

Borsa İstanbul'da Muhasebe Hilesi Tespiti: Beneish Modeli ve SMOTE Tabanlı Lojistik Regresyon Yaklaşımı*

Suat KARA**

Pelin ÖZCAN***

ÖZET

Muhasebe hileleri, sermaye piyasalarında yatırımcı güvenini sarsan ve finansal istikrarı tehdit eden ciddi bir sorundur. Bu çalışmanın amacı, Türkiye'deki imalat sanayi firmalarında muhasebe hilesi riskini tespit etmek üzere Beneish Modeli temelli lojistik regresyon skorları geliştirmektir. Çalışmada, 2018–2022 dönemine ait Finnet verileri kullanılmıştır. Manipülasyon yaptığı düşünülen firmalar, bağımsız denetim raporlarında yer alan olumsuz ve şartlı görüş ifadelerine göre; manipülasyon riski düşük olanlar ise Kurumsal Yönetim Endeksi firmaları dikkate alınarak belirlenmiştir. Ham veri ile oluşturulan modelin başarısı sınırlı kalmış, bu durum azınlık sınıf verisinin yetersizliğine bağlanmıştır. Bu sorunu aşmak için Sentetik Azınlık Aşırı Örnekleme (SMOTE) yöntemi uygulanmış ve iki modelin performansları karşılaştırılmıştır. SMOTE yöntemiyle oluşturulan modelin doğruluk (%85), F1 skoru (%82) ve AUC değeri (0.89) gibi ölçütlerde anlamlı şekilde daha başarılı olduğu tespit edilmiştir. Bu çalışma, Beneish modelinin SMOTE ile desteklenerek Türkiye'de sektör bazlı uygulanmasının hile tespitinde etkili olduğunu ortaya koyarak literatüre özgün katkı sunmaktadır. Politika yapıcılara, finansal tablo denetim süreçlerinde istatistiksel temelli erken uyarı sistemlerinin desteklenmesi ve regülasyonlara entegre edilmesi önerilmektedir.

Anahtar Kelimeler: Muhasebe Hileleri, Beneish Modeli, SMOTE

Jel Sınıflandırması: M41, M42

Detecting Accounting Fraud in Borsa Istanbul: A Logistic Regression Approach Based on the Beneish Model and SMOTE

ABSTRACT

Accounting fraud is a serious problem that undermines investor confidence in capital markets and threatens financial stability. The aim of this study is to develop logistic regression scores based on the Beneish Model to detect the risk of accounting fraud in manufacturing industry companies in Türkiye. Finnet data for the period 2018–2022 was used in the study. Companies thought to have committed manipulation were determined according to the negative and conditional opinion statements in the independent audit reports; those with a low risk of manipulation were determined by considering the Corporate Governance Index companies. The success of the model created with raw data was limited, and this was attributed to the insufficiency of minority class data. To overcome this problem, the Synthetic Minority Oversampling (SMOTE) method was applied and the performances of the two models were compared. It was determined that the model created with the SMOTE method was significantly more successful in criteria such as accuracy (85%), F1 score (82%) and AUC value (0.89). This study provides an original contribution to the literature by demonstrating that the sector-based application of the Beneish model supported by SMOTE in Turkey is effective in fraud detection. It is recommended that policy makers support statistical-based early warning systems in financial statement audit processes and integrate them into regulations.

Keywords: Accounting Fraud, Beneish Model, SMOTE

Jel Classification: M41, M42

* Published by The Journal of Accounting and Finance. This article is published under the Creative Commons Attribution (CC BY 4.0) licence.

Makale Gönderim Tarihi: 29.04.2025, **Makale Kabul Tarihi:** 07.07.2025, **Makale Türü:** Araştırma Makalesi
Bu çalışma, Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü'nde yürütülen 'Muhasebe Hilelerinde Algoritmik Yaklaşım' başlıklı doktora tezinden elde edilen bulgular temel alınarak hazırlanmış ve Balıkesir Üniversitesi Bilimsel Araştırma Projeleri Birimi (BAP) tarafından 2022/032 numaralı proje kapsamında desteklenen doktora tezinden türetilmiştir.

** Prof. Dr., Balıkesir Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve Lojistik Bölümü, suatkara@balikesir.edu.tr, ORCID: 0000-0001-7818-2551

*** Balıkesir Üniversitesi, SBE., pbekret@gmail.com, ORCID: 0000-0002-6507-9763

1. GİRİŞ

Finansal raporlar, şirketlerin ekonomik faaliyetlerini doğru, güvenilir ve şeffaf bir şekilde yansıtmaları açısından hayati öneme sahiptir. Bu raporlar; yatırımcılar, kredi verenler, düzenleyici kurumlar ve kamuoyu tarafından alınacak kararların temelini oluşturur. Ancak zaman zaman bazı şirketlerin, çeşitli gerekçelerle bilançolarını ve gelir tablolarını olduğundan daha olumlu göstermek amacıyla gerçeğe aykırı finansal bilgiler sundukları görülmektedir. Bu tür uygulamalar “muhasabe hilesi” ya da “finansal tablo manipülasyonu” olarak adlandırılmaktadır (Kara ve Özcan, 2020:158). Muhasebe hileleri, şirketleri kısa vadede daha karlı veya mali olarak güçlü gösterme amacı taşısa da, bu durum orta ve uzun vadede şirketin sürdürülebilirliğine zarar vermekte ve finansal sistemin bütünlüğünü tehdit etmektedir.

Muhasebe hileleri yalnızca şirket yönetimlerini değil, aynı zamanda piyasadaki yatırımcı güvenini, ekonomik istikrarı ve sermaye piyasalarının işleyişini doğrudan etkileyebilmektedir. Dolayısıyla ekonomiye de ek maliyet getirmektedir (Uğurlu, 2011:42). 2000’li yılların başında ortaya çıkan Enron, WorldCom ve Xerox gibi büyük muhasabe skandalları, yalnızca ilgili ülke ekonomilerine ciddi zararlar vermekle kalmamış, aynı zamanda muhasabe standartları ile denetim süreçlerinin küresel ölçekte yeniden gözden geçirilmesine yol açmıştır.

Literatürde muhasabe hilelerinin tespiti için çeşitli modeller (Healy, DeAngelo, Jones vb.) bulunmaktadır. Bunlar arasında en yaygın olarak kullanılan model Beneish Modeli’dir. Bu model, finansal rasyolardan yararlanarak şirketlerin manipülasyon yapıp yapmadığını istatistiksel olarak ölçmeye çalışır. Diğer modeller ise, daha çok kazanç yönetimini ölçmeyi amaçlamaktadır ve genellikle olağan dışı tahakkukların hesaplanmasına dayanır. Ancak bu modeller genellikle ABD gibi gelişmiş ekonomilerin verilerine dayanılarak oluşturulmuştur. Bu sebeple Küçüksözen (2004), doktora tezinde Beneish Modeli’ni uyarlayarak bağımsız değişkenlere satışlardaki büyüme endeksi ve finansman giderlerinin satışlara oranı değişkenlerini eklemiştir.

Bu çalışma, Türkiye’de Borsa İstanbul’da (BİST) işlem gören imalat sanayi şirketlerinin finansal oranları üzerinden muhasabe hilesi riskini öngörebilecek istatistiksel skor modelleri geliştirmeyi amaçlamaktadır. Bu kapsamda, Beneish Modeli temel alınarak iki ayrı lojistik

regresyon skoru oluşturulmuştur. Şirketlerin finansal raporlama kapsamındaki farklılıkları dikkate alınarak, nakit akış tablosu yayımlamayan firmalar için dokuz değişkenli; bu tabloyu yayımlayan firmalar için ise on değişkenli ayrı model yapıları kurgulanmıştır. Böylece veri setinde yer alan firmaların raporlama düzeylerine uygun olarak daha tutarlı ve karşılaştırılabilir modeller geliştirilmesi hedeflenmiştir. Veri setinde manipülasyon yaptığı kabul edilen firma sayısının oldukça sınırlı olması, tahmin modellerinin azınlık sınıfını doğru sınıflama yeteneğini zayıflatmaktadır. Bu sınıf dengesizliğini gidermek ve modelin performansını artırmak amacıyla, her iki model SMOTE (Sentetik Azınlık Örnekleme) yöntemi ile yeniden yapılandırılarak dengelenmiş veri setleriyle de test edilmiştir.

SMOTE yöntemi ile azınlık sınıfa ait örneklerin artırılmasıyla oluşturulan dengeli veri setleri sayesinde, lojistik regresyon modellerinin performansı dört farklı yapı altında test edilmiştir. Bu yapılar, SMOTE uygulanmamış ve uygulanmış olmak üzere hem dokuz değişkenli hem de on değişkenli modelleri kapsamaktadır. Modellerin performansları doğruluk oranı, duyarlılık, özgüllük, kesinlik ve F1 skoru gibi ölçütlerle karşılaştırılmış; özellikle SMOTE'un model başarımına etkisi ampirik olarak ortaya konulmuştur.

Bu çalışma, Beneish modeline dayalı muhasebe hilesi tespitinde, veri dengesizliğinin etkisini sorgulamakta ve örnekleme temelli bir yaklaşım olan SMOTE'un uygulanabilirliğini imalat sektörü özelinde test etmektedir. Geliştirilen skor modelleri özellikle dengesiz veri yapılarının SMOTE gibi yöntemlerle dengelenmesinin model başarımını anlamlı ölçüde artırabileceğini göstermektedir.

2. LİTERATÜR TARAMASI

Beneish Modeli, Türkiye'de muhasebe hilelerinin tespitine yönelik olarak çeşitli dönemlerde ve farklı sektörlerde birçok çalışmada kullanılmıştır. Bu çalışmaların büyük bölümü, modeli olduğu gibi uygulamakta ya da Türkiye'ye uyarlamaya çalışmakta; ancak örnekleme yapısı, veri dengesizliği gibi metodolojik sorunları çoğu zaman dikkate almamaktadır.

