştirildiğinde en iyi sınıflandırıcı olduğu, ayrıca veri dengeleme yöntemlerinin özellikle hassasiyeti ve MCC metriğini ciddi ölçüde artırmada faydalı olduğu gözlemlenmektedir. Onu takip eden SMOTE yönteminin en çok CatBoost sınıflandırıcısının tahmin performansına katkı sağladığı gözlenmektedir.

Sınıflandırıcı performansı açısından benzer sonuçlar Hussin Adam Khatir ve Bee (2022) ile Trivedi (2020) tarafından da elde edilmiştir. Bu çalışmaların sonuçları da benzer şekilde ROS yönteminin en çok RF sınıflandırıcıyı desteklediği göstermektedir. Öte yandan CatBoost sınıflandırıcıya ilişkin bir araştırmaya literatürde rastlanmamıştır. Bu çalışmanın sonuçlarına göre her iki veri setinde de CatBoost sınıflandırıcısının genelleme kabiliyetine en çok katkı yapan yöntemin SMOTE olduğu görülmektedir. Yapılacak bir diğer kıyaslama, sınıf etiketi oranları arasındaki farkın German veri setinde, Australian veri setine göre daha düşük olmasıdır. Literatürle benzer şekilde elde edilen bulgular gösteriyor ki aşırı dengesizlik durumunda RUS ve TOMEK daha kötü sonuçlar elde edilmesine neden olmaktadır.

Gelecekteki araştırmalarda, farklı kredi kategorileri üzerinde modellerin test edilmesi ve kayıp fonksiyonları temelli dengeleme stratejilerinin analiz kapsamına alınması planlanmaktadır. Son olarak, sınıf dengesizliğinin daha yüksek olduğu veri setlerinde genellenmiş sonuçlara ihtiyaç olduğundan daha geniş sayıda dengesiz veri setleri üzerinde araştırmaların yapılması ilgili literatüre katkı sağlayacaktır.

## KAYNAKÇA

- Alam, T. M., Shaukat, K., Hameed, I. A., Luo, S., Sarwar, M. U., Shabbir, S., ... and Khushi, M. (2020). An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access*, 8, 201173–201198.
- Aly, S., Alfonse, M., Roushdy, M. I., & Salem, A. B. M. (2022). Developing an intelligent system for predicting bankruptcy. *Journal of Theoretical and Applied Information Technology*, 100(7), 2068–2088.
- Anis, M., & Ali, M. (2017). Investigating the performance of smote for class imbalanced learning: a case study of credit scoring datasets. *Eur. Sci. J*, 13(33), 340–353.
- Anggoro, D.A. & Mukti, S.S. (2021). Performance comparison of grid search and random search methods for hyperparameter tuning in extreme gradient boosting algorithm to predict chronic kidney failure. *International Journal of Intelligent Engineering and Systems*, 14(6), 198–207.
- Aruleba, I., & Sun, Y. (2024). Effective credit risk prediction using ensemble classifiers with model explanation. *IEEE Access*, 12, 115015–115025.
- Breiman, L. (2001). Random forests. *Machine learning*. 45(1), 5–32.
- Bunkhumpornpat, C. & Sinapiromsaran, K. (2014). Safe level graph for majority under-sampling techniques. *Chiang Mai Journal of Science*, 41(5.2), 1419–1428.
- Chawla, N.V., Bowyer, K.W., Hall, L.O. & Kegelmeyer, W.P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Chen, M.-C., & Shih-Hsien Huang. (2003). Credit scoring and rejected instances reassigning through evolutionary computation techniques. *Expert Systems with Applications*, 24, 433–41.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. *In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).
- Chen, J., Xie, L., Liu, D.H. & Xiao, J. (2017). Effect analysis of resampling techniques on the performance of customer credit scoring models. *DEStech Transactions on Computer Science and Engineering*, 12, 375–380.
- Dong, H., Liu, R., & Tham, A. W. (2024). Accuracy comparison between five machine learning algorithms for financial risk evaluation. *Journal of Risk and Financial Management*, 17(2), 50.
- Dorogush, A. V., Ershov, V. & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.
- Esenogho, E., Mienye, I. D., Swart, T. G., Aruleba, K., & Obaido, G. (2022). A neural network ensemble with feature engineering for improved credit card fraud detection. *IEEE Access*, 10, 16400–16407.
- Fernández A., García S., Galar M., Prati R.C., Krawczyk B. & Herrera F. (2018). *Learning from imbalanced data sets*. Springer, Cham
- Glučina, M., Lorencin, A., Anđelić, N., & Lorencin, I. (2023). Cervical cancer diagnostics using machine learning algorithms and class balancing techniques. *Applied Sciences*, 13(2), 1061.
- Gupta, S.C. & Goel, N. (2023). Predictive modeling and analytics for diabetes using hyperparameter tuned machine learning techniques. *Procedia Computer Science*, 218, 1257–1269.
- Han, X., Cui, R., Lan, Y., Kang, Y., Deng, J., & Jia, N. (2019). A Gaussian mixture model based combined resampling algorithm for classification of imbalanced credit data sets. *International Journal of Machine Learning and Cybernetics*, 10, 3687–3699.
- Hussin Adam Khatir, A. A., & Bee, M. (2022). Machine learning models and data-balancing techniques for credit scoring: What is the best combination?. *Risks*, 10(9), 169.
- Johnson, J., & Khoshgoftaar, T. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6, 1-54. <https://doi.org/10.1186/s40537-019-0192-5>.
- Kecman, V. (2005). Support vector machines: theory and applications. In L.Wang (Eds). *Support Vector Machines-An Introduction* (pp. 1–47). Springer, Berlin Heidelberg. <https://doi.org/10.1007/b95439>.

- Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress Artif Intell*, 5(4), 221–232.
- Kubat, M. & Matwin, S. (1997). Addressing the curse of imbalanced training sets: one-sided selection. *In Proceedings of the Fourteenth International Conference on Machine Learning*, 97(1), 179–186.
- Kumari, P., & Mishra, S. P. (2019). Analysis of credit card fraud detection using fusion classifiers. *In Computational Intelligence in Data Mining: Proceedings of the International Conference on CIDM 2017*, (pp. 111–122). Springer Singapore.
- Louis, L. D., Dunton, A., Saklecha, S. R., Sivakumar, S. N. K., Ahmed, A. S., Sheth, S., & Chang, S. Y. (2024). Mitigating risk in P2P lending network: enhancing predictions with GenAI and SMOTE. *Journal of Networking and Network Applications*, 4(2), 48–59.
- Milli, M. E. F., Aras, S., & Kocakoç, İ. D. (2024). Investigating the effect of class balancing methods on the performance of machine learning techniques: credit risk application. *İzmir Yönetim Dergisi*, 5(1), 55–70.
- Moguerza, J., & Muñoz, A. (2006). Support Vector Machines with Applications. *Statistical Science*, 21, 322–336. <https://doi.org/10.1214/088342306000000493>. Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Mooijman, P., Catal, C., Tekinerdogan, B., Lommen, A., & Blokland, M. (2023). The effects of data balancing approaches: A case study. *Applied Soft Computing*, 132, 109853.
- Oreski, G. (2023). Synthesizing credit data using autoencoders and generative adversarial networks. *Knowledge-Based Systems*, 274, 110646.
- Pall, R., Gauthier, Y., Auer S. & Mowaswes, W. (2023). Predicting drug shortages using pharmacy data and machine learning. *Health Care Manag. Sci.*, 26(3), 395–411. <https://doi.org/10.1007/s10729-022-09627-y>.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.
- Rofik, R., Aulia, R., Musaadah, K., Ardyani, S. S. F., & Hakim, A. A. (2024). The optimization of credit scoring model using stacking ensemble learning and oversampling techniques. *Journal of Information System Exploration and Research*, 2(1).
- Saini, M., & Susan, S. (2023). Tackling class imbalance in computer vision: a contemporary review. *Artificial Intelligence Review*, 56(Suppl 1), 1279–1335.
- Saleh, R., & Fleyeh, H. (2022). Using supervised machine learning to predict the status of road signs. *Transportation research procedia*, 62, 221–228.
- Seera, M., Lim, C. P., Kumar, A., Dhamotharan, L., & Tan, K. H. (2024). An intelligent payment card fraud detection system. *Annals of Operations Research*, 334(1), 445–467.
- Shen, L., Liu, W., Chen, X., Gu, Q. & Liu, X. (2020). Improving machine learning-based code smell detection via hyper-parameter optimization. *27th Asia-Pacific Software Engineering Conference (APSEC)*, 276–285.
- Thabtah, F., Hammoud, S., Kamalov, F., & Gonsalves, A. (2020). Data imbalance in classification: Experimental evaluation. *Inf. Sci.*, 513, 429–441. <https://doi.org/10.1016/j.ins.2019.11.004>.
- Tomek, I. (1976). Two modifications of CNN. *IEEE Transactions on Systems, Man, and Cybernetics*, 11, 769–772.
- Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, 63, 101413.
- Waititu, H.W., Arap Koskei, J.K. & Onyango, N.O. (2020). Determinants of under five child mortality from kdhs data: a balanced random survival forests (BRSF) technique. *International Journal of Statistics and Applications*, 10(5), 118–130.
- Wang, A.X., Chukova, S. S. & Nguyen, B.P. (2023). Synthetic minority oversampling using edited displacement-based  $k$ -nearest neighbors. *Applied Soft Computing Journal*, 148, 1–12.

