

T.C.
Marmara Üniversitesi
Sosyal Bilimler Enstitüsü

BANKACILIKTA KREDİ DERECELENDİRME
MODELİ

Yüksek Lisans Tezi

Murat Ceylan

Danışman: Doç. Dr., Enes Eryarsoy

İSTANBUL - 2022

TEZ ONAY SAYFASI



ÖZET

BANKACILIKTA KREDİ DERECELENDİRME MODELİ

MURAT CEYLAN

Bir kredi derecelendirme modeli, her başvuru sahibi veya belirli koşulları karşılayan mevcut her müşteri için temerrüt olasılığına (*probability of default* - PD) sahiptir. Bu, bankaların daha objektif, tutarlı, hızlı ve verimli borç verme kararları almalarını ve daha az idari çalışma ve düşük risk ile kredi operasyonlarını yönetmelerini sağlar. Ayrıca, bankacılık sektörünün yönetim kurumlarının da bir gereğidir. Bu amaçla oluşturulan skorkart da açıklayıcı değişkenler ile başvuranın temerrüt olasılığı arasındaki ilişkiyi yansıtır. Uygulamadan sonra, eğitilmiş model, yeni bir başvuruya veya her başvuru sahibine veya belirli koşulları karşılayan mevcut her müşteriye temerrüt olasılığı kullanılarak elde edilen bir skor atar.

Bu tezin amacı, özel banka müşterilerinin kredi derecelerinin geliştirilen bir model ile tespit edilmesi ve bu modelin gerçekleşen sonuçlar ile ne derecede uyumlu olduğunun, bir başka ifade ile model tutarlılığı ve güvenilirliğinin test edilmesidir. Kredi derecelendirmesi için çeşitli son teknoloji makine öğrenimi ve gelişmiş istatistiksel yöntemler kullanılmaktadır ve çalışmamızda, çıktılarda yeterli doğruluk, verimlilik ve yorumlanabilirlik dengesi sağladığı için endüstride en yaygın kullanılan tahmin yöntemlerinden birisi olan lojistik regresyon yöntemi tercih edilmiştir. Tezin bundan sonraki bölümlerinin yapısı şu şekildedir: 2. Bölüm’de kredi derecelendirme kavramına genel bir bakış sunulmuş, tarihsel gelişim süreci, uygulama alanları ve yeni trendler hakkında bilgiler paylaşılmıştır. Bölümün devamında bankacılık sisteminde kredi derecelendirme üzerine yapılmış çalışmalardan bir kesit sunularak yaygın olarak kullanılmış modellerin literatürdeki karşılaştırılmasına yer verilmiştir. Bölümün sonunda ise kredi derecelendirme modellerinin teknik ayrıntıları paylaşılmış, çalışmanın ana odağında yer alan lojistik regresyon modellerinin üzerinde durulmuştur. 3. Bölüm ampirik sonuçların paylaşıldığı bölümdür. Burada çalışmada kullanılan veri sözlüğü paylaşılmış,

devamında bağımlı ve bağımsız deęişkenlerin seçim süreci hakkında detaylar verilmiştir. Modelin geliştirilmesi ve performans analizlerinin ardından model geçerliliğinin test edilmesi ile ampirik bölüm tamamlanmıştır. Sonuç bölümünde ise çalışmanın amacı, kullanılan veri seti ve elde edilen bulgular özetlenmekte; son olarak çalışmaya dair çeşitli kısıtlardan bahsederek gelecek araştırmalar için tavsiyelerde bulunmaktadır.

Anahtar Kelimeler: temerrüt olasılığı (PD), kredi derecelendirme modelleri, kredi puanlama modelleri, lojistik regresyon, kredi riski, risk yönetimi, Türk bankaları.

ABSTRACT

CREDIT RATING MODEL IN BANKING

MURAT CEYLAN

A credit rating model has the probability of default (PD) for every applicant or every existing customer who meets certain conditions. This enables banks to make more objective, consistent, fast and efficient lending decisions and manage their credit operations with less administrative work and lower risk. It is also a requirement of the management institutions of the banking sector. The scorecard created for this purpose also reflects the relationship between the explanatory variables and the default probability of the applicant. After implementation, the trained model assigns a score using the probability of default to a new application or to each applicant, or to each existing customer that meets certain conditions.

The aim of this thesis is to determine the credit ratings of private bank customers with a developed model and to test the consistency and reliability of the model, in other words, to what extent this model is compatible with the actual results. Various state-of-the-art machine learning and advanced statistical methods are used for credit rating, and in our study, logistic regression method, which is one of the most widely used estimation methods in the industry, was preferred because it provides a sufficient balance of accuracy, efficiency and interpretability in the outputs. The structure of the next parts of the thesis is as follows: In the 2nd part, an overview of the concept of credit rating is presented, information about the historical development process, application areas and new trends are shared. In the continuation of the section, a section of the studies on credit rating in the banking system is presented and a comparison of widely used models in the literature is given. At the end of the chapter, the technical details of the credit rating models are shared and the logistic regression models, which are the main focus of the study, are emphasized. Section 3 is the section where empirical results are shared. Here, the data dictionary used in the study is shared, and then the details about the selection

process of dependent and independent variables are given. After the development of the model and performance analysis, the empirical part was completed by testing the validity of the model. In the conclusion part, the purpose of the study, the data set used and the findings obtained are summarized; Finally, recommendations for future research are made by mentioning various limitations of the study.

Keywords: probability of default (PD), credit rating models, credit scoring models, logistic regression, credit risk, risk management, Turkish banks.

ÖNSÖZ

Çalışma sürecinde her türlü yol gösterici olan, olumlu tavrıyla beni cesaretlendiren, bilgi birikimi ile çalışmama farklı açıdan bakmamı sağlayan, bu çalışmanın gerçekleşmesindeki destek ve katkılarından dolayı dolaylı dolaylı danışmanım Doç. Dr. Enes ERYARSOY'a teşekkürlerimi sunarım.

Son olarak aldığım kararları her zaman destekleyen, sadece bu çalışma sürecinde değil her zaman beni cesaretlendiren ve moral veren eşim Sinem Ceylan ve kızım Elif Ece Ceylan'a sonsuz şükranlarımı sunuyorum.

İÇİNDEKİLER

| | |
|---|-----|
| TEZ ONAY SAYFASI..... | i |
| ÖZET..... | ii |
| ABSTRACT..... | iv |
| ÖNSÖZ..... | vi |
| İÇİNDEKİLER..... | vii |
| TABLO LİSTESİ..... | x |
| ŞEKİL LİSTESİ..... | xi |
| KISALTMALAR..... | xii |
| 1. BÖLÜM: GİRİŞ VE MOTİVASYON..... | 1 |
| 2.BÖLÜM: KREDİ DERECELENDİRME KAVRAMI VE BANKALARDA KREDİ DERECELENDİRME ÜZERİNE LİTERATÜR TARAMASI..... | 4 |
| 2.1. Bankalarda Kredi Derecelendirme: Literatür Taraması..... | 7 |
| 2.2. Kredi Derecelendirmede Kullanılan Modeller..... | 15 |
| 2.2.1. Doğrusal Olasılık Modelleri..... | 15 |
| 2.2.2. Lineer Diskriminant Modelleri..... | 16 |
| 2.2.3. Lojistik Regresyon Modelleri..... | 17 |
| 2.2.4. Yeni Nesil Modeller..... | 20 |
| 2.2.4.1. Merton Tabanlı Modeller..... | 21 |
| 2.2.4.2. Sermayenin Risk Ayarlı Getirisi (RAROC) Modeli..... | 21 |
| 2.2.4.3. CreditMetrics™..... | 23 |
| 2.2.4.4. Tarihsel Temerrüt Oranı Yaklaşımı..... | 23 |
| 2.2.4.5. KMV Portfolio Manager Modeli..... | 23 |
| 2.2.4.6. Kredi Portföyü Görünümü (Credit Portfolio View)..... | 24 |

| | |
|--|----|
| 2.2.4.7. CreditRisk+..... | 26 |
| 3.BÖLÜM: UYGULAMA..... | 27 |
| 3.1. Veri Sözlüğü..... | 27 |
| 3.2. Veri Kontrolü ve Kullanıcı Kabul Testleri..... | 29 |
| 3.3. Bağımlı Değişken Veri Hazırlığı ve Seçimi..... | 29 |
| 3.4. Bağımsız Değişken Veri Hazırlığı ve Seçimi..... | 31 |
| 3.4.1. Kayıp Veri Oranlarına Göre Alanların Değerlendirilmesi ve Seçimi..... | 31 |
| 3.4.2. İstatistikî Sınamalara Göre Alanların Değerlendirilmesi ve Seçimi (Gini, Korelasyon)..... | 33 |
| 3.4.2.1. Gini Katsayısı..... | 33 |
| 3.4.2.2. Korelasyon Analizi..... | 34 |
| 3.5. Modelin Geliştirilmesi..... | 35 |
| 3.5.1. Weight of Evidence (WoE)..... | 36 |
| 3.5.2. Information Value (IV)..... | 36 |
| 3.5.3. Lojistik Regresyon..... | 37 |
| 3.5.4. Skorlama..... | 37 |
| 3.6. Model Performans Değerlendirmesi ve Validasyonu..... | 40 |
| 3.6.1. Model Performansı..... | 40 |
| 3.6.1.1. ROC Eğrisi ve Gini Katsayısı..... | 40 |
| 3.6.2. Model Validasyonu..... | 43 |
| 3.6.2.1. Hata Matrisi (Confusion Matrix)..... | 43 |
| 3.6.2.2. Popülasyon İstikrarlılık Gösterisi (PSI)..... | 44 |

| | |
|---|----|
| 3.6.2.3. Temerrüt Oranı (Default Rate) Dağılımı | 46 |
| 3.6.2.4. Herfindahl-Hirschman Endeksi (HHI) | 48 |
| 3.6.2.5. Binom Testi..... | 50 |
| 4.SONUÇ | 53 |
| KAYNAKÇA | 57 |
| EKLER..... | 62 |

TABLO LİSTESİ

| | |
|--|----|
| Tablo 1. Değişken Yapısına Göre Lojistik Regresyon Örnekleri | 19 |
| Tablo 2. Veri Sözlüğü | 28 |
| Tablo 3. Değişken Bazında Kayıp Veri Oranları..... | 31 |
| Tablo 4. Tüm Data – Canlı & Temerrüt Yüzdeleri..... | 35 |
| Tablo 5. Train Data – Canlı & Temerrüt Yüzdeleri..... | 35 |
| Tablo 6. Validasyon Datası – Canlı & Temerrüt Yüzdeleri..... | 35 |
| Tablo 7. Model Değişkenleri ve Açıklamaları | 37 |
| Tablo 8. Skor Aralığı – Tüm Data..... | 39 |
| Tablo 9. Skor Aralığı – Train Data..... | 39 |
| Tablo 10. Skor Aralığı – Validasyon Datası | 40 |
| Tablo 11. Sınıflandırma Tablosu | 41 |
| Tablo 12. ROC Eşik Değerleri | 42 |
| Tablo 13. Model Performans Sonuçları | 43 |
| Tablo 14. Hata Matrisi Eşik Değerleri | 43 |
| Tablo 15. Hata Matrisi Sonucu..... | 44 |
| Tablo 16. Popülasyon İstikrarlılık Göstergesi Eşik Değerleri..... | 44 |
| Tablo 17. PSI Test Sonuçları – Tüm Data..... | 45 |
| Tablo 18. PSI Test Sonuçları – Train Data..... | 45 |
| Tablo 19. PSI Test Sonuçları – Validasyon Datası | 46 |
| Tablo 20. Herfindahl-Hirschman Endeksi Eşik Değerleri | 48 |
| Tablo 21. HHI Testi Sonucu – Tüm Data..... | 49 |
| Tablo 22. HHI Testi Sonucu – Train Data..... | 49 |
| Tablo 23. HHI Testi Sonucu – Validasyon Datası..... | 50 |
| Tablo 24. Binom Testi Sonucu – Tüm Data | 51 |
| Tablo 25. Binom Testi Sonucu – Train Data..... | 51 |
| Tablo 26. Binom Testi Sonucu – Validasyon Datası | 52 |

ŞEKİL LİSTESİ

| | |
|---|----|
| Şekil 1. İki Değişken Arasındaki İlişkinin Grafıksel Görünümü | 34 |
| Şekil 2. Örnek ROC Eğrisi..... | 42 |
| Şekil 3. Dönem Bazlı Temerrüt Oranları Değişimi..... | 47 |
| Şekil 4. Temerrüt Oranı Grafiğı | 48 |



KISALTMALAR

| | |
|--------------|--|
| KOBİ | : Küçük ve Orta Büyüklükteki İşletmeler |
| PD | : Temerrüt Olasılığı (Probability of Default) |
| DR | : Temerrüt Oranı (Default Rate) |
| RAROC | : Sermayenin Risk Ayarlı Getirisi |
| WoE | : Weight of Evidence |
| IV | : Information Value |
| PSI | : Popülasyon İstikrarlılık Göstergesi (Population Stability Index) |
| HHI | : Herfindahl-Hirschman Index |

1. BÖLÜM: GİRİŞ VE MOTİVASYON

İçinde yaşadığımız bilgi çağının ya da bir başka yaygın ifade ile dijital çağın anahtar kelimelerinden olan “veri”, sürekli olarak işlenerek yaşamın her alanında kullanılacak nihai bilginin çıkarılabilmesi adına başvurulmuş en değerli kaynaklardan bir tanesidir. Buna paralel olarak dünyada her an daha fazla kişi, kurum ya da cihaz birbirine bağlanmakta ve bunun sonucunda küresel veri alanının boyutu üssel olarak büyümektedir. International Data Corporation (IDC) tarafından yapılan tahminlere göre her iki yılda mevcut veri stoku iki katına çıkmaktadır ve 2025'te oluşturulmuş ve çoğaltılan veri hacminin 2016'da üretilen verilerin boyutundan on kat daha fazla olacağı beklenmektedir (Reinsel vd., 2017).

Oluşturulan veri hacmi hızla büyüdükçe, bilgiyi elde etmek için verilerin kullanılması, matematik, istatistik ve bilgisayar bilimi için önemli ve önemli bir ilgi alanı haline gelmiştir. Sonuç olarak, veri bilimi disiplinler arası bir alan olarak ortaya çıkmıştır. Finans sektörü, büyük miktarda veri toplayabilen, bunları bir işletme varlığı olarak depolayabilen ve işleyebilen başlıca sektörlerden birisi olması nedeniyle veri bilimi açısından oldukça önemli bir noktadadır. Sadece bankalar değil, diğer finansal kurumlar da verilere büyük önem vermekte ve bu doğrultuda iş birimleri kurarak onları veriye dayalı teknoloji çağına uyum sağlamak üzere eğitmektedir. Pazarlama analizi departmanları kampanya yönetimi, denetim birimleri sahtekarlık tespiti, kredi izleme ve tahsilat birimleri gelecekte potansiyel ödeme sorunlarının öngörülebilmesi amacıyla sistemler oluşturarak iş modellerini bunun üzerine kurmaktadır.

Ticari bankaların aracı finansal kurumlar olarak ana faaliyetlerinden biri, finansal kaynakları artırmak ve bunları üçüncü şahıslara, farklı büyüklükteki şirket ya da diğer ekonomik birimlere borç vermektir. Çeşitli hizmetlerine ilişkin ücret ve komisyonların yanı sıra, bankaların gelirlerinin büyük kısmı sağladıkları kredilerden gelmektedir. banka kredileri temel olarak tüketici kredileri, mikro, küçük ve orta ölçekli işletmelere (KOBİ) verilen krediler ile ticari ve kurumsal krediler olarak kategorize edilebilir. Tüketici kredileri çok çeşitli farklı kredileri içerir ve nakit kredilerden ipoteklere kadar bireylerin

ihtiyaç duyabilecekleri birçok alana hitap eder. Ticari ve kurumsal firmalara verilen krediler, KOBİ'lerden daha büyük şirketler içindir ve örnek olarak proje finansmanı verilebilir.

Her borçlunun geri ödeme yetenekleri bağlamında farklı özellikleri vardır. Bu nedenle bankalar borçlulara ilişkin olarak kendi finansal yükümlülüklerini yerine getirmeme riskini üstlenir ve uygun bir faiz oranı belirlerler. Riskli müşteriler için, banka hiç kredi vermemeyi veya daha yüksek bir faiz oranı belirlemeyi seçebilir. Böylece, kredinin faiz oranı ile bankaya ödünç verilen paranın maliyeti arasındaki fark, yayılma (spread) oranını artırmakta ve bankanın riskini azaltmaktadır. Bu noktada kredi derecelendirme modelleri, muhtemelen kredilerini geri ödeyemeyecek olan adayları veya müşterileri belirlemek için kullanılır. Kredi başvurusunun onaylanması sadece derecelendirme modelinin çıktısına bağlı değildir; birlikişinin görüşü, iş kuralları ve politikaları, bir kredinin değerlendirilmesinde belirleyici diğer faktörlerdir. Kredi derecelendirmesi, geçmişteki verilerin gelecekteki davranışı tahmin etmek için kullanılabilceğini varsayar. Bu nedenle, gelecekteki kredi değeri olan adaylar geçmiş kredi değeri olan başvuru sahiplerine benzeme eğiliminde olacaktır. Bir kredi derecelendirme modeli, her başvuru sahibi veya belirli koşulları karşılayan mevcut her müşteri için temerrüt olasılığına (*probability of default* - PD) sahiptir. Bu, bankaların daha objektif, tutarlı, hızlı ve verimli borç verme kararları almalarını ve daha az idari çalışma ve düşük risk ile kredi operasyonlarını yönetmelerini sağlar. Ayrıca, bankacılık sektörünün yönetim kurumlarının da bir gereğidir. Bu amaçla oluşturulan skorkart da açıklayıcı değişkenler ile başvuranın temerrüt olasılığı arasındaki ilişkiyi yansıtır. Uygulamadan sonra, eğitimli model, yeni bir başvuruya veya her başvuru sahibine veya belirli koşulları karşılayan mevcut her müşteriye temerrüt olasılığı kullanılarak elde edilen bir skor atar.

Bu tezin amacı, özel banka müşterilerinin kredi derecelerinin geliştirilen bir model ile tespit edilmesi ve bu modelin gerçekleşen sonuçlar ile ne derecede uyumlu olduğunun, bir başka ifade ile model tutarlılığı ve güvenilirliğinin test edilmesidir. Kredi derecelendirmesi için çeşitli son teknoloji makine öğrenimi ve gelişmiş istatistiksel

yöntemler kullanılmaktadır ve çalışmamızda, çıktılarda yeterli doğruluk, verimlilik ve yorumlanabilirlik dengesi sağladığı için endüstride en yaygın kullanılan tahmin yöntemlerinden birisi olan lojistik regresyon yöntemi tercih edilmiştir. Tezin bundan sonraki bölümlerinin yapısı şu şekildedir: 2. Bölüm’de kredi derecelendirme kavramına genel bir bakış sunulmuş, tarihsel gelişim süreci, uygulamala alanları ve yeni trendler hakkında bilgiler paylaşılmıştır. Bölümün devamında bankacılık sitemnde kredi derecelendirme üzerine yapılmış çalışmalardan bir kesit sunularak yaygın olarak kullanılmış modellerin literatürdeki karşılaştırılmasına yer verilmiştir. Bölümün sonunda ise kredi derecelendirme modellerinin teknik ayrıntıları paylaşılmış, çalışmanın ana odağında yer alan lojistik regresyon modellerinin üzerinde durulmuştur. 3. Bölüm ampirik sonuçların paylaşıldığı bölümdür. Burada çalışmada kullanılan veri sözlüğü paylaşılmış, devamında bağımlı ve bağımsız değişkenlerin seçim süreci hakkında detaylar verilmiştir. Modelin geliştirilmesi ve performans analizlerinin ardından model geçerliliğinin test edilmesi ile ampirik bölüm tamamlanmıştır. Sonuç bölümünde ise çalışmanın amacı, kullanılan veri seti ve elde edilen bulgular özetlenmekte; son olarak çalışmaya dair çeşitli kısıtlardan bahserek gelecek araştırmalar için tavsiyelerde bulunmaktadır.