Varıcı ve Er (2013), Beneish Modeli'ni BIST 100 endeksindeki 39 imalat firması için uygulamış ve firmaların yaklaşık yarısının finansal manipülasyon riski taşıdığı sonucuna

ulaşmıştır. Benzer şekilde Fındık ve Öztürk (2016), model aracılığıyla finansal manipülasyonun tahakkuk esaslı uygulamalardan kaynaklandığını belirtmişlerdir. Bu çalışmalar, Beneish Modeli'nin temel göstergelerini kullanmakla birlikte, istatistiksel modelleme içermemektedir.

Tepeli ve Kayıhan (2016) ise Türkiye'ye özgü uyarlamalar içeren bir versiyonu kullanarak gıda sektöründe firma bazlı analiz yapmış; ancak sınıflandırma başarımı ya da modelin tahmin gücüne yönelik ölçütler sunmamışlardır. Kara vd. (2016), denetim raporlarındaki şartlı görüşleri dikkate alarak manipülatör firmaları belirlemiş ve sınıflandırma doğruluğunu analiz etmiştir. Bu çalışma, örnekleme seçiminde literatüre katkı sunsa da veri dengesizliği gibi önemli bir istatistiksel sorunla ilgilenmemiştir.

Tekin (2017), SPK bültenleri ve denetim raporlarını bir arada değerlendirerek probit regresyon uygulamış; ancak yine sınıf dağılımı problemi göz ardı edilmiştir. Kara ve Özcan (2020) ise Beneish Modeli'ni yapay sinir ağları ile test etmiş ve %100 doğruluk oranı bildirmiştir. Ancak, yüksek doğruluk değerleri modelin çoğunluk sınıfa aşırı duyarlı olabileceğine işaret etmektedir.

Uluslararası bağlamda, Hołda (2020), Beneish Modeli'nin Varşova Borsası'ndaki şirketlerde uygulanabilirliğini test etmiş ve 8 değişkenli modelin daha gerçekçi sonuçlar sunduğunu ortaya koymuştur. Benligiray ve Onay (2021) ise modeli Türkiye koşullarına uyarlamış ve "Beneish TR" versiyonunun daha başarılı olduğunu göstermiştir. Her iki çalışma da modelin yerel uyarlamalarıyla ilgili önemli katkılar sunsa da, sınıf dengesizliğine yönelik sistematik bir çözüm önermemiştir.

Mevcut literatürde öne çıkan ortak sınırlılık, finansal manipülasyonun tespitinde kullanılan bu modellerin çoğunlukla dengesiz veri yapısını dikkate almaması ve sınıflandırma performansını iyileştirecek ileri düzey tekniklere (örneğin SMOTE gibi) başvurmamış olmasıdır.

Bu bağlamda, bu çalışma literatürdeki önemli bir boşluğu doldurmaktadır. Borsa İstanbul'da işlem gören imalat sanayi firmaları üzerinden yürütülen araştırmada, Beneish Modeli'nin iki varyasyonu geliştirilmiş, ham veri ile oluşturulan skor modelleri, veri dengesizliği problemi SMOTE yöntemi ile giderilerek tekrar kurgulanmıştır. Dokuz ve on

değişkenli olarak yapılandırılan bu modeller, firmaların raporlama düzeylerine uygun biçimde ayrıştırılmış ve sınıflandırma başarımları karşılaştırmalı olarak değerlendirilmiştir. Bu yönüyle çalışma, yalnızca Beneish Modeli'nin sektörel düzeyde uygulanabilirliğini değil, aynı zamanda veri dengesizliğine karşı geliştirilen istatistiksel çözümlerin etkisini de ortaya koymaktadır. Bu katkılar, hem akademik araştırmalar hem de uygulayıcı kurumlar için erken uyarı sistemlerinin tasarımına ışık tutacak niteliktedir.

3. BENEİSH MODELİ

Beneish (1999), muhasebe manipülasyonu yapan şirketler ile yapmayan şirketleri finansal tablo verilerini kullanarak oluşturduğu değişkenleri kullanarak bir model oluşturmuştur. 2013 yılında yapılan bir çalışmada ise, kazanç manipülasyonuna yönelik geliştirilen modelin, hisse senedi yatırımcıları üzerinde nasıl bir etki yarattığı incelenmiştir (Kara vd.,2023:29). Bu model aşağıdaki gibidir:

$$M_i = \beta X_i + \varepsilon_i$$

M_i : Bağımlı değişkeni (Manipülatör şirketler için 1, kontrol şirketleri için 0)

X_i : Bağımsız değişken

ε_i : Hata terimi

Beneish (1999) finansal tablo verilerinden elde edilmiş değişkenleri probit regresyona tabi tutarak aşağıdaki denkleme ulaşmıştır.

$$M = -4,84 + 0,920 * DSRI + 0,528 * GMI + 0,404 * AQI + 0,892 * SGI + 0,115 * DEPI - 0,172 * SGAI + 4,67 * Acr - 0,327 * LEVI$$

Modelde bulunan bağımsız değişkenler aşağıdaki tabloda gösterilmiştir.

Tablo 1. Türkiye’ye Uyarlanmış Beneish Modeli Bağımsız Değişkenleri

Kısaltma	Değişken Adı	Formül	Açıklama
SGI	Satışların Büyüme Oranı	$\text{Brüt Satışlar (t)} / \text{Brüt Satışlar (t-1)}$	Şirketin yıllık satış büyümesini gösterir.
DSRI	Alacakların Satışlara Oranı	$(\text{Alacaklar (t)} / \text{Satışlar (t)}) \div (\text{Alacaklar (t-1)} / \text{Satışlar (t-1)})$	Alacaklardaki artış, gelir tahakkuk riskini gösterebilir.
GMI	Brüt Kar Marjı Endeksi	$\text{Brüt Kar Marjı (t-1)} / \text{Brüt Kar Marjı (t)}$	Marjdaki düşüş, yöneticilerin manipülasyon yapma ihtimalini artırabilir.
AQI	Aktif Kalite Endeksi	$[1 - (\text{Dönen Varlıklar (t)} / \text{Toplam Aktifler (t)})] + (\text{Maddi Duran Varlıklar (t)} / \text{Toplam Aktifler (t-1)})$	Maddi olmayan varlıkların oranındaki artışı yansıtır.
ACR	Nakit Olmayan Faaliyetlerden Sağlanan Nakit Akışı Oranı	$(\text{Faaliyet Kârı} - \text{Nakit Akışı}) / \text{Toplam Aktifler}$	Kâr ile gerçek nakit akışı arasındaki farkı gösterir.
DEPI	Amortisman Oranı Endeksi	$\text{Amortisman (t-1)} / [\text{Amortisman (t)} + \text{Amortisman (t-1)}]$	Amortisman yöntemindeki değişiklikleri veya yavaşlatılmış amortismanı yansıtır.
SGAI	Satış, Genel ve İdari Gider Endeksi	$(\text{Genel Giderler (t)} / \text{Satışlar (t)}) \div (\text{Genel Giderler (t-1)} / \text{Satışlar (t-1)})$	Gider artışlarının satışlarla uyumunu analiz eder.
LEVI	Kaldıraç Endeksi	$(\text{Borçlar (t)} / \text{Aktifler (t)}) \div (\text{Borçlar (t-1)} / \text{Aktifler (t-1)})$	Borçluluk düzeyindeki değişimi ölçer.
SSE	Stokların Brüt Satışlara Oranı (Küçüksözen, 2004)	$\text{Stoklar (t)} / \text{Satışlar (t-1)}$	Stokların satışlara oranla artışını gösterir.
FSE	Finansman Giderlerinin Maddi Duran Varlıklara Oranı (Küçüksözen, 2004)	$\text{Finansman Giderleri (t)} / \text{Maddi Duran Varlıklar (t)}$	Finansman yükünün sabit kıymetlere göre büyüklüğünü ölçer.

Beneish, 1999 yılında yapmış olduğu çalışmasında Acr değişkenini, nakit akım tablosu verileri kullanılmadan hesaplamıştır. Ancak, Beneish ve arkadaşları 2013 yılında yaptıkları güncellemede, bu değişkenin hesaplanma yönteminde bazı değişiklikler yapmışlardır. (Öcal vd., 2017:6).

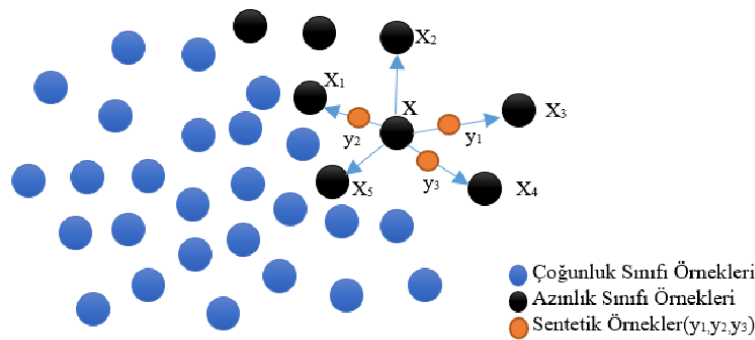
Beneish, aynı zamanda çeşitli hesaplamalarla modelin eşik değerini belirlemiş ve bir şirketin manipülatör olduğunu ya da olmadığını bu eşik değerlere göre yorumlamıştır. Aşağıdaki şekilde bu eşik değerler gösterilmiştir.