- Wang, H. & Liu, X. (2021). Undersampling bankruptcy prediction: Taiwan bankruptcy data. *Plos One*, 16(7), 1–17.
- Xiao, J., Wang, Y., Chen, J., Xie, L., & Huang, J. (2021). Impact of resampling methods and classification models on the imbalanced credit scoring problems. *Information Sciences*, 569, 508–526.
- Zhang, W., Yang, D., & Zhang, S. (2021). A new hybrid ensemble model with voting-based outlier detection and balanced sampling for credit scoring. *Expert Systems with Applications*, 174, 114744.
- Zhou, Y., Shamsu Uddin, M., Habib, T., Chi, G., & Yuan, K. (2021). Feature selection in credit risk modeling: an international evidence. *Economic Research-Ekonomska Istraživanja*, 34(1), 3064–3091.
- University of California Irvine. Machine Learning Repository. Eriřim: 06.11.2024. <https://archive.ics.uci.edu/>

#### **YAZAR BEYANI**

Bu makale Arařtırma ve Yayın Etiđine uygundur. Beyan edilecek herhangi bir ıkar atıřması yoktur. Arařtırmanın ortaya konulmasında herhangi bir mali destek alınmamıřtır. **Mevcut alıřma iin mevzuat geređi etik izni alınmaya ihtiya yoktur. Bu konuda yazarlar tarafından dergiye “Etik İznine Gerek Olmadıđına Dair Beyan Formu” gnderilmiřtir.**

#### **ARAřTIRMACILARIN MAKALEYE KATKI ORANI BEYANI**

- 1. yazar katkı oranı: %50**
- 2. yazar katkı oranı: %50**

## EKLER

Ek 1. German ve Australian Veri Seti için Hiperparametre Seçim Tablosu

|                                 |                   | German Hiperparametre Seçimi |       |      |      |       | Australian Hiperparametre Seçimi |        |        |       |       |
|---------------------------------|-------------------|------------------------------|-------|------|------|-------|----------------------------------|--------|--------|-------|-------|
| Makine Öğrenme Sınıflandırıcısı | Hiperparametre    | DÜZEY                        | SMOTE | ROS  | RUS  | TOMEK | DÜZEY                            | SMOTE  | ROS    | RUS   | TOMEK |
| SVM                             | C                 | 1000                         | 1     | 10   | 0.5  | 1     | 128                              | 10     | 10     | 128   | 128   |
|                                 | gamma             | 0.0001                       | 0.1   | 0.1  | 0.1  | 0.01  | 1e-05                            | 0.0001 | 0.0001 | 1e-05 | 1e-05 |
|                                 | kernel            | rbf                          | rbf   | rbf  | rbf  | rbf   | rbf                              | rbf    | rbf    | rbf   | rbf   |
| RF                              | bootstrap         | True                         | True  | True | True | True  | True                             | True   | True   | True  | True  |
|                                 | max_depth         | 20                           | 30    | 20   | None | None  | 5                                | None   | 20     | 10    | 30    |
|                                 | min_samples_leaf  | 3                            | 1     | 1    | 3    | 2     | 1                                | 2      | 1      | 3     | 1     |
|                                 | min_samples_split | 2                            | 2     | 3    | 3    | 2     | 4                                | 6      | 4      | 4     | 6     |
|                                 | n_estimators      | 100                          | 400   | 400  | 400  | 300   | 100                              | 300    | 100    | 300   | 100   |
| XgBoost                         | max_depth         | 5                            | 10    | 10   | 3    | 2     | 10                               | 10     | 3      | 5     | 1     |
|                                 | eta               | 0.01                         | 0.01  | 0.1  | 0.01 | 0.5   | 0.01                             | 0.1    | 0.01   | 0.01  | 0.5   |
|                                 | n_estimators      | 400                          | 400   | 100  | 500  | 100   | 300                              | 100    | 1000   | 400   | 100   |
|                                 | gamma             | 1                            | 0.5   | 0.5  | 0.5  | 0.5   | 0.5                              | 0.5    | 0.5    | 1     | 5     |
| CatBoost                        | max_depth         | 5                            | 7     | 9    | 3    | 5     | 3                                | 7      | 9      | 3     | 5     |
|                                 | learning_rate     | 0.1                          | 0.1   | 0.1  | 0.1  | 0.01  | 0.1                              | 0.1    | 0.01   | 0.1   | 0.01  |
|                                 | n_estimators      | 100                          | 500   | 500  | 100  | 1500  | 300                              | 400    | 1500   | 300   | 1500  |
|                                 | l2_leaf_reg       | 10                           | 10    | 10   | 10   | 20    | 5                                | 5      | 1      | 1     | 10    |