2. BÖLÜM: KREDİ DERECELENDİRME KAVRAMI VE BANKALARDA KREDİ DERECELENDİRME ÜZERİNE LİTERATÜR TARAMASI

Kredi derecelendirme, bir devletin, finansal kuruluşun veya bir bireyin, geçmişteki ve günümüzdeki nitel ve nicel olarak gözlemlenebilen verilerini kullanarak, bu ekonomik birimlerin finansal yükümlülüklerini zamanında yerine getirip getiremeyeceklerini tahmin etmek amacıyla oluşturulmuş bir sınıflandırma sistemidir (Yazıcı, 2009). Bir başka ifade ile kredi Derecelendirme, bir finansal birimin mevcut ya da olası borcunu ödeyebilme kabiliyetinin ölçülmesi olarak tanımlanabilir. Kredi Derecelendirme Kuruluşları (KDK) ise, kredi derecelendirme faaliyetlerini yürüten; verdikleri kredi notlarıyla hem borç verenin hem de borçlunun finansal durumları hakkında bilgi veren özel kuruluşlardır. Bu kuruluşlar finansal döngü içerisinde yer alan kuruluşların, geçmişte yaşadıkları ve hâlihazırda içerisinde buldukları ekonomik koşullara ve ekonomilerini etkileyen diğer faktörlere göre kredibilitelerinin puanlandırılmasından ve kredi risklerinin değerlendirilmesinden sorumludur (Al-Sakka R., 2010). Bireysel müşterilerin kredi derecelendirmeleri ise genelde sistemin kendi içerisinde yerine getirilir. Bu derecelendirme sisteminde, kredi için başvuran kişinin geçmişte kullandığı tüm krediler ile ilgili bilgiler; ne kadar kredi kullandığı, ne kadar zamanda ödediği, bunları öderken herhangi bir düzensizlik yapıp yapmadığı; gecikmeler, takibe düşmüş krediler ve bunlarla ilgili bütün bilgiler kullanılmaktadır. Bu derecelendirme sonucu elde edilen olasılıklar veya puanlar, bankalar tarafından karar vermede destek olarak kullanılır. En basit haliyle, bir derecelendirme, iyi ve kötü davranışları tahminlerken ve birbirinden ayırırken kullanılan, istatistiksel olarak açıklayıcı olduğu belirlenen bir grup değişkenden oluşan bir tür değerlendirme aracıdır. Bu noktada bireylere, şirketlere ve devletlere verilen kredi notlarının ayrımı, derecelendirmenin İçsel Derecelendirme ve Dışsal Derecelendirme olmak üzere iki türe ayrılması ile tanımlanmıştır: *İçsel derecelendirme* çoğunlukla bankaların kendi müşterileri ile aralarındaki “güven ilişkisine” dayalıdır. Bankalar, müşterilerinin kredibilitelerini derecelendirmek amacıyla onlardan bir takım doküman

ve veriler isterler. Bu belgelerin standart ve anlaşılır olması beklenir zira bu belgeler aracılığı ile müşterilerin sonraki periyotlar için borçlarını ödeyebilme kapasiteleri belirlenmektedir.. *Dışsal derecelendirme* ise büyük ölçekli firmalar ya da devletler için S&P, Moody's ve Fitch gibi KDK'ların verdiği kredi notlarıdır. Buradaki ana hedef büyük firmaların sermaye piyasasından borçlanabilmesi için gereklilikleri yerine getirip getiremediği hakkındaki bilgileri sağlamaktır (Ayaz, 2016). Dolayısıyla aslında kredi derecelendirme faaliyetleri en küçüğünden en büyüğüne kadar, finansal süreçlerin içerisinde yer alan birçok birey ya da kurum için yapılmaktadır.

Günümüz Kredi derecelendirme kuruluşlarına benzer yapılar, piyasalardaki belirsizliklerin öngörülmesi ve genel anlamda finansal risklerin elimine edilebilmesi ihtiyaçlarına karşılık kurulmuştur. 1800' lü yıllarda ABD' de yaşanan ekonomik çöküş birçok yatırımcının mağdur olmasına sebep olmuş, piyasalarda yatırımcıların risk sorununa yönelik arayışların ortaya çıkmasına yol açmıştır. Kavram olarak kredi derecelendirme, bir tekstilci olan Lewis Pappan' ın, ilk olarak 1837 yılında, müşteri çevresinin ekonomik durumlarını kendi belirlediği istatistiki bir yöntemle tespit etmesi şeklinde ortaya çıkmıştır. (Ban vd., 2003; Yazıcı, 2009) Lewis Pappan, Amerika' da yaşanan krizin etkisi ve ABD' nin doğusu ile batısı arasındaki ekonomik farklılık ve ticaretin yönünün doğudan batıya olması sebebiyle, batıdaki tüccarların borçluluk durumlarını değerlendirmiş ve kredi derecelendirme faaliyetlerinin temelini atmıştır. Tüccarlar Ajansı adıyla bilinen şirketin toptancılardan, tekstilcilerden, tüccarlardan ve sigorta şirketlerinden oluşan takipçilerinin sayısı, 1880 yılında 7000 kişi iken, 1890 yılında 40000 kişiye ulaşmış ve 1900 yılı itibarıyla de 1 milyon kişiyi aşmıştır. Tüccarlar Ajansı 1933 yılında, bir avukat olan ve müvekkillerinin finansal durumları hakkındaki bilgileri tüccarlara satan John Bradstreet' in kurduğu ve 1857 yılında ilk kez ticari rating kitabını yayınlayan Bradstreet Company ile birleşerek Don & Bradstreet adını almıştır. Ardından söz konusu şirket 1962 yılında Moody's Investors Service tarafından satın alınmıştır (Sylla R., 2001). Bununla birlikte ekonomide yaşanan gelişmeler ve piyasalardaki risk ve güven sorunu, benzer derecelendirme faaliyetlerini beraberinde getirmiştir. 1868 yılında Henry Varnum Poor, ilk kez, büyük demiryolu şirketleri için

finansal istatistikler içeren ve şirketlere bağımsız bilgi sağlayan Manuel of the Railroads of the United States' i yayınlamıştır (Setty, G., vd., 2003).

20. yy ile birlikte günümüzde de varlıklarını devam ettiren üç büyük KDK kurulmuştur. 1900 yılında John Moody, bugünkü bilinen şekliyle ilk derecelendirme kuruluşu olan Moody's Investors Service' i kurmuş ve 1909 yılında demiryolu tahvillerine ait reytingleri yayınlamıştır. Poor's Publishing rating çalışmalarına 1916 yılında başlamış, 1922 yılında rating vermeye başlayan Standard Statistics ile 1941 yılında birleşerek Standard & Poor' s ismini almışlardır. 1913 yılında şirket performanlarına dair verilerin analizlerini yayınlayan Fitch, 1924 yılında Fitch Publishing Company olarak rating sektörüne girmiş ve 1997 yılında IBCA ile birleşmiştir (Orhan M., 2011).

Bankaların uyguladığı derecelendirme modellerinde derecelendirme değişkenleri, başvuru sırasında borç verenin sahip olduğu herhangi bir veri kaynağından seçilebilir. Bu değişkenlere örnek olarak demografik bilgiler (yaş, ikamet süresi, posta kodu), mevcut ilişki (müşterilik yaşı, ürün sayısı, ödeme performansı, önceki talepler), kredi bürosu verileri (diğer banka bilgileri, gecikmeleri, kamu kayıtları), emlak verileri, vb. verilebilir. Başvuru sahibinin toplam puanı, o başvuru sahibinin derecelendirmesinde mevcut olan her bir değişkene ilişkin nitelik puanlarının toplamını ifade etmektedir. Elde edilen risk skoru bilgisi, kârı maksimize edecek ve zararı da en aza indirecek yeni değerlendirme stratejileri geliştirmek için kullanılabilir. Yüksek riskli adaylar için stratejilerden bazıları aşağıdakilerden bir veya bir kaç olabilir:

- Risk seviyesi çok yüksekse, kredileri / ürünleri azaltmak
- Daha düşük limitli ürünler tayin etmek
- Başvuru sahibinden daha yüksek değerli teminat talep etmek
- Bir kredi için daha yüksek kar payı oranı belirlemek
- Sigorta poliçelerinden daha yüksek primler almak
- Potansiyel riskli başvuruları “yakın izleme listesine” koymak

Buna karşılık, yüksek puan alan başvuru sahiplerine daha yüksek oranlar ve kredi limitleri verilebilir, altın veya platin kartlar gibi premium ürünlere ya da şirket tarafından sunulan ek ürünlere yükseltme teklif edilebilir.

Bu nedenle, risk puanlaması, alacaklılara, elde edilen bilgilere dayanarak tutarlı ve objektif karar verme fırsatı sunar. İş bilgisi ile birleştirildiğinde, öngörücü modelleme teknolojileri, risk yöneticilerine risk yönetimi süreci üzerinde ek verimlilik ve kontrol sağlar (Siddiqi, 2012). Diğer yandan daha büyük ölçekteki kredi derecelendirme faaliyetleri (devletleri ya da büyük şirketleri konu alan) hem kantitatif hem de kalitatif değerlendirmelere dayalı olabilir. Kredi Derecelendirme Kuruluşları finansal birimlerin kredi riskini, borçlunun finansal yükümlülüklerini zamanında yerine getirme olasılığını, borcunun ana para ve faizini ödeme kabiliyet ve isteğini; çoğunlukla kendilerine ait bir metodoloji ile, nitel ve nicel (qualitative and quantitative) değerlendirmeler ile ölçerler.

2.1. Bankalarda Kredi Derecelendirme: Literatür Taraması

Her bir borçluya farklı bir PD atayarak borçlunun risk düzeyini yani temerrüt olasılığını tahmin etmek artık birçok bankada yaygın olarak kullanılmaktadır. PD, belirli bir karşı tarafın yükümlülüklerini yerine getiremeyeceğini veya karşılamaya istekli olmayacağını belirtir. PD'nin yanlış tahmin edilmesi, makul olmayan derecelendirmeye, ve finansal araçların yanlış fiyatlandırılmasına sebep olmaktadır bu nedenle son küresel finansal krizin nedenlerinden birisi de yanlış derecelendirme faaliyetleri olmuştur (Gurný ve Gurný, 2013; Saka and Orhan, 2018).

Bu nedenlerden dolayı, finansal konuların temerrüt olasılığının tahmininin uzun süredir finansal araştırma alanında güncel ve ilgi çeken bir konudur. Finans dünyasında temerrüt olasılığını tahmin etmek için birçok farklı model kullanılmaktadır. En yaygın kullanılanlar arasında genellikle kredi derecelendirme (ya da puanlama) modelleri olarak bilinen modeller yer almaktadır ve bu modeller daha sonraki bölümlerde detaylı olarak açıklanmıştır. Bu modeller temelinde, bir şirketin (ya da finansal birimin) ana finansal göstergelerini girdi olarak kullanan ve her birine, temerrüt tahminindeki göreceli önemini yansıtan bir ağırlık veren çok değişkenli modellerdir. Bu modeller genellikle üç kategoriye ayrılır: a) diskriminant analizleri (doğrusal, ikinci dereceden); b) regresyon modelleri

(doğrusal, logit, probit); c) yeni nesin modeller (yapay sinir ağları). Bu modellerin çıkış noktası, çalışma prensipleri ve uygulama alanları ile ilgili ayrıntılı çalışmalar arasında, Resti ve Sironi (2007), Green (2008) ile Engelmann ve Rauhmeier (2006) gibi çalışmalar gösterilebilir.

Bu tez, bankalarda kredi puanlama modelleri analizine odaklanmaktadır ve dolayısıyla literatür analizi de bu doğrultuda yapılmıştır. Bankalar tarafından homojen bir grup borçlunun kredibilitesini değerlendirmek için kullanılan en yaygın yöntem kredi puanlamasıdır (Emel vd., 2003). Kredi derecelendirme (ya da puanlama) modellerinin altında yatan teknikler, 1930'larda Fischer (1936) ve Durand (1941) gibi yazarlar tarafından geliştirilmiş olsa da bu modellerin gelişimi ve uygulama açısından yaygınlaşmasındaki hızlı yükselme, 1960'larda, başarılı ve başarısız firmaların eşleştirilmiş örneklerini karşılaştırarak, iş başarısızlıklarının finansal verilerle tahmin edilebileceğini gösteren Beaver (1967) ve Altman (1968) çalışmalarıyla yaşanmıştır¹. Kredi puanlama, geliştirilen en eski finansal risk yönetimi araçlarından biridir ve bu açıdan 1950'lerde ABD'li perakendeciler ve postayla sipariş veren firmalar tarafından kullanımı, yatırım portföylerinde var olan riski yönetmek ve çeşitlendirmek için portföy analizinin ilk uygulamalarıyla eş zamanlıdır. Ek olarak, kredi puanlaması, tüketici davranışına ilişkin verilerin en erken kullanımlarından biri olduğu için veri madenciliğinin (data mining) atası olarak da kabul edilebilir (Thomas vd., 2002). Bu alandaki önemli katkılar arasında McFadden (1976), Altman vd., (1981), Queen and Roll (1987), Shumway (2001) ve Balthazar (2006) gösterilebilir. Halihazırda önerilen kredi puanlama modellerinin büyük çoğunluğu, temel olarak finansal kurumların temerrütlerinin göreceli olarak düşük sıklıkta meydana gelmesi ve tüm verilerin kamuya açık olmaması nedeniyle

¹ Muhtemelen, kredi puanlamasına çoklu diskriminant analizi uygulamanın en erken kullanımı, araba kredisi başvurularını inceleyen Durand'ın (1941) çalışmasıdır. Kurumsal iflas tahmininde iyi bilinen bir uygulama, kurumsal mali tablolardan sekiz değişkenden alınan beş finansal orana dayalı ilk operasyonel puanlama modelini geliştiren Altman'a (1968) aittir ve bu çalışma, finansal oranların doğrusal bir birleşimi olan bir Z-Skorunu üretmiştir.

finansal olmayan kurumlardan oluşan bir örneklemeden türetilmiştir Gurný ve Gurný (2013). Bununla birlikte, sağlıklı finansal kurumlar için temel faktörleri belirlemeye yönelik, finansal tabloların kullanıldığı Peresetsky ve Karminsky (2008) ile Gurný ve Gurný (2010) gibi önemli çalışmalar da bulunmaktadır.

Kredi derecelendirme, herhangi bir bankanın işleyişinde, gelecekteki finansal sonuçlar üzerinde gerçek bir etkisi olan kilit bir unsur olarak kabul edilir. Başlangıçta, kredi puanlama kavramı sadece basit uzman puanlama kartlarına atıfta bulunsa da, teknolojinin gelişmesiyle birlikte basit puanlama, yüksek derecede ilerleme ile tahmine dayalı modellere dönüşmüştür. Kredi puanlama, kredi başvurusunda bulunan bir müşterinin kredi riski tutarını değerlendirme yöntemi olarak tanımlanmıştır ve bu doğrultuda puanlama sistemleri, geçmiş verileri ve çeşitli istatistiksel teknikleri kullanarak, başvuranların yükümlülüklerinin zamanında ödenmesini önemli ölçüde etkileyen bireysel özelliklerinin etkisini tespit için dizayn edilmiştir. Bu yöntem, bankanın başvuranları veya borçluları farklı bir risk seviyesinde gösteren gruplara kolayca sınıflandırabileceği belirli bir nokta (kesme noktası) sağlar. Etkili bir puanlama modeli oluşturmak için, geçmiş veriler ve borcun geri ödenmesini tahmin etmede yardımcı olan seçilmiş özellikler analiz edilir. İyi yapılandırılmış bir puanlama modeli, gelecekte yükümlülüklerini sözleşme şartlarına uygun olarak yerine getirecek müşterilere daha yüksek, bununla ilgili zorluk yaşayan borçlulara ise daha düşük bir derecelendirme vermelidir (Mester, 1997; Kil vd., 2021).

Kredi puanlamanın sadece borç verenler için değil, borç alanlar için de birçok avantajı vardır. Çoğu zaman, kredi başvurularıyla ilgili olarak kredi kararları vermek için puanlama modelleri kullanılır ve dahası, kredi limitleri belirlenirken, mevcut hesapları yönetirken ve müşterilerin finansal güvenilirliğini tahmin ederken yeni nesil teknikler de kullanılmaktadır (Fensterstock, 2005). Kredi derecelendirme sisteminin önemli bir avantajı, objektifliği ve kredi karar verme sürecindeki ölçülebilir kriterleri standart hale getirebilmesidir. Modelin nesnelligi, borç alanlar arasında ırk, cinsiyet, din ve diğer ayrımcı faktörlerden bağımsız olarak her birine eşit muamele edilmesini garanti altına alır. Ayrıca, bu sistem, bankanın puanlama algoritmasında uygulaması olan hemen hemen her

kredi politikasını kullanmasına izin veren yüksek esneklik ile karakterize edilir (Wysinski, 2013). Doğru oluşturulmuş bir puanlama modeli, aynı zamanda kişinin bankadaki sorunlu kredilerin sayısını önemli ölçüde sınıflandırmasını sağlayarak kredi portföyünün iyileşmesine ve gelişmesine katkı sağlar. Kredi puanlama sisteminin kullanımında olduğu gibi, değer kaybı olan kredilerin sayısı, kredi kararının geleneksel bir müşteri doğrulama yöntemiyle, genellikle çeşitli niteliklere sahip çalışanlar tarafından verildiği duruma göre daha düşüktür (Blöchlingle ve Leippold, 2006). Ponicki'ye (1996) göre, kredi puanlaması, bankacılık sektörü genelinde kredileri değerlendirmek için standart bir teknik, işlemleri yürütmek için etkin bir yol sağlar ve müşterilerden alacakların tahsil edilmesini kolaylaştırır. Puanlama modelleri aynı zamanda basit bir başvuru süreci sunarak kredi kuruluşlarının müşterilerine de fayda sağlar, müşterilerin kredi başvurularına zamanında yanıt almalarını ve ihtiyaç duyduklarında krediye erişmelerini sağlar.

Kredi puanlama modeli, finans ve bankacılıkta araştırma modellemesinin en başarılı uygulamalarından biridir ve sektördeki puanlama analistlerinin sayısı sürekli artmaktadır. Ancak kredi puanlaması, bazı finansal türevlerin fiyatlandırılması veya portföy analizi ile aynı karakteristiğe sahip olmadığından, konuyla ilgili literatürün sınırlı olduğu söylenebilir. Bununla birlikte, kredi puanlaması, son kırk yılda tüketici kredisindeki olağanüstü büyümeye izin vermede hayati önem taşımaktadır. Doğru, tutarlı ve otomatik olarak çalışan bir risk değerlendirme aracının varlığı, tüketici kredisini verenlerin kredi hacimlerini bugünkü şekliyle genişletmelerine ciddi bir katkı sağlamıştır (Bailey, 2001; Mays, 2001; Siddiqi, 2006).

Aynı zamanda, literatür, kredi itibarının incelenmesi için bir araç olarak kredi notunun kullanımına ilişkin belirli sorun ve tehditlere de işaret etmektedir. Bir bankada kullanılan puanlama sistemi genellikle hızlı bir şekilde eski haline gelir ve bu da yönetimi sürekli olarak yeni borçluların verilerini girmeye ve kullanılan puanlama kartını sistematik olarak güncellemeye zorlar. Kredi derelendirme yönteminin kendisinin de yeni kredi alan kişi sayısını olumsuz etkileyebileceği belirtilmektedir. Bunun nedeni, kredi geçmişi olmayan müşterilerin olumsuz bir kredibilite değerlendirmesi alma ve bunun sonucunda kredi vermeyi reddetme olasılığının yüksek olmasıdır (Kil vd., 2021). Ayrıca, daha düşük

bir borçlunun notu, kredinin maliyetini artırarak krediyi daha az cazip hale getirir (Kinda ve Achonu, 2012). Samreen ve Zaidi (2012), puanlama modellerinin kullanımıyla ilişkili risklere rağmen, ticari bankaların kredi sürecinin bir parçası olarak kredi değerliliğini değerlendirmek için bir puanlama modeli kullanmalarının önemini belirtmektedir. Bankalar için bu modeli benimseyerek sorunlu kredi sayısını azaltmak, içerdiği problemlere rağmen güncel ve geçerli bir yöntem olmaya devam etmektedir. Bu öneriler, kredi sürecinde kredi notu kullanmanın avantajlarının, uygulamalarının olumsuz yönlerinden daha ağır bastığını vurgulamaktadır (Kil vd., 2021).

Kredi notu kullanımının kredilendirme politikası parametreleri ve finans dışı sektörden alacak portföyünün kalitesi üzerindeki etkisi birçok ülkede araştırma konusudur. Çerçeve vd. (2001), araştırmalarında borçluların kredi riskini tahmin etmek için puanlama modellerinin kullanılması sayesinde, bireysel müşterilere (aynı zamanda küçük ve orta ölçekli işletmelere) verilen kredi sayısının da arttığını göstermiştir. Berger vd. (2005) makalelerinde, kendi puanlama modellerini kullanan bankaların aksine, üçüncü şahısların puanlama sonuçlarını satın alan bankalarda kredi vermede daha fazla artış görülebileceğini belirtmektedir. Deyoung vd. (2006), kendi kredi riski değerlendirme modellerini kullanan bankalardan alınan kredilerin daha düşük kalitede olduğuna dair ek kanıtlar sunmaktadır. Blöchlinge ve Leippold (2006) kredi notu kullanmanın ekonomik faydaları üzerine yaptıkları çalışmada, derecelendirme modelinin kârlılığının diğer rakiplerin ayrıştırma gücüne de bağlı olduğunu göstermiştir. Bankalar tarafından kredi sürecinde "zayıf" puanlama modellerinin kullanılması, daha fazla "kötü" borçluyu cezbeder ve bu nedenle, düşük ayırım gücüne sahip bir rakip, daha düşük gelirlere ve daha büyük kayıplara neden olur. Böyle bir durumda, borçluları zayıf ve güvenilmez bir kredi derecelendirme sistemi temelinde değerlendiren bankalar, kredilerini geri ödemeyenlere daha fazla kredi verirken, aynı zamanda kredibilite açısından sağlam potansiyel borçluları reddederler (Abdou vd., 2007).

Koh vd. (2006), makalelerinde farklı puanlama modellerini bir araya getirmek için veri madenciliği tekniklerinin uygulanmasını tartışmış ve örneklemiştir, kredi riskini tahmin etmede daha etkili olan tek bir modelin oluşturulmasına olanak tanınmasına rağmen, bu

sürecin de kendi içerisinde tuzaklarının olduğunu belirtmişlerdir. Bunlardan biri, kredi riski tahmin sonuçlarının dayandığı bireysel modellerin gösterdiği sonuçları önemli ölçüde aşacak şekilde birkaç puanlama modelini birleştirmenin zor ve bazen imkansız olması olarak gösterilmiştir.