Şekil 1. Beneish Modeli Eşik Değer Görseli
(Kaynak: www.stablebread.com, 2025)

4. SMOTE (SENTETİK AZINLIK ÖRNEKLEME) TEKNIĞİ

Sınıflandırma problemlerinde, azınlık sınıfına ait örneklerin veri setinde sayıca yetersiz olması, modellerin bu sınıfı yeterince öğrenememesine ve dolayısıyla genel sınıflama başarımının düşmesine neden olabilmektedir. Bu tür dengesiz veri yapılarında, algoritmalar çoğu zaman baskın sınıfın örüntülerini öğrenmeye eğilim gösterir ve tahminlerinde bu sınıfa öncelik verir. Bu durum, muhasebe hilesi tespiti gibi nadir ancak kritik vakaların analizinde olduğu gibi; kanser teşhisi gibi düşük görülme sıklığına sahip sınıflarda da önemli bir sorun yaratmaktadır. Bu problemi aşmak amacıyla geliştirilen SMOTE (Synthetic Minority Over-sampling Technique) yöntemi, dengesiz veri setlerinin dengelenmesinde en yaygın kullanılan örnekleme tekniklerinden biridir. İlk olarak Chawla vd. tarafından 2002 yılında önerilen bu yöntem, farklı alanlardaki birçok dengesiz sınıflandırma problemi için başarılı bir şekilde uygulanmıştır. Geleneksel rastgele çoğaltma yöntemlerinden farklı olarak SMOTE, azınlık sınıfına ait her gözlemin en yakın k komşusu temel alınarak yeni ve sentetik örnekler üretir (Yavaş vd., 2020:261). Bu sayede veri seti, daha dengeli ve model öğrenimine elverişli hâle getirilir. Ayrıca, SMOTE bir örnekleme tekniği olmakla beraber, bu örnekleme işlemini bir algoritma mantığıyla yaptığı için literatürde yaygın olarak “SMOTE Algoritması” olarak kullanılmaktadır. Şekil 2’de SMOTE algoritmasının çalışma sistemi gösterilmektedir.



Şekil 2. SMOTE algoritması çalışma sistemi

Kaynak: Yavaş vd., 2020:261

5. YÖNTEM

Bu çalışmada, Türkiye'de Borsa İstanbul'da (BİST) işlem gören imalat sanayi firmalarının finansal verileri kullanılarak, muhasebe hilelerinin tespiti amacıyla Beneish modeline dayalı iki ayrı lojistik regresyon modeli geliştirilmiştir. Veri seti, firmaların finansal tabloları arasında nakit akış tablosunun bulunup bulunmama durumuna göre ikiye ayrılmış; nakit akış tablosu bulunmayan firmalar için 9 değişkenli (Toplam Tahakkukların Toplam Varlıklara Oranı hariç tutulmuştur), nakit akış tablosu bulunan firmalar için ise 10 değişkenli model yapıları oluşturulmuştur.

Çalışmanın bağımlı değişkeni olan “manipülasyon yapma durumu”, bağımsız denetim raporlarında yer alan şartlı görüş, olumsuz görüş veya görüş bildirmeme ifadelerine dayanılarak belirlenmiştir. Bu tür görüşlere sahip firmalar manipülatör (1), Kurumsal Yönetim Endeksi'nde yer alan firmalar ise manipülatör olmayan (0) olarak sınıflandırılmıştır. Bağımsız değişkenler, Beneish'in orijinal modelinde yer alan 8 finansal rasyo ile, Küçüksözen (2004) tarafından modele eklenen “Satışlardaki Büyüme Endeksi” ve “Finansman Giderlerinin Satışlara Oranı” değişkenlerinden oluşmaktadır.

Modelleme sürecine geçilmeden önce, bağımsız değişkenler arasında çoklu doğrusal bağlantı (multicollinearity) riskini azaltmak amacıyla korelasyon analizi yapılmıştır. Değişkenler arası ikili ilişkiler incelenmiş ve yüksek korelasyon gösteren değişkenler dikkatle değerlendirilmiştir.

Veri setinde sınıflar arasında gözlenen ciddi dengesizlik nedeniyle, azınlık sınıfa (manipülatör firmalar) ait örnek sayısı sınırlı kalmıştır. Bu durumun modelin öğrenme sürecini ve tahmin performansını olumsuz etkileyebileceği düşüncesinden yola çıkılarak, Synthetic Minority Over-sampling Technique (SMOTE) yöntemi kullanılarak dengeli veri setleri oluşturulmuştur.

Bu kapsamda, hem orijinal (SMOTE uygulanmamış) hem de SMOTE ile dengelenmiş veri setleri üzerinden 9 ve 10 değişkenli olmak üzere toplam dört farklı lojistik regresyon modeli oluşturulmuştur. Her bir modelin sınıflandırma performansı, Doğruluk (Accuracy),

Duyarlılık (Recall), Kesinlik (Precision), F1 Skoru, ROC Eğrisi Altındaki Alan (AUC) gibi istatistiksel ölçütler kullanılarak değerlendirilmiştir.

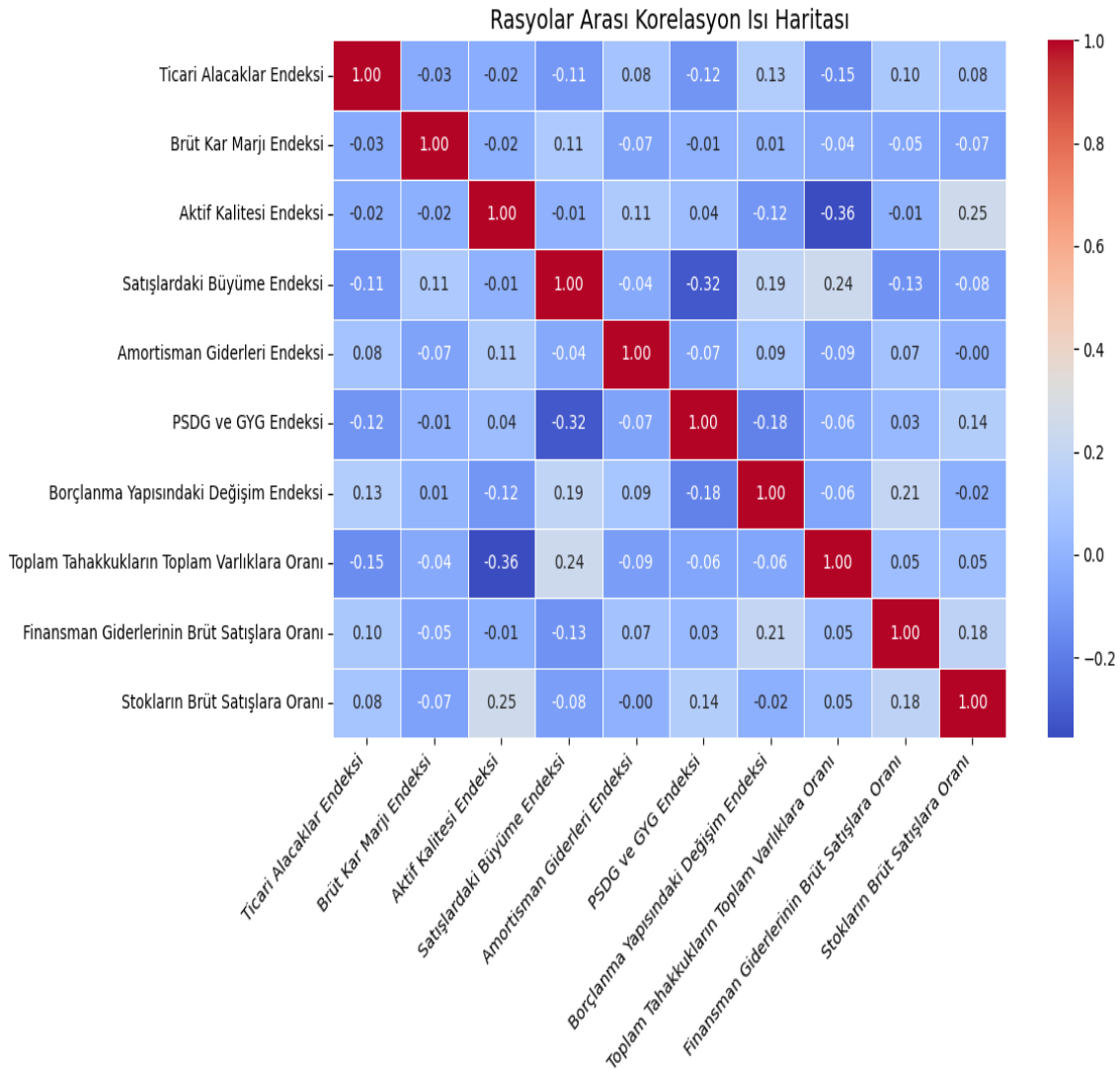
Model tahminlerinin sınıflandırılmasında en uygun eşik değerler, Youden J istatistiği temel alınarak belirlenmiştir. Böylece, modellerin özellikle azınlık sınıfı olan manipülatör firmaları ne ölçüde doğru tespit edebildiği karşılaştırmalı olarak analiz edilmiştir.

Tüm analizler, veri işleme, görselleştirme ve modelleme adımları dahil olmak üzere Python programlama dili ve ilgili kütüphaneler (pandas, scikit-learn, seaborn, matplotlib) kullanılarak gerçekleştirilmiştir.

6. BULGULAR

Modellemede kullanılan bağımsız değişkenler arasında yüksek korelasyon bulunup bulunmadığını tespit etmek amacıyla öncelikle korelasyon analizi yapılmıştır. Bu analiz, değişkenler arası çoklu doğrusal bağlantı (multicollinearity) riskinin kontrol edilmesi açısından önem arz etmektedir. Aksi takdirde, yüksek korelasyon gösteren değişkenlerin aynı anda modele dahil edilmesi, tahmin gücünü bozabilir ve modelin kararlılığını olumsuz etkileyebilir.

Şekil 1’de yer alan ısı haritası, çalışmada kullanılan 10 bağımsız değişken arasındaki Pearson korelasyon katsayılarını görsel olarak sunmaktadır. Görselde dikkat çeken en yüksek negatif korelasyon, Aktif Kalitesi Endeksi ile Toplam Tahakkukların Toplam Varlıklara Oranı arasında olup -0.36 düzeyindedir. Diğer değişken çiftleri arasında ise ± 0.50 ’yi aşan bir ilişki gözlenmemiştir. Korelasyon katsayılarının bu düzeyde kalması, değişkenler arasında yüksek düzeyli doğrusal ilişkinin bulunmadığını göstermekte; bu da çoklu doğrusal bağlantı riskinin düşük olduğunu ve değişkenlerin aynı anda modele dahil edilebileceğini ortaya koymaktadır.



Şekil 3. Rasyolar Arası Korelasyon Isı Haritası

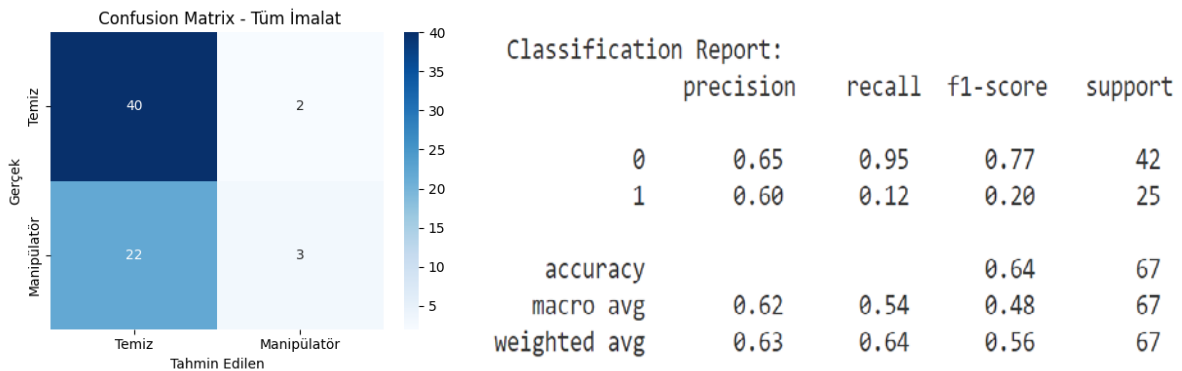
Bu çalışmada korelasyon analizleri Python programlama dili kullanılarak gerçekleştirilmiş; yalnızca açıklayıcı rasyolardan oluşan veri seti üzerinden korelasyon matrisi oluşturulmuş ve seaborn kütüphanesi aracılığıyla görselleştirme yapılmıştır. Kullanılan kod yapısı ve görselleştirme örneği Ek 2’de sunulmuştur.

6.1. Dokuz Değişkenli SMOTE Uygulanmamış Modelin Bulguları

SMOTE uygulanmaksızın geliştirilen dokuz değişkenli lojistik regresyon modeline ait bulgular modelin sınıflandırma başarımında ciddi bir dengesizlik olduğunu göstermektedir.

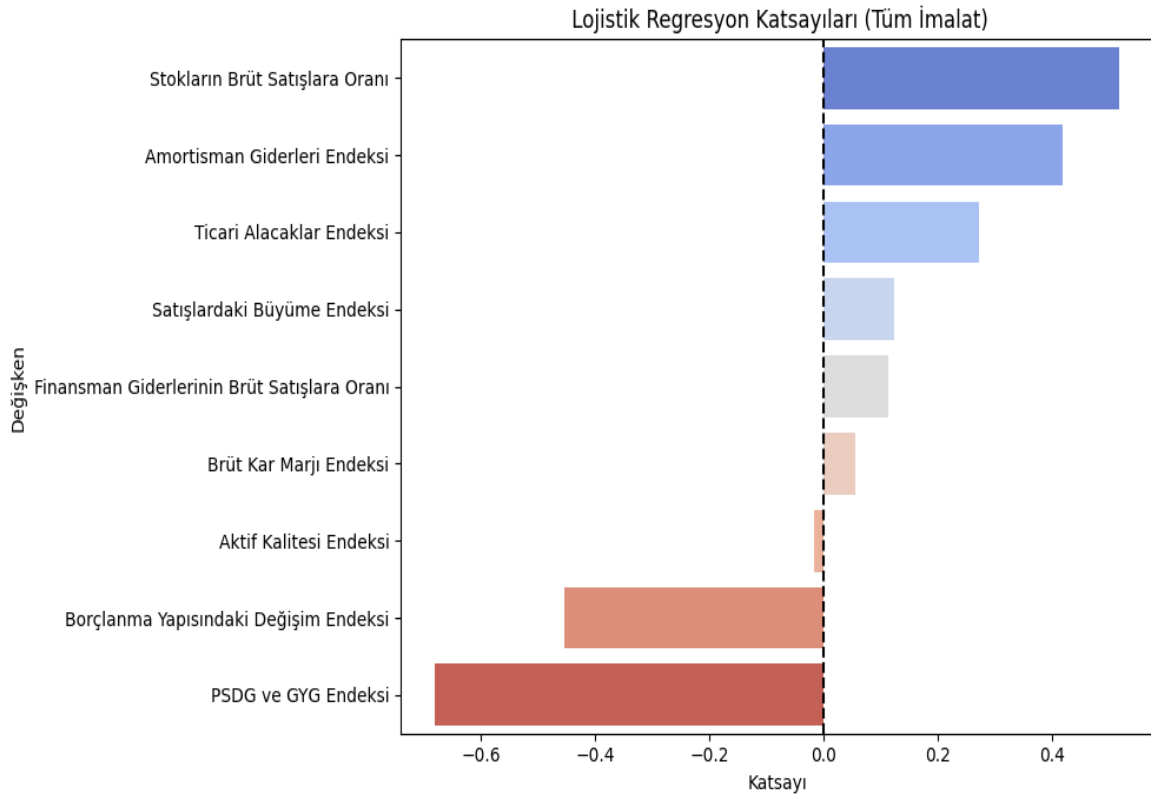
Elde edilen performans ölçütleri, modelin özellikle azınlık sınıf olan manipülatör firmaları tanımlamadaki başarısını değerlendirmek açısından önemlidir. Bu bağlamda, precision, modelin manipülasyon riski taşıyan firmalara yönelik tahminlerinde ne ölçüde isabetli olduğunu gösterirken; recall, gerçek manipülatör firmaların ne kadarının doğru tespit edildiğini ifade eder. Bu iki ölçümün dengeli bir değerlendirmesini sunan F1 skoru ise, özellikle sınıflar arası dengesizlik barındıran veri setlerinde modelin genel başarımını yansıtmaya açısından kritik öneme sahiptir.

Sınıflandırma raporuna göre genel doğruluk (precision) oranı %65 olmakla birlikte, model manipülasyon yapmayan firmaları ayırt etmede oldukça başarılı iken, manipülasyon yapan firmaları tespit etmede zayıf kalmıştır. Modelin sınıf 0 (kontrol) için duyarlılığı (recall) %95, F1 skoru ise %77 iken; sınıf 1 (manipülatör) için duyarlılık yalnızca %12, F1 skoru ise %20 seviyesindedir. Bu durum, modelin azınlık sınıf olan hileli firmaları büyük ölçüde gözden kaçırdığını göstermektedir.



Şekil 4. Dokuz Değişkenli Modelin Tahmin Sonuçları (SMOTE Uygulanmamış)

Confusion matrix incelendiğinde, 25 adet manipülatör firmadan yalnızca 3'ünün doğru sınıflandırılabilirdiği, 22'sinin ise temize atandığı görülmektedir. Temiz firmalardan 42 tanesinden 40'ı doğru tahmin edilmiştir. Bu durum modelin ağırlıklı olarak çoğunluk sınıfına (temiz firmalara) odaklandığını ve öğrenme sürecinde azınlık sınıfı yeterince temsil edemediğini ortaya koymaktadır.



Şekil 5. Dokuz Değişkenli Modelin Lojistik Regresyon Katsayıları (SMOTE Uygulanmamış)

Şekil 6’ da görülen regresyon katsayılarına göre, modelde pozitif yönlü en belirgin katkı sağlayan değişkenler sırasıyla stokların brüt satışlara oranı, amortisman giderleri endeksi ve ticari alacaklar endeksi olmuştur. Diğer yandan, negatif katsayıya sahip olan borçlanma yapısındaki değişim endeksi ve özellikle pazarlama satış dağıtım giderleri ve genel yönetim giderleri endeksi, manipülasyon riskinin azalmasıyla ilişkilendirilmiştir. Bu durum, belirli oranların manipülasyon ihtimaliyle pozitif ya da negatif yönlü ilişkisinin yorumlanmasında önemli ipuçları sunmaktadır.