Bir ticari kredi bilgi paylaşım sistemi kullanmanın bu kuruluşlarla ilişkili kredi riski üzerindeki etkisini belirlemeye çalışan Dierkes vd. (2013) tarafından yürütülen çalışmanın sonuçları, bankaların paylaşım sistemleri tarafından kredi bilgilerinin uygulanmasının, hem toplu olarak hem de bireysel firmalar için temerrüt tahminlerinin doğruluğunda önemli bir artışı etkilediğini belirtmektedir. Kerage ve Jagongo'nun (2014) bulguları da potansiyel borçluların kredi geçmişine ilişkin mevcut bilgilerin kredi şirketleri tarafından kullanımı ile bankacılık sektörü performansı arasında pozitif bir ilişki olduğunu kanıtlamaktadır. Bankaların, müşterilerinin kredi geçmişleri hakkında bilgi paylaşımını artırdıkça, sorunlu kredi oranlarında bir azalma da dahil olmak üzere performansları artmıştır. Giannopoulos (2018), KOBİ'lerin temerrütünü tahmin etmede yapay kredi puanlama modellerinin etkinliğini incelemiştir. Bir sinir ağı modelini ve bir karar ağacı modelini banka tarafından uygulanan kredi puanlama modeliyle karşılaştırarak, bankanın modelinin kredilerin temerrütünü tahmin etmede nispeten daha kötü bir performansa sahip olduğunu bulmuştur. Berger vd. (2011) tarafından sunulan ampirik analiz, topluluk bankaları tarafından tüketici kredisi puanlamasının kullanılmasının, bankaların kredi portföyünün kalitesinde önemli bir değişiklik olmaksızın küçük işletme kredilerinde bir artış ile ilişkili olduğunu ileri sürmektedir. Ayrıca, küçük işletme kredileri için topluluk bankaları tarafından kredi puanlarının kullanılması, karlılıkta zamanla ılımlaşan ilk düşüşle ilişkilidir.

Literatürde, bankacılık sektörünün kendi içerisinde kredi derecelendirme modellerinin oluşturulmasında lojistik regresyon başta olmak üzere diskriminant analizi, regresyon analizi, probit analizi gibi istatistiksel teknikler yaygın olarak incelenmiştir ve birbirleriyle karşılaştırılmıştır (Boyes vd., 1989; Steenackers ve Goovaerts, 1989; Banasik vd., 2001; ve Sarlija vd., 2004). Diğer yandan “iyi” veya “kötü” olmak üzere sadece iki grup (ikili) müşteri kredisinin kullanılması da kredi puanlama modellerine uygundur ve

halen kredi derecelendirme uygulamalarında yaygın olarak kullanılmaktadır (Banasik vd., 2001; Lee vd., 2002; ve Kim ve Sohn, 2004). Öte yandan, üç grup tüketici kredisinin kullanılması da kredi puanlama modellerinde sınıflandırma yaklaşımlarından biri haline gelmiştir. Bu çalışmalardan bazıları “iyi”, “kötü” ya da “reddedildi” şeklinde kategorilendirme yaparken (Steenackers ve Goovaerts, 1989), diğerleri “iyi”, “yoksul” ya da “kötü” şeklinde gruplamaya gitmiştir (Sarlija vd., 2004). Kredi puanlama modellerinin oluşturulmasında diğer istatistiksel tekniklerin yanı sıra probit analizi (Guillen ve Artis, 1992; Banasik vd., 2003) de yaygın olarak kullanılmıştır. Genel olarak, kredi puanlama modellerinin oluşturulmasında kullanılan genel bir “en iyi istatistiksel teknik/yöntem” yoktur. Bu yöntem, kredi türü, veri setinin karakteristiği, müşteri profili, dışsal ekonomik çevre ve daha birçok farklı parametreye göre değişiklik gösterebilir. Diğer bir ifade ile neyin en iyi olduğu, problemin ayrıntılarına, veri yapısına, kullanılan verinin ve hedef kitlenin özelliklerine bağlıdır (Hand ve Henley, 1997). Diğer yandan farklı teknikler arasında karşılaştırma yapan çalışmalarda sinir ağları ve bulanık mantık algoritmaları gibi yeni nesil istatistiksel tekniklerin geleneksel olanlardan daha iyi olduğu iddia edilmiştir fakat ortalama doğru sınıflandırma yüzdesi veya isabet oranı açısından farklı istatistiksel teknikler arasında belirgin bir fark yoktur (Blochlinger ve Leippold, 2006; Hoffmann vd., 2007; Abdou vd., 2007). Eddy ve Engku Abu Bakar (2017), geliştirilen yapay zeka tabanlı modeller ve makine öğrenme yöntemlerin arasından bazı tekniklerin (özellikle yapay zekaya dayalı olanlar) daha iyi kredi puanlama modellerinin uygulanmasına izin vermesine rağmen yine de istatistiksel olarak hala geride kaldığını, zira bu tekniklerin henüz kullanıcı dostu olmadığını ve kullanımlarının halen zor olduğunu belirtmektedir. Abdou vd. (2007), Mısır örnekleme özelinde probit, lojistik ve diskriminant analizlerini karşılaştırmış ve modellerin sıralamasının karar kriterine göre değiştiğini, en yüksek ortalamanın doğru sınıflandırma oranını kullanan lojistik regresyon yönteminden elde edildiğini fakat en düşük tahmini yanlış sınıflandırma maliyetini için diskriminant analizinin en optimum yöntem olduğunu belirtmiştir. Dolayısıyla yazarlar önemli olanın, bankanın karar vericisinin bakış açısına bağlı olduğunu, model seçiminin yanlış sınıflandırma maliyetinden kaçınma veya bir pazar araştırmasından güçlü bir müşteri

kitlesine sahip olduklarını bilme durumlarına göre deđiŒeceđini belirtmiŒlerdir. Yakın dnem alıŒmaları arasında,  farklı modeli karŒılaŒtırmasıyla ne ıkan Gurny and Gurny (2013), Abdou (2007)'de olduđu gibi  tr kredi puanlama modeli (logit model, probit model ve lineer diskriminant analizi) aracılıđıyla PD'nin ABD ticari bankalarını tahmin etme olasılıđını araŒtırmıŒtır. alıŒmada 298 Amerikan ticari bankası rneđinden PD tahmini iin  model treten yazarlar, daha sonra en uygun modeli belirlemek iin sonraki 100 Amerikan ticari bankasını kontrol grubu olarak belirlemiŒtir. Yazarlar her  tahmin modeli ile de finansal kriz zamanında elde edilen veri setinden tahminde bulunmuŒlar ve bu nedenle kullanımlarının da piyasa evriminin bu zel aŒamasıyla sınırlı olduđunu belirtmiŒlerdir. Bu aıdan, tahmini modeller, olduka kısa bir temerrt tahmini (1-2 yıl) ile tek dnemli tahmin modelleridir. Bu sınırlamalar dikkate alındıđında yazarlar, elde edilen sonulardan tahmin edilen logit modelinin bankaların temerrt tahmini iin en uygun model olduđunu belirtmiŒlerdir. Diđer yandan daha yksek temerrt olasılıđının tespiti iin en aıklayıcı gce sahip finansal gstergelerin hesaplanma tarihi ile finansal kuruluŒlar iin iflas tarihi arasındaki gecikme hakkında ampirik bilgi eksikliđi bulunduđu belirtilmiŒtir. Buradaki alıŒmalarda yapılan analizlerin, kabul edilen ve bazılarının daha sonra kt olduđu ortaya ıkan kredi baŒvurularına dayandıđını belirtmek gerekir. Elbette, belirtilen puanlama modelleri kullanılmıŒ olsaydı, baŒlangıta reddedilen bazı baŒvurular kabul nerilerine yol aabilirdi. Bununla birlikte, bankanın kendi alternatif puanlama sistemleri, istatistiksel modeller tarafından reddedilecek olan baŒvuruların hepsini olmasa da birođunu filtreleyebilir.

Sonu olarak, istatistiksel tabanlı puanlama modeli teknikleri, bankalar iin daha yaygın olarak tercih edilen yntemler olmaya devam etmektedir. Bu teknikler arasında lojistik regresyon halen en popler olanıdır (Kil vd., 2021) ve bu tezde de uygulama blmnn ana yntemi olarak tercih edilmiŒtir. Sonraki blmde, yukarıda literatrdeki uygulamalarından kısaca bahsedilen kredi derelendirme modellerinin teknik ayrıntıları baŒlanılmaktadır. Tezin ana ynteminin lojistik regresyon olması nedeniyle, bu yntemde ayrıntıya dada fazla girilmiŒtir.

2.2. Kredi Derecelendirmede Kullanılan Modeller

Bu bölümde yukarıda bahsedilen kredi derecelendirme modellerinde kullanılan yöntemlerin detayları ve değişken seçme süreci kısaca anlatılmaktadır. Kredi puanlama modelleri, borçlunun temerrüde düşme olasılığını tahmin etmek için kullanılan en yaygın araçlardır. Finansal ve ekonomik göstergelere ağırlık verilmesi prensibi ile çalışırlar ve ağırlıklar, borçlunun temerrüdü tahmininde bu göstergelerin önemini ifade eder.

2.2.1. Doğrusal Olasılık Modelleri

Doğrusal olasılık modelleri, 0 ile 1 arasında bir değer aldığı kabul edilen ve bu doğrultuda tanımlanan bağımlı değişkenin test edildiği bir regresyon modelidir. Bu yöntem, PD'yi etkileyen unsurlar ile temerrüt olayının arasında doğrusal bir bağlantı olduğunu varsaymaktadır (Koç, 2018). Bu modellemede bankanın kredi kullandırdığı müşterilerin geçmiş yıllardaki bilgileri modelin kurulumunda bağımsız değişken olarak ele alınır. Kredi verilen birimlerin geri ödeme noktasında temerrüde düşme ihtimallerini açıklama özelliğindeki bağımsız değişkenler ve onların katsayıları, doğrusal regresyon ile tespit edilmeye çalışılarak ilgili birimin temerrüde düşüp düşmeyeceği tahmin edilir. Doğrusal olasılık modeli kurulurken en başta veri seti belirlenir ve ardından bankalar tarafından önceki senelerde verilen kredilere ait veriler göz önünde bulundurularak müşteriler iki gruba bölünür. Birinci grup müşteriler, kredileri geri ödemede sorun yaşamayan, düzenli ve finansal açıdan sağlam müşterilerden oluşur. İkinci gruptaki müşteriler ise kredisini düzenli olarak geri ödemeyen, temerrüde düşmüş kişilerdir. Modelin teknik olarak kurulurken temerrüde düşmüş olan kişiler $Y_i = 1$, temerrüde düşmemiş olanlar ise $Y_i = 0$ şeklinde tanımlanır. Bu noktada Y , temerrüde düşme ihtimalini göstermektedir ve olasılık teorisi gereği 0-1 aralığında bir değer alır. Modelin tamamlanması adınan bağımlı değişkeni açıklauabileceği düşünülen n tane açıklayıcı (ya da bağımsız) değişken belirlenir ve bu bağımsız değişkenler, müşterinin demografik bilgilerinden, gelir düzeyinden, borçluluk durumundan, piyasa istihbarat verilerinden, karlılık, likidite, cari oran, stok devir hızı, nakit dönüş süresi gibi birçok finansal rasyodan oluşabilir. En nihayetinde bağımsız değişkenler de atandıktan sonra doğrusal regresyon

analizi vasıtası ile ve en küçük kareler yöntemi ile (1) numaralı denklem kullanılarak model tahmin edilir (Korkmaz, 2004).

$$Y_i = \beta_0 + \sum \beta_j X_{ij} + e_i \quad \text{Denklem 1.}$$

Burada, β_j regresyon katsayılarını, e_i hata terimini ve X_{ij} bağımsız değişkenleri temsil eder. İfade edilen denklem ile müşteriye ait X_{ij} datası gözlemnebebildiği takdirde PD hesaplanabilir. Ancak burada aktarılan doğrusal olasılık modelinin en önemli dezavantajlarından bir tanesi tahmin edilen PD 0-1 aralığı dışına da taşabilmesidir. Bu noktada, eğer PD oranı negatif veya %100'den büyük olarak hesaplanırsa daha gelişmiş olan logit modelin kullanılması tavsiye edilmektedir. Çoğunlukla benzer bir yaklaşıma ve sürece dayanan logit modelde PD %0 ile %100 arasında bir değer olarak tahmin edilir.

2.2.2. Lineer Diskriminant Modelleri

Lineer diskriminant analizi, seçili tüm açıklayıcı değişkenlerden, bağımlı değişkendeki çeşitliliği en iyi gösterebilen tek bir bağımsız değişken üretmektedir ve kredi risklerinin hesaplanması amacıyla, Altman tarafından 1968'de bu model için hazırlanan ilk Z-skor modeli (2)'deki şekliyle formülize edilmiştir:

$$Z = (1,2)X_1 + (1,4)X_2 + (3,3)X_3 + (0,6)X_4 + (1,0)X_5 \quad \text{Denklem 2.}$$

Yukarıdaki formülasyon ile tahmin edilen kredi skoru, Altman tarafından oluşturulmuş olan bir ölçekle karşılaştırılmakta ve alınan skora göre “çok iyi”, “iyi”, “orta”, “düşük”, “zayıf” şeklindeki bir borçlu gruplandırması yapılmaktadır. Bağımlı değişken şeklinde tanımlanan z, temerrüt riski için temel ölçüdür ve burada borçlunun geçmiş zaman bilgilerinden elde edilen açıklayıcı değişkenlerin değerine ve bu değişkenlerin geçmişteki temerrüde düşme ve düşmeme durumlarına olan etkilerine göre ağırlıklandırılmış öneminden etkilenir. Bağımlı değişken olan z arttıkça, müşterinin PD ihtimali azalmaktadır. Bu modele doğrultusunda $Z > 1,81$ ise, düşük PD ve $Z < 1,81$ ise, yüksek PD şeklinde karar verilir (Kutman, 2001).

2.2.3. Lojistik Regresyon Modelleri

Lojistik regresyon, ikili bir sonucun (sıfır veya bir) olasılığının, veri setindeki bir dizi potansiyel tahmin değişkeniyle ilişkili olup olmadığını analiz eden ve oldukça yaygın olarak kullanılan bir istatistiksel modelleme tekniğidir. Bu model, logaritmik dönüşümler sonucu doğrusal forma kavuşturulan ve bağımlı değişkeni kategorik olan ve bu değişkeni “yapay değişken” olarak adlandırılan bir yöntemdir. Bu yöntemde bağımlı ve bağımsız değişkenler arasında logit bir ilişki olduğunu varsayılır ve dolayısıyla lojistik regresyonun doğrusal olmayan modeller üretebildiği, bunun da doğrusal modellerin içerdiği dezavantajlara yönelik ciddi bir çözüm anlamına geldiği söylenebilir. Bu yöntem daha önce de bahsedildiği gibi oldukça yaygın olarak kullanılmaktadır ve bunun nedenleri şu şekilde özetlenebilir:

- İlk olarak kategorik bağımlı değişken seçilebilirken açıklayıcı değişkenlerin sürekli ya da süreksiz olmalarına yönelik kısıtlama yoktur. Dolayısıyla model, her iki şekilde de kullanıma açıktır.
- Bu yaklaşımda kurulan model doğrusal hale getirilir ve bu da tahminlemeyi kolaylaştırır.
- Yöntem, SPSS, STATA ve SAS gibi birçok istatistik paket programında oldukça kolay bir şekilde yönetilebilir.
- Yöntem, açıklayıcı değişkenlerin olasılık fonksiyonlarının ne şekilde dağılımları gerektiği ile ilgili bir önşart sunmaz, dolayısıyla varsayım ihlali riski düşüktür.
- Lojistik regresyonda, doğrusal modellerde olduğu gibi negatif olasılıkla karşılaşma ihtimali yoktur, zira olasılık dağılımı 0-1 arasında tahmin edilir.
- Lojistik regresyon yöntemi, bağımsız değişken ile bağımlı değişken arasındaki ilişkinin doğrusal olmasını şart koşmaz. Dolayısıyla bu değişkenler arasındaki bağlantı üssel ya da polinomik yapıda olabilir.

Bağımlı değişkenin yapay olduğu modellerin temeli olan doğrusal olasılık modelinin sorunlarını aşmak için bazı dönüşümlerle doğrusal hale gelen fonksiyonlara ihtiyaç vardır. Lojistik dağılım fonksiyonunu incelemek için iki kategorili bağımlı değişken ele alınabilir (Baydemir, 2014).

$$P_i = E(Y = 1/X_i) = 1/(1 + e^{-(\beta_0 + \beta_1 X_i)}) \quad \text{Denklem 3.}$$

Burada (3) numaralı denkleme $Z_i = \beta_0 + \beta_1 X_i$ dönüşümü yapıldığında $P_i = 1/(1 + e^{-Z_i})$ elde edilir. Bu eşitlik, lojistik dağılım fonksiyonu olarak adlandırılır ve eşitlikte Z_i , $-\infty$ ile $+\infty$ arasında bir değer alırken P_i 'nin de 0 ile 1 arasında olması beklenir. Z_i 'nin dolayısıyla da X_i 'nin P_i ile ilişkisinin doğrusal olmadığı kurulan modelden anlaşılabilir. Bu iki karakteristiğin sağlanabilmesi ile doğrusal olasılık modelinin problemlerinin üstesinden gelinmiş olur (Baydemir, 2014). Bu eşitliklerden hareketle, tanımlı bir X^* değerinde belirli bir olayın gerçekleşme olasılığı lojistik dağılım fonksiyonu ile tahmin edilir. Lojistik modeldeki β_1 katsayısı, X 'te meydana gelen 1 birimlik artışın L 'de sebep olacağı değişimi, yani “*istenilen olayın gerçekleşme lehine fark oranındaki artışı*” gösterir (Baydemir, 2014). Dolayısıyla burada P_i yeniden ifade edilirse “gerçekleşme ihtimalinin denklemi” olarak adlandırılır:

$$P_i = \frac{1}{(1 + e^{-Z_i})} \quad \text{Denklem 4.}$$

Eşitliğin tekrar ele alınıp doğrusal hale getirilmesinden hareketle P_i , kategorik değişkene ait istenilen olayın gerçekleşmesi olasılığı anlamına geldiği için olayın gerçekleşmeme olasılığı $1 - P_i$ olarak ifade edilebilir.

$$1 - P_i = \frac{1}{(1 + e^{Z_i})} \quad \text{Denklem 5.}$$

Diğer yandan, olayın gerçekleşme olasılığının gerçekleşmeme olasılığına oranı ise (6) ile gösterilmiştir.

$$\frac{P_i}{1 - P_i} = \frac{1 + e^{Z_i}}{1 + e^{-Z_i}} = e^{Z_i} \quad \text{Denklem 6.}$$

Bu oran, söz konusu olayın meydana gelme ihtimalinin meydana gelmeme ihtimaline fark oranı olarak isimlendirilmektedir. Bu eşitlikte son olarak her iki tarafın doğal logaritması alınır ve lojistik model elde edilir (Baydemir, 2014).

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \ln e^{Z_i} \quad \text{Denklem 7.}$$

Bu denklemde Z_i yerine $\beta_0 + \beta_1 X_i$ yazılırsa $L_i = \ln(e^{\beta_0 + \beta_1 X_i}) = \beta_0 + \beta_1 X_i$ elde edilir. Bu dönüşümün ardından L , hem ağırlıklar hem de değişkenler açısından doğrusallaştırılmış olur ve “ L ”, modelin “logit” olarak adlandırılmasının da sebebidir (Baydemir, 2014). Lojistik regresyon yönteminde modelin uygun olup olmadığı “*model ki-kare*” yöntemi ile, her açıklayıcı değişken için ayrı ayrı anlamlılık testleri ise Wald istatistiği ile kontrol edilir. Lojistik regresyon analizlerinde bilinmesi gereken önemli yabancı terimler bulunmaktadır. Bunlardan ilki olan *Odds*, bir olayın görülme ihtimalinin, görülmemeye ihtimaline oranı iken *oddsratio* (OR), iki *odds*’nin birbirlerine bölünmeleri ile elde edilen orandır ve iki değişken arasındaki ilişkiyi özetler. *Logit* ise hesaplanan *oddsratio*’nun doğal logaritması alınarak elde edilen değerdir. *Oddsratio* asimetriktir fakat doğal logaritması alındığında simetrik hale gelir. Burada aslında logit olarak ifade edilen parametre, doğrusal regresyon analizlerinde yer alan β katsayısının karşılığı olarak bilinir.

Lojistik regresyon yöntemi çoğunlukla bağımlı iki kategoriye sahip bağımlı değişken durumlarında kullanılsa da bağımlı değişkenin ikiden fazla kategoriye sahip olduğu Tablo 1’de yer alan şartlar altında da uygulanabilir.

Tablo 1. Değişken Yapısına Göre Lojistik Regresyon Örnekleri

| Bağımlı Değişkenin Yapısı | Örnek |
|---|-----------------------------|
| Kategori sayısı iki (Binominal) | Temerrüt, canlı |
| İkiden fazla kategorili fakat sırasız (Multinomial) | İşsiz, çalışan, emekli |
| İkiden fazla kategorili fakat sıralı (Ordinal) | Düşük, orta, yüksek gelirli |

Lojistik regresyon analizi yöntemini kullanmadan önce dikkat edilmesi gereken bir varsayım bulunmasa da araştırmalarda genel istatistik kural ve yaklaşımları elbette önemlidir. Bu noktada ilk olarak uygun tüm açıklayıcı değişkenlerin modelde yer alması önem arz etmektedir zira bazı değişkenlerin unutulması ya da gözardı edilmesi hata teriminin büyümesine neden olan sebepler arasındadır. Bunun yanı sıra gereksiz tüm bağımsız değişkenler de dışlanmalı ve modele dahil edilmemelidir. Herhangi mantıksal bir ilişki dahilinde olmayan açıklayıcı değişkenin modele dahil edilmesi, hem daha

karmaşık bir yapı sunar hem de modelin yorumlanması zorlaştırır. Diğer yandan anlamlı bir istatistiksel sonuç, ilişkinin hatalı yorumlanmasına ve nedensellik olmadığı halde varmışçasına çıkarım yapılmasına neden olabilir. Üçüncü olarak aynı gözlemden bir kez veri alınmalı ve tekrarlayan ölçümlerden kaçınılmalıdır. Dördüncü olarak bağımsız değişkenlerde ölçüm hatası küçük olmalıdır ve bunu azaltmak için eksik veriden olabildiğince kaçınılmalıdır. Çünkü yüksek hatalar, yanlış katsayı tahminlerine ve modelin yetersiz yapılandırılmasına neden olur. Beşinci olarak bağımsız değişkenler arasında çoklu bağlantı bulunmamalı, diğer bir ifade ile açıklayıcı değişkenlerin birbirleriyle ilişkili olmamasına özen gösterilmelidir. Altıncı olarak uç değerler elimine edilmelidir zira aşırı değerler tahmin sonucunu önemli derecede yönlendirebilir. Yedinci olarak örneklem büyüklüğü yeterli olmalıdır, az sayıda gözlem güvenilir tahmin yapılmasını zorlaştırır. Son olarak beklenen ve gözlenen varyanslar arasındaki fark küçük olmalıdır çünkü bağımlı değişkenin beklenen varyansı ile gözlenen varyansı arasındaki büyük bir fark, yine modelin kalitesizliğini ve eksik tahmin gücünü gösterir. Varyansların homojen olmaması problemi genellikle örneklemin rastgele seçilmesinden veya araştırılan konuda yeterli akademik bilgi olmamasından kaynaklanmaktadır (Sümbüloğlu ve Sümbüloğlu, 2007).

2.2.4. Yeni Nesil Modeller

Yakın tarihte özellikle bankacılık ve finans sektöründe kredi derecelendirme yöntemlerinde meydana gelen yenilik ve gelişmeler, yeni nesil modelleri beraberinde getirmiştir. Özellikle belirsiz durumlar karşısında rastlanılan riskler için yeni nesil yöntemlerle belirlenecek sermaye düzeyi, olası kötü senaryolar karşısında kurumun ödeme gücüne düşmemesi için oldukça kritiktir. Bu yöntemler arasında, Merton tabanlı modeller, Sermayenin Risk Ayarlı Getirisi Modeli (RAROC), CreditMetrics™, Tarihsel Temerrüt Oranı Yaklaşımı, KMV Portfolio Manager Modeli, Kredi Portföyü Görünümü (Credit Portfolio View) ve CreditRisk+ gibi yaklaşımlar öne çıkmakta ve hem ulusal hem de uluslararası düzeyde kabul görmektedir. Yeni nesillere dair Koç (2018)'de faydalı bir özet sunulmuştur.

2.2.4.1. Merton Tabanlı Modeller

Robert Merton tarafından 1974 yılında önerilen Merton modeli, Black-Scholes-Merton opsiyon fiyatlama yöntemlerini kullanmakta ve temerrüt riski ile firmanın sermaye yapısı arasında bir ilişki sağlamaktadır. Bu modeller genel olarak temerrüt olasılığı, sermaye yapısındaki dalgalanmalar ve varlık fiyatları gibi makroekonomik göstergelere odaklanmaktadır (Engelman and Pahm, 2020). Dayanak varlıklar üzerindeki bir alım opsiyonu, opsiyonun kullanım fiyatına ulaştıktan sonra varlıklar üzerinde talepte bulunan bir arayan ile aynı özelliklere sahiptir. Bu durumda, opsiyonun kullanım fiyatı, firmanın yükümlülüklerinin defter değerine eşittir. Varlıkların değeri firmanın yükümlülüklerini karşılamaya yetmiyorsa, alım opsiyonuna sahip hissedarlar opsiyonlarını kullanmazlar ve firmayı alacaklılarına bırakırlar. Genellikle, kredi vadesi daha uzun olsa bile, yalnızca bir yılı analiz eden tek dönemlik bir modelde uygulanır (Crouhy vd., 1999). Merton modelleri PD'yi rasyolara göre değil, varlık fiyatına göre tahmin etmektedir (Kavcıoğlu, 2011). Moody's KMV tarafından geliştirilen çeşitli yazılım ürünleri Merton veri tabanını kullanarak risk ölçümleri yapmaktadır (Canbolat ve Gümrah, 2015).

2.2.4.2. Sermayenin Risk Ayarlı Getirisi (RAROC) Modeli

Piyasa verilerini baz alarak kredi riskini ölçmek amacıyla kullanılan yaygın modellerden bir tanesi de RAROC (Sermayenin Risk Ayarlı Getirisi - Risk Adjusted Return on Capital) modelidir ve bu model uygulayıcılar tarafından geliştirildiği 70'lere dayanmaktadır. RAROC modeli ilk olarak Bankers Trust tarafından uygulanmıştır ve günümüzde de dünyanın birçok gelişmiş finansal kuruluşları tarafından kendi bünyelerine entegre edilerek kullanılmaktadır. RAROC modelinin temelinde yatan mantık, bankanın verdiği krediden beklenen ve $(net\ faiz + ücretler)/kredi\ tutarı$ şeklinde gösterilebilecek varlık getirisini değerlendirmek yerine, beklenen faiz ve ücretlerin kredi riskine oranını göz önünde bulundurmadan geçer. Diğer bir ifade ile bankaların kredi gelirlerini tahsis edilen kredi tutarına oranlamak yerine, kredi riskini gösteren başka bir ölçüme oranlamanın daha tutarlılığı olacağı beklenir (Saunders, 1999). Bu yaklaşım (8) numaralı denklem ile gösterilmiştir.

$$RAROC = \frac{\text{Krediden elde edilen bir yıllık gelir}}{\text{Kredi riski ya da risk sermayesi}} \quad \text{Denklem 8.}$$

RAROC katsayısının benchmark sermaye maliyetine göre yeterince yüksek olması, kredinin tahsisine onay verilmesi anlamına gelir. Bunun yanı sıra, mevcuttaki aktif kredilerin RAROC katsayısının benchmark sermaye maliyetinin altına inmesi, bankanın krediyi kârlı hale getirecek şekilde bir ayarlama yapmasını gerektirir (Korkmaz, 2004). Bu noktada belirtmek gerekir ki, RAROC yönteminde hesaplamamanın zorluğu, kredi riskinin ölçümünün zorluğundan kaynaklanır. Öte yandan bu yöntem işleyişinde, bir finansal varlığın ortalama süresi tahmin edilebilir. Bir kredinin değerindeki değişim " $\Delta L/L$ ", kredinin süresine ve faiz oranındaki değişimin büyüklüğüne " $\Delta R/(1 + R)$ " bağlıdır ve bu ilişki, kredinin değerindeki değişimin formülü olarak (9) ile gösterilmiştir.

$$\frac{\Delta L}{L} = -\Delta L * \Delta R / (1 + R) \quad \text{Denklem 9.}$$

Benzer yaklaşım, sermaye riski ya da kayıp miktarının formülize edilmesinde de uygulanabilir fakat burada bir fark, faiz oranındaki şok yerine, kredi kalitesindeki şokun dikkate alınmasıdır.

$$\Delta L = -DL * L * \Delta R / (1 + R) \quad \text{Denklem 10.}$$

Bu denklemde ΔL , sermaye riski ya da kayıp miktarı, DL kredinin süresi, L risk miktarı ya da kredi büyüklüğü ve son olarak $\Delta R/(1+R)$ ise kredi risk priminde beklenen maksimum değişim anlamına gelmektedir. Bu yaklaşımda kredi süresi ve miktarı kolayca tahmin edilebilir ancak takip eden yıl kredi risk primindeki en büyük değişimin kestirilmesi zordur. Bu noktada krediye dair risk bilgileri kamu ile paylaşılmadığından şirket tahvillerinin işlem gördüğü tahvil piyasası, risk primlerinin bulunması için incelenmektedir. S&P tarafından yayınlanan kredi derecelendirme sınıflarından (AAA, AA, A, BBB, BB, B, CCC, CC, C ve D) hangi gruba girdiği belirlenen borçlu tarafın sonraki adımda o kredi derecelendirme sınıfında bulunan tüm tahvillerin bir önceki seneye göre gerçekleşen risk primi değişimleri ele alınacaktır. Risk primi en çok değişen tahvilin

risk prim deęiřimi, olabilecek en kötü senaryonun incelenebilmesi için hesaba katılmaktadır.

2.2.4.3. CreditMetrics™

Temerrüt durumunda doğabilecek tahmini zararlar ölçülmesi amacıyla yaygın olarak kullanılan modellerden bir tanesi CreditMetrics™ Modelidir. 1997 yılında riske maruz deęer (VaR) ile birlikte dizayn edilmiş olan bu model (Canbolat ve Gümrah, 2015) günümüzde Chase Grubuna ait eski JPMorgan Bankası'nın yapılarından bir tanesi olan Risk Metrics™ Group'un bir alt markasıdır. Modelin yaygınlığı ve uygulamadaki başarısı, model geliştiricilerinin uyguladıkları kredi riski metodolojisini tamamen şeffaf bir şekilde paylaşımlarından gelmektedir (Hacıyeva, 2019). CreditMetrics™ yönteminin çalışma prensibinde bir borçlunun geçmiş verileri baz alınarak bir kredi notundan dięerine geçiři olasılıęını (ρ) gösteren matrisin hesaplanması yatmaktadır. Bu matriste her satırdaki olasılıkların toplamı 1 olduğundan dolayı tüm olasılık durumları yansıtılır. Geçiři olasılık tabloları, Fitch, Moody's ve S&P gibi önde gelen kredi derecelendirme kurumları tarafından yayınlanmaktadır (Allen ve Powell, 2011).

2.2.4.4. Tarihsel Temerrüt Oranı Yaklařımı

Tarihsel Temerrüt Oranı Yaklařımı geçmiş verilere öngörüde bulunmaktadır. Bu model, derecelendirme skorlarına göre gruplandırılmış tahvillerin temerrüt oranları ve vadeye kalan süreleri baz alınarak kurulmaktadır. Bu model vasıtası ile bankalar, aynı ya da bezner karakteristięe sahip olan varlıkların tarihsel temerrüt oranını analiz ederek bu doğrultuda bir politika izleyebilirler (Öker, 2007).

2.2.4.5. KMV Portfolio Manager Modeli

Kredi portföy yönetiminin KMV modeli ilk kez 1993 yılında detaylandırılmıştır. Bu model, kredi riskinin nicelleřtirilmesi için çeřitli modellerin geliştirilmesine olanak saęlamıştır. Bu modelde krediler için etkin sınır çizilmek istenir ve portföy içindeki farklı borçluların kredilerin daęılımında optimal aęırlığı (X_i) tahmin edilmeye çalışılır. Bu amaçla başlangıçta üç deęiřken belirlenir ki bu deęiřkenler sırasıyla borçlu i 'ye verilmiş kredinin beklenen getirisi R_i , borçlu i 'ye verilen kredinin riski D_i borçlu i ve j 'ye verilen kredilerin temerrüt risklerinin birbirleri ile olan korelasyonu p_{ij} olarak tanımlanır ve bu

kapsamda bir kredinin beklenen getirisi (10)'daki, beklenmeyen kaybı ise (11)'deki gibi hesaplanır (Saunders, 1999).

$$R_i = AIS_i - E(L_i) = AIS_i - [EDF_i * LGD_i] \quad \text{Denklem 11.}$$

$$\delta_i = UL_i = \delta(D_i) * LGD_i = [EDF_i * (1 - EDF_i)V_1 * LGD] \quad \text{Denklem 12.}$$

Bu yaklaşımda **kredinin getirisi (R_i)**, "*krediden kazanılan yıllık ücretler*" + *bankanın fonlama maliyeti ile borçlunun ödediği faiz oranı arasındaki fark*" hesaplanılarak bulunur (all -in spread - AIS). Hesaplanan değerden, krediden beklenen kayıp $E(L_i)$ çıkarılır ki beklenen kayıp, borçlunun beklenen temerrüt sıklığının (Expected Default Frequency - EDF_i) temerrüt durumunda yaşanacak kaybın şiddeti (Loss given default - LGD_i) ile çarpımına eşittir. **Kredinin Riski (δ_i)** ise kredinin beklenen değeri ile temerrüt nedeniyle kaybın şiddetinin (LGD_i) çarpımı etrafındaki temerrüde düşme oranının volatilitisini göstermektedir. Kredinin beklenmeyen kaybı (unexpected loss - UL_i) ise temerrüt oranının değişimi ile LGD_i 'nin çarpılmasıyla hesaplanır. Hesaplanan beklenmeyen kayıp (UL_i) değeri, kredinin riskine eşittir. Son olarak KVM Portfolio Manager modeli, hisse senedi getirilerinin getiri bileşenlerini ele alarak iki müşteri arasındaki gözlemlenmemiş temerrüt riski korelasyonunu ölçmek amacı ile, o hisse senetlerinin önceki yıllarda birlikte hareketlerine dayalı bir korelasyon (ρ_{ij}) hesaplar. Bu modele göre eğer temerrüt korelasyonları 0.002 ile 0.15 arasında hesaplanmış ise, temerrüt oranı düşük olarak kabul edilir ki bu oldukça mantıklı bir çıkarımdır zira oldukça büyük iki şirketin, Apple ve Microsoft gibi, birlikte iflas etme ihtimalleri oldukça düşüktür. Gelecek yılda aynı anda her iki şirketin varlıklarının toplam tutarı, borçlarının tutarından daha düşük olmalıdır. Burada hesaplanan korelasyon katsayılarının düşük olması, bankalar için de oldukça iyi bir netice anlamına gelecektir zira tahsis edilen kredilerin birçok borçlu arasında paylaştırılması ile portföy riskini azaltılabilir.

2.2.4.6. **Kredi Portföyü Görünümü (Credit Portfolio View)**

Kredi portföyü görünümü modeli Wilson (1997) tarafından geliştirilmiş ve McKinsey Consulting Company tarafından önerilmiştir. McKinsey, yalnızca temerrüt riskini ölçen ayrı bir çok dönemli ekonometrik model önermektedir. Temerrüt olasılıkları

makroekonomik deęişkenlere baęlıdır. Varsayılan olasılıkların bir logit işlevi tarafından üretildięi varsayılır. Makroekonomik deęişkenler her ülke için belirtilir. Bu baęımsız deęişkenlerin her birinin, 2. dereceden (AR(2)) bir otoregresif modeli takip ettięi varsayılır (Derbali ve Hallara, 2013). Ortak ülke veya sektör temerrüt oranları kalibre edilir ve Choleski ayrıştırması, tüm segmentler arasında ortak temerrüt olasılıklarının dağılımının simülasyonunu sağlamak için kullanılır. Ardından, ülke/endüstri segmentlerinin her biri için benzersiz koşulsuz Markov geçiş matrisi çıkarılır. Son adım, kümülatif temerrüt ve geçiş olasılıklarının dağılımını üretmektir. Model, gerçek ve yakın zamanda gözlemlenen makroekonomik deęişkenlere baęlı olarak, sonraki dönemler için varsayılan oranların bir dağılımını, makroekonomik deęişkenlerin belirli bir otomatik gerileme sürecini takip etmesinden yararlanarak üretir. Bu yaklaşım, derecelendirme deformasyonundan kaynaklı zararları da içeren CreditRisk+ metodolojisini de geliştirmektedir. Bu yaklaşım bir yandan da Monte Carlo simülasyonları gerektirmesi nedeniyle karmaşık teknik odaklar da gerektirir (Witzany, 2017).

Kredi Portföyü Görünümü modelinde temerrüt ihtimalleri bir logit fonksiyon olarak tasarlanmıştır. Modelin açıklayıcı deęişkeni tarihsel ve güncel makroekonomik deęişkenler, baęımlı deęişkenlerini ise ülke ve sektör kredi derecelendirme endeksi verileri oluşturur. Bu bağlamda, endekse göre spekülatif dereceye sahip borçlular için PD hesaplanmaktadır (Gönenç, 2002). Bu makroekonomik model, belirli bir sektör j 'de ve spekülatif bir borçlunun t döneminde PD_{jt} olarak ifade edilen temerrüt olasılığını açıklayan standart bir yöntemdir ve (12) olarak gösterilmiştir.

$$Y_{jt} = \sum \beta_j X_{jt} + \varepsilon_{jt} \quad \text{Denklem 13.}$$

X_{jt} belirlenen sektör için temel makroekonomik deęişkenleri temsil eder ve β_j ise regresyon katsayısıdır. Dięer yandan ε_{jt} hata terimlerinin baęımsız oldukları ve normal bir dağılımla eşit varyansla dağıldıkları ve aynı zamanda bir kovaryans matrisi ile karakterize edilen ortak normal dağılıma sahip oldukları varsayılır. Varsayımların simülasyon için kullanılması, model kalibre edildikten sonra tarihsel makroekonomik deęişkenlere ve temerrüt oranlarına dayanır (Witzany, 2017).

2.2.4.7. *CreditRisk+*

CreditRisk+ Credit Suisse Finansal Hizmetler tarafından geliştirilmiştir. Riske maruz değer çerçevesi çizmek isteyen CreditMetrics'in aksine, CreditRisk+ belli bir miktarın üzerindeki kayıpların karşılanması için bankanın sermaye gereğini hesaplamaya odaklanarak, krediden beklenen kaybı ve bu kayıpların olasılık dağılımını tahmin etmeye çalışmaktadır (Saunders, 1999). CreditRisk+'da temerrüt, firmanın temellerine bağlı değildir, ancak stokastik yoğunluk parametrelili bir Poisson dağılımını izleyen dışsal bir değişken olarak modellenmiştir. Bu yoğunluk parametresi, küçük bir zaman aralığında varsayılan oranı veya temerrüt olasılığını temsil eder. Bunun, bir Gama dağılımını takip eden bilinmeyen bir "ekonomik faktör" tarafından yönlendirildiği varsayılır, dolayısıyla varsayılan oran da Gama dağılımıdır. Temerrüt halinde, alacaklı, borç miktarından tahsilat oranı düşülerek eşit miktarda zarara uğrar. CreditMetrics'in aksine, borcun değeri CreditRisk+'da modellenmez. Portföyleri ve borçlular arasındaki korelasyonları ele almak için bunlar alt portföyler halinde sınıflandırılır. Her alt portföy, belirli bir ekonomik faktörden etkilenir. Bu sınıflandırma, aynı alt portföydeki borçlular arasında bağımsızlığın varsayılmasına izin verir. Bu, modelin matematiğini kolaylaştırır. Ayrıca, her bir alt portföy içinde, borçlular kredi risklerine göre gruplara ayrılmaktadır. Her bantta, her bir kredi riskinin boyutu ayarlanır, böylece her bant ortak bir maruziyet V_i ile karakterize edilir (Korkmaz, 2004).

3. BÖLÜM: UYGULAMA

Ampirik olarak hazırlanan skorkarne modellemesi kapsamında SAS programı kullanılarak bir çalışma yapılmış ve literatürde de sıkça kullanılan «Lojistik Regresyon» yöntemi kullanılmıştır. Burada amaç derecelendirme modelinde yer alacak girdi (input) kombinasyonunun seçimi sağlayabilmektir. Lojistik regresyon metodu ile temerrüt durumunu tahmin etmekte kullanılacak inputların optimal kombinasyonu ve önem derecesi yani ağırlıkları belirlenir.