Lojistik Regresyon Modeli:

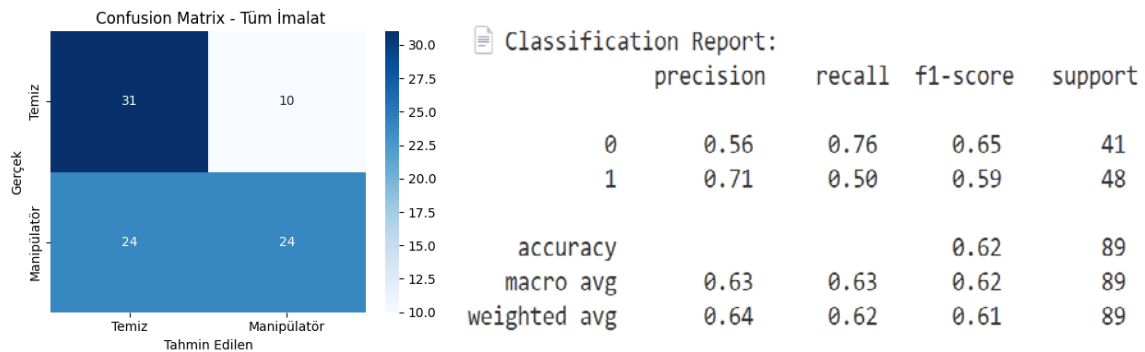
$$M = -1.4113 + (0.272 \times \text{Ticari Alacaklar Endeksi}) + (0.0549 \times \text{Brüt Kar Marjı Endeksi}) + (-0.0168 \times \text{Aktif Kalitesi Endeksi}) + (0.1243 \times \text{Satışlardaki Büyüme Endeksi}) + (0.4186 \times \text{Amortisman Giderleri Endeksi}) + (-0.6797 \times \text{PSDG ve GYG Endeksi}) + (-0.4524 \times \text{Borçlanma Yapısındaki Değişim Endeksi}) + (0.1125 \times \text{Finansman Giderlerinin Brüt Satışlara Oranı}) + (0.5174 \times \text{Stokların Brüt Satışlara Oranı})$$

Sonuç olarak, bu model yalnızca çoğunluk sınıf için anlamlı tahminlerde bulunabilmiş; manipülasyon tespiti gibi kritik bir konuda etkinlik seviyesi oldukça düşük kalmıştır. Bu bağlamda, sınıf dengesizliği probleminin ortadan kaldırılması amacıyla bir sonraki aşamada SMOTE uygulanarak veri seti dengelenmiş ve model performansının nasıl değiştiği analiz edilmiştir.

6.2. Dokuz Değişkenli SMOTE Uygulanmış Modelin Bulguları

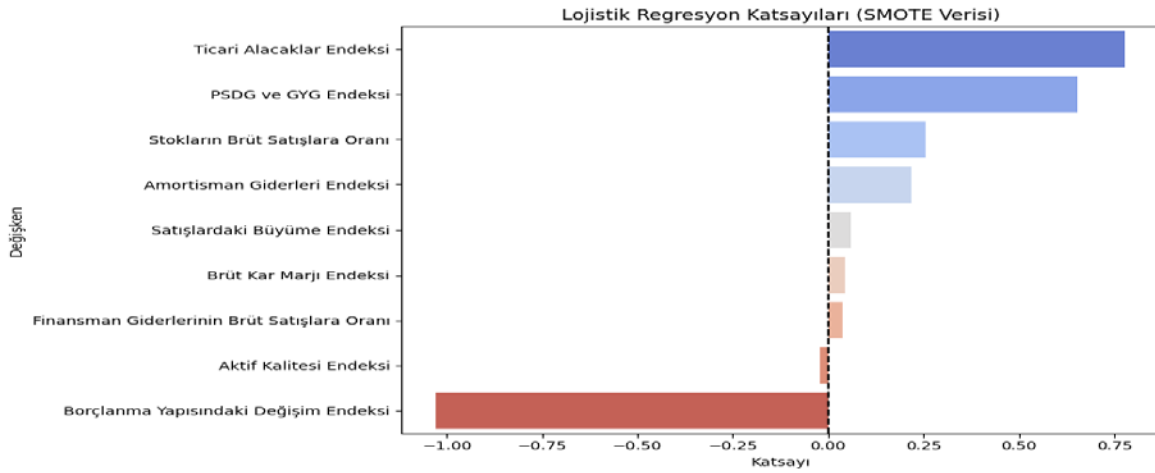
Dengesiz sınıf dağılımının model başarımı üzerindeki etkisini gidermek amacıyla, veri setine SMOTE (Synthetic Minority Over-sampling Technique) uygulanarak azınlık sınıf olan manipülatör firmalara ait sentetik gözlemler oluşturulmuştur. Bu işlem sonrasında elde edilen dengelenmiş veri setiyle kurulan 9 değişkenli lojistik regresyon modeli, önceki modele kıyasla manipülasyon tespiti açısından daha dengeli ve tatmin edici sonuçlar üretmiştir.

Sınıflandırma çıktısına göre, modelin genel doğruluk oranı %62 olup sınıf 0 (temiz) için duyarlılık %76, sınıf 1 (manipülatör) için ise %50 olarak gerçekleşmiştir. Özellikle azınlık sınıf olan manipülatör firmalarda recall değerinin SMOTE uygulanmamış modele kıyasla önemli ölçüde yükseldiği gözlemlenmektedir. Bu durum, SMOTE'un modeli azınlık sınıfı tanımaya yönelik daha duyarlı hâle getirdiğini ortaya koymaktadır.



Şekil 6. Dokuz Değişkenli Modelin Tahmin Sonuçları (SMOTE Uygulanmış)

Confusion matrix'e göre, model 48 manipülatör firmadan 24'ünü doğru tahmin etmiş, 24'ünü ise temiz olarak sınıflandırmıştır. Bu oran, SMOTE uygulanmamış modele göre önemli bir gelişme göstermektedir (önceki model yalnızca 3 manipülatörü tanıyabilmişti). Temiz firmalarda ise 31 doğru, 10 yanlış sınıflandırma yapılmıştır.



Şekil 7. Dokuz Değişkenli Modelin Lojistik Regresyon Katsayıları (SMOTE Uygulanmış)

Katsayı analizine göre, modele pozitif katkı sağlayan en güçlü değişken ticari alacaklar endeksi (0.778) olup, onu PSDG ve GYG endeksi (0.653) ve stokların brüt satışlara oranı (0.256) izlemektedir. Negatif katsayıya sahip en etkili değişken ise borçlanma yapısındaki değişim endeksi (-1.029) olmuştur. Bu bulgular, bazı finansal göstergelerin manipülasyon eğilimi ile güçlü biçimde ilişkili olduğunu ortaya koymaktadır.

Lojistik Regresyon Modeli:

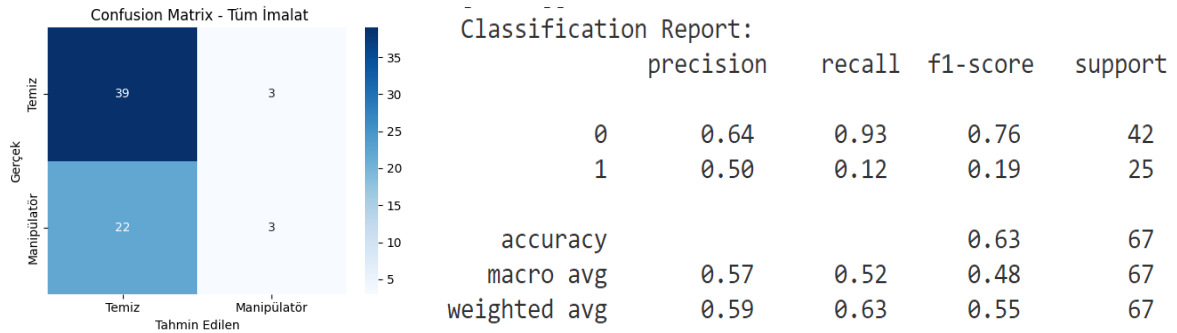
$$M = -1.3299 + (0.778 \times \text{Ticari Alacaklar Endeksi}) + (0.0455 \times \text{Brüt Kar Marjı Endeksi}) + (-0.0203 \times \text{Aktif Kalitesi Endeksi}) + (0.0595 \times \text{Satışlardaki Büyüme Endeksi}) + (0.2173 \times \text{Amortisman Giderleri Endeksi}) + (0.6526 \times \text{PSDG ve GYG Endeksi}) + (-1.0286 \times \text{Borçlanma Yapısındaki Değişim Endeksi}) + (0.0385 \times \text{Finansman Giderlerinin Brüt Satışlara Oranı}) + (0.2557 \times \text{Stokların Brüt Satışlara Oranı})$$

Sonuç olarak, SMOTE uygulanmış bu model, dengesizliğin giderilmesiyle birlikte manipülatör firmaları tanımada önemli ölçüde daha başarılı sonuçlar üretmiştir. Duyarlılık ve F1 skorundaki yükselişler, modelin her iki sınıfı da daha dengeli şekilde öğrendiğini göstermektedir. Bu durum, özellikle finansal hile tespiti gibi azınlık sınıfa odaklanan çalışmalarda veri dengeleme tekniklerinin etkili bir yöntem olduğunu göstermektedir.

6.3. On Değişkenli SMOTE Uygulanmamış Modelin Bulguları

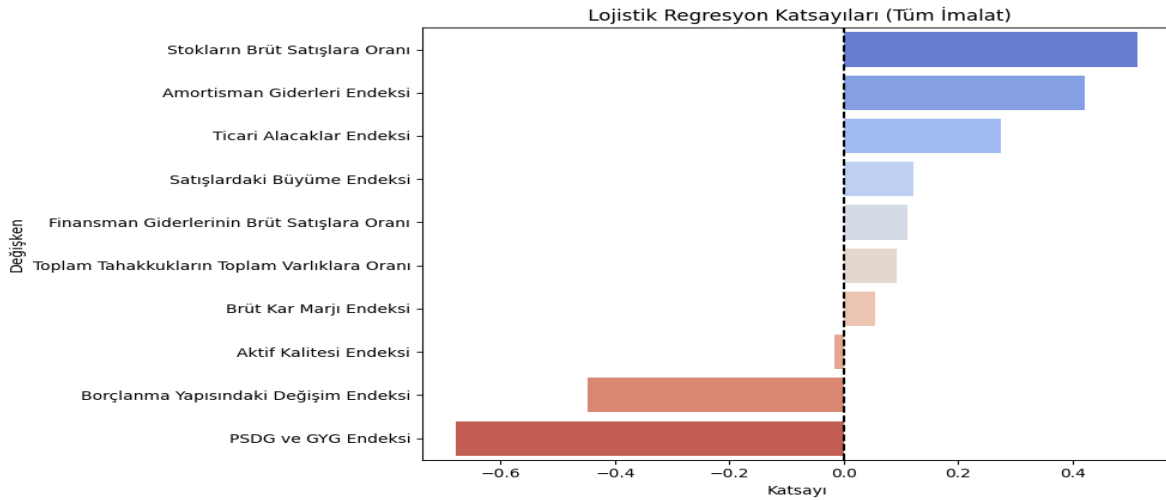
Nakit akış tablosu yayımlayan firmalara ait finansal oranlar kullanılarak oluşturulan 10 değişkenli lojistik regresyon modeline SMOTE uygulanmadan yapılan analizde, modelin genel doğruluk oranı %63 olarak hesaplanmıştır. Ancak bu doğruluk, önceki 9 değişkenli modele benzer şekilde yine çoğunluk sınıfına (temiz firmalar) odaklanan bir yapıya sahiptir.

Sınıflandırma performans metriklerine göre modelin sınıf 0 (kontrol) için duyarlılık oranı %93 gibi yüksek bir seviyede gerçekleşmiş; buna karşın sınıf 1 (manipülatör) için duyarlılık yalnızca %12 düzeyinde kalmıştır. Bu sonuç, modelin azınlık sınıfı öğrenmede yetersiz kaldığını ve manipülatör firmaların büyük bölümünü gözden kaçırdığını göstermektedir.



Şekil 8. On Değişkenli Modelin Tahmin Sonuçları (SMOTE Uygulanmamış)

Yukarıdaki confusion matrix'te görüldüğü üzere, model 25 manipülatör firmadan yalnızca 3'ünü doğru sınıflandırabilmiş, 22 tanesini yanlışlıkla temiz olarak etiketlemiştir. Buna karşılık 42 temiz firmadan 39'u doğru tanımlanmıştır.



Şekil 9. On Değişkenli Modelin Lojistik Regresyon Katsayıları (SMOTE Uygulanmamış)

Regresyon katsayılarına göre, modele en yüksek pozitif katkı sağlayan değişken Stokların Brüt Satışlara Oranı (0.47) olurken, bunu Amortisman Giderleri Endeksi (0.42) ve Ticari Alacaklar Endeksi (0.29) izlemiştir. En güçlü negatif etkiye sahip değişken ise PSDG ve GYG Endeksi (-0.64) olmuştur. Bu bulgular, modelin hile olasılığına ilişkin tahminlerini belirli finansal oranlara dayalı olarak oluşturduğunu göstermektedir.

Lojistik Regresyon Modeli:

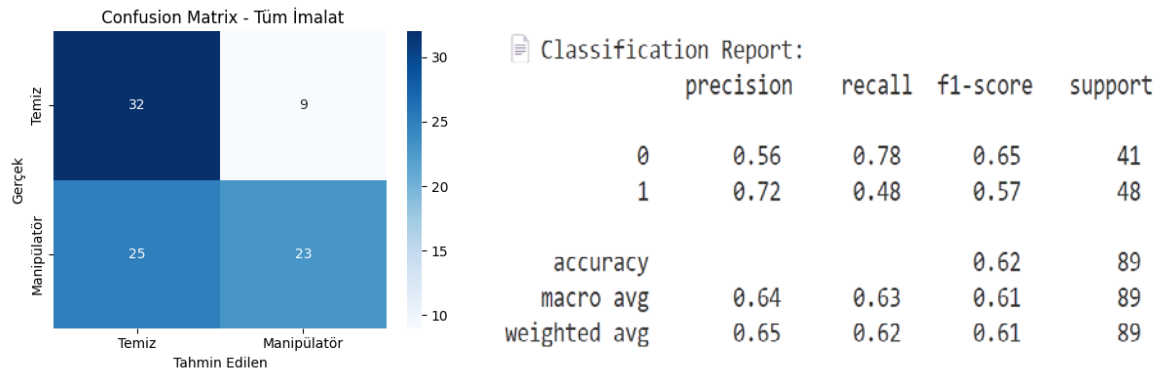
$$M = -1.414 + (0.2752 \times \text{Ticari Alacaklar Endeksi}) + (0.0552 \times \text{Brüt Kar Marjı Endeksi}) + (-0.0162 \times \text{Aktif Kalitesi Endeksi}) + (0.1222 \times \text{Satışlardaki Büyüme Endeksi}) + (0.4207 \times \text{Amortisman Giderleri Endeksi}) + (-0.6776 \times \text{PSDG ve GYG Endeksi}) + (-0.4489 \times \text{Borçlanma Yapısındaki Değişim Endeksi}) + (0.0925 \times \text{Toplam Tahakkukların Toplam Varlıklara Oranı}) + (0.1117 \times \text{Finansman Giderlerinin Brüt Satışlara Oranı}) + (0.5127 \times \text{Stokların Brüt Satışlara Oranı})$$

Sonuç olarak, SMOTE uygulanmaksızın oluşturulan on değişkenli model, yalnızca çoğunluk sınıfa yüksek doğrulukla tahmin yapabilmiş; ancak hile tespiti gibi kritik bir konuda azınlık sınıfa yönelik yetersiz bir performans göstermiştir. Bu durum, bir sonraki aşamada veri setine SMOTE uygulanmasının zorunluluğunu ortaya koymaktadır.

6.4. On Değişkenli SMOTE Uygulanmış Modelin Bulguları

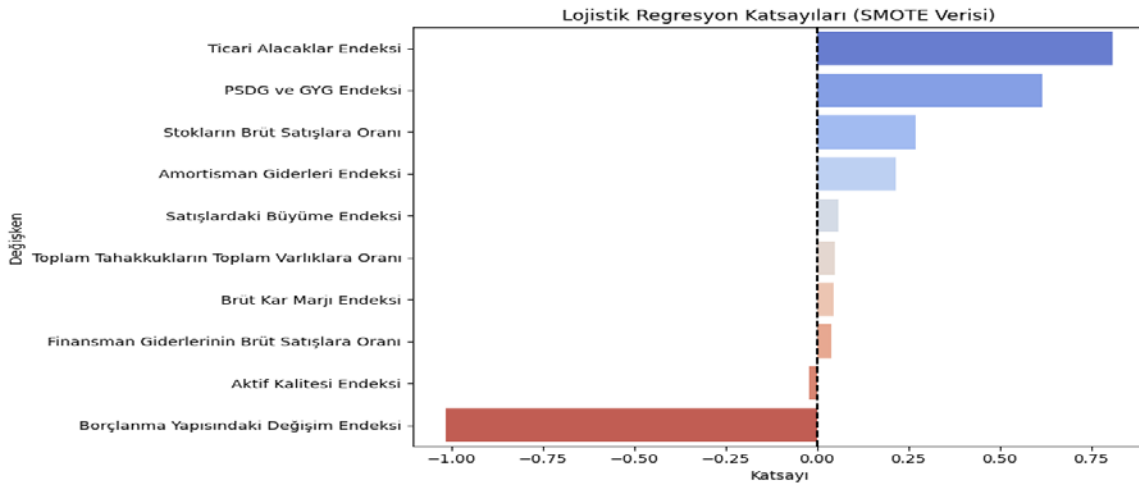
Veri setindeki sınıf dengesizliğinin azaltılması amacıyla, 10 değişkenli modelde SMOTE yöntemi uygulanarak azınlık sınıfa (manipülatör firmalar) ait sentetik gözlemler oluşturulmuştur. Bu sayede modelin hem sınıf 0 (kontrol) hem de sınıf 1 (manipülatör) üzerinde daha dengeli bir öğrenme yapısı geliştirmesi hedeflenmiştir.

Sınıflandırma performansı açısından modelin genel doğruluk oranı %62 olarak gerçekleşmiştir. Ancak bu modelin öne çıkan başarısı, manipülatör sınıfı tanımadaki iyileşmiş performansdır. Bu sınıf için precision değeri 0.72, recall değeri 0.48 ve F1 skoru 0,57 olarak hesaplanmıştır. Bu sonuçlar, SMOTE uygulanmamış modele göre önemli bir gelişmeye işaret etmektedir.



Şekil 10. On Değişkenli Modelin Tahmin Sonuçları (SMOTE Uygulanmış)

Confusion matrix'te de görülebileceği üzere, model manipülatör sınıfa ait 48 firmadan 23'ünü doğru sınıflandırmış; kalan 25'ini yanlışlıkla temiz olarak tanımlamıştır. Temiz sınıfta ise 32 doğru, 9 yanlış sınıflandırma yapılmıştır. Bu yapı, sınıf dengesinin sağlanmasının model performansına olumlu katkı sunduğunu ortaya koymaktadır.



Şekil 11. On Değişkenli Modelin Lojistik Regresyon Katsayıları (SMOTE Uygulanmış)

Katsayı grafiğine göre, modelde en güçlü pozitif etkiye sahip değişken ticari alacaklar endeksi (yaklaşık 0.78) olmuştur. Bunu PSDG ve GYG Endeksi (yaklaşık 0.65) ve stokların brüt satışlara oranı (yaklaşık 0.26) izlemiştir. Negatif yönlü en etkili değişken ise borçlanma yapısındaki değişim endeksi (yaklaşık -1.03) olarak öne çıkmaktadır.