Tüm modelleme kapsamında izlenen adımlar aşağıdaki gibidir:

- 1- *Veri sözlüğünün oluşturulması*
- 2- *Veri kontrolü ve kullanıcı kabul testleri*
- 3- *Bağımlı değişken veri hazırlığı ve seçimi*
- 4- *Bağımsız değişken veri hazırlığı ve seçimi*
- 5- *Modelin Geliştirilmesi*
- 6- *Model performans değerlendirmesi ve validasyonu*

3.1. Veri Sözlüğü

Modelleme aşamasından önce yapılması gereken ilk adım bir veri sözlüğü oluşturmaktır. Veri sözlüğünün oluşturulmasındaki amaç, modeli etkileyebilecek tüm alanların oluşturulup açıklamaları ile birlikte dokümente edilmesi ve verilerin içerik olarak genel bir tanımının oluşturulmasıdır. Böylece değişken açıklamaları ve nasıl hesaplandıkları açıkça ifade edilmiş olunur. Oluşturulan 48 değişkenin bulunduğu veri sözlüğü Tablo 2'deki gibidir:

Tablo 2. Veri Sözlüğü

| Değişken Adı | Açıklaması |
|----------------------------------|---|
| ALACAK_TAHSIL_SURESI | Alacak Tahsil Süresi (gün) |
| BORC_DEVIR_SURESI_G | Borç Devir Süresi (gün) |
| BRUT_KAR_MARJI_INSA | Brüt Kar Marjı |
| CARI_ORAN | Cari Oran |
| LIKIDITE_ORANI | Likidite Oranı |
| TPLM_BORC_TOPLAM_AKTIF | Toplam Borç / Toplam Aktif |
| DONEM_VOK_MARJI | VÖK Marjı |
| DONEM_YIL_KV_UV_FINANSAL | KV + UV Finansal Borçlar / Net Satışlar |
| TOP_BORC_TOP_AKTIF | Toplam Borç / Toplam Aktif |
| ALACAK_TAHSIL_S | Alacak Tahsil Süresi (gün) |
| CARI_ORAN | Cari Oran |
| KARLILIK_KALDIR | Karlılık Kaldırıcı (ROE / ROA) |
| KV_UV_FINANSA | KV + UV Finansal Borçlar / Net Satışlar |
| BIR_ONCEKI_YIL_LIKIDITE_ORANI | Likidite Oranı |
| BR_ONCKI_DNM_NT_STS_REL_BYME | Net Satışlardaki Reel Büyüme |
| DM_FD_12AyGcOdS | Gayri Nakdi Kredi komisyonlarında, son 12 ayda 6 işgünü ve üzerinde gecikerek yapılmış ödeme sayısı |
| FIRMANIN_FAALİYET_SURESI | Firmanın faaliyet süresi? (yıl) |
| FM_MI_BrcSatOrn | Kısa Vadeli Ticari Borç / Net Satışlar |
| FN_MV_MVrGnDgOr_MIVr_Borc | (Firmanın Mal Varlıklarının Güncel Değeri + Hakim Ortak / % 25 ve daha fazla paylı ortakların Mal Varlıklarının Net Güncel Değeri (Borçlar Hariç)) / Firma Borçları (Leasing , Faktoring Dahil) |
| FN_MZ_KVNRkLtOr | TCMB Kısa Vadeli Nakit Risk / TCMB Kısa Vadeli Nakit Limit (Kredi Tarihine en yakın ay sonu itibariyle) |
| FN_MZ_SnNkRsOr | (TCMB Nakit Risk + KKB Bireysel Krediler Kalan Borç Taksidi*) / (TCMB Nakit Limit (Kredi Tarihine en yakın ay sonu itibariyle) + KKB Bireysel Krediler Toplam Meblağ) |
| FRMNN_DEVAM_EDEN_YATIRIM_TUTARI | Firmanın devam eden yatırımı |
| HAKIM_ORTAGIN_YASI | Hakim Ortağın Yaşı |
| IKI_ONCEKI_TOP_BORC_TOP_AKTIF | Toplam Borç / Toplam Aktif |
| IKI_ONCEKI_YIL_ALACAK_TAHSIL_S | Alacak Tahsil Süresi (gün) |
| IKI_ONCEKI_YIL_BORC_DEVIR_SURE | Borç Devir Süresi (gün) |
| IKI_ONCEKI_YIL_CARI_ORAN | Cari Oran |
| IKI_ONCEKI_YIL_KV_UV_FINANSA | KV + UV Finansal Borçlar / Net Satışlar |
| IKI_ONCEKI_YIL_LIKIDITE_ORANI | Likidite Oranı |
| IKI_ONCEKI_YIL_VOK_MARJI | VÖK Marjı |
| ISLRI_YRTMDN_SRML_ORTA_SKTR_DNYM | İşleri Yürütmeden Sorumlu Ortağın kredi talep eden firmanın içinde bulunduğu sektördeki deneyimi kaç yıldır? |
| MVCT_ADRSTKI_FAALİYET_SURESI | Mevcut adreste faaliyete süresi (yıl) |
| NET_SATISLARDAKI_REEL_BUYUME | Net Satışlardaki Reel Büyüme |

| Değişken Adı | Açıklaması |
|----------------------------------|--|
| ONCEKI_YIL_BORC_DEVIR_SURESI | Borç Devir Süresi (gün) |
| ONCEKI_YIL_BRUT_KAR_MARJI_INS | Brüt Kar Marjı |
| ONCEKI_YIL_VOK_MAR | VÖK Marjı |
| SON_AY_ODENEN_BANKAMIZ_CEKLERI | Son ay ödenen Bankamız Çekleri Toplam Tutarı / Son 12 ay ödenen Bankamız Çekleri Ortalama Tutarı oranı |
| SON_YIL_ALACAK_TAHSIL_SURESI | Alacak Tahsil Süresi (gün) |
| SON_YIL_BORC_DEVIR_SURESI_GUN | Borç Devir Süresi (gün) |
| SON_YIL_BRUT_KAR_MARJI_INSAAT | Brüt Kar Marjı |
| SON_YIL_CARI_ORAN | Cari Oran |
| SON_YIL_KARLILIK_KALDIRACI_R | Karlılık Kaldırıcı (ROE / ROA) |
| SON_YIL_KV_UV_FINANSAL_BORCL | KV + UV Finansal Borçlar / Net Satışlar |
| SON_YIL_LIKIDITE_ORANI | Likidite Oranı |
| SON_YIL_TOPLAM_BORC_TOPLAM_AKTIF | Toplam Borç / Toplam Aktif |
| SON_YIL_VOK_MARJI | VÖK Marjı |
| SR_FHGB_MEVFAALSUR | Mevcut Faaliyet Süresi |
| TCMB_TPLM_RISK_TCMB_TOPLAM_LIMIT | TCMB Toplam Risk / TCMB Toplam Limit |

3.2. Veri Kontrolü ve Kullanıcı Kabul Testleri

Modelleme çalışmaları öncesinde gerekli olan verinin ihtiyaçlar ve gereksinimler açısından uygunluğunun temin edilmesini amaçlayan adımdır.

3.3. Bağımlı Değişken Veri Hazırlığı ve Seçimi

Çalışmamızda skorkarne hazırlanırken kullanılan temel yöntem “lojistik regresyon” yöntemidir. Daha önceki bölümlerde detayları paylaşıldığı gibi, lojistik regresyon, Y bağımlı değişkenininin X açıklayıcı değişkenlerinin katsayıları ile doğrusal ilişkide olmasının yanı sıra, Y 'nin gerçekleşme ihtimalinin doğal logaritması ile de doğrusal ilişkide olduğu ön kabulünden hareketle çalışır. Kredi derecelendirmede geleneksel lineer regresyon yöntemi yerine lojistik regresyonun kullanılmasının nedeni, bağımlı değişken olan Y 'nin olasılık belirtmesi nedeniyle sadece 0-1 arasında bir değer alması gerekliliğidir. Y 'nin 1 olması kredinin temerrüde düştüğünü gösterirken bu değer 0 olması, kredinin sorunlu kredi sınıfına girmediğini, aksine borçlunun finansal açıdan en sağlam olarak nitelendirilebileceğini gösterir. Bu yöntem daha önce de bahsedildiği gibi oldukça yaygın olarak kullanılmaktadır ve bunun nedenleri şu şekilde özetlenebilir:

- İlk olarak kategorik bağımlı değişken seçilirken açıklayıcı değişkenlerin sürekli ya da süreksiz olmalarına yönelik kısıtlama yoktur. Dolayısıyla model, her iki şekilde de kullanıma açıktır.
- Bu yaklaşımda kurulan model doğrusal hale getirilir ve bu da tahminlemeyi kolaylaştırır.
- Yöntem, SPSS, STATA ve SAS gibi birçok istatistik paket programında oldukça kolay bir şekilde yönetilebilir.
- Yöntem, açıklayıcı değişkenlerin olasılık fonksiyonlarının ne şekilde dağılımları gerektiği ile ilgili bir ön şart sunmaz, dolayısıyla varsayım ihlali riski düşüktür.
- Lojistik regresyonda, doğrusal modellerde olduğu gibi negatif olasılıkla karşılaşma ihtimali yoktur, zira olasılık dağımı 0-1 arasında tahmin edilir.
- Lojistik regresyon yöntemi, bağımsız değişken ile bağımlı değişken arasındaki ilişkinin doğrusal olmasını şart koşmaz. Dolayısıyla bu değişkenler arasındaki bağlantı üssel ya da polinomik yapıda olabilir.

Bağımlı değişkenin yapay olduğu modellerin temeli olan doğrusal olasılık modelinin sorunlarını aşmak için bazı dönüşümlerle doğrusal hale gelen fonksiyonlara ihtiyaç vardır. Lojistik dağılım fonksiyonunu incelemek için iki kategorili bağımlı değişken ele alınabilir (Baydemir, 2014).

Bu bölümde, modelleme perspektifinden mevcut durumu açıklanmıştır. Çalışma kapsamında model datası hazırlanırken aşağıdaki adımlar uygulanmıştır.

- Ampirik yapılan çalışmada, rating notu ve bu notu oluşturmaya aday bir veri seti oluşturulmuştur. İlgili veri seti dönemi 2015/07-2019/01 dönemleri arasında kapsamaktadır.
- Ampirik çalışmada rating notu bulunan müşterilere ait kredi bilgileri bulunmaktadır. Müşteriye ait kredisinin olup olmadığı ile ilgili flag bir değişken oluşturulmuştur.
- Rating notuna ait risk eşleştirme kriterleri aşağıdaki gibi listelenmiştir.
 - Rating notu bulunan müşteri rating tarihinden sonra maksimum 6 ay içerisinde kredi kullanmış ise risk flagi atanmıştır.

- Kredi kullanan müşterinin kredi kullanım tarihinden sonra rating notu oluşmuş ve kredi vadesi rating tarihinden büyük veya eşitse risk flagi atanmıştır.
- Riski ve rating notu eşleştirilen listeden aynı ay içerisinde birden fazla rating notu bulunan müşteriler tekilleştirilmiştir.
- Yukarıdaki maddeler uygulandıktan sonra rating tarihinden itibaren 1 yıl içerisinde temerrüde düşen kayıtlar “Bad” bunun dışında kalan kayıtlar ise “Good” olarak tanımlanmıştır.

3.4. Bağımsız Değişken Veri Hazırlığı ve Seçimi

3.4.1. Kayıp Veri Oranlarına Göre Alanların Değerlendirilmesi ve Seçimi

Modelleme projesi kapsamında paylaşılan veri sözlüğü doğrultusunda oluşturulan değişkenler kayıp veri oranlarına göre değerlendirilmiştir. Bu değişkenlerden %70 ve daha fazla kayıp oranına sahip değişkenler değişken havuzundan çıkartılmıştır. Daha sonra bu elenen değişkenler uzman bakış açısıyla değerlendirilecektir. Kayıp veri oranı; her bir değişken bazında boş olan satır sayısının, toplam satır sayısına bölünmesi ile bulunmuştur. Toplam veri seti 23,863 satırdan oluşmaktadır.

Hazırlanan veri setindeki kayıp veri oranları Tablo 3’deki gibidir:

Tablo 3. Değişken Bazında Kayıp Veri Oranları

| Değişken Adı | Dolu Veri | Kayıp Veri | Kayıp Veri Oranı |
|----------------------------------|-----------|------------|------------------|
| ARA_DONEM_ALACAK_TAHSIL_SURESI | 23,387 | 476 | 1.99% |
| ARA_DONEM_BORC_DEVIR_SURESI_G | 23,294 | 569 | 2.38% |
| ARA_DONEM_BRUT_KAR_MARJI_INSA | 23,852 | 11 | 0.05% |
| ARA_DONEM_CARI_ORAN | 23,718 | 145 | 0.61% |
| ARA_DONEM_LIKIDITE_ORANI | 23,759 | 104 | 0.44% |
| ARA_DONEM_TPLM_BORC_TOPLAM_AKTIF | 23,849 | 14 | 0.06% |
| ARA_DONEM_VOK_MARJI | 23,812 | 51 | 0.21% |
| ARA_DONEM_YIL_KV_UV_FINANSAL | 23,576 | 287 | 1.20% |
| BIR_ONCEKI_TOP_BORC_TOP_AKTIF | 23,841 | 22 | 0.09% |
| BIR_ONCEKI_YIL_ALACAK_TAHSIL_S | 23,442 | 421 | 1.76% |
| BIR_ONCEKI_YIL_CARI_ORAN | 23,540 | 323 | 1.35% |

| Değişken Adı | Dolu Veri | Kayıp Veri | Kayıp Veri Oranı |
|----------------------------------|------------------|-------------------|-------------------------|
| BIR_ONCEKI_YIL_KARLILIK_KALDIR | 23,863 | - | 0.00% |
| BIR_ONCEKI_YIL_KV_UV_FINANSA | 23,646 | 217 | 0.91% |
| BIR_ONCEKI_YIL_LIKIDITE_ORANI | 23,657 | 206 | 0.86% |
| BR_ONCKI_DNM_NT_STS_REL_BYME | 23,863 | - | 0.00% |
| DM_FD_12AyGcOdS | 23,863 | - | 0.00% |
| FIRMANIN_FAALİYET_SURESI | 23,807 | 56 | 0.23% |
| FM_MI_BrcSatOrn | 23,714 | 149 | 0.62% |
| FN_MV_MVrGnDgOr_MlVr_Borc | 23,863 | - | 0.00% |
| FN_MZ_KVNRkLtOr | 23,863 | - | 0.00% |
| FN_MZ_SnNkRsOr | 23,801 | 62 | 0.26% |
| FRMNN_DEVAM_EDEN_YATIRIM_TUTARI | 23,863 | - | 0.00% |
| HAKIM_ORTAGIN_YASI | 23,863 | - | 0.00% |
| IKI_ONCEKI_TOP_BORC_TOP_AKTIF | 23,850 | 13 | 0.05% |
| IKI_ONCEKI_YIL_ALACAK_TAHSIL_S | 23,483 | 380 | 1.59% |
| IKI_ONCEKI_YIL_BORC_DEVIR_SURE | 23,401 | 462 | 1.94% |
| IKI_ONCEKI_YIL_CARI_ORAN | 23,550 | 313 | 1.31% |
| IKI_ONCEKI_YIL_KV_UV_FINANSA | 23,654 | 209 | 0.88% |
| IKI_ONCEKI_YIL_LIKIDITE_ORANI | 23,676 | 187 | 0.78% |
| IKI_ONCEKI_YIL_VOK_MARJI | 23,427 | 436 | 1.83% |
| ISLRI_YRTMDN_SRML_ORTA_SKTR_DNYM | 23,818 | 45 | 0.19% |
| MVCT_ADRSTKI_FAALİYET_SURESI | 23,863 | - | 0.00% |
| NET_SATISLARDAKI_REEL_BUYUME | 23,863 | - | 0.00% |
| ONCEKI_YIL_BORC_DEVIR_SURESI | 23,243 | 620 | 2.60% |
| ONCEKI_YIL_BRUT_KAR_MARJI_INS | 23,825 | 38 | 0.16% |
| ONCEKI_YIL_VOK_MAR | 23,526 | 337 | 1.41% |
| SON_AY_ODENEN_BANKAMIZ_CEKLERI | 23,863 | - | 0.00% |
| SON_YIL_ALACAK_TAHSIL_SURESI | 23,425 | 438 | 1.84% |
| SON_YIL_BORC_DEVIR_SURESI_GUN | 23,237 | 626 | 2.62% |
| SON_YIL_BRUT_KAR_MARJI_INSAAT | 23,850 | 13 | 0.05% |
| SON_YIL_CARI_ORAN | 876 | 22,987 | 96.33% |
| SON_YIL_KARLILIK_KALDIRACI_R | 23,863 | - | 0.00% |
| SON_YIL_KV_UV_FINANSAL_BORCL | 23,663 | 200 | 0.84% |
| SON_YIL_LIKIDITE_ORANI | 23,665 | 198 | 0.83% |
| SON_YIL_TOPLAM_BORC_TOPLAM_AKTIF | 23,851 | 12 | 0.05% |
| SON_YIL_VOK_MARJI | 23,741 | 122 | 0.51% |
| SR_FHGB_MEVFAALSUR | 23,787 | 76 | 0.32% |
| TCMB_TPLM_RISK_TCMB_TOPLAM_LIMIT | 23,863 | - | 0.00% |

Tablo 3’den de görüleceği üzere kayıp veri oranı 96.33% olduğundan dolayı SON_YIL_CARI_ORAN değişkeni elenmiştir.

3.4.2. İstatistikî Sınamalara Göre Alanların Değerlendirilmesi ve Seçimi (Gini, Korelasyon)

3.4.2.1. Gini Katsayısı

Gini katsayısı, İtalyan ekonomist Corrado Gini tarafından sunulan ve Gini katsayısı, ekonomik eşitsizliğin en sık kullanılan ölçülerinden biridir. Katsayı, 0 ile 1 (veya %0 ile %100) arasında herhangi bir değer alabilir ve mutlak eşitlik doğrusu ile Lorenz eğrisi altında kalan alanın, mutlak eşitlik doğrusu altındaki üçgenin alanına bölünmesi ile bulunur.. Sıfır katsayısı, bir nüfus içinde tamamen eşit bir gelir veya servet dağılımını gösterir. Bir nüfustaki bir kişi tüm geliri alırken diğer insanlar hiçbir şey kazanmıyorsa, bir katsayısı mükemmel bir eşitsizliği temsil eder. Ayrıca bazı nadir durumlarda katsayı %100’ü geçebilir. Bu, teorik olarak, bir nüfusun geliri veya serveti negatif olduğunda ortaya çıkabilir. Ancak bu bahsedilen senaryolar gerçek dünyada oldukça nadirdir. Veriler, katsayının genel olarak %24 ile %63 arasında değiştiğini göstermektedir².

Gini katsayısı başlangıçta gelir dağılımı adaletsizliğini hesaplamak için kullanılmasına rağmen kredi derecelendirme modellerinde, modelin tahmin gücünü hesaplamak için de uygulanan yaygın ve güvenilir bir yöntemdir. Gini katsayısının formülü (14) ile gösterilmiştir (Thomas, Edelman, Crook, 2002):

$$Gini\ Index = \left(1 - \frac{2 \sum_{i=2}^m (n_i^{event} \sum_{j=1}^{i-1} n_j^{event}) + \sum_{k=1}^m (n_k^{event} n_k^{non-event})}{N^{event} N^{non-event}} \right) * 100 \quad \text{Denklem 14.}$$

Skorkarne modellemesi bağlamında, Gini Endeksi, modellemeye dahil edilecek değişken havuzunun incelenmesi ve dağılımı uygun olmayan değişkenlerin elemesinin gerçekleştirilmesi için kullanılmıştır. Son olarak, Gini katsayısının %50’den büyük bir değerde olması, tahminlerin tatmin edici düzeyde olduğunu, %30’un altında kalması ise sonuçların yeterince iyi olmadığını gösterir. Gini katsayısı ile ROC eğrisi arasında

² Kaynak: CFI Education Inc., (2021). “Gini Coefficient - A statistical measure of economic inequality in a population”, <https://corporatefinanceinstitute.com/resources/knowledge/economics/gini-coefficient/> (Erişim tarihi, 29 Kasım 2021).

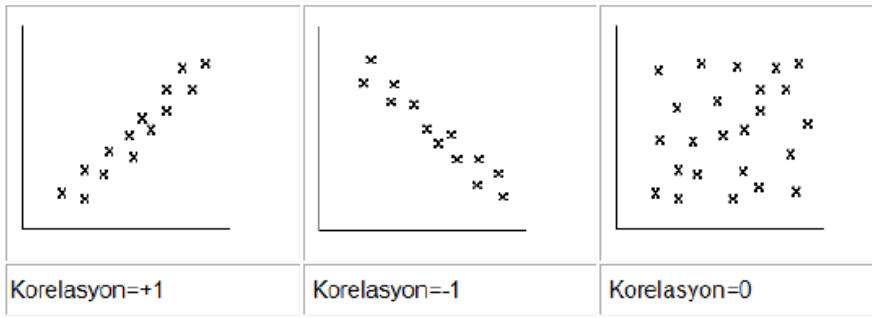
yaklaşık bir ilişki söz konusudur ve bu katsayı da ROC eğrisi altında kalan alana benzer şekilde hesaplanır. Bu ilişki (15) ile gösterilmiştir (Thomas, Edelman, Crook, 2002):

$$ROC \approx (D + 1)/2$$

Denklem 15.

3.4.2.2. Korelasyon Analizi

Korelasyon analizi, iki (veya daha fazla) nicel değişken arasındaki ilişkiyi belirtmek için kullanılan bir terimdir. Bu analiz temel olarak nicel değişkenler arasında doğrusal bir (lineer) ilişki varsayımına dayanmaktadır. İkili değişkenler için ilişki ölçülerine benzer şekilde, değişkenler arasındaki ilişkinin “gücünü” veya “büyüklüğünü” ve aynı zamanda yönünü ölçer. Bir korelasyon analizinin nihai sonucu, değerleri -1 ile +1 arasında değişen bir korelasyon katsayısıdır. Bu durumda +1'lik bir korelasyon katsayısı, iki değişkenin pozitif bir şekilde mükemmel düzeyde ilişkili olduğunu, -1'lik bir korelasyon katsayısı, iki değişkenin negatif bir şekilde mükemmel düzeyde ilişkili olduğunu gösterirken, sıfır bir korelasyon katsayısı, incelenen iki değişken arasında doğrusal bir ilişki olmadığını ifade eder. Bu üç korelasyon örnekleri Şekil 1’de gösterilmiştir (Gogtay ve Thatte, 2017).



Şekil 1. İki Değişken Arasındaki İlişkinin Grafıksel Görünümü

Doğrusal korelasyonun hesaplanmasında, Pearson Momentler Çarpımı korelasyonu yaygın bir şekilde kullanılır ve bu değer, en az bir aralık ölçeğinde ölçülen iki değişken arasındaki ilişkinin gücünün ve yönünün bir ölçüsüdür. Bu hesaplamaların yapılabilmesi için veriler en az aralıklı ölçekle toplanmalı ve süreklilik gösteren nicel bir değişken bulunmalıdır. Korelasyon katsayısının formülü, (16) olarak aşağıda sunulmuştur.

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right) \left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}}$$

Denklem 16.

3.5. Modelin Geliştirilmesi

Modelleme kapsamında ihtiyaç duyulan hedef ve bağımsız değişkenler belirlendikten sonra modelleme adımları uygulanmıştır. Skorkarne modellemesi kapsamında en tutarlı algoritma seçilebilmesi için, verinin % 70'i modelleme verisi (train), %30'u validasyon verisi olarak kullanılmaktadır. Bu bilgiler sırasıyla Tablo 4, Tablo 5 ve Tablo 16'da özetlenmiştir.

Tablo 4. Tüm Data – Canlı & Temerrüt Yüzdeleri

| Data = DATA | | | |
|-------------|---------------|-----------------|---------|
| Variable | Numeric Value | Formatted Value | Percent |
| BAD_GOOD | 0 | 0 | 97.98% |
| BAD_GOOD | 1 | 1 | 2.01% |

Tablo 5. Train Data – Canlı & Temerrüt Yüzdeleri

| Data=TRAIN | | | |
|------------|----------|----------|----------|
| Variable | Variable | Variable | Variable |
| BAD_GOOD | 0 | 0 | 97.98% |
| BAD_GOOD | 1 | 1 | 2.01% |

Tablo 6. Validasyon Datası – Canlı & Temerrüt Yüzdeleri

| Data=VALIDATE | | | |
|---------------|----------|----------|----------|
| Variable | Variable | Variable | Variable |
| BAD_GOOD | 0 | 0 | 97.96% |
| BAD_GOOD | 1 | 1 | 2.03% |

Modelleme adımları;

- 1- Modelleme kapsamında belirlenen verinin içeri alınması
- 2- Model ve validasyon veri kümelerinin ayrılması
- 3- Veri gruplama

Başlangıçta otomatik gruplama en iyi gruplamayı sağlayabilir. Ek olarak Veri Gruplama, değişkenleri kullanıcı bakış açısıyla yeniden gruplandırmanıza olanak tanır. Ayrıca değişkenleri tarama veya seçme hakkı da tanımaktadır. Her özellik için aldığı değere göre Weight of Evidence (WoE)'yi hesaplar.