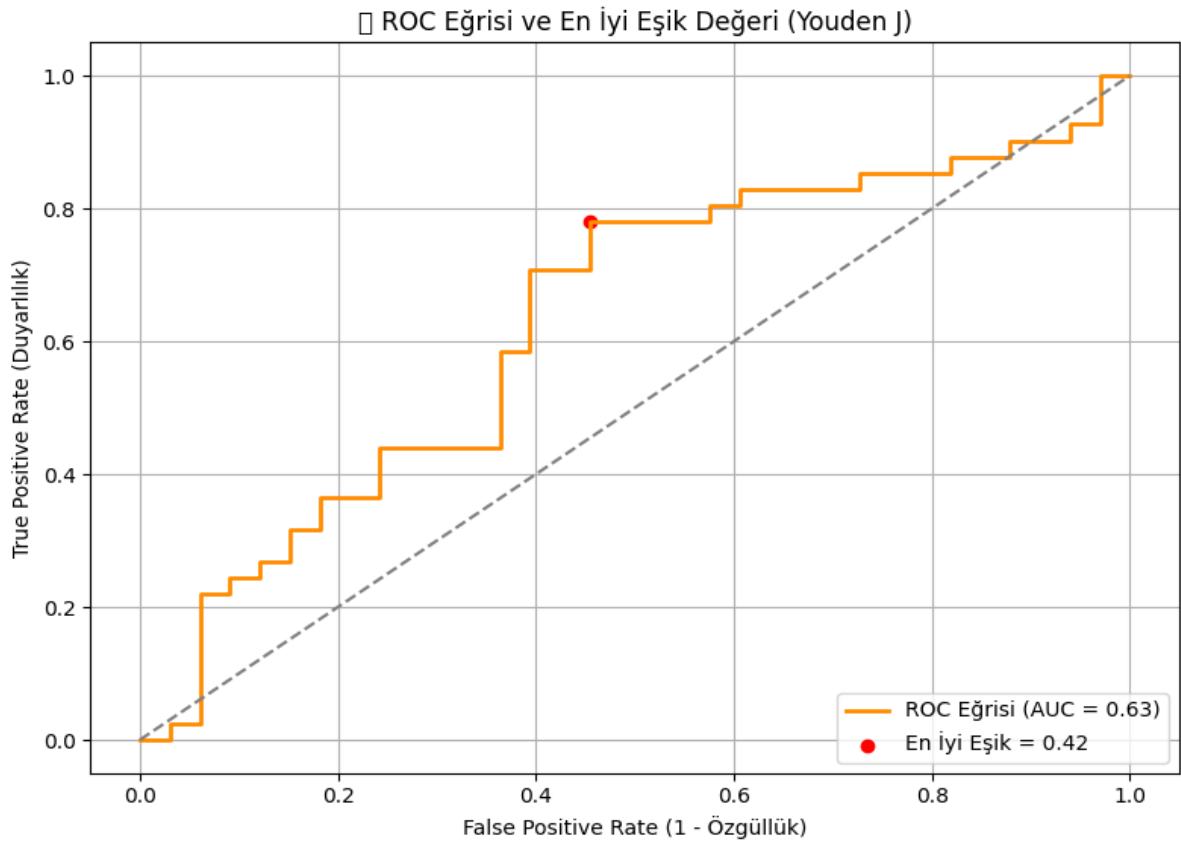
Lojistik Regresyon Modeli:

$$M = -1.352 + (0.8076 \times \text{Ticari Alacaklar Endeksi}) + (0.0464 \times \text{Brüt Kar Marjı Endeksi}) + (-0.0202 \times \text{Aktif Kalitesi Endeksi}) + (0.0584 \times \text{Satışlardaki Büyüme Endeksi}) + (0.216 \times \text{Amortisman Giderleri Endeksi}) + (0.6153 \times \text{PSDG ve GYG Endeksi}) + (-1.0135 \times \text{Borçlanma Yapısındaki Değişim Endeksi}) + (0.0494 \times \text{Toplam Tahakkukların Toplam Varlıklara Oranı}) + (0.0383 \times \text{Finansman Giderlerinin Brüt Satışlara Oranı}) + (0.2703 \times \text{Stokların Brüt Satışlara Oranı})$$

Genel olarak, SMOTE uygulanmış 10 değişkenli model, hem sınıf dengesini sağlamada hem de manipülatör sınıfa karşı daha yüksek duyarlılık göstermede anlamlı bir gelişme sunmuştur. Bu sonuçlar, muhasebe hilelerini tespit etmek amacıyla geliştirilen finansal modellerde SMOTE gibi veri dengeleme tekniklerinin etkili bir strateji olduğunu doğrulamaktadır.

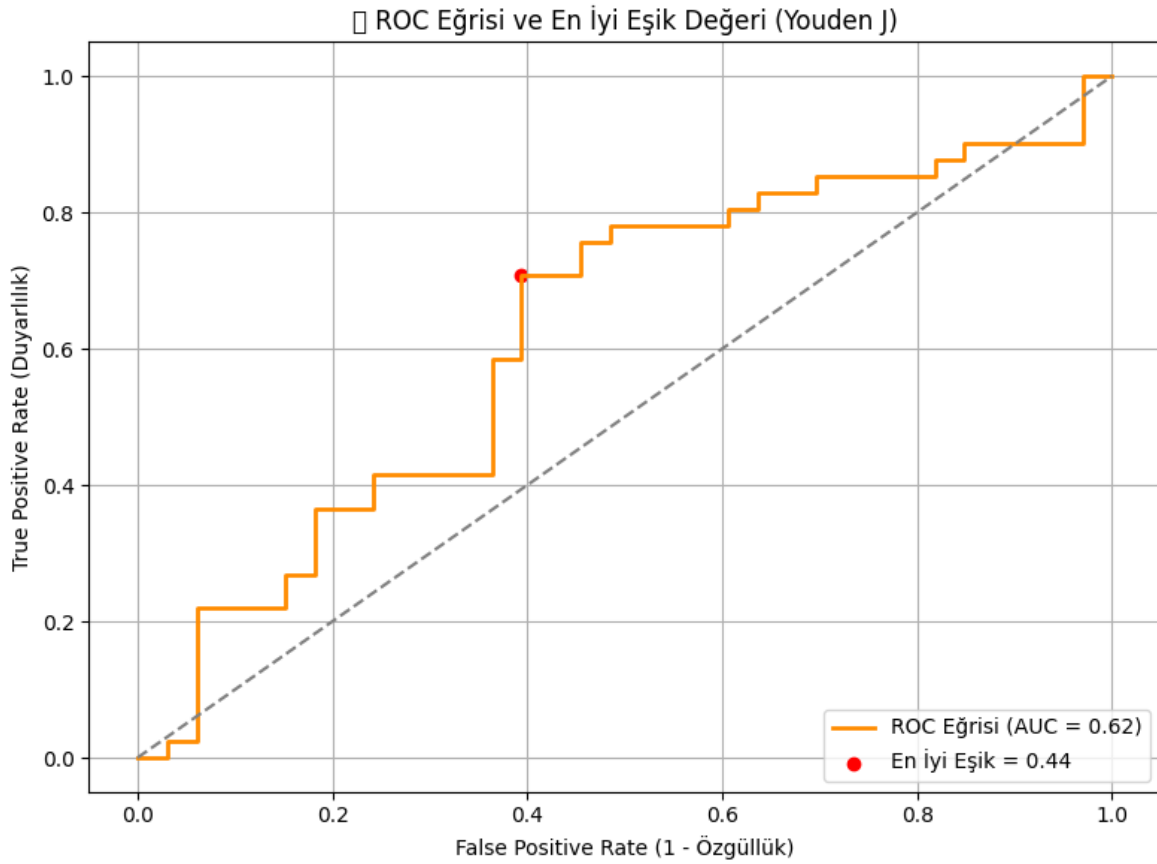
6.5. Eşik Değeri Tespiti ve ROC Eğrisi Analizi

Finansal manipülasyon riskini öngören sınıflandırma modellerinde, doğruluk oranı kadar en uygun eşik değerinin belirlenmesi de büyük önem taşımaktadır. Bu nedenle bu çalışmada, hem 9 hem de 10 değişkenli ve SMOTE uygulanmış iki farklı model için, ROC eğrisi altında Youden J istatistiği hesaplanarak en uygun sınıflama eşik değeri tespit edilmiştir. Elde edilen sonuçlar aşağıda sunulmaktadır:



Şekil 12. Dokuz Değişkenli Model – ROC Eğrisi ve Youden J Eşik Değeri

ROC eğrisi altında hesaplanan 0.63 AUC değeri, modelin sınıflandırma performansının makul düzeyde olduğunu göstermektedir. 0.42 olarak belirlenen eşik değerin üzerinde kalan firmalar “manipülatör” sınıfı (1) olarak değerlendirilmiştir.



Şekil 13. On Değişkenli Model – ROC Eğrisi ve Youden J Eşik Değeri

ROC eğrisi altında hesaplanan 0.62 AUC değeri, modelin sınıflandırma performansının makul düzeyde olduğunu göstermektedir. 0.44 olarak belirlenen eşik değerinin üzerinde kalan firmalar “manipülatör” sınıfı (1) olarak değerlendirilmiştir.

Her iki modelde de eşik değerinin belirlenmesi, modelin manipülatör sınıfını daha güvenilir şekilde tespit etmesini sağlamıştır. Youden J istatistiği ile elde edilen bu eşikler, finansal raporlarda olası hile ihtimali taşıyan firmaları belirleme sürecinde karar vericilere pratik bir araç sunmaktadır.

6. SONUÇ

Bu çalışmada, Türkiye’de Borsa İstanbul’da (BIST) işlem gören imalat sanayi firmalarına ait finansal oranlar kullanılarak, muhasebe hilelerini tespit etmeye yönelik dört farklı lojistik regresyon modeli geliştirilmiştir. Modellerin iki tanesi SMOTE uygulanmamış

orijinal verilerle, diğer iki tanesi ise SMOTE uygulanarak dengelenmiş veri setleriyle oluşturulmuştur. Firma grupları, nakit akış tablosu yayımlayıp yayımlamama durumlarına göre 9 ve 10 değişkenli olarak ayrıştırılmış; böylece veri yapısına uygun farklı modelleme yapıları kurgulanmıştır.