3.5.1. Weight of Evidence (WoE)

Weight of Evidence (WoE), bir nitelik veya grup seviyesinin görel riskini ölçer. Bu değer, "iyi" (hedef değişken = 0) veya "kötü" (hedef değişken = 1) olan binary (iki terimli) hedef değişkeninin değerine bağlıdır.

Weight of Evidence (WoE), iyi ve kötü başvuruları ayırt etmedeki bir özellik niteliğini ölçer. Weight of Evidence, her grup düzeyinde iyi başvuru sahiplerinin kötü başvuru sahiplerine oranına dayanmaktadır. Bir değişken için WoE aşağıdaki şekilde hesaplanır:

$$WoE = \ln \left(\frac{\text{Distribution of Goods}}{\text{Distribution of Bads}} \right) \quad \text{Denklem 17.}$$

Negatif değerler, hesaplanan değişken ve değişkenin değerleri için, kötü başvuranların oranının iyi başvuru oranından daha fazla gözlemlendiğini göstermektedir. Diğer bir deyişle, negatif WoE değerleri, bu gruptaki başvuru sahiplerinin daha büyük bir kredi riski oluşturması bakımından daha kötüdür. Varsayılan olarak, eksik değerler kendi gruplarına atanır. WoE eğrisinin şekli, skorkarnedeki noktaların nasıl atandığını gösterir. WoE eğrisinin hesaplanma metodolojisi kapsamında monotonik davranması beklenmektedir.

3.5.2. Information Value (IV)

Bir değişkenin öngörücü gücü (yüksek riskli başvuru sahiplerini düşük riskli olanlardan ayırma kabiliyeti) Information Value veya Gini Endeksi ile değerlendirilir.

Information Value, deęişkenlerin aldıkları deęer aralıklarının WoE'sinin aęırlıklı bir toplamıdır. Information Value aşıęıdaki gibi hesaplanır.

$$IV = \sum_{i=1}^m (P(attribute_i | non - event) - P(attribute_i event)) * WoE$$

Denklem 18.

3.5.3. Lojistik Regresyon

Modelleme alıřmaları kapsamında kademeli (stepwise) yntemi uygulanmıřtır. Bu yntem ile deęişkenlerin anlamlılık dzeyeleri ve korelasyon analizleri yapılmıřtır. Daha sonraki srete uzman grüşü ile model deęerlendirme yapılarak deęişkenler tekrar deęerlendirilmiřtir. Uzman grüşünden sonra model sonu deęişkenleri son halini almıřtır.

3.5.4. Skorlama

Gruplanmış (binlenmiş) deęişkenleri lojistik regresyon modelinde girdi olarak kullanarak modelleme gerekleřtirmektedir. Ayrıca, skor puanlarını ve modelleme sonucunda elde edilen skorkartını hesaplamak iin regresyon parametrelerini leklendirir. Son olarak skorkarne, elde edilen skorkartını anlamada yardımcı olan puanlama ve karakteristik (deęişken) analizi gerekleřtirir ve puanlamaya dayalı bir strateji oluřturulmasına yardımcı olur.

Veri Gruplamada deęişkenler WoE hesaplaması sonunda elde ettięi information value ve monotoniklik durumlarına gre deęerlendirilmiřtir. Weight of Evidence (WoE), iyi ve kt bařvuruları ayırt etmedeki bir zellik nitelięini ler. Weight of Evidence, her grup dzeyinde iyi bařvuru sahiplerinin kt bařvuru sahiplerine oranına dayanmaktadır.

Sonraki ařamada kademeli lojistik regresyon metodolojisi uygulanmıřtır. Bu yntem ile deęişkenlerin anlamlılık dzeyeleri ve korelasyon analizleri yapılmıřtır. Bu metodoloji kapsamında modele dahil olan 12 deęişken aşıęıda yer alan Tablo 7'deki gibidir:

Tablo 7. Model Deęişkenleri ve Aıklamaları

| Deęişken İsimleri | Deęişken Aıklamaları | Model Seimi / Uzman Grüşü |
|-----------------------|-----------------------|-----------------------------|
| BRUT_KAR_MARJI_INSAAT | Brüt Kar Marjı | Model Seimi |

| Değişken İsimleri | Değişken Açıklamaları | Model Seçimi / Uzman Görüşü |
|----------------------------------|---|-----------------------------|
| CARI_ORAN | Cari Oran | Model Seçimi |
| FIRMANIN_FAALİYET_SURESI | Firmanın Faaliyet Süresi (YIL) | Model Seçimi |
| FN_MZ_KVNRkLtOr | TCMB Nakit Risk / TCMB Nakit Limit (Kredi Tarihine en yakın ay sonu itibariyle) | Korelasyondan Elendi |
| FN_MZ_SnNkRsOr | Son Ay Nakit Risk / Son 12'nci Ay Nakit Risk Ortalama Aylık Artış Oranı | Model Seçimi |
| FRMNN_SHPLK_DURUMU_51_KIME_AIT | Firmanın Sahiplik Durumu (% 51 kime ait) | Model Seçimi |
| HAKIM_ORTAGIN_YASI | Hakim Ortağın Yaşı | Model Seçimi |
| KS_VDL_FNNSL_BRCLR_TPLM_NKT | Net KV Finansal Borçlar / Toplam Nakit Açığı | Uzman Görüşü |
| FLYTLRDN_YRTLN_NKT | Faaliyetlerden Yaratılan Nakit / (Faiz + Anapara + Leasing Borcu) | Model Seçimi |
| BORC_DEVIR_SURESI_GUN | Borç Devir Süresi (GÜN) | Model Seçimi |
| KV_UV_FINANSAL_BORCLAR | KV + UV Finansal Borçlar / Net Satışlar | Model Seçimi |
| TOPLAM_BORC_TOPLAM_AKTIF | Toplam Borç / Toplam Aktif | Uzman Görüşü |
| TCMB_TPLM_RISK_TCMB_TOPLAM_LIMIT | TCMB Toplam Risk / TCMB Toplam Limit (Kredi Tarihine en yakın ay sonu itibariyle) | Model Seçimi |

Bu 12 değişkenden 2 adedi uzman görüşü ile modelleme sürecine dahil edilmiştir. Daha sonra 2 değişkenin hem diğer değişkenler ile korelasyonu hem de temerrüt oranı (default rate) içinde monotonik bir dağılım gösterip göstermediği kontrol edilmiştir. Sonrasında diğer değişkenler ile korelasyonun yüksek olmadığı ve monotonik dağıldığı gözlemlenmiştir.

Nihai değişkenler her bir grup için skorlanarak nihai rating skor skalası oluşmaktadır. Skor skalası tüm data, model kümesi ve validasyon kümesi için 3 ayrı model şeklinde oluşur ve karşılaştırması yapılır. Karşılaştırmalı küme tablosu, Tablo 8, Tablo 9 ve Tablo 10'da yer almaktadır.

Tablo 8. Skor Aralığı – Tüm Data

| Rating | Skor Aralığı | Gerçekleşen Temerrüt Oranı (%) | Model Ort. Temerrüt Olasılığı (%) |
|---------------|---------------------|---------------------------------------|--|
| AAA | Score >= 885 | 0,27% | 0,28% |
| AA | 822 <= Score < 884 | 0,67% | 0,51% |
| A | 791 <= Score < 821 | 0,78% | 0,73% |
| BBB | 731 <= Score < 790 | 0,98% | 0,96% |
| BB | 684 <= Score < 730 | 1,24% | 1,29% |
| B | 654 <= Score < 683 | 1,74% | 1,72% |
| CCC | 636 <= Score < 653 | 2,60% | 2,39% |
| CC | 573 <= Score < 635 | 2,76% | 2,43% |
| C | 494 <= Score < 572 | 3,10% | 3,38% |
| D | Score < 493 | 6,78% | 7,03% |

Tablo 9. Skor Aralığı – Train Data

| Rating | Skor Aralığı | Gerçekleşen Temerrüt Oranı (%) | Model Ort. Temerrüt Olasılığı (%) |
|---------------|---------------------|---------------------------------------|--|
| AAA | Score >= 885 | 0,11% | 0,28% |
| AA | 822 <= Score < 884 | 0,54% | 0,51% |
| A | 791 <= Score < 821 | 0,59% | 0,73% |
| BBB | 731 <= Score < 790 | 0,83% | 0,96% |
| BB | 684 <= Score < 730 | 1,19% | 1,29% |
| B | 654 <= Score < 683 | 1,83% | 1,72% |
| CCC | 636 <= Score < 653 | 2,95% | 2,39% |
| CC | 573 <= Score < 635 | 3,02% | 2,43% |
| C | 494 <= Score < 572 | 3,17% | 3,38% |
| D | Score < 493 | 6,81% | 7,03% |

Tablo 10. Skor Aralığı – Validasyon Datası

| Rating | Skor Aralığı | Gerçekleşen Temerrüt Oranı (%) | Model Ort. Temerrüt Olasılığı (%) |
|--------|--------------------|--------------------------------|-----------------------------------|
| AAA | Score >= 885 | 0,63% | 0,28% |
| AA | 822 <= Score < 884 | 0,95% | 0,51% |
| A | 791 <= Score < 821 | 1,27% | 0,73% |
| BBB | 731 <= Score < 790 | 1,33% | 0,96% |
| BB | 684 <= Score < 730 | 1,36% | 1,29% |
| B | 654 <= Score < 683 | 1,51% | 1,72% |
| CCC | 636 <= Score < 653 | 1,78% | 2,39% |
| CC | 573 <= Score < 635 | 2,12% | 2,43% |
| C | 494 <= Score < 572 | 2,94% | 3,38% |
| D | Score < 493 | 6,71% | 7,03% |

Model kurulduktan sonra modelin performansının ölçülmesi amacıyla validasyon testleri uygulanmıştır. Uzman görüşü ile seçilen alanlar performans testleri ile tekrar değerlendirilecek ve gerekli görülmesi durumunda bu alanların çıkarılarak sadece model sonucunda anlamlı çıkan değişkenler ile model revize edilecektir.

3.6. Model Performans Değerlendirmesi ve Validasyonu

3.6.1. Model Performansı

Modelleme sonrası hazırlanan skorkarnenin güvenilirliği ROC eğrisi ve Gini Katsayısı gibi istatistiksel yöntemler yardımı ile test edilmiştir.

3.6.1.1. ROC Eğrisi ve Gini Katsayısı

Derecelendirme modeli ile yapılan tahminlerin doğru sınıflandırılıp sınıflandırılmadığını gösteren Gini katsayısını hesaplamanın yöntemlerinden bir tanesi, daha önce de bahsedildiği gibi ROC eğrisidir. ROC eğrisinin altındaki alan, aynı zamanda kredi puanı modellerinin performansını değerlendirmek ve karşılaştırmak için popüler bir ölçümdür. ROC eğrisi, karışıklık matrisinden iki oranı özetler, bunlar Gerçek Pozitif Oran (TPR veya Geri Çağırma) ve Yanlış Pozitif Oran (FPR)'dır (Anderson, 2007). Gerçek Pozitif Oran (TPR), modelimizin verilerimizdeki toplam temerrüde düşen borçlu sayısı üzerinden yakaladığı temerrüde düşen borçluların sayısı olarak tanımlanır. Yanlış Pozitif

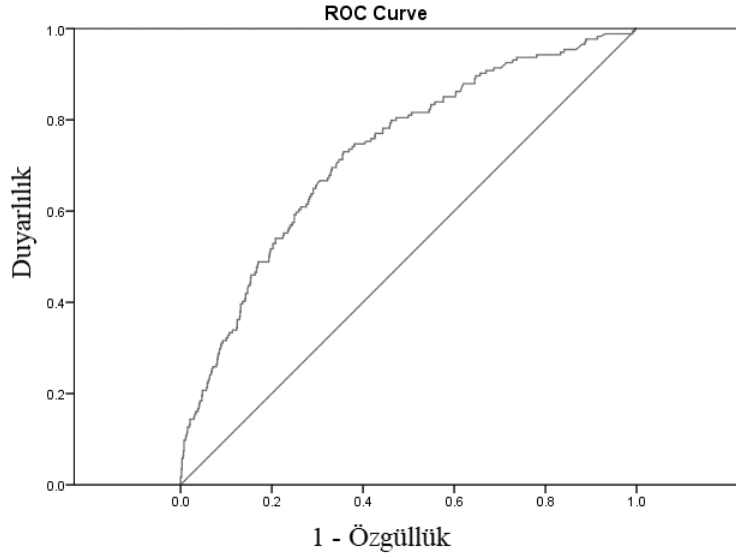
Oran (FPR), modelin temerrüt olmayan örneklerin toplam sayısı üzerinden bir temerrüdü yanlış tahmin ettiği durumların sayısı olarak hesaplanır.

ROC eğrisi, 1 ile 1000 arasındaki eşiklerden kaynaklanan ve TPR ve FPR'lerini süren karışıklık matrisleri kullanılarak oluşturulur. ROC eğrisinin y eksenini TPR değerlerini temsil eder ve x eksenini FPR değerlerini temsil eder (bkz. Şekil 2). ROC eğrisi altında kalan alan 0 ile 1 arasında bir değer alır. Eğri altında kalan alan arttıkça, modelin tahmin gücü yükselir, bu da modelin hedef değişkenleri sağlıklı bir şekilde sınıflandırabildiği anlamına gelir.

Tablo 11. Sınıflandırma Tablosu

| | | Tahmin Edilen | |
|-------------|--------|---------------|-------|
| | | Y = 0 | Y = 1 |
| Gözlemlenen | Y = 0 | A | B |
| | Y = 1 | C | D |
| | Toplam | A + C | B + D |

Sınıflandırma tablosundaki verilerle hesaplanan duyarlılık $A/A+C$, özgüllük ise $D/(B+D)$ şeklinde formülize edilir. Belirlenecek olan kesim noktasına göre bu değerler farklılık göstermekle birlikte genellikle 0.50 olarak tercih edilmektedir. Hesaplanan özgüllük ve duyarlılık ile ROC eğrisi çizilirken y eksenine duyarlılık, x eksenine ise (1-özgüllük) yer alır. ROC eğrisi için örnek bir gösterim aşağıda yer almaktadır. Duyarlılık ve 1-Özgüllük oranları kullanılarak oluşturulan eğri ile x eksenini arasında kalan alan, AUR'yi verir. Bu alanın nasıl yorumlanacağı ile ilgili yaklaşımlar, Tablo 12'de verilmiştir ve genel kurala göre $ROC < 0.50$ ise bir ayırım olmadığı, yani modelin başarısının düşük olduğu çıkarımı yapılır (Anderson, 2007).



Şekil 2. Örnek ROC Eğrisi

Hesaplamalarda ROC eğrisinin altında kalan alan da (Area Under Curve=AUC) dikkate alınır. Modelde ROC eğrisinin altında kalan alan Tablo 7’de görüldüğü üzere train veri setinde %76 ve validasyon veri setinde %68 çıktığından dolayı modelin tahmin gücünün iyi olduğunu göstermektedir.

Tablo 12. ROC Eşik Değerleri

| Eşik Değerleri | Açıklama |
|------------------------|--|
| $ROC \leq 0.50$ | Kabul edilebilir düzeyde bir fark ya da ayırım yoktur. |
| $0.50 \leq ROC < 0.70$ | Zayıf düzeyde bir fark ya da ayırım vardır. |
| $0.70 \leq ROC < 0.80$ | Kabul edilebilir düzeyde bir fark ya da ayırım vardır. |
| $0.80 \leq ROC < 0.90$ | İyi seviyede bir fark ya da ayırım vardır. |
| $ROC \geq 0.90$ | Mükemmel bir fark ya da ayırım vardır. |

Uygulayıcılar, model doğrulama raporlarında Gini katsayısına ek olarak AUR (Area Under ROC)'yi açıklama eğilimindedir. Ancak bu metrikler doğrusal bir ilişkiye sahip olduğundan, bu metriklerin birlikte açıklanması modelin kalitesinin değerlendirilmesine herhangi bir değer katmaz. Genel formülü $2 * AUR - 1$ olduğundan dolayı %0 ile %100

arasında bir değer elde edilir. %0 sonucu, iyi ve kötü kredi ayırımının başarısız yapıldığını, %100 ise ayırımın mükemmel başarıya yapıldığını gösterir. Çalışmamızda kullanılan modelde, Gini Katsayısı train veri setinde %52 ve validasyon veri setinde %41 olarak elde edilmiştir ve bu sonuçlar Tablo 13’de paylaşılmıştır.

Tablo 13. Model Performans Sonuçları

| Fit Statistics | Statistics Label | Train | Validation |
|----------------|------------------|-------|------------|
| _AUR_ | Area Under ROC | 0,760 | 0,680 |
| _Gini_ | Gini Coefficient | 0,520 | 0,410 |

3.6.2. Model Validasyonu

Model validasyonu hata matrisi, popülasyon istikrarlılık göstergesi (PSI), temerrüt oranı dağılımı, Herfindahl-Hirschman endeksi (HHI) ve binom testi gibi istatistiksel yöntemler yardımı ile test edilmiştir.

3.6.2.1. Hata Matrisi (Confusion Matrix)

Gerçekleşen canlı ve temerrüt sayılarının model tahminindeki canlı ve temerrüt sayılarının arasındaki hata payıdır. Model tahminindeki canlı ve temerrüt sayılarını belirlerken CCC ve altındaki notlar temerrüt, CCC üzerindeki notlar canlı olarak kabul edilmiştir.

Tablo 14. Hata Matrisi Eşik Değerleri

| Eşik Değerleri | Açıklamaları |
|--|-------------------------|
| Sensitivity, Specificity, Accuracy < 0.50 | Çok düşük tahmin gücü |
| 0.50 > = Sensitivity, Specificity, Accuracy < 0.60 | Düşük tahmin gücü |
| 0.60 > = Sensitivity, Specificity, Accuracy < 0.70 | Orta seviye tahmin gücü |
| Sensitivity, Specificity, Accuracy > = 0.70 | İyi seviye tahmin gücü |

Hata matrisi gerçekleşen ve tahmin edilen temerrüt durumları kullanılarak karşılaştırılmıştır. Karşılaştırma yapılabilmesi amacıyla temerrüt oranları ikili değerlere (0-1), ortalama temerrüt oranı sınır değer alınarak çevrilmiştir. Hata matrisi

incelendiğinde model duyarlılığının ve gösterim gücünün “Orta Seviye” tahmin gücüne sahip olduğu gözlenmiştir.

Tablo 15. Hata Matrisi Sonucu

| Değişkenler | Değerleri |
|----------------|-----------|
| Accuracy Ratio | 64.45% |
| Specifity | 70.54% |
| Sensitivity | 64.32% |

3.6.2.2. Popülasyon İstikrarlılık Gösterisi (PSI)

Skorkarne sisteminin odağında, yeni gelen başvuruların skorkarne geliştirilirken yani önceki verilere dayanan başvurulardan farklı olmayacağı varsayımı yatmaktadır. Popülasyon istikrar hesaplaması, skorkarne kullanılmaya başlandıktan sonra gelen müşteri verileri ile kullanılan örnek kütleyle karşılaştırarak popülasyon istikrar indeksini tahmin eder. PSI Formülü (19)'da paylaşılmıştır.

$$PSI = \sum \left((Actual \% - Expected \%) * \left(\ln \frac{Actual \%}{Expected \%} \right) \right) \quad \text{Denklem 19.}$$

Yapılan hesaplama neticesinde şayet iki kütle arasında büyük bir fark ortaya çıkarsa modelde geliştirmeler yapma ihtiyacı oluşur. Bunun en önemli sebebi, portföy yapısında yaşanan değişimin skorkarnenin performansını olumsuz yönde etkileyecek olmasıdır.

Tablo 16. Popülasyon İstikrarlılık Göstergesi Eşik Değerleri

| Eşik Değerleri | Açıklamaları |
|--------------------|---|
| PSI <= 0.10 | Dağılımda önemli bir değişim yoktur. |
| 0.10 < PSI <= 0.25 | Dağılımda orta seviye değişim vardır. |
| PSI > 0.25 | Dağılımda önemli derecede değişim vardır. |

Sistemdeki değişime ilave olarak bu değişime sebep olan faktörlerin de açıklanması gerekmektedir ve değişken analiz raporu, model inşaa edilirken kullanılan her bir

değişkene ait değerlerin varyasyonunu incelemektedir ve bu raporun modelde yer alan tüm değişkenler için yapılması gerekir. Çalışmanın PSI raporları Tablo 17, Tablo 18 ve Tablo 19’da paylaşılmıştır.

Tablo 17. PSI Test Sonuçları – Tüm Data

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | DR | PD | Tahmini Temerrüt Sayısı | Tahmini DR | PSI |
|---------------|---------------|------------|----------------------|-------|-------|-------------------------|------------|---------------|
| AAA | 2.585 | 7 | 2.592 | 0,27% | 0,28% | 7,28 | 0,28% | 0,0000 |
| AA | 2.389 | 16 | 2.405 | 0,67% | 0,51% | 12,27 | 0,51% | 0,0004 |
| A | 1.653 | 13 | 1.666 | 0,78% | 0,73% | 12,09 | 0,73% | 0,0000 |
| BBB | 3.626 | 36 | 3.662 | 0,98% | 0,96% | 35,00 | 0,96% | 0,0000 |
| BB | 2.87 | 36 | 2.906 | 1,24% | 1,29% | 37,37 | 1,29% | 0,0000 |
| B | 1.916 | 34 | 1.95 | 1,74% | 1,72% | 33,61 | 1,72% | 0,0000 |
| CCC | 1.649 | 44 | 1.693 | 2,60% | 2,39% | 40,45 | 2,39% | 0,0002 |
| CC | 2.294 | 65 | 2.359 | 2,76% | 2,43% | 57,28 | 2,43% | 0,0004 |
| C | 2.186 | 70 | 2.256 | 3,10% | 3,38% | 76,24 | 3,38% | 0,0002 |
| D | 2.213 | 161 | 2.374 | 6,78% | 7,03% | 166,79 | 7,03% | 0,0001 |
| Toplam | 23.381 | 482 | 23.863 | | | | | 0,0014 |

Tablo 18. PSI Test Sonuçları – Train Data

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | DR | PD | Tahmini Temerrüt Sayısı | Tahmini DR | PSI |
|---------------|---------------|------------|----------------------|-------|-------|-------------------------|------------|---------------|
| AAA | 1.801 | 2 | 1.803 | 0,11% | 0,28% | 5,07 | 0,28% | 0,0016 |
| AA | 1.661 | 9 | 1.670 | 0,54% | 0,51% | 8,52 | 0,51% | 0,0000 |
| A | 1.185 | 7 | 1.192 | 0,59% | 0,73% | 8,65 | 0,73% | 0,0003 |
| BBB | 2.514 | 21 | 2.535 | 0,83% | 0,96% | 24,23 | 0,96% | 0,0002 |
| BB | 1.999 | 24 | 2.023 | 1,19% | 1,29% | 26,02 | 1,29% | 0,0001 |
| B | 1.394 | 26 | 1.420 | 1,83% | 1,72% | 24,47 | 1,72% | 0,0001 |
| CCC | 1.153 | 35 | 1.188 | 2,95% | 2,39% | 28,38 | 2,39% | 0,0012 |
| CC | 1.603 | 50 | 1.653 | 3,02% | 2,43% | 40,14 | 2,43% | 0,0013 |
| C | 1.525 | 50 | 1.575 | 3,17% | 3,38% | 53,22 | 3,38% | 0,0001 |
| D | 1.532 | 112 | 1.644 | 6,81% | 7,03% | 115,50 | 7,03% | 0,0001 |
| Toplam | 16.367 | 336 | 16.703 | | | | | 0,0049 |

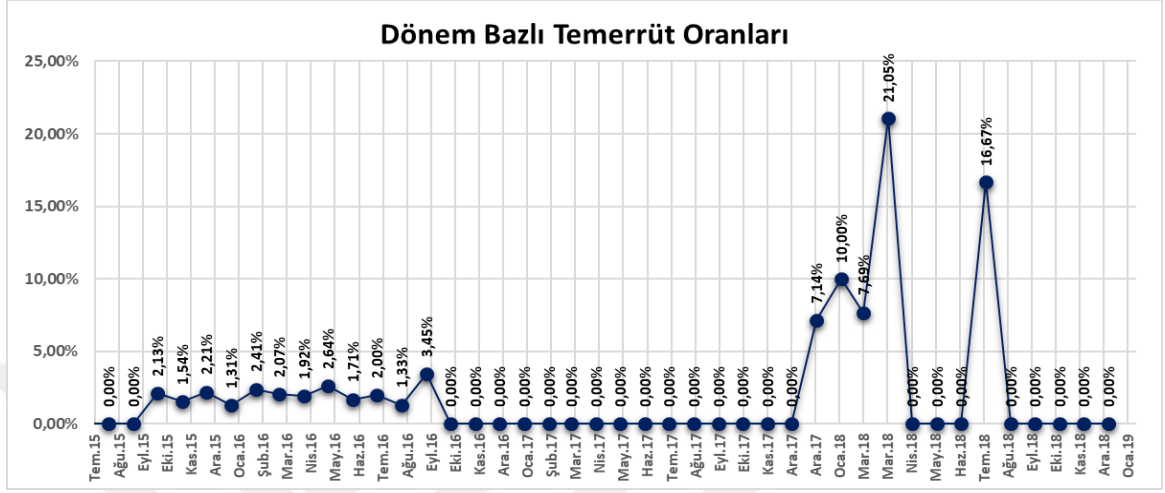
Tablo 19. PSI Test Sonuçları – Validasyon Datası

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | DR | PD | Tahmini Temerrüt Sayısı | Tahmini DR | PSI |
|--------------------|--------------|-----------------|-----------------------------|-----------|-----------|--------------------------------|-------------------|---------------|
| AAA | 784 | 5 | 789 | 0,63% | 0,28% | 2,22 | 0,28% | 0,0029 |
| AA | 728 | 7 | 735 | 0,95% | 0,51% | 3,75 | 0,51% | 0,0028 |
| A | 468 | 6 | 474 | 1,27% | 0,73% | 3,44 | 0,73% | 0,0030 |
| BBB | 1112 | 15 | 1127 | 1,33% | 0,96% | 10,77 | 0,96% | 0,0012 |
| BB | 871 | 12 | 883 | 1,36% | 1,29% | 11,36 | 1,29% | 0,0000 |
| B | 522 | 8 | 530 | 1,51% | 1,72% | 9,13 | 1,72% | 0,0003 |
| CCC | 496 | 9 | 505 | 1,78% | 2,39% | 12,07 | 2,39% | 0,0018 |
| CC | 691 | 15 | 706 | 2,12% | 2,43% | 17,14 | 2,43% | 0,0004 |
| C | 661 | 20 | 681 | 2,94% | 3,38% | 23,01 | 3,38% | 0,0006 |
| D | 681 | 49 | 730 | 6,71% | 7,03% | 51,29 | 7,03% | 0,0001 |
| Toplam | 7.014 | 146 | 7.160 | | | | | 0,0131 |

Tablo 17, Tablo 18 ve Tablo 19’da yer alan, tüm data, train ve validasyon datalarının PSI test sonuçları ayrı ayrı incelendiğinde rating notu dağılımlarında değişim olmadığı tespit edilmiştir.

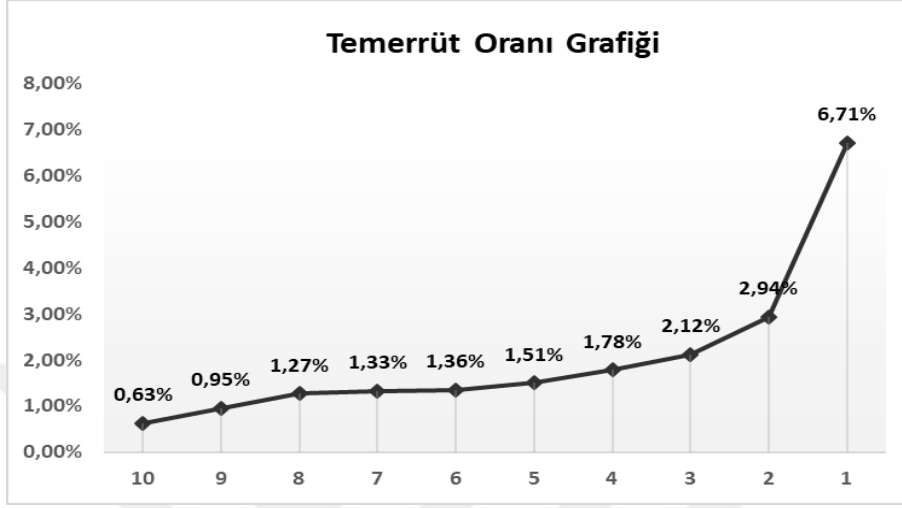
3.6.2.3. Temerrüt Oranı (Default Rate) Dağılımı

Her bir kredi derecelendirme notunun temerrüt adedinin o kredi derecesindeki toplam müşteri adedine bölünmesi ile bulunur. Temerrüt Oranı Dağılımı’nın iyi dereceden kötü dereceye doğru giderken monotonik bir artış göstermesi beklenmektedir.



Şekil 3. Dönem Bazlı Temerrüt Oranları Değişimi

Model verisindeki temerrüt oranları aylık frekans aralıklarında incelenmiş olup, temerrüt oranlarında müşteri adetlerinde ciddi sapmaların olduğu modelde 2016 yılına ait daha fazla veri bulunurken diğer yıllar için daha az bulunması sebebiyle önemli bir farklılık görülmemiştir. Ayrıca, rating bazında temerrüt oranlarının dağılımına bakıldığında iyi nottan kötü nota doğru gittikçe, beklenildiği gibi temerrüt oranlarının monotonik artış gösterdiği görülmektedir.



Şekil 4. Temerrüt Oranı Grafiği

3.6.2.4. Herfindahl-Hirschman Endeksi (HHI)

Literatürde Herfindahl Endeksi (HHI) olarak bilinen ve Orris Herfindahl ve Albert Hirschman adlı ekonomistler tarafından formül haline getirilmiş olan bu ölçüm sistemi, bir sektördeki rekabet düzeyini tespit etmek için uygulanmaktadır. Bu yöntem ile genellikle kredi riski arasındaki ilişki analiz edilir ve burada bir segmente ait payların kareleri toplanarak çıkan toplama göre sektördeki rekabet düzeyi belirlenmektedir (bkz. Denklem 20). HHI endeksi her bir rating içerisindeki ratinglerin karesinin toplanması ile hesaplanmaktadır. Herfindahl-Hirschman Endeksi (HHI) Formülü aşağıda verilmiştir.

$$H = \sum_i^N s_i^2$$

Denklem 20.

Tablo 20. Herfindahl-Hirschman Endeksi Eşik Değerleri

| Eşik Değerleri | Açıklamaları |
|-------------------|------------------|
| HHI ≤ 0.20 | Yoğunlaşma Yok |
| 0.20 < HHI ≤ 0.30 | Kısmi Yoğunlaşma |
| HHI > 0.30 | Aşırı Yoğunlaşma |

Yapılan çalışma sonucunda train, validasyon ve tüm veri HHI testinden geçer not almaktadır. İlgili metodoloji ile ilgili özet sonuçlar Tablo 15, Tablo 16 ve Tablo 17’de bulunmaktadır.

Tablo 21. HHI Testi Sonucu – Tüm Data

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | % | HHI |
|---------------|---------------|------------|----------------------|--------|---------------|
| AAA | 2.585 | 7 | 2.592 | 10,86% | 1,18% |
| AA | 2.389 | 16 | 2.405 | 10,08% | 1,02% |
| A | 1.653 | 13 | 1.666 | 6,98% | 0,49% |
| BBB | 3.626 | 36 | 3.662 | 15,35% | 2,35% |
| BB | 2.870 | 36 | 2.906 | 12,18% | 1,48% |
| B | 1.916 | 34 | 1.950 | 8,17% | 0,67% |
| CCC | 1.649 | 44 | 1.693 | 7,09% | 0,50% |
| CC | 2.294 | 65 | 2.359 | 9,89% | 0,98% |
| C | 2.186 | 70 | 2.256 | 9,45% | 0,89% |
| D | 2.213 | 161 | 2.374 | 9,95% | 0,99% |
| Toplam | 23.381 | 482 | 23.863 | | 10,55% |

Tablo 22. HHI Testi Sonucu – Train Data

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | % | HHI |
|---------------|---------------|------------|----------------------|--------|---------------|
| AAA | 1.801 | 2 | 1.803 | 10,79% | 1,17% |
| AA | 1.661 | 9 | 1.670 | 10,00% | 1,00% |
| A | 1.185 | 7 | 1.192 | 7,14% | 0,51% |
| BBB | 2.514 | 21 | 2.535 | 15,18% | 2,30% |
| BB | 1.999 | 24 | 2.023 | 12,11% | 1,47% |
| B | 1.394 | 26 | 1.420 | 8,50% | 0,72% |
| CCC | 1.153 | 35 | 1.188 | 7,11% | 0,51% |
| CC | 1.603 | 50 | 1.653 | 9,90% | 0,98% |
| C | 1.525 | 50 | 1.575 | 9,43% | 0,89% |
| D | 1.532 | 112 | 1.644 | 9,84% | 0,97% |
| Toplam | 16.367 | 336 | 16.703 | | 10,51% |

Tablo 23. HHI Testi Sonucu – Validasyon Dataseti

| Rating Notu | Canlı | Temerrüt | Toplam Müşteri Adedi | % | HHI |
|---------------|--------------|------------|----------------------|--------|---------------|
| AAA | 784 | 5 | 789 | 11,02% | 1,21% |
| AA | 728 | 7 | 735 | 10,27% | 1,05% |
| A | 468 | 6 | 474 | 6,62% | 0,44% |
| BBB | 1.112 | 15 | 1.127 | 15,74% | 2,48% |
| BB | 871 | 12 | 883 | 12,33% | 1,52% |
| B | 522 | 8 | 530 | 7,40% | 0,55% |
| CCC | 496 | 9 | 505 | 7,05% | 0,50% |
| CC | 691 | 15 | 706 | 9,86% | 0,97% |
| C | 661 | 20 | 681 | 9,51% | 0,90% |
| D | 681 | 49 | 730 | 10,20% | 1,04% |
| Toplam | 7.014 | 146 | 7.160 | | 10,67% |

Oluşturulan derecelendirme cetvelindeki derecelendirme havuzlarında yoğunlaşma olup olmadığının kontrolü amacıyla, hem tüm veri, hem model verisi hem de validasyon verisi üzerinde HHI testi gerçekleştirilmiştir. Test sonucunda, derecelendirme havuzlarında herhangi bir yoğunlaşmanın olmadığı tespit edilmiştir.

3.6.2.5. Binom Testi

Gerçekleşen ve tahmin edilen PD'lerin binom dağılımında karşılaştırılarak tahmin edilen PD'lerin doğruluğunun ölçüldüğü testtir. Binom testi için kurulan ana hipotez aşağıda belirtilmiştir.

H0: "İlgili gecikme gün sayısı kategorisi için tahmin edilen PD doğrudur."

İlgili kategoride temerrüt eden müşterilerin sayısının (k), belirlenen güven aralığı içerisindeki kritik değere (k') eşit veya bu değerden büyük olması durumunda ana hipotez reddedilir. Kritik değer binom dağılımına ilişkin formülü aşağıda verilmiştir (Bank for International Settlements, 2005). : Binom Dağılımı Formülü aşağıda paylaşılmıştır.

$$k^* = \min \left\{ k \mid \sum_{j=k}^n \binom{n}{j} PD^j (1 - PD)^{n-j} \leq 1 - q \right\} \quad \text{Denklem 21.}$$

Yukarıdaki formülün merkezi sınır teoremine göre revize edilmesi sonucunda aşağıdaki formüle ulaşılmıştır.

$$p^* \approx \phi^{-1}(q) \sqrt{\frac{PD(1 - PD)}{n}} + PD \quad \text{Denklem 22.}$$

Yapılan binom testi sonucunda her rating'e ait tahmin edilen ve gerçekleşen PD oranları karşılaştırılmış hem de tüm veri üzerinden gerçekleşen ve tahmin edilen PD oranları binom testine tabi tutulmuştur. Sonuçlar Tablo 18, Tablo 19 ve Tablo 20'de belirtilmiştir.

Tablo 24. Binom Testi Sonucu – Tüm Data

| Rating Notu | Obs. Defaults | Total Obs. | PD | % Obs. DR | Binom % 95 | Binom %97,5 | Binom %99 |
|-------------|---------------|------------|-------------|-----------|--------------|--------------|--------------|
| AAA | 7 | 2.592 | 0,00280924 | 0,0027 | Not Rejected | Not Rejected | Not Rejected |
| AA | 16 | 2.405 | 0,00510146 | 0,0067 | Not Rejected | Not Rejected | Not Rejected |
| A | 13 | 1.666 | 0,007259492 | 0,0078 | Not Rejected | Not Rejected | Not Rejected |
| BBB | 36 | 3.662 | 0,009558856 | 0,0098 | Not Rejected | Not Rejected | Not Rejected |
| BB | 36 | 2.906 | 0,012860615 | 0,0124 | Not Rejected | Not Rejected | Not Rejected |
| B | 34 | 1.950 | 0,017235151 | 0,0174 | Not Rejected | Not Rejected | Not Rejected |
| CCC | 44 | 1.693 | 0,023891623 | 0,0260 | Not Rejected | Not Rejected | Not Rejected |
| CC | 65 | 2.359 | 0,024280098 | 0,0276 | Not Rejected | Not Rejected | Not Rejected |
| C | 70 | 2.256 | 0,03379284 | 0,0310 | Not Rejected | Not Rejected | Not Rejected |
| D | 161 | 2.374 | 0,070258021 | 0,0678 | Not Rejected | Not Rejected | Not Rejected |

Tablo 25. Binom Testi Sonucu – Train Data

| Rating Notu | Obs. Defaults | Total Obs. | PD | % Obs. DR | Binom % 95 | Binom %97,5 | Binom %99 |
|-------------|---------------|------------|-------------|-----------|--------------|--------------|--------------|
| AAA | 2 | 1.803 | 0,00280924 | 0,0011 | Not Rejected | Not Rejected | Not Rejected |
| AA | 9 | 1.670 | 0,00510146 | 0,0054 | Not Rejected | Not Rejected | Not Rejected |
| A | 7 | 1.192 | 0,007259492 | 0,0059 | Not Rejected | Not Rejected | Not Rejected |
| BBB | 21 | 2.535 | 0,009558856 | 0,0083 | Not Rejected | Not Rejected | Not Rejected |
| BB | 24 | 2.023 | 0,012860615 | 0,0119 | Not Rejected | Not Rejected | Not Rejected |
| B | 26 | 1.420 | 0,017235151 | 0,0183 | Not Rejected | Not Rejected | Not Rejected |
| CCC | 35 | 1.188 | 0,023891623 | 0,0295 | Not Rejected | Not Rejected | Not Rejected |
| CC | 50 | 1.653 | 0,024280098 | 0,0302 | Not Rejected | Not Rejected | Not Rejected |
| C | 50 | 1.575 | 0,03379284 | 0,0317 | Not Rejected | Not Rejected | Not Rejected |
| D | 112 | 1.644 | 0,070258021 | 0,0681 | Not Rejected | Not Rejected | Not Rejected |

Tablo 26. Binom Testi Sonucu – Validasyon Datası

| Rating Notu | Obs. Defaults | Total Obs. | PD | % Obs. DR | Binom % 95 | Binom %97,5 | Binom %99 |
|--------------------|----------------------|-------------------|-------------|------------------|-------------------|--------------------|------------------|
| AAA | 2 | 1.803 | 0,00280924 | 0,0011 | Not Rejected | Not Rejected | Not Rejected |
| AA | 5 | 789 | 0,00280924 | 0,0063 | Not Rejected | Not Rejected | Not Rejected |
| A | 7 | 735 | 0,00510146 | 0,0095 | Not Rejected | Not Rejected | Not Rejected |
| BBB | 6 | 474 | 0,007259492 | 0,0127 | Not Rejected | Not Rejected | Not Rejected |
| BB | 15 | 1.127 | 0,009558856 | 0,0133 | Not Rejected | Not Rejected | Not Rejected |
| B | 12 | 883 | 0,012860615 | 0,0136 | Not Rejected | Not Rejected | Not Rejected |
| CCC | 8 | 530 | 0,017235151 | 0,0151 | Not Rejected | Not Rejected | Not Rejected |
| CC | 9 | 505 | 0,023891623 | 0,0178 | Not Rejected | Not Rejected | Not Rejected |
| C | 15 | 706 | 0,024280098 | 0,0212 | Not Rejected | Not Rejected | Not Rejected |
| D | 20 | 681 | 0,03379284 | 0,0294 | Not Rejected | Not Rejected | Not Rejected |

Binom Testi %95 güven aralığında rating bazlı temerrüt adedi dağılımlarına uygulanmıştır. Tahmin edilen temerrüt adedinin, gözlemlenen temerrüt adedi üzerinden hesaplanan belirli bir aralıkta olması beklenmektedir. Modele ait tüm rating notlarının binom testinden geçtiği görülmüştür.

4. SONUÇ

İçinde yaşadığımız bilgi çağının ya da bir başka yaygın ifade ile dijital çağın anahtar kelimelerinden olan “veri”, sürekli olarak işlenerek yaşamın her alanında kullanılacak nihai bilginin çıkarılabilmesi adına başvuru en değerli kaynaklardan bir tanesidir. Buna paralel olarak dünyada her an daha fazla kişi, kurum ya da cihaz birbirine bağlanmakta ve bunun sonucunda küresel veri alanının boyutu üssel olarak büyümektedir. International Data Corporation (IDC) tarafından yapılan tahminlere göre her iki yılda mevcut veri stoku iki katına çıkmaktadır ve 2025'te oluşturulan ve çoğaltılan veri hacminin 2016'da üretilen verilerin boyutundan on kat daha fazla olacağı beklenmektedir (Reinsel vd., 2017).

Oluşturulan veri hacmi hızla büyüdükçe, bilgiyi elde etmek için verilerin kullanılması, matematik, istatistik ve bilgisayar bilimi için önemli ve önemli bir ilgi alanı haline gelmiştir. Sonuç olarak, veri bilimi disiplinler arası bir alan olarak ortaya çıkmıştır. Finans sektörü, büyük miktarda veri toplayabilen, bunları bir işletme varlığı olarak depolayabilen ve işleyebilen başlıca sektörlerden birisi olması nedeniyle veri bilimi açısından oldukça önemli bir noktadadır. Sadece bankalar değil, diğer finansal kurumlar da verilere büyük önem vermekte ve bu doğrultuda iş birimleri kurarak onları veriye dayalı teknoloji çağına uyum sağlamak üzere eğitmektedir. Pazarlama analizi departmanları kampanya yönetimi, denetim birimleri sahtekârlık tespiti, kredi izleme ve tahsilat birimleri gelecekte potansiyel ödeme sorunlarının öngörülebilmesi amacıyla sistemler oluşturarak iş modellerini bunun üzerine kurmaktadır.