Tablo 2. Model Karşılaştırmaları

Model	Accuracy	Recall (1)	F1 (1)	AUC	Youden Eşik
9 Değişkenli – SMOTE’siz	0.64	0.12	0.20	0.63	—
9 Değişkenli – SMOTE’li	0.62	0.50	0.59	0.63	0.42
10 Değişkenli – SMOTE’siz	0.63	0.12	0.19	0.62	—
10 Değişkenli – SMOTE’li	0.62	0.48	0.57	0.62	0.44

SMOTE uygulanmayan modeller, yüksek doğruluk oranları sunsalar da bu performans büyük ölçüde çoğunluk sınıfı (temiz firmalara) odaklanmalarından kaynaklanmaktadır. Özellikle azınlık sınıf olan manipülatör firmaların doğru sınıflandırılma oranı (recall) oldukça düşüktür (%12). Bu durum, modelin asıl amacı olan hileli firmaları tespit etmede yetersiz kaldığını göstermektedir.

Öte yandan, SMOTE algoritması ile dengelenmiş veri setleri ile oluşturulan modeller, manipülasyon sınıfı için çok daha tatmin edici duyarlılık (recall) ve F1 skoru değerlerine ulaşmıştır. Bu, SMOTE’un yalnızca veri dengesizliğini azaltmakla kalmayıp, aynı zamanda modelin gerçek risk gruplarını tanıma becerisini artırdığını ortaya koymuştur.

Ayrıca, ROC eğrisi altında elde edilen AUC değerlerinin her iki nihai modelde (SMOTE uygulanmış) 0.60’ın üzerinde olması, modellerin istatistiksel olarak anlamlı sınıflama performansına sahip olduğunu göstermektedir. Youden J istatistiği ile belirlenen eşik değerleri (9 değişkenli için 0.42, 10 değişkenli için 0.44), karar vericiye somut bir sınırlama sunmakta ve riskli firmaların belirlenmesinde rehberlik etmektedir.

Bu çalışma, finansal tablo manipülasyonlarının tespitinde kullanılan istatistiksel modellerin başarısının, sadece modelleme tekniğine değil, aynı zamanda kullanılan verinin yapısal özelliklerine de bağlı olduğunu ortaya koymuştur. Özellikle azınlık sınıf olarak

tanımlanan manipölatör firmaların sayıca düşük olması, standart sınıflandırma modellerinde bu firmaların doğru şekilde tanımlanmasını zorlaştırmaktadır.

Bu bağlamda, Model çıktılarında AUC, F1 skoru ve recall gibi metriklerde elde edilen gelişmeler, özellikle kamu gözetim otoritesi, bağımsız denetim firmaları ve sermaye piyasası düzenleyicileri için kritik öneme sahiptir. Bu kapsamda, algoritmik risk skorlarının mevzuata entegre edilmesi, yüksek riskli firmaların erken tespitine katkı sağlayabilir. Denetim süreçlerine yapay zekâ destekli ön tanılama sistemleri entegre edilerek denetim kaynaklarının daha verimli kullanılması mümkündür.

Çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle analiz yalnızca Borsa İstanbul'daki imalat firmalarıyla sınırlıdır. Gelecek çalışmalar, farklı sektörleri kapsayacak şekilde modelin genellenebilirliğini test etmeli ve alternatif makine öğrenmesi algoritmalarının performansını değerlendirmelidir. Ayrıca, şirket içi kurumsal yönetim yapısı, yönetici değişimleri ve piyasa tepkileri gibi finansal olmayan göstergelerin modele entegre edilmesi, daha kapsamlı risk değerlendirmeleri yapılmasına olanak sağlayabilir.

Sonuç olarak, bu çalışma muhasebe hilelerinin tespiti alanında mevcut literatürdeki yöntemleri Türkiye'deki imalat sanayi firmaları bağlamında uygulayarak, veri temelli analizlerin bu alandaki potansiyel katkılarını ortaya koymuştur. Çalışma, geliştirilen modellerin sağladığı bulgular doğrultusunda, finansal rapor manipölasyonlarının erken dönemde belirlenmesine yönelik yöntemsel katkı olarak, veri dengesizliği sorununa karşı SMOTE destekli regresyon modellerinin Türkiye bağlamında etkinliğini göstermektedir.

KAYNAKLAR

- Benligiray, S. - Onay, A. (2021). “Beneish modelinin Türkiye'ye uyarlanması: Tespit başarımını geliştiren bir uygulama”, Muhasebe ve Finansman Dergisi, (91), 511–528.
- Beneish, M. D. (1999). “The detection of earnings manipulation”. Financial Analysts Journal, 55(5), 24–36.
- Fındık, H. - Öztürk, E. (2016). “Finansal bilgi manipülasyonunun Beneish Modeli yardımıyla ölçülmesi: BIST imalat sanayi üzerine bir araştırma”. İşletme Araştırmaları Dergisi, 8(1), 483-499.
- Hołda, A. (2020). “Using the Beneish M-score model: Evidence from non-financial companies listed on the Warsaw Stock Exchange”. Investment Management & Financial Innovations, 17(4), 389–398.
- Kara, S. - Sakarya, Ş. - Aksu, M. (2016). “Beneish Modeli ile kazanç manipülasyonunun tespit edilmesi: BIST şirketleri üzerine ampirik bir uygulama”. Uluslararası Alanya İşletme Fakültesi Dergisi, 16(3), 85–102.
- Kara, S. - Özcan, P. (2020). “Muhasebe manipülasyonlarında yapay sinir ağlarının önemi ve bir uygulama”. Muhasebe ve Denetime Bakış, 20(60), 155–176.
- Kara, S. - Toraman, Ö. E. - Toraman, A. M. (2023). “Kazanç yönetimi ve muhasebe uygulamaları”. Bursa: Ekin Yayınevi.
- Küçüksözen, C. (2004). “Finansal bilgi manipülasyonu: nedenleri, yöntemleri, amaçları, teknikleri, sonuçları ve IMKB şirketleri üzerine ampirik bir çalışma (yayımlanmamış doktora tezi)”. Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Öcal, N. - Atasoy, Y. - Öcal, Ö. (2017). “Muhasebe bazlı kazanç manipülasyonunun tespitinde kullanılan Beneish Modeli'nin test edilmesi”. 21. Finans Sempozyumu, Balıkesir.
- Tekin, E. (2017). “2010-2014 Yılları arasında Türkiye'de halka açık şirketlerde manipülasyon üzerine Beneish Modeli ile ampirik çalışma (yayımlanmamış doktora tezi)”. Başkent Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Tepeli, Y. - Kayıhan, B. (2016). “Muhasebe manipülasyonunun Beneish Modeli ile tespit edilmesi: BIST gıda maddeleri sanayi sektöründe bir uygulama”. Yönetim ve Ekonomi Araştırmaları Dergisi, 9(1), 39–60.
- Uğurlu, M. (2011). “Finansal tablolardaki hile riskinin belirlenmesi: yapay sinir ağı modeliyle bir bankada uygulama (yayımlanmamış doktora tezi)”. Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü, Kütahya.
- Varıcı, İ. - Er, B. (2013). “Muhasebe manipülasyonu ve firma performansı ilişkisi: İMKB uygulaması”. Ege Akademik Bakış, 50(579), 79–92.

Yavaş, M. - Güran, A. - Uysal, M. (2020). “COVID-19 veri kümesinin SMOTE tabanlı örnekleme yöntemi uygulanarak sınıflandırılması”. Avrupa Bilim ve Teknoloji Dergisi, (Special Issue), 258–264.

www.stablebread.com. (2025). “How to use the Beneish M-Score to detect earnings manipulation”. Erişim adresi: <https://www.stablebread.com>

Ek 1 : Lojistik Regresyon Modelleme Kodları

```

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

# Excel dosyasını oku
df = pd.read_excel("MODELLEME TÜM İMALAT 9 DEĞİŞKENLİ.xlsx")

# Bağımsız ve bağımlı değişkenleri ayır
X = df.drop(columns=["ŞİRKETLER", "YIL", "1/0"])
y = df["1/0"]

# Eğitim / test seti ayır
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Model kur ve eğit
model = LogisticRegression(max_iter=1000)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

# 4. Lojistik regresyon modeli oluştur
model = LogisticRegression(max_iter=1000)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

# Değerlendirme
print(" Doğruluk Oranı:", accuracy_score(y_test, y_pred))
print(" Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
print(" Classification Report:\n", classification_report(y_test, y_pred))

# Katsayıları DataFrame olarak hazırla
coef_df = pd.DataFrame({
    "Değişken": X.columns,
    "Katsayı": model.coef_[0]
}).sort_values("Katsayı", ascending=False)

# Katsayıları görselleştir
plt.figure(figsize=(10, 6))
sns.barplot(x="Katsayı", y="Değişken", data=coef_df, palette="coolwarm")
plt.title(" Lojistik Regresyon Katsayıları (Tüm İmalat)")
plt.axvline(0, color="black", linestyle="--")
plt.tight_layout()

```

Ek 2 : Korelasyon Analizi İçin Kullanılan Python Kod Yapısı

```
#korelasyon analizi ve görselleştirme
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Excel dosyasını oku (örnek olarak bir dosya adı verdim)
df = pd.read_excel("1.KORELASYON HESAPLAMA.xlsx")

# Korelasyon için sadece sayısal rasyoları al
X = df.drop(columns=["ŞİRKETLER", "YIL", "1/0"]) # hedef değişken ve açıklayıcı olmayanlar çıkarılır

# Korelasyon matrisini oluştur
correlation_matrix = X.corr()

# Korelasyon matrisini yazdır
print("\n🔍 Korelasyon Matrisi:\n")
print(correlation_matrix)

# Korelasyon matrisini görselleştir (heatmap)
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5)
plt.title("Rasyolar Arası Korelasyon Isı Haritası", fontsize=14)
plt.xticks(rotation=45, ha='right', fontsize=10)
plt.yticks(fontsize=10)
plt.tight_layout()
plt.show()
```