Ticari bankaların aracı finansal kurumlar olarak ana faaliyetlerinden biri, finansal kaynakları artırmak ve bunları üçüncü şahıslara, farklı büyüklükteki şirket ya da diğer ekonomik birimlere borç vermektir. Çeşitli hizmetlerine ilişkin ücret ve komisyonların yanı sıra, bankaların gelirlerinin büyük kısmı sağladıkları kredilerden gelmektedir. Banka kredileri temel olarak tüketici kredileri, mikro, küçük ve orta ölçekli işletmelere (KOBİ) verilen krediler ile ticari ve kurumsal krediler olarak kategorize edilebilir. Tüketici kredileri çok çeşitli farklı kredileri içerir ve nakit kredilerden ipoteklere kadar bireylerin ihtiyaç duyabilecekleri birçok alana hitap eder. Ticari ve kurumsal firmalara verilen

krediler, KOBİ'lerden daha büyük şirketler içindir ve örnek olarak proje finansmanı verilebilir.

Her borçlunun geri ödeme yetenekleri bağlamında farklı özellikleri vardır. Bu nedenle bankalar borçlulara ilişkin olarak kendi finansal yükümlülüklerini yerine getirmeme riskini üstlenir ve uygun bir faiz oranı belirlerler. Riskli müşteriler için, banka hiç kredi vermemeyi veya daha yüksek bir faiz oranı belirlemeyi seçebilir. Böylece, kredinin faiz oranı ile bankaya ödünç verilen paranın maliyeti arasındaki fark, yayılma (spread) oranını artırmakta ve bankanın riskini azaltmaktadır. Bu noktada kredi derecelendirme modelleri, muhtemelen kredilerini geri ödeyemeyecek olan adayları veya müşterileri belirlemek için kullanılır. Kredi başvurusunun onaylanması sadece derecelendirme modelinin çıktısına bağlı değildir; bilirkişinin görüşü, iş kuralları ve politikaları, bir kredinin değerlendirilmesinde belirleyici diğer faktörlerdir. Kredi derecelendirmesi, geçmişteki verilerin gelecekteki davranışı tahmin etmek için kullanılabilirliğini varsayar. Bu nedenle, gelecekteki kredi değeri olan adaylar geçmiş kredi değeri olan başvuru sahiplerine benzeme eğiliminde olacaktır. Bir kredi derecelendirme modeli, her başvuru sahibi veya belirli koşulları karşılayan mevcut her müşteri için temerrüt olasılığına (*probability of default* - PD) sahiptir. Bu, bankaların daha objektif, tutarlı, hızlı ve verimli borç verme kararları almalarını ve daha az idari çalışma ve düşük risk ile kredi operasyonlarını yönetmelerini sağlar. Ayrıca, bankacılık sektörünün yönetim kurumlarının da bir gereğidir. Bu amaçla oluşturulan skorkart da açıklayıcı değişkenler ile başvuranın temerrüt olasılığı arasındaki ilişkiyi yansıtır. Uygulamadan sonra, eğitilmiş model, yeni bir başvuruya veya her başvuru sahibine veya belirli koşulları karşılayan mevcut her müşteriye temerrüt olasılığı kullanılarak elde edilen bir skor atar.

Bankaların uyguladığı derecelendirme modellerinde derecelendirme değişkenleri, başvuru sırasında borç verenin sahip olduğu herhangi bir veri kaynağından seçilebilir. Bu değişkenlere örnek olarak demografik bilgiler (yaş, ikamet süresi, posta kodu), mevcut ilişki (müşterilik yaşı, ürün sayısı, ödeme performansı, önceki talepler), kredi bürosu verileri (diğer banka bilgileri, gecikmeleri, kamu kayıtları), emlak verileri, vb. verilebilir.

Başvuru sahibinin toplam puanı, o başvuru sahibinin derecelendirmesinde mevcut olan her bir değişkene ilişkin nitelik puanlarının toplamını ifade etmektedir. Elde edilen risk skoru bilgisi, kârı maksimize edecek ve zararı da en aza indirecek yeni değerlendirme stratejileri geliştirmek için kullanılabilir. Yüksek riskli adaylar için stratejilerden bazıları aşağıdakilerden bir veya bir kaçını olabilir:

- Risk seviyesi çok yüksekse, kredileri / ürünleri azaltmak
- Daha düşük limitli ürünler tayin etmek
- Başvuru sahibinden daha yüksek değerli teminat talep etmek
- Bir kredi için daha yüksek kar payı oranı belirlemek
- Sigorta poliçelerinden daha yüksek primler almak
- Potansiyel riskli başvuruları “yakın izleme listesine” koymak

Buna karşılık, yüksek puan alan başvuru sahiplerine daha yüksek oranlar ve kredi limitleri verilebilir, altın veya platin kartlar gibi premium ürünlere ya da şirket tarafından sunulan ek ürünlere yükseltme teklif edilebilir.

Bu nedenle, risk puanlaması, alacaklılara, elde edilen bilgilere dayanarak tutarlı ve objektif karar verme fırsatı sunar. İş bilgisi ile birleştirildiğinde, öngörücü modelleme teknolojileri, risk yöneticilerine risk yönetimi süreci üzerinde ek verimlilik ve kontrol sağlar (Siddiqi, 2012). Kredi derecelendirme, herhangi bir bankanın işleyişinde, gelecekteki finansal sonuçlar üzerinde gerçek bir etkisi olan kilit bir unsur olarak kabul edilir. Başlangıçta, kredi puanlama kavramını sadece basit uzman puanlama kartlarına atıfta bulunsa da, teknolojinin gelişmesiyle birlikte basit puanlama, yüksek derecede ilerleme ile tahmine dayalı modellere dönüşmüştür. Diğer yandan daha büyük ölçekteki kredi derecelendirme faaliyetleri (devletleri ya da büyük şirketleri konu alan) hem kantitatif hem de kalitatif değerlendirmelere dayalı olabilir. Kredi Derecelendirme Kuruluşları finansal birimlerin kredi riskini, borçlunun finansal yükümlülüklerini zamanında yerine getirme olasılığını, borcunun ana para ve faizini ödeme kabiliyet ve isteğini; çoğunlukla kendilerine ait bir metodoloji ile, nitel ve nicel (qualitative and quantitative) değerlendirmeler ile ölçerler.

Çalışmanın amacı, Özel banka müşterilerinin kredi derecelerinin geliştirilen bir model ile tespit edilmesi ve bu modelin gerçekleşen sonuçlar ile ne derecede uyumlu olduğunun, bir başka ifade ile model tutarlılığı ve güvenilirliğinin test edilmesidir. Kredi derecelendirmesi için çeşitli son teknoloji makine öğrenimi ve gelişmiş istatistiksel yöntemler kullanılmaktadır ve çalışmamızda, çıktılarda yeterli doğruluk, verimlilik ve yorumlanabilirlik dengesi sağladığı için endüstride en yaygın kullanılan tahmin yöntemlerinden birisi olan lojistik regresyon yöntemi tercih edilmiştir. İstatistiksel tabanlı puanlama modeli teknikleri, bankalar için daha yaygın olarak tercih edilen yöntemler olmaya devam etmektedir. Bu teknikler arasında lojistik regresyon halen en popüler olanıdır (Kil vd., 2021) ve bu tezde de uygulama bölümünün ana yöntemi olarak tercih edilmiştir. SAS programı kullanılarak lojistik regresyon yöntemi kullanılarak kurulan derecelendirme modelinin performans ve validasyon sonuçlarına göre model kullanılabilir olarak değerlendirilmiştir. Ancak, ampirik bir çalışma olduğundan bağımlı değişkenlerinin sayısı ve satır sayısı azdır. Bu nedenle, veri sayısının ve değişken sayısının çoğaltılarak modelin kurulması gerektiği düşünülmektedir.

Gelecekteki çalışmalar, mevcut çalışmada kullanılan geleneksel puanlama modelinin yanı sıra sinir ağları gibi ileri istatistiksel puanlama tekniklerini kullanmayı ve belki de bulanık algoritmalar gibi diğer tekniklerle entegre etmeyi amaçlamalıdır. Buna ek olarak, daha fazla veri toplanması ve puanlama modellerinin doğruluğunu artıracak daha fazla değişken kullanılması planlanmalıdır. Son olarak, gelecekteki araştırmaların, sonuçları genelleştirmek amacıyla birden fazla bankanın veri setini kullanması, tutarlılık ve daha uzun bir zaman dilimi ile daha geniş bir hedef kitle üzerindeki geçerlilik açısından önem arz etmektedir.

KAYNAKÇA

- Abdou, H., El-Masry, A., & Pointon, J. (2007). On the applicability of credit scoring models in Egyptian banks. *Banks and Bank systems*, 2(1).
- Allen, D.E., & Powell, R., (2011). *Credit Risk Measurement Methodologies. The 19th International Congress on Modelling and Simulation*, Australian Mathematical Sciences Institute, Australia.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.
- Altman, E. I., Avery, R. B., Sinkey Jr, J. F., & Eisenbeis, R. A. (1981). *Application of classification techniques in business, banking, and finance* (Vol. 3). Jai Press.
- Altman, E. I., Rijken, H., Watt, M., Balan, D., Forero, J., & Mina, J. (2010). The Z-Metrics™ methodology for estimating company credit ratings and default risk probabilities. *RiskMetrics Group*, 18-21.
- Altman, E.I., (2002). *Revisiting Credit Scoring Models in a BASEL 2 Environment*, London Risk Books, London.
- Anderson, R., (2007). *The Credit Scoring Toolkit*, Oxford University Press.
- Ayaz, B. H., (2016). *Uluslararası Kredi Derecelendirme Kuruluşları Ve Türkiye Ekonomisi Üzerindeki Rollerini*, Çukurova Üniversitesi, İktisadi Ve İdari Bilimler Fakültesi, Adana.
- Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the operational research society*, 54(6), 627-635.
- Bailey, M. (Ed.). (2004). *Credit scoring: The principles and practicalities*. White Box Publ.
- Banasik, J., Crook, J., & Thomas, L. (2001). Scoring by usage. *Journal of the Operational Research Society*, 52(9), 997-1006.

- Banasik, J., Crook, J., & Thomas, L. (2003). Sample selection bias in credit scoring models. *Journal of the Operational Research Society*, 54(8), 822-832.
- Bank for International Settlements, (2005). Working Paper No.14: Studies on the Validation of Internal Rating Systems.
- Baydemir, M. B., (2014). *Lojistik Regresyon Analizi Üzerine Bir İnceleme*, İnönü Üniversitesi, Sosyal Bilimler Enstitüsü, Malatya.
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111.
- Blöchlinger, A., & Leippold, M. (2006). Economic benefit of powerful credit scoring. *Journal of Banking & Finance*, 30(3), 851-873.
- Canbolat, M. A., Gümrah, A., (2015). *Analysis of Credit Risk Measurement Methods in the Evaluation of Credit Demands*, Universal Journal of Accounting and Finance.
- Derbali, A., & Hallara, S. (2013). Analysis of Default Probability: A Comparative Theoretical Approach between the Credit Portfolio View Model and the Creditrisk Model. *International Journal of Business Management & Research (IJBMR)*, 3(1), 157-170.
- Durand, D. (1941). *Risk elements in consumer installment financing*. National Bureau of Economic Research, New York.
- Erdoğan, S., (2005). *Mali Tahlil Ve İstihbarat Uzman Programı*, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü.
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2), 179-188.
- Gogtay, N. J., & Thatte, U. M. (2017). Principles of correlation analysis. *Journal of the Association of Physicians of India*, 65(3), 78-81.
- Greene, W. (1998). Sample selection in credit-scoring models. *Japan and the world economy*, 10(3), 299-316.
- Guillen, M., & Artis, M. (1994). *Count data models for a credit scoring system* (No. 021).

- Gurný, P., & Gurný, M. (2009). Estimation of PD of financial institutions within linear discriminant analysis. *Mathematical Methods in Economics*.
- Gurný, P., & Gurný, M. (2013). Comparison of credit scoring models on probability of default estimation for us banks.
- Hacıyeva, N., (2019). *Makroekonomik Ve Bankaya Özgü Faktörlerin Ticari Bankalarda Kredi Riski Üzerindeki Etkisi – Türkiye Örneği*, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Hand, D. J. (1981). Discrimination and classification. *Wiley Series in Probability and Mathematical Statistics*.
- Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 160(3), 523-541.
- Hoffmann, F., Baesens, B., Mues, C., Van Gestel, T., & Vanthienen, J. (2007). Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms. *European journal of operational research*, 177(1), 540-555.
- Joseph, M., (2005). *A PD Validation Framework for Basel II Internal Ratings - Based Systems*.
- Kavcıoğlu, Ş., (2011). *Ticari Bankacılıkta Kredi Riskinin Ve Kredi Riski Ölçüm Modellerinin Değerlendirilmesi*. Finansal Araştırmalar Ve Çalışmalar Dergisi.
- Kil, K., Ciukaj, R., & Chrzanowska, J. (2021). Scoring Models and Credit Risk: The Case of Cooperative Banks in Poland. *Risks*, 9(7), 132.
- Kim, Y. S., & Sohn, S. Y. (2004). Managing loan customers using misclassification patterns of credit scoring model. *Expert Systems with Applications*, 26(4), 567-573.
- Koç, S., (2018). *Lojistik Regresyon Yöntemi İle Kredi Skorlama Uygulaması, Yayınlanmış Yüksek Lisans Tezi*, Kadir Has Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

- Korkmaz, T.K., (2004). *Bankalarda Kredi Riski Ölçümünde Alternatif Yöntemler*, Active Bankacılık ve Finans Dergisi.
- Kutman, Ö., (2001). *Türkiye 'deki Şirketlerde Erken Uyarı Göstergelerinin Araştırılması*, Doğu Üniversitesi Dergisi, Sayı.4, İstanbul.
- Lewis, E. M. (1992). *An introduction to credit scoring*. Fair, Isaac and Company.
- Moralas, B., (2019). *An Application Of Credit Scoring Based On Logistic Regression*, Boğaziçi Üniversitesi.
- Orgler, Y. E. (1971). *Evaluation of bank consumer loans with credit scoring models*. Tel-Aviv University, Department of Environmental Sciences.
- Öker, A., (2007). *Ticari Bankalarda Kredi Ve Kredi Riski Yönetimi-Bir Uygulama*, *Yayınlanmamış Doktora Tezi*, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, Finansman Bilim Dalı, İstanbul.
- Özgür, S., (2019). *Bankacılık Sektöründe Kredi Riski Ve Kredi Türevleri: Ampirik Bir Uygulama*, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Pindyck, R. S., & Rubinfeld, D. L. (1976). *Econometric models and business forecasts*. McGraw-Hill.
- Reinsel, D., J. Gantz and J. Rydning, (2017). *Data Age 2025: The Evolution of Data to Life-Critical*, International Data Corporation (IDC), Massachusetts.
- Saka H., & Orhan M. (2018). Are Sovereign Ratings by CRAs Consistent? *Panoeconomicus*, 65(1), 95-115. doi:10.2298/PAN150311002S
- Salman, N., (2018). *Kredi Riski Yönetimi Ve İller Bankası Uygulaması*, *Uzmanlık Tezi*.
- Šarlija, N., Benšić, M., & Bohaček, Z. (2004). Multinomial model in consumer credit scoring. In *10th International Conference on Operational Research (10)*.
- Saunders, A., (1999). *Financial Institutions*, Mc Graw-Hill.

- Siddiqi, N. (2012). *Credit risk scorecards: developing and implementing intelligent credit scoring* (Vol. 3). John Wiley & Sons.
- Steenackers, A., & Goovaerts, M. (1989). A credit scoring model for personal loans. *Insurance: Mathematics & Economics*, 8(1), 31-34.
- Sullivan, A. C. (1981). Consumer finance. *EI Altman, Financial Handbook (9.3-9.27)*, New York: John Wiley & Sons.
- Sümbülođlu, K., & Sümbülođlu, V. Biyoistatistik,(2007). *Hatiboglu Yayınevi, Ankara*, 76-86.
- Şimşek, K.Ç., (2007). *Bankacılıkta Risk Ve Risk Ölçüm Yöntemleri, Yayınlanmış Yüksek Lisans Tezi*, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Tatlıdil, H., Özel, M., (2005). *Firma Derecelendirme Çalışmaları Konusunda Çok Değişkenli İstatistiksel Analize Dayalı Karar Destek Sistemlerinin Kullanımı*, Bankacılar Dergisi, Sayı 54.
- Thomas, L., Crook, J., & Edelman, D. (2017). *Credit scoring and its applications*. Society for industrial and Applied Mathematics.
- Thomas, L.C., Edelman, D.B., Crook, J.N., (2002). *Credit Scoring and Its Applications*, Society For Industrial and Applied Mathematics, Philedelphia.
- Thompson, P. (1998). Bank lending and the environment: policies and opportunities. *International Journal of Bank Marketing*.
- Witzany, J., (2017). *Credit Risk Management Pricing, Measurement, and Modeling*.
- Yazıcı, M., (2009). *Kredi Derecelendirme Kuruluşlarının Önemi ve Denetimi*, Maliye Finans Yazıları, Yıl 22 Sayı 82.

EKLER

EK 1: Kullanılan SAS Kodları

EK 1A: Her bir değişken bazında kayıp (missing) değerlerini saydırmak için kullanılan SAS kodları

```
proc means data=Ampirik_Calisma NMISS N; run;
```

EK 1B: Kayıp oranı % 70'ten fazla olan SON_YIL_CARI_ORAN değişkenin veri setinden çıkarılması için kullanılan SAS kodları

```
data Ampirik_Calisma_1 (drop= SON_YIL_CARI_ORAN);  
set Ampirik_Calisma;  
run;
```

EK 1C: Lojistik Regresyon için kullanılan SAS kodları

```
proc logistic data= Ampirik_Calisma_1 out= Ampirik_Calisma_2 descending;  
class rank / param=ref ;  
model default_flag = ARA_DONEM_ALACAK_TAHSIL_SURESI  
ARA_DONEM_BORC_DEVIR_SURESI_G  
ARA_DONEM_BRUT_KAR_MARJI_INSA  
ARA_DONEM_CARI_ORAN  
ARA_DONEM_LIKIDITE_ORANI  
ARA_DONEM_TPLM_BORC_TOPLAM_AKTIF  
ARA_DONEM_VOK_MARJI  
ARA_DONEM_YIL_KV_UV_FINANSAL  
BIR_ONCEKI_TOP_BORC_TOP_AKTIF  
BIR_ONCEKI_YIL_CARI_ORAN  
BIR_ONCEKI_YIL_KARLILIK_KALDIR  
BIR_ONCEKI_YIL_KV_UV_FINANSA  
BIR_ONCEKI_YIL_LIKIDITE_ORANI  
BR_ONCKI_DNM_NT_STS_REL_BYME  
DM_FD_12AyGcOdS  
FIRMANIN_FAALİYET_SURESI
```

```
FM_MI_BrcSatOrn
FN_MV_MVrGnDgOr_MlVr_Borc
FN_MZ_KVNRkLtOr
FN_MZ_SnNkRsOr
FRMNN_DEVAM_EDEN_YATIRIM_TUTARI
HAKIM_ORTAGIN_YASI
IKI_ONCEKI_TOP_BORC_TOP_AKTIF
IKI_ONCEKI_YIL_ALACAK_TAHSIL_S
IKI_ONCEKI_YIL_BORC_DEVIR_SURE
IKI_ONCEKI_YIL_CARI_ORAN
IKI_ONCEKI_YIL_KV_UV_FINANSA
IKI_ONCEKI_YIL_LIKIDITE_ORANI
IKI_ONCEKI_YIL_VOK_MARJI
ISLRI_YRTMDN_SRML_ORTA_SKTR_DNYM
MVCT_ADRSTKI_FAALIYET_SURESI
NET_SATISLARDAKI_REEL_BUYUME
ONCEKI_YIL_BORC_DEVIR_SURESI
ONCEKI_YIL_BRUT_KAR_MARJI_INS
ONCEKI_YIL_VOK_MAR
SON_AY_ODENEN_BANKAMIZ_CEKLERI
SON_YIL_ALACAK_TAHSIL_SURESI
SON_YIL_BORC_DEVIR_SURESI_GUN
SON_YIL_BRUT_KAR_MARJI__INSAAT
SON_YIL_KARLILIK_KALDIRACI_R
SON_YIL_KV_UV_FINANSAL_BORCL
SON_YIL_LIKIDITE_ORANI
SON_YIL_TOPLAM_BORC_TOPLAM_AKTIF
SON_YIL_VOK_MARJI
SR_FHGB_MEVFAALSUR
TCMB_TPLM_RISK_TCMB_TOPLAM_LIMIT
Rank ;
run;
```

EK 1D: Lojistik regresyon sonucunda modelin seçmiş olduğu değişkenlerin aralarındaki korelasyona bakmak için kullanılan SAS kodları

```
PROC CORR DATA= Ampirik_Calisma_2;
VAR BRUT_KAR_MARJI_INSAAT
CARI_ORAN
```

```
FIRMANIN_FAALİYET_SURESI  
FN_MZ_KVNRkLtOr  
FN_MZ_SnNkRsOr  
FRMNN_SHPLK_DURUMU_51_KIME_AIT  
HAKIM_ORTAGIN_YASI  
KS_VDL_FNNSL_BRCLR_TPLM_NKT  
FLYTLRDN_YRTLN_NKT  
BORC_DEVIR_SURESI_GUN  
KV_UV_FINANSAL_BORCLAR  
TOPLAM_BORC_TOPLAM_AKTIF  
TCMB_TPLM_RISK_TCMB_TOPLAM_LIMIT;  
RUN;
